Extensions to the Regression Discontinuity Design with Applications in Biostatistics

Mariam Olaide Adeleke

A thesis submitted for the degree of
Doctor of Philosophy

Department of Statistical Science
University College London

I, Mariam Olaide Adeleke confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The regression discontinuity (RD) design is a method for estimating a treatment effect in an observational study where there is a treatment allocation guideline that can be linked to the value of a continuous assignment variable and a pre-determined threshold. Typically, treatment is offered to patients whose assignment variable values lie above (or below) the threshold. Patients whose assignment variable values lie close to the threshold can be seen as exchangeable and typically, treatment effect estimation in an RD design involves comparing patients above and below the threshold. For a continuous outcome, estimating a treatment effect usually entails fitting local linear regression models for patients above and below the threshold. We propose the use of the thin plate regression spline to fit flexible regression models for patients above and below the threshold. Limited research has been done on an RD design for binary and time-to-event outcomes. For the binary outcome, we focused on the estimation of the risk ratio. The Wald and multiplicative structural mean models are approaches for estimating the risk ratio that can be applied to an RD design, however, they require additional assumptions. In this thesis, we have proposed an alternative approach for the estimation of the risk ratio that is based on the assumptions of the RD design. For the time-to-event outcome, the accelerated failure time (AFT) model was considered because it has some desirable properties in terms of interpreting causal effects. We propose an estimator of the acceleration factor that is based on the assumptions of an RD design. In addition to this, the structural AFT, a common approach for estimating the acceleration factor in observation studies, was discussed. Simulation studies were carried out to compare the proposed approaches with the existing ones, the results show that the proposed approaches compete favourably with, and in some cases, perform better than the existing methods. In addition, we have provided Bayesian alternatives to the three proposed approaches. Finally, we demonstrated these methods by applying them to real datasets on statin and metformin prescriptions.

Impact statement

With the increasing availability of large databases of observational datasets in medicine, developing methods for treatment effect estimation is of interest. Generally, treatment effect estimation from observational studies can be difficult because of the potential effect of unobserved confounding that can lead to a biased estimate of the treatment effect. As such, many methods of treatment effect estimation for observational studies are based on the untestable assumption of no unobserved confounding. In this thesis, we have explored methods for estimating treatment effect in an observational study using the regression discontinuity (RD) design, which can be used to estimate a treatment effect in the presence of unobserved confounding. An RD design is applicable where is there is a guideline for treatment allocation that can be linked to the value of a continuous assignment variable and a pre-specified threshold value. In many cases, treatments are prescribed according to pre-defined external guidelines. For instance, in the UK primary care, the National Institute for Health and Care Excellence (NICE) is the body that provides guidelines for the treatment of diseases and conditions for medical practitioners. As such, the RD design is applicable for treatment effect estimation in many data that are routinely collected as part of primary care. RD designs have been primarily used in the econometric literature, where methods for treatment effect estimation for a continuous outcome have been well researched, many of which involve fitting local linear regression models. As an alternative to the linear models, we propose a flexible data-driven approach for the estimation of treatment effect. In medical studies, in addition to a continuous outcome, binary and time-to-event outcomes are also of interest. Limited research have been carried out to develop methods for treatment effect estimation for binary and time-to-event outcomes in an RD design. We have proposed methods for treatment effect estimation for binary and time-to-event outcomes. We have illustrated how these methods can be applied to real datasets using data extracted from the THIN database to investigate the effect of statin prescription on low-density lipoprotein (LDL) cholesterol level and the effect of metformin prescription on type II diabetes-related complications in patients at risk of type II diabetes. Finally, we have presented Bayesian alternatives to the proposed approaches for continuous, binary and time-to-event outcomes. An advantage of a Bayesian approach is that findings from similar studies, for example, a clinical trial that has investigated the treatment of interest, can be incorporated in the treatment effect estimation process.

Acknowledgements

I am intensely grateful to my supervisor Dr Aidan O'Keeffe, for his guidance, support and timely constructive criticism during my PhD programme and the completion of this thesis. His continuous encouragements have been highly invaluable to my development as a researcher.

I express my appreciation to my second supervisor Prof Gianluca Baio, his advice and suggestions have been indispensable for my PhD research. I will like to thank Prof Giampiero Marra for his helpful comments and contributions to this thesis during my upgrade examination.

My special thanks to UCL for providing financial support for my PhD program. I appreciate all staff members and students of UCL Department of Statistical Science for creating a welcoming and supportive environment.

My special and immeasurable thanks to my parents and siblings for their prayers, undying love, continuous encouragement and support throughout my endeavours. I express my endless thanks to my husband Femi, for his continuous support and understanding throughout my PhD programme. I will like to extend my appreciation to my family members in Nigeria and the United Kingdom for their support during and beyond this study.

To the friends I made during my PhD programme, it has been a pleasure getting to know you. I wish you the best in your future goals and career plans.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Often, a main aim in medical research is to estimate the effect of a treatment (or intervention) on an outcome of interest. The gold standard approach for treatment effect estimation is to carry out a randomised controlled trial (Altman, 1991). Typically, this involves randomly allocating patients to one of two (or more) groups that either receive the treatment/ intervention under investigation or not. When randomisation is successful, it is expected that the patients assigned to each group are exchangeable and their characteristics are balanced in terms of other variables that are potentially related to both the outcome and the intervention. Therefore, any difference in the outcome observed between the two groups is attributable to the treatment and the effect of the treatment can be estimated by simply comparing the two groups, for instance, by comparing means of the outcome. However, in many situations, randomisation might not be possible. What we may have is simply an observation that some subjects received the treatment and others did not, where no-one has "intervened" and state of nature is observed. Such a scenario is referred to as an observational study.

In an observational study, comparing the treated and untreated might not give a reliable indication of the treatment effect. This is because the underlying relationship between the treatment and outcome might be distorted by other variables that are related to both the outcome and treatment, known as confounding variables, because the patients in the study are not randomly allocated to receive or not to receive treatment. As a result, treatment effect estimation in an observational study may require different methods that account for these features.

In this thesis, we consider an approach to treatment effect estimation that falls under the category of natural experiments, known as a regression discontinuity (RD)

design. These methods use quasi-randomisation that occurs in the treatment allocation process, and as a result, can produce a treatment effect estimate that accounts for unobserved confounders.

An RD design is an approach to treatment effect estimation in observational studies. It is used in scenarios where there is a guideline (rule) for allocating an intervention to subjects based on a continuous assignment variable and threshold. Generally, the intervention is offered to a subject whose assignment variable value lies above (or below) a specified value, known as the threshold.

For example, consider the prescription of anti-hypertensive drugs where a drug is prescribed depending on a patient's systolic blood pressure. The anti-hypertensive drug is to be prescribed to patients whose systolic blood pressure is greater than or equal to 140mmHg. Here, the assignment variable is the systolic blood pressure while the threshold is at 140mmHg.

In an RD design, a region around the threshold, called a bandwidth, is pre-specified to indicate assignment variable values that are considered "close" to the threshold, such that patients whose assignment variable values lie "close" to the threshold (above and below) can be seen as similar or exchangeable. That is, their characteristics will be similar to each other and balanced with respect to confounding variables, such as age, gender, etc., as we would have in a randomised controlled trial. For example, if we have a group of patients whose systolic blood pressure is just below the threshold of 140mmHg and another group of patients whose systolic blood pressure measurements lie just above the threshold of 140mmHg, these two groups might be considered similar in terms of important variables (such as diabetes status, body mass index, age etc.).

Conversely, if we have a group of patients whose systolic blood pressure is high, say around 155mmHg, and another group of patients with a low systolic blood pressure, say around 115mmHg, these two groups of patients will be different from each other and may differ with respect to body mass index, diabetes status etc., clearly, these groups may not be directly comparable. As a result, the threshold may be seen as a randomising device that allocates treatment to patients whose systolic blood pressure values lie above the threshold and withholds treatment for patients

whose systolic blood pressure values lie below the threshold, amongst patients whose systolic blood pressure values lie close to the threshold. It is therefore important that patients cannot manipulate the value of their assignment variable to ultimately determine their position with respect to the pre-determined and externally defined threshold, even when they are aware of the threshold rule, for the threshold to be a valid randomising device.

In medical studies, the manipulation of a treatment rule might not be an issue of concern as, in many cases, treatment rules are externally imposed and the assignment variable is computed based on inherent patient characteristics over which they do not have control. For example, in UK health care, the National Institute for Health and Care Excellence (NICE) is the body that provides guidelines on the treatment of diseases and conditions. NICE has recommended that statins, a class of cholesterol lowering drugs, be prescribed to patients whose risk of developing cardiovascular disease in 10 years is greater than or equal to 20% (NICE, 2008). The risk score is computed using risk prediction algorithms, such as the Framingham risk score (Wilson et al., 1998) or Q-RISK score (Hippisley-Cox et al., 2008) each of which uses patient information such as age, sex, BMI, cholesterol level, blood pressure etc. to compute the probability of developing cardiovascular disease in 10 years. Therefore, patients cannot manipulate their risk score even when they are aware of the guideline for statin prescription.

Depending on how well the treatment allocation guideline is adhered to, there are two types of RD designs. A sharp RD design occurs when treatment assignment is a deterministic function of the assignment variable, that is, every patient whose assignment variable value lies above the threshold will receive treatment while every patient whose assignment variable value lies below the threshold will not receive treatment. An example is depicted in Figure 1.1(a) where all patients with assignment variable values above the threshold received treatment while all patients whose assignment variable values lie below the threshold did not receive treatment. On the other hand, Figure 1.1(b) depicts a scenario where some patients whose assignment variable values lie below the threshold received treatment, while some patients whose assignment variable values lie above the threshold did not receive treatment. Here, there is a partial adherence to the treatment allocation guideline and the design is

called a fuzzy RD design.

In medical studies, a fuzzy RD design is more applicable, in general, than a sharp design. In some cases, a clinician might offer treatment to a patient whose assignment variable value lies below the threshold based on other characteristics of the patient if the clinician thinks the patient might benefit from the treatment and, likewise, some patients whose assignment variable values lie above the threshold might prefer to pursue other alternatives, such as lifestyle changes, rather than receive treatment or intervention. As a result, in this thesis, we focus on developing methods for treatment effect estimation in a fuzzy RD design. In addition, the sharp RD design can be seen as a special case of the fuzzy RD design and, as such, methods developed for the fuzzy RD design can be applied to the sharp RD design.

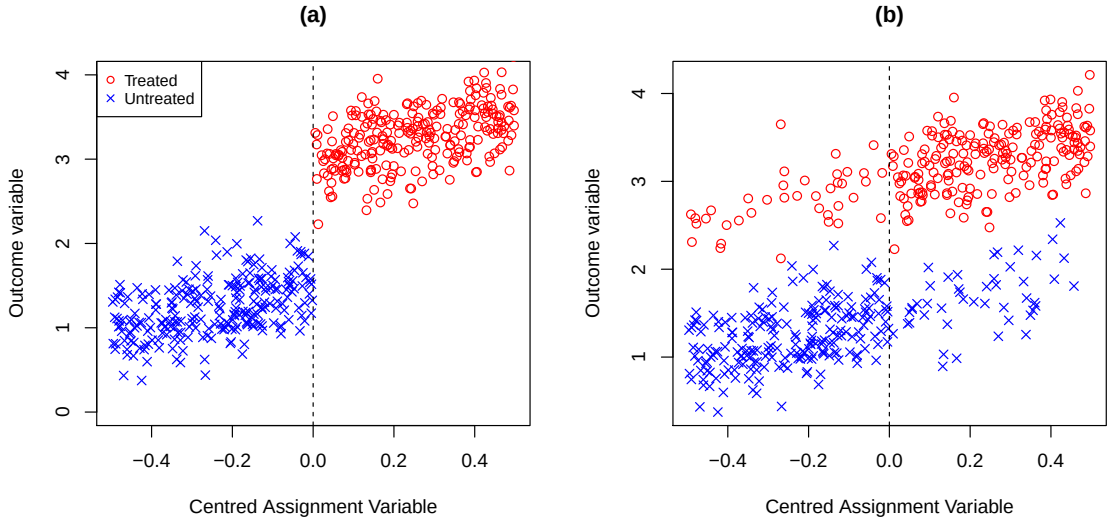

Figure 1.1: Plot representing the types of regression discontinuity designs. Threshold indicated by the dashed line, "treated" are patients that received the intervention while "untreated" are patients that did not receive the intervention (a) a sharp regression discontinuity design, (b) a fuzzy regression discontinuity design.

Consider the example on the anti-hypertensive drug introduced earlier. We might be interested in the effect of the anti-hypertensive drug on

(1) the blood pressure of patients after a certain time, or

(2) whether or not the blood pressure of patients is less than some value, say 120mmHg, after a certain time, or

(3) how long it takes for the blood pressure of patients to be 120mmHg or lower.

This illustrates three types of outcomes that might be of interest in medical studies: (1), (2) and (3) are examples of continuous, binary and time-to-event outcomes, respectively. In this thesis, we will explore new approaches to treatment effect estimation for these types of outcomes in a fuzzy RD design.

First, we consider an estimator of the treatment effect when the outcome of interest is continuous. The RD design has been widely applied and developed for this case (Cook, 2008; Angrist et al., 1996; Bor et al., 2014; van Leeuwan et al., 2016; Cattaneo et al., 2020b) and the local average treatment effect (LATE) has been established as an unbiased estimator of the treatment effect at the threshold. Estimating the LATE involves fitting regression models with the outcome of interest as the response variable and the assignment variable as the predictor. Typically, linear models might be used to model the relationship between the outcome and the assignment variable, although, the estimate from such models could be biased if the relationship between the outcome and assignment variable is not linear. Alternatively, especially in the Econometrics literature, local polynomial models (linear or quadratic are recommended) are fitted for patients above and below the threshold (Imbens and Lemieux, 2007; Calonico et al., 2014, 2020a; Cattaneo et al., 2020b). These methods might also produce a biased estimate if the underlying relationship can not be adequately modelled by the chosen degree of polynomial. We propose the use of thin plane regression splines (TPRS) for modelling the relationship between the outcome and the assignment variable. The TPRS approach is a completely data-driven method of fitting flexible models that captures non-linearity where it exists. That is, the TPRS does not require an assumption about the underlying relationship between the outcome and the assignment variables.

We further explore methods that have been used for the estimation of treatment effect when the outcome of interest is binary in the instrumental variable (IV) framework, namely, the Wald IV and multiplicative structural mean model (MSMM) estimators (Hernán and Robins, 2006; Didelez et al., 2010; Geneletti et al., 2019). These two methods are used for the estimation of the risk ratio and they can be applied to an RD design. It has been shown that the Wald IV estimator is not always consistent for the risk ratio and, therefore, this estimator is not always appropriate (Palmer et al., 2011). On the other hand, the MSMM estimator is consistent for the

risk ratio but sometimes, the estimate could be negative which is counterintuitive for an estimator of the risk ratio (Geneletti et al., 2019). In this work, we propose an estimator of the risk ratio that is based on the assumptions of the RD design alone.

Similarly, treatment effect estimation for time-to-event outcomes has not been widely explored in an RD design. Bor et al. (2014) proposed an estimator of the hazard ratio in a fuzzy RD design, but the estimator is only valid where the event of interest is rare and the follow-up time is short. As such, this estimator is not always applicable. In this thesis, we focus on the estimation of the treatment effect under the accelerated failure time (AFT) assumption, which is an alternative to the popular proportional hazard assumption for modelling time-to-event outcomes. One of the attractive properties of the treatment effect estimate under AFT assumption is that it can be interpreted directly in terms of the time-to-event outcome.

A widely used method for estimating the acceleration factor in observational studies is the structural accelerated failure time (S-AFT) approach, which has been shown to produce an unbiased estimate of the treatment effect (Hernán et al., 2005; Yang et al., 2020). However, this estimation approach requires a no unobserved confounding assumption which is a strong assumption that cannot be tested and could result in misleading estimates if it is violated. The effect of unobserved confounding remains a major challenge for the treatment effect estimation in most observational studies (Rosenbaum and Rubin, 1984; Nørgaard et al., 2017). One of the advantages of the RD design is that it does not rely on a no unobserved confounding assumption. As such, a treatment effect in an RD design may be recovered in the presence of unobserved confounding. Therefore, we propose an estimator of the acceleration factor that relies only on the assumptions of RD design.

Finally, we provide Bayesian alternatives for the proposed methods for continuous, binary and time-to-event outcomes. One advantage of the Bayesian approach to inference is that it provides a straight-forward way to incorporate additional information of some parameters of interest in the model. For instance, there might have been a similar study which might provide additional information about some parameters in the model being fitted or knowledge from an expert that might be useful, and such information could be adapted in the form of prior distributions.

In the next chapter, we give a formal introduction to the RD design and its assumptions are clearly stated. In Chapter 3, we propose the use of the TPRS approach for the estimation of the LATE as an alternative to the linear and local polynomial regressions. We compare the traditional methods and the proposed approach using simulation studies. The traditional methods and the TPRS method are then applied to a real data to estimate the effect of statin prescription on LDL cholesterol level. Chapter 4 focuses on the estimation of a risk ratio for a binary event in a fuzzy RD design. We describe the WALD IV and MSMM methods as the existing approaches to estimate the risk ratio and then propose an estimator that is based on the RD design assumptions that we term the RDD-RR estimator. These three methods are then compared using simulation studies and they are applied to real data on statin prescription. In Chapter 5, we present methods for estimating the acceleration factor in an RD design: the S-AFT estimator, an existing approach for estimating the acceleration factor in observational studies and a new estimator for the acceleration factor in a fuzzy RD design is proposed. We compare the new estimator and the S-AFT method using simulation studies. The two methods are then applied to a real data to estimate the effect of metformin prescription on Type II diabetes-related complications. Finally, in Chapter 6, we present Bayesian alternatives to the proposed methods presented in Chapters 3, 4 and 5. We compare the Bayesian and frequentist approaches using simulation studies.

Chapter 2

Regression discontinuity design and potential outcomes framework

In this chapter, we shall formally introduce the RD design. A review of the history of the RD design as well as the assumptions required to estimate a treatment effect in an RD design will be discussed.

2.1 Introduction

A regression discontinuity (RD) design is a method for estimating a treatment effect in observational studies where treatment is allocated based on a pre-specified rule. Specifically, a treatment or intervention will be prescribed to an individual (patient) if his/her value of a continuous assignment variable is above (or below) a pre-determined, externally-defined threshold and the treatment is not prescribed when the value of their assignment variable lies below (or above) the threshold. Since the threshold is external and not specific to individual patients, we might expect that patients whose assignment variable values are 'just above' and 'just below' the threshold will be balanced in terms of the distribution of potential confounders and comparable (Oldenburg et al., 2016).

The RD design was first introduced by Thistlethwaite and Campbell (1960). Their paper sought to determine the effect of public recognition for achievement on students' likelihood of receiving other scholarships and students' future career plans. The assignment variable was based on a student's score on a scholarship qualifying test. The scores were divided into 20 intervals and the corresponding threshold for receiving the scholarship was set at interval 11. They noted that students with

scores above the threshold received more public recognition, in that their names were published in newspapers and booklets distributed to colleges, universities and other scholarship granting agencies. The students answered a questionnaire approximately six months after the award that measured the outcome variables; number of students that obtained other scholarships and information on their future career plans. They found that public recognition for achievement increased students' chances of obtaining other scholarships but did not affect their future career plans. Following this, the RD design has been widely used in Economics (Imbens and Kalyanaraman, 2009; Hahn et al., 2001; Imbens and Angrist, 1994; Angrist et al., 1996; Xu, 2017) and in Political Science (Caughey and Sekhon, 2011; Erikson et al., 2015). In medicine, many treatments are prescribed according to externally-defined rules and, as such, the use of RD design may seem appealing in medical research. However, the RD design has been under-used in medical research until recently. One of the first papers to outline the use of the RD design in medicine was Linden et al. (2006), where a guideline of how to use an RD design was given. In this research paper, it was stated that RD design can be seen as one of the best alternatives to randomised trials (among other similar methodologies for treatment effect in observational studies).

In recent years, more researchers have begun to investigate the use of the RD design in medicine and epidemiology (Bor et al., 2014; Geneletti et al., 2015; O'Keeffe and Baio, 2016; van Leeuwan et al., 2016; Smith et al., 2017). O'Keeffe et al. (2014) gave an introduction to the RD design and its potential application in medical research. Bor et al. (2014) looked at the effect of early initiation of anti-retro viral treatment (ART) on the survivor experience of HIV patients. The assignment variable in the study was the CD4 count of patients. Patients whose CD4 count was below 200 cells μL were eligible for ART. They compared patients with CD4 count close to the threshold and found that patients just below the threshold have a lower hazard of death than patients just above the threshold.

Geneletti et al. (2015) presented a Bayesian approach to treatment effect estimation for a continuous, non-time-to-event, outcome. In their paper, the formal assumptions of the RD design were stated. Moscoe et al. (2015) presented an overview of papers published in epidemiology that considered an RD design as an approach to treatment effect estimation, thereby highlighting the under-utilisation of RD designs

in medical research. Basta and Halloran (2019) have highlighted the features of the RD design and how it can be used to evaluate the effectiveness of vaccines.

The majority of RD designs in the literature have focused on continuous outcomes that are often assumed to be normally distributed. When the outcome is binary, van Leeuwen et al. (2018) showed that the RD design provides valid estimates of the treatment effect (odds ratio) when used to validate the results of three randomised controlled trials (RCT). To mimic an RD design, age was used as an assignment variable in two of the RCTs; one RCT looked at the effect of corticosteroids on 14-day mortality after head injury, the second RCT compared the effect of two treatments for acute myocardial infarction on 30-day mortality. Baseline total cholesterol was used as assignment variable for the third RCT that aimed to estimate the effect of provastatin on the risk of coronary disease in elderly individuals. In all cases, the median values of the assignment variables were used as the threshold value. The odds ratios were estimated from logistic regression models. Geneletti et al. (2019) proposed a Bayesian approach to estimate the treatment effect (risk ratio) of statin prescription on the probability of low-density lipoprotein cholesterol levels reaching recommended levels. For a time-to-event outcome, Bor et al. (2014) explored the estimation of a hazard ratio on the effect of early initiation of ART on time-to-mortality in HIV patients for a sharp RD design.

Before further discussion of the methods of estimation, we consider the assumptions underlying an RD design.

2.2 Notations and Assumptions

Having described the RD design, we outline the mathematical notation required in this thesis. Consider a study with N patients, we define the outcome of interest for patient $i,\ i = 1, ..., N$ as Y_i. The assignment variable on which the treatment guideline is based is defined as X_i with the externally defined threshold value denoted by x_0. In addition, we define the centred assignment variable as $X_i^c = X_i - x_0$ which will have a value of 0 at the threshold. Furthermore, Z_i is defined as the threshold indicator which takes value 1 if the value of the assignment variable for patient i

is above the threshold and 0 otherwise. A_i is the treatment indicator, which takes the value 1 if patient i receives the treatment and 0 otherwise. The set of potential confounding variables (observed and unobserved) are defined as $\mathcal{C} = \mathcal{O} \cup \mathcal{U}$.

In addition, we define a bandwidth $h \in \mathbb{R}$ such that patients whose assignment variable values fall within the range $[x_0 - h, x_0 + h]$ are included in the analysis. The RD design assumes that patients with assignment variable values close to the threshold are similar and can be seen as exchangeable in order to be able to estimate a treatment effect and this bandwidth defines the area around the threshold in which patients are considered to be similar. In applications of the RD design for continuous outcomes, several studies have looked into determining the optimal bandwidth. The Imbens-Kalyanaraman (IK) optimal bandwidth, which was developed by Imbens and Kalyanaraman (2009), is commonly used in the econometrics literature. Calonico et al. (2014) noted that the confidence interval computed based on the IK optimal bandwidth yields a poor coverage of the treatment effect and they proposed another optimal bandwidth. Both the IK optimal bandwidth and the optimal bandwidth proposed by Calonico et al. (2014) are based on minimising the asymptotic mean square error of the treatment effect. Recently, Ricciardi et al. (2020) proposed a Dirichlet process mixture model to identify patients that are similar, above and below the threshold, such that those patients will be included in the analysis. In this thesis, we shall carry out a sensitivity analysis to check the behaviour of estimates as the bandwidth changes. We desire that estimates are not too sensitive to bandwidth changes and if severe changes is observed, care should be taken in interpreting an estimate. It has also been noted that the smaller the bandwidth, the smaller the bias in the treatment effect estimate at the threshold (Bor et al., 2014).

We now state a set of assumptions that are necessary for the identification of a treatment effect in an RD design (Imbens and Angrist, 1994; Hahn et al., 2001; Imbens and Lemieux, 2007; Geneletti et al., 2015), the treatment effect of interest here is defined as $\mathbb{E}(Y_i|A_i = 1) - \mathbb{E}(Y_i|A_i = 0)$. The assumptions are stated using the language of conditional independence (Dawid, 1979): if a random variable, A, is independent of another random variable B, conditional on C that is: $\mathbb{P}(A|B, C) = \mathbb{P}(A|C)$, this is represented as $A \perp\!\!\!\perp B|C$. On the other hand, $A \not\perp\!\!\!\perp B$ means that random variables A and B are not independent.

Assumption 1: The probability of receiving the treatment should be discontinuous at the threshold:

$$\lim_{x \downarrow x_0} \mathbb{P}(A_i = 1 | X_i = x) \neq \lim_{x \uparrow x_0} \mathbb{P}(A_i = 1 | X_i = x).$$

This assumption is fundamental in an RD design and is the basis of determining whether or not an RD design is applicable. We expect that the probability of receiving treatment for patients whose assignment variable values lie below the threshold should be different from the probability of receiving treatment for patients whose assignment variable values lie above the threshold. In a sharp RD design, the probability jumps from 0 to 1 for patients whose assignment variable values are below and above the threshold, respectively, a pictorial representation of this, using simulated data, is presented in Figure 2.1 (a). For the fuzzy design, as depicted in Figure 2.1 (b), although the probability does not jump from 0 to 1 at the threshold, there is still a clear discontinuity in the probability of receiving treatment at threshold.

Assumption 2: The threshold and treatment indicators are not independent.

$$A_i \not\perp\!\!\!\perp Z_i.$$

This assesses the validity of the threshold indicator as a determinant of treatment allocation. If Z_i and A_i are independent, an RD design cannot be used because treatment allocation occurs because of a reason other than the threshold rule. We note that for a sharp RD design, $A_i \equiv Z_i$, and therefore this assumption is satisfied where the design is sharp. When the treatment allocation guideline is not strictly adhered to, as in a fuzzy RD design, a causal effect can still be identified (Imbens and Angrist, 1994; Angrist et al., 1996). However, we must be satisfied that Z_i is still used, at least partially, in the allocation of treatment. For a fuzzy design, we would expect a reasonable level of correlation between A_i and Z_i

Assumption 3: The expectation of the outcome variable is continuous at the threshold conditional on the treatment.

$\mathbb{E}(Y_i | X_i = x, A_i = a, \mathcal{C}_i)$ is continuous at $x = x_0$ for $a = 0$ and $a = 1$.

The causal effect in an RD design is estimated from the change observed in the

expected outcome at the threshold. Therefore, this change should be due to the fact that the treatment allocation is different on either side of the threshold (or at least that the probability of receiving the treatment is discontinuous at the threshold). However, conditional on the treatment assignment being the constant on both sides of the threshold, we expect that no change would occur in the expected outcome. Otherwise, the threshold indicator (or other variables apart from the treatment) might be responsible for any observed change.

Assumption 4: The threshold indicator is independent of confounders conditional on the assignment variable.

$$Z_i \perp\!\!\!\perp \mathcal{C}_i | X_i.$$

This assumption ensures that a patient is placed above or below the threshold based on the value of the assignment variable only. Therefore, even if patients are able to manipulate the values of their characteristics, this does not necessarily have any effect on determining their position with regard to the threshold. For example, in the statin prescription example introduced in Chapter 1, although the risk score is computed using potential confounders, the determination of whether a patient is above or below the threshold is based solely on their 10-year risk score. That is, once the risk score has been computed, knowledge of the potential confounders is not required to determine treatment allocation.

Assumption 5: The threshold indicator and outcome variable are independent, conditional on the other variables.

$$Y \perp\!\!\!\perp Z | (X, A, \mathcal{C}).$$

This assumption indicates that the threshold indicator does not confound the relationship between the outcome and the treatment and the threshold indicator does not have a direct effect on the outcome. This assumption is important when patients have knowledge of the treatment guideline and are able to manipulate their outcome variables. This ensures that even with the knowledge of the treatment guideline, patients cannot manipulate their outcome variables to ultimately change the threshold indicator.

Assumption 6: There is no systematic non-adherence to the treatment guideline. This is applicable in a fuzzy RD design to ensure that non-adherence is not a deliberate action to defy the treatment allocation guideline. We are assuming that the person administering the treatment does not intentionally give the treatment to patients based on a guideline that is opposite of the recommended treatment guideline. For example, we let S_a and S_b be the guideline that an administrator uses to prescribe treatment to patients above and below the threshold, respectively with S_a and S_b equal to 0 (when the treatment is not given) and equal to 1 (when the treatment is given). We can write this assumption as:

$$\Pr(S_a = 0, S_b = 1) = 0.$$

That is, no one will intentionally give the treatment to all patients below the threshold and deny treatment to all patients above the threshold. This assumption, also known as "no defiers" assumption, ensures that a local treatment effect for compliers may be identified at the threshold.

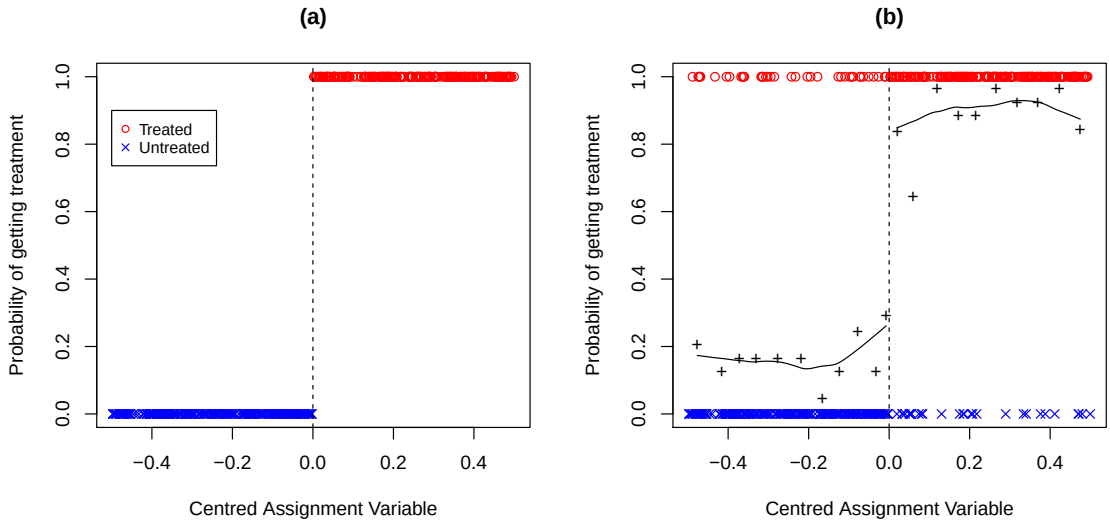

Figure 2.1: The probabilities of getting treated below and above the threshold in (a) sharp design and (b) fuzzy design. The probabilities jump from 0 to 1 in a sharp design. The black crosses are the expected probabilities calculated in bins.

The assumptions stated above are for continuous (non-time-to-event) outcomes. Further modifications or additions to the assumptions will be provided for the estimation of treatment effects for binary and time-to-event outcomes.

2.3 Potential outcomes framework

In some RD design literature, the design is introduced in terms of potential outcomes (Imbens and Angrist, 1994; Hahn et al., 2001) and in this section, we shall introduce the potential outcome framework. We recall that the outcome for patient i is given as Y_i and the treatment indicator is A_i. We define Y_i^0 as the outcome that would be observed for patient i if patient i did not receive the treatment and Y_i^1 is the outcome that would be observed if patient i received the treatment. $Y_i^a,\ a \in \{0,\ 1\}$ is known as the potential or counterfactual outcome.

Using the definition of a potential outcome, it is straightforward to define the treatment effect. For instance, let $\theta_i = Y_i^1 - Y_i^0$, then, θ_i is the individual treatment effect for patient i. That is, we compare what would happen if patient i is treated to what would happen if patient i is not treated. In the same vein, the average treatment effect could be stated as $\theta = \mathbb{E}(Y_i^1 - Y_i^0)$. However, we note that, for each patient, it is not possible to observe both potential outcomes Y_i^0 and Y_i^1. That is, if patient i receives the treatment, the observed outcome will be $Y_i = Y_i^1$. On the other hand, the observed outcome is $Y_i = Y_i^0$ if patient i does not receive treatment. This is known as the consistency assumption, which links the observed outcome with the potential outcomes (Pearl, 2010). We shall refer to the potential outcome framework in Chapters 4 and 5 when discussing estimators of treatment effect for binary and time-to-event outcomes.

In the next chapter, we shall discuss the local average treatment effect (LATE) estimator, an estimator of the treatment effect in a fuzzy RD design when the outcome of interest is continuous (non-time-to-event). We shall introduce a data-driven flexible approach to estimate the LATE using thin plate regression splines.

Chapter 3

Regression discontinuity design applied to a continuous outcome

In this chapter, we shall introduce an estimator of a treatment effect in an RD design where the outcome of interest is continuous (and non-time-to-event). We shall present traditional approaches for treatment effect estimation which involve fitting linear or local polynomial regression models and compare these to a novel thin plate regression spline approach. The traditional approaches require pre-specifying the form of relationship between the outcome and the assignment variable, which could produce misleading estimates if the true assignment variable - outcome relationship differs from the pre-specified one. We propose an alternative approach which involves fitting flexible, data-driven regression models called thin plate regression splines, which do not require pre-specification of the relationship between the outcome and assignment variable. These traditional approaches and the proposed method will then be applied to real data on the effect of statin prescription on LDL cholesterol level.

3.1 Introduction

The treatment effect estimate in an RD design is "local", in that the RD design focuses on estimating the treatment effect at the threshold. For a sharp RD design, the treatment effect at the threshold is estimated using the average treatment effect (ATE) (O'Keeffe and Baio, 2016). This may be estimated by fitting regression models (often linear) above and below the threshold. The ATE is then the difference

in means at the threshold and it is given in Equation 3.1 below:

$$\text{ATE} = \lim_{x \downarrow x_0} \mathbb{E}(Y_i | X_i = x) - \lim_{x \uparrow x_0} \mathbb{E}(Y_i | X_i = x). \tag{3.1}$$

Typically, suitable models are fitted to data above and below the threshold and the ATE is estimated as the difference in the expected value of the outcome at the threshold (i.e. conditional on $x_i^c = 0$).

For a sharp RD design, Equation 3.1 is equivalent to comparing treated and untreated patients because all patients above the threshold are treated and all patients below the threshold are untreated. However, for a fuzzy RD design, the ATE is a biased estimate of the treatment effect because patients above the threshold will include untreated patients and patients below the threshold will include treated patients. In a fuzzy design, the ATE estimate is analogous to the intention to treat estimate in a randomised trial where there is non-compliance to the treatment allocation (Sheiner and Rubin, 1995). The local average treatment effect (LATE) has been derived as an estimator of the treatment effect in a fuzzy RD design (Imbens and Angrist, 1994; Hahn et al., 2001). The LATE, also called a complier average causal effect, is 'local' because it is the average treatment effect for compliers, that is, patients who will take the treatment if they are offered it and will not take the treatment if it is not offered to them. The LATE is expressed as

$$\text{LATE} = \frac{\lim_{x \downarrow x_0} \mathbb{E}(Y_i | X_i = x) - \lim_{x \uparrow x_0} \mathbb{E}(Y_i | X_i = x)}{\mathbb{P}(A_i = 1 | Z_i = 1) - \mathbb{P}(A_i = 1 | Z_i = 0)}. \tag{3.2}$$

In this chapter, we shall explore methods that can be used to estimate the components of the LATE. In some applications, linear models are fitted above and below the threshold with the outcome as a predictor for the numerator terms and the treatment indicator as a predictor for the denominator terms (Imbens and Lemieux, 2007; Oldenburg et al., 2016; O'Keeffe and Baio, 2016). Alternatively, and what is often found in Econometric literature, local polynomial (linear or quadratic) models can be fitted for patients above and below the threshold (Imbens and Lemieux, 2007; Calonico et al., 2014, 2020a; Cattaneo et al., 2020a). As we have noted earlier, a linear model assumes that the assignment variable - outcome relationship is linear while the local polynomial model assumes that this relationship is linear or quadratic. These methods may produce biased estimates if the underlying relation-

ship is not the same as that assumed one. To alleviate this issue, an approach that can fit flexible models without pre-specifying the underlying relationship might be appealing. As such, we propose the use of the thin plate regression spline which can fit flexible regression models without prior assumption of the form of the underlying relationship. Next, we discuss in detail the traditional methods and the proposed approach.

3.2 Linear models

We begin by describing the simplest approach which entails fitting linear models to observations above and below the threshold that fall within the bandwidth (h) to estimate the components of the LATE (Imbens and Lemieux, 2007). Without loss of generality, we assume that $X_i \in [0,1]$ and we define $\mathcal{A} = \{i|Z_i = 1 \ \cap \ X_i \in [x_0, x_0+h]\}$ and $\mathcal{B} = \{i|Z_i = 0 \ \cap \ X_i \in [x_0-h, x_0)\}$ to be the sets of patients whose assignment variable values are above and below the threshold respectively that fall within the bandwidth.

The numerator of the LATE at the threshold is the difference in the intercepts of the two models that are presented in Equation 3.3 below.

$$\begin{aligned} \mathbb{E}(Y_i|Z_i = 1) &= \beta_{01} + \beta_{11}x_i^c \quad \text{for } i \in \mathcal{A}, \\ \mathbb{E}(Y_i|Z_i = 0) &= \beta_{00} + \beta_{10}x_i^c \quad \text{for } i \in \mathcal{B}. \end{aligned} \tag{3.3}$$

The denominator of the LATE is the difference in the probabilities of receiving treatment above and below the threshold and can be obtained as the difference of the intercepts of the two models in Equation 3.4

$$\begin{aligned} \mathbb{E}(A_i|Z_i = 1) &= \gamma_{01} + \gamma_{11}x_i^c \quad \text{for } i \in \mathcal{A}, \\ \mathbb{E}(A_i|Z_i = 0) &= \gamma_{00} + \gamma_{10}x_i^c \quad \text{for } i \in \mathcal{B}. \end{aligned} \tag{3.4}$$

Defining λ_{L} as the LATE estimator that is obtained from fitting linear models, the LATE is written

$$\lambda_{\mathrm{L}} = \frac{\beta_{\mathrm{L}}}{\gamma} = \frac{\beta_{01} - \beta_{00}}{\gamma_{01} - \gamma_{00}}.$$

Alternatively, the linear model LATE can be estimated using a two-stage least squares (TSLS) approach (Imbens and Lemieux, 2007). The estimate of the TSLS and the λ_L are equivalent, but the TSLS estimate may not enjoy the distributional assumptions used when fitting linear regression models, that might be useful for the construction of confidence intervals.

A challenge of LATE estimation is that it is not straightforward to estimate the variance of the LATE because it is a ratio of two random variables. O'Keeffe and Baio (2016) proposed an approach to estimate the variance of the LATE estimator using a Taylor approximation:

$$\begin{aligned}\mathrm{Var}(\hat{\lambda}_{\mathrm{L}}) = \frac{\hat{\sigma}^2}{\hat{\gamma}^2}\left(\sum_{i\in\mathcal{A}}\alpha_i^2 + \sum_{i\in\mathcal{B}}\beta_i^2\right) + \frac{\hat{\beta}_{\mathrm{L}}^2}{\hat{\gamma}^4}\left(\hat{\phi}_1^2\sum_{i\in\mathcal{A}}\alpha_i^2 + \hat{\phi}_0^2\sum_{i\in\mathcal{B}}\beta_i^2\right) \\ - \frac{2\hat{\beta}_{\mathrm{L}}}{\hat{\gamma}^3}\left(\hat{\rho}_1\sum_{i\in\mathcal{A}}\alpha_i + \hat{\rho}_0\sum_{i\in\mathcal{B}}\beta_i\right).\end{aligned}$$

$$\begin{aligned}\text{where}\quad \bar{x}_1 &= \frac{1}{n_1}\sum_{i\in\mathcal{A}}x_i, \quad \bar{x}_0 = \frac{1}{n_0}\sum_{i\in\mathcal{B}}x_i,\\ \alpha_i &= \frac{1}{n_1} + \frac{\bar{x}_1^2 - \bar{x}_1 x_i}{\sum_{i\in\mathcal{A}}(x_i-\bar{x}_1)^2},\\ \beta_i &= \frac{1}{n_0} + \frac{\bar{x}_0^2 - \bar{x}_0 x_i}{\sum_{i\in\mathcal{B}}(x_i-\bar{x}_0)^2},\\ \hat{\rho}_1 &= \frac{1}{n_1-1}\sum_{i\in\mathcal{A}}(y_i-\hat{y}_i)(a_i-\hat{a}_i),\\ \hat{\rho}_0 &= \frac{1}{n_0-1}\sum_{i\in\mathcal{B}}(y_i-\hat{y}_i)(a_i-\hat{a}_i),\\ \hat{\phi}_1 &= \frac{1}{n_1-2}\sum_{i\in\mathcal{A}}(a_i-\hat{a}_i)^2\\ \hat{\phi}_0 &= \frac{1}{n_0-2}\sum_{i\in\mathcal{B}}(a_i-\hat{a}_i)^2\\ s_1^2 &= \frac{1}{n_1-2}\sum_{i\in\mathcal{A}}(y_i-\hat{y}_i)^2,\\ s_0^2 &= \frac{1}{n_0-2}\sum_{i\in\mathcal{B}}(y_i-\hat{y}_i)^2,\\ \hat{\sigma}^2 &= \frac{(n_1-2)s_1^2 + (n_0-2)s_0^2}{n_1+n_0-4},\end{aligned}$$

n_1 and n_0 the represents number of patients above and below the threshold respectively.

The approach we have described will be appropriate when the underlying relationship between the outcome and the assignment variable is linear above and below the threshold. We will now discuss more flexible methods that might be better where the underlying relationship is not linear.

3.3 Robust local polynomial regression

Calonico et al. (2014) and Calonico et al. (2020a) have proposed a robust bias-corrected approach for estimating the LATE in an RD design. This approach is a commonly-used approach for estimating the LATE in econometric literature which we shall now describe.

The estimate of the LATE from a local polynomial regression of degree p that is fitted within a bandwidth of h is given as

$$\hat{\tau}_{\text{FRD}}(h) = \frac{\hat{\mu}_{Y1,\ p}(h) - \hat{\mu}_{Y0,\ p}(h)}{\hat{\mu}_{A1,\ p}(h) - \hat{\mu}_{A0,\ p}(h)},$$

where for $z \in \{0,\ 1\}$, $\hat{\mu}_{Yz,\ p}(h)$ and $\hat{\mu}_{Az,\ p}(h)$ are intercepts of the local polynomial regression of degree p with the outcome Y and treatment A as the response variables respectively. Typically, a local linear regression (that is, $p = 1$) is recommended in an RD design (Imbens and Lemieux, 2007). In this case, for $z \in \{0,\ 1\}$ and $\mathcal{J} \in \{\mathcal{A},\ \mathcal{B}\}$, the coefficients of the local linear regression are estimated as follows

$$\left(\hat{\mu}_{Yz,\ 1}(h),\ \hat{\mu}^{(1)}_{Yz,\ 1}(h)\right)^{\top} = \arg\min_{a_0,\ a_1} \sum_{i \in \mathcal{J}} (Y_i - a_0 - a_1 X_i^c)^2 K\left(\frac{X_i^c}{h}\right),$$
$$\left(\hat{\mu}_{Az,\ 1}(h),\ \hat{\mu}^{(1)}_{Az,\ 1}(h)\right)^{\top} = \arg\min_{b_0,\ b_1} \sum_{i \in \mathcal{J}} (A_i - b_0 - b_1 X_i^c)^2 K\left(\frac{X_i^c}{h}\right),$$

where $\hat{\mu}_{Yz,\ 1}(h)$, $\hat{\mu}^{(1)}_{Yz,\ 1}(h)$, $\hat{\mu}_{Az,\ 1}(h)$ and $\hat{\mu}^{(1)}_{Az,\ 1}(h)$ are the estimated values for a_0, a_1, b_0, and b_1 respectively, and $K(u)$ represents a kernel function. Typically, in an RD design, the triangular kernel is recommended because of its boundary

properties (Cattaneo et al., 2020b), and the triangular kernel defined as

$$K(u) = \begin{cases} 1 - |u| & \text{if } |u| \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

It has been shown that estimates from a local polynomial regression are biased (Imbens and Angrist, 1994; Hahn et al., 2001), therefore the LATE estimate $\hat{\tau}_{\text{FRD}}(h)$ will also be biased. The bias of $\hat{\tau}_{\text{FRD}}(h)$ can be estimated as

$$\begin{aligned} \hat{B}_{\text{FRD}}(h,\ r) = &\left(\frac{1}{\hat{\tau}_{A,\ \text{SRD}}(h)} \frac{\hat{\mu}^{(2)}_{Y1,\ q}(r)}{2} - \frac{\hat{\tau}_{Y,\ \text{SRD}}(h)}{\hat{\tau}^2_{A,\ \text{SRD}}(h)} \frac{\hat{\mu}^{(2)}_{A1,\ q}(r)}{2} \right) \mathcal{R}_{1,\text{FRD}}(h) \\ &- \left(\frac{1}{\hat{\tau}_{A,\ \text{SRD}}(h)} \frac{\hat{\mu}^{(2)}_{Y0,\ q}(r)}{2} - \frac{\hat{\tau}_{Y,\ \text{SRD}}(h)}{\hat{\tau}^2_{A,\ \text{SRD}}(h)} \frac{\hat{\mu}^{(2)}_{A0,\ q}(r)}{2} \right) \mathcal{R}_{0,\text{FRD}}(h). \end{aligned}$$

Where for $z \in \{0,\ 1\}$

$\hat{\tau}_{A,\ \text{SRD}}(h) = \hat{\mu}_{A1,\ p}(h) - \hat{\mu}_{A0,\ p}(h)$,

$\hat{\tau}_{Y,\ \text{SRD}}(h) = \hat{\mu}_{Y1,\ p}(h) - \hat{\mu}_{Y0,\ p}(h)$

$\hat{\mu}^{(2)}_{Yz,\ q}(r)$ and $\hat{\mu}^{(2)}_{Az,\ q}(r)$ are the coefficients of $(X^c_i)^2$ in a local polynomial regression of degree $q \geq 2$ and $q > p$ fitted within bandwidth r with the outcome and treatment indicator as response variable respectively.

$\mathcal{R}_{0,\text{FRD}}(h)$ and $\mathcal{R}_{1,\text{FRD}}(h)$ are asymptotically bounded observed quantities given in Calonico et al. (2014).

It is recommended that the bandwidth (r) for estimating the bias should be greater than the bandwidth (h) for the LATE estimate (Calonico et al., 2014). Therefore, the bias corrected estimate of the LATE is given as

$$\hat{\tau}^{\text{bc}}_{\text{FRD}}(h,\ r) = \hat{\tau}_{\text{FRD}}(h) - \hat{B}_{\text{FRD}}(h,\ r).$$

In addition, Calonico et al. (2020a) also derived an optimal bandwidth for estimating $\hat{\tau}^{\text{bc}}_{\text{FRD}}(h,\ r)$ as well as a robust standard error estimate that provides an accurate coverage of the confidence interval of $\hat{\tau}^{\text{bc}}_{\text{FRD}}(h,\ r)$. The robust bias-corrected approach, as well as the optimal bandwidth and the robust standard error, is implemented in the `rdrobust` package in `R` (Calonico et al., 2020b).

We shall explore the use of thin plate regression splines for estimating the LATE. Rather than fitting local kernel regression models above and below the threshold, where the kernel type and polynomial degree need to be specified, a thin plate regression spline may represent a more flexible model with fewer assumptions to capture the relationship between the assignment variable and the outcome of interest.

3.4 Flexible regression models

We have described the linear and robust bias-corrected approaches of estimating the LATE in a fuzzy RD design. The linear model is based on the assumption that the relationship between the outcome and the assignment variable is linear above and below the threshold. However, when the relationship is not linear, the estimator derived from the linear models may yield a biased estimate of the treatment effect. The robust bias-corrected approach also requires a specification of the degree of polynomial to be fitted, if the choice of polynomial degree cannot adequately model the relationship between the outcome and the assignment variable, it might also lead to a biased estimate. In this section, we focus on the numerator of the LATE while the denominator will be fitted using linear models in Equation 3.4. We propose the use of a flexible regression spline model that is completely data driven and therefore does not assume the form of the underlying relationship between the outcome and assignment variable, called thin plate regression spline (TPRS). Before we describe the TPRS, we shall first describe regression splines.

3.4.1 Regression splines

Regression splines provide a flexible method to estimate a functional relationship between variables without specifying the form of the relationship (Hastie and Tibshirani, 1999). Piece-wise polynomials are fitted between internal cut-points, known as knots. The knots are defined in ascending order within the range of values of X as $a < t_1 < t_2 < ... < t_K < b$, where a and b are the minimum and maximum values of X respectively. The number of knots (K) in a regression spline determines how smooth or wiggly the function will be. A regression spline of degree ν or order $\nu + 1$

is given as

$$s(x) = \beta_0 + \sum_{j=1}^{\nu} \beta_j \, x^j + \sum_{k=1}^{K} \theta_k \, (x - t_k)_+^{\nu}, \tag{3.5}$$

with $x_+ = \max(0, x)$. A popular choice for ν is 3 which yields a piece-wise cubic polynomial with continuous derivatives up to order 2 at each knot. van Leeuwan et al. (2016) has used a cubic regression spline for the estimation of the numerator of the LATE. A potential drawback of the standard regression spline approach is that the number of knots and the locations at which knots are placed must be chosen before fitting the model. As depicted in Figure 3.1, the higher the number of knots, the more wiggly the fitted line. This is particularly relevant to our context because different estimates of the treatment effect at the threshold may be obtained, depending on the number and location of knots. As a result, a TPRS model (Duchon, 1977) may be a more attractive approach to flexible modelling within an RD design.

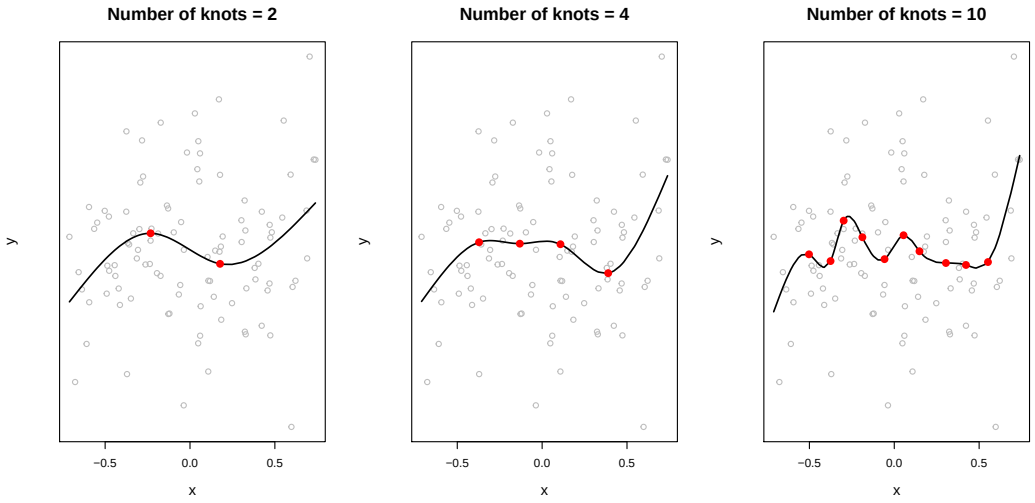

Figure 3.1: Fitted lines of cubic regression splines with varying number of knots to illustrate how the number and location of knots influences the fit of a regression spline model. The red dots are the knot positions.

3.4.2 Thin plate regression spline

The thin plate spline is an alternative which can be used to fit a flexible regression spline, but there is no requirement of knot location and number of knots and the fit is completely data driven. It uses all data points as knots and the smoothness

of the spline is penalised using a smoothing parameter (Duchon, 1977). First, we describe the thin plate spline model in a general setting and then describe how it will be applied to estimate the numerator of the LATE.

It should be noted that the thin plate spline can be used to model flexible relationships between an outcome and more than one predictor. In a RD design, however, we focus on fitting the thin plate spline with one predictor (the assignment variable) and note that additional predictors can be added if it is necessary. The thin plate spline with predictor x_i and outcome variable y_i is given as

$$y_i = f(x_i) + \epsilon_i$$

where

$$f(x) = \sum_{j=1}^{n} \delta_j \eta(|x - x_j|) + \sum_{k=1}^{M} \alpha_k \phi_k(x). \tag{3.6}$$

n is the number of unique values of x, M is defined to be equal to $\binom{m+d-1}{d}$, where d is the number of predictors in the model and m is chosen to be the smallest integer that satisfies $2m > d + 1$. In the RD design example we are considering, $d = 1$ and $m = 2$, therefore, $M = 2$.

The basis function $\eta(|x - x_j|)$ is defined as

$$\eta(|x-x_j|) = \begin{cases} \dfrac{(-1)^{m+1+d/2}}{2^{2m-1}\pi^{d/2}(m-1)!(m-d/2)!}|x - x_j|^{2m-d}\log(|x - x_j|) & \text{if } d \text{ is even} \\ \dfrac{\Gamma(d/2-m)}{2^{2m}\pi^{d/2}(m-1)!}|x - x_j|^{2m-d} & \text{if } d \text{ is odd.} \end{cases}$$

In the case where $d = 1$ and $m = 2$, as we have in the RD design example, we have that

$$\eta\left(|x - x_j|\right) = \frac{\Gamma(0.5)}{12\sqrt{\pi}}\left(|x - x_j|\right)^3.$$

The smooth functions $\phi_k(x)$ are defined as $\phi_k(x) = (\phi_1(x), \phi_2(x))^\top = (1, x)^\top$. The parameters $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top$ and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$ are coefficients of the model that need to be estimated. Because there are more parameters $(n + 2)$ than the number of unique values of x (n), the estimation process is subject to an additional identifiability constraint: $\mathbf{G}^\top\boldsymbol{\delta} = \mathbf{0}$, where $G_{ij} = \phi_j(x_i)$.

The parameters of the thin plate spline model are estimated by minimising

$$||\mathbf{y} - \mathbf{E}\boldsymbol{\delta} - \mathbf{G}\boldsymbol{\alpha}||^2 + \lambda\boldsymbol{\delta}^\top\mathbf{E}\boldsymbol{\delta} \tag{3.7}$$

subject to $\mathbf{G}^\top\boldsymbol{\delta} = \mathbf{0}$, with respect to $\boldsymbol{\delta}$ and $\boldsymbol{\alpha}$. Here $\mathbf{E}$ is an $n \times n$ matrix with $E_{ij} = \eta(|x_i - x_j|)$ and λ is a smoothing parameter (with $\lambda > 0$) that controls the trade-off between the fit of f to the set of data and the smoothness of f. A small value of λ will lead to a more wiggly fit while a high λ results in a smooth fit as shown in Figure 3.2. In this thesis, the smoothing parameter is estimated by minimising the generalised cross-validation criterion which is readily available in the `mgcv` package in `R` (Wood, 2019).

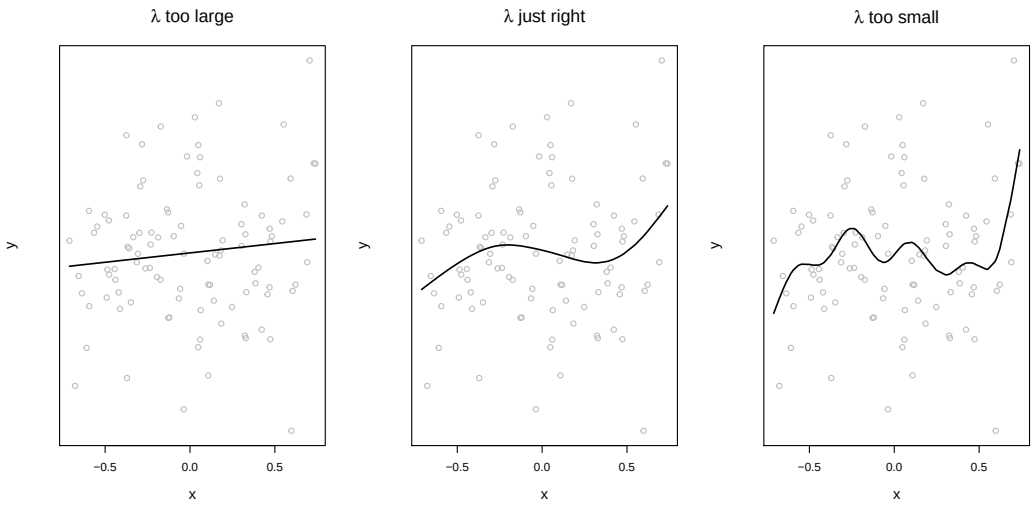

Figure 3.2: Fitted lines of three TPRS models with varying values of λ to illustrate how the value of λ influences the smoothness of TPRS model.

Fitting of the thin plate spline by minimising the expression in Equation 3.7 could be computationally intensive. This is because the number of parameters to be estimated is the same as the number of unique values of x in the dataset. A way around this is to fit what is called the thin plate regression spline (Wood, 2003, 2017). This involves reducing the dimension of the parameters of the basis functions $\boldsymbol{\delta}$ from n to $k < n$.

We define $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{U}^\top$ to be the eigen decomposition of $\mathbf{E}$ where $\mathbf{U}$ is the $n \times n$ matrix whose columns represent the eigenvectors and $\mathbf{D}$ is the $n \times n$ diagonal matrix of the eigenvalues. We define $\mathbf{U}_k$ to be the $n \times k$ matrix that consists of the first k columns of $\mathbf{U}$ and $\mathbf{D}_k$ be the upper left $k \times k$ submatrix of $\mathbf{D}$. The dimension of $\boldsymbol{\delta}$ is truncated from n to k by replacing $\mathbf{U}\mathbf{D}\mathbf{U}^\top$ with $\mathbf{U}_k\mathbf{D}_k\mathbf{U}_k^\top$ and we define

$\boldsymbol{\delta} = \mathbf{U}_k\boldsymbol{\delta}_k$. Here, $\boldsymbol{\delta}_k$ is the vector of parameters of the thin plate regression spline with dimension k. The minimisation problem in Equation 3.7 is now reduced to the minimisation of

$$||\mathbf{y} - \mathbf{U}_k\mathbf{D}_k\boldsymbol{\delta}_k - \mathbf{G}\boldsymbol{\alpha}||^2 + \lambda\boldsymbol{\delta}_k^\top\mathbf{D}_k\boldsymbol{\delta}_k$$

subject to $\mathbf{G}^\top\mathbf{U}_k\boldsymbol{\delta}_k = \mathbf{0}$, with respect to $\boldsymbol{\delta}_k$ and $\boldsymbol{\alpha}$. This constrained minimisation problem can be changed to an unconstrained one by accounting for the identifiability constraint $\mathbf{G}^\top\mathbf{U}_k\boldsymbol{\delta}_k = \mathbf{0}$ in the unconstrained minimisation problem. This can be done by determining the orthogonal column basis $\mathbf{Z}_k$ such that $\mathbf{G}^\top\mathbf{U}_k\mathbf{Z}_k = \mathbf{0}$; setting $\boldsymbol{\delta}_k = \mathbf{Z}_k\tilde{\boldsymbol{\delta}}$ results in the unconstrained problem - minimise

$$||\mathbf{y} - \mathbf{U}_k\mathbf{D}_k\mathbf{Z}_k\tilde{\boldsymbol{\delta}} - \mathbf{G}\boldsymbol{\alpha}||^2 + \lambda\tilde{\boldsymbol{\delta}}^\top\mathbf{Z}_k^\top\mathbf{D}_k\mathbf{Z}_k\tilde{\boldsymbol{\delta}} \tag{3.8}$$

with respect to $\tilde{\boldsymbol{\delta}}, \boldsymbol{\alpha}$. The parameters of the full thin plate spline can be retrieved using the following equation $\boldsymbol{\delta} = \mathbf{U}_k\mathbf{Z}_k\tilde{\boldsymbol{\delta}}_k$.

To estimate the numerator of the LATE, we need to determine the predicted value of y at the threshold from the regression model. We shall now describe how to extract the predicted value of y at a point $x = x_0$, that is, $f(x_0)$ from Equation 3.6. The solution for $\boldsymbol{\beta}_{\text{tpr}} = \left(\tilde{\boldsymbol{\delta}}, \boldsymbol{\alpha}\right)^\top$ from the minimisation problem in Equation 3.8 is given as

$$\hat{\boldsymbol{\beta}}_{\text{tpr}} = \left(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}} + \mathbf{S}\right)^{-1}\tilde{\mathbf{X}}^\top\mathbf{y},$$

where

$$\tilde{\mathbf{X}} = \begin{bmatrix}\mathbf{U}_k\mathbf{D}_k\mathbf{Z}_k & \mathbf{G}\end{bmatrix},$$

$$\text{and} \qquad \mathbf{S} = \lambda\begin{bmatrix}\mathbf{Z}_k^\top\mathbf{D}_k\mathbf{Z}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}.$$

From here, the estimated parameters of the full thin plate spline in Equation 3.6 are then recovered as

$$\hat{\boldsymbol{\beta}}_{\text{tp}} = \mathbf{B}\hat{\boldsymbol{\beta}}_{\text{tpr}} = \left(\hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\alpha}}\right)^\top,$$

where

$$\mathbf{B} = \begin{bmatrix}\mathbf{U}_k\mathbf{Z}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2\end{bmatrix}.$$

and

$$\mathbf{I}_2 \text{ is a } 2 \times 2 \text{ identity matrix.}$$

To predict the value of y at a point $x = x_0$, we define the corresponding design matrix for $x = x_0$ as follows

$$\tilde{\boldsymbol{x}}_0 = \left[\eta\left(|x_0 - x_1|\right), \ldots, \eta\left(|x_0 - x_n|\right),\ 1,\ x_0\right]^\top.$$

Therefore, the predicted value at point x_0 is given as

$$\hat{f}(x_0) = \tilde{\boldsymbol{x}}_0^\top \hat{\boldsymbol{\beta}}_{\text{tp}}.$$

The TPRS model can be fitted using the `gam` function in the `mgcv` package in R and predicted values are obtained by applying the `predict` function to a `gam` object.

3.4.3 Thin plate spline in an RD design

We will now discuss the LATE estimator where the numerator is estimated using thin plate regression spline models described in Section 3.4.2. First, we define the following:

The outcome of interest $\mathbf{y} = (y_1, \ldots, y_n)^\top$ is partitioned into $\mathbf{y} = (\mathbf{y}_0, \mathbf{y}_1)^\top$ such that $\mathbf{y}_0$ and $\mathbf{y}_1$ represent the observed outcomes for patients whose assignment variable values lie below and above the threshold, respectively.

The treatment indicator $\mathbf{a} = (a_1, \ldots, a_n)^\top$ is partitioned as $\mathbf{a} = (\mathbf{a}_0, \mathbf{a}_1)^\top$ such that $\mathbf{a}_0$ and $\mathbf{a}_1$ represent the treatment indicators for patients whose assignment variable values lie below and above the threshold, respectively.

The assignment variable $\mathbf{x} = (x_1, \ldots, x_n)^\top$ is partitioned into $\mathbf{x} = (\mathbf{x}_0, \mathbf{x}_1)^\top$ such that $\mathbf{x}_0$ and $\mathbf{x}_1$ represents the assignment variable for patients below and above the threshold respectively and n_0 and n_1 as the number of patients whose assignment variable values lies below and above the threshold respectively. We define the design matrix $\mathbf{X}_0 = (\mathbf{1}_0 \ \ \mathbf{x}_0)$ and $\mathbf{X}_1 = (\mathbf{1}_1 \ \ \mathbf{x}_1)$ for patients below and above the threshold

respectively, where, for $j \in \{0,\ 1\}$, $\mathbf{1}_j$ is a $n_j \times 1$ matrix of 1's.

The LATE estimator is then defined as follows

$$\lambda_{\text{tprs}} = \frac{\beta}{\gamma} = \frac{f_1(x_0) - f_0(x_0)}{\gamma_{01} - \gamma_{00}},$$

and for $j \in \{0,\ 1\}$,

$$\hat{f}_j(x_0) = \tilde{\boldsymbol{x}}_{0j}^{\top} \mathbf{B}_j \hat{\boldsymbol{\beta}}_j,$$
$$\hat{\gamma}_{0j} = \mathbf{e}^{\top} \hat{\boldsymbol{\gamma}}_j,$$

where $\tilde{\boldsymbol{x}}_{0j}^{\top}$ and $\mathbf{B}_j$ are as described in Section 3.4.2 and $\hat{\boldsymbol{\beta}}_j = \left(\hat{\tilde{\boldsymbol{\delta}}}_j, \hat{\boldsymbol{\alpha}}_j\right)$ are the parameters from the unconstrained thin plate regression splines as given in Equation 3.8. We define $\mathbf{e} = (1 \ \ x_0)^{\top}$ and $\hat{\boldsymbol{\gamma}}_j = \left(\hat{\boldsymbol{\gamma}}_{0j},\ \hat{\boldsymbol{\gamma}}_{1j}\right)$ are the parameters of the linear regression models with treatment as the response variable and the assignment variable as a predictor presented in Equation 3.4. We will now proceed to discuss the variance of the TPRS approach to LATE estimation.

3.4.3.1 Variance of estimator

We recall that the TPRS estimator of the LATE is given as

$$\hat{\lambda}_{\text{tprs}} = g(\hat{\beta}, \hat{\gamma}) = \frac{\hat{\beta}}{\hat{\gamma}},$$

where $\hat{\beta}$ is estimated from fitting TPRS models whose parameters are asymptotically normally distributed (Wood, 2003). Similarly, $\hat{\gamma}$ is estimated from linear models in Equation 3.4 whose parameters are also normally distributed (Seber and Lee, 2012). Therefore, we can apply the multivariate delta method to estimate the variance of $\hat{\lambda}_{\text{tprs}}$:

$$\begin{aligned} V(\hat{\lambda}) &\approx \Delta \left(g(\hat{\beta}, \hat{\gamma})\right)^{\top} \ \hat{\Sigma}_{\gamma,\beta} \ \Delta \left(g(\hat{\beta}, \hat{\gamma})\right) \\ &= \begin{pmatrix} \frac{1}{\hat{\gamma}} & -\frac{\hat{\beta}}{\hat{\gamma}^2} \end{pmatrix} \hat{\Sigma}_{\gamma,\beta} \begin{pmatrix} \frac{1}{\hat{\gamma}} & -\frac{\hat{\beta}}{\hat{\gamma}^2} \end{pmatrix}^{\top}. \end{aligned}$$

Here $\hat{\Sigma}_{\beta,\gamma}$ is an estimate of the variance-covariance matrix of $\left(\hat{\beta}\ \hat{\gamma}\right)^\top$ where

$$\begin{aligned}
\left[\hat{\Sigma}_{\beta,\gamma}\right]_{11} &= \tilde{\boldsymbol{x}}_{01}^\top \mathbf{B}_1 \text{Var}(\hat{\boldsymbol{\beta}}_1)\mathbf{B}_1^\top \tilde{\boldsymbol{x}}_{01} + \tilde{\boldsymbol{x}}_{00}^\top \mathbf{B}_0 \text{Var}(\hat{\boldsymbol{\beta}}_0)\mathbf{B}_0^\top \tilde{\boldsymbol{x}}_{00};\\
\left[\hat{\Sigma}_{\beta,\gamma}\right]_{12} &= \tilde{\boldsymbol{x}}_{01}^\top \mathbf{B}_1 \left(\tilde{\mathbf{X}}_1^\top \tilde{\mathbf{X}}_1 + \mathbf{S}_1\right)^{-1} \tilde{\mathbf{X}}_1^\top \hat{\text{Cov}}(\mathbf{Y}_1, \mathbf{A}_1)\mathbf{X}_1 \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} \mathbf{e}\\
&\quad + \tilde{\boldsymbol{x}}_{00}^\top \mathbf{B}_0 \left(\tilde{\mathbf{X}}_0^\top \tilde{\mathbf{X}}_0 + \mathbf{S}_0\right)^{-1} \tilde{\mathbf{X}}_0^\top \hat{\text{Cov}}(\mathbf{Y}_0\mathbf{A}_0)\mathbf{X}_0 \left(\mathbf{X}_0^\top \mathbf{X}_0\right)^{-1} \mathbf{e};\\
\left[\hat{\Sigma}_{\beta,\gamma}\right]_{22} &= \mathbf{e}^\top \left(\text{Var}\left(\hat{\boldsymbol{\gamma}}_1\right) + \text{Var}\left(\hat{\boldsymbol{\gamma}}_0\right)\right)\mathbf{e}.
\end{aligned}$$

We note that, for $j \in \{0,\ 1\}$ and n_j being the number of patients with $Z_i = j$

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\beta}}_j) &= \hat{\sigma}_j^2 \left(\tilde{\mathbf{X}}_j^\top \tilde{\mathbf{X}}_j + \mathbf{S}_j\right)^{-1},\\
\hat{\text{Cov}}(\mathbf{Y}_j, \mathbf{A}_j) &= \frac{(\mathbf{y}_j - \hat{\mathbf{y}}_j)^\top (\mathbf{a}_j - \hat{\mathbf{a}}_j)}{n_j},\\
\text{Var}\left(\hat{\boldsymbol{\gamma}}_j\right) &= \left(\mathbf{X}_j^\top \mathbf{X}_j\right)^{-1} \mathbf{X}_j^\top \text{Var}(\mathbf{A}_j)\\
\text{Var}(\mathbf{A}_j) &= \frac{(\mathbf{a}_j - \hat{\mathbf{a}}_j)^\top (\mathbf{a}_j - \hat{\mathbf{a}}_j)}{n_j - 2}.
\end{aligned}$$

Proof. Here, we derive the elements of the variance covariance matrix $\hat{\Sigma}_{\beta,\gamma}$.

1. $\left[\hat{\Sigma}_{\beta,\gamma}\right]_{11}$ is the variance of $\hat{\beta}$. It can be expressed as follows.

$$\begin{aligned}
\text{Var}(\hat{\beta}) =& \text{Var}\left(\tilde{\boldsymbol{x}}_{01}^\top \mathbf{B}_1 \hat{\boldsymbol{\beta}}_1 - \tilde{\boldsymbol{x}}_{00}^\top \mathbf{B}_0 \hat{\boldsymbol{\beta}}_0\right)\\
=& \tilde{\boldsymbol{x}}_{01}^\top \mathbf{B}_1 \text{Var}(\hat{\boldsymbol{\beta}}_1)\mathbf{B}_1^\top \tilde{\boldsymbol{x}}_{01} + \tilde{\boldsymbol{x}}_{00}^\top \mathbf{B}_0 \text{Var}(\hat{\boldsymbol{\beta}}_0)\mathbf{B}_0^\top \tilde{\boldsymbol{x}}_{00}.
\end{aligned} \tag{3.9}$$

   The variance of parameters from the thin plate regression spline models are given as:

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\beta}}_1) =& \hat{\sigma}_1^2 \left(\tilde{\mathbf{X}}_1^\top \tilde{\mathbf{X}}_1 + \mathbf{S}_1\right)^{-1},\\
\text{Var}(\hat{\boldsymbol{\beta}}_0) =& \hat{\sigma}_0^2 \left(\tilde{\mathbf{X}}_0^\top \tilde{\mathbf{X}}_0 + \mathbf{S}_0\right)^{-1},
\end{aligned}$$

   where $\hat{\sigma}_1$ and $\hat{\sigma}_0$ are estimates of the variance of the outcomes $\mathbf{y}_1$ and $\mathbf{y}_0$ respectively.

2. $\left[\hat{\Sigma}_{\beta,\gamma}\right]_{22}$ is the variance of $\hat{\gamma}$ and is expressed as:

$$\begin{aligned}\text{Var}(\hat{\gamma}) &= \text{Var}\left(\mathbf{e}^\top\hat{\boldsymbol{\gamma}}_1 - \mathbf{e}^\top\hat{\boldsymbol{\gamma}}_0\right)\\ &= \text{Var}\left(\mathbf{e}^\top\hat{\boldsymbol{\gamma}}_1\right) + \text{Var}\left(\mathbf{e}^\top\hat{\boldsymbol{\gamma}}_0\right)\\ &= \mathbf{e}^\top\left(\text{Var}\left(\hat{\boldsymbol{\gamma}}_1\right) + \text{Var}\left(\hat{\boldsymbol{\gamma}}_0\right)\right)\mathbf{e}\end{aligned} \tag{3.10}$$

The variance of the parameters from the linear model are given as:

$$\begin{aligned}\text{Var}\left(\hat{\boldsymbol{\gamma}}_1\right) &= \left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\mathbf{X}_1^\top\text{Var}(\mathbf{A}_1)\\ \text{Var}\left(\hat{\boldsymbol{\gamma}}_0\right) &= \left(\mathbf{X}_0^\top\mathbf{X}_0\right)^{-1}\mathbf{X}_0^\top\text{Var}(\mathbf{A}_0),\end{aligned}$$

where the variance of $\mathbf{A}$ is calculated as

$$\begin{aligned}\text{Var}(\mathbf{A}_1) &= \frac{\left(\mathbf{a}_1 - \hat{\mathbf{a}}_1\right)^\top\left(\mathbf{a}_1 - \hat{\mathbf{a}}_1\right)}{n_1 - 2}\\ \text{Var}(\mathbf{A}_0) &= \frac{\left(\mathbf{a}_0 - \hat{\mathbf{a}}_0\right)^\top\left(\mathbf{a}_0 - \hat{\mathbf{a}}_0\right)}{n_0 - 2}\end{aligned}$$

3. Now, we proceed to estimate of the covariance of β and γ which is equal to $\left[\hat{\Sigma}_{\beta,\gamma}\right]_{12}$.

$$\begin{aligned}\text{Cov}\left(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\gamma}}\right) =& \text{Cov}\left(\left\{\tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\hat{\boldsymbol{\beta}}_1 - \tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\hat{\boldsymbol{\beta}}_0\right\},\left\{\boldsymbol{e}^\top\hat{\boldsymbol{\gamma}}_1 - \boldsymbol{e}^\top\hat{\boldsymbol{\gamma}}_0\right\}\right)\\ =& \text{Cov}\left(\tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\hat{\boldsymbol{\beta}}_1, \boldsymbol{e}^\top\hat{\boldsymbol{\gamma}}_1\right) + \text{Cov}\left(\tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\hat{\boldsymbol{\beta}}_0, \boldsymbol{e}^\top\hat{\boldsymbol{\gamma}}_0\right)\\ =& \mathbb{E}\left[\tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\hat{\boldsymbol{\beta}}_1\cdot\hat{\boldsymbol{\gamma}}_1^\top\boldsymbol{e} - \mathbb{E}\left(\tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\hat{\boldsymbol{\beta}}_1\right)\mathbb{E}\left(\hat{\boldsymbol{\gamma}}_1^\top\boldsymbol{e}\right)\right] +\\ & \mathbb{E}\left[\tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\hat{\boldsymbol{\beta}}_0\cdot\hat{\boldsymbol{\gamma}}_0^\top\boldsymbol{e} - \mathbb{E}\left(\tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\hat{\boldsymbol{\beta}}_0\right)\mathbb{E}\left(\hat{\boldsymbol{\gamma}}_0^\top\boldsymbol{e}\right)\right]\\ =& \tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\mathbb{E}\left[\hat{\boldsymbol{\beta}}_1\hat{\boldsymbol{\gamma}}_1^\top - \mathbb{E}\left(\hat{\boldsymbol{\beta}}_1\right)\mathbb{E}\left(\hat{\boldsymbol{\gamma}}_1^\top\right)\right]\boldsymbol{e} +\\ & \tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\mathbb{E}\left[\hat{\boldsymbol{\beta}}_0\hat{\boldsymbol{\gamma}}_0^\top - \mathbb{E}\left(\hat{\boldsymbol{\beta}}_0\right)\mathbb{E}\left(\hat{\boldsymbol{\gamma}}_0^\top\right)\right]\boldsymbol{e}\\ =& \tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\left(\tilde{\mathbf{X}}_1^\top\tilde{\mathbf{X}}_1 + \mathbf{S}_1\right)^{-1}\tilde{\mathbf{X}}_1^\top\mathbb{E}\left[\mathbf{Y}_1\mathbf{A}_1^\top - \mathbb{E}\left(\mathbf{Y}_1\right)\left(\mathbf{A}_1^\top\right)\right]\\ & \left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\mathbf{X}_1^\top\boldsymbol{e} +\\ & \tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\left(\tilde{\mathbf{X}}_0^\top\tilde{\mathbf{X}}_0 + \mathbf{S}_0\right)^{-1}\tilde{\mathbf{X}}_0^\top\mathbb{E}\left[\mathbf{Y}_0\mathbf{A}_0^\top - \mathbb{E}\left(\mathbf{Y}_0\right)\left(\mathbf{A}_0^\top\right)\right]\\ & \left(\mathbf{X}_0^\top\mathbf{X}_0\right)^{-1}\mathbf{X}_0^\top\boldsymbol{e}\\ =& \tilde{\boldsymbol{x}}_{01}^\top\mathbf{B}_1\left(\tilde{\mathbf{X}}_1^\top\tilde{\mathbf{X}}_1 + \mathbf{S}_1\right)^{-1}\tilde{\mathbf{X}}_1^\top\text{Cov}\left(\mathbf{Y}_1,\mathbf{A}_1\right)\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\mathbf{X}_1^\top\boldsymbol{e} +\\ & \tilde{\boldsymbol{x}}_{00}^\top\mathbf{B}_0\left(\tilde{\mathbf{X}}_0^\top\tilde{\mathbf{X}}_0 + \mathbf{S}_0\right)^{-1}\tilde{\mathbf{X}}_0^\top\text{Cov}\left(\mathbf{Y}_0,\mathbf{A}_0\right)\left(\mathbf{X}_0^\top\mathbf{X}_0\right)^{-1}\mathbf{X}_0^\top\boldsymbol{e},\end{aligned} \tag{3.11}$$

where the covariance of $\mathbf{Y}$ and $\mathbf{A}$ is computed as follows

$$\begin{aligned}\text{Cov}(\mathbf{Y}_a, \mathbf{A}_b) =& \frac{(\mathbf{y}_1 - \hat{\mathbf{y}}_1)^\top (\mathbf{a}_1 - \hat{\mathbf{a}}_1)}{n_1}, \\ \text{Cov}(\mathbf{Y}_b, \mathbf{A}_b) =& \frac{(\mathbf{y}_0 - \hat{\mathbf{y}}_0)^\top (\mathbf{a}_0 - \hat{\mathbf{a}}_0)}{n_0}.\end{aligned}$$

□

So far, we have described two traditional methods and one novel approach to LATE estimation. In the next section, we shall compare these three methods using simulation studies.

3.5 Simulation studies

We carried out simulation studies to compare the performance of the three methods that we have described in Section 3.2, 3.3 and 3.4.2 for estimating the LATE at the threshold. Data were simulated to represent a fuzzy RD design with N patients, with different relationships between the outcome and assignment variable. We considered four scenarios with varying relationships (both linear and non-linear) between the outcome and assignment variable. Data simulation and analyses were performed using `R` (R Core Team, 2018). The robust biased-corrected estimate was computed using the `rdrobust` package in `R` (Calonico et al., 2020b) and thin plate regression spline models were fitted using the `gam` function in the `mgcv` package in R (Wood, 2019). First, we describe the data simulation.

3.5.1 Description of simulation study

The following steps are carried out to simulate the data for $N = 2500$ patients and the process was repeated $M = 2000$ times to create 2000 datasets with 2500 observations each.

Step 1: Simulate the assignment variable X_i, $i = 1, 2, ..., N$ from a Uniform (0.2, 0.8) distribution. The threshold x_0 is set at 0.5.

Step 2: Compute the threshold indicator $Z_i = \mathbb{1}(X_i \geq x_0)$.

Step 3: The probability receiving the treatment p_i is calculated from the equation below

$$\log\left(\frac{p_i}{1-p_i}\right) = \begin{cases} -0.13 + 4.66X_i & \text{if } Z_i = 1 \\ -3.72 + 4.66X_i & \text{if } Z_i = 0 \end{cases}$$

Here, the probabilities of receiving the treatment are 0.9 and 0.2 above and below the threshold respectively.

Step 4: The treatment indicator A_i is then simulated from a Bernoulli distribution as follows

$$A_i \sim \text{Bernoulli } (p_i)$$

Step 5: The outcome variable Y_i is simulated as follows

$$Y_i \sim \begin{cases} \text{Normal}(g_1(X_i), 0.2) & \text{if } A_i = 1 \\ \text{Normal}(g_0(X_i), 0.2) & \text{if } A_i = 0 \end{cases}$$

Where $g_0(X_i)$ and $g_1(X_i)$ represent the different simulation scenarios to depict the varying relationships between the outcome and assignment variable for untreated and treated patients respectively. These simulation scenarios are defined below.

Step 6: The steps listed above are repeated 2000 times to form 2000 datasets.

In the simulation studies, we considered four scenarios with varying relationships between the outcome and the assignment variable and they are presented below. Figure 3.3 is a visual representation of the underlying functional forms for the four scenarios considered. In addition, a sample of the simulated data under each scenario is presented in Figure 3.4. In all cases, the treatment effect at the threshold is set to -2.

Scenario 1: This scenario represents a situation where the underlying relationship between the outcome and assignment variable is linear above

and below the threshold.

$$g_0(X_i) = 4.8 + 0.4X_i$$
$$g_1(X_i) = 2.85 + 0.3X_i$$

Scenario 2: In this scenario, a non-linear relationship was simulated between X and Y below the threshold and with linearity above the threshold.

$$g_0(X_i) = 4.8 + 0.4X_i$$
$$g_1(X_i) = 3.58 - \left(\sin\{2\pi(X_i + 0.2)^3\}\right)^3$$

Scenario 3: This represents a linear relationship between Y and X below the threshold but the underlying relationship is non-linear above the threshold.

$$g_0(X_i) = 3.75 + 0.5X_i^2 + \frac{1.2}{1 + \exp\{24.75 - 55X_i\}}$$
$$g_1(X_i) = 2.85 + 0.3X_i$$

Scenario 4: This scenario has a non-linear relationship between X and Y above and below the threshold.

$$g_0(X_i) = 4 + 0.8X_i^2 + \frac{1}{1 + \exp\{33.6 - 70X_i\}}$$
$$g_1(X_i) = 3.58 - \left(\sin\{2\pi(X_i + 0.2)^3\}\right)^3$$

3.5.2 Assessment of the performance of methods

Here, we describe the measures that will be used to compare the three methods for estimating the LATE that we have discussed earlier. First, we define the following:

- $\text{Estimate}_i^{(m)}$ is the estimate of the LATE obtained when method m is applied to the i^{th} simulated sample.
- $\text{SE}_i^{(m)}$ is the standard error (square root of the variance) of the LATE estimate obtained when method m is applied to the i^{th} simulated sample.

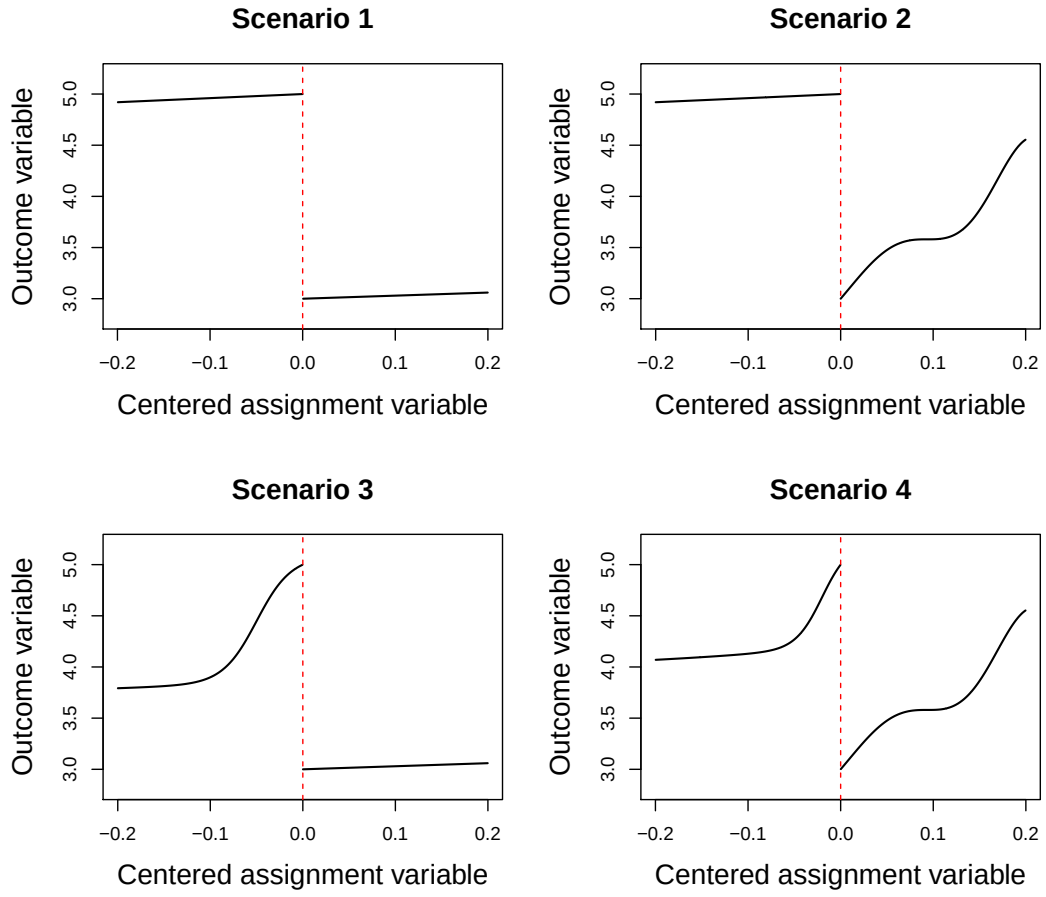

Figure 3.3: The functional forms of the relationship between the outcome and assignment variable that are used in the simulation studies in Section 3.5.

The measures that will be used to compare the methods and are reported in the tables of results are:

Estimate: We report the sample mean of the estimates obtained from the simulated samples. This gives information on how well the methods estimate the true value of the treatment effect on average.

$$\text{Estimate}^{(m)} = \frac{1}{M}\sum_{i=1}^{M} \text{Estimate}_i^{(m)}$$

Bias: The bias measures the deviation of the estimated value from the true value of the treatment effect. We desire to have an unbiased estimate, so an estimator with less bias is preferred.

$$\text{Bias}^{(m)} = -2 - \text{Estimate}^{(m)}$$

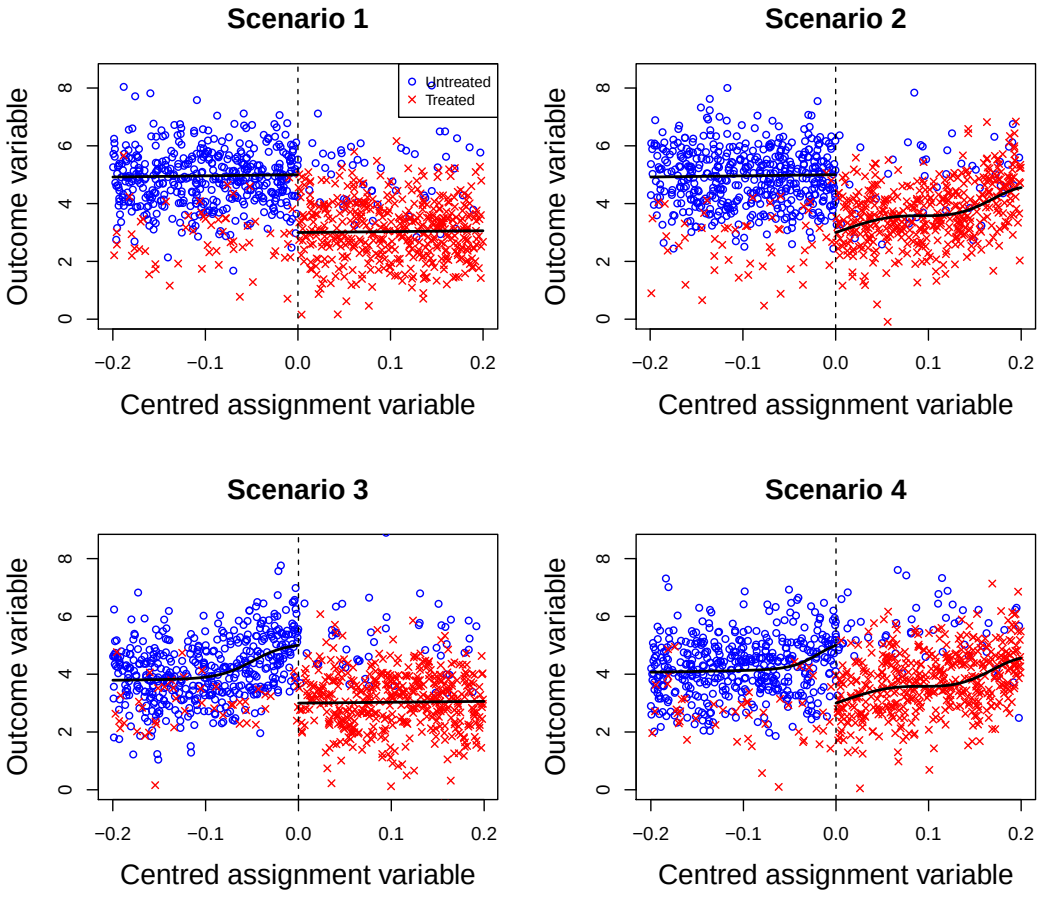

Figure 3.4: The functional forms (as well as a sample of the data) of the relationship between the outcome and assignment variable that are used in the simulation studies in Section 3.5.

Empirical standard error (ESE): The ESE is a measure of variation of the estimator. It gives an estimate of the standard deviation of the sampling distribution of the estimator.

$$\text{ESE}^{(m)} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}\left(\text{Estimate}_i^{(m)} - \text{Estimate}^{(m)}\right)^2}$$

Average standard error (ASE): This is also an estimate of the variation of the estimator. Ideally, we want the ASE to be close to the value of the ESE. This is because that the ESE approximates the true sampling variation of the estimate. Therefore, an ASE that is lower than the ESE indicates an underestimation of the standard error and a higher ASE means an overestimation of the standard error.

$$\text{ASE}^{(m)} = \frac{1}{M}\sum_{i=1}^{M}\text{SE}_i^{(m)}$$

95% Coverage: The coverage measures the proportion of times that the confidence intervals cover the true value of the treatment effect. In this case, we are constructing a 95% confidence interval, therefore, the coverage is expected to be approximately 95%. If the coverage is lower than the nominal level (95%), this might be as a result of a biased estimate or an underestimation of the standard error. A coverage that is above the nominal level is an indication of overestimation of the standard error. The 95% confidence interval and 95% coverage are defined as follows:

$$95\%\ \text{CI}_i^{(m)} = \text{Estimate}_i^{(m)} \pm 1.96\ \text{SE}_i^{(m)}$$

$$95\%\ \text{Cov}^{(m)} = \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}\left(-2 \in 95\%\ \text{CI}_i^{(m)}\right)$$

3.5.3 Results of simulation study

We have conducted the simulation studies as described above and a sensitivity analysis is carried out for the bandwidths. We considered five bandwidth sizes: 0.2, 0.15, 0.1 and 0.05, as well as the optimal bandwidth of the robust bias-corrected approach to check how the estimates vary across bandwidths. In general, an estimator that produces stable estimates across the bandwidths is preferable (Lee and Lemieux, 2010). In the case where an estimator is sensitive to the bandwidth, care needs to be taken to interpret the estimates.

Figure 3.5 shows a graphical representation of the estimates of the LATE for the four scenarios in the form of box plots. The horizontal red line represents the value of the true treatment effect. Table 3.1 is the table of the numerical summaries; estimate, bias, empirical and average standard errors and the 95% coverage of the LATE from the linear, robust bias-corrected and thin plate regression spline methods for Scenarios 1 and 2. Table 3.2 is the table of numerical results summaries of the results obtained for Scenarios 3 and 4. For the robust bias-corrected approach, we set the degrees of polynomial of the local kernel regressions for the point estimator and bias estimator to 2 and 3 respectively, that is, $p = 2$ and $q = 3$.

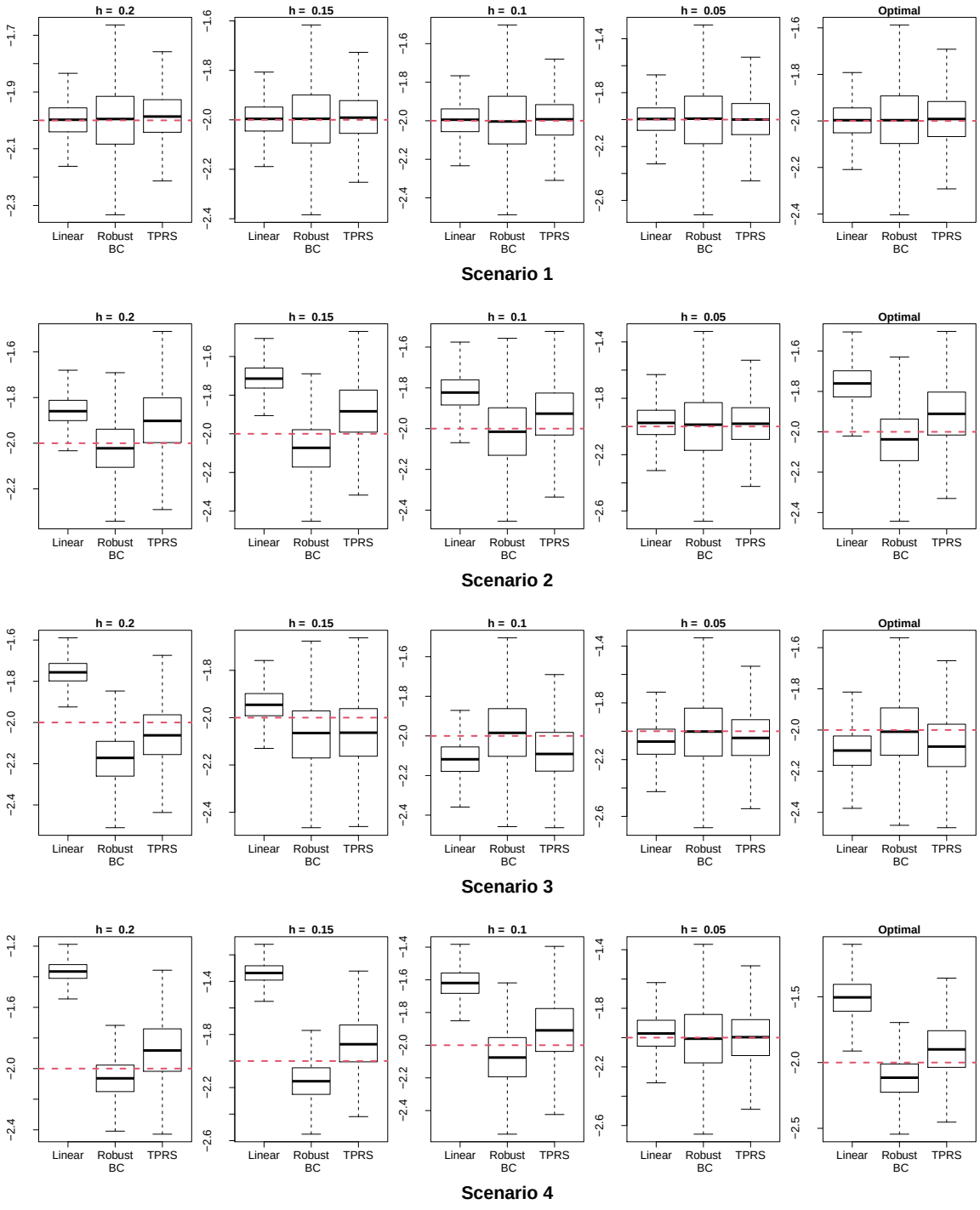

Figure 3.5: Boxplots of the estimates of LATE from the simulation study to compare the linear, robust bias-corrected (BC) and thin plate regression spline (TPRS) approaches. The red dashed line denotes the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 3.1: Estimates, biases, empirical and average standard errors (ESE and ASE) and 95% Coverage of the LATE using traditional (Linear), robust bias-corrected (BC) and Thin plate regression spline (TPRS) methods of estimation from 2000 repeated simulated samples for Scenarios 1 and 2.

	Scenario 1					*Scenario 2*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
					Bandwidth=0.2, Effect size = -2, Sample size = 2500					
Linear	-2.00	0.00	0.06	0.06	95.0	-1.86	-0.14	0.07	0.07	44.4
Robust BC	-2.00	0.00	0.13	0.13	94.7	-2.02	0.02	0.12	0.13	95.3
TPRS	-1.98	-0.02	0.10	0.10	94.5	-1.90	-0.10	0.15	0.14	87.1
					Bandwidth=0.15, Effect size = -2, Sample size = 2500					
Linear	-2.00	0.00	0.07	0.07	95.4	-1.71	-0.29	0.08	0.08	3.0
Robust BC	-2.00	0.00	0.15	0.15	95.5	-2.08	0.08	0.14	0.15	91.8
TPRS	-1.99	-0.01	0.12	0.11	95.2	-1.89	-0.11	0.16	0.14	78.6
					Bandwidth=0.1, Effect size = -2, Sample size = 2500					
Linear	-2.00	0.00	0.09	0.09	95.4	-1.82	-0.18	0.09	0.09	52.8
Robust BC	-2.00	0.00	0.19	0.18	95.4	-2.01	0.01	0.18	0.18	95.2
TPRS	-1.99	-0.01	0.15	0.14	94.8	-1.93	-0.07	0.16	0.16	88.3
					Bandwidth=0.05, Effect size = -2, Sample size = 2500					
Linear	-2.00	0.00	0.13	0.13	95.6	-1.97	-0.03	0.13	0.13	96.0
Robust BC	-2.00	0.00	0.27	0.26	94.8	-1.99	-0.01	0.26	0.26	95.6
TPRS	-2.00	0.00	0.21	0.20	95.4	-1.98	-0.02	0.21	0.20	96.5
					Optimal bandwidth, Effect size = -2, Sample size = 2500					
Linear	-2.00	0.00	0.08	0.08	95.6	-1.76	-0.24	0.10	0.09	25.8
Robust BC	-2.00	0.00	0.16	0.15	94.2	-2.04	0.04	0.16	0.16	94.8
TPRS	-1.99	-0.01	0.15	0.13	94.5	-1.92	-0.08	0.17	0.15	84.7

For Scenario 1, the underlying relationship between the outcome and assignment variable is linear, and we observe that the linear, robust bias-corrected (BC) and thin plate regression spline (TPRS) methods yield unbiased estimates of the treatment effect across the bandwidths, although the standard error estimates of the robust BC and TPRS methods are larger than that of the linear method. For the TPRS method, this is because there is extra uncertainty in determining the relationship between

the outcome and assignment variable. In contrast, with the linear estimator, the form of relationship between the outcome and assignment variable has already been pre-specified. We note that, the three methods provide adequate coverage of the treatment effect of around 95%. As expected, the standard error estimates increase as the bandwidth size reduces. This is as a result of a reduction in the number of observations included in the analysis for each bandwidth size as the bandwidth decreases.

Under Scenario 2, the relationship between the outcome and assignment variable is linear below the threshold and non-linear above the threshold. We see that for the larger bandwidths, the linear and TPRS methods appear to produce biased estimates of the treatment effect, although the bias of the TPRS method is less than that of the linear method. In addition, the TPRS method yields a better coverage of the true treatment effect than the linear method. However, the estimates obtained from the robust BC method are closer to the value of the treatment effect compared to other two methods. Similarly, the coverage of the robust BC method is closer to the nominal 95% for larger and optimal bandwidths. As observed under Scenario 1, the robust BC and TPRS methods also have higher standard error estimates when compared to the linear method.

Under Scenario 3, the relationship between the outcome and assignment variable is linear below the threshold and non-linear above the threshold, and we observe that, for bandwidth 0.2, the linear and robust BC methods produce biased estimates of the treatment effect and poor coverage of the treatment effect. In contrast, the TPRS method yields estimates that are close to the value of the true treatment effect for all bandwidths considered and provides good coverage of the treatment effect.

For Scenario 4, where the relationship between the outcome and assignment variable is non-linear above and below the threshold, the linear method only produced an unbiased estimator and adequate coverage of the treatment effect for bandwidth of 0.05. The estimates of the LATE from the robust BC and TPRS methods are closer to the value of the true treatment effect as well as better coverage of the treatment effect.

Table 3.2: Estimates, biases, empirical and average standard errors (ESE and ASE) and 95% Coverage of the LATE using traditional (Linear), robust bias-corrected and Thin plate regression spline (TPRS) methods of estimation from 2000 repeated simulated samples for Scenarios 3 and 4.

	Scenario 3					*Scenario 4*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
					Bandwidth=0.2, Effect size = -2, Sample size = 2500					
Linear	-1.76	-0.24	0.06	0.07	4.6	-1.37	-0.63	0.07	0.07	0.0
Robust BC	-2.17	0.17	0.13	0.13	73.1	-2.06	0.06	0.13	0.13	91.6
TPRS	-2.05	0.05	0.16	0.13	91.2	-1.88	-0.12	0.20	0.17	85.2
					Bandwidth=0.15, Effect size = -2, Sample size = 2500					
Linear	-1.95	-0.05	0.07	0.08	91.6	-1.34	-0.66	0.08	0.08	0.0
Robust BC	-2.07	0.07	0.15	0.15	92.7	-2.15	0.15	0.15	0.15	81.2
TPRS	-2.05	0.05	0.17	0.15	91.7	-1.87	-0.13	0.21	0.17	81.2
					Bandwidth=0.1, Effect size = -2, Sample size = 2500					
Linear	-2.12	0.12	0.09	0.10	81.1	-1.62	-0.38	0.09	0.09	1.5
Robust BC	-1.98	-0.02	0.18	0.18	94.7	-2.08	0.08	0.18	0.18	93.1
TPRS	-2.07	0.07	0.17	0.14	87.0	-1.91	-0.09	0.20	0.18	87.0
					Bandwidth=0.05, Effect size = -2, Sample size = 2500					
Linear	-2.07	0.07	0.13	0.13	94.0	-1.97	-0.03	0.13	0.13	95.3
Robust BC	-2.00	0.00	0.27	0.26	95.1	-2.01	0.01	0.26	0.26	95.2
TPRS	-2.03	0.03	0.23	0.20	94.8	-2.00	0.00	0.21	0.20	95.4
					Optimal bandwidth, Effect size = -2, Sample size = 2000					
Linear	-2.10	0.10	0.10	0.10	84.0	-1.51	-0.49	0.15	0.09	2.3
Robust BC	-2.01	0.01	0.17	0.16	93.4	-2.12	0.12	0.16	0.16	89.2
TPRS	-2.06	0.06	0.18	0.15	88.7	-1.90	-0.10	0.21	0.18	84.7

For Scenarios 2, 3 and 4 where there exists a non-linear relationship between the outcome and assignment variable, the linear method is able to produce unbiased estimates for a bandwidth of 0.05. This is expected because from Figures 3.3 and 3.4, we observe that for the smaller bandwidth the relationship between the outcome and assignment variable becomes approximately linear and hence the relationship can be modelled adequately by a linear model.

It has been noted that the coverage of the TPRS model might not be accurate for point-wise estimation compared to situations in which the coverage is calculated for the entire curve (Marra and Wood, 2012; Wood, 2020). A reason for this is because the estimated variance of the parameters of the TPRS model does not take into account the uncertainty of the smoothing parameter. Wood (2020) has provided a bootstrap approach to tackle this problem in order to improve the point-wise coverage of the TPRS approach. This method is readily available in the `mgcv` package in `R` and we implemented it to estimate the variance of the treatment effect estimate when fitting TPRS models.

Overall, we see that the robust BC and TPRS approaches outperform the linear approach in that estimates are less biased and provide better coverage of the true treatment effect. The robust BC and TPRS approaches seem to be comparable, except in scenario 2, in terms of producing estimates that are close to the treatment effect. We note that for the robust BC approach, we have to specify the degrees of polynomial of the local regressions models for the point estimator and bias estimator. On the other hand, the TPRS approach does not require such specification as it is data driven, removing this choice from the user, which may be attractive to clinicians.

We have now compared three methods of estimating the LATE in simulation studies. In the next section, we shall apply the three methods to a real dataset on the effect of statin prescription on low density lipoprotein cholesterol level.

3.6 Example on Prescription of Statins in UK Primary Care

The United Kingdom's National Institute for Health and Care Excellence (NICE) set a guideline that statins - a class of cholesterol lowering drugs - be prescribed to adults aged under 75 years whose risk of developing a cardiovascular event is greater than 20% (NICE, 2008) (the guideline was later revised in July 2014 from 20% to 10%). The 10-year risk score is calculated by a general practitioner (GP) using risk prediction algorithms, such as the Framingham risk score (Wilson et al.,

1998) or Q-RISK score (Hippisley-Cox et al., 2008). The data considered in this example were collected between 2007 and 2014 when the NICE guideline was set at 20%, therefore, the assignment variable is the 10-year risk score and threshold is at 20%. The outcome variable is the low lipodensity (LDL) cholesterol level, sometimes known as 'bad cholesterol' because a high LDL cholesterol level can increase the chance of having heart problems or a stroke (Duncan et al., 2019; NHS, 2019a) and the treatment is the prescription of statins.

We aim to estimate the effect statin prescription on LDL cholesterol levels using data extracted from The Health Improvement Network (THIN) database, a large source of UK primary care data. The database contains anonymised information of patient records collected from over 500 British GP practices and is generally representative of UK population (Bourke et al., 2004; Blak et al., 2011). We extracted data for 1386 male patients aged between 50 and 70 years who are non-diabetic, non-smokers and had never experienced a cardiovascular event (stroke or myocardial infarction), for whom a 10-year CVD risk score was calculated by a GP at some point between January 2007 and December 2008. We obtained the following information: patient risk of developing CVD in 10 years (X), treatment status (A), initial low density lipoprotein (LDL) cholesterol level, at the time the risk score was calculated, and the outcome was measured about 6 weeks after statin prescription for patients that receive statin and 6 weeks after the first measurement of LDL cholesterol level for untreated patients.

Of the 1386 patients, 831 (60%) patients had risk score values above the threshold while 555 (40%) patients had risk score values below the threshold. Of those patients with risk score values above the threshold, 605 (73%) received a statin prescription whereas, of those patients with risk score values below the threshold, 100 (18%) received a statin prescription.

Figure 3.6 (a) shows a plot of the probability of receiving a statin prescription against the 10-year risk of developing a CVD event. The dashed vertical line is represents the threshold value. This figure suggests that patients above the threshold have a higher probability of getting a statin prescription compared to patients below the threshold and that a discontinuity may exist in the probability of getting a statin prescription at threshold. This indicates that the use an RD design appears to be

suitable for this data.

Figure 3.6 (b) shows a scatter plot of the outcome (LDL cholesterol level) against the 10-year risk of developing a CVD event. The red and blue dots represent patients that received and did not receive a statin prescription respectively. This plots shows some discontinuity in the distribution of the blue and red dots. That is, it seems that, on average, patients that did not receive a statin prescription have a higher LDL cholesterol level compared to patients that received a statin prescription.

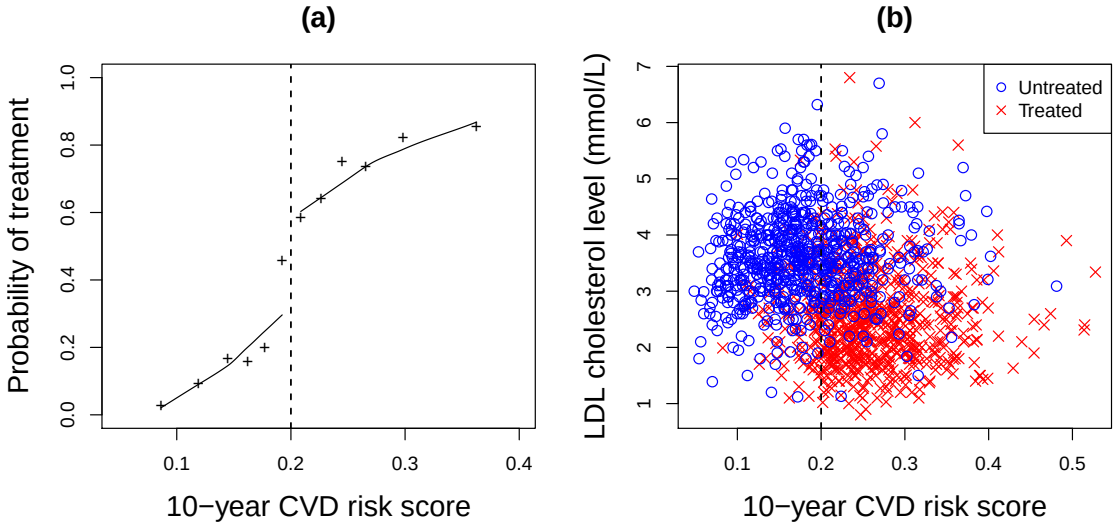

Figure 3.6: (a) Plot of probabilities of getting treated for patients above and below the threshold, (b) the scatter plot of the LDL cholesterol level and 10-year risk score. The black crosses are the expected probabilities calculated in bins.

We now proceed to apply the three methods; linear, robust BC and thin plate regression spline, to the data to estimate the effect of statin prescription on the reduction of LDL cholesterol level. For all methods, we adjusted for the initial LDL cholesterol level, by including it in the regression models for the three approaches:

We define R_i to be the initial LDL cholesterol level for patient i, for the linear approach, the outcome and treatment models are given as

$$\begin{aligned}\mathbb{E}(Y_i|Z_i = z_i, X_i = x_i, R_i = r_i) &= \beta_0 + \beta_1 z_i + \beta_2 x_i + \beta_3 z_i x_i + \beta_4 r_i, \\ \mathbb{E}(A_i|Z_i = z_i, X_i = x_i, R_i = r_i) &= \gamma_0 + \gamma_1 z_i + \gamma_2 x_i + \gamma_3 z_i x_i + \gamma_4 r_i.\end{aligned} \tag{3.12}$$

So that the LATE is estimated as

$$\frac{\hat{\beta}_1}{\hat{\gamma}_1}.$$

For the robust BC approach, local polynomial regression models are fitted for the outcome and treatment variables with r_i included in the models similar to Equation 3.12. Details of this approach are given by Calonico et al. (2019).

For the TPRS approach, the outcome model is given as

$$y_i = f(\boldsymbol{x}_i) + \beta z_i + e_i,$$

where, β measures the discontinuity in the expected value of the outcome at the threshold and therefore, it represents the numerator of the LATE, $\boldsymbol{x}_i = (x_i, r_i)^\top$ and

$$f(\boldsymbol{x}) = \sum_{j=1}^{n} \delta_j \eta(|\boldsymbol{x} - \boldsymbol{x}_j|) + \sum_{k=1}^{M} \alpha_j \phi_j(\boldsymbol{x}),$$

where $\eta(.)$ and $\phi(.)$ are as defined in Section 3.4.2. For the denominator of the LATE, we use the treatment model in Equation 3.12. A similar approach can be used to adjust for other covariates in the models when necessary.

We calculated the treatment effect within five bandwidths; 0.05, 0.1, 0.15, 0.2 and the optimal bandwidth for the robust BC method. We set the degrees of the polynomials of the local regression models for the point estimator and bias estimator of the robust BC approach to 2 and 3, respectively. Table 3.3 shows the estimates of the LATE with corresponding standard errors obtained using these three methods.

From Table 3.3, it is observed that the treatment estimates obtained from the three methods are negative. This implies that the LDL cholesterol level of the patients that receive a statin prescription is lower than that of patients who did not receive statins, perhaps suggesting a beneficial effect of statins in reducing LDL cholesterol level. We observe that the estimates of the treatment effect obtained from the linear and thin plate regression spline methods are similar to each other, perhaps, because the relationship between the risk score and the LDL cholesterol level is linear. The estimates from the robust BC approach are noted to be higher and very unstable than the other two approaches across the bandwidths. More notable is the fact that the standard error estimates of the robust BC approach are quite large, compared to the other approaches, which makes the estimates from the robust BC approach imprecise. The large standard error estimates from the robust BC approach are perhaps due to a bias-variance trade-off.

Table 3.3: Estimates and associated standard errors for the LATE at the threshold for the THIN data example on the prescription of statins based on 10-year CVD risk score.

Bandwidth:	*0.05*		*0.1*		*0.15*		*Optimal*	
N:	676		1103		1309		792	
Method	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Linear	-0.89	0.85	-1.32	0.34	-1.44	0.25	-1.15	0.58
Robust BC	-1.12	9.73	-1.67	1.83	-1.32	1.04	-2.19	11.10
TPRS	-1.04	0.83	-1.30	0.34	-1.42	0.25	-1.30	0.58

We note that the treatment effect estimates vary for different bandwidths sizes. Typically, if there is a treatment effect, we will expect that all estimates of the treatment effect across bandwidths will be in the same direction. In this example, all the estimates across bandwidths suggest a beneficial effect of statin prescription in lowering LDL cholesterol level. But since it has been established that the smaller the bandwidth, the lower the bias, we will usually prefer the estimate from the smallest bandwidth provided the sample size for such bandwidth is not too small.

Figure 3.7 is the plot of the fitted models for the linear, TPRS and triangular regression models for the numerator of the LATE. The robust BC approach entails fitting a triangular kernel model, then the LATE estimate from the triangular model is bias-corrected. Figure 3.8 is the plot of the fitted models for the linear and triangular regression models for the denominator of the LATE. We used a linear model to estimate the denominator of the TPRS approach. We observe that the fitted models for the three approaches are really similar to each other, this explains why the linear and TPRS estimates of the treatment effect are similar. The estimate from the robust BC approach is quite different from the other two because, although the fitted regression lines for the three methods are similar, an additional bias correction is carried out for the robust BC approach.

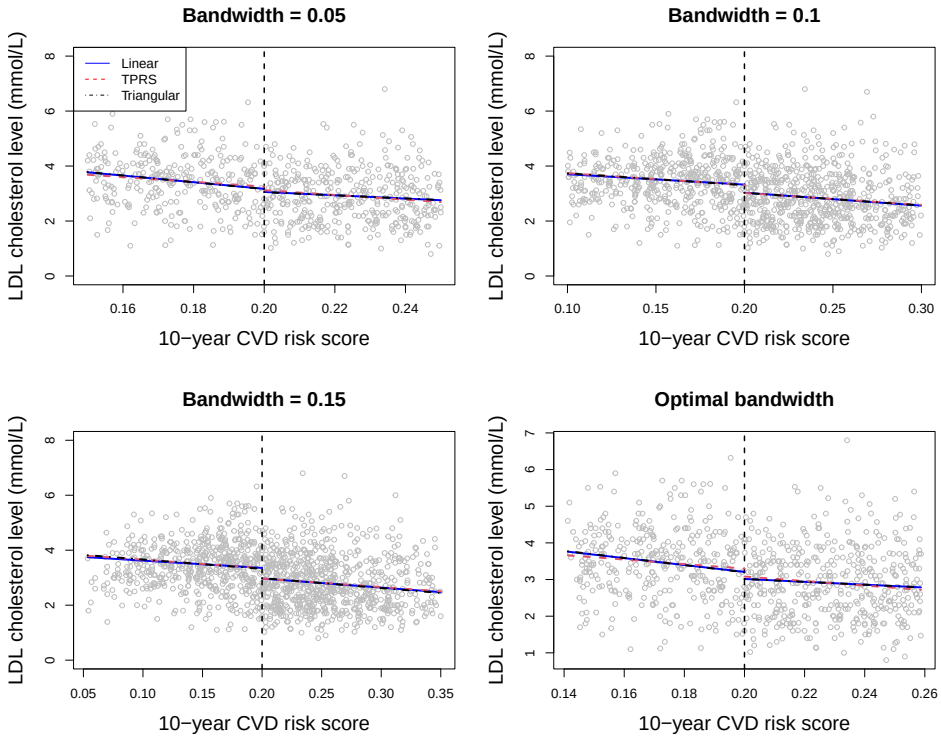

Figure 3.7: Fitted lines for linear, TRPS and triangular regression models for estimating the numerator of the LATE.

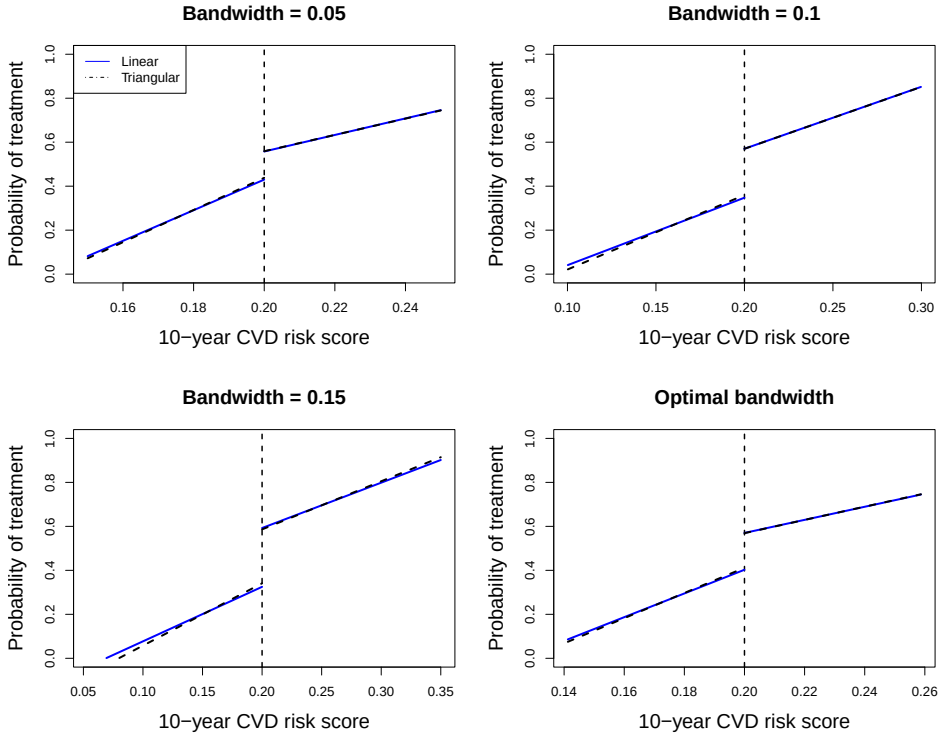

Figure 3.8: Fitted lines for linear and triangular regression models for estimating the denominator of the LATE.

3.7 Conclusions

In this chapter, we have discussed three methods for estimating the local average treatment estimate (LATE), an estimator of the treatment effect in a fuzzy RD design for a continuous outcome. We discussed a traditional approach which involves fitting linear models to data above and below the threshold. A more sophisticated approach, known as the robust bias-corrected (BC) approach, was also discussed. The robust BC approach involves fitting local polynomial regression models below and above the threshold and then corrects for the bias of the local polynomial regression models. In addition, we proposed fitting thin plate regression spline (TPRS) models below and above the threshold as an alternative. The TPRS method is a data-driven regression approach to fit flexible regression models that does not require any strong assumption about the underlying relationship between the response and predictor variables.

We carried out simulation studies with varying relationships between the outcome and assignment variable to compare the performance of the three methods. The methods produced estimates close to the treatment effect when the underlying relationship between the outcome and the assignment variable is linear. For scenarios where the simulated underlying relationship is not linear, the linear method produced biased estimates of the treatment effect. The thin plate regression spline and robust BC methods yielded unbiased estimates of the treatment effect as we would expect. For scenarios 3 and 4, where the underlying relationships between the outcome and the assignment variable are non-linear, the robust BC and TPRS approaches produced estimates that are closer to the value of the true treatment effect.

The three methods were applied to a real data on the prescription of statins in UK primary care. The estimates obtained from the methods suggest a beneficial effect of statin prescription for reducing LDL cholesterol level.

In this chapter, we described methods for the treatment effect estimation for a continuous outcome. However, binary outcomes are also of importance in medical studies. In the next chapter, we shall discuss methods to the treatment effect estimation in an RD design when the outcome of interest is binary.

Chapter 4

Regression discontinuity designs where the outcome is binary

In Chapter 3, we discussed treatment effect estimation in a fuzzy RD design when the outcome of interest is a continuous variable. In practice, situations often arise where the outcome of interest is a binary variable. For instance, whether or not death occurred after surgery, relapse of a disease after a given period of time, relief of pain after taking a medicine are all examples of possible binary outcomes of interest. Here, the binary outcome takes value 1 when a patient experiences an event of interest during the study and it takes value 0 otherwise. In this chapter, we will consider methods for an RD design where the outcome is binary and, in particular, we will focus on risk ratio estimation. We explore methods for risk ratio estimation in the literature, and a new approach that is based on the RD design assumptions will be given.

4.1 Introduction

Modelling binary outcomes usually focuses on estimating the risk or odds of an event of interest occurring. Suppose Y_i is a binary outcome for patient i which is equal to 1 if patient i experiences an event of interest and $Y_i = 0$ if patient i did not experience the event of interest. The risk is defined as the probability that the event of interest will occur, that is,

$$\text{Risk} = \mathbb{P}(Y_i = 1).$$

The odds of the event of interest is defined as a ratio of the probability of the event occurring and the probability of the event not occurring:

$$\text{Odds} = \frac{\mathbb{P}(Y_i = 1)}{\mathbb{P}(Y_i = 0)}.$$

In studies where interest lies in estimating a treatment effect (with the treatment indicator denoted as A_i) where the outcome if binary, the treatment effect is often measured by the risk difference, risk ratio or odds ratio; defined as follows:

$$\begin{aligned}
\text{Risk difference } &= \mathbb{P}(Y_i = 1|A_i = 1) - \mathbb{P}(Y_i = 1|A_i = 0), \\
\text{Risk ratio } &= \frac{\mathbb{P}(Y_i = 1|A_i = 1)}{\mathbb{P}(Y_i = 1|A_i = 0)} \text{ or} \\
\text{Odds ratio } &= \frac{\mathbb{P}(Y_i = 1|A_i = 1)}{\mathbb{P}(Y_i = 0|A_i = 1)} \frac{\mathbb{P}(Y_i = 0|A_i = 0)}{\mathbb{P}(Y_i = 1|A_i = 0)}
\end{aligned}$$

In this chapter, we focus on treatment effect estimation in an RD design where the outcome of interest is binary. Xu (2017) looked at estimating risk differences in a sharp RD design for a categorical outcome variable. They proposed the use of multinomial logistic regression models for patients whose assignment variable values lie above and below the threshold. In the case where the outcome is binary, logistic regression models were fitted and the risk difference to compare those above and below the threshold were estimated. The method proposed in this study is applicable for a sharp RD design and will be analogous to an intention to treat analysis for a fuzzy RD design. To our knowledge, the only instance where an estimator of the treatment effect was proposed for a binary outcome in a fuzzy RD design is in Geneletti et al. (2019). They proposed a Bayesian approach of the multiplicative structural mean model (MSMM) to estimate the risk ratio in a fuzzy RD design. The MSMM is a popular approach for estimating the risk ratio in the instrumental variable framework (Didelez et al., 2010). The MSMM approach requires additional assumptions for the identification of the treatment effect in an RD design. In the next section, we shall discuss some estimators of the risk ratio, including the MSMM, that can be applied in a fuzzy RD design. In addition, we will present a new approach to estimating the risk ratio that is based on the assumptions of the RD design. In this chapter, we focus on estimation of risk ratio, though, in many epidemiology applications, odds ratio is more popular because it can be estimated directly from a

logistic regression model. However, the estimators of causal odds ratio are generally more biased than estimators of causal risk ratios (Palmer et al., 2011; Geneletti et al., 2019). Consequently, we stick to estimating the risk ratio.

4.2 Estimators of the risk ratio

As we have noted earlier, there is limited research on methods to estimate the risk ratio in a fuzzy RD design. As such, we explore methods applicable in the instrumental variable framework. The instrumental variable framework provides methodology for the estimation of treatment effect in an observational study, which relies on the existence of variable(s) called the instrument(s) (Bowden and Turkington, 1990; Angrist et al., 1996; Wooldridge, 2010). An instrument is a variable that is related to the treatment but is not affected by the treatment. Formally, a variable is an instrument if it satisfies the following conditions:

- It is not affected by the treatment but has a causal relationship with the treatment.
- It does not have a direct effect on the outcome.

A common application of the instrumental variable literature in epidemiology is in Mendelian randomisation, where some genetic markers are used as instruments. For example, Larsson et al. (2020), carried out a study to understand the relationship between alcohol consumption and cardiovascular disease using Mendelian randomisation. Some genetic variants have been identified that are associated with alcohol consumption. These genetic variants were consequently used as instruments to identify the effect of alcohol consumption on the risk of experiencing a cardiovascular disease.

The genetic markers can have an effect on the treatment (alcohol consumption) and not the other way around, and it is believed that the presence (or absence) of these genetic variants does not influence the outcome (risk of experiencing a cardiovascular disease). Hence, the genetic markers represent example of instruments.

Now we compare the definition of an instrument in the instrumental variable framework to the threshold indicator in a fuzzy RD design. In the RD design, we stated that the threshold indicator can be seen as a randomising device for treatment allocation. This implies that the treatment is a function of the threshold indicator, also implied by Assumptions 1 and 2 of the RD design stated on Section 2.2. Therefore, the threshold indicator satisfies the first assumption of an instrument stated above.

Assumption 5 of the RD design states that the threshold indicator is independent of the outcome conditional on other variables. This implies that the threshold indicator does not have a direct casual relationship with the outcome and is in line with the second property of an instrument stated above.

As a result, the threshold indicator in an RD design is similar to an instrument in the instrumental variable framework (Geneletti et al., 2015). Therefore, for the estimation of the treatment effect in an RD design, we shall consider some existing methods from the instrumental variable framework and attempt to apply these to an RD design.

Next, we will describe some methods that can be used for estimating the risk ratio in a fuzzy RD design.

4.2.1 Wald estimator

The first method we will consider for the estimation of the risk ratio in an RD design is the Wald approach, also known as the Wald IV approach. This was introduced as an adaptation of the Wald estimator for continuous outcomes (Didelez et al., 2010; Palmer et al., 2011). The WALD estimator of the risk ratio, on the log-scale, is given as:

$$\log(\text{WALD}) = \frac{\log \mathbb{E}(Y_i|Z_i = 1) - \log \mathbb{E}(Y_i|Z_i = 0)}{\mathbb{E}(A_i|Z_i = 1) - \mathbb{E}(A_i|Z_i = 0)}.$$

This expresses the log of risk ratio in the form of the LATE estimator. That is, the numerator is estimated as the difference of the log of the expected value of the outcome between the patients above and below the threshold, and the denominator remains the difference in the probability of receiving treatment for patients above

and below the threshold. Didelez et al. (2010) has shown that this estimator is consistent for the risk ratio when the causal effect is small and the event is rare. Therefore, in cases where the effect of the treatment is large or the event of interest is common in the study population, the WALD estimator might be inappropriate.

In addition, the following modelling assumptions are required for this estimator to identify the causal risk ratio (Didelez et al., 2010; Palmer et al., 2011):

W1 $\log \mathbb{E}(Y_i|A_i = a)$ is linear in a.

W2 $\mathbb{E}(A_i|Z_i = z)$ is linear in z.

Owing to assumption W1 above, a log-linear model is recommended for estimating the numerator of the log(WALD-RR) estimator. However, because Y_i is a binary variable, it might be more natural to fit a logistic regression model because this ensures that the estimated expected values of Y_i will always lie between 0 and 1.

Remark 4.2.1. *For a binary treatment, the logit link in a logistic regression model satisfies assumption W1*

Proof. Assumption W1 states that

$$\log \mathbb{E}(Y_i|A_i = a) \text{ is linear in } a.$$

Fitting a logistic regression model implies that

$$\log \mathbb{E}(Y_i|A_i = a) - \log(1 - \mathbb{E}(Y_i|A_i = a)) \text{ is linear in } a.$$

That is,

$$\log \mathbb{E}(Y_i|A_i = a) - \log(1 - \mathbb{E}(Y_i|A_i = a) = \alpha_0 + \alpha_1 a.$$

Now, we want to show that $\log \mathbb{E}(Y_i|A_i = a)$ in the equation above can be expressed as a linear function of a. From the equation above, we have that,

$$\mathbb{E}(Y_i|A_i = a) = \frac{\exp(\alpha_0 + \alpha_1 a)}{1 + \exp(\alpha_0 + \alpha_1 a)},$$

$$\implies \log \mathbb{E}(Y_i|A_i = a) = \alpha_0 + \alpha_1 a - \log(1 + \exp(\alpha_0 + \alpha_1 a)).$$

For $a \in \{0,\ 1\}$, we have that

$$\mathbb{E}(Y_i|A_i = a) = \alpha_0^* + \alpha_1^* a,$$

where $\alpha_0^* = \alpha_0 - \log(1 + \exp(\alpha_0))$ and $\alpha_1^* = \alpha_1 - \log(\frac{1+\exp(\alpha_0+\alpha_1)}{1+\exp(\alpha_0)})$ □

In remark 4.2.1, we show that, for a binary treatment, fitting a logistic regression model satisfies Assumption W1. As such, to estimate the numerator of log(WALD-RR) we fit a logistic regression model for Y_i. In an RD design, since we are interested in the treatment effect at the threshold, the Wald estimator of the risk ratio at the threshold is given as:

$$\log(\text{WALD-RR}) = \frac{\lim_{x \downarrow x_0} \log \mathbb{E}(Y_i|X_i = x) - \lim_{x \uparrow x_0} \log \mathbb{E}(Y_i|X_i = x)}{\mathbb{E}(A_i|Z_i = 1) - \mathbb{E}(A_i|Z_i = 0)}.$$

4.2.2 Multiplicative Structural Mean Model

Secondly, we shall discuss the multiplicative structural mean model (MSMM). The MSMM is commonly used to estimate the risk ratio in the instrumental variables framework (Hernán and Robins, 2006; Clarke and Windmeijer, 2010; Didelez et al., 2010) and has been used in an RD design paper (Geneletti et al., 2019). The MSMM is defined in terms of counterfactual outcomes. As described in Section 2.3, counterfactual outcomes, also called potential outcomes, are a set of outcomes that represent what would have been observed depending on the treatment that a patient receives. In the case of a binary outcome, the counterfactual outcomes will also be binary. We define Y_i^1 and Y_i^0 as the counterfactual outcomes of having been treated and not having received the treatment, respectively, for patient i.

The multiplicative structural mean model (MSMM) compares the log of the expectation of the counterfactual outcomes for the treated group and is given in Equation 4.1

$$\log \mathbb{E}(Y_i^1|Z_i, A_i) - \log \mathbb{E}(Y_i^0|A_i = a, Z_i = z) = \psi_0 a + \psi_1 z \tag{4.1}$$

Using an MSMM, the risk ratio for the treated can be estimated consistently subject to the following additional assumptions (Hernán and Robins, 2006; Didelez et al.,

2010; Geneletti et al., 2019):

M1 There is no interaction between the threshold indicator and treatment on the multiplicative level (no effect modification).

M2 $\log \mathbb{E}(Y_i^1 | A_i = a, Z_i = z) - \log \mathbb{E}(Y_i^0 | A_i = a, Z_i = z)$ is linear in a.

From Equation 4.1, $\exp(\psi_0)$ represents the risk ratio for the treated (RRT) for patients below the threshold and $\exp(\psi_0 + \psi_1)$ is the RRT for patients above the threshold. Assumption M1 implies that the threshold does not modify the relationship between the outcome and treatment, that is the treatment effect is not different for patients above and below the threshold. This is in line with Assumption 5 of the RD design that the threshold indicator is independent of the outcome conditional on the treatment. Therefore, under Assumption M1, $\psi_1 = 0$ and, as such, the treatment effect of interest, RRT, is $\exp(\psi_0)$.

The relationship between the RRT and the overall risk ratio, which we will refer to as the causal risk ratio (CRR) is given below

$$\begin{aligned} \frac{\mathbb{E}(Y^1|A=1)}{\mathbb{E}(Y^0|A=1)} &= \frac{\mathbb{E}(Y^1|A=1)}{\mathbb{E}(Y^0|A=0)} \frac{\mathbb{E}(Y^0|A=0)}{\mathbb{E}(Y^0|A=1)} \\ \text{RRT} &= \text{CRR} \times \text{SB}. \end{aligned}$$

SB represents the selection bias, where selection bias refers to when the expectation of the counterfactual outcome for treated patients ($\mathbb{E}(Y^0|A=1)$) is not equal to the expectation of the observed outcome of untreated patients ($\mathbb{E}(Y^0|A=0)$). This occurs when there is no randomisation in treatment allocation and the treatment groups are not exchangeable.

Therefore, the risk ratio for the treated includes both the causal risk ratio and the effect due to selection bias. If there is no selection bias, the risk ratio for the treated and causal risk ratio are equal. In an RD design, for patients that are close to the threshold, we expect that patients above and below the threshold are similar in terms of potential confounders and exchangeable, in which case, we should expect that there is no selection bias.

The analytic expression of the RRT derived from MSMM was provided by Hernán

and Robins (2006) and is presented below:

$$\text{RRT} = 1 - \frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\mathbb{E}(Y_i(1 - A_i)|Z_i = 1) - \mathbb{E}(Y_i(1 - A_i)|Z_i = 0)}. \tag{4.2}$$

In practice, because of Assumption M2, generalised linear models with a log link function are generally fitted to estimate the components of Equation 4.2. However, since the outcome is binary, using a logit link function might be a more natural approach because it ensures that estimates of probabilities are strictly between 0 and 1. It has been shown that fitting a logistic regression model leads to an estimate that is similar to that obtained when fitting a log-linear model (Geneletti et al., 2019).

Remark 4.2.2. *For a binary treatment, the logit link in a logistic regression model satisfies Assumption M2.*

Proof. Assumption M2 implies that

$\log \mathbb{E}(Y_i^1|A_i = a, Z_i = z)$ and $\log \mathbb{E}(Y_i^0|A_i = a, Z_i = z)$ are both linear in a.

However, by fitting a logistic regression model, we have that

$\log \mathbb{E}(Y_i^1|A_i = a, Z_i = z) - \log(1 - \mathbb{E}(Y_i^1|A_i = a, Z_i = z))$ and

$\log \mathbb{E}(Y_i^0|A_i = a, Z_i = z) - \log(1 - \mathbb{E}(Y_i^0|A_i = a, Z_i = z))$ are linear in a.

That is,

$$\log \mathbb{E}(Y_i^1|A_i = a, Z_i = z)) - \log(1 - \mathbb{E}(Y_i^1|A_i = a, Z_i = z)) = \alpha_0 + \alpha_1 a, \text{ and}$$
$$\log \mathbb{E}(Y_i^0|A_i = a, Z_i = z)) - \log(1 - \mathbb{E}(Y_i^0|A_i = a, Z_i = z)) = \beta_0 + \beta_1 a.$$

The task here is to check that $\log \mathbb{E}(Y_i^1|A_i = a, Z_i = z))$ and $\log \mathbb{E}(Y_i^0|A_i = a, Z_i = z))$ in the equations above can be expressed as linear function of a. We have that

$$\mathbb{E}(Y_i^1|A_i = a, Z_i = z) = \frac{\exp(\alpha_0 + \alpha_1 a)}{1 + \exp(\alpha + \alpha_1 a)} \text{ and}$$
$$\mathbb{E}(Y_i^0|A_i = a, Z_i = z) = \frac{\exp(\beta_0 + \beta_1 a)}{1 + \exp(\beta_0 + \beta_1 a)}.$$

Therefore,

$$\log \mathbb{E}(Y_i^1|A_i = a, Z_i = z)) = \alpha_0 + \alpha_1 a - \log(1 + \exp(\alpha_0 + \alpha_1 a)) \text{ and}$$
$$\log \mathbb{E}(Y_i^0|A_i = a, Z_i = z)) = \beta_0 + \beta_1 a - \log(1 + \exp(\beta_0 + \beta_1 a))$$

This expressions above are non linear in a. However, for $a \in \{0, 1\}$, they can be re-written in form of a linear function of a as follows:

$$\log \mathbb{E}(Y_i^1|A_i = a, Z_i = z)) = \alpha_0^* + \alpha_1^* a \text{ and}$$
$$\log \mathbb{E}(Y_i^0|A_i = a, Z_i = z)) = \beta_0^* + \beta_1^* a$$

Where $\alpha_0^* = \alpha_0 - \log(1 + \exp(\alpha_0))$, $\alpha_1^* = \alpha_1 - \log\left(\frac{1+\exp(\alpha_0+\alpha_1)}{1+\exp(\alpha_0)}\right)$,
$\beta_0^* = \beta_0 - \log(1 + \exp(\beta_0))$ and $\beta_1^* = \beta_1 - \log\left(\frac{1+\exp(\beta_0+\beta_1)}{1+\exp(\beta_0)}\right)$. □

Therefore, owing to Remark 4.2.2, for binary treatment, the logistic model satisfies Assumption M2 of the MSMM.

A downside of the MSMM estimator is that the estimate could be negative, which is counter-intuitive since a risk ratio should be strictly positive. Geneletti et al. (2019) proposed a Bayesian method of estimating RRT by specifying a Gamma distribution as the prior of RRT. Since the support of the Gamma distribution is non-negative, this ensures that the estimate of the RRT will be non-negative. However, this approach cannot be replicated in the non-Bayesian setting that we consider in this chapter.

So far we have described two existing methods that can be used to estimate the risk ratio in an RD design. We shall now describe a new method for estimating the risk ratio based on the RD design assumptions that we stated in Chapter 2.

4.2.3 RD design method

In this section, we shall describe a new approach to estimate the risk ratio in a fuzzy RD design. The Wald estimator has a desirable property that its estimate of the risk ratio is always positive. However, the Wald estimator is not consistent when the value of the treatment effect is large or the probability of the event of interest

occurring is not small. The MSMM estimate requires additional assumptions in order to estimate the risk ratio. Here, we propose a new alternative for risk ratio estimation in an RD design based on the RD design assumptions stated in Chapter 2. The treatment effect of interest is the risk ratio at the threshold and is given below:

$$\lim_{x \to x_0} \frac{\mathbb{E}(Y_i|A_i = 1, X_i = x)}{\mathbb{E}(Y_i|A_i = 0, X_i = x)} \tag{4.3}$$

For a fuzzy RD design, the components (numerator and denominator) of the expression in Equation 4.3 cannot be estimated directly (e.g from a logistic regression model) owing to the potential effect of unobserved confounders, which might lead to a biased estimate of the treatment effect. Rather, we will exploit the fact that we have information about treatment allocation, that is based on the value of an assignment variable and a pre-determined threshold, to derive an estimator of the risk ratio in a fuzzy RD design. Using the assumptions of an RD design, we derive the following estimator of the risk ratio:

RDD-RR =

$$1 - \lim_{x \to x_0} \frac{\mathbb{E}(Y_i|Z_i = 1, X_i = x) - \mathbb{E}(Y_i|Z_i = 0, X_i = x)}{\mathbb{E}(Y_i|Z_i = 1, X_i = x)\mathbb{E}(A_i|Z_i = 0) - \mathbb{E}(Y_i|Z_i = 0, X_i = x)\mathbb{E}(A_i|Z_i = 1)}$$

Proof. The bandwidth is chosen such that the subjects included in the data are balanced and exchangeable with respect to confounders.

Based on Assumption 4; Z is independent of confounders conditional on X, we can obtain unbiased estimates of the following (the effect of Z on Y):

$$\lim_{x \to x_0} \mathbb{E}(Y_i|Z_i = z, X_i = x) \text{ for } z \in \{0, 1\}$$

These can be obtained by fitting logistic regression models for patients above and below the threshold.

For simplicity, we drop $\lim_{x \to x_0}$ and X so that we have

$$\lim_{x \to x_0} \mathbb{E}(Y_i|Z_i = 1, X_i = x) \equiv \mathbb{E}(Y_i|Z_i = 1) \text{ and}$$

$$\lim_{x \to x_0} \mathbb{E}(Y_i|Z_i = 0, X_i = x) \equiv \mathbb{E}(Y_i|Z_i = 0)$$

Using the law of total probability:

$$\begin{aligned}\mathbb{E}(Y_i|Z_i=1) =& \mathbb{E}(\mathbb{E}(Y_i|Z_i=1)|A_i)\\ =& \mathbb{E}(Y_i|Z_i=1,A_i=1)\mathbb{P}(A_i=1|Z_i=1)\\ &+\mathbb{E}(Y_i|Z_i=1,A_i=0)\mathbb{P}(A_i=0|Z_i=1)\end{aligned}$$

Applying Assumption 5 - Conditional independence of Y and Z:

$$\mathbb{E}(Y_i|Z_i=1) = \mathbb{E}(Y_i|A_i=1)\mathbb{P}(A_i=1|Z_i=1) + \mathbb{E}(Y_i|A_i=0)\mathbb{P}(A_i=0|Z_i=1) \tag{4.4}$$

Also,

$$\mathbb{E}(Y_i|Z_i=0) = \mathbb{E}(Y_i|A_i=1)\mathbb{P}(A_i=1|Z_i=0) + \mathbb{E}(Y_i|A_i=0)\mathbb{P}(A_i=0|Z_i=0) \tag{4.5}$$

Solving Equations 4.4 and 4.5 simultaneously yields

$$\mathbb{E}(Y_i|A_i=1) = \frac{\mathbb{E}(Y_i|Z_i=1)\mathbb{P}(A_i=0|Z_i=0) - \mathbb{E}(Y_i|Z_i=0)\mathbb{P}(A_i=0|Z_i=1)}{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)} \tag{4.6}$$

$$\mathbb{E}(Y_i|A_i=0) = \frac{\mathbb{E}(Y_i|Z_i=0)\mathbb{P}(A_i=1|Z_i=1) - \mathbb{E}(Y_i|Z_i=1)\mathbb{P}(A_i=1|Z_i=0)}{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)} \tag{4.7}$$

Therefore, the estimator for the risk ratio based on the RD design assumptions is given in Equation 4.8 as the ratio of Equations 4.6 and 4.7

$$\text{RDD-RR} = \frac{\mathbb{E}(Y_i|Z_i=1)\mathbb{P}(A_i=0|Z_i=0) - \mathbb{E}(Y_i|Z_i=0)\mathbb{P}(A_i=0|Z_i=1)}{\mathbb{E}(Y_i|Z_i=0)\mathbb{P}(A_i=1|Z_i=1) - \mathbb{E}(Y_i|Z_i=1)\mathbb{P}(A_i=1|Z_i=0)} \tag{4.8}$$

By substituting $\mathbb{P}(A_i=0|Z_i=z) = 1-\mathbb{P}(A_i=1|Z_i=z)$ and inserting $\lim_{x\to x_0}$ and X, we have

$$\text{RDD-RR} = 1 - \lim_{x\to x_0} \frac{\mathbb{E}(Y_i|Z_i=1,X_i=x) - \mathbb{E}(Y_i|Z_i=0,X_i=x)}{\mathbb{E}(Y_i|Z_i=1,X_i=x)\mathbb{E}(A_i|Z_i=0) - \mathbb{E}(Y_i|Z_i=0,X_i=x)\mathbb{E}(A_i|Z_i=1)} \tag{4.9}$$

□

Equation 4.9 is the estimator for risk ratio that we derived based on RD design

assumptions. The performance of this estimator will be compared with the Wald and MSMM estimators. In the next section, we shall describe the bootstrapping approach that we use to estimate the variance of the three methods described above, when fitting models.

4.2.4 Non-parametric bootstrap

So far, we have described three methods that can be used to estimate the risk ratio in an RD design. Two are existing methods in the instrumental variable framework that can be adapted to the RD design and the third one is a new approach that we have proposed. In order to compare these three methods, it is also important to measure the variability of the estimates from each method. We use a bootstrapping approach to estimate the variance of the three methods of estimating the risk ratio in an RD design.

Bootstrapping is a re-sampling technique that can be used to obtain the properties of unknown population parameters (Efron, 1981). This is similar to the use of sampling distributions to obtain the properties of the estimators of the population parameters. In theory, sampling distributions are obtained by drawing all samples from the population and calculating the estimates of interest from the samples drawn. In reality, we usually have only one sample from the population. Therefore, bootstrapping involves taking samples with replacement from the observed sample, where the samples from the observed sample are known as bootstrap samples. The accuracy of this approach largely depends on the assumption that the observed sample is representative of the population (Berrar, 2019).

Given a dataset of size N from which we have computed the estimates of the causal risk ratio using the methods described above, the steps taken to calculate the variance of the risk ratio estimates are outlined as follows (Efron and Tibshirani, 1993):

Step 1: Take a random sample (with replacement) of size N from the original dataset.

Step 2: Calculate the Wald-RR, RDD-RDD and MSMM estimates of the risk ratio from the bootstrap sample generated in Step 1.

Step 3: Repeat Steps 1 and 2 B times to obtain B estimates of the risk ratio for each of the three methods.

Step 4: The variance of the estimator is computed as the sample variance of B estimates for each of the methods.

In this chapter, the focus is on the estimation of the causal risk ratio. However, another popular treatment effect measure for a binary outcome is the odds ratio. Next, we shall briefly discuss the odds ratio and why it might not be suitable for treatment effect estimation in the context of an RD design.

4.3 Non collapsibility of odds ratio

The odds ratio is another measure that is used to estimate treatment effect for binary outcomes. Consider a binary outcome variable Y and a treatment indicator A. In addition, we define B as a variable that is associated with Y but is not a confounder, that is, B is not associated with A. Without loss of generality, we assume that B is a binary variable. We define the marginal odds ratio to be:

$$\text{OR}_{YA} = \frac{\mathbb{P}(Y=1|A=1)}{\mathbb{P}(Y=0|A=1)} \frac{\mathbb{P}(Y=0|A=0)}{\mathbb{P}(Y=1|A=0)},$$

and the conditional odds ratio to be:

$$\text{OR}_{YA|B} = \frac{\mathbb{P}(Y=1|A=1,B=b)}{\mathbb{P}(Y=0|A=1,B=b)} \frac{\mathbb{P}(Y=0|A=0,B=b)}{\mathbb{P}(Y=1|A=0,B=b)}.$$

A measure is said to be collapsible if the marginal measure is a weighted average of the conditional measures, conversely, a measure is non-collapsible if the marginal measure is not a weighted average of the conditional measures (Hernán et al., 2011; Huitfeldt et al., 2019). It has been established that the odds ratio is a non-collapsible measure (Burgess, 2016), and we will illustrate a consequence of non-collapsibility of the odds ratio using an example similar to one used in Greenland et al. (1999) of an hypothetical population depicted in Table 4.1. We observe that the probability of receiving treatment is 0.5 for the two levels of B, that is, A and B are not associated

with each other, therefore, B is not a confounder. From Table 4.1, we observe that the conditional odds ratio for the two levels of B are both equal 3.86. However, despite the conditional odds ratios being equal, the marginal odds ratio is not equal to 3.86. This is a consequence of the non-collapsibility of the odds ratio where the marginal odds ratio not bounded by the conditional odds ratios. Typically, we expect that the marginal measure is bounded by the conditional measures, however, this intuition does not apply for non-collapsible measures.

Table 4.1: Example of non-collapsibility of the odds ratio

	$B=1$		$B=0$		Marginal	
	$A=1$	$A=0$	$A=1$	$A=0$	$A=1$	$A=0$
$Y=1$	90	70	30	10	120	80
$Y=0$	10	30	70	90	80	120
$\mathbb{P}(Y=1)$	0.9	0.7	0.3	0.1	0.6	0.4
Odds ratio	3.86		3.86		2.25	

While collapsibility is a desirable property, this is does not mean that non-collapsible measures should be avoided. For a non-collabsible measure, like the odds ratio, the marginal odds ratio and conditional odds ratio are measuring different quantities, that is the marginal odds ratio is targeting the overall treatment effect while the conditional odds ratio is targeting the treatment effect for a subgroup of the population. Therefore, neither the marginal nor the conditional measure is wrong, the choice of which one to report simply depends on the research question. However, we note that for a non-collapsible measure, it might be difficult to distinguish between disparity due to non-collapsibity and disparity due to confounding bias when assessing the performance of estimators. Therefore, in this thesis, we stick to estimating the risk ratio because it is a collapsible measure (Burgess, 2016).

In the next section, we shall compare the performance of the three methods for estimating the risk ratio; Wald-RR, MSMM and RDD-RR using simulation studies.

4.4 Simulation study

We have discussed three approaches for estimating the causal risk ratio in an RD design which are

- the Wald-RR method,
- the multiplicative structural mean model (MSMM) and
- the RD design approach that we have derived: RDD-RR.

In this section, we will conduct simulation studies to evaluate the performance of the three approaches and examine how the novel RDD-RR approach compares to the Wald IV and MSMM methods. We compare the methods by varying the level of fuzziness in the design and we also vary the degree of unobserved confounding. A description of the simulation scenarios considered is given below.

4.4.1 Description of simulation study

The data simulation in this section were based on the real dataset on statin prescription in the UK primary care that was described in Section 3.6. As such, we use the original values of the risk scores as assignment variable in the simulated data. We recall that the guideline for statin prescription states that statins should be prescribed if a patient's risk of developing CVD in 10 years is greater or equal to 20%, therefore, the threshold is set to be 20%.

For each patient i, $i = 1, ..., 1384$, we describe the steps of simulating the treatment indicator and outcome variable as well as other variables as we will expect from a fuzzy RD design.

Step 1: The centred assignment variable (X_i^c) and threshold indicator (Z_i) are defined as:

$$X_i^c = X_i - 0.2,$$
$$Z_i = \mathbb{1}(X_i^c \geq 0).$$

Step 2: A confounding variable U_i is simulated from a uniform distribution.

$$U_i \sim \text{Uniform}(0, 1)$$

Step 3: We calculate p_i, the probability of receiving the treatment from Equation 4.10:

$$\text{logit}(p_i) = \beta_0 + \beta_1 Z_i + \beta_2 U_i + \beta_3 X_i^c \tag{4.10}$$

The parameters of the equation above are set to reflect the level of fuzziness of the design and the effect of confounding with respect to the treatment. The lower the value of β_1, the more fuzzy the design. β_2 reflects effect of confounding, if $\beta_2 = 0$, this indicates there is no relationship between A_i and U_i and, therefore, no confounding.

Step 4: The treatment indicator (A_i) is simulated from a Bernoulli distribution where p_i is as given above:

$$A_i \sim \text{Bernoulli}(p_i).$$

Step 5: The expectation of the binary outcome Y_i (p_i^y) is calculated from

$$\text{logit}(p_i^y) = \beta_4 + \beta_5 A_i + \beta_6 U_i \tag{4.11}$$

β_5 reflects the treatment effect, the treatment effect (risk ratio) is calculated as

$$\text{RR} = \frac{\int_0^1 \text{expit}(\beta_4 + \beta_5 + \beta_6 u_i) f_U(u_i) \text{d}u_i}{\int_0^1 \text{expit}(\beta_4 + \beta_6 u_i) f_U(u_i) \text{d}u_i},$$

where $\text{expit}(x) = \frac{\exp(x)}{1+\exp(x)}$. β_6 reflects the effect of confounding with respect to the outcome. If $\beta_6 = 0$, this implies there is no relationship between Y_i and U_i, which means U_i is not a confounder.

Step 6: The binary outcome (Y_i) is then simulated from a Bernoulli distribution. In our example, the outcome represents whether patient i experiences at least a 1mmol/L reduction in LDL cholesterol level or not.

$$Y_i \sim \text{Bernoulli}(p_i^y).$$

Step 7: Steps 1 to 6 above are repeated 2000 times to produce 2000 datasets.

The values of the parameters of Equations 4.10 and 4.11 are defined as follows.

i. β_0 is set so that the overall probability of receiving the treatment is 0.5. This is in line with what was observed in the real dataset that will be described in Section 4.5, later. The value of β_0 is determined as the solution of the following non-linear equation:

$$\frac{1}{n}\sum_{i=1}^{n} \text{expit}(\beta_0 + \beta_1 Z_i + \beta_2 U_i + \beta_3 X_i^c) = 0.5.$$

ii. $\beta_3 = 2$ to reflect the positive relationship between the treatment indicator and assignment variable.

iii. β_4 is set so that the overall $\mathbb{P}(Y_i = 1)$ is 0.44. Again this value is in line with what was observed from the real dataset that we will be discussing later.

Simulation Scenarios

Here, we describe the specified values for β_1 and β_2 in Equation 4.10 and β_6 in Equation 4.11. This simulation approach and choice of values of the parameters is similar to the one used by Geneletti et al. (2015). In Table 4.2, the values of the parameters are specified and there are six scenarios in total which represent varying levels of fuzziness and levels of confounding in the simulated data.

For each of 2000 simulated datasets, we computed the correlation coefficient between the confounder (U) and the treatment indicator, and between U and the outcome. The average of the estimated correlation coefficients is reported in Table 4.2 as $\rho_{A,U}$ and $\rho_{Y,U}$ respectively. The values of the correlation coefficient give an insight about the level of (linear) association between the confounder and the treatment and the confounder and the outcome. For scenarios where there are no unobserved confounders, the estimate of the correlation coefficient between the outcome and the confounder is close to zero. For scenarios where the effect of the confounder is low, the values of the correlation coefficient are slightly larger than when there are no confounders. The correlations become higher when the effect of the unobserved confounder is high.

Likewise, we calculated the probability of complying to the treatment guideline, that is, $\mathbb{P}(A_i = 1|Z_i = 1) - \mathbb{P}(A_i = 1|Z_i = 0)$ for the 2000 simulated datasets in each scenario. The average of the probabilities is reported as P.C. in Table 4.2. We define this based on the level of fuzziness - if there is a high compliance to the treatment guideline, the fuzziness is weak, and if the probability of compliance is low then fuzziness is strong. The probability of compliance for the scenarios with weak fuzziness is between 0.8 and 0.9 and for the strong fuzziness scenarios the probability of compliance to the treatment guideline is around 0.53.

Table 4.2: Values of parameters in Equations 4.10 and 4.11 for the simulation scenarios with the corresponding probability of compliance (P.C.) and estimates of correlation coefficients between Y and U ($\rho_{Y,U}$) and A and U ($\rho_{A,U}$)

Scenario	Parameters			P.C.	$\rho_{Y,U}$	$\rho_{A,U}$
Weak Fuzziness, No Confounding	$\beta_1 = 6$	$\beta_2 = 0$	$\beta_6 = 0$	0.91	0.00	0.00
Weak Fuzziness, Low Confounding	$\beta_1 = 8$	$\beta_2 = 6.5$	$\beta_6 = 1$	0.90	0.16	0.12
Weak Fuzziness, High Confounding	$\beta_1 = 8$	$\beta_2 = -9$	$\beta_6 = 2.5$	0.81	0.29	-0.24
Strong Fuzziness, No Confounding	$\beta_1 = 2$	$\beta_2 = 0$	$\beta_6 = 0$	0.53	0.00	0.00
Strong Fuzziness, Low Confounding	$\beta_1 = 2$	$\beta_2 = 1.5$	$\beta_6 = 1$	0.52	0.16	0.15
Strong Fuzziness, High Confounding	$\beta_1 = 2.5$	$\beta_2 = -3.5$	$\beta_6 = 2$	0.53	0.21	-0.31

4.4.2 Results of simulation studies

In this section, we will now apply the three methods described for the estimation of the risk ratio in an RD design to the datasets obtained from the simulation studies. A sensitivity analysis was carried out using different bandwidths and, as such, we considered four bandwidth sizes; 0.2, 0.15, 0.1 and 0.05. This is to check how the estimates vary across different bandwidth sizes.

Figures 4.1 and 4.2 show boxplots of the estimates of the risk ratio obtained from the simulation studies under the weak and strong fuzziness scenarios respectively. The figures include the estimates for the four bandwidth sizes considered and for the no, low and high confounding scenarios. Tables 4.3, 4.4 and 4.5 present the numerical summaries of the results from the simulation studies under the no, low and high confounding scenarios, respectively. Results are summarised using the mean of the estimates, bias, empirical and average standard errors and the 95%

coverage. Because the risk ratio is generally skewed, the numerical summaries given in the tables are for the log of the risk ratio estimates.

Under the no confounding and weak fuzziness scenario, as presented in the left panel of Table 4.3, we observed that the three methods yield estimates that are close to true treatment effect value. Where the bandwidth is 0.2, the biases are the same for the three methods, however, we observe that only for the RDD-RR approach does the bias reduce as the bandwidth size becomes smaller, as we should expect in the estimates. In addition, we observe that the MSMM approach seems to have higher variability, looking at the ESE and ASE when compared to the RDD-RR and WALD-RR approaches. In terms of 95% coverage, the three methods provide a good coverage of the treatment effect.

For the no confounding and strong fuzziness scenario, in the right panel of Table 4.3, we observe a similar pattern to that to under the no confounding and weak fuzziness scenario. The estimates obtained from the three approaches are close to the true treatment effect and the biases of the RDD-RR approach reduce as the bandwidth size reduces. However, we observed that the variability (ESE and ASE) is higher under the strong fuzziness scenario compared to the weak fuzziness scenario. This is not unexpected as the RD design is based on the relationship between the threshold indicator and the treatment. The strong fuzziness scenario indicates a weak relationship between the threshold indicator and treatment and as a result, this leads to an increase in uncertainty in the estimates.

Also, under the strong fuzziness scenario, especially for the smaller bandwidths, the coverage is above the nominal level of 95%, this may be because in these scenario, the ASE is greater than the ESE. The ESE measures the standard deviation of the estimates which approximates the correct standard error of the sampling variation of the estimate. As a result, the ASE targets the ESE and therefore, in this example, where the ASE is greater than than ESE, this implies an overestimation of the standard error which in turn leads to coverage being above the nominal level.

Table 4.4 shows numerical summaries of estimates from the simulation studies conducted under the low confounding scenario. The results obtained under this scenario are similar to those observed under the no confounding scenario. Therefore,

the points made above are applicable to this scenario also. The fact that the results obtained under the no confounding scenario are similar to those observed under the low confounding scenario suggests that the performance of the three methods is not overly affected by confounding. This suggests that the RD design may be useful as a methodology for treatment effect estimation in an observational studies in the presence of unobserved confounding.

Under the scenario where confounding is high, as shown in Table 4.5, the RDD-RR and Wald-RR methods continue to produce unbiased estimates with coverage close to the nominal level and are not affected by the presence of unobserved confounding. However, the MSMM approach yields biased estimates and it continues to have a higher variability resulting in a high coverage that is above the nominal value.

Looking at the results across the bandwidths, as expected, the variability of estimates becomes larger as the bandwidth size reduces. This is because the number of observations that is used to estimate the treatment effect becomes smaller as the bandwidth size decreases. In terms of bias, we see that the bias observed in the estimates from RDD-RR approach reduces as the bandwidth reduces which would also be expected in an RD design. However, the bias of the MSMM and Wald-RR approaches does not seem reduce with the bandwidth. This may be because these methods were developed under the instrumental variable framework where bandwidths are not considered as in an RD design.

From the simulation studies carried out, we observe that the three methods are comparable in terms of estimating the treatment effect with little to no bias under the no and low confounding scenarios. However, we observe that the MSMM method has a higher variability compared to the other two methods which results in a coverage higher than the nominal level. Also, it is observed that, under the high confounding scenario, the MSMM method produced biased estimates of the treatment effect at the threshold. In addition, the computational time of the MSMM approach is about twice that of the RDD-RR and Wald-RR approaches. This suggests that the RDD-RR and WALD-RR approaches should be preferred over the MSMM approaches.

From the results of the simulation studies, it seems that the proposed RDD-RR approach and the Wald-RR approach are comparable. However, it has been noted

that the Wald-RR approach is not consistent for the causal risk ratio when the treatment effect is large. In order to check this, we carried out additional simulation studies where the value of the true risk ratio is set to be 4, is considered to be a large treatment effect.

Figures 4.3 and 4.4 are the boxplots of the estimates of the risk ratio under the weak and strong fuzziness scenarios, respectively, when the risk ratio set to 4. Tables 4.6, 4.7 and 4.8 show numerical summaries of the estimates of the risk ratio under the no, low and high confounding scenarios, respectively, when the risk ratio is set to 4.

Under the weak fuzziness scenarios, examining Figure 4.3 and the left panels of Tables 4.6, 4.7 and 4.8, we observe that the WALD-RR approach yields biased estimates of the treatment effect. On the other hand, the RDD-RR approach continues to yield estimates close the true treatment effect and its coverage is close to the nominal value. This is a confirmation of the claim that the WALD-RR method performs poorly when the treatment effect is high and we observed that estimates obtained from the WALD-RR approach do not accurately estimate the true treatment effect. On the other hand, the RDD-RR approach estimates are unbiased for the weak scenario across the bandwidths considered.

When the level of fuzziness is strong, we observe that both the RDD-RR and WALD-RR approaches yield biased estimates of the treatment effect. We note, however, that the RDD-RR approach still has a lower bias, in general, and it continues to provide better coverage of the treatment effect compared to the WALD-RR approach. As such, the RDD-RR approach performs better than the WALD-RR approach and should be preferred over the WALD-RR approach.

Based on the results that we have obtained from the simulation studies, the proposed RDD-RR approach appears to be a suitable estimator of the risk ratio in an RD design. It produces estimates close to the value of the true treatment effect under the no, low and high confounding scenarios which is a desirable characteristic for a method to estimate treatment effect in an observational study data may be subject to unobserved confounding.

We have compared the three methods for estimating risk ratio in a fuzzy RD design using simulation studies. In the next section, we shall apply the three methods to a

real dataset on the prescription of statins in UK primary care.

Table 4.3: Estimates, biases, empirical standard errors, average standard errors and 95% coverage for the log of the risk ratio under the no confounding scenario. The true value of the log of the risk ratio is $\log(1.5) = 0.405$. The sample size is 1384 in each simulated dataset and simulations were repeated 2000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405									
RDD-RR	0.40	0.01	0.13	0.13	94.9	0.38	0.02	0.19	0.20	96.2
MSMM	0.41	-0.01	0.15	0.15	95.6	0.42	-0.02	0.22	0.24	97.2
Wald-RR	0.41	-0.01	0.14	0.14	94.8	0.39	0.02	0.20	0.20	95.7
	Bandwidth = 0.15 , Treatment effect = 0.405									
RDD-RR	0.40	0.01	0.14	0.14	95.0	0.38	0.02	0.20	0.20	95.8
MSMM	0.41	-0.01	0.15	0.16	95.9	0.42	-0.01	0.23	0.25	97.1
Wald-RR	0.41	-0.01	0.15	0.15	95.2	0.39	0.02	0.20	0.20	95.4
	Bandwidth = 0.1 , Treatment effect = 0.405									
RDD-RR	0.40	0.00	0.16	0.16	95.2	0.39	0.02	0.23	0.24	96.2
MSMM	0.41	-0.01	0.18	0.18	96.3	0.42	-0.02	0.26	0.30	97.6
Wald-RR	0.42	-0.01	0.17	0.17	95.0	0.40	0.01	0.23	0.24	95.5
	Bandwidth = 0.05 , Treatment effect = 0.405									
RDD-RR	0.41	0.00	0.22	0.22	95.5	0.40	0.01	0.31	0.34	98.2
MSMM	0.43	-0.02	0.25	0.26	97.0	0.44	-0.03	0.38	0.44	98.1
Wald-RR	0.42	-0.02	0.23	0.23	95.4	0.40	0.00	0.32	0.33	96.7

Table 4.4: Estimates, biases, empirical standard errors, average standard errors and 95% coverage for the log of the risk ratio under the low confounding scenario. The true value of the log of the risk ratio is log(1.5) = 0.405. The sample size is 1384 in each simulated dataset and simulations were repeated 2000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
Bandwidth = 0.2 , Treatment effect = 0.405										
RDD-RR	0.40	0.00	0.13	0.13	95.5	0.38	0.03	0.20	0.20	95.2
MSMM	0.40	0.01	0.14	0.14	95.3	0.41	-0.01	0.23	0.24	96.8
Wald-RR	0.41	-0.01	0.14	0.14	95.5	0.38	0.02	0.20	0.20	94.7
Bandwidth = 0.15 , Treatment effect = 0.405										
RDD-RR	0.40	0.00	0.14	0.14	95.6	0.38	0.02	0.21	0.21	95.3
MSMM	0.40	0.01	0.14	0.15	95.7	0.41	-0.01	0.24	0.25	96.5
Wald-RR	0.41	-0.01	0.14	0.15	96.0	0.39	0.02	0.21	0.21	94.8
Bandwidth = 0.1 , Treatment effect = 0.405										
RDD-RR	0.40	0.00	0.15	0.16	95.9	0.39	0.02	0.24	0.24	95.7
MSMM	0.40	0.00	0.16	0.17	96.1	0.42	-0.01	0.28	0.30	97.1
Wald-RR	0.42	-0.01	0.16	0.17	95.9	0.39	0.01	0.24	0.24	94.6
Bandwidth = 0.05 , Treatment effect = 0.405										
RDD-RR	0.40	0.00	0.22	0.22	95.0	0.40	0.01	0.33	0.35	98.2
MSMM	0.41	0.00	0.23	0.24	96.3	0.44	-0.03	0.40	0.46	98.5
Wald-RR	0.42	-0.02	0.23	0.23	94.9	0.40	0.00	0.33	0.34	96.2

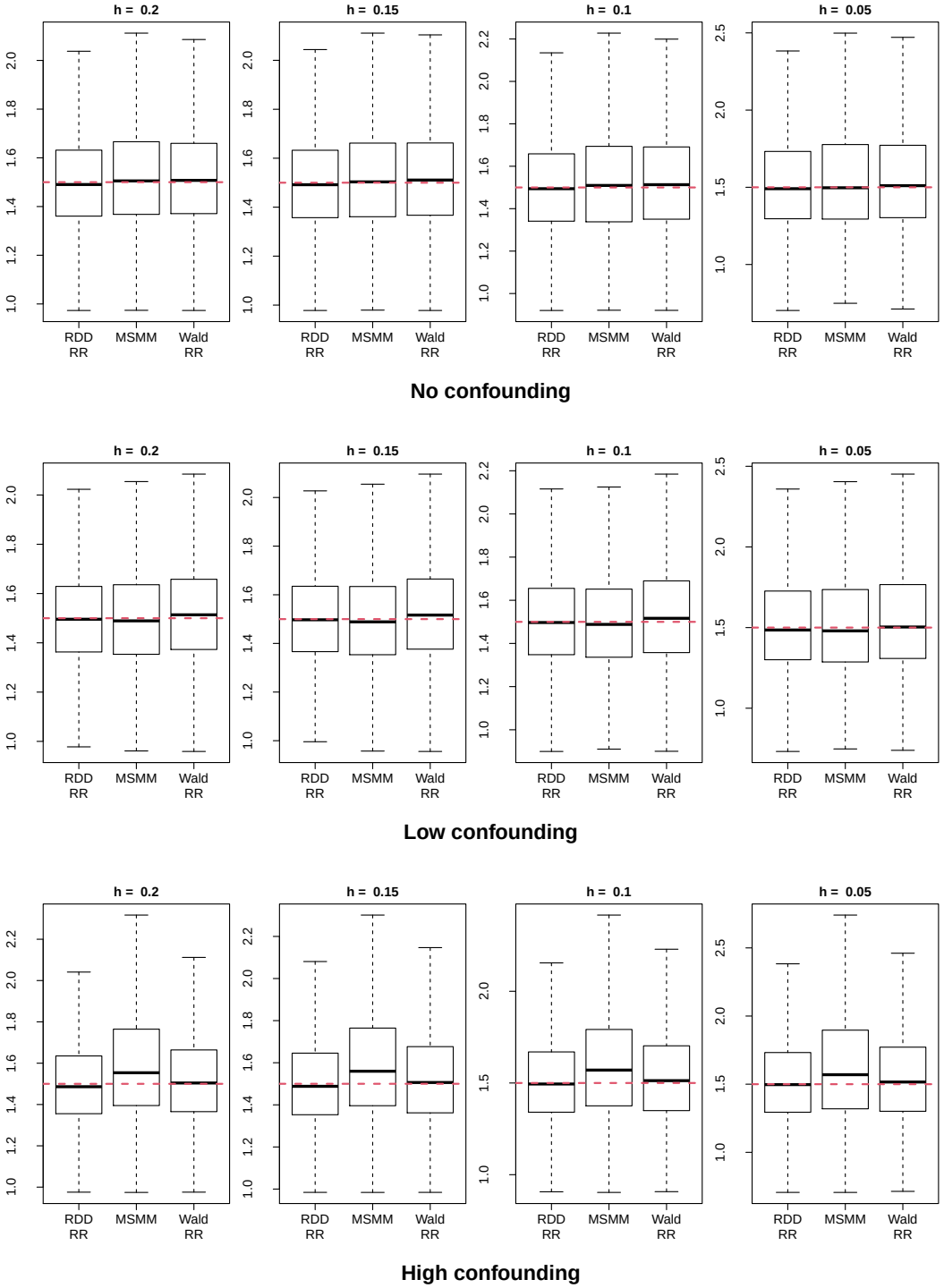

Figure 4.1: Boxplots of the simulation study estimates to compare the RDD-RR, MSMM and Wald-RR methods of estimating the risk ratio under weak fuzziness scenario. The red dashed line denotes the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

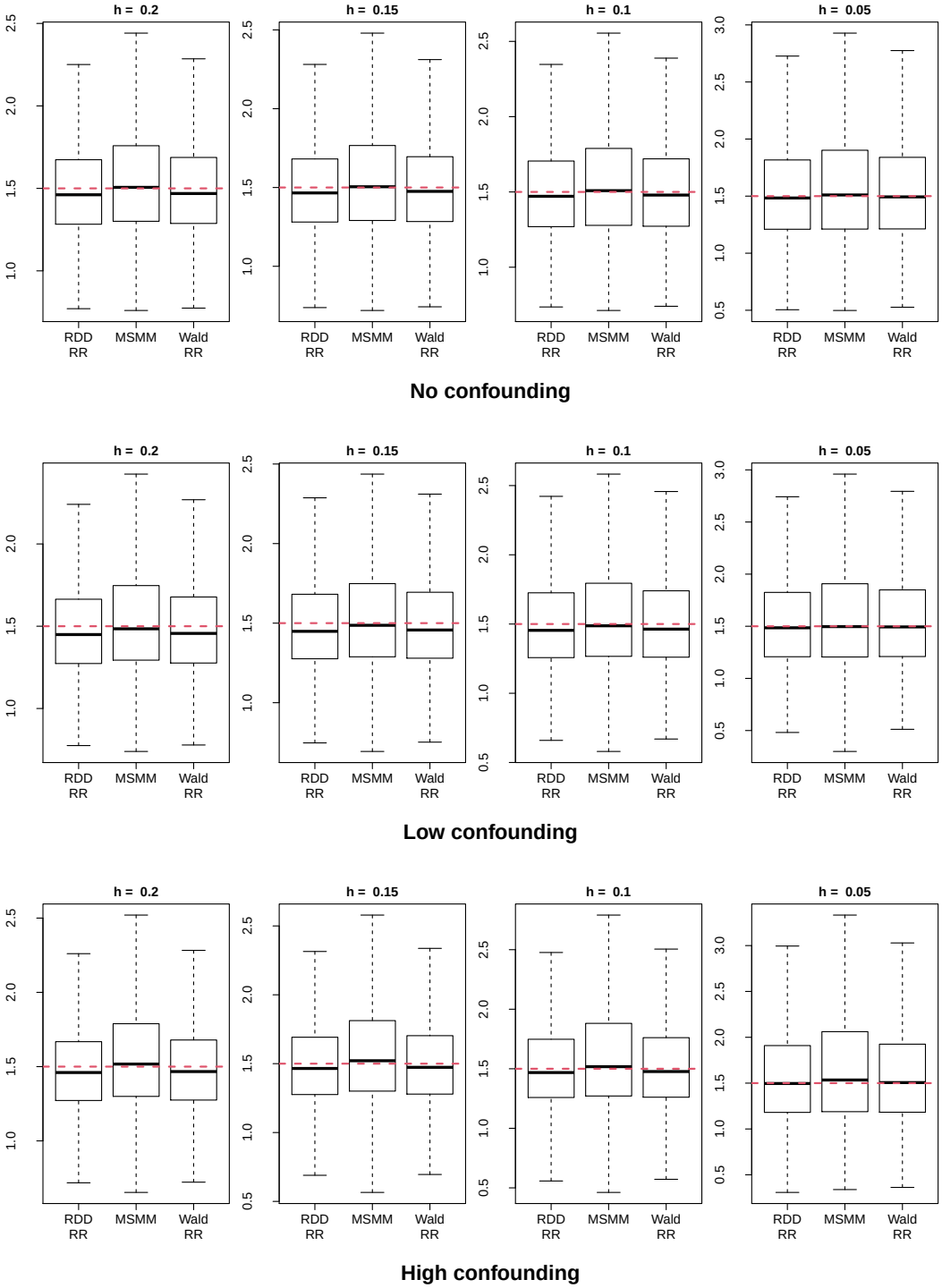

Figure 4.2: Boxplots of the simulation study estimates to compare the RDD-RR, MSMM and Wald-RR methods of estimating the risk ratio under strong fuzziness scenario. The red dashes line denotes the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 4.5: Estimates, biases, empirical standard errors, average standard errors and 95% coverage for the log of the risk ratio under the high confounding scenario. The true value of the log of the risk ratio is log(1.5) = 0.405. The sample size is 1384 in each simulated dataset and simulations were repeated 2000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
					Bandwidth = 0.2 , Treatment effect = 0.405					
RDD-RR	0.40	0.01	0.14	0.14	94.3	0.39	0.02	0.21	0.21	95.3
MSMM	0.45	-0.05	0.17	0.17	96.5	0.44	-0.04	0.27	0.29	97.5
Wald-RR	0.41	-0.01	0.14	0.14	94.5	0.39	0.02	0.21	0.21	94.7
					Bandwidth = 0.15 , Treatment effect = 0.405					
RDD-RR	0.40	0.00	0.14	0.14	94.5	0.39	0.02	0.22	0.22	96.0
MSMM	0.45	-0.05	0.18	0.18	96.7	0.44	-0.04	0.28	0.30	97.9
Wald-RR	0.41	-0.01	0.15	0.15	94.5	0.39	0.01	0.22	0.22	95.2
					Bandwidth = 0.1 , Treatment effect = 0.405					
RDD-RR	0.40	0.00	0.16	0.16	95.7	0.39	0.01	0.25	0.26	97.0
MSMM	0.46	-0.05	0.20	0.21	97.8	0.45	-0.04	0.31	0.36	98.3
Wald-RR	0.42	-0.01	0.17	0.17	95.8	0.40	0.01	0.25	0.26	96.0
					Bandwidth = 0.05 , Treatment effect = 0.405					
RDD-RR	0.41	0.00	0.22	0.22	95.6	0.41	0.00	0.35	0.38	98.6
MSMM	0.47	-0.07	0.28	0.31	97.9	0.47	-0.06	0.44	0.53	98.9
Wald-RR	0.42	-0.02	0.23	0.23	95.5	0.41	-0.01	0.35	0.36	96.2

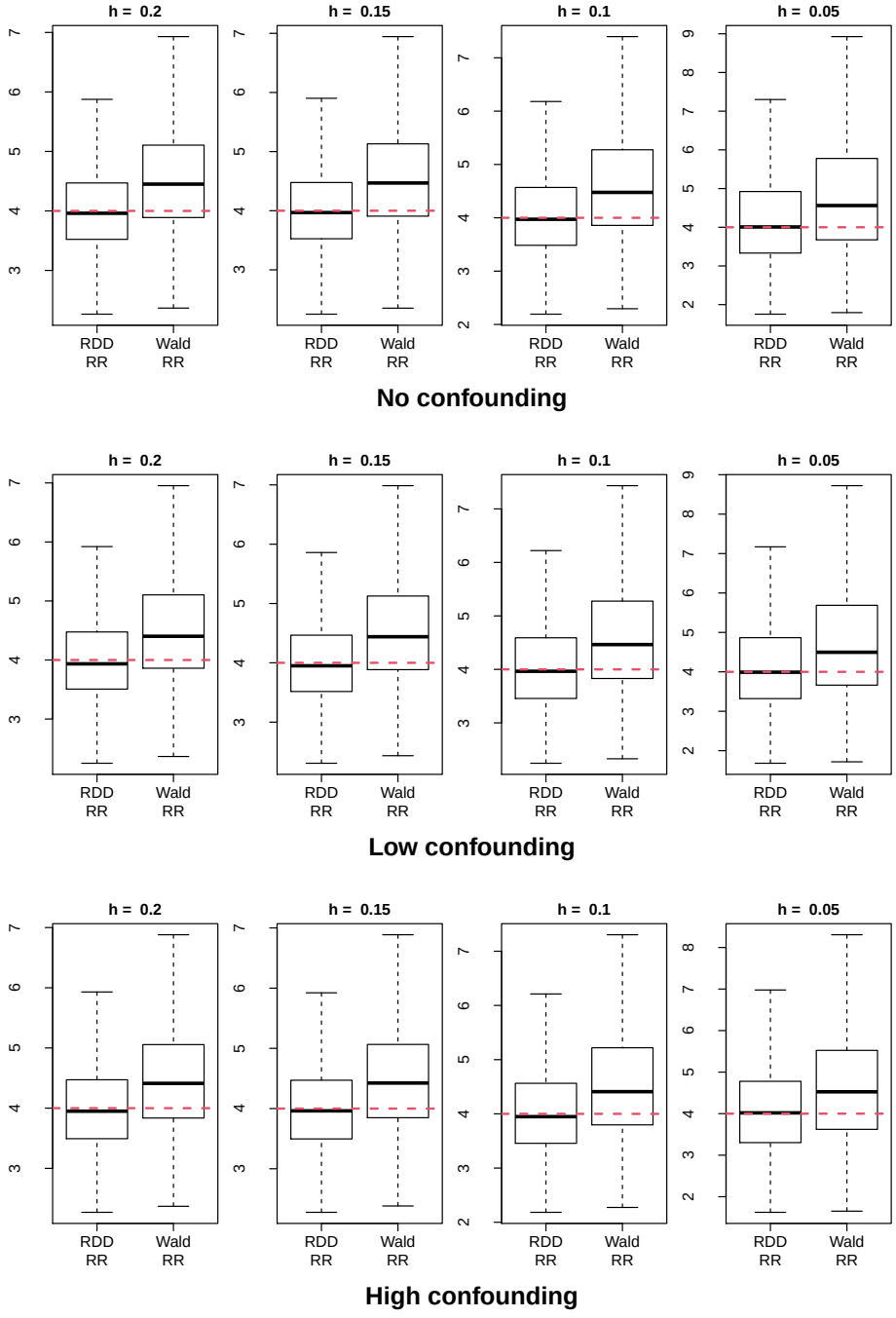

Figure 4.3: Boxplots the simulation study estimates to compare the RDD-RR and Wald-RR methods of estimating the risk under weak fuzziness scenario for large treatment effect. The red dashed line denotes the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

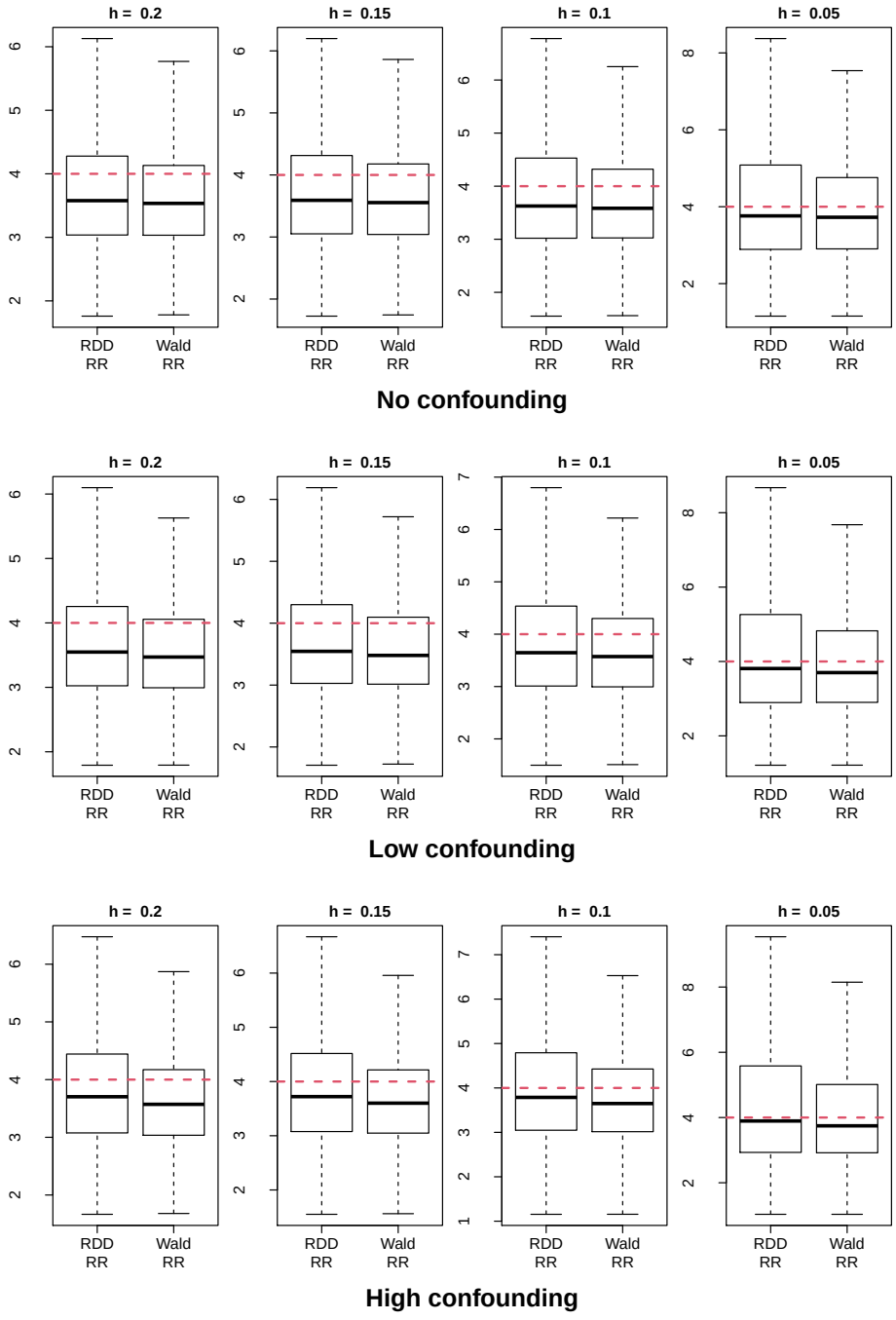

Figure 4.4: Boxplots of the simulation study estimates to compare the RDD-RR and Wald-RR methods of estimating the risk under strong fuzziness scenario for large treatment effect. The red dashed line denotes the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 4.6: Estimates, biases, empirical standard errors, average standard errors and 95% coverage for the log of the risk ratio under the no confounding scenario for large treatment effect. The true value of the log of the risk ratio is $\log(4) = 1.387$. The sample size is 1384 in each simulated dataset and simulations were repeated 2000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 1.387									
RDD-RR	1.38	0.01	0.18	0.18	94.2	1.30	0.09	0.27	0.27	91.8
WALD-RR	1.50	-0.11	0.21	0.20	92.9	1.27	0.12	0.23	0.23	90.6
	Bandwidth = 0.15 , Treatment effect = 1.387									
RDD-RR	1.38	0.00	0.18	0.18	94.3	1.30	0.09	0.27	0.28	92.7
WALD-RR	1.50	-0.11	0.21	0.21	92.2	1.27	0.11	0.24	0.24	91.4
	Bandwidth = 0.1 , Treatment effect = 1.387									
RDD-RR	1.39	0.00	0.21	0.20	94.5	1.32	0.07	0.31	0.33	93.4
WALD-RR	1.51	-0.13	0.24	0.24	93.0	1.29	0.10	0.27	0.27	92.6
	Bandwidth = 0.05 , Treatment effect = 1.387									
RDD-RR	1.41	-0.02	0.29	0.29	95.0	1.37	0.01	0.47	0.50	94.5
WALD-RR	1.54	-0.15	0.34	0.33	94.5	1.32	0.06	0.38	0.38	94.0

Table 4.7: Estimates, biases, empirical standard errors, average standard errors and 95% coverage for the log of the risk ratio under the low confounding scenario for large treatment effect. The true value of the log of the risk ratio is log(4) = 1.387. The sample size is 1384 in each simulated dataset and simulations were repeated 2000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 1.387									
RDD-RR	1.38	0.01	0.18	0.18	94.7	1.29	0.10	0.27	0.28	91.8
WALD-RR	1.50	-0.11	0.20	0.20	93.0	1.25	0.13	0.23	0.23	89.7
	Bandwidth = 0.15 , Treatment effect = 1.387									
RDD-RR	1.38	0.00	0.18	0.18	94.9	1.29	0.10	0.27	0.29	92.6
WALD-RR	1.50	-0.11	0.21	0.21	93.6	1.26	0.13	0.23	0.24	90.5
	Bandwidth = 0.1 , Treatment effect = 1.387									
RDD-RR	1.39	0.00	0.20	0.20	95.0	1.32	0.07	0.32	0.34	94.1
WALD-RR	1.51	-0.12	0.23	0.23	93.8	1.28	0.11	0.27	0.28	93.1
	Bandwidth = 0.05 , Treatment effect = 1.387									
RDD-RR	1.40	-0.01	0.28	0.29	95.7	1.39	-0.01	0.51	0.52	94.2
WALD-RR	1.52	-0.14	0.32	0.33	96.0	1.32	0.06	0.39	0.39	94.0

Table 4.8: Estimates, biases, empirical standard errors, average standard errors and 95% coverage for the log of the risk ratio under the high confounding scenario for large treatment effect. The true value of the log of the risk ratio is log(4) = 1.387. The sample size is 1384 in each simulated dataset and simulations were repeated 2000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 1.387									
RDD-RR	1.38	0.01	0.18	0.18	95.4	1.33	0.06	0.29	0.32	94.1
WALD-RR	1.49	-0.10	0.20	0.21	94.2	1.28	0.11	0.24	0.25	92.6
	Bandwidth = 0.15 , Treatment effect = 1.387									
RDD-RR	1.38	0.01	0.18	0.18	95.8	1.33	0.05	0.30	0.32	94.3
WALD-RR	1.49	-0.11	0.20	0.21	94.3	1.28	0.11	0.25	0.25	92.6
	Bandwidth = 0.1 , Treatment effect = 1.387									
RDD-RR	1.38	0.00	0.21	0.21	95.1	1.36	0.02	0.36	0.39	95.2
WALD-RR	1.50	-0.11	0.24	0.24	93.9	1.30	0.09	0.29	0.29	94.0
	Bandwidth = 0.05 , Treatment effect = 1.387									
RDD-RR	1.39	-0.01	0.29	0.29	95.0	1.45	-0.06	0.56	0.57	94.9
WALD-RR	1.51	-0.12	0.33	0.33	95.2	1.35	0.04	0.41	0.42	95.3

4.5 Example: Prescription of Statins in UK Primary Care

In this section, the three methods for estimating the risk ratio in a fuzzy RD design that we have described in Sections 4.2.1, 4.2.2 and 4.2.3 are applied to the real dataset on statin prescription that was introduced in Section 3.6.

The dataset consists of 1384 male patients who are non-diabetic and non-smokers. From the data, information on patients' risk of developing CVD in 10 years (X), treatment status (A), initial low density lipoprotein (LDL) cholesterol level and LDL cholesterol level after treatment allocation are obtained. According to the NICE guideline, the threshold x_0 is set to be 0.2. Therefore, the centred assignment variable is calculated as $X^c = X - 0.2$, and the threshold indicator (Z) is computed using the usual formula.

A 1mmol/L reduction in LDL cholesterol level has been linked to a reduction in major cardiovascular events (Cholesterol Treatment Trialists' (CTT) Collaborators et al., 2012). As a result, we define a binary outcome variable as the indicator function that takes the value 1 if the LDL cholesterol level is reduced by at least 1mmol/L and 0 if otherwise. The initial measurement of the LDL cholesterol level was taken at the time the risk score was computed. The final LDL cholesterol level was measured about 6 weeks after statin prescription (for treated patients) and for untreated patients, within 6 weeks after the initial measurement was taken. We aim to estimate the effect (risk ratio) of statin prescription in reducing LDL cholesterol level by at least 1mmol/L.

Out of the 1384 patients, 657 (49%) patients experienced an LDL cholesterol level reduction of at least 1mmol/L. Figure 4.5 shows a plot of the outcome against the assignment variable and we observe that the probability of experiencing a reduction of at least 1mmol/L in the LDL cholesterol level is lower in patients below the threshold compared to patients above the threshold which suggests a possible beneficial effect of statin prescription.

WALD-RR, MSMM and RDD-RR methods are applied to the THIN dataset described above to estimate the risk ratio. Table 4.9 shows the estimates of the risk

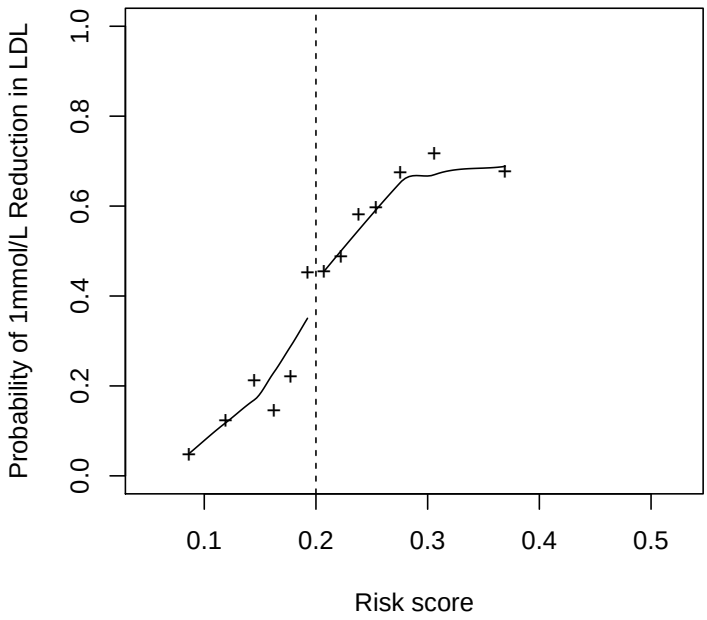

Figure 4.5: Plot of the probability of altleast 1mmol/L reduction in LDL cholesterol level against the risk of developing CVD in 10 years. The black crosses are the expected probabilities calculated in bins.

ratio obtained from the three methods along with the corresponding 95% confidence intervals. As observed in the simulation studies, the MSMM approach yields estimates that have a higher variability compared to those from the other two methods. For bandwidths 0.2, 0.15 and 0.1, the estimates of the risk ratio are greater than one, which suggests patients that receive statins are more likely to have their LDL cholesterol level reduced by at least 1mmol/L compared to patients that did not receive statins. However, for all bandwidths, the treatment effects are not statistically significant (at the 5% level) as all 95% confidence intervals contain 1.

We note that for a bandwidth of 0.05, the treatment effect estimate is lower than 1. Owing to the varying values of the treatment effect across the bandwidths, we should be cautious in interpreting the estimates. However, since for all bandwidths, the 95% confidence intervals contain 1, this implies that there is insufficient evidence to suggest that statin prescription reduces LDL cholesterol level by at least 1mmol/L. In addition, we note that the outcome here is originally a continuous outcome that was categorised, this is generally not advised because dichotomising a continuous variable may lead to a loss of information (Altman and Royston, 2006). Therefore,

the analyses in this section are to illustrate how WALD-RR, MSMM and RDD-RR methods can be applied to a real dataset.

Table 4.9: Estimates of the effect of statins prescription on reduction of LDL cholesterol level.

Method	Risk ratio	95% CI
Bandwidth = 0.2; N = 1368		
RDD-RR	1.43	0.91, 2.23
MSMM	2.43	0.92, 6.41
Wald-RR	1.43	0.91, 2.25
Bandwidth = 0.15; N = 1307		
RDD-RR	1.26	0.79, 2.00
MSMM	1.81	0.52, 6.31
Wald-RR	1.26	0.79, 2.01
Bandwidth = 0.1; N = 1101		
RDD-RR	1.17	0.65, 2.11
MSMM	1.51	0.32, 7.18
Wald-RR	1.17	0.65, 2.10
Bandwidth = 0.05; N = 676		
RDD-RR	0.71	0.24, 2.08
MSMM	0.38	0.04, 3.78
Wald-RR	0.72	0.29, 1.80

Figure 4.6 is the plot of the fitted models with the outcome (experiencing an at least 1 mmol/L reduction in LDL cholesterol level) as response and assignment variable as predictor across the four bandwidths considered. We observe that the fitted lines reflect the results we have in Table 4.9 because for bandwidths 0.2, 0.15 and 0.1, the probability of experiencing an at least 1mmol/L reduction in LDL cholesterol level is higher in the patients above the threshold compared to patients below the threshold. However, the reverse is the case for bandwidth 0.05.

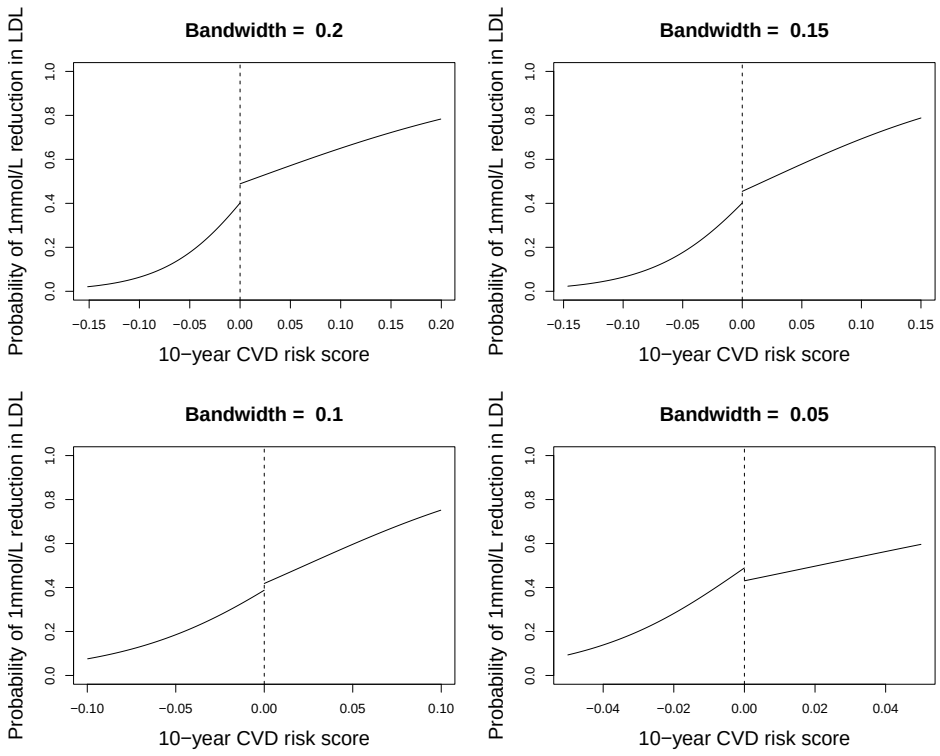

Figure 4.6: Fitted lines for models with the outcome (experiencing an at least 1 mmol/L reduction in LDL cholesterol level) as response and assignment variable as predictor across the bandwidths

4.6 Conclusions

In this chapter, we have explored methods for estimating the risk ratio in an RD design where the outcome of interest is binary. First, we discussed the Wald and MSMM approaches, which are methods originally developed in the instrumental variable framework but can be adapted for use in an RD design. In addition, we proposed the RDD-RR estimator for the risk ratio based on the assumptions of the RD design.

We carried out simulation studies to compare these three methods. Under no and low confounding scenarios, the three methods are comparable; they produce estimates close to the true treatment effect and yield a good coverage of the treatment effect. However, the estimates of MSMM approach are noted to have higher variability compared to estimates from the WALD-RR and RDD-RR approaches. When the effect of unobserved confounding is high, the MSMM approach yields biased estimates of the treatment effect, however, the WALD-RR and RDD-RR approaches continue to yield unbiased estimates of the treatment effect.

We investigated the performance of the WALD-RR and RDD-RR approaches when the treatment effect is large, as it has been noted in literature that the WALD-RR approach is not consistent for the risk ratio for a large treatment effect. The results of the simulation studies confirm this claim.

Overall, the proposed RDD-RR method seems to be preferable for estimating the risk ratio in a fuzzy RD design compared to the other two methods.

We have explored the methods of estimating treatment effect in a fuzzy RD design for continuous and binary outcomes. A time-to-event outcome is another outcome that is often of interest in medical studies and, in the next chapter, we shall explore treatment effect estimation for a time-to-event outcome in a fuzzy RD design.

Chapter 5

Treatment effect estimation for time-to-event data under the accelerated failure time assumption.

In the previous chapters, we have investigated the use of RD design for cases where the outcomes of interest are continuous (non-time-to-event) and binary. In this chapter, we explore the use of RD design for time-to-event outcomes. We propose a method for treatment effect estimation in a fuzzy RD design under the accelerated failure time (AFT) assumption for a time-to-event outcome. This method will then be compared to the structural AFT approach (Hernán et al., 2005), an approach that has been routinely used to estimate the treatment effect in observational studies under the AFT assumption.

5.1 Introduction

An outcome of interest is said to be a time-to-event or survival outcome if it represents the time it takes until an event of interest occurs in an individual. For example, in a study that involves cancer patients, the outcome could be the time to death; or, perhaps, the time until pain relief in a study of a pain-relieving medication. The time-to-event is measured from a pre-defined time origin for all patients, for instance, the time at which patients are enrolled into a study.

A time-to-event outcome is non-negative and usually exhibits right skewness as depicted in the histogram (plotted from a simulated data) in Figure 5.1 (a). Another important feature of time-to-event data is that the actual time of the event may not be observed for some patients. It might be that some patients have not experienced

the event of interest at the end of the study. For such patients, we may only know that they might experience the event at a time beyond the time when study ended. In addition, some patients may drop out of a study before they experienced the event for various reasons. Again, for these patients, we are able to observe the time at which a patient dropped out of the study and that the patient perhaps experiences the event of interest at a later unknown point in time. This is known as right censoring.

Figure 5.1 (b) depicts the observed time-to-event in a study that ends at a hypothetical time point 20. Therefore, the observed time-to-event, as opposed to the actual time-to-event, for any patient that has not experienced the event at the end of the study will be 20. As depicted in Figure 5.1, the mean of the observed time-to-event gives a biased estimate of the measure of location of the time-to-event, whereas, the medians of the actual and observed outcomes are equal. As a result, for a time-to-event outcome, distributional quantiles (such as median) may be more interpretable and are often used for description or comparison of time-to-event outcomes. Because of these characteristics (skewness and censoring) of time-to-event outcomes, standard statistical inference and estimation approaches are not usually appropriate. Non-parametric, semi-parametric and parametric methods have been developed to handle the characteristics of time-to-event data (Kalbfleisch and Prentice, 2002; Collett, 2003).

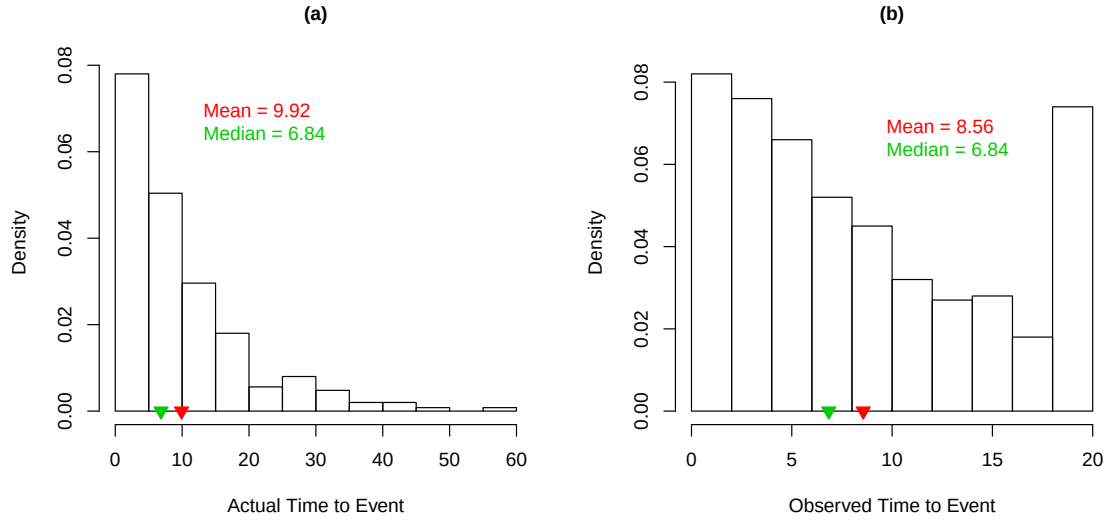

Figure 5.1: Histogram for (a) Actual time-to-event and (b) Observed time-to-event where there is administrative censoring after time point 20.

Suppose we have a population of N patients, we define T_i as the actual time that patient i, $(i = 1, \ldots, N)$, experienced the event of interest. If patient i is right censored, the right censoring time is represented by C_i. Since right censoring implies that the event of interest has not yet occurred and might occur at a later time, that is, $C_i < T_i$, the observed time-to-event is defined as $T_i^* = \min\{T_i, C_i\}$ which will be equal to T_i if patient i experiences the event and C_i if patient i is right censored. Finally, we define $\delta_i = \mathbb{1}(T_i < C_i)$ as the event indicator which takes value 1 if patient i experienced the event and 0 if patient i is right censored.

The outcome (time-to-event) is a random variable, and so, it has an associated probability density function $f(t)$ and cumulative distribution function $F(t) = \mathbb{P}(T < t) = \int_0^t f(u)\mathrm{d}u$. Similarly, the following functions of the time-to-event outcome can be defined

- Survivor function $S(t) = \mathbb{P}(T \geq t) = 1 - F(t)$, is the probability that a patient will experience the event after some time t.
- The hazard rate is the instantaneous rate of experiencing the event of interest and is defined as $h(t) = \lim_{\delta t \downarrow 0} \frac{1}{\delta t}\mathbb{P}(t \leq T < t + \delta t | T \geq t)$.
- The cumulative hazard is defined as $H(t) = \int_0^t h(u)\mathrm{d}u$.

The following relationships exist between these functions and are useful in estimation:

- $S(t) = \exp(-H(t))$
- $h(t) = -\frac{d}{dt}\log S(t)$
- $H(t) = -\log(S(t))$
- $f(t) = h(t)S(t)$.

In this chapter, we focus on treatment effect estimation where the outcome of interest is time-to-event and the treatment assignment can be linked to the value of a continuous assignment variable and a pre-defined threshold value as we have in an RD design.

For a sharp RD design, treatment effect estimation can be fairly straightforward and is usually done by comparing the outcomes between groups above and below the threshold. The comparison can be based on a hazard ratio or acceleration factor (or other appropriate measures) that are estimated from parametric or semi-parametric survival models (Collett, 2003). However, where there is only partial adherence to the rule, as we have in a fuzzy RD design, it becomes more complicated to estimate the treatment effect as method under a sharp design results will give an estimate that is analogous to the intention to treat estimate in a randomised controlled trial with non-compliance, which may be biased for the true treatment effect.

Treatment effect estimation in an RD design when the outcome is a time-to-event has not been widely examined in the literature. The first instance where we see this is in Bor et al. (2014), where the authors investigated the effect of anti-retro viral treatment (ART) on the survivor experience of HIV patients. Patients with CD4 count below 200μL are eligible for ART, therefore, in their study, CD4 count is the assignment variable with the threshold set to 200μL. They fitted a Cox proportional hazard model to compare patients just below the threshold and patients just above the threshold. The result of the analysis showed that patients just below the threshold have a lower hazard of death compared to patients just above the threshold. The hazard ratio estimate described above would be the treatment effect in a sharp RD design. Bor et al. (2014) provided an extension on how to estimate the hazard ratio for a fuzzy design, but the extension relies on the assumption that the event of interest is rare. Therefore, this method will not be applicable in all cases. Cho et al. (2019) proposed an approach to a estimate treatment effect in a fuzzy RD design for right censored data. In their work, censoring was handled using the censoring unbiased transformation approach which is not readily available in standard statistical software, and this makes the method difficult to implement in many clinical settings.

In this chapter, the aim is to develop a methodology for treatment effect estimation in a fuzzy RD design where the outcome is time-to-event under the accelerated failure time (AFT) assumption that can be implemented in standard statistical software. We propose a modification of the LATE estimator that is used to estimate the treatment effect for a continuous outcome (Imbens and Lemieux, 2007;

O'Keeffe and Baio, 2016) in a fuzzy RD design. In addition, we compare the proposed methodology to the structural models framework in survival analysis, which is a well established method for estimating a treatment effect in observational studies (Hernán and Robins, 2006; Didelez et al., 2010; Clarke and Windmeijer, 2010).

5.2 Assumptions

As mentioned in Chapter 2, when the outcome is a time-to-event, additional assumptions are necessary for the identification of a treatment effect in an RD design. Assumptions 1, 2, 4, 5 and 6 as stated in Section 2.2 are required with no modification. In addition to these five assumptions, we present these two assumptions.

T1 We present a modification of Assumption 3:

$$f(t|X = x, Z, A = a) \text{ is continuous at } x = x_0 \text{ for } a = 0, 1.$$

This modification of Assumption 3 is necessary because the interest here is no longer only the expectation of the outcome at the threshold, rather other properties of the time-to-event, such as hazard, median survival time etc are also of interest.

T2 Censoring is non-informative. We assume the reason for censoring is unrelated to the event of interest.

5.3 Accelerated failure time assumption

Below, we describe two common models that can be used for modelling a time-to-event outcome and explain why we take we take an accelerated failure time approach for the modelling of time-to-event outcomes in RD design.

Proportional hazards (PH) model: This is the most commonly used model for a time-to-event outcome. The PH model assumes that the hazard rate of the treated group is proportional to the hazard rate of the untreated group: $h(t|A = 1) = \theta\, h(t|A = 0)$,

where θ is the treatment effect known as the hazard ratio. The effect of the treatment is modelled through the hazard ratio, therefore, the proportional hazards model can be written as $h(t) = \exp(A\alpha)\, h_0(t)$, where $\theta = \exp(\alpha)$

Accelerated failure time (AFT) model: This model assumes that the time-to-event in the treated group is accelerated (or decelerated) by a constant compared to the time-to-event in the untreated group, that is,

$$S(t|A = 1) = S(\phi^{-1}t|A = 0),$$

with ϕ being the treatment effect and it is known as an acceleration factor. If the acceleration factor is greater than 1, it implies a beneficial effect of the treatment, that is the time-to-event in the treated group is higher than the time-to-event in the untreated group. Likewise, an acceleration factor value less than 1 implies that the time-to-event in the treated group is lower compared to the untreated group. The assumption of the AFT model can be rewritten as

$$\mathbb{P}(T > t|A = 1) = \mathbb{P}(\phi T > t|A = 0).$$

As such, the AFT assumption implies that T_1 and ϕT_0 have the same distribution, where T_a, $a \in \{0, 1\}$ represents the time-to-event for a patient that receives treatment a. Therefore, the times-to-event in the two groups can be expressed as $T_1 = \phi T_0$. As a result, a log-linear model can be fitted for the survival time,

$$T = \exp(A\beta)T_0,$$

and β expresses the effect of the treatment on the log of the survival time and the acceleration factor is $\phi = \exp(\beta)$.

It has been shown that the hazard ratio is a non-collapsible measure (Sjölander et al., 2016). We recall that a measure is non-collapsible if the marginal measure is not a weighted average of the conditional measures, which implies that conditioning on a covariate related to the outcome, even if it is unrelated to the treatment, changes the size of the hazard ratio. This can make it difficult to assess the performance of an estimator of the causal hazard ratio. However, the acceleration factor is a collapsible measure (see Remark 5.3.1 below) and as a result, we focus on the estimation of the

acceleration factor in a fuzzy RD design.

Remark 5.3.1. *The acceleration factor is a collapsible measure.*

Proof. We define a variable B that is associated with the outcome only, that is, B is not a confounder. The log linear model with A (the treatment indicator) and B as the covariates is given as

$$\log(T_i) = \beta_1 A_i + \beta_2 B_i + \epsilon, \quad i = 1, \dots n.$$

From the Equation above, $\exp(\beta_1)$ is the effect of A (the acceleration factor) on the outcome conditional B.

We now derive the marginal effect of A by marginalising over B. Without loss of generality, we assume B is a continuous random variable.

$$\begin{aligned}
T_i &= \exp(\beta_1 A_i + \beta_2 B_i + \epsilon) \\
\mathbb{E}(T_i|A_i, B_i) &= \exp(\beta_1 A_i + \beta_2 B_i)\, \mathbb{E}(\exp(\epsilon_i)) \\
\mathbb{E}(T_i|A_i) &= \int_b \exp(\beta_1 A_i + \beta_2 B_i)\, \mathbb{E}(\exp(\epsilon_i)) f_B(b_i) \mathrm{d}b_i \\
&= \exp(\beta_1 A_i)\, \mathbb{E}(\exp(\epsilon_i)) \int_b \exp(\beta_2 B_i) f_B(b_i) \mathrm{d}b_i \\
&= \exp(\beta_1 A_i)\, \mathbb{E}(\exp(\epsilon_i^*)),
\end{aligned}$$

where $\epsilon_i^* = \beta_2 B_i + \epsilon_i$ and because $A_i \perp\!\!\!\perp \epsilon_i^*$, $\exp(\beta_1)$ is the marginal effect of A on the outcome. Since we have shown that the marginal and conditional effects coincide, this implies that the acceleration factor is a collapsible measure. □

We focus on the estimation of the acceleration factor in an RD design and, we shall consider a Weibull parametric AFT model. In practice, other distributions such as a log-normal or log-logistic distribution could be fitted and an optimality criterion based on the model residuals, such as the R^2, might be used to select the distribution that best fits an observed set of data (Chan et al., 2018). Alternatively, semi parametric AFT models have been developed (Wei, 1992; Buckley and James, 1979) and these could be considered.

The AFT assumption can be represented using a log linear model as follows

$$\log(T_i) = \mathbf{x}_i^T \boldsymbol{\beta} + \epsilon_i. \tag{5.1}$$

Here, ϵ_i represents the log of the event time for patients with $\mathbf{x}_i = 0$, where $\mathbf{x}_i$ is a vector of covariates. The model defined in Equation 5.1 is generally used in semi-parametric estimation because, when the distribution of ϵ_i is not specified, only $\boldsymbol{\beta}$ can be estimated.

However, in parametric estimation, when the distribution of ϵ_i is specified, ϵ_i can be rewritten in terms of its location and scale parameters and as a linear function of another variable W_i; $\epsilon_i = \mu + \sigma W_i$. Here, μ and σ are the location and scale parameters of ϵ_i respectively. Therefore, Equation 5.1 can be re-written to include an intercept term, as presented in Equation 5.2 below:

$$\log(T_i) = \mu + \mathbf{x}_i^T \boldsymbol{\beta} + \sigma W_i \tag{5.2}$$

Estimation is usually done using a maximum likelihood approach and the likelihood function is given by

$$L(\beta, \sigma, \mu | t, \delta) = \prod_i f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}.$$

This can be expressed in terms of the density and survivor function of W as

$$L(\beta, \sigma, \mu | t, \delta) = \prod_i (\sigma t_i)^{-\delta_i} f_W(w_i)^{\delta_i} S_W(w_i)^{1-\delta_i}, \tag{5.3}$$

where $w_i = \{\log(t_i) - \mu - \mathbf{x}_i^T \boldsymbol{\beta}\}/\sigma$.

We will now describe how the likelihood function for the Weibull AFT model is derived.

5.3.1 Weibull Parametric AFT model

In this thesis, we consider the use of a parametric AFT Weibull model. As such, we re-parameterize the Weibull distribution so that the form of the model shown in Equation 5.2 may be applied.

The baseline hazard function for a time-to-event variable that follows a Weibull distribution with scale parameter λ and shape parameter k ($T \sim \text{Weibull}(\lambda, k)$) is given as: $h_0(t) = \lambda k t^{k-1}$. Using this, we can derive expressions for the corresponding cumulative hazard, survivor and density functions as follows

$$H_0(t) = \lambda t^k;$$
$$S_0(t) = \exp(-\lambda t^k);$$
$$f_0(t) = \lambda k t^{k-1} \exp(-\lambda t^k).$$

Now, we let $\epsilon = \log(t_0)$, where t_0 represents the time-to-event for patients with $\mathbf{x} = 0$, as defined in Equation 5.1, ϵ has a Gumbel distribution with density function, expectation and variance given as:

$$f_\epsilon(\epsilon) = \lambda k \exp\{-\lambda(e^{\epsilon k} - \epsilon k)\}, \quad -\infty < \epsilon < \infty$$
$$\mathbb{E}(\epsilon) = -\frac{\log(\lambda)}{k} - \frac{\gamma}{k},$$
$$\text{where } \gamma = \int_0^\infty \log(t)\exp(-t)\mathrm{d}t \text{ is the Euler-Mascheroni constant,}$$
$$\text{Var}(\epsilon) = \frac{\pi^2}{6k^2}.$$

Comparing the expressions for expectation and variance of ϵ from the equations above with the expectation and variance of the error term in Equation 5.1, we obtain:

$$\mathbb{E}(\epsilon) = \mu + \sigma\ \mathbb{E}(W),$$
$$\text{Var}(\epsilon) = \sigma^2\ \text{Var}(W).$$

We can express the parameters of Equation 5.2 in terms of the parameters of the Weibull distribution as:

$$\mu = -\frac{\log(\lambda)}{k}; \quad \sigma = k^{-1}.$$

W also has a Gumbel distribution with density and survivor functions as given below:

$$f_W(w) = \exp\{-(e^w - w)\}, \quad -\infty < w < \infty; \quad \mathbb{E}(W) = -\gamma,\ \text{Var}(W) = \frac{\pi^2}{6},$$
$$S_W(w) = \exp\{-e^w\}.$$

The likelihood function for the Weibull AFT model is then

$$L(\beta, \sigma, \mu) = \prod_i (\sigma t_i)^{-\delta_i} \exp\{-e^{w_i} + \delta_i w_i\}. \tag{5.4}$$

5.4 Estimators of Acceleration Factor in an RD design

We will now discuss methods to estimate the treatment effect under the AFT assumption in a fuzzy RD design. We shall use the assumptions of the RD design to derive an estimator of the acceleration factor.

5.4.1 RDD-AFT approach

In this section, we present the proposed method for the estimation of the acceleration factor in a fuzzy RD design. As we have discussed in Chapter 3, the local average treatment effect (LATE) is an estimator of the treatment effect in a fuzzy RD design when the outcome of interest is continuous (Angrist et al., 1996; O'Keeffe and Baio, 2016). It involves fitting regression models to data above and below the threshold, and estimating the difference in the expected values of the outcome for patients above and below the threshold observed at the threshold. The LATE is derived by scaling the estimate at the threshold by the difference in probability of receiving treatment above and below the threshold.

With the log linear representation of the AFT given in Equation 5.5 below, we can estimate the discontinuity at the threshold on the log scale of the outcome. That is, the difference in the expected values of the log of the outcome for patients above and below the threshold that is observed at the threshold.

$$\log(T_i) = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_i^c + \alpha_3 X_i^c Z_i + \sigma w_i. \tag{5.5}$$

α_3 can be set to zero if we believe the slopes above and below the threshold are equal. α_1 measures the difference in $\mathbb{E}\{\log(T_i)\}$ between patients above and below

the threshold that is observed at the threshold. If the distribution of W_i is specified, the parameters of the model in Equation 5.5 are estimated parametrically, otherwise semi parametric estimation can be carried out. In this thesis, a Weibull distribution is assumed for $\exp(W_i)$ and estimation is carried out by maximising the likelihood given in Equation 5.4.

Assuming the model is correct, and Assumptions 4 (threshold indicator is independent of confounders conditional on assignment variable) and T2 (censoring is uninformative) are satisfied, α_1 will be an unbiased estimator of the treatment effect in a sharp RD design. However, in a fuzzy RD design, this estimate needs to be adjusted, just like in the standard LATE estimate, in order to produce an unbiased estimator of the treatment effect as presented below.

$$\beta_{\text{RDD}} = \frac{\alpha_1}{\pi_1 - \pi_0}. \tag{5.6}$$

Where π_z is the probability of being assigned treatment when $Z = z$. Therefore, the acceleration factor that is reported is corresponding to $\exp(\beta_{\text{RDD}})$.

Proof. Consider the model:

$$\log(T_i) = \beta A_i + \epsilon_i. \tag{5.7}$$

β is the treatment effect of interest. Here, the expectation of ϵ is not necessarily zero as it represents the expected value of $\log(T_i)$ when treatment is not received.

Fitting the marginal model in Equation 5.7 might lead to a biased estimate of the treatment effect due to the potential effect of confounding variables. Therefore, we estimate treatment effect by exploiting the fact that we have partial information on how the treatment is allocated.

We can represent

$$\lim_{x \downarrow x_0} \mathbb{E}\{\log(T_i)|X_i = x\} = \lim_{x \downarrow x_0} \mathbb{E}\{\log(T_i)|X_i = x, Z_i = 1\} \text{ and}$$
$$\lim_{x \uparrow x_0} \mathbb{E}\{\log(T_i)|X_i = x\} = \lim_{x \uparrow x_0} \mathbb{E}\{\log(T_i)|X_i = x, Z_i = 0\}$$

For simplicity, we will drop the limits in subsequent derivations: $\mathbb{E}\{\log(T_i)|Z_i = 1\}$ will appear instead of $\lim_{x \downarrow x_0} \mathbb{E}\{\log(T_i)|X_i = x)\}$ and

$\mathbb{E}\{\log(T_i)|Z_i = 0\}$ instead of $\lim_{x\uparrow x_0} \mathbb{E}\{\log(T_i)|X_i = x\}$.

From Equation 5.7, it follows that

$$\begin{aligned}
&\mathbb{E}\{\log(T_i)|Z_i = 1\} - \mathbb{E}\{\log(T_i)|Z_i = 0\} \\
&\quad = \mathbb{E}\{\beta A_i + \epsilon_i|Z_i = 1\} - \mathbb{E}\{\beta A_i + \epsilon_i|Z_i = 0\} \\
&\quad = \beta[\mathbb{E}\{A_i|Z_i = 1\} - \mathbb{E}\{A_i|Z_i = 0\}] - [\mathbb{E}\{\epsilon_i|Z_i = 1\} - \mathbb{E}\{\epsilon_i|Z_i = 0\}]
\end{aligned}$$

Putting $\epsilon_i = \mathbb{E}\{\log(T_i)|A_i = 0)\}$ gives

$$\begin{aligned}
&= \beta[\mathbb{E}\{A_i|Z_i = 1\} - \mathbb{E}\{A_i|Z_i = 0\}] \\
&- [\mathbb{E}\{\log(T_i)|A_i = 0, Z_i = 1\} - \mathbb{E}\{\log(T_i)|A_i = 0, Z_i = 0\}]
\end{aligned}$$

Under the modified version of Assumption 3, that is, Assumption T1, we have that

$$\mathbb{E}\{\log(T_i)|A_i = 0, Z_i = 1\} - \mathbb{E}\{\log(T_i)|A_i = 0, Z_i = 0\} = 0$$

Therefore,

$$\mathbb{E}\{\log(T_i)|Z_i = 1\} - \mathbb{E}\{\log(T_i)|Z_i = 0\} = \beta[\mathbb{E}\{A_i|Z_i = 1\} - \mathbb{E}\{A_i|Z_i = 0\}]$$

By rearranging the equation above and adding the limits, we have

$$\beta = \frac{\lim_{x\downarrow x_0} \mathbb{E}\{\log(T_i)|X_i = x\} - \lim_{x\uparrow x_0} \mathbb{E}\{\log(T_i)|X_i = x\}}{\mathbb{E}\{A_i|Z_i = 1\} - \mathbb{E}\{A_i|Z_i = 0\}}.$$

The numerator term is equivalent to the estimate of α_1 in Equation 5.5 while the denominator term is the probability of compliance as presented in Equation 5.6.

□

In the next section, we discuss how the variance of the proposed estimator of the acceleration factor in a fuzzy RD design is estimated.

5.4.1.1 Variance Estimation

We have derived an estimator of the acceleration factor in an RD design, and we will now describe how the variance of the estimator may be estimated. To do this, we shall begin by deriving the variance of the log of the acceleration factor:

$$\beta_{\text{RDD}} = \frac{\alpha_1}{\gamma_1}.$$

The numerator is estimated from Equation 5.5 while the denominator represents the difference in the probability of receiving treatment for patients above and below the threshold, that is $\gamma_1 = \pi_1 - \pi_0$. The denominator can be estimated from the regression model in the equation given below

$$A_i = \gamma_0 + \gamma_1 Z_i + e_i, \quad e_i \sim \mathcal{N}(0, \sigma_e) \tag{5.8}$$

$\boldsymbol{\alpha} = (\alpha_0,\ \alpha_1,\ \alpha_2,\ \alpha_3)^\top$ and $\boldsymbol{\gamma} = (\gamma_0,\ \gamma_1)^\top$ are estimated by maximising the likelihoods of their corresponding models. Therefore, they are maximum likelihood estimators and they are asymptotically unbiased and normally distributed (Newey and McFadden, 1994). That is,

$$\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \boldsymbol{\alpha} \\ \\ \boldsymbol{\gamma} \end{pmatrix}, \begin{pmatrix} \Sigma_\alpha & \Sigma_{\alpha,\gamma} \\ \\ \Sigma_{\alpha,\gamma} & \Sigma_\gamma \end{pmatrix} \right],$$

where Σ_α and Σ_γ are the variance-covariance matrices of $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\gamma}}$ respectively and they are also estimated using the maximum likelihood approach. $\Sigma_{\alpha,\gamma}$ is the covariance matrix of $\hat{\alpha}$ and $\hat{\gamma}$, which will be estimated using a copula approach that will be described below.

To estimate the variance of $\hat{\beta}$, we apply the multivariate delta method. We are interested in estimating the variance of the ratio of two normally distributed variables:

$$\hat{\beta}_{\text{RDD}} = g(\alpha_1, \gamma_1) = \frac{\hat{\alpha_1}}{\hat{\gamma_1}}.$$

By applying the multivariate delta method, the variance of $\hat{\beta}$ is approximated as:

$$\begin{aligned} \text{Var}(\hat{\beta}_{\text{RDD}}) &\approx \Delta\left(g\left(\hat{\alpha}_1,\ \hat{\gamma}_1\right)\right)^\top \hat{\Sigma}_{\alpha_1,\ \gamma_1}\ \Delta\left(g\left(\hat{\alpha}_1, \hat{\gamma}_1\right)\right) \\ &= \begin{pmatrix} \frac{1}{\hat{\gamma}_1} \\ \\ -\frac{\hat{\alpha}_1}{\hat{\gamma}_1^2} \end{pmatrix}^\top \begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha,\gamma} \\ \\ \sigma_{\alpha,\gamma} & \sigma_\gamma^2 \end{pmatrix} \begin{pmatrix} \frac{1}{\hat{\gamma}_1} \\ \\ -\frac{\hat{\alpha}_1}{\hat{\gamma}_1^2} \end{pmatrix} \end{aligned} \tag{5.9}$$

Where $\sigma_\alpha^2 = \Sigma_{\alpha[2,2]}$, $\sigma_\gamma^2 = \Sigma_{\gamma[2,2]}$ and $\sigma_{\alpha,\gamma} = \Sigma_{\alpha,\gamma[1,2]}$.

Now, we discuss the copula approach that we used to estimate $\Sigma_{\alpha,\gamma}$. A copula is

defined as a cumulative distribution function whose marginals are uniformly distributed on [0, 1]. That is, given n random variables $U_1, \ldots U_n$ that are uniformly distributed; $U_i \sim \text{Uniform}(0, 1)$, then a copula $\tilde{C}$ is defined as:

$$\tilde{C} = \mathbb{P}(U_1 \leq u_i \ldots U_n \leq u_n).$$

Below, we state the Sklar's theorem that has been attributed to copulas being widely used to model the dependence structure between two or more variables (Sklar, 1959; Nelsen, 2006).

Sklar's theorem: *Let H be a two-dimensional distribution function with marginal distribution functions F and G. Then there exists a copula $\tilde{C}$ such that $H(x, y) = \tilde{C}(F(x), G(y))$. Conversely, for any distribution functions F and G and any copula $\tilde{C}$, the function H defined above is a two-dimensional distribution function with marginals F and G. Furthermore, if F and G are continuous, $\tilde{C}$ is unique.*

According to the Sklar's theorem, the joint cumulative distribution function of a set of random variables can be expressed in terms of the marginal distribution functions and a copula $\tilde{C}$ (Nelsen, 2006). For example, in the case we are considering in this chapter, the joint cumulative function of T_i and A_i ($F(t_i,\ a_i)$) can be expressed as

$$F(t_i,\ a_i) = \tilde{C}(F_1(t_i), F_2(a_i)),$$

where $F_1(t_i)$ and $F_2(a_i)$ are the marginal distribution functions of T_i and A_i respectively. The marginals can be modelled separately via $F_1(t_i)$ and $F_2(a_i)$ while the information on the dependence of T_i and A_i is contained in the copula $\tilde{C}$.

In survival analysis, rather than using the cumulative function, modelling is generally done using the survivor function. Sklar's theorem can also be extended to the survivor function, that is, the joint survivor function can be expressed in terms of the marginal survivor functions and a copula $\mathbb{C}$ (Marra and Radice, 2020):

$$S(t_i, a_i) = \mathbb{C}\left(S_1(t_i), S_2(a_i)\right),$$

where $S(t_i,\ a_i)$ is the joint survivor function of T_i and A_i, $S_1(t_i) = 1 - F_1(t_i)$ and $S_2(a_i) = 1 - F_2(a_i)$ are the marginal survivor functions of T_i and A_i respectively.

In this thesis, we use the Gaussian copula function because it is a commonly used copula and it is easily constructed (Meyer, 2013). The Gaussian copula is given as:

$$\mathbb{C}\left(S_1(t_i), S_2(a_i)\right) = \Phi_R\left(\Phi^{-1}(u_{1i})\Phi^{-1}(u_{2i})\right),$$

where $u_{1i} = S_1(t_i)$, $u_{2i} = S_2(a_i)$, Φ^{-1} is the inverse cumulative distribution function of a standard normal distribution and Φ_R is the joint cumulative distribution function of a bivariate normal distribution with mean vector equal to zero and covariance matrix equal to the correlation matrix R and

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

We will proceed to discuss how the likelihood function is derived, the joint density function of the time-to-event outcome and the treatment indicator is given as:

$$f(t_i,\ a_i) = \frac{\partial^2 F(t_i, a_i)}{\partial t_i \partial a_i}$$

For an uncensored patient,

$$\begin{aligned}
F(t_i,\ a_i) &\propto \mathbb{P}(T_i \leq t_i, A_i \leq a_i) \\
&= \int_0^{t_i}\int_0^{a_i} f(x_i, y_i)dx_i dy_i \\
&= \int_0^{t_i}\left[\int_0^{\infty} f(x_i, y_i)dy_i - \int_{a_i}^{\infty} f(x_i, y_i)dy_i\right]dxi \\
&= \int_0^{t_i}\left[f_1(x_i) - \int_{a_i}^{\infty} f(x_i, y_i)dy_i\right]dx_i \\
&= \int_0^{t_i} f_1(x_i)dx_i - \left[\int_0^{\infty}\int_{a_i}^{\infty} f(x_i, y_i)dy_i dx_i - \int_{t_i}^{\infty}\int_{a_i}^{\infty} f(x_i, y_i)dy_i dx_i\right] \\
&= 1 - \int_{t_i}^{\infty} f_1(x_i)dx_i - \int_{a_i}^{\infty} f(y_i)dy_i + \int_{t_i}^{\infty}\int_{a_i}^{\infty} f(x_i, y_i)dy_i dx_i \\
&= 1 - S_1(t_i) - S_2(a_1) + S(a_i, t_i) \\
&= 1 - S_1(t_i) - S_2(a_i) + \mathbb{C}\left(S_1(t_i), S_2(a_i)\right).
\end{aligned}$$

Therefore, the joint density function for an uncensored patient is

$$f(t_i,\ a_i) \propto \frac{\partial^2 \mathbb{C}\left(S_1(t_i), S_2(a_i)\right)}{\partial t_i \partial a_i}$$

For a censored patient, we let $g(c_i)$ represent the density function of the censoring time, then,

$$\begin{aligned}
F(t_i,\ a_i) &\propto \mathbb{P}(C_i \leq t_i, A_i \leq a_i, T_i \geq C_i) \\
&= \int_0^{t_i} \int_0^{a_i} \int_{c_i}^{\infty} f(x_i, y_i) g(c_i) dx_i dy_i dc_i \\
&= \int_0^{t_i} \left[\int_0^{\infty} \int_{c_i}^{\infty} f(x_i, y_i) dx_i dy_i - \int_{a_i}^{\infty} \int_{c_i}^{\infty} f(x_i, y_i) dx_i dy_i \right] g(c_i) dc_i \\
&= \int_0^{t_i} \left[\int_{c_i}^{\infty} f(x_i) dx_i - \int_{a_i}^{\infty} \int_{c_i}^{\infty} f(x_i, y_i) dx_i dy_i \right] g(c_i) dc_i \\
&= \int_0^{t_i} \Big[S_1(c_i) - (S(c_i, a_i) \Big] g(c_i) dc_i \\
&= \int_0^{t_i} \Big[S_1(c_i) - \mathbb{C}\left(S_1(c_i), S_2(a_i)\right) \Big] g(c_i) dc_i.
\end{aligned}$$

Therefore, the joint density function for a censored patient is

$$f(t_i,\ a_i) \propto -\frac{\partial \mathbb{C}\left(S_1(t_i), S_2(a_i)\right) g(t_i)}{\partial a_i}$$

As such, the log likelihood function is derived based on the contributions to the likelihood of the censored and uncensored patients, and it is given as

$$\begin{aligned}
\ell &= \sum_{i=1}^{n} \left\{ \delta_i \log \left[\frac{\partial^2 \mathbb{C}(u_{1i}, u_{2i})}{\partial t_i \partial a_i} \right] + (1 - \delta_i) \log \left[-\frac{\partial \mathbb{C}(u_{1i}, u_{2i})}{\partial a_i} \right] \right\} \\
&= \sum_{i=1}^{n} \left\{ \delta_i \log \left[\frac{\partial^2 \mathbb{C}(u_{1i}, u_{2i})}{\partial u_{1i} \partial u_{2i}} f_1(t_i) f_2(a_1) \right] + (1 - \delta_i) \log \left[\frac{\partial \mathbb{C}(u_{1i}, u_{2i})}{\partial u_{2i}} f_2(a_i) \right] \right\}
\end{aligned} \tag{5.10}$$

where $u_{1i} = S_1(t_i)$ and $u_{2i} = S_2(a_i)$.

The components of the log likelihood function in Equation 5.10 are given below:

$$\frac{\partial^2 \mathbb{C}(u_{1i}, u_{2i})}{\partial u_{1i} \partial u_{2i}} = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{ -\frac{\left(\rho\Phi^{-1}(u_{1i})\right)^2 + \left(\rho\Phi^{-1}(u_{2i})\right)^2 - 2\rho\Phi^{-1}(u_{1i})\Phi^{-1}(u_{2i})}{2(1-\rho^2)} \right\},$$

$$\frac{\partial \mathbb{C}(u_{1i}, u_{2i})}{\partial u_{2i}} = \Phi\left(\frac{\Phi^{-1}(u_{1i}) - \rho\Phi^{-1}(u_{2i})}{\sqrt{1-\rho^2}} \right),$$

$$f_1(t_i) = \frac{1}{\sigma t_i} \exp\{-\exp\{w_i\} + w_i\},$$

$$f_2(a_i) = \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\left\{ -\frac{e_i^2}{2} \right\},$$

$$u_{1i} = S_1(t_i) = \exp\{-\exp\{w_i\}\},$$

$$u_{2i} = S_2(a_i) = \Phi(-e_i),$$

$$w_i = \frac{\log(t_i) - \alpha_0 - \alpha_1 z_i - \alpha_2 x_i - \alpha_3 x_i z_i}{\sigma},$$

$$e_i = \frac{a_i - \gamma_0 - \gamma_1 z_i}{\sigma_e},$$

where $\Phi(.)$ is the cumulative function of the standard normal distribution and $\Phi^{-1}(.)$ is the inverse of the cumulative function of the standard normal distribution.

The parameters to be estimated are $\boldsymbol{\alpha} = (\alpha_0,\ \alpha_1,\ \alpha_2,\ \alpha_3)^\top$, σ, $\boldsymbol{\gamma} = (\gamma_0,\ \gamma_1)^\top$, σ_e and ρ. We estimate $\boldsymbol{\alpha} = (\alpha_0,\ \alpha_1,\ \alpha_2,\ \alpha_3)^\top$ and σ from the marginal model in Equation 5.5 and $\boldsymbol{\gamma} = (\gamma_0,\ \gamma_1)^\top$ and σ_e are estimated from the marginal model in Equation 5.8 then their estimated values are substituted into the log likelihood function in Equation 5.10, which is maximised in order to estimate ρ.

The variance covariance matrix can be approximated from the Hessian matrix which is equal to the negative observed information ($\mathcal{H} = -\mathcal{I}$). The expression for the Hessian is presented below. The Hessian is calculated using the estimated values of the parameters.

$$\mathcal{H} = \begin{bmatrix} \frac{\partial^2 \ell}{\partial \boldsymbol{\alpha}^2} & \frac{\partial^2 \ell}{\partial \boldsymbol{\alpha} \partial \sigma} & \frac{\partial^2 \ell}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\gamma}} & \frac{\partial^2 \ell}{\partial \boldsymbol{\alpha} \sigma_e} & \frac{\partial^2 \ell}{\partial \boldsymbol{\alpha} \partial \rho} \\ \frac{\partial^2 \ell}{\partial \sigma \partial \boldsymbol{\alpha}} & \frac{\partial^2 \ell}{\partial \sigma^2} & \frac{\partial^2 \ell}{\partial \sigma \partial \boldsymbol{\gamma}} & \frac{\partial^2 \ell}{\partial \sigma \partial \sigma_e} & \frac{\partial^2 \ell}{\partial \sigma \partial \rho} \\ \frac{\partial^2 \ell}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\alpha}} & \frac{\partial^2 \ell}{\partial \boldsymbol{\gamma} \partial \sigma} & \frac{\partial^2 \ell}{\partial \boldsymbol{\gamma}^2} & \frac{\partial^2 \ell}{\partial \boldsymbol{\gamma} \partial \sigma_e} & \frac{\partial^2 \ell}{\partial \boldsymbol{\gamma} \partial \rho} \\ \frac{\partial^2 \ell}{\partial \sigma_e \partial \boldsymbol{\alpha}} & \frac{\partial^2 \ell}{\partial \sigma_e \partial \sigma} & \frac{\partial^2 \ell}{\partial \sigma_e \partial \boldsymbol{\gamma}} & \frac{\partial^2 \ell}{\partial \sigma_e^2} & \frac{\partial^2 \ell}{\partial \sigma_e \partial \rho} \\ \frac{\partial^2 \ell}{\partial \rho \partial \boldsymbol{\alpha}} & \frac{\partial^2 \ell}{\partial \rho \partial \sigma} & \frac{\partial^2 \ell}{\partial \rho \partial \boldsymbol{\gamma}} & \frac{\partial^2 \ell}{\partial \rho \partial \sigma_e} & \frac{\partial^2 \ell}{\partial \rho^2} \end{bmatrix}$$

The variance covariance matrix can be approximated as (Collett, 2003):

$$\Sigma = \text{Cov}(\hat{\boldsymbol{\alpha}}, \hat{\sigma}, \hat{\boldsymbol{\gamma}}, \hat{\sigma_e}, \hat{\rho}) = -\mathcal{H}^{-1}.$$

From the estimated variance-covariance matrix, we can extract the required covariance of α_1 and γ_1 as required to estimate the variance of the estimator of the treatment effect as given in Equation 5.9. Alternatively, the variance of RDD-AFT can be estimated using a bootstrap approach. However, the copula approach described above is not as time consuming as a bootstrap approach.

We have described the proposed estimator for the acceleration factor, based on the assumptions of the RD design. In the next section, the structural AFT estimator will be discussed. This method was developed to estimate a treatment effect (acceleration factor) in observational studies. It is not an RD design approach so it does not depend on the assumptions of an RD design, we shall describe the assumptions of this method separately.

5.4.2 Structural AFT model

Structural models have been well established in causal effect estimation when treatment is not randomly assigned (Hernán and Robins, 2006; Didelez et al., 2010; Clarke and Windmeijer, 2010). They have been applied to causal effect estimation with continuous (Hernán and Robins, 2006), binary (Clarke and Windmeijer, 2010; Geneletti et al., 2019) and time-to-event outcomes (Hernán et al., 2005; Martinussen

et al., 2017). We recall the definition of counterfactual outcomes and how they are used to form structural models. The observed time-to-event for patient i is given as T_i, then we define T_i^a as the potential outcome that would be observed for patient i if treatment a is received. Therefore, T_i^1 and T_i^0 represent the potential/ counterfactual outcomes for patient i when he/she receives the treatment and when treatment is not received respectively.

Under the AFT assumption, relationship between the counterfactual outcomes can be expressed as:

$$T_i^1 = \exp(\beta) T_i^0. \tag{5.11}$$

Models that are defined in terms of the counterfactual outcomes are called structural models. Therefore, Equation 5.11 is referred to as a structural AFT (S-AFT) model. The ratio between the counterfactual outcomes, $\exp(\beta)$, is the treatment effect (that is, the acceleration factor).

In order to obtain a valid estimate of the treatment effect, the following assumptions must be satisfied.

S1 There are no unobserved confounders or treatment allocation is strongly ignorable conditional on observed variables:

$$T_i^1,\ T_i^0 \perp\!\!\!\perp A_i \mid \mathcal{O}_i.$$

This assumption implies that the potential time-to-event outcome for patient i does not depend on the treatment they actually receive conditional on the observed variables.

S2 The probability of receiving treatment is greater than 0 conditional on observed variables, that is,

$$\mathbb{P}(A_i = 1 | \mathcal{O}_i) > 0 \text{ for all } i.$$

The S-AFT model is not strictly an RD design approach. Instead, this method compares the treated and untreated groups while correcting for the effect of confounding using the observed confounders and relying on the assumptions presented above. In contrast, the RDD-AFT method estimates treatment at the threshold by comparing

the patients above and below the threshold and adjusting the result of this comparison with the probability of compliance to the treatment guideline. The S-AFT model would be applicable for an RD design, if there is no unmeasured confounding in the region around the threshold or, alternatively, if any confounding variables present in a region around the threshold were accounted-for in the S-AFT model. Where the assumptions mentioned above are satisfied, β can be estimated using G-estimation or inverse probability treatment weighting (IPTW) (Hernán et al., 2005). The two methods will be discussed below.

G-estimation: First, we consider estimation of β through G-estimation when there is no censoring. The counterfactual outcome of not receiving the treatment is estimated from the relation in Equation 5.11 as

$$T_i^0 = \exp(-\beta)T_i^1.$$

Under the consistency assumption, that is, $T_i^1 = T_i$ if patient i is treated and $T_i^0 = T_i$ if patient i is untreated, we can rewrite the equation above as

$$T_i^0 = \exp(-\beta A_i)T_i.$$

This ensures that $T_i^0 = T_i$ if a patient is untreated and $T_i^0 = \exp(-\beta)T_i$ for treated patients. We consider T^0 as a function of β and we write it as $T^0(\beta)$.

A logistic regression model with the treatment indicator as response and observed confounders and the estimated counterfactual outcome as predictors is given by:

$$\text{logit}\{\mathbb{P}(A_i = 1|\mathcal{O}_i, T_i^0(\beta^*))\} = \theta_0 + \theta_1\mathcal{O}_i + \theta_2 T_i^0(\beta^*).$$

When Assumption S1 holds, $T^0(\beta)$ is independent of treatment conditional on $\mathcal{O}$, then $\theta_2 = 0$. Therefore, the value of β^* that leads to the failure to reject the null hypothesis that $\theta_2 = 0$ is the g-estimate of β. The g-estimate can be obtained by minimising the score test statistic for $\theta_2 = 0$ which is equivalent to finding the solution to the estimating equation below (Hernán et al., 2005):

$$U(\beta^*) = 0,$$

where

$$U(\beta^*) = \sum_i T_i^0(\beta^*)\{A_i - \mathbb{P}(A_i = 1|\mathcal{O}_i)\} \tag{5.12}$$

Hence the value of β^* that solves the estimating equation above is reported as the g-estimate of β. Below, we describe the algorithm for obtaining the g-estimate of β.

1. Fit a logistic regression model for the treatment, conditional on the observed confounders:

$$\text{logit}\{\mathbb{P}(A_i = 1|\mathcal{O}_i)\} = \theta_0 + \theta_1\mathcal{O}_i.$$

2. Obtain the predicted probabilities ($\mathbb{P}(A_i = 1|\mathcal{O}_i)$) from the fitted logistic regression above.

3. Solve the estimating equation below the obtain the g-estimate of β.

$$U(\beta^*) = 0,$$

   where $U(\beta^*)$ is as given in Equation 5.12. The `uniroot` function in `R` (R Core Team, 2018) can be used to solve the estimating equation.

Inverse Probability weighting: The AFT model assumes that the time-to-event in the treated group is accelerated (or decelerated) compared to the time-to-event in the untreated group by a constant: $T(1) = \phi T(0)$, where $T(a)$ is the survival time of the group that receives treatment a. Under this assumption, a log-linear model can be fitted for the survival time. The log-linear model is generally of the form:

$$\log T_i = \beta A_i + \epsilon_i \tag{5.13}$$

ϵ_i is the log of the event time for patients with $A_i = 0$ and β is the coefficient of A and $\phi = \exp(\beta)$.

The idea behind inverse probability weighting is to create a pseudo population that represents a population where treatment is randomly allocated as we would have in a randomised trial. In order to do this, subjects are weighted according to their probability of getting treated. The weight is defined as:

$$W_i^I = \begin{cases} [\mathbb{P}(A_i = 1|\mathcal{O}_i)]^{-1} & \text{for patients who received treatment} \\ [\mathbb{P}(A_i = 0|\mathcal{O}_i)]^{-1} & \text{for patients who did not receive treatment} \end{cases}$$

These weights are then used to fit a weighted log-linear regression model presented in Equation 5.13. This is achieved by including the weights (W_i^I) in the likelihood function in Equation 5.3:

$$L(\beta, \sigma, \mu) = \prod W^I (\sigma t)^{-\delta} f_W(w)^{\delta} S_W(w)^{1-\delta}.$$

In this thesis, we will estimate treatment effect using g-estimation approach owing to the fact for some simulation samples, the weighted log-linear model for the inverse probability weighting approach run in `R` did not converge.

Accommodating censoring in the estimation methods

Censoring of a time-to-event outcome may introduce bias when estimating a treatment effect because censored observations provide incomplete information about the actual time-to-event. Standard analysis of time-to-event outcomes has been used to handle censoring through the likelihood (or partial likelihood) as seen the Section 5.3. However, for the g-estimation procedure, the approach described earlier has not accounted for censoring. To handle censoring using g-estimation, we need to specify the type of censoring; whether the patient dropped out during the study or the patient did not experience the event by the end of the study (administrative censoring).

When censoring occurs because of drop out, inverse probability weighting for censoring is employed. The purpose is to create a pseudo population that represents a situation where no drop out is observed. Patients that dropped out will be assigned a weight of zero while patients who remain in the study are weighted based on their similarities to patients who dropped out. We define an indicator variable D_i that takes value 0 if patient i drops out and 1 if otherwise. The inverse probability weight for censoring is calculated as

$$W_i^D = \frac{D_i}{\mathbb{P}(D_i = 1|X_i, A_i)}.$$

This weight is then included in the score function in Equation 5.12.

For administrative censoring, $T_i^0(\beta^*)$ in Equation 5.12 is replaced with an indicator variable $\Delta_i(\beta^*)$, a function of $T_i^0(\beta^*)$. The idea is to assign a score of zero to patients that are administratively censored and those that would have been administratively censored. As a result, some patients who experienced the event would also have a score of zero. The intuition for this is explained further in Hernán et al. (2005). Let K specify the end of the study such that any patient with $T_i^a \geq K$ would be administratively censored. Define $K(\beta^*)$, the maximum observable time-to-event to event as

$$K(\beta^*) = \begin{cases} K & \text{if } \beta^* \leq 0 \\ K\exp(-\beta^*) & \text{if } \beta^* > 0 \end{cases}$$

Therefore, $\Delta_i(\beta^*) = \mathbb{1}(T_i^0(\beta^*) \geq K(\beta^*))$.

We will report β^* that solves the estimating equation

$$U(\beta^*) = 0,$$

where

$$U(\beta^*) = \sum_i W_i^D \Delta_i(\beta^*)\{A_i - \mathbb{P}(A_i = 1|\mathcal{O}_i)\}. \tag{5.14}$$

As such, the estimating equation in Step 3 of the algorithm for obtaining the g-estimate of β is replaced by Equation 5.14.

We have described the proposed RDD-AFT approach and S-AFT approach as methods for estimating the acceleration factor in a fuzzy RD design. In the next session, we shall carry out simulation studies to evaluate these methods.

5.5 Simulation Studies

Simulation studies were conducted to compare the performance of S-AFT and the RDD-AFT approaches. Data were simulated to represent fuzzy RD designs with varying levels of confounding and different levels of fuzziness. Details of the simulation process and results obtained are provided below.

5.5.1 Description of simulation study

Below, we describe the steps used in the simulation of the data sets to compare the RDD-AFT and S-AFT approaches for estimating the acceleration factor in an RD design. For each simulated dataset, a sample size of 2000 subjects was simulated and a total of 1000 datasets were created for each scenario.

Step 1: For each subject, an assignment variable, X_i, is simulated from a continuous uniform distribution.

$$X_i \sim \text{Uniform}(0,1) \quad i = 1, ..., 2000.$$

Step 2: We set the threshold to be equal to 0.5 and define the centred assignment variable, X_i^C, and threshold indicators, Z_i, as

$$X_i^C = X_i - 0.5;$$
$$Z_i = \mathbb{1}\{X_i^c \geq 0\}.$$

Step 3: A confounding variable, U_i, is simulated from a standard normal distribution

$$U_i \sim \mathcal{N}(0,1).$$

Step 4: The probability that the i^{th} subject receives treatment is given by

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 Z_i + \beta_2 X_i^C + \beta_3 U_i. \tag{5.15}$$

The parameters of this model are specified to reflect the level of fuzziness of the design and the level of confounding with regard to treatment allocation. The lower the value of β_1, the more fuzzy the design, β_3 reflects the correlation between the treatment indicator and the confounder, if $\beta_3 = 0$, this indicates there is no relationship between the treatment and the confounding variable.

Step 5: The treatment indicator, A_i, is simulated as follows

$$A_i \sim \text{Bernoulli}(p_i).$$

Step 6: The 'true' time-to-event is simulated as

$$\log T_i = \beta_4 A_i + \beta_5 U_i + \epsilon_i. \tag{5.16}$$

Here, β_4 defines the treatment effect (on the log-scale), β_5 reflects the correlation between the time-to-event outcome and the confounder U_i and ϵ_i is the logarithm of the true time-to-event when $A_i = 0$ and $U_i = 0$.

Step 7: To incorporate right censoring, we specify a probability of drop-out, p^c, and an 'end-of-study' time K. For $U_i^C \sim \text{Uniform}(0, 1)$, the censoring time ($C_i$) is defined as follows

$$C_i = \begin{cases} K & \text{if } U_i^C > p^c \\ \sim \text{Uniform}(0, K) & \text{if } U_i^C \leq p^c. \end{cases}$$

Step 8: The observed time-to-event and event indicators are defined as follows

$$T_i^* = \min(T_i, C_i),$$
$$\delta_i = \mathbb{1}\{T_i < C_i\}.$$

Step 9: Steps 1–8 are repeated $N = 2000$ times, to create a dataset with 2000 subjects.

Step 10: Steps 1–9 are repeated $M = 1000$ times and $M = 1000$ datasets, each with $N = 2000$ subjects, are obtained.

The values of some of the parameters in the simulation steps above are explained here. We set $\beta_0 = -2$, which makes the probability of receiving treatment at the threshold for patients below the threshold equal to 0.2 for $U_i = 0$ and $\beta_2 = 2$, which establishes a (positive) relationship between the assignment variable and the treatment indicator. We set $\beta_4 = \log(1.5)$ so that the acceleration factor is 1.5 and the median survival time was chosen to be 7 years for untreated patients when $U_i = 0$, as such, we set

$$\exp(\epsilon_i) \sim \text{Weibull}\left(\frac{\log(2)}{7^2}, 2\right).$$

Details of the remaining parameters that reflects the level of fuzziness and effects of

confounding are given below.

Simulation Scenarios

As we had under the binary outcome, the probability of compliance and level of confounding were varied by specifying the values of β_1 and β_3 in Equation 5.15 and β_5 in Equation 5.16. Table 5.1 gives the values of these parameters for the six scenarios considered.

For each of 1000 simulated datasets, we computed the correlation coefficient between the confounder (U) and the treatment indicator and between U and the observed time-to-event. The averages of the estimated correlation coefficients are reported in Table 5.1 as $\rho_{A,U}$ and $\rho_{T,U}$ respectively. The values of the correlation coefficients provide an insight into the level of (linear) association between the confounder and the treatment and the confounder and the outcome. For scenarios where there are no unobserved confounders, the estimate of the correlation coefficient between the outcome and the confounder is zero. For scenarios where the effect of the confounder is low, the values of the correlation coefficient are slightly larger than when there are no confounders. The correlations become higher when the effect of the unobserved confounder is high.

Likewise, we calculated the probability of compliance for the 1000 simulated datasets in each scenario. The average of these probabilities is reported as P.C. in Table 5.1. The level of fuzziness of the design is defined based on the probability of compliance, if there is a high compliance to the treatment guideline, the fuzziness is weak and fuzziness is strong where the probability of compliance is low. The probability of compliance for the scenarios with weak fuzziness is between 80% and 90%. For the strong fuzziness scenarios, the probability of compliance to the treatment guideline is between 54% and 57%.

5.5.2 Results of simulation studies

The RDD-AFT and S-AFT approaches were applied to the data simulated to estimate the treatment effect under varying levels of confounding and fuzziness. Additionally, since the RDD-AFT method exploits "exchangeability" of subjects close to

Table 5.1: Values of parameters in Equations 5.15 and 5.16 for the simulation scenarios with the corresponding probability of compliance (P.C.) and estimates of correlation coefficients between T^* and U ($\rho_{T,U}$) and A and U ($\rho_{A,U}$).

Scenario	Parameters			P.C.	$\rho_{T,U}$	$\rho_{A,U}$
Weak fuzziness, No Confounding	β_1=10	β_3=0	β_5=0	0.90	0.00	0.00
Weak fuzziness, Low Confounding	β_1=10	β_3=-1	β_5=0.3	0.87	0.16	-0.10
Weak fuzziness, High Confounding	β_1=10	β_3=2	β_5=0.7	0.79	0.40	0.22
Strong fuzziness, No Confounding	β_1=2.5	β_3=0	β_5=0	0.57	0.00	0.00
Strong fuzziness, Low Confounding	β_1=2.5	β_3=-0.5	β_5=0.3	0.55	0.16	-0.16
Strong fuzziness, High Confounding	β_1=2.5	β_3=0.7	β_5=0.7	0.54	0.41	0.22

the threshold, the closeness to the threshold is measured based on the bandwidth h. Patients whose assignment variable falls in the range $[X^c - h, X^c + h]$, where $X^c = X - x_0$ is the centered assignment variable, are included in the analysis. The bandwidth was varied to check the sensitivity of estimates to different bandwidth sizes.

For the S-AFT estimator, the data were considered to be from an observational study where we observe the time-to-event (T), event indicator (δ), treatment indicator (A) and assignment variable (X) and X was considered as the only observed confounder. We note that the S-AFT estimator is not an RD design approach, hence it does not rely on the information about the treatment guideline. As such one of its assumptions is that confounders are observed. In the simulation scenarios where effect of confounding is present, since the confounding variable is known, we have included an additional estimator called S-AFT adj. This is an S-AFT estimator but the confounding variable treated as an observed variable. This may give further give insight into how much the presence of unobserved confounders affects the S-AFT estimator.

In all cases, the probability of being censored because of drop out is set to 15% and we assume that the study ended after 10 years. Therefore, patients with a time-to-event greater than 10 years are also censored. Overall, the proportion of censorship is about 55% owing to drop out and loss to follow-up. The acceleration factor for all the scenarios is set to be 1.50.

Tables 5.2, 5.3 and 5.4 show numerical summaries of the logarithm of the acceleration

factor obtained from the simulation studies under the no, low and high confounding scenarios, respectively. Each table contains results under weak and strong fuzziness. Visual representations of the estimates of the acceleration factor from the simulation studies are also provided. Figures 5.2 and 5.3 show boxplots of the estimates of the acceleration factor for the weak and strong fuzziness scenarios, respectively.

As seen in Table 5.2, where there are no unobserved confounders, both the RDD-AFT and S-AFT approaches produce unbiased estimates of the treatment effect for both the weak and strong fuzziness cases and the coverage for the two methods is close to the nominal level of 95% in each case. We observe that the ESE and ASE are mostly similar to each other, which implies that the ASE is not over or under-estimating the variability of the estimators. We also observe that as the level of fuzziness increases, the uncertainty in the estimate of the RDD-AFT approach increases. This is expected because the RDD-AFT estimator is sensitive to the level of compliance to the treatment assignment rule.

When the effect of the unobserved confounding is low, as shown in Table 5.3, some bias is noticed in the estimate of the S-AFT approach. The bias observed in the S-AFT approach has translated in it having coverage that is quite low compared to the nominal level of 95%. This implies that, under this scenario, the 95% confidence interval of the S-AFT approach does not adequately cover the value of the true treatment effect.

On the other hand, the RDD-AFT produces estimates that are close to the value of the true treatment effect under both the weak and strong fuzziness scenarios. In terms of coverage, the RDD-AFT approach is able to consistently produce confidence intervals that cover the treatment effect under the weak fuzziness and strong fuzziness scenarios as its coverage is close to the 95% nominal value. As we have observed when there are no unobserved confounders, the RDD-AFT approach is more accurate under the weak fuzziness than the strong fuzziness scenario and the ESE and ASE are smaller under the weak fuzziness scenario.

The bias observed in the S-AFT estimator becomes even larger when the effect of confounding is high, shown in Table 5.4, as a result, the coverages of the value of the treatment effect by the associated confidence intervals are quite low. In contrast, the

RDD-AFT approach continues to produce estimates that are close to the treatment effect and yields coverage levels that are close to 95%. In all the scenarios considered, we observe that the variance of the RDD-AFT estimator increases when the level of fuzziness is strong. This expected because it reflects the uncertainty of the estimate of the treatment effect when the treatment allocation is not strongly associated with the threshold indicator.

In the simulation above, we observe that the S-AFT approach produces biased estimates in the low and high confounding scenarios. This is not unexpected because an important assumption of the S-AFT approach is that confounders are observed. In the simulation study, since the confounder is known, we can adjust for the confounder in the S-AFT model. In Tables 5.3 and 5.4, an additional estimator called S-AFT adj. is included which is the S-AFT estimator but with the confounder adjusted for. As expected, this estimator does not suffer from the bias that was observed with the unadjusted S-AFT estimator.

Overall, when there are no unobserved confounders, both the S-AFT and RDD-AFT estimators produce unbiased estimates and they both have good coverage of the value of the true treatment effect. However, the S-AFT approach becomes biased and its coverage level is low as the effect of confounding increases. The RDD-AFT approach also exhibits some bias as the effect of confounding increases and level of fuzziness increases, but the bias in the RDD-AFT approach is lower than that of the S-AFT approach. In addition, the coverage of the RDD-AFT approach in most cases is around 95%.

Hence, we conclude that the RDD-AFT approach may be more desirable than the S-AFT estimator and that the RDD-AFT approach should be preferred when we have (partial) information about how the treatment assigned.

We have proposed a method for estimating the treatment effect when the outcome of interest is time-to-effect for an RD design. In addition, the structural AFT method was discussed as the popular approach of estimating treatment effect in observational studies when the outcome of interest is time-to-event under the AFT assumption. These two methods were compared using simulation studies under varying levels of compliance to the treatment guideline and confounding. Next, we shall apply the

two methods to real dataset on the effect of metformin prescription on all-cause mortality and time to a cardiovascular event in patients at risk of type II diabetes.

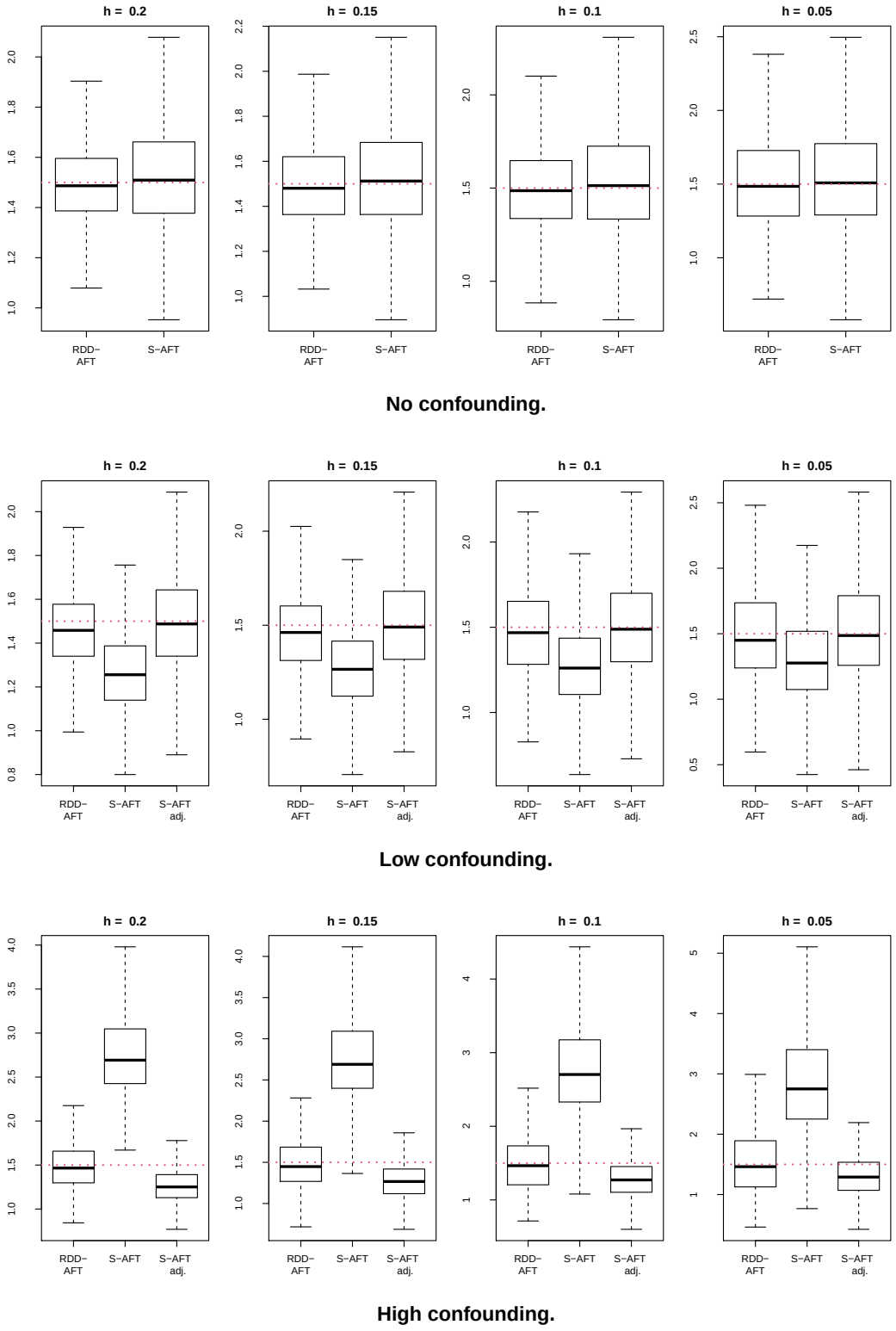

Figure 5.2: Boxplots of the results from the simulation study to compare the RDD-AFT and S-AFT methods under weak fuzziness scenario. The red dash line is the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

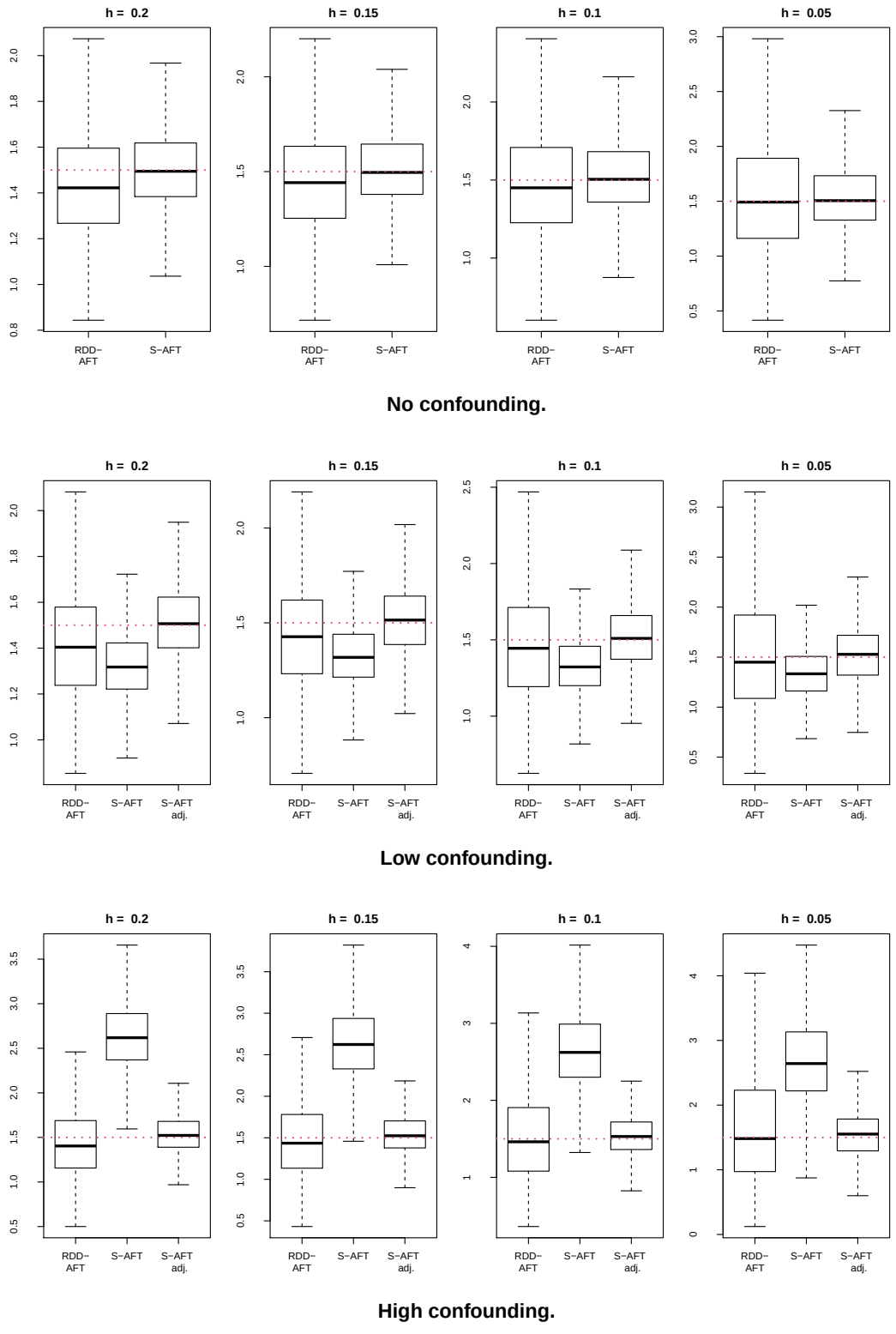

Figure 5.3: Boxplots of the results from the simulation study to compare the RDD-AFT and S-AFT methods under strong fuzziness scenario. The red dash line is the true treatment effect. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 5.2: Estimates, biases, empirical standard errors (ESE) , average standard errors (ASE) and 95% Coverage of the log of the acceleration factor under the no confounding scenario. The true value of the log of the acceleration factor is log(1.5) = 0.405. The sample size was 2000 in each simulated dataset and simulations were repeated 1000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.40	0.01	0.11	0.11	94.3	0.35	0.05	0.17	0.17	93.6
S-AFT	0.41	-0.01	0.19	0.15	94.9	0.40	0.00	0.14	0.11	94.4
	Bandwidth = 0.15 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.40	0.01	0.13	0.13	96.1	0.36	0.04	0.20	0.20	94.7
S-AFT	0.42	-0.01	0.19	0.16	95.9	0.41	0.00	0.16	0.13	94.5
	Bandwidth = 0.1 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.40	0.01	0.16	0.16	95.6	0.37	0.04	0.25	0.26	94.8
S-AFT	0.42	-0.01	0.23	0.19	95.4	0.41	-0.01	0.19	0.15	93.6
	Bandwidth = 0.05 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.40	0.01	0.23	0.23	96.3	0.40	0.01	0.39	0.39	95.0
S-AFT	0.42	-0.01	0.31	0.26	96.0	0.42	-0.02	0.25	0.19	94.6

Table 5.3: Estimates, biases, empirical standard errors (ESE) , average standard errors (ASE) and 95% Coverage of the log of the acceleration factor under the low confounding scenario. The true value of the log of the acceleration factor is log(1.5) = 0.405. The sample size was 2000 in each simulated dataset and simulations were repeated 1000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.03	0.12	0.12	94.0	0.34	0.07	0.18	0.19	94.5
S-AFT	0.23	0.17	0.18	0.15	71.8	0.27	0.13	0.15	0.11	75.0
S-AFT adj.	0.40	0.01	0.18	0.15	94.9	0.40	0.00	0.15	0.11	94.4
	Bandwidth = 0.15 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.03	0.14	0.14	95.3	0.35	0.06	0.22	0.22	94.3
S-AFT	0.23	0.17	0.20	0.16	74.3	0.27	0.13	0.16	0.13	78.9
S-AFT adj.	0.40	0.01	0.21	0.16	95.6	0.40	0.00	0.16	0.13	94.3
	Bandwidth = 0.1 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.03	0.18	0.18	95.7	0.37	0.04	0.28	0.28	95.2
S-AFT	0.23	0.17	0.24	0.19	80.6	0.27	0.13	0.19	0.15	83.5
S-AFT adj.	0.40	0.01	0.25	0.19	95.0	0.40	0.00	0.19	0.15	94.9
	Bandwidth = 0.05 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.02	0.26	0.25	94.9	0.38	0.02	0.42	0.42	96.0
S-AFT	0.24	0.16	0.32	0.25	85.9	0.27	0.13	0.25	0.20	87.9
S-AFT adj.	0.41	0.00	0.32	0.25	94.1	0.41	0.00	0.25	0.19	95.2

Table 5.4: Estimates, biases, empirical standard errors (ESE) , average standard errors (ASE) and 95% Coverage of the log of the acceleration factor under the high confounding scenario. The true value of the log of the acceleration factor is log(1.5) = 0.405. The sample size was 2000 in each simulated dataset and simulations were repeated 1000 times.

	Weak fuzziness					*Strong fuzziness*				
Method	Estimate	Bias	ESE	ASE	Coverage	Estimate	Bias	ESE	ASE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.02	0.18	0.18	95.3	0.34	0.06	0.28	0.28	94.6
S-AFT	1.01	-0.60	0.19	0.16	6.3	0.97	-0.56	0.18	0.15	6.9
S-AFT adj.	0.22	0.18	0.18	0.15	74.0	0.42	-0.02	0.17	0.13	93.0
	Bandwidth = 0.15 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.02	0.21	0.21	94.8	0.35	0.05	0.34	0.34	94.7
S-AFT	1.01	-0.60	0.22	0.18	9.7	0.97	-0.56	0.20	0.16	11.0
S-AFT adj.	0.23	0.18	0.21	0.17	78.2	0.42	-0.02	0.19	0.15	92.0
	Bandwidth = 0.1 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.03	0.26	0.26	95.2	0.37	0.04	0.45	0.43	94.3
S-AFT	1.00	-0.60	0.26	0.21	18.0	0.97	-0.56	0.22	0.19	16.7
S-AFT adj.	0.23	0.17	0.24	0.20	82.7	0.42	-0.01	0.21	0.17	93.4
	Bandwidth = 0.05 , Treatment effect = 0.405, Sample size = 2000									
RDD-AFT	0.38	0.02	0.38	0.37	95.0	0.38	0.02	0.67	0.64	95.2
S-AFT	1.02	-0.61	0.32	0.28	37.8	0.98	-0.57	0.30	0.25	31.8
S-AFT adj.	0.24	0.16	0.30	0.25	87.9	0.43	-0.02	0.29	0.23	92.4

5.6 Example: Prescription of metformin in patients at risk of Type II Diabetes

In this section, we apply the RDD-AFT and S-AFT approaches to a real dataset on metformin prescription in patients at risk of type II diabetes. The aim of the analysis is to investigate the effect of a metformin prescription on time to all-cause mortality and time to a cardiovascular event in patients who are at risk of type II diabetes. Metformin is a drug prescribed to reduce blood glucose level in patients with type II diabetes and those at risk of developing type II diabetes (NHS, 2019b). For the diagnosis of type II diabetes in UK primary care, glycated haemoglobin level (HbA1c), a measure of the average blood glucose level in the body over a duration of 3 months, is used. The National Institute for Health and Care Excellence (NICE)'s guideline on metformin prescription states that standard release metformin should be prescribed to patients whose HbA1c level is greater than or equal to 48mmol/mol (NICE, 2015). The data used in this section was extracted from The Health Improvement Network (THIN) database.

In this example, the HbA1c level is the continuous assignment variable and the threshold is set to be 48mmol/mol. The outcomes of interest are time to all-cause mortality and time to a cardiovascular event, with the time origin defined to be the time of first HbA1c measurement. We extracted data for 4532 male patients aged between 40 and 80 years who had their first HbA1c measurement in 2010, who had not been diagnosed with diabetes previously and whose body mass index (BMI) was less than 30kg/m^2.

Of these 4532 patients, 643 patients had HbA1c values above the threshold and 3889 patients had HbA1c values below the threshold. Of those patients with HbA1c values above the threshold, 453 (70%) were prescribed metformin whereas, of those patients with HbA1c values below the threshold, 27 (1%) were prescribed metformin. This indicates that the use of an RD design may be suitable for this data as the probability of receiving the treatment appears to differ substantially for patients above and below the threshold.

We compare the distribution of potential confounders for patients above and below the threshold. This is necessary to check that the distributions of confounders

are similar within the bandwidths considered. The potential confounders that are compared are age, body mass index, LDL and HDL cholesterol levels, and these confounders are chosen based on our discussions with epidemiologists. Table 5.5 shows the mean and standard deviation of these variables for patients above and below the threshold. We see that the BMI, LDL and HDL cholesterol levels are similar for patients above and below the threshold within bandwidths considered. However, for all bandwidths, we observe that patients above the threshold are older than those below the threshold. Based on this, we included age in the models for estimating the treatment effect to control the potential bias that could be introduced because of the possible imbalance in the distribution of age.

Table 5.5: Sample means and standard deviations for potential confounding variables above ($Z = 1$) and below ($Z = 0$) the threshold, for various HbA1c bandwidths (h).

Factors		$h = 10$		$h = 8$		$h = 6$		$h = 5$	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age at origin (years)	$Z = 0$	60.47	10.27	61.10	10.22	61.46	10.19	61.70	10.22
	$Z = 1$	62.75	9.80	63.04	9.63	62.90	9.85	63.41	9.83
BMI (kg/m^2)	$Z = 0$	25.98	2.66	26.14	2.63	26.23	2.59	26.29	2.62
	$Z = 1$	26.50	2.42	26.50	2.46	26.44	2.56	26.47	2.60
LDL Cholesterol ($mmol/L$)	$Z = 0$	3.20	0.93	3.18	0.94	3.14	0.92	3.15	0.92
	$Z = 1$	3.15	0.87	3.17	0.85	3.19	0.85	3.19	0.86
HDL Cholesterol ($mmol/L$)	$Z = 0$	1.30	0.37	1.27	0.35	1.25	0.36	1.24	0.36
	$Z = 1$	1.20	0.37	1.20	0.37	1.19	0.35	1.20	0.36

The probabilities of dying and experiencing a CVD event as well as the median time-to-event for patients who died and patients who experienced a CVD event are given Table 5.6. The probability of dying is similar in patients above and below the threshold across the bandwidths. The median times to death are broadly similar but slightly higher for subjects below the threshold, especially as the bandwidth increases.

For CVD event, the probability of a CVD event appears to be higher in patients below the threshold compared to those above the threshold. The median time-to-event is also higher for patients below the threshold across all bandwidths. This

would suggest that CVD event rates and times may be different for patients above and below the threshold.

Table 5.6: Empirical probabilities of death and a CVD event, together with sample median time-to-event (in years) for a variety of bandwidths.

	Bandwidth:	*h=10*		*h=8*		*h= 6*		*h= 5*	
	Z	0	1	0	1	0	1	0	1
	N	1956	202	1172	184	860	155	626	133
Death	Probability	0.06	0.06	0.07	0.07	0.08	0.08	0.08	0.09
	Median time-to-event (years)	2.97	2.31	2.92	2.31	2.99	2.31	2.72	2.31
CVD Event	Probability	0.07	0.04	0.08	0.04	0.07	0.03	0.07	0.03
	Median time-to-event (years)	2.58	1.34	2.71	1.18	2.38	1.34	2.38	1.62

Finally, the RDD-AFT and S-AFT approaches are applied to the data described above. Four bandwidths (h =10, 8, 6 and 5) were considered to check the sensitivity of the estimates to changes in bandwidth. Since age appears to be imbalanced for patients above and below the threshold, age was included as a covariate in the models for both approaches. Table 5.7 shows the results obtained from the analysis of the data and presents the estimated acceleration factor and the corresponding 95% confidence intervals. The standard error that is used to compute the confidence interval for the RDD-AFT approach was estimated using the variance estimation method described in Section 5.4.1.1 while a bootstrapping approach is used to estimate the standard error for the S-AFT approach.

The effect of metformin prescription on time to all-cause mortality and time to a cardiovascular event is the interest in this example. The point estimates for both the RDD-AFT and S-AFT models, across the bandwidths, are above 1. This may - at first sight - suggest that the median survival times to mortality and cardiovascular event for patients that receive a metformin prescription are higher than for patients that do not receive a metformin prescription. However, in all cases, the confidence intervals include 1, implying that the treatment effect estimates are not statistically significant and we do not have sufficient evidence to suggest a beneficial effect of metformin prescription (at the 5% level). The confidence intervals are generally wide, this may be because of a reduced sample size when the data are sub-sampled

so that only patients whose HbA1c values lie close to the threshold are included in the RD design analysis.

Table 5.7: Estimates and 95% confidence intervals of the acceleration factor for RDD-AFT and S-AFT approaches across varying bandwidths.

Bandwidth:	*10*	*8*	*6*	*5*
Outcome: Time to all cause mortality				
Method	Estimate (95% C.I.)	Estimate (95% C.I.)	Estimate (95% C.I.)	Estimate (95% C.I.)
RDD-AFT	2.37 (0.75, 7.45)	2.09 (0.53, 8.19)	1.32 (0.37, 4.70)	1.21 (0.26, 5.65)
S-AFT	1.11 (0.16, 7.93)	1.22 (0.20, 7.40)	2.59 (0.43, 15.44)	2.59 (0.41, 16.17)
Outcome: Time to a cardiovascular event				
Method	Estimate (95% C.I.)	Estimate (95% C.I.)	Estimate (95% C.I.)	Estimate (95% C.I.)
RDD-AFT	2.47 (0.57, 10.59)	1.25 (0.25, 6.25)	1.12 (0.13, 9.61)	1.22 (0.09, 15.92)
S-AFT	2.51 (0.34, 18.72)	2.69 (0.34, 21.16)	6.53 (0.92, 46.25)	3.51 (1.25, 9.88)

5.7 Conclusions

In this chapter, we have explored methods for estimating the acceleration factor in a fuzzy RD design for a time-to-event outcome. We proposed the RDD-AFT approach to estimate the acceleration factor based on the RD design assumptions, and we compared this to an S-AFT approach, a popular approach for estimating the acceleration factor in observational studies using simulation studies. When there is no unobserved confounding, the two methods yielded unbiased estimates of the treatment effect. However, for low and high confounding scenarios, the estimates from the S-AFT approach were biased for the treatment effect. The RDD-AFT approach continued to yield estimates close to the true treatment effect, even in the presence of unobserved confounding.

The two methods were applied to a real dataset to estimate the effect of a metformin prescription on time to death and time to a cardiovascular event. The results suggest a beneficial effect of metformin prescription, however, the estimates are not statistically significant.

In the next chapter, we present Bayesian alternatives to the methods that we have proposed for continuous, binary and time-to-event outcomes.

Chapter 6

Bayesian alternatives to proposed methods

In this chapter, we shall consider Bayesian alternatives to the proposed methods outlined in Chapters 3, 4 and 5 for treatment effect estimation in the RD design for continuous, binary and time-to-event outcomes. The Bayesian approach can provide a straightforward way to incorporate prior information about the parameter of interest.

6.1 Introduction

Given observed data $\mathbf{x}$, we assume that we are interested in estimating a parameter of interest, say θ. In the frequentist approach, the parameter is usually estimated using the information contained in the data only. Using a Bayesian approach, we combine prior information about the parameter, such as expert opinion or results from a different study, with the information from the data to obtain an updated information about the parameter. In an RD design, the Bayesian approach might be useful because we can obtain prior information about the parameter of interest from similar studies or clinical knowledge.

Typically, a prior distribution is defined to capture the prior information about the parameter, $p(\theta)$, the information from the data is defined in terms of the likelihood, $f(\mathbf{x}|\theta)$). The likelihood and prior are combined using Bayes theorem to obtain the posterior distribution of the parameter, $p(\theta|\mathbf{x})$):

$$p(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta)p(\theta).$$

Once the posterior distribution has been obtained, the parameter of interest can then be described in terms of summary measures, such as expected value and variance, of the posterior distribution. In some cases, the posterior distribution has a closed form, therefore, we can easily derive such summaries from the posterior distribution. In most cases, however, it might be difficult to derive the posterior distribution, in which case, Markov chain Monte Carlo (MCMC) approaches can be used to obtain samples from the posterior distribution (Hastings, 1970; Gelman and Rubin, 1992). The samples obtained from the posterior distribution can be used to estimate measures such as the expected value and variance of the parameter. In this thesis, Gibbs sampling, an MCMC approach, will be used to obtain samples from the posterior distributions that we require (Geman and Geman, 1984; Lunn et al., 2012).

Below, we describe the typical algorithm to obtain samples from the posterior distribution.

1. Choose an initial value for the parameter of interest.
2. Use an MCMC approach to obtain $M + N$ samples from the posterior distribution, in this thesis, we use Gibbs sampling.
3. Discard the first M samples (iterations) as burn-in. Typically M is decided after it appears that the chain has converged. For instance, in Figure 6.1 (a), M is chosen to be 150 as we see that it appears the iterations have converged at the 150^{th} iteration.
4. Repeat steps 1 to 3 above p times, with different initial values of the parameter of interest. p represents the number of chains and, it is recommended to run more than one chain, as it serves as a useful check for convergence because it is important that all the chains converge to the same area irrespective of the starting value. This is depicted in Figure 6.1 (b) where the two chains have different starting values, Chain 1 appears to have converged before Chain 2 but M is chosen at the iteration where both chains have converged.
5. The remaining $N \times p$ samples from the posterior distribution are summarised and presented.

An advantage of the Bayesian approach over the frequentist approach is that using

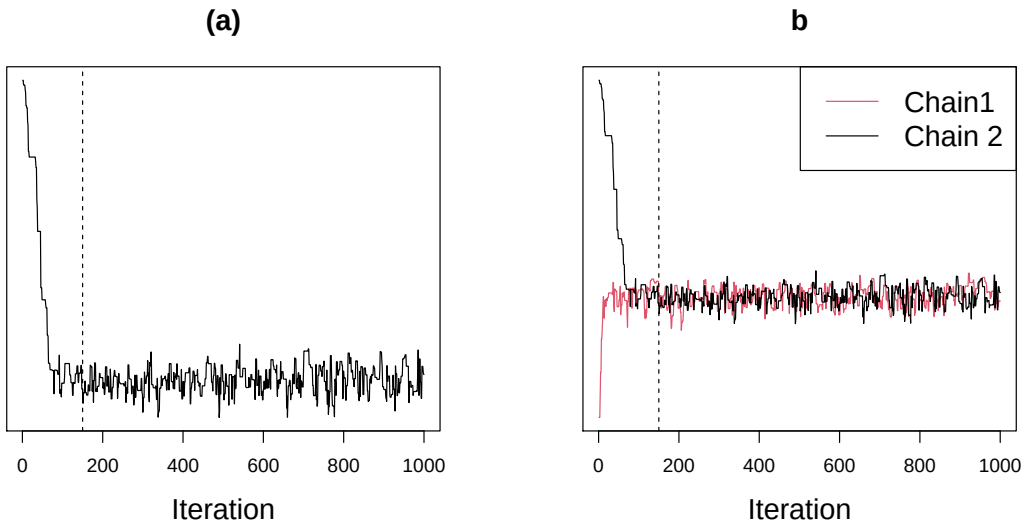

Figure 6.1: Trace plot to illustrate how convergence is checked when (a) one chain is run and (b) two chains are run. The dotted lines represents potential burn-in point.

the Bayesian approach provides full information on the distribution of the parameter of interest. As such, all properties of the parameter that are required can be easily computed. For instance, in the frequentist approach, bootstrapping was employed to estimate the variance of the estimator. With a Bayesian approach, the variability in the parameter of interest is easily estimated from samples obtained from the posterior distribution of the parameter of interest.

We shall now proceed to discuss the Bayesian alternatives to the RD design estimation methods discussed earlier in the thesis. In the next section, we present a Bayesian approach to the thin plate regression spline method for estimating the local average treatment effect (LATE) discussed in Section 3.6.

6.2 Continuous outcome: Bayesian Thin Plate Regression Spline

In Chapter 3, the LATE was discussed as an estimator of the treatment effect for a continuous outcome in a fuzzy RD design. Furthermore, the use of the thin plate regression spline as a flexible approach for the estimation of the numerator of the LATE was proposed. Here, we present a Bayesian alternative for LATE estimation using thin plate regression spline.

We recall that the thin plate regression spline estimator of the LATE is given as:

$$\hat{\lambda}_{\text{tprs}} = \frac{\hat{f}_1(x_0) - \hat{f}_0(x_0)}{\text{expit}(\hat{\pi}_{01}) - \text{expit}(\hat{\pi}_{00})},$$

where for $j \in \{0, 1\}$, the numerator terms are estimated from thin plate spline models given below:

$$\begin{aligned} y_i &= f(x_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2), \text{ and} \\ \hat{f}_j(x_0) &= \boldsymbol{\eta}_j^\top \hat{\boldsymbol{\delta}}_j + \hat{\alpha}_{1j} + \hat{\alpha}_{2j} x_0, \end{aligned}$$

and the denominator terms are estimated from logistic regression models given below:

$$\begin{aligned} \text{logit}(\mathbb{P}(A_i = 1 | X_i^c = x_i^c)) &= \pi_{01} + \pi_{11} x_i^c \quad i \in \mathcal{A}, \\ \text{logit}(\mathbb{P}(A_i = 1 | X_i^c = x_i^c)) &= \pi_{00} + \pi_{10} x_i^c \quad i \in \mathcal{B}. \end{aligned}$$

The corresponding Bayesian TPRS estimator of the LATE is denoted by λ_{btprs} and we specify a normal prior distribution for λ_{btprs}:

$$\lambda_{\text{btprs}} \sim \text{Normal}(\mu_\lambda, \sigma_\lambda^2).$$

The prior information about the treatment effect can then be incorporated via the prior mean (μ_λ) and prior variance (σ_λ^2).

We recall that the minimisation problem of the thin plate regression spline is a penalised regression problem. One of the ways to fit a penalised regression model is to define the regression model in the form of a mixed regression model (Crainiceanu et al., 2005; Wood, 2017). Therefore, we define the mixed regression model as

$$\mathbf{y}_j = \tilde{\mathbf{E}}_j \boldsymbol{\delta}_j^* + \mathbf{G}_j \boldsymbol{\alpha}_j + \boldsymbol{\epsilon}_j, \quad \text{Cov}(\boldsymbol{\epsilon}_j, \boldsymbol{\delta}_j^*) = \begin{bmatrix} \sigma_{\epsilon j}^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_{\delta j}^2 \mathbf{I} \end{bmatrix}. \tag{6.1}$$

where $\tilde{\mathbf{E}}_j = \mathbf{U}_{kj} \mathbf{D}_k \mathbf{Z}_{kj} \left(\mathbf{Z}_{kj}^\top \mathbf{D}_{kj} \mathbf{Z}_{kj} \right)^{-\frac{1}{2}}$, $\hat{\boldsymbol{\delta}}_j = \mathbf{U}_{kj} \mathbf{Z}_{kj} \left(\mathbf{Z}_{kj}^\top \mathbf{D}_{kj} \mathbf{Z}_{kj} \right)^{-\frac{1}{2}} \hat{\boldsymbol{\delta}}_j^*$, $\boldsymbol{G}_j$ is the matrix of smooth functions as described in Section 3.4.2 and $\boldsymbol{\delta}_j^*$ and $\boldsymbol{\alpha}_j = (\alpha_{1j}, \alpha_{2j})^\top$ are the parameters to be estimated. The distributions for model fitting are specified

as follows:

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\mu_{i1}, \sigma^2_{\epsilon 1}), \qquad \mu_{i1} = \tilde{\mathbf{E}}_{i1}\boldsymbol{\delta}^*_1 + \mathbf{G}_{i1}\boldsymbol{\alpha}_1 \quad \text{for } i \in \mathcal{A}, \\
Y_i &\sim \text{Normal}(\mu_{i0}, \sigma^2_{\epsilon 0}), \qquad \mu_{i0} = \tilde{\mathbf{E}}_{i0}\boldsymbol{\delta}^*_0 + \mathbf{G}_{i0}\boldsymbol{\alpha}_0 \quad \text{for } i \in \mathcal{B}, \\
A_i &\sim \text{Binomial}(1, p_{i1}), \quad \text{logit}(p_{i1}) = \pi_{01} + \pi_{11}x^c_i \quad \text{for } i \in \mathcal{A}, \\
A_i &\sim \text{Binomial}(1, p_{i0}), \quad \text{logit}(p_{i0}) = \pi_{00} + \pi_{10}x^c_i \quad \text{for } i \in \mathcal{B}.
\end{aligned}
\tag{6.2}
$$

Where $\tilde{\mathbf{E}}_{ij}$ and $\mathbf{G}_{ij}$ are the ith row of $\tilde{\mathbf{E}}_j$ and $\mathbf{G}_j$ respectively. Since we have placed a prior distribution on λ_{btprs}, we will express α_{11} as a function of the other parameters:

$$
\alpha_{11} = \lambda_{\text{btprs}}[\text{expit}(\pi_{01}) - \text{expit}(\pi_{00})] + f_0(x_0) - \boldsymbol{\eta}_1\boldsymbol{\delta}_1 - \alpha_{21}x_0.
$$

The prior distributions of the remaining parameters are given below:

$$
\begin{aligned}
\boldsymbol{\delta}^*_j &\sim \text{Normal}(0, \sigma^2_{\delta j}), \\
\boldsymbol{\alpha}_0 &\sim \text{Normal}(0, \sigma^2_{\alpha 0}), \\
\alpha_{21} &\sim \text{Normal}(0, \sigma^2_{\alpha 1}), \\
\sigma^{-2}_{\epsilon j} &\sim \text{Gamma}(a_{\epsilon j}, b_{\epsilon j}), \\
\sigma^{-2}_{\delta j} &\sim \text{Gamma}(a_{\delta j}, b_{\delta j}), \\
\pi_{00} &\sim \text{Normal}(-2, 1), \\
\pi_{01} &\sim \text{Normal}(2, 1), \\
\pi_{10} &\sim \text{Normal}(0, \sigma^2_{\pi 1}), \\
\pi_{11} &\sim \text{Normal}(0, \sigma^2_{\pi 0}).
\end{aligned}
$$

The values of the parameters of the prior distributions will incorporate prior information about the parameters. The models stated in Equation 6.2 may be run using the `R2jags` package in `R` (Plummer, 2019), which uses Gibbs sampling to obtain samples from the posterior distributions of the parameters of interest. The posterior mean and the corresponding standard deviation of λ_{btprs} are then reported as the LATE estimate and its standard error.

We carried out simulation studies to compare the performance of the Bayesian TPRS estimator of the LATE to the non-Bayesian version. Datasets that were simulated in Section 3.5 are used for this purpose. For the simulation studies, the prior distri-

bution of the treatment effect is $\lambda_{\text{btprs}} \sim \text{Normal}(-2, 0.2)$. The prior distributions of the remaining parameters are:

$$\begin{aligned}
\boldsymbol{\delta}_z^* &\sim \text{Normal}(0, \sigma_{\delta z}^2), & \pi_{00} &\sim \text{Normal}(-2, 1),\\
\boldsymbol{\alpha}_0 &\sim \text{Normal}(0, 10), & \pi_{01} &\sim \text{Normal}(2, 1),\\
\alpha_{21} &\sim \text{Normal}(0, 10), & \pi_{10} &\sim \text{Normal}(0, 10),\\
\sigma_{\epsilon z}^{-2} &\sim \text{Gamma}(1, 1), & \pi_{11} &\sim \text{Normal}(0, 10),\\
\sigma_{\delta z}^{-2} &\sim \text{Gamma}(0.1, 0.1).
\end{aligned}$$

We ran 2 chains with 3000 iterations each, the first 1000 iterations were discarded as burn-in. Overall, there were 4000 samples from the posterior distribution in total, 2000 from each chain.

We recall that four scenarios with varying relationships between the outcome and assignment variables were considered. Figure 6.2 shows boxplots for the posterior mean of the LATE obtained from the simulation studies for Scenarios 1, 2, 3 and 4. Table 6.1 contains the numerical results from the simulation studies to compare the frequentist and Bayesian TPRS approaches for Scenarios 1 and 2, while Table 6.2 contains numerical summaries for Scenarios 3 and 4.

For Scenario 1, both the frequentist and Bayesian approaches produce unbiased estimates of the true treatment effect. We note that for the Bayesian approach, coverage is quite high, which suggests that the Bayesian approach might be over-estimating the standard error. For both frequentist and Bayesian approaches, the 95% confidence interval (equal tail credible interval for the Bayesian approach) contains the treatment effect.

Under Scenario 2, we observe that the estimates from the Bayesian approach are close to the value of the true treatment effect across bandwidths, and we note that biases of estimates from the frequentist approach are larger. Similarly, in this scenario, we observe that the coverage for the Bayesian approach is higher than the nominal value which might imply that the Bayesian approach is over-estimating the standard error estimate.

For Scenarios 3 and 4, we observe that biases of the estimates from the Bayesian

approach are smaller compared to those from the frequentist approach. Similarly, the coverage of the credible intervals of the Bayesian approach is above the nominal level.

Overall, it seems that by specifying the prior distribution, the Bayesian approach produces estimates that are closer to the true value of the treatment effect compared to the frequentist approach that is only based on the information from the observed data. However, the Bayesian approach seem to over-estimate the standard error, as such, coverage of the credible intervals of the Bayesian approach are quite high.

Table 6.1: Estimates, biases, empirical standard errors and the 95% coverage of the LATE to compare the Bayesian and frequentist approaches for estimating the LATE using the thin plate spline for Scenarios 1 and 2.

	Scenario 1				*Scenario 2*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
	Bandwidth = 0.2 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-1.98	-0.02	0.10	94.3	-1.90	-0.10	0.15	86.2
Bayes TPRS	-2.01	0.01	0.17	99.3	-1.96	-0.04	0.18	99.2
	Bandwidth = 0.15 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-1.99	-0.01	0.12	94.8	-1.89	-0.11	0.16	77.9
Bayes TPRS	-2.01	0.01	0.17	99.7	-1.96	-0.04	0.17	99.9
	Bandwidth = 0.1 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-1.99	-0.01	0.15	94.3	-1.93	-0.07	0.16	87.8
Bayes TPRS	-2.01	0.01	0.16	100.0	-2.00	0.00	0.16	100.0
	Bandwidth = 0.05 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-2.00	0.00	0.21	95.1	-1.98	-0.02	0.21	96.0
Bayes TPRS	-2.01	0.01	0.15	100.0	-2.01	0.01	0.14	100.0

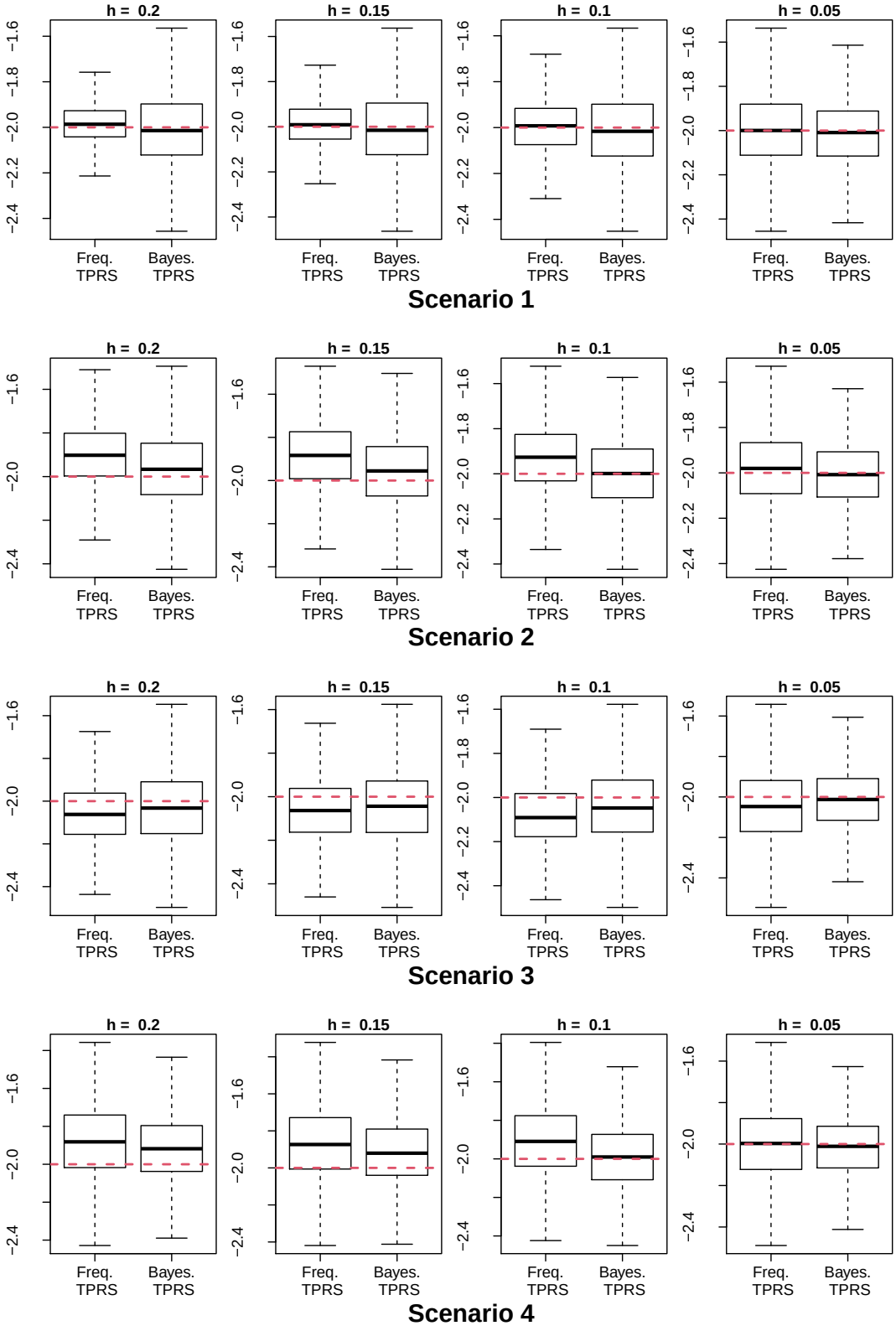

Figure 6.2: Boxplots of the estimate of the LATE to compare the frequentist and Bayesian thin plate regression spline methods for estimating the LATE. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 6.2: Estimates, biases, empirical standard errors and the 95% coverage of the LATE to compare the Bayesian and frequentist approaches for estimating the LATE using the thin plate spline for Scenarios 3 and 4.

	Scenario 3				*Scenario 4*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
	Bandwidth = 0.2 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-2.05	0.05	0.16	90.3	-1.88	-0.12	0.20	84.6
Bayes TPRS	-2.03	0.03	0.18	98.7	-1.91	-0.09	0.18	97.7
	Bandwidth = 0.15 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-2.05	0.05	0.17	90.6	-1.87	-0.13	0.21	80.0
Bayes TPRS	-2.04	0.04	0.18	99.2	-1.92	-0.08	0.18	98.9
	Bandwidth = 0.1 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-2.07	0.07	0.17	86.1	-1.91	-0.09	0.20	86.4
Bayes TPRS	-2.04	0.04	0.18	99.8	-1.99	-0.01	0.17	99.9
	Bandwidth = 0.05 , Treatment effect = -2, Sample size = 2000							
Freq TPRS	-2.03	0.03	0.23	94.2	-2.00	0.00	0.21	94.8
Bayes TPRS	-2.01	0.01	0.15	100.0	-2.01	0.01	0.15	100.0

6.2.1 Example on Statin Prescription in UK Primary Care

In this section, we shall apply the Bayesian approach described above to the data described in Section 3.6. We recall that the aim of the analysis is to estimate the effect of statin prescription on LDL cholesterol level. The relevant variables from the data are defined as follows:

- Y_i is the outcome, the LDL cholesterol level (mmol/L).
- X_i is the assignment variable, the risk of developing a cardiovascular disease in 10 years.
- $x_0 = 0.2$ is the threshold for the treatment guideline: patients are to receive a statin prescription if the value of their assignment variable is greater or equal to x_0.
- A_i is the treatment indicator that takes value 1 if patient i receives a statin prescription, 0 otherwise.

Baigent et al. (2005) carried out a meta-analysis of 14 randomised trials on the safety and efficacy of statin prescription where it was noted that the average reduction in LDL cholesterol level is -1.09, with the values ranging from 0.35 to 1.77. This informed our choice of prior distribution for the treatment effect of interest in the example we are considering:

$$\lambda_{\text{btprs}} \sim \text{Normal}(-1, 1).$$

The prior distributions of the remaining parameters are given below

$$\begin{aligned}
\beta = \lambda_{\text{btprs}} (\pi_{01} - \pi_{00}), \qquad & \pi_{00} \sim \text{Normal}(-2, 1), \\
\boldsymbol{\delta}^* \sim \text{Normal}(0, \sigma_\delta^2), \qquad & \pi_{01} \sim \text{Normal}(2, 1), \\
\boldsymbol{\alpha} \sim \text{Normal}(0, 10), \qquad & \pi_{10} \sim \text{Normal}(0, 10), \\
\sigma_\epsilon^{-2} \sim \text{Gamma}(1, 1), \qquad & \pi_{11} \sim \text{Normal}(0, 10), \\
\sigma_\delta^{-2} \sim \text{Gamma}(0.1, 0.1). &
\end{aligned}$$

We carried out the Bayesian analysis in `R` using the `r2jags` package (Plummer, 2019).

We ran 4 chains with 5500 iterations each, the first 2000 iterations of each chain were discarded as burn-in, so that for each chain, 2500 iterations were saved. We checked the trace plots of the samples from the posterior distribution of the treatment effect to confirm that the chains converged. The traceplots for the four bandwidths are presented in Figure 6.3, we observe that the there is sufficient mixture of the chains which also suggests that the chains have converged. As such, we present the results using the properties of the posterior distribution of the parameter of interest.

Results obtained from the Bayesian analysis are presented in Table 6.3, we also show the results from the frequentist method here. The results of the Bayesian TPRS indicate that statin prescription is associated with a reduction in the LDL cholesterol level. This is in agreement with results from frequentist approach and also is inline with the purpose of statin prescription. The standard error estimates of the Bayesian approach are larger than the standard error estimates of the frequentist approach, this is similar to what we observed in the simulation studies where the Bayesian approach over-estimates the standard error estimate. Next, we shall present the Bayesian alternative of the RDD-RR estimator that was developed for estimating the risk ratio in an RD design presented in Chapter 4.

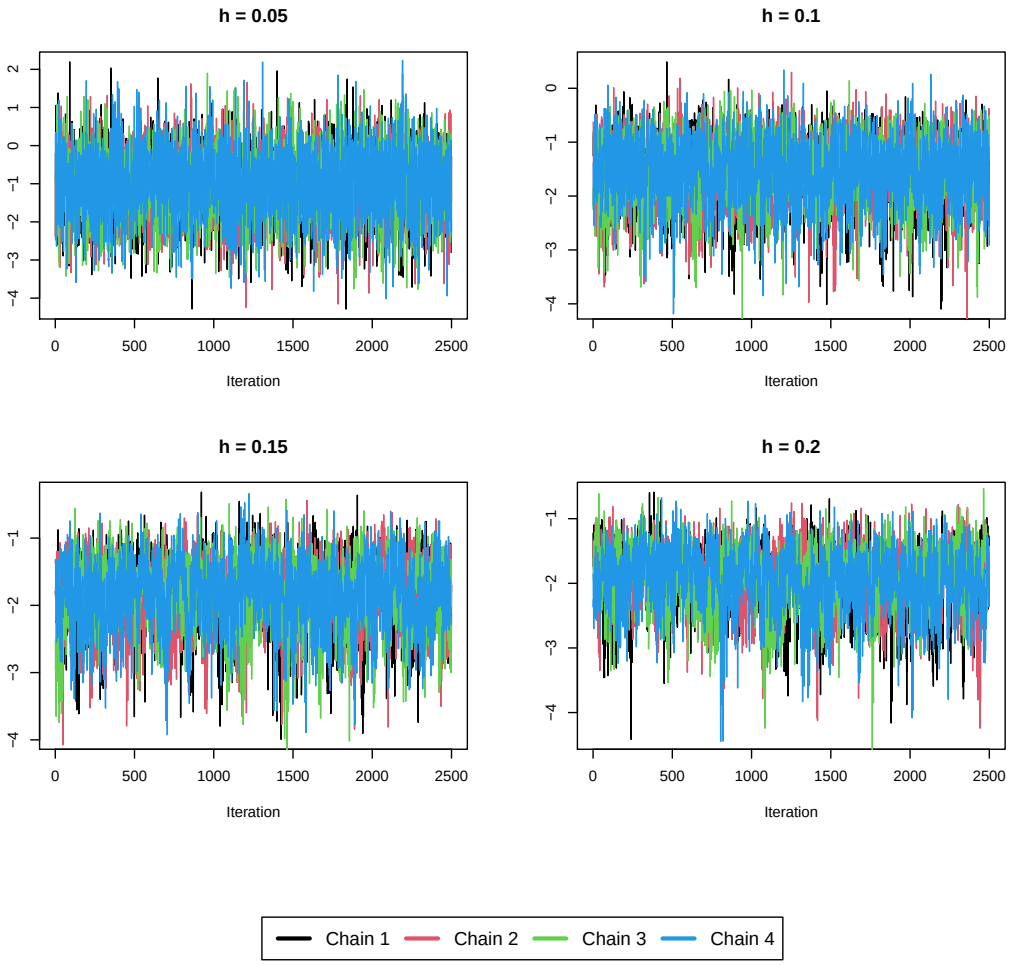

Figure 6.3: Trace plots of λ_{btprs} for the four bandwidths considered.

Table 6.3: The posterior mean (estimates) and associated standard deviation (SE) of the Bayesian approach of estimating the LATE for the THIN data example on the prescription of statins based on 10-year CVD risk score. Results from the frequentist approach are also included.

Bandwidth:	*0.05*		*0.1*		*0.15*		*0.2*	
Method	Estimate	SD	Estimate	SD	Estimate	SD	Estimate	SD
Freq TPRS	-1.04	0.83	-1.30	0.34	-1.42	0.25	-1.48	0.22
Bayes TPRS	-1.01	0.84	-1.57	0.59	-1.87	0.52	-1.95	0.49

LATE: Local average treatment effect;

6.3 Bayesian methods for binary outcomes

In Chapter 4, the RDD-RR estimator was proposed for estimating the risk ratio in an RD design. Simulation studies indicated that this approach is preferable to the WALD-RR and MSMM approaches that can be used to estimate the risk ratio in an RD design. Here, we propose a Bayesian approach to estimate the RDD-RR.

The RDD-RR estimator is given as

$$\text{RDD}_{\text{RR}} = 1 - \lim_{x \to x_0} \frac{\mathbb{E}(Y|Z=1, X=x) - \mathbb{E}(Y|Z=0, X=x)}{\mathbb{E}(Y|Z=1, X=x)\mathbb{E}(A|Z=0) - \mathbb{E}(Y|Z=0, X=x)\mathbb{E}(A|Z=1)}$$

We denote the Bayesian RDD-RR as λ_{RR} and we specify the prior distribution of λ_{RR} as follows:

$$\lambda_{\text{RR}} \sim \text{Log-normal}(\mu_{\text{RR}}, \sigma^2_{\text{RR}}).$$

We specify a log-normal distribution for the prior distribution of λ_{RR} to ensure that the estimate of λ_{RR} will be non-negative. Prior information about λ_{RR} can be incorporated in the values of μ_{RR} and σ^2_{RR}. The distributions for model fitting are specified as follows:

$$\begin{aligned}
&Y_i \sim \text{Binomial}(1, \mu_{i1}) \quad \text{logit}(\mu_{i1}) = \gamma_{01} + \gamma_{11} x_i^c \quad \text{for } i \in \mathcal{A}, \\
&Y_i \sim \text{Binomial}(1, \mu_{i0}) \quad \text{logit}(\mu_{i0}) = \gamma_{00} + \gamma_{10} x_i^c \quad \text{for } i \in \mathcal{B}, \\
&A_i \sim \text{Binomial}(1, \pi_a) \quad \text{for } i \in \mathcal{A}, \\
&A_i \sim \text{Binomial}(1, \pi_b) \quad \text{for } i \in \mathcal{B}.
\end{aligned} \tag{6.3}$$

Similar to what we have done in Section 6.2, since we have specified a prior distribution for λ_{RR}, we will express γ_{01} as a function of the other parameters:

$$\text{expit}(\gamma_{01}) = \frac{\text{expit}\left(\gamma_{00}\right)\left(\pi_1 - \pi_1\lambda_{\text{RR}} - 1\right)}{\pi_0 - \pi_0\lambda_{\text{RR}} - 1},$$

where $\text{expit}(x) = \frac{\exp(x)}{1+\exp(x)}$. The prior distributions for the remaining parameters are

given below:

$$\gamma_{00} \sim \text{Normal}(0, \sigma^2_{00}),$$
$$\gamma_{10} \sim \text{Normal}(0, \sigma^2_{10}),$$
$$\gamma_{11} \sim \text{Normal}(0, \sigma^2_{11}),$$
$$\text{logit}(\pi_1) \sim \text{Normal}(2, 1),$$
$$\text{logit}(\pi_0) \sim \text{Normal}(-2, 1).$$

The Bayesian estimator of RDD-RR is calculated as the posterior mean of the samples obtained from the posterior distribution of λ_{RR}.

We carried out simulation studies using the datasets that were simulated in Section 4.4 to compare the Bayesian and frequentist approaches of estimating the RDD-RR. For the Bayesian approach, the prior distribution of the treatment effect is specified as: $\lambda_{\text{RR}} \sim \text{Log-normal}(\log(1.5), 0.5)$, the values specified for the parameters of the prior distributions for the remaining parameters are $\sigma^2_{00} = \sigma^2_{10} = \sigma^2_{11} = 10$. As before, six simulation scenarios were considered varying the levels of confounding and fuzziness.

Figures 6.4 and 6.5 show boxplots of the posterior means of the treatment effect obtained from Bayesian approach and estimates obtained using frequentist approach from the simulation studies under the weak and strong fuzziness scenarios, respectively. Tables 6.4, 6.5 and 6.6 show the numerical summaries of the estimates of the simulation studies under the no, low and high confounding scenarios respectively. As we observed under the frequentist approach, results presented in these tables are for the log of the risk ratio.

Comparing the boxplots in Figures 6.4 and 6.5 and the estimates in Tables 6.4, 6.5 and 6.6, we observe that the results from the frequentist and Bayesian approaches are very similar to each other. Therefore, the performances of the two approaches are comparable in terms of bias and standard error estimates. An advantage of the Bayesian approach, however, is that rather than estimating the standard error of RDD-RR estimate with the bootstrapping approach as we did for the frequentist approach, the standard error of the estimate from the Bayesian approach is obtained directly from the samples obtained from the posterior distribution of the treatment

effect.

Under the weak fuzziness scenario, we see that the two methods of estimating RDD-RR produced very little or no bias across the bandwidths considered. In addition, the performance of the estimators does not seem to be affected by the level of confounding, as the estimates under the no, low and high confounding scenarios are similar. This conforms with the advantage of using an RD design, that is, the RD design can identify a treatment effect in observational data in the presence of unobserved confounding. In terms of standard errors, the ESE and ASE are equal, which indicates the standard errors of the estimates for the two approaches are correctly estimated.

Under the strong fuzziness scenario, as shown in Figure 6.5 and the right panels of Tables 6.4, 6.5 and 6.6, we see that estimates from the frequentist and Bayesian approaches are similar and we observe that estimates from both approaches are biased where the bandwidth size is large. However, as the bandwidth reduces the bias also reduces, but, owing to the fact that for a small bandwidth, the number of observations is also small, estimates tend to be more variable. Furthermore, as we observed under the weak fuzziness scenario, the performance of the estimators is not affected by the level of confounding in the design.

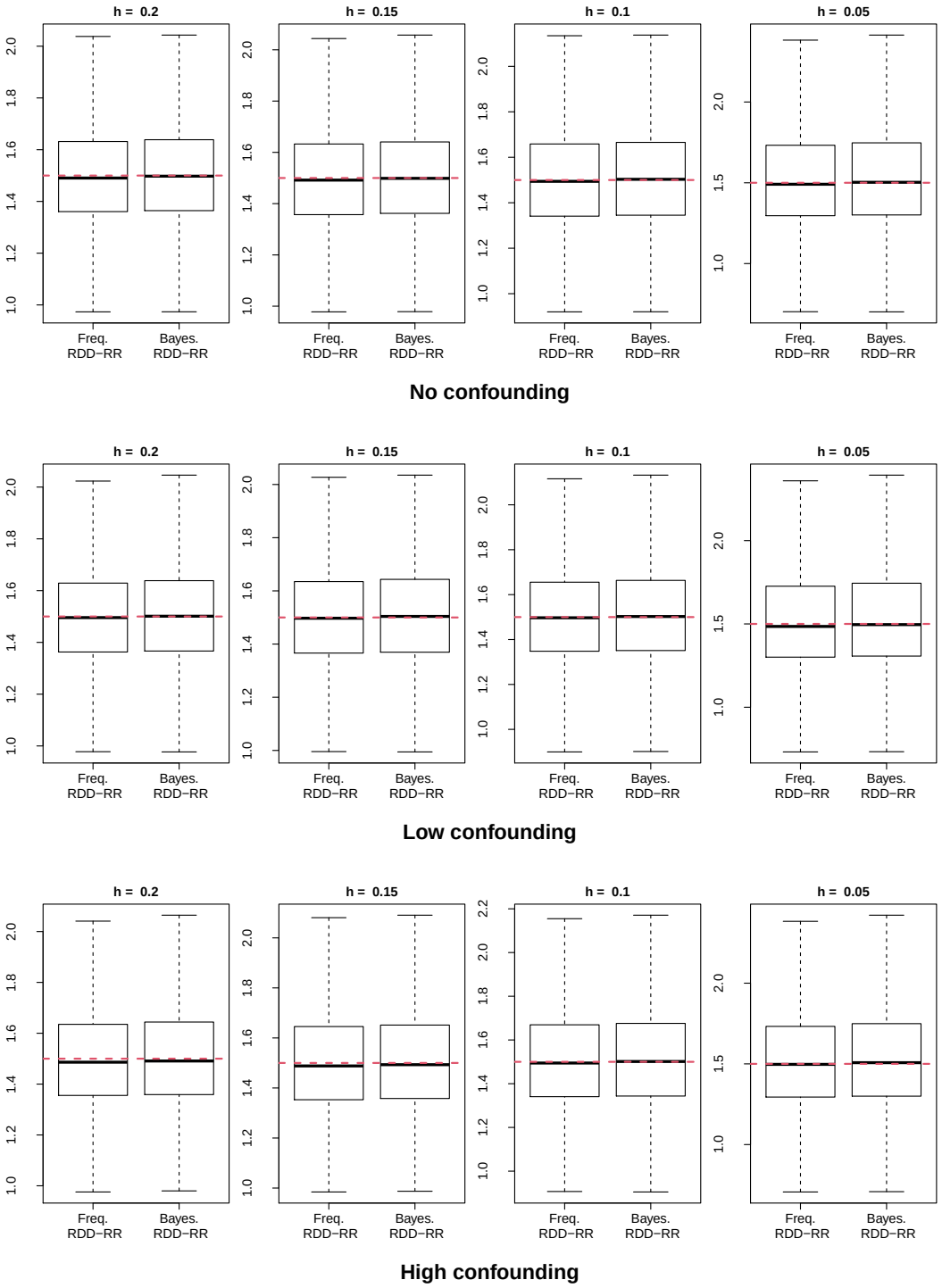

Figure 6.4: Boxplots of the estimates from simulation studies to compare the Bayesian and frequentist approaches for estimating the RDD-RR under the weak fuzziness scenario. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

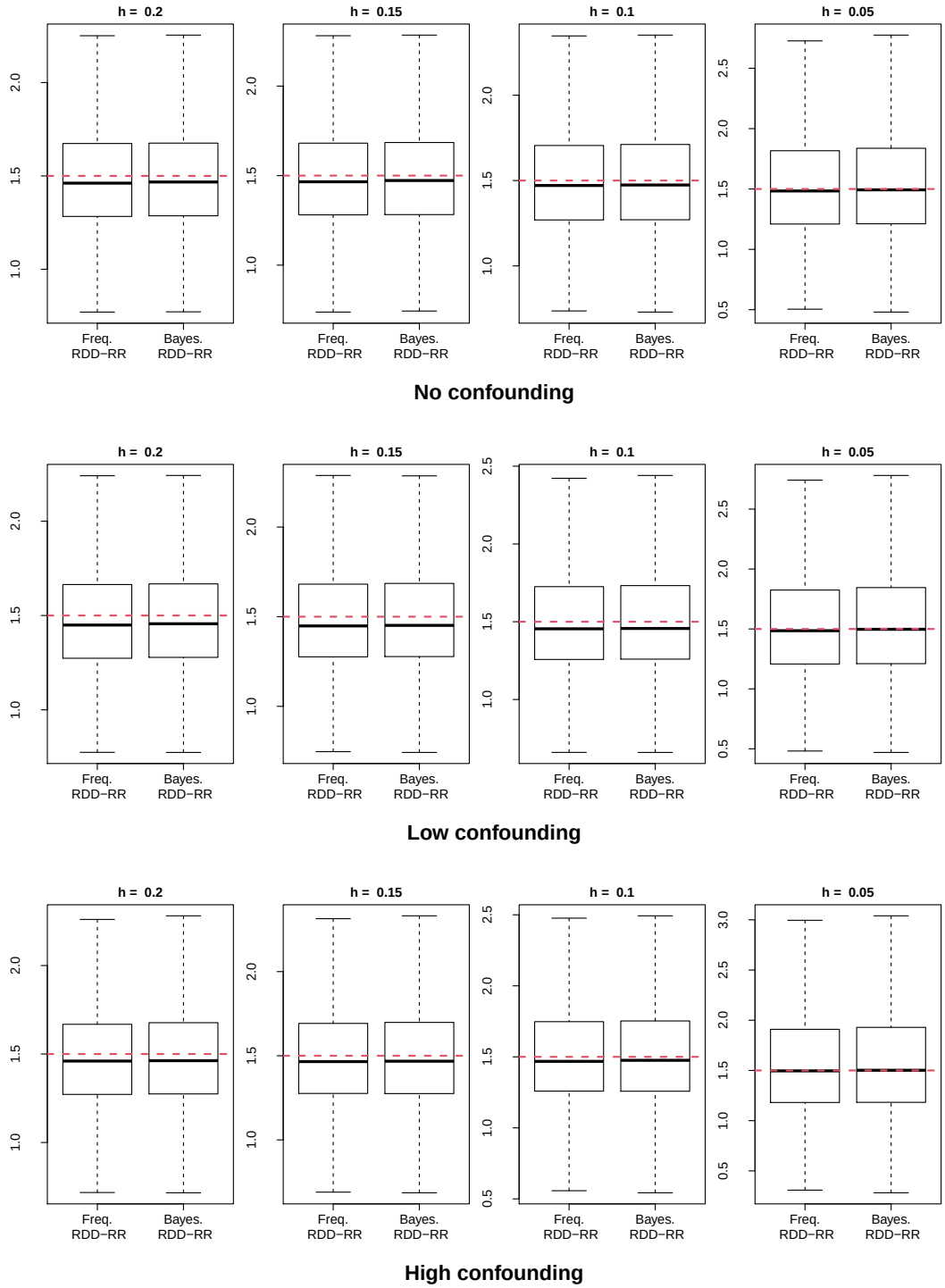

Figure 6.5: Boxplots of the estimates from simulation studies to compare the Bayesian and frequentist approaches for estimating the RDD-RR under the strong fuzziness scenario. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 6.4: Estimates, biases, empirical standard errors and the 95% coverage of RDD-RR from simulation studies to compare the Bayesian and frequentist approaches under the no confounding scenario.

	Weak fuzziness				*Strong fuzziness*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.01	0.13	94.9	0.38	0.02	0.19	96.2
Bayes RDD-RR	0.40	0.01	0.13	94.9	0.38	0.02	0.18	95.2
	Bandwidth = 0.15 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.01	0.14	95.0	0.38	0.02	0.20	95.8
Bayes RDD-RR	0.40	0.01	0.13	94.6	0.38	0.03	0.19	95.2
	Bandwidth = 0.1 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.16	95.2	0.39	0.02	0.23	96.2
Bayes RDD-RR	0.40	0.01	0.16	94.3	0.38	0.02	0.21	95.6
	Bandwidth = 0.05 , Treatment effect = 0.405							
Freq RDD-RR	0.41	0.00	0.22	95.4	0.40	0.01	0.31	97.7
Bayes RDD-RR	0.40	0.01	0.21	94.6	0.38	0.02	0.27	97.0

Table 6.5: Estimates, biases, empirical standard errors and the 95% coverage of RDD-RR from simulation studies to compare the Bayesian and frequentist approaches under the low confounding scenario.

	Weak fuzziness				*Strong fuzziness*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.13	95.6	0.38	0.03	0.20	95.2
Bayes RDD-RR	0.40	0.01	0.13	95.6	0.38	0.03	0.19	94.9
	Bandwidth = 0.15 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.14	95.6	0.38	0.02	0.21	95.3
Bayes RDD-RR	0.40	0.01	0.13	95.7	0.38	0.03	0.20	95.0
	Bandwidth = 0.1 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.15	95.9	0.39	0.02	0.24	95.6
Bayes RDD-RR	0.40	0.01	0.15	95.4	0.38	0.02	0.22	94.9
	Bandwidth = 0.05 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.22	95.0	0.40	0.01	0.33	97.6
Bayes RDD-RR	0.39	0.01	0.20	95.0	0.38	0.02	0.28	96.4

Table 6.6: Estimates, biases, empirical standard errors and the 95% coverage of RDD-RR from simulation studies to compare the Bayesian and frequentist approaches under the high confounding scenario.

	Weak fuzziness				*Strong fuzziness*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.01	0.14	94.3	0.39	0.02	0.21	95.2
Bayes RDD-RR	0.40	0.01	0.14	94.4	0.38	0.02	0.20	94.7
	Bandwidth = 0.15 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.14	94.6	0.39	0.02	0.22	96.0
Bayes RDD-RR	0.40	0.01	0.14	93.3	0.38	0.02	0.20	94.8
	Bandwidth = 0.1 , Treatment effect = 0.405							
Freq RDD-RR	0.40	0.00	0.16	95.7	0.39	0.01	0.25	97.0
Bayes RDD-RR	0.40	0.01	0.16	94.7	0.39	0.02	0.23	95.7
	Bandwidth = 0.05 , Treatment effect = 0.405							
Freq RDD-RR	0.41	0.00	0.22	95.6	0.41	0.00	0.35	98.1
Bayes RDD-RR	0.40	0.01	0.20	94.8	0.39	0.01	0.29	97.2

6.3.1 Example on Statin Prescription in UK Primary Care

We applied the Bayesian approach to estimating the RDD-RR to the real data on statin prescription, described in Section 4.5. We recall the variables from the dataset are defined as:

- The outcome: Y_i equals 1 if the LDL cholesterol level of patient i is reduced by at least 1mmol/L and 0 otherwise.
- Treatment indicator: A_i equals 1 if patient i receives a statin prescription and 0 otherwise.
- The assignment variable: X_i is the risk that patient i develops a cardiovascular disease in 10 years.
- The threshold: $x_0 = 0.2$, the treatment guideline states that patients with value of assignment variable greater or equal to x_0 should receive a statin prescription and vice versa.
- The threshold indicator: Z_i equals 1 if the value of patient i's assignment variable is greater or equal to x_0 and 0 otherwise.

We place a vaguely informative prior on the treatment effect of interest - $\lambda_{\text{RR}} \sim \text{Lognormal}(\log(1.5), 2)$. The values specified for the parameters of the prior distributions for the remaining parameters are $\sigma^2_{00} = \sigma^2_{10} = \sigma^2_{11} = 10$.

For the Bayesian analysis, we ran 4 MCMC chains with 5500 iterations each. The first 2000 iterations from each chain was discarded as burn-in so that 2500 iterations was saved per chain. To check for convergence, the trace plots of the parameter of interest are provided in Figure 6.6 for the four bandwidths considered. The trace plots show that the four chains have converged across the four bandwidths considered and we present a summary of the samples from the posterior distribution of the parameter of interest.

Table 6.7 presents estimates obtained from the RDD-RR method for the frequentist and Bayesian approaches for bandwidths 0.05, 0.1, 0.15 and 0.2 as well as the 95% confidence intervals for the frequentist approach and 95% equal tail credible

intervals for the Bayesian approach. As we have observed from the results from the simulation studies, the estimates from the Bayesian approach are similar to those from the frequentist approach.

The estimates from the Bayesian and frequentist approach suggest that there may be a beneficial effect of statin prescription in reducing LDL cholesterol level for bandwidths 0.1, 0.15 or 0.2. However, for bandwidth of 0.05, estimates indicate that patients who receive a statin prescription are less likely to have their LDL cholesterol level reduced by at least 1mmol/L compared to patients that did not receive a statin prescription. But, since the credible and confidence intervals contain 1, we do not have sufficient, significant, evidence about the potential benefit of statin prescription.

In the next section, we shall describe the Bayesian alternative to the RDD-AFT estimator that is used to estimate the acceleration factor in an RD design that was proposed in Chapter 5.

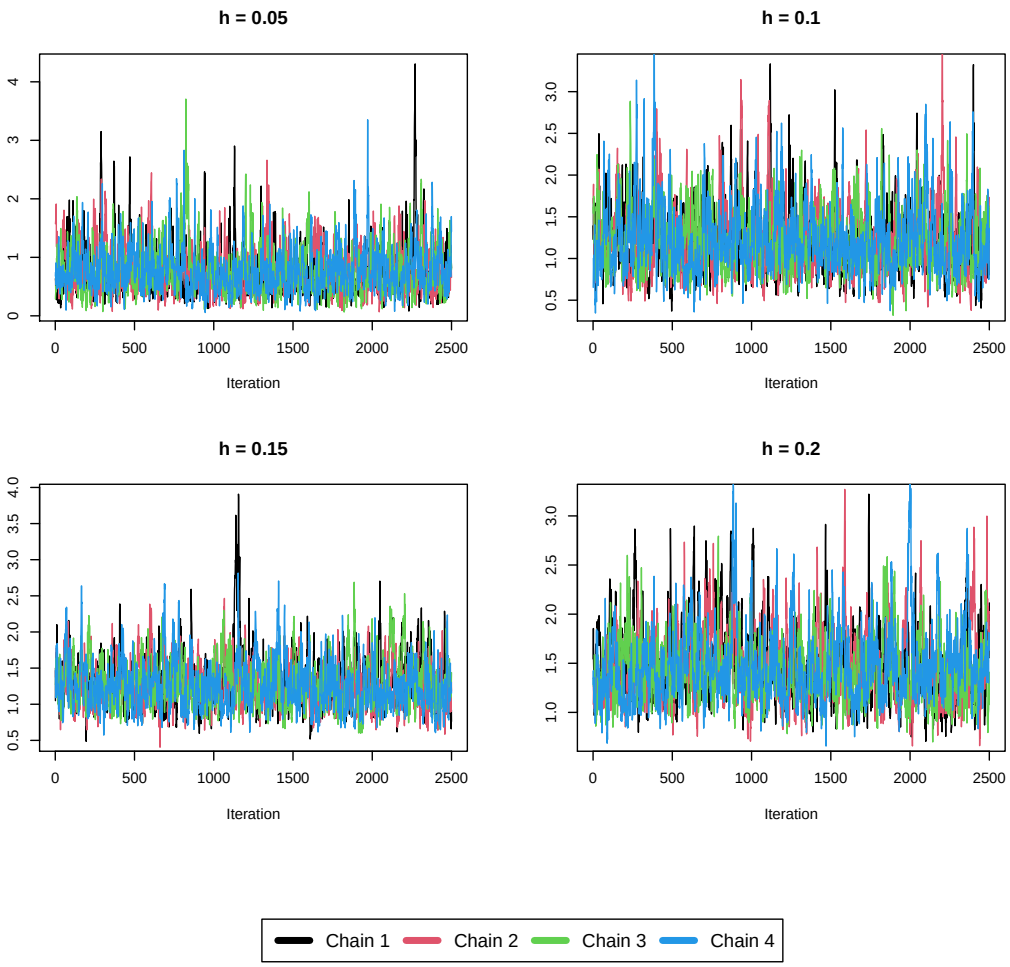

Figure 6.6: Trace plots of the Bayesian RDD-RR for the four bandwidths considered.

Table 6.7: Estimates and 95% credible (and confidence) intervals of RDD-RR estimate of the effect of statin prescription on reducing LDL cholesterol level.

Bandwidth:	*0.05*		*0.1*		*0.15*		*0.2*	
Method	Estimate	95% C.I.	Estimate	95% C.I.	Estimate	95% C.I.	Estimate	95% C.I.
Freq RDD-RR	0.71	(0.24, 2.08)	1.17	(0.65, 2.11)	1.26	(0.79, 2.01)	1.43	(0.91, 2.23)
Bayes RDD-RR	0.78	(0.23, 1.72)	1.21	(0.64, 2.12)	1.27	(0.77, 2.03)	1.45	(0.93, 2.30)

6.4 Bayesian methods for time-to-event outcomes

In this section, we present a Bayesian alternative to the estimator of the acceleration factor in an RD design that was proposed in Chapter 5. We recall that the logarithm of the RDD-AFT estimator that has been derived is

$$\log(\text{RDD-AFT}) = \frac{\alpha_1}{\pi_1 - \pi_0}, \tag{6.4}$$

where α_1 is obtained from the AFT model

$$\log(T_i) = \alpha_1 Z_i + \alpha_2 X_i + \alpha_3 X_i Z_i + \epsilon_i, \tag{6.5}$$

and the elements of the denominator π_z represents the probability of receiving treatment when $Z = z\ , z \in \{0,\ 1\}$, that is, below and above the threshold.

In this thesis, we focused on the Weibull parametric model, however, as mentioned in Chapter 5, other parametric distributions can be considered. To fit a Weibull AFT model, it may be more straightforward to write the model using the generalised gamma distribution. In the `jags` manual (Plummer, 2017), the probability density function of a random variable T that has a generalised gamma distribution is given as

$$f(t) = \frac{b\lambda^{br-1}\exp\{-(\lambda t)^b\}}{\Gamma(r)}, \quad b, \lambda, r > 0.$$

If $r = 1$, the generalised gamma distribution is a Weibull distribution with probability density function:

$$f(t) = b\lambda^{b-1}\exp\{-(\lambda t)^b\} \quad b, \lambda, r > 0.$$

The survivor function of this parameterisation of the Weibull distribution is

$$\mathbb{P}(T \geq t) = \exp\{-(\lambda t)^b\}.$$

Here, λ represents the factor by which the time-to-event t is accelerated. Therefore, under the AFT assumption, the relationship between the covariates and the outcome is modelled through λ.

We denote the Bayesian RDD-AFT estimator as λ_{AFT} and its prior distribution can be specified as:

$$\lambda_{\text{AFT}} \sim \text{Lognormal}(\mu_{\text{AFT}}, \sigma^2_{\text{AFT}}).$$

Here, we have specified a log-normal distribution as the prior distribution of λ_{AFT} to ensure that the estimate of the acceleration factor will be non-negative. Next, we describe the distributions for modelling the outcome variable and treatment indicator.

$$\begin{aligned}
T_i &\sim \text{Generalised Gamma}(1, \lambda_i, b), \\
\log(\lambda_i) &= \alpha_0 + \alpha_1 Z_i + \alpha_2 X_i + \alpha_3 X_i Z_i, \\
A_i &\sim \text{Binomial}(1, \pi_1) \quad \text{for } i \in \mathcal{A}, \\
A_i &\sim \text{Binomial}(1, \pi_0) \quad \text{for } i \in \mathcal{B}.
\end{aligned}$$

Since we have placed a prior on λ_{AFT}, we will express α_1 as a function of the remaining parameters:

$$\alpha_1 = \lambda_{\text{AFT}}(\pi_1 - \pi_0).$$

The prior distributions for the remaining parameters are given as:

$$\begin{aligned}
\alpha_0 &\sim \text{Normal}(0, \sigma^2_{\alpha 0}) \\
\alpha_2 &\sim \text{Normal}(0, \sigma^2_{\alpha 2}) \\
\alpha_3 &\sim \text{Normal}(0, \sigma^2_{\alpha 3}), \\
b &\sim \text{Gamma}(c_b, d_b), \\
\text{logit}(\pi_1) &\sim \text{Normal}(2, 1), \\
\text{logit}(\pi_0) &\sim \text{Normal}(-2, 1).
\end{aligned}$$

The models specified above can be run using the `rjags` package (Plummer, 2019) in `R` (R Core Team, 2018) and the samples from the posterior distribution of the parameter of interest (λ_{AFT}) saved after convergence has been achieved. The treatment effect estimate can be obtained by computing the posterior mean and the standard error of the estimate is the posterior standard deviation.

Simulation studies were carried out to check the performance of the Bayesian approach to RDD-AFT estimation that has been described above and compare this

to the non-Bayesian RDD-AFT approach seen in Chapter 5. The prior distribution for the treatment effect is specified as $\lambda_{\text{AFT}} \sim \text{Lognormal}(\log(1.5), 0.5)$. Details of the parameters of the prior distributions for the remaining parameters are: $\sigma^2_{\alpha 0} = \sigma^2_{\alpha 2} = \sigma^2_{\alpha 3} = 10$, $c_b = d_b = 10^{-2}$. The datasets used for the simulation studies are those that were used in the simulation studies under the frequentist approach that were described in Section 5.5. As with the binary outcome, six simulation scenarios were considered by varying the probabilities of compliance with the treatment guideline (level of fuzziness) and the level of confounding. Figures 6.7 and 6.8 show boxplots of the estimates obtained from the simulation studies under the weak and strong fuzziness scenarios, respectively. Tables 6.8, 6.9 and 6.10 present numerical summaries of the estimates of log(RDD-AFT) under the no, low and high confounding scenarios, respectively.

Under the weak fuzziness scenario, as shown in Figure 6.7 and the left panels of Tables 6.8, 6.9 and 6.10, we observe that the frequentist and Bayesian approaches of estimating the RDD-AFT yield little to no bias when there is no confounding across the bandwidths. As the level of confounding increases, some bias is observed for larger bandwidths, however, the bias reduces as the bandwidth reduces. Overall, the estimates and standard errors obtained from the frequentist and Bayesian approaches are similar to one another.

For the strong fuzziness scenario, as shown in Figure 6.8 and the right panels of Tables 6.8, 6.9 and 6.10, we observe that, for larger bandwidths, estimates from the Bayesian and frequentist approaches are biased. However, the bias reduces as the bandwidth size reduces. In addition, we observe that the standard error estimates under this scenario are larger than those under the weak fuzziness scenario. The probability of compliance with the treatment guideline is smaller under the strong fuzziness scenario, which causes a higher uncertainty about the estimate, thereby resulting in a larger standard error estimate under this scenario. Notably, we observe that the standard error estimates of the Bayesian approach are smaller compared to the frequentist approach and this might in turn lead to the width of the credible intervals of estimates from the Bayesian approach to be narrower than the width of the confidence intervals of estimates from the frequentist approach.

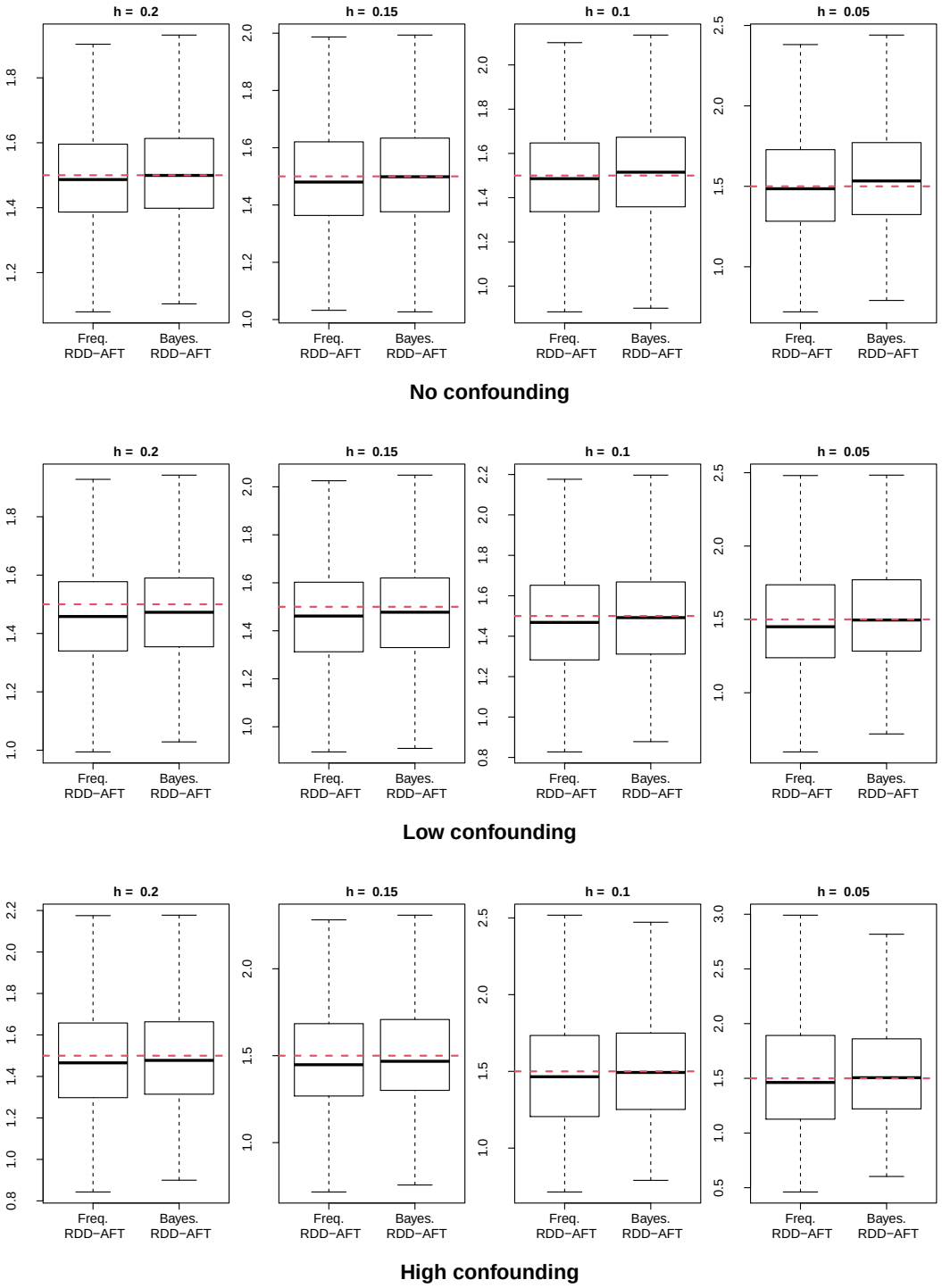

Figure 6.7: Boxplots of the estimates from simulation studies to compare the Bayesian and frequentist approaches for estimating the RDD-AFT under the weak scenario. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

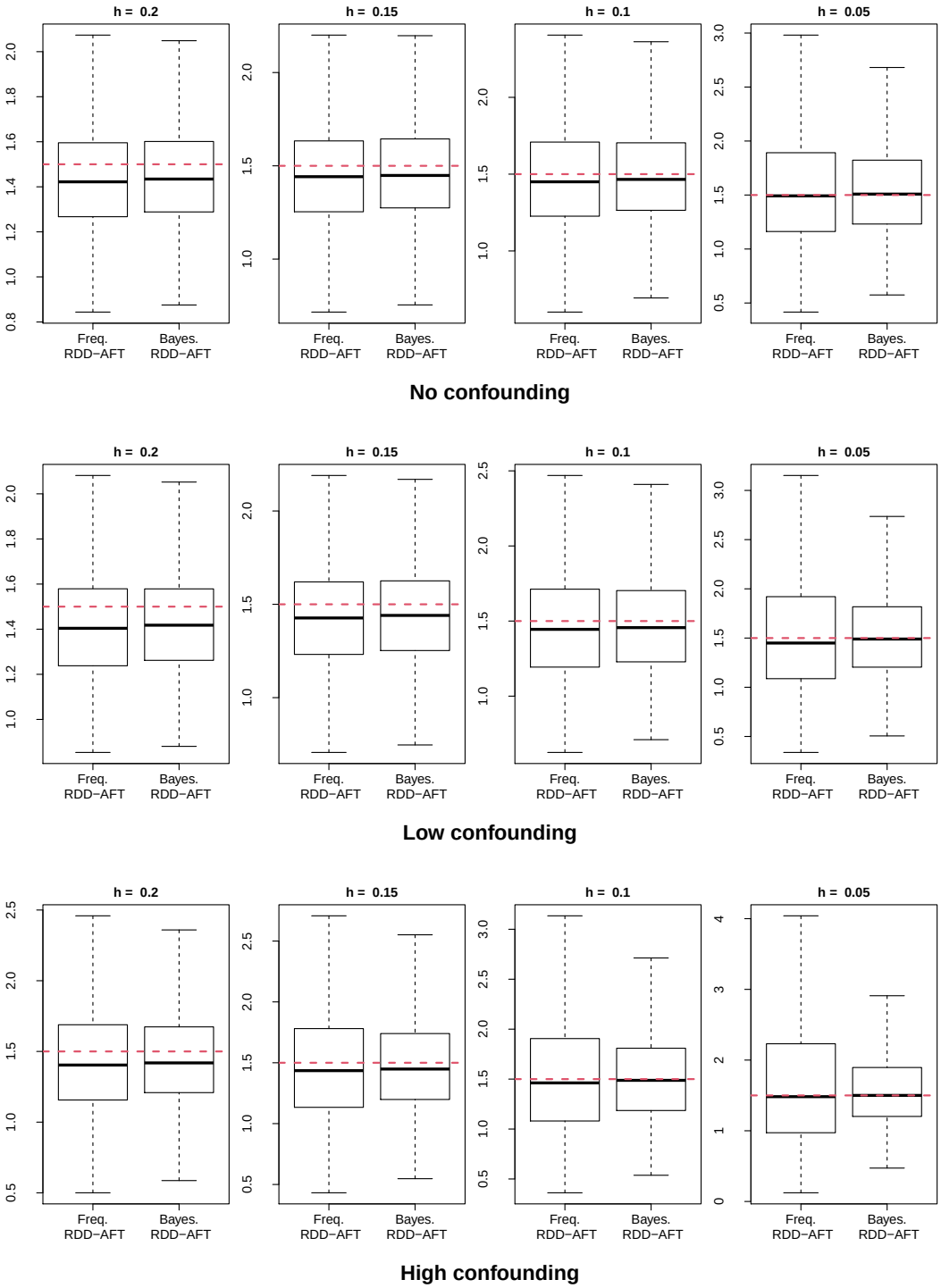

Figure 6.8: Boxplots of the estimates from simulation studies to compare the Bayesian and frequentist approaches for estimating the RDD-AFT under the strong scenario. The central line and limit of the boxplots represent the median and inter-quartile range of the estimates respectively.

Table 6.8: Estimates, biases, empirical standard errors and the 95% coverage of RDD-AFT from simulation studies to compare the Bayesian and frequentist approaches under the no confounding scenario.

	Weak fuzziness				*Strong fuzziness*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
	Bandwidth = 0.2 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.40	0.01	0.11	94.6	0.36	0.05	0.17	94.2
Bayes RDD-AFT	0.41	0.00	0.11	94.0	0.36	0.04	0.16	94.3
	Bandwidth = 0.15 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.39	0.01	0.12	95.9	0.37	0.04	0.20	94.8
Bayes RDD-AFT	0.41	0.00	0.13	95.7	0.37	0.03	0.19	95.2
	Bandwidth = 0.1 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.40	0.01	0.15	95.3	0.37	0.04	0.25	95.2
Bayes RDD-AFT	0.41	-0.01	0.16	95.9	0.38	0.03	0.22	95.3
	Bandwidth = 0.05 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.40	0.01	0.22	96.8	0.39	0.01	0.38	95.0
Bayes RDD-AFT	0.42	-0.02	0.21	96.0	0.41	0.00	0.30	96.7

Table 6.9: Estimates, biases, empirical standard errors and the 95% coverage of RDD-AFT from simulation studies to compare the Bayesian and frequentist approaches under the low confounding scenario.

	Weak fuzziness				*Strong fuzziness*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
Bandwidth = 0.2 , Treatment effect = 0.405, Sample size = 2000								
Freq RDD-AFT	0.37	0.03	0.12	94.3	0.34	0.07	0.18	94.3
Bayes RDD-AFT	0.38	0.02	0.12	94.5	0.35	0.06	0.17	93.9
	Bandwidth = 0.15 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.37	0.03	0.14	94.6	0.35	0.06	0.22	94.4
Bayes RDD-AFT	0.39	0.02	0.14	95.4	0.36	0.05	0.20	95.1
	Bandwidth = 0.1 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.38	0.03	0.17	95.9	0.36	0.04	0.28	95.2
Bayes RDD-AFT	0.40	0.01	0.17	95.9	0.38	0.03	0.24	95.6
	Bandwidth = 0.05 , Treatment effect = 0.405, Sample size = 2000							
Freq RDD-AFT	0.38	0.02	0.25	95.4	0.38	0.03	0.42	96.4
Bayes RDD-AFT	0.41	0.00	0.24	95.4	0.40	0.01	0.31	97.6

Table 6.10: Estimates, biases, empirical standard errors and the 95% coverage of RDD-AFT from simulation studies to compare the Bayesian and frequentist approaches under the high confounding scenario.

	Weak fuzziness				*Strong fuzziness*			
Method	Estimate	Bias	ESE	Coverage	Estimate	Bias	ESE	Coverage
Bandwidth = 0.2 , Treatment effect = 0.405, Sample size = 2000								
Freq RDD-AFT	0.38	0.03	0.18	95.5	0.34	0.06	0.28	95.1
Bayes RDD-AFT	0.39	0.01	0.17	94.8	0.36	0.05	0.25	95.5
Bandwidth = 0.15 , Treatment effect = 0.405, Sample size = 2000								
Freq RDD-AFT	0.38	0.02	0.21	95.1	0.35	0.05	0.34	94.8
Bayes RDD-AFT	0.39	0.01	0.20	95.7	0.37	0.04	0.27	95.4
Bandwidth = 0.1 , Treatment effect = 0.405, Sample size = 2000								
Freq RDD-AFT	0.38	0.03	0.26	95.6	0.36	0.04	0.44	94.4
Bayes RDD-AFT	0.40	0.01	0.24	95.9	0.38	0.02	0.33	97.0
Bandwidth = 0.05 , Treatment effect = 0.405, Sample size = 2000								
Freq RDD-AFT	0.38	0.02	0.38	94.3	0.38	0.02	0.67	94.9
Bayes RDD-AFT	0.41	0.00	0.30	96.5	0.40	0.00	0.36	98.1

6.4.1 Example on Metformin Prescription in UK Primary Care

We applied the Bayesian approach to RDD-AFT estimation to the data on effect of metformin prescription on type II diabetes related complications: all-cause mortality and cardiovascular disease. In the dataset, we defined two outcome variables of interest, the time to all-cause mortality and the time to a cardiovascular disease event. The remaining variables in the data are defined as

- The treatment indicator: A_i equals 1 if patient i receives a metfomin prescription and 0 otherwise.
- The assignment variable: X_i is the HbA1c value of patient i
- The threshold: $x_0 = 48$mmol/mol, patients are to receive a metformin prescription if their HbA1c value is greater or equal to x_0.
- The threshold indicator: Z_i equals 1 if patient's i HbA1c value is greater or equal to x_0 and 0 otherwise.

Here, we are interested in the estimating the effect (acceleration factor) of metformin prescription on time to all-cause-mortality and time to a CVD event. To specify the prior distribution of the treatment effect, we looked at the results from Han et al. (2019), a meta-analysis of 40 studies that looked at the effect of metformin prescription on all-cause-mortality and CVD event. They reported the hazard ratio for all-cause-mortality and CVD event as 0.67 and 0.83 respectively. Since acceleration factor and hazard ratio are inversely related, that is, a beneficial effect of as treatment would result in a hazard ratio less than 1 and an acceleration factor greater than 1, we specify the following prior distribution that for the acceleration factors of interest for both outcomes.

$$\lambda_{\text{AFT}} \sim \text{Lognormal}(0.2, 0.1).$$

Figure 6.9 is the density of the prior distribution of the treatment effect of interest. The parameters of prior distributions of the other parameters are specified as: $\sigma^2_{\alpha 0} = \sigma^2_{\alpha 2} = \sigma^2_{\alpha 3} = 10$, $c_b = d_b = 10^{-2}$.

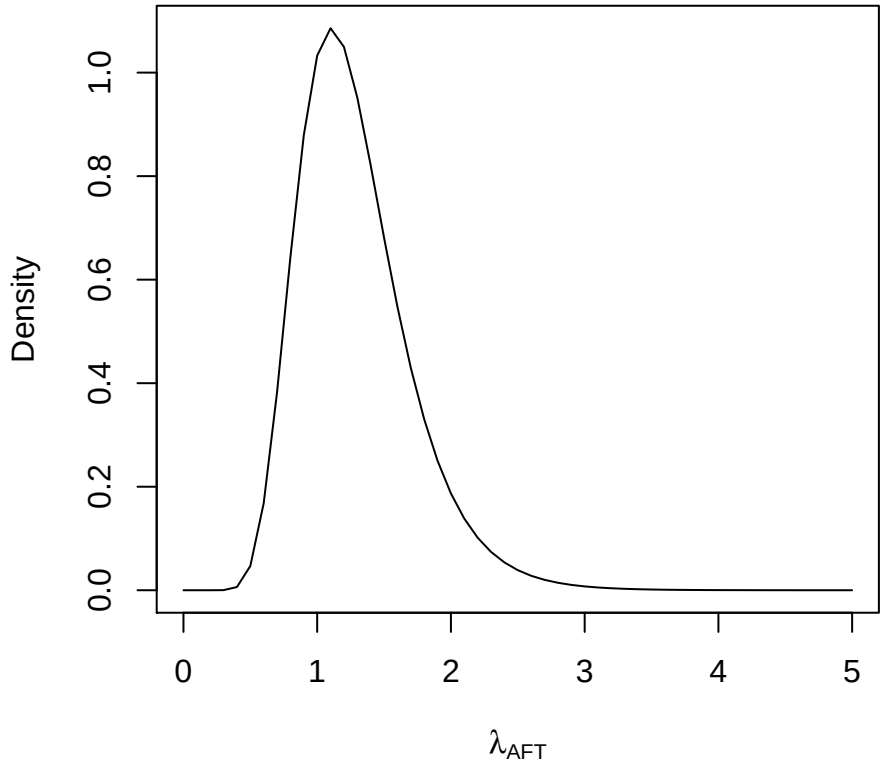

Figure 6.9: Density of the prior distribution of λ_{AFT}

We estimated RDD-AFT using the Bayesian approach that has been described for the two outcomes of interest. As we have done for the continuous and binary outcomes earlier, we ran 4 MCMC chains for each outcome of interest. The first 2000 iterations from each chain were discarded as burn-in and the remaining 2500 iterations were saved. To check for convergence, the trace plots of log(RDD-AFT) for the two outcomes of interest are provided in Figures 6.10 and 6.11 for the four bandwidths considered. The trace plots show that the four chains have converged across the four bandwidths considered and so, we present a summary of the samples from the posterior distribution of the log(RDD-AFT).

We present estimates and 95% credible intervals of the RDD-AFT for the effect of a metformin prescription on time to all-cause mortality and time to a cardiovascular disease event for patients at risk of type II diabetes in Table 6.11. We observe that the estimated values of the effect of metformin prescription from the Bayesian approach appear to be stable across bandwidths when compared to estimates from

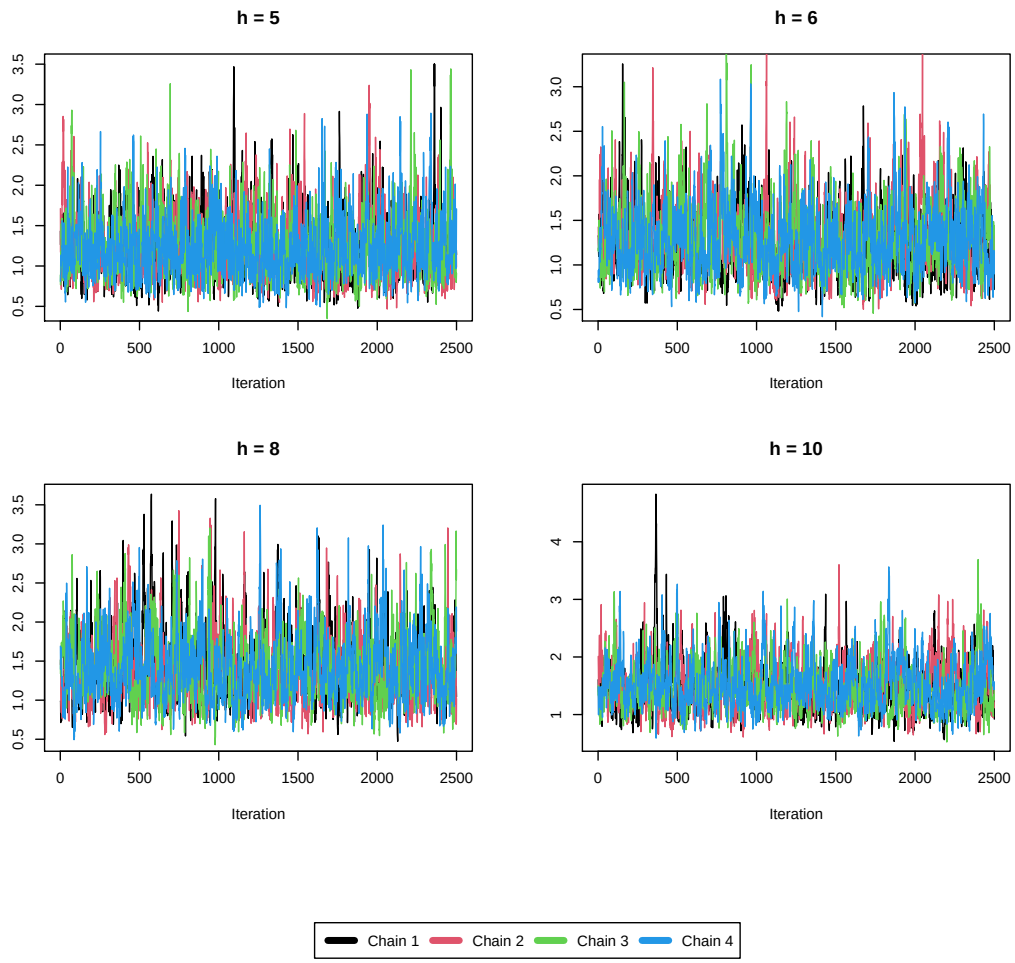

Figure 6.10: Trace plots of log(RDD-AFT) for the four bandwidths considered when event of interest is all-cause mortality.

the frequentist approach. However, estimates from both methods suggest that the median time to all-cause mortality and the median time to a cardiovascular disease event are higher in patients that receive a metformin prescription compared to patients that do not receive a prescription. This implies a potential benefit of receiving a prescription of metformin. However, we note that the credible (and confidence) intervals contain 1, which implies there is insufficient evidence to suggest a beneficial effect of metformin prescription. We observe that the estimates from the Bayesian approach are more precise as the 95% credible intervals are narrower compared to the 95% confidence intervals of the frequentist approach.

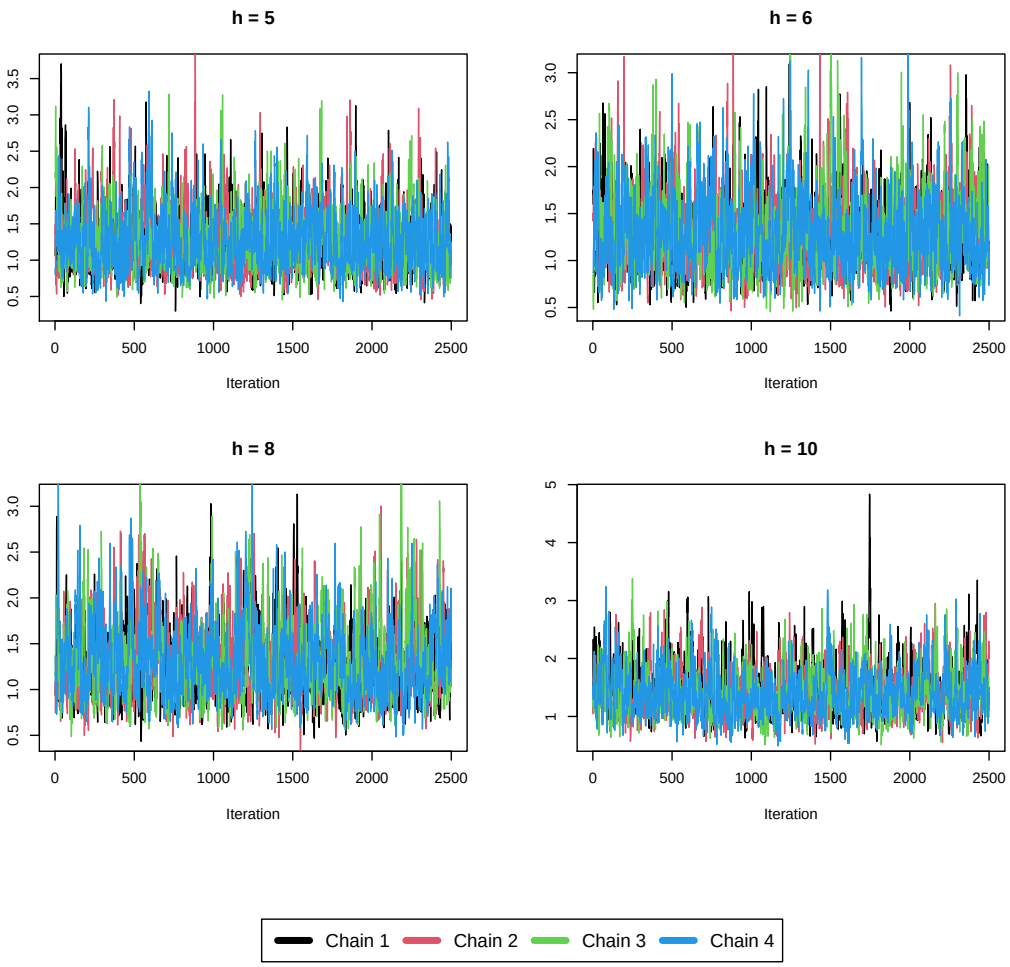

Figure 6.11: Trace plots of log(RDD-AFT) for the four bandwidths considered when event of interest is a cardiovascular event.

Table 6.11: Estimates and 95% credible (and confidence) intervals of RDD-AFT estimate of the effect of metformin prescription on type II diabetes complications.

Bandwidth:	*10*		*8*		*6*		*5*	
Outcome: Time to all cause mortality								
Method	Estimate	(95% C.I.)	Estimate	(95% C.I.)	Estimate	(95% C.I.)	Estimate	(95% C.I.)
Freq RDD-AFT	2.37	(0.75, 7.45)	2.09	(0.53, 8.19)	1.32	(0.37, 4.70)	1.21	(0.26, 5.65)
Bayes RDD-AFT	1.48	(0.85, 2.47)	1.41	(0.77, 2.42)	1.29	(0.70, 2.17)	1.25	(0.68, 2.14)
Outcome: Time to a cardiovascular event								
Method	Estimate	(95% C.I.)	Estimate	(95% C.I.)	Estimate	(95% C.I.)	Estimate	(95% C.I.)
Freq RDD-AFT	2.47	(0.57, 10.59)	1.25	(0.25, 6.25)	1.12	(0.13, 9.61)	1.22	(0.09, 15.92)
Bayes RDD-AFT	1.44	(0.79, 2.38)	1.30	(0.70, 2.26)	1.29	(0.69, 2.28)	1.27	(0.65, 2.24)

6.5 Conclusions

In this chapter, we presented Bayesian alternatives to the proposed methods for treatment effect estimation in Chapters 3, 4 and 5.

For continuous outcomes, a Bayesian alternative to the TPRS approach to LATE estimation was presented and we carried out simulation studies using the datasets that were simulated in Chapter 3 to compare the frequentist and Bayesian approaches to LATE estimation. Results from the simulation studies show that under the scenario where the underlying relationship between the outcome and the assignment variable is linear, both the frequentist and Bayesian methods yield estimates close to the value of the true treatment effect. For other scenarios, the relationships between the outcome and assignment variable are non-linear, the frequentist approach produces biased estimates for larger bandwidths. The Bayesian approach, however, produced estimates close to the value of the treatment effect for all bandwidths considered. We proceeded to apply the Bayesian approach to the real dataset on statin prescription. The estimates from the Bayesian and frequentist approaches suggest a beneficial effect of statin prescription on reducing LDL cholesterol level.

For binary outcomes, we presented a Bayesian alternative to RDD-RR estimation. Simulation studies, with varying levels of confounding and fuzziness, to compare the Bayesian and frequentist approaches to RDD-RR estimation were carried out and the estimates from the Bayesian and frequentist approaches are similar. We also observed that the two approaches produced similar estimates when they were applied to the real data on statin prescription.

For a time-to-event outcome, we presented a Bayesian alternative to RDD-AFT estimation and we carried out simulation studies, with varying levels of confounding and fuzziness, to compare Bayesian and frequentist methods. Under the weak fuzziness scenario, estimates and standard errors obtained from the Bayesian and frequentist approaches were similar. However, under the strong fuzziness scenario, estimates from the Bayesian approach were more precise than those from the frequentist approach as the standard errors from the Bayesian approach were smaller. The two approaches were applied to real data on metformin prescription to investigate the effect of metformin prescription on time to death and time to a CVD event. The

estimates from the Bayesian and frequentist approaches suggest a beneficial effect of metformin prescription. However, we note that, in general, estimates from the Bayesian approach were more precise because the 95% credible intervals from the Bayesian approach are narrower than the 95% confidence interval intervals from the frequentist approach, and this could have resulted from the prior distribution specified for the treatment effect.

In all cases, we have illustrated how to incorporate findings from other studies, in the form of prior distributions, in the implementation of the Bayesian alternatives of the frequentist methods.

Chapter 7

Final Discussion

This thesis focused on methods for treatment effect estimation from observational data using regression discontinuity designs, when the outcome of interest is continuous, binary or a time-to-event. Treatment effect estimation in an RD design when the outcome is continuous has been widely researched and the LATE is an established estimator of treatment effect. In contrast, treatment effect estimation in an RD design has not been widely explored for cases where the outcome of interest is either binary or a time-to-event, despite the extensive use of such outcomes in medical studies. We introduced new approaches for estimating treatment effects in an RD design for binary and time-to-event outcomes. We shall give a summary of the work that has been presented in this thesis and explore possible future extensions and directions.

7.1 Summary of work done

In Chapter 2, we provided a formal introduction to the RD design and presented the assumptions required to identify a treatment effect using an RD design. These assumptions were used for deriving the estimators that we proposed.

In Chapter 3, we introduced the treatment effect estimator when the outcome is continuous, known as the local average treatment effect (LATE) estimator. We described the linear and robust bias-corrected (BC) approaches of estimating the LATE. These two methods require the prior assumption/ specification of the underlying relationship between the outcome and the assignment variable. As such, the accuracy of the estimates from these methods might depend on whether or not the underlying relationship between the outcome and the assignment variable is cor-

rectly specified. We proposed an alternative approach to LATE estimation using thin plate regression spline (TPRS) that can model flexible, non-linear relationships where such relationships exist.

We carried out simulation studies to compare the three methods with varying relationships between the outcome and the assignment variable. The TPRS approach was able to yield estimates close to the value of the treatment effect for all scenarios. Overall, the TPRS approach can serve as a useful check to other methods, especially where it is suspected that the relationship between the outcome and the assignment variable is not linear.

The three methods were applied to a real dataset to estimate the effect of statin prescription on LDL cholesterol level. Results for the linear and TPRS models were similar and suggest a beneficial effect of statin prescription. The results from the robust BC approach also indicate a beneficial effect of statin prescription, but the estimates are imprecise as the standard error estimates are quite large.

In Chapter 4, we explored methods for treatment effect estimation when the outcome of interest is a binary variable with a focus on estimating the risk ratio. We discussed the Wald-RR and multiplicative structural mean model (MSMM) methods, which are existing methods that can be used for risk ratio estimation in an RD design. We proposed an estimator of the risk ratio based on the assumptions of the RD design that we termed RDD-RR.

Simulation studies were carried out to compare the performance of these methods for estimating the risk ratio with varying levels of confounding and fuzziness. The results from the simulation studies suggest that the three methods performed similarly for scenarios with no or low effect of confounding. However, it was observed that MSMM approach has a higher variability as its standard error estimates are consistently larger than the standard error estimates of the WALD-RR and RDD-RR methods. The MSMM approach also seemed to be affected by confounding as the estimates obtained from this approach under the high confounding scenario were biased.

Further simulation studies were carried out to check the performance of the WALD-RR approach when the treatment effect is large. The results confirm that the WALD-

RR approach yields biased estimates of the treatment effect. As such, the proposed RDD-RR approach appears to be a preferable method of estimating the risk ratio in a fuzzy RD design.

Additionally, the three methods discussed for estimating the risk ratio were applied to a real data on statin prescription. The outcome was whether or not the reduction in LDL cholesterol level is at least 1mmol/L. The estimates from the three methods suggest that patients that receive statins prescription are more likely to have their LDL cholesterol level reduced by at least 1mmol/L. However, for all methods, the risk ratio estimates were not statistically significant at the 5% level.

In Chapter 5, we proceeded to explore methods for treatment effect estimation when the outcome is a time-to-event. We focused on estimating the acceleration factor because it can be interpreted directly in terms of the outcome. We discussed the structural AFT approach, a method that has been used for the estimation of the acceleration factor in observational studies. Under the RD design framework, no estimator has been developed for the estimation of the acceleration factor. Therefore, we derived the RDD-AFT estimator for the acceleration factor based on the RD design assumptions.

To compare the performance of the S-AFT and RDD-AFT approaches under varying levels of confounding and fuzziness, simulation studies were carried out. The RDD-AFT yielded estimates that are close to the value of the true treatment effect for all scenarios considered. However, for scenarios where there is confounding, the S-AFT approach yields biased estimates of the treatment effect. The results from the simulation studies are not surprising because one of the assumptions of the S-AFT approach is that there are no unobserved confounders. As a result, we expect that the S-AFT approach will yield unbiased estimates of the treatment effect when there is no confounding but will yield biased estimates when the unobserved confounding is ignored.

This implies that for S-AFT approach to produce an unbiased estimate, the assumption of no unobserved confounders must be satisfied. On the other hand, the proposed RDD-AFT approach is able to yield estimates that are close to the value of the true treatment effect and a good coverage of the treatment effect in the presence

of unobserved confounding.

We proceeded to apply the two methods to the real data on metformin prescription with the outcomes time to all-cause-mortality and time to a CVD event in patients at risk of type II diabetes. Although the estimates from the methods are imprecise, as the 95% confidence intervals include 1, the point estimates of the acceleration factor suggest that the median times to all-cause-mortality and a CVD event in patients that receive metformin prescription are higher than in patients that did not receive metformin prescription.

In Chapter 6, we provided Bayesian alternatives to the methods we have proposed for treatment effect estimation in Chapters 3, 4 and 5 for the continuous, binary and time-to-event outcomes, respectively. Bayesian analysis provides a straightforward way of incorporating results from similar studies about the parameters of interest. In this thesis, we focus on treatment effect estimation from observational data, therefore, it is often the case that there have been previous studies that have looked at the effect of the treatment being investigated. As such, in the Bayesian analysis, we might incorporate the results from such previous studies via the prior distribution specification.

With the increasing availability of large databases in medicine, and the fact that many treatments are prescribed according to pre-specified, government guidelines, the use of RD designs for treatment effect estimation is appealing. The treatment effect estimation approaches we have proposed in this thesis can be estimated using standard statistical software, they are easily accessible and applicable to a wide variety of problems in medicine.

7.2 Future directions and extensions

In Chapter 3, we have explored the use of a flexible modelling approach using thin plate regression spline models that is completely data driven, and we have shown that this approach performs reasonably well. As such, it might be worth extending the flexible modelling approach to the estimators we have developed for binary and time-to-event outcomes.

For the binary outcome, it might be straightforward to fit a data-driven flexible model because the `mgcv` package (Wood, 2019) in `R` provides a generalised linear model option. As such, TPRS models can be fitted using logit links to estimate the components of the RDD-RR. For the time-to-event outcome, in addition to modelling the relationship between the outcome and covariates flexibly, the underlying hazard function could also be modelled flexibly as well. Marra et al. (2021) have proposed a flexible modelling approach for a time-to-event outcome under proportional hazards and proportional odds assumptions. It might be useful to extend the idea to the accelerated failure time assumption.

Typically, in clinical databases, patients are followed over a period of time, resulting in large longitudinal datasets. As a result, we may have some data in which the assignment variable (and other covariates) have multiple measurements over time. The treatment could also vary with time, which may yield a dynamic treatment strategy. Another extension to the work that has been presented in this thesis is to explore methods where threshold is measured at multiple points over time and model the assignment variable as a time-varying covariate to explore how changes in the values of the assignment variable over time may change a treatment effect estimate.

This extension might be particularly appealing for time-to-event outcomes because the structural AFT approach, which the RDD-AFT approach was compared with, has been extended to handle cases of dynamic treatment strategies and time varying confounders (Hernán et al., 2005). As we have established in the simulation studies, when there are unobserved confounders the S-AFT approach fails to produce reliable estimates of the treatment effect. The proposed RDD-AFT approach, however, provides a reasonable estimate and coverage of the treatment effect. Therefore, it may be useful to extend the RDD-AFT approach to handle dynamic treatment strategies and multiple measurements of the assignment variable, in both Bayesian and frequentist settings.

Appendix A

Derivation of the LATE estimator

The local average treatment effect estimator has been developed as an estimator of the treatment effect in a fuzzy RD design. We shall derive the LATE estimator below.

We define the variables in an RD design within a pre-defined bandwidth h:

- Assignment variable X,
- Threshold indicator Z,
- Treatment indicator A,
- Continuous outcome of interest Y.

Based on Assumption 4 that Z is independent of confounders conditional on X, we can obtain unbiased estimate of effect of Z on Y at the threshold:

$$\lim_{x \to x_0} \mathbb{E}(Y_i | Z_i = 1, X_i = x) - \lim_{x \to x_0} \mathbb{E}(Y_i | Z_i = 0, X_i = x).$$

For simplicity, we drop $\lim_{x \to x_0}$ and X in the equation above so that

$$\lim_{x \to x_0} \mathbb{E}(Y_i | Z_i = 1, X_i = x) \equiv \mathbb{E}(Y_i | Z_i = 1), \text{ and}$$
$$\lim_{x \to x_0} \mathbb{E}(Y_i | Z_i = 0, X_i = x) \equiv \mathbb{E}(Y_i | Z_i = 0).$$

Using the law of total probability, we have that

$$\mathbb{E}(Y_i | Z_i = 1) - \mathbb{E}(Y_i | Z_i = 0) =$$
$$\mathbb{E}(Y_i | Z_i = 1, A_i = 1)\mathbb{P}(A_i = 1 | Z_i = 1) + \mathbb{E}(Y_i | Z_i = 1, A_i = 0)\mathbb{P}(A_i = 0 | Z_i = 1) -$$
$$\left\{\mathbb{E}(Y_i | Z_i = 0, A_i = 1)\mathbb{P}(A_i = 1 | Z_i = 0) + \mathbb{E}(Y_i | Z_i = 0, A_i = 0)\mathbb{P}(A_i = 0 | Z_i = 0)\right\}.$$

Applying Assumption 5 - conditional independence of Y_i and Z_i:

$$
\begin{aligned}
&\mathbb{E}(Y_i|Z_i=1) - \mathbb{E}(Y_i|Z_i=0) \\
&= \mathbb{E}(Y_i|A_i=1)\mathbb{P}(A_i=1|Z_i=1) + \mathbb{E}(Y_i|A_i=0)\mathbb{P}(A_i=0|Z_i=1) \\
&- \{\mathbb{E}(Y_i|A_i=1)\mathbb{P}(A_i=1|Z_i=0) + \mathbb{E}(Y_i|A_i=0)\mathbb{P}(A_i=0|Z_i=0)\} \\
&= \mathbb{E}(Y_i|A_i=1)\mathbb{P}(A_i=1|Z_i=1) + \mathbb{E}(Y_i|A_i=0)\{1-\mathbb{P}(A_i=1|Z_i=1)\} \\
&- \{\mathbb{E}(Y_i|A_i=1)\mathbb{P}(A_i=1|Z_i=0) + \mathbb{E}(Y_i|A_i=0)\{1-\mathbb{P}(A_i=0|Z_i=0)\}\} \\
&= \mathbb{E}(Y_i|A_i=1)\{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)\} \\
&- \mathbb{E}(Y_i|A_i=0)\{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)\} \\
&= \{\mathbb{E}(Y_i|A_i=1) - \mathbb{E}(Y_i|A_i=0)\}\{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)\}.
\end{aligned}
$$

By rearranging, we have

$$
\mathbb{E}(Y_i|A_i=1) - \mathbb{E}(Y_i|A_i=0) = \frac{\mathbb{E}(Y_i|Z_i=1) - \mathbb{E}(Y_i|Z_i=0)}{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)}.
$$

By inserting $\lim_{x\to x_0}$ and X_i back into the equation, we have

$$
\begin{aligned}
&\lim_{x\to x_0} \mathbb{E}(Y_i|A_i=1, X_i=x) - \lim_{x\to x_0} \mathbb{E}(Y_i|A_i=0, X_i=x) \\
&= \lim_{x\to x_0} \frac{\mathbb{E}(Y_i|Z_i=1, X_i=x) - \lim_{x\to x_0}\mathbb{E}(Y_i|Z_i=0, X_i=x)}{\mathbb{P}(A_i=1|Z_i=1) - \mathbb{P}(A_i=1|Z_i=0)}.
\end{aligned}
$$

Hence, the LATE is recovered.

Appendix B

Probability of compliance

In this part, we explain why the denominator of the LATE is referred to as the probability of compliance. The probability of compliance can be defined as the probability that

$$\mathbb{P}(A = 1|Z = 1) - \mathbb{P}(A = 1|Z = 0) = 1.$$

That is, the probability of compliance is

$$\begin{aligned}
&\mathbb{P}\{\mathbb{P}(A = 1|Z = 1) - \mathbb{P}(A = 1|Z = 0) = 1\} \\
=&\mathbb{E}\{\mathbb{P}(A = 1|Z = 1) - \mathbb{P}(A = 1|Z = 0) \\
=&\mathbb{P}(A = 1|Z = 1) - \mathbb{P}(A = 1|Z = 0)
\end{aligned}$$

Therefore, the probability of compliance can be expressed as $\mathbb{P}(A = 1|Z = 1) - \mathbb{P}(A = 1|Z = 0)$.

References

Altman, D. G. (1991) Randomisation. *BMJ*, **302**, 1481 – 1482. Doi: 10.1136/bmj.302.6791.1481.

Altman, D. G. and Royston, P. (2006) The cost of dichotomising continuous variables. *BMJ*, **332**, 1080. Doi: 10.1136/bmj.332.7549.1080.

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

Baigent, C., Keech, A., Kearney, P., Blackwell, L., Buck, G., Pollicino, C., Kirby, A., Sourjina, T., Peto, R., Collins, R., Simes, R. and Cholesterol Treatment Trialists' (CTT) Collaborators (2005) Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet*, 1267–1278. Doi: 10.1016/S0140-6736(05)67394-1.

Basta, N. E. and Halloran, M. E. (2019) Evaluating the effectiveness of vaccines using a regression discontinuity design. *American journal of epidemiology*, **188**, 987 – 990.

Berrar, D. (2019) Introduction to the non-parametric bootstrap. In *Encyclopedia of Bioinformatics and Computational Biology* (eds. S. Ranganathan, M. Gribskov, K. Nakai and C. Schönbach), 766 – 773. Oxford: Academic Press. http://www.sciencedirect.com/science/article/pii/B9780128096338203506.

Blak, B. T., Thompson, M., Dattani, H. and Bourke, A. (2011) Generalisability of the health improvement network (thin) database: demographics, chronic disease prevalence and mortality rates. *Informatics in Primary Care*, **19**, 251 – 255.

Bor, J., Moscoe, E., Mutevedzi, P., Newell, M. L. and Bärnighausen., T. (2014) Regression discontinuity designs in epidemiology: Causal inference without randomized trials. *Epidemiology*, **25**, 729 –737.

Bourke, A., Dattani, H. and M, R. (2004) Feasibility study and methodology to create a quality-evaluated database of primary care data. *Informatics in Primary Care*, **12**, 171 – 177.

Bowden, R. J. and Turkington, D. A. (1990) *Instrumental Variables*. Cambridge University Press.

Buckley, J. and James, I. (1979) Linear regression with censored data. *Biometrika*, **66**, 429–436.

Burgess, S. (2016) Estimating and contextualizing the attenuation of odds ratios due to non-collapsibility. *Communications in Statistics - Theory and Methods*. `https://doi.org/10.1080/03610926.2015.1006778`.

Calonico, S., Cattaneo, M. D. and Farrell, M. H. (2020a) Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *Econometrics Journal*, **23**, 192 – 210.

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R. (2019) Regression Discontinuity Designs Using Covariates. *The Review of Economics and Statistics*, **101**, 442–451.

— (2020b) *rdrobust: Robust Data-Driven Statistical Inference in Regression-Discontinuity Designs*. R package version 0.99.9. `https://CRAN.R-project.org/package=rdrobust`.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014) Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, **82**, 2295–2326.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2020a) *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Cattaneo, M. D., Titiunik, R. and Vazquez-Bare, G. (2020b) The regression discontinuity design. In *The SAGE handbook of research methods in political science and international relations* (eds. L. Curini and R. Franzese), chap. 44, 835–857. London: SAGE Publications Ltd.

Caughey, D. and Sekhon, J. S. (2011) Elections and the regression discontinuity design: Lessons from close u.s. house races, 1942-2008. *Political Analysis*, **19**, 385–408.

Chan, P. H., xu, R. and Chambers, C. D. (2018) A study of r^2 measure under the accelerated failure time models. *Commun Stat Simul Comput*, **47**, 380–391.

Cho, Y., Hu, C. and Ghosh, D. (2019) Analysis of regression discontinuity designs using censored data. *ArXiv e-prints*.

Cholesterol Treatment Trialists' (CTT) Collaborators, Mihaylova, B., Emberson, J., Blackwell, L., Keech, A., Simes, J., Barnes, E. H., Voysey, M., Gray, A., Collins, R. and Baigent, C. (2012) The effects of lowering ldl cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet*, **380**, 581 – 590.

Clarke, P. S. and Windmeijer, F. (2010) Identification of causal effects on binary outcomes using structural mean models. *Biostatistics*, **11**, 756–770.

Collett, D. (2003) *Modelling Survival Data in Medical Research, Second Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.

Cook, T. D. (2008) Waiting for life to arrive: A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, **142**, 636 – 654.

Crainiceanu, C. M., Ruppert, D. and Wand, M. P. (2005) Bayesian analysis for penalized spline regression using winbugs. *Journal of Statistical Software*, **14**, 1–24. doi: 10.18637/jss.v014.i14.

Dawid, A. P. (1979) Conditional independence in statistical theory. *Journal of the Royal Statistical Society. Series B (Methodological)*, **41**, 1–31.

Didelez, V., Meng, S. and Sheehan, N. A. (2010) Assumptions of IV methods for observational epidemiology. *Statistical Science*, **25**, 22–40.

Duchon, J. (1977) Splines minimizing rotation-invariant semi-norms in sobolev spaces. In *Constructive Theory of Functions of Several Variables* (eds. W. Schempp and K. Zeller), 85–100. Berlin, Heidelberg: Springer.

Duncan, M. S., Vasan, R. S. and Xanthakis, V. (2019) Trajectories of blood lipid concentrations over the adult life course and risk of cardiovascular disease and all cause mortality: Observations from the framingham study over 35 years. *Journal of the American Heart Association*, **8**. 10.1161/JAHA.118.011433.

Efron, B. (1981) Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, **68**, 589–599. URL: https://doi.org/10.1093/biomet/68.3.589.

Efron, B. and Tibshirani, R. (1993) *An Introduction to the Boostrap*. Boca Raton, FL: Chapman & Hall/CRC.

Erikson, R. S., Folke, O. and Snyder, J. M. (2015) A gubernatorial helping hand? how governors affect presidential elections. *The Journal of Politics*, **77**, 491–504.

Gelman, A. and Rubin, D. B. (1992) Inference from iterative simulation using multiple sequences. *Statistical Science*, **1**, 451–511.

Geman, S. and Geman, D. (1984) Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-6**, 721–741.

Geneletti, S., O'Keeffe, A. G., Sharples, L. D., Ricahrdson, S. and Baio, G. (2015) Bayesian regression discontinuity designs: incorporating clinical knowledge in the

causal analysis of primary care data. *Statistics in Medicine*.

Geneletti, S., Ricciardi, F., O'Keeffe, A. G. and Baio, G. (2019) Bayesian modelling for binary outcomes in the regression discontinuity design. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **182**, 983–1002.

Greenland, S., Robins, J. M. and Pearl, J. (1999) Confounding and collapsibility in causal inference. *Statistical Science*, **14**, 26 – 46.

Hahn, J., Todd, P. and Van der Klaauw, W. (2001) Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, **69**, 201–209.

Han, Y., Xie, H., Liu, Y., Gao, P., Yang, X. and Shen, Z. (2019) Effect of metformin on all-cause and cardiovascular mortality in patients with coronary artery diseases: a systematic review and an updated meta-analysis. *Cardiovascular Diabetology*, **18**. https://doi.org/10.1186/s12933-019-0900-7.

Hastie, T. J. and Tibshirani, R. J. (1999) *Generalized Additive Models*. Chapman and Hall.

Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.

Hernán, M. A., Cole, S. R., Margolick, J. B., Cohen, M. H. and Robins, J. M. (2005) Structural accelerated failure time models for survival analysis in studies with time-varying treatments. *Pharmacoepidemiology and drug safety*, **14**, 477–491.

Hernán, M. A., Clayton, D. and Keiding, N. (2011) The simpson's paradox unraveled. *International Journal of Epidemiology*, **40**, 780–785.

Hernán, M. A. and Robins, J. M. (2006) Instruments for causal inference: an epidemiologist's dream? *Epidemiology*, **4**, 360–372.

Hippisley-Cox, J., Coupland, C., Vinogradova, Y., Robson, J., Minhas, R., Sheikh, A. and Brindle, P. (2008) Predicting cardiovascular risk in england and wales: prospective derivation and validation of qrisk2. *BMJ*, **336**, 1475–1482. Doi: 10.1136/bmj.39609.449676.25.

Huitfeldt, A., Stensrud, M. J. and Suzuki, E. (2019) On the collapsibility of measures of effect in the counterfactual causal framework. *Emerging Themes in Epidemiology*, **16**. https://doi.org/10.1186/s12982-018-0083-9.

Imbens, G. and Kalyanaraman, K. (2009) Optimal bandwidth choice for the regression discontinuity estimator. *Working Paper 14726*, National Bureau of Economic Research.

Imbens, G. and Lemieux, T. (2007) Regression discontinuity designs: A guide to practice. *Working Paper 337*, National Bureau of Economic Research.

Imbens, G. W. and Angrist, J. D. (1994) Identification and estimation of local average treatment effects. *Econometrica*, **62**, 467–475.

Kalbfleisch, J. D. and Prentice, R. L. (2002) *The statistical analysis of failure time data*. New York: John Wiley & Sons.

Larsson, S. C., Burgess, S., Mason, A. M. and Michaëlsson, K. (2020) Alcohol consumption and cardiovascular disease. *Circulation: Genomic and Precision Medicine*, **13**. doi: 10.1161/CIRCGEN.119.002814.

Lee, D. S. and Lemieux, T. (2010) Regression discontinuity designs in economics. *Journal of Economic Literature*, **48**, 281 – 355.

Linden, A., Adams, J. L. and Roberts, N. (2006) Evaluating disease management programme effectiveness: an introduction to the regression discontinuity design. *Journal of Evaluation in Clinical Practice*, **12**, 124–131.

Lunn, D., Jackson, C., Best, N., Thomas, A. and Spiegelhalter, D. (2012) *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press.

Marra, G., Farcomeni, A. and Radice, R. (2021) Link-based survival additive models under mixed censoring to assess risks of hospital-acquired infections. *Computational Statistics & Data Analysis*, **155**.

Marra, G. and Radice, R. (2020) Copula link-based additive models for right-censored event time data. *Journal of the American Statistical Association*, **115**, 886–895. Doi: 10.1080/01621459.2019.1593178.

Marra, G. and Wood, S. N. (2012) Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, **39**, 53 – 74.

Martinussen, T., Nørbo Sørensen, D. and Vansteelandt, S. (2017) Instrumental variables estimation under a structural cox model. *Biostatistics*.

Meyer, C. (2013) The bivariate normal copula. *Communications in Statistics - Theory and Methods*, **42**, 2402–2422. Doi: 10.1080/03610926.2011.611316.

Moscoe, E., Bor, J. and Bärnighausen, T. (2015) Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *Journal of Clinical Epidemiology*, **68**, 132 – 143.

Nelsen, R. (2006) *An Introduction to Copulas*. New York: Springer, 2 edn.

Newey, W. and McFadden, D. (1994) Chapter 36 Large sample estimation and hypothesis testing. vol. 4 of *Handbook of Econometrics*, 2111 – 2245. Elsevier. Doi: https://doi.org/10.1016/S1573-4412(05)80005-4.

NHS (2019a) High cholesterol. https://www.nhs.uk/conditions/high-cholesterol. Accessed August 11, 2020.

— (2019b) Metformin. https://www.nhs.uk/medicines/metformin/. Accessed April 10, 2021.

NICE (2008) Quick reference guide: Statins for the prevention of cardiovascular events.

— (2015) Type 2 diabetes in adults: management.

Nørgaard, M., Ehrenstein, V. and Vandenbroucke, J. P. (2017) Confounding in observational studies based on large health care databases: problems and potential solutions - a primer for the clinician. *Clinical Epidemiology*, **9**, 185–193.

O'Keeffe, A. G. and Baio, G. (2016) Approaches to the estimation of the local average treatment effect in a regression discontinuity design. *Scandinavian Journal of Statistics*.

O'Keeffe, A. G., Geneletti, S., Baio, G., Sharples, L. D., Nazareth, I. and Peterson, I. (2014) Regression discontinuity designs: an approach to the evaluation of treatment efficacy in primary care using observational data. *BMJ*.

Oldenburg, C. E., Moscoe, E. and Bärnighausen, T. (2016) Regression discontinuity for causal effect estimation in epidemiology. *Current Epidemiology Reports*, **3**, 233–241.

Palmer, T. M., Sterne, J. A. C., Harbord, R. M., Lawlor, D. A., Sheehan, N. A., Meng, S., Granell, R., Smith, G. D. and Didelez, V. (2011) Instrumental variable estimation of causal risk ratios and causal odds ratios in mendelian randomization analyses. *American Journal of Epidemiology*, **173**, 1392–1403.

Pearl, J. (2010) On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology*, **21**, 872–875.

Plummer, M. (2017) *JAGS Version 4.3.0 User Manual*. https://people.stat.sc.edu/hansont/stat740/jags_user_manual.pdf.

— (2019) *rjags: Bayesian Graphical Models using MCMC*. R package version 4-10 https://CRAN.R-project.org/package=rjags.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Ricciardi, F., Liverani, S. and Baio, G. (2020) Dirichlet process mixture models for regression discontinuity designs.

Rosenbaum, P. R. and Rubin, D. B. (1984) Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, **79**, 516–524.

Seber, G. A. F. and Lee, A. J. (2012) *Linear Regression Analysis*. Wiley Series in Probability and Statistics. Wiley. URL: `https://books.google.co.uk/books?id=X2Y6OkXl8ysC`.

Sheiner, L. B. and Rubin, D. B. (1995) Intention-to-treat analysis and the goals of clinical trials. *Clinical Pharmacology & Therapeutics*, **57**, 6–15.

Sjölander, A., Dahlqwist, E. and Zetterqvist, J. (2016) A note on the noncollapsibility of rate differences and rate ratios. *Epidemiology*, **27**, 356–359. Doi: `10.1097/EDE.0000000000000043`.

Sklar, A. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.

Smith, L. M., Lévesque, L. E., Kaufman, J. S. and Strumpf, E. C. (2017) Strategies for evaluating the assumptions of the regression discontinuity design: a case study using a human papillomavirus vaccination programme. *International Journal of Epidemiology*, **46**, 939–949.

Thistlethwaite, D. L. and Campbell, D. T. (1960) Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, **51**, 309–317.

van Leeuwan, N., Lingman, H. F., de Crean, A. J. M., Nieboer, D., Mooijaart, S. P., Richard, E. and Steyerberg, E. W. (2016) Regression discontinuity design: Simulation and application in two cardiovascular trials with continuous outcomes. *Epidemiology*, **27**, 503–511.

van Leeuwen, N., Lingsma, H. F., Mooijaart, S. P., Nieboer, D., Trompet, S. and Steyerberg, E. W. (2018) Regression discontinuity was a valid design for dichotomous outcomes in three randomized trials. *Journal of Clinical Epidemiology*, **98**, 70 – 79.

Wei, L. J. (1992) The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Statistics in Medicine*, **11**, 1871–1879.

Wilson, P. W. F., D'Agostino, R. B., Belanger, A. M., Silbershatz, H. and Kannel, W. B. (1998) Prediction of coronary heart disease using risk factor categories. *Circulation*, **97**, 1837–1847.

Wood, S. N. (2003) Thin-plate regression splines. *Journal of the Royal Statistical Society (B)*, **65**, 95–114.

— (2017) *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edn.

— (2019) *Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*. Version 1.8-31. `https://cran.r-project.org/web/packages/mgcv`.

— (2020) Inference and computation with generalized additive models and their extensions. *TEST*, **29**, 307 – 339. URL: `https://doi.org/10.1007/s11749-020-00711-5`.

Wooldridge, J. M. (2010) *Econometric analysis of cross section and panel data*. Cambridge, Mass.: MIT Press, second edition. edn.

Xu, K.-L. (2017) Regression discontinuity with categorical outcomes. *Journal of Econometrics*, **201**, 1 – 18.

Yang, S., Pieper, K. and Cools, F. (2020) Semiparametric estimation of structural failure time models in continuous-time processes. *Biometrika*, **107**, 123–136.