Value Search Estimators of Individualized Treatment Regimes Using a New Class of Weights

Yuxin Fan

Master of Science

Department of Epidemiology, Biostatistics and Occupational Health

McGill University

Montreal, Quebec

April 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Yuxin Fan

DEDICATION

I would like to dedicate this thesis to my dearest parents, whose unconditionally emotional and financial support and constant encouragement have always been, and will always be, my greatest source of strength and inspiration.

ACKNOWLEDGEMENTS

I would like to express my most sincere gratitude and thanks to my supervisor, Professor Erica Moodie, for proposing this interesting and challenging project, and for her vast guidance, support, and help throughout the work of this thesis. I am particularly grateful to Dr. Moodie for holding weekly meetings for this past year, for always being very patient and suggesting viable avenues to solve the numerous problems I encountered, for raising constructive criticisms that were absolutely essential in progressing my thesis, and for shaping me into the statistician I am today. Her professional suggestions always tackling problems and she solves these problems in a very organized and detailed fashion will definitely have a huge influence on my future career. I could not be more happier that I have the luck to meet her and choose her as my supervisor.

I also wish to thank my friends Mei Shuyan, Dr. Luo Yu who has been sources of consistent support and great comfort during the past two years.

My special thanks goes to my friend Dr. Pang Menglan who provide a lot of needed guide and computational expertise and being super nice and patient for a very long period of time.

I would like to send my thanks to my dearest friend Dr. Chen Chen for always being my great support for past twenty years.

I am very appreciated to my friend Luc Villandre for his suggestions and translation of my French abstract.

I would like to express my appreciation to my thesis commitee Dr. Nema Dean for

her valuable suggestions for this thesis.

I want to send my special thanks to my best friend Xu Ying, who has always enlightened me by her wisdom, creativity, strength, and keen logic for the past six years.

I am also grateful to all the professors who have taught me.

ABSTRACT

Personalized medicine is a rapidly growing field of health research. Dynamic treatment regimes (DTRs) are a way of formalizing the sequence of decisions that are made based on the personal medical history. Value search estimators such as inverse probability weighted estimators (IPWE) and augmented inverse probability weighted estimators (AIPWE) are frequently used for estimating DTRs. These estimators directly specify a restricted class of regimes and find the optimal regime by maximizing the expected outcome under each of the regimes in the class. The IPWE is a singly robust estimator which requires the correct specification of the treatment model, however, the AIPWE enjoys double robustness properties: an unbiased estimator is obtained provided at least one of the outcome regression model or treatment model is correctly specified. Recently, a new method of estimating DTRs was proposed, dynamic weighted ordinary least squares (dWOLS) that combines two established methods: Q-learning and G-estimation. In this thesis, instead of using the original inverse probability weights, I propose the use of dWOLS-style weights in singly- and doubly-robust value-search estimators to estimate the optimal DTRs. The new singly-robust estimators with the dWOLS weights are proven to possess the consistency property, whereas the doubly-robust estimators are shown not to achieve consistency. I illustrate the performance of the newly proposed estimation methods through simulation studies and further illustrate them in an analysis of the United States National Health and Nutrition Examination Survey.

RÉSUMÉ

La médecine personnalisée est un domaine de recherche en pleine croissance. Les programmes dynamiques de traitement (PDT) permettent de formaliser et d'adapter une série de décisions en fonction de l'historique médicale d'un patient. L'estimation des PDT dépend souvent d'estimateurs pondérés selon la probabilité inverse (EPPI) ou d'estimateurs augmentés pondérés selon la probabilité inverse (EAPPI). Ces deux familles d'estimateurs spécifient non seulement une classe restreinte de programmes, elles permettent aussi d'identifier le programme optimal au moyen de la maximisation de la réponse espérée pour chaque programme. Les EPPI sont des estimateurs robustes simples: ils requièrent une spécification correcte du modèle de traitement. Les EAPPI bénéficient toutefois d'une double robustesse: ils sont non biaisés à condition que soient correctement spécifiés le modèle de traitement ou le modèle de régression pour la réponse. Récemment, une nouvelle méthode d'estimation des PDT a été proposée: la méthode des moindres carrés ordinaires, pondérés et dynamiques (MCOPD). Elle combine deux méthodes communes: le Q-learning et le G-estimation. Dans ce mémoire, plutôt que la pondération selon la probabilité inverse conventionnelle, je propose la pondération de style MCOPD pour les estimateurs robustes simples et doubles dans le contexte des PDT. Je démontre que le nouvel estimateur robuste simple est cohérent, contrairement à l'estimateur robuste double. J'illustre la performance des méthodes proposées par l'intermédiaire de simulations et de l'analyse de données tirées du National Health and Nutrition Survey (NHANES) des États Unis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 General Overview of Personalized Medicine

Personalized medicine (PM), also known as precision medicine, is becoming a more and more important topic. PM is a medical model that proposes the customization of healthcare for each patient. The basic idea behind personalized medicine is that medical treatment is tailored for each individual patient.

The primary and essential motivation of personalized medicine is the well-established fact that patients often respond in different ways to a specific medical treatment, both in terms of the primary outcome as well as side-effects. Benefits of personalized medicine include increased compliance or adherence to medical treatment, having the option of enhanced patient care by selecting the optimal treatment, and reduction of the overall cost of healthcare since practices or products are tailored to the individual patient's needs. In this model, diagnostic testing is often employed for selecting appropriate and optimal treatments (Chakraborty and Moodie, 2013). The term "personalized treatment" is often used in the context of individualization on genetic information, as well as on other measures such as symptoms or other clinical responses.

The primary reason that statisticians are interested in this increasingly important topic is the growing interest in making the personalized treatment more evidence-based or data-driven, hence posing new methodological challenges that are

often beyond the scope of traditional quantitative tools. Much of statistical literature has focused on personalized treatment for chronic diseases, for example both Tsuang and Woolson (1977) and Jorm et al. (1997) focus on depression and schizophrenia; I will describe various study designs and statistical analysis methods that aid in developing evidence-based personalized treatments for chronic diseases. For effective long-term care of the patients, many of these chronic conditions require ongoing medical intervention, following the chronic care model (CCM) (Wagner et al., 2001). This poses additional challenges for the paradigm of personalized medicine since the personalization has to happen through multiple stages of intervention. In this context, dynamic treatment regimes (Murphy, 2003; Robins, 2004; Lavori and Dawson, 2004) offer a vehicle to operationalize the sequential decision-making process involved in the personalized clinical practice consistent with the CCM, and thereby a potential way to improve it.

In this work, I will restrict my research to a single treatment decision rather than a longitudinal sequence of decisions.

1.2 Individualized Treatment Regimes and Dynamic Treatment Regimes

An individualized treatment regime (ITR) is a mathematical function that takes in covariates and outputs a decision for a single stage of treatment; a dynamic treatment regime (DTR) is simply a sequence of ITRs for treatment given longitudinally in sequence, i.e. a DTR is a set of sequential decision rules in which decisions on treatment are made for a given state (e.g. demographics, case history, genetic information, etc.) of the patient that can change over time; therefore a DTR could

involve changing the treatment based on the patient response to the previous treatment. In comparison, for traditional treatments, all subjects are assigned the same level and type of treatment, and the assigned treatments do not depend on or update according to the status of patients. ITRs allow treatment to be individualized to the patient through a systematic set of rules.

1.3 Thesis Aims and Structure

There are a variety of approaches to estimating an ITR, all of which fall into two broad classes: regression-based methods and value search (classification-based) estimators.

The regression-based approach to estimating an optimal ITR, which indirectly estimates the optimal ITR, seeks to classify subjects into different risk levels estimated by a parametric or semi-parametric regression model using prognostic factors and then to assign therapy according to risk level. For instance, one could model the conditional expectation of the mean outcome given history and intermediate decisions via regression, and then find the difference in outcomes from treatments and optimize the estimated mean or contrast model at each stage, eventually leading to an optimal personalized treatment sequence. However, the parametric or semi-parametric regression model assumptions may not be valid due to the complexity of the disease mechanism and individual heterogeneity. If the model specification is not correct, these methods run the risk of providing a biased estimator of the optimal decision rule.

As an alternative, the second method of estimating optimal ITRs, value search estimation, proposes an approach that directly maximizes the value function or

marginal mean outcome of each member in a pre-specified class of regimes, often indexed by some parameter $\eta \in \Theta$, and then picks the regimes that maximize the estimated mean outcome of the population amongst each of the candidate regimes. This method emphasizes prediction accuracy of the clinical response model instead optimizing the decision rule directly. Value search estimators include familiar approaches such as inverse probability weighting.

Recently, a new method of estimating ITRs was proposed, called dynamic weighted ordinary least squares. The approach introduced a new class of weights to be used in a regression-based approach to ITR estimation. The objective of the research undertaken in this thesis is to investigate the use of this class of weights in value search estimation of ITRs.

The thesis is structured as follows. In Chapter 2, I review some popular methods to estimate ITRs via both regression-based approaches and value search estimators. In Chapter 3, I outline the proposed method of using the new family of weights in value search estimators and consider asymptotic properties of the estimators. The proposed approaches are then evaluated via simulations in Chapter 4, and applied to the United States' National Health and Nutrition Examination Survey in Chapter 5. Chapter 6 summarizes the contributions to the statistical literature made in this thesis and concludes with future work.

CHAPTER 2
Methods to Estimate Optimal Individualized Treatment Regimes

In this Chapter, I will explain the necessary data, and possible data sources, for estimation of an optimal ITR. I will introduce the notation that will be used throughout the thesis, and provide a literature review of existing methods used to estimate ITRs. An ITR provides an algorithm for treatment, and this algorithm may resemble clinical care provided by physicians, who typically attempt to summarize all relevant patient information to obtain a treatment recommendation that is best for a specific individual. To estimate a regime using data, the following information is needed:

1. treatment options, including different medications, drugs, dose-level, modes of delivery, time schedules, etc;
2. critical decision points at which treatment is assessed and decisions are made (i.e. when are decisions made to continue, stop, add, or subtract treatment);
3. how treatment decisions are currently made (measurements on those covariates used to make treatment decisions, and information on how are they used);
4. measurements on other covariates that might also be useful for making treatment-tailored decisions;
5. and possibly also measurements on any additional prognostic variables that would help to predict the outcome.

Within the ITR framework, I can use data to formalize clinical decision making such that treatment rules are data-based. An optimal treatment regime is that leading to the greatest benefit (biggest value of a positive outcome) overall in the target patient population. As we shall soon see, information on items (1), (2), and (4) are essential for any statistical analysis of a data source that aims to estimate an ITR. All methods of estimation additionally require at least one of (3) or (5) to be available to the analyst as well.

2.1 Data Sources

The data for discovering optimal ITRs can be obtained from either observational studies (non-experimental data) or randomized trials (experimental data).

2.1.1 Observational Data

Non-experimental data, i.e. observational data, may arise from a variety of sources, such as cohort studies, electronic medical records, etc. The upside of this kind of data is that it is often possible to obtain data on a large number of individuals at a low cost. However, some of the important variables may not be available from the data set and it is then not easy to learn about their potential to improve through personalization (Chakraborty and Moodie, 2013). There has been vigorous research recently on estimating optimal treatment regimes based on the data from observational studies by Sterne et al. (2005), Petersen et al. (2014), etc.

Drawing a causal interpretation from observational data relies heavily on several assumptions which I will be discussing in Section 2.1.4.

I know that estimation from observational data can be tricky owing to confounding and various hidden biases; thus, when randomized data are available, they

are often preferred (Rubin, 1974; Holland, 1986; Rosenbaum, 1991). It is crucial for researchers to generate meaningful data for discovering optimal ITRs, which is a developmental procedure rather than a confirmatory procedure. Thus, the usual confirmatory randomized trials are no longer especially helpful and a special class of designs, called sequential multiple assignment randomized trial (SMART) designs, tailor-made for the purpose of discovering optimal DTRs, has been developed.

2.1.2 Sequential Multiple Assignment Randomized Trial

Randomized trials are the gold standard in study design, as randomization coupled with compliance allows causal interpretations to be drawn from the statistical association (relation) (Chakraborty and Moodie, 2013).

The sequential multiple assignment randomized trial (SMART) design framework was proposed by Murphy (2005). A SMART design is a special kind of multistage randomized clinical trial design in which the assumption of "sequentially ignorable treatment" holds by design. The assumption of sequential ignorability asserts that the assigned component is independent of future potential outcomes, conditional on the subject history, but is untestable in data from observational studies (Kosorok and Moodie, 2015).

A SMART starts with an initial randomization of patients to possible treatments options, followed by re-randomization at each subsequent treatment decision stage of all patients to another treatment at each stage.

The main and very important goal of SMARTs is to provide data that allow estimation of DTRs through a series of trials and mimics that of a usual complete

randomized trial in order to perform a more standard analysis and it allows comparison of two or more DTRs by using hypothesis testing which can lead to discovering the optimal outcome. The secondary and important goal of SMARTs is to provide data to measure the detailed tailoring or personalization of treatment, i.e. by estimating the effects of DTRs; the optimal DTR is likely to be some form of "assign the first treatment A and if subjects respond, stay on this treatment A; if subjects do not respond, switch to treatment B". I thus can discover more personalized regime by exploring the effects on particular subgroups based on the baseline or intermediate outcomes.

The rerandomization at each stage could depend on information collected after previous treatments, such as how well the patient responded to the previous treatment, before assigning the new treatment. For the purpose of time and feasibility, SMARTs are mostly designed with two stages and two or three treatment options at each stage but can be applied with designs that require more than two stages and three treatment options at a specific stage.

One of the earliest SMART trials was conducted in the field of mental health. CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness) was funded by National Institute of Mental Health (NIHM) and was designed to investigate the treatment sequences of antipsychotic drugs. The CATIE study was designed to mimic real-life practice in order to inform future clinical care. CATIE is a multistage SMART design with five treatment options at the first stage, four treatment options at the second stage First and eight treatment options at the third stage. First, patients were randomized to one of five treatments: perphenazine, olanzapine,

quietiapine, risperidone, and ziprasidone and were assessed for psychotic symptoms and side effects. Within 18 months, subjects could choose to switch medication if this treatment was not effective. If the first treatment was effective and tolerable then the patient could choose to stay on this treatment. Subjects who discontinue their assigned treatment an the first stage are recommended to one of the two treatment assignment pathways and subjects who discontinue the treatment assignment at the second stage can choose the treatment designed for third stage, etc. (Stroup et al., 2003)

Many variables were collected at study entry such as demographics, disease background; and a number of longitudinal outcome variables including schizophrenia symptom measurements, medication adherence and side effects (Kosorok and Moodie, 2015). Now consider a hypothetical example of a SMART design, that is similar in context to one of the analyses that will be presented in Chapter 5. I consider a population of sedentary adults and randomize people into one of two groups (i) reduce TV to at most 1 hour per day or (ii) continue watching at their usual level. If participants exhibit a high BMI after 6 months, re-randomize the people in arm (i) to attend an aerobics class once per week or to cut out TV altogether, and re-randomize those in (ii) to reduce their TV to 1 hour per day or to an aerobics class once per week without reducing their TV. I present the SMART in Figure 2-1 .

Figure 2–1: Diagram of the hypothetical two-stage SMART design mimicking NHANES the general context of the analyses in Chapter 5.

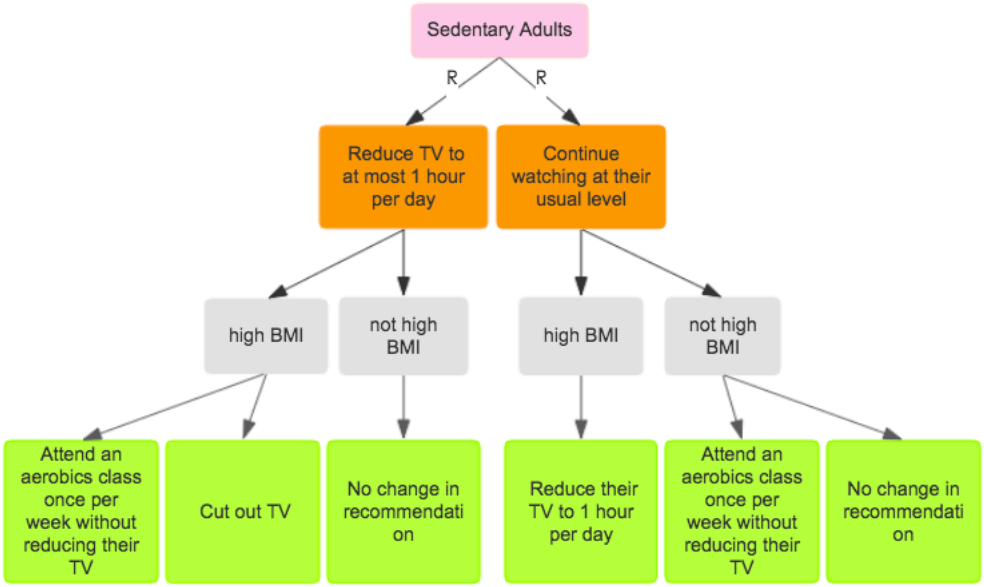


Figure 2-1: Pink boxes indicate baseline variables defining the (sub)populations of interest. Grey boxes are intermediate variables defining the subpopulations of interest, and on which treatment choices (and hence randomization probabilities) depend. Orange boxes denote first-stage randomization, whereas green boxes denote second stage treatment, which is randomized only for the subpopulations who exhibit high BMI at the intermediate measurement.

2.2 Preliminaries

2.2.1 Notation

I use the following notation through out this thesis. Suppose we have k distinct treatment intervals, $j = 1, ..., K$, then:

- y: Patient outcome, typically defined such that larger values are preferred;
- a_j: The jth treatment decision. I assume treatments are binary with $a_j = \{0, 1\}$, referred to as 'reference or control' and 'treatment', respectively;
- x_j: Covariate matrix containing non-treatment information (e.g. age, disease severity) available prior to the jth treatment decision. Note that x_j could also be a confounding variable, i.e. a common "cause" of both treatment and outcome. Those variables x_j that interact with treatment are called tailoring or prescriptive variables;
- h_j: Covariate matrix containing patient information or history prior to the jth treatment decision and denoted $h_j = \{x_1, a_1, x_2, a_2, ..., x_{j-1}, a_{j-1}, x_j\}$;
- A treatment rule is a function $d_j(h_j)$, $d_j : h_j \to a_j$;
- A DTR is a collection of treatment rules $(d_1, d_2, ..., d_k)$.

Note for one-interval setting ($k = 1$), the data would only contain (X_i, A_i, Y_i), $i = 1, 2, .., n$.

2.2.2 The Potential Outcomes Framework

Many of the existing methods for observational data rely on the notion of potential outcomes, also called counterfactuals, defined as a patient's outcome if he followed a particular treatment regime, possibly different from the regime that he was actually observed to follow. The individual-level causal effect of a regime may be

viewed as the difference in outcomes if a person had followed that regime as compared to a reference regime. The fundamental problem of causal inference lies in the definition of causal parameters at an individual level. Suppose we are interested in the causal effect of taking treatment a instead of a' in a simple single-stage randomized trial, denoted as $Y_i(a) - Y_i(a')$. Obviously, it is not possible to observe the outcomes of taking both treatments a and a'; thus the individual-level causal effect can never be observed. However, population-level causal parameters (or average causal effects) can be identified under randomization with perfect compliance, or bounded under randomization with non-compliance. Without randomization, i.e. in observational studies or indeed randomized trials with imperfect compliance, more assumptions are needed to estimate the population-level effects.

2.2.3 Assumptions

The axiom of consistency is the fundamental requirement in the potential outcome framework I have discussed above. This axiom states that the observed outcome under treatment a is equal to the counterfactual $Y(a)$, thus linking the unobservable counterfactuals with observable data.

I will now state the assumptions required for unbiased estimation of the treatment effects, and hence of an ITR:

- Stable unit treatment value assumption (SUTVA): A subject's outcome $Y(a)$ is not influenced by other subjects' treatment allocation (Rubin, 1980).
- Exchangeability, i.e. no unmeasured confounding (NUC)

$$A \perp \{Y(A)\}|X$$

- No extrapolation (Positivity): Subjects are to be exposable to both treatments (there is no contra indication). The mathematical definition of strong positivity is as follows

$$\inf_{a \in \mathcal{A}} P(A = a|X) > 0, a.e.$$

- Well-defined exposure (cannot have multiple versions of treatment).

The first assumption may be called no interference (or no interaction) between units (Cox, 1958) and is often reasonable, particularly in the context of randomized trials where study subjects are drawn from a large population. If one subjects' treatment affects other subject's outcome, then this assumption does not hold. There are some special circumstances in which SUTVA may fail, such as vaccinations for contagious disease where the phenomenon of "herd immunity" may lead to protection of unvaccinated individuals or in the context of group therapy where the interpersonal dynamic between group members could influence outcomes (Chakraborty and Moodie, 2013).

The second assumption always holds under randomization and may also be (approximately) true in observational settings where all relevant confounders have been measured.

Many confounding-adjustment methods in ITR estimation rely on the propensity score (PS). This approach requires the construction of a treatment model - often a predicted probability resulting from a logistic regression of treatment on covariates. The PS is defined to be the coarsest function

$$\pi(x) = P(A = 1|X = x)$$

where A is a binary treatment and X is a collection of measured covariates (Rosenbaum and Rubin, 1983). The PS is a balancing score such that, conditional on the propensity score, the distribution of measured covariates is similar between treated and untreated subjects, i.e. treatment received is independent of measured covariates X used to construct the PS.

The third assumption, positivity or experimental treatment assignment (ETA) states that each possible treatment level occurs with some positive probability within each stratum of covariate X; in particular, it requires that there are both treated and untreated individuals at every level of the covariate history. Positivity may be violated either theoretically or practically. A theoretical or structural violation occurs if the study design prevents certain individuals receiving a specific treatment (failure of one type of drug may preclude the prescription of other drugs in that class). A practical violation of this assumption may occur when a particular stratum of subjects has a very low probability of getting treatment (Chakraborty and Moodie, 2013).

Having now described the necessary notation and assumptions, I will next review existing approaches to the ITR estimation.

2.3 Regression-based Approaches to Estimation

Regression-based approaches suggest that estimation of an optimal dynamic treatment regime can be achieved by proposing and fitting regression models for the outcome based on history and treatment received to obtain an estimator for the optimal treatment regime. The most common regression-based methods in the literature are Q-learning (Watkins, 1989; Watkins and Dayan, 1992), and G-estimation

(Robins, 2004) with dynamic weighted ordinary least squares (Wallace and Moodie, 2015) a new addition.

2.3.1 Q-learning

Q-learning (Watkins, 1989; Watkins and Dayan, 1992) is a relatively simple method, which focuses on functions of quality of treatment, i.e. Q-functions, defined as

$$Q_K(h_K, a_K) = E[Y|H_K = h_K]$$

and

$$Q_j(h_j, a_j) = E[\max_{a_{j+1}} Q_{j+1}(H_{j+1}, A_{j+1})|H_j = h_j, A_j = a_j)].$$

In a simple one-stage setting, the Q-function is simply

$$Q(x, a) = E[Y|X = x].$$

In practice, when in a multi-interval setting, the Q-functions are not known, and thus we need to estimate them from the existing data, and we often model the Q-functions linearly as

$$Q_j(h_j, a_j; \beta_j, \psi_j) = \beta_j^T h_j^\beta + \psi_j^T a_j h_j^\psi$$

where h_j^β and h_j^ψ are sets of covariates that have a predictive effect on the outcome that is not modified by the treatment and the covariates that interact with treatment respectively. I give the proposed algorithm which calculates the estimates of ψ_j recursively.

1. Interval K parameter estimation: we need to find β_K and ψ_K by proposing a model, leading to the following estimates

$$(\hat{\beta}_k, \hat{\psi_k}) = \arg\min \frac{1}{n} \sum_{i=1}^{n} (Y_i - Q_k(h_{ki}, A_{ki}; \beta_k, \psi_k))^2.$$

2. Interval K optimal ITR is simply

$$a_K^{opt} : \arg\max_{a_K} Q_K(h_K, a_K; \hat{\beta_K}, \hat{\psi_K}).$$

3. Interval $j-1$ pseudo outcome:

$$\hat{Y}_{j-1,i} = \max_{a_j} Q_j(h_{ji}, a_j; \hat{\beta}_j, \hat{\psi_j}).$$

4. Interval $j-1$ parameter estimation:

$$(\hat{\beta}_{j-1}, \hat{\psi}_{j-1}) = \arg\min \frac{1}{n} \sum_{i=1}^{n} (Y_{j-1,i} - Q_{j-1}(h_{j-1i}, A_{j-1i}; \beta_{j-1}, \psi_{j-1}))^2.$$

5. Interval $j-1$ parameter rule:

$$a_{j-1}^{opt} : \arg\max_{a_{j-1}} Q_{j-1}(h_{j-1}, a_{-1j}; \hat{\beta}_{j-1}, \hat{\psi_{j-1}}).$$

6. Repeat this procedure until we get the estimates of all β_j's and ψ_j's, $j = 1, ...K$. In a one-interval setting, only steps 1-2 are required, so that a linear regression is performed and the optimal treatment is that which maximizes the expected response.

The Q-learning approach has the advantage of ease in computing, however, lack of robustness in the estimators could be problematic since Q-learning depends heavily on being able to correctly specify the model for the Q-functions. Hence, this approach needs to be undertaken with caution.

2.3.2 G-estimation

The G-estimation approach that Robins (1992) proposed is a recursive approach like Q-learning, and starts by estimating the optimal final-stage treatment and then using the data and this final-stage information, moves backward from the last stage to the first stage to estimate each previous optimal decision. It requires additional modeling, however it can offer stronger robustness to potential model misspecification. The goal of G-estimation is to estimate the parameters ψ of the optimal contrast function by a combination of regression models and estimating equations. The expected outcome is written as

$$E[Y|H = h] = \beta_j^T h_j^\beta + \sum_{j=1}^{k} \gamma_j(h_j, a_j; \psi_j). \tag{2.1}$$

The term $\beta_j^T h_j^\beta$ is the same as the one I presented in the Q-functions and the sum corresponds to the interaction terms in the Q-function. The G-estimation approach is based on semiparametric models and applies sequentially across intervals. The conditional expectation contains two parts.

The first part of (2.1) is the treatment-free contribution, and the second part of (2.1) is the sum of blip-to-reference functions to be defined as

$$\gamma_j(h_j, a_j) = E[Y^{\bar{a}_j, \underline{a}_{j+1}^{opt}} - Y^{\bar{a}_{j-1}, \underline{a}_j = 0, \underline{a}_{j+1}^{opt}} | H_j = h_j]$$

The blip function $\gamma_j(h_j, a_j)$, at any stage j is defined to be the expected difference in outcome when using a reference regime $a_j = 0$ instead of $a_j = 1$ at stage j. The blip functions compare the difference in expected outcome between a subject who receives treatment a at stage j, and one who receives the reference treatment,

assuming identical history up to stage j and optimal treatment thereafter.
The regret function is defined as

$$\mu_j(h_j, a_j) = E[Y^{\bar{a}_{j-1}, \underline{a}_j^{opt}} - Y^{\bar{a}_j, \underline{a}_{j+1}^{opt}} | H_j = h_j].$$

The regret function compares the difference between expected outcome under optimal treatment and expected outcome under assigned treatment. The term "regret" indicates how much we would regret by choosing a instead of a^{opt}. If treatment A is binary we have a relation between the blip and regret functions:
$\mu_j(h_j, a_j) = \gamma_j(h_j, a_j^{opt}) - \gamma_j(h_j, a_j)$. The blips and regret functions are essential to many ITRs estimating methods.

The procedure of G-estimation is as follows

1. Propose a model for the blip function $\gamma_j(h_j, a_j, \psi_j)$ and let $S(A_j) = \frac{d\gamma_j}{d\psi_j}$.
2. Define the G-function $G_j(\psi) = Y - \gamma_j(h_j, a_j, \psi_j) + \sum_{k=j+1}^{K} \mu_k(h_k, a_k, \psi_k)$
3. For expected treatment outcome model: propose a model for $\mathrm{E}[G_j(\psi)|H_j = h_j; \beta_j]$ and use the data to estimate the parameters and obtain estimates of β_j in terms of ψ_j.
4. For the treatment model: propose a model for $E[S(A_j)|H_j = h_j; \alpha_j]$, and use the data to get the estimates of $\hat{\alpha}_j$.
5. Construct the function
   $U_j(\psi_j, \beta_j, \alpha_j) = (G_j(\psi) - E[G_j(\psi)|H_j = h_j; \beta_j])(S(A_j) - E[S(A_j)|H_j = h_j; \alpha_j])$
   and by substituting parameter estimates from previous steps estimate the blip parameter by solving the equation system $0 = \sum_{i=1}^{n} U_j(\psi_j, \hat{\beta}_j(\psi_j), \hat{\alpha}_j)$.

In G-estimation, the goal of the treatment model is to eliminate imbalance in the distribution of covariates between treatment and untreated subjects.

2.3.3 Dynamic Weighted Ordinary Least Squares Approach

The dynamic weighted ordinary least squares (dWOLS) approach (Wallace and Moodie, 2015) builds on the method of G-estimation and Q-learning, and is also a recursive approach. The method involves doing a sequence of weighted ordinary least squares regressions in order to obtain blip parameter estimates. Here our assumptions are that the blip functions at each interval are linear in ψ_j and our treatment is binary. Then the algorithm of dWOLS is as follows

1. Define the outcome $\tilde{Y}_j = Y_j + \sum_{k=j+1}^{K} \mu_k(h_k, a_k, \psi_k)$.
2. For the treatment model:

   propose a model of $E[S(A_j)|H_j = h_j; \alpha_j]$ and use the data to obtain estimates $\hat{\alpha}_j$.
3. Choose a weight function $w_j(a_j, \mathbf{x}_j, \alpha_j)$, and use the estimates $\hat{\alpha}_j$ to obtain the estimates of weights $\hat{w}_j$.
4. Conduct a weighted least square regression of $\tilde{Y}_j$ on $(\mathbf{x}_j^{\beta}, a_j \mathbf{x}_j^{\psi})$ with weights $\hat{w}_j$ where $\mathbf{x}_j^{\beta}$ and $\mathbf{x}_j^{\psi}$ are variables and functions of variables to be included in the linear, treatment-free model.

Note that there are several different weights that I can potentially choose,

(1) absolute value weights: $w_{1i}(a_i, x_i) = |a_i - P(A_i = 1|\mathbf{X_i} = x_i)$

(2) inverse probability weights (IPW): $w_{2i}(a_i, x_i) = \frac{1}{P(A_i = a_i|\mathbf{X}=\mathbf{x})}$

(3) alternative dWOLS weights: $w_{3i}(a_i, x_i) = 1_{\{a_i=1\}} + 1_{\{a_i=0\}} \frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$ and $w_{4i}(a_i, x_i) = 1_{\{a_i=0\}} + 1_{\{a_i=1\}} \frac{1-P(A_i=1|\mathbf{X_i})}{P(A_i=1|\mathbf{X_i})}$, where $1_{\{a_i\}}$ is the indicator function that takes values in $\{0, 1\}$.

All of above weights satisfy $\pi(\mathbf{x})w(1, \mathbf{x}) = (1-\pi(\mathbf{x}))w(0, \mathbf{x})$, which provide consistent estimators of ψ (see proof of theorem 1 in the supplementary materials of Wallace and Moodie (2015)).

2.4 Value Search Approaches to Estimation

The regression based estimators have been criticized for parameterizing the ITR only indirectly, and often relying on linear specifications of the treatment rule. In contrast, value search estimators are more direct in that this approach first parameterizes the regimes and then finds the value, i.e. the expected outcome, associated with each regime in the set of regimes under consideration.

The most well-known value search estimation methods from the literature include marginal structural models (MSM), proposed by Robins et al. (2000); Hernán et al. (2000); Murphy et al. (2001), estimated by the inverse probability weighting estimator (IPWE) proposed by Robins et al. (2000), or the augmented inverse probability weighting estimator (AIPWE) of Zhang et al. (2012).

The value of a specific treatment regime, d, is denoted as $V^d = E[Y(d)]$, which is the expected outcome if all subjects were to follow this regime. The optimal regime is the one regime that maximizes the expected outcome. Many candidate regimes, for example, take the form "Treat subject i if their initial information X_i is greater than some value η", and all regimes are then defined by a set of thresholds, i.e. $\eta \in \Theta$.

The goal of value search estimators is to identify the optimal threshold, η^{opt}, from the defined set of thresholds, Θ.

These methods are more "direct" in that they first parameterize the regimes and then see the value associated with the set of regimes under consideration, in contrast to the regression-based approaches which parameterize the regime indirectly.

2.4.1 Inverse Probability Weighted Estimator

The inverse probability weighting method (Robins et al., 2000) lies in the second class of value search estimation strategy which rather than taking a regression-based approach where one needs to propose a parametric model, instead directly estimates the value function or marginal mean outcome of each member in a pre-specified class of regimes $\mathcal{D}$, often indexed by some parameter $\eta \in \Theta$, where $d \in \mathcal{D}$ is a function of covariates X and parameters η. Then picks the regime that maximizes the estimated value, i.e. the estimated optimal is

$$d^{opt} = \arg\max_{d \in \mathcal{D}} \hat{V}^d \equiv \arg\max_{\eta \in \Theta} \hat{V}^d$$

where $\hat{V}^d$ (or $\hat{V}^{d(\eta)}$) is the estimated value function of the regime d. Estimating the value of the candidate regimes can be done by collecting all the subjects whose realized treatment experiences are consistent with each candidate ITR, and computing the weighted average of the primary outcome.

In a single interval setting, I can write the value function as

$$V^d = E_d Y = \int Y dP_d = \int Y(\frac{dP_d}{dP_\pi})dP_\pi.$$

Note that changing the probability measure is the same idea as importance sampling (Hammersley and Handscomb; Hastings, 1970) in Monte Carlo simulation, hence the ratio $\frac{dP_d}{dP_\pi}$ can be simplified to

$$\frac{1_{\{A=d(X)\}}}{\pi(A|X)} \tag{2.2}$$

where $\pi(A|X)$ is the propensity score I have defined in Section 2.2.3. Thus

$$V^d = E_d Y = \int Y dP_d = \int Y \frac{1_{\{A=d(X)\}}}{\pi(A|X)} dP_\pi = \int w_{d,\pi} Y dP_\pi$$

where $w_{d,\pi} = \frac{1_{\{A=d(X)\}}}{\pi(A|X)}$ is a weight function depending on data. By using these weights, subjects whose treatment is "unusual" (i.e. subjects with a low probability of getting their specific treatment given X_i) are given a larger weight, thus balancing the confounding distribution in the treated and untreated groups.

An intuitively natural way to estimate V^d is $\hat{V}^d$,

$$\hat{V}^d = P_n[w_{d,\pi} Y]$$

where P_n denotes the empirical average over a sample size of n. Even though the expectation of the weight function $w_{d,\pi}$ is 1, it is preferable to normalize the weights by their sample mean to obtain a more stable estimate. The resulting estimator, called the inverse probability of treatment weighted estimator (Robins, 2000), is given by

$$\hat{V}^d_{IPWE} = \frac{P_n[w_{d,\pi} Y]}{P_n[w_{d,\pi}]} = \frac{1}{n} \sum_{i=1}^{n} \frac{1_{\{A=d(X)\}} y_i}{f_{A|X}(A|X)}.$$

Zhang et al. (2012) noted the $IPWE(\eta)$ is a value search estimator for $E[Y(d_\eta)]$ and equivalent to estimating the contrast function at each of the observed data points by

$$IPWE(\eta) = \frac{1}{n}\sum_{i=1}^{n}\frac{C_{\eta,i}Y_i}{\pi_c(X_i;\eta)}$$

where $\hat{\pi}(X_i)$ is a proposed model for the propensity score $\pi(X)$, and

$$\begin{aligned}\pi_c(X_i;\eta) &= \hat{\pi}(X_i)d(X,\eta) + [1-\hat{\pi}(X_i)][1-d(X,\eta)]\\ C_\eta &= Ad(X,\eta) + (1-A)\{1-d(X,\eta)\}\end{aligned}$$

and where

$$\hat{C}_{IPWE}(X_i) = \frac{A_i}{\hat{\pi}(X_i)}Y_i - \frac{1-A_i}{1-\hat{\pi}(X_i)}Y_i.$$

Note that there are two approaches to IPWE: a non-parametric approach as I described above and a parametric marginal structural model (Robins et al., 2000; Hernán et al., 2000), where the weighted average is computed but a regression model is proposed for V^d. The MSM estimation procedure via inverse probability weights is as follows:

1. Identify a set of candidate thresholds that we are interested in, denote the set of thresholds as Θ.
2. Fit treatment models: fit a logistic regression model for the probability of being treated at each interval including all confounding covariates.
3. Generate an augmented dataset defined by $\eta \in \Theta$.
4. Calculate the IPWEs by using fitted values of model in step 1 by (2.2).

5. Fit a model for the mean outcome by regressing on some function of parameters η, denoted by $f(\eta)$, e.g $E[Y|\eta] = \beta_0 + \beta_1\eta + \beta_2\eta^2$ and estimate its parameters by weighted ordinary least squares regression on this augmented dataset using the weights from last step.
6. Obtain η^{opt} as the value of η that maximizes the parametric model posited for $E[Y|\eta]$.

2.4.2 Augmented Inverse Probability Weighted Estimator

Zhang et al. (2012) proposed a doubly robust augmented inverse probability weighted estimator, which maximizes the mean outcome across all regimes in a pre-specified class and is more robust to misspecification of the mean regression models.

The expected outcome can be expressed as

$$E[Y(d)] = E_X[\mu(1,X)d(X) + \mu(0,X)(1-d(X))]$$

Note that IPWE can become doubly robust by augmentation as follows

$$\begin{aligned} E[Y(d)] &= E_X\left[\{Y(1_{\{A=d\}}) - \mu(d,X)\} + \mu(d,X)\right] \\ &= E_X[Y(1_{\{A=d\}}) - \mu(d,X)] + E_X\left[\mu(d,X)\right] \end{aligned}$$

where $\mu(a,X) = E[Y|A=a, X=x]$, and d^{opt} represents the regime leading to the largest value of $E[Y(d)]$ ($\hat{V}^d$) among $d \in \mathcal{D}$. i.e $d^{opt} = \arg\max_{d\in\mathcal{D}} E[Y(d)]$ ($\arg\max_{d\in\mathcal{D}} \hat{V}^d$).

Now I posit a parametric regression model $\mu(A,X;\beta)$ for $E[Y|A,X] = \mu(A,X)$. If I correctly specified the model, then $\mu(A,X;\beta) \to \mu(A,X)$.

I can view the proposed model $\mu(A, X; \beta)$ as defining the class of all treatment regimes indexed by β, $\mathcal{D}$, and Zhang et al. (2012) suggest a way to directly estimate the optimal regime.

Based on the above considerations to avoid the misspecification of proposed regression models, the AIPWE approach is viewed as an optimization problem which needs to identify an estimator for $E[Y(d_\eta)]$ and to maximize it directly in η to obtain an estimator $\hat{\eta}^{opt}$ for η^{opt} (Zhang et al., 2012). For fixed η, Zhang et al. (2012) have shown that the AIPWE for $E[Y(d_\eta)]$ is given by (The second term is the augmentation term after some algebra)

$$\begin{aligned} AIPWE(\eta) &= \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{C_{\eta,i}Y_i}{\pi_c(X_i;\eta,\hat{\gamma})} - \frac{C_{\eta,i}-\pi_c(X_i;\eta,\hat{\gamma})}{\pi_c(X_i;\eta,\hat{\gamma})}m(X_i;\eta,\hat{\beta})\right\} \\ &= \frac{1}{n}\sum_{i=1}^{n}\left\{d(X_i,\eta)\hat{C}_{AIPWE}(X_i)\right\} \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1-A_i}{1-\pi(X_i,\hat{\gamma})} - \frac{A-\pi(X_i,\hat{\gamma})}{1-\pi(X_i,\hat{\gamma})}\mu(0,X_i,\hat{\beta})\right\} \end{aligned}$$

where $\pi(X,\gamma)$ is a proposed model for the propensity score $\pi(X)$, $\hat{\gamma}$ is the maximum likelihood (ML) estimator for γ and

$$\begin{aligned} \pi_c(X_i;\eta,\hat{\gamma}) &= \pi(X_i;\hat{\gamma})d(X,\eta) + [1-\pi(X_i;\hat{\gamma})][1-d(X,\eta)] \\ C_\eta &= Ad(X,\eta) + (1-A)\{1-d(X,\eta)\} \\ m(X_i;\eta,\hat{\beta}) &= \mu(1,X_i,\hat{\beta})d(X_i,\eta) + \mu(0,X_i,\hat{\beta})\{1-d(X_i,\eta)\} \end{aligned}$$

and

$$\begin{aligned}\hat{C}_{AIPWE}(X_i) &= \frac{A_i}{\pi(X_i,\hat{\gamma})}Y_i - \frac{1-A_i}{1-\pi(X_i,\hat{\gamma})}Y_i \\ &\quad -\frac{A_i-\pi(X_i,\hat{\gamma})}{\pi(X_i,\hat{\gamma})}\mu(1,X_i,\hat{\beta}) - \frac{A_i-\pi(X_i,\hat{\gamma})}{1-\pi(X_i,\hat{\gamma})}\mu(0,X_i,\hat{\beta}).\end{aligned}$$

The AIPWE borrows information from the specified parametric regression model for the outcome whereas IPWE estimator makes no use of a regression model for the outcome. The AIPWE has the double robustness property and this estimator is still consistent for $E[Y(d_\eta)]$ if at most one of the propensity score $\pi(X;\gamma)$ or outcome regression model $\mu(A,X;\beta)$ is not incorrectly specified. Thus, AIPWE approach increases IPWE approach's robustness and often also its asymptotic efficiency by the augmented term.

2.4.3 Outcome Weighted Learning

Zhao et al. (2012) proposed an approach for estimating the optimal ITR within a weighted classification framework, where the weights are determined from the clinical outcomes. This particular method emphasizes prediction accuracy of the clinical response model instead of optimizing the decision rule directly. The method then reduces to a computational problem by substituting the 0-1 loss in the classification with a convex surrogate loss as is done with the support vector machine.

I assume our treatment assignments are binary and denoted by $A \in \{-1,1\}$, and following our previously-defined notation, x_j is the covariate matrix containing non-treatment information, and Y is the observed clinical outcome. By maximizing the expected outcome, I can obtain an optimal ITR. Note while the 0/1 coding of treatment is widely used in the causal inference literature, the $-1/1$ coding is more

common in Q-learning, SMART design, and OWL literature, thus while I shall use $-1/1$ coding in the section, I will adopt the $0/1$ coding in the rest of this thesis.

I assume that I observe i.i.d. data (X_i, A_i, Y_i) , $i = 1, 2, ..., n$ from a randomized trial. The approaches I have discussed previously indirectly estimate the optimal ITR, and may actually produce suboptimal ITR's if the model for Y given (X, A) is overfitted. Thus, Zhao et al. (2012) proposed a nonparametric approach that directly maximizes the value function, denoted by V^d given as follows, based on an OWL method, which relies on

$$V^d = E_d Y = \int Y dP_d = \int Y \left\{ \frac{dP_d}{dP_\pi} \right\} dP_\pi.$$

Recall the above integral is the same as the IPWE's value function of a regime d is,

$$V^d_{IPWE} = \frac{P_n[w_{d,\pi} Y]}{P_n[w_{d,\pi}]}.$$

Naturally, an optimal treatment regime, d^{opt}, is a regime that maximized V^d,

$$\begin{aligned} d^{opt} &= \arg\max_d V^d_{IPWE} \\ &= \arg\max_d E\left\{ \frac{1_{\{A=d(H)\}}}{f_{A|H}(A|H)} Y \right\}. \end{aligned}$$

Zhao et al. (2012) noticed that maximizing V^d is the same as finding the optimal ITR by minimizing the following

$$\begin{aligned} d^{opt} &= \arg\min_d E[Y|A=1] + E[Y|A=-1] - V^d \\ &= \arg\min_d E\left\{ \frac{1_{\{A\neq d(X)\}}}{A\pi + (1-A)/2} Y \right\}, \end{aligned}$$

which can be viewed as a weighted classification error and the estimator (viewed as 0-1 loss in machine learning) of this is defined by

$$\hat{d}^{opt} = \arg\min_{d} n^{-1} \sum_{i=1}^{n} \frac{Y_i}{A_i\pi + (1-A_i)/2} 1_{\{A \neq d(X)\}} \tag{2.3}$$

where $d(X_i)$ can be shown to represent $sign(f(x))$, which can be either positive or negative. However, equation (2.2) is very difficult to minimize because of the discontinuity and non convexity. Hence, Zhao et al. (2012) used a convex surrogate loss called hinge loss (Cortes and Vapnik 1995), which penalizes the complexity of the decision function to avoid overfitting. The following is their function to minimize

$$n^{-1} \sum_{i=1}^{n} \frac{Y_i}{A_i\pi + (1-A_i)/2} (1 - A_i f(X_i))^{+} \lambda_n ||f||^2 \tag{2.4}$$

where $x^{+} = max(x, 0)$, $\|f\|$ is some norm of f to penalize more complicated functions, and the term $\lambda_n ||f||^2$ is a tuning parameter.

Thus, they can carry out this minimization problem as a weighted classification problem via standard optimization techniques. Outcome weighted learning could be argued to be the most complex method of all the methods I have discussed thus far. Further, it is singly robust. Thus, OWL is the least preferred of the value-search estimation approaches as it is more difficult to understand and implement than IPWE and less robust than AIPWE.

2.5 Summary

In this chapter, I have reviewed of the most commonly used regression-based and value search approaches to estimation of optimal dynamic treatment regimes. These methods can be used in both single-stage setting and longitudinal settings.

Table 2–1: Comparison of ITR estimation methods

Method	Robustness	Specification Requirement	The Form of Decision Rule can be Estimated
Q − learning	singly-robust	outcome model	linear rules of the form "treat when $\eta_0 + \eta_1 X_1 + \ldots + \eta_k X_k > 0$"
G − estimation	doubly-robust	outcome model and propensity score model	linear rules of the form "treat when $\eta_0 + \eta_1 X_1 + \ldots + \eta_k X_k > 0$"
dWOLS	doubly-robust	outcome model and propensity score model	linear rules of the form "treat when $\eta_0 + \eta_1 X_1 + \ldots + \eta_k X_k > 0$"
IPWE	singly-robust	propensity score model	linear form of rules or more complicated form i.e. "treat when $I(\eta_1 > X_1) * ... * I(\eta_k > X_k)$"
AIPWE	doubly-robust	outcome model and propensity score model	linear form of rules or more complicated form i.e. "treat when $I(\eta_1 > X_1) * ... * I(\eta_k > X_k)$"
OWL	singly-robust	propensity score model	linear form of rules or more complicated form i.e. "treat when $I(\eta_1 > X_1) * ... * I(\eta_k > X_k)$"

In general, the value search estimators target the parameters of the treatment rule itself instead of the parameters of the mean outcome model or the contrast model. Value search estimators are often non-parametric and thus require milder assumptions about the outcome model specifications. Hence, the value search estimators are more robust. One significant shortcoming is that the estimators of the value functions and the corresponding estimated treatment regimes can be highly variable due to the weighting by inverse probabilities.

Now I discuss how well these methods perform and present the advantages and disadvantages of each method in the Table 2-1.

The Q-learning method is easy to implement because the idea and implementation are quite clear and simple, but it depends very much on correctly specifying models for Q-functions. Thus, the potential for model misspecification is high which may yield estimated regimes that are far from the true optimal regime.

The G-estimation method is based on a semi-parametric estimating function. G-estimation provides the double robustness property, but has traditionally been unpopular due to its unintuitive presentation in the literature.

The value search approach can handle more complicated rule forms, however, it can be difficult to implement when the number of tailoring variables is even moderately large.

Both the dWOLS and AIPWE approaches enjoy the double robustness property as well and are rather simple in computation and concept. AIPWE is the most promising and attractive of the value search estimators since it is doubly robust, whereas IPWE and OWL are not, at the modest cost of requiring specification of an

outcome model. Also, the AIPWE can provide more flexibility than the regression-based approaches of Q-learning, G-estimation, and dWOLS.

Although AIPWE offers many nice properties, there is still room for improvement and my goal is to fill this gap in the next chapter by proposing a more efficient, doubly robust value search estimator that incorporates the weights proposed for dWOLS by Wallace and Moodie (2015).

CHAPTER 3
Using dWOLS Weights in Value Search Estimators

3.1 Proposed Approach

In the last chapter, I noted that IPWE and AIPWE could be used with more complicated regime rule forms (non-linear ITRs); moreover, AIPWE enjoys the doubly robust property. I also notice that one specific approach from the literature of dynamic treatment regimes; namely the dynamic weighted ordinary least square regression proposed by Wallace and Moodie (2015), also provides double robustness. In particular, I observed that some weights in this framework provided greater efficiency in the decision rule parameter estimation compared to the conventional inverse probability weights.

The main reason that I prefer the absolute value weights rather than inverse probability weights is that I anticipate smaller variance owing to the variability of the weights themselves. The absolute value weights are bounded by $|a_i - P(a_i|X)|$ which is always less than or equal to 1. Based on this boundary, I would expect the absolute weights w_{1i} to have a smaller variance compared to the inverse probability weighting which can become arbitrarily large.

As for the third type of alternative dWOLS weight, w_{3i}, I also suspect that this will lead to a smaller weight variance than the inverse probability weights. Like IPW, it may be unbounded for subjects who do not receive treatment (subjects with $a = 0$) but will be bounded for all treated subjects. Likewise for w_{a_i}.

Our hypothesis is that the weights of dWOLS could be used in value search estimation of individualized treatment regimes. In this thesis, in order to explore the development of ITR estimation methods, I will propose some new estimators that are defined by combining the weighting schemes in dWOLS with the value search approaches together to estimate $E[Y(d_\eta)]$. The new estimators are called simple (s) and augmented (a) value search estimators (VSEs), i.e. s-VSE and a-VSE.

The formula for the simple estimators of the value function is written as

$$\hat{V}^d = \hat{E}[Y(d_\eta)] = n^{-1} \sum_{i=1}^{n} w_i y_i 1_{\{a_i = d\}}. \tag{3.1}$$

The formula for the augmented estimators of the value function is written as

$$\hat{V}^d = \hat{E}[Y(d_\eta)] = n^{-1} \sum_{i=1}^{n} [w_i(y_i - \mu(a_i, x_i; \beta))1_{\{a_i = d\}} + \mu(a_i, x_i; \beta)] \tag{3.2}$$

where d is a candidate regime indexed by η and $\mu(a_i, x_i; \beta)$ is a proposed parametric model for the mean outcome given treatment and covariates.

In this chapter, I will prove the consistency of the new estimators, demonstrate their performance via simulation. Further, in Chapter 5, I apply these to National Health and Nutrition Examination Survey data.

3.2 Properties of the Newly Proposed Estimators

THEOREM 1*: The s-VSE as defined in (3.1) is consistent for a quantity proportional to the value function, yielding a consistent estimator for the optimal threshold,* η^{opt}.

Proof: The estimator is written as

$$\hat{E}[Y(d_\eta)] = n^{-1}\sum_{i=1}^{n} w_i y_i 1_{\{a_i=d\}}.$$

Then

$$\begin{aligned}
\hat{V}^d &= n^{-1}\sum_{i=1}^{n} w_i y_i 1_{\{a_i=d\}} \\
&\xrightarrow{P} E\{E[w_i Y_i 1_{\{a_i=d\}}|X_i]\} \\
&= E\{w_i E[Y_i(d)]\} \\
&= kV^d
\end{aligned}$$

where $k = E[w_i 1_{\{a_i=d\}}] \neq 1$, however this estimator will still be unbiased since k is a constant, thus

$$\arg\max_{d\in\mathcal{D}} kV^d = \arg\max_{d\in\mathcal{D}} V^d.$$

Note that if I use the inverse probability weights, I would have $k = 1$. So even if I obtain a biased estimator of V^d, I can still have an unbiased estimator of η^{opt}.

□

The dWOLS weights remove confounding bias and therefore are useful in performing regression on a weighted dataset. For a weighted data set (y_w, x_w, a_w), where weights are in the form of dWOLS weights, Wallace and Moodie (2015) showed that $x_w \perp a_w$ i.e. $E[X_w|A_w = 0] = E[X_w|A_w = 1]$, is achieved with dWOLS weights which satisfy

$$\frac{P(A_w = 0|X_w = x)}{P(A_w = 0)} = \frac{P(A_w = 1|X_w = x)}{P(A_w = 1)}.$$

THEOREM 2*: a-VSE as defined in (3.2) is consistent for η^{opt} only when inverse probability of treatment weights are used.*

Proof: I now prove the consistency of the a-VSEs. Suppose we have a-VSEs with inverse probability weights, that is the AIPWE. Now, if $\mu(A, X; \beta)$ is correctly specified, then

$$n^{-1} \sum_{i=1}^{n} [(y_i - \mu(a_i, x_i; \beta)] \xrightarrow{P} 0.$$

Thus the first augmented term equals 0 asymptotically and we have

$$n^{-1} \sum_{i=1}^{n} [w_i (y_i - \mu(a_i, x_i; \beta)) 1_{\{a_i = d\}} + \mu(a_i, x_i; \beta)] \rightarrow n^{-1} \sum_{i=1}^{n} \mu(a_i, x_i; \beta) \rightarrow V^d.$$

Hence if $\mu(A, X; \beta)$ is correctly specified, the AIPWE is a consistent estimator for $E[Y(d_\eta)]$.

Similarly, if $\pi(A = a|X; \alpha)$ is a parametric model for the propensity score which is correctly specified, we have

$$n^{-1} \sum_{i=1}^{n} w_i 1_{\{a_i = d\}} \quad = \quad n^{-1} \sum_{i=1}^{n} \left\{ \frac{1_{\{a_i = d = 0\}}}{1 - \pi(A = 1|X; \alpha)} + \frac{1_{\{a_i = d = 1\}}}{\pi(A = 1|X; \alpha)} \right\} \rightarrow 1$$

and so

$$\sum_{i=1}^{n} [-w_i \mu(a_i, x_i; \beta) 1_{\{a_i = d\}} + \mu(a_i, x_i; \beta)] \rightarrow 0.$$

We now have

$$n^{-1}\sum_{i=1}^{n}[w_i(y_i-\mu(a_i,x_i;\beta))1_{\{a_i=d\}}+\mu(a_i,x_i;\beta)]$$
$$=n^{-1}[\sum_{i=1}^{n}w_iy_i-\sum_{i=1}^{n}(w_i\mu(a_i,x_i;\beta)1_{\{a_i=d\}}+\mu(a_i,x_i;\beta))]\to V^d.$$

Hence, the AIPWE is doubly robust and therefore consistent if $\mu(A,X;\beta)$ or $\pi(A|X;\alpha)$ is correctly specified.

Note if we substitute "absolute value weights" and "alternative dWOLS weights" to the above value function of a-VSEs, then

$$n^{-1}\sum_{i=1}^{n}w_i1_{\{a_i=d\}}\nrightarrow 1.$$

and so

$$\sum_{i=1}^{n}[-w_i\mu(a_i,x_i;\beta)1_{\{a_i=d\}}+\mu(a_i,x_i;\beta)]\nrightarrow 0$$

Hence, it turns out that the AIPWE with dWOLS weights does not lead to an unbiased estimator of V^d, however, a-VSEs is consistent provided the outcome model is correctly specified.

□

3.3 Implementation

3.3.1 Parametric Approach

The procedures to construct the new estimators parametrically are as follows:

1. Identify a set of candidate thresholds that we are interested in searching over, denote the set of thresholds as Θ. In this thesis, I will consider regimes of the form "treat when $x>\eta$", and Θ contains the list of all candidate thresholds η.

Table 3–1: Hypothetical dataset

subject	X	Y	A
1	-1.2	-0.5	1
2	0.3	1.5	0
3	-0.8	-0.4	1

Table 3–2: Augmented dataset corresponding to the hypothetical data in Table 3-1

subject	X	Y	A	η
1	-1.2	-0.5	1	-2
1	-1.2	-0.5	1	-1.4
2	0.3	1.5	0	0.4
3	-0.8	-0.4	1	-2
3	-0.8	-0.4	1	-1.4

2. Propose a model $\pi(A|X;\alpha)$ for the propensity score $\pi(A|X)$ and fit the model to estimate the probability of being treated for each subject at each interval (e.g fitting a logistic regression model of A on X).
3. Generate a full augmented dataset by $\eta \in \Theta$. For example, if we have a hypothetical dataset (Table 3-1) and I define the thresholds $\eta \in \{-2, -1.4, 0.4\}$, then the corresponding augmented dataset is shown as Table 3-2.
4. Construct weights by transforming $\pi(A|X;\alpha)$ according to the desired form, e.g. inverse probability of treatment, absolute value weights or alternative dWOLS weights.
5. Propose a model for the mean outcome as a function of the threshold η, and possibly also baseline covariates. It is typical to choose a parameterization for dependence of the outcome on η that is quadratic, allowing for the possibility that the mean outcome is maximized as a function of η at an interior point of the set Θ.

6. Estimate the parameters of the regression model proposed in step (5) using a weighted linear regression.
7. Obtain η^{opt} as the value of η that maximizes the expected outcome.

3.3.2 Non-parametric Approach

The procedures to construct the new estimators non-parametrically are as follows:

1. Begin using steps 1-2 above.
2. We define the thresholds $\eta \in \Theta$, and keep rows of data such that "$X > \eta$ and $A = 1$" and "$X < \eta$ and $A = 0$".
3. Construct weights by transforming $\pi(A|X;\alpha)$ according to the desired form, e.g. inverse probability of treatment, absolute value weights or alternative dWOLS weights.
4. For each candidate threshold, compute the s-VSR or a-VSE according to equation (3.1) or (3.2), respectively.
5. Use the maximum s-VSE and a-VSE estimates to obtain an estimate of the optimal regime threshold, η^{opt}.

3.4 Summary

In this chapter, I have outlined a new class of value search estimators for the optimal ITR derived by incorporating weights from the regression-based dWOLS approach into a value search estimator. I proved that the singly robust estimators yield a consistent estimator of the optimal regime parameters, derived from a (not necessarily unbiased) estimator of the value function. I also showed that the augmented version of the estimator does not possess the property of double-robustness

except in the special case where weighting is by the inverse probability of treatment. I then provided a step-by-step algorithm for implementing both the s- and a-VSEs. In the next chapter, the performance of the estimators is explored.

CHAPTER 4
Simulations

In the previous chapter, I proposed the use of dWOLS weights in combination with value search estimation. I now investigate the performance of several implementations of the s-VSE and a-VSE. In particular, I examine the following seven estimators:

1. s-VSE with IPWs, i.e. $s - VSE_{IPW}(\eta)$
2. s-VSE with absolute value weights, i.e. $s - VSE_{abs}(\eta)$
3. s-VSE with first alternative dWOLS weights (w_{3i}), i.e. $s - VSE_{alt}(\eta)$
4. s-VSE with no weights (an incorrectly specified propensity score model), i.e. $s - VSE_{nowt}(\eta)$
5. a-VSE with IPWs, i.e. $a - VSE_{IPW}(\eta)$
6. a-VSE with absolute value weights, i.e. $a - VSE_{abs}(\eta)$
7. a-VSE with first alternative dWOLS weights (w_{3i}), i.e. $a - VSE_{alt}(\eta)$

As proven in the last chapter, the first three of these estimators are consistent provided the propensity score is correctly specified. The fifth is doubly robust, and hence consistent provided either the propensity score or the outcome regression model is correctly specified, whereas the final two estimators will only be consistent with the outcome regression model is correctly specified, which is typically difficult to achieve.

I now demonstrate procedures and simulation results of both parametric and non-parametric approaches to constructing the above proposed estimators.

4.1 Parametric Estimation

I now illustrate and evaluate the performance of all seven estimators by simulation studies. I use a single-stage and one tailoring variable simulated example with data generated for each subject as follows:

- Covariate $X_i \sim U(0,1)$
- Treatment $P(A_i = 1|X_i = x_i) = x_i$
- Outcome $Y_i \sim N(e^{X_i} - X_i^3 + A_i(-1 + 4X_i), 1)$.

I refer this is the first data generating scenario. It is straightforward to obtain the true optimal treatment regime $d^{opt}(\mathrm{X})=1_{\{-1+4X>0\}}=1_{\{X>\frac{1}{4}\}}$.

I refer the following data generation scheme as the second data generating scenario:

- Covariates $X_i \sim N(1,2)$
- Treatment $logit[P(A_i = 1|X_i = x_i)] = \frac{e^{-0.5+0.6X}}{1+e^{-0.5+0.6X}}$
- Outcome $Y_i \sim N(e^{X_i} - X_i^3 + A_i(-1 + 4X_i), 1)$.

It is still the truth that the optimal regime $d^{opt}(\mathrm{X})=1_{\{-1+4X>0\}}=1_{\{X>\frac{1}{4}\}}$. I now obtain $E[Y(d^{opt})] = 12.07$, and η^{opt}=0.25.

I define the η's as a sequence of values in the interval $[0,1]$ for the parametric approach.

For the parametric approach, I generate the full augmented dataset using the following the criteria:

1. if $x > \eta$ and $a = 1$, or
2. if $x < \eta$ and $a = 0$.

I perform the one-dimensional searching for different η values from the interval $[0, 1]$ in order to add η as an extra variable to the fully augmented dataset. Now, based on this fully augmented data set, I use the correctly specified propensity score $\pi(X_i) = x_i$ (I use $\pi(X_i) = 1 - x_i$ as the incorrect treatment model) and perform the regression method with the proposed outcome model: $Y \sim 1 + \eta + \eta^2$ with the IPWs and other weights. Then I extract the coefficients from the outcome regression model and treat them as the coefficients of a quadratic function and maximize this function to get the optimal η as $\hat{\eta^{opt}}$. For those quadratic terms that have negative coefficient values, the function is concave and thus, the optimal η value is the one corresponding to the maximum function value. For those quadratic terms that have positive coefficients, I examined which boundary η value gave the maximum value of this function (in this case it would be either 0 or 1). The true value of η is 0.25 as we discussed previously. Under the data-generating mechanism, I obtain $E[Y(d^{opt})] = 2.37$. Then I use the seven estimators I proposed earlier to do the following analysis:

- **Analysis 1** s-VSEs with IPWs: treatment model correct

  $w(a_i, x_i) = \frac{1}{P(A_i=a_i|\mathbf{X_i}=\mathbf{x_i})}$

  Regress y on (1,η,η^2) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 2** s-VSEs with absolute value weights: treatment model correct

  $w(a_i, x_i) = |a_i - P(A_i = 1|\mathbf{X_i} = \mathbf{x_i})|$

  Regress y on (1,η,η^2) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 3** s-VSEs with dWLOS weights: treatment model correct

  $w(a_i, x_i) = 1_{a_i=1} + 1_{a_i=0}\frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$

  Regress y on $(1,\eta,\eta^2)$ with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 4** s-VSEs with no weights: treatment model incorrect

  $w(a_i, x_i) = \frac{1}{P(A_i=a_i|\mathbf{X_i}=\mathbf{x_i})}$

  Regress y on $(1,\eta,\eta^2)$ with $P[A_i = 1|X_i = x_i] = 1 - x_i$

- **Analysis 5** a-VSEs with IPWs: treatment model correct and outcome model incorrect

  $w(a_i, x_i) = \frac{1}{P(A_i=a_i|\mathbf{X_i}=\mathbf{x_i})}$

  Regress y on $(1,\eta,\eta^2)$ with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 6** a-VSEs with absolute value weights: treatment model correct and outcome model incorrect

  $w(a_i, x_i) = |a_i - P(A_i = 1|\mathbf{X_i} = \mathbf{x_i})|$

  Regress y on $(1,\eta,\eta^2)$ with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 7** a-VSEs with alternative dWLOS weights: treatment model correct and outcome model incorrect

  $w(a_i, x_i) = 1_{a_i=1} + 1_{a_i=0}\frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$

  Regress y on $(1,\eta,\eta^2)$ with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 8** a-VSEs with alternative dWOLS weights: treatment model incorrect and outcome model correct

  $w(a_i, x_i) = 1_{a_i=1} + 1_{a_i=0}\frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$

  Regress y on $(1,\eta,\eta^2)$ with $P[A_i = 1|X_i = x_i] = 1 - x_i$

I generate the boxplots in Figures 4-1 to 4-3 of $\hat{\eta^{opt}}$ estimates for our estimators and the results of 1000 simulated datasets of size 100, 200, 500. Dashed lines indicate true optimal threshold. As I expected from our theory, $s-VSE_{IPW}$, $s-VSE_{abs}$, and $s-VSE_{alt}$ appear to be consistent if our treatment model is correctly specified. For the single-robust estimator without weight, the boxplot shows lack of consistency since the estimates deviate away from the true η, 0.25.

The performance of $a-VSE_{IPW}$ appears consistent and the variance of this estimator seems to be smaller comparing to all three $s-VSEs$. The $a-VSE_{abs}$ and $a-VSE_{alt}$ are appeared to be biased for true $\eta^{opt}=0.25$, except for Analysis 8, in which the outcome model is correctly specified. The $a-VSE_{alt}$ also is consistent in the setting where the treatment model is incorrectly specified and the outcome regression model is correctly specified.

An out of sample estimate of the performance of the estimated ITR was found by applying the estimated ITR to a new sample of size 100,000 individuals; the estimated values of V^d and its standard errors are presented in Table 4-1 to Table 4-4.

I present in Table 4-1 the estimated values of $E[Y(\hat{d}^{opt})]$ for the above eight analysis. This indicates how well each approach performs in terms of optimizing the outcome under the estimated optimal regime. I obtain similar results as the regression approach estimators. The unbiased $s-VSEs$ and $a-VSEs$ have better performance. Analyses 6 and 7 produce worse prediction for the average outcomes by using the inconsistent estimates of $\hat{d}^{opt}=1_{\{x>\hat{\eta}^{opt}\}}$.

Figure 4–1: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 100 using parametric s- and a-VSEs under the first data-generating scenario.

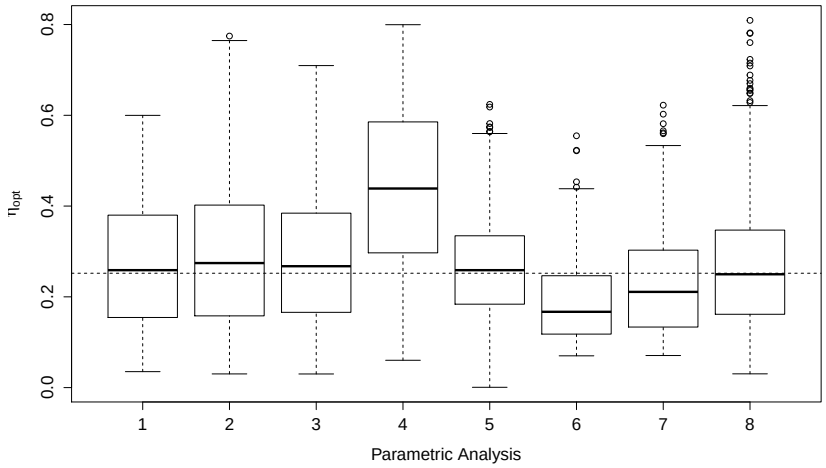


Figure 4–2: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 200 using parametric s- and a-VSEs under the first data-generating scenario.

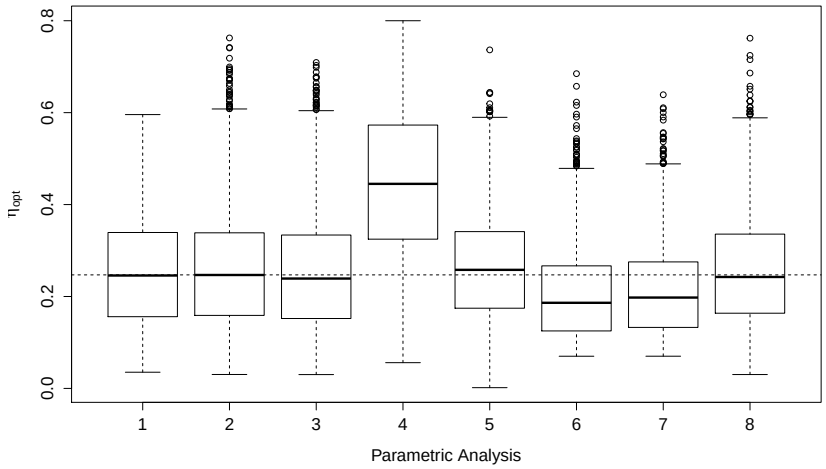
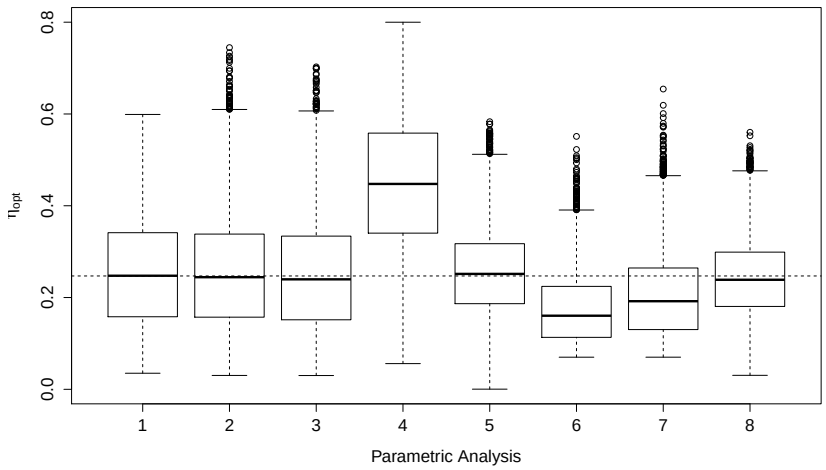


Table 4–1: The mean outcome estimates under the estimated optimal rule for the first data generation scenario via parametric approach.

Subjects	Analysis	Estimate of $V^{\hat{d}}$	SE of $V^{\hat{d}}$	Bias of $\hat{\eta}^{opt}$	SE of $\hat{\eta}^{opt}$
n=100	**Analysis 1**	2.319	0.010	0.007	0.097
	Analysis 2	2.365	0.015	0.015	0.071
	Analysis 3	2.441	0.012	0.012	0.081
	Analysis 4	2.142	0.198	0.213	0.191
	Analysis 5	2.167	0.076	0.002	0.073
	Analysis 6	2.081	0.238	0.124	0.113
	Analysis 7	2.023	0.201	0.105	0.114
	Analysis 8	2.241	0.091	0.034	0.081
n=200	**Analysis 1**	2.451	0.010	0.005	0.069
	Analysis 2	2.315	0.014	0.031	0.071
	Analysis 3	2.194	0.013	0.030	0.064
	Analysis 4	2.112	0.078	0.204	0.145
	Analysis 5	2.270	0.033	0.075	0.051
	Analysis 6	2.021	0.212	0.101	0.101
	Analysis 7	1.916	0.178	0.107	0.089
	Analysis 8	2.263	0.076	0.022	0.054
n=500	**Analysis 1**	2.291	0.007	0.003	0.067
	Analysis 2	2.515	0.015	0.035	0.052
	Analysis 3	2.357	0.013	0.031	0.051
	Analysis 4	2.112	0.178	0.203	0.091
	Analysis 5	2.401	0.024	0.078	0.041
	Analysis 6	2.011	0.106	0.112	0.109
	Analysis 7	2.031	0.108	0.127	0.092
	Analysis 8	2.351	0.061	0.012	0.036

Figure 4–3: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 500 using parametric s- and a-VSEs under the first data-generating scenario.



I applied our estimators to another mechanism of data generation as well

- Covariates $X_i \sim N(1, 2)$
- Treatment $logit[P(A_i = 1|X_i = x_i)] = \frac{e^{-0.5+0.6X}}{1+e^{-0.5+0.6X}}$
- Outcome $Y_i \sim N(e^{X_i} - X_i^3 + A_i(-1 + 4X_i), 1)$.

It is still the truth that the optimal regime d^{opt}(X)=$I(-1 + 4X > 0)$=$I(X > \frac{1}{4})$. I now obtain $E[Y(d^{opt})] = 12.07$, and η^{opt}=0.25.

The estimates of η^{opt} as shown in Figures 4-4 to 4-6 based on the second data generating scenario, and lead to similar conclusion as our last simulation.

I present in Table 4-2 the estimated values of $E[Y(\hat{d}^{opt})]$ for the above eight analysis. As in the first set of simulations, estimators perform as expected: all s-VSEs are consistent when treatment is modelled correctly. Of the augmented estimators,

Figure 4–4: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 100 using parametric s- and a-VSEs under the second data-generating scenario.

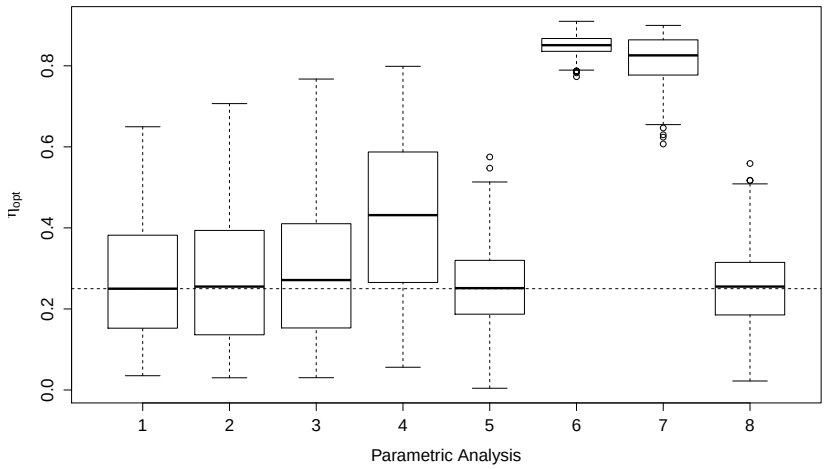


Figure 4–5: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 200 using parametric s- and a-VSEs under the second data-generating scenario.

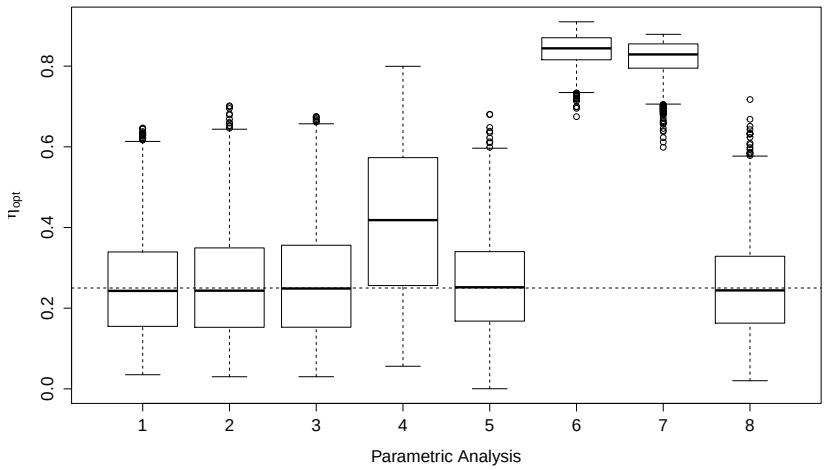
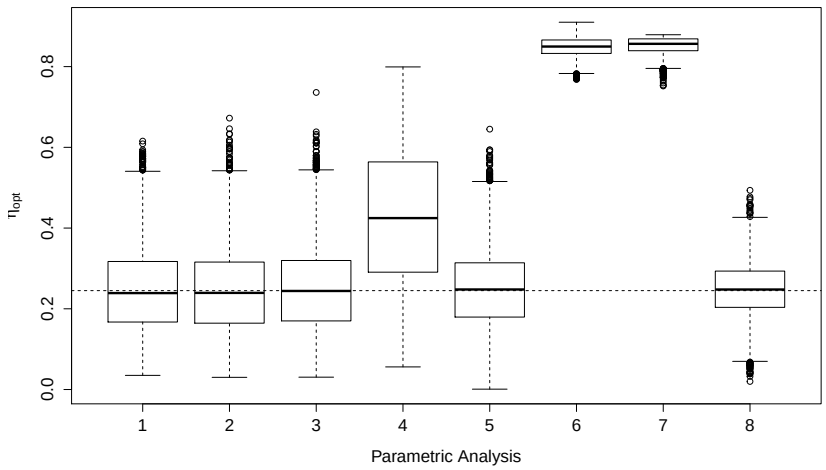


Figure 4–6: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 500 using parametric s- and a-VSEs under the second data-generating scenario.



a-VSE with a correct outcome model (for any weighting scheme) is consistent, and only a-VSE with IPW is doubly robust.

All consistent estimators provide decision thresholds that lead to population outcomes that are closer to the true optimal outcome (see Table 4-2) as compared to the first data generating scenario (Table 4-1).

Table 4–2: The mean outcome estimates under the estimated optimal rule for the second data generation scenario via parametric approach

Subjects	Analysis	Estimate of $V^{\hat{d}}$	SE of $V^{\hat{d}}$	Bias of $\hat{\eta}^{opt}$	SE of $\hat{\eta}^{opt}$
n=100	**Analysis 1**	12.325	19.282	0.017	0.097
	Analysis 2	12.165	9.662	0.013	0.055
	Analysis 3	12.314	9.969	0.021	0.071
	Analysis 4	11.027	33.017	0.171	0.195
	Analysis 5	12.121	11.048	0.009	0.071
	Analysis 6	11.228	10.782	0.624	0.183
	Analysis 7	11.158	15.698	0.605	0.124
	Analysis 8	12.231	12.993	0.034	0.081
n=200	**Analysis 1**	12.215	11.104	0.025	0.063
	Analysis 2	12.273	9.361	0.018	0.073
	Analysis 3	12.071	8.153	0.078	0.061
	Analysis 4	11.671	23.091	0.181	0.145
	Analysis 5	12.157	10.030	0.005	0.052
	Analysis 6	11.727	9.126	0.626	0.145
	Analysis 7	11.411	11.324	0.602	0.190
	Analysis 8	12.169	10.015	0.012	0.054
n=500	**Analysis 1**	12.273	9.655	0.013	0.022
	Analysis 2	12.165	8.818	0.035	0.052
	Analysis 3	12.375	7.111	0.031	0.050
	Analysis 4	11.213	15.641	0.173	0.093
	Analysis 5	12.126	9.177	0.018	0.045
	Analysis 6	11.171	10.711	0.628	0.109
	Analysis 7	11.312	10.172	0.637	0.102
	Analysis 8	12.194	7.610	0.017	0.042

4.2 Non-parametric Estimation

For the non-parametric approach, the simulated data and augmented data set generation was the same as in Section 4.1. I define the η's as a sequence of values in the interval $[0, 1]$ by increments of 0.01 for the non-parametric approach.

In order to investigate these weights, I now apply non-parametric VSE by using three different weights satisfying equation (4.1) with the same augmented dataset generation method as the first approach. Note that here, unlike previously, I directly construct seven estimators: $s - VSE_{IPW}$, $s - VSE_{abs}$,$s - VSE_{alt}$, $s - VSE_{nowt}$ $a - VSE_{IPW}$, $s - VSE_{abs}$, $s - VSE_{alt}$ and all the weights are of the form I described earlier in the literature review Section 2.2.3, which involves proposing models for both $\mu(A, X; \beta)$ and the propensity scores $\pi(X)$. I use both correctly and incorrectly specified propensity score models. For sample sizes 100, 200 and 500, I generate 1000 simulated datasets according to the first and second data generating scenario and conduct in total eight analyses:

- **Analysis 1** s-VSEs with IPWs: treatment model correct

  $w(a_i, x_i) = \frac{1}{P(A_i=a_i|\mathbf{X_i}=\mathbf{x_i})}$

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 2** s-VSEs with absolute value weights: treatment model correct

  $w(a_i, x_i) = |a_i - P(A_i = 1|\mathbf{X_i} = \mathbf{x_i})|$

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 3** s-VSEs with alternative dWLOS weights: treatment model correct

  $w(a_i, x_i) = 1_{a_i=1} + 1_{a_i=0}\frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 4** s-VSEs with no weights: treatment model incorrect,

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = 1 - x_i$

- **Analysis 5** a-VSEs with IPWs: treatment model correct and outcome model incorrect, $w(a_i, x_i) = \frac{1}{P(A_i=a_i|\mathbf{X_i}=\mathbf{x_i})}$

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 6** a-VSEs with absolute value weights: treatment model correct and outcome model incorrect, $w(a_i, x_i) = |a_i - P(A_i = 1|\mathbf{X_i} = \mathbf{x_i})|$

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 7** a-VSEs with alternative dWOLS weights: treatment model correct and outcome model incorrect, $w(a_i, x_i) = 1_{a_i=1} + 1_{a_i=0}\frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$

  Regress y on (1,a,x,ax) with $P[A_i = 1|X_i = x_i] = x_i$

- **Analysis 8** a-VSEs with alternative dWOLS weights: treatment model incorrect and outcome model correct, $w(a_i, x_i) = 1_{a_i=1} + 1_{a_i=0}\frac{P(A_i=1|\mathbf{X_i})}{1-P(A_i=1|\mathbf{X_i})}$

  Regress y on (1,a,x,e^x,x^3,ax) with $P[A_i = 1|X_i = x_i] = x_i + x_i^2$

I generated the boxplots (Figures 4-7 to 4-9) of estimated η^{opt} values for the eight analyses for 1000 simulated datasets for all $s - VSEs$ and $a - VSEs$ with the three different weight schemes and both correct and incorrect propensity score models. Dashed lines indicate true η^{opt}, which is 0.25. Results are as expected, and consistent with the parametric estimators.

Figure 4–7: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 100 using parametric s- and a-VSEs under the first data-generating scenario.

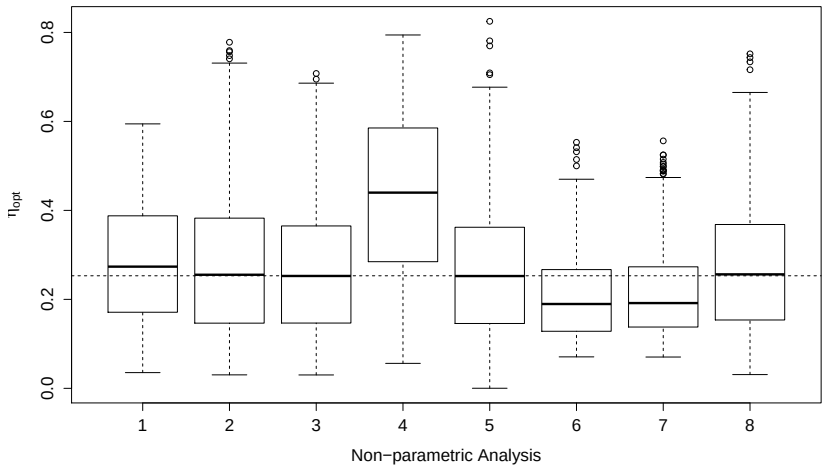


Figure 4–8: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 200 using parametric s- and a-VSEs under the first data-generating scenario.

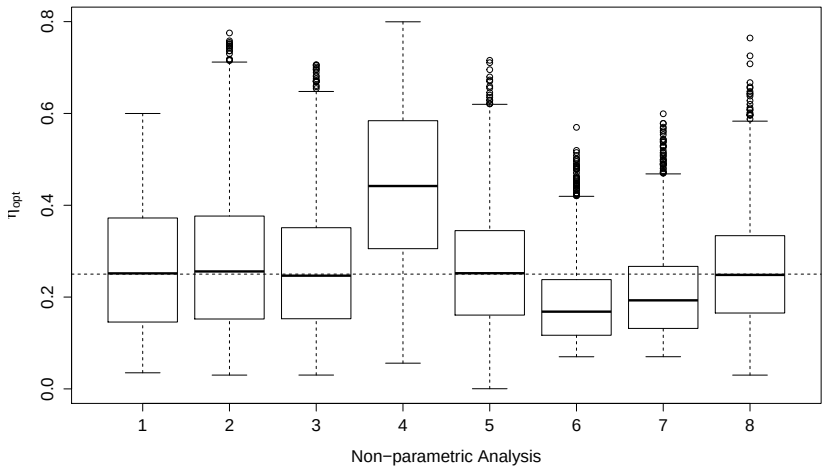
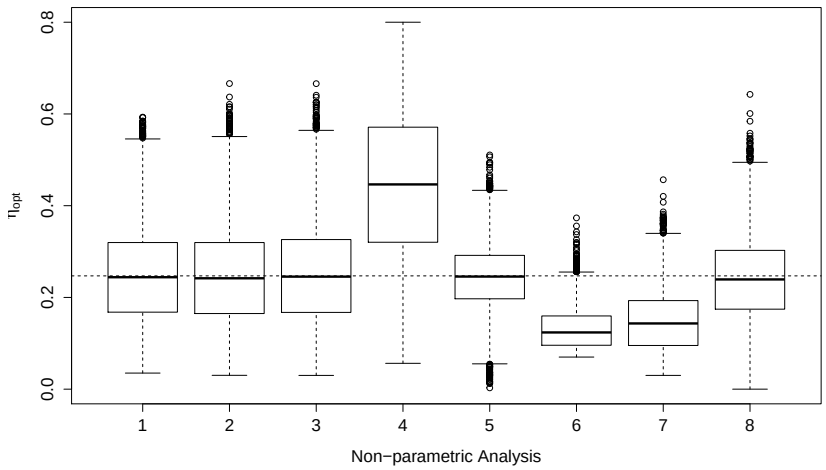


Figure 4–9: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 500 using parametric s- and a-VSEs under the first data-generating scenario.



I found that in the single-stage setting, $s-VSEs$ with all three weight schemes provides unbiased and almost equally larger variances compared to the original $a-VSE_{IPW}$. The $a-VSE_{abs}$ and $a-VSE_{alt}$ seem to return smaller variances than both of the $s-VSEs$ and $a-VSE_{IPW}$, however, these two estimators where biased. The variance relationship between $a-VSE_{abs}$ and $a-VSE_{alt}$ is not very clear and thus, further investigation may be needed in order to combine these new weights with other estimators that could enjoy consistency properties.

I present in Table 4-3 the estimates of $E[Y(\hat{d}^{opt})]$ for the above eight analysis. Again, I obtain similar results as the parametric VSEs. Except for Analysis 4, 6, and 7, the $s-VSEs$ and $a-VSEs$ have similar performance as measured by the expected outcome in population treated under the estimated optimal ITR. Analyses

6 and 7 produce worse prediction for the average outcomes by using the inconsistent estimates of $\hat{d}^{opt} = 1_{\{x>\hat{\eta}^{opt}\}}$.

Table 4–3: The mean outcome estimates under the estimated optimal rule for the first data generation scenario via non-parametric approach

Subjects	Analysis	Estimate of $V^{\hat{d}}$	SE of $V^{\hat{d}}$	Bias of $\hat{\eta}^{opt}$	SE of $\hat{\eta}^{opt}$
n=100	**Analysis 1**	2.251	0.031	0.007	0.096
	Analysis 2	2.352	0.022	0.015	0.061
	Analysis 3	2.264	0.027	0.012	0.065
	Analysis 4	2.133	0.198	0.203	0.191
	Analysis 5	2.391	0.076	0.002	0.076
	Analysis 6	1.951	0.238	0.124	0.113
	Analysis 7	2.051	0.201	0.105	0.114
	Analysis 8	2.464	0.110	0.034	0.081
n=200	**Analysis 1**	2.251	0.027	0.005	0.099
	Analysis 2	2.435	0.021	0.031	0.071
	Analysis 3	2.561	0.023	0.030	0.064
	Analysis 4	2.131	0.178	0.206	0.145
	Analysis 5	2.353	0.033	0.071	0.089
	Analysis 6	2.017	0.212	0.100	0.106
	Analysis 7	2.041	0.178	0.105	0.082
	Analysis 8	2.363	0.103	0.023	0.054
n=500	**Analysis 1**	2.313	0.019	0.003	0.057
	Analysis 2	2.500	0.013	0.034	0.042
	Analysis 3	2.317	0.015	0.033	0.061
	Analysis 4	2.011	0.178	0.202	0.069
	Analysis 5	2.313	0.024	0.078	0.043
	Analysis 6	2.015	0.106	0.112	0.467
	Analysis 7	1.921	0.108	0.125	0.061
	Analysis 8	2.371	0.061	0.011	0.031

The estimates of η^{opt} for second data generation scheme is shown in Figures 4-10 to 4-12 and similar conclusion can be reached as in the parametric modeling simulation. Dashed lines indicate true η^{opt}, which is 0.25.

Figure 4–10: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 100 using parametric s- and a-VSEs under the second data-generating scenario.

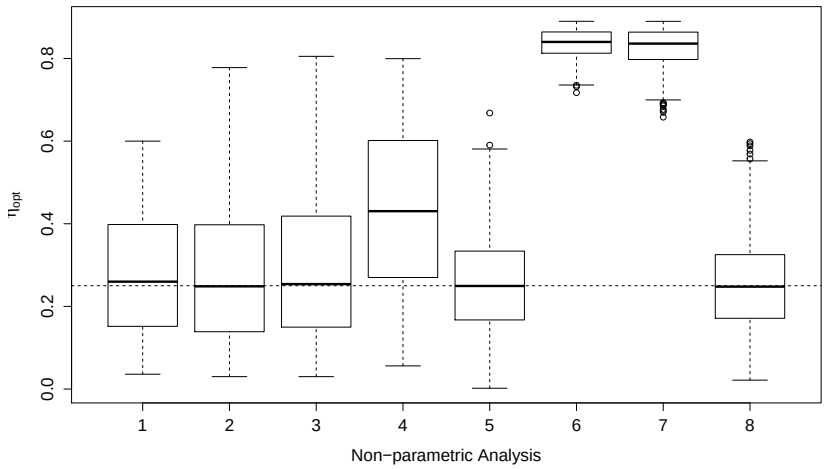


Figure 4–11: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 200 using parametric s- and a-VSEs under the second data-generating scenario.

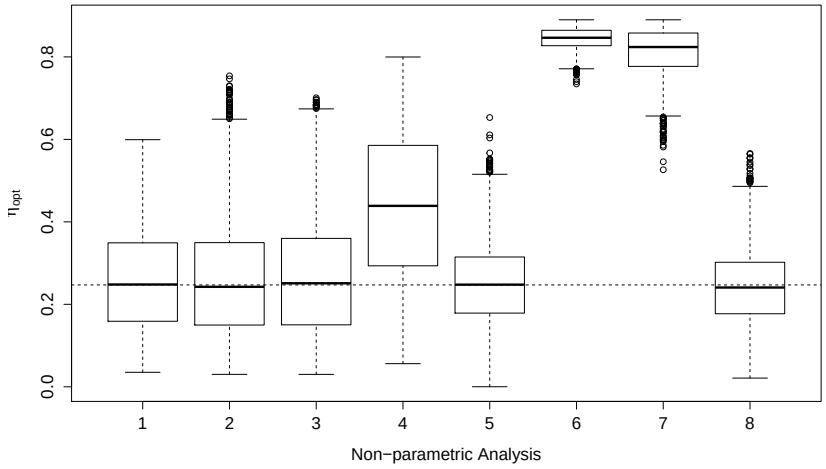
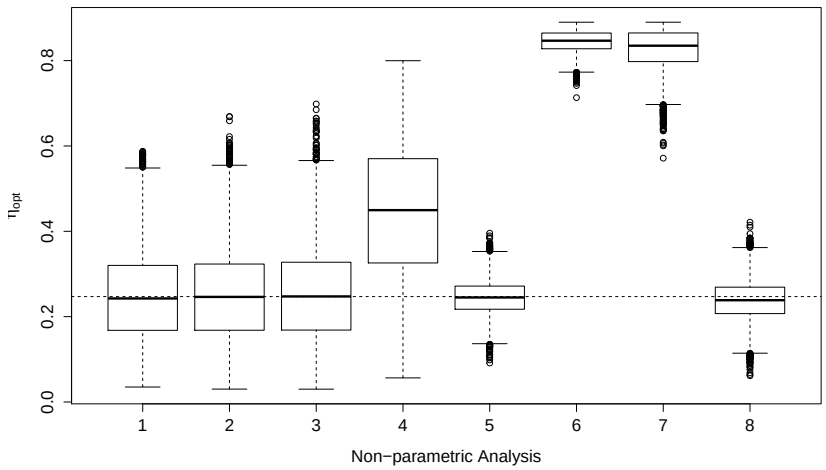


Figure 4–12: Distribution of the estimated optimal threshold across 1000 simulated datasets of size 500 using parametric s- and a-VSEs under the second data-generating scenario.



I present in Table 4-4 the estimates of $E[Y(\hat{d}^{opt})]$ for the above eight analyses under the second data generation scenario. As before, the unbiased $s-VSEs$ and $a-VSEs$ have similar performance as measured by the expected outcome in a new population treated under the estimated optimal ITR. Analyses 6 and 7 produce worse prediction for the average outcomes by using inconsistent estimates of $\hat{d}^{opt} = 1_{\{x>\hat{\eta}^{opt}\}}$.

All the simulation results that I have shown for the non-parametric value search approach demonstrate great similarity to the results for the parametric value search estimators.

4.3 Summary

In this chapter, I performed simulations which demonstrated the consistency of the s-VSE approach under correct specification of the treatment model and consistency of the a-VSE approach under correct specification of the outcome model, as

Table 4–4: The mean outcome estimates under the estimated optimal rule for the second data generation scenario via non-parametric approach

Subjects	Analysis	Estimate of $V^{\hat{d}}$	SE of $V^{\hat{d}}$	Bias of $\hat{\eta}^{opt}$	SE of $\hat{\eta}^{opt}$
n=100	**Analysis 1**	12.375	17.122	0.017	0.087
	Analysis 2	12.053	10.212	0.024	0.053
	Analysis 3	12.154	9.661	0.021	0.072
	Analysis 4	11.052	29.217	0.091	0.196
	Analysis 5	12.081	10.781	0.009	0.076
	Analysis 6	11.122	11.698	0.638	0.193
	Analysis 7	11.151	12.698	0.615	0.124
	Analysis 8	12.231	12.112	0.034	0.088
n=200	**Analysis 1**	12.175	10.114	0.005	0.063
	Analysis 2	12.132	9.251	0.021	0.071
	Analysis 3	12.371	8.850	0.018	0.064
	Analysis 4	11.022	20.280	0.184	0.145
	Analysis 5	12.313	11.371	0.075	0.051
	Analysis 6	11.237	10.819	0.655	0.151
	Analysis 7	11.415	10.181	0.637	0.189
	Analysis 8	12.135	11.821	0.022	0.054
n=500	**Analysis 1**	12.353	7.077	0.003	0.067
	Analysis 2	12.645	8.125	0.035	0.052
	Analysis 3	12.615	8.127	0.031	0.050
	Analysis 4	11.813	13.271	0.173	0.093
	Analysis 5	12.076	8.041	0.078	0.045
	Analysis 6	11.723	7.914	0.612	0.109
	Analysis 7	11.321	9.101	0.627	0.102
	Analysis 8	12.102	8.093	0.012	0.041

well as double-robustness in the special case of a-VSE where the inverse probability of treatment weighting is used. As expected, the estimator of (a quantity proportional to) the value function was made more efficient by the use of absolute value weights; however this did not translate into smaller variability in the estimator of the optimal threshold. I am therefore left with the rather discouraging conclusion that improvements on the s-VSE that corresponds to the traditional IPWE cannot easily be made through a judicious choice of weights.

CHAPTER 5
Analysis of the National Health and Nutrition Examination Survey Data

5.1 Background

It is commonly recognized that exercise and physical activity have wide-ranging health benefits, from decreased weight to improved mood (Taylor et al., 2004). Increasingly, leisure time activities are more sedentary, with television watching, video games, and personal computing among the most popular pastimes. Further, people in industrialized countries are expending less energy in activities of daily living, and at work. A change in the volume of daily physical activity may account for this apparent discrepancy (United States. Department of Health, 1996).

The prevalence of overweight diagnoses continues to increase in the US adult population (Must et al., 1999). Mokdad et al. (2003) state that overweight and obesity diagnoses were significantly associated with diabetes, high blood pressure, high cholesterol, asthma, arthritis, and poor health status. Obesity and diabetes are major causes of morbidity and mortality in the United States (Must et al., 1999) as well as throughout the world. Evidence from several studies indicates that obesity and weight gain are associated with an increased risk of diabetes (Ford et al., 1997).

Hypertension affects 65 million adult Americans and is a major risk factor for myocardial infarction, stroke, heart failure, and renal failure. The prevalence of hypertension has not declined although the improvement in blood pressure control

is encouraging. The control of blood pressure is crucial in the prevention of these adverse outcomes (Ong et al., 2007).

It has been suggested that even adults who are unwilling or unable to meet recommended levels of exercise will benefit from undertaking some physical activity (Garber et al., 2011). It may even be the case that there are psychological benefits to setting lower but achievable targets. I therefore set out to investigate whether the amount of physical activity needed to optimize three different measures of health (body mass index (BMI), systolic blood pressure, and self-reported general health) should be tailored to individual characteristics. In particular, I investigated whether it would be advantageous to tailor physical activity to how sedentary a person is, as measured by the number of TV hours watched, or by self-reported level of depression. I apply unbiased s- and a-VSEs to data from the United States' National Health and Nutrition Examination Survey Data to examine this question.

5.2 Methods

5.2.1 The NHANES data

The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations (Centers for Disease Control and Prevention).

The NHANES data has multiple waves, and in each wave, the NHANES interview includes demographic, socioeconomic, dietary, and health-related questions.

The examination component consists of medical, dental, and physiological measurements, as well as laboratory tests administered by highly trained medical personnel (Centers for Disease Control and Prevention).

According to NHANES's description file, 'The NHANES data is collected on the prevalence of chronic conditions in the population. Estimates for previously undiagnosed conditions, as well as those known to and reported by respondents, are produced through the survey'.[1]

For my analysis, in order to investigate whether access to physical activity could be tailored to individual characteristics, namely how sedentary a person is, which is measured by the number of TV hours, and depression level. I apply the proposed new estimators on the Third National Health and Nutrition Examination Survey data (NHANES III), restricted to adults with age greater than or equal to 18 and excluding pregnant women.

5.2.2 Analyses

I begin by outlining the variables used for the analyses, then proceed to describe the estimators.

Outcome variables I considered three outcomes: BMI, self-reported general health, and systolic blood pressure. BMI is a tool used to measure weight. As a high BMI is indicative of greater weight (e.g., a BMI greater than or equal to 30 defines obesity, while a BMI less than 30 but at least 25 indicates overweight), I transformed the outcome, since ITR estimation methods are traditionally set up seek to maximize the

[1] `http://www.cdc.gov/nchs/nhanes/about_nhanes.htm`

outcome. I recoded the outcome BMI as $BMI = |max(BMI) - BMI|$. The second outcome that I considered was self-reported general health, which is measured on a 5 point scale (poor, fair, good, very good, or excellent); these categories were assigned values from 1 to 5 and the variable was treated as continuous. Finally, systolic blood pressure (SBP) was transformed as follows: $SBP = max(SBP) - SBP$. Note that SBP is measured as the average of three measurements of systolic blood pressure.

Treatment variable I consider physical activity as the treatment variable. Physical activity is being record as "Yes" if a participant does moderate or vigorous-intensity sports, fitness or recreational activities; if not, physical activity is recorded as "No". I recode "Yes" as 1 and "No" as 0 for convenience.

Tailoring variables I consider two tailoring variables: number of TV hours watched and self-reported level of depression. The hours of TV watched is measured by number of hours per day on average that a participant watched TV over the past 30 days and a 7 level scale (0 hours, 0 to 1 hour, 1 hour, 2 hours, 3 hours, 4 hours, More than 4 hours). These categories were assigned values -0.5, 0.5, 1, 2, 3, 4, and 5 and the variable was treated as continuous. The self-reported level of depression is measured by number of days where participant felt down, depressed or hopeless. The levels are none, several, majority (more than half the days), or almost all. I combined two levels (majority and almost all) together due to sparse numbers and recoded the 3 level as -0.5, 1, and 2.

Confounding variables Several potentia confounding variables were included in the analysis: age, gender, poverty, income, race, and smoking status. Poverty is a ratio of family income to poverty guidelines. Smaller numbers indicate more poverty.

Smoking status is measured by self-report of having smoked 100 or more cigarettes in their life time.

Estimators

1. s-VSE (IPWE) with IPWs, i.e. $s - VSE_{IPW}$
2. s-VSE with absolute value weights, i.e. $s - VSE_{abs}$
3. s-VSE with alternative dWOLS weights, i.e. $s - VSE_{alt}$
4. a-VSE with IPW (AIPWE), i.e. $a - VSE_{IPW}$
5. a-VSE with absolute value weights, i.e. $a - VSE_{abs}$
6. a-VSE with alternative dWOLS weights, i.e. $a - VSE_{alt}$

I used a logistic regression model to adjust for confounding, and a linear model for the mean outcome required for the augmented approaches. The non-parametric implementation of the estimators was used in all cases. Candidate regimes were "be physically active if hours of TV watched $> \eta_{TV}$" for $\eta_{TV} \in \{-0.05, 0.25, 0.75, 1.5, 2.5, 3.5, 4.2\}$ and "Be physically active if depression $> \eta_{dep}$" for $\eta_{dep} \in \{-0.1, 1.1, 1.6\}$.

Standard errors were computed by bootstrap resampling, taking 1000 samples with replacement. I also assessed the stability of the estimated treatment rule across resamples by calculating the frequency with which the optimal threshold in the original sample was optimal across the 1000 bootstraps.

5.3 Results

I obtained the estimates of η^{opt} and then the estimates of standard error of η^{opt} from 1000 bootstraps, and present these in Table 5-1. These regimes suggest that everyone should do moderate or vigorous-intensity sports, fitness or recreational activities, and there is no advantage to tailoring.

Results of fitting the mean model using linear models are presented in Tables 5-4, 5-5, and 5-6 and show that interactions are not statistically significant, which is supportive of the conclusion that the VSE approaches reached.

The estimates of $s - VSE_{IPW}(\eta)$ and $a - VSE_{AIPW}(\eta)$ for the value function $E(Y)$ and the corresponding standard errors are presented in Figures 5-1 to 5-6, and the optimal regimes η for each tailoring variable suggest there is no evidence of tailoring and everyone should do moderate or vigorous-intensity sports, fitness or recreational activities. From Table 5-2 and Table 5-3, the percentage of times

Table 5–1: Estimated optimal regimes $\hat{\eta}^{opt}$ and expected outcomes for Self-reported General Health, SBP and BMI for the tailoring variables TV hours watched and depression

Tailoring Variable: TVhours						
Outcome	BMI		Self-reported General Health		SBP	
Estimators	$E[Y]$	$\hat{\eta}^{opt}$	$E[Y]$	$\hat{\eta}^{opt}$	$E[Y]$	$\hat{\eta}^{opt}$
$s - VSE_{IPW}(\eta)$	-5.086	0.75	2.380	-0.05	94.175	0.52
$s - VSE_{w_1}(\eta)$	N/A	0.75	N/A	-0.04	N/A	1.11
$s - VSE_{w_3}(\eta)$	N/A	0.73	N/A	-0.05	N/A	0.29
$a - VSE_{IPW}(\eta)$	-5.085	0.03	2.355	-0.05	94.796	1.01
$a - VSE_{w_1}(\eta)$	N/A	-0.05	N/A	-0.04	N/A	-0.04
$a - VSE_{w_3}(\eta)$	N/A	-0.04	N/A	-0.04	N/A	0.16
Tailoring Variable: Depressed						
Outcome	BMI		Self-reported General Health		SBP	
Estimators	$E[Y]$	$\hat{\eta}^{opt}$	$E[Y]$	$\hat{\eta}^{opt}$	$E[Y]$	$\hat{\eta}^{opt}$
$s - VSE_{IPW}(\eta)$	-5.053	-0.10	2.095	-0.10	91.447	0.24
$s - VSE_{w_1}(\eta)$	N/A	-0.09	N/A	-0.10	N/A	-0.069
$s - VSE_{w_3}(\eta)$	N/A	-0.10	N/A	-0.09	N/A	0.29
$a - VSE_{AIPW}(\eta)$	-5.053	-0.10	2.114	-0.10	92.401	0.02
$a - VSE_{w_1}(\eta)$	N/A	-0.08	N/A	-0.10	N/A	0.01
$a - VSE_{w_3}(\eta)$	N/A	-0.10	N/A	-0.10	N/A	0.44

Figure 5–1: $E(Y_{BMI})$ vs. $\hat{\eta}_{TVhours}$ with 95% bootstrap confidence intervals

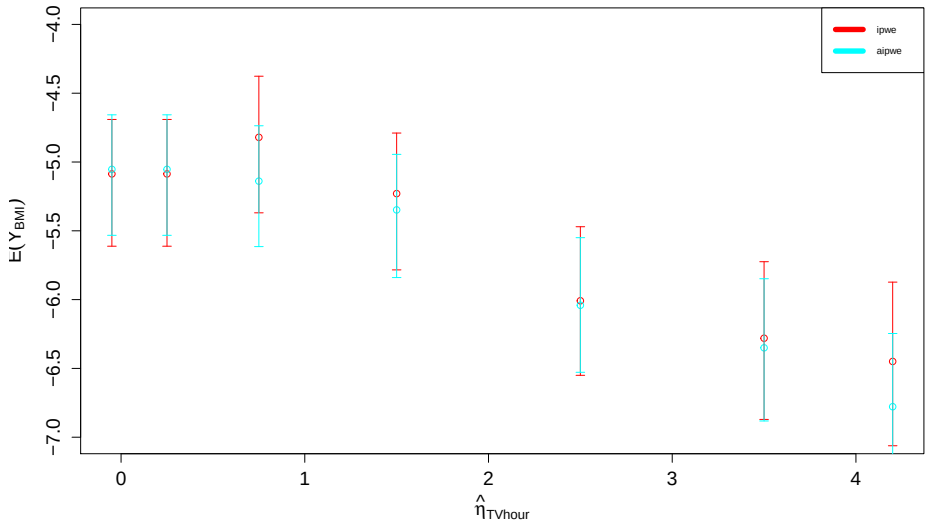


Figure 5–2: $E(Y_{BMI})$ vs. $\hat{\eta}_{Depressed}$ with 95% bootstrap confidence intervals

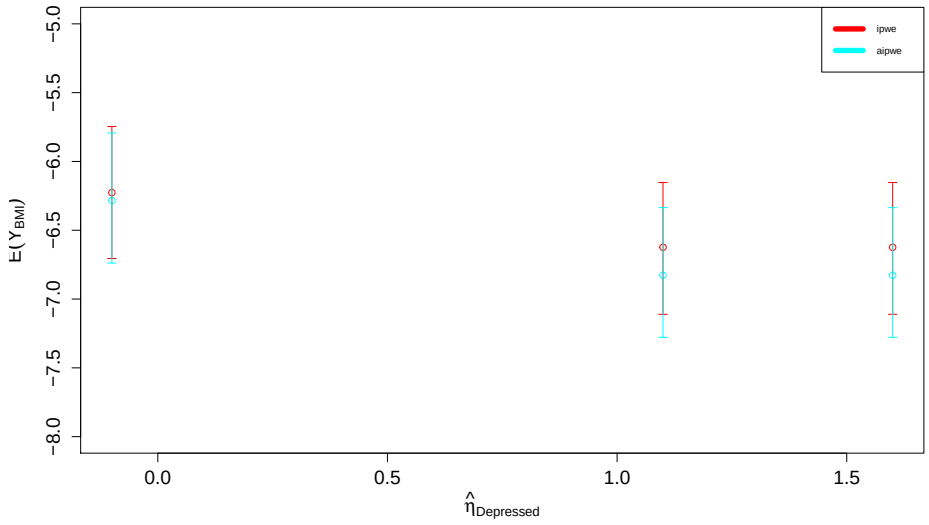


Figure 5–3: $E(Y_{\textit{Self-reported General Health}})$ vs. $\hat{\eta}_{TVhours}$ with 95% bootstrap confidence intervals

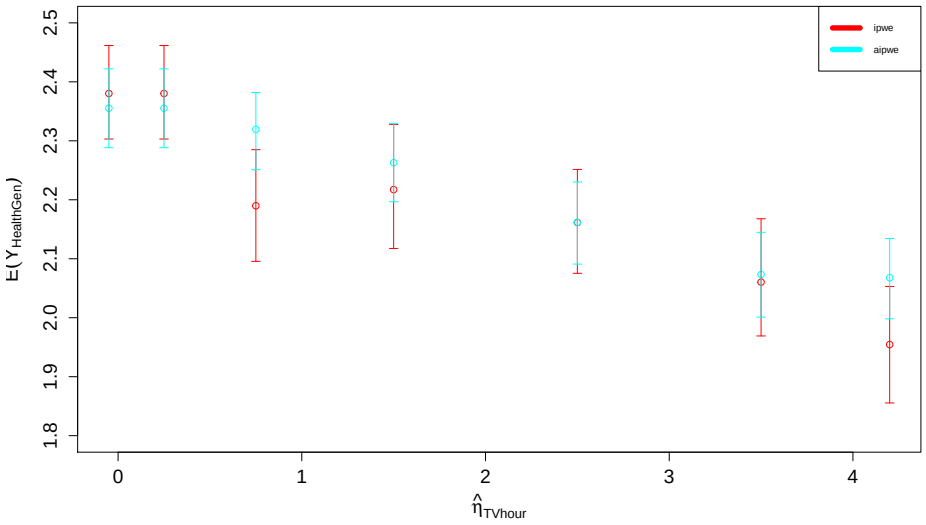


Figure 5–4: $E(Y_{\textit{Self-reported General Health}})$ vs. $\hat{\eta}_{Depressed}$ with 95% bootstrap confidence intervals

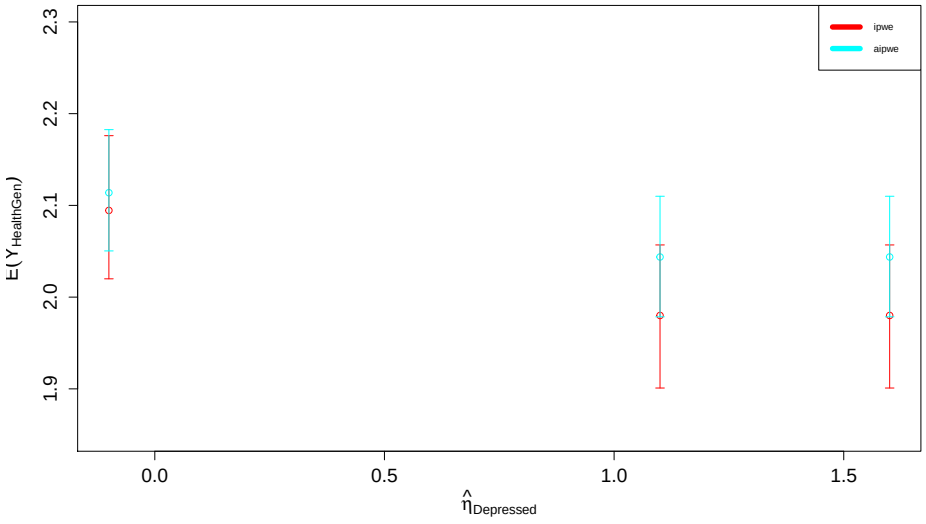


Figure 5–5: $E(Y_{SBP})$ vs. $\hat{\eta}_{TVhours}$ with 95% bootstrap confidence intervals

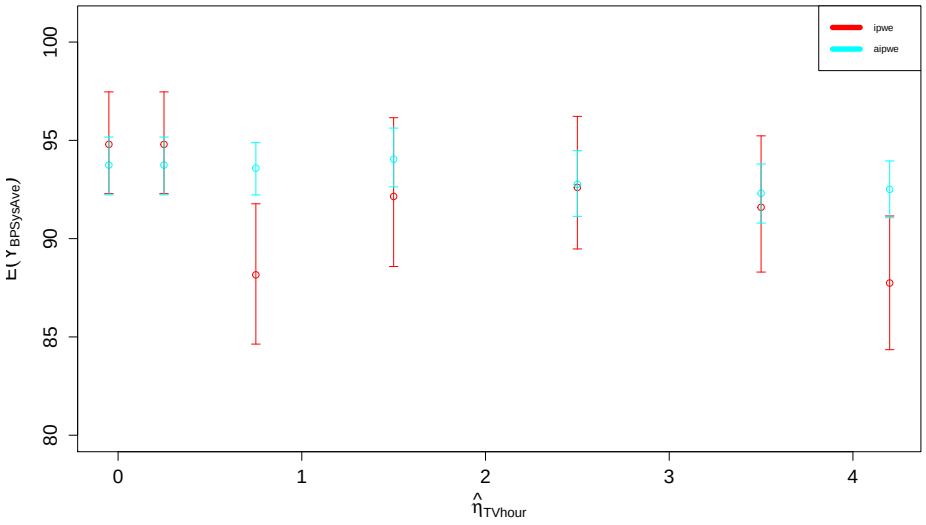
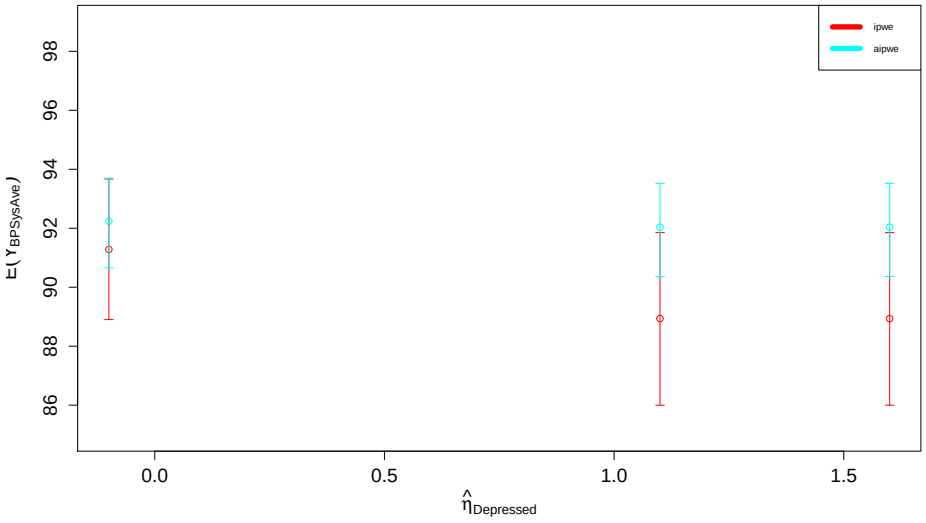


Figure 5–6: $E(Y_{SBP})$ vs. $\hat{\eta}_{Depressed}$ with 95% bootstrap confidence intervals



estimated optimal regimes η^{opt} chosen show the strong robustness of our regimes for the proposed $s-VSE_{IPW}$, $s-VSE_{w_1}$, $s-VSE_{w_3}$, and $a-VSE_{AIPW}$.

Table 5–2: Percentage of times estimated optimal regimes $\hat{\eta}^{opt}$ chosen over all treatment regimes and 95% CI of the expected outcomes for Self-reported General Health, SBP and BMI for the tailoring variable TV hours watched

Tailoring Variable: TVhour	Outcome: BMI		
Estimator	% times η^{opt} chosen	$E[Y]$	95% CI of $E[Y]$
$s-VSE_{IPW}$	0.007	-5.086	(-5.344,-4.369)
$s-VSE_{W_1}$	0.006	N/A	N/A
$s-VSE_{W_3}$	0.024	N/A	N/A
$a-VSE_{IPW}$	0.874	-5.085	(-5.494,-4.636)
	Outcome: Self-reported General Health		
Estimator	% times η^{opt} chosen	$E[Y]$	95% CI of $E[Y]$
$s-VSE_{IPW}$	1.000	2.380	(2.306,2.455)
$s-VSE_{w_1}$	0.996	N/A	N/A
$s-VSE_{w_3}$	0.999	N/A	N/A
$a-VSE_{IPW}$	0.996	2.355	(2.283,2.424)
	Outcome: SBP		
Estimator	% times η^{opt} chosen	$E[Y]$	95% CI of $E[Y]$
$s-VSE_{IPW}$	0.764	94.175	(92.611,97.738)
$s-VSE_{w_1}$	0.534	N/A	N/A
$s-VSE_{w_3}$	0.819	N/A	N/A
$a-VSE_{IPW}$	0.316	94.796	(92.813,95.545)

Table 5–3: Percentage times estimated optimal regimes $\hat{\eta}^{opt}$ chosen over all treatment regimes and 95% CI of the expected outcomes for Self-reported General Health, SBP and BMI for tailoring variable depression

Tailoring Variable: Depressed	Outcome: BMI		
Estimator	% times η^{opt} chosen	$E[Y]$	95% CI of $E[Y]$
$s-VSE_{IPW}$	0.995	-5.053	(-6.714,-5.780)
$s-VSE_{W_1}$	0.990	N/A	N/A
$s-VSE_{W_3}$	0.965	N/A	N/A
$a-VSE_{IPW}$	1.000	-5.053	(-6.744,-5.835)
	Outcome: Self-reported General Health		
Estimator	% times η^{opt} chosen	$E[Y]$	95% CI of $E[Y]$
$s-VSE_{IPW}$	0.995	2.095	(2.018,2.170)
$s-VSE_{w_1}$	1.000	N/A	N/A
$s-VSE_{w_3}$	0.991	N/A	N/A
$a-VSE_{IPW}$	0.998	2.114	(2.048,2.179)
	Outcome: SBP		
Estimator	% times η^{opt} chosen	$E[Y]$	95% CI of $E[Y]$
$s-VSE_{IPW}$	0.905	91.447	(89.210,93.829)
$s-VSE_{w_1}$	0.974	N/A	N/A
$s-VSE_{w_3}$	0.898	N/A	N/A
$a-VSE_{IPW}$	0.677	92.401	(90.889,93.881)

Table 5–4: Linear regression of outcome BMI with tailoring TV hours watched and depression

Variables	Estimate	Std. Error	p-value
(Intercept)	-8.736	0.990	< 0.001
Physical activity	2.509	0.559	< 0.001
Age	0.042	0.001	< 0.001
Gender (male)	-0.150	0.291	0.607
Poverty	0.106	0.152	0.485
Race(Black)	-1.790	0.892	0.045
Race(Hispanic)	-0.265	0.993	0.789
Race(Mexican)	-2.668	0.934	0.004
Race(White)	-1.010	0.761	0.185
Race(Other)	-1.728	1.092	0.114
Income	9.723^{-6}	6.764^{-6}	0.151
Smoke status	2.033	0.306	< 0.001
Depressed	0.183	0.2.235	0.413
TVhour	-0.231	0.1281	0.071
Physical activity*depressed	-0.185	0.341	0.058
Physical activity*TVhours	-0.202	0.179	0.257

Table 5–5: Linear regression of outcome Self-reported General Health with tailoring TV hours watched and depression

Variables	Estimate	Std. Error	p-value
(Intercept)	1.943	0.160	< 0.001
Physical activity	0.377	0.090	< 0.001
age	-0.003	0.002	0.043
Gender(male)	-0.037	0.047	0.429
Poverty	0.137	0.025	< 0.001
Race(Black)	-0.222	0.144	0.123
Race(Hispanic)	-0.131	0.160	0.412
Race(Mexican)	-0.187	0.151	0.213
Race(White)	0.016	0.123	0.895
Race(Other)	0.134	0.176	0.447
Income	-3.385^{-7}	1.092^{-6}	0.756
Smoke status	-0.149	0.049	0.003
Depressed	-0.199	0.036	< 0.001
TVhours	0.014	0.021	0.476
Physical activity*depressed	0.073	0.055	0.194
Physical activity*TVhours	-0.034	0.029	0.227

Table 5–6: Linear regression of outcome SBP with tailoring TV hours watched and depression

Variables	Estimate	Std. Error	p-value
(Intercept)	117.1	3.199	< 0.001
Physical activity	3.934	1.802	0.029
Age	-0.504	0.032	< 0.001
Gender(male)	-0.968	0.938	0.302
Poverty	4.168	0.487	< 0.001
Race(Black)	-7.086	2.871	0.014
Race(Hispanic)	-3.374	3.197	0.292
Race(Mexican)	-3.728	3.007	0.215
Race(White)	-1.755	2.450	0.474
Race(Other)	2.951	3.516	0.401
Income	-1.710^{-4}	2.177^{-5}	< 0.001
Smoke status	-0.872	0.986	0.377
Depressed	1.550	0.719	0.031
TVhours	0.471	0.412	0.254
Physical activity*depressed	-2.420	1.096	0.057
Physical activity*TVhours	-0.956	0.576	0.101

5.4 Summary

Using the proposed VSEs, I found no evidence that physical activity recommendations should be tailored; as demonstrated by the plots, where the expected outcome for different values of tailoring threshold are very flat, with no significant peak that might indicate an optimal ITR out of all the candidate rules. The results from my non-parametric VSEs are supported by the regression-based approach, which found a significant effect of physical activity but no significant interaction with either hours of TV watched or level of depression. My findings are in line with current medical recommendations that physical activity benefits all (able) adults.

CHAPTER 6
Discussion and Conclusion

I have proposed new simple and augmented value search estimators to estimate the optimal individualized treatment regime over a specified class of regimes in a single stage setting, where the class must be defined prior to analysis based on clinical considerations such as cost and availability of the tailoring variable of interest. The new class of VSEs may be applied to data arising from either a clinical trial or an observational study; in the latter case, under the assumption of no unmeasured confounding, removal of confounding bias is achieved through modeling of the propensity score and creating appropriate weights such as inverse probability of treatment weights or absolute value weights. I proved consistency of the optimal threshold estimator (despite bias in the estimator of the value function) for all simple VSEs and augmented VSEs under the assumption of correct outcome model specification. In general, however, the a-VSEs were not doubly robust.

In this work, I have focused only on the one-interval setting, to estimate optimal ITRs. It is possible to extend this to a multiple-interval setting so as to examine DTRs, i.e. treatment sequences. However, as my results have suggested there is little benefit to using less variable weights to improving estimation of the parameters of ultimate interest, that is the ITRs; it is not clear that there is significant benefit to using weights other than the traditional and hence familiar, inverse probability of treatment weights.

A rather surprising result was observed in my simulations for the first scheme of data generation: unlike the estimator of the value function, the variance of the estimator for the optimal treatment threshold, η^{opt}, appears to be somewhat insensitive to large changes in the sample size. An important avenue for further research is understanding the convergence rate of the VSEs, as this will be crucial to understanding power and sample size considerations in future analyses.

The drive to personalize medicine continues, and the need for methods that are robust and easily explained to clinical researchers is a high priority. The statistical challenges are significant and exciting in a field that is less than two decades old.

References

Centers for Disease Control and Prevention. About the National Health and Nutrition Examination Survey. URL `http://www.cdc.gov/nchs/nhanes/about_nhanes.htm`.

B. Chakraborty and E. Moodie. *Statistical Methods for Dynamic Treatment Regimes: Reinforcement Learning, Causal Inference, and Personalized Medicine*. Springer, 2013.

D. R. Cox. Planning of experiments. 1958.

E. S. Ford, D. F. Williamson, and S. Liu. Weight change and diabetes incidence: findings from a national cohort of U.S. adults. *American Journal of Epidemiology*, 146(3):214–222, 1997.

C. E. Garber, B. Blissmer, M. R. Deschenes, B. A. Franklin, M. J. Lamonte, I.-M. Lee, D. C. Nieman, and D. P. Swain. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise. *Medicine and Science in Sports and Exercise*, 43(7):1334–1359, 2011.

J. Hammersley and D. Handscomb. Monte Carlo Methods (Methuen, London, 1964).

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

M. Á. Hernán, B. Brumback, and J. M. Robins. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11(5):561–570, 2000.

P. W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986.

A. Jorm, A. Korten, P. Jacomb, H. Christensen, B. Rodgers, and P. Pollitt. Public beliefs about causes and risk factors for depression and schizophrenia. *Social Psychiatry and Psychiatric Epidemiology*, 32(3):143–148, 1997.

M. R. Kosorok and E. E. Moodie. *Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine*. SIAM, 2015.

P. W. Lavori and R. Dawson. Dynamic treatment regimes: practical design considerations. *Clinical Trials*, 1(1):9–20, 2004.

A. H. Mokdad, E. S. Ford, B. A. Bowman, W. H. Dietz, F. Vinicor, V. S. Bales, and J. S. Marks. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *Journal of the American Medical Association*, 289(1):76–79, 2003.

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

S. A. Murphy. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):331–355, 2003.

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

A. Must, J. Spadano, E. H. Coakley, A. E. Field, G. Colditz, and W. H. Dietz. The disease burden associated with overweight and obesity. *Journal of the American*

Medical Association, 282(16):1523–1529, 1999.

K. L. Ong, B. M. Cheung, Y. B. Man, C. P. Lau, and K. S. Lam. Prevalence, awareness, treatment, and control of hypertension among United States adults 1999–2004. *Hypertension*, 49(1):69–75, 2007.

M. Petersen, C. T. Yiannoutsos, A. Justice, and M. Egger. Observational research on NCDs in HIV-positive populations: conceptual and methodological considerations. *Journal of Acquired Immune Deficiency Syndromes (1999)*, 67(01):S8, 2014.

J. Robins. Estimation of the time-dependent accelerated failure time model in the presence of confounding factors. *Biometrika*, 79(2):321–334, 1992.

J. M. Robins. Marginal structural models versus structural nested models as tools for causal inference. In *Statistical models in Epidemiology, the Environment, and Clinical Trials*, pages 95–133. Springer, 2000.

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

J. M. Robins, M. A. Hernan, and B. Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560, 2000.

P. R. Rosenbaum. Discussing hidden bias in observational studies. *Annals of Internal Medicine*, 115(11):901–905, 1991.

P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688, 1974.

D. B. Rubin. Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American Statistical Association*, 75(371): 591–593, 1980.

J. A. Sterne, M. A. Hernán, B. Ledergerber, K. Tilling, R. Weber, P. Sendi, M. Rickenbach, J. M. Robins, M. Egger, S. H. C. Study, et al. Long-term effectiveness of potent antiretroviral therapy in preventing aids and death: a prospective cohort study. *The Lancet*, 366(9483):378–384, 2005.

T. S. Stroup, J. P. McEvoy, M. S. Swartz, M. J. Byerly, I. D. Glick, J. M. Canive, M. F. McGee, G. M. Simpson, M. C. Stevens, and J. A. Lieberman. The national institute of mental health clinical antipsychotic trials of intervention effectiveness (catie) project: Schizophrenia trial design and protocol development. *Schizophrenia Bulletin*, 29(1):15–31, 2003.

A. Taylor, N. Cable, G. Faulkner, M. Hillsdon, M. Narici, and A. Van Der Bij. Physical activity and older adults: a review of health benefits and the effectiveness of interventions. *Journal of Sports Sciences*, 22(8):703–725, 2004.

M. T. Tsuang and R. F. Woolson. Mortality in patients with schizophrenia, mania, depression and surgical conditions. a comparison with general population mortality. *The British Journal of Psychiatry*, 130(2):162–166, 1977.

United States. Department of Health. *Physical activity and health: a report of the Surgeon General*. Diane Publishing, 1996.

E. H. Wagner, B. T. Austin, C. Davis, M. Hindmarsh, J. Schaefer, and A. Bonomi. Improving chronic illness care: Translating evidence into action. *Health Affairs*, 20(6):64–78, 2001.

M. P. Wallace and E. E. Moodie. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*, 71(3):636–644, 2015.

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

C. J. C. H. Watkins. *Learning from Delayed Rewards*. PhD thesis, University of Cambridge England, 1989.

B. Zhang, A. A. Tsiatis, E. B. Laber, and M. Davidian. A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018, 2012.

Y. Zhao, D. Zeng, A. J. Rush, and M. R. Kosorok. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118, 2012.