ABSTRACT

FAN, AILIN. New Statistical Methods for Precision Medicine: Variable Selection for Optimal Dynamic Treatment Regimes and Subgroup Detection. (Under the direction of Dr. Wenbin Lu and Dr. Rui Song.)

Due to patients' heterogeneity and a growing number of specifically targeted treatments, precision medicine draws attentions for customization of therapies and medical decisions for individual patient. In this dissertation, we investigate statistical methods to address two problems in precision medicine. The first problem is variable selection for optimal dynamic treatment regime. Variable selection is gaining more attention because it plays an important role in deriving practical and reliable optimal treatment regimes, especially when there are a large number of predictors. The second problem is subgroup detection of patients with enhanced treatment effects. By assessing heterogeneous treatment effects based on a variety of covariates, subgroup detection helps narrow down the target population of a treatment.

In Chapter 2, we develop a sequential advantage selection method for variable selection for optimal treatment regime. Variables that have qualitative interactions with treatment are of clinical importance for treatment decision-making. A qualitative interaction of a variable with treatment arises when the treatment effect changes direction as the value of the variable varies. Our sequential advantage selection method sequentially selects variables with a qualitative interaction and can be applied in multiple-decision-point settings. Numerical studies suggest that the proposed method is useful in identifying important variables under various underlying true models.

In Chapter 3, we propose a penalized A-learning method for deriving the optimal dynamic treatment regime when the number of covariates is of the non-polynomial order of the sample size. To preserve the double robustness property of the A-learning method, we adopt the Dantzig selector which directly penalizes the A-learning estimating equations. Simulation studies show that the proposed method achieves good performance in terms of both variable selection and estimated optimal treatment regimes.

In Chapter 4, we propose a systematic method for subgroup detection. Our method first tests the existence of a subgroup, and then identifies the subgroup if the null hypothesis on non-existence of such a subgroup is rejected. A semiparametric model for response is considered for model flexibility, and a doubly-robust test statistic is constructed based on this model. Moreover, a sample size calculation method for subgroup detection is developed based on the proposed statistic. Simulation studies are provided to illustrate the empirical performance of the proposed methods for subgroup detection and sample size calculation.

New Statistical Methods for Precision Medicine: Variable Selection for Optimal Dynamic Treatment Regimes and Subgroup Detection

by
Ailin Fan

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Statistics

Raleigh, North Carolina

2016

APPROVED BY:

Dr. Howard D. Bondell	Dr. Eric B. Laber
Dr. Wenbin Lu Co-chair of Advisory Committee	Dr. Rui Song Co-chair of Advisory Committee

DEDICATION

To my family.

BIOGRAPHY

The author was born in Huichang, Jiangxi, China in January 1991, where she spent the first thirteen years of her life. She graduated from Linchuan first high school in 2007 and attended Peking University afterwards. After receiving her bachelor's degree in statistics in 2011, she continued her study in the Department of Statistics at North Carolina State University. In 2013, she was granted the M.S. degree in Statistics en route to getting her Ph.D. degree. Under the direction of Drs. Lu and Song, she will complete her Ph.D. in Statistics in May 2016.

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude to my advisors, Dr. Wenbin Lu and Dr. Rui Song, for their excellent guidance and enormous support. They help me get started in statistical research, and show integrity and discretion in every step of research work. I have the highest regard for their passion in research and broad knowledge in statistics. They are able to express wonderful ideas in brief words, and are always accommodating and generous with their time. I could not have chosen two better mentors to guide me through my statistical research, and this dissertation would not be possible without their inspiration and encouragement.

I would also like to thank Dr. Howard Bondell and Dr. Eric Laber for serving on my dissertation committee. They have provided constructive comments that help me improve my work presented in this dissertation. I also thank Dr. James Wilson, who agreed to serve on my committee as the graduate school representative out of his busy schedule.

I appreciate the study opportunity provided by the Department of Statistics at North Carolina State University. I am grateful for all the faculty members for offering great courses on various topics that broaden my knowledge of statistics. I thank Dr. Daowen Zhang for being my academic advisor during my first two years' study. His valuable guidance helped me have a better plan of my Ph.D. study and research. I also enjoyed random talks with Dr. John Monahan and Dr. Dennis Boos in the hallways. I thank Dr. Charles Smith for his snacks and toys when I stay late in SAS Hall. I also thank DGPs and the department staff for their patience when I bug them with questions on requirements of the program. Special thanks to all my friends who share laughs and tears with me in these years. You have made graduate school enjoyable and memorable.

I am also thankful to the valuable graduate industrial traineeship and internship experiences at Bioventus, GlaxoSmithKline and SAS. Special thanks to my wonderful supervisors, John Jones, Rick Lewis, Alex Chien and Rain Xie. John is nice and encouraging, and I was impressed with his well-documented reports and good coding style. It was a pleasure to work with Rick;

he showed me how to collaborate and communicate with scientists from other fields of study, and made my experience at GSK very fruitful. Alex and Rain were energetic and inspiring, and gave me valuable advices on how to be a good statistical consultant and a good presenter when I was at SAS. I appreciate all their mentorship that sharpened my career aspirations.

Most importantly, none of this would have been made possible without the love and support of my family. My Mom and Dad have always been caring and supportive with their whole hearts. They have aided and encouraged me throughout this endeavor. I would also like to thank my boyfriend, Linbo, for sharing all the ups and downs in this journey. For everything they have done, I am extremely grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Precision medicine is emerging as a new medical paradigm for treatment that is tailored to an individual's genes, environment and lifestyle. It is of interest to develop statistical methods used in discovery and characterization of heterogeneity of subject responses to treatments. We tackle the problems in precision medicine applications under two frameworks: the first framework is to find the right treatment for each subject, which is conceptualized as optimal dynamic treatment regime; the second framework is to find the right subject for a given treatment, which is conceptualized as subgroup detection.

Under the first framework, we are interested in deriving the optimal treatment regime for the considered population of subjects, which could involve one treatment decision or a sequence of treatment decisions at multiple stages, and the latter one is referred to as the optimal dynamic treatment regime. Optimal dynamic treatment regime is a sequence of decision rules tailored through time to individual's information, and will maximize the final expected response when implemented. In practice, the sequential multiple assignment randomized trial (SMART) is an experimental design useful for deriving optimal dynamic treatment regimes (Murphy, 2005; Qian et al., 2013). The SMART designs have been used for some chronic or relapsing diseases. One example is the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial

(Fava et al., 2003; Rush et al., 2004). A large number of works have been developed to derive optimal dynamic treatment regimes based on data from clinical trials and observational studies. For example, marginal structure models (Robins, 1997; Murphy et al., 2001) allow estimation of the mean response under a dynamic treatment regime. Q-learning (Watkins, 1989; Watkins and Dayan, 1992; Murphy, 2005; Zhao et al., 2011; Chakraborty et al., 2010; Song et al., 2015) and A-learning (Murphy, 2003; Robins, 2004) are two popular backward induction methods for deriving optimal dynamic treatment regimes: the former builds regression models for the so-called Q functions while the latter is based on modeling contrast functions. In particular, A-learning has the double robustness property, i.e. when either the baseline mean function or the propensity score model is correctly specified, the resulting A-learning estimating equation for the contrast function is consistent.

A major challenge in deriving an optimal dynamic treatment regime arises when an extraordinary large number of prognostic factors are available, but not all of them are necessary for making treatment decision. This makes variable selection an emerging need in personalized medicine. Most existing variable selection techniques focus on selecting variables that are important for prediction. With such methods, some variables that are poor in prediction but are critical for treatment decision-making may be ignored.

Along this line, we explore and study the variable selection problem for optimal dynamic treatment regime in Chapter 2 and 3. We develop a sequential advantage selection method in Chapter 2 for selecting important variables for treatment decision making. The goal of this method is to select variables that have qualitative interactions with treatment. If the treatment effect changes direction as the value of the variable varies, this variable has qualitative interaction with treatment. Gunter et al. (2011) proposed the S-score method to characterize the magnitude of qualitative interaction of a variable with treatment. Our method is based on a modified S-score, which is named as sequential advantage. By selecting variables sequentially with the largest sequential advantage, and our method is guaranteed to pick variables with

qualitative interactions. To select the best candidate subset of variables for decision-making, we also propose a BIC-type criterion that is based on the sequential advantage. We evaluated the empirical performance of the proposed method in both single-decision-point and multiple-decision-point settings, as well as in an application to depression data from a clinical trial. In Chapter 3, we propose a penalized A-learning method for deriving the optimal dynamic treatment regime when the number of covariates is of the non-polynomial order of the sample size. To preserve the double robustness property of the A-learning method, we adopt the Dantzig selector which directly penalizes the A-leaning estimating equations. Empirical performance of the proposed approach is evaluated by simulations and illustrated with an application to a data from the STAR*D study.

Under the second framework, the interest lies in identifying a subgroup of subjects with an enhanced treatment effect. It is common to perform subgroup analysis in clinical trials. However, without careful designs and good practices, subgroup analysis is likely to suffer from false positive findings. It is important to develop a pre-specified data-driven approach to handle this problem. A number of data-driven approaches have been developed for the subgroup identification. Song and Pepe (2004) considered the binary response case and proposed using the selection impact curve (SIC) to evaluate treatment policies dictated by a single covariate. Then, based on the SIC, an optimal division of the population for assigning treatments can be obtained. Bonetti and Gelber (2004) grouped patients by values of a single covariate and estimated treatment effects on overlapping subsets of patients using a moving average procedure. Foster et al. (2011) developed a "Virtual Twins" method which first predicted the probabilities of response to treatment and control, and then used tree methods to obtain the subgroups with an enhanced treatment effect. Cai et al. (2011) and Zhao et al. (2013) proposed using parametric scoring systems based on multiple baseline covariates to rank treatment effects and then identified patients who benefit more from the new treatment. More recently, Shen and He (2015) considered a linear logistic-normal mixture model for the response and developed a

likelihood-based test for the existence of a subgroup.

In Chapter 4, we propose a systematic method for testing and identifying a subgroup with an enhanced treatment effect. We adopt a change-plane technique to first test the existence of a subgroup, and then identify the subgroup if the null hypothesis on non-existence of such a subgroup is rejected. A semiparametric model is considered for the response with an unspecified baseline function and an interaction between a subgroup indicator and treatment. A doubly-robust test statistic is constructed based on this model, and asymptotic distributions of the test statistic under both null and local alternative hypotheses are derived. Moreover, a sample size calculation method for subgroup detection is developed based on the proposed statistic. The finite sample performance of the proposed test is evaluated via simulations. Finally, the proposed methods for subgroup identification and sample size calculation are applied to a data from an AIDS study.

Chapter 2

VARIABLE SELECTION FOR OPTIMAL TREATMENT REGIME

2.1 Introduction

Personalized medicine is emerging as a new strategy for treatment that takes individual heterogeneity in background characteristics, clinical measurements, and genetic information into consideration. In this paradigm, treatment duration, dose, and type are adjusted over time and are tailored according to an individual's information with the aim of optimizing the effectiveness of treatment. This approach is different from the traditional "one-size-fits-all" treatment, which ignores the long-term benefits and individual heterogeneities. Great interest lies in finding optimal treatment regimes based on data from clinical trials and observational studies (e.g. Murphy, 2003; Robins, 2004; Moodie et al., 2007).

As the amount of information able to be collected on individuals continues to increase, more and more covariates are measured and are available in clinical studies. For example, a clinical trial may collect a large amount of information on a patient's demographics, medical history, intermediate outcomes, and side effects. However, it may be expensive or time-consuming to

collect all of this information in clinical practice, and redundancy in covariate information may impair the accuracy of optimal treatment decisions as well as its interpretation. Thus, a natural problem that arises in the estimation of optimal treatment regimes is how to identify the important covariates for treatment decision making.

Although variable selection is an important area in modern statistical research, current variable selection techniques mainly focus on selecting variables for prediction. Such approaches may not be able to adequately predict the interactions of variables with treatment and thus may neglect variables that are vital for decision making. In medical decision-making settings, variables that have qualitative interactions with treatments are clinically important (Peto, 1982). These variables are called prescriptive variables, which help prescribe the optimal treatment regimes. These variables should be distinguished from predictive variables, which help to increase prediction accuracy.

Scarce research has been done to study variable selection techniques for decision making. Qualitative interaction tests (Gail and Simon, 1985; Piantadosi and Gail, 1993; Yan, 2004) have been used to test a small number of expert determined pre-specified interactions. However, many of the tests were designed to test only qualitative interactions between categorical variables and treatments. Moreover, when the number of covariates is large, these tests are too conservative when controlling the error rate for multiple testing. Penalized methods have also been studied to identify variables important for making treatment decisions. Among others, Qian and Murphy (2011) developed a two-step procedure, where they first estimate the conditional mean response using the penalized least squares regression with L_1 penalty and then derive the estimated optimal treatment regimes from this estimated conditional mean. Lu et al. (2013) proposed a penalized least squares regression in an A-learning framework, which does not require the correct specification of the baseline mean model and directly selects variables with nonzero interactions with treatment. However, both methods do not directly target prescriptive variables that are important for treatment decision-making. Gunter et al. (2011) proposed a variable-ranking

measure that characterizes the qualitative interaction of an individual variable with treatment, namely the S-score. Then, a hybrid algorithm that combines S-score ranking and weighted LASSO was used to select variables for treatment decision-making.

Our work was motivated from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study (Fava et al., 2003; Rush et al., 2004). The STAR*D study was a sequential, multiple-assignment, randomized trial (SMART, see Murphy (2005); Qian et al. (2013)) for patients with non-psychotic major depressive disorder. This study aimed to determine which antidepressant medications, in what order, and what combination, should be given to patients to yield the optimal treatment effect. A large number of covariates were collected at baseline, such as patient demographic characteristics and medical history. In addition, several intermediate medical measurements were taken to assist in treatment decision making at the second or higher treatment decision points. It is hard to select covariates useful for making decisions from such a large number of covariates based on experts' opinions only. Thus, variable selection is crucial for deriving the optimal treatment regimes in the STAR*D trial.

We propose a variable selection method to identify prescriptive variables for deriving optimal treatment regimes with single-stage and multi-stage treatment decisions. In particular, we develop a quantity named *sequential advantage*, which can be viewed as a sequential S-score. This quantity characterizes additional information provided by a new variable to treatment decision-making, conditional on the effects of the covariates that are included from previous steps. We also propose a BIC-type criterion that is based on sequential advantage to choose the best candidate model for treatment decision-making. As sequential advantage measures the potential for a qualitative interaction with treatment, our method targets prescriptive variables.

Compared to the S-score method, our method is more accurate in the sense that it tends to select more prescriptive variables but selects fewer variables overall. This behavior is due to the sequential advantage selection, which can incorporate the correlation among variables. Thus, our method can largely exclude spurious variables that are marginally important but jointly

unimportant. The proposed method has satisfactory performance in each stage of dynamic treatment regimes. Because the proposed method starts from the null model, the implementation is feasible in high-dimensional settings provided that the true model is sufficiently sparse.

The remainder of the chapter is organized as follows. In Section 2.2, we introduce the framework for deriving optimal dynamic treatment regimes. In Section 2.3, S-score ranking for selecting prescriptive variables is introduced. Section 2.4 provides the proposed sequential advantage selection method for variable selection in optimal treatment decision-making. We demonstrate the method's performance in Section 2.5 by simulation studies in various scenarios and illustrate the method using data from the STAR*D clinical trial in Section 2.6.

2.2 Overview for Dynamic Treatment Regime

Potential outcome model is the framework for developing optimal dynamic treatment regimes. It was first introduced by Neyman et al. (1990) to analyze the causal effect of time-independent treatments in randomized studies. Rubin (1978) extended Neyman's work to the analysis of causal effects of time-independent treatments from observational data. Robins (1986) proposed a formal theory of causal inference based on their works, which extended Rubin's "point treatment" theory to longitudinal studies with direct and indirect effects and time-varying treatments and confounder. In this section, we will first introduce the notation and data structures for the dynamic treatment regime problem, then potential outcome models and assumptions will be provided to make causal inference based on the observed data. Lastly, the optimal dynamic treatment regimes will be defined using the potential outcomes.

2.2.1 Data Structure

Suppose treatment decisions are made at a finite number of times which we'll denote by $t_1, ..., t_K$, where t_1 is the baseline where treatment is first initiated. Here the number of treatments K

may be fixed or random, depending on the characteristic of the treatment. The data for a single individual can be summarized as a time-ordered sequence of variables

$$(\mathbf{X}_1, A_1, ..., \mathbf{X}_K, A_K, Y),$$

where $\mathbf{X}_1, ..., \mathbf{X}_K$ are covariate information measured prior to treatment at the beginning of each time point $t_1, .., t_K$, whose possible value sets are $\mathcal{X}_1, ..., \mathcal{X}_K$; and $A_1, ..., A_K$ are the treatments given at each time point $t_1, ..., t_K$ among a possible set of treatments denote by $\mathcal{A}_1, ..., \mathcal{A}_K$. The outcome of interest is Y with larger values indicating better response. The sets $\mathcal{X}_k$ $(k = 1, ..., K)$ and $\mathcal{A}_k$ $(k = 1, ..., K)$ could be either finite or infinite, and both can be multidimensional.

Overbar notation is used to denote the history of time-dependent variables. That is, $\bar{\mathbf{X}}_K = (\mathbf{X}_1, ..., \mathbf{X}_K), \bar{A}_k = (A_1, ..., A_k), k = 1, 2, ..., K-1$, and $\bar{X} = (\mathbf{X}_1, .., \mathbf{X}_K), \bar{A} = (A_1, ..., A_K)$. Upper case Roman letters denote random variables and lower case Roman letters denote the realizations of the random variables. Then the observed data are summarized as

$$O_i = (X_{1i}, A_{1i}, ..., X_{Ki}, A_{Ki}, Y_i), i = 1, ..., n,$$

where i indicates the ith individual in the sample. $\{O_i, i = 1, .., n\}$ are independent and identically distributed across i.

2.2.2 Potential Outcomes and Assumptions

To establish causal relationship between treatments and outcomes, we introduce the concepts of potential outcomes. Intuitively, the potential outcomes are the responses of a randomly chosen individual if they were given a particular treatment $\bar{a}_k$. Let $\bar{\mathcal{A}}_k = \mathcal{A}_1 \times \cdots \times \mathcal{A}_k$ be the collection of all possible combinations of treatments at time points $(t_1, ..., t_k)$; and let $\bar{\mathcal{A}} = \bar{\mathcal{A}}_K$. The set

of all potential outcomes is defined as

$$W = \left\{ \{ \mathbf{X}_1, \mathbf{X}_2^*(a_1), \mathbf{X}_3^*(\bar{a}_2), ..., \mathbf{X}_K^*(\bar{a}_{K-1}), Y^*(\bar{a}) \}, \text{ for } a_1 \in \mathcal{A}_1, ..., \bar{a}_{K-1} \in \bar{\mathcal{A}}_{K-1}, \bar{a} \in \bar{\mathcal{A}} \right\},$$

where $\mathbf{X}_k^*(\bar{a}_{k-1}), k = 2, ..., K$ are potential intermediate covariates that would accrue between t_{k-1} and t_k given the treatment history $\bar{a}_{k-1}$ $(k = 2, ..., K)$ and $Y^*(\bar{a})$ is potential final response that would result if treated according to $\bar{a}_K$. Note that the potential outcomes are defined for all possible treatments. They are also named as counter-factual variables because only the outcomes corresponding to the treatment regime given to the individual could be observed in reality; the other potential outcomes are corresponding to hypothetical values if the individual actually receives a treatment different from the one considered.

Suppose we have a treatment $\bar{a}$ which is corresponding to no interventions at all decision points, the causal effect of another treatment $\bar{a}'$ on the final outcomes at the individual level can be written as $Y^*(\bar{a}') - Y^*(\bar{a})$. Because only the potential outcomes at the particular treatment regime that is given to the individual can be observed, it is not possible to estimate the individual causal treatment effect. However, the treatment causal effect at population level, which is $E(Y^*(\bar{a}') - Y^*(\bar{a}))$, is possible to be estimated under the following assumptions on the potential outcomes:

(1) Stable Unit Treatment Value Assumption (SUTVA): an individual's outcome is the same as the potential outcome for the assigned treatment, and is not influenced by other individual's treatment allocation (Rubin, 1978). That is, if an individual receives treatment $\bar{A}$, we have $\mathbf{X}_k = \mathbf{X}_k^*(\bar{A}_{k-1}) = \sum_{\bar{a}_{k-1} \in \bar{\mathcal{A}}_{k-1}} \mathbf{X}_k^*(\bar{a}_{k-1}) I(\bar{A}_{k-1} = \bar{a}_{k-1}), k = 2, ..., K$, and $Y = Y^*(\bar{A}) = \sum_{\bar{a} \in \bar{\mathcal{A}}} Y^*(\bar{a}) I(\bar{A} = \bar{a})$.

(2) No unmeasured confounders: the treatment assignment is independent of the potential outcomes conditional on past information (Robins, 1997). That is, $A_k \perp W | \bar{\mathbf{X}}_k, \bar{A}_{k-1}$

holds for $k = 1, ..., K$, where A_0 is null.

The SUTVA is usually reasonable but cannot be verified generally, such as in a vaccine intervention situation, where the individual's outcome is clearly affected by the treatments of others. In epidemiology, the confounders are referred to as the variables that may be related both to the outcomes and the individuals who may get treatment. The no unmeasured confounders assumption states that treatment assignment is based on past recorded information only. This assumption holds in a sequentially randomized experiment, but may be questionable in an observational study. Under these two assumptions, the average treatment causal effect $E(Y^*(\bar{a}') - Y^*(\bar{a}))$ could be estimated by $\hat{E}(Y|\bar{A} = \bar{a}') - \hat{E}(Y|\bar{A} = \bar{a})$, which only depends on the observed data.

2.2.3 Optimal Dynamic Treatment Regimes

A dynamic treatment regime is a set of rules that shows how to treat an individual over time based on the past information. We denote the dynamic treatment regime as $d = (d_1, ..., d_K)$, where $d_k : \Gamma_k \to \mathcal{A}_k$ is a map of current available information at time t_k to the possible treatment decisions that could be made at t_k. Here $\Gamma_k = \{(\bar{x}_k, \bar{a}_{k-1}) \in \bar{\mathcal{X}}_k \times \bar{\mathcal{A}}_{k-1}\}$ is the set of historical information including covariates and treatments. The goal is to find a treatment regime that maximizes $E(Y^*(d))$, the mean potential outcome of a population if all the individuals are given treatments indicated by the treatment regime d.

To find the optimal dynamic treatment regime, we need to derive the distribution of the potential outcomes $\{\mathbf{X}_1, \mathbf{X}_2^*(d), ..., \mathbf{X}_K{}^*(d), Y^*(d)\}$ based on the distribution of the observed data $\{O_1, ..., O_n\}$. The two key assumptions in section 2.2.2 are required to estimate the effect of a dynamic regime. Under these assumptions, the joint density of the potential outcomes $\{\mathbf{X}_1, \mathbf{X}_2^*(d), ..., \mathbf{X}_K{}^*(d), Y^*(d)\}$ for any treatment regime d could be obtained as $P_{\mathbf{X}_1, ..., \mathbf{X}_K{}^*(d), Y^*(d)}(\mathbf{X}_1, ..., \mathbf{X}_K, y) = p_{\mathbf{X}_1}(\mathbf{x}_1) \times \cdots \times p_{Y|\bar{\mathbf{X}}_K, \bar{A}_k}\{y|\bar{\mathbf{x}}_K, \bar{d}_K(\bar{\mathbf{x}}_K)\}$. Hence the mean

response $E(Y^*(d))$ could be expressed as:

$$E\left[Y^*(\bar{a})_{a_1=d_1(\mathbf{X}_1),\ldots,a_K=d_K\{\bar{\mathbf{X}}_K(\bar{a}_{K-1}),\bar{a}_{K-1}\}}\right] = E[E[\cdots E[Y|\bar{\mathbf{X}}_K, \bar{A}_{K-1}, A_K = d_K]\cdots|\mathbf{X}_1, A_1 = d_1]]. \tag{2.1}$$

This is Robins's G-computation (Robins, 1986, 1997; Gill and Robins, 2001). See Gill and Robins (2001) and Murphy (2003) for more general conditions and proofs.

Suppose the class of regimes we consider is $\mathcal{D}$. To find the optimal dynamic treatment regime among $\mathcal{D}$, the treatments indicated by regime $d \in \mathcal{D}$ should be represented by the observed data. This requirement leads to another important assumption for indicating the optimal dynamic treatment regime: the positivity assumption (Robins, 1994). That is, the treatment patterns which are consistent with the regime $d \in \mathcal{D}$ can occur in the longitudinal data with positive probability: $P_{\bar{A}_k|\bar{\mathbf{X}}_K,\bar{A}_{k-1}}(a_k|\bar{s}_k, \bar{a}_{k-1}) > 0$, for $(\bar{s}_k, \bar{a}_{k-1})$ satisfies $P_{\bar{S}_k,\bar{A}_{k-1}}(\bar{s}_k, \bar{a}_{k-1}) > 0$ and $a_k = d_k(\bar{s}_k, \bar{a}_{k-1}), k = 1, ..., K$. In practice, optimal dynamic regimes could be applied to observational data or data from clinical trials. Recently, the experimental designs that are specifically suited for finding the optimal dynamic treatment regimes are of interest (Murphy, 2005; Collins et al., 2007), and the Sequential Multiple Assignment Randomized Trials (SMART, Murphy, 2005) design is widely used for constructing the optimal dynamic treatment regime (Thall et al., 2000; Rush et al., 2004; Auyeung et al., 2009; Chakraborty et al., 2010).

2.3 S-Score Ranking

When deriving optimal treatment regimes, only variables that have qualitative interaction effects with the treatment play a role. Gunter et al. (2011) pointed out two factors that affect the degree of a qualitative interaction: the magnitude of interaction between the variable and the treatment and the proportion of patients for whom the optimal treatment changes given the knowledge of the variable. Based on these two factors, they proposed the S-score, which

characterizes the degree of qualitative interaction of a variable. For single treatment decision A, the S-score for the jth covariate, X_j, is defined as:

$$S_j = \sum_{i=1}^{n} \left[\max_a \{\hat{\mathbb{E}}(Y_i | X_{ij} = x_{ij}, A_i = a)\} - \hat{\mathbb{E}}(Y_i | X_{ij} = x_{ij}, A_i = \hat{a}) \right], \tag{2.2}$$

where $\hat{\mathbb{E}}(Y_i | X_{ij} = x_{ij}, A_i = a)$ is an estimator of $\mathbb{E}(Y_i | X_{ij} = x_{ij}, A_i = a)$, and $\hat{a} = \text{argmax}_a \, \hat{\mathbb{E}}(Y | A = a)$, i.e., the treatment that leads to the largest treatment-specific mean response. The S-score is always non-negative, and a higher valued S-score indicates a greater potential for the covariate to have a qualitative interaction with treatment.

To show that the S-score captures both the magnitude of interaction and the proportion of subjects whose optimal treatment changes, we illustrate with an example. Consider the model $\mathbb{E}(Y | X_j, A) = \beta_0 + \beta_1 X_j + \beta_2 A + \beta_3 X_j A$, and let $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)^T$ denote the estimates of $(\beta_0, \beta_1, \beta_2, \beta_3)^T$. The S-score for X_j is then given by

$$S_j = \sum_{i=1}^{n} (\hat{\beta}_2 + \hat{\beta}_3 x_{ij}) \left[\mathbf{1}(\hat{\beta}_2 + \hat{\beta}_3 x_{ij} \geq 0) - \hat{a} \right]. \tag{2.3}$$

In equation (2.3), $(\hat{\beta}_2 + \hat{\beta}_3 x_{ij})$ represents the magnitude of the treatment effect as a function of X_{ij}, and $\mathbf{1}(\hat{\beta}_2 + \hat{\beta}_3 x_{ij} \geq 0) - \hat{a}$ indicates whether the optimal treatment for patient i changes given the knowledge of X_{ij}. Therefore, both factors are reflected in the S-score.

Although the S-score has very appealing properties for characterizing qualitative interaction of an individual covariate, there are some limitations with the S-score ranking. First, when the number of covariates is large, the S-score is not very effective for selecting qualitative interactions; variables that have no qualitative interaction with treatment can have non-zero S-scores due to correlations among covariates. In the algorithm proposed by Gunter et al. (2011), a weighted LASSO is used to select important interactions based on a linear model built on variables with non-zero S-scores, and the inverses of individual S-scores are used as the weights

in the weighted LASSO selection. In addition, an adjusted gain in value criterion is used to select the best subset of variables along the solution path for those selected non-zero S-scores. This hybrid algorithm helps to pick variables among the pool of variables with non-zero S-scores. Second, because the S-score evaluates each variable individually, some variables that are jointly crucial for optimal treatment decision-making may be neglected. Third, the S-score method proposed by Gunter et al. (2011) is only studied for a single-stage treatment decision. These limitations motivate us to develop a forward-selection procedure based on a modified S-score, named *sequential advantage*, for selecting variables having qualitative interactions with treatment for both single-stage and multi-stage treatment decisions.

2.4 Sequential Advantage Selection

In this section, we introduce sequential advantage and describe sequential advantage selection algorithms for both single-stage and multi-stage treatment decisions.

2.4.1 Sequential Advantage

We introduce sequential advantage in a single-stage treatment decision study. Let $\mathcal{M} = \{j^1, .., j^k\}$ denote an arbitrary model with $X_{j^1}, ..., X_{j^k}$ as the selected covariates and $\mathcal{F} = \{1, ..., p\}$ denote the full model. In addition, let $\mathbf{X}_i$ denote the covariate for subject i and $\mathbf{X}_{i(\mathcal{M})} = \{X_{ij} : j \in \mathcal{M}\}$ denote the associated covariates corresponding to model $\mathcal{M}$. The sequential advantage of variable X_j, $j \in \mathcal{F} \setminus \mathcal{M}^{(k-1)}$, is defined as:

$$\begin{aligned} S_j^{(k)} =& \frac{1}{n} \sum_{i=1}^{n} \Big[\max_a \{ \hat{\mathbb{E}}(Y | \mathbf{X}_{\mathcal{M}_j^{(k)}} = \mathbf{x}_{i\mathcal{M}_j^{(k)}}, A = a) \} \\ & - \hat{\mathbb{E}}(Y | \mathbf{X}_{\mathcal{M}_j^{(k)}} = \mathbf{x}_{i\mathcal{M}_j^{(k)}}, A = a_{opt}^{(k-1)}(\mathbf{x}_{i\mathcal{M}^{(k-1)}})) \Big], \end{aligned} \tag{2.4}$$

where $\mathcal{M}^{(k-1)} = \{j^1, ..., j^{k-1}\}$ is the model selected at the $(k-1)$th step, $\mathcal{M}_j^{(k)} = \mathcal{M}^{(k-1)} \cup \{j\}$, $\hat{\mathbb{E}}(Y|\mathbf{X}_{\mathcal{M}_j^{(k)}} = \mathbf{x}_{i\mathcal{M}_j^{(k)}}, A = a)$ is the estimated conditional mean response based on an assumed model with predictors $\mathbf{X}_{\mathcal{M}_j^{(k)}}$ and A, and $a_{opt}^{(k-1)}(\mathbf{x}_{i\mathcal{M}^{(k-1)}})$ is the optimal treatment regime obtained based on the variables in $\mathcal{M}^{(k-1)}$. In practice, a linear model with main effects of $\mathbf{X}_{\mathcal{M}_j^{(k)}}$ and A as well as interaction effects between $\mathbf{X}_{\mathcal{M}_j^{(k)}}$ and A can be used to obtain $\hat{\mathbb{E}}(Y|\mathbf{X}_{\mathcal{M}_j^{(k)}} = \mathbf{x}_{i\mathcal{M}_j^{(k)}}, A = a)$. Similarly, $a_{opt}^{(k-1)}(\mathbf{x}_{i\mathcal{M}^{(k-1)}})$ can be obtained based on the fitted model $\hat{\mathbb{E}}(Y|\mathbf{X}_{\mathcal{M}^{(k-1)}} = \mathbf{x}_{i\mathcal{M}_{(k-1)}}, A = a)$. The sequential advantage defined in (2.4) is similar to the S-score in spirit, but represents the additional benefit of including variable X_j to improve the optimal treatment regime estimated based on previously selected variables.

2.4.2 Sequential Advantage Selection Algorithm

In this section, we propose a variable selection method based on sequential advantage in a forward selection manner. We first describe the sequential advantage selection (SAS) algorithm for selecting variables that have a qualitative interaction with treatment in a single treatment decision study, and then extend SAS to accommodate multiple treatment decisions using Q-learning in the next section. The SAS algorithm for a single-stage treatment decision is given as follows.

(i) **Initialization**. Set $\mathcal{M}^{(0)} = \emptyset$. Compute $a_{opt}^{(0)} = \text{argmax}_a \hat{\mathbb{E}}(Y|A = a)$, and let $S^{(0)} = \hat{\mathbb{E}}(Y|A = a_{opt}^{(0)}) - \hat{\mathbb{E}}(Y)$.

(ii) **Sequential Advantage Selection**. In the kth step ($k \geq 1$), we have $\mathcal{M}^{(k-1)}$. For every $j \in \mathcal{F} \setminus \mathcal{M}^{(k-1)}$, we consider candidate covariates $\mathcal{M}_j^{(k)} = \mathcal{M}^{(k-1)} \cup \{j\}$ and compute the sequential advantage (2.4) corresponding to the jth covariate in the kth step. The kth variable to be selected is the one with the largest sequential advantage in this step: $j^k = \text{argmax}_{j \in \mathcal{F} \setminus \mathcal{M}^{(k-1)}} \{S_j^{(k)}\}$. Update $\mathcal{M}^{(k)} = \mathcal{M}^{(k-1)} \cup \{j^k\}$ and the estimated optimal treatment regime based on the first k selected variables $\mathbf{X}_{\mathcal{M}^{(k)}}$, i.e.,

$a_{opt}^{(k)}(\mathbf{x}_{\mathcal{M}^{(k)}}) = \text{argmax}_a \hat{\mathbb{E}}(Y|\mathbf{X}_{\mathcal{M}^{(k)}} = \mathbf{x}_{\mathcal{M}^{(k)}}, A = a)$. Let $S^{(k)} = S_{j^k}^{(k)}$.

(iii) **Selection of Best Subset**. Iterate Step 2 to obtain a solution path for the first m selected variables: $\mathcal{M}^{(m)} = \{j^1, ..., j^m\}$, where m is a predefined integer that is usually chosen to be less than $n/2$. We use a BIC-type criterion to select the best subset of variables:

$$\text{BIC}(l) = -\log(\sum_{i=0}^{l} S^{(i)}) + l\log(n)/n.$$

Let $\hat{m} = \text{argmin}_{0 \le l \le m} \text{BIC}(l)$. Then, $\mathcal{M}^{(\hat{m})}$ is the set of selected important variables for the treatment decision, and $a_{opt}^{(\hat{m})}(\mathbf{x}_{\mathcal{M}^{(\hat{m})}})$ is the estimated optimal treatment regime obtained based on the selected variables $\mathbf{X}_{\mathcal{M}^{(\hat{m})}}$.

In the SAS algorithm, $S^{(k)}$ is the sequential advantage based on the kth selected variable, and the proposed BIC-type criterion balances between the accumulated sequential advantages for making the optimal treatment decision and the size of the model.

2.4.3 Extension to Multi-Stage Treatment Decisions

For a study with multiple treatment decisions that has the data structure as shown in Section 2.2.1, we use a modified Q-learning algorithm to estimate the optimal dynamic treatment regime via backwards induction. We apply the SAS algorithm at each stage to select important variables for treatment decision-making and use these variables to model Q-functions. The sequential advantage selection algorithm for multiple treatment decisions is given as follows.

(i) At the Kth stage, the response is Y and the covariates are $H_K = \{\mathbf{X}_1, A_1, ..., A_{K-1}, \mathbf{X}_K\}$. Following the SAS algorithm, $\hat{m}_K$ variables are selected, and the set of indexes of selected variables is denoted by $\widehat{\mathcal{M}}_K$. The Q-function at the Kth stage based on the selected variables is

$$Q_K(h_{K,\widehat{\mathcal{M}}_K}, a_K) = \mathbb{E}(Y|H_{K,\widehat{\mathcal{M}}_K} = h_{K,\widehat{\mathcal{M}}_K}, A_K = a_K).$$

In addition, the contrast function is $C_K(h_{K,\widehat{\mathcal{M}}_K}) = Q_K(h_{K,\widehat{\mathcal{M}}_K}, 1) - Q_K(h_{K,\widehat{\mathcal{M}}_K}, 0)$. Then, the corresponding optimal treatment regime and value function at the Kth stage are:

$$d_K^{opt}(h_{K,\widehat{\mathcal{M}}_K}) = I\{C_K(h_{K,\widehat{\mathcal{M}}_K}) \geq 0\},$$

$$V_K(h_{K,\widehat{\mathcal{M}}_K}) = Y + C_K(h_{K,\widehat{\mathcal{M}}_K}) \left\{ d_K^{opt}(h_{K,\widehat{\mathcal{M}}_K}) - a_K \right\}.$$

(ii) At the kth stage ($k = K-1, ..., 1$), use $V_{k+1}(h_{k+1,\widehat{\mathcal{M}}_{k+1}})$ from the previous stage as the response, and the covariates at this stage are $H_k = \{\mathbf{X}_1, A_1, ..., A_{k-1}, \mathbf{X}_k\}$. Following the SAS algorithm, $\hat{m}_k$ variables are selected. Similar to the Kth stage, we can define $\widehat{\mathcal{M}}_k$ and derive the Q-function $Q_k(h_{k,\widehat{\mathcal{M}}_k}, a_k)$ and contrast function $C_k(h_{k,\widehat{\mathcal{M}}_k})$ based on the selected variables. Then, the corresponding optimal treatment regime and value function at the kth stage are:

$$d_k^{opt}(h_{k,\widehat{\mathcal{M}}_k}) = I\{C_k(h_{k,\widehat{\mathcal{M}}_k}) \geq 0\},$$

$$V_k(h_{k,\widehat{\mathcal{M}}_k}) = V_{k+1}(h_{k+1,\widehat{\mathcal{M}}_{k+1}}) + C_k(h_{k,\widehat{\mathcal{M}}_k}) \left\{ d_k^{opt}(h_{k,\widehat{\mathcal{M}}_k}) - a_k \right\}.$$

In the above algorithm, the value function is estimated based on the contrast function, which is different from the classical Q-learning algorithm where the value function is estimated based on the Q-function directly. Compared with the Q-function based value estimation, the contrast-function based value estimation is more robust in the sense that the baseline model is not required to be specified correctly. As the method targets prescriptive variables that have qualitative interaction with treatment, the contrast-function based value estimation is more suitable.

2.5 Simulation Studies

In this section, we conducted simulation studies to evaluate the performance of the proposed method in both single-stage and multi-stage treatment decisions studies.

2.5.1 Single-Stage Treatment Decision Study

The performance of the proposed SAS method is evaluated and compared with the S-score method and the method proposed by Lu et al. (2013) under various settings. The S-score method was implemented as follows: we first identified all variables with non-zero S-scores and ranked the importance of variables based on their S-scores in decreasing order. Finally, we selected the variables with the largest k non-zero S-scores, where k was chosen as the number of important prescriptive variables selected by our SAS method for easy comparison. Note that the focus here is to compare the performance of sequential advantage and S-score in terms of variable ranking. Therefore, the S-score method considered here is different from the original S-score method proposed by Gunter et al. (2011), which is a hybrid algorithm that combines S-score ranking and weighted LASSO selection.

For the method of Lu et al. (2013), we considered LASSO selection based on a least square loss with constant baseline, i.e.,

$$\min_{\alpha,\boldsymbol{\beta}} \sum_{i=1}^{n} \left[Y_i - \alpha - \{A_i - \pi(\mathbf{X}_i)\}\boldsymbol{\beta}^T \tilde{\mathbf{X}}_i \right]^2 + \lambda \sum_{j=0}^{p} |\beta_j|,$$

where $\pi(\mathbf{X}_i) = P(A_i = 1|\mathbf{X}_i)$ is the propensity score, $\tilde{\mathbf{X}}_i = (1, \mathbf{X}_i^T)^T$, and $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_p)^T$. In our simulations, $\pi(\mathbf{X}_i)$ is constant and is estimated by the sample proportion. This method was implemented using R-package `glmnet`, and the tuning parameter λ was chosen by the built-in cross validation. We refer this method as LASSO.

We consider the following four models to generate simulation data:

- Model I: $Y = 1+\gamma_1^T\mathbf{X}+A\boldsymbol{\beta}^T\tilde{\mathbf{X}}+\epsilon$ with $\gamma_1 = (1, -1, \mathbf{0}_{p-2})^T$, $\boldsymbol{\beta} = (0.1, 1, \mathbf{0}_7, -0.9, 0.8, \mathbf{0}_{p-10})$;

- Model II: $Y = 1 + 0.5\sin(\pi\gamma_1^T\mathbf{X}) + 0.25(1 + \gamma_2^T\mathbf{X})^2 + A\boldsymbol{\beta}^T\tilde{\mathbf{X}} + \epsilon$ with $\gamma_1 = (1, -1, \mathbf{0}_{p-2})^T$, $\gamma_2 = (1, \mathbf{0}_2, -1, \mathbf{0}_5, 1, \mathbf{0}_{p-10})^T$, and $\boldsymbol{\beta}$ being the same as in Model I;

- Model III: $Y = 1+\gamma_1^T\mathbf{X}+A\boldsymbol{\beta}^T\tilde{\mathbf{X}}+\epsilon$ with $\gamma_1 = (1, -1, \mathbf{0}_{p-2})^T$, $\boldsymbol{\beta} = (0.1, 1, \mathbf{0}_7, -0.9, 0.8, \mathbf{0}_{10}, 1, 0.8, -1, \mathbf{0}_5, 1, -0.8, \mathbf{0}_{p-30})$;

- Model IV: $Y = 1 + 0.5\sin(\pi\gamma_1^T\mathbf{X}) + 0.25(1 + \gamma_2^T\mathbf{X})^2 + A\boldsymbol{\beta}^T\tilde{\mathbf{X}} + \epsilon$ with γ_1 and γ_2 being the same as in Model II, and $\boldsymbol{\beta}$ being the same as in Model III.

Although all four models have linear interaction forms between covariates and treatment, they have different functional forms for the baseline effects. In our SAS method, the forward selection is based on the working model: $\mathbb{E}(Y) = \boldsymbol{\gamma}^T\tilde{\mathbf{X}}+A\boldsymbol{\beta}^T\tilde{\mathbf{X}}$, which is correctly specified under Models I and III but is misspecified under Models II and IV. Models I and II have three important prescriptive variables (X_1, X_9, X_{10}), while Models III and IV have eight important prescriptive variables $(X_1, X_9, X_{10}, X_{21}, X_{22}, X_{23}, X_{29}, X_{30})$. Covariates $\mathbf{X} = (X_1, ..., X_p)^T$ are generated from a multivariate normal distribution: each entry is normal with mean zero, variance one, and the correlation between covariates is $\text{Corr}(X_j, X_k) = \rho^{|j-k|}$, for $j \neq k$, $j, k = 1, ..., p$. Here, ρ is chosen to be 0.2, 0.5, and 0.8, representing weak, moderate, and strong correlations. We considered randomized trials, where A is generated from a Bernoulli distribution with the success probability of 0.5. The error term, ϵ, is normally distributed with mean zero and variance 0.25. We ran 500 simulations for each scenario with $n = 200$ and $p = 1000$.

Because the generative models are complex, it becomes rather difficult to evaluate the degree of qualitative interaction of each variable with treatment. As an illustration, we show in Figure 2.1 the marginal interaction plots of variables X_1, X_9, and X_{10} with treatment under two scenarios: $\rho = 0.2$ and $\rho = 0.8$. These marginal plots are for one simulated data under model I, where X_1, X_9, and X_{10} are important prescriptive variables. Based on Figure 2.1, when the

correlation is weak ($\rho = 0.2$), all variables show clear qualitative interaction with treatment; when the correlation is strong ($\rho = 0.8$), either variable X_9 or X_{10} (here is X_9) has nearly no qualitative interaction with treatment. This is possibly due to the fact that these two variables have strong positive correlation but opposite covariate effects. This result implies that the S-score method may fail to identify one of the variables because the method relies on the measures for the marginal qualitative interaction.

Table 2.1 summarizes simulation results for variable selection and estimated optimal treatment regimes of the three methods. For variable selection, we report size and true positive (TP), which are the average numbers of selected variables and correctly identified prescriptive variables over 500 simulations, respectively. For assessing estimated optimal treatment regimes, we compute the mean value ratio between the value following the estimated optimal treatment regime, $Q(\hat{g}^{opt})$, and the value following the true optimal treatment regime, $Q(g^{opt})$, denoted by VR $= Q(\hat{g}^{opt})/Q(g^{opt})$. Here, the value of a given treatment regime is computed by averaging outcomes generated from the true model with the treatment dictated by the considered regime using Monte Carlo simulations with 10,000 replicates. In addition, we report the mean error rates of the estimated optimal treatment regimes for treatment decision-making compared with the true optimal treatment regimes, denoted by ER. The numbers given in parentheses are the associated sample standard deviations.

The results in Table 2.1 show that the SAS method selects more true prescriptive variables in most cases but with fewer selected variables. For example, under model I, SAS has size $=$ 6.7 and TP $= 2.98$, whereas S-score has TP $= 1.61$ and LASSO has TP $= 1.75$ with size $=$ 21.94. In addition, it is observed that there are too many variables with non-zero S-scores and that the LASSO method tends to select more variables than the SAS method, especially when $\rho = 0.2$ and 0.5. Compared to the marginal S-score method, the SAS method includes more true positives in most cases, which indicates that the sequential advantage is a better characterization of prescriptive variables than the S-score. Under Model I with weak and moderate correlations,

the SAS method can recover almost all of the important variables. However, for the other three models, all three methods missed a few important variables due to the weak signals of these variables and/or model misspecification.

Based on results about values and error rates on estimated optimal treatment regimes in Table 2.1, the SAS method provides good estimates of optimal treatment regimes with values close to the true optimal values and low error rates among all three methods. The error rates provided by the LASSO method are high in most cases; this is partly because the LASSO estimates tend to have large bias due to shrinkage. As correlation increases, the values of estimated treatment regimes are less affected by the quality of the variable selection because some variables may be good surrogates for the true variables when estimating the optimal treatment regime.

We also compare solution paths of the three methods in Figure 2.2. Here, we define a solution path as the trajectory of the number of identified important variables as the number of selected variables increases according to the selection order. For demonstration purpose, we only plot the solution paths for the first 30 selected variables. The SAS method has a natural order of selected variables. For the S-score method, we ranked the variables in descending order of the S-scores of these variables. The LASSO method has a solution path of $\boldsymbol{\beta}$, which can be used to determine the order of variables entering the model. The solution path plots allow us to evaluate the ability of each method to identify important variables given that the same number of variables is selected.

Figure 2.2 indicates that when the size is fixed, the SAS method includes the largest number of important variables in most cases. When the data are generated under model I, the SAS method can include all of the important variables quickly under weak and moderate correlations. However, under strong correlations, they are likely to be missed by all three methods because some important variables are highly correlated. When the model is misspecified, the SAS method is slightly better in Model II, while all three methods do not differ significantly in Model IV.

Overall, the SAS method performs well on both the aspect of variable selection and the aspect of estimating optimal treatment regime. The SAS method can select most important variables at a moderate size of the selected variables. When the model is correctly specified and the correlations between covariates are not too high, the SAS method is able to identify all important variables. Moreover, the error rates of the optimal treatment regime based on the SAS method are low, and the estimated values are close to the true optimal value.

2.5.2 Multi-Stage Treatment Decisions Study

To illustrate the sequential advantage selection algorithm for multi-stage treatment decisions (MTD) study, we applied the SAS algorithm to simulated data with two-stage treatment decisions based on the following generative model for the final response:

$$Y = A_1 A_2 + A_2(a + \boldsymbol{\beta}_{12}^T \mathbf{X}_1 + \boldsymbol{\beta}_{21}^T \mathbf{X}_2) + A_1(a + \boldsymbol{\beta}_{11}^T \mathbf{X}_1) + \epsilon, \tag{2.5}$$

where A_k, the treatment at stage k, follows a Bernoulli distribution with parameter 0.5 for $k = 1$ and 2. The covariates collected at baseline, $\mathbf{X}_1$, include $p_1 = 500$ variables and are denoted as $\mathbf{X}_1 = (X_{1,1}, X_{1,2}, ..., X_{1,p_1})^T$. We generate $\mathbf{X}_1$ from a multivariate normal distribution with mean zero, variance one, and correlation $\text{corr}(X_{1,j}, X_{1,l}) = 0.2^{|j-l|}, j \neq l$. The intermediate covariates collected at the second stage are denoted by $\mathbf{X}_2$. For demonstration purposes, we consider a one-dimensional intermediate covariate X_2 and assume that $X_2 = c_0 + c_1 X_{1,1} + c_2 A_1 + c_3 A_1 X_{1,1} + e$, where the normal random error e has mean zero and variance σ_2^2. The random error for response Y, ϵ, is normally distributed with mean zero and variance σ_1^2.

The parameter values for the above two-stage model are chosen as follows: $\boldsymbol{\beta}_{12} = (0, 0, 1, -1, \mathbf{0}_{p_1-4})^T$, $\beta_{21} = 1$, $\boldsymbol{\beta}_{11} = (\mathbf{0}_4, 1, -1, \mathbf{0}_{p_1-6})^T$, $a = 0$. For the standard deviations of two random errors, we choose $\sigma_1 = \sigma_2 = 0.5$. For the parameter $\mathbf{c} = (c_0, c_1, c_2, c_3)^T$ in the model for X_2, we consider three sets of values to evaluate the carry-on effects of baseline variables through

intermediate covariates: $\mathbf{c} = (0,1,0,0)^T$, $(0,0,1,0)^T$ and $(0,1,1,1)^T$.

Based on the generative model (2.5), it is clear that the optimal treatment regime at stage 2 is $g_2^{opt}(\mathbf{x}_1, a_1, x_2) = 1(a_1 + \boldsymbol{\beta}_{12}^T \mathbf{x}_1 + \beta_{21} x_2 \geq 0)$. Thus, four variables $(X_2, A_1, X_{1,3}, X_{1,4})$ determine the optimal treatment regime at stage 2. At stage 1, the Q-function is

$$\begin{aligned} Q_1(\mathbf{X}_1, A_1) &= E\left\{|A_1 + \boldsymbol{\beta}_{12}^T \mathbf{X}_1 + \beta_{21} X_2|_+ \big| \mathbf{X}_1, A_1\right\} + A_1(\boldsymbol{\beta}_{11}^T \mathbf{X}_1) \\ &= \sigma_2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{\mu_1^2}{2\sigma_2^2}\} + \mu_1[1 - \Phi(-\mu_1/\sigma_2)] + A_1(\boldsymbol{\beta}_{11}^T \mathbf{X}_1), \end{aligned}$$

where $\mu_1 = A_1 + \boldsymbol{\beta}_{12}^T \mathbf{X}_1 + \boldsymbol{\beta}_{21}[c_0 + c_1 X_{1,1} + c_2 A_1 + c_3 A_1 X_{1,1}]$. The optimal treatment regime at stage 1 is $g_1^{opt}(\mathbf{x}_1) = 1\{Q_1(\mathbf{x}_1, 1) > Q_1(\mathbf{x}_1, 0)\}$. There are five important variables $(X_{1,1}, X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6})$ for determining the optimal treatment regime at stage 1 when $\mathbf{c} = (0,1,0,0)^T$ and $(0,1,1,1)^T$, and four important variables $(X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6})$ when $\mathbf{c} = (0,0,1,0)^T$. Table 2.2 summarizes the optimal treatment regimes and the important variables in this simulation study. Although the optimal treatment regime and the corresponding important variables at stage 2 are explicitly defined, the optimal treatment regime at stage 1 takes a complex non-linear form. Therefore, the effects of the important variables are difficult to evaluate.

We applied the SAS algorithm to the simulated data for sample sizes n = 100, 200, and 400 over 100 replications. Simulation results are summarized in Tables 2.3 and 2.4. Table 2.3 presents results on variable selection and estimated optimal treatment regimes at both stages 1 and 2, where the same statistics as in Table 2.1 are reported (Size, TP, VR and ER). For the mean value ratio at stage 2 (VR2), we adopt random treatment regimes at stage 1 to calculate the outcome values because the optimal treatment regimes at stage 1 have not been estimated at this stage. Table 2.4 reports the proportions of each important variable being selected for all scenarios.

According to Table 2.3, the numbers of true positives at stage 2 increase and get close to the true number (four) when the sample size gets large; at stage 1, the numbers of true positives

also increase, but 1-2 variables are missed. We will examine which variables are missed when analyzing results in Table 2.4. Based on the results for values and error rates, SAS provides good estimated optimal treatment regimes at both stages 1 and 2 when the sample size is large. We note that performances for $\mathbf{c} = (0, 1, 0, 0)^T$ at stage 1 are worse than the other scenarios. This indicates that the manner in which the intermediate variable depends on covariates from the last stage also affects the quality of the estimated optimal treatment regime. It is not apparent why the case with $\mathbf{c} = (0, 1, 0, 0)^T$ performs worse, and results in Table 2.4 partially explain this phenomenon.

Table 2.4 shows more detailed variable selection results. At stage 2, all but A_1 among the four important variables can almost always be selected when the sample size is large. A_1 can be selected more often for $\mathbf{c} = (0, 1, 0, 0)^T$ than for the other two scenarios; this may be because X_2 depends on A_1 in these cases, which partially eliminates the effects of A_1 on the final response, Y. At stage 1, only variables $X_{1,5}$ and $X_{1,6}$ can always be selected for all three scenarios when $n = 400$. These two variables appear in the optimal treatment regime at stage 1 in a linear form. On the contrary, variables $X_{1,3}$, $X_{1,4}$, and $X_{1,1}$ that present in a non-linear form are not always selected. The probabilities of selecting $X_{1,3}$ and $X_{1,4}$ for all three scenarios are low. This may be because these two variables do not have substantial effects on the optimal treatment regime; the high values and low error rates at stage 1 in Table 2.3 also verify this argument. The probabilities of including $X_{1,1}$ differ between the first two scenarios. A possible explanation is that $X_{1,1}$ interacts with A_1 in μ_1 for the first scenario, which makes its sequential advantage for being selected large. This may also explain why the scenario with $\mathbf{c} = (0, 1, 1, 1)^T$ performs better than the scenario with $\mathbf{c} = (0, 1, 0, 0)^T$ in Table 2.3.

Based on these results, the SAS algorithm performs well on both variable selection and optimal treatment regime estimation. The complex form of the optimal treatment regime at stage 1 makes it more difficult to identify important variables and brings a challenge for variable selection.

2.6 Application to STAR*D Study

We apply the proposed method to data from the STAR*D study, which was conducted to determine the effectiveness of different treatments for patients with major depressive disorder (MDD) who had not been adequately benefiting from initial treatment with an antidepressant. There were 4041 participants (age 18-75) with nonpsychotic MDD enrolled in this study. Initially, these participants were treated with citalopram (CIT) up to 14 weeks. Subsequently, 3 more levels of treatments were provided for participants without a satisfactory response to CIT. At Level 2, participants were eligible for seven treatment options, which may be conceptualized as two treatment strategies: medication or psychotherapy switch, and medication or psychotherapy augmentation. Available treatments for participants to switch were: sertraline (SER), venlafaxine (VEN), bupropion (BUP) and cognitive therapy (CT); available treatments for patients to augment were: augmenting CIT with bupropion (CIT+BUP), buspirone (CIT+BUS) or cognitive therapy (CIT+CT). Participants without a satisfactory response to CT were provided additional medication treatments, which is called Level 2A. All participants who did not respond satisfactorily at Level 2 or 2A were eligible for Level 3, where possible treatments were medication switch to mirtazapine (MIRT) or nortriptyline (NTP), and medication augmentation with either lithium (Li) or thyroid hormone (THY). Participants without satisfactory response to Level 3 were re-randomized at Level 4 to either tranylcypromine (TCP) or a combination of mirtazapine and venlafaxine (MIRT+VEN). Participants who responded satisfactorily were followed up to 1 year. See Fava et al. (2003) and Rush et al. (2004) for more detailed description of this STAR*D design.

For illustration, we focus on a subset of participants who were given treatment BUP or SER at Level 2, did not receive satisfactory responses, and were randomized to treatment MIRT or NTP at Level 3. There were 73 participants who meet this condition. Among these participants, 36 were treated with BUP and 37 were treated with SER at Level 2, and 33 were

treated with NTP and 40 were treated with MIRT at Level 3. Our goal is to identify relevant prescriptive predictors and estimate optimal dynamic treatment regimes at Levels 2 and 3 that maximize the mean response at the end of Level 3. We consider 381 covariates as possible relevant predictors, which are listed in Table 2.5. These covariates include participant features such as age, gender, socioeconomic status, and ethnicity; illness features such as medication history and family history of mood disorders; and care features such as clinician type. Intermediate medical conditions from Levels 1 and 2, such as degree of symptom improvement and side effect burden, are also considered. For treatment regime at Level 3, all 381 covariates and the treatment at Level 2 are considered as possible predictors. For the treatment regime at Level 2, the intermediate medical conditions at Level 2 are no longer available, thus there are only 305 covariates considered for treatment decision making. We used negative 16-item Quick Inventory of Depressive Symptomatology-Clinician-Rated (QIDS-C_{16}) at the end of Level 3 as the final response, which is a measurement of symptomatic status. Because low QIDS-C_{16} stands for remission, the negative QIDS-C_{16} was used such that larger value indicates better response.

We apply the SAS algorithm to this data set. The results are as follows. At Level 3, there are four covariates selected based on the BIC criterion: "ringing in ears" in patient rated inventory of side effects at Level 2 (EARNG-Level2), "hard to control worrying" in psychiatric diagnostic screening questionnaire at baseline (WYCRL), "feeling of worthlessness or guilt" in baseline protocol eligibility (DSMFW), and "fatigue or loss of energy" in baseline protocol eligibility (DSMLE). All four covariates are binary covariates with 1 indicating "Yes" and 0 indicating "No". The estimated optimal treatment regime is $I(-18.57 + 13.79 \times (\text{EARNG-Level2}) - 8.46 \times \text{WYCRL} + 6.36 \times \text{DSMFW} + 16.88 \times \text{DSMLE} \geq 0)$, where 1 stands for treatment NTP and 0 stands for treatment MIRT. This optimal treatment regime assigns 25 participants to NTP and the remaining 48 participants to MIRT. At Level 2, there are seven covariates selected based on the BIC criterion: "TE flashbacks of traumatic event" in the psychiatric diagnostic screening questionnaire at baseline (TEFSH), "EM worry saying something stupid" in the psychiatric

diagnostic screening questionnaire at baseline (EMSTP), "QIDS psychomotor agitation" in the quick inventory of depressive symptomatology - clinician at Level 1 (CAGIT), "think drink too much" in the psychiatric diagnostic screening questionnaire at baseline (DKMCH), "QIDS outlook (self)" in the quick inventory of depressive symptomatology - clinician at Level 1 (CVWSF), "IM convinced others spying" in the psychiatric diagnostic screening questionnaire at baseline (IMSPY), and "sleep at least 1-2 hours less 2 weeks" in the psychiatric diagnostic screening questionnaire at baseline (LSL2W). Among these seven covariates, TEFSH, EMSTP, DKMCH, IMSPY, and LSL2W are binary, with 1 indicating "Yes" and 0 indicating "No"; CAGIT and CVWSF are categorical covariates with 4 levels indicated by 0 to 3. The estimated optimal treatment regime is $I(-5.50 + 3.91 \times \text{TEFSH} + 11.17 \times \text{EMSTP} + 3.76 \times \text{CAGIT} - 4.65 \times \text{DKMCH} + 4.29 \times \text{CVWSF} + 6.57 \times \text{IMSPY} - 8.48\text{LSL2W} \geq 0)$, where 1 stands for treatment BUP and 0 stands for treatment SER. This optimal treatment regime assigns 39 participants to BUP and the remaining 34 participants to SER.

To further examine the estimated optimal dynamic treatment regime, we estimate the value of the estimated optimal dynamic treatment regime, that is, the mean outcome following the estimated optimal treatment regime, using the inverse probability weighted estimator proposed by Zhang et al. (2013), defined as

$$\text{IPW} = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i I(A_{i,1} = g_1(\mathbf{X}_i), A_{i,2} = g_2(\mathbf{X}_i))}{\pi(A_{i,1})\pi(A_{i,2})}.$$

Here Y_i is the outcome for ith individual, $A_{i,1}$ and $A_{i,2}$ are the treatments given to the ith individual at stage 1 and stage 2, respectively, $g_1(\mathbf{X}_i)$ and $g_2(\mathbf{X}_i)$ are the estimated treatment regimes at stages 1 and 2, and $\pi(A_{i,1})$ and $\pi(A_{i,2})$ are the probabilities of receiving treatment $A_{i,1}$ at stage 1 and treatment $A_{i,2}$ at stage 2, respectively. The estimated value for the estimated optimal dynamic treatment regime based on the SAS algorithm is compared to the estimated values when all subjects are treated with the non-dynamic treatment regimes: BUP+NTP,

BUP+MIRT, SER+NTP and SER+MIRT. The estimated values are shown in Table 2.6. We also report the 95% confidence intervals for the differences between values of the estimated optimal dynamic treatment regime and the four non-dynamic treatment regimes based on 1,000 bootstrap samples. The results show that the value of the estimated optimal dynamic treatment regime based on the SAS algorithm is significantly larger than those of the non-dynamic treatment regimes.

2.7 Discussion

In this article, we propose a forward-stepwise variable selection method based on sequential advantage for deriving optimal treatment regimes in both single-stage and multi-stage treatment decision studies. Our method generalizes S-score ranking and directly targets prescriptive variables that are important for decision making. We also propose a BIC-type criterion to select the number of important prescriptive variables needed for treatment decision making. The proposed method can be extended to other types of outcomes, such as categorical or censored survival data.

A two-step procedure may be used for selecting important prescriptive variables. For example, in the first step, we fit a flexible regression model of Y given A and X using some tree-based methods, such as BART or GBM. Let $\hat{Q}(X)$ denote the estimated interaction effects of covariates X and treatment indicator A. In the second step, we can consider a classification problem with responses $\text{sign}\{\hat{Q}(X)\}$ and covariates X, and select important covariates based on high-dimensional classification methods such as penalized logistic regression or support vector machine (SVM). Such a two-step procedure can potentially decrease the chance of missing important covariates as compared to a one-step approach such as the proposed method. Although a two-step procedure looks appealing, it also has limitations. First, when p is much larger than n, the estimation of the interaction effects using a tree-based method is usually quite challeng-

ing, especially when the effects are only small to moderate. Second, if the interaction effects are badly estimated in the first step, the resulting classification and selection in the second step can be erroneous.

In addition, inspired by the composite algorithm proposed by Gunter, Zhu and Murphy (2011), a two-step hybrid procedure can be built based on the SAS method. Specifically, in the first step, we use a penalized regression method to select important variables both in the main effects and in the interaction effects based on an assumed model. In the second step, we apply the SAS method based on selected variables from the first step. Such a hybrid procedure generally may have better selection performance than a single-step selection method.

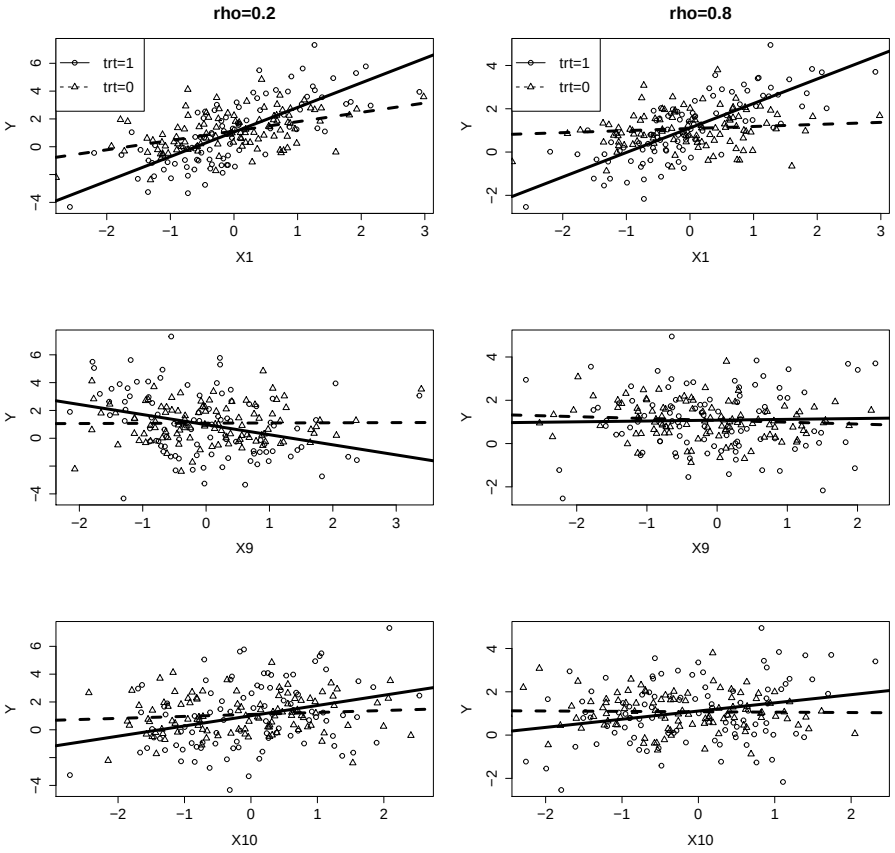

Figure 2.1 Plots of the marginal interaction of covariates X_1, X_9, and X_{10} with treatment (triangles are for treatment 1, and circles are for treatment 0). The fitted lines for treatment 1 (dashed) and treatment 0 (dotted) are from simple linear regression. The left panel is for $\rho = 0.2$; the right panel is for $\rho = 0.8$.

Table 2.1 Simulation results of sequential advantage selection (SAS), S-score and LASSO methods in the single-stage treatment decision study.

	SAS				S-score				LASSO			
ρ	Size	TP	VR	ER	Size $S \neq 0$	TP	VR	ER	Size	TP	VR	ER
Model I												
0.2	6.70	2.98	99.1	5.6	765.28	1.61	84.7	27.5	21.94	1.75	79.5	32.8
	(.06)	(.01)	(.1)	(.1)	(5.16)	(.02)	(.2)	(.3)	(.74)	(.03)	(.4)	(.3)
0.5	7.56	2.91	98.5	7.4	757.18	1.31	86.7	26.6	15.04	1.37	84.2	29.4
	(.08)	(.01)	(.1)	(.2)	(5.24)	(.02)	(.1)	(.2)	(.59)	(.03)	(.3)	(.3)
0.8	8.21	1.76	94.2	17.2	738.44	1.04	94.2	18.8	11.44	1.10	93.0	20.6
	(.07)	(.03)	(.1)	(.2)	(5.48)	(.01)	(.0)	(.1)	(.45)	(.01)	(.1)	(.2)
Model II												
0.2	11.14	2.24	89.8	28.3	765.27	1.73	87.7	31.5	15.84	1.48	86.3	34.3
	(.10)	(.03)	(.2)	(.3)	(5.23)	(.02)	(.2)	(.3)	(.68)	(.03)	(.2)	(.3)
0.5	11.81	1.82	88.9	30.8	758.89	1.30	87.3	33.7	13.36	1.10	87.5	33.5
	(.09)	(.03)	(.2)	(.2)	(5.34)	(.02)	(.1)	(.2)	(.70)	(.03)	(.2)	(.3)
0.8	10.84	1.36	90.4	29.5	749.84	1.09	92.2	26.7	11.65	0.98	92.1	26.0
	(.09)	(.02)	(.1)	(.2)	(5.38)	(.01)	(.1)	(.2)	(.48)	(.02)	(.2)	(.3)
Model III												
0.2	11.73	5.13	84.2	18.3	783.39	3.38	74.7	29.1	27.09	4.03	73.4	30.9
	(.13)	(.12)	(.7)	(.5)	(7.13)	(.05)	(.4)	(.3)	(1.15)	(.10)	(.4)	(.3)
0.5	10.41	4.67	87.0	18.6	776.15	2.88	78.3	28.1	25.05	3.24	77.6	29.3
	(.11)	(.10)	(.5)	(.4)	(7.38)	(.04)	(.3)	(.2)	(1.17)	(.08)	(.3)	(.3)
0.8	7.74	3.01	90.0	19.6	760.68	2.93	90.9	20.3	17.37	2.49	88.5	22.2
	(.10)	(.05)	(.1)	(.2)	(7.62)	(.03)	(.1)	(.2)	(.67)	(.04)	(.2)	(.2)
Model IV												
0.2	11.85	3.29	81.4	29.1	779.14	3.21	81.2	30.5	23.80	3.50	80.5	32.6
	(.11)	(.09)	(.4)	(.4)	(7.56)	(.05)	(.2)	(.3)	(1.12)	(.10)	(.3)	(.4)
0.5	11.68	2.80	82.9	29.6	769.07	2.67	82.3	31.1	18.83	2.51	82.4	32.2
	(.11)	(.07)	(.3)	(.3)	(7.65)	(.04)	(.2)	(.2)	(1.01)	(.07)	(.3)	(.3)
0.8	9.68	2.47	88.6	25.7	758.69	2.82	90.8	23.7	15.72	2.18	89.2	25.6
	(.11)	(.05)	(.2)	(.2)	(7.58)	(.03)	(.2)	(.2)	(.64)	(.04)	(.2)	(.3)

Size: the average number of selected variables; Size $S \neq 0$: the average number of variables with nonzero S-scores; TP: the average number of correctly identified prescriptive variables (the true value is three under models I and II, and is eight under models III and IV); VR: the mean value ratio; ER: the mean error rate. Sample standard deviations are shown in parentheses.

Table 2.2 Optimal Treatment Regimes and Corresponding Important Variables in Multi-Stage Simulation Study

	Optimal Treatment Regime		Important Variables
Stage 2	$1(a_1 + \boldsymbol{\beta}_{12}^T \mathbf{x}_1 + \beta_{21} x_2 \geq 0)$		$(X_2, A_1, X_{1,3}, X_{1,4})$
Stage 1	$1\{Q_1(\mathbf{x}_1, 1) > Q_1(\mathbf{x}_1, 0)\}$	$\mathbf{c} = (0,1,0,0)^T$ and $(0,1,1,1)^T$	$(X_{1,1}, X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6})$
		$\mathbf{c} = (0,0,1,0)^T$	$(X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6})$

Table 2.3 Simulation Results of SAS-MTD method for Two-Stage Treatment Decisions Study.

	Stage 2				Stage 1			
n	Size	TP	VR2	ER	Size	TP	VR1	ER
$\mathbf{c} = (0,1,1,1)^T$								
100	5.22 (.46)	2.08 (.09)	85.3	17.8	6.29 (.36)	0.70 (.07)	73.1	26.9
200	4.08 (.11)	3.25 (.06)	94.3	8.8	5.43 (.25)	2.34 (.10)	93.4	14.5
400	4.02 (.06)	3.75 (.04)	97.5	4.4	3.61 (.11)	3.14 (.04)	98.2	8.5
$\mathbf{c} = (0,1,0,0)^T$								
100	6.70 (.28)	1.80 (.11)	67.7	24.9	8.88 (.35)	0.45 (.07)	49.1	39.8
200	6.38 (.21)	3.48 (.06)	89.7	13.5	11.80 (.27)	1.78 (.08)	78.2	26.6
400	5.82 (.19)	3.88 (.03)	96.1	7.8	13.01 (.36)	2.41 (.07)	92.7	16.6
$\mathbf{c} = (0,0,1,0)^T$								
100	5.93 (.29)	1.94 (.10)	80.2	21.6	4.96 (.41)	0.63 (.08)	71.4	25.0
200	5.13 (.15)	3.15 (.04)	94.8	11.5	5.75 (.34)	1.89 (.08)	91.0	15.6
400	4.70 (.16)	3.53 (.05)	97.8	7.0	4.12 (.23)	2.91 (.09)	97.4	8.5

Size: the average number of selected variables; TP: the number of correctly identified important variables; VR2: the mean value ratio, where random treatment regimes are adopted at stage 1; VR1: the mean value ratio between the estimated and true treatment regimes; ER: the mean error rate. VR1, VR2, and ER are presented in the percentage scale. Sample standard errors are shown in parenthesis.

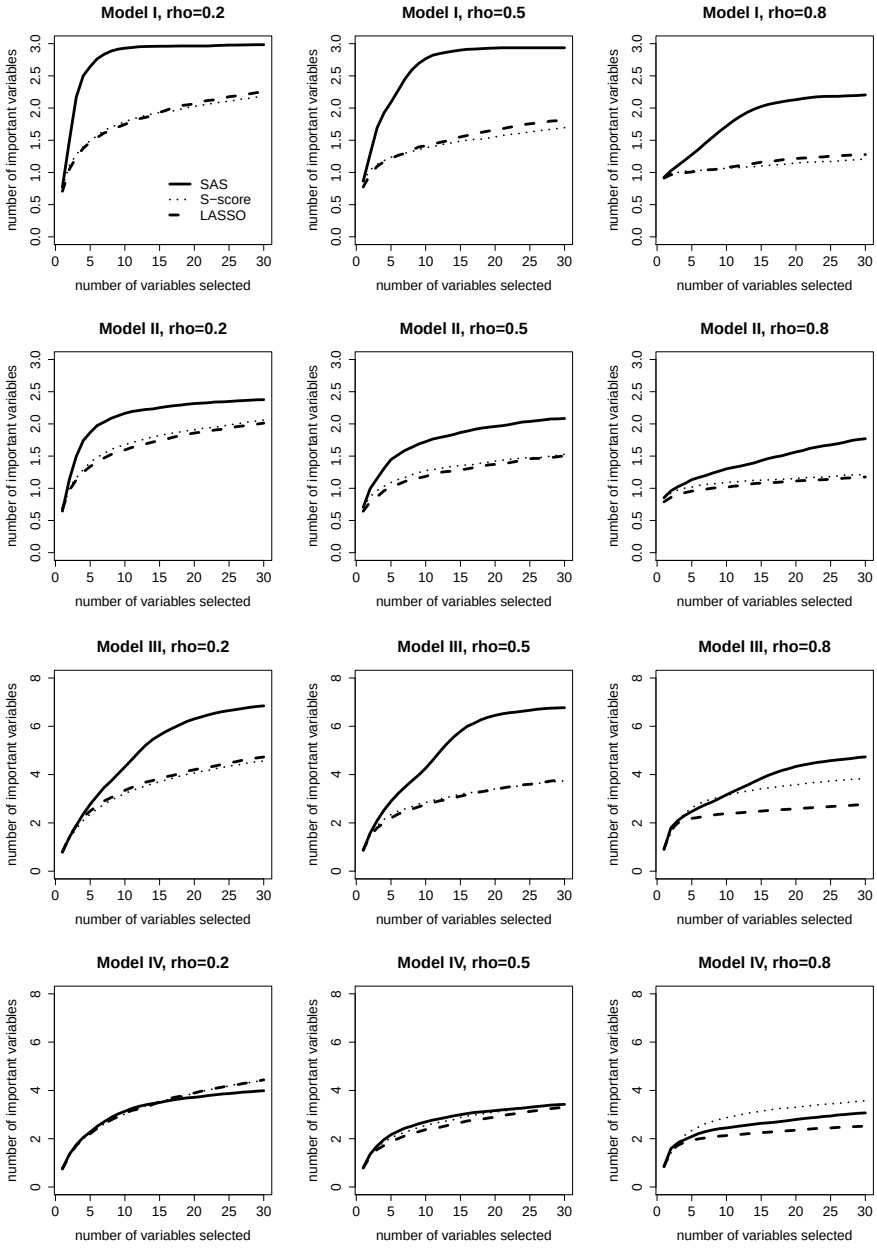

Figure 2.2 Solution paths of sequential advantage selection (SAS, solid line), S-score (dotted line), and LASSO (dashed line) methods for the single-stage treatment decision simulation study.

Table 2.4 Proportion of Each Important Variable Being Selected at Stage 2 and Stage 1.

	Stage 2				Stage 1				
n	$X_{1,3}$	$X_{1,4}$	X_2	A_1	$X_{1,3}$	$X_{1,4}$	$X_{1,5}$	$X_{1,6}$	$X_{1,1}$
	1	-1	1	1	*	*	1	-1	*/-
$\mathbf{c} = (0,1,1,1)^T$									
100	0.58	0.49	0.85	0.16	0.01	0.01	0.13	0.08	0.47
200	0.95	0.96	0.97	0.37	0.02	0.03	0.72	0.71	0.86
400	1.00	1.00	1.00	0.75	0.07	0.07	1.00	1.00	1.00
$\mathbf{c} = (0,1,0,0)^T$									
100	0.59	0.51	0.55	0.15	0.01	0.00	0.24	0.16	0.04
200	0.99	1.00	0.92	0.57	0.03	0.04	0.82	0.84	0.05
400	1.00	1.00	0.98	0.90	0.17	0.07	1.00	1.00	0.17
$\mathbf{c} = (0,0,1,0)^T$									
100	0.57	0.49	0.55	0.33	0.05	0.03	0.26	0.29	-
200	1.00	1.00	0.79	0.36	0.13	0.10	0.82	0.84	-
400	1.00	1.00	0.96	0.57	0.54	0.39	0.99	0.99	-

Important variables at Stage 2 are $X_{1,3}, X_{1,4}, X_2$ and A_1, and the corresponding coefficients in the true optimal treatment regime at Stage 2 are $(1, -1, 1, 1)$. Important variables at Stage 1 are $X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6}, X_{1,1}$ for the first two scenarios and $X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6}$ for the third scenario. The coefficients for $X_{1,5}$ and $X_{1,6}$ are 1 and -1 in the true optimal treatment regime at Stage 1; $X_{1,3}, X_{1,4}$ and/or $X_{1,1}$ appear in the true optimal treatment regime at Stage 1 with a nonlinear form. '*' means this variable is important for the treatment decision and the coefficient is unknown, while '-' means this variable is not important for the treatment decision.

Table 2.5 List of covariates used in the analysis of STAR*D study.

PARTICIPANT FEATURES	
1 Gender	2-6 Ethnicity
7 Economic study consent	8 Depressed mood
9 Diminished interest or pleasure	10 Weight loss while not dieting
11 Insomnia or hypersomnia	12 Psychomotor agitation or retardation
13 Fatigue or loss of energy	14 Feelings of worthlessness or guilt
15 Diminished ability to concentrate	16 Recurrent thoughts of death or suicide
17 Age	18 Number of relatives living with patient
19 Number of friends living with patient	20 Total number of persons in household
21 Years of schooling completed	22 Highest degree received
23 On medical or psychiatric leave	24 Medicare
25 Medicaid	26 Private insurance
27 Better able to make important decisions	28 Better able to enjoy things
29 Impact of your family and friends	30-35 Current marital status
36-41 Current employment status	42-44 Currently a student
45-46 Currently do volunteer work	
ILLNESS FEATURES	
47-60 Cumulative Illness Rating Scale	61-78 Hamilton rating scale for depression
79-82 Medication history	83-221 Psychiatric diagnostic screening questionnaire
222 Baseline Axis I psychiatric condition	
223 Baseline Axis II psychiatric condition	224 Family hx depression
225 Family hx bipolar disorder	226 Family hx alcohol abuse
227 Family hx drug abuse	228 Family hx suicide
CARE FEATURES	
229 Type of clinical site	
INTERMEDIATE MEDICAL CONDITIONS AT LEVEL 1	
230 QIDS-C score change rate	231 AIDS-C percent improvement
232 QIDS-SR score change rate	233 FISER frequency score change rate
234 FISER intensity score change rate	235 GRSEB score change rate
236 CGII score change rate	237 Patient presently a suicide risk
238 Patient in remission	239 Study medical daily dose
240-290 Patient rated inventory of side effects	291-305 Quick Inventory of Depressive Symptomatology
INTERMEDIATE MEDICAL CONDITIONS AT LEVEL 2	
306 QIDS-C score change rate	307 AIDS-C percent improvement
308 QIDS-SR score change rate	309 FISER frequency score change rate
310 FISER intensity score change rate	311 GRSEB score change rate
312 CGII score change rate	313 Patient presently a suicide risk
314 Patient in remission	315 Study medical daily dose
316-366 Patient rated inventory of side effects	367-381 Quick Inventory of Depressive Symptomatology

Table 2.6 Estimated Values of Different Treatment Regimes and Confidence Intervals for the Differences of Values.

Treatment Regime	Estimated Value	Diff	95% CI on Diff
optimal regime from SAS-MTD	-5.26		
BUP + NTP	-13.27	8.01	[2.83,13.64]
BUP + MIRT	-11.71	6.45	[1.83,11.02]
SER + NTP	-13.15	7.89	[1.94,13.72]
SER + MIRT	-12.63	7.37	[2.88,11.86]

Chapter 3

PENALIZED A-LEARNING FOR OPTIMAL TREATMENT DECISIONS WITH NP-DIMENSIONALITY

3.1 Introduction

Personalized medicine is a medical paradigm that focuses on finding the most effective treatment decision based on individual patient information. For many chronic diseases, such as cancer, cardiovascular disease and diabetes, treatment decisions need to be tailored over time according to patients' responses to previous treatments. Such an adaptive treatment strategy is referred as an dynamic treatment regime. Formally speaking, a dynamic treatment regime is a sequence of decision rules, dictating how the treatment will be tailored through time to individual's status.

Various methods have been proposed to estimate the optimal dynamic treatment regime,

including Q-learning (Watkins and Dayan, 1992; Chakraborty et al., 2010) and A-learning (Robins et al., 2000; Murphy, 2003). Both Q-learning and A-learning rely on a backward induction algorithm to find the optimal dynamic treatment regime, however, Q-learning models the conditional mean of the outcome given predictors and treatment while A-learning directly models the contrast function that is sufficient for treatment decision. In particular, A-learning has the so-called double robustness property, i.e. when either the baseline mean function or the propensity score model is correctly specified, the resulting A-learning estimating equation for the contrast function is consistent.

The rapid advances and breakthrough in technology and communication systems make it possible to gather an extraordinary large number of prognostic factors for each individual, such as patient's genetic information, demographic characteristics, medical history and clinical measurements over time. With such data gathered at hand, it is of significant importance to organize and integrate information that is relevant to make optimal individualized treatment decisions, which makes variable selection as an emerging need for implementing personalized medicine with big data. In addition, variable selection is an essential tool in making inference for problems in which the number of covariates is comparable or much larger than the sample size. There have been extensive developments of penalized methods for variable selection in prediction, for example, LASSO (Tibshirani, 1996), SCAD (Fan and Li, 2001) and Dantzig selector (Candès and Tao, 2007), to name a few. In contrast to most penalized regression methods, which adds a penalty term to an objective function, the Dantzig selector focused directly on estimating equations.

Although there is a large amount of work on developing variable selection methods for prediction, variable selection tools for deriving optimal individualized treatment regimes have been less studied, especially when the number of predictors is much larger than the sample size. Qian and Murphy (2011) proposed to estimate the conditional mean response using a L_1-penalized regression and studied the error bound of the value function for the estimated treatment regime.

However, the associated variable selection properties, such as selection consistency, convergence rate and oracle distribution, are not studied. Lu et al. (2013) introduced a new penalized least squares regression framework, which is robust against the misspecification of the conditional mean function. However, they only studied the case when the number of covariates is fixed and the propensity score model is known as in randomized clinical trials.

In this chapter, we propose a penalized A-learning method for deriving the optimal dynamic treatment regime when the number of covariates is of the non-polynomial (NP) order of the sample size. To preserve the double robustness property of the A-learning method, we adopt the Dantzig selector Candès and Tao (2007) which directly penalizes the A-leaning estimating equations.

The remainder of this chapter is organized as follows. We introduce our proposed penalized A-learning method in Section 3.2. Some implementation issues are addressed in Section 3.3, followed by simulation results in Section 3.4. We apply our method to a data from the STAR*D study in Section 3.5.

3.2 Penalized A-Learning

For simplicity of presentation, we only consider a two-stage study where binary treatment decisions are made at time points t_1 and t_2. The data of a subject can be summarized as

$$O = (S^{(1)}, A^{(1)}, S^{(2)}, A^{(2)}, Y), \tag{3.1}$$

where $S^{(1)}$ denotes the covariates collected prior to t_1, $A^{(1)} = 1/0$ is the treatment received at time t_1, $S^{(2)}$ denotes intermediate covariates collected between time points t_1 and t_2, $A^{(2)} = 1/0$ is the treatment received at time t_2, and Y is the final outcome of interest. As usual, it is assumed that a larger value of Y stands for a better clinical outcome. Our goal is to find a dynamic

treatment regime to maximize the mean outcome. Throughout this chapter, we make the stable unit treatment value assumption and sequential randomization assumption (Murphy, 2003) as standard for studying dynamic treatment regimes.

The observed data of n subjects can be summarized as

$$O_i = (S_i^{(1)}, A_i^{(1)}, S_i^{(2)}, A_i^{(2)}, Y_i), i = 1, \ldots, n,$$

which are assumed to be independently and identically distributed copies of O. We assume the following semiparametric regression model for Y:

$$Y_i = h^{(2)}(X_i) + A_i^{(2)}(X_i^T \beta_2) + e_i, \tag{3.2}$$

where $X_i = ((S_i^{(1)})^T, A_i^{(1)}, (S_i^{(2)})^T)^T$ is the vector of covariates for the ith patient, whose first element is 1, $h^{(2)}(\cdot)$ is an unspecified baseline mean function and e_i is an independent error with mean 0. The design matrix is denoted as $X = (X_1, \ldots, X_n)^T$.

Define

$$V_i^{(2)} = \max_{A_i^{(2)}} \mathrm{E}(Y_i | S_i^{(1)}, A_i^{(1)}, S_i^{(2)}, A_i^{(2)}) = h^{(2)}(X_i) + X_i^T \beta_2 I(X_i^T \beta_2 > 0),$$

where $I(\cdot)$ stands for the indicator function. At the first stage, we consider the following model for $V^{(2)}$:

$$V_i^{(2)} = h^{(1)}(S_i^{(1)}) + A_i^{(1)} C(S_i^{(1)}) + \epsilon_i, \tag{3.3}$$

where $E(\epsilon_i | A_i^{(1)}, S_i^{(1)}) = 0$, and $h^{(1)}(\cdot)$ and $C(\cdot)$ are functions of the baseline covariates. To simplify the notation, we use a shorthand S_i for $S_i^{(1)}$ and let $S = (S_1, \ldots, S_n)^T$, the design matrix at the baseline.

It can be shown that the optimal dynamic treatment regime is given by $d^{opt} = (d_1^{opt}, d_2^{opt})$, where d_1^{opt} and d_2^{opt} take the form

$$d_1^{opt}(S_i) = I(C(S_i) > 0) \quad \text{and} \quad d_2^{opt}(X_i) = I(X_i^T \beta_2 > 0). \tag{3.4}$$

To estimate d_1^{opt} and d_2^{opt}, we posit the following models for $C(\cdot)$, $h^{(1)}(\cdot)$, $h^{(2)}(\cdot)$, $\pi^{(1)}(\cdot)$, and $\pi^{(2)}(\cdot)$:

$$\pi^{(1)}(s, \alpha_1) = \exp(s^T \alpha_1)/\{1 + \exp(s^T \alpha_1)\}, \tag{3.5}$$

$$\pi^{(2)}(x, \alpha_2) = \exp(x^T \alpha_2)/\{1 + \exp(x^T \alpha_2)\}, \tag{3.6}$$

$$h^{(1)}(s) = s^T \theta_1, \quad h^{(2)}(x) = x^T \theta_2, \quad C(s) = s^T \beta_1. \tag{3.7}$$

where the first elements of x and s are 1,

$$\pi^{(1)}(s) = \Pr(A_i^{(1)} = 1 | S_i = s) \quad \text{and} \quad \pi^{(2)}(x) = \Pr(A_i^{(2)} = 1 | X_i = x).$$

Models in (3.5)-(3.7) can be misspecified, however, we require that either $h^{(j)}$ or $\pi^{(j)}$ is correct for $j = 1, 2$. As in A-learning, we use backward induction to estimate the optimal dynamic treatment regime. At the second decision point, we first estimate the parameters in the posited propensity score and baseline mean models using penalized regressions. Specifically, define

$$\hat{\alpha}_2 = \arg \min_{\alpha_2 \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^{n} [\log\{1 + \exp(X_i^T \alpha_2)\} - A_i^{(2)} X_i^T \alpha_2] + \sum_{j=1}^{p} \rho_1^{(2)}(|\alpha_2^j|, \lambda_{1n}^{(2)}),$$

and

$$\hat{\theta}_2 = \arg\min_{\theta_2 \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^{n} (1 - A_i^{(2)})(Y_i - X_i^T \theta_2)^2 + \sum_{j=1}^{p} \rho_2^{(2)}(|\theta_2^j|, \lambda_{2n}^{(2)}),$$

where $\alpha_2 = (\alpha_2^1, \cdots, \alpha_2^p)^T$ and $\theta_2 = (\theta_2^1, \cdots, \theta_2^p)^T$, and $\rho_1^{(2)}$ and $\rho_2^{(2)}$ belong to the class of concave penalty functions (Lv and Fan, 2009).

Next, we estimate β_2 in (3.2) using the Dantzig selector based on A-learning estimating function (Murphy, 2003), defined by

$$\hat{\beta}_2 = \arg\min_{\beta_2 \in \Lambda^{(2)}} ||\beta_2||_1, \tag{3.8}$$

where

$$\Lambda^{(2)} = \left\{ \beta_2 \in \mathbb{R}^p : ||\frac{1}{n} X^T \mathrm{diag}(A^{(2)} - \hat{\pi}^{(2)})\{Y - X\hat{\theta}_2 - A^{(2)} \circ (X\beta_2)\}||_\infty \leq \lambda_{3n}^{(2)} \right\},$$
$$A^{(2)} = (A_1^{(2)}, \ldots, A_n^{(2)})^T \quad \text{and} \quad \hat{\pi}^{(2)} = (\pi(X_1, \hat{\alpha}_2), \ldots, \pi(X_n, \hat{\alpha}_2))^T.$$

To estimate the regime at the first decision point, we estimate $V_i^{(2)}$ using the advantage function (Murphy, 2003) by

$$\hat{V}_i = Y_i + X_i^T \hat{\beta}_2 \{I(X_i^T \hat{\beta}_2 > 0) - A_i^{(2)}\}. \tag{3.9}$$

Similarly, define

$$\hat{\alpha}_1 = \arg\min_{\alpha_1 \in \mathbb{R}^q} \frac{1}{n} \sum_{i=1}^{n} [\log\{1 + \exp(S_i^T \alpha_1)\} - A_i^{(1)} S_i^T \alpha_1] + \sum_{j=1}^{q} \rho_1^{(1)}(|\alpha_1^j|, \lambda_{1n}^{(1)}),$$

and

$$\hat{\theta}_1 = \arg\min_{\theta_1 \in \mathbb{R}^q} \frac{1}{n}\sum_{i=1}^{n}(1 - A_i^{(1)})(\hat{V}_i - S_i^T\theta_1)^2 + \sum_{j=1}^{q}\rho_2^{(1)}(|\theta_1^j|, \lambda_{2n}^{(1)}),$$

where $\alpha_1 = (\alpha_1^1, \cdots, \alpha_1^q)^T$ and $\theta_1 = (\theta_1^1, \cdots, \theta_1^q)^T$, and $\rho_1^{(1)}$ and $\rho_2^{(1)}$ are concave penalty functions. Then, we estimate β_1 in (3.3) by

$$\hat{\beta}_1 = \arg\min_{\beta_1 \in \Lambda^{(1)}} ||\beta_1||_1, \tag{3.10}$$

where

$$\Lambda^{(1)} = \left\{\beta_1 \in \mathbb{R}^q : ||\frac{1}{n}S^T\mathrm{diag}(A^{(1)} - \hat{\pi}^{(1)})\{Y - S\hat{\theta}_1 - A^{(1)} \circ (S\beta_1)\}||_\infty \leq \lambda_{3n}^{(1)}\right\},$$
$$A^{(1)} = (A_1^{(1)}, \ldots, A_n^{(1)})^T \quad \text{and} \quad \hat{\pi}^{(1)} = (\pi(S_1, \hat{\alpha}_1), \ldots, \pi(S_n, \hat{\alpha}_1))^T.$$

The estimated optimal dynamic treatment regime is given by

$$\hat{d}_1(S_i) = I(\hat{\beta}_1^T S_i > 0) \quad \text{and} \quad \hat{d}_2(X_i) = I(\hat{\beta}_2^T X_i > 0). \tag{3.11}$$

3.3 Some Implementation Issues

When the tuning parameters in optimization problems (3.8) and (3.10) are fixed, the Dantzig selector can be solved by a standard linear programming algorithm. One issue for implementing Dantzig selector is the choice of the tuning parameters. We use a BIC criterion for selecting tuning parameters. For Dantzig selector (3.8), $\lambda_{3n}^{(2)}$ is chosen as the minimizer of

$$\mathrm{BIC}(\lambda) = n\log(RSS(\lambda)/n) + d(\lambda)\{\log(n) + \log(p+1)\}, \tag{3.12}$$

where $RSS(\lambda) = \sum_{i=1}^{n} \left[\{A_i^{(2)} - \pi^{(2)}(X_i, \hat{\alpha}_2)\}(Y_i^{(2)} - X_i^T\hat{\theta}_2 - A_i^{(2)}X_i^T\hat{\beta}_2) \right]^2$, and $d(\lambda)$ is the number of nonzero components in $\hat{\beta}_2$. A similar BIC criterion was proposed by Chen and Chen (2008). We use a similar criterion for choosing $\lambda_{3n}^{(1)}$.

It was observed that the Dantzig estimators may underestimate the true values of parameters due to the shrinkage estimation(Candès and Tao, 2007). Therefore, we use a two-step procedure for practical implementation, which is referred as Gauss-Dantzig selector in Candès and Tao (2007). Specifically, in the first step, we apply the proposed penalized A-learning to select important variables for making an optimal decision, i.e. those variables with non-zero estimated coefficients. Then, in the second step, their corresponding coefficients are re-calculated by solving the unpenalized A-learning estimating equations with important variables only.

3.4 Simulation Studies

3.4.1 Settings

To evaluate the numerical performance of the proposed penalized A-learning method, we consider simulation studies with two treatment decision points, based on the following model:

$$Y = A_1A_2 + A_2(\beta_{12}^T S^{(1)} + \beta_{21} S^{(2)}) + A_1(\beta_{11}^T S^{(1)}) + \epsilon, \tag{3.13}$$

where $A_j, j = 1, 2$, is the treatment given at the jth stage, $S^{(j)}, j = 1, 2$, denote the covariate information collected before the jth treatment is given, and Y is the final response of interest. Random error ϵ follows a normal distribution with mean 0 and variance σ_1^2. Treatment A_j takes two values: 0 and 1, and two models are considered for the propensity scores: constant model, $P(A_j = 1) = 0.5$, and probit models. Here, covariates $S^{(1)} = (S_1^{(1)}, ..., S_q^{(1)})^T$ follow a multivariate normal distribution with mean 0 and variance I_q. In addition, the intermediate covariate $S^{(2)}$ is a scalar and generated as $S^{(2)} = b_0 + b_1S_1^{(1)} + b_2A_1 + b_3A_1S_1^{(1)} + e$, where e

follows a normal distribution with mean 0 and variance σ_2^2.

Based on model (3.13), the optimal treatment regime at stage 2 is $I(A_1+\beta_{12}^T S^{(1)}+\beta_{21}S^{(2)}>0)$. Following this optimal treatment regime at stage 2, the Q-function at stage 1 is given by

$$\begin{aligned} Q_1(S^{(1)},A_1) &= E\left\{(A_1+\beta_{12}^T S^{(1)}+\beta_{21}S^{(2)})_+ \big| S^{(1)},A_1\right\} + A_1(\beta_{11}^T S^{(1)}) \\ &= \sigma_2\frac{\beta_{21}}{\sqrt{2\pi}}\exp\left(-\frac{\mu_1^2}{2\sigma_2^2}\right) + \beta_{21}\mu_1\{1-\Phi(-\mu_1/\sigma_2)\} + A_1(\beta_{11}^T S^{(1)}), \end{aligned}$$

where $\mu_1 = (A_1+\beta_{12}^T S^{(1)})/\beta_{21} + (b_0+b_1 S_1^{(1)}+b_2 A_1+b_3 A_1 S_1^{(1)})$ and $a_+ = (|a|+a)/2$. Therefore, the contrast function $C(S^{(1)}) = Q_1(S^{(1)},1) - Q_1(S^{(1)},0)$ and thus the optimal treatment regime at stage 1 is $I\{C(S^{(1)})>0\}$.

To evaluate the double robustness of the proposed method, we consider a variety of scenarios with correctly specified and misspecified baseline mean and/or propensity score models. At stage 2, a linear model with covariates $S^{(1)}$, $S^{(2)}$ and A_1 is fitted for the baseline mean function, while the true baseline mean function is $h^{(2)}(X) = A_1(\beta_{11}^T S^{(1)})$. We choose $\beta_{11} = 0_q$, for which the baseline mean function is correctly specified, and $\beta_{11} = (0_4, 1, -1, 0_{q-6})^T$, for which the baseline mean function is misspecified. At stage 1, a linear model with covariates $S^{(1)}$ is fitted for the baseline mean function, which is always misspecified. Logistic models are used for estimating the propensity scores, which are correctly specified for the constant model but misspecified for the probit model. The following four settings are considered:

Setting 1: $\beta_{11} = 0_q$, $P(A_2=1) = 0.5$;

Setting 2: $\beta_{11} = (0_4, 1, -1, 0_{q-6})^T$, $P(A_2=1) = 0.5$;

Setting 3: $\beta_{11} = 0_q$, $P(A_2=1) = \Phi(S^T\gamma)$;

Setting 4: $\beta_{11} = (0_4, 1, -1, 0_{q-6})^T$, $P(A_2=1) = \Phi(S^T\gamma)$,

where $S = ((S^{(1)})^T, S^{(2)})^T$ and $\Phi(\cdot)$ is the cumulative distribution function of standard nor-

mal. For other parameters, we choose $P(A_1 = 1) = 0.5$, $\beta_{12} = (0, 0, 1, -1, 0_{q-4})^T$, $\beta_{21} = 1$, $\sigma_1 = \sigma_2 = 0.5$, $b = (b_0, b_1, b_2, b_3)^T = (0, 1, 1, 1)^T$, and $\gamma = (0_{q-2}, 1, -1, 1)^T$. Table 3.1 summarizes the information of model misspecification for the baseline mean and propensity score models and associated important variables under different settings. In next section, we show simulation results of the four settings with $q = 50/100/200$ and sample size $n = 100/200$ over 100 replications.

3.4.2 Results

Table 3.2 summarizes variable selection results for optimal treatment decisions and the empirical performance of the estimated optimal treatment regime compared with the true optimal regime. Specifically, it reports the false negative (FN) rate (the percentage of important variables that are missed) and false positives (FP) rate (the percentage of unimportant variables that are selected), the ratio of value functions (denoted by VR) calculated using the value function of the estimated optimal treatment regime divided by that of the true optimal regime, and the error rates (ER) of the estimated optimal treatment regimes for treatment decision making, in both stages. Here, the ER at stage 2 is calculated as the mean of $n^{-1}\sum_{i=1}|I(\hat{\beta}_2^T X_i > 0) - I(\beta_{2,0}^T X_i > 0)|$ and at stage 1 as the mean of $n^{-1}\sum_{i=1}|I(\hat{\beta}_1^T S_i > 0) - I(C(S_i) > 0)|$. The value function of a given treatment regime is calculated using Monte Carlo simulations based on 10,000 replications. The VR at stage 2 (devoted by VR*) is to compare the estimated optimal treatment regime at stage 2 and a randomly assigned treatment at stage 1 as in simulated data with the true optimal dynamic treatment regime for both stages. The VR at stage 1 is to compare the estimated optimal dynamic treatment regime with the true optimal dynamic treatment regime for both stages.

We make the following observations. First, the FN rates are higher than the FP rates. This suggests that the Danzig selector tends to have conservative variable selection results, which is commonly seen in the literature. Second, the variable selection results and the error rates

of the estimated optimal treatment regime at stage 2 are generally better than those at stage 1, which is expected since the optimal linear treatment decision rule is correctly specified at stage 2 but not at stage 1. At stage 2, for $n = 200$, about 3.5%-6% important variables are not selected when both the baseline and propensity score models are correctly specified (setting 1), and the number gets to around 6%-20% when one of the two models is misspecified (settings 2 and 3), and goes up to around 21%-37% when none of the models is correctly specified (setting 4). Hence, model misspecification deteriorates variable selection performance for optimal treatment decisions making. Similar trends are found for the ratio of value functions and error rates of the estimated treatment regimes. Third, the estimation and selection performance of the estimated optimal dynamic treatment regimes improves as the sample size increase. In particular, when $n = 200$, the VR's are all above 92% and some are close to 100%, which implies that the estimated optimal treatment regimes nearly maximize the value functions, even under the misspecification of the optimal decision rule at stage 1.

Table 3.3 presents the selection frequencies of important variables in both stages. We make the following observations. First, at stage 2, variables $S_3^{(1)}$, $S_4^{(1)}$ and $S^{(2)}$ are almost always selected in settings 1-3, which shows the double robustness of the proposed method. Second, at stage 1, variables $S_1^{(1)}$, $S_3^{(1)}$ and $S_4^{(1)}$ appear nonlinearly in the Q-function at stage 1, and thus their effects on optimal treatment decision making are unclear due to the complex form of the contrast function. For these three variables, only $S_1^{(1)}$ is often selected, while $S_3^{(1)}$ and $S_4^{(1)}$ are almost always missed. Third, at stage 1, variables $S_5^{(1)}$ and $S_6^{(1)}$ are important variables for settings 2 and 4, which appear in the Q-function as a linear interaction term with A_1. These two variables have a clear influence on optimal treatment decision making at stage 1, and thus are almost always selected when $n = 200$.

3.5 Application to STAR*D Study

We applied the proposed method to a data from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, which was conducted to compare different treatments for patients with major depressive disorder (MDD). There were 4041 participants (age 18-75) with nonpsychotic MDD enrolled in this study. At first level, all participants were treated with citalopram (CIT) up to 14 weeks. Subsequently, 3 more levels of treatments were provided for participants without a satisfactory response to CIT. At each level, participants were randomly assigned to treatment options acceptable to them. At Level 2, participants were eligible for seven treatment options: sertraline (SER), venlafaxine (VEN), bupropion (BUP), cognitive therapy (CT), and augmenting CIT with bupropion (CIT+BUP), buspirone (CIT+BUS) or cognitive therapy (CIT+CT). Participants without a satisfactory response to CT were proceeded to Level 2A for additional medication treatments. All participants who did not respond satisfactorily at Level 2 or 2A were eligible for four treatments at Level 3: medication switch to mirtazapine (MIRT) or nortriptyline (NTP), and medication augmentation with either lithium (Li) or thyroid hormone (THY). Participants without satisfactory response to Level 3 were re-randomized at Level 4 to either tranylcypromine (TCP) or a combination of mirtazapine and venlafaxine (MIRT+VEN). See Fava et al. (2003) and Rush et al. (2004) for more details of the STAR*D study. One goal of the study is to determine which treatment strategies, in what order or sequence, provide the optimal treatment effect.

As an illustration, we focused on a subset of participants who were given treatment BUP or SER at Level 2 and did not receive satisfactory responses, and then were randomized to treatment MIRT or NTP at Level 3. For this study, we considered 381 covariates collected at baseline and intermediate levels as possible relevant predictors. For treatment regime at Level 3, all the 381 covariates as well as the assigned treatment at Level 2 were considered as possible predictors for making optimal treatment decision. For treatment regime at Level 2, 305 covari-

ates that were collected before giving treatment at Level 2 were considered for making optimal treatment decision. Negative 16-item Quick Inventory of Depressive Symptomatology-Clinician-Rated (QIDS-C_{16}) was used as the final response, which is a measurement of symptomatic status of depression. There are 73 participants who had complete records in the subset of data we are interested in. Among these participants, 36 were treated with BUP and 37 were treated with SER at Level 2, and 33 were treated with NTP and 40 were treated with MIRT at Level 3.

The selection an estimation results are summarized as follows. At Level 3, our method selected three covariates: "study med daily dose 1" in clinic visit clinical record form at Level 2 (ADDS1-Level2), "age" in baseline demographics (AGE), and "QIDS-C percent improvement" in clinic visit clinical record form at Level 1 (QCIMP). The estimated optimal treatment regime is $I(-0.006 \times (\text{ADDS1-Level2}) - 0.020 \times \text{AGE} + 0.080 \times \text{QCIMP} \geq 0)$, where 1 represents treatment NTP and 0 represents treatment MIRT. This optimal treatment regime assigns 23 participants to NTP and the rest 50 participants to MIRT. At Level 2, our method also selected three covariates: "Total number of persons in household" in baseline demographics (THOUS), "QIDS-C current score" in clinic visit clinical record form at baseline (QCCUR), and "FISER frequency score" in clinic visit clinical record form at baseline at baseline (FISERFRQ). The estimated optimal treatment regime is $I(0.026 \times \text{THOUS} + 0.073 \times \text{QCCUR} - 0.058 \times \text{FISERFRQ} \geq 0)$, where 1 stands for treatment BUP and 0 stands for treatment SER. This optimal treatment regime assigns 44 participants to BUP and the rest 29 participants to SER.

Table 3.1 Simulation Settings.

	Stage	Baseline	Propensity Score	Important Variables
Setting 1	Stage 2	right	right	$(S^{(2)}, A_1, S_3^{(1)}, S_4^{(1)})$
	Stage 1	wrong	right	$(S_1^{(1)}, S_3^{(1)}, S_4^{(1)})$
Setting 2	Stage 2	wrong	right	$(S^{(2)}, A_1, S_3^{(1)}, S_4^{(1)})$
	Stage 1	wrong	right	$(S_1^{(1)}, S_3^{(1)}, S_4^{(1)}, S_5^{(1)}, S_6^{(1)})$
Setting 3	Stage 2	right	wrong	$(S^{(2)}, A_1, S_3^{(1)}, S_4^{(1)})$
	Stage 1	wrong	right	$(S_1^{(1)}, S_3^{(1)}, S_4^{(1)})$
Setting 4	Stage 2	wrong	wrong	$(S^{(2)}, A_1, S_3^{(1)}, S_4^{(1)})$
	Stage 1	wrong	right	$(S_1^{(1)}, S_3^{(1)}, S_4^{(1)}, S_5^{(1)}, S_6^{(1)})$

Table 3.2 Variable Selection Simulation Results (%).

		Stage 2				Stage 1			
n	q	FN	FP	VR*	ER	FN	FP	VR	ER
Setting 1									
100	50	12.5	2.6	60.8	6.5	61.3	1.4	93.1	22.6
	100	18.8	1.3	60.2	7.5	72.3	0.5	92.4	24.4
	200	22.8	0.4	59.6	8.2	78.3	0.2	91.9	25.1
200	50	3.5	3.6	61.8	4.2	46.7	1.1	94.5	22.4
	100	2.8	1.9	61.7	4.1	48.0	0.7	94.4	22.4
	200	6.0	0.7	61.5	4.5	57.0	0.2	94.1	23.1
Setting 2									
100	50	32.8	1.3	58.2	12.5	58.2	2.0	92.8	21.8
	100	41.3	0.6	57.1	14.6	71.6	0.5	88.9	25.8
	200	46.0	0.2	56.5	15.2	84.0	0.2	86.9	28.4
200	50	13.0	2.2	60.9	7.2	32.2	2.8	99.8	15.0
	100	16.0	0.9	60.6	7.4	34.4	1.3	99.1	14.8
	200	20.0	0.3	60.1	7.8	38.6	0.5	98.2	14.6
Setting 3									
100	50	24.0	7.7	57.5	13.2	64.3	1.4	89.6	21.2
	100	31.8	3.3	56.9	14.1	66.7	0.7	89.4	22.3
	200	40.3	1.4	55.4	15.5	75	0.3	86.9	23.6
200	50	5.8	10.2	61.2	8.0	49.0	1.4	93.8	21.7
	100	7.5	5.1	60.7	9.1	53.7	0.6	93.2	21.6
	200	15.3	2.5	60.1	9.6	61.3	0.2	92.7	21.9
Setting 4									
100	50	47.8	4.5	51.9	20.2	60.0	2.0	85.3	22.4
	100	57.8	2.1	49.9	23.5	72.6	0.5	79.2	27.2
	200	61.3	0.7	49.6	22.9	83.4	0.2	78.9	28.3
200	50	21.0	7.8	58.2	13.4	33.8	2.8	96.8	15.3
	100	30.5	2.9	57.4	15.5	37.0	1.3	96.0	15.0
	200	37.3	1.5	56.0	17.1	41.6	0.5	93.9	16.0

FN: proportion of related variables with zero coefficients
FP: proportion of unrelated variables with nonzero coefficients
VR: value ratio between estimated and true treatment regimes
ER: error rate of estimated treatment regimes

Table 3.3 Proportion of Each Important Variable Being Selected at Stage 2 and Stage 1.

		Stage 2				Stage 1				
n	q	$S_3^{(1)}$	$S_4^{(1)}$	$S^{(2)}$	A_1	$S_1^{(1)}$	$S_3^{(1)}$	$S_4^{(1)}$	$S_5^{(1)}$	$S_6^{(1)}$
		1	-1	1	1	*	*	*	1	-1
Setting 1										
100	50	0.98	1.00	1.00	0.52	0.91	0.13	0.12	-	-
	100	0.97	0.95	1.00	0.33	1.00	0.06	0.07	-	-
	200	0.94	0.92	1.00	0.23	0.57	0.05	0.03	-	-
200	50	1.00	1.00	1.00	0.86	1.00	0.29	0.31	-	-
	100	1.00	1.00	1.00	0.89	1.00	0.29	0.27	-	-
	200	1.00	1.00	1.00	0.76	1.00	0.17	0.12	-	-
Setting 2										
100	50	0.79	0.82	1.00	0.08	0.66	0.09	0.06	0.64	0.64
	100	0.68	0.61	1.00	0.06	0.48	0.04	0.05	0.41	0.44
	200	0.58	0.54	1.00	0.04	0.29	0	0	0.29	0.22
200	50	1.00	1.00	1.00	0.48	0.99	0.23	0.17	1.00	1.00
	100	1.00	1.00	1.00	0.36	0.96	0.19	0.17	0.97	0.97
	200	0.99	0.99	1.00	0.22	0.96	0.06	0.08	0.97	1.00
Setting 3										
100	50	0.84	0.86	0.97	0.37	0.95	0.08	0.04	-	-
	100	0.78	0.78	0.98	0.19	0.80	0.11	0.09	-	-
	200	0.72	0.65	0.95	0.07	0.67	0.03	0.05	-	-
200	50	1.00	1.00	1.00	0.77	1.00	0.24	0.29	-	-
	100	0.98	1.00	1.00	0.72	1.00	0.20	0.19	-	-
	200	0.97	0.95	1.00	0.47	1.00	0.10	0.06	-	-
Setting 4										
100	50	0.57	0.60	0.86	0.06	0.72	0.04	0.05	0.64	0.55
	100	0.35	0.45	0.84	0.05	0.50	0.05	0.06	0.36	0.40
	200	0.34	0.34	0.85	0.02	0.37	0	0	0.23	0.23
200	50	0.94	0.92	0.98	0.32	0.99	0.16	0.19	0.97	1.00
	100	0.83	0.83	0.98	0.14	0.99	0.12	0.08	0.97	0.99
	200	0.75	0.68	0.97	0.11	0.96	0.06	0.06	0.90	0.95

Chapter 4

CHANGE-PLANE ANALYSIS FOR SUBGROUP DETECTION AND SAMPLE SIZE CALCULATION

4.1 Introduction

Classical clinical trials are designed to assess therapeutic benefits of treatments for the whole population that has been considered. However, due to patients' heterogeneity in response to treatments, it is likely that a new treatment is effective or has an enhanced effect compared to a standard treatment only for a specific subpopulation. By making use of patient-specific baseline covariates, subgroup analysis aims to identify subgroups of patients with enhanced treatment effects, which can help to narrow down the target population of a treatment. Hence, it provides an important tool for assessing treatment effects and selecting target populations for future studies.

A number of data-driven approaches have been developed for the subgroup identification.

Song and Pepe (2004) considered the binary response case and proposed using the selection impact curve (SIC) to evaluate treatment policies dictated by a single covariate. Then, based on the SIC, an optimal division of the population for assigning treatments can be obtained. Bonetti and Gelber (2004) grouped patients by values of a single covariate and estimated treatment effects on overlapping subsets of patients using a moving average procedure. Kuk et al. (2010) used recursive subsetting algorithm for identifying subgroups who respond to treatment with high prediction accuracy for clinical outcomes. Foster et al. (2011) developed a "Virtual Twins" method which first predicted the probabilities of response to treatment and control, and then used tree methods to obtain the subgroups with an enhanced treatment effect. Cai et al. (2011) and Zhao et al. (2013) proposed using parametric scoring systems based on multiple baseline covariates to rank treatment effects and then identified patients who benefit more from the new treatment. There are, however, well known risks for undertaking subgroup analysis (Assmann et al., 2000; Wang et al., 2007). For example, subgroup identification may suffer from false positive findings without being performed with a sound statistical hypothesis testing procedure.

Recently, Shen and He (2015) considered a linear logistic-normal mixture model for the response and developed a likelihood-based test for the existence of a subgroup. If a subgroup exists as indicated by the test, the fitted logistic regression model for the subgroup indicator can be used to score patients for treatment selection. The method proposed in Shen and He (2015) provides a valid test for detecting the subgroup with the following two limitations. First, the method relies on some parametric assumptions, such as linear covariate effects and a logistic-normal mixture model for the response, which may be restrictive in applications. Second, since the subgroup is defined by a latent binary variable, the fitted logistic probability for the subgroup indicator is used for treatment selection. This requires selecting a proper threshold parameter, which can be subjective.

In this chapter, we consider change-plane analysis for subgroup detection and sample size

calculation. Our contribution over the literature can be summarized in the following three folds. First, we consider a semiparametric model with an unspecified baseline function and an interaction between a subgroup indicator and the treatment for the mean response, which greatly improves the flexibility of the response models considered in the literature. In addition, the subgroup indicator is explicitly defined by a change-plane as a function of covariates. Second, adopting techniques similar as those in change-point analysis (Liang et al., 1990; Andrews, 1993; Bai, 1997), we propose a doubly-robust score-type statistic for testing the existence of a subgroup with an enhanced treatment effect. The proposed test is doubly-robust in the sense that it is valid when either the baseline function or the propensity score model is correctly specified. If the null hypothesis that a subgroup does not exist is rejected, the change-plane that defines the subgroup can be estimated by maximizing the score-type statistic. Third, we derive the asymptotic distributions of the proposed statistic under both the null and the local alternative hypotheses. A resampling method is developed to approximate the asymptotic null distribution of the test statistic. Based on the derived asymptotic distributions, we also propose a sample size calculation procedure to design a randomized clinical trial for subgroup detection, which has been seldom studied in the literature.

The remainder of this chapter is organized as follows. Section 4.2 introduces the considered semiparametric model and the proposed doubly-robust score test statistic for subgroup detection. The asymptotic distributions of the test statistic under both the null and local alternative hypotheses are also presented. Section 4.3 presents a sample size calculation procedure based on the proposed test. The numerical performance of the proposed test and the associated sample size calculation method are evaluated by simulation studies in Section 4.4. An application of the proposed method to a data from the AIDS Clinical Trials Group protocol 175 is illustrated in Section 4.5. This chapter is concluded with some discussions in Section 4.6. All the technical derivations are given in the Appendix.

4.2 Change-Plane Analysis

4.2.1 The Proposed Model

Let $\boldsymbol{X}$ denote the baseline covariates collected for a subject in an experimental or observational study, A denote the treatment received by the subject, and Y denote his or her response of interest. Here we restrict our attention to a dichotomous treatment coded as 0 and 1, and a continuous response. Let $\boldsymbol{Z} = (\boldsymbol{X}^T, A, Y)^T$. The observed data consist of $\left\{\boldsymbol{Z}_i = (\boldsymbol{X}_i^T, A_i, Y_i)^T, i = 1, ..., n\right\}$, which are n independent and identically distributed (i.i.d.) copies of $\boldsymbol{Z}$. Consider the following semiparametric model

$$Y_i = \mu(\boldsymbol{X}_i) + \tau A_i \mathbf{1}(\boldsymbol{\theta}^T \boldsymbol{X}_i \geq 0) + \epsilon_i, \tag{4.1}$$

where $\mu(\mathbf{X})$ is an unknown baseline mean function for patients in treatment 0, $\mathbf{1}(\cdot)$ is the indicator function, and $E(\epsilon_i | A_i, \boldsymbol{X}_i) = 0$. We assume that the first element of $\boldsymbol{X}$ is 1, $\boldsymbol{X}$ is a $(p+1)$-dimensional vector $(1, X_1, ..., X_p)^T$, and $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_p)^T$ is a $(p+1)$-dimensional vector of parameters. For the identifiability of $\boldsymbol{\theta}$, let $\|\boldsymbol{\theta}\| = 1$, where $\|\cdot\|$ is the ℓ_2-norm. When $\tau = 0$, treatments do not have an effect on the response and thus there are no subgroups with enhanced treatment effects. When $\tau \neq 0$, a subgroup of patients with an enhanced treatment effect exists and is defined by the change-plane $\mathbf{1}(\boldsymbol{\theta}^T \boldsymbol{X} \geq 0)$.

The proposed model is flexible since it puts no assumptions on the baseline mean function. On the other hand, it places constraints on the form of subgroup and the treatment effect for the subgroup, which are directly related to our goal of subgroup detection and identification. Semiparametric models analogous to model (4.1) have been considered in the literature for deriving optimal treatment regimes (Murphy, 2003; Robins, 2004). The difference is the way that the interaction between treatment A and covariates $\boldsymbol{X}$ is modeled. For the subgroup identification problem, we consider the interaction term $A\mathbf{1}(\boldsymbol{\theta}^T \boldsymbol{X} \geq 0)$. To test whether there

exists a subgroup with an enhanced treatment effect, it is equivalent to test the hypothesis

$$H_0 : \tau = 0 \text{ versus } H_a : \tau \neq 0. \tag{4.2}$$

4.2.2 A Doubly-Robust Test

When $\boldsymbol{\theta}$ is known, model (4.1) fits in the class of semiparametric models considered in Robins and Rotnizky (2001). Based on the semiparametric theory (Tsiatis, 2007), a class of doubly-robust estimating equations for τ is given by

$$\sum_{i=1}^{n} \lambda(\boldsymbol{X}_i) \{A_i - \pi(\boldsymbol{X}_i)\} \left\{Y_i - h(\boldsymbol{X}_i) - \tau A_i \mathbf{1}\left(\boldsymbol{\theta}^T \boldsymbol{X}_i \geq 0\right)\right\} = 0, \tag{4.3}$$

where $\lambda(\boldsymbol{X})$ and $h(\boldsymbol{X})$ are arbitrary functions, and $\pi(\boldsymbol{X}) = P(A = 1|\boldsymbol{X})$ is the propensity score. It can be shown that when either the baseline mean function $h(\boldsymbol{X})$ or the propensity score model $\pi(\boldsymbol{X})$ is correctly specified, (4.3) is a consistent estimating equation for τ.

Under the assumption that the random errors ϵ_i's are homoscedastic, the most efficient doubly-robust estimating equation is obtained by setting $\lambda(\boldsymbol{X}) = \mathbf{1}(\boldsymbol{\theta}^T \boldsymbol{X} \geq 0)$ and $h(\boldsymbol{X}) = \mu(\boldsymbol{X})$. As the true baseline function $\mu(\boldsymbol{X})$ and propensity score model $\pi(\boldsymbol{X})$ may not be known in practice, we posit parametric models $h(\boldsymbol{X}, \boldsymbol{\beta})$ and $\pi(\boldsymbol{X}, \boldsymbol{\gamma})$ for $h(\boldsymbol{X})$ and $\pi(\boldsymbol{X})$, respectively. For example, a linear model can be used for $h(\boldsymbol{X}, \boldsymbol{\beta})$ while a logistic model for $\pi(\boldsymbol{X}, \boldsymbol{\gamma})$. Define $\boldsymbol{\eta} = (\boldsymbol{\beta}^T, \boldsymbol{\gamma}^T)^T$. Here, although two parametric forms are considered for fitting the baseline and propensity score model, we only need assume one of them is correct. We consider the following score test statistic for testing $H_0 : \tau = 0$:

$$\sum_{i=1}^{n} \psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta}) \equiv \sum_{i=1}^{n} \mathbf{1}\left(\boldsymbol{\theta}^T \boldsymbol{X}_i \geq 0\right) \{A_i - \pi(\boldsymbol{X}_i, \tilde{\boldsymbol{\gamma}})\} \left\{Y_i - h(\boldsymbol{X}_i, \tilde{\boldsymbol{\beta}})\right\},$$

where $\tilde{\boldsymbol{\eta}} = (\tilde{\boldsymbol{\beta}}^T, \tilde{\boldsymbol{\gamma}}^T)^T$, $\tilde{\boldsymbol{\beta}}$ is an estimator of $\boldsymbol{\beta}$ under the null, and $\tilde{\boldsymbol{\gamma}}$ is an estimator of $\boldsymbol{\gamma}$.

Specifically, $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\gamma}}$ are solutions to the following equations

$$\boldsymbol{\Psi}_{2n}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\psi_2(\boldsymbol{Z}_i,\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}D_{\boldsymbol{\beta}}(\boldsymbol{X}_i)\{Y_i - h(\boldsymbol{X}_i,\boldsymbol{\beta})\} = 0,$$

$$\boldsymbol{\Psi}_{3n}(\boldsymbol{\gamma}) = \frac{1}{n}\sum_{i=1}^{n}\psi_3(\boldsymbol{Z}_i,\boldsymbol{\gamma}) = \frac{1}{n}\sum_{i=1}^{n}D_{\boldsymbol{\gamma}}(\boldsymbol{X}_i)\{A_i - \pi(\boldsymbol{X}_i,\boldsymbol{\gamma})\} = 0,$$

where $D_{\boldsymbol{\beta}}(\boldsymbol{X}_i) = \partial h(\boldsymbol{X}_i,\boldsymbol{\beta})/\partial\boldsymbol{\beta}$ and $D_{\boldsymbol{\gamma}}(\boldsymbol{X}_i) = [\pi(\boldsymbol{X}_i,\boldsymbol{\gamma})\{1-\pi(\boldsymbol{X}_i,\boldsymbol{\gamma})\}]^{-1}\partial\pi(\boldsymbol{X}_i,\boldsymbol{\gamma})/\partial\boldsymbol{\gamma}$.

Note that model (4.1) does not depend on $\boldsymbol{\theta}$ when $\tau = 0$, hence the parameter $\boldsymbol{\theta}$ is identifiable only under the alternative hypothesis. This makes the testing problem given in (4.2) non-regular and standard asymptotic testing framework are not directly applicable. Davies (1977, 1987) consider tests when a nuisance parameter appears under the alternative hypothesis. Andrews (2001) studied such non-regular testing problems for a number of likelihood-based testing procedures under a variety of parametric models. Similar testing problems have also been widely studied for detecting change-points. However, to our knowledge, it remains uninvestigated for detecting the existence of a change-plane based on a semiparametric model. We consider a supremum of squared score test statistics:

$$T_n = \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{\{\sum_{i=1}^{n}\psi_1(\boldsymbol{Z}_i,\tilde{\boldsymbol{\eta}};\boldsymbol{\theta})\}^2}{n\tilde{V}_S(\boldsymbol{\theta})}, \tag{4.4}$$

where $\boldsymbol{\Theta} = \{\boldsymbol{\theta}\in\mathbb{R}^{p+1} : \|\boldsymbol{\theta}\| = 1\}$ and $\tilde{V}_S(\boldsymbol{\theta})$ is a consistent estimator for the asymptotic variance of $n^{-1/2}\sum_{i=1}^{n}\psi_1(\boldsymbol{Z}_i,\tilde{\boldsymbol{\eta}};\boldsymbol{\theta})$ under the null hypothesis. The definition of $\tilde{V}_S(\boldsymbol{\theta})$ is given in the next section.

To compute the test statistic T_n, we need to find the supremum of squared score test statistics over a unit ball in $\mathbb{R}^{p+1}$. Since it is infeasible to get the supremum explicitly, we use a numerical method to find the maximum over the space $\boldsymbol{\Theta}$. To incorporate the unit ball constraint, it is natural to consider a sphere coordinates transformation $\boldsymbol{\phi} = (\phi_1,...,\phi_p)^T \mapsto \boldsymbol{\theta}$,

where ϕ_p ranges over $[0, 2\pi)$ and other elements of $\boldsymbol{\phi}$ range over $[0, \pi]$. The transformation is given as follows

$$\begin{cases} \theta_0 = \cos(\phi_1), \\ \quad \cdots \\ \theta_{p-1} = \sin(\phi_1)\sin(\phi_2)\cdots\cos(\phi_p), \\ \theta_p = \sin(\phi_1)\sin(\phi_2)\cdots\sin(\phi_p). \end{cases}$$

We consider a set of grid points of $\boldsymbol{\phi}$ over $[0, \pi]^{p-1} \times [0, 2\pi)$ and compute the maximum of squared score statistics over the set of grid points to approximate T_n.

In the next section, we establish the asymptotic distributions of T_n under both the null and the local alternative hypotheses. In addition, we propose a resampling method to compute the critical values of the limiting null distribution. When the null hypotheses is rejected, the change-plane parameter $\boldsymbol{\theta}$ can be estimated by

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\sup} \frac{\{\sum_{i=1}^n \psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})\}^2}{n\tilde{V}_S(\boldsymbol{\theta})}. \tag{4.5}$$

Thus the estimated subgroup with an enhanced treatment effect is $\mathbf{1}(\hat{\boldsymbol{\theta}}^T\boldsymbol{X} \geq 0)$.

4.2.3 Asymptotic Distributions of T_n

Define $\boldsymbol{\Psi}_2(\boldsymbol{\beta}) = E\{\boldsymbol{\Psi}_{2n}(\boldsymbol{\beta})\}$ and $\boldsymbol{\Psi}_3(\boldsymbol{\gamma}) = E\{\boldsymbol{\Psi}_{3n}(\boldsymbol{\gamma})\}$. To establish the asymptotic distributions of T_n, we make the following assumptions:

A1. Equations $\boldsymbol{\Psi}_2(\boldsymbol{\beta}) = 0$ and $\boldsymbol{\Psi}_3(\boldsymbol{\gamma}) = 0$ have unique solutions $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$, respectively, and the solutions $\boldsymbol{\eta}_0 = (\boldsymbol{\beta}_0^T, \boldsymbol{\gamma}_0^T)^T$ are in a compact set of the parameter space.

A2: We have

$$\sqrt{n}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) = -C_1^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_2(\boldsymbol{Z}_i, \boldsymbol{\beta}_0) + o_p(1),$$

$$\sqrt{n}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) = -C_2^{-1} \sum_{i=1}^{n} \psi_3(\boldsymbol{Z}_i, \boldsymbol{\gamma}_0) + o_p(1),$$

where $C_1 = E\{\psi_2'(\boldsymbol{Z}, \boldsymbol{\beta}_0)\}$, $C_2 = E\{\psi_3'(\boldsymbol{Z}, \boldsymbol{\gamma}_0)\}$, and both of them are finite and positive definite deterministic matrices.

A3: The function $\psi_1(\boldsymbol{Z}, \boldsymbol{\eta}; \boldsymbol{\theta})$ is twice continuously differentiable with respect to $\boldsymbol{\eta}$, and has bounded first and second derivatives.

A4: The function $E[\{Y - h(\boldsymbol{X}, \boldsymbol{\beta})\}^2]$ is uniformly bounded in $\boldsymbol{\beta}$.

A5: We have $0 < P(\boldsymbol{\theta}^T \boldsymbol{X} \geq 0) < 1$ for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$.

Assumptions A1 and A2 ensure the consistency and asymptotic normality of $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\gamma}}$. These assumptions are satisfied for many commonly used parametric models under mild conditions, such as a linear model for $h(\boldsymbol{X}, \boldsymbol{\beta})$ and a logistic model for $\pi(\boldsymbol{X}, \boldsymbol{\gamma})$. The asymptotic distributions of $\tilde{\boldsymbol{\beta}}$ under the null and local alternative hypotheses are similar to those established in Le Cam's third lemma (Van der Vaart (2000, p. 90)). Assumptions A3-A5 are assumed to establish the weak convergence of the process $n^{-1/2} \sum_{i=1}^{n} \psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})$ indexed by $\boldsymbol{\theta}$.

Theorem 1. *Suppose that either the baseline mean function $h(\boldsymbol{X}, \boldsymbol{\beta})$ or the propensity model $\pi(\boldsymbol{X}, \boldsymbol{\gamma})$ is correctly specified, but not necessarily both. If Assumptions A1-A5 hold, T_n converges in distribution to $\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} G^2(\boldsymbol{\theta})$ under H_0 as n goes to infinity, where $\{G(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ is a mean zero Gaussian process with the covariance function*

$$\Sigma(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = E\{\psi_{1*}(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}_1)\psi_{1*}(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}_2)\} \Big/ \sqrt{E\psi_{1*}^2(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}_1) E\psi_{1*}^2(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}_2)}$$

for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, where

$$\psi_{1*}(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}) = \psi_1(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}) - K_1^T C_1^{-1} \psi_2(\boldsymbol{Z}, \boldsymbol{\beta}_0) - K_2^T C_2^{-1} \psi_3(\boldsymbol{Z}, \boldsymbol{\gamma}_0),$$

$K_1 = E\{\partial \psi_1(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}) / \partial \boldsymbol{\beta}\}$, and $K_2 = E\{\partial \psi_1(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}) / \partial \boldsymbol{\gamma}\}$.

Next, we establish the asymptotic distribution of T_n under a sequence of local alternatives $H_{1n} : \tau = n^{-1/2}\delta$.

Theorem 2. *Suppose that either the baseline mean function $h(\boldsymbol{X}, \boldsymbol{\beta})$ or the propensity model $\pi(\boldsymbol{X}, \boldsymbol{\gamma})$ is correctly specified, but not necessarily both. If Assumptions A1-A5 hold, T_n converges in distribution to $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} G_\delta^2(\boldsymbol{\theta})$ under H_{1n} as n goes to infinity, where $\{G_\delta(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ is a Gaussian process with the mean function*

$$\mu(\boldsymbol{\theta}) = \delta E\left[\{\mathbf{1}(\boldsymbol{\theta}_0^T\boldsymbol{X} \geq 0, \boldsymbol{\theta}^T\boldsymbol{X} \geq 0)\pi_0(\boldsymbol{X})\{1 - \pi(\boldsymbol{X}, \boldsymbol{\gamma}_0)\}\right] \Big/ \sqrt{E\{\psi_{1*}^2(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta})\}}$$

and the covariance function $\Sigma(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, where $\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$ and $\pi_0(\boldsymbol{X})$ is the true propensity score model.

To calculate the critical values for the test, we use a resampling method to approximate the limiting null distribution of the test statistic. Define

$$\hat{\psi}_{1*}(\boldsymbol{Z}, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta}) = \psi_1(\boldsymbol{Z}, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta}) - \hat{K}_1^T\hat{C}_1^{-1}\psi_2(\boldsymbol{Z}, \tilde{\boldsymbol{\beta}}) - \hat{K}_2^T\hat{C}_2^{-1}\psi_3(\boldsymbol{Z}, \tilde{\boldsymbol{\gamma}}),$$

where $\hat{K}_1$, $\hat{K}_2$, $\hat{C}_1$ and $\hat{C}_2$ are the empirical estimates of their population counterparts. Specifically, $\hat{K}_1 = \frac{1}{n}\sum_{i=1}^n \partial\psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})/\partial\boldsymbol{\beta}$, $\hat{K}_2 = \frac{1}{n}\sum_{i=1}^n \partial\psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})/\partial\boldsymbol{\gamma}$, and $\hat{C}_1 = \frac{1}{n}\sum_{i=1}^n \partial\psi_2(\boldsymbol{Z}_i, \tilde{\boldsymbol{\beta}})/\partial\boldsymbol{\beta}$, $\hat{C}_2 = \frac{1}{n}\sum_{i=1}^n \partial\psi_3(\boldsymbol{Z}_i, \tilde{\boldsymbol{\gamma}})/\partial\boldsymbol{\gamma}$. Then, $\tilde{V}_S(\boldsymbol{\theta}) = n^{-1}\sum_{i=1}^n\{\hat{\psi}_{1*}(\boldsymbol{Z}, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})\}^2$. We consider the following perturbed test statistic

$$T_n^* = \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{\left\{\sum_{i=1}^n \xi_i\hat{\psi}_{1*}(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})\right\}^2}{n\tilde{V}_S(\boldsymbol{\theta})},$$

where $\xi_1, ..., \xi_n$ are i.i.d. standard normal random variables independent of data. By generating a large number of perturbed test statistics, we can use the empirical distribution of T_n^* to compute the critical value C_α, the upper α quantile of the empirical distribution. Then an α-level test

reject the null hypothesis when $T_n > C_\alpha$.

4.3 Sample Size Calculation

Since most clinical trials are designed to detect the overall treatment effect, they may lack power to detect a subgroup with an enhanced treatment effect (Yusuf et al., 1991; Rothwell, 2005). For example, Brookes et al. (2004) has shown that a trial with 80% power for the overall effect had only 29% power to detect an interaction effect of the same magnitude. To appropriately conduct a subgroup analysis with targeted power, a careful design and predefined statistical analysis protocol are important (Assmann et al., 2000; Cui et al., 2002). In this section, we provide a sample size calculation method based on the proposed test for subgroup detection in a randomized clinical trial.

To derive the sample size formula, we first calculate the asymptotic power of the test under the local alternatives $H_{1n} : \tau = n^{-1/2}\delta$, where n is the sample size. The sample size formula can then be derived at a pre-specified power $1-\beta$. Here we are interested in sample size calculation for a randomized trial, therefore the propensity score is given and there is no need to estimate $\boldsymbol{\gamma}$. In addition, we assume that the errors ϵ_i's in model (4.1) are i.i.d. with mean 0 and variance σ^2. Under this case, the asymptotic covariance function of the test statistic T_n can be simplified as

$$\Sigma(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \frac{E\{\mathbf{1}(\boldsymbol{\theta}_1^T \boldsymbol{X} \geq 0, \boldsymbol{\theta}_2^T \boldsymbol{X} \geq 0) g(\boldsymbol{X})\}}{\sqrt{E\{\mathbf{1}(\boldsymbol{\theta}_1^T \boldsymbol{X} \geq 0) g(\boldsymbol{X})\} E\{\mathbf{1}(\boldsymbol{\theta}_2^T \boldsymbol{X} \geq 0) g(\boldsymbol{X})\}}},$$

where $g(\boldsymbol{X}) = \pi(1-\pi)[\{\mu(\boldsymbol{X}) - h(\boldsymbol{X}, \boldsymbol{\beta}_0)\}^2 + \sigma^2]$ and $\pi = P(A = 1)$. In addition, under the local alternatives, the asymptotic mean function of T_n is given by

$$\mu(\boldsymbol{\theta}) = \delta\pi(1-\pi)\frac{E\{\mathbf{1}(\boldsymbol{\theta}_0^T \boldsymbol{X} \geq 0, \boldsymbol{\theta}^T \boldsymbol{X} \geq 0)\}}{\sqrt{E\{\mathbf{1}(\boldsymbol{\theta}^T \boldsymbol{X} \geq 0) g(\boldsymbol{X})\}}}.$$

For an α-level test to have $1-\beta$ power in detecting an enhanced treatment effect of size τ_0,

we need to find δ_0 such that $P(\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} G^2_{\delta_0}(\boldsymbol{\theta}) > q_\alpha) = 1-\beta$, where q_α is the upper α-quantile of the distribution of $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} G^2(\boldsymbol{\theta})$, and $G(\boldsymbol{\theta})$ is the Gaussian process defined in Theorem 1. Based on the relationship $\tau_0 = n^{-1/2}\delta_0$, the required sample size is given by $n = (\delta_0/\tau_0)^2$. To find δ_0, we take the following three steps. In Step 1, we compute the mean function $\mu(\boldsymbol{\theta})$ and the covariance function $\Sigma(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ via numerical integration, for which we need to specify the true value $\boldsymbol{\theta}_0$ in the change-plane, the distribution of covariates $\boldsymbol{X}$, the difference between the true baseline mean function and the posited mean function, $\mu(\boldsymbol{X}) - h(\boldsymbol{X}, \boldsymbol{\beta}_0)$, and σ^2, the variance of ϵ. These quantities can be estimated from historical data or a pilot study. In Step 2, for any given δ, we compute the probability $P(\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} G^2_\delta(\boldsymbol{\theta}) > q_\alpha)$ via Monte Carlo simulations detailed as follows. We first approximate $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} G^2_\delta(\boldsymbol{\theta})$ by $\max_{k=1,...,K} G^2_\delta(\boldsymbol{\theta}_k)$, where $\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_K$ is a set of fine grids of $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. This can be done by the same sphere coordinates transformation used previously. Next, we generate $(W_1, ..., W_K)^T$ from a multivariate normal distribution with the mean $(\mu(\boldsymbol{\theta}_1), ..., \mu(\boldsymbol{\theta}_K))^T$ and the variance-covariance matrix $\{\Sigma(\boldsymbol{\theta}_{k_1}, \boldsymbol{\theta}_{k_2}) : \ k_1, k_2 = 1, ..., K\}$. Finally, we compute the probability $P(\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} G^2_\delta(\boldsymbol{\theta}) > q_\alpha)$ based on the empirical distribution of $\max_{k=1,...,K} W_k^2$. Note that q_α can be calculated similarly by generating $(W_1, ..., W_K)^T$ from a multivariate normal distribution with mean 0 and the same variance-covariance matrix. In practice, we generate a large set, say $10,000$ of $\max_{k=1,...,K} W_k^2$ to compute the probability. In Step 3, we find δ_0 via a grid search.

4.4 Simulation Studies

4.4.1 Test and Estimation

We conducted extensive simulation studies to investigate the empirical performance of the proposed test for subgroup detection and the estimation for the change-plane parameter $\boldsymbol{\theta}$ under the alternative hypotheses. In particular, we considered various settings to examine the robustness of the test against the misspecification of the baseline mean function in both randomized

and observational studies.

Simulated data with sample sizes $n = 500$ and 1000 were generated based on model (4.1), where two covariates $\boldsymbol{X} = (X_1, X_2)^T$ were considered. Here, X_1 follows a Bernoulli distribution with the success probability 0.5 and X_2 follows a uniform distribution on $(-1, 1)$. The random noise ϵ is normally distributed with mean zero and variance 0.25. For the treatment assignment indictor A, we considered the following two settings for the propensity score model $\pi(\boldsymbol{X})$ (In short as P-Model hereinafter):

- P-Model I: $\pi(\boldsymbol{X}) = 0.5$;

- P-Model II: $\pi(\boldsymbol{X}) = \frac{\exp(\gamma_0+\gamma_1 X_1+\gamma_2 X_2)}{1+\exp(\gamma_0+\gamma_1 X_1+\gamma_2 X_2)}$, $\gamma_0 = 0, \gamma_1 = \gamma_2 = 0.5$.

The two settings represent a randomized clinical trial and an observational study, respectively. We also considered three baseline mean functions for $\mu(\boldsymbol{X})$ (In short as B-Model hereinafter):

- B-Model I: $\mu(\boldsymbol{X}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$, $\beta_0 = \beta_1 = \beta_2 = 1$;

- B-Model II: $\mu(\boldsymbol{X}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_2^2$, $\beta_0 = 1, \beta_1 = 0.5, \beta_2 = 0, \beta_3 = 1$;

- B-Model III: $\mu(\boldsymbol{X}) = \beta_0 + \beta_1 \sin(\beta_2 X_1 + \beta_3 \pi X_2)$, $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 1$.

The proposed test was implemented on each simulated dataset. When calculating the test statistic in (4.4), we fit a linear model $h(\boldsymbol{X}, \boldsymbol{\beta})$ for the baseline mean function and a logistic model $\pi(\boldsymbol{X}, \boldsymbol{\gamma})$ for the propensity score. Therefore, the baseline mean function is correctly specified for the setting with B-Model I and is misspecified for the settings with B-Models II and III, while the propensity score model is correctly specified for both P-Model I and II. When calculating the test statistic, we used the spherical coordinates transformation and searched the supremum over $K = 100 \times 100$ grid points, with 100 grid values for each angular coordinate. For each test, we used 1000 resamplings to obtain the critical values of the test. We reported the empirical type I errors and powers of the test. Simulation results are summarized below for both the null ($\tau = 0$) and the alternative ($\tau \neq 0$), respectively.

4.4.1.1 Type I Errors

For each setting, we simulated 5000 data sets to compute type I errors of the test with the significance level of 0.05 and 0.1. The results in Table 4.1 show that the empirical type I errors are all close to their nominal values, which demonstrate the validity and robustness of the proposed test for subgroup detection.

4.4.1.2 Powers and Estimates of Change-Plane Parameters

Under alternative hypotheses, the enhanced treatment effect in the subgroup is set to be $\tau = \pm 0.1$, ± 0.25 and ± 0.5. In addition, the true change-plane parameter is chosen as $\boldsymbol{\theta}_0 = (-0.15, 0.3, 0.942)^T$ for all settings. With this choice of $\boldsymbol{\theta}_0$, the subgroup with an enhanced treatment effect contains approximately 50% of the population, and includes subjects with $X_1 = 1$ and $X_2 \geq -0.159$ or $X_1 = 0$ and $X_2 \geq 0.159$.

Empirical powers based on 1000 simulated datasets for all settings are shown in Table 4.2. As expected, the powers for detecting the subgroup increase as the sample size n or the magnitude of treatment effect τ increases. When the magnitude of treatment effect τ increases to 0.5, the powers are almost 100% for all settings. The powers for B-Model II are comparable to those for B-Model I, while the powers for B-Model III are relatively smaller than those of B-Models I and II. One explanation may be that since B-Model III has a very nonlinear baseline mean function, a posited linear model may not be a good fit and thus lost some efficiency.

Next, we estimated the change-plane parameter $\boldsymbol{\theta}$ by (4.5). We report the bias and the empirical standard deviation of the estimates $\hat{\boldsymbol{\theta}}$ in Figure 4.1. We also report the misclassification rate for identifying the true subgroup in Table 4.3. The misclassification rate is the proportion of subjects who are misidentified either as members in the subgroup or as members not in the subgroup, and is calculated by $\frac{1}{n}\sum_{i=1}^{n} |\mathbf{1}(\hat{\boldsymbol{\theta}}^T \boldsymbol{X}_i \geq 0) - \mathbf{1}({\boldsymbol{\theta}_0}^T \boldsymbol{X}_i \geq 0)|$.

Based on the results, it is observed that the biases and standard deviations of the estimates

decrease as the sample size or the magnitude of treatment effect increase. In particular, when the magnitude of treatment effect increases to 0.5, all the estimate are nearly unbiased. For small treatment effects, the estimators for $\boldsymbol{\theta}$ are underestimated. This may be because that the true $\theta_2 = 0.942$, which is likely to be underestimated due to the upper limit of 1 for θ. Similar to the power results, the estimates for B-Model III have larger biases and standard deviations compared to those for B-Model I and II. In addition, Table 4.3 shows that the misclassification rates also decrease as the sample size n or the magnitude of treatment effect τ increases. When the magnitude of the treatment effect increases to 0.5, most misclassification rates are less than 5% except for those under B-Model III.

For practical use, we also report the computation time for conducting the test based on our method with various sample sizes n and number of grid points K on unit ball $\boldsymbol{\Theta}$. Simulation setting with B-Model I, P-Model I, $\tau = 0.5$ and $M = 1000$ is used for calculating the average time cost and its standard deviation, as shown in Table 4.4.

Furthermore, we compare the empirical powers with the powers based on method in Shen and He (2015), and the misclassification rates of identified subgroups with those based on method in Zhao et al. (2013).

We apply the EM test in Shen and He (2015) to simulated data under our settings with P-Model I and B-Model I/II/III at $\tau = 0$, 0.1 and 0.5. Following their method, the covariate associated with the subgroup mean is $\boldsymbol{Z} = (X_1, X_2, A)^T$ and the covariate associated with the subgroup membership is $\boldsymbol{X}$. Other parameters for implementing their method are chosen as following: number of EM iteration step $K = 3$, number of boostrap samples $B = 1000$, and set of $\boldsymbol{\gamma}$ values is $\boldsymbol{\Gamma} = \{(0.1, 1, 1)^T, (0.1, 1, -1)^T, (0.1, -1, 1)^T, (0.1, -1, -1)^T\}$. The results are reported in Table 4.5. When baseline model is linear (B-Model I), their method pertains correct type I errors, and the empirical powers are slightly smaller compared to those in Table 4.2. When baseline model is not linear (B-Model II/III), their method do not have the correct type I errors (100%), and hence is not valid. This shows that their method is very sensitive to

the baseline model.

Based on Zhao et al. (2013), the subgroup of interest is defined by $\{\hat{D}(\boldsymbol{X}) \geq c\}$, where $\hat{D}(\boldsymbol{X})$ is the estimated treatment difference, and c is some fixed constant. We used a single model for both treatment groups $E(Y|\boldsymbol{X}, A) = \boldsymbol{\beta}_1\boldsymbol{X} + \boldsymbol{\beta}_2\boldsymbol{X} \times A$, and thus the estimated treatment difference is $\hat{D}(\boldsymbol{X}) = \hat{\boldsymbol{\beta}}_2\boldsymbol{X}$. Here, we are interested in subgroup defined by $\{\hat{D}(\boldsymbol{X}) \geq c\}$, and the average treatment difference for this subgroup $\widehat{\text{AD}}(c)$ can be estimated as in Zhao et al. (2013). Cut-off value for c was chosen such that $\widehat{\text{AD}}(c) = \tau$. Misclassification rates of these subgroups were calculated under simulation settings with P-Model I and B-Model I/II/III at $\tau = 0.1$ and 0.5, and corresponding results are reported in Table 4.6. Subgroups based on the method in Zhao et al. (2013) have higher misclassification rates.

In summary, the proposed test can maintain correct type I errors under the null and reasonable powers to detect subgroup under alternative hypotheses. It also provides consistent estimates of the change-plane parameter for subgroup identification.

4.4.2 Sample Size Calculation Examples

In this section, we conducted a simulation study to evaluate the proposed sample size calculation procedure for a randomized trial with the equal treatment assignment probability, i.e. $\pi = 0.5$. In this simulation study, we considered a single covariate X from a uniform distribution on (-1,1) in all settings. The subgroup of interest is defined by $X \geq \theta_0$, where θ_0 was chosen as: -0.5, 0 and 0.5, corresponding to the scenarios that 75%, 50% and 25% of the population are in the subgroup with an enhanced treatment effect. The variance of the random noise ϵ is set as $\sigma^2 = 0.25$. We considered three levels of treatment effects: $\tau = 0.1$, 0.25 and 0.5, which represent small, medium and large effects, respectively. In addition, we considered three baseline mean functions: $\mu(X) = 1 + X$, $1 - X^2$ and $1 + \sin(\tilde{\pi}X)$, where $\tilde{\pi}$ is the circumference to diameter ratio. In our test statistics, we fit a linear model $h(X, \boldsymbol{\beta})$ for $\mu(X)$. After some calculation, it can be shown that when $\mu(X) = 1 - X^2$, $h(X, \boldsymbol{\beta}_0) = 2/3$; while when $\mu(X) = 1 + \sin(\tilde{\pi}X)$,

$h(X, \boldsymbol{\beta}_0) = 1 + (3/\tilde{\pi})X$. Therefore, the difference $\mu(X) - h(X, \boldsymbol{\beta}_0)$ can be calculated accordingly. We calculate the required sample size n for the test with the level $\alpha = 0.05$ and power $1 - \beta = 90\%$.

Table 4.7 summarizes the results. For each setting, based on the calculated sample size n, we generated 1000 data sets and computed the empirical powers of the proposed test statistic, which are also given in Table 4.7. The empirical powers under all settings are close to the nominal level 90%.

In addition, we compared the sample sizes with those based on the method in Brookes et al. (2004), where they inflate the original sample size for overall treatment effect such that the interaction test between treatment and subgroup can achieve the nominal power. For interactions of the same magnitude as the overall effect, the inflation factor of sample size is approximately fourfold. That is, the required sample size for examining whether treatment effects differ between subgroups is $n = 4n_{overall}$, where $n_{overall}$ is the required sample size for overall treatment effect. It's needed to clarify that their method is limited to the situation of two equally sized subgroups. Under our simulation settings, $\theta_0 = 0$ is corresponding to the case of two equally sized subgroups. Based on their formula, we need to calculate the sample size $n_{overall}$ for testing the overall treatment effect with model $Y = \beta_0 + \beta_1 X + \tau A + \epsilon$. We use the sample size formula for analysis of covariance following Borm et al. (2007): $n_{overall} = (1 - \rho^2)n_t$, where ρ is the correlation between X and Y, and n_A is the required sample size in a two-sample t-test: $n_t = \frac{4(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\tau^2}$.

Table 4.8 shows the calculated sample size using the above formula and the corresponding empirical power for our test. Based on these results, the sample sizes are all smaller compared to those in Table 4.7, and have insufficient powers. The reason may be that our test is used for detecting the existence of a subgroup with enhanced treatment effect, while Brookes et al. (2004) only considers the treatment and subgroup interaction effect for a specific subgroup.

4.5 Application to AIDS Data

We illustrated the proposed method with a data from the AIDS Clinical Trials Group (ACTG) protocol 175 (Hammer et al., 1996), a study that randomized subjects to four different daily regimens: zidovudine (ZDV) monotherapy, ZDV + didanosine (ddI), ZDV + zalcitabine (zal) and ddI monotherapy. We focused on comparing two treatments: ZDV+ddI (treatment 1) and ZDV+zal (treatment 0). There are 522 subjects in treatment 1 and 524 subjects in treatment 0. Following Lu et al. (2013), we considered the CD4 counts (cells/mm^3) at 20 ± 5 weeks after randomization as the response and used two covariates for subgroup identification: age (years) and homosexual activity (0=no, 1=yes), denoted as homo.

We applied the proposed method to detect whether there is a subgroup with an enhanced treatment effect. A linear model was used for the baseline function. The value of the test statistic is 21.25, which is calculated based on 200×50 grid points of the sphere coordinates. The p-value based on 1000 resamplings is less than 0.001, showing a strong evidence for the existence of a subgroup with an enhanced treatment effect. The estimated change-plane parameter $\hat{\boldsymbol{\theta}} = (-0.576, 0.037, -0.816)^T$. The identified subgroup includes subjects with $age > 37.64$ if $homo = 1$ or with $age > 15.58$ if $homo = 0$. There are 622 subjects included in this subgroup, among which 315 subjects received treatment 1 and 307 received treatment 0. Given the estimated change-plane and the fitted linear model for the baseline mean function, we can estimate the enhanced treatment effect τ, which is $\hat{\tau} = 41.6$. Therefore, for patients in the identified subgroup, treatment 1 is better than treatment 0. This agrees with the findings in Lu et al. (2013) that treatment 1 is better than treatment 0 for older patients.

Next, based on the AIDS data, we calculated the required sample size for subgroup detection in future balanced randomized trials. As an illustration, we considered the test with the size $\alpha = 0.05$ and power $1 - \beta = 0.9$. From the AIDS data, we estimated the standard deviation of ϵ as $\hat{\sigma} = 145.9$. Therefore, we set $\sigma = \hat{\sigma}$ and the true change-plane parameter as $\boldsymbol{\theta}_0 =$

$(-0.576, 0.037, -0.816)^T$. Covariate *age* is assumed from a normal distribution with estimated mean 35.33 and standard deviation 8.75 and covariate *homo* is from a Bernoulli distribution with the success probability 0.66, similar to those in the AIDS data. For simplicity, we set the difference $\mu(\boldsymbol{X}) - h(\boldsymbol{X}, \boldsymbol{\beta}_0) = 0$. The estimated sample sizes for different treatment effect sizes are given in Table 4.9. For this AIDS study, there were 1046 subjects receiving either treatment 1 or treatment 0. Therefore, the proposed test can approximately achieve 90% power at treatment effect $\tau = 60$ for this study to identify a subgroup with an enhanced treatment effect.

4.6 Discussion

In this chapter, based on a change-plane analysis technique, we developed a doubly-robust testing procedure for detecting a subgroup with an enhanced treatment effect. We established the asymptotic distributions of the proposed test statistic under both the null and the local alternative hypotheses. We also developed its associated sample size calculation method, which is useful for sizing a clinical trial with desired power for subgroup detection.

In our current work, the subgroup with an enhanced treatment effect is defined by a change-plane, which may be restrictive sometimes. It is feasible to extend the way for defining a subgroup from a change-plane to more general forms, for example, a combination of multiple change-planes $1(\theta_{10} + \theta_{11}X_1 + \theta_{12}X_2 \geq 0)$ and $1(\theta_{20} + \theta_{21}X_3 \geq 0)$. However, a more complicated form for defining a subgroup requires a more comprehensive way of searching the supremum of squared score-type statistics over the possible space, which may be challenging. One assumption made in the considered semiparametric model is that the enhanced treatment effect in the subgroup is constant. It is also interesting to study a more general situation that the magnitude of the enhanced treatment effect varies for subjects in the subgroup. Lastly, it is likely that many covariates are collected at the baseline but not all of them are useful for subgroup

detection. Therefore, a built-in variable selection for subgroup detection will be very helpful, which warrants further investigation.

Table 4.1 Type I errors of the proposed test based on resampling. (Corresponding standard errors for type I errors with size 0.05 and 0.1 are 0.003 and 0.004.)

		B-Model I		B-Model II		B-Model III	
n	P-Model	size 0.05	size 0.1	size 0.05	size 0.1	size 0.05	size 0.1
100	I	0.052	0.104	0.054	0.107	0.050	0.099
	II	0.050	0.105	0.052	0.110	0.051	0.106
500	I	0.052	0.100	0.045	0.102	0.047	0.092
	II	0.055	0.105	0.051	0.106	0.054	0.101
1000	I	0.050	0.101	0.049	0.100	0.053	0.108
	II	0.051	0.105	0.044	0.102	0.053	0.108

Table 4.2 Power (%) of the proposed test at 0.05 and 0.1 levels. (Standard errors are shown in parenthesis.)

			B-Model I		B-Model II		B-Model III	
n	P-Model	τ	size 0.05	size 0.1	size 0.05	size 0.1	size 0.05	size 0.1
500	I	0.1	21.2 (1.3)	31.1 (1.5)	17.5 (1.2)	27.5 (1.4)	9.9 (0.9)	16.1 (1.2)
		0.25	90.3 (0.9)	95.1 (0.7)	81.3 (1.2)	88.4 (1.0)	45.9 (1.6)	57.7 (1.6)
		0.5	100 (0)	100 (0)	100 (0)	100 (0)	97.5 (0.5)	99.1 (0.3)
		−0.1	19.9 (1.3)	30.1 (1.5)	16.9 (1.2)	27.0 (1.4)	9.1 (0.9)	16.5 (1.2)
		−0.25	88.7 (1.0)	94.0 (0.7)	74.5 (1.4)	84.3 (1.2)	47.6 (1.6)	60.4 (1.5)
		−0.5	100 (0)	100 (0)	100 (0)	100 (0)	99.4 (0.2)	99.7 (0.2)
	II	0.1	18.8 (1.2)	29.5 (1.4)	21.4 (1.3)	30.8 (1.5)	11.1 (1.0)	18.6 (1.2)
		0.25	84.6 (1.1)	90.2 (0.9)	76.6 (1.3)	83.3 (1.2)	42.9 (1.6)	58.2 (1.6)
		0.5	100 (0)	100 (0)	99.9 (0.1)	100 (0)	97.1 (0.5)	98.8 (0.3)
		−0.1	20.1 (1.3)	29.6 (1.4)	18.3 (1.2)	26.4 (1.4)	12.7 (1.1)	20.8 (1.3)
		−0.25	84.5 (1.1)	91.5 (0.9)	73.9 (1.4)	82.5 (1.2)	46.2 (1.6)	58.0 (1.6)
		−0.5	100 (0)	100 (0)	99.9 (0.1)	100 (0)	98.4 (0.4)	98.8 (0.3)
1000	I	0.1	41.3 (1.6)	52.5 (1.6)	30.4 (1.5)	43.3 (1.6)	19.5 (1.3)	27.1 (1.4)
		0.25	99.8 (0.1)	99.9 (0.1)	99.0 (0.3)	99.5 (0.2)	77.1 (1.3)	86.0 (1.1)
		0.5	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)
		−0.1	39.3 (1.5)	51.3 (1.6)	29.9 (1.4)	43.1 (1.6)	17.3 (1.2)	24.3 (1.4)
		−0.25	99.7 (0.2)	99.8 (0.1)	98.2 (0.4)	99.4 (0.2)	78.8 (1.3)	86.4 (1.1)
		−0.5	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)
	II	0.1	36.2 (1.5)	48.5 (1.6)	29.7 (1.4)	42.1 (1.6)	15.5 (1.1)	23.7 (1.3)
		0.25	99.2 (0.3)	99.8 (0.1)	97.4 (0.5)	99.4 (0.2)	71.9 (1.4)	81.2 (1.2)
		0.5	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)
		−0.1	37.6 (1.5)	50.3 (1.6)	27.9 (1.4)	40.1 (1.6)	18.9 (1.2)	29.9 (1.4)
		−0.25	99.6 (0.2)	99.7 (0.2)	96.4 (0.6)	98.0 (0.4)	79.5 (1.3)	88.2 (1.0)
		−0.5	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)

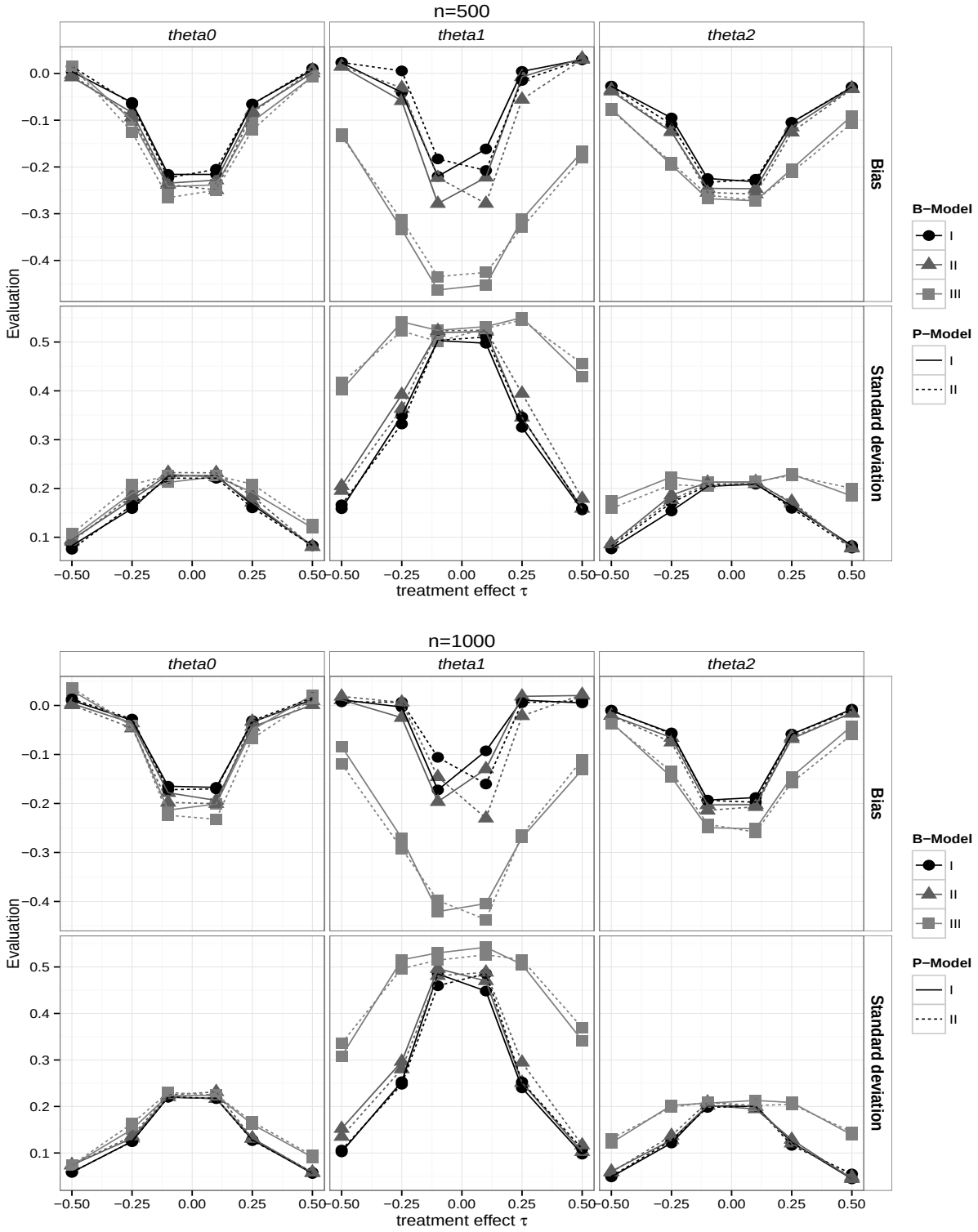

Figure 4.1 Bias and standard deviation of estimated change-plane parameter $\boldsymbol{\theta}$.

Table 4.3 Misclassification rate (%) of identified subgroup based on estimated change-plane parameter $\boldsymbol{\theta}$. (Standard errors are shown in parentheses.)

		P-Model I			P-Model II		
n	τ	B-Model I	B-Model II	B-Model III	B-Model I	B-Model II	B-Model III
500	0.1	26.3 (1.4)	28.3 (1.4)	33.3 (1.5)	26.2 (1.4)	30.6 (1.5)	33.1 (1.5)
	0.25	11.8 (1.0)	12.7 (1.1)	22.1 (1.3)	12.2 (1.0)	14.6 (1.1)	22.9 (1.3)
	0.5	4.8 (0.7)	4.7 (0.7)	12.4 (1.0)	4.8 (0.7)	5.2 (0.7)	12.3 (1.0)
	−0.1	26.6 (1.4)	29.2 (1.4)	33.0 (1.5)	26.6 (1.4)	29.0 (1.4)	34.1 (1.5)
	−0.25	12.0 (1.0)	14.1 (1.1)	21.7 (1.3)	11.8 (1.0)	14.2 (1.1)	23.6 (1.3)
	−0.5	4.6 (0.7)	6.3 (0.8)	9.9 (0.9)	4.9 (0.7)	6.1 (0.8)	11.2 (1.0)
1000	0.1	20.9 (1.3)	23.0 (1.3)	30.0 (1.5)	22.9 (1.3)	25.9 (1.4)	32.2 (1.5)
	0.25	7.6 (0.8)	8.5 (0.9)	17.4 (1.2)	8.1 (0.9)	9.4 (0.9)	18.0 (1.2)
	0.5	2.6 (0.5)	3.0 (0.2)	9.1 (0.9)	3.0 (0.5)	3.3 (0.5)	9.3 (0.9)
	−0.1	22.6 (1.3)	23.9 (1.3)	31.0 (1.5)	21.2 (1.3)	24.0 (1.4)	31.0 (1.5)
	−0.25	7.9 (0.9)	9.6 (0.9)	17.1 (1.2)	7.7 (0.8)	9.2 (0.9)	17.4 (1.2)
	−0.5	2.9 (0.5)	4.7 (0.7)	7.2 (0.8)	2.8 (0.5)	4.0 (0.6)	7.9 (0.9)

Table 4.4 Average timing performance (in seconds).

K	$n = 500$	$n = 1000$	$n = 2000$
1000	4.10 (0.29)	7.14 (0.79)	10.20 (1.22)
10000	36.55 (5.40)	62.74 (1.24)	102.72 (12.57)

Table 4.5 Type I error and power (%) of the EM test (Shen and He, 2015) and our method for B-Model I at 0.05 and 0.1 levels. (Standard errors are shown in parenthesis.)

		EM Test		Proposed Method	
n	τ	size 0.05	size 0.1	size 0.05	size 0.1
500	0	4.6 (0.7)	9.4 (0.9)	5.2 (0.3)	10.0 (0.4)
	0.1	6.9 (0.8)	12.6 (1.0)	21.2 (1.3)	31.1 (1.5)
	0.5	89.6 (0.9)	93.3 (0.8)	100 (0)	100 (0)
1000	0	5.4 (0.7)	10.4 (1.0)	5.0 (0.3)	10.1 (0.4)
	0.1	10.4 (1.0)	18.6 (1.2)	41.3 (1.6)	52.5 (1.6)
	0.5	99.8 (0.1)	100 (0)	100 (0)	100 (0)

Table 4.6 Misclassification rate (%) of identified subgroup based on Zhao et al. (2013) with different cut-off point c such that $\widehat{\mathrm{AD}}(c) = \tau$ and our method. Results are summarized under scenarios with P-Model I and B-Model I/II/III.

		Zhao et al. (2013)			Proposed Method		
n	τ	B-Model I	B-Model II	B-Model III	B-Model I	B-Model II	B-Model III
500	0.1	36.3 (0.6)	39.4 (0.6)	42.6 (0.6)	26.3 (1.4)	28.3 (1.4)	33.3 (1.5)
	0.5	13.5 (0.3)	13.5 (0.3)	17.6 (0.4)	4.8 (0.7)	4.7 (0.7)	12.4 (1.0)
1000	0.1	32.2 (0.5)	34.8 (0.6)	37.0 (0.5)	20.9 (1.3)	23.0 (1.3)	30.0 (1.5)
	0.5	10.9 (0.2)	11.2 (0.2)	13.3 (0.3)	2.5 (0.5)	3.0 (0.2)	9.1 (0.9)

Table 4.7 Results for sample size calculation. Here, n is the required sample size given by our procedure. Empirical power of the test with the calculated sample size is reported based on 1000 data replications.

$\mu(X)$		$1+X$		$1-X^2$		$1+\sin(\tilde{\pi}X)$	
τ	θ_0	n	Power	n	Power	n	Power
0.1	0	2992	91.3	4054	90.0	5440	91.1
	0.5	6034	91.8	8972	94.3	10924	90.6
	-0.5	2042	90.8	2726	90.2	3514	91.7
0.25	0	480	91.4	650	88.9	872	88.5
	0.5	966	91.8	1436	94.2	1748	89.7
	-0.5	328	90.8	436	89.1	564	88.7
0.5	0	120	87.2	164	85.6	218	85.4
	0.5	242	88.4	360	92.7	438	88.5
	-0.5	82	87.6	110	85.3	142	89.0

Table 4.8 Results for sample size calculation based on the method in Brookes et al. (2004) and Borm et al. (2007). Here, n is the required sample size. Empirical power of the test with the calculated sample size is reported based on 1000 data replications.

$\mu(X)$		$1+X$		$1-X^2$		$1+\sin(\tilde{\pi}X)$	
τ	θ_0	n	Power	n	Power	n	Power
0.1	0	1580	63.9	1840	57.0	2318	52.3
0.25	0	254	61.1	296	51.0	372	48.2
0.5	0	64	52.4	74	41.6	94	38.2

Table 4.9 Required sample sizes for detecting a subgroup with an enhanced treatment effect τ based on the AIDS study data.

treatment effect τ	sample size n
40	2392
60	1064
80	598
100	384

BIBLIOGRAPHY

Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, 61(4):821–856.

Andrews, D. W. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica*, 69(3):683–734.

Assmann, S. F., Pocock, S. J., Enos, L. E., and Kasten, L. E. (2000). Subgroup analysis and other (mis) uses of baseline data in clinical trials. *The Lancet*, 355(9209):1064–1069.

Auyeung, S. F., Long, Q., Royster, E. B., Murthy, S., McNutt, M. D., Lawson, D., Miller, A., Manatunga, A., and Musselman, D. L. (2009). Sequential multiple-assignment randomized trial design of neurobehavioral treatment for patients with metastatic malignant melanoma undergoing high-dose interferon-alpha therapy. *Clinical Trials*, 6(5):480–490.

Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 79(4):551–563.

Bonetti, M. and Gelber, R. D. (2004). Patterns of treatment effects in subsets of patients in clinical trials. *Biostatistics*, 5(3):465–481.

Borm, G. F., Fransen, J., and Lemmens, W. A. (2007). A simple sample size formula for analysis of covariance in randomized clinical trials. *Journal of clinical epidemiology*, 60(12):1234–1238.

Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A., and Peters, T. J. (2004). Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *Journal of clinical epidemiology*, 57(3):229–236.

Cai, T., Tian, L., Wong, P. H., and Wei, L. (2011). Analysis of randomized comparative clinical trial data for personalized treatment selections. *Biostatistics*, 12(2):270–282.

Candès, E. and Tao, T. (2007). Rejoinder: "The Dantzig selector: statistical estimation when p is much larger than n" [Ann. Statist. **35** (2007), no. 6, 2313–2351; mr2382644]. *Ann. Statist.*, 35(6):2392–2404.

Chakraborty, B., Murphy, S., and Strecher, V. (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Stat. Methods Med. Res.*, 19(3):317–343.

Chen, J. and Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, 95(3):759–771.

Collins, L. M., Murphy, S. A., and Strecher, V. (2007). The multiphase optimization strategy (most) and the sequential multiple assignment randomized trial (smart): new methods for more potent ehealth interventions. *American Journal of Preventive Medicine*, 32(5):S112–S118.

Cui, L., James Hung, H., Wang, S. J., and Tsong, Y. (2002). Issues related to subgroup analysis in clinical trials. *Journal of biopharmaceutical statistics*, 12(3):347–358.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64(2):247–254.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 74(1):33–43.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.*, 96(456):1348–1360.

Fava, M., Rush, A. J., Trivedi, M. H., Nierenberg, A. A., Thase, M. E., Sackeim, H. A., Quitkin, F. M., Wisniewski, S., Lavori, P. W., Rosenbaum, J. F., et al. (2003). Background and rationale for the sequenced treatment alternatives to relieve depression (star* d) study. *Psychiatric Clinics of North America*, 26(2):457–494.

Foster, J. C., Taylor, J. M., and Ruberg, S. J. (2011). Subgroup identification from randomized clinical trial data. *Statistics in medicine*, 30(24):2867–2880.

Gail, M. and Simon, R. (1985). Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*, pages 361–372.

Gill, R. D. and Robins, J. M. (2001). Causal inference for complex longitudinal data: the continuous case. *Annals of Statistics*, pages 1785–1811.

Gunter, L., Zhu, J., and Murphy, S. (2011). Variable selection for qualitative interactions. *Statistical methodology*, 8(1):42–55.

Hammer, S. M., Katzenstein, D. A., Hughes, M. D., Gundacker, H., Schooley, R. T., Haubrich, R. H., Henry, W. K., Lederman, M. M., Phair, J. P., Niu, M., et al. (1996). A trial comparing nucleoside monotherapy with combination therapy in hiv-infected adults with cd4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*, 335(15):1081–1090.

Kuk, A., Li, J., and Rush, A. J. (2010). Recursive subsetting to identify patients in the star* d: a method to enhance the accuracy of early prediction of treatment outcome and to inform personalized care. *The Journal of clinical psychiatry*, 71(11):1502–1508.

Liang, K.-Y., Self, S. G., and Liu, X. (1990). The cox proportional hazards model with change point: An epidemiologic application. *Biometrics*, 46(3):783–793.

Lu, W., Zhang, H. H., and Zeng, D. (2013). Variable selection for optimal treatment decision. *Statistical methods in medical research*, 22(5):493–504.

Lv, J. and Fan, Y. (2009). A unified approach to model selection and sparse recovery using regularized least squares. *Ann. Statist.*, 37(6A):3498–3528.

Moodie, E. E., Richardson, T. S., and Stephens, D. A. (2007). Demystifying optimal dynamic treatment regimes. *Biometrics*, 63(2):447–455.

Murphy, S., Van Der Laan, M., and Robins, J. (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association*, 96(456):1410–1423.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 65(2):331–366.

Murphy, S. A. (2005). An experimental design for the development of adaptive treatment strategies. *Statistics in medicine*, 24(10):1455–1481.

Neyman, J., Dabrowska, D., and Speed, T. (1990). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 5(4):465–472.

Peto, R. (1982). Statistical aspects of cancel trials. In Halnan, K., editor, *Treatment of Cancer*, pages 867–871. Chapman, London, UK.

Piantadosi, S. and Gail, M. (1993). A comparison of the power of two tests for qualitative interactions. *Statistics in Medicine*, 12(13):1239–1248.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of statistics*, 39(2):1180.

Qian, M., Nahum-Shani, I., and Murphy, S. A. (2013). Dynamic treatment regimes. In *Modern Clinical Trial Analysis*, pages 127–148. Springer.

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure periodapplication to control of the healthy worker survivor effect. *Mathematical Modelling*, 7(9):1393–1512.

Robins, J., Hernan, M., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiol.*, 11:550–560.

Robins, J. M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics-Theory and methods*, 23(8):2379–2412.

Robins, J. M. (1997). Causal inference from complex longitudinal data. In *Latent variable modeling and applications to causality*, pages 69–117. Springer.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In *Proceedings of the second seattle Symposium in Biostatistics*, pages 189–326. Springer.

Robins, J. M. and Rotnizky, A. (2001). Comment on the Bickel and Kwon article, "Inference for semiparametric models: some questions and an answer". *Statistica Sinica*, 11(4):920–936.

Rothwell, P. M. (2005). Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *The Lancet*, 365(9454):176–186.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, pages 34–58.

Rush, A. J., Fava, M., Wisniewski, S. R., Lavori, P. W., Trivedi, M. H., Sackeim, H. A., Thase, M. E., Nierenberg, A. A., Quitkin, F. M., Kashner, T. M., et al. (2004). Sequenced treatment alternatives to relieve depression (star* d): rationale and design. *Controlled clinical trials*, 25(1):119–142.

Shen, J. and He, X. (2015). Inference for subgroup analysis with a structured logistic-normal mixture model. *Journal of the American Statistical Association*, 110(509):303–312.

Song, R., Wang, W., Zeng, D., and Kosorok, M. R. (2015). Penalized q-learning for dynamic treatment regimens. *Statistica Sinica*, 25(3):901–920.

Song, X. and Pepe, M. S. (2004). Evaluating markers for selecting a patient's treatment. *Biometrics*, 60(4):874–883.

Thall, P. F., Millikan, R. E., Sung, H.-G., et al. (2000). Evaluating multiple treatment courses in clinical trials. *Statistics in Medicine*, 19(8):1011–1028.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 73(3):273–282.

Tsiatis, A. (2007). *Semiparametric theory and missing data*. Springer Science & Business Media.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

Wang, R., Lagakos, S. W., Ware, J. H., Hunter, D. J., and Drazen, J. M. (2007). Statistics in medicinereporting of subgroup analyses in clinical trials. *New England Journal of Medicine*, 357(21):2189–2194.

Watkins, C. J. (1989). *Learning from delayed rewards*. PhD thesis, University of Cambridge, England.

Watkins, C. J. and Dayan, P. (1992). Q-learning. *Machine learning*, 8(3-4):279–292.

Yan, X. (2004). Test for qualitative interaction in equivalence trials when the number of centres is large. *Statistics in medicine*, 23(5):711–722.

Yusuf, S., Wittes, J., Probstfield, J., and Tyroler, H. A. (1991). Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *Jama*, 266(1):93–98.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2013). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika*, page ast014.

Zhao, L., Tian, L., Cai, T., Claggett, B., and Wei, L.-J. (2013). Effectively selecting a target population for a future comparative study. *Journal of the American Statistical Association*, 108(502):527–539.

Zhao, Y., Zeng, D., Socinski, M. A., and Kosorok, M. R. (2011). Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*, 67(4):1422–1433.

APPENDIX

Supplementary Materials for Chapter 4

Proof of Theorem 1. By Taylor expansion and assumptions A1-A2, we have

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta}) \\ = & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \psi_1(\boldsymbol{Z}_i, \boldsymbol{\eta}_0; \boldsymbol{\theta}) - K_1^T C_1^{-1} \psi_2(\boldsymbol{Z}, \boldsymbol{\beta}_0) - K_2^T C_2^{-1} \psi_3(\boldsymbol{Z}, \boldsymbol{\gamma}_0) \right\} + o_p(1) \qquad (6) \\ = & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_{1*}(\boldsymbol{Z}_i, \boldsymbol{\eta}_0; \boldsymbol{\theta}) + o_p(1), \end{aligned}$$

where $\psi_{1*}(\boldsymbol{Z}_i, \boldsymbol{\eta}_0; \boldsymbol{\theta})$'s are i.i.d. with mean 0 under the null when either the propensity score model or the baseline mean function is correctly specified.

In addition, by assumptions A3-A5, we can show that the class $\mathcal{F} = \{\psi_{1*}(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}) : \|\boldsymbol{\theta}\| = 1\}$ is P-Donsker. Therefore, $n^{-1/2} \sum_{i=1}^{n} \psi_1(\boldsymbol{Z}_i, \tilde{\boldsymbol{\eta}}; \boldsymbol{\theta})$ converges weakly to a mean zero Gaussian process with the covariance function $E\{\psi_{1*}(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}_1)\psi_{1*}(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta}_2)\}$, where $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$. Finally, it is easy to show that the variance estimator $\tilde{V}_S(\boldsymbol{\theta})$ converges uniformly to $E\{\psi_{1*}^2(\boldsymbol{Z}, \boldsymbol{\eta}_0; \boldsymbol{\theta})\}$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ under both the null and the local alternative hypotheses. Therefore, the results established in Theorem 1 hold.

□

Proof of Theorem 2. Under the local alternatives, we have the same asymptotic representation (6). In addition,

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_1(\boldsymbol{Z}_i, \boldsymbol{\eta}_0; \boldsymbol{\theta}) \\ = & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{1}\left(\boldsymbol{\theta}^T \boldsymbol{X}_i \geq 0\right) \{A_i - \pi(\boldsymbol{X}_i, \boldsymbol{\gamma}_0)\} \left\{ Y_i - h(\boldsymbol{X}_i, \boldsymbol{\beta}_0) - \frac{\delta}{\sqrt{n}} A_i \mathbf{1}\left(\boldsymbol{\theta}_0^T \boldsymbol{X}_i \geq 0\right) \right\} \\ & + \frac{1}{n} \sum_{i=1}^{n} \delta \mathbf{1}\left(\boldsymbol{\theta}^T \boldsymbol{X}_i \geq 0, \boldsymbol{\theta}_0^T \boldsymbol{X}_i \geq 0\right) A_i \{A_i - \pi(\boldsymbol{X}_i, \boldsymbol{\gamma}_0)\}. \end{aligned}$$

The terms in the first summation are i.i.d. with mean 0 under the local alternatives when either the propensity score mode or the baseline mean function is correctly specified. As in Theorem 1, it can be show that the first summation term converges weakly to the same mean zero Gaussian process as $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_1(\boldsymbol{Z}_i, \boldsymbol{\eta}_0; \boldsymbol{\theta})$ does under the null. In addition, it can be shown that the second summation term converges uniformly to $\delta E\left[\left\{\mathbf{1}(\boldsymbol{\theta}_0^T \boldsymbol{X} \geq 0, \boldsymbol{\theta}^T \boldsymbol{X} \geq 0)\pi_0(\boldsymbol{X})\{1 - \pi(\boldsymbol{X}, \boldsymbol{\gamma}_0)\right\}\right]$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Therefore, the results established in Theorem 2 hold.

□