2022

Causal Inference Methods For Joint Censored Cost And Effectiveness Outcomes

Nicholas Andrew Illenberger
University of Pennsylvania

Causal Inference Methods For Joint Censored Cost And Effectiveness Outcomes

Abstract

Informed healthcare policy decisions must be driven by consideration of an intervention's effectiveness as well as its cost. Cost-effectiveness analyses provide a framework for decision making that balances these joint outcomes in some optimal way. However, because these studies often use data from observational sources, results may be biased due to unmeasured or time-varying confounding, informative cost censoring, and skewed or zero-inflated data. The goals of this dissertation are two-fold; we aim to (1) elucidate the conditions under which causal conclusions can be drawn from cost-effectiveness data, and (2) develop novel statistical methods for identifying cost-effective treatments while accounting for confounding and other data irregularities. We discuss three such developments: regression methodology for a novel probabilistic measure of cost-effectiveness, interpretable Q-learning based methods for identifying cost-effective treatment strategies, and a flexible and efficient influence function based estimator of average treatment cost that is robust to unmeasured confounding given a valid instrumental variable. We evaluate the operating characteristics of our proposed methods under several realistic data scenarios through simulation studies. We also illustrate usage by identifying cost-effective adjuvant treatments for early-stage endometrial cancer patients as well as assessing differences in costs between surgical and non-surgical interventions for gallstones and hemorrhaging using observational data.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Epidemiology & Biostatistics

First Advisor

Nandita Mitra

Subject Categories

Biostatistics

CAUSAL INFERENCE METHODS FOR JOINT CENSORED COST AND EFFECTIVENESS OUTCOMES

Nicholas Andrew Illenberger

A DISSERTATION

in

Epidemiology and Biostatistics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2022

Supervisor of Dissertation

_____

Nandita Mitra, Professor of Biostatistics

Graduate Group Chairperson

_____

Nandita Mitra, Professor of Biostatistics

Dissertation Committee

Kristin A. Linn, Assistant Professor, Biostatistics

Wei Yang, Associate Professor, Biostatistics

Justin Bekelman, Professor, Radiation Oncology

Andrew J. Spieker, Assistant Professor, Biostatistics, Vanderbilt University Medical Center

CAUSAL INFERENCE METHODS FOR JOINT CENSORED COST AND EFFECTIVENESS OUTCOMES

ACKNOWLEDGEMENT

This dissertation could not have been completed without the support, guidance, and generosity of the mentors and friends I've met throughout my graduate studies. I would first like to thank my advisor, Nandita Mitra, for her unrelenting support and for her friendship. For each new research direction, conference presentation, and job application, you have always been there to provide your sincere advice. Thank you in advance for the questions I will be sure to ask and for the support you will be sure to give in the future.

Thank you to the members of my committee: Kristin Linn, Andrew Spieker, Wei Yang, and Justin Bekelman. Your feedback and advice has been critical in shaping this dissertation. Thank you also to my collaborators: Austin Kilaru, Zachary Frosch, Sameed Khatana, Julia Scialla, Jennie Ma, Indika Mallawaarachchi, Joe Sobanko, Brian Cheng, and Shannon Nugent. I am grateful for our conversations, many of which have inspired the developments within this body of work and many more of which have inspired ideas for future works.

Thank you to all the students in the department. To those within my cohort —Carolyn, Justin, Lily, Angela, Jianqiao, Ken, Hayley, and Jill —for commiserating during difficult times and for celebrating when things have inevitably worked out. To those without: Mac, Arman, Joanna, Gary, Ali, Sarah, Danni, Jenny, Caleb, Benny, and everyone else who deserves mention. My life is brighter for having met you all.

Thank you to the faculty and staff I've met while working in the department. In particular, thank you to Pamela Shaw, who mentored my first rotation project and helped orient me on this road. Thank you to Andrew Spieker, whose vision for uniting complex methodologies with clinical utility helped provide the foundation for this work. And thank you to Eli Elliot and Catherine Vallejo for their assistance with all things great and small.

Thank you to my parents who have loved and supported me through everything I've done. I could not be happier to see myself becoming more and more like you both everyday. Thank you to my sister, Jessica, for being a constant role model. And finally, thank you to Carolyn. You are my best friend and there are not words to express how grateful I am to have met you.

ABSTRACT

CAUSAL INFERENCE METHODS FOR JOINT CENSORED COST AND EFFECTIVENESS OUTCOMES

Nicholas Andrew Illenberger

Nandita Mitra

Informed healthcare policy decisions must be driven by consideration of an intervention's effectiveness as well as its cost. Cost-effectiveness analyses provide a framework for decision making that balances these joint outcomes in some optimal way. However, because these studies often use data from observational sources, results may be biased due to unmeasured or time-varying confounding, informative cost censoring, and skewed or zero-inflated data. The goals of this dissertation are two-fold; we aim to (1) elucidate the conditions under which causal conclusions can be drawn from cost-effectiveness data, and (2) develop novel statistical methods for identifying cost-effective treatments while accounting for confounding and other data irregularities. We discuss three such developments: regression methodology for a novel probabilistic measure of cost-effectiveness, interpretable Q-learning based methods for identifying cost-effective treatment strategies, and a flexible and efficient influence function based estimator of average treatment cost that is robust to unmeasured confounding given a valid instrumental variable. We evaluate the operating characteristics of our proposed methods under several realistic data scenarios through simulation studies. We also illustrate usage by identifying cost-effective adjuvant treatments for early-stage endometrial cancer patients as well as assessing differences in costs between surgical and non-surgical interventions for gallstones and hemorrhaging using observational data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

Due to rising healthcare costs, there has been growing interest in improved methods for cost-effectiveness analyses. Cost-effectiveness research is concerned with identifying policies that can balance treatment effectiveness and overall costs. Optimizing decisions with respect to effectiveness may result in prohibitively expensive treatment strategies; on the other hand, optimizing with respect to costs may result in poor patient outcomes. Data for cost-effectiveness analyses are often drawn from large observational databases, such as insurance claims data. Because patients in these databases often have claims data for long periods of time, censoring and time-varying confounding is often present. The major complications of analysing cost-effectiveness data are (1) joint cost and effectiveness outcomes, (2) informative cost censoring, (3) measured and unmeasured confounding that may vary over time, and (4) irregular cost and effectiveness distributions. Irregularity in the distributions of these outcomes may include right-skewed or zero-inflated costs, and censored survival data.

This dissertation is focused on developing causal inference methodologies for determining the cost-effectiveness of medical interventions. We propose three major contributions: (1) regression methods for a probabilistic measure of cost-effectiveness, (2) a two-step procedure built upon Q-learning methodology for identifying optimal cost-effective treatment strategies, and (3) an instrumental variable estimator of expected treatment costs that can be used with censored cost data. The methods proposed have been constructed to allow for flexible modelling of joint cost and effectiveness outcomes, corrections for informative cost censoring, and adjustments for confounding- measured and unmeasured, static and time-varying. While developed for use in cost-effectiveness analyses, these methods are applicable in any complex settings with a ”cost-benefit” framework or other joint outcomes. For example, they may be used to find optimal treatments that can maximize patient health outcomes survival while reducing the risk of adverse events associated with aggressive treatments.

In Chapter 2, we develop regression methodology to assess heterogeneity in the cost-effectiveness of a treatment due to measured covariates. This work focuses on the Net Benefit Separation (NBS), a probabilistic measure of cost-effectiveness that describes the probability that a randomly selected

patient receiving treatment will experience greater treatment benefit than another patient receiving control. Because regression methodologies are traditionally used to estimate the relationship between measured covariates and the expected value of an outcome, certain modifications are required to estimate the NBS, a measure of stochastic ordering. Additionally, because data in cost-effectiveness settings are typically drawn from observational databases and may be censored, we must consider methods to account for both exposure-outcome confounding and informative cost censoring.

Regression approaches for alternative measures of average cost-effectiveness, namely the Incremental Cost-Effectiveness Ratio (ICER) and the Net Monetary Benefit (NMB) have been explored (Nixon and Thompson, 2005; Willan, Briggs, and Hoch, 2004). Because the distribution of medical costs is often skewed, cost-effectiveness comparators that do not rely on the average values of cost or measures of effectiveness can be useful. Pepe (1997) developed regression methods for measures of stochastic ordering for use in diagnostic testing settings. Their methodology cannot directly be used in cost-effectiveness settings because it does not account for exposure-outcome confounding, nor can it accommodate censored outcomes. In their paper introducing the NBS, Spieker et al. (2021) describe a Monte Carlo standardization procedure combined with Inverse Probability of Censoring Weighting (IPCW) approaches to account for both concerns when the estimand of interest is the population-level NBS. Because the cost-effectiveness of a treatment may vary according to patient covariates modifications to their approach are required to estimate conditional variants of the NBS.

Our proposed regression methodology can account for variability in the NBS using observed covariates. Because the NBS can be characterized as a measure of stochastic ordering, we can embed the approach of Pepe (1997) within our proposed methodology. We implement a Monte Carlo Standardization procedure to account for observed differences between patients receiving treatment and those receiving control. To adjust for informative cost censoring, models for the cost distribution used within the standardization procedure are estimated using IPCW approaches. Additionally, in this chapter we propose a hypothesis testing framework based upon the non-parametric bootstrap for determining the strength of evidence for the effects of covariates on NBS.

In Chapter 3, we develop methodology for identifying optimal and interpretable dynamic treatment regimes under a constraint on overall treatment costs. In addition, we describe a novel cost-

effectiveness analysis that can be used to select an optimally cost-effective regime from among a set of candidate regimes corresponding to different cost-constraints. Due to heterogeneity in how patients respond to treatments as well as in their cost accrual rates, individualized treatment strategies which can identify which treatments are most beneficial for which patients are necessary. Estimation of optimal treatment rules within this setting are complicated by time-varying confounding and censoring. When possible, the form of treatment rules should be restricted to ensure that they are interpretable to health policy professionals.

Recently, Lakkaraju and Rudin (2017) and Xu et al. (2020) have developed methodology to learn optimal cost-effective individualized treatment rules in settings with point-exposures. Both propose treatment rules that maximize treatment effectiveness while placing a penalty on overall treatment cost. For settings with time-varying treatments, Laber et al. (2018) proposed policy search methodology to identify optimal treatment regimes under a pre-determined constraint on overall cost (defined by number of safety events). While this approach can accommodate time-varying treatments, the proposed decision rules take the form of linear decision boundaries, which may be difficult to interpret. Alternatively, Zhang et al. (2018) argue that treatment rules based on decision lists provide both flexibility and interpretability.

In this work, we propose an efficient algorithm for identifying optimal list-based decision rules under fixed cost constraints. We use Q-learning with policy search methodology to maximize treatment effectiveness within the class of list-based regimes with cost constraints. We modify an algorithm proposed by Zhang et al. (2018) to efficiently estimate list-based decision rules while placing a constraint on overall cost. Additionally, we propose a novel iterative cost-effectiveness analysis that uses ICER to select an optimally cost-effective regime from a set of candidate regimes corresponding to different cost-constraints. Use of an iterative cost-effectiveness analysis allows us to (1) characterize how the optimal regime changes as cost-restrictions vary, and (2) avoid complications arising from non-sequential cost-effectiveness analyses in settings with two or more experimental treatments.

In Chapter 4, we provide instrumental variable methodology for estimating differences in treatment cost. Studies aimed at estimating medical costs accrued under different treatments are necessary for making informed healthcare policy decisions. These methods are particularly useful when analysing treatments with similar clinical effectiveness. In a recent series of papers, Keele et al.

(2018, 2019) and Fogarty et al. (2021) assessed the effect of treatment by emergency general surgery (EGS) on patient mortality for a set of acute conditions for which non-operative alternatives exist. While more aggressive forms of care often exhibit protective effects in adjacent medical settings (Lorch et al., 2012), these analyses showed a lack of beneficial evidence for EGS for many of the conditions under study. Here, knowledge of how EGS affects accrued medical costs is critical for making informed policy recommendations. Data for performing these analyses are often pulled from complex observational sources such as claims databases. Common complications in this setting are: (1) measured and unmeasured confounding of the exposure-outcome relationship, (2) informative cost censoring, and (3) irregular cost distributions. Methodologies that are robust to both sources of confounding and that are flexible enough to account for irregularities in the cost distribution and censoring have not previously been explored.

Popular IV approaches such as two stage least squares (2SLS) or two stage predictor substitution (2SPS) have become more common in medical and epidemiologic research (Keele et al., 2018, 2019; Lorch et al., 2012). Because these methods often rely on parametric assumptions, implementation may not be valid in settings with highly skewed cost data. Additionally, extensions of these methods to settings with informatively censored cost outcomes have not been adequately developed. Influence function based approaches have recently been proposed by Díaz (2019) and Lee, Kennedy, and Mitra (2020) to identify treatment effects on patient survival in the presence of censoring. These methods are doubly robust, efficient, and can incorporate semiparametric modeling approaches. However, extensions to informatively censored cost data have not been explored.

We propose two instrumental variable estimators of the complier average treatment effect on costs. Given a valid instrument, these estimators are unbiased in the presence of unmeasured confounding. Additionally, the use of a partitioned cost estimator allows us to address informative cost censoring and improve efficiency by utilizing data from patients with censored medical costs. The proposed estimators are based on influence functions and hence provide doubly robust, semiparametric, and efficient estimation of treatment effects as well as valid asymptotic inference. We present results from simulation studies assessing the performance of the proposed estimator under varying degrees of censoring, IV strength, and nuisance function misspecification. We also apply the proposed approach to a study assessing the costs of surgical and non-surgical interventions for gallstones and hemorrhaging using observational data.

CHAPTER 2

A REGRESSION FRAMEWORK FOR A PROBABILISTIC MEASURE OF COST-EFFECTIVENESS

2.1. Introduction

Cost-effectiveness analyses are a useful tool for aggregating information on the differences in cost and clinical effectiveness between two comparator health interventions. Typically, a cost-effectiveness analysis relies on defining a particular kind of summary measure, which may be used to optimize the tradeoff between treatment costs and effectiveness. Historically, most analyses have involved either the incremental cost-effectiveness ratio (ICER) or the equivalent net monetary benefit (NMB). Each of these compares the mean difference in cost and the mean difference in some clinical measure between treatments. Specifically, the ICER is defined as the ratio of these two quantities and therefore has units of cost per unit of the clinical measure (for instance, $USD per year, if the clinical outcome is marked by post-treatment survival time). The experimental treatment is declared cost-effective if the ICER lies below a certain threshold, known as the willingness-to-pay (WTP). The NMB was introduced by Stinnett and Mullahy (1998) as a refined, yet equivalent measure of cost-effectiveness that suffers from neither a singularity in the denominator nor the known poor statistical properties associated with ratio quantities.

In recent work, Spieker et al. (2021) introduced a novel probabilistic measure of cost-effectiveness, the net benefit separation (NBS). The NBS characterizes the stochastic ordering of individual net benefits (INB) between treated and untreated populations. The INB itself is not a comparative measure; rather, the NBS characterizes the difference in the distribution of INBs between treated and untreated groups and can serve as an alternative measure of a treatment's cost-effectiveness. Compared with NMB and ICER, which are functions of the average difference in cost and effectiveness measures between treatments, NBS measures the probability that a patient receiving treatment will experience greater treatment benefit than a patient receiving control. Because the distribution of medical costs, and consequently INBs, are often skewed, cost-effectiveness comparators that do not rely on the average values of cost or measures of effectiveness can be useful. In settings where treatment is randomized, NBS can be estimated nonparametrically using a scaled

variant of the Wilcoxon rank-sum statistic. As with NMB and ICER, extensions of NBS to settings with informative cost censoring and confounding have been proposed. Spieker et al. (2021) introduce a semiparametric Monte Carlo standardization procedure to estimate NBS while adjusting for differences in the distribution of confounding variables across treatment arms. Models used within the standardization procedure may be fit using inverse probability of censoring weighting (IPCW) techniques if there are censored cost data (Bang and Tsiatis, 2000).

In practice, the extent of cost-effectiveness associated with a treatment may vary across levels of an observed covariate. As an example, consider the setting of endometrial cancer, in which treatment often involves total hysterectomy. Following surgery, there is a large degree of variation in the treatment a patient receives, with some receiving adjuvant radiation, some receiving adjuvant chemotherapy, and some receiving neither (Latif et al., 2014). While costs associated with radiation and chemotherapy are certain to exceed those associated with a control condition, the survival benefit associated with either radiation or chemotherapy relative to neither could reasonably be expected to vary by patient covariates, e.g., cancer stage. Specifically, those with a lower cancer stage (and therefore, a lower overall risk of recurrence) may not experience the same degree of survival benefit from adjuvant treatment as compared to those with a higher cancer stage. In addition, because NBS and NMB are complementary but not equivalent measures of cost-effectiveness, they can provide different conclusions regarding the cost-effectiveness of a treatment. For example, if costs are highly right-skewed, then the expected monetary benefit of a treatment may be negative even if treated patients tend to observe greater monetary benefit than untreated ones. Understanding how the cost-effectiveness of a treatment changes with respect to measured covariates can help reconcile this heterogeneity.

In previous work, regression methods for NMB and ICER have been developed to compare cost-effectiveness across subgroups defined by observed covariates. Willan, Briggs, and Hoch (2004) show how NMB may be characterized as a function of the parameters in a regression of cost and effectiveness outcomes on treatment and observed covariates. Under the assumption of bivariate normal errors, the authors proposed a hypothesis test for the effects of covariates on NMB. Because cost distributions are typically skewed, Nixon et al. presented an alternative approach which could accommodate more flexible error structures (Nixon and Thompson, 2005). While positing regression models similar to those of Willan, Briggs, and Hoch (2004), their method implements

Markov Chain Monte Carlo methods to accommodate gamma and log-normal distributed errors. These approaches are useful if we are interested in how the cost-effectiveness of an intervention (as measured by NMB or ICER) varies across levels of measured covariates, however no methods currently exist which can explain variability in NBS using covariates.

In this paper, we introduce regression methodology to account for variability in the NBS using observed covariates. In previous work, Pepe et al. developed regression methods for measures of stochastic ordering for use in diagnostic testing settings (Pepe, 1997). Alonzo and Pepe later extended this work to improve computational efficiency in large samples (Alonzo and Pepe, 2002). Because the NBS is defined as the stochastic ordering of INBs between treatment groups, we are able to embed their approach within our proposed methodology. We also describe how Monte Carlo standardization and IPCW methods may be used to account for censoring and confounding when performing NBS regression. Additionally, we propose a hypothesis testing framework for determining the strength of evidence for the effects of covariates on NBS.

The remainder of this paper is organized as follows: In Section 2.2 we review the net benefit separation as proposed by Spieker et al. (2021). In Section 2.3 we define a conditional variant of the NBS and introduce our regression-based estimation procedure. In Section 2.4 we conduct simulation studies to examine the finite sample properties of this estimator. In Section 2.5, we use our proposed approach to analyze the cost-effectiveness of adjuvant treatments for endometrial cancer. In the final section, we summarise the utility of our proposed cost-effectiveness parameters and discuss limitations of our methodology.

2.2. Net benefit separation

Assume we have individual patient data on costs and treatment effectiveness. Let A be a binary treatment indicator, Z be a measure of effectiveness, Y be a measure of medical cost, and let $i = 1, \ldots, n$ index sample units. In a typical setting, Z and Y may represent survival time and medical cost in USD($). For predefined WTP, λ, INB is defined as $B(\lambda) = \lambda Z - Y$. We assume without loss of generality that larger values of Z are desirable, so that greater INB indicates greater treatment benefit. If we let $B_a(\lambda)$ represent a randomly sampled INB from a hypothetical population

receiving treatment $A = a$, then NBS is a measure of stochastic ordering that is defined as

$$\theta(\lambda) = P\big(B_1(\lambda) > B_0(\lambda)\big).$$

NBS can be interpreted as the probability that a randomly sampled unit from a population receiving treatment will have a greater INB than an independently sampled unit from a control population. If the only distributional difference between INBs in treated and untreated populations is a location shift, then the NBS may be used to describe differences in the median INB, although this is not generally true. For a particular λ, if $\theta(\lambda) = 0.5$, then the treatments are equally cost-effective with respect to NBS. To be specific, $\theta(\lambda) = 0.5$ implies that randomly selected INBs from a treated population do not tend to be larger than those from an untreated distribution and vice versa. Note also that $\theta(\lambda) > 0.5$ does not imply $\mathbb{E}\left[B_1(\lambda)\right] > \mathbb{E}\left[B_0(\lambda)\right]$, i.e. cost-effectiveness with respect to NBS does not imply greater expected INB. For example, if the distribution of INBs is skewed, then the expected benefit under control may be larger than that under treatment even if the majority of patients receiving treatment experience greater benefit than those receiving control. Thus, NBS can help provide a more complete view of the range of patient experiences, particularly when cost and effectiveness metrics are skewed.

For simplicity, we first consider estimation in a setting where treatment is randomized and with no censoring. Here, the observed distribution of INBs in the treatment and control arms form nonparametric estimates of the distributions of $B_1(\lambda)$ and $B_0(\lambda)$. Thus, NBS can be consistently estimated using a scaled variant of the Wilcoxon rank-sum statistic

$$\widehat{\theta}(\lambda) = \frac{1}{2N_0}\left(\frac{1}{N_1}\sum_{i=1}^{N} 2A_i R_i(\lambda) - N_1 - 1\right),$$

where N_a is the number of participants receiving treatment $A = a$ and $R_i(\lambda)$ is the pooled rank of the INB for the i^{th} individual. Spieker et al. proposed a flexible standardization approach to estimate the NBS in the presence of measured confounders (Spieker et al., 2021). By modelling the joint distribution of observed variables, they are able to apply Monte Carlo methodology to sample INBs from treated and untreated populations which have the same underlying distribution of confounders. Applying the scaled Wilcoxon rank-sum statistic to the generated INBs provides an estimate of NBS

that accounts for measured confounding. If data are censored, then IPCW can be incorporated in the modeling step to account for censored cost data. In the section that follows, we will extend this methodology to evaluate the association between measured covariates and the NBS.

2.3. Regression of net benefit separation

In this section we introduce regression methodology for NBS. Section 2.3.1 defines a conditional variant of NBS and describes procedures for estimation and hypothesis testing when treatment is randomized. In Section 2.3.2 we propose a Monte Carlo standardization procedure to address measured confounding. Section 2.3.3 discusses how to incorporate IPCW to account for informative cost censoring.

2.3.1. Conditional net benefit separation

Suppose X is some covariate of interest which may explain variation in the NBS. In an endometrial cancer setting, X may represent cancer stage or age at diagnosis. We want to assess how the cost-effectiveness of a treatment changes with respect to X. Specifically, we may seek to (1) characterize the association between X and NBS, (2) assess the NBS across levels of X, and (3) evaluate the strength of evidence for association. When treatment is randomized and X is discrete, the first two goals can be accomplished by estimating $\theta(\lambda|X = x)$ using the scaled Wilcoxon rank-sum statistic within subsets defined by $X = x$. On the other hand, if X is continuous, then this is not possible without first discretizing X. In either case, there are currently no methods to assess the strength of association between X and the NBS. In this section, we introduce a procedure that can accomplish all three of the stated goals for continuous or discrete X.

For a given λ, the NBS conditional on X is defined as

$$\theta(\lambda|X) = P(B_1(\lambda) > B_0(\lambda)|X). \tag{2.1}$$

It can be shown that the conditional NBS is equivalent to

$$\theta(\lambda|X) = \int_0^1 S_{1|X}(S_{0|X}^{-1}(\omega))d\omega \tag{2.2}$$

where $S_{a|X}(u) = P(B_a(\lambda) > u|X)$ denotes the conditional survivor function for INB in a hypothetical

population receiving treatment $A = a$ and $S^{-1}_{a|X}(u)$ is the quantile function for the same population. To estimate $\theta(\lambda|X)$, we build methodology parallel to that used by Alonzo and Pepe (2002) for ROC regression. Earlier work by Pepe (2000) noted that if we define $U_{ij} = \mathbb{1}(B_{1i}(\lambda) > B_{0j}(\lambda))$ where i and j index treated and untreated patients, then $S_{1|x}(S^{-1}_{0|x}(\omega_j))$ is equivalent to the expectation of U_{ij} conditional on $S_{0|x}(B_{0j}) = \omega_j$ and $X = x$. Given estimated quantiles $\widehat{\omega}_j$, standard GLM procedures can be used to estimate $S_{1|x}(S^{-1}_{0|x}(\omega))$. Because U_{ij} must be defined for every pair of treated and untreated units, this approach is intractable for large samples. Instead, we use an alternative approach proposed by Alonzo and Pepe (2002) which improves on the efficiency of the original estimator. Let Ω be a finite set of values between 0 and 1. For each $\omega \in \Omega$ and each unit i receiving treatment define $U_{i\omega} = \mathbb{1}(B_{1i}(\lambda) > S^{-1}_{0|X_i}(\omega))$. That is, $U_{i\omega}$ is the indicator of whether the i-th treated unit's INB is greater than the ω-th quantile of the distribution of INBs among the untreated units. It follows that $\mathbb{E}[U_{i\omega}] = S_{1|X_i}(S^{-1}_{0|X_i}(\omega))$. For a given probability, $\omega \in \Omega$, this definition of $U_{i\omega}$ gives the indicator of whether a treated unit's observed INB is greater than the ω-quantile of the distribution of INBs in an untreated population. In practice, because $S^{-1}_{0|X}(\omega)$ is unknown, $U_{i\omega}$ is defined using a consistent estimator of the quantile function $U_{i\omega} = \mathbb{1}(B_i(\lambda) > \widehat{S}^{-1}_{0|X_i}(\omega))$. Alonzo and Pepe (2002) suggest using empirical estimates of the quantiles when possible (i.e. for discrete X) and quantile regression (Koenker and Bassett Jr, 1978) otherwise.

The conditional expectation of $U_{i\omega}$ can be modeled using any flexible regression approach. One sensible modeling choice is the probit form: $\mathbb{E}[U_\omega|X] = \Phi(\beta_0 + \beta_1 X + \beta_2 \Phi^{-1}(\omega))$ where $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ are the cumulative distribution function and the quantile function of the standard normal distribution respectively. This is the form that would arise if $B_1(\lambda)$ and $B_0(\lambda)$ were normally distributed or if they could be monotonically rescaled to be normally distributed. Hanley (1988) shows that when fitting ROC curves using the binormal form, estimates of AUC exhibit low bias even when the model is misspecified. Because we characterize NBS as the area under the fit binormal model, it can be seen as analogous to AUC. Once the conditional expectation has been modeled, NBS can be estimated through numerical integration using standard software (Piessens et al., 2012). Under the probit model, our estimate is given by

$$\widehat{\theta}(\lambda|X) = \int_0^1 \Phi(\widehat{\beta}_0 + \widehat{\beta}_1 X + \widehat{\beta}_2 \Phi^{-1}(\omega)) d\omega. \tag{2.3}$$

Testing whether X influences $\theta(\lambda|X)$ is equivalent to testing whether the coefficient(s) associated with X are equal to zero. For example, under the probit model this is a test of if $\beta_1 = 0$. Because multiple values $U_{i\omega}$ are defined for each individual (one for each $\omega \in \Omega$), model-based variance estimates that assume independent observations are invalid. Instead, we propose the nonparametric bootstrap hypothesis test. An important consideration in the regression of NBS is the choice of Ω. For sets consisting of N_ω equally spaced points, $\Omega = \{j/(N_\omega + 1); j = 1, ..., N_\omega\}$, selecting $N_\omega = N_0 - 1$ is equivalent to the estimator originally proposed by Pepe (2000) and is maximally efficient. However, Alonzo and Pepe (2002) showed that smaller choices of N_ω were able to achieve close to maximal efficiency while mitigating computational burden.

2.3.2. Adjustment for measured confounding

Suppose that there are measured covariates, L, which affect the probability of receiving treatment as well as the cost or effectiveness outcomes. Our proposed estimation procedure consists of two steps: a Monte Carlo standardization step, and a regression step. The first step constructs two pseudo-populations which differ in treatment status, but where the distribution of confounding variables, L, is equivalent. Let $\widehat{f}(z, y|A, X, L)$ denote some model for the conditional distribution of cost and effectiveness measures. The pseudo-population under treatment $A = a$ is built by drawing M observations from the empirical distribution of the baseline covariates, $\widehat{F}(L, X)$, and then using the sampled values of L and X to sample from the fit model $\widehat{f}(z, y|A = a, X, L)$. From these pseudo-populations, we may sample from the joint distributions of the INBs and X among treated and untreated units, $(B_1(\lambda), X)$ and $(B_0(\lambda), X)$ respectively. This methodology ensures that INBs are sampled from populations which receive different treatments, but have the same underlying distribution of measured confounders. This procedure is semiparametric because, while parametric models may be used to fit the distributions of Z and Y, the joint distribution $F(X, L)$ is estimated nonparametrically. The Monte Carlo standardization procedure is summarized below:

1. Model the conditional distributions $f(Z|A, X, L)$ and $f(Y|A, X, L, Z)$ where $f(\cdot)$ and $F(\cdot)$ denote the probability density function and the cumulative density function of a random variable. Let $\widehat{f}(Z|A, X, L)$ and $\widehat{f}(Y|A, X, L, Z)$ denote the estimated distributions.

2. For $a = 0, 1$:

    - Sample $\widetilde{l}_m$ and $\widetilde{x}_m$ for $m = 1, ..., M$ from the empirical distribution $\widehat{F}(L, X)$.

- For $m = 1, ..., M$ sample $\widetilde{z}_m$ and $\widetilde{y}_m$ from $\widehat{f}(Z|A = a, \widetilde{x}_m, \widetilde{l}_m)$ and $\widehat{f}(Y|A = a, \widetilde{x}_m, \widetilde{l}_m, \widetilde{z}_m)$
- Calculate $\widetilde{B}_{a,m}(\lambda)$ for each pair $(\widetilde{z}_m, \widetilde{y}_m)$

The goal of the Monte Carlo standardization procedure is to sample INBs from their true underlying distribution while adjusting for differences in measured covariates between treatment groups. Because of this, it is important to incorporate residual variance when drawing samples from $\widehat{f}(Z|A, X, L)$ and $\widehat{f}(Y, A, X, L, Z)$. This can be accomplished by adding randomly sampled residuals to predicted values arising from the fit models. Additionally, the models for $f(Z|A, X, L)$ and $f(Y|A, X, L, Z)$ can be flexibly estimated to reduce bias. For example, zero-inflated gamma regression models can be used to accommodate skewed cost distributions with structural zeros. By including effectiveness as a predictor of cost, this procedure allows for correlation between Z and Y. In estimating the conditional distributions $f(Z|A, X, L)$ and $f(Y|A, X, L, Z)$, we aim to properly specify the joint distribution $f(Z, Y|A, X, L, Z)$. Alternatively, this joint distribution may be decomposed into $f(Y|A, X, L)$ and $f(Z|A, X, L, Y)$. Because these approaches are equivalent, the choice to model one distribution prior to the other is flexible. However, this decision should be driven by beliefs regarding correct model specification that relies on knowledge of the data and question of interest. Using flexible methodology to estimate the conditional distributions of cost and effectiveness may reduce the risk of model misspecification.

The second step of our estimation procedure entails regressing the NBS using the sampled INBs. Sampled values from the joint distributions of $(B_1(\lambda), X)$ and $(B_0(\lambda), X)$ are independent from the distribution of measured confounders. This independence allows us to perform NBS regression as described in Section 2.3.1. The regression step is outlined below:

1. For each $\omega \in \Omega$, use sampled values $\widetilde{B}_{0,m}(\lambda)$ to estimate $S^{-1}_{0|X}(\omega)$.
2. For each generated $\widetilde{B}_{1,m}(\lambda)$ and $\omega \in \Omega$, define $U_{m\omega} = \mathbb{1}(\widetilde{B}_{1,m}(\lambda) > \widehat{S}^{-1}_{0|X}(\omega))$.
3. Fit a model for $\mathbb{E}[U_{m\omega}|X]$
4. For $X = x$, estimate the subgroup-specific NBS by numerically integrating:

$$\widehat{\theta}(\lambda|X = x) = \int_0^1 \widehat{\mathbb{E}}[U_{i\omega}|X = x]$$

2.3.3. Informative Censoring

Up until this point, we have developed methods which are suitable for cases with no censoring, e.g. when Z is a continuous outcome with no missing data. Often, the effectiveness outcome of greatest clinical interest is time to death post treatment. Because patients followed over time may be lost to follow-up, we introduce methodology to account for censoring when estimating NBS. Suppose we are interested in estimating costs up until a maximum time, τ. For observation i, let C_i be the censoring time, T_i be the survival time, and $T_i^* = \text{min}(T_i, \tau)$. Define $\delta_i = \mathbb{1}(C_i < T_i)$ and $\delta_i^* = \mathbb{1}(C_i < T_i^*)$ be indicators of censored survival time and censored costs. The observable data are δ, δ^*, and $Z = min(T_i, C_i)$. Because costs at time of censoring can be correlated with costs at event time, it is necessary to use IPCW to estimate $f(Y|A, X, L, Z)$. Let $G(t) = P(t \leq C)$ denote the probability that the censoring time is beyond t. If we assume T and C are independent, then $G(t)$ can be modeled within each treatment group using the Kaplan-Meier product limit estimator based on the data (C, δ^*). If instead we assume that survival and censoring times are independent only once we condition on measured discrete and continuous variables, V and W, then we can use the stratified cox model to estimate $G(t)$. Under this model, we assume that the hazard function for censoring is given by:

$$h(t|V, W) = \text{exp}(\eta^T W) h_V(t)$$

where η^T is the transposed vector of regression coefficients, η. Under the stratified cox model the baseline hazard function, $h_V(t)$ may vary over strata defined by V. This enables adjustment for predictors that satisfy the proportional hazards assumption, W, as well as for those that do not, V. Importantly, selection of W and V is driven by beliefs regarding the censoring mechanism and may not be related to confounding variables, L, or effect modifiers, X. The coefficients in this model, η, can be estimated using the Cox partial likelihood (Cox, 1975):

$$\widehat{G}(T_i^*|V_i, W_i) = \text{exp}\left(- \sum_{j=1}^{n} \frac{\delta_j \mathbb{1}(V_i = V_j, T_i^* > X_j) \text{exp}(\widehat{\eta}^T W_i)}{\sum_{k=1}^{n} \mathbb{1}(V_i = V_k, X_k \geq X_j) \text{exp}(\widehat{\eta}^T W_k)} \right)$$

For each individual we define the IPCW as

$$w_i = \frac{\delta_i^*}{\widehat{G}(T_i^*)}$$

Note that for censored individuals $\delta_i^* = 0$, so that $w_i = 0$. The weights, w_i, can be incorporated into the Monte Carlo procedure when modeling $f(Y|A, X, L, Z)$. For example, suppose we assume that costs follow a log-normal distribution. To fit $f(Y|A, X, L, Z)$, we can perform weighted log-normal regression using only patients with fully observed cost data (i.e. those with $\delta_i^* = 1$). Any modeling procedure that incorporates weighting may be used to estimate the cost distribution. The resulting estimate, $\widehat{f}(Y|A, X, L, Z)$, can be used in the standardization procedure described in Section 2.3.2.

2.4. Simulations

We perform a simulation study to determine the finite-sample properties of the proposed regression estimator. Although data used for cost-effectiveness analyses often have thousands of observations, it is beneficial to understand how our methodology performs in both small and large sample settings. We consider simulation settings with low sample size ($n = 500$) and high sample size ($n = 5000$) as well as low (10%), medium (30%), and high (50%) rates of censoring. Finally, we consider the case where the stratifying variable is and is not associated with $\theta(\lambda|X)$. For each simulation, we estimate the NBS under two WTP thresholds, $\lambda = 2$ and 12. These represent the cost a payer is willing to incur for a unit change in effectiveness.

We generate the data as follows: $L \sim N(0, 1)$, $X \sim$ Bernoulli$(p_x = 0.25)$, $A \sim$ Bernoulli$(p_a =$ expit$(L))$, $T \sim$ Weibull$(k = 2, \lambda =$ exp$(4.05 + 0.15A + 0.2L + \beta_x X + \beta_{ax} A \times X))$, $C \sim$ Weibull$(k = 2, \lambda =$ exp$(\gamma + 0.5A))$, and $Y \sim$ Lognormal$(\mu = 4.2 + 0.002T + 0.5A, \sigma^2 = 0.16)$. In simulations where X is associated with $\theta(\lambda|X)$ we set $\beta_x = 0.1$ and $\beta_{ax} = 0.5$. Under these parameter values $\theta(2|X = 0) = 0.353$, $\theta(2|X = 1) = 0.588$, $\theta(12|X = 0) = 0.527$ and $\theta(12|X = 1) = 0.746$.

Treated subjects are expected to have greater costs and survival times than those which are untreated. Patients with $X = 1$ benefit more from treatment (with respect to survival) than those with $X = 0$. The separation between the distribution of INBs between treated and untreated subjects is larger in a population where $X = 1$ than in one with $X = 0$. In the setting where $\theta(\lambda|X)$ does not depend on X, both β_x and β_{ax} are set to zero. In these simulations, $\theta(2|X) = 0.353$ and $\theta(12|X) = 0.527$ regardless of the value of X. In simulations where X affects NBS, we set γ to

0.5119, 0.4410, and 3.9600 to induce 10%, 30%, and 50% censoring. When X does not affect NBS, these values are 5.007, 4.315, and 3.8760, respectively.

To estimate $\theta(\lambda|X)$, we use the standardization procedure described in Section 2.3.1. A Weibull regression model is used to estimate the distribution of T and a log-normal model is used for costs. Because Y is informatively censored, we use IPCW when estimating the cost model. To estimate the probability of being uncensored, we use the Cox proportional hazards model. In the standardization step, we draw $M = 5000$ observations within each level of A. For the regression of $S_{1|X}(S^{-1}_{0|X}(\omega))$, we obtain $\widehat{S}^{-1}_{0|X}(\omega))$ for $\omega \in \Omega$ using the empirical quantiles within levels of X. The previously described probit link is used to model $U_{i\omega}$ conditional on X. Standard errors are estimated as the standard deviation of $B = 300$ bootstrap replicates. The nominal type-I error rate is set to 0.05 for our hypothesis tests. Hypothesis tests are performed using symmetric 95% CIs based on the bootstrap replicates.

For each setting, we simulate 1000 datasets, estimate the NBS conditional on X, and perform the nonparametric bootstrap hypothesis test. We report the mean estimate of the NBS across simulations, the mean bootstrap estimated standard error of the NBS, the standard deviation of estimated NBS across simulations, and the proportion of hypothesis tests in which the null hypothesis was rejected. Tables 2.1 and 2.2 provide simulation results for the settings where X affects and does not affect the NBS, respectively.

Average point estimates of NBS are close to their true values for all tested values of X and λ. Bootstrap estimated standard errors closely approximate the empirical standard errors observed accross simulations. Due to a lower effective sample size, standard errors increase as the proportion of censored units increases. Similarly, higher variability in the estimates of $\theta(\lambda|X = 0)$ can be attributed to the fact that there is a lower prevalence of $X = 0$ in our simulations. In simulations where X affects NBS, power is lowest when there is a sample size of 500, 50% censoring, and a WTP of 2. All other observed rejection rates exceed 0.80. When X does not affect cost-effectiveness, the probability of rejecting the null is close to the desired 0.05. All simulations are performed in R 3.6.0. Code to replicate simulation results and for the following data example are provided in our supplemental materials.

Cens.	Sample Size	λ	X	$\theta(\lambda\|X)$	Est.	$\widehat{\text{SE}}$	ESE	Pr. Reject
10%	500	2	0	0.353	0.352	0.028	0.030	0.986
10%	500	2	1	0.588	0.589	0.049	0.049	–
10%	500	12	0	0.527	0.526	0.028	0.030	0.988
10%	500	12	1	0.746	0.745	0.038	0.038	–
10%	5000	2	0	0.353	0.354	0.011	0.010	1.000
10%	5000	2	1	0.588	0.587	0.019	0.020	–
10%	5000	12	0	0.527	0.527	0.011	0.010	1.000
10%	5000	12	1	0.746	0.745	0.015	0.015	–
30%	500	2	0	0.353	0.351	0.032	0.032	0.924
30%	500	2	1	0.588	0.591	0.061	0.065	–
30%	500	12	0	0.527	0.526	0.032	0.032	0.964
30%	500	12	1	0.746	0.746	0.045	0.046	–
30%	5000	2	0	0.353	0.352	0.012	0.012	1.000
30%	5000	2	1	0.588	0.588	0.023	0.023	–
30%	5000	12	0	0.527	0.526	0.012	0.012	1.000
30%	5000	12	1	0.746	0.745	0.017	0.017	–
50%	500	2	0	0.353	0.353	0.041	0.040	0.694
50%	500	2	1	0.588	0.580	0.083	0.087	–
50%	500	12	0	0.527	0.527	0.037	0.036	0.842
50%	500	12	1	0.746	0.742	0.057	0.055	–
50%	5000	2	0	0.353	0.353	0.015	0.016	1.000
50%	5000	2	1	0.588	0.586	0.034	0.033	–
50%	5000	12	0	0.527	0.527	0.013	0.014	1.000
50%	5000	12	1	0.746	0.745	0.020	0.019	–

Table 2.1: Results for setting where X has an effect on the NBS. Columns represent the probability of being censored, sample size, WTP (λ), value of X, true NBS ($\theta(\lambda|X)$), mean point estimate (Est.), mean estimated standard error ($\widehat{\text{SE}}$), empirical standard error (ESE), and proportion of simulations in which the null hypothesis of no effect of X is rejected (Pr. Reject).

Cens.	Sample Size	λ	X	$\theta(\lambda\|X)$	Est.	$\widehat{\text{SE}}$	ESE	Pr. Reject
10%	500	2	0	0.353	0.355	0.028	0.029	0.054
10%	500	2	1	0.353	0.355	0.047	0.049	–
10%	500	12	0	0.527	0.528	0.029	0.029	0.048
10%	500	12	1	0.527	0.530	0.047	0.049	–
10%	5000	2	0	0.353	0.354	0.011	0.010	0.048
10%	5000	2	1	0.353	0.353	0.018	0.017	–
10%	5000	12	0	0.527	0.528	0.011	0.010	0.046
10%	5000	12	1	0.527	0.527	0.018	0.018	–
30%	500	2	0	0.353	0.353	0.033	0.034	0.040
30%	500	2	1	0.353	0.356	0.055	0.054	–
30%	500	12	0	0.527	0.526	0.032	0.035	0.046
30%	500	12	1	0.527	0.529	0.053	0.051	–
30%	5000	2	0	0.353	0.352	0.012	0.013	0.054
30%	5000	2	1	0.353	0.353	0.021	0.020	–
30%	5000	12	0	0.527	0.525	0.012	0.012	0.054
30%	5000	12	1	0.527	0.527	0.020	0.020	–
50%	500	2	0	0.353	0.354	0.043	0.041	0.046
50%	500	2	1	0.353	0.356	0.069	0.068	–
50%	500	12	0	0.527	0.528	0.038	0.037	0.040
50%	500	12	1	0.527	0.529	0.062	0.061	–
50%	5000	2	0	0.353	0.354	0.017	0.016	0.036
50%	5000	2	1	0.353	0.354	0.027	0.027	–
50%	5000	12	0	0.527	0.527	0.014	0.014	0.054
50%	5000	12	1	0.527	0.527	0.023	0.023	–

Table 2.2: Results for setting where X has no effect on the NBS. Columns represent the probability of being censored, sample size, WTP (λ), value of X, true NBS ($\theta(\lambda|X)$), mean point estimate (Est.), mean estimated standard error ($\widehat{\text{SE}}$), empirical standard error (ESE), and proportion of simulations in which the null hypothesis of no effect of X is rejected (Pr. Reject).

2.4.1. Sensitivity to unmeasured confounding

In settings where we estimate NBS using data from observational databases, we must consider the potential effects of unmeasured confounding. Because the Monte Carlo procedure described in Section 2.3.2 only ensures that INBs are sampled from populations with the same underlying distribution of measured covariates, it is necessary to understand how estimates of NBS may be affected by unmeasured confounding. We perform simulations to determine how sensitive the proposed regression estimator of NBS is to unmeasured confounding. We simulate two standard normal variates, U_1 and U_2. The confounder U_1 influences treatment probability and survival time, while U_2 affects treatment and cost. We consider two settings, (1) where only the exposure-survival relationship is confounded and (2) where only the exposure-cost relationship is confounded. Variables L, X, and C are simulated as in previous simulations, $A \sim$ Bernoulli (expit$(L + \gamma_1 U_1 + \gamma_2 U_2)$), $T \sim$ Weibull$(k = 2, \lambda =$ exp$(4.05 + 0.15A + 0.2L - \eta_1 U_1 + 0.1X + 0.5A \times X))$, and $Y \sim$ Lognormal$(\mu = 4.2 + 0.002T + 0.5A + \eta_2 U_2, \sigma^2 = 0.16)$. The parameters γ_1, γ_2, η_1 and η_2 control the strength of unmeasured confounding. In simulations where the exposure-survival relationship is confounded, we examine the effect of low ($\gamma_1 = 0.5$, $\eta_1 = 0.05$), medium ($\gamma_1 = 0.75$, $\eta_1 = 0.15$), and high ($\gamma_1 = 1$, $\eta_1 = 0.3$) levels of confounding. The parameters γ_2 and η_2 equal zero in this setting. When the exposure-cost relationship is confounded, $\gamma_1 = \eta_1 = 0$. Low, medium, and high levels of confounding are simulated by setting ($\gamma_2 = 0.5$, $\eta_2 = 0.17$), ($\gamma_2 = 0.75$, $\eta_2 = 0.5$), and ($\gamma_2 = 1$, $\eta_2 = 1$) respectively

We simulate 1000 datasets with 5000 observations and a 30% censoring rate. We do not account for U_1 or U_2 when estimating NBS using the Monte Carlo standardization procedure. Simulation results for these simulations are provided in Table 2.3. As the level of confounding increases, the mean estimate of the NBS becomes more biased. When confounding is strongest, percent bias exceeds 10%. Because cost is given less weight in INB for large WTP, the bias from unmeasured exposure-cost confounding decreases as WTP increases. Similarly, bias increases with WTP when there is unmeasured confounding of the exposure-survival relationship. Bootstrap estimated standard errors are able to capture the true variability in estimates of NBS regardless of the strength of confounding.

			Survival Confounded				Cost Confounded			
Conf.	λ	X	$\theta(\lambda\|X)$	Est.	$\widehat{\text{SE}}$	ESE	$\theta(\lambda\|X)$	Est.	$\widehat{\text{SE}}$	ESE
low	2	0	0.354	0.346	0.012	0.012	0.358	0.329	0.011	0.012
low	2	1	0.588	0.578	0.024	0.023	0.586	0.559	0.024	0.024
low	12	0	0.527	0.516	0.011	0.012	0.527	0.519	0.012	0.012
low	12	1	0.746	0.735	0.018	0.017	0.745	0.740	0.017	0.017
med.	2	0	0.356	0.323	0.012	0.011	0.384	0.285	0.011	0.011
med.	2	1	0.585	0.546	0.024	0.023	0.571	0.475	0.026	0.026
med.	12	0	0.526	0.482	0.012	0.012	0.522	0.490	0.012	0.012
med.	12	1	0.739	0.700	0.019	0.020	0.741	0.718	0.018	0.018
high	2	0	0.362	0.287	0.011	0.011	0.423	0.251	0.011	0.011
high	2	1	0.576	0.488	0.024	0.024	0.556	0.386	0.027	0.027
high	12	0	0.523	0.425	0.012	0.012	0.512	0.427	0.012	0.012
high	12	1	0.722	0.631	0.020	0.020	0.720	0.651	0.020	0.021

Table 2.3: Results concerning sensitivity to unmeasured confounding of the exposure-survival and the exposure-cost relationship. Columns represent the WTP (λ), value of X, true NBS ($\theta(\lambda|X)$), mean point estimate (Est.), mean estimated standard error ($\widehat{\text{SE}}$), and empirical standard error (ESE).

2.5. Endometrial Cancer Study

The standard treatment for early-stage endometrial cancer is complete hysterectomy. Patients sometimes receive additional adjuvant radiation or chemotherapy to decrease the risk of recurrence (Latif et al., 2014). However, there is insufficient evidence that adjuvant radiation therapy increases overall survival (Shaeffer and Randall, 2005). Because more than half of the cost associated with the treatment of endometrial cancer is accrued in the period after initial treatment (Mariotto et al., 2011), it is important to ensure that only those patients who are likely to benefit from adjuvant therapies receive additional treatment. We illustrate how to determine the cost-effectiveness of adjuvant radiation and chemotherapy using empirically-based synthetic data developed by Spieker et al. (2021). This dataset was developed based on data from a large observational endometrial cancer database. Followup for patients in the original database is insufficient to perform a meaningful cost-effectiveness analysis. Covariates in the synthetic data are drawn from their empirical distribution in the cancer database. Cost and survival outcomes are simulated based on relations observed in the database and from literature surrounding the prognoses of stage I and stage II endometrial cancer (Shaeffer and Randall, 2005; Spieker et al., 2020a; Susumu et al., 2008). The data have covariate, treatment, and outcome distributions similar to what would be observed in practice. The use of synthetic data is motivated by a desire to make the code and data for this analysis available

as well as to ensure sufficient followup.

We consider three possible courses of treatment post-hysterectomy: (1) receipt of adjuvant radiation therapy (RT), (2) receipt of adjuvant chemotherapy (CT), and (3) receipt of neither therapy, hereafter referred to as control. In our first analysis we compare those receiving RT to those receiving control, and in our second we compare CT to control. Because the data are simulated, the results of our analysis are to be used solely for illustrating how NBS regression can be used in a practical setting. This work does not make use of data from human subjects and does not require IRB approval.

The data contain monthly follow-up information on $N = 13526$ endometrial cancer patients. The upper bound on follow-up time is ten years, at which point all observations are administratively censored. Our goal is to evaluate the association between two measures of patient health, cancer stage and Charlson comorbidity index (Charlson et al., 1994), and the cost-effectiveness of adjuvant RT or CT compared to control. We define cost-effectiveness using medical costs in USD($) and survival time in years. We assume that age at diagnosis, cancer stage, Charlson index, baseline receipt of RT or CT, and number of hospitalizations in the first-month post-surgery are confounders of the exposure-outcome relationship. Treatment status is defined over the period from two-to-four months post-hysterectomy. For example, if a patient received adjuvant RT during this period, then they are included in the RT group in our analysis regardless of their treatment status throughout the remainder of the study period. If a patient received both RT and CT during this period, they are included in both treatment groups. The mean age at diagnosis is 73.70 (SD = 6.59), and most patients have stage I cancer (93.6%). Charlson comorbidity indices are between zero and five, with 51.32% of patients having an index of zero. In the first month post-hysterectomy, the baseline period, 74 (0.55%) patients received RT and 810 (5.99%) received CT. From two-to-four months post-surgery, these figures were 3440 (25.43%) and 305 (2.25%), respectively.

2.5.1. Analysis

We wish to estimate NBS comparing each adjuvant therapy to control conditional on cancer stage and Charlson comorbidity index. We consider a range of WTP (from $\$50k$ to $\$120k$ per year),cancer stages (stages I and II), and categorized Charlson indices (0, 1, and $2+$). Because baseline covariates affect both treatment status and patient outcomes, we apply the standardization procedure

Treatment	λ	Charlson(0)	Charlson(1)	Charlson(2+)
Stage I				
RT	50	0.54 (0.51, 0.55)	0.52 (0.50, 0.54)	0.51 (0.48, 0.53)
	120	0.55 (0.52, 0.62)	0.53 (0.52, 0.56)	0.53 (0.50, 0.55)
CT	50	0.63 (0.56, 0.68)	0.60 (0.53, 0.66)	0.56 (0.49, 0.63)
	120	0.67 (0.60, 0.71)	0.64 (0.58, 0.70)	0.60 (0.54, 0.67)
Stage II				
RT	50	0.59 (0.54, 0.64)	0.57 (0.54, 0.63)	0.56 (0.52, 0.62)
	120	0.62 (0.56, 0.65)	0.60 (0.56, 0.65)	0.60 (0.54, 0.64)
CT	50	0.58 (0.37, 0.69)	0.54 (0.34, 0.67)	0.49 (0.30, 0.63)
	120	0.64 (0.45, 0.74)	0.61 (0.43, 0.72)	0.58 (0.40, 0.70))

Table 2.4: Estimated NBS comparing RT to control and CT to control conditional on Charlson index and cancer stage. Non-parametric bootstrap-based 95% CIs for each estimated NBS are provided in parentheses.

described in Section 2.3.2. We assume survival time can be modelled using a Weibull distribution. To account for structural zeros in cost data, we use a two-stage zero-inflated cost model where the probability of having zero costs is estimated using a logistic model and the distribution of nonzero costs is fit using a log-normal model. All models are fit conditional on treatment status and confounders. For each treatment, we draw $M = 10000$ Monte Carlo observations from the fit models for survival and medical costs and use a probit model when finding the regression estimator of the NBS. For our analysis of RT versus control, we model the NBS using the form:

$$\theta(\lambda|\text{Charlson, Stage}) = \int_0^1 \Phi\Big(\widehat{\beta}_0^R + \widehat{\beta}_1^R \text{ Stage II } + \widehat{\beta}_2^R \text{ Charlson(1)} + \widehat{\beta}_3^R \text{ Charlson(2+) } + \widehat{\beta}_4^R \Phi^{-1}(\omega)\Big) d\omega$$

The variables Charlson(1) and Charlson(2+) are indicators of whether a patient's Charlson score is equal to one or at least two, respectively. Analogous models are used when estimating the NBS comparing CT to control conditional on cancer stage and Charlson index. Parameters for the regression model comparing CT to control are denoted by the superscript C. Hypothesis testing is based on $B = 1000$ nonparametric bootstrap replicates.

2.5.2. Results

In Table 2.4, we provide estimates of the NBS for each of the treatments as a function of Charlson index and cancer stage. In almost every setting, the estimated NBS is greater than 0.5, indicating

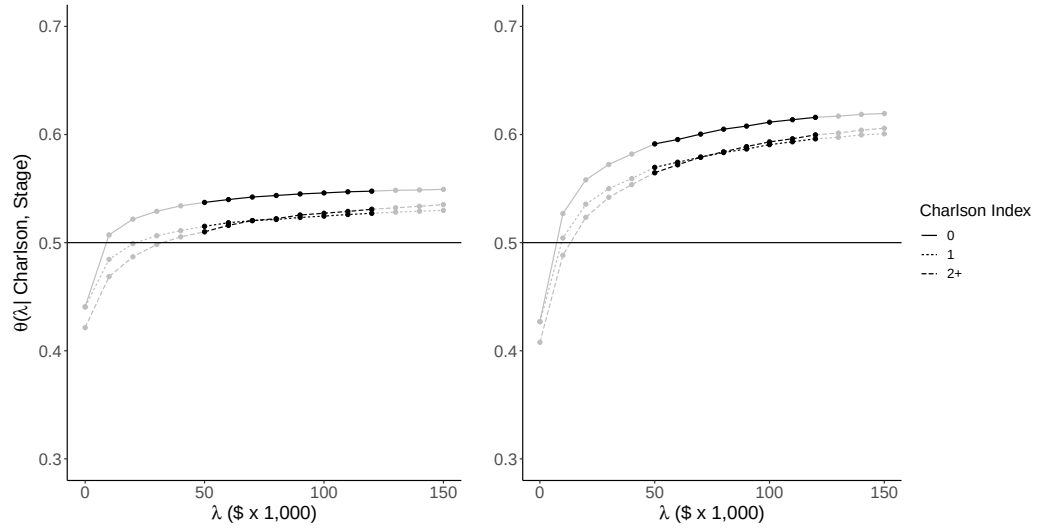

Figure 2.1: Estimates of the NBS comparing RT to control as a function of the WTP, λ. CED curves are provided for patient cohorts conditional on cancer stage and Charlson comorbidity index. Results for patients with stage I cancer are provided on the left, while those for patients with stage II cancer are on the right. Estimated NBS within the range of primary interest are denoted in black. Gray indicates estimated NBS outside of this region, provided to observe the behavior of the NBS.

that patients are likely to benefit from treatment regardless of their cancer stage or number of comorbidities at diagnosis. Because there are few subjects who have stage II cancer and receive adjuvant chemotherapy ($n = 45$), estimates of the NBS within this cohort too uncertain to claim cost-effectiveness. For WTP of $\$50k$, the estimated NBS comparing RT to control among patients with stage I cancer and a comorbidity index of zero is 0.54 with 95% CI $(0.51, 0.55)$. We conclude that a randomly selected patient receiving RT will have a greater INB than a patient receiving control 54% of the time. Adjuvant radiation is cost-effective for this patient cohort. The 95% CI for $\theta(\lambda = 50|$Stage I, Charlson(2+)$)$ comparing RT to control is $(0.48, 0.53)$. Thus, there is inconclusive evidence on the cost-effectiveness of RT compared to control for patients with stage I cancer and multiple comorbidities. Figures 2.1 and 2.2 provide cost-effectiveness determination curves (CED) comparing adjuvant radiation to control and adjuvant chemotherapy to control, respectively. The CED is defined by Spieker et al. (2021) as a graphical tool which shows how NBS is affected by changes in WTP. The bold segment of the plots represent estimated NBS at WTP between $\$50$k and $\$120$k, the range of primary interest (Shiroiwa et al., 2010). Estimates of NBS outside of this range are in gray to illustrate the behavior of NBS at extreme WTP.

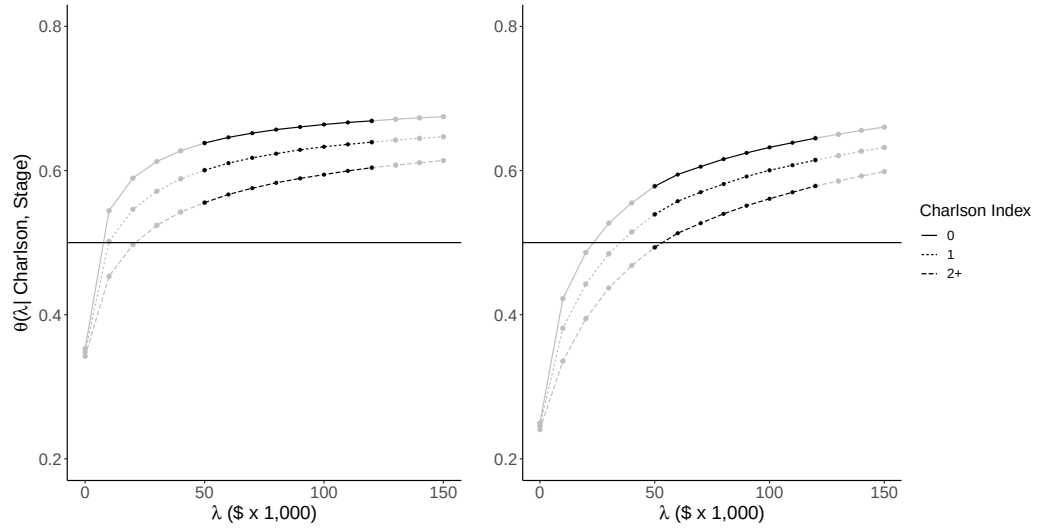

Figure 2.2: Estimates of the NBS comparing CT to control as a function of the WTP, λ. CED curves are provided for patient cohorts conditional on cancer stage and Charlson comorbidity index. Results for patients with stage I cancer are provided on the left, while those for patients with stage II cancer are on the right. Estimated NBS within the range of primary interest are denoted in black. Gray indicates estimated NBS outside of this region, provided to observe the behavior of the NBS.

For each treatment and WTP, we conduct a hypothesis test to evaluate if there is an association between either cancer stage or Charlson comorbidity index and NBS. As per Section 2.3.1, we can determine whether cancer stage is associated with NBS after controlling for comorbidities by testing the null hypotheses $H_0^R : \beta_1^R = 0$ and $H_0^C : \beta_1^C = 0$ against their alternatives. At $\lambda = \$50$k, the p-values for these tests are 0.019 and 0.366, respectively. There is evidence that, after adjusting for Charlson comorbidity indices, the cost-effectiveness of adjuvant radiation may depend on cancer stage. However, there is insufficient evidence to say same for adjuvant chemotherapy at this WTP. At a WTP of $\$120$k, the p-values for the tests of the same hypotheses are 0.007 and 0.543. Again, there is evidence that cancer stage is associated with the cost-effectiveness of adjuvant radiation therapy but not of chemotherapy. Among patients with a Charlson comorbidity score of zero, the estimated NBS at a WTP of $\$120$k is 0.55 for patients with stage I cancer and 0.62 for those with stage II cancer. This indicates that patients with more advanced cancers may benefit more from receiving adjuvant radiation therapy than those with stage I cancer. A similar conclusion is drawn at a WTP of $\$50$k.

To test whether patient's comorbidity indices are associated with NBS after accounting for cancer

stage, we perform a test of the hypotheses $H_0^R : \beta_2^R = \beta_3^R = 0$ and $H_0^C : \beta_2^C = \beta_3^C = 0$ against their respective alternatives. At a WTP of $\$50$k, the p-value for the test of parameters in the adjuvant radiation model is 0.022, while that for chemotherapy less than 0.005. At a WTP of $\$120$k, the test for association between the Charlson comorbidity index and the cost-effectiveness of adjuvant radiation results in a p-value of 0.076. At the same WTP the test for adjuvant chemotherapy, again, has a p-value less than 0.005. At both considered WTP, there is evidence that Charlson comorbidity score can help determine which patients may most benefit from adjuvant chemotherapy. However, while there is evidence that the cost-effectiveness of adjuvant radiation is affected by Charlson comorbidity index at a WTP of $\$50$k, there is insufficient evidence at the larger WTP. Because medical costs are given greater weight for lower WTP, this suggests that comorbidities may have a greater effect on medical costs than on patient survival. For patients being treated with adjuvant chemotherapy, point estimates of the NBS in Table 2.4 indicate that patients with greater Charlson comorbidity indices may not receive as much benefit as those with fewer comorbidities, regardless of cancer stage.

2.6. Discussion

In this paper, we introduce a regression framework for NBS that allows us to explain variability in cost-effectiveness arising from covariates. The proposed method is the first to enable estimation of NBS within levels of measured covariates and allow for testing the effects of covariates on NBS. Regression of NBS is done in three steps: (1) estimate the distribution of cost and effectiveness outcomes conditional on treatment and measured covariates, (2) implement Monte Carlo standardization to sample INBs from treated and untreated populations with the same underlying distribution of confounders, and (3) apply ROC regression techniques developed by Alonzo and Pepe (2002) to estimate NBS. Our proposed methodology generalizes the work of Spieker et al. (2021) on estimating NBS to allow for covariate adjustment and hypothesis testing. Understanding how patient characteristics influence cost-effectiveness can help policy makers allocate resources towards groups which are most likely to benefit from treatment.

In cost-effectiveness analyses where patients are lost to follow-up, estimators that assume independence between costs at time of censoring and those at time of death are known to be inconsistent (Lin et al., 1997). We provide methodology for incorporating IPCW in models for the conditional dis-

tribution of treatment costs. When the probability of censoring depends on measured covariates, we describe how Cox proportional hazards models can be used to determine covariate-specific IPCW. In simulations, we find that our method is able to attain low bias and adequate power in settings with small sample size and up to 50% censoring. Moreover, estimates of standard error obtained using the nonparametric bootstrap are representative of empirical standard errors across various levels of censoring.

Our proposed estimation procedure assumes that models for the conditional distributions of cost and effectiveness are properly specified. Flexible modeling techniques present a potential route to reduce the risk of model misspecification. Bayesian nonparametric methods for zero-inflated data have been developed by Oganisian, Mitra, and Roy (2018) to estimate cost distributions, but extensions of this methodology to cost-effectiveness analyses have not been explored. Similarly, ensemble learning algorithms such as Super Learner (Laan, Polley, and Hubbard, 2007) could be used to model complex cost and effectiveness distributions. For example, in their doubly robust estimator of NMB, Li et al. (2018) successfully use Super Learner to estimate both propensity score and outcome models. In future work, it would be useful to determine the best ways to incorporate these methods into regression of NBS. After the standardization step, parametric regression models for NBS may also be misspecified. Prior work by Hanley (1988) has examined the robustness of the binormal form for modeling ROC curves. Their results suggest that estimates of AUC arising from the binormal form, and analogously NBS, may exhibit low bias even when the form is misspecified. Additionally, because the underlying distribution of INBs are left unspecified, the proposed NBS regression framework can be described as "distribution-free" (Alonzo and Pepe, 2002). Optimally efficient selection of Ω is an open area of research. In their paper concerning ROC regression, Alonzo and Pepe (2002) perform simulations illustrating close to maximal efficiency can be achieved even if the number of selected quantiles in Ω is 10% of the maximum length (i.e. the number of unique values among untreated units). Further work that explores how to tune selection of Ω to the distribution of INBs may improve efficiency.

The methods proposed in this article are suited for settings where all confounding variables are measured, such as in cost-effectiveness analyses run in conjunction with randomized clinical trials or for certain covariate rich observational databases. When unmeasured confounding is present, estimates of NBS may be affected. In simulations with low levels of unmeasured confounding, our

regression estimator of NBS exhibits small to moderate bias (between 2% and 8% bias). As expected, we found that as the strength of confounding increases, estimates of NBS become more biased. In any studies involving observational databases, researchers should consider potential sources of unmeasured confounding. Methods to assess sensitivity of NBS to unmeasured confounding may also be useful and will be a subject of future work.

Note that NBS is defined in terms of baseline treatment status. In our endometrial cancer example, the comparison groups are those who received some adjuvant therapy in the period from two to four months after surgery and those who did not. Because treatment effects may depend on the duration or frequency of treatment, conclusions concerning cost-effectiveness may change depending on whether treatment is considered time-stable, as in our definition, or time-dependent. To estimate cumulative medical costs under a time-dependent treatment strategy, Spieker, Roy, and Mitra (2018) developed a nested g-computation procedure. Similar methodologies that can incorporate time-dependent treatment may be useful when we are interested in comparing the cost-effectiveness of multiple treatment strategies and are the subject of future work.

Because patient covariates may influence the cost-effectiveness of a treatment, our proposed regression framework for NBS provides a useful approach to identify which groups may benefit the most from treatment. The results from our synthetic data analysis illustrate how we can evaluate the effects of cancer stage or number of comorbidities on the cost-effectiveness of adjuvant therapies for endometrial cancer patients. This information may be useful to policy makers aiming to better allocate resources.

CHAPTER 3

IDENTIFYING OPTIMALLY COST-EFFECTIVE DYNAMIC TREATMENT REGIMES WITH A Q-LEARNING APPROACH

3.1. Introduction

Due to rising healthcare costs, there has been growing interest in improved methods for cost-effectiveness analyses. Cost-effectiveness research is concerned with identifying policies that can balance treatment effectiveness and overall costs. Adjuvant radiation or chemotherapy is occasionally recommended to endometrial cancer patients undergoing hysterectomy because these therapies can reduce the risk of locoregional recurrence (Heerik et al., 2021). However, these therapies are also associated with higher toxicity rates and greater treatment costs (Randall et al., 2006). Identifying treatment sequences that are both effective and cost-efficient is complicated by heterogeneous treatment responses across individuals. Patients at high-risk of cancer recurrence are likely to benefit more from adjuvant therapies than are patients at low-risk of recurrence (Heerik et al., 2021). These complications underscore the need for treatment regime methodology that can optimally allocate limited resources towards patients that are most likely to benefit.

Traditionally, regression approaches have been used to address treatment response heterogeneity and its effects on cost-effectiveness. Willan, Briggs, and Hoch (2004) show that if cost and effectiveness metrics are linear functions of treatment status and covariates, then the incremental cost-effectiveness ratio (ICER), one measure of cost-effectiveness, can be represented as a function of the parameters in a least-squares regression and standard inferential procedures can be used to assess cost-effectiveness within subgroups. Alternatively, Nixon and Thompson (2005) develop Bayesian Markov Chain Monte Carlo methods for inference that are more robust to skewed costs. While these methods can describe cost-effectiveness within subgroups, neither allows for estimation of an optimal treatment strategy. Xu et al. (2020) developed methodology to learn optimal cost-effective treatment rules by maximizing a composite measure of cost-effectiveness, the individual net benefit. For a given willingness-to-pay, this method targets the true optimally cost-effective regime. However, because estimates of the expected cost and effectiveness under this regime can not be determined, its use is difficult in policy recommendation settings. Laber et al. (2018)

provides policy search methodology to identify a maximally effective treatment regime from among regimes with a pre-determined threshold on overall cost (defined by number of safety events). This approach allows for direct estimation of the cost and effectiveness of a regime fit under a given cost constraint. However, the decision rules within these regimes take the form of linear decision boundaries and may be difficult to interpret. Recent work by Zhang et al. (2018) argues that treatment rules based on decision lists provide both flexibility and interpretability.

In this paper, we propose an efficient algorithm for identifying optimal list-based decision rules under fixed cost constraints. We use Q-learning with policy search methodology to maximize treatment effectiveness within the class of list-based regimes with cost constraints. We modify the algorithm proposed by Zhang et al. (2018) to efficiently estimate our list-based decision rules. Within our data example, we propose an iterative cost-effectiveness analysis that uses ICER to select an optimally cost-effective regime for the adjuvant treatment of endometrial cancer from a set of candidate regimes corresponding to different cost-constraints.

The remainder of this manuscript is organized as follows. In Section 3.2 we introduce methodology for estimating optimal list-based treatment regimes under a constraint on overall expected treatment costs. We explore the operating characteristics of our proposed regime identification approach through simulations in Section 3.3. In Section 3.4 we perform a cost-effectiveness analysis to identify an optimal cost-effective treatment strategy for assigning endometrial cancer patients to adjuvant radiation and chemotherapy. We describe additional considerations and avenues for future developments in Section 3.5.

3.2. Methods

3.2.1. Set-up and notation

Assume that treatment decisions are made at the beginning of K distinct intervals indexed by $k = 1, ..., K$. For each individual $i = 1, ..., n$ in interval k, we observe $(W_{ik}, A_{ik}, C_{ik}, Z_{ik}, Y_{ik})$. Here, W_k denotes confounding variables collected at the beginning of interval k, A_k denotes treatment status, C_k is a censoring indicator, and Z_k and Y_k are clinical effectiveness and cost outcomes collected at the end of interval k. Within each interval, variables are observed in the order: $W_k \longrightarrow A_k \longrightarrow C_k \longrightarrow (Z_k, Y_k)$. We use overbar notation to denote covariate history, e.g.

$\overline{A}_k = (A_1, A_2, ..., A_{k-1}, A_k)$, and underbar notation to denote future values of a covariate, e.g. $\underline{A}_k = (A_{k+1}, A_{k+2}, ..., A_K)$. Let $H_k = (\overline{W}_k, \overline{A}_{k-1}, \overline{Z}_{k-1}, \overline{Y}_{k-1})$ denote a patient's covariate and treatment history at prior to decision k.

A dynamic treatment regime is a set of decision rules $d = \{d_1, ..., d_K\}$ where d_k is a mapping $d_k : \mathcal{H}_k \longrightarrow \mathcal{A}_k$ from the space of all possible covariate histories into that of treatment decisions. Under a potential outcomes framework, if $\mathcal{D}$ is a class of treatment regimes, then the optimal dynamic treatment regime within this class, d^{opt}, satisfies the condition $\mathbb{E}[Z_K^{d^{\text{opt}}, c_K=0}] \geq \mathbb{E}[Z_K^{d, c_K=0}]$ for all $d \in \mathcal{D}$. Because a common goal in health policy is making treatment decisions within resource-limited settings, we restrict consideration to the class of list-based treatment rules with $\mathbb{E}[Y_K^{d, c_K=0}] < \tau$ for a predetermined cost-constraint, τ. Note that τ constrains the expected cost for the entire population. Individual costs under this regime may exceed this constraint. This restricted search sacrifices regime flexibility in favor of interpretability. Following Zhang et al. (2018), a list-based regime is a regime for which each treatment rule d_k for $k = 1, ..., K$ consists of a series of if-else statements:

If $H_k \in R_{k1}$ then $A_k = a_{k1}$;

else if $H_k \in R_{k2}$ then $A_k = a_{k2}$;

$\vdots$

else if $H_k \in R_{kL_k}$ then $A_k = a_{kL_k}$

Here, L_k is the preset maximal list length for a decision at interval k, and R_{kl} is a subset of $\mathcal{H}_k$ for $l = 1, ..., L_k$. The maximal list length should be chosen to balance flexibility and interpretability, with longer lists being more flexible but more difficult to interpret. For simplicity and to aid interpretability, we restrict R_{kl} to clauses involving thresholding of a single covariate (e.g. $R_{kl} = \{h_k \in \mathcal{H}_k : h_{kj} \leq \theta\}$ for $1 \leq j \leq \dim(H_k),\ \theta \in \mathbb{R}$).

Estimation of the optimal cost-restricted regime involves identifying optimal values for $\{(R_{kl}, a_{kl}) : l = 1, ..., L_k\}_{k=1}^K$. We employ an integrated Q-learning and policy search approach to define and

estimate these values. Define the K^{th} stage Q-functions for Z and Y:

$$Q_K^Z(a_K, h_K) = \mathbb{E}[Z_K | A_K = a_K, H_K = h_K, C_K = 0]$$
$$Q_K^Y(a_K, h_K) = \mathbb{E}[Y_K | A_K = a_K, H_K = h_K, C_K = 0]$$

For a treatment regime d, we may recursively define Q-functions for intervals $k = K-1, ..., 1$:

$$Q_k^Z(a_k, h_k; d) = \mathbb{E}[Z_k + Q_{k+1}^Z(d_{k+1}(H_{k+1}), H_{k+1}) | A_k = a_k, H_k = h_k, C_k = 0],$$
$$Q_k^Y(a_k, h_k; d) = \mathbb{E}[Y_k + Q_{k+1}^Y(d_{k+1}(H_{k+1}), H_{k+1}) | A_k = a_k, H_k = h_k, C_k = 0].$$

By the principles of dynamic programming, the optimal treatment regime can be identified by optimizing over the Q-functions at each individual decision point (Bellman, 1966). In the policy search context, the process of finding an optimal regime within a pre-specified class reduces to finding the optimal decision rule within this class at each decision point. Because the k^{th} stage Q-function for cost denotes the expected cost accrued in the k^{th} interval and in future intervals under a pre-specified regime, decompose the overall cost constraint τ into K components $\tau_1, ..., \tau_K$ representing the interval specific cost constraints where $\sum_{k=1}^K \tau_k = \tau$. The optimal choices of (R_K, a_K) within the class of list-based and cost constrained decision rules are given by:

$$\{R_K^{\text{opt}}, a_K^{\text{opt}}\} = \underset{R_K, a_K}{\arg\max}\, \mathbb{E}\left[Q_K^Z\left(d_K(H_k; R_K, a_K), H_K\right)\right]$$
$$\text{subject to } \mathbb{E}\left[Q_K^Y(d(H_K; R_K, a_K), H_K)\right] < \tau_K.$$

For decision intervals $k = K-1, ..., 1$ the optimal choices (R_k, a_k) are given by:

$$\{R_k^{\text{opt}}, a_k^{\text{opt}}\} = \underset{R_k, a_k}{\arg\max}\, \mathbb{E}\left[Q_k^Z\left(d_K(h_k; R_k, a_k), h_k; \underline{R}_k^{\text{opt}}, \underline{a}_k^{\text{opt}}\right)\right]$$
$$\text{subject to } \mathbb{E}\left[Q_k^Y(d(H_k; R_k, a_k), H_k; ; \underline{R}_k^{\text{opt}}, \underline{a}_k^{\text{opt}})\right] < \sum_{j=k}^K \tau_k.$$

To connect the distribution of the observed covariates to that of the potential outcomes we invoke the following identification assumptions:

(A1) (Positivity) If $P(H_k = h_k) > 0$, then $P(A_k = a_k | H_k = h_k) > 0$ for all a_k

(A2) (Consistency) $Z_k = Z_k^{\overline{A}_k}$ and $Y_k = Y_k^{\overline{A}_k}$ for $k = 1, ..., K$

(A3) (Sequentially ignorable treatment assignment) $(Z_k^{\overline{a}_k}, Y_k^{\overline{a}_k}, L_{k+1}^{\overline{a}_k}) \perp\!\!\!\perp A_k | H_k$ for $k = 1, ..., K$

(A4) (Sequentially ignorable censoring) $(Z_k^{\overline{a}_k}, Y_k^{\overline{a}_k}, L_{k+1}^{\overline{a}_k}) \perp\!\!\!\perp C_k | A_k, H_k$ for $k = 1, ..., K$

Under these assumptions, Schulte et al. (2014) and Laber et al. (2018) show:

$$Q_K^Z(a_K, h_K) = \mathbb{E}[Z_K^{\overline{a}_{K-1}, a_K, c_K=0} | H_K^{\overline{a}_{K-1}, a_K, c_{K-1}=0} = h_K]$$
$$Q_K^Y(a_K, h_K) = \mathbb{E}[Y_K^{a_{K-1}, a_K, c_K=0} | H_K^{\overline{a}_{K-1}, a_K, c_{K-1}=0} = h_K]$$

and, for $k = 1, ..., k-1$:

$$Q_k^Z(a_k, h_K; d) = \mathbb{E}\left[Z_k^{\overline{a}_{k-1}, a_k, c_k=0} + \sum_{j=k+1}^{K} Z_j^{\overline{a}_{k-1}, a_k, \underline{c}_k=0, \underline{d}_{k+1}} \middle| H_k^{\overline{a}_{k-1}, a_k, c_{k-1}=0} = h_k \right]$$
$$Q_k^Y(a_k, h_K; d) = \mathbb{E}\left[Y_k^{\overline{a}_{k-1}, a_k, c_k=0} + \sum_{j=k+1}^{K} Y_j^{\overline{a}_{k-1}, a_k, \underline{c}_k=0, \underline{d}_{k+1}} \middle| H_k^{\overline{a}_{k-1}, a_k, , c_{k-1}=0} = h_k \right]$$

It follows that identifying optimal choices of $\{(R_{kl}, a_{kl}) : l = 1, ..., L_k\}_{k=1}^K$ maximizes the potential treatment effectiveness under a constraint on potential cost.

In the next subsection, we propose a modification of the algorithm developed by Zhang et al. (2018) for identifying globally optimal list-based regimes. Our extension allows for regimes to be fit under a preset cost constraints.

3.2.2. Estimating the Optimal Decision Rules

To estimate the optimal list-based and cost constrained regime, we use a backwards recursive procedure. The optimal decision rule for the final interval is estimated first, and earlier timepoints are estimated assuming optimal decisions are made at all future timepoints. For each decision point k, we sequentially estimate the pairs (R_{kl}, a_{kl}) for clauses $l = 1, ..., L_k$.

We illustrate estimation of the final decision rule, d_K, before describing estimation for earlier intervals. For a unit entering the final interval with covariate history h_K, define the unconstrained optimal rule $\widetilde{d}_K(h_K) = \arg\max_{a_K} Q_K^Z(a_K, h_K)$. We want to approximate the optimal rule within the class of interpretable list-based and cost constrained treatment rules. To estimate the first clause

of the constrained rule, define the intermediate list-based decision rule:

If $h_K \in R_{K1}$ then $A_K = a_{K1}$;

else if $H_K \in \mathcal{H}_K$ then $A_K = \widetilde{d}_K(h_K)$

Given estimates of the K^{th} stage Q-functions, the estimated mean effectiveness and cost measures under this regime are given by:

$$\Psi^Z_{K1}(R_{K1}, a_{K1}) = \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{I}(h_{Ki} \in R_{K1})\widehat{Q}^Z_K(a_{K1}, h_{Ki}) + \mathbb{I}(h_{Ki} \notin R_{K1})\widehat{Q}^Z_K(\widetilde{d}_K(h_{Ki}), h_{Ki})\right]$$
$$\Psi^Y_{K1}(R_{K1}, a_{K1}) = \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{I}(h_{Ki} \in R_{K1})\widehat{Q}^Y_K(a_{K1}, h_{Ki}) + \mathbb{I}(h_{Ki} \notin R_{K1})\widehat{Q}^Y_K(\widetilde{d}_K(h_{Ki}), h_{Ki})\right].$$

We search for the values of (R_{K1}, a_{K1}) which maximize Ψ^Z_{K1} while ensuring that the expected cost, Ψ^Y_{K1}, is less than or equal to the cost constraint, τ_K. As Zhang et al. (2018) point out, maximizing the objective function, Ψ^Z_{K1}, is equivalent to minimizing the difference between this intermediate rule and the optimal rule. Imposing the constraint $\Psi^Y_{K1} < \tau_K$ ensures that we are minimizing within the class of cost constrained rules. Different choices of (R_{K1}, a_{K1}) may lead to equivalent expected effectiveness and cost estimates. For example, for units where the posited treatment, a_{K1}, is equal to the optimal treatment, $\tilde{d}_K(h_{Ki})$, inclusion in R_{K1} will not influence expected outcomes. In these cases, we opt to reward decision regions that assign treatment to a greater number of patients by adding a complexity term, η, to our objective function: $\Psi^Z_{K1}(R_{K1}, a_{K1}) + \eta\{\sum_{i=1}^{n}\mathbb{I}(h_{Ki} \in R_{K1})\}$. Because η may reward "larger" regions at the expense of mean effectiveness, cross-validation can be used to select this parameter. Ideally, η should be large enough to prevent searching over trivial clauses but small enough to retain maximal effectiveness. The optimal choices of (R_{K1}, a_{K1}) for the intermediate rule are given by:

$$(\widehat{R}_{K1}, \widehat{a}_{K1}) = \underset{R_{K1},\ a_{K1}}{\arg\max}\ \Psi^Z_{K1}(R_{K1}, a_{K1}) + \eta\left\{\sum_{i=1}^{n}\mathbb{I}(h_{Ki} \in R_{K1})\right\}$$
$$\text{subject to } \Psi^Y_{K1}(R_{K1}, a_{K1}) < \tau_K.$$

This procedure can be generalized to estimate optimal regions and treatment choices for each of the L_K clauses in the list-based rule. The algorithm can be summarized as follows:

1. Let $l = 1$.

2. Define $\widehat{G}_{Kl} = \mathcal{H}_K \setminus \left(\bigcup_{s<l} \widehat{R}_{Ks}\right)$. If $l = L_K$, force $R_{Kl} = \mathcal{H}_K$. The estimated mean of $X = Z$ or Y under the l^{th} intermediate rule is defined as:

$$\Psi^X_{Kl}(R_{Kl}, a_{Kl}) = \frac{1}{n}\sum_{i=1}^{n}\Big\{\mathbb{I}(h_{Ki} \in R_{Kl}, h_{Ki} \in \widehat{G}_{Kl})\widehat{Q}^X_K(a_{Kl}, h_{Ki}) + \mathbb{I}(h_{Ki} \notin R_{Kl}, h_{Ki} \in \widehat{G}_{Kl})\widehat{Q}^X_K(\widetilde{d}_K(h_{Ki}), h_{Ki}) + \sum_{j=1}^{l-1}\mathbb{I}(h_{Ki} \in \widehat{R}_{Kj}, h_{Ki} \notin \widehat{G}_{Kj})\widehat{Q}^X_K(a_{Kj}, h_{Ki})\Big\}. \tag{3.1}$$

3. Define $(\widehat{R}_{Kl}, \widehat{a}_{Kl})$ as:

$$(\widehat{R}_{Kl}, \widehat{a}_{Kl}) = \underset{R_{Kl}, a_{Kl}}{\arg\max}\ \Psi^Z_{Kl}(R_{Kl}, a_{Kl}) + \eta\left(\frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(h_{Ki} \in R_{Kl}, h_{Ki} \in \widehat{G}_{Kl})\right);$$
$$\text{subject to } \Psi^Y_{Kl}(R_{Kl}, a_{Kl}) < \tau_K.$$

   If $l < L_K$ set $l = l + 1$ and return to Step 2, otherwise stop.

The algorithm above can be used to estimate $\left\{(\widehat{R}_{Kl}, \widehat{a}_{Kl}) : l = 1, ..., L_K\right\}$. These estimates completely determine the optimal list-based and cost constrained decision rule for the K^{th} interval, $\widehat{d}_K$. To estimate treatment rules for decisions $k = 1, ..., K-1$, let $Q^Z_k(h_k, a_k; \widehat{d}\,)$ and $Q^Y_k(h_k, a_k; \widehat{d}\,)$ denote the Q-functions for interval k assuming that the optimal list-based and cost constrained rule is followed in future intervals. The unconstrained optimal treatment decision at interval k is $\widetilde{d}_k = \arg\max_{a_k} Q^Z_k(h_k, a_k; \widehat{d}\,)$. To approximate this rule within the class of list-based and cost constrained treatment rules, we may apply the steps listed above replacing the K^{th} stage Q-functions with their k^{th} stage equivalents, and the cost constraint τ_K with $\sum_{j=k}^{K}\tau_j$.

The proposed algorithm allows estimation of an optimal list-based regime under a set of interval-specific cost constraints. This work requires two important modifications of the algorithm described by Zhang et al. (2018) for identifying globally optimal list-based regimes. First, our procedure requires that the optimal values of (R_{kl}, a_{kl}) result in a rule which satisfies some pre-specified cost constraint. To achieve this, we must be able to estimate the expected cost accrued under each intermediate decision rule. This motivates our second departure from the original algorithm.

Equation 3.1 can be viewed as the sum of expected outcomes in three exhaustive groups: (1) those satisfying the current clause but not past clauses ($\mathbb{I}(h_{ki} \in R_{kl}, h_{ki} \in \widehat{G}_{kl}) = 1$), (2) those satisfying neither the current clause nor past clauses ($\mathbb{I}(h_{ki} \notin R_{kl}, h_{ki} \in \widehat{G}_{kl}) = 1$), and (3) those satisfying at least one previous clause ($\sum_{j=1}^{l-1} \mathbb{I}(h_{ki} \in R_{kj}, h_{ki} \notin \widehat{G}_{kj}) = 1$). In order to estimate the overall expected outcomes, we must consider the expected outcome in each of these groups. Because Zhang et al. (2018) are interested in finding an optimal list-based regime without a cost constraint, at each clause they need only find (R_{kl}, a_{kl}) that maximize effectiveness among unassigned units, i.e. those in groups (1) and (2). Because the outcomes for those in group (3) are fixed with respect to the choice of (R_{kl}, a_{kl}), these two methods will result in equivalent rules in cases where there is no cost constraint. This modification is necessary for ensuring the overall expected cost is below the cost constraint. A final important clarification concerns adjustments for censored individuals. While Q-functions for a given interval must be estimated using only those patients with observed outcomes, all patients who are uncensored at the beginning of a decision interval may be used to estimate optimal clauses within the decision list. This is true because the objective functions described in Equation 3.1 only require covariate histories, H_k, for each individual. This information is available for all units who have neither died nor been censored prior to interval k.

3.2.3. Efficient optimization of treatment rules

We propose an algorithm for identifying optimal choices of R_{kl} and a_{kl}, at clause l of decision rule k. Let $\{\widehat{R}_{kj}, \widehat{a}_{kj} : j = 1, ..., l-1\}$ denote estimated optimal regions and treatments for clauses prior to l. Recall that regions R_{kl} are defined in terms of a threshold, θ, on a variable within H_k. Let H_{kp}, for $1 \leq p \leq \dim(H_k)$, denote a covariate contained within H_k. First, consider decision regions of the form $R_{kl} = \{h_k : h_{kp} \leq \theta\}$ where patients with $h_k \in R_{kl}$ are given treatment a_{kl}. For $l < L_k$ the intermediate decision rule is given by:

If $h_k \in \widehat{R}_{k1}$ then $A_k = \widehat{a}_{k1}$;

else if $h_k \in \widehat{R}_{k2}$ then $A_k = \widehat{a}_{k2}$;

$\vdots$

else if $h_k \in \widehat{R}_{k,l-1}$ then $A_k = \widehat{a}_{k,l-1}$;

else if $h_{kp} \leq \theta$ then $A_k = a_{kl}$;

else $A_k = \widetilde{d}(h_k)$

For each unit, define:

$$U_i^Z = \mathbb{I}(h_i \in \widehat{G}_l)\widehat{Q}_k^Z(a_{kl}, h_i) + \sum_{j=1}^{l-1} \mathbb{I}(h_i \in \widehat{R}_j, h_i \notin \widehat{G}_j)\widehat{Q}_k^Z(a_{kj}, h_i)$$
$$V_i^Z = \mathbb{I}(h_i \in \widehat{G}_l)\widehat{Q}_k^Z(\widetilde{d}(h_i), h_i) + \sum_{j=1}^{l-1} \mathbb{I}(h_i \in \widehat{R}_j, h_i \notin \widehat{G}_j)\widehat{Q}_k^Z(a_{kj}, h_i)$$

The values U_i^Z and V_i^Z represent expected effectiveness outcomes for patients included in and excluded from R_{kl}, respectively. Note that if a patient satisfies a previous clause, then $U_i^Z = V_i^Z$ so that the patient's expected outcome is independent of the choice of θ and a_{kl}. It can be seen that $\Psi_{kl}^Z(R_{kl}, a_{kl}) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(h_{kpi} \leq \theta)U_i^Z + \mathbb{I}(h_{kpi} > \theta)V_i^Z$. If we define U_i^Y and V_i^Y similarly, it follows that $\Psi_{kl}^Y(R_{kl}, a_{kl}) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(h_{kpi} \leq \theta)U_i^Y + \mathbb{I}(h_{kpi} > \theta)V_i^Y$.

Although θ may take any real value, the unique values of Ψ_{kl}^Z and Ψ_{kl}^Y can be obtained by setting θ to the order statistics of h_{kp}. Let $\theta_1^*, ..., \theta_m^*$ denote the unique values of H_{kp}, where $m \leq n$. The regimes using θ_j^* and θ_{j-1}^* differ only in the treatments received by previously unassigned units with $h_{kpi} = \theta_j^*$. Under the first regime these units receive treatment a_{kj}, while under the second they receive $\widetilde{d}(h_i)$. It can be seen that Ψ^Z and Ψ^Y follow the recursive relationships:

$$\Psi_{kl}^Z\big(\theta = \theta_j^*, a_{kl})\big) = \Psi_{kl}^Z\big(\theta = \theta_{j-1}^*, a_{kl}\big) + \sum_{i=1}^n \mathbb{I}(h_{kpi} = \theta_j^*)(U_i^Z - V_i^Z)$$
$$\Psi_{kl}^Y\big(\theta = \theta_j^*, a_{kl})\big) = \Psi_{kl}^Y\big(\theta = \theta_{j-1}^*, a_{kl}\big) + \sum_{i=1}^n \mathbb{I}(h_{kpi} = \theta_j^*)(U_i^Y - V_i^Y).$$

These relationships allow us to quickly enumerate all possible values of Ψ^Z and Ψ^Y for regions of the form $R_{kl} = \{h_k : h_{kp} \leq \theta\}$ and assigned treatment a_{kl}. Analogous relationships can be defined for regions of the form $R_{kl} = \{h_k : h_{kp} > \theta\}$. Optimal choices for R_{kl} and a_{kl} can be obtained by iterating over each variable in H_k and each possible treatment option, and identifying which choices lead to maximal values of $\Psi^Z(R_{kl}, a_{kl}) + \eta\sum_{i=1}^n \mathbb{I}(H_{ki} \in R_{kl})$ under the restriction $\Psi_{R_{kl}, a_{kl}}^Y < \sum_{j=k}^K \tau_k$. As a result of this recursive relationship, we may enumerate all possible values of Ψ^Z and Ψ^Y corresponding to a specific treatment a_{kj} and regions of the form $R_{kl} = \{h_k : h_{kp} \leq \theta\}$ in $O(n)$ time. Thus, if there are p candidate covariates and s different treatment options, then each clause can be optimized in $O(nps)$ time.

Table 3.1: Mean value of Monte Carlo estimates of the mean survival (MC. Surv.) and cost (MC. Cost), and mean estimated survival ($\widehat{\text{Surv.}}$) and cost ($\widehat{\text{Cost}}$) arising from averaging over the Q-functions. Results are provided for simulations with 30% censoring and with low, medium, and high levels of correlation and sample sizes (n).

			Linear Model				SuperLearner			
Corr.	τ	n	$\widehat{\text{Cost}}$	$\widehat{\text{Surv.}}$	MC. Cost	MC. Surv.	$\widehat{\text{Cost}}$	$\widehat{\text{Surv.}}$	MC. Cost	MC. Surv.
Low	26	500	25.73	1.46	25.25	1.42	25.23	1.62	25.57	1.62
		1000	25.81	1.43	24.15	1.43	25.36	1.65	25.80	1.70
		5000	25.98	1.39	23.70	1.42	25.90	1.70	26.21	1.77
	28	500	27.69	1.50	26.89	1.47	27.78	1.65	27.11	1.67
		1000	27.79	1.46	25.70	1.46	27.82	1.70	27.36	1.76
		5000	28.02	1.42	25.72	1.44	27.87	1.76	28.05	1.84
	30	500	29.71	1.54	28.92	1.50	29.42	1.69	28.69	1.72
		1000	29.86	1.49	27.76	1.48	29.79	1.74	29.24	1.80
		5000	29.99	1.45	27.52	1.47	29.83	1.80	29.89	1.88
Med.	26	500	25.57	1.44	25.59	1.41	25.76	1.56	25.24	1.56
		1000	25.71	1.39	24.26	1.40	25.84	1.59	25.81	1.64
		5000	25.97	1.36	23.55	1.40	25.94	1.63	26.40	1.69
	28	500	27.65	1.46	27.02	1.42	27.26	1.60	27.57	1.62
		1000	27.78	1.42	25.87	1.42	27.78	1.63	27.44	1.68
		5000	27.99	1.38	25.32	1.41	27.91	1.68	28.26	1.75
	30	500	29.71	1.48	28.70	1.45	29.53	1.63	28.89	1.65
		1000	29.74	1.44	27.20	1.45	29.79	1.68	29.42	1.74
		5000	30.00	1.41	27.16	1.43	29.89	1.74	30.11	1.80
High	26	500	25.56	1.41	25.69	1.37	25.77	1.52	25.66	1.51
		1000	25.65	1.37	24.52	1.38	25.64	1.53	25.75	1.56
		5000	25.87	1.33	23.55	1.38	25.95	1.56	26.51	1.62
	28	500	27.44	1.43	27.23	1.40	27.72	1.54	27.40	1.55
		1000	27.66	1.39	25.91	1.40	27.70	1.58	27.66	1.62
		5000	27.95	1.36	25.18	1.39	27.95	1.62	28.44	1.68
	30	500	29.56	1.45	29.12	1.42	29.71	1.58	29.09	1.59
		1000	29.63	1.41	27.21	1.42	29.74	1.62	29.57	1.66
		5000	29.99	1.38	26.73	1.41	29.92	1.66	30.40	1.72

3.3. Simulation study

We perform a simulation study to (1) demonstrate that our approach is able to fit valid cost constrained regimes in realistic settings and (2) illustrate the utility of flexible ensemble learning approaches for improving regime performance. We simulate data with $K = 3$ decision points and $n = 500$, 1000, and 5000 sample units. At each time point, we simulate a vector of independent standard normal confounding variables, W_k; a binary treatment decision, A_k; an indicator of whether a unit is censored during an interval, C_k; an indicator of whether a unit has survived to the end of the interval, S_k; and cost accrued within the interval, Y_k. Simulations are run at various sample sizes and levels of correlation between survival and cost outcomes. Details for how data are simulated can be found in the appendix. We compare two approaches for modeling the Q-functions: (1) linear

models, and (2) SuperLearner. SuperLearner is an ensemble learner, combining predictions from a set of candidate learners to minimize cross-validated risk (Laan and Rose, 2011). The candidate learners within our SuperLearner are random forests, neural nets, elastic net, and generalized linear models.

For each simulation setting, we generate 500 datasets and estimate an optimal cost constrained regimes under overall cost constraints of $\tau = 26$, 28, and 30. We set the maximal list length, L_k, to 5 for all decision intervals $k = 1, 2, 3$. After regimes are identified, we simulate 10^6 new units, treat them according to the estimated treatment rules, and record the mean survival and cost among these units. This procedure results in two estimates of the mean cost and survival under each fit regime; one obtained by averaging over the estimated Q-functions for the initial interval, and another based on Monte Carlo simulation. The Monte Carlo approach allows us to determine how each regime performs when applied to new data and approximates the true cost and survival under that regime. If the Q-functions are appropriately modelled, then estimates of the expected cost and survival for a fit regime should be similar to Monte Carlo approximations of the truth.

Table 3.1 provides the mean Q-function-based estimate of cost and survival as well as the mean Monte Carlo estimated cost and survival across regimes fit under varying cost constraints and levels of correlation between cost and effectiveness outcomes. Results in this table are from settings with approximately 30% censoring over the course of the three treatment intervals. Results at 0% and 60% censoring are provided in the appendix. In nearly all settings, the mean Monte Carlo estimated cost fall below the specified cost threshold. Only in scenarios where SuperLearner was used with large sample datasets, was the cost constraint not satisfied. In these settings, the mean Q-function based estimate of the total cost underestimated the true Monte-Carlo estimated cost (e.g. true vs. estimated cost of 30.40 vs. 29.92, when $n = 5000$, $\tau = 30$, and under high correlation). Due to misspecification, regimes identified using linear models consistently overestimated the costs observed among new units. Because total regime costs are overestimated when using linear models, some regimes which would truly satisfy the cost constraint may be excluded from consideration. This results in more conservative treatment strategies and decreased survival when compared to SuperLearner-based regimes. In the simulation setting with $n = 1000$, $\tau = 28$, and medium correlation between outcomes, both Linear-model and SuperLearner-based regimes have a mean estimated treatment cost of 27.78. However, because the SuperLearner-based models

more accurately estimate the true regime cost (25.87 and 27.44 for linear-model and SuperLearner-based regimes, respectively), resources are better utilized and units achieve greater survival (1.42 vs. 1.68). Additionally, because SuperLearner may improve our ability to identify patients that are likely to benefit from treatment, Monte Carlo estimates of survival in SuperLearner-based regimes exceed those of Linear-model-based regimes even when costs are similar (e.g. Linear model vs. SuperLearner-based survival of 1.37 vs. 1.51 and costs of 25.69 vs. 25.66 when $n = 500$, $\tau = 26$, and under high correlation).

3.4. The cost-effectiveness of adjuvant therapies for endometrial cancer

The standard treatment for patients with early-stage endometrial cancer is complete hysterectomy. Throughout the post-surgical period, patients may receive adjuvant radiation or chemotherapy to decrease the risk of recurrence (Latif et al., 2014). The decision to provide adjuvant therapy will depend on individual patient characteristics. In particular, younger patients or those with low-grade histology have low risk of recurrence and may not benefit from adjuvant therapies (Creutzberg et al., 2000). Additionally, the best treatment decision for a given patient will change over time. Hogberg et al. (2010) found that certain high risk patients who initially receive adjuvant chemotherapy may exhibit greater survival if switched to radiation therapy. However, the optimal sequencing of adjuvant therapies remains a subject of controversy (Heerik et al., 2021). In this analysis, we use data from the linked SEER-Medicare database to identify an optimally cost-effective regime for assigning adjuvant treatments to endometrial cancer patients. In doing so, we aim to explore how different treatment strategies effect patient survival and costs. Patients in this database were diagnosed with endometrioid histology cancer between 2000 and 2011, with follow-up until 2013.

Each patient in the database is followed for 24 months. We divide this period into four intervals (months 1-6, 7-12, 13-18, and 19-24), wherein treatment is assigned within the first month of each interval. At the beginning of each interval, patients are assigned one of three treatments: (1) adjuvant radiation therapy (RT), (2) adjuvant chemotherapy (CT), or (3) monitoring alone (neither radiation nor chemotherapy), hereafter referred to as control. The data contain followup information on 13722 patients, 714 of whom have censored outcomes. Of the total study population, 27 patients were excluded because they were treated with both adjuvant RT and CT. At baseline, data are available on patient's age, race, and cancer stage and grade. Additionally, Charlson comorbidity

indices and the number of hospitalizations are recorded at every month of follow-up. The mean age at diagnosis is 73.72 (SD = 6.58), and most patients have stage I cancer (94.03%). Charlson comorbidity indices are between zero and five, with 54.39% of patients having an index of zero. The 24-month-restricted mean survival among patients is 22.86 months and the mean accumulated cost over the study period is 33830.89.

Due to sensitivity from random seeds when using SuperLearner, we fit Q-functions using a generalized additive model. Generalized additive models provide a compromise model that is both more flexible than linear models and more robust to random seed than SuperLearner. Our outcomes of interest were total cost over the 24 month period and restricted mean survival. Covariates used as predictors include: age at diagnosis, race, cancer stage, cancer grade, number of hospitalizations over the previous six months, maximum Charlson comorbidity index over the previous six months, total cost accrued over the previous six months, and previous treatment assignment. Treatment rules for each interval contain a maximum of five clauses and assign adjuvant therapies based upon a patient's age at diagnosis, Charlson comorbidity index, number of hospitalizations over the past 6 months, previous treatments, and their cancer stage and grade.

Because medical costs do not accrue at the same rate throughout the post-surgical period, we allow the interval specific cost constraints for each candidate regime to change with time. To select the cost-constraints of our candidate regimes, we first identify a maximally effective treatment regime by optimizing decision rules without resource constraints (i.e. $\tau = \infty$). Next, identify the cost accrual pattern under both the optimally effective regime and some "lower bar" regime- here, the standard of care or what is observe in the data. The cost constraints for the candidate regimes are defined as mixtures of the lower and upper bound cost accrual patterns. For $s \in [0, 1]$, the vector of interval-specific cost constraints corresponding to s is given by $\boldsymbol{\tau}_s = (1-s)(18500, 5500, 5500, 5000)^T + s(18500, 5500, 9000, 8000)^T$. For $s = 0$, the interval-specific constraints approximate the pattern of cost-accrual observed among patients within the database, while for $s = 1$, the constraints mirror the cost-accrual pattern under an optimal regime fit without resource constraints. Candidate regimes are fit under the vector of cost constraints corresponding to $s = 0, 0.25, 0.5, 0.75, 1$.

To identify an optimally cost-effective regime, we sequentially compare each candidate regime using the ICER. Given two candidate regimes, d_1 and d_2, ICER is a comparative measure defined as the ratio of the difference in expected cost to the difference in expected effectiveness between the

two regimes:

$$\text{ICER}(d_2, d_1) = \frac{\mathbb{E}[Y^{d_2}] - \mathbb{E}[Y^{d_1}]}{\mathbb{E}[Z^{d_2}] - \mathbb{E}[Z^{d_1}]}.$$

ICER can be interpreted as the cost per unit change in effectiveness obtained by switching from treatment regime d_1 to d_2. Let λ denote a pre-selected willingness-to-pay (WTP) parameter. The WTP represents the maximum cost a payer is willing to incur for a unit change in effectiveness. Adopting regime d_2 over the comparator regime, d_1, is considered cost-effective if ICER is less than the chosen WTP. Given a set of regimes corresponding to different cost-constraints, $\widehat{d}(\tau_1)$, ..., $\widehat{d}(\tau_J)$, we provide an algorithm that identifies the optimally cost-effective regime by sequentially comparing each candidate regime to the best existing alternative.

1. Define the current most cost-effective regime: $\widehat{d}_{\text{CE}} \coloneqq \widehat{d}(\tau_1)$ and set $j = 2$
2. If $\Big(\text{ICER}(\widehat{d}(\tau_j), \widehat{d}_{\text{CE}}) < \lambda\Big)$, then $\widehat{d}_{\text{CE}} \coloneqq \widehat{d}(\tau_j)$, otherwise do not update $\widehat{d}_{\text{CE}}$.
3. If $j < J$, set $j = j + 1$ and repeat from Step 2, otherwise $\widehat{d}_{\text{CE}}$ is the optimally cost-effective regime.

We perform this cost-effectiveness analysis to select an optimally cost-effective regime from the set of regimes with cost-constraints given by τ_s for $s = 0, 0.2, 0.4, ..., 1$. Results of this analysis under a WTP of $\$4{,}166$/month or equivalently, $\$50{,}000$/year are provided in Table 3.2. Because none of the regimes have ICER less than this WTP when compared with Regime I, Regime I (the least expensive regime) is the optimally cost-effective list-based treatment regime. The treatment rules for this regime are provided within the appendix. Under this regime, patients are expected to accrue approximately $\$34{,}380$ in medical costs and to survive for 22.840 months within the two years following hysterectomy. Estimated restricted mean survival and cost under this regime are similar to those observed among patients in our database.

We may characterize the differences between the optimal regime and other candidate regimes is by comparing treatment patterns under each regime. Figure 3.1 provides the most frequent treatment patterns under the optimal, a moderately expensive (regime III), and the most expensive regimes. Under the optimal regime the three most common treatment patterns are: (1) Neither treatment for 2.0 years (68%), (2) Neither treatment for the first 1.0 years and chemotherapy for 1.0 years (21%), and (3) adjuvant chemotherapy for 0.5 years followed by neither therapy for 1.5 years (4%).

Table 3.2: Estimated mean survival (months) and cost (USD$) for patients treated according to each candidate regime. Incremental cost-effectiveness ratios comparing subsequent candidate regimes. Results for optimally cost-effective treatment regime at WTP of $\$4{,}166$/month bolded.

Regime	s	Cost	Survival	ICER	Comparator
1	**0**	**34379.86**	**22.840**	**NA**	**NA**
2	0.25	37623.31	23.018	18250.62	1
3	0.50	40976.36	23.079	27663.32	1
4	0.75	44447.11	23.157	31801.15	1
5	1.00	46590.46	23.175	36464.93	1

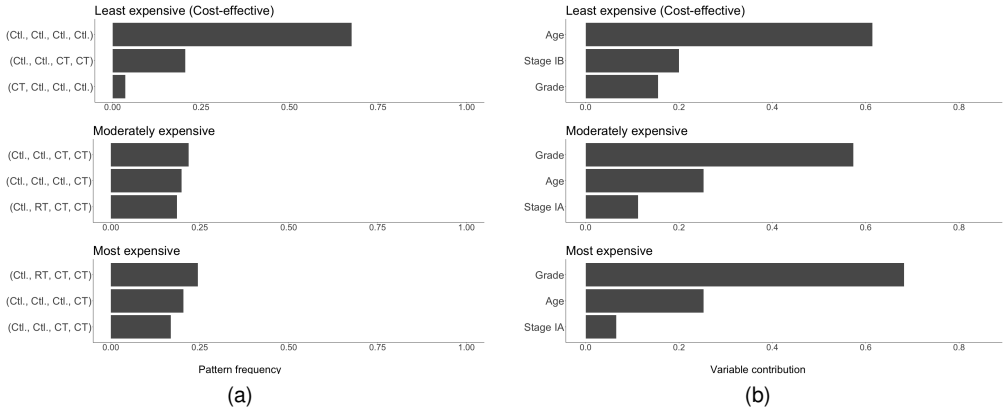

Figure 3.1: Most common treatment patterns (a) and variables (b) for the optimally cost-effective, a moderately expensive, and most expensive treatment regimes. Estimated by determining which treatments the observed units would be assigned to based on their covariate history.

This differs from the most common treatment patterns under both the moderately expensive and the most expensive regimes, under which less than 5% of patients receive neither adjuvant therapy for the duration of study. In practice, treatment with radiation or chemotherapy after the initial monitoring period often indicates cancer recurrence. Because our data do not contain information on recurrence, other covariates must be used as proxies to determine whether recurrence has occurred and additional treatment is required. The most common treatment patterns in the more expensive regimes entail treatment after a period of monitoring, suggesting that additional costs are incurred due to more aggressive predictions of recurrence. We may also compare which variables are most frequently selected to assign treatment between regimes using Figure 3.1. For all three considered regimes, treatment decisions were most commonly based on patient's age, cancer grade, and cancer stage with variation in the relative importance of these factors.

3.5. Discussion

In this paper, we present a two-step procedure for identifying an optimally cost-effective dynamic treatment regime with interpretable, list-based decision rules. In the first step, we use a novel Q-learning and policy search-based approach to estimate optimal list-based regimes that maximize treatment effectiveness under a predefined constraint on treatment costs. Through simulations we show the validity of our cost constrained regimes and illustrate how flexible ensemble learners can improve regime performance. The second step consists of a cost-effectiveness analysis that can select an optimally cost-effective regime from a set of candidate regimes characterized by variable cost constraints. This procedure identifies the most cost-effective treatment strategy by iteratively comparing each candidate regime with the best existing alternative.

While our method works for general measures of effectiveness, in our simulations and data example we focus on restricted mean survival. This is a traditional measure of effectiveness within cost-effectiveness studies (Li et al., 2018; Spieker et al., 2019). Recently, Linn, Laber, and Stefanski (2017) developed methods for optimizing specified quantiles of a distribution. Future extensions may explore how this methodology could be used to allow for maximization of median survival time while placing a constraint on overall cost. In contrast with previous work on identifying cost-effective regimes which have maximized effectiveness while penalizing costs Lakkaraju and Rudin (2017) and Xu et al. (2020), our proposed cost-effectiveness analysis is based upon iterative comparisons of candidate regimes. This allows us to better characterize how decisions change as the allowable cost increases and avoids complications arising from non-iterative cost-effectiveness analyses (Cohen and Reynolds, 2008). For example, suppose two or more experimental treatments are considered cost-effective when compared with the standard of care. Additional iterative comparisons between the experimental treatments are necessary to determine which treatment is preferred under a pre-specified budgetary constraint.

An important consideration in the design of cost-effectiveness analyses is in the selection of cost constraints. In general, if too few cost constraints (and thus, too few candidate regimes) are selected, then we may be unable to approximate a true optimally cost-effective regime. If cost constraints are too similar, then ICER may become unstable due to similar estimated effectiveness. This consideration is further complicated by difficulties in quantifying uncertainty. Constrained esti-

mation results in nonstandard asymptotic theory, and methods for performing inference have not yet been explored (Laber et al., 2018). Additionally, the modeling choice for Q-functions may influence estimated treatment rules. Flexible methods like SuperLearner may improve regime performance, but are also subject to variability across random seeds. For estimates of treatment effects, Benkeser (2020) propose combining estimates arising from different seeds. However, work is required to determine the best procedure for combining decision-list based rules.

CHAPTER 4

INFLUENCE FUNCTION BASED INSTRUMENTAL VARIABLE ESTIMATOR OF CENSORED MEDICAL COSTS

4.1. Introduction

Studies aimed at estimating medical costs accrued under different treatments are necessary for making informed healthcare policy decisions. These methods are particularly useful when analysing treatments with similar clinical effectiveness. In a recent series of papers, Keele et al. (2018, 2019) and Fogarty et al. (2021) assessed the effect of treatment by emergency general surgery (EGS) on patient mortality for a set of acute conditions for which non-operative alternatives exist. While more aggressive forms of care often exhibit protective effects in adjacent medical settings(Lorch et al., 2012), these analyses showed a lack of beneficial evidence for EGS for many of the conditions under study. Here, knowledge of how EGS affects accrued medical costs is critical for making informed policy recommendations. Data for performing these analyses are often pulled from complex observational sources such as claims databases. Common complications in this setting are: (1) measured and unmeasured confounding of the exposure-outcome relationship, (2) informative cost censoring, and (3) irregular cost distributions. Methodologies that are robust to both sources of confounding and that are flexible enough to account for irregularities in the cost distribution and censoring have not previously been explored.

Drawing causal conclusions from observational databases requires careful consideration of the effects of exposure-outcome confounding. If all sources of confounding are measured, then we may assess the effect of different treatments on overall medical costs using inverse probability weighting (Li et al., 2016), g-computation (Spieker et al., 2020b) , or doubly robust approaches (Li et al., 2018). Many of these approaches build upon the inverse probability of censoring weighting (IPCW) framework for accounting for informative cost censoring (Bang and Tsiatis, 2000). This problem was first noted by Lin et al. (1997), who discussed how correlation between accrued costs at time of censoring and those at time of event can induce bias in estimates of overall treatment cost. However, these methods do not extend to settings with unmeasured confounding. Sensitivity analyses to assess how robust cost estimates may be to unmeasured confounding have been proposed,

namely by Handorf et al. (2013, 2018). Alternatively, the instrumental variable (IV) framework can be used to directly estimate causal effects in settings with unmeasured confounding.

Popular IV approaches such as two stage least squares (2SLS) or two stage predictor substitution (2SPS) have become more common in medical and epidemiologic research (Keele et al., 2018, 2019; Lorch et al., 2012). Because these methods often rely on parametric assumptions, implementation may not be valid in settings with highly skewed cost data. Additionally, extensions of these methods to settings with informatively censored cost outcomes have not been adequately developed. Influence function based approaches have recently been proposed by Díaz (2019) and Lee, Kennedy, and Mitra (2020) to identify treatment effects on patient survival in the presence of censoring. These methods are doubly robust, efficient, and can incorporate semiparametric modeling approaches. However, extensions to informatively censored cost data have not been explored.

In this paper, we propose two instrumental variable estimators of the complier average treatment effect on costs. Given a valid instrument, these estimators are unbiased in the presence of unmeasured confounding. Additionally, the use of a partitioned cost estimator allows us to address informative cost censoring and improve efficiency by utilizing data from patients with censored medical costs. The proposed estimators are based on influence functions and hence provide doubly robust, semiparametric, and efficient estimation of treatment effects as well as valid asymptotic inference. We present results from simulation studies assessing the performance of the proposed estimator under varying degrees of censoring, IV strength, and nuisance function misspecification. We also apply the proposed approach to a study assessing the costs of surgical and non-surgical interventions for gallstones and hemorrhaging using observational data.

In Section 4.2 we describe instrumental variable methodology for estimating the causal effect of an exposure on accrued medical costs using a simple and partitioned weighted estimator. Details concerning asymptotic inference are also provided. Section 4.3 provides the results of simulation studies exploring the operating characteristics of our two proposed estimators in various realistic settings.To conclude, Section 4.4 will first discuss the utility of our proposed methodologies and the implications of our results before describing challenges and areas for future work.

4.2. Methods

4.2.1. Notation and Identification

In this section, we present novel methodology for estimating the effect of an exposure on potentially censored medical costs using an instrumental variable. Specifically, if a study follows patients up to some maximal time, τ, then we aim to estimate the treatment effect on costs over the interval $[0, \tau]$. For each unit $i = 1, ..., n$ we observe the random vector $O_i = (L_i, Z_i, A_i, T_i, \Delta_i, Y_i)$. Here, L is a vector of confounding variables, Z is a binary or continuous instrumental variable, A denotes exposure status, T denotes time of censoring or event, Δ is an indicator of whether a patient has fully observed cost outcomes, and Y represents total accrued medical costs up to time τ. If patient cost data is recorded at multiple points during the study duration, say at times $0 = t_0 < t_1 < t_2 < ... < t_K = \tau$, then we may also represent the observed data vector as $O_i = (L_i, Z_i, A_i, T_{1i}, \Delta_{1i}, Y_{1i}, T_{2i}, \Delta_{2i}, Y_{2i}, ..., T_{Ki}, \Delta_{Ki}, Y_{Ki})$ where $T_k = \min(T, t_k)$ denotes survival up to time t_k, Δ_k indicates whether costs are fully observed up to time t_k, and Y_k represents costs accrued over the interval $(t_{k-1}, t_k]$. Note that the sum of interval-specific costs for each individual is equal to the total accrued cost, i.e. $\sum_{k=1}^{K} Y_{ki} = Y_i$.

We want to estimate the causal effect of the treatment, A, on overall medical costs, Y. Specifically, we want to know the treatment effect on the restricted mean cost (cost up to the end of interval K, $\sum_{k=1}^{K} Y_k = Y$). Our target parameter is the local average treatment effect (LATE):

$$\Psi = E[Y^{a=1} - Y^{a=0} | A^{z=1} > A^{z=0}]$$

In the absence of censoring, the LATE may be identified using the observed data under the following set of assumptions:

(A1) IV relevance: The IV must be associated with the treatment of interest, $\mathbb{E}[A|L, Z = 1] \neq \mathbb{E}[A|L, Z = 0]$

(A2) There must be no unmeasured confounding between the IV and the outcome, $Z \perp\!\!\!\perp A^z | L$ and $Z \perp\!\!\!\perp Y^{a,z} | L$

(A3) Exclusion Restriction: The IV can only affect the outcome through the treatment. $Y^z =$

$Y^{A^z,z} = Y^{A^z}$

(A4) No defiers: There are no units that would take treatment under $Z = 0$ that would not have given $Z = 1$, $P(A^{z=0} > A^{z=1}) = 0$

Theorem 4.2.1. *Under assumptions (A1)-(A4) the LATE, Ψ, may be identified as:*

$$\Psi = \frac{E\Big[E[Y|Z=1,L] - E[Y|Z=0,L]\Big]}{E\Big[E[A|Z=1,L] - E[A|Z=0,L]\Big]} \tag{4.1}$$

where $E[Y|Z=z,L] = \sum_{a=0}^{1} E[Y|A=a,Z=z,L]P(A=a|Z=z,L)$

When cost data are censored, additional consideration must be given to avoid bias due to informative cost censoring (Lin et al., 1997). This problem arises due to correlation between costs at time of event and costs at time of censoring. Under an additional assumption of non-informative censoring, $(A5)\ T \perp\!\!\!\perp C|A,Z,L$, we may identify the LATE using an inverse probability of censoring weighted (IPCW) estimator. This is accomplished by noting the following equivalencies:

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[Y|Z=z,L]] &= \mathbb{E}\Big[\sum_{a=0}^{1}\mathbb{E}[Y|A=a,Z=z,L]P(A=a|Z=z,L)\Big] \\
&= \mathbb{E}\Big[\sum_{a=0}^{1}\mathbb{E}\Big[\frac{Y\Delta}{G(T|a,z,L)}|A=a,Z=z,L\Big]P(A=a|Z=z,L)\Big] \qquad (4.2)\\
&= \mathbb{E}\Big[\sum_{a=0}^{1}\sum_{k=1}^{K}\mathbb{E}\Big[\frac{Y_k\Delta_k}{G(T_k|a,z,L)}|A=a,Z=z,L\Big]P(A=a|Z=z,L)\Big] \qquad (4.3)
\end{aligned}$$

where $G(T|a,1,L) = P(C > T|A=a,Z=1,L)$ represents the survivor function for the censoring distribution evaluated at time T. Based on the representations in Equations 4.2 and 4.3, we provide two methods for identifying the LATE based upon IPCW estimation: (1) a simple estimator corresponding to settings where costs are measured at event or end-of-study and (2) a partitioned estimator for use in settings where costs are measured over time. Unlike the simple weighted estimator, which only uses information from units which are fully observed, the partitioned representation allows utilization of cost data for as long as that data is available on a given unit. Define the following nuisance functions:

1. $\mu_{a,z}(l) = E\Big[\frac{Y\Delta}{G(T)}\Big|A=a,Z=z,L=l\Big]$

2. $\mu^2_{a,z}(l) = E\left[\frac{Y\Delta}{G(T)^2}\middle| A = a, Z = z, L = l\right]$

3. $\pi_{a,z}(l) = P(A = a|Z = z, L = l)$

4. $\omega_z(l) = P(Z = z|L = l)$

5. $G(t|a, z, l) = P(C > t|A = a, Z = z, L = l)$

Let Θ denote the set of nuisance functions. Using these nuisance functions, we define the two IPCW-based identifications of the LATE below:

$$\Psi_s = \frac{E\left[\sum_{a=0}^{1} \mu_{a,1}(L)\pi_{a,1}(L) - \mu_{a,0}(L)\pi_{a,0}(L)\right]}{E[\pi_{1,1}(L) - \pi_{1,0}(L)]} \tag{4.4}$$

And

$$\Psi_w = \frac{E\left[\sum_{a=0}^{1}\sum_{k=1}^{K} \mu_{a,1,k}(L)\pi_{a,1}(L) - \mu_{a,0,k}(L)\pi_{a,0}(L)\right]}{E[\pi_{1,1}(L) - \pi_{1,0}(L)]} \tag{4.5}$$

4.2.2. Influence Function Estimators

We want to define an influence function based estimator of the LATE using both the simple and partitioned cost estimators. An influence function based estimator will allow us to perform valid asymptotic inference while (1) incorporating flexible modeling approaches, (2) being doubly robust, and (3) attaining parametric rates of convergence if models are correctly specified. We first consider the influence function estimator based upon the simple weighted representation in Equation 4.4.

Theorem 4.2.2. *The efficient influence function for the simple weighted representation, Ψ_s, is given by:*

$$IF(\Psi) = \frac{\big(M_s(Z = 1) - M_s(Z = 0)\big) - \Psi\big(\Pi(Z = 1) - \Pi(Z = 0)\big)}{E[\pi_{1,1}(L) - \pi_{1,0}(L)]}$$

Where

$$M_s(Z = z) = \sum_{a=0}^{1} \frac{I(a, z)}{\omega_z(L)}\left(\frac{Y\Delta}{G(T|a, z, L)} - \mu_{a,z}(L) - \mu^2_{a,z}(L)(\Delta - G(T|a, z, L)\right) + \mu_{a,z}(L)\left(\frac{I(Z = z)}{\omega_z(L)}\big(I(A = a) - \pi_{a,z}(L)\big) + \pi_{a,z,L}\right)$$

and

$$\Pi(Z=z) = \frac{I(Z=z)}{\omega_z(L)}(A - \pi_{1,z}(L)) + \pi_{1,z}(L)$$

We may similarly prove that the EIF for the partitioned representation, Ψ_p, is equivalent to that of Ψ_s except replacing $M_s(Z=z)$ with $M_p(Z=z)$ where

$$M_p(Z=z) = \sum_{a=0}^{1} \frac{I(a,z)}{\omega_z(L)} \left\{ \sum_{k=1}^{K} \left(\frac{Y_k \Delta_k}{G(T_k|a,z,L)} - \mu_{a,z,k}(L) - \mu_{a,z,k}^2(L)(\Delta - G(T_k|a,z,L) \right) \right\} + \left\{ \sum_{k=1}^{K} \mu_{a,z,k}(L) \right\} \left(\frac{I(Z=z)}{\omega_z(L)} \big(I(A=a) - \pi_{a,z}(L) \big) + \pi_{a,z,L} \right)$$

Under true values of Θ, it can be shown that $E[IF(\Psi;\Theta)] = 0$. Given estimates of the nuisance functions, $\widehat{\Theta}$, this implies the construction of an estimator for Ψ that solves the estimating equation $P_n(IF(\Psi,\widehat{\Theta})) = 0$. Define the simple and partitioned influence function based estimators of the LATE:

$$\widehat{\Psi}_s = \frac{P_n(M_s(1;\widehat{\Theta}) - M_s(0;\widehat{\Theta}))}{P_n(\Pi(1;\widehat{\Theta}) - \Pi(0;\widehat{\Theta}))} \tag{4.6}$$

and

$$\widehat{\Psi}_p = \frac{P_n(M_p(1;\widehat{\Theta}) - M_p(0;\widehat{\Theta}))}{P_n(\Pi(1;\widehat{\Theta}) - \Pi(0;\widehat{\Theta}))} \tag{4.7}$$

4.2.3. Asymptotic Results

In this section we describe asymptotic results for our simple and partitioned influence function based estimators. The asymptotic distribution of the simple weighted influence function estimator, $\widehat{\Psi}_s$ is described by the following theorem:

Theorem 4.2.3. *Assume the following conditions:*

(C1) The nuisance functions, Θ, are in the Donsker class.

(C2) For some constant $\epsilon > 0$, $P(\epsilon < \widehat{\omega}_z < \infty) = 1$, $P(\epsilon < \pi_{az} < \infty) = 1$, $P(\epsilon < \widehat{\mu}_{az}^2 \pi_{az} < \infty) = 1$, and $P(\epsilon < \widehat{\mu}_{az} < \infty) = 1$ for all z and a

Under (C1), (C2), and (A1)-(A6), we can show the following:

$$\widehat{\Psi}_s - \Psi = O_p\Big(\sum_{z,a\in\{0,1\}} \|\omega_z - \widehat{\omega}_z\| \, \|\mu_{az} - \widehat{\mu}_{az}\| + \|\omega_z - \widehat{\omega}_z\| \, \|\pi_{az} - \widehat{\pi}_{az}\| + \left\|G(T) - \widehat{G}(T)\right\| \Big) +$$
$$\zeta_n^{-1}(P_n - P)(\phi_1(\Theta) - \Psi\phi_2(\Theta)) + o_p(n^{-1/2})$$

Where:

$$\zeta_n = P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}))$$

Theorem 4.2.3 describes the asymptotic behavior of our proposed simple weighted estimator of the LATE. It can be seen that the first error component goes to zero if either (1) the censoring mechanism, G, and the IV prevalence model, ω_z, or (2) the censoring mechanism and both the mean model, μ_{az}, and the propensity score model, π_{az}, are correctly specified. If neither of these conditions are true, then the simple weighted estimator will exhibit bias. The second term, $(P_n - P)(\phi_1(\Theta) - \Psi\phi_2(\Theta))$, is asymptotically normal under the central limit theorem. This suggests that we may estimate the variance of the simple weighted estimator as:

$$\widehat{V}(\widehat{\Psi}_s) = (n-1)^{-1} \sum_{i=1}^{n} \left(\frac{\phi_{1si}(\widehat{\Theta}) - \widehat{\Psi}_s\phi_{2si}(\widehat{\Theta})}{\phi_{2si}(\widehat{\Theta}) + \psi_{2si}(\widehat{\Theta})} - \left[n^{-1} \sum_{i=1}^{n} \frac{\phi_{1si}(\widehat{\Theta}) - \widehat{\Psi}_s\phi_{2si}(\widehat{\Theta})}{\phi_{2si}(\widehat{\Theta}) + \psi_{2si}(\widehat{\Theta})} \right] \right)^2$$

Under a similar set of assumptions, we may draw similar conclusions regarding the partitioned weighted estimator.

Theorem 4.2.4. *Assume the following conditions:*

(C1) The nuisance functions, Θ, are in the Donsker class.

(C2) For some constant $\epsilon > 0$, $P(\epsilon < \widehat{\omega}_z < \infty) = 1$, $P(\epsilon < \pi_{az} < \infty) = 1$, $P(\epsilon < \sum_{k=1}^{K}(\widehat{\mu}_{azk}^2)\pi_{az} < \infty) = 1$, and $P(\epsilon < \sum_{k=1}^{K}(\widehat{\mu}_{az}) < \infty) = 1$ for all z and a

Under (C1), (C2), and (A1)-(A6), we can show the following:

$$\widehat{\Psi}_p - \Psi = O_p\Big(\sum_{z,a\in\{0,1\}} \|\omega_z - \widehat{\omega}_z\| \left\| \sum_{k=1}^{K} (\mu_{azk} - \widehat{\mu}_{azk}) \right\| + \|\omega_z - \widehat{\omega}_z\| \, \|\pi_{az} - \widehat{\pi}_{az}\| + \left\|G(T) - \widehat{G}(T)\right\| \Big) +$$
$$\zeta_n^{-1}(P_n - P)(\phi_1(\Theta) - \Psi\phi_2(\Theta)) + o_p(n^{-1/2})$$

Where:

$$\zeta_n = P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}))$$

This suggests similar robustness properties as the simple weighted estimator, and the following estimator of the asymptotic variance:

$$\widehat{V}(\widehat{\Psi}_p) = (n-1)^{-1} \sum_{i=1}^{n} \frac{\phi_{1pi}(\widehat{\Theta}) - \widehat{\Psi}_p \phi_{2pi}(\widehat{\Theta})}{\phi_{2pi}(\widehat{\Theta}) + \psi_{2pi}(\widehat{\Theta})}$$

The asymptotic results for both the simple and partitioned weighted estimators rely on the condition that the nuisance functions, Θ, are within the Donsker class. This condition controls how close functions of the estimated nuisance functions can get to their limiting versions (e.g. the difference between $(P_n - P)\phi_1(\widehat{\Theta})$ and $(P_n - P)\phi_1(\Theta)$) (Kennedy, 2016). This condition can be restrictive if estimators with unbounded variation (such as random forests) are used, as these are generally not Donsker (Díaz, 2019). This condition can be removed by implementing sample-split versions of the estimators. Partition the index set $\{1, ..., n\}$ into M mutually exclusive groups $V_1, ..., V_M$ such that $\bigcup_{m=1}^{M} V_m = \{1, ..., n\}$ and $Vm \bigcap V_{m'} = \emptyset$ for $m \neq m^{'}$. Let $\widehat{\Psi}_{sm}$ and $\widehat{\Psi}_{pm}$ be the simple and partitioned estimators obtained by using units $\{1, ..., n\} \notin V_m$ to estimate the nuisance functions and units V_m to estimate the causal effect. Define the sample-split simple and partitioned estimators,

$$\widehat{\Psi}_s = M^{-1} \sum_{m=1}^{M} \widehat{\Psi}_{sm} \qquad \widehat{\Psi}_p = M^{-1} \sum_{m=1}^{M} \widehat{\Psi}_{pm},$$

and their associated variances,

$$\widehat{V}(\widehat{\Psi}_s) = M^{-1} \sum_{m=1}^{M} \widehat{V}(\widehat{\Psi}_{sm}) \qquad \widehat{V}(\widehat{\Psi}_p) = M^{-1} \sum_{m=1}^{M} \widehat{V}(\widehat{\Psi}_{pm}).$$

Note that within each validation set, V_m, the estimated nuisance functions $\widehat{\Theta}$ are fixed. This fact allows us to apply the empirical process properties that our asymptotic results are based upon without invoking Donsker class conditions.

4.3. Simulation studies

We perform simulation studies to evaluate the operating characteristics of our proposed influence function based instrumental variable estimators. This study will compare three instrumental variable estimators: (1) a 2SPS estimator based upon complete cases, (2) the simple weighted influence function estimator, and (3) the partitioned weighted influence function estimator. Our estimators' ability to capture the true treatment effect may be affected by instrument strength, censoring, or by mispecification of the nuisance functions described in Section 4.2. Because of this, we explore simulation settings with both weak and strong instruments, high and low levels of censoring, and different degrees of model mispecification.

Within each simulation setting, we construct 500 datasets of $n = 5000$ units. For each unit, we simulate the covariates (L, U, Z, A, C, T, Y) where L and U are random normal covariates with means $\mu_L = 2$ and $\mu_U = 1$ shared variance of $\sigma^2 = 1$, Z is a binary instrumental variable with $P(Z = 1|L) = \text{expit}(0.5L)$, A is a binary exposure with $P(A = 1|Z, L, U) = \text{expit}(-0.4+\nu Z+0.25L+0.5U-0.5(U^2))$, C and T are exponentially distributed censoring and survival times with rate parameters of $\lambda_C = \exp(\gamma_c+0.3A)$ and $\lambda_T = \exp(0.2-A-0.25L-0.25U)$, and Y is a right-skewed cost variable. Units are followed for until a maximal time, $\tau = 4$. Costs $(Y_0, Y_1, Y_2, Y_3, Y_4)$ are simulated at baseline, and at times $t = 1, 2, 3$, and 4 for each unit depending on their observed survival outcomes. In addition, costs associated with death, Y_D, are simulated for units if they have an observed event. Baseline costs are simulated from a gamma distribution with shape parameter $\alpha = 2.5$ and scale parameter $\theta = 1/\beta$ where $\beta = \exp(-1A - 0.5L - 0.2U)$. Follow-up costs and death costs are also simulated with shape α, but have scale parameters of $0.1/\beta$ and $0.2/\beta$ respectively. Under this setup the difference in total expected treatment costs $Y = \sum_{k=0}^{4} Y_k + Y_D$ between exposed and non-exposed units is $\Psi = 22.987$. The parameters ν and γ_c control the instrument strength and degree of censoring. In weak instrument settings (10% difference in treatment prevalence between instrument levels), $\nu = 0.35$ and for strong instrument settings (45% difference in prevalence), $\nu = 2.5$. Low levels of censoring (30%) correspond to $\gamma_c = -2.5$ and high levels of censoring (50%) to $\gamma_c = -1.25$.

Table 4.1 provides simulation results across simulated datasets at varying levels of IV strength and degrees of censoring. We may notice the following across all simulation settings: (1) the simple

IV Strength	Censoring	Estimator	Mean	% Bias	SE	$\widehat{SE}$	Coverage
Weak	30%	Naive	14.852	-0.354	7.377	8.538	0.874
		Simple	23.404	0.018	16.144	18.097	0.968
		Partitioned	22.808	-0.008	12.764	13.533	0.978
	50%	Naive	14.522	-0.369	8.677	9.724	0.914
		Simple	27.537	0.197	30.962	29.325	0.924
		Partitioned	25.672	0.116	17.964	17.756	0.950
Strong	30%	Naive	18.484	-0.196	1.164	1.536	0.108
		Simple	23.340	0.015	1.950	1.894	0.926
		Partitioned	22.539	-0.020	1.620	1.619	0.940
	50%	Naive	17.924	-0.221	1.241	1.613	0.084
		Simple	26.937	0.171	3.298	3.307	0.704
		Partitioned	24.810	0.079	1.847	1.836	0.792

Table 4.1: Mean point estimate, percent bias, standard error (SE), mean estimated standard error, and empirical coverage of the three instrumental variable estimators of the LATE under different degrees of IV strength and censoring. All nuisance functions correctly specified.

and partitioned estimators exhibit smaller bias than the naive estimator, (2) the standard error of the partitioned estimator is less than that of the simple weighted estimator, and (3) the standard errors of all estimators are appropriately estimated. When we compare simulations based on weak and strong instruments, we notice that observed coverage is greater when the instrument is weak. This appears to be driven by large increases in the estimated standard errors of these estimators in this setting. For example, the standard errors of the partitioned estimator when there is 50% censoring are 17.756 and 1.836 in weak and strong IV settings, respectively. Additionally, the bias of all estimators increases as the level of censoring increases. These two results lead to decreased coverage for our proposed simple and partitioned estimators when there is a strong IV and high levels of censoring.

Table 4.2 provides the performance of each of the instrumental variable estimators under different modeling specifications for the nuisance functions. Here, model misspecification is defined as follows: models for ω are incorrect if they do not adjust for L, models for π are incorrect if they account for Z but not L, models for G are incorrect if they do not adjust for A, Z, or L, and those for μ are incorrect if they adjust for A and Z but not L. Because the complete-case 2SPS estimator does not rely on the estimated nuisance functions, we only provide results for this estimator in the first scenario, wherein all models are correctly specified. The bias of the naive estimator is nearly

Scenario	Estimator	Mean	% Bias	SE	$\widehat{SE}$	Coverage
(1) Correct	Naive	18.644	-0.189	2.670	2.683	0.650
	Simple	23.091	0.004	5.406	4.686	0.940
	Partitioned	22.484	-0.022	4.149	3.851	0.950
(2) Incorrect ω	Simple	24.706	0.074	4.097	3.383	0.860
	Partitioned	23.038	0.002	3.215	2.869	0.910
(3) Incorrect π, μ	Simple	23.133	0.006	6.212	5.707	0.950
	Partitioned	22.009	-0.043	4.874	4.963	0.983
(4) Incorrect G	Simple	23.463	0.020	4.869	4.396	0.930
	Partitioned	22.642	-0.016	4.118	3.836	0.947
(5) Incorrect ω, μ	Simple	37.433	0.628	4.309	4.338	0.073
	Partitioned	36.136	0.571	3.740	3.811	0.050
(6) Incorrect ω, π	Simple	23.945	0.041	4.089	3.238	0.860
	Partitioned	22.524	-0.021	3.138	2.742	0.910
(7) Incorrect ω, π, μ	Simple	36.702	0.596	4.110	4.038	0.073
	Partitioned	35.543	0.545	3.578	3.533	0.040

Table 4.2: Mean point estimate, percent bias, standard error (SE), mean estimated standard error, and empirical coverage of the three instrumental variable estimators of the LATE under different degrees of nuisance function misspecification. Simulations based on strong IV and 30% censoring.

20%, while both the simple and partitioned estimators exhibit less than 10% bias in settings (1), (2), (3), (4), and (6). We notice the largest increase in bias when both the IV prevalence model, ω, and the outcome model, μ, are misspecified (e.g. scenarios (5) and (7)). In both of these settings, the observed bias of the simple and partitioned estimators exceed that of the naive estimator ($>$50% bias vs. 18.9% bias). As in the previous simulations, estimated standard errors are good approximations of the empirically observed standard errors.

4.4. Discussion

In this paper we propose two influence function based estimators of the local average treatment effect of an exposure on restricted mean medical costs. Our estimators are built upon inverse probability of censoring weighting techniques, which can be used to eliminate bias associated with informative cost censoring. In studies where medical cost information is recorded at multiple points, our partitioned estimator can improve efficiency by utilizing available cost information on all units rather than on only those with fully observed cost data. In addition, our proposed simple and

partitioned weighted estimators are built upon influence function theory. Under a set of entropy conditions, this allows for flexible and efficient semiparametric estimation of treatment effects and valid asymptotic inference based upon the central limit theorem.

Through simulation studies, we are able to illustrate the bias reduction and efficiency gains of our simple and partitioned estimators in a range of realistic settings. Due to informative cost censoring and skewed cost data, the common two-stage predictor substitution model based upon complete case data results in large degrees of bias and poor coverage. In all simulation settings except those in which both the IV prevalence and outcome models were misspecified, the influence function based estimators resulted in large reductions in bias compared with this naive estimator. Both of the proposed estimators exhibited small bias in settings with high levels of censoring. This may be due to difficulty in modeling the outcome mechanism when the censoring weights have a large effect. More flexible modeling procedures, such as stacking methods (Xing, Lesperance, and Zhang, 2020), may be able to further reduce bias in these scenarios.

In future work, this methodology could potentially be used to conduct flexible and efficient cost-effectiveness analyses in settings with unmeasured confounding. The development of methodology for estimating the effectiveness and cost of a treatment using instrumental variables can lead to straightforward estimation of measures of cost-effectiveness such as the Net Monetary Benefit. Given two proposed interventions, the net monetary benefit (NMB) is obtained by subtracting the observed difference in medical costs from the product of a willingness-to-pay parameter and the difference in observed effectiveness. The NMB measures net monetary value associated with a treatment. Combination of the proposed methods with similar methods for estimating treatment effects (Lee, Kennedy, and Mitra, 2020) may be able to facilitate efficient cost-effectiveness analyses that are robust to unmeasured confounding using observational data.

CHAPTER 5

DISCUSSION

In this dissertation, we propose three novel methodologies for performing cost-effectiveness analyses. These developments are motivated by the need for methods that can (1) assess censored joint outcomes data, (2) account for confounding arising from non-randomization, and (3) accomodate flexible modeling approaches for zero-inflated or skewed cost and effectiveness data. These methods may be used to help guide the optimal allocation of limited resources in medical settings.

In our first project, we introduce a regression framework for NBS that allows us to explain variability in cost-effectiveness arising from covariates. The proposed method is the first to enable estimation of NBS within levels of measured covariates and allow for testing the effects of covariates on NBS. Regression of NBS is done in three steps: (1) estimate the distribution of cost and effectiveness outcomes conditional on treatment and measured covariates, (2) implement Monte Carlo standardization to sample INBs from treated and untreated populations with the same underlying distribution of confounders, and (3) apply ROC regression techniques developed by Alonzo and Pepe (2002) to estimate NBS. Understanding how patient characteristics influence cost-effectiveness can help policy makers allocate resources towards groups which are most likely to benefit from treatment. In the future, flexible modeling techniques present a potential route to reduce the risk of model misspecification. Bayesian nonparametric methods for zero-inflated data have been developed by Oganisian, Mitra, and Roy (2018) to estimate cost distributions, but extensions of this methodology to cost-effectiveness analyses have not been explored. Similarly, ensemble learning algorithms such as Super Learner (Laan, Polley, and Hubbard, 2007) could be used to model complex cost and effectiveness distributions. For example, in their doubly robust estimator of NMB, Li et al. (2018) successfully use Super Learner to estimate both propensity score and outcome models. In future work, it would be useful to determine the best ways to incorporate these methods into regression of NBS. Additionally, when unmeasured confounding is present, estimates of NBS may be affected. In our simulation studies, we found that as the strength of confounding increases, estimates of NBS become more biased. Methods to assess the sensitivity of NBS to unmeasured confounding may also be useful avenue for future research.

For our second project, we present a two-step procedure for identifying an optimally cost-effective dynamic treatment regime with interpretable, list-based decision rules. In the first step, we use a novel Q-learning and policy search-based approach to estimate optimal list-based regimes that maximize treatment effectiveness under a predefined constraint on treatment costs. Through simulations we show the validity of our cost constrained regimes and illustrate how flexible ensemble learners can improve regime performance. The second step consists of a cost-effectiveness analysis that can select an optimally cost-effective regime from a set of candidate regimes characterized by variable cost constraints. This procedure identifies the most cost-effective treatment strategy by iteratively comparing each candidate regime with the best existing alternative. Recently, Linn, Laber, and Stefanski (2017) developed methods for optimizing specified quantiles of a distribution. Future extensions of our work may explore how this methodology could be used to allow for maximization of median survival time while placing a constraint on overall cost. Additionally, the modeling choice for Q-functions may influence estimated treatment rules. Flexible methods like SuperLearner may improve regime performance, but are also subject to variability across random seeds. For estimates of treatment effects, Benkeser (2020) propose combining estimates arising from different seeds. However, work is required to determine the best procedure for combining decision-list based rules.

For our third project, we propose two instrumental variable estimators of the local average treatment effect of an exposure on restricted mean medical costs. Our estimators are built upon inverse probability of censoring weighting techniques, which can be used to eliminate bias associated with informative cost censoring. In studies where medical cost information is recorded at multiple points, our partitioned cost estimator can improve efficiency by utilizing available cost information on all units rather than on only those with fully observed cost data. In addition, our proposed simple and partitioned weighted estimators are built upon influence function theory. Under a set of entropy conditions, this allows for flexible and efficient semiparametric estimation of treatment effects and valid asymptotic inference based upon the central limit theorem. In future work, this methodology could potentially be used to conduct flexible and efficient cost-effectiveness analyses in settings with unmeasured confounding. The development of methodology for estimating the effectiveness and cost of a treatment using instrumental variables can lead to straightforward estimation of measures of cost-effectiveness such as the Net Monetary Benefit. Given two proposed interventions, the net monetary benefit (NMB) is obtained by subtracting the observed difference in medical costs

from the product of a willingness-to-pay parameter and the difference in observed effectiveness. The NMB measures net monetary value associated with a treatment. Combination of the proposed methods with similar methods for estimating treatment effects (Lee, Kennedy, and Mitra, 2020) may be able to facilitate efficient cost-effectiveness analyses that are robust to unmeasured confounding using observational data.

This work develops novel statistical methods for performing cost-effectiveness analyses using complex observational data. The studies included in these works illustrate how our methodologies can be applied to help guide the allocation of limited resources towards patients who are most likely to benefit. These methods are tailored to be robust to confounding, informative cost censoring, and allow flexible analysis of joint outcomes data. Moving forward, we believe that this work could seed the development of nonparametric machine-learning approaches for joint outcomes such as those that arise in cost-effectiveness analyses. Because cost and effectiveness data often exhibit complex nonlinear relationships, influence function based estimators of NMB and ICER can provide robust evidence for health policy implementation.

APPENDIX A

INFLUENCE FUNCTION BASED INSTRUMENTAL VARIABLE ESTIMATOR OF CENSORED MEDICAL COSTS.

A.1. Identification of the LATE

Proof. Identification of the LATE. Recall that the LATE is defined as:

$$\begin{aligned}\psi &= \mathbb{E}[Y^{a=1} - Y^{a=0} | A^{z=1} > A^{z=0}] \\ &= \frac{\mathbb{E}[(Y^{a=1} - Y^{a=0})(A^{z=1} - A^{z=0})]}{P(A^{z=1} > A^{z=0})} \\ &= (A)/(B)\end{aligned}$$

The numerator of this expression can be identified as follows:

$$\begin{aligned}(A) &= \mathbb{E}[(Y^{a=1} - Y^{a=0})(A^{z=1} - A^{z=0})] \\ &= \mathbb{E}[(Y^{a=1} - Y^{a=0})A^{z=1} - (Y^{a=1} - Y^{a=0})A^{z=0} + Y^{a=0} - Y^{a=0}] \\ &= \mathbb{E}[(Y^{a=1} - Y^{a=0})A^{z=1} + Y^{a=0} - (Y^{a=1} - Y^{a=0})A^{z=0} - Y^{a=0}] \\ &= \mathbb{E}[Y^{A^{z=1}} - Y^{A^{z=0}}] \\ &= \mathbb{E}[Y^{z=1} - Y^{z=0}] \\ &= \mathbb{E}[\mathbb{E}[Y^{z=1}|L] - \mathbb{E}[Y^{z=0}|L]] \\ &= \mathbb{E}[\mathbb{E}[Y^{z=1}|Z=1, L] - \mathbb{E}[Y^{z=0}|Z=0, L]] \\ &= \mathbb{E}[\mathbb{E}[Y|Z=1, L] - \mathbb{E}[Y|Z=0, L]]\end{aligned}$$

Here, the fourth equality follows from the equality $Y = (Y^{a=1} - Y^{a=0})A + Y^{a=0}$, he fifth follows from the exclusion restriction, the seventh from no unmeasured confounding, and the eighth follows from

consistency. The denominator is given by:

$$\begin{aligned}(B) &= P(A^{z=1} > A^{Z=0}) \\ &= \mathbb{E}[A^{z=1} - A^{z=0}] \\ &= \mathbb{E}[\mathbb{E}[A|Z=1, L] - E[A|Z=0, L]]\end{aligned}$$

□

A.2. EIF for LATE

We wish to find the EIF for the simple weighted estimator of the LATE. Consider the following:

$$IF\left(\frac{M_s(Z=1) - M_s(Z=0)}{\Pi(Z=1) - \Pi(Z=0)}\right) = \frac{IF(M_s(Z=1) - M_s(Z=0))(\Pi(Z=1) - \Pi(Z=0)) - (M_s(Z=1) - M_s(Z=0))IF(\Pi(Z=1) - \Pi(Z=0)))}{(\Pi(Z=1) - \Pi(Z=0))^2}$$

Where

$$IF(M_s(Z=1) - M_s(Z=0)) = IF(M_s(Z=1)) - IF(M_s(Z=0))$$

and

$$IF(\Pi(Z=1) - \Pi(Z=0)) = IF(\Pi(Z=1)) - IF(\Pi(Z=0)).$$

This implies we need only find the influence functions for $M_s(Z=z)$ and $\Pi(Z=z)$. First consider the influence function for $\Pi(Z=z)$.

$$
\begin{aligned}
IF(\Pi(Z=z)) &= IF\bigg(\int \mathbb{E}[A|Z=z,L=l]p(l)dl\bigg) \\
&= \int IF(E[A|Z=z,L=l])p(l)dl + \\
&\quad \int E[A|Z=z,L=l]IF(p(l))dl \\
&= \int IF(E[A|Z=z,L=l])p(l)dl + \\
&\quad \int E[A|Z=z,L=l](\mathbb{I}(L=l)-p(l))dl \\
&= \int IF(E[A|Z=z,L=l])p(l)dl + \\
&\quad \int E[A|Z=z,L=l]\mathbb{I}(L=l))dl + \\
&\quad \int E[A|Z=z,L=l]p(l)dl \\
&= \int IF(E[A|Z=z,L=l])p(l)dl - \\
&\quad E[A|Z=z,L] - \Pi(Z=z) \\
&= \int \frac{I(Z=z,L=l)}{p(z|l)p(l)}(A-E(A|z,l)p(l)dl + \\
&\quad E[A|Z=z,L] - \Pi(Z=z) \\
&= \frac{I(Z=z)}{p(z|L)}(A-E[A|z,L]) + E[A|Z=z,L] - \Pi(Z=z)
\end{aligned}
$$

Finally, we can identify the IF for $M_s(Z=z)$ as follows:

$$
\begin{aligned}
IF(M_s(Z=z)) &= IF\bigg(\int \sum_{a=0}^{1} E\bigg[\frac{Y\delta}{G(T|a,z,l)}\bigg|a,z,l\bigg]P(A=a|z,l)p(l)dl\bigg) \\
&= \int \sum_{a=0}^{1} IF\bigg(E\bigg[\frac{Y\delta}{G(T|a,z,l)}\bigg|a,z,l\bigg]\bigg)P(A=a|z,l)p(l)dl + \qquad \text{(A.1)} \\
&\quad \int \sum_{a=0}^{1} E\bigg[\frac{Y\delta}{G(T|a,z,l)}\bigg|a,z,l\bigg]IF\bigg(P(A=a|z,l)\bigg)p(l)dl + \qquad \text{(A.2)} \\
&\quad \int \sum_{a=0}^{1} E\bigg[\frac{Y\delta}{G(T|a,z,l)}\bigg|a,z,l\bigg]P(A=a|z,l)IF\bigg(p(l)\bigg)dl \qquad \text{(A.3)}
\end{aligned}
$$

We break down this derivation into parts (1), (2), and (3). Consider (3):

$$
\begin{aligned}
(3) &= \int \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] P(A=a|z,l) IF\big(p(l)\big) dl \\
&= \int \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] P(A=a|z,l)(I(L=l)-p(l)) dl \\
&= \int \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] P(A=a|z,l) I(L=l) dl - \\
&\quad \int \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] P(A=a|z,l) p(l) dl \\
&= \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,L)}\middle|a,z,L\right] P(A=a|z,L) - M_s(Z=z)
\end{aligned}
$$

Consider (2):

$$
\begin{aligned}
(2) &= \int \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] IF\Big(P(A=a|z,l)\Big) p(l) dl \\
&= \int \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] \frac{I(Z=z,L=l)}{p(z|l)p(l)} \big(A - E[A|Z=z,l]\big) p(l) dl \\
&= \sum_{a=0}^{1} E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right] \frac{I(Z=z,L=l)}{p(z|l)} \big(A - E[A|Z=z,l]\big)
\end{aligned}
$$

Finally, consider (1):

$$
(1) = \int \sum_{a=0}^{1} IF\left(E\left[\frac{Y\delta}{G(T|a,z,l)}\middle|a,z,l\right]\right) P(A=a|z,l) p(l) dl
$$

Where

$$
\begin{aligned}
IF\Big(E\Big[\frac{Y\delta}{G(T|a,z,l)}\Big|a,z,l\Big]\Big) &= IF\Big(\int\int\int \frac{y\delta}{G(t|a,z,l)}p(y,\delta,t|a,z,l)d(y,\delta,t)\Big)\\
&= \int\int\int IF(\frac{y\delta}{G(t|a,z,l}p(y,\delta,t|a,z,l)d(y,\delta,t) +\\
&\quad \int\int\int \frac{y\delta}{G(t|a,z,l)}IF\big(p(y,\delta,t|a,z,l)\big)d(y,\delta,t)\\
&= \int\int\int \frac{-y\delta}{G(t|a,z,l)^2}IF(G(t|a,z,l))p(y,\delta,t|a,z,l)d(y,\delta,t) +\\
&\quad \int\int\int \frac{y\delta}{G(t|a,z,l)}\Big(\frac{I(a,z,l)}{p(a,z,l)}(I(y,\delta,t)-p(y,\delta,t|a,z,l))\Big)d(y,\delta,t)\\
&= \int\int\int \frac{-y\delta}{G(t|a,z,l)^2}\Big(\frac{I(a,z,l)}{p(a,z,l)}I(T=t)(\Delta-G(T|a,z,l)\Big)p(y,\delta,t|a,z,l)d(y,\delta,t) +\\
&\quad \int\int\int \frac{y\delta}{G(t|a,z,l)}\Big(\frac{I(a,z,l)}{p(a,z,l)}(I(y,\delta,t))\Big)d(y,\delta,t) -\\
&\quad \int\int\int \frac{y\delta}{G(t|a,z,l)}\Big(\frac{I(a,z,l)}{p(a,z,l)}(p(y,\delta,t|a,z,l)\Big)d(y,\delta,t)\\
&= \Big(\frac{I(a,z,l)}{p(a,z,l)}(\Delta-G(T|a,z,l)))E\Big[\frac{-Y\Delta}{G(T|a,z,l)^2}\Big]\\
&\quad \Big(\frac{Y\Delta}{G(T|a,z,l)}-E\Big[\frac{Y\Delta}{G(T|a,z,l)}\Big|a,z,l\Big]\Big)\Big(\frac{I(a,z,l)}{p(a,z,l)}\Big)\\
&= \frac{I(a,z,l)}{p(a,z,l)}\Big(\frac{Y\Delta}{G(T|a,z,l)}-E\Big[\frac{Y\Delta}{G(T|a,z,l)}\Big|a,z,l\Big]-E\Big[\frac{Y\Delta}{G(T|a,z,l)^2}\Big|a,z,l\Big](\Delta-G(T|a,z,l)\Big)
\end{aligned}
$$

So that,

$$
\begin{aligned}
(1) &= \int\sum_{a=0}^{1}\frac{I(a,z,l)}{p(a,z,l)}\Big(\frac{Y\Delta}{G(T|a,z,l)}-E\Big[\frac{Y\Delta}{G(T|a,z,l)}\Big|a,z,l\Big]-E\Big[\frac{Y\Delta}{G(T|a,z,l)^2}\Big|a,z,l\Big](\Delta-G(T|a,z,l)\Big)P(A=a|z,l)p(l)dl\\
&= \int\sum_{a=0}^{1}\frac{I(a,z,l)}{p(a,z|l)}\Big(\frac{Y\Delta}{G(T|a,z,l)}-E\Big[\frac{Y\Delta}{G(T|a,z,l)}\Big|a,z,l\Big]-E\Big[\frac{Y\Delta}{G(T|a,z,l)^2}\Big|a,z,l\Big](\Delta-G(T|a,z,l)\Big)P(A=a|z,l)dl\\
&= \sum_{a=0}^{1}\frac{I(a,z,L)}{p(a,z|L)}\Big(\frac{Y\Delta}{G(T|a,z,L)}-E\Big[\frac{Y\Delta}{G(T|a,z,L)}\Big|a,z,L\Big]-E\Big[\frac{Y\Delta}{G(T|a,z,L)^2}\Big|a,z,L\Big](\Delta-G(T|a,z,L)\Big)P(A=a|z,L)
\end{aligned}
$$

Because of this, we have the following result:

$$
\begin{aligned}
IF(M_s(Z=z)) &= (1)+(2)+(3)\\
&= \sum_{a=0}^{1}\frac{I(a,z)}{p(a,z|L)}\Big(\frac{Y\Delta}{G(T|a,z,L)}-E\Big[\frac{Y\Delta}{G(T|a,z,L)}\Big|a,z,L\Big]-E\Big[\frac{Y\Delta}{G(T|a,z,L)^2}\Big|a,z,L\Big](\Delta-G(T|a,z,L)\Big)P(A=a|z,L)+\\
&\quad \sum_{a=0}^{1}E\Big[\frac{Y\delta}{G(T|a,z,L)}\Big|a,z,L\Big]\frac{I(Z=z)}{p(z|L)}(I(A=a)-P(A|Z=z,l))+\\
&\quad \sum_{a=0}^{1}E\Big[\frac{Y\delta}{G(T|a,z,L)}\Big|a,z,L\Big]P(A=a|z,L)-M_s(Z=z)\\
&= \sum_{a=0}^{1}\frac{I(a,z)}{p(z|L)}\Big(\frac{Y\Delta}{G(T|a,z,L)}-E\Big[\frac{Y\Delta}{G(T|a,z,L)}\Big|a,z,L\Big]-E\Big[\frac{Y\Delta}{G(T|a,z,L)^2}\Big|a,z,L\Big](\Delta-G(T|a,z,L)\Big)+\\
&\quad E\Big[\frac{Y\delta}{G(T|a,z,L)}\Big|a,z,L\Big]\Big(\frac{I(Z=z)}{p(z|L)}(I(A=a)-P(A|Z=z,l))+P(A=a|z,L)\Big)-M_s(Z=z)
\end{aligned}
$$

A.3. Asymptotic Results

Consider the following:

$$\begin{aligned}\widehat{\Psi}(\widehat{\Theta}) - \Psi(\Theta) &= \frac{P_n(\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}))}{P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}))} - \frac{P(\psi_1(\Theta))}{P(\psi_2(\Theta))} \\ &= \zeta^{-1}\left\{P_n(\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \frac{P(\psi_1(\Theta))}{P(\psi_2(\Theta))}P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}))\right\}\end{aligned}$$

Where:

$$\zeta = P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}))$$

From there:

$$\begin{aligned}\widehat{\Psi}(\widehat{\Theta}) - \Psi(\Theta) &= \zeta^{-1}\left\{P_n(\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \frac{P(\psi_1(\Theta))}{P(\psi_2(\Theta))}P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}))\right\} \\ &= \zeta^{-1}\left\{P_n(\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \Psi(\Theta)P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta})) + P(\psi_1(\Theta)) - P(\psi_1(\Theta))\right\} \\ &= \zeta^{-1}\left\{P_n(\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - P(\psi_1(\Theta)) - \Psi(\Theta)\left(P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta})) - P(\psi_2(\Theta))\right)\right\} \\ &= \zeta^{-1}\Big\{P_n(\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \phi_1(\Theta) - \psi_1(\theta)) + P_n\big(\phi_1(\Theta) + \psi_1(\Theta)\big) - P(\psi_1(\Theta)) \\ &\qquad - \psi(\Theta)\left(P_n(\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}) - \phi_2(\Theta) - \psi_2(\Theta)) + P_n(\phi_2(\Theta) + \psi_2(\Theta)) - P(\psi_2(\Theta))\right)\Big\} \\ &= \zeta^{-1}\Big\{(P_n - P)[\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \phi_1(\Theta) - \psi_1(\Theta)] + P[\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \phi_1(\Theta) - \psi_1(\Theta)] \\ &\qquad - \psi(\Theta)(P_n - P)[\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}) - \phi_2(\Theta) - \psi_2(\Theta)] - \psi(\Theta)P[\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}) - \phi_2(\Theta) - \psi_2(\Theta)] \\ &\qquad + P_n[\phi_1(\Theta) + \psi_1(\Theta)] - P[\psi_1(\Theta)] + \psi(\Theta)(P_n[\phi_2(\Theta) + \psi_2(\Theta)] - P[\psi_2(\Theta)])\Big\}\end{aligned}$$

This difference can be broken down into three components as follows:

$$\widehat{\Psi}(\widehat{\Theta}) - \Psi(\Theta) = \zeta^{-1}\Big\{(P_n - P)[\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \phi_1(\Theta) - \psi_1(\Theta)] - \psi(\Theta)(P_n - P)[\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}) - \phi_2(\Theta) - \psi_2(\Theta)]\Big\} \tag{A.4}$$

$$+ \zeta^{-1}\Big\{P[\phi_1(\widehat{\Theta}) + \psi_1(\widehat{\Theta}) - \psi_1(\Theta)] - \psi(\Theta)P[\phi_2(\widehat{\Theta}) + \psi_2(\widehat{\Theta}) - \psi_2(\Theta)]\Big\} \tag{A.5}$$

$$+ \zeta^{-1}\Big\{(P_n - P)[\psi_1(\Theta) + \phi_1(\Theta)] + \psi(\Theta)(P_n - P)[\psi_2(\Theta) + \phi_2(\Theta)]\Big\} \tag{A.6}$$

BIBLIOGRAPHY

Alonzo, TA and Pepe, MS (2002). Distribution-free ROC analysis using binary regression techniques. *Biostatistics* 3.3, 421–432.

Bang, H and Tsiatis, AA (2000). Estimating Medical Costs with Censored Data. *Biometrika* 87.2, 329–343.

Bellman, R (1966). Dynamic programming. *Science* 153.3731, 34–37.

Benkeser, D (2020). *Webinar: Practical issues in targeted learning*. `https://www.youtube.com/watch?v=t0ggpi-G_Cs&t=2564s&ab_channel=PutnamDataSciences`. Accessed: 2020-06-29.

Charlson, M, Szatrowski, TP, Peterson, J, and Gold, J (1994). Validation of a Combined Comorbidity Index. *Journal of Clinical Epidemiology* 47.11, 1245–1251.

Cohen, DJ and Reynolds, MR (2008). Interpreting the results of cost-effectiveness studies. *Journal of the American College of Cardiology* 52.25, 2119–2126.

Cox, DR (1975). Partial Likelihood. *Biometrika* 62.2, 269–276.

Creutzberg, CL, Putten, WL van, Koper, PC, Lybeert, ML, Jobsen, JJ, Wárlám-Rodenhuis, CC, De Winter, KA, Lutgens, LC, Bergh, AC van den, Steen-Banasik, E van de, et al. (2000). Surgery and postoperative radiotherapy versus surgery alone for patients with stage-1 endometrial carcinoma: multicentre randomised trial. *The Lancet* 355.9213, 1404–1411.

Díaz, I (2019). Statistical inference for data-adaptive doubly robust estimators with survival outcomes. *Statistics in medicine* 38.15, 2735–2748.

Fogarty, CB, Lee, K, Kelz, RR, and Keele, LJ (2021). Biased encouragements and heterogeneous effects in an instrumental variable study of emergency general surgical outcomes. *Journal of the American Statistical Association* 116.536, 1625–1636.

Handorf, EA, Bekelman, JE, Heitjan, DF, and Mitra, N (2013). Evaluating costs with unmeasured confounding: a sensitivity analysis for the treatment effect. *The Annals of Applied Statistics* 7.4, 2062.

Handorf, EA, Heitjan, DF, Bekelman, JE, and Mitra, N (2018). Estimating cost-effectiveness from claims and registry data with measured and unmeasured confounders. *Statistical Methods in Medical Research*, 0962280218759137.

Hanley, JA (1988). The robustness of the "Binormal" Assumptions used in Fitting ROC Curves. *Medical Decision Making* 8.3, 197–203.

Heerik, ASV van den, Horeweg, N, Boer, SM de, Bosse, T, and Creutzberg, CL (2021). Adjuvant therapy for endometrial cancer in the era of molecular classification: radiotherapy, chemoradiation and novel targets for therapy. *International Journal of Gynecologic Cancer* 31.4.

Hogberg, T, Signorelli, M, De Oliveira, CF, Fossati, R, Lissoni, AA, Sorbe, B, Andersson, H, Grenman, S, Lundgren, C, Rosenberg, P, et al. (2010). Sequential adjuvant chemotherapy and ra-

diotherapy in endometrial cancer–results from two randomised studies. *European Journal of Cancer* 46.13, 2422–2431.

Keele, L, Sharoky, CE, Sellers, MM, Wirtalla, CJ, and Kelz, RR (2018). An instrumental variables design for the effect of emergency general surgery. *Epidemiologic Methods* 7.1.

Keele, L, Zhao, Q, Kelz, RR, and Small, D (2019). Falsification tests for instrumental variable designs with an application to tendency to operate. *Medical care* 57.2, 167.

Kennedy, EH (2016). Semiparametric theory and empirical processes in causal inference. In: *Statistical causal inferences and their applications in public health research*. Springer, 141–167.

Koenker, R and Bassett Jr, G (1978). Regression Quantiles. *Econometrica: Journal of the Econometric Society*, 33–50.

Laan, MJ Van der, Polley, EC, and Hubbard, AE (2007). Super learner. *Statistical applications in genetics and molecular biology* 6.1.

Laan, MJ Van der and Rose, S (2011). *Targeted learning: causal inference for observational and experimental data*. Springer Science & Business Media.

Laber, EB, Wu, F, Munera, C, Lipkovich, I, Colucci, S, and Ripa, S (2018). Identifying optimal dosage regimes under safety constraints: An application to long term opioid treatment of chronic pain. *Statistics in Medicine* 37.9, 1407–1418.

Lakkaraju, H and Rudin, C (2017). "Learning cost-effective and interpretable treatment regimes". In: *Artificial Intelligence and Statistics*, 166–175.

Latif, NA, Haggerty, A, Jean, S, Lin, L, and Ko, E (2014). Adjuvant Therapy in Early-stage Endometrial Cancer: a Systematic Review of the Evidence, Guidelines, and Clinical Practice in the US. *The Oncologist* 19.6, 645–653.

Lee, Y, Kennedy, EH, and Mitra, N (2020). Doubly robust nonparametric instrumental variable estimators for survival outcomes. *arXiv preprint arXiv:2007.12973*.

Li, J, Handorf, E, Bekelman, J, and Mitra, N (2016). Propensity Score and Doubly Robust Methods for Estimating the Effect of Treatment on Censored Cost. *Statistics in Medicine* 35.12, 1985–1999.

Li, J, Vachani, A, Epstein, A, and Mitra, N (2018). A Doubly Robust Approach for Cost–Effectiveness Estimation from Observational Data. *Statistical Methods in Medical Research* 27.10, 3126–3138.

Lin, D, Feuer, E, Etzioni, R, and Wax, Y (1997). Estimating medical costs from incomplete follow-up data. *Biometrics*, 419–434.

Linn, KA, Laber, EB, and Stefanski, LA (2017). Interactive Q-learning for quantiles. *Journal of the American Statistical Association* 112.518, 638–649.

Lorch, SA, Baiocchi, M, Ahlberg, CE, and Small, DS (2012). The differential impact of delivery hospital on the outcomes of premature infants. *Pediatrics* 130.2, 270–278.

Mariotto, AB, Yabroff, KR, Shao, Y, Feuer, EJ, and Brown, ML (2011). Projections of the Cost of Cancer Care in the United States: 2010–2020. *Journal of the National Cancer Institute* 103.2, 117–128.

Nixon, RM and Thompson, SG (2005). Methods for Incorporating Covariate Adjustment, Subgroup Analysis and Between-Centre Differences into Cost-Effectiveness Evaluations. *Health Economics* 14.12, 1217–1229.

Oganisian, A, Mitra, N, and Roy, JA (2018). A Bayesian Nonparametric Model for Zero-inflated Outcomes: Prediction, Clustering, and Causal Estimation. *Biometrics*.

Pepe, MS (1997). A Regression Modelling Framework for Receiver Operating Characteristic Curves in Medical Diagnostic Testing. *Biometrika* 84.3, 595–608.

Pepe, MS (2000). An Interpretation for the ROC Curve and Inference using GLM Procedures. *Biometrics* 56.2, 352–359.

Piessens, R, Doncker-Kapenga, E de, Überhuber, CW, and Kahaner, DK (2012). *Quadpack: A Subroutine Package for Automatic Integration*. Vol. 1. Springer Science & Business Media.

Randall, ME, Filiaci, VL, Muss, H, Spirtos, NM, Mannel, RS, Fowler, J, Thigpen, JT, and Benda, JA (2006). Randomized phase III trial of whole-abdominal irradiation versus doxorubicin and cisplatin chemotherapy in advanced endometrial carcinoma: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 24.1, 36–44.

Schulte, PJ, Tsiatis, AA, Laber, EB, and Davidian, M (2014). Q-and A-learning methods for estimating optimal dynamic treatment regimes. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 29.4, 640.

Shaeffer, DT and Randall, ME (2005). Adjuvant Radiotherapy in Endometrial Carcinoma. *The Oncologist* 10.8, 623–631.

Shiroiwa, T, Sung, YK, Fukuda, T, Lang, HC, Bae, SC, and Tsutani, K (2010). International Survey on Willingness-to-Pay (WTP) for one Additional QALY Gained: What is the Threshold of Cost Effectiveness? *Health Economics* 19.4, 422–437.

Spieker, AJ, Illenberger, N, Roy, JA, and Mitra, N (2019). *Net benefit separation and the determination curve: a probabilistic framework for cost-effectiveness estimation*. arXiv: `1912.00039 [stat.ME]`.

Spieker, AJ, Illenberger, N, Roy, JA, and Mitra, N (2021). Net benefit separation and the determination curve: A probabilistic framework for cost-effectiveness estimation. *Statistical methods in medical research* 30.5, 1306–1319.

Spieker, AJ, Ko, EM, Roy, JA, and Mitra, N (2020a). Nested g-computation: A causal approach to analysis of censored medical costs in the presence of time-varying treatment. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 69.5, 1189–1208.

Spieker, AJ, Ko, EM, Roy, JA, and Mitra, N (2020b). Nested g-computation: a causal approach to analysis of censored medical costs in the presence of time-varying treatment. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 69.5, 1189–1208.

Spieker, AJ, Roy, J, and Mitra, N (2018). Analyzing Medical Costs with Time-dependent Treatment: The Nested g-formula. *Health Economics* 27.7, 1063–1073.

Stinnett, AA and Mullahy, J (1998). Net Health Benefits: A new Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis. *Medical Decision Making* 18.2_suppl, S68–S80.

Susumu, N, Sagae, S, Udagawa, Y, Niwa, K, Kuramoto, H, Satoh, S, and Kudo, R (2008). Randomized Phase III Trial of Pelvic Radiotherapy versus Cisplatin-based Combined Chemotherapy in Patients with Intermediate and High-risk Endometrial Cancer: A Japanese Gynecologic Oncology Group Study. *Gynecologic Oncology* 108.1, 226–233.

Willan, AR, Briggs, AH, and Hoch, JS (2004). Regression Methods for Covariate Adjustment and Subgroup Analysis for Non-Censored Cost-Effectiveness Data. *Health Economics* 13.5, 461–475.

Xing, L, Lesperance, ML, and Zhang, X (2020). Simultaneous prediction of multiple outcomes using revised stacking algorithms. *Bioinformatics* 36.1, 65–72.

Xu, Y, Greene, TH, Bress, AP, Sauer, BC, Bellows, BK, Zhang, Y, Weintraub, WS, Moran, AE, and Shen, J (2020). Estimating the optimal individualized treatment rule from a cost-effectiveness perspective. *Biometrics*.

Zhang, Y, Laber, EB, Davidian, M, and Tsiatis, AA (2018). Interpretable dynamic treatment regimes. *Journal of the American Statistical Association* 113.524, 1541–1549.