Title
Towards Understanding Treatment Effect Heterogeneity

Author
Wei, Linqing

Publication Date
2023

Peer reviewed|Thesis/dissertation

Towards Understanding Treatment Effect Heterogeneity

by

Linqing Wei

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biostatistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jingshen Wang, Chair
Professor Alan E. Hubbard
Professor Zeyu Zheng

Spring 2023

Towards Understanding Treatment Effect Heterogeneity

Abstract

Towards Understanding Treatment Effect Heterogeneity

by

Linqing Wei

Doctor of Philosophy in Biostatistics

University of California, Berkeley

Professor Jingshen Wang, Chair

Understanding treatment effect heterogeneity has been an increasingly important task in various fields. Treatment effect heterogeneity not only adds granularity to the understanding of everyday matters, but also assists better-informed decision making on many scientific frontiers. In biomedical studies, learning treatment effect heterogeneity helps clinicians to apply personalized treatments to patient subpopulations with different genetic profiles. Instead of prescribing one drug for all, refined prescription strategies can potentially improve patients' overall welfare. In social science studies, evaluating the treatment effect heterogeneity of candidate policies provides guidance for policy makers to implement future social programs. In technology companies, understanding treatment effect heterogeneity helps decision makers to depict market segregation, so that advertisement budgets can be strategically allocated to particular consumer subpopulations among which a new product is more likely to earn profits.

This dissertation provides a set of statistical methodologies for understanding treatment effect heterogeneity and is organized into three chapters with three separate aims: (1) estimating treatment effect heterogeneity, (2) confirming treatment effect heterogeneity, and (3) designing adaptive experiments toward learning treatment effect heterogeneity

Chapter 1 introduces a statistical methodology aiming to estimate treatment effect heterogeneity efficiently. We take a model-free semiparametric perspective and aim to efficiently evaluate the heterogeneous treatment effects of multiple subgroups simultaneously under the one-step targeted maximum-likelihood estimation framework. When the number of subgroups is large, we further expand this path of research by looking at a variation of the one-step TMLE that is robust to the presence of small estimated propensity scores in finite samples.

Chapter 2 proposes a statistical methodology for confirming the estimated heterogeneous treatment effects. Understanding the impact of the most effective treatments on an outcome variables is crucial in various disciplines. Due to the widespread winner's curse phenomenon, conventional statistical inference assuming that the top policies are chosen independent of the random sample may lead to overly optimistic evaluations of the best policies. In addition,

given the increased availability of large datasets, such an issue can be further complicated when researchers include many covariates to estimate the policy or treatment effects in an attempt to control for potential confounders. To simultaneously address the above-mentioned issues, we propose a resampling-based procedure that not only lifts the winner's curse in evaluating the best policies observed in a random sample, but also is robust to the presence of many covariates. The proposed inference procedure yields accurate point estimates and valid frequentist confidence intervals that achieve the exact nominal level as the sample size goes to infinity for multiple best policy effect sizes.

Chapter 3 provides an alternative perspective of studying the treatment effect heterogeneity. While much of the existing work in this research area has focused on either analyzing observational data based on untestable causal assumptions or conducting post hoc analyses of existing randomized controlled trial data, little work has gone into designing randomized experiments specifically for uncovering treatment effect heterogeneity. In this chapter, we develop a unified adaptive experimental design framework towards better learning treatment effect heterogeneity by efficiently identifying subgroups with enhanced treatment effects from a frequentist viewpoint. The adaptive nature of our framework allows practitioners to sequentially allocate experimental efforts adapting to the accrued evidence during the experiment. The resulting design framework can not only complement A/B tests in e-commerce but also unify enrichment designs and response adaptive randomization designs in clinical settings. Our theoretical investigations illustrate the trade-offs between complete randomization and our adaptive experimental algorithms.

To my parents and friends.

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to thank Professor Jingshen Wang for her mentorship and guidance. Working with Jingshen has been a great journey. I have been consistently impressed by the way she approaches scientific problems. Her creativity, work ethics, and scientific rigor have influenced me in all the positive ways and made the past four years an incredibly fulfilling period of time. I have received tremendous amount of support and learned a ton from Jingshen, and I feel extremely fortunate to be her student. I want to thank Professor Alan Hubbard, who offered me the first research opportunity at Berkeley. With the knowledge and support from Alan, I gained invaluable experience in applied research and had the opportunity of collaborating with exceptional health care professionals. Additionally, I am grateful to Professor Zeyu Zheng for his expertise and professionalism in operations research. His guidance and suggestions have been invaluable to me, and I have benefitted considerably from.

I am also grateful for the incredible faculties and researchers I crossed paths with. Professor Sandra McCoy and her team's expertise in HIV research have deeply impressed me, and I feel fortunate to have the opportunity to have worked with them, and I have learned so much along the way. I want to thank Professor Xuming He for his guidance for junior researchers. In addition, I cherish the valuable opportunities to have worked with all the talented collaborators. I would like to express my sincere gratitude to Mark van der Laan, Maya Petersen, Chong Wu, Xinzhou Guo, Molei Liu, Tianxi Cai, and Yuqing Zhou for their generous help and contributions.

Furthermore, I want to thank my colleagues and friends for their unwavering support throughout my graduate studies. I am fortunate to have worked alongside so many wonderful colleagues. I want to thank Yue You, Wilson Cai, and Jonathan Levy for their valuable advice and encouragement during the early years of graduate school. I want to thank Yutong Wang for all the discussions we had and the fun times we spent together. Besides my biostatistics friends, I am also thankful for my friends who work in other fields. I appreciate all the opportunities, help, and suggesstions they offered at various stages of graduate school.

Lastly, I want to thank my parents and Yida for their unconditional support throughout the years and for making this journey possible.

Chapter 1

Efficient Estimation of Heterogeneous Treatment Effects for Multiple Subgroups

1.1 Introduction

Motivation and our contribution

In biomedical studies with observational data, investigators often aim to assess the heterogeneity of treatment effects in subpopulations of patients. Such analyses may provide useful information for patient care and for future medical research. For example, existing studies suggest that statins–a class of commonly prescribed coronary artery disease (CAD) drugs for lowering low-density lipoprotein cholesterol concentration–may reduce Alzheimer's disease (AD) risk in some, but not all population ([188]). Understanding the heterogeneous treatment effects of statin usage may provide new insights for personalizing drug prescriptions to prevent AD.

In this chapter, we aim to make valid inference on heterogeneous treatment effects in a user-supplied family of subgroups after adjusting for potential confounding factors with state-of-the-art machine learning algorithms. Motivated by our case study (Section 1.7), we work under the setting that the treatment and outcome variables are binary. The extension of our method to continuous outcomes is discussed in Appendix 1.9. Our parameter of interest includes relative risk under a treatment versus a control in d pre-specified subgroups of interest: $\boldsymbol{\alpha}_{\mathbf{RR}} = (\alpha_{\mathrm{RR},1}, \dots \alpha_{\mathrm{RR},d})^{\intercal}$, $\alpha_{\mathrm{RR},j} = \frac{P\left(Y(1)=1|X\in\mathcal{A}_j\right)}{P\left(Y(0)=1|X\in\mathcal{A}_j\right)}$, $j = 1, \dots, d$, where $P\big(Y(1) = 1|X \in \mathcal{A}_j\big)$ (or $P\big(Y(0) = 1|X \in \mathcal{A}_j\big)$) is the conditional expectations of the potential outcome under treatment (or control) evaluated in the subgroup $\mathcal{A}_j$. We denote $X \in \mathbb{R}^p$ as the potential confounders, and denote $\{\mathcal{A}_j\}_{j=1}^d$ as pre-specified possibly overlapped subgroups. We work under the classical semi-parametric inference framework, in which we aim to make

inference on the low-dimensional target parameter $\boldsymbol{\alpha}_{\mathbf{RR}}$ in the presence of high-dimensional nuisance parameters (see Section 1.4 for rigorous statements).

In this context, two potential issues emerge when one evaluates the treatment effects for multiple subgroups. On the one hand, while a commonly used method is to serially divide individuals into subgroups based on relevant pre-treatment characteristics and then estimate the treatment effect in each subgroup with either the (augmented) inverse propensity score weighting ([137]) or the targeted maximum likelihood estimator (TMLE) ([104]), this "one-group-at-a-time" approach can be computationally costly (see Section 1.3 for a concrete example). On the other hand, when the estimated propensity scores or subgroup proportions are close to zero or one in finite samples (a phenomenon refereed to as "practical positivity violation" in [127]), such approaches can be numerically unstable due to the inverse propensity score or inverse subgroup proportion weights tending to infinity.

To address such potential issues, we work with a one-step targeted maximum likelihood estimator that "targets" multiple subgroup treatment effects simultaneously. The so-called "targeting" step here involves fluctuating the initial plug-in estimator of the nuisance parameters in semiparametric models in directions which maximally adjust those initial estimates per change in the log-likelihood. Furthermore, we propose a variation of the one-step TMLE that not only targets multiple subgroups simultaneously but is also robust to the presence of small estimated propensity scores in finite samples. Deviating from the mainstream literature on the targeted learning, we also look into the problem from an optimization point of view, where we further demonstrate that such a variation of the one-step TMLE can be viewed as a reparametrized dual formulation of the primal optimization problem.

From our theoretical investigations, we show that the proposed estimator for multiple subgroup treatment effects attains the semiparametric efficiency bound, and it converges in distribution to a multivariate Gaussian distribution when the sample size becomes large. This result thus allows us to construct valid simultaneous confidence intervals and develop powerful multiple testing procedures fully utilizing the joint dependence among the subgroup specific test statistics. In addition to these large sample guarantees, through simulation studies, we demonstrate that the proposed estimator has substantial finite sample improvements relative to either applying the classical targeted learning approach [103] or the "double machine learning" frequently adopted in the econometrics literature [33]. From an application point of view, leveraging the observational data collected from the UK Biobank study, we analyze the differential effects of inheriting rs12916-T allele (a proxy for statin usage) in decreasing AD risk across multiple subgroups.

Related literature

The proposed method builds on the foundation of the targeted learning framework which is, broadly speaking, a meta-learning framework allowing various machine learning algorithms to enter the process of estimating desired target parameters ([103]). [104] propose the original version of TMLE, which uses maximum likelihood in a least favourable direction and then performs k-step updates using the estimated scores, in an effort to better estimate the target

parameter. [186] introduce the cross-validated TMLE, which relaxes the stringent Donsker condition via sample splitting for the initial estimation of the nuisance parameters. [100] further advances the original TMLE by designing different sets of candidate scores. A recent advancement in the targeted learning framework is the one-step TMLE ([99]), which adopts a "universal least favorable submodel" to avoid excessive data fitting in the locally least favorable submodel. In terms of estimating a vector of multi-dimensional parameters with TMLE, seminal works by [103] and [99] develop a universal canonical one-dimensional submodel such that the one-step TMLE, only maximizing the log-likelihood over a univariate parameter, solves the multivariate efficient influence curve equation. A recent work ([110]) adopts this general TMLE approach for estimating the variance of the stratum-specific treatment effect functions. We also note that the general strategy of TMLE that targets multi-dimensional parameters have also been discussed for estimating survival curves (see, [102], Chapter 5 for example).

Our proposal contributes to the semiparametric statistics literature. Early work on semiparametric statistics ([123]) provides general efficiency results for the development of semiparametric estimators. Based on these efficiency results, [135] propose a general estimating equation approach that solves for the parameter of interest by setting the efficient score equations to zero. The estimating equation approach is further discussed in [101]. [21] develop a one-step estimator that adds the empirical average of the efficient influence function to an initial estimator. [124] advances the semiparametric efficiency results by accounting for the nonparametric estimation of nuisance parameters. [161] discusses the use of maximum likelihood estimator and parametric submodel in semiparametric estimation.

Our work is also tied to the literature on heterogeneous treatment effect estimation in causal inference. Different from our parameter of interest, [31], building on the debiased double machine learning framework ([33]), propose to estimate the average treatment effect conditional on a small subset of the potential confounders. [98] and [3] propose meta-learning frameworks that estimates the average treatment conditional on all possible confounders. Unlike our approach, which efficiently evaluates the treatment effects in pre-specified subgroups, [86] formulate the problem on heterogeneous treatment effect identification from a variable selection perspective. In this thread on heterogeneity identification, [154] propose a recursive partitioning tree approach to identify treatment heterogeneity across subgroups. [162] provide a nice overview of subgroup selection problems encountered in practice.

1.2 Causal Framework and Identification

Let $\{O_i\}_{i=1}^n = \{(Y_i, T_i, X_i)\}_{i=1}^n$ be an independent and identically distributed (i.i.d.) random sample of the observed binary response variable Y, the treatment indicator variable T, and potential confounders $X \in \mathbb{R}^p$. In accordance with the Neyman-Rubin causal model ([125, 148]), we define the potential outcome $Y(T)$ as the outcome we would have observed under the treatment assignment T. The observed outcome is thus the potential outcome variable corresponding to the received treatment, i.e., $Y = TY(1) + (1-T)Y(0)$. This framework

allows us to characterize the multi-subgroup disease risk under different treatment arms as: $\boldsymbol{\alpha}_t = (\alpha_{t,1}, \dots \alpha_{t,d})^\intercal$, $\alpha_{t,j} = P\big(Y(t) = 1 | X \in \mathcal{A}_j\big)$, $t \in \{0,1\}$, $\quad j = 1, \dots, d$, where $\mathcal{A}_j$ denotes a pre-specified subgroup j. Here, we allow different subgroup to overlaps, and we assume that the variables used to define the subgroups of interest are based on X. When comparing disease risks between two treatment arms, our framework allows practitioners to estimate three popular causal effect measures: relative risk, odds ratio, and absolute risk difference, across different subgroups, defined as $\boldsymbol{\alpha}_{\mathbf{RR}} = (\alpha_{\text{RR},1}, \dots \alpha_{\text{RR},d})^\intercal$, $\alpha_{\text{RR},j} = \alpha_{1,j}/\alpha_{0,j}$, $\boldsymbol{\alpha}_{\mathbf{OR}} = (\alpha_{\text{OR},1}, \dots \alpha_{\text{OR},d})^\intercal$, $\alpha_{\text{OR},j} = \big(\alpha_{1,j}/(1-\alpha_{1,j})\big)/\big(\alpha_{0,j}/(1-\alpha_{0,j})\big)$, and $\boldsymbol{\alpha}_{\mathbf{ARD}} = (\alpha_{1,1} - \alpha_{0,1}, \dots, \alpha_{1,d} - \alpha_{0,d})$ (Section 1.3).

The three causal quantities described above are not observable because the potential outcomes are subject to missingness, meaning that for each individual we observe either the potential outcome under the control, $Y(0)$, or the potential outcome under the treatment, $Y(1)$, but never both. Following the mainstream literature in causal inference, we impose the unconfoundedness, positivity, and stable unit treatment value assumptions (SUTVA) below to identify our causal parameters of interest:

Assumption 1 (Unconfoundedness)**.** *Conditional on X, the treatment assignment is as good as random, that is $T \perp Y(1), Y(0) | X$.*

Assumption 2 (Positivity)**.** *For any $x \in X$, $t \in \{0,1\}$, there exists a constant $c \in (0,1)$ such that $c < P(T = t | X = x, X \in \mathcal{A}_j) < 1 - c$ and $c < P(\mathcal{A}_j) < 1 - c$, for $j = 1, \dots, d$.*

Assumption 3 (SUTVA)**.** *If unit i receives treatment T_i, the observed outcome Y_i equals the potential outcome $Y_i(T_i)$, meaning that the potential outcome for unit i under treatment T_i is unrelated to the treatment received by other units.*

Under Assumption 1-3, we are able to identify $\alpha_{t,j}$ as $\alpha_{t,j} = P\big(Y(1) = 1 | X \in \mathcal{A}_j\big) = E_X\big[P(Y = 1 | T = t, X \in \mathcal{A}_j)\big]$. Here, by "identify" we mean that under Assumption 1, the causal effect involving unobserved potential outcomes can be first written as a function of observed data. Then, within an i.i.d. sample $\{(Y_i, T_i, X_i)\}_{i=1}^n$, under Assumption 2 and 3, the causal parameter can be estimated (or point identified) at a regular parametric root-n rate [94].

Notation We use P to denote the probability operator and E to denote the expectation operator. We use capitalized letters to denote random variables, e.g. T, and lower case letters to denote the realizations of random variables, e.g. t. For $t \in \{0,1\}$, we denote $p_t(X) = P(Y = 1 | T = t, X)$ as the conditional probability of $Y = 1$ given $T = t$ and X. $e_t(X) = P(T = t | X)$ denotes the conditional probability of $T = t$ given X. Lastly, we define $\text{expit}(x) = \frac{1}{1+e^{-x}}$ and $\text{logit}(x) = \log(\frac{x}{1-x})$.

1.3 Multiple Subgroup Targeted Learning

In this section, to simplify presentation, we first introduce our method on estimating the conditional average risk $\boldsymbol{\alpha}_t$ for group $t \in \{0,1\}$ and defer the estimation for other causal

parameters to Section 1.3 and Appendix 1.9. We shall review the classical one-step targeted maximum likelihood estimator (TMLE) ([99]) in a single subgroup case, followed by discussing its limitations when naively generalizing it to the multi-subgroup case. We then introduce the one-step TMLE that directly targets the multi-subgroup treatment effects simultaneously.

Limitation of the classical one-step TMLE

To estimate $\boldsymbol{\alpha}_t$, a natural choice is to apply the one-step TMLE in each subgroup separately. For a subgroup j, one-step TMLE starts with some initial estimates of $p_t(X)$ and $e_t(X)$ using the observations in the subgroup $\mathcal{A}_j$, denoted as $\hat{p}_{tj}^{\texttt{Init}}(X)$ and $\hat{e}_{tj}(X)$. These initial estimates can be obtained from any state-of-art machine learning methods–such as random forest, gradient boosting ([22]), or Highly Adaptive Lasso (HAL) ([20])–as long as they are not too far away from the target estimands (see Assumption 5 in Section 1.4 for rigorous specifications). Within a random sample, because $\hat{p}_{tj}^{\text{Init}}(X)$ and $\hat{e}_{tj}(X)$ may substantially deviate from the truth, the targeted learning approach identifies a correction term, $\hat{\varepsilon} \cdot \hat{S}_{tj}(X)$, that pushes the initial estimates to "concentrate/target" on the estimand: $\hat{p}_{tj}(X_i) = \text{expit}\Big(\text{logit}\big(\hat{p}_{tj}^{\text{Init}}(X_i)\big) + \hat{\varepsilon} \cdot \hat{S}_{tj}(X_i)\Big)$, $\hat{S}_{tj}(X_i) = \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}_{tj}(X_i)}$. Here, $\hat{P}(\mathcal{A}_j) = \frac{\sum_{i=1}^n \mathbb{1}(X_i \in \mathcal{A}_j)}{n}$, $\hat{\varepsilon}$ captures the magnitude of the correction $\hat{S}_{tj}(X_i)$ (so called "clever covariate" in [104]), and it is the estimated coefficient of $\hat{S}_{tj}(X_i)$ in the logistic regression:

$$Y_i \sim \text{logit}\big(\hat{p}_{tj}^{\text{Init}}(X_i)\big) + \varepsilon \hat{S}_{tj}(X_i), \quad i \in \mathcal{A}_{tj}, \tag{1.1}$$

that regresses Y_i on $\text{logit}\big(\hat{p}_{tj}^{\text{Init}}(X_i)\big)$ and $\hat{S}_{tj}(X_i)$ with a fixed coefficient 1 for $\text{logit}\big(\hat{p}_{tj}^{\text{Init}}(X_i)\big)$. Here $\mathcal{A}_{tj} = \mathcal{A}_j \cap \{i : T_i = t\}$ contains the subjects with $T_i = t$ in the subgroup $\mathcal{A}_j$. After this one-step correction, the final estimate $\hat{\alpha}_{t,j}^{\text{one-step}}$ takes the empirical average of $\hat{p}_{tj}(X_i)$: $\hat{\alpha}_{t,j}^{\text{one-step}} = \frac{1}{n_{tj}} \sum_{i=1}^n \hat{p}_{tj}(X_i)$, where n_{tj} is the cardinality of the set $\mathcal{A}_{tj}$.

The regression problem defined in Eq (1.1) is the essence of the one-step TMLE. Such a regression problem adaptively learns the difference between $\hat{p}_{tj}^{\text{Init}}(\cdot)$ and $p_{tj}(\cdot)$ from the data, aiming to find an $\hat{\varepsilon}$ that locally improves the empirical fit of the initial estimator $\hat{p}_{tj}^{\text{Init}}(\cdot)$. We choose $\hat{\varepsilon}$ in a data adaptive fashion because when the initial estimate of the conditional probability is identical to the true conditional probability, we hope to set $\hat{\varepsilon} = 0$. It is only when the initial estimate $\hat{p}_{tj}^{\text{Init}}(\cdot)$ drifts away from $p_{tj}(\cdot)$, $\hat{\varepsilon}$ accounts for their difference and updates $\hat{p}_{tj}^{\text{Init}}(\cdot)$ accordingly. Furthermore, because our goal is to estimate $\alpha_{t,j}$, the clever covariate $S_{tj}(X_i)$ specifies the updating direction of the initial estimator that yields a maximal change (or maximal information gain) in the target parameter. Benefiting from such an update, the final estimator $\hat{\alpha}_{t,j}^{\text{one-step}}$ attains the semiparametric efficiency bound under the regularity conditions in Section 1.4. In addition, because the one-step TMLE applies an "expit" transformation on the sum of $\text{logit}\big(\hat{p}_{tj}^{\text{Init}}(X_i)\big)$ and the inverse propensity score, the estimated conditional risk $\hat{\alpha}_{t,j}^{\text{one-step}}$ never falls out of the range between 0 and 1 regardless of how small $\hat{e}_{tj}(\cdot)$ is (see Section 1.6 for numerical verification).

Table 1.1: Computational time (in seconds) of the conventional TMLE and the proposed method with sample size $n = 228,466$ on a Lenovo NeXtScale nx360m5 node (24 cores per node) equipped with Intel Xeon Haswell processor. The core frequency is 2.3 Ghz and supports 16 floating-point operations per clock period.

Classical one-step TMLE	iTMLE
1441.36	924.51

Nevertheless, naively carrying out the above procedure one subgroup at a time can be computationally inefficient in the presence of many subgroups. In a simple comparison provided in Table 1.1, our proposed estimator directly targeting the multi-subgroup parameter $\boldsymbol{\alpha}_t$ as a whole improves the computational speed by about 35% compared to this one-group-at-a-time approach, when the initial estimator $\hat{p}_{tj}^{\text{Init}}(\cdot)$ and the estimated propensity scores $\hat{e}_{tj}(\cdot)$ are obtained via GLMs.

One-step TMLE targeting multiple subgroups

Procedure overview

To avoid the discussed potential problems of the conventional one-step TMLE, we amend the one-step TMLE estimator so that it directly targets $\boldsymbol{\alpha}_t$. A natural idea is to replace the univariate clever covariate with a multi-dimensional vector of clever covariates $\big(\hat{S}_{t1}(X_i), \ldots, \hat{S}_{td}(X_i)\big)^\intercal$ in the logistic regression

$$Y_i \sim \text{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{d} \varepsilon_{t,j} \cdot \hat{S}_{tj}(X_i), \quad i \in \{i : T_i = t\}, \tag{1.2}$$

where $\hat{S}_{tj}(X_i) = \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}_t(X_i)}$. Note that here we generate the initial estimates $\hat{p}_t^{\text{Init}}(X_i)$ and $\hat{e}_t(X_i)$ with the entire available sample. We then construct the estimator for $\boldsymbol{\alpha}_t$ with

$$\hat{\boldsymbol{\alpha}}_t^{\text{one-step}} = \big(\frac{1}{n_{t1}} \sum_{i=1}^{n} \hat{p}_{t1}(X_i), \ldots, \frac{1}{n_{td}} \sum_{i=1}^{n} \hat{p}_{td}(X_i)\big)^\intercal, \tag{1.3}$$

where $\hat{p}_{tj}(X_i) = \text{expit}\Big(\text{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \hat{\varepsilon}_{t,j} \cdot \hat{S}_{t,j}(X_i)\Big)$.

In the presence of multiple subgroups with large d, we may observe small $\hat{P}(A_j)$ or $\hat{e}_t(X_i)$ within a random sample. In this situation, given that $\hat{P}(A_j)$ and $\hat{e}_t(X_i)$ enter the regression problem in Eq (1.2) as denominators, the above procedure can potentially produce numerically unstable estimates, which may inflate the variance of $\hat{\boldsymbol{\alpha}}_t^{\text{one-step}}$. We hope to further robustify the above procedure by considering a simple variation, where we shall also demonstrate that the algorithm proposed below is a reparametrized dual problem of the above (primal) problem defined in Eq (1.2).

Our proposed procedure operates as follows, for each iteration k,

$$
\begin{aligned}
&Y_i \sim \text{logit}\big(\hat{p}_t^{(k-1)}(X_i)\big) + \gamma \tilde{S}_t^{(k-1)}(X_i), \\
&\hat{p}_t^{(k)}(X_i) = \text{expit}\Big(\text{logit}\big(\hat{p}_t^{(k-1)}(X_i)\big) + \hat{\gamma}^{(k)} \cdot \tilde{S}_t^{(k-1)}(X_i)\Big), \quad i \in \{i : T_i = t\}, \quad k = 1, \ldots, K,
\end{aligned} \tag{1.4}
$$

where $\hat{\gamma}^{(k)}$ is the estimated regression coefficient obtained in the logistic regression (1.4). $\hat{p}_t^{(1)}(X_i)$ denotes the initial estimate. $\hat{p}_t^{(k-1)}(X_i)$ denotes the estimate from the previous iteration, and $\tilde{S}_t^{(k-1)}(X_i)$ is the customized "clever covariate" that directly targets $\boldsymbol{\alpha}_t$:

$$
\tilde{S}_t^{(k-1)}(X_i) = \frac{\sum_{j=1}^{d} \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}_t(X_i)} \cdot \Big(\sum_{l=1}^{n} \hat{\phi}_j^{(k-1)}(Y_l, T_l, X_l) \Big)}{\sqrt{\sum_{j=1}^{d} \Big(\sum_{l=1}^{n} \hat{\phi}_j^{(k-1)}(Y_l, T_l, X_l) \Big)^2}}, \tag{1.5}
$$

where $\hat{\phi}_j^{(k-1)}(Y_i, T_i, X_i) = \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}_t(X_i)} (Y_i - \hat{p}_t^{(k-1)}(X_i))$. The intuition of $\tilde{S}_t^{(k-1)}(X_i)$ shall be explained in the next section. When the maximum number of iterations K is reached or when $\hat{\gamma}$ is sufficiently close to 0, we take the final estimate $\hat{p}_t(X_i) = \hat{p}_t^{(K)}(X_i)$ and estimate $\boldsymbol{\alpha}_t$ again with:

$$
\hat{\boldsymbol{\alpha}}_t = \Big(\frac{\sum_{i \in \mathcal{A}_1} \hat{p}_t(X_i)}{n_{t1}}, \ldots, \frac{\sum_{i \in \mathcal{A}_d} \hat{p}_t(X_i)}{n_{td}} \Big)^{\intercal}, \tag{1.6}
$$

where $n_{tj} = \sum_{i=1}^{n} \mathbb{1}(T_i = t) \mathbb{1}(X_i \in \mathcal{A}_j)$ denotes the subgroup j's sample size in the arm t.

We refer to the estimator in Eq (1.6), which is obtained from Eq (1.4), as the iterative version of the one-step TMLE (iTMLE) targeting multiple subgroups of interest.

Intuitive explanation of our proposal

Note that although the proposed estimators in Eq (1.3) and Eq (1.6) are asymptotically equivalent as $n \to \infty$, we provide some heuristic explanations of the benefits of adopting our procedure defined in Eq (1.4) compared to the procedure defined in Eq (1.2) in finite samples.

First, given that the performance of the one-step TMLE defined by Eq (1.2) depends on the initial estimator $\hat{p}_t^{\text{Init}}(X_i)$, our revised procedure in Eq (1.4) works with an improved initial estimator in each iteration. Concretely, in Eq (1.4), the initial estimator entering each iteration is constantly being updated, leading to increased estimation efficiency and reduced estimation bias compared to the procedure defined in Eq (1.2). Such improvements can be rather prominent in finite samples (See Appendix 1.9 for simulation comparisons).

Second, the form of the clever covariate $\tilde{S}_t(X_i)$ in Eq (1.4) may have the added benefit of being robust to the presence of small estimated propensity scores, because the estimated propensity scores only enter the estimation process after being self-normalized in $\tilde{S}_t(X_i)$. Small propensity scores are often encountered in datasets with unbalanced covariate distribution across the treatment and control groups. Such an imbalance can lead to conventional

estimators having substantial biases and large variances [37, 127]. Many numerical studies have found that similar self-normalization of propensity scores provides much more stable estimates of the treatment effects in finite samples ([70]). While the original formulation of the primal problem in Eq (1.2) involves a sum over d inverse propensity score weighted clever covariates, its performance can be sensitive to the presence of small propensity scores in finite samples. Even though the estimator obtained by Eq (1.4) and the estimator obtained by Eq (1.2) are asymptotically equivalent, the estimator obtained by Eq (1.4) may have finite sample improvements when the estimated propensity scores are small.

Third, the estimator obtained from Eq (1.4) not only remains semi-parametric efficient and "doubly robust," but also solves the direct sample analogue of the efficient influence function. To see why it is semiparametric efficient, we set the derivative of the objective function of the logistic regression in (1.2) with respect to ε to zero, which reduces to (see Appendix 1.9 for detailed derivations)

$$\sum_{j=1}^{d}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)}\frac{T_i}{\hat{e}_t(X_i)}(Y_i-\hat{p}_t(X_i))\right)^2=0. \tag{1.7}$$

This indicates that our estimator $\hat{\boldsymbol{\alpha}}_t=(\hat{\alpha}_{t,1},\ldots,\hat{\alpha}_{t,d})^\intercal$ solves the direct sample analogue of the efficient influence function: $\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)}\left\{\frac{T_i}{\hat{e}_t(X_i)}(Y_i-\hat{p}_t(X_i))+\hat{p}_t(X_i)\right\}-\hat{\alpha}_{t,j}=0,\ j=1,\ldots,d$. Therefore, it attains the semiparametric efficiency bound ([21]) under appropriate conditions imposed on the nuisance parameter estimators (Theorem 1). Regarding the "doubly robustness," for any model-based estimators $\hat{e}_t(\cdot)$ and $\hat{p}_t(\cdot)$, our estimator combines regression imputation and inverse propensity score weighting, and remains consistent if either the model $e_t(\cdot)$ or $p_t(\cdot)$ is misspecified (see Section 1.6 for simulation results). In Section 1.4, we shall provide further heuristic explanations of the targeted maximum likelihood estimator from a semiparametric inference point of view.

Extension to relative risk, odds ratio, and absolute risk difference estimations

Given that $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_0}$ are the building blocks of the multi-subgroup relative risk and odds ratio, estimation for these two parameters of interest largely follows our proposal in Section 1.3. The iterative version of the one-step TMLE needs a slight modification in that at each iteration k, we adopt the following logistic regression problem: $Y_i\sim\text{logit}\big(\hat{p}^{(k-1)}(T_i,X_i)\big)+\gamma_1\tilde{S}_1^{(k-1)}(X_i)+\gamma_0\tilde{S}_0^{(k-1)}(X_i),\ k=1,\ldots,K$, and perform the updating as $\hat{p}^{(k)}(T_i,X_i)=\text{expit}\Big(\text{logit}\big(\hat{p}^{(k-1)}(T_i,X_i)\big)+\hat{\gamma}_1^{(k)}\cdot\tilde{S}_1^{(k-1)}(X_i)+\hat{\gamma}_0^{(k)}\cdot\tilde{S}_0^{(k-1)}(X_i\Big)$. Then we estimate $\boldsymbol{\alpha_{RR}}$, $\boldsymbol{\alpha_{OR}}$, and $\boldsymbol{\alpha_{ARD}}$ with $\hat{\boldsymbol{\alpha}}_{\mathbf{RR}}=\left(\frac{\hat{\alpha}_{1,1}}{\hat{\alpha}_{0,1}},\ldots,\frac{\hat{\alpha}_{1,d}}{\hat{\alpha}_{0,d}}\right)$, $\hat{\boldsymbol{\alpha}}_{\mathbf{OR}}=\left(\frac{\hat{\alpha}_{1,1}}{1-\hat{\alpha}_{1,1}}\Big/\frac{\hat{\alpha}_{0,1}}{1-\hat{\alpha}_{0,1}},\ldots,\frac{\hat{\alpha}_{1,d}}{1-\hat{\alpha}_{1,d}}\Big/\frac{\hat{\alpha}_{0,d}}{1-\hat{\alpha}_{0,d}}\right)$, and $\hat{\boldsymbol{\alpha}}_{\mathbf{ARD}}=\left(\hat{\alpha}_{1,1}-\hat{\alpha}_{0,1},\ldots,\hat{\alpha}_{1,d}-\hat{\alpha}_{0,d}\right)$

As for constructing simultaneous confidence intervals, we apply the Delta method on $(\boldsymbol{\alpha_1},\boldsymbol{\alpha_0})$ to estimate the sample covariance matrices of the relative risk and the odds ratio

estimators following a recipe similar to Section 1.5. To avoid redundancy, we leave the detailed descriptions to Appendix 1.9.

1.4 Theoretical Investigations

Properties of the proposed estimator

In this section, we introduce the main theoretical results and some necessary notation. We defer additional notation and regularity conditions to Section 1.4. Recall that $\{O_i\}_{i=1}^n := \{(Y_i, T_i, X_i)\}_{i=1}^n$ is an i.i.d. random sample defined on the space $\mathcal{O}$ with respect to a probability measure P. Denote $o = (y, t, x)$ as a realized data point, $o \in \mathcal{O}$.

Theorem 1. *Under Assumptions 1-5, we define the vector of the efficient influence function* $\boldsymbol{\varphi}_t = (\varphi_{t,1}, \ldots, \varphi_{t,d})^\intercal$, *where* $\varphi_{t,j}$ *is the efficient influence function (as given in Eq (8)) measured at a realized data point* $o = (y, t, x)$ *for the subgroup* j. *The proposed conditional risk estimator* $\hat{\boldsymbol{\alpha}}_t = (\alpha_{t,1}, \ldots, \alpha_{t,d})^\intercal \in \mathbb{R}^d$, *after scaling by* $\sqrt{n}$, *converges to a multivariate Gaussian random variable with mean 0 and covariance matrix* $P[\boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\intercal]$ *when* $n \to \infty$, *that is* $\sqrt{n}\big(\hat{\boldsymbol{\alpha}}_t - \boldsymbol{\alpha}_t\big) \rightsquigarrow \mathcal{N}\Big(0, P[\boldsymbol{\varphi}_t \boldsymbol{\varphi}_t^\intercal]\Big)$. *(See the precise definition of* $\varphi_{t,j}$ *in Section 1.4).*

Theorem 1 says that our conditional risk estimator converges in distribution to a multivariate Gaussian distribution. For any subgroups under consideration, the variance of our conditional risk estimator attains the semiparametric efficiency bound. Theorem 1 also justifies the validity of the simultaneous confidence interval provided in Eq to be presented (1.18) in Section 1.5.

Derivations of the efficient influence functions for relative risk, odds ratio and absolute risk difference estimators are provided in Appendix 1.9. We summarize the large sample properties of $\boldsymbol{\alpha}_{\mathbf{RR}}$, $\boldsymbol{\alpha}_{\mathbf{OR}}$, and $\boldsymbol{\alpha}_{\mathbf{ARD}}$ in the following Proposition 1, which demonstrates that the variance of the proposed causal effect estimators attains the semiparametric efficiency bound. The proof of the proposition below can be found in Appendix 1.9.

Proposition 1. *Under Assumptions 1 - 5, define the vector of the efficient influence function*

$$\boldsymbol{\varphi}_{RR} = (\varphi_{RR,1}, \ldots, \varphi_{RR,d})^\intercal,$$

the vector of the efficient influence function

$$\boldsymbol{\varphi}_{OR} = (\varphi_{OR,1}, \ldots, \varphi_{OR,d})^\intercal,$$

and the vector of the efficient influence function

$$\boldsymbol{\varphi}_{ARD} = (\varphi_{ARD,1}, \ldots, \varphi_{ARD,d})^\intercal,$$

where $\varphi_{RR,j}$, $\varphi_{OR,j}$, *and* $\varphi_{ARD,j}$ *are the efficient influence functions (as given in Eq (9)-(11)) measured at a realized data point* $o = (y, t, x)$. *The proposed causal effect estimators satisfy that as* $n \to \infty$, $\sqrt{n}\big(\hat{\boldsymbol{\alpha}}_{RR} - \boldsymbol{\alpha}_{RR}\big) \rightsquigarrow \mathcal{N}\Big(0, P[\boldsymbol{\varphi}_{RR} \boldsymbol{\varphi}_{RR}^\intercal]\Big)$, $\sqrt{n}\big(\hat{\boldsymbol{\alpha}}_{OR} - \boldsymbol{\alpha}_{OR}\big) \rightsquigarrow$

$\mathcal{N}\Big(0, P[\boldsymbol{\varphi}_{OR}\boldsymbol{\varphi}_{OR}^{\intercal}]\Big)$ *and* $\sqrt{n}\big(\hat{\boldsymbol{\alpha}}_{ARD} - \boldsymbol{\alpha}_{ARD}\big) \rightsquigarrow \mathcal{N}\Big(0, P[\boldsymbol{\varphi}_{ARD}\boldsymbol{\varphi}_{ARD}^{\intercal}]\Big)$ *(See the precise definitions of* $\varphi_{RR,j}$*,* $\varphi_{OR,j}$*, and* $\varphi_{ARD,j}$ *in Section 1.4).*

Regularity conditions

In this section, we introduce additional notation and assumptions adopted in the theoretical results. Recall that $\{O_i\}_{i=1}^n := \{(Y_i, T_i, X_i)\}_{i=1}^n$ are i.i.d. random variables defined on the space $\mathcal{O}$ with respect to a probability measure P. If $\mathcal{F}$ is a collection of real-valued functions defined on $\mathcal{O}$, we assume that $Pf = \int f \mathrm{d}P$ exists for each $f \in \mathcal{F}$. Note that such a notation can be more helpful as it allows us to conveniently work with random functions. We use $E_X[f(X)]$ to denote the expectation taken with respect to the random variable X when it is more convenient to simplify notation. Given the probability measure P, our target parameter $\boldsymbol{\alpha}_t$ can also be written as a statistical function of P, denoted as $\boldsymbol{\alpha}_t(P)$. Let $\mathcal{H}$ be a convex set of functions such that the true nuisance parameter $\boldsymbol{\eta}_0 \triangleq (e(x), p_1(x), p_0(x), P(\mathcal{A}_1), \ldots, P(\mathcal{A}_d)) \in \mathcal{H}$. Let $\mathcal{H}_n \subset \mathcal{H}$ denote the nuisance estimator realization set, i.e., the estimator of the nuisance parameters satisfy $\hat{\boldsymbol{\eta}} = (\hat{e}_t(x), \hat{p}_1(x), \hat{p}_0(x), \hat{P}(\mathcal{A}_1), \ldots, \hat{P}(\mathcal{A}_d)) \in \mathcal{H}_n$.

Let c, q, and C be fixed strictly positive constants, where $q > 2$. Let $(\xi_n)_{n=1}^{\infty}$ and $(\Delta_n)_{n=1}^{\infty}$ be sequences of positive constants approaching 0. Denote the l_q-norm with respect to a probability measure P as $||\cdot||_{P,q}$, e.g. $||f(X)||_{P,q} := (\int |f(x)|^q dP(x))^{1/q}$. For $o \in \mathcal{O}$, we define $\boldsymbol{\varphi}_t(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0) \triangleq \big(\varphi_{t,1}, \ldots, \varphi_{t,d}\big)^{\intercal}$ as the vector of the efficient influence function for estimating $\boldsymbol{\alpha}_t$, $\boldsymbol{\varphi}_{\mathrm{RR}}(o; \boldsymbol{\alpha}_{\mathrm{RR}}, \boldsymbol{\eta}_0) \triangleq \big(\varphi_{\mathrm{RR},1}, \ldots, \varphi_{\mathrm{RR},d}\big)^{\intercal}$ as the vector of the efficient influence function for estimating $\boldsymbol{\alpha}_{\mathrm{RR}}$, $\boldsymbol{\varphi}_{\mathrm{OR}}(o; \boldsymbol{\alpha}_{\mathrm{OR}}, \boldsymbol{\eta}_0) \triangleq \big(\varphi_{\mathrm{OR},1}, \ldots, \varphi_{\mathrm{OR},d}\big)^{\intercal}$ as the vector of the efficient influence function for estimating $\boldsymbol{\alpha}_{\mathrm{OR}}$, and $\boldsymbol{\varphi}_{\mathrm{ARD}}(o; \boldsymbol{\alpha}_{\mathrm{ARD}}, \boldsymbol{\eta}_0) \triangleq \big(\varphi_{\mathrm{ARD},1}, \ldots, \varphi_{\mathrm{ARD},d}\big)^{\intercal}$ as the vector of the efficient influence function for estimating $\boldsymbol{\alpha}_{\mathrm{ARD}}$, where for $j = 1, \ldots, d$,

$$\varphi_{t,j} \triangleq \varphi_{t,j}(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0) = \frac{\mathbb{1}(x \in \mathcal{A}_j)}{P(\mathcal{A}_j)}\Big[\Big(y - p_t(x)\Big)\frac{\mathbb{1}(T = t)}{e_t(x)} + p_t(x) - \alpha_{t,j}\Big], \tag{1.8}$$

$$\begin{aligned}\varphi_{\mathrm{RR},j} \triangleq \varphi_{\mathrm{RR},j}(o; \boldsymbol{\alpha}_{\mathrm{RR}}, \boldsymbol{\eta}_0) &= \frac{\mathbb{1}(x \in \mathcal{A}_j)}{P(\mathcal{A}_j)}\Big[\frac{1}{\alpha_{0,j}}\Big(\big(y - p_1(x)\big)\frac{t}{e_1(x)} + p_1(x) - \alpha_{1,j}\Big) \\ &\quad + \frac{\alpha_{1,j}}{\alpha_{0,j}^2}\Big(\frac{1-t}{e_0(x)}\big(y - p_0(x)\big) + p_0(x) - \alpha_{0,j}\Big)\Big],\end{aligned} \tag{1.9}$$

$$\begin{aligned}\varphi_{\mathrm{OR},j} \triangleq \varphi_{\mathrm{OR},j}(o; \boldsymbol{\alpha}_{\mathrm{OR}}, \boldsymbol{\eta}_0) &= \frac{\mathbb{1}(x \in \mathcal{A}_j)}{P(\mathcal{A}_j)}\Big[\frac{1 - \alpha_{0,j}}{\alpha_{0,j}(1-\alpha_{1,j})^2}\Big(\big(y - p_1(x)\big)\frac{t}{e_1(x)} + p_1(x) - \alpha_{1,j}\Big) \\ &\quad - \frac{\alpha_{1,j}}{\alpha_{0,j}^2(1-\alpha_{1,j})}\Big(\frac{1-t}{e_0(x)}\big(y - p_0(x)\big) + p_0(x) - \alpha_{0,j}\Big)\Big].\end{aligned} \tag{1.10}$$

$$\begin{aligned}\varphi_{\mathrm{ARD},j} \triangleq \varphi_{\mathrm{ARD},j}(o; \boldsymbol{\alpha}_{\mathrm{ARD}}, \boldsymbol{\eta}_0) &= \frac{\mathbb{1}(x \in \mathcal{A}_j)}{P(\mathcal{A}_j)}\Big[\Big(\big(y - p_1(x)\big)\frac{t}{e_1(x)} + p_1(x) - \alpha_{1,j}\Big) \\ &\quad - \Big(\frac{1-t}{e_0(x)}\big(y - p_0(x)\big) + p_0(x) - \alpha_{0,j}\Big)\Big].\end{aligned} \tag{1.11}$$

Assumption 4. *The function class* $\{\boldsymbol{\varphi}(o;\boldsymbol{\alpha}_t,\boldsymbol{\eta}),\boldsymbol{\eta}\in\mathcal{H}\}$ *is a Donsker class.*

Assumption 5. *The nuisance parameter estimator* $\hat{\boldsymbol{\eta}}$ *satisfies that* $\sup_{\boldsymbol{\eta}\in\mathcal{H}_n}||\boldsymbol{\eta}-\boldsymbol{\eta}_0||_2 = o_P(1)$ *and* $||\hat{e}(X)-e(X)||_{P,2}\times||\hat{p}_t(X)-p_t(X)||_{P,2}\leq\xi_n n^{-1/2}$ *holds with probability 1 when* n *tends to infinity.*

Assumption 2 assumes that all units have non-zero probabilities of being assigned to the treatment or the control arm. Such an assumption has been frequently considered in the causal inference literature. In addition, because we estimate the treatment effects across multiple subgroups, we require each subgroup to satisfy the positivity condition as well. Assumption 4 assumes the Donsker class condition for the class of efficient influence functions. This Donsker class condition can be weakened by conducting cross-fitting (see Appendix 1.9 for implementation details) and at the expense of more complicated proofs (see [186], for example). Additionally, [20] propose the highly adaptive lasso (HAL) estimator which guarantees $\sqrt{n}$-rate of convergence in the initial estimation step. Assumption 5 imposes regularity conditions on the nuisance parameter estimator. The second part in Assumption 5 bounds the product of errors of the nuisance parameter estimators $\hat{p}_t(X)$ and $\hat{e}(X)$.

Duality theory

In this section, we provide an alternative explanation of the iterative one-step TMLE from a duality theory perspective. Recall that in the discussed method, we replace the univariate clever covariate $\hat{S}_{tj}(X_i)$ with a multi-dimensional vector of clever covariate in the logistic regression in Eq (1.2). Because we hope to limit the sum of the squared influences of the clever covariates on updating $p_t(\cdot)$ when d is a large number, we impose a constraint that $\|\boldsymbol{\varepsilon}\|_2\leq\delta$, where $\boldsymbol{\varepsilon}=(\varepsilon_1,\ldots,\varepsilon_d)^\intercal$. Thus, solving for $\boldsymbol{\varepsilon}$ can be equivalently reformulated as iteratively solving the constraint optimization problem below

$$\hat{\boldsymbol{\varepsilon}}^{(k)} = \underset{||\boldsymbol{\varepsilon}||\leq\delta}{\arg\min} -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{(k-1)}(X_i)\big)+\sum_{j=1}^{d}\varepsilon_j\cdot\hat{S}_{t,j}(X_i)\Big) - \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{(k-1)}(X_i)\right)+\sum_{j=1}^{d}\varepsilon_j\cdot\hat{S}_{t,j}(X_i)}\Big)\Big]. \tag{1.12}$$

where $\hat{S}_{t,j}(X_i)=\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i=t)}{\hat{e}_t(X_i)}$. We refer to the above problem as the primal problem. We show that the above optimization problem has the following dual:

$$\hat{\lambda}^{(k)} = \underset{\lambda\geq0}{\arg\max} -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{(k-1)}(X_i)\big)+\tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}\Big) - \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{(k-1)}(X_i)\right)+\tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}}\Big)\Big]-\lambda\delta, \tag{1.13}$$

where $\tilde{S}_t(X_i)$ is defined in Section 1.3. Providing that the strong duality holds, this primal-dual relationship says that we can estimate the regression coefficient $\hat{\boldsymbol{\varepsilon}}^{(k)}$ in the primal

problem by either solving the primal problem (1.12) or by solving the dual problem (1.13) and then exploiting the primal-dual relationship. After reparametrizing (1.13) with $\gamma = \frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}$, in Lemma 1, we demonstrate that the unconstrained optimization problem:

$$\hat{\gamma}^{(k)} = \underset{\gamma \geq 0}{\arg\min} \; -\frac{1}{n} \sum_{i:T_i=t} \Big[Y_i \Big(\text{logit}\big(\hat{p}_t^{(k-1)}(X_i)\big) + \tilde{S}_t(X_i) \cdot \gamma \Big) - \log\Big(1 + \exp^{\text{logit}\left(\hat{p}_t^{(k-1)}(X_i)\right) + \tilde{S}_t(X_i)\cdot\gamma}\Big)\Big], \tag{1.14}$$

which is the reparametrized dual problem of the primal problem, and the updated estimate obtained from primal problem with large d yields the same estimate as what we propose whenever δ is sufficiently close to zero.

Lemma 1 (Primal and dual relationship)**.** *The optimization problems* (1.12) *and* (1.13) *form a primal-dual pair, and their solutions satisfy* $\varepsilon_j^{(k)} = \frac{\hat{\phi}_j^{(k-1)}(Y_i,T_i,X_i)}{\lambda}$, $j = 1, \ldots, d$, *where recall* $\hat{\phi}_j^{(k-1)}(Y_i, T_i, X_i)$ *is defined as*

$$\hat{\phi}_j^{(k-1)}(Y_i, T_i, X_i) = \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)} \frac{\mathbb{1}\mathbb{1}(T_i = t)}{\hat{e}_t(X_i)} (Y_i - \hat{p}_1^{(k-1)}(X_i)).$$

Reparametrizing $\gamma = \frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}$ *in the dual problem* (1.13)*. Whenever* $||\hat{\phi}(X_i)||_2 \cdot \delta/\gamma \to 0$*, the dual problem can be represented as*

$$\underset{\gamma > 0}{\arg\min} -\frac{1}{n} \sum_{i=1}^{n} \Big[Y_i \Big(\mathit{logit}\big(\hat{p}_1^{(k-1)}(X_i)\big) + \tilde{S}_1(X_i) \cdot \gamma \Big) - \log\Big(1 + \mathit{exp}^{\mathit{logit}\left(\hat{p}_1^{(k-1)}(X_i)\right) + \tilde{S}_1(X_i)\cdot\gamma}\Big)\Big], \tag{Reparametrized dual problem}$$

which yields the same solution as the primal problem.

The dual formulation in the discussed method has several benefits. In the reparametrized dual problem (1.14), because the estimated propensity score only enters the estimation process after being self-normalized in $\tilde{S}_t(X_i)$, the discussed approach has the added benefit of being robust to the presence of small estimated propensity scores. Small propensity scores are often encountered in datasets with unbalanced covariate distribution across the treatment and control groups, and such unbalancedness can lead to conventional estimators having substantial bias and large variances. Many numerical studies have found that similar self-normalization of propensity scores provides much more stable estimates of the treatment effects in finite samples. While the original formulation of the primal problem involves a sum over d inverse propensity score weighted clever covariates, its performance can be sensitive to the presence of small propensity score in finite samples.

Heuristics of TMLE from a semiparametric inference perspective

To fully appreciate the one-step targeted maximum likelihood estimator introduced in Section 1.3 from a semiparametric perspective, we start with introducing some basics of the classical semiparametric inference framework. To facilitate the discussion, we slightly abuse notations in using $\alpha_0 = E_X[E_{Y|X,T=t}[Y|X,T=t]]$ (the mean outcome under the treatment arm t) to denote the target parameter and in using P_0 to denote the true probability measure. Specifically, in a semiparametric model, we observe an i.i.d. sample $\{O_i\}_{i=1}^n = \{(Y_i, T_i, X_i)\}_{i=1}^n$ defined on the space $\mathcal{O}$ with a probability measure P_0 that possesses a density. The density belongs to the class $\mathcal{M} = \Big\{p\big(o;\alpha,\eta\big),\ \alpha \in \mathcal{A},\ \eta \in \mathcal{H}\Big\}$, with respect to some dominating measure ν, where $\mathcal{A} \subset \mathbb{R}$, and $\mathcal{H}$ is an infinite-dimensional set. We denote the true density that generates the data by $p_0\big(o;\alpha_0,\eta_0\big) \in \mathcal{M}$. Then, the parameter of interest is the finite-dimensional parameter α_0, and the nuisance parameter is the infinite-dimensional parameter $\eta_0 = \big(p_t(x), e(x), f(x)\big)$, where $p_t(x)$ is the conditional density function of Y given $T=t$ and $X=x$, $e(x)$ is the propensity score, and $f(x)$ is the marginal density of X evaluated at x. Given the probability measure P_0, we can also write the target parameter α_0 as a statistical function of P_0, denoted as $\alpha_0(P_0)$.

Under the above semiparametric statistics framework, a natural question raised here: how to evaluate the statistical efficiency of estimators for α_0 in a semiparametric model? As is often the case in semiparametric statistics, infinite-dimensional problems are tackled by first working with a finite-dimensional problem as an approximation and then taking limit to infinity ([161]). Therefore, the first step in a semiparametric model is to consider a simpler finite-dimensional "parameteric submodel" contained in $\mathcal{M}$, and use the theory and methods developed in classical parametric models to obtain an efficient estimator (that typically attains the Cramèr-Rao lower bound) of α_0. Similar to the Cramèr-Rao lower bound in parametric models, we use "semiparametric efficiency lower bound" as a metric for evaluating the asymptotic behavior of the semiparametric estimators. Heuristically, the semiparametric efficiency lower bound is simply the supremum of the Cramèr–Rao bounds for all "parametric submodels" for estimating α_0 ([158]).

Formally, we define a parametric submodel as $\mathcal{M}_{\alpha,\eta_{\varepsilon,S_h}} = \Big\{p\big(o;\alpha,\eta_{\varepsilon,S_h}\big),\ \alpha \in \mathcal{A},\ \varepsilon \in \mathcal{E} \subset \mathbb{R}\Big\}$, where S_h is the score function indexed by h which is, intuitively, the direction we perturb the nuisance parameter, and ε is the perturbing magnitude. In this parametric submodel, the true density is obtained by setting $\alpha = \alpha_0$ and $\varepsilon = 0$. Following the above definition, [99] have shown that in a parametric submodel with sufficient smoothness, the Cramèr-Rao lower bound $\mathrm{CR}(\alpha_0, S_h)$ for estimating α_0 in the submodel $\mathcal{M}_{\alpha,\eta_{\varepsilon,S_h}}$ satisfies

$$\mathrm{CR}(\alpha_0, S_h) = \lim_{\varepsilon \to 0} \frac{\big(\alpha(P_{\varepsilon,S_h}) - \alpha(P_0)\big)^2}{-2E_O\big[\log p(O;\alpha_0,\eta_{\varepsilon,S_h}) - \log p(O;\alpha_0,\eta_0)\big]}. \tag{1.15}$$

This suggests that $\mathrm{CR}(\alpha_0, S_h)$ captures the square change in the target parameter divided by the change in the log-likelihood at an infinitesimal ε.

Because the semiparametric efficiency bound (SPEB) is the supremum of the Cramèr-Rao lower bound for all parametric submodels, in order to find an estimator of α_0 that attains the SPEB, the targeted learning approach looks for a submodel in which a small change in the log-likelihood yields the maximal change in target parameter ([104]). Such a parametric submodel, which maximizes the Cramèr-Rao lower bound defined in Eq (1.15), is known as the "least favorable submodel" in the semiparametric literature ([161, 103]). One can thus view the classical targeted learning approach as a principled approach of constructing the least favorable submodel. The one-dimensional least favorable submodel constructed by TMLE allows us to directly work with the conditional likelihood related to the target parameter and perform the usual updating step as in MLE. While the classical MLE method cannot be extended to semiparametric models due to the infinite-dimensional nuisance parameter component, TMLE makes the MLE method feasible in semiparametric models.

Nevertheless, the least favorable submodel satisfying Eq (1.15) is only "locally least favorable" because the density $p(O;\alpha_0,\eta_{\varepsilon,S_h})$ maximizes the Cramèr-Rao lower bound locally at $\varepsilon = 0$. This suggests a practical drawback presents if an initial estimate $\hat{p}_0^{\text{Init}}(o)$ is far away from $p_0(o)$, yielding a large $\hat{\varepsilon}$ as more calibration is needed to push $\hat{p}_0^{\text{Init}}(o)$ towards the truth. A larger calibration yields a larger denominator in Eq (1.15) and thus smaller change in the target parameter. The consequence is that even though we can iteratively update the initial estimate until $\hat{\varepsilon} \approx 0$, the sample variance has been inflated because each updating step fails to maximize the change in the target parameter while maintaining minimal change in the log-likelihood.

To resolve the issue that the maximal change in the target parameter is only attained at $\varepsilon = 0$, [99] introduce the concept of "the universal least favorable submodel," denoted as $\mathcal{M}_{\alpha,\eta_{\varepsilon,S_h}} = \Big\{p\big(o;\alpha,\eta_{\varepsilon,S_h}\big),\ \alpha\in\mathcal{A},\ \varepsilon\in\mathcal{E}\subset\mathbb{R}\Big\}$. $\mathcal{M}_{\alpha,\eta_\varepsilon}$ is the universal least favorable submodel if

$$\frac{\partial}{\partial\varepsilon}\Big[\log p(o;\alpha_0,\eta_{\varepsilon,S_h}) - \log p(o;\alpha_0,\eta_0)\Big] = \varphi(o;\alpha_0,\eta_{\varepsilon,S_h}),\quad \forall\varepsilon\in\mathcal{E}\subset\mathbb{R}.$$

The universal least favorable submodel defined above achieves maximal change in the target parameter as long as $\varepsilon\in\mathcal{E}\subset\mathbb{R}$. The practical benefit of adopting the universal least favorable submodel is that we can avoid inflating the sample variance while performing TMLE updates, especially when the sample size is small.

In our example for estimating α_0, the universal least favorable submodel takes the form $\mathcal{M}_{\alpha,\eta_{\varepsilon,S_h}} = \Big\{p\big(o;\alpha,\eta_{\varepsilon,S_h}\big),\ \alpha\in\mathcal{A},\ \varepsilon\in\mathcal{E}\subset\mathbb{R}\Big\}$, where the nuisance parameter is

$$\eta_{\varepsilon,S_h} = \big(p_t^{\varepsilon,S_h}(x), e(x), f(x)\big),\quad p_t^{\varepsilon,S_h}(x) = \text{expit}\Big(\text{logit}\big(p_t(x)\big) + \varepsilon\cdot S_h\Big),\quad S_h = t/e_t(x). \tag{1.16}$$

Therefore, provided with an initial estimate of the nuisance parameter

$$\hat{\eta}_{\varepsilon,S_h} = (\hat{p}_t^{\text{Init}}(x), \hat{e}_t(x), \hat{f}(x))$$

obtained from an i.i.d. sample, the one-step TMLE tries to find $\hat{\varepsilon}$ that minimizes the denominator of the ratio defined in (1.15), or equivalently, maximizes the likelihood based on the observed data:

$$\begin{aligned}\hat{\varepsilon} &= \arg\min_{\varepsilon} \; -\mathbb{E}_n\big[\log p(O;\alpha_0,\hat{\eta}_{\varepsilon,S_h}) - \log p(O;\alpha_0,\eta_0)\big] \\ &= \arg\min_{\varepsilon} \; -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\operatorname{logit}\big(\hat{p}_t(X_i)\big)+\varepsilon\cdot\hat{S}_t(X_i)\Big) - \log\Big(1+\exp^{\operatorname{logit}\left(\hat{p}_t(X_i)\right)+\varepsilon\cdot\hat{S}_t(X_i)}\Big)\Big],\end{aligned}$$

where $\hat{S}_t(X_i) = \mathbb{1}(T_i = t)/\hat{e}_t(X_i)$. Then, we see that $\hat{\varepsilon}$ is equivalent to the estimated coefficient from running the logistic regression discussed in Section 1.3.

1.5 Simultaneous Confidence Intervals

To construct a level-q confidence interval for a single subgroup j, we work with $\hat{\alpha}_{t,j} \pm \Phi^{-1}(1-q/2)\cdot\left(\frac{\hat{\Sigma}_{t,jj}}{n}\right)^{1/2}$, where $\hat{\Sigma}_t$ is the estimated covariance matrix with

$$\hat{\Sigma}_t = \big(\hat{\Sigma}_{t,jk}\big)_{j,k=1}^{d} = \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\varphi}}_{t,i}\hat{\boldsymbol{\varphi}}_{t,i}^{\intercal}, \quad \hat{\boldsymbol{\varphi}}_{t,i} = \big(\hat{\varphi}_{t,1}(Y_i,T_i,X_i),\ldots,\hat{\varphi}_{t,d}(Y_i,T_i,X_i)\big)^{\intercal},$$

$$\hat{\varphi}_{t,j}(O_i) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)}\Big[\Big(\frac{T_i}{\hat{e}_t(X_i)}\big(Y_i-\hat{p}_t(X_i)\big)+\hat{p}_t(X_i)-\alpha_{t,j}\Big). \tag{1.17}$$

To construct a simultaneous level-q confidence interval though, let $\hat{\kappa}(q,\tilde{\Sigma}_t)$ be a consistent estimate of the $(1-q)$-th quantile of $\max_{j\in1,\ldots,d}|Z_j|$, where $(Z_1,\ldots,Z_d)^{\intercal}\sim N\big(0,\tilde{\Sigma}_t\big)$ with $\tilde{\Sigma}_t = \big(\tilde{\Sigma}_{t,jk}\big)_{j,k=1}^{d}$ and $\tilde{\Sigma}_{t,jk} = \frac{\hat{\Sigma}_{t,jk}}{\sqrt{\hat{\Sigma}_{t,jj}\hat{\Sigma}_{t,kk}}}$. Then, the constructed simultaneous confidence interval satisfies

$$\lim_{n\to\infty} P\left(\hat{\alpha}_{t,j}\pm\hat{\kappa}(q,\tilde{\Sigma}_t)\cdot\left(\frac{\hat{\Sigma}_{t,jj}}{n}\right)^{1/2}, j=1,\ldots,d\right) = 1-q. \tag{1.18}$$

Such a simultaneous confidence interval ensures that all the confidence intervals cover the corresponding true subgroup parameter at the same time.

1.6 Simulation Studies

To demonstrate the merit of the proposed method (iTMLE), we compare it with some conventional estimators under overlapping and non-overlapping subgroups cases. We compare the proposed method with a doubly robust (DR) estimator and a generalized linear model estimator (GLM), and we compare the cross-fitted version of iTMLE with the double machine learning (DML) method, since DML also utilizes cross-fitting. Before we present our

simulation results, we summarize two main takeaways from the simulation studies for our readers: (1) The proposed method has smaller bias, smaller variance, and lower family-wise error rate (FWER) compared to the considered estimators in finite samples. Recall that FWER refers to the probability of at least one constructed simultaneous confidence interval excluding the truth; (2) With cross-fitting, the proposed method shows enhanced finite sample performance in terms of smaller bias than the implementation without cross-fitting.

We measure the performance of various estimators according to their $\sqrt{n}$-scaled biases (computed as the root-n sum of mean differences between the Monte Carlo estimates and the true parameter across multiple subgroups), standard deviations (computed as the root-n sum of standard deviations of the Monte Carlo estimates across multiple subgroup), and FWER (computed as the proportion of Monte Carlo samples in which at least one constructed confidence interval for multiple subgroups excluding the truth). We scale the bias and variance by the sample size as they converge to zero as n goes to infinity.

Simulation design

Our simulation design mimics observational studies where treatments are assigned based on covariates. We simulate 1000 random Monte Carlo samples from: $X = (X_1, \ldots, X_5)^\intercal \sim N(0, \Sigma), \Sigma_{ij} = 0.5^{|i-j|}, T \sim \text{Bernoulli}\Big(\text{expit}(X_1 - 0.5 \cdot X_2 + 0.25 \cdot X_3 + 0.1 \cdot X_4)\Big)$, and $Y|T, X \sim \text{Bernoulli}\Big(\text{expit}(21 + T + 27.4 \cdot X_1 + 13.7 \cdot X_2 + 13.7 \cdot X_3 + 13.7 \cdot X_4)\Big)$. We consider this specific simulation design because the design has been frequently adopted in the causal inference literature ; see [91, 85] for example. This enables us to better compare our approach with existing methods. Kindly pointed out by an anonymous reviewer, the above simulation design produces rather deterministic outcomes, and we thus provide additional simulation results under an alternative simulation design in Appendix 1.9.

We consider two types of subgroups: overlapping subgroups and non-overlapping subgroups. Overlapping subgroups with moderate d, $d = 4$, are generated by $\mathcal{A}_1 = \{X_1 > \Phi^{-1}(0.1)\}, \mathcal{A}_2 = \{\Phi^{-1}(0.1) < X_2 < \Phi^{-1}(0.9)\}, \mathcal{A}_3 = \{X_3 + X_4 > -2\}, \mathcal{A}_4 = \{\mathbb{1}_{X_4 > 0.5} > -1\}$. Non-overlapping subgroups with large d, $d = 10$, are generated by $\mathcal{A}_j = \big\{Q_{X_1}(j/10) < X_1 < Q_{X_1}((j+1)/10)\big\}, j = 1, \ldots, 10$. For simplicity, in the following simulation studies, the considered parameter is $\boldsymbol{\alpha}_1 = (\alpha_{1,1}, \ldots, \alpha_{1,d})^\intercal$.

Comparison with conventional estimators

We generate initial estimates of $e_t(\cdot)$ and $p_t(\cdot)$ through logistic regression, random forest, or gradient boosting, implemented in `R` packages `stats`, `ranger` ([176]), and `xgboost` ([28]). We compare the proposed iterative one-step TMLE method (iTMLE) with the doubly robust estimator, a simple regression adjustment estimator, and the inverse propensity score estimator, which are defined as $\hat{\alpha}_{t,j}^{\text{DR}} = \frac{1}{n_j} \sum_{i \in \mathcal{A}_j} \left[\frac{T_i}{\hat{e}_t(X_i)} (Y_i - \hat{p}_t^{\text{Init}}(X_i)) + \hat{p}_t^{\text{Init}}(X_i) \right]$, $\hat{\alpha}_{t,j}^{\text{GLM}} = \frac{1}{n_j} \sum_{i \in \mathcal{A}_j} \hat{p}_t^{\text{Init}}(X_i)$, $\hat{\alpha}_{t,j}^{\text{IPW}} = \frac{1}{n_j} \sum_{i \in \mathcal{A}_j} \frac{T_i}{\hat{e}_t(X_i)} Y_i$. Simultaneous confidence intervals for these

estimators are constructed using standard large sample theory adopted in the literature (see [68] for the doubly robust estimator and [170] for the IPW estimator). We provide finite-sample comparisons through Figure 1.1(A) – (C) for overlapping subgroups, and Figure 1.1(D) – (E) for non-overlapping subgroups. As the IPW estimator has much larger variance than the other estimators, we exclude its results from these figures. From Figure 1.1, we observe that the iTMLE estimator outperforms the others for bias, standard deviation, and FWER, regardless of how $e_1(\cdot)$ and $p_1(\cdot)$ are estimated in the first stage. This is in-line with our theoretical results because the proposed estimator consists of a data-adaptive bias correction term (Section 1.3), which largely improves its finite sample performance. In addition, among all three initial estimators, random forest overall seems to be a winner.

Comparison with the double machine learning

In this part of the simulation study, we compare the performance of the cross-validated version of iterated one-step TMLE for multiple parameters with the double machine learning (DML) method ([33]). DML also involves the estimations of the propensity score model and the conditional mean model, and it is a meta-learning method that relies on Neyman orthogonal score and cross-fitting to generate debiased estimates for the causal estimands. The simulation results of the three-fold cross-validated iTMLE and DML (implemented with the `R` package `DoubleML` ([16])) are presented in Figure 1.2. There are two takeaways from the summarized results in Figure 1.2. First, the performance of CV-iTMLE surpasses DML. Although DML is rather robust compared to the doubly robust estimator, it still yields larger bias and variance than CV-iTMLE. Second, compared to the iTMLE implementation without cross-fitting (Figure 1.1), CV-iTMLE shows a faster convergence rate. We conjecture that the sample splitting step allows the non-parametric estimators in the initial stage to converge faster and thus shows more robust performance (smaller bias, smaller standard deviation, and smaller FWER).

1.7 Case Study in UK Biobank Data

Statins are the most commonly prescribed cholesterol-lowering medications in the United States. Cholesterol's role in β-amyloid processing and the potential link between serum cholesterol levels and AD pathology ([132]) have led to the argument that cholesterol-moderating drugs such as statins could reduce the risk of AD onset and progression. However, this argument is controversial by current evidence. Several cohort studies found a negative association between statin usage and AD ([188]), while others have failed to replicate those findings. These inconsistent findings might be due to the effect of statins on AD varying across gender, age, and other subgroups ([188]). Thus, we hypothesize that statin usage has significant benefits of reducing AD risk in some (but not all) subgroups. To test this hypothesis, we analyzed data in the UK Biobank to investigate the heterogeneous treatment effect of inheriting rs12916-T allele, a proxy for statin usage, on AD risk in the White British subpop-

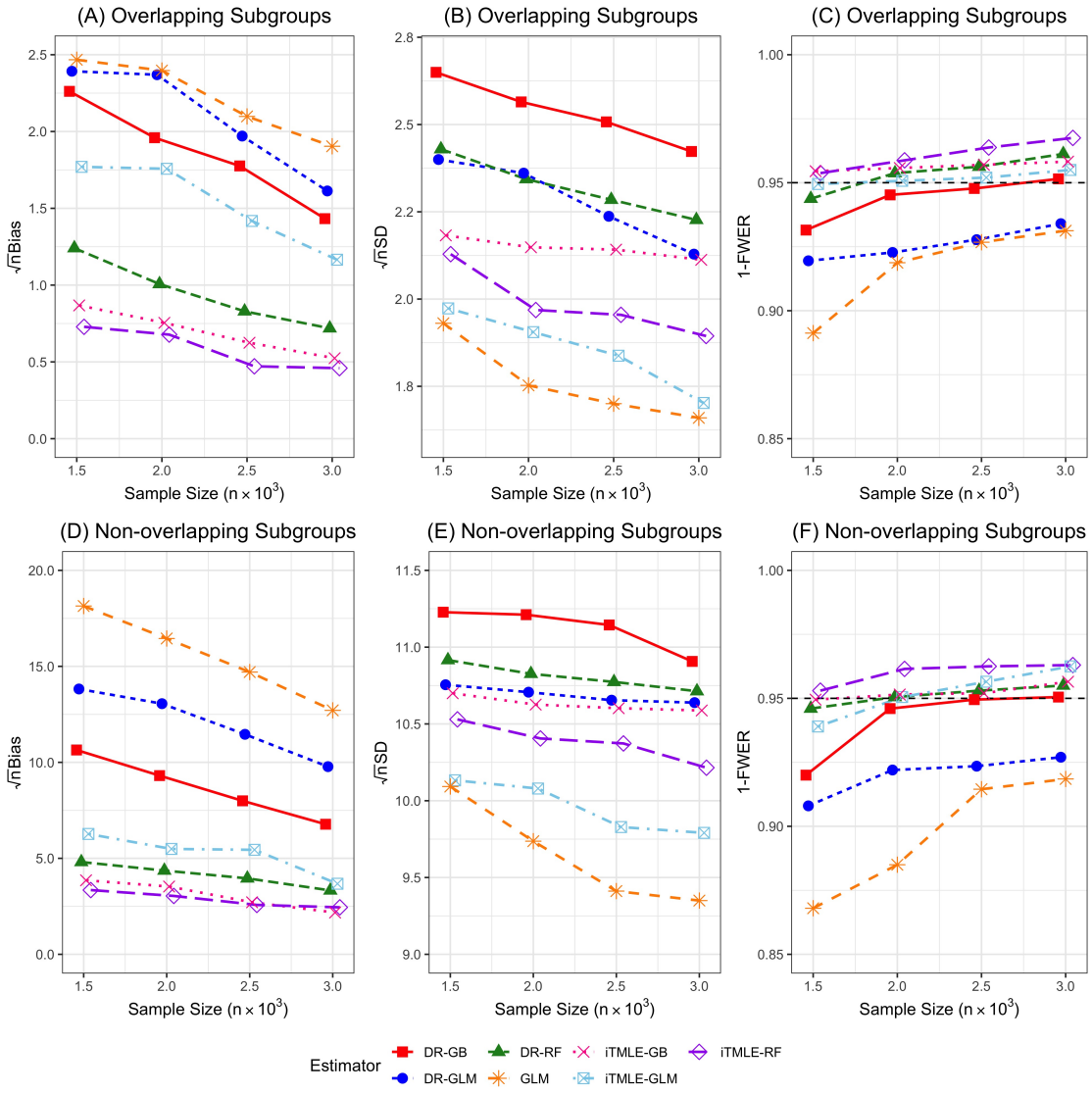


Figure 1.1: Comparison of bias, standard deviation (scaled by root-n), and (1-FWER) in overlapping and non-overlapping subgroups. "iTMLE" denotes the proposed estimator. "DR" denotes the doubly robust estimator. "GLM" denotes the generalized linear models. The maximum Monte Carlo standard error is 0.026 for iTMLE, 0.028 for DR, and 0.022 for GLM. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

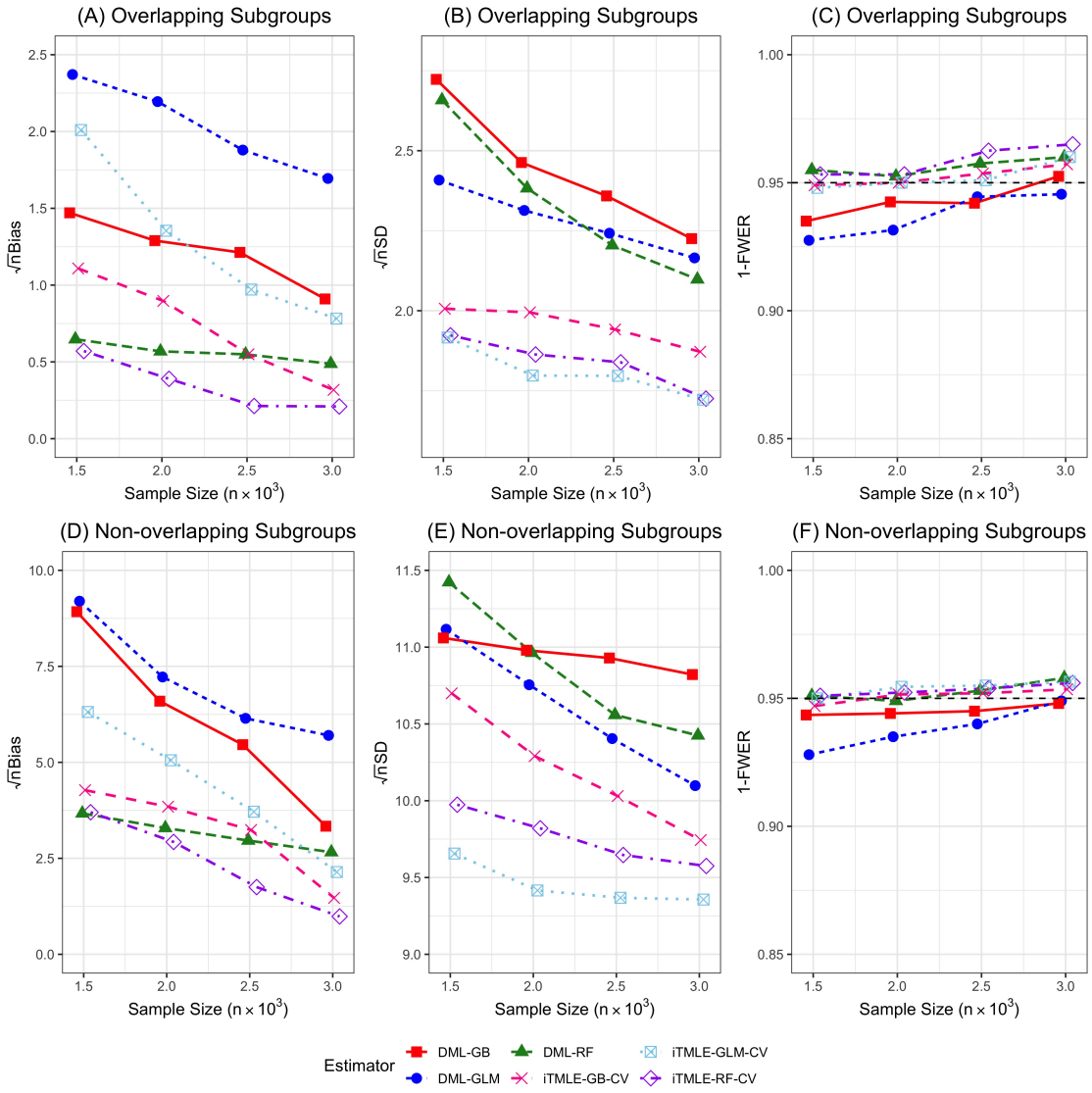


Figure 1.2: Comparison of the cross-validated iTMLE implementation and the double machine learning method. "iTMLE-CV" denotes the proposed method with cross-fitting. "DML" denotes the double machine learning method. The maximum Monte Carlo standard error of (1-FWER) is 0.024 for CV-iTMLE and 0.026 for DML. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

ulations. We considered a cross-sectional study design by looking at the disease prevalence at the end of year 2021.

Study design

The UK Biobank study recruited 502,536 participants aged from 40 to 69 in the United Kingdom from 2006 to 2010. We defined AD status by integrating information provided by Hospital Episode Statistics, death registries, and self-reported diagnoses (see details in Appendix 1.9). We restricted our study to 293,929 White British individuals. These individuals are unrelated and had passed standard quality control steps.

Instead of directly adopting statin usage as a treatment variable, we adopted a genetic variant rs12916-T as a surrogate treatment variable. This means that if the subject carries the variant rs12916-T, the treatment indicator variable is set to be $T = 1$; otherwise, T is set to be zero. We adopted this genetic surrogate biomarker as the treatment variable for two reasons. On the one hand, the rs12916-T allele only affects the LDL cholesterol concentration through HMGCR inhibition, and it is thus functionally equivalent to statin usage ([155, 65]). More specifically, the decreased LDL cholesterol level associated with statin usage is similar to the association pattern with *rs12916-T* ($R^2 = 0.94$) [177], thus rs12916-T is a sensible surrogate treatment variable for statin usage. On the other hand, given that genetic variants are randomly inherited from parents, our treatment variable (whether or not the individual carries rs12916-T) is thus independent of unmeasured confounding factors such as lifestyle modifications after statin usage, potentially making Assumption 1 more plausible.

To account for genetic pleiotropy, we adjusted for SNPs that are associated with LDL. Briefly, we selected 385 independent genome-wide significant SNPs (with p-values less than 5×10^{-8} and $R^2 < 0.01$) associated with LDL according to the published genome-wide association study (GWAS) results harmonized in GWAS Catalogue ([116]). We further adjusted for age and gender variables, which may improve estimation efficiency given their associations with the outcome. It was our hope that this study design could increase the plausibility of Assumption 1.

We investigated the effect of inheriting rs12916-T allele on AD risk in the following subgroups: (1) males, (2) females, (3) age < 65, (4)age ≥ 65, (5) individuals with high AD genetic risk, and (6) individuals with low AD genetic risk. Notably, "high AD genetic risk" was defined as either a subject's parents or siblings being diagnosed with AD, while "Low AD genetic risk" was defined as neither a subject's parents nor siblings being diagnosed with AD. We compared the performance of the proposed method (CV-iTMLE) with the double machine learning (DML) method and the widely used generalized linear models (GLM). We used the random forest as our first stage estimator as it provides the most robust results in our simulation studies.

Because statin usage may increase the risk of T2D ([155]), as a secondary analysis, we further investigated the effect of inheriting rs12916-T allele on T2D to evaluate the potential heterogeneous side effects. The study design and study results of this secondary analysis can be found in Appendix 1.9.

Results

Figure 1.3 summarizes the effect of inheriting rs12916-T (a proxy for statin usage) on AD risk in considered subgroups. As the GLM was applied to each subgroup separately and the sample size was much smaller, leading to non-significant associations for all the subgroups. The DML method also did not find any significant effects in all subgroups. This might be caused by small estimated propensity scores, leading to large variability in finite samples. In contrast, by targeting all subgroups simultaneously, the proposed method suggested that carrying rs12916-T allele is protective against AD in the subgroup younger than 65 (RR: 0.92, 95% CI: 0.86–0.98). In sum, our proposed method showed shortened confidence intervals with improved statistical power in detecting significant subgroups, while the GLM and DML methods tend to lose power.

Lastly, we acknowledge that the current study design has several potential limitations. First, our study only investigated the treatment effect of carrying rs12916-T allele or not. Although this genetic variant is a sensible proxy for statin usage, the findings from this study need to be interpreted cautiously. Second, our study was based on UK Biobank data and only focused on White British population. UK Biobank participants were healthier (e.g., fewer self-reported health conditions) than the general population. Thus, our findings may not be generalizable to other populations.

1.8 Discussions

In this chapter, we propose a semiparametric efficient method for simultaneous heterogeneous treatment effect estimation across multiple subgroups. The proposed method allows us to construct a powerful multiple testing procedure leveraging the subgroup dependence structure. In our empirical studies, the proposed method demonstrates finite sample improvements compared to other conventional methods, including the doubly robust estimator, the classical one-step TMLE, and the double machine learning method. This chapter opens a variety of possibilities for future research. From a methodological perspective, our current method can be extended to work with other types of outcomes. For example, if the outcome is continuous, one can either modify the updating step [63], or dichotomize a continuous outcome into binary values (more details can be found in Appendix 1.9). From an application perspective, the proposed method can be flexibly adapted to various clinical studies and assist the evaluation of other subpopulations of interest.

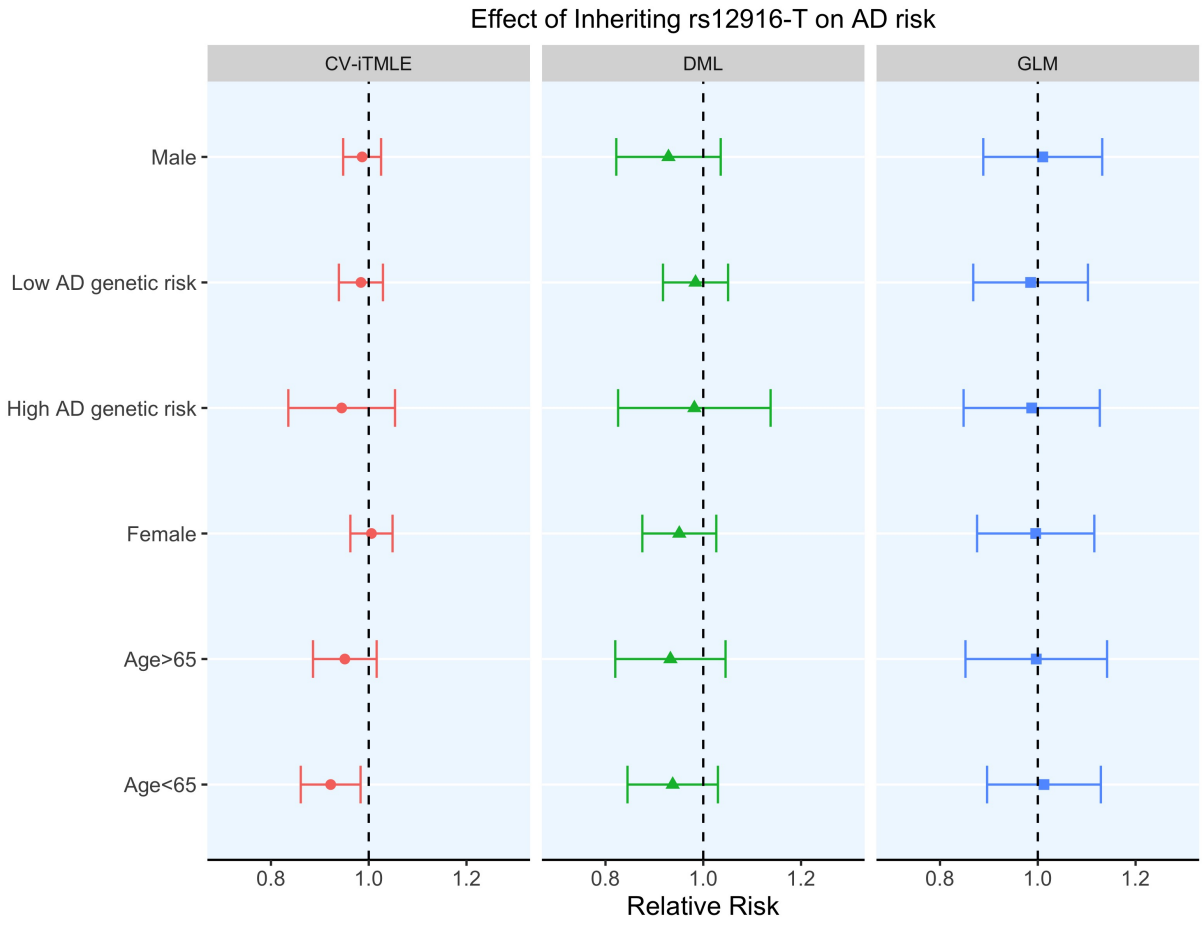


Figure 1.3: The effect of inheriting rs12916-T allele (a proxy for statin usage) on the risk of developing Alzheimer's disease (AD) in the UK Biobank white British population ($n = 293,929$). "DML" denotes the double machine learning method. "GLM" denotes the generalized linear models. GLM is used for association test and does not imply causal relationships. "CV-iTMLE" denotes the cross-validated iTMLE method.

1.9 Supplementary Materials

Proof of Theorems

Proof of Theorem 1

Suppose that $\{O_i\}_{i=1}^n := \{(Y_i, T_i, X_i)\}_{i=1}^n$ are i.i.d. random variables defined on the space $\mathcal{O}$ with probability measure P. For a real-valued function f on $\mathcal{X}$, the empirical measure $\mathbb{P}_n$ is defined by $\mathbb{P}_n(f) := \frac{1}{n}\sum_{i=1}^n f(O_i)$, and the empirical mean is defined as $\mathbb{E}_n[f(X)] := \frac{1}{n}\sum_{i=1}^n f(X_i)$. If $\mathcal{F}$ is a collection of real-valued functions defined on $\mathcal{X}$, then $\{\mathbb{P}_n f : f \in \mathcal{F}\}$ is the empirical measure indexed by $\mathcal{F}$. We assume that $Pf = \int f \mathrm{d}P$ exists for each $f \in \mathcal{F}$. Note that such a notation can be more helpful as we can treatment random functions. We use $E_X[f(X)]$ to denote expectation taken with respect to the random variable X when it is more convenient to simplify notations. Our procedure estimates the joint distribution of (Y_i, T_i, X_i), and the estimated density evaluated at $(y, t, x) \in \mathcal{X}$ is $\hat{p}_t(Y_i = y|X_i = x) \cdot \hat{e}_t(x) \cdot \hat{p}(x)$. We denote the joint density of (Y_i, T_i, X_i) as $p(y, t, x)$ which can be decomposed into the product $p(y, t, x) = p_t(x)e_t(x)p(x)$. Here $e_t(x) = te(x) + (1 - t)(1 - e(x))$. We denote the probability measure defined by such an estimated density as $\mathbb{P}_n^*$. Given the probability measure P, our target parameter $\boldsymbol{\alpha}_t$ can also be written as a statistical function of P, denoted as $\boldsymbol{\alpha}_t(P)$. Similarly, our proposed estimator can be written as $\boldsymbol{\alpha}_t(\mathbb{P}_n^*)$. The vector of efficient influence function of our target parameter $\boldsymbol{\alpha}_t(P)$ is denoted as

$$\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big) := \Big(\psi_1\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big), \ldots, \boldsymbol{\psi}_J\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big)\Big)^\top,$$

where $\boldsymbol{\eta}$ contains the nuisance parameters. We decompose the efficient influence function into two parts:

$$\begin{aligned}
\psi_j^{(1)}\big(o; \boldsymbol{\eta}(P)\big) &= \frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j)}\Big[\frac{\mathbb{1}\{t = 1\}}{e(x)} + \frac{\mathbb{1}\{t = 0\}}{1 - e(x)}\Big]\big(y - p_t(x)\big), \\
\psi_j^{(2)}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big) &= \frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j)}\big(p_t(x) - \alpha_t\big),
\end{aligned}$$

where the first part does not depend on the target parameter, and $\boldsymbol{\alpha}_t = \int p_t \mathrm{d}P = \mathbb{E}_X[p_t(X)]$ is a scalar. Similarly, we then decompose the vector of efficient influence into

$$\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big) = \boldsymbol{\psi}^{(1)}\big(o; \boldsymbol{\eta}(P)\big) + \boldsymbol{\psi}^{(2)}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big).$$

Our estimator satisfies the following expansion:

$$\hat{\boldsymbol{\alpha}}_t(\mathbb{P}_n^*) - \boldsymbol{\alpha}_t(P) = (\mathbb{P}_n - P)\boldsymbol{\psi}(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)) + I_{n1} + I_{n2}, \tag{1.19}$$

where the remainder terms are derived in Section 1.9

$$\begin{aligned}
I_{n1} &= R_{n2} + R_{n3} = (\mathbb{P}_n - P)\big(\boldsymbol{\psi}(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(\mathbb{P}_n^*)) - \boldsymbol{\psi}(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P))\big), \\
I_{n2} &= R_{n4} = P\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(\mathbb{P}_n^*)\big) = P\big[\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(\mathbb{P}_n^*)\big) - \boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t(P), \boldsymbol{\eta}(P)\big)\big].
\end{aligned}$$

Let $\boldsymbol{\alpha}_t(P) = \boldsymbol{\alpha}_t$ and $\boldsymbol{\eta}(P) = \boldsymbol{\eta}_0$, and let the estimated parameters $\hat{\boldsymbol{\alpha}}_t(\mathbb{P}_n^*) = \hat{\boldsymbol{\alpha}}_t$ and $\boldsymbol{\eta}(\mathbb{P}_n^*) = \hat{\boldsymbol{\eta}}$. The above remainder terms simplify to

$$\begin{aligned} I_{n1} &= (\mathbb{P}_n - P)\big(\boldsymbol{\psi}(o; \boldsymbol{\alpha}_t, \hat{\boldsymbol{\eta}}) - \boldsymbol{\psi}(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0)\big), \\ I_{n2} &= P\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \hat{\boldsymbol{\eta}}\big) = P\big[\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \hat{\boldsymbol{\eta}}\big) - \boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0\big)\big]. \end{aligned}$$

For any $\boldsymbol{\eta}$ in the nuisance estimator realization set $\mathcal{H}_n$, the term $I_{n1} = o_P(1)$ uniformly over $P \in \mathcal{P}_n$ under Assumption 4.2. To bound I_{n2}, we introduce the function

$$f(r) = P\Big[\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0 + r(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0)\big)\Big].$$

By Taylor expansion and the fact that $f(0) = 0$, we have

$$f(1) = f'(0) + f''(\tilde{r})/2, \quad \text{for some } \tilde{r} \in (0, 1).$$

We will next verify in Step 1 that $f'(0) = 0$ and verify in Step 2 that $\sup_{r\in[0,1)} ||f''(r)|| = o_P(1/\sqrt{n})$. This finishes our proof.

Step 1. For any $\boldsymbol{\eta} \in \mathcal{H}_n$, the first order derivative is equal to

$$\begin{aligned} f'(0) &= \partial_\eta P\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0\big)[\boldsymbol{\eta} - \boldsymbol{\eta}_0] \\ &= \partial_r\{P\big[\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0 + r(\boldsymbol{\eta} - \boldsymbol{\eta}_0)\big)\big]\}\big|_{r=0}, \end{aligned}$$

First, we want to show $\boldsymbol{\psi}$ is Neyman orthogonal, such that $f'(0) = 0$.

We denote $\boldsymbol{\eta} = (\check{e}(X), \check{p}_1(X), \check{p}_0(X), \check{\mathbb{P}}(\mathcal{A}_1), \ldots, \check{\mathbb{P}}(\mathcal{A}_J))$. The Gateaux derivative in the direction of $\boldsymbol{\eta} - \boldsymbol{\eta}_0 = \big(\check{e}(X) - e(X), \check{p}_1(X) - p_1(X), \check{p}_0(X) - p_0(X), \check{\mathbb{P}}(\mathcal{A}_1) - P(\mathcal{A}_1), \ldots, \check{\mathbb{P}}(\mathcal{A}_J) - P(\mathcal{A}_J)\big)$ is

$$\begin{aligned} &\partial_r\{E\big[\boldsymbol{\psi}\big(o; \boldsymbol{\alpha}_t, \boldsymbol{\eta}_0 + r(\boldsymbol{\eta} - \boldsymbol{\eta}_0)\big)\big]\} \\ &= \begin{pmatrix} \partial_r E\left[\frac{\mathbb{1}\{x\in\mathcal{A}_1\}}{P(\mathcal{A}_1) + r\left(\check{\mathbb{P}}(\mathcal{A}_1) - P(\mathcal{A}_1)\right)}\left[\left(\frac{\mathbb{1}\{t=1\}}{e(x) + r\left(\check{e}(x) - e(x)\right)} + \frac{\mathbb{1}\{t=0\}}{1 - e(x) - r\left(\check{e}(x) - e(x)\right)}\right)\right.\right. \\ \qquad \left.\left. \cdot\Big(y - p_t(x) - r\big(\check{p}_t(x) - p_t(x)\big)\Big) + \big(p_t(x) + r(\check{p}_t(x) - p_t(x)) - \alpha_t\big)\right]\right], \\ \qquad\qquad \vdots \\ \partial_r E\left[\frac{\mathbb{1}\{x\in\mathcal{A}_J\}}{P(\mathcal{A}_J) + r\left(\check{\mathbb{P}}(\mathcal{A}_J) - P(\mathcal{A}_J)\right)}\left[\left(\frac{\mathbb{1}\{t=1\}}{e(x) + r\left(\check{e}(x) - e(x)\right)} + \frac{\mathbb{1}\{t=0\}}{1 - e(x) - r\left(\check{e}(x) - e(x)\right)}\right)\right.\right. \\ \qquad \left.\left. \cdot\Big(y - p_t(x) - r\big(\check{p}_t(x) - p_t(x)\big)\Big) + \big(p_t(x) + r(\check{p}_t(x) - p_t(x)) - \alpha_t\big)\right]\right] \end{pmatrix} \end{aligned}$$

The Gateaux derivative for each subgroup j is

$$
\begin{aligned}
&= \partial_r E\Bigg[\frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}\Big[\Big(\frac{\mathbb{1}\{t=1\}}{e(x) + r\big(\check{e}(x) - e(x)\big)} + \frac{\mathbb{1}\{t=0\}}{1 - e(x) - r\big(\check{e}(x) - e(x)\big)}\Big) \\
&\qquad \cdot \Big(y - p_t(x) - r\big(\check{p}_t(x) - p_t(x)\big)\Big) + \big(p_t(x) + r(\check{p}_t(x) - p_t(x)) - \alpha_t\big)\Big]\Bigg], \\
&= \partial_r E\Bigg[\frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} \cdot (p_t(x) + r(\check{p}_t(x) - p_t(x)) - \alpha_t) \\
&\quad + \frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} \cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big) \\
&\quad \cdot \Big[\frac{\mathbb{1}\{t=1\}}{e(x) + r\big(\check{e}(x) - e(x)\big)} + \frac{\mathbb{1}\{t=0\}}{1 - e(x) - r\big(\check{e}(x) - e(x)\big)}\Big]\Bigg], \\
&= E\Bigg[\frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}(\check{p}_t(x) - p_t(x)) \\
&\quad - \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2}\big(p_t(x) + r(\check{p}_t(x) - p_t(x)) - \alpha_t\big) \\
&\quad + \frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} \cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big) \cdot \\
&\quad \cdot \Big[-\frac{\mathbb{1}\{t=1\}\big(\check{e}(x) - e(x)\big)}{\big(e(x) + r\big(\check{e}(x) - e(x)\big)\big)^2} + \frac{\mathbb{1}\{t=0\}\big(\check{e}(x) - e(x)\big)}{\big(1 - e(x) - r\big(\check{e}(x) - e(x)\big)\big)^2}\Big] \\
&\quad + \Big[\frac{-\mathbb{1}\{x \in \mathcal{A}_j\}(\check{p}_t(x) - p_t(x))}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} - \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2} \\
&\quad \cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big)\Big] \\
&\quad \cdot \Big[\frac{\mathbb{1}\{t=1\}}{e(x) + r\big(\check{e}(x) - e(x)\big)} + \frac{\mathbb{1}\{t=0\}}{1 - e(x) - r\big(\check{e}(x) - e(x)\big)}\Big]\Bigg].
\end{aligned}
$$

Set $r = 0$,

$$
\begin{aligned}
&= E\Bigg[\frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j)}(\check{p}_t(x) - p_t(x)) - \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j)\big)^2}\big(p_t(x) - \alpha_t\big) \\
&\quad + \frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j)}\Big(y - p_t(x)\Big)\Big[-\frac{\mathbb{1}\{t=1\}\big(\check{e}(x) - e(x)\big)}{\big(e(x)\big)^2} + \frac{\mathbb{1}\{t=0\}\big(\check{e}(x) - e(x)\big)}{\big(1-e(x)\big)^2}\Big] \\
&\quad + \Big[\frac{-\mathbb{1}\{x \in \mathcal{A}_j\}(\check{p}_t(x) - p_t(x))}{P(\mathcal{A}_j)} - \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j)\big)^2} \cdot \Big(y - p_t(x)\Big)\Big] \\
&\quad \cdot \Big[\frac{\mathbb{1}\{t=1\}}{e(x)} + \frac{\mathbb{1}\{t=0\}}{1-e(x)}\Big]\Bigg].
\end{aligned}
$$

Given that

$$
\begin{aligned}
&E[p_t(x) - \alpha_t|x] = 0, \quad E[t=1|X] = e(x), \\
&E[t(y - p_t(x))|x] = 0, \quad E[(1-t)(y - p_t(x))|x] = 0,
\end{aligned}
$$

the first-order Gateaux derivative for each subgroup j is 0. Thus $f'(0) = 0$ for all the subgroups.

Step 2. The second order remainder term satisfies

$$||f''(\tilde{r})/2|| \leq \sup_{r \in (0,1)} ||f''(r)/2||.$$

For each subgroup j,

$$
\begin{aligned}
&\partial_r^2 f_j(r) \\
&= E\Bigg[- \frac{\mathbb{1}\{x \in \mathcal{A}_j\}(\check{p}_t(x) - p_t(x))\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2} - \\
&\quad \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2}\big(\check{p}_t(x) - p_t(x)\big) \\
&\quad + 2\frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)^2}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^3}\big(p_t(x) + r(\check{p}_t(x) - p_t(x)) - \alpha_t\big) \\
&\quad + \frac{\mathbb{1}\{x \in \mathcal{A}_j\}}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} \cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big) \cdot \\
&\qquad \cdot 2\Big[\frac{\mathbb{1}\{t = 1\}\big(\check{e}(x) - e(x)\big)^2}{\big(e(x) + r\big(\check{e}(x) - e(x)\big)\big)^3} - \frac{\mathbb{1}\{t = 0\}\big(\check{e}(x) - e(x)\big)^2}{\big(1 - e(x) - r\big(\check{e}(x) - e(x)\big)\big)^3}\Big] \\
&\quad + \Big[\frac{-\mathbb{1}\{x \in \mathcal{A}_j\}(\check{p}_t(x) - p_t(x))}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} - \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2} \\
&\cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big)\Big] \\
&\qquad \cdot \Big[- \frac{\mathbb{1}\{t = 1\}\big(\check{e}(x) - e(x)\big)}{\big(e(x) + r\big(\check{e}(x) - e(x)\big)\big)^2} + \frac{\mathbb{1}\{t = 0\}\big(\check{e}(x) - e(x)\big)}{\big(1 - e(x) - r\big(\check{e}(x) - e(x)\big)\big)^2}\Big] \\
&\quad + \Big[\frac{-\mathbb{1}\{x \in \mathcal{A}_j\}(\check{p}_t(x) - p_t(x))}{P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)} - \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2} \\
&\qquad \cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big)\Big] \\
&\qquad \cdot \Big[\frac{-\mathbb{1}\{t = 1\}\big(\check{e}(x) - e(x)\big)}{\big(e(x) + r\big(\check{e}(x) - e(x)\big)\big)^2} + \frac{\mathbb{1}\{t = 0\}\big(\check{e}(x) - e(x)\big)}{\big(1 - e(x) - r\big(\check{e}(x) - e(x)\big)\big)^2}\Big]. \\
&\quad + \Big[\frac{\mathbb{1}\{x \in \mathcal{A}_j\}(\check{p}_t(x) - p_t(x))\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2} \\
&\quad + \frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big) \cdot (\check{p}_t(x) - p_t(x))}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^2} \\
&\quad + 2\frac{\mathbb{1}\{x \in \mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)^2}{\big(P(\mathcal{A}_j) + r\big(\check{\mathbb{P}}(\mathcal{A}_j) - P(\mathcal{A}_j)\big)\big)^3} \cdot \Big(y - p_t(x) - r(\check{p}_t(x) - p_t(x))\Big)\Big] \\
&\qquad \cdot \Big[\frac{\mathbb{1}\{t = 1\}}{e(x) + r\big(\check{e}(x) - e(x)\big)} + \frac{\mathbb{1}\{t = 0\}}{1 - e(x) - r\big(\check{e}(x) - e(x)\big)}\Big]\Bigg].
\end{aligned}
$$

Simplify the second-order Gateaux derivative we have

$$
\begin{aligned}
\partial_r^2 f(r) = E\Bigg[& -2\frac{\mathbb{1}\{x\in\mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)\big)^2}\big(\check{p}_t(x)-p_t(x)\big)\Bigg] \\
& + E\Bigg[2\frac{\mathbb{1}\{x\in\mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)^2}{\big(P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)\big)^3}\big(p_t(x)+r(\check{p}_t(x)-p_t(x))-\alpha_t\big)\Bigg] \\
& + E\Bigg[\frac{\mathbb{1}\{x\in\mathcal{A}_j\}}{P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)}\cdot\Big(y-p_t(x)-r(\check{p}_t(x)-p_t(x))\Big)\cdot \\
& \quad \cdot 2\Big[\frac{\mathbb{1}\{t=1\}\big(\check{e}(x)-e(x)\big)^2}{\big(e(x)+r\big(\check{e}(x)-e(x)\big)\big)^3}-\frac{\mathbb{1}\{t=0\}\big(\check{e}(x)-e(x)\big)^2}{\big(1-e(x)-r\big(\check{e}(x)-e(x)\big)\big)^3}\Big]\Bigg] \\
& + E\Bigg[2\Big[\frac{-\mathbb{1}\{x\in\mathcal{A}_j\}(\check{p}_t(x)-p_t(x))}{P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)}-\frac{\mathbb{1}\{x\in\mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)\big)^2} \\
& \cdot\Big(y-p_t(x)-r(\check{p}_t(x)-p_t(x))\Big)\Big] \\
& \quad \cdot\Big[-\frac{\mathbb{1}\{t=1\}\big(\check{e}(x)-e(x)\big)}{\big(e(x)+r\big(\check{e}(x)-e(x)\big)\big)^2}+\frac{\mathbb{1}\{t=0\}\big(\check{e}(x)-e(x)\big)}{\big(1-e(x)-r\big(\check{e}(x)-e(x)\big)\big)^2}\Big]\Bigg] \\
& + E\Bigg[\Big[\frac{2\mathbb{1}\{x\in\mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)}{\big(P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)\big)^2}\big(\check{p}_t(x)-p_t(x)\big) \\
& \quad +\frac{2\mathbb{1}\{x\in\mathcal{A}_j\}\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)^2}{\big(P(\mathcal{A}_j)+r\big(\check{\mathbb{P}}(\mathcal{A}_j)-P(\mathcal{A}_j)\big)\big)^3}\cdot\Big(y-p_t(x)-r(\check{p}_t(x)-p_t(x))\Big)\Big] \\
& \quad \cdot\Big[\frac{\mathbb{1}\{t=1\}}{e(x)+r\big(\check{e}(x)-e(x)\big)}+\frac{\mathbb{1}\{t=0\}}{1-e(x)-r\big(\check{e}(x)-e(x)\big)}\Big]\Bigg].
\end{aligned}
$$

Set $r=\tilde{r}$, by Assumption 4.3, for some constants C and ξ_n,

$$
\begin{aligned}
&|f_j''(\tilde{r})| \le C||\check{e}(x)-e(x)||_2||\check{p}_t(x)-p_t(x)||_2 \le \delta_n n^{-1/2}, \quad j=1,\dots,J, \\
&||f''(\tilde{r})||_\infty \le \sup_{r\in[0,1)}||f''(r)||_\infty = o_P(1/\sqrt{n}).
\end{aligned}
$$

In sum,

$$
f(1)=f'(0)+f''(\tilde{r})/2=o_P(1/\sqrt{n}), \quad r\in[0,1).
$$

Derivation of the expansion (1.19)

It holds trivially true that our estimator satisfies

$$\begin{aligned}
\hat{\boldsymbol{\alpha}}_t - \boldsymbol{\alpha}_t =& -\int \boldsymbol{\psi}^{(1)}\big(o;\hat{\boldsymbol{\eta}}\big)\mathrm{d}P + \Big(\hat{\boldsymbol{\alpha}}_t - \boldsymbol{\alpha}_t + \int \boldsymbol{\psi}^{(1)}\big(o;\hat{\boldsymbol{\eta}}\big)\mathrm{d}P\Big) \\
:=& \int \boldsymbol{\psi}^{(1)}\big(o;\hat{\boldsymbol{\eta}}\big)\mathrm{d}\mathbb{P}_n - \int \boldsymbol{\psi}^{(1)}\big(o;\hat{\boldsymbol{\eta}}\big)\mathrm{d}P + R_{n1} \\
=& \mathbb{P}_n\boldsymbol{\psi}^{(1)}(o;\hat{\boldsymbol{\eta}}) - P\boldsymbol{\psi}^{(1)}(o;\hat{\boldsymbol{\eta}}) + R_{n1} \\
=& (\mathbb{P}_n - P)\boldsymbol{\psi}^{(1)}(o;\boldsymbol{\eta}_0) + (\mathbb{P}_n - P)\big(\boldsymbol{\psi}^{(1)}(o;\hat{\boldsymbol{\eta}}) - \boldsymbol{\psi}^{(1)}(o;\boldsymbol{\eta}_0)\big) + R_{n1} \\
=& (\mathbb{P}_n - P)\boldsymbol{\psi}^{(1)}(o;\boldsymbol{\eta}_0) + R_{n2} + R_{n1},
\end{aligned}$$

where the second equality is guaranteed by our proposed procedure, and

$$\begin{aligned}
R_{n1} =& \hat{\boldsymbol{\alpha}}_t - \boldsymbol{\alpha}_t + \int \boldsymbol{\psi}^{(1)}\big(o;\hat{\boldsymbol{\eta}}\big)\mathrm{d}P \\
=& \mathbb{E}_n\big[\hat{p}_t(X)\big] - E_X\big[p_t(X)\big] \\
&+ E_{Y,T,X}\left[\frac{\mathbb{1}\{X\in\mathcal{A}_j\}}{\hat{\mathbb{P}}(\mathcal{A}_j)}\Big(\frac{\mathbb{1}\{T=1\}}{\hat{e}(X)} + \frac{\mathbb{1}\{T=0\}}{1-\hat{e}(X)}\Big)\big(Y - \hat{p}_T(X)\big)\right] \\
=& \mathbb{E}_n\big[\hat{p}_t(X)\big] - E_X\big[\hat{p}_t(X)\big] + E_X\big[\hat{p}_t(X)\big] - E_X\big[p_t(X)\big] \\
&+ E_{Y,T,X}\left[\frac{\mathbb{1}\{X\in\mathcal{A}_j\}}{\hat{\mathbb{P}}(\mathcal{A}_j)}\Big(\frac{\mathbb{1}\{T=1\}}{\hat{e}(X)} + \frac{\mathbb{1}\{T=0\}}{1-\hat{e}(X)}\Big)\big(Y - \hat{p}_T(X)\big)\right] \\
=& (\mathbb{P}_n - P)\hat{p}_t + P\boldsymbol{\psi}\big(o;\boldsymbol{\alpha}_t,\hat{\boldsymbol{\eta}}\big) \\
=& (\mathbb{P}_n - P)(\hat{p}_t - p_t) + (\mathbb{P}_n - P)p_t + P\boldsymbol{\psi}\big(o;\boldsymbol{\alpha}_t,\hat{\boldsymbol{\eta}}\big) \\
=& (\mathbb{P}_n - P)\boldsymbol{\psi}^{(2)}(o;\boldsymbol{\alpha},\boldsymbol{\eta}_0) + (\mathbb{P}_n - P)(\hat{p}_t - p_t) + (\mathbb{P}_n - P)p_t + P\boldsymbol{\psi}\big(o;\boldsymbol{\alpha}_t,\hat{\boldsymbol{\eta}}\big).
\end{aligned}$$

To this end, we have the final expansion:

$$\hat{\boldsymbol{\alpha}}_t - \boldsymbol{\alpha}_t = (\mathbb{P}_n - P)\boldsymbol{\psi}(o;\boldsymbol{\alpha},\boldsymbol{\eta}_0) + R_{n2} + R_{n3} + R_{n4},$$

where

$$\begin{aligned}
R_{n2} &= (\mathbb{P}_n - P)\big(\boldsymbol{\psi}^{(1)}(o;\hat{\boldsymbol{\eta}}) - \boldsymbol{\psi}^{(1)}(o;\boldsymbol{\eta}_0)\big), \\
R_{n3} &= (\mathbb{P}_n - P)(\hat{p}_t - p_t) \\
R_{n4} &= P\boldsymbol{\psi}\big(o;\boldsymbol{\alpha}_t,\hat{\boldsymbol{\eta}}\big).
\end{aligned}$$

Proof of semiparametric efficiency results with delta method

[161] shows that the asymptotic variance obtained by the delta method indeed achieves the semiparametric efficiency bound (Chapter 25.7). Since the results in [161] are formulated

using notation different from ours, in what follows, we provide a justification to show why the asymptotic variances of $\hat{\alpha}_{\text{RR}}$ and $\hat{\alpha}_{\text{OR}}$ attain the semiparametric efficiency In what follows, we justify why the asymptotic variances of $\hat{\alpha}_{ARD}$, $\hat{\alpha}_{RR}$, and $\hat{\alpha}_{OR}$ attain the semiparametric efficiency bound. For simplicity, we shall not discuss the subgroup case, but the justifications below can be easily extended to subgroup $\hat{\alpha}_{ARD}$, $\hat{\alpha}_{RR}$, and $\hat{\alpha}_{OR}$.

First, denote the joint density of (Y,T,X) as

$$f(y,t,x) = \big(f_1(y|x)e(x)\big)^t \big(f_0(y|x)(1-e(x))\big)^{1-t} f(x),$$

where $e(x) := f(t|x)$ denotes the propensity score, $f_1(y|x)$ denotes the conditional density of y given x under treatment, and $f_0(y|x)$ denotes the conditional density of y given x under control. Assume the parametric submodel indexed by parameter θ is

$$f(y,t,x;\theta) = \big(f_1(y|x;\theta)e(x;\theta)\big)^t \big(f_0(y|x;\theta)(1-e(x;\theta))\big)^{1-t} f(x;\theta),$$

where θ is indexed a finite-dimensional parameter β and an infinite-dimensional parameter η, i.e. $\theta = (\beta, \eta)$. The score function of the above parametric submodel is

$$s(y,t,x|\theta) = ts_1(y|x;\theta) + (1-t)s_0(y|x;\theta) + \frac{t-e(x;\theta)}{e(x;\theta)(1-e(x;\theta))}e'(x;\theta) + p(x;\theta), \quad (1.20)$$

where $s_1(y|x;\theta) = \frac{d}{d\theta}\log f_1(Y|X;\theta)$, $s_0(y|x;\theta) = \frac{d}{d\theta}\log f_0(Y|X;\theta)$, $p(x;\theta) = \frac{d}{d\theta}\log f(X;\theta)$, and $e'(x;\theta) = \frac{d}{d\theta}e(x;\theta)$. From the score function Eq (1.20) we obtain the tangent space $\mathcal{T}$ spanned by the score function $ts_1(y|x) + (1-t)s_0(y|x) + \frac{t-e(x)}{e(x)(1-e(x))}e'(x) + p(x)$. Now assume θ can be parametrized by β_1 and β_0, that is

$$\beta_1(\theta) = \int\int y f_1(y|x;\theta) f(x;\theta) dy dx,$$
$$\beta_0(\theta) = \int\int y f_0(y|x;\theta) f(x;\theta) dy dx.$$

$\beta_1(\theta)$ and $\beta_0(\theta)$ represent the conditional mean of outcome under treatment and control, respectively. The pathwise derivatives of $\beta_1(\theta)$ and $\beta_0(\theta)$ are

$$\frac{\partial\beta_1(\theta_0)}{\partial\theta} = \int\int y s_1(y|x;\theta_0) f_1(y|x) f(x) dy dx + \int\int \beta_1(x) p(x;\theta_0) f(x) dx,$$
$$\frac{\partial\beta_0(\theta_0)}{\partial\theta} = \int\int y s_0(y|x;\theta_0) f_0(y|x) f(x) dy dx + \int\int \beta_0(x) p(x;\theta_0) f(x) dx,$$

where θ_0 denotes the true parameter. If the target parameter of interest is the absolute risk

difference (ARD),

$$\begin{aligned}\frac{\partial \beta^{\mathrm{ARD}}(\theta_0)}{\partial \theta} &= \frac{\partial \beta^{\mathrm{ARD}}(\theta_0)}{\partial\big(\beta_1(\theta_0),\beta_0(\theta_0)\big)}\frac{\partial\big(\beta_1(\theta_0),\beta_0(\theta_0)\big)}{\partial\theta}\\ &= (1\ \ -1)\begin{pmatrix}\frac{\partial\beta_1(\theta_0)}{\partial\theta}\\ \frac{\partial\beta_0(\theta_0)}{\partial\theta}\end{pmatrix} = \frac{\partial\beta_1(\theta_0)}{\partial\theta} - \frac{\partial\beta_0(\theta_0)}{\partial\theta},\\ &= \int\int y s_1(y|x;\theta_0) f_1(y|x) f(x) dy dx + \int\int \beta_1(x) p(x;\theta_0) f(x) dx\\ &\quad - \int\int y s_0(y|x;\theta_0) f_0(y|x) f(x) dy dx - \int\int \beta_0(x) p(x;\theta_0) f(x) dx.\end{aligned}$$

Similarly, if the target parameter of interest is the relative risk,

$$\begin{aligned}\frac{\partial \beta^{\mathrm{RR}}(\theta_0)}{\partial \theta} &= \frac{\partial \beta^{\mathrm{RR}}(\theta_0)}{\partial\big(\beta_1(\theta_0),\beta_0(\theta_0)\big)}\frac{\partial\big(\beta_1(\theta_0),\beta_0(\theta_0)\big)}{\partial\theta}\\ &= \begin{pmatrix}\frac{1}{\beta_0(\theta_0)} & \frac{-\beta_1(\theta_0)}{\beta_0^2(\theta_0)}\end{pmatrix}\begin{pmatrix}\frac{\partial\beta_1(\theta_0)}{\partial\theta}\\ \frac{\partial\beta_0(\theta_0)}{\partial\theta}\end{pmatrix} = \frac{1}{\beta_0(\theta_0)}\frac{\partial\beta_1(\theta_0)}{\partial\theta} - \frac{\beta_1(\theta_0)}{\beta_0^2(\theta_0)}\frac{\partial\beta_0(\theta_0)}{\partial\theta},\\ &= \frac{1}{\beta_0(\theta_0)}\Big(\int\int y s_1(y|x;\theta_0) f_1(y|x) f(x) dy dx + \int\int \beta_1(x) p(x;\theta_0) f(x) dx\Big)\\ &\quad - \frac{\beta_1(\theta_0)}{\beta_0^2(\theta_0)}\Big(\int\int y s_0(y|x;\theta_0) f_0(y|x) f(x) dy dx + \int\int \beta_0(x) p(x;\theta_0) f(x) dx\Big).\end{aligned}$$

If the target parameter of interest is the odds ratio,

$$\begin{aligned}\frac{\partial \beta^{\mathrm{OR}}(\theta_0)}{\partial \theta} &= \frac{\partial \beta^{\mathrm{OR}}(\theta_0)}{\partial\big(\beta_1(\theta_0),\beta_0(\theta_0)\big)}\frac{\partial\big(\beta_1(\theta_0),\beta_0(\theta_0)\big)}{\partial\theta}\\ &= \begin{pmatrix}\frac{1-\beta_0(\theta_0)}{\beta_0(\theta_0)(1-\beta_1(\theta_0))^2} & \frac{-\beta_1(\theta_0)}{\beta_0^2(\theta_0)(1-\beta_1(\theta_0))}\end{pmatrix}\begin{pmatrix}\frac{\partial\beta_1(\theta_0)}{\partial\theta}\\ \frac{\partial\beta_0(\theta_0)}{\partial\theta}\end{pmatrix},\\ &= \frac{1-\beta_0(\theta_0)}{\beta_0(\theta_0)(1-\beta_1(\theta_0))^2}\frac{\partial\beta_1(\theta_0)}{\partial\theta} - \frac{\beta_1(\theta_0)}{\beta_0^2(\theta_0)(1-\beta_1(\theta_0))}\frac{\partial\beta_0(\theta_0)}{\partial\theta},\\ &= \frac{1-\beta_0(\theta_0)}{\beta_0(\theta_0)(1-\beta_1(\theta_0))^2}\Big(\int\int y s_1(y|x;\theta_0) f_1(y|x) f(x) dy dx\\ &+ \int\int \beta_1(x) p(x;\theta_0) f(x) dx\Big)\\ &\quad - \frac{\beta_1(\theta_0)}{\beta_0^2(\theta_0)(1-\beta_1(\theta_0))}\Big(\int\int y s_0(y|x;\theta_0) f_0(y|x) f(x) dy dx\\ &+ \int\int \beta_0(x) p(x;\theta_0) f(x) dx\Big).\end{aligned}$$

Now let

$$\varphi^{\mathrm{ARD}}(Y,T,X) = \frac{T}{e(X)}\big(Y-\beta_1(X)\big) - \frac{1-T}{1-e(X)}\big(Y-\beta_0(X)\big) + \beta_1(X) - \beta_0(X) - (\beta_1-\beta_0),$$

$$\begin{aligned}\varphi^{\mathrm{RR}}(Y,T,X) = {}& \frac{1}{\beta_0}\Big(\frac{T}{e(X)}\big(Y-\beta_1(X)\big) + \beta_1(X) - \beta_1\Big) \\ & - \frac{\beta_1}{\beta_0^2}\Big(\frac{1-T}{1-e(X)}\big(Y-\beta_0(X)\big) + \beta_0(X) - \beta_0\Big),\end{aligned}$$

$$\begin{aligned}\varphi^{\mathrm{OR}}(Y,T,X) = {}& \frac{1-\beta_0}{\beta_0(1-\beta_1)^2}\Big(\frac{T}{e(X)}\big(Y-\beta_1(X)\big) + \beta_1(X) - \beta_1\Big) \\ & - \frac{\beta_1}{\beta_0^2(1-\beta_1)}\Big(\frac{1-T}{1-e(X)}\big(Y-\beta_0(X)\big) + \beta_0(X) - \beta_0\Big).\end{aligned}$$

Taking the product of $\varphi(Y,T,X)$ and the score function in Eq (1.20), we observe that

$$\begin{aligned}\frac{\partial \beta^{\mathrm{ARD}}(\theta_0)}{\partial \theta_0} &= \mathbb{E}[\varphi^{\mathrm{ARD}}(Y,T,X)\cdot s(Y,T,X|\theta_0)], \\ \frac{\partial \beta^{\mathrm{RR}}(\theta_0)}{\partial \theta_0} &= \mathbb{E}[\varphi^{\mathrm{RR}}(Y,T,X)\cdot s(Y,T,X|\theta_0)], \\ \frac{\partial \beta^{\mathrm{OR}}(\theta_0)}{\partial \theta_0} &= \mathbb{E}[\varphi^{\mathrm{OR}}(Y,T,X)\cdot s(Y,T,X|\theta_0)].\end{aligned}$$

The above derivations suggest that $\varphi^{\mathrm{ARD}}(Y,T,X)$, $\varphi^{\mathrm{RR}}(Y,T,X)$, and $\varphi^{\mathrm{OR}}(Y,T,X) \in \mathcal{T}$, and thus the semiparametric efficiency bounds of ARD, RR, and OR can be computed as $\mathbb{E}\big[\big(\varphi^{\mathrm{ARD}}(Y,T,X)\big)^2\big]$, $\mathbb{E}\big[\big(\varphi^{\mathrm{RR}}(Y,T,X)\big)^2\big]$, and $\mathbb{E}\big[\big(\varphi^{\mathrm{OR}}(Y,T,X)\big)^2\big]$, respectively.

The above derivations suggest that the semiparametric efficiency results in Theorem 1 can be generalized to various estimators of interest.

Proof of Lemma 1

To simplify notation, suppose we work with a single update and denote the initial estimate as $\hat{p}_t^{\mathrm{Init}}(\cdot)$. The primal optimization problem is:

$$\begin{aligned}\min_{\boldsymbol{\varepsilon}\in\mathbb{R}^J} \quad & -\frac{1}{n}\sum_{i=1}^{n}\Big[Y_i\Big(\mathrm{logit}\big(\hat{p}_t^{\mathrm{Init}}(X_i)\big) + \sum_{j=1}^{J}\varepsilon_j\cdot\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i=t)}{\hat{e}(X_i)}\Big) \\ & \qquad - \log\Big(1+\exp^{\mathrm{logit}\left(\hat{p}_t^{\mathrm{Init}}(X_i)\right)+\sum_{j=1}^{J}\varepsilon_j\cdot\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i=t)}{\hat{e}(X_i)}}\Big)\Big], \\ \text{s.t.} \quad & ||\boldsymbol{\varepsilon}||_2 - \delta \le 0.\end{aligned}$$

The Lagrangian associated with the primal problem is

$$L(\boldsymbol{\varepsilon}, \lambda) = -\frac{1}{n}\sum_{i=1}^{n}\Big[Y_i\Big(\operatorname{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\varepsilon_j \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big) \\ - \log\Big(1 + \exp^{\operatorname{logit}\left(\hat{p}_t^{\text{Init}}(X_i)\right) + \sum_{j=1}^{J}\varepsilon_j \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}}\Big)\Big] + \lambda(||\boldsymbol{\varepsilon}||_2 - \delta),$$

where $\lambda \geq 0$ is the Lagrange multiplier. The Lagrangian primal problem is defined as

$$\min_{\boldsymbol{\varepsilon}} L_{\texttt{primal}}(\boldsymbol{\varepsilon}, \lambda) = \min_{\boldsymbol{\varepsilon}} \max_{\lambda \geq 0} L(\boldsymbol{\varepsilon}, \lambda).$$

The Lagrangian dual function is thus defined as

$$L_{\texttt{dual}}(\lambda) = \min_{\boldsymbol{\varepsilon}} L(\boldsymbol{\varepsilon}, \lambda),$$

and the Lagrangian dual problem is

$$\max_{\lambda \geq 0} L_{\texttt{dual}}(\lambda) = \max_{\lambda \geq 0} \min_{\boldsymbol{\varepsilon}} L(\boldsymbol{\varepsilon}, \lambda).$$

Given our optimization problem is a convex problem, there is no duality gap between the primal and dual problems.

Next, we solve for $\hat{\varepsilon}_j$ by taking derivative of the dual function.

$$\begin{aligned}
\frac{\partial L_{\texttt{dual}}(\hat{\boldsymbol{\varepsilon}}, \lambda)}{\partial \varepsilon_j} &= -\frac{1}{n}\sum_{i:T_i=t}\Big\{Y_i\frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)} \\
&\quad - \frac{\exp\Big(\operatorname{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\hat{\varepsilon}_j \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big)}{1 + \exp\Big(\operatorname{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\hat{\varepsilon}_j \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big)} \\
&\quad \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big\} + \lambda\varepsilon_j, \\
&= -\frac{1}{n}\sum_{i:T_i=t}\Big\{\Big[Y_i - \frac{\exp\Big(\operatorname{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\hat{\varepsilon}_j \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big)}{1 + \exp\Big(\operatorname{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\hat{\varepsilon}_j \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big)}\Big] \\
&\quad \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)}\Big\} + \lambda\varepsilon_j, \\
&= -\frac{1}{n}\sum_{i:T_i=t}\Big\{\Big(Y_i - \hat{p}_t(X_i)\Big) \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\frac{1}{\hat{e}(X_i)}\Big\} + \lambda\varepsilon_j, \\
\frac{\partial L_{\texttt{dual}}(\hat{\boldsymbol{\varepsilon}}, \lambda)}{\partial \lambda} &= -\hat{\phi}_j(X_i) + \lambda\varepsilon_j = 0, \\
&\implies \hat{\phi}_j(X_i) = \lambda\varepsilon_j, \quad \varepsilon_j = \frac{\hat{\phi}_j(X_i)}{\lambda}.
\end{aligned}$$

By plugging in $\varepsilon_j = \frac{\hat{\phi}_j(X_i)}{\lambda}$, we have the Lagrangian dual function equals to

$$L_{\texttt{dual}}(\lambda) = -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\frac{\hat{\phi}_j(X_i)\hat{S}_j(X_i)}{\lambda}\Big) - \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{\text{Init}}(X_i)\right)+\sum_{j=1}^{J}\frac{\hat{\phi}_j(X_i)\hat{S}_j(X_i)}{\lambda}}\Big)\Big] + \lambda(\frac{||\hat{\boldsymbol{\phi}}(\boldsymbol{X_i})||_2}{\lambda} - \delta).$$

Hence, the dual problem reduces to

$$\max_{\lambda\geq 0}\ -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \sum_{j=1}^{J}\frac{\hat{\phi}_j(X_i)\hat{S}_j(X_i)}{\lambda}\Big) - \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{\text{Init}}(X_i)\right)+\sum_{j=1}^{J}\frac{\hat{\phi}_j(X_i)\hat{S}_j(X_i)}{\lambda}}\Big)\Big] + \lambda(\frac{||\hat{\boldsymbol{\phi}}(\boldsymbol{X_i})||_2}{\lambda} - \delta).$$

With the following substitution,

$$\sum_{j=1}^{J}\frac{\hat{\phi}_j(X_i)\hat{S}_j(X_i)}{\lambda} = \frac{\sum_{j=1}^{J}\hat{\phi}_j\hat{S}_j(X_i)}{||\hat{\boldsymbol{\phi}}(X_i)||_2}\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda} = \tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda},$$

the dual problem satisfies

$$\begin{aligned}
&\arg\max_{\lambda\geq 0}\ -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}\Big) \\
&\qquad - \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{\text{Init}}(X_i)\right)+\tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}}\Big)\Big] + \lambda(\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda} - \delta), \\
=&\arg\max_{\lambda\geq 0}\ -\frac{1}{n}\sum_{i:T_i=t}\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{\text{Init}}(X_i)\big) + \tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}\Big) \\
&\qquad - \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{\text{Init}}(X_i)\right)+\tilde{S}_t(X_i)\cdot\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2}{\lambda}}\Big)\Big] - \lambda\delta.
\end{aligned}$$

Through the reparametrization $\gamma = \frac{||\hat{\phi}(X_i)||_2}{\delta}$, the dual problem can be reformulated as

$$\begin{aligned}
\arg\min_{\gamma>0} -\frac{1}{n}\sum_{i:T_i=t}&\Big[Y_i\Big(\text{logit}\big(\hat{p}_t^{(k-1)}(X_i)\big) + \tilde{S}_t(X_i)\cdot\gamma\Big) \\
&- \log\Big(1+\exp^{\text{logit}\left(\hat{p}_t^{(k-1)}(X_i)\right)+\tilde{S}_t(X_i)\cdot\gamma}\Big)\Big] - \frac{||\hat{\boldsymbol{\phi}}(X_i)||_2\cdot\delta}{\gamma},
\end{aligned}$$

where $\frac{||\hat{\boldsymbol{\phi}}(X_i)||_2\cdot\delta}{\gamma}$ is sufficiently close to 0.

Additional Remarks

Remark 1. *The multivariate local least favorable submodel defined in Eq (2) implies a one-dimensional universal least favorable submodel. The resulting one-step TMLE along this universal least favorable submodel corresponds with iteratively maximizing log likelihood for the multivariate least favorable submodel in $\boldsymbol{\varepsilon}$ under a constraint that $||\boldsymbol{\varepsilon}|| = \gamma$ for some small γ, and noting that this MLE $\boldsymbol{\varepsilon}$ is known in closed form and equals $\gamma \cdot \frac{P_n\varphi}{||P_n\varphi||}$. One stops the iteration when the log-likelihood reaches its maximum. Our proposed method in Eq (3) can be viewed as a variation of this in the targeted learning literature.*

Remark 2. *The above theoretical results can be readily extended to infinitely many subgroups cases. Concretely, we denote $\boldsymbol{\alpha}_t(\boldsymbol{\nu})$ as a set of subgroup parameters indexed by a continuous vector $\boldsymbol{\nu}$ defined on a compact parameter space, where $\boldsymbol{\alpha}_t(\boldsymbol{\nu}) = P\big(Y(t) = 1 | X \in \mathcal{A}(\boldsymbol{\nu})\big)$, and we define the vector of efficient influence functions as $\boldsymbol{\varphi}_t(\boldsymbol{\nu})$. As long as $\boldsymbol{\alpha}_t(\boldsymbol{\nu})$ is a smooth function of $\boldsymbol{\nu}$ the function class $\{\boldsymbol{\varphi}(o; \boldsymbol{\alpha}_t(\boldsymbol{\nu}), \boldsymbol{\eta}), \boldsymbol{\eta} \in \mathcal{H}\}$ is a Donsker class, our theoretical results suggest that $\sqrt{n}(\hat{\boldsymbol{\alpha}}_t(\boldsymbol{\nu}) - \boldsymbol{\alpha}_t(\boldsymbol{\nu}))$ converges to a Gaussian process.*

Extension of the Proposed Method

Extension to continuous outcomes

There are two options to adapt our proposed method to continuous outcomes. The first option is to use a different updating procedure for continuous outcomes [63]. For example, $\hat{p}_t^{(k)}$ can be obtained through a linear update:

$$\hat{p}_t^{(k)}(X_i) = \hat{p}_t^{(k-1)}(X_i) + \hat{\gamma}^{(k)} \cdot \tilde{S}_t^{(k-1)}(X_i), \quad i \in \{i : T_i = t\},$$

where

$$\tilde{S}_t^{(k-1)}(X_i) = \frac{\sum_{j=1}^{d} \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}_t(X_i)} \cdot \Big(\sum_{l=1}^{n} \hat{\phi}_j^{(k-1)}(Y_l, T_l, X_l) \Big)}{\sqrt{\sum_{j=1}^{d} \Big(\sum_{l=1}^{n} \hat{\phi}_j^{(k-1)}(Y_l, T_l, X_l) \Big)^2}},$$

and $\hat{\phi}_j^{(k-1)}(Y_i, T_i, X_i) = \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{P}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}_t(X_i)} (Y_i - \hat{p}_t^{(k-1)}(X_i))$. $\hat{\gamma}^{(k)}$ can then be obtained by minimizing a user specified loss function $L(\cdot)$,

$$\hat{\gamma}^{(k)} = \arg\min_{\gamma \geq 0} \frac{1}{n} \sum_{i:T_i = t} L\Big(Y_i,\ \hat{p}_t^{(k-1)}(X_i), \gamma,\ \tilde{S}_t^{(k-1)}(X_i)\Big).$$

For example, one can consider the l_2-loss [63], that is

$$\hat{\gamma}^{(k)} = \arg\min_{\gamma \geq 0} \frac{1}{n} \sum_{i:T_i = t} ||Y_i - \hat{p}_t^{(k)}(X_i)||_2^2.$$

The second option is to dichotomize a continuous outcome into a binary outcome [111, 103]. Then the proposed methodology in the main chapter can be directly applied.

Simultaneous confidence intervals of absolute risk difference, relative risk, and odds ratio

Let $\hat{\kappa}(q, J)$ be a consistent estimate of the $(1-q)$-th quantile of $\max_{j\in 1,\ldots,J}|Z_j|$, where $Z_1, \ldots, Z_J$ are i.i.d. standard normal random variables. Then,

$$\hat{\alpha}_{\mathrm{ARD},j} \pm \hat{\kappa}_{\mathrm{ARD},q/2}\Big(\frac{\hat{\Sigma}_{\mathrm{ARD},jj}}{n}\Big)^{1/2}, \quad \hat{\alpha}_{\mathrm{RR},j} \pm \hat{\kappa}_{\mathrm{RR},q/2}\Big(\frac{\hat{\Sigma}_{\mathrm{RR},jj}}{n}\Big)^{1/2}, \quad \hat{\alpha}_{\mathrm{OR},j} \pm \hat{\kappa}_{\mathrm{OR},q/2}\Big(\frac{\hat{\Sigma}_{\mathrm{OR},jj}}{n}\Big)^{1/2}$$

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{\alpha}_{\mathrm{ARD},j} \pm \hat{\kappa}(q,J)\cdot\Big(\frac{\hat{\Sigma}_{\mathrm{ATE},jj}}{n}\Big)^{1/2}, j=1,\ldots,J\right) = 1-q,$$
$$\lim_{n\to\infty} \mathbb{P}\left(\hat{\alpha}_{\mathrm{RR},j} \pm \hat{\kappa}(q,J)\cdot\Big(\frac{\hat{\Sigma}_{\mathrm{RR},jj}}{n}\Big)^{1/2}, j=1,\ldots,J\right) = 1-q,$$
$$\lim_{n\to\infty} \mathbb{P}\left(\hat{\alpha}_{\mathrm{OR},j} \pm \hat{\kappa}(q,J)\cdot\Big(\frac{\hat{\Sigma}_{\mathrm{OR},jj}}{n}\Big)^{1/2}, j=1,\ldots,J\right) = 1-q,$$

where $\hat{\Sigma}_{\mathrm{ARD}} = \big(\hat{\Sigma}_{\mathrm{ARD},jk}\big)_{j,k=1}^{J} = \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\varphi}}_{\mathrm{ARD},i}\hat{\boldsymbol{\varphi}}'_{\mathrm{ARD},i}$,

$$\hat{\boldsymbol{\varphi}}_{\mathrm{ARD},i} = \big(\hat{\varphi}_{\mathrm{ARD},1}(Y_i,T_i,X_i),\ldots,\hat{\varphi}_{\mathrm{ARD},J}(Y_i,T_i,X_i)\big)'$$

. Similarly, we can construct the covariance matrix $\hat{\Sigma}_{\mathrm{RR}}$ and $\hat{\Sigma}_{\mathrm{OR}}$. The plug-in estimates of the efficient influence functions are,

$$\begin{aligned}
\hat{\varphi}_{\mathrm{ARD},j}(O_i) &= \frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)}\Big[\Big(\frac{T_i}{\hat{e}_1(X_i)}\big(Y_i-\hat{p}_1(X_i)\big)+\hat{p}_1(X_i)-\hat{\alpha}_1\Big)\\
&\quad-\Big(\frac{1-T_i}{\hat{e}_0(X_i)}\big(Y_i-\hat{p}_0(X_i)\big)+\hat{p}_0(X_i)-\hat{\alpha}_0\Big)\Big],\\
\hat{\varphi}_{\mathrm{RR},j}(O_i) &= \frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)}\Big[\frac{1}{\hat{\alpha}_0}\Big(\frac{T_i}{\hat{e}_1(X_i)}\big(Y_i-\hat{p}_1(X_i)\big)+\hat{p}_1(X_i)-\hat{\alpha}_1\Big)\\
&\quad-\frac{\hat{\alpha}_1}{\hat{\alpha}_0^2}\Big(\frac{1-T_i}{\hat{e}_0(X_i)}\big(Y_i-\hat{p}_0(X_i)\big)+\hat{p}_0(X_i)-\hat{\alpha}_0\Big)\Big],\\
\hat{\varphi}_{\mathrm{OR},j}(O_i) &= \frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)}\Big[\frac{1-\hat{\alpha}_0}{\hat{\alpha}_0(1-\hat{\alpha}_1)^2}\Big(\frac{T_i}{\hat{e}_1(X_i)}\big(Y_i-\hat{p}_1(X_i)\big)+\hat{p}_1(X_i)-\hat{\alpha}_1\Big)\\
&\quad-\frac{\hat{\alpha}_1}{\hat{\alpha}_0^2(1-\hat{\alpha}_1)}\Big(\frac{1-T_i}{\hat{e}_0(X_i)}\big(Y_i-\hat{p}_0(X_i)\big)+\hat{p}_0(X_i)-\hat{\alpha}_0\Big)\Big].
\end{aligned}$$

We obtain the efficient influence functions for $\boldsymbol{\alpha}_{\mathbf{ARD}}$, $\boldsymbol{\alpha}_{\mathbf{RR}}$, and $\boldsymbol{\alpha}_{\mathbf{OR}}$ by applying multivariate delta method on $(\boldsymbol{\alpha}_{\mathbf{1}}, \boldsymbol{\alpha}_{\mathbf{0}})$.

Proof of Eq (7) in Section 1.3

In this part, we aim to derive the score function under the iterative procedure. For simplicity, denote the final update as $\hat{p}_1(X)$, $\hat{\varepsilon}$. The conditional likelihood function of Y given (T, X) is:

$$L(Y|T,X) = p(T,X)^Y \cdot \big(1 - p(T,X)\big)^{1-Y}.$$

Proof.

$$\begin{aligned}
\frac{\partial \log L(Y|T,X;\varepsilon)}{\partial \varepsilon} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{J}\Big\{Y_i\tilde{S}_1(X_i) \\
&- \frac{\operatorname{expit}\Big(\operatorname{logit}\big(\hat{p}_1^{(K-1)}(X_i)\big) + \hat{\varepsilon}\cdot\tilde{S}_1(X_i)\Big)}{1+\operatorname{expit}\Big(\operatorname{logit}\big(\hat{p}_1^{(K-1)}(X_i)\big) + \hat{\varepsilon}\cdot\tilde{S}_1(X_i)\Big)}\cdot\tilde{S}_1(X_i)\Big\} = 0, \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{J}\Big\{\Big[Y_i \\
&- \frac{\operatorname{expit}\Big(\operatorname{logit}\big(\hat{p}_1^{(K-1)}(X_i)\big) + \hat{\varepsilon}\cdot\tilde{S}_1(X_i)\Big)}{1+\operatorname{expit}\Big(\operatorname{logit}\big(\hat{p}_1^{(K-1)}(X_i)\big) + \hat{\varepsilon}\cdot\tilde{S}_1(X_i)\Big)}\Big]\cdot\tilde{S}_1(X_i)\Big\} = 0, \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{J}\Big\{\Big(Y_i - \hat{p}_1(X_i)\Big)\cdot\tilde{S}_1(X_i)\Big\} = 0, \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{J}\Big\{\Big(Y_i - \hat{p}_1(X_i)\Big)\cdot\frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)}\frac{T_i}{\hat{e}_1(X_i)} \\
&\cdot \frac{\frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_j(Y_i,T_i,X_i)}{\sqrt{\sum_{m=1}^{J}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_m(Y_i,T_i,X_i)\Big)^2}}\Big\}, \\
&= \frac{\sum_{j=1}^{J}\Big\{\frac{1}{n}\sum_{i=1}^{n}\big\{\hat{\phi}_j(Y_i,T_i,X_i)\big\}\frac{1}{n}\sum_{i=1}^{n}\big\{\hat{\phi}_j(Y_i,T_i,X_i)\big\}\Big\}}{\sqrt{\sum_{m=1}^{J}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_m(Y_i,T_i,X_i)\Big)^2}}, \\
&= \frac{\sum_{j=1}^{J}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_j(Y_i,T_i,X_i)\Big)^2}{\sqrt{\sum_{m=1}^{J}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_m(Y_i,T_i,X_i)\Big)^2}} \\
&= \sqrt{\sum_{j=1}^{J}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_j(Y_i,T_i,X_i)\Big)^2} = 0.
\end{aligned}$$

□

Implementation Details

Implementation details of cross-fitted TMLE

As mentioned in the main chapter, the Donsker class condition on the efficient influence function can be relaxed by cross-fitting. Here, we briefly discuss the implementation details of the cross-fitted iterative version of the one-step TMLE of the multivariate dimensional parameters. The non-iterative version can be carried out similarly.

Step 1. Randomly split the sample into V equal-sized subsamples.

Step 2. For $v \leftarrow 1$ to V:

(a) Use subsample v as the validation data and the rest as training data. Generate initial estimates of $p_t(X)$ and $e_t(X)$ by fitting the model on the training set, and predict on the validation set, denoted as $\hat{p}_{t,v}^{(0)}(X)$ and $\hat{e}_{t,v}(X)$.

(b) For $k \leftarrow 1, \ldots, K$ (or until converge):

$$\hat{\varepsilon}_v^{(k)} = \arg\max_{\varepsilon \in \mathbb{R}} \Bigg\{ \frac{1}{n_v} \sum_{i:T_i=t, i \in v} \Big[Y_i \Big(\text{logit}\big(\hat{p}_{t,v}^{(k-1)}(X_i)\big) + \varepsilon \tilde{S}_{t,v}^{(k-1)}(X_i) \Big) - \log\Big(1 + \exp^{\text{logit}\left(\hat{p}_{t,v}^{(k-1)}(X_i)\right) + \varepsilon \tilde{S}_{t,v}^{(k-1)}(X_i)} \Big) \Big] \Bigg\},$$

where

$$\tilde{S}_{t,v}^{(k-1)}(X_i) = \frac{\sum_{j=1}^{d} \frac{1}{n_v} \sum_{l \in v} \hat{\phi}_j^{(k-1)}(Y_l, T_l, X_l)}{\sqrt{\sum_{m=1}^{d} \left(\frac{1}{n_v} \sum_{i \in v} \hat{\phi}_m^{(k-1)}(Y_i, T_i, X_i) \right)^2}} \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)} \frac{T_i}{\hat{e}(X_i)},$$

and $\hat{\phi}_{j,v}^{(k-1)} = \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)} \frac{T_i}{\hat{e}_t(X_i)} (Y_i - \hat{p}_{t,v}^{(k-1)}(X_i))$.

(c) Update the conditional risk via:

$$\hat{p}_{t,v}^{(k)}(X_i) = \text{expit}\left(\text{logit}\big(\hat{p}_{t,v}^{(k-1)}(X_i)\big) + \varepsilon_v^{(k)} \cdot \tilde{S}_{t,v}^{(k-1)}(X_i) \right).$$

(d) Estimate $\alpha_{j,v}$ by:

$$\hat{\alpha}_{j,v} = \frac{\sum_{i \in v} \mathbb{1}(X_i \in \mathcal{A}_j) \cdot \hat{p}_{t,v}^{(K)}(X_i)}{\sum_{i \in v} \mathbb{1}(X_i \in \mathcal{A}_j)}.$$

Step 3. Aggregate estimates from the validation sets by:

$$\hat{\alpha}_j = \frac{1}{V}\sum_{v=1}^{V} \hat{\alpha}_{j,v}.$$

Implementation details of the relative risk and odds ratio estimators

Step 1. Randomly split the sample into V equal-sized subsamples.

Step 2. For $v \leftarrow 1$ to V:

(a) Use subsample v as the validation data and the rest as training data. Generate initial estimates of $p_1(X)$, $p_0(X)$, $e_1(X)$ and $e_0(X)$ by fitting the model on the training set, and predict on the validation set, denoted as $\hat{p}^{(0)}_{1,v}(X)$, $\hat{p}^{(0)}_{0,v}(X)$, $\hat{e}_{1,v}(X)$, and $\hat{e}_{0,v}(X)$.

(b) For $k \leftarrow 1, \ldots, K$ (or until converge):

$$\hat{\varepsilon}^{(k)}_{t,v} = \arg\max_{\varepsilon\in\mathbb{R}} \Bigg\{ \frac{1}{n_v}\sum_{i\in v} \Big[Y_i \Big(\text{logit}\big(\hat{p}^{(k-1)}_{t,v}(X_i)\big) + \varepsilon \tilde{S}^{(k-1)}_{t,v}(X_i) \Big) - \log\Big(1 + \exp^{\text{logit}\left(\hat{p}^{(k-1)}_{t,v}(X_i)\right) + \varepsilon \tilde{S}^{(k-1)}_{t,v}(X_i)}\Big)\Big]\Bigg\},$$

where

$$\tilde{S}^{(k-1)}_{t,v}(X_i) = \frac{\sum_{j=1}^{d} \frac{1}{n_v}\sum_{l\in v} \hat{\phi}^{(k-1)}_j(Y_l, T_l, X_l)}{\sqrt{\sum_{m=1}^{d}\left(\frac{1}{n_v}\sum_{i\in v}\hat{\phi}^{(k-1)}_m(Y_i, T_i, X_i)\right)^2}} \cdot \frac{\mathbb{1}(X_i \in \mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)} \frac{\mathbb{1}(T_i = t)}{\hat{e}(X_i)},$$

and $\hat{\phi}^{(k-1)}_{j,v} = \frac{\mathbb{1}(X_i\in\mathcal{A}_j)}{\hat{\mathbb{P}}(\mathcal{A}_j)} \frac{T_i}{\hat{e}_t(X_i)}(Y_i - \hat{p}^{(k-1)}_{t,v}(X_i))$.

(c) Update the conditional risk via:

$$\hat{p}^{(k)}_{t,v}(T_i, X_i) = \text{expit}\left(\text{logit}\big(\hat{p}^{(k-1)}_{t,v}(X_i)\big) + \varepsilon^{(k)}_{t,v} \cdot \tilde{S}^{(k-1)}_{1,v}(X_i)\right).$$

(d) Estimate $\alpha_{j,v}$ by:

$$\hat{\alpha}_{tj,v} = \frac{\sum_{i\in v}\mathbb{1}(X_i \in \mathcal{A}_j)\cdot \hat{p}^{(K)}_{t,v}(X_i)}{\sum_{i\in v}\mathbb{1}(X_i\in\mathcal{A}_j)}.$$

Step 3. Aggregate estimates from the validation sets by:

$$\hat{\alpha}_{tj} = \frac{1}{V}\sum_{v=1}^{V}\hat{\alpha}_{tj,v}.$$

Step 4. Estimate $\boldsymbol{\alpha}_{\text{RR}}$ and $\boldsymbol{\alpha}_{\text{OR}}$ as

$$\hat{\boldsymbol{\alpha}}_{\mathbf{RR}} = \Big(\frac{\hat{\alpha}_{1,1}}{\hat{\alpha}_{0,1}}, \ldots, \frac{\hat{\alpha}_{1,d}}{\hat{\alpha}_{0,d}}\Big), \quad \hat{\boldsymbol{\alpha}}_{\mathbf{OR}} = \Big(\frac{\hat{\alpha}_{1,1}}{1-\hat{\alpha}_{1,1}} \Big/ \frac{\hat{\alpha}_{0,1}}{1-\hat{\alpha}_{0,1}}, \ldots, \frac{\hat{\alpha}_{1,d}}{1-\hat{\alpha}_{1,d}} \Big/ \frac{\hat{\alpha}_{0,d}}{1-\hat{\alpha}_{0,d}}\Big).$$

Additional Simulation Results

Comparison of the targeted learning approaches discussed in Section 1.3

In this simulation study, we compare the proposed iTMLE method with the conventional targeted learning approach without targeting multiple subgroups. That is, we compare the performances for four estimators: (1) "TMLE-single," which is the TMLE without targeting multiple subgroups as discussed in Section 1.3, (2) "TMLE-multiple," which is the one-step TMLE estimator that targets multiple subgroups discussed in Eq (2), (3) "TMLE-multiple-ulfm," which is the same method as (2) but operates under the universal least favorable submodel, and (4) "iTMLE," which is an iterative version of the one-step TMLE estimator discussed in Eq (3).

Figure 1.4 demonstrates that when $d = 4$, iTMLE has smaller bias, smaller variance, and lower FWER. When the number of subgroups is not too large ($d = 4$), targeting one subgroup at a time ("TMLE-single") yields smaller bias than targeting multiple subgroups ("TMLE-multiple"), while "TMLE-single" loses control over bias, variance, and FWER when the number of subgroups is large. "TMLE-multiple-ulfm" shows similar performance to iTMLE, but with slightly larger bias and variance.

Misspecified propensity score model

In this section, we compare the performance of the discussed method with other conventional estimators under the mis-specified propensity score model. The results from Figure 1.5 and Figure 1.6 are in-line with simulation studies in Section 6 of the main chapter.

An alternative simulation design

Kindly pointed out by an anonymous reviewer, the simulation design adopted in our main chapter produces rather deterministic outcomes. Therefore, we provide additional simulation results under an alternative simulation design:

$$\begin{aligned} X = (X_1, \ldots, X_5)^\intercal \sim N(0, \Sigma), \quad \Sigma_{ij} = 0.5^{|i-j|}, \\ T \sim \text{Bernoulli}\Big(\text{expit}(X_1 - 0.5 \cdot X_2 + 0.25 \cdot X_3 + 0.1 \cdot X_4)\Big), \\ Y|T, X \sim \text{Bernoulli}\Big(\text{expit}(T + \cdot X_1 + \cdot X_2 + \cdot X_3 + \cdot X_4)\Big). \end{aligned}$$

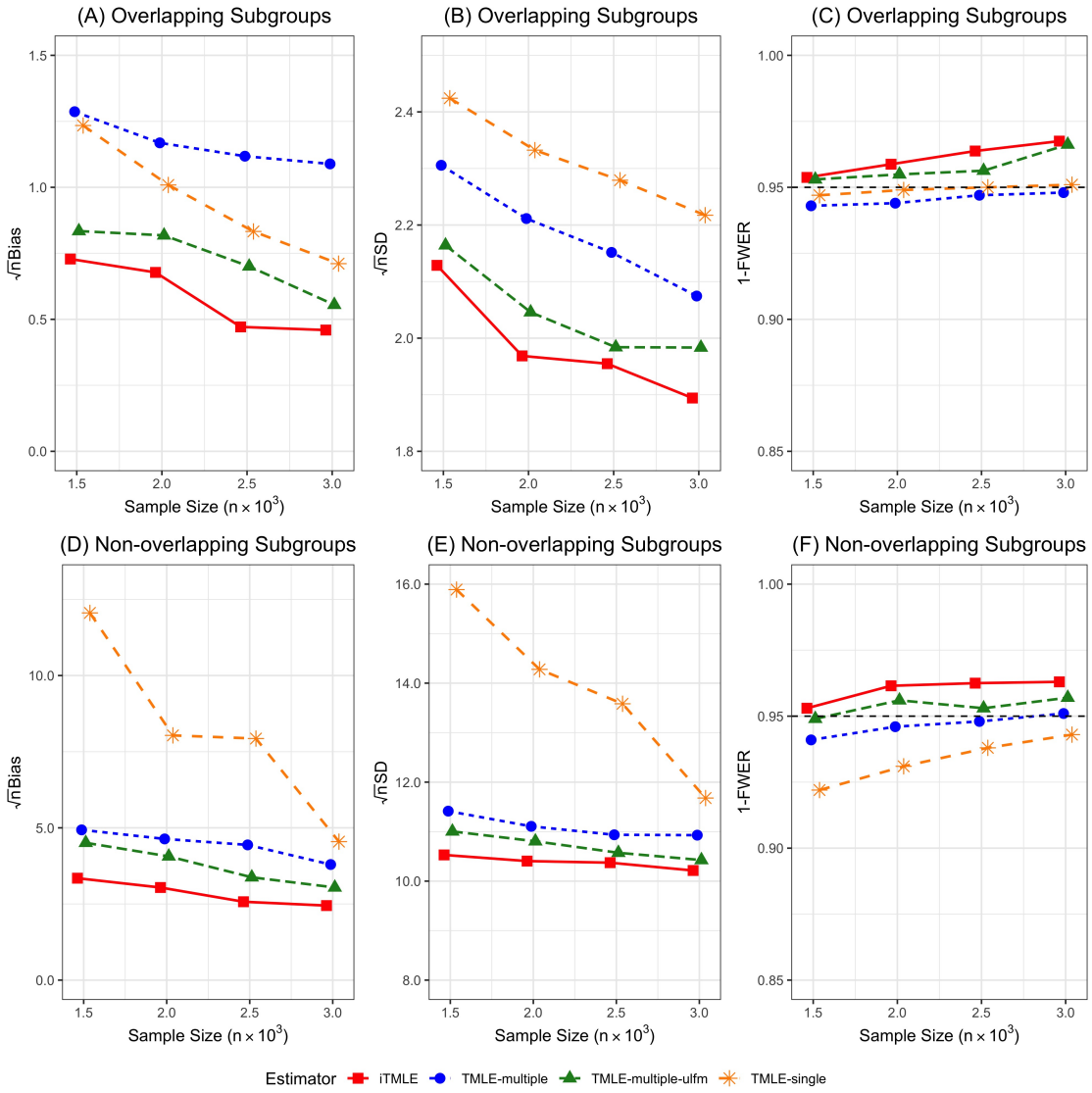


Figure 1.4: Comparison of four TMLE estimators under overlapping ($d = 4$) and non-overlapping subgroups ($d = 10$). The considered methods include (1) iTMLE; (2) TMLE-single; (3) TMLE-multiple; (4) TMLE-multiple-ulfm. All four methods use the random forest for the initial estimations. The maximum Monte Carlo standard error of (1-FWER) across the four estimators is 0.027. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

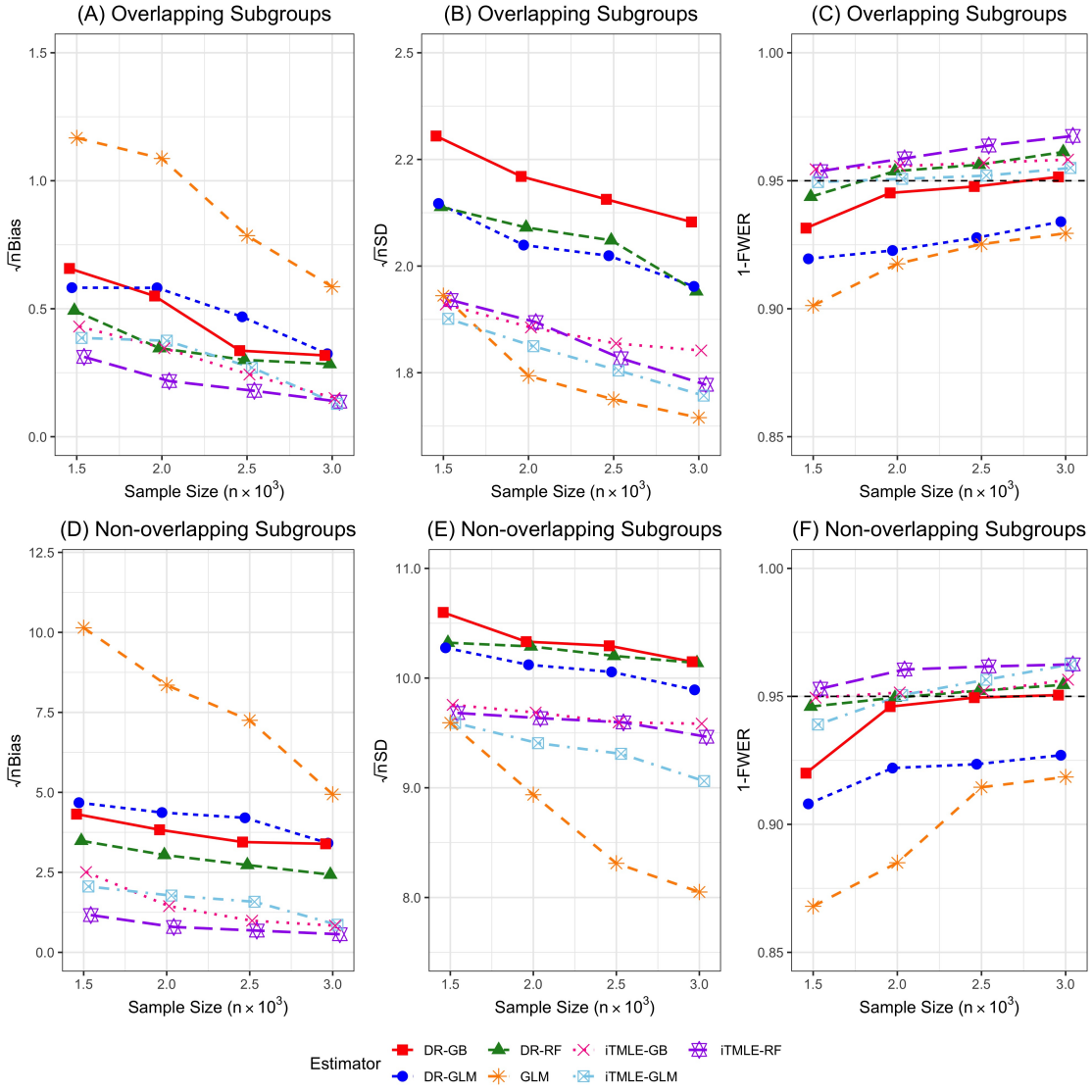


Figure 1.5: Comparison of iTMLE with the conventional methods. "iTMLE" denotes the discussed method. "DR" denotes the doubly robust estimator. "GLM" denotes the generalized linear models. The maximum Monte Carlo standard error of (1-FWER) across the four estimators is 0.030. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

We provide the distribution of $p_t(X) = P(Y = 1|T = t, X)$ under the alternative simulation design in Figured 1.7 (A) and (B), and under the simulation design adopted in the main chapter in Figured 1.7 (C) and (D).

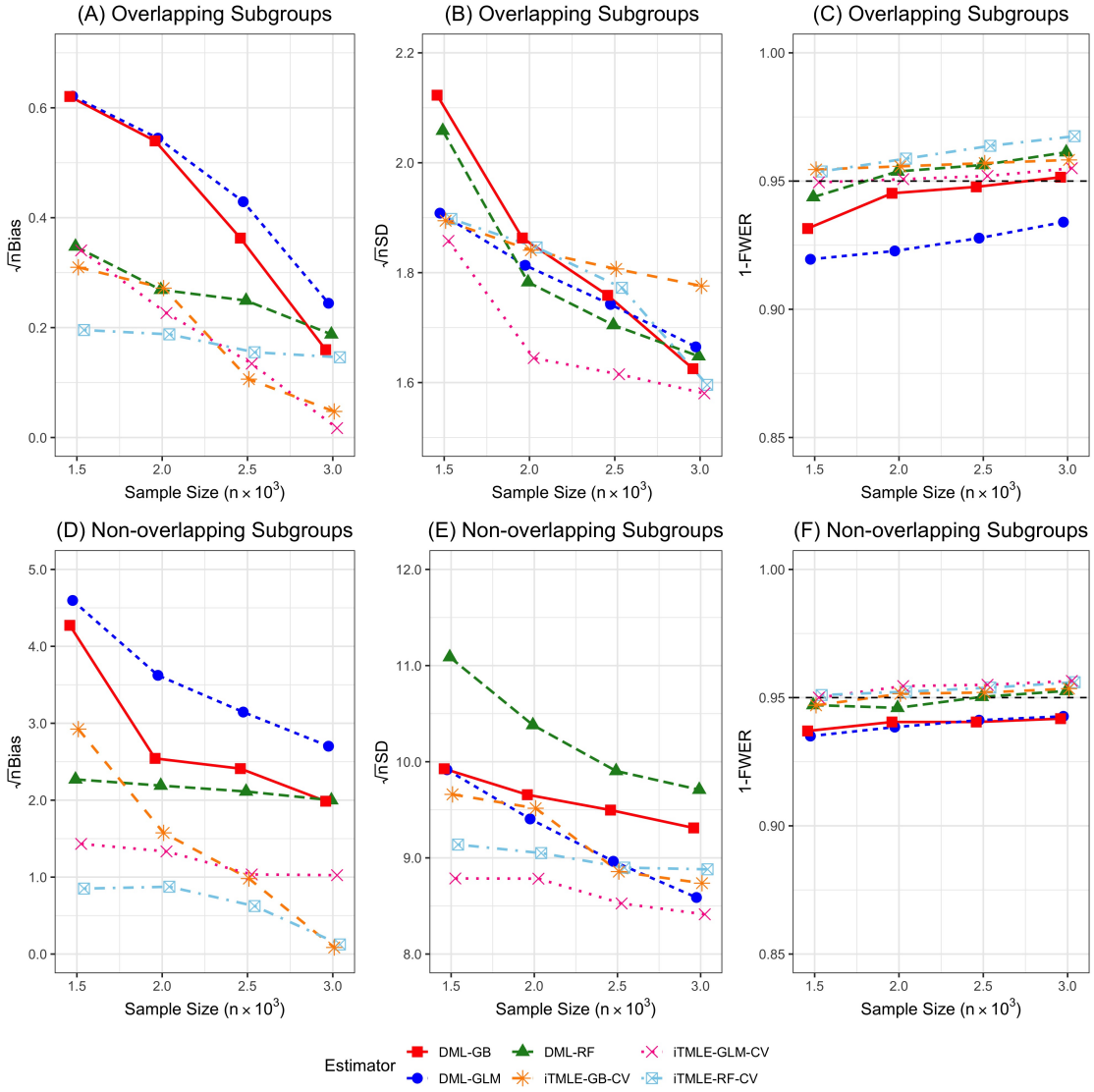


Figure 1.6: Comparison of CV-iTMLE with the double machine learning method. "CV-iTMLE" denotes the discussed method with cross-fitting. "DML" denotes the double machine learning method. The maximum Monte Carlo standard error of (1-FWER) across the four estimators is 0.028. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

The results are summarized in Figure 1.8 and 1.9. From these new simulation results, we observe that the conclusions overall do not substantially differ from our previous simulation results. The $\sqrt{n}$-scaled standard deviation of all the estimators increase compared to the

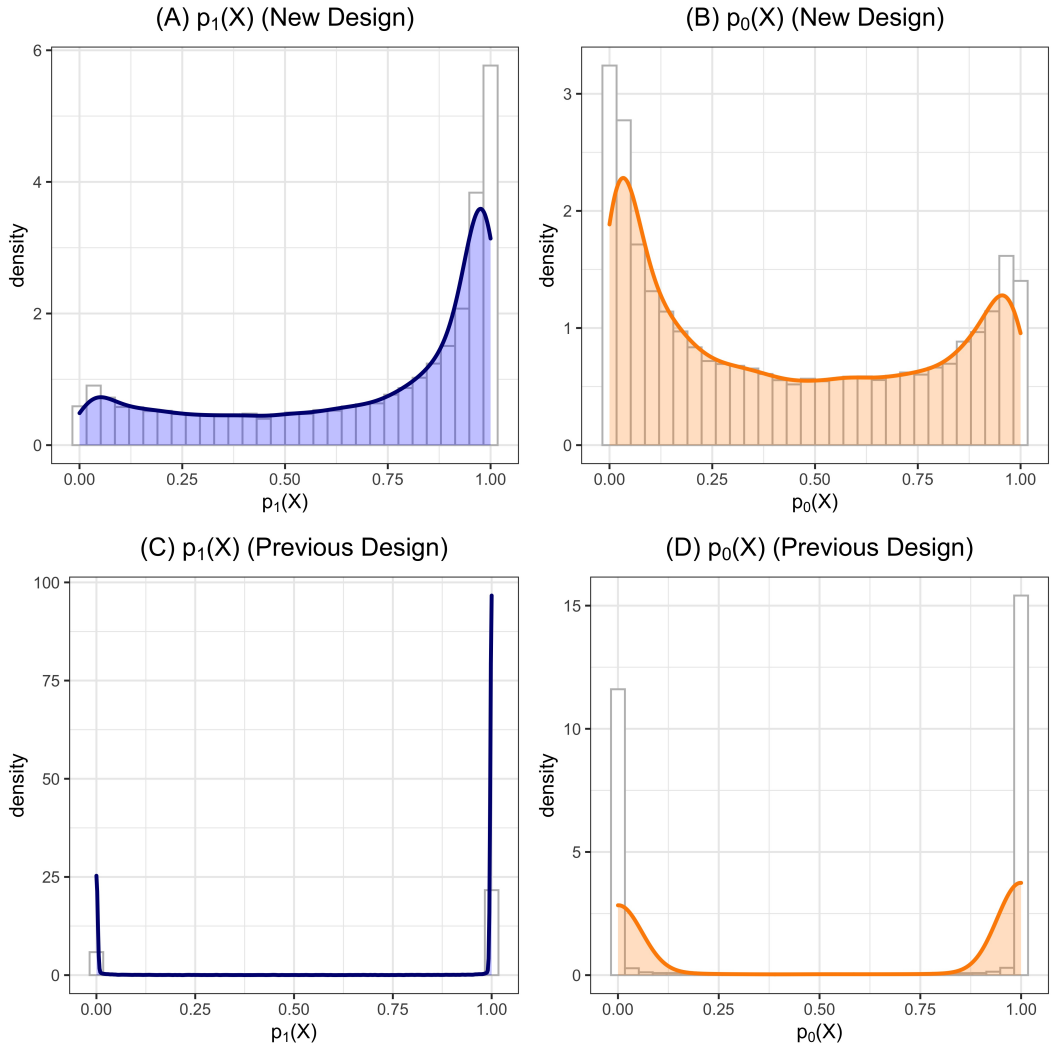


Figure 1.7: Panels (A) and (B) provide the distribution of $p_t(X)$ under the new simulation design, and Panels (C) and (B) provide the distribution of $p_t(X)$ under the original simulation design.

previous simulation design. We conjecture that the increased variances are due to the higher variability of the outcome variables under the new simulation setup.

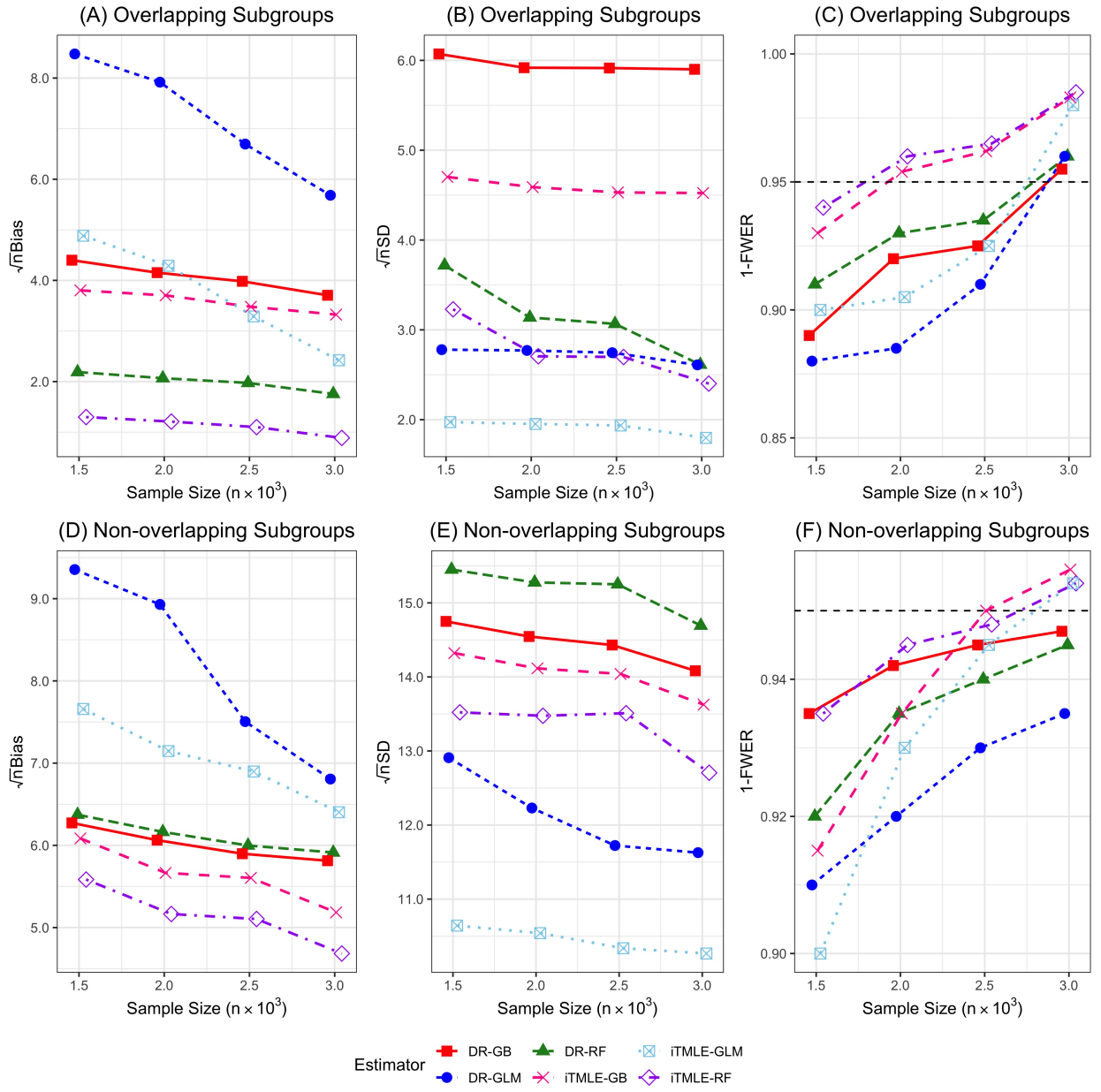


Figure 1.8: Comparison of bias, standard deviation (scaled by root-n), and (1-FWER) in overlapping and non-overlapping subgroups. "iTMLE" denotes the proposed estimator. "DR" denotes the doubly robust estimator. The maximum Monte Carlo standard error of (1-FWER) across the four estimators is 0.030. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

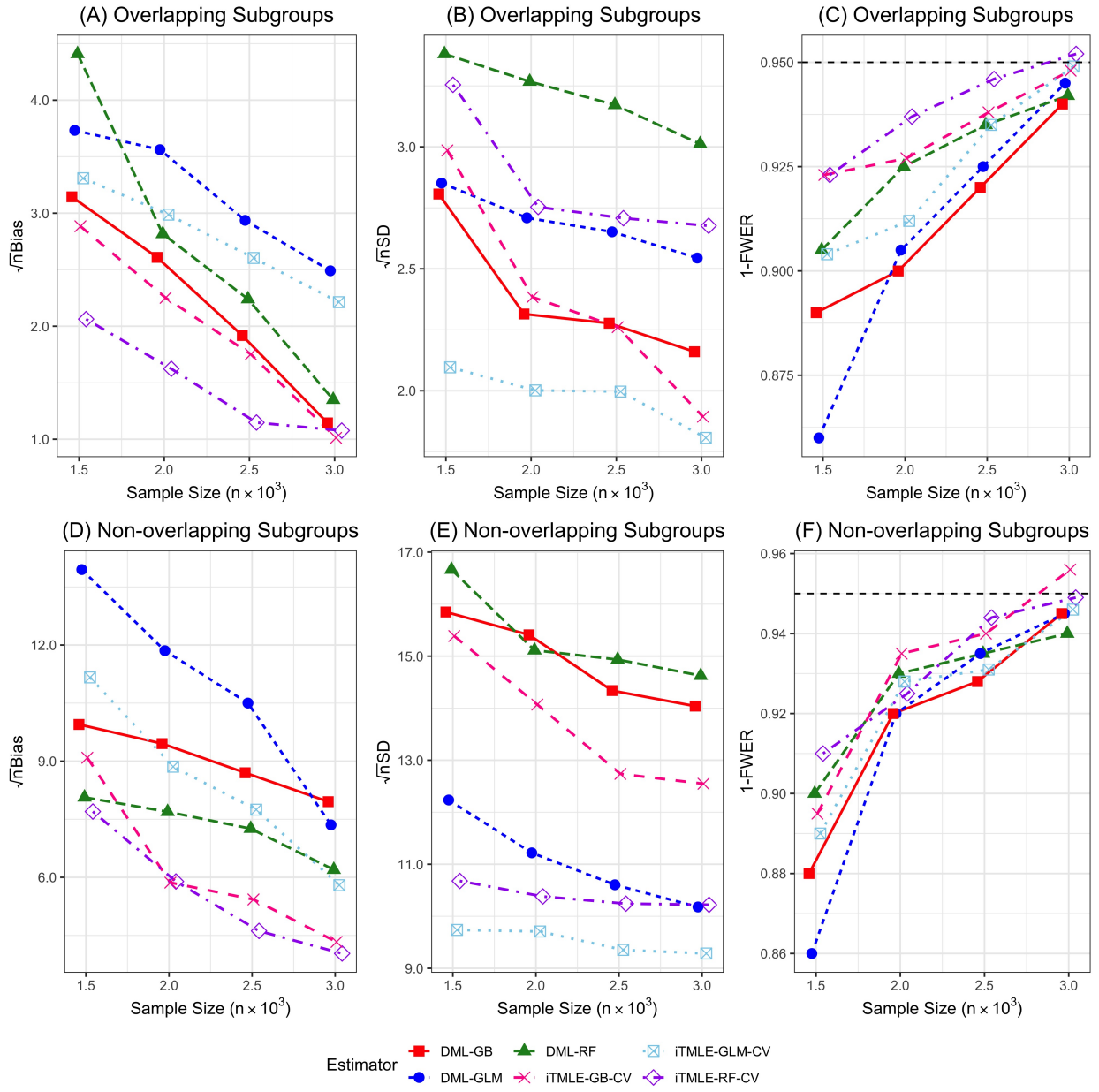


Figure 1.9: Comparison of bias, standard deviation (scaled by root-n), and (1-FWER) in overlapping and non-overlapping subgroups. "iTMLE" denotes the proposed estimator. "DML" denotes the double machine learning method. The maximum Monte Carlo standard error of (1-FWER) across the four estimators is 0.031. "The maximum Monte Carlo standard error of (1-FWER)" refers to the largest standard error of (1-FWER) (out of all three considered estimators for the propensity score and the conditional expectation of the outcome based on logistic regression, random forest, and gradient boosting) computed from Monte Carlo samples.

Details on the Case Studies

UK Biobank data preprocessing details

In UK Biobank study, participants provided lifestyle, medical history, and other health-related information through electronic questionnaires and physical measurements at one of the 22 assessment centers. Blood samples were also collected for genotyping. The UK Biobank study gained approval from the National Health Service's National Research Ethics Service North West (11/NW/0382).

The individuals we investigated are unrelated and had passed standard quality control steps, including removal of outliers for heterozygosity or genotype missing rate, withdrawal of informed consent, and mismatch between reported and inferred sex by genotypes.

We obtain phenotype and genotype data from UK Biobank with the following steps. For phenotype data, first, we download encoded data in *.enc* format from UK Biobank's Access Management System (AMS). To decrypt the encoded data, we download three helper programs: `ukb_md5`, `ukb_unpack`, and `ukb_conv`. Note that the helper programs are only supported by Windows and Linux systems. Second, we verify the integrity of the encoded data via `ukb_md5` and unpack them into *.enc_ukb* format with `ukb_unpack`. To convert the data into readible format, we use `ukb_conv` to convert the *.enc_ukb* data into *.csv* format (other options include txt, r, sas stata or bulk format). The data dictionary can be obtained using `ukb_conv` with *docs* option. After decrypting and converting the encoded data, we obtain a dataset with sample size n = 502, 481 and 20, 502 variables. In our study, since we only work with the phenotypes at the baseline to avoid confounding issues, we extract baseline variables with the suffix "-0.0". The phenotype data we extract include gender, age at recruitment, AD family history, International Statistical Classification of Disease 9th revision (ICD 9) and 10th revision (ICD 10) codes and self-report for T2D and AD.

For genotype data, first, we download imputed genotypes and associated sample information for 23 chromosomes. *Imputation BGEN* and *Imputation sample* can be obtained via `ukb_gene` program. *Imputation BGI* and *Imputation MAF+info* can be downloaded directly from UK Biobank resources 1965 and 1967. Second, we use the *Imputation sample* file to remove individuals without genotype information, which yields sample size $n = 407,057$. Finally, we read in BGEN files with `snp_readBGEN` function in `R` package `bigsnpr` using *HapMap3* as the reference genomes. `snp_readBGEN` converts the BGEN files into an `R` object comprising of two elements: *genotype* and *map*, where *genotype* represents the imputed genotypes in a matrix format and *map* contains the features of SNPs (chromosome, rsid, physical position, major and minor alleles and allele frequency). We only extract the genotype matrix from *genotype* and *rsid* from *map* as our genotype data. We restrict our sample to subjects used for the principle components (PCs) computation, since those individuals are unrelated. From the extracted genotype data, we obtain the treatment variable: rs12916 (on chromosome 5), a functionally equivalent SNP of statins.

Since rs12916 and T2D are associated with some other SNPs due to pleiotropy, we adjust for low-density lipoprotein (LDL) and T2D related SNPs in our study. To find disease-

associated SNPs, we rely on the published GWAS studies from the GWAS catalogue. In our study, we define the disease-associated SNPs as SNPs associated with LDL or T2D with p-values less than 5×10^{-8}. To determine the $p-$values for multiple correlated SNPs in the same locus, we use the linkage disequilibrium clumping procedure with $R^2 < 0.01$. Our filtration criteria yield 385 disease-associated SNPs.

Disease status definition

We identify T2D cases from three sources: doctor diagnosis, and self-reports. If one's self-reported T2D status is missing, we define the self-reported T2D as the following: self-reported diabetes =1 and self-reported gestational diabetes = 0 and self-reported type 1 diabetes = 0. To identify AD cases, we reply on self-reports , family-reports, and ICD codes. We use ICD-10 codes: G309, G301, F002, F000, G308, G300, F009, F001, and ICD-9 codes: 3310.

Case study: T2D as the outcome

Because statin usage may increase the risk of T2D ([155]), as a secondary analysis, we further investigated the effect of rs12916-T allele on T2D under the same considered subgroups to evaluate the potential heterogeneous side effects. We still considered the "high AD genetic risk" subgroup and the "low AD genetic risk" subgroup in this setting because existing studies suggest that insulin resistance links T2D and AD [27]. As some genetic variants are shared between T2D and AD [61], people with high AD genetic risk may be more vulnerable to T2D risk and thus is more sensitive to the side effects of statins use. Therefore, we hypothesized that the effect of carrying rs12916-T allele on T2D risks could be heterogeneous in subgroups with different AD genetic risks and evaluated our hypothesis through subgroup analysis. We compared the performance of the proposed method (CV-iTMLE) with the double machine learning (DML) method and the widely used generalized linear models (GLM). We used the random forest as our first stage estimator as it provides the most robust results in our simulation studies.

Figure 1.10 demonstrated that the treatment effect of carrying rs12916-T allele on T2D risk is heterogeneous across considered subgroups. Both the proposed method and the double machine learning (DML) method suggested that carrying rs12916-T allele increases T2D risk in females and in individuals under 65. For the significant subgroup, the confidence interval of the proposed method was much shorter than that of DML, indicating that the proposed method is more efficient. Furthermore, our results showed that the effect of inheriting rs12916-T allele on T2D risk is heterogeneous in subgroups with different AD genetic risks, which potentially implies that subjects with higher AD genetic risk can be more vulnerable to statin usage. Such findings could be partially explained by the similar pathological pathways shared between T2D and AD [112]. Some of our above findings are in-line with current believes in the existing literature. For example, the results from one randomized clinical trial (JUPITER trial [120]) indicate that statins may increase T2D risk more signifi-

cantly in females than in males. The significant adverse effect of carrying rs12916-T allele (a proxy for statin usage) in individuals with high AD genetic risk is a rather novel finding. We conjecture that this is because insulin resistance links T2D and AD [27], and some genetic variants are also shared between T2D and AD [61]. These findings warrant further clinical studies.

In this secondary analysis, our proposed method again showed shortened confidence intervals and hence improved power in detecting significant subgroups, while the GLM and the double machine learning method tend to lose power. We conjectured that the double machine learning method failed to detect the adverse effect of rs12916-T allele on the high AD genetic risk subpopulation because of the large variance caused by small estimated propensity scores. In contrast, the proposed method is rather robust to the small estimated propensity scores.

Efficient Influence Function and Delta Method

Efficient influence function derivation for subgroup conditional risk

Suppose our parameter of interest is $\alpha = E[Y|T = 1, X \in \mathcal{A}_j]$.

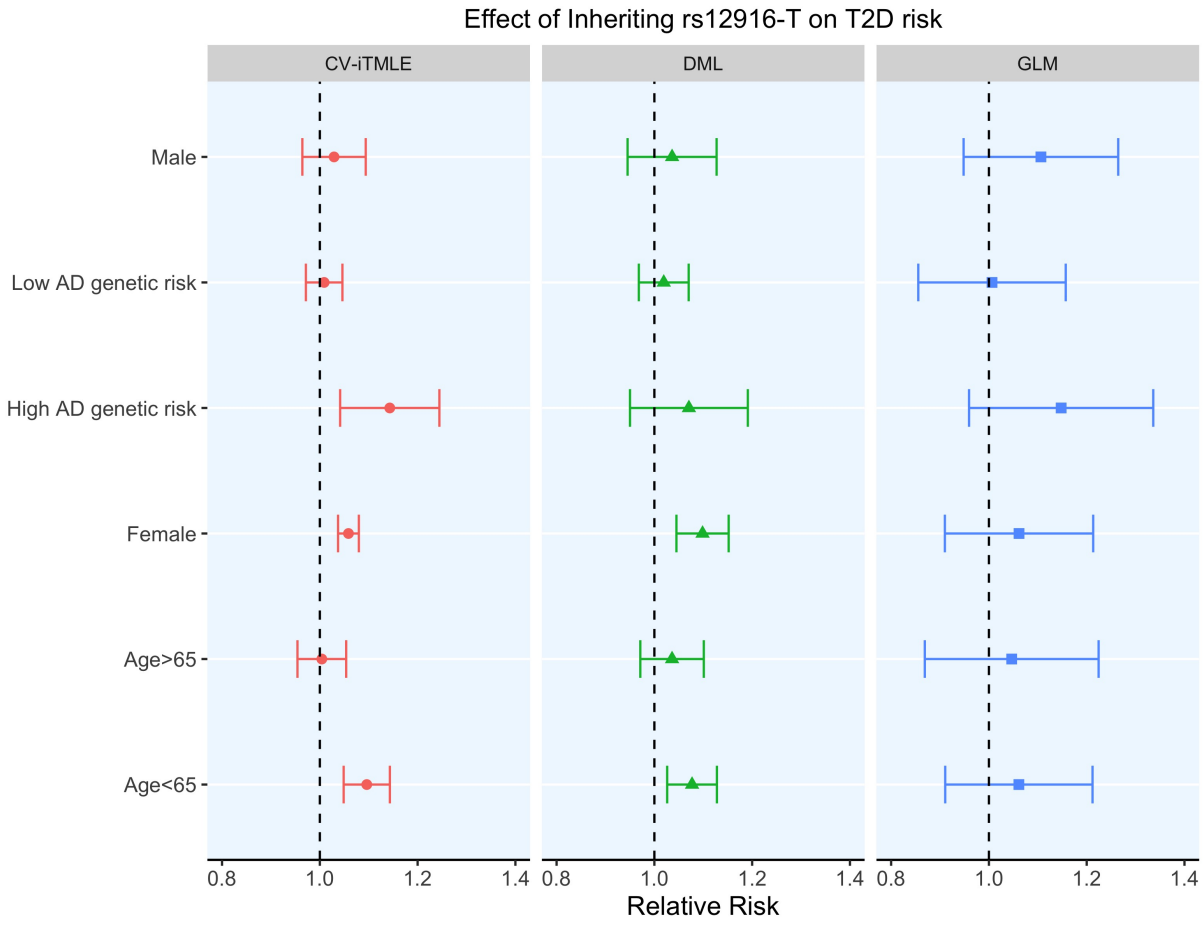


Figure 1.10: The effect of inheriting rs12916-T allele (a proxy for statin usage) on the type 2 diabetes (T2D) risk in the UK Biobank white British population ($n = 293,929$). "DML" denotes the double machine learning method. "GLM" denotes the generalized linear models. GLM is used for association test and does not imply causal relationships. "CV-iTMLE" denotes the cross-validated iTMLE method.

Proof. Given $\alpha(P_0) = \mathbb{E}[Y(1)|X \in \mathcal{A}_j]$ and the ε-perturbed distribution is P_ε^h,

$$\begin{aligned}
\frac{\partial}{\partial\varepsilon}\alpha(P_\varepsilon^h)\big|_{\varepsilon=0} =& \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big\{\int_{y\in\mathbb{R}} t\cdot y\cdot \mathrm{d}\alpha(P_\varepsilon^h) - \alpha(P)\Big\},\\
=& \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big\{\int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} t\cdot y\cdot \mathbb{1}(x\in\mathcal{A}_j)\cdot f(x,y,t;\alpha)(1+\varepsilon s(x,y,t;\alpha))\mathrm{d}x\mathrm{d}y,\\
&\qquad -\int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} t\cdot y\cdot \mathbb{1}(x\in\mathcal{A}_j)\cdot f(x,y,t;\alpha)\mathrm{d}x\mathrm{d}y\Big\},\\
=& \int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} t\cdot y\cdot \mathbb{1}(x\in\mathcal{A}_j)\cdot f(x,y,t;\alpha)s(x,y,t;\alpha)\mathrm{d}x\mathrm{d}y,\\
=& \int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} t\cdot y\cdot \mathbb{1}(x\in\mathcal{A}_j)\cdot\Big(s(y|x\cdot\mathbb{1}(x\in\mathcal{A}_j),t;\alpha)\\
&\cdot f(y|x\cdot\mathbb{1}(x\in\mathcal{A}_j),t;\alpha)\cdot f(t|x\cdot\mathbb{1}(x\in\mathcal{A}_j);\alpha)\cdot f(x\cdot\mathbb{1}(x\in\mathcal{A});\alpha\Big)\mathrm{d}x\mathrm{d}y,\\
=& \int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} t\cdot y\cdot \mathbb{1}(x\in\mathcal{A}_j)\cdot\Big(s(y|x\cdot\mathbb{1}(x\in\mathcal{A}_j),t;\alpha)\\
&\cdot\frac{f(y,x\cdot\mathbb{1}(x\in\mathcal{A}_j),t;\alpha)}{f(t|x\cdot\mathbb{1}(x\in\mathcal{A}_j);\alpha)f(x\cdot\mathbb{1}(x\in\mathcal{A}_j);\alpha)}\Big)dxdy,\\
=& \int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}} t\cdot y\cdot \mathbb{1}(x\in\mathcal{A}_j)\\
&\cdot\Big(s(y|x\cdot\mathbb{1}(x\in\mathcal{A}_j),t;\alpha)\cdot\frac{f(y,x\cdot\mathbb{1}(x\in\mathcal{A}_j),t;\alpha)}{e(x;\alpha)\cdot f(x\cdot\mathbb{1}(x\in\mathcal{A}_j);\alpha)}dxdy,
\end{aligned}$$

given $\int_{x\in\mathbb{R}} f(x\cdot\mathbb{1}(x\in\mathcal{A}_j;\alpha)dx = P(\mathcal{A}_j;\alpha)$,

$$\begin{aligned}
=&\frac{\mathbb{1}(x\in\mathcal{A}_j)}{P(\mathcal{A}_j;\alpha)}\int_{y\in\mathbb{R}}\int_{x\in\mathbb{R}}\frac{t\cdot y}{e(x;\alpha)}\Big(s(y|x\cdot\mathbb{1}(x\in\mathcal{A}_j);\alpha)\cdot f(y,x\cdot\mathbb{1}(x\in\mathcal{A}_j;\alpha)dxdy,\\
=&E\Big[\frac{\mathbb{1}(x\in\mathcal{A}_j)}{P(\mathcal{A}_j;\alpha)}\frac{T}{e(X;\alpha)}Y\cdot s(x,y;\alpha)\Big].
\end{aligned}$$

Next, we follow the standard technique and obtain the efficient influence function as

$$\begin{aligned}
\varphi_{1,j}(Y,D,X) &= \frac{\mathbb{1}(X\in\mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\Big(\frac{T}{e(X)}\big(Y-p_1(X)\big)+p_1(x)-\alpha_{1,j}\Big),\\
\varphi_{0,j}(Y,D,X) &= \frac{\mathbb{1}(X\in\mathcal{A}_j)}{\mathbb{P}(\mathcal{A}_j)}\Big(\frac{1-T}{1-e(X)}\big(Y-p_0(X)\big)+p_0(X)-\alpha_{0,j}\Big).
\end{aligned}$$

□

Delta method

The efficient influence functions of ATE, the relative risk and the odds ratio can be derived by applying the delta method on the efficient influence functions of $(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})$.

$$
\begin{aligned}
\boldsymbol{\varphi_{\mathrm{ATE}}} &= (\boldsymbol{\varphi_1}, \boldsymbol{\varphi_0}) \cdot \frac{\partial f_{\mathrm{ATE}}(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})}{\partial (\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})} \\
&= \begin{pmatrix} \varphi_{1,1} & \varphi_{0,1} \\ \vdots & \vdots \\ \varphi_{1,J} & \varphi_{0,J} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \varphi_{1,1} - \varphi_{0,1} \\ \vdots \\ \varphi_{1,J} - \varphi_{0,J} \end{pmatrix}, \\
\boldsymbol{\varphi_{\mathrm{OR}}} &= (\boldsymbol{\varphi_1}, \boldsymbol{\varphi_0}) \cdot \frac{\partial f_{\mathrm{OR}}(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})}{\partial (\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})} \\
&= \begin{pmatrix} \varphi_{1,1} & \varphi_{0,1} \\ \vdots & \vdots \\ \varphi_{1,J} & \varphi_{0,J} \end{pmatrix} \begin{pmatrix} \frac{1-\alpha_0}{\alpha_0(1-\alpha_1)^2} \\ -\frac{\alpha_1}{\alpha_0^2(1-\alpha_1)} \end{pmatrix} = \begin{pmatrix} \frac{1-\alpha_0}{\alpha_0(1-\alpha_1)^2} \cdot \varphi_{1,1} - \frac{\alpha_1}{\alpha_0^2(1-\alpha_1)} \cdot \varphi_{0,1} \\ \vdots \\ \frac{1-\alpha_0}{\alpha_0(1-\alpha_1)^2} \cdot \varphi_{1,J} - \frac{\alpha_1}{\alpha_0^2(1-\alpha_1)} \varphi_{0,J} \end{pmatrix}, \\
\boldsymbol{\varphi_{\mathrm{RR}}} &= (\boldsymbol{\varphi_1}, \boldsymbol{\varphi_0}) \cdot \frac{\partial f_{\mathrm{RR}}(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})}{\partial (\boldsymbol{\alpha_1}, \boldsymbol{\alpha_0})} \\
&= \begin{pmatrix} \varphi_{1,1} & \varphi_{0,1} \\ \vdots & \vdots \\ \varphi_{1,J} & \varphi_{0,J} \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha_0} \\ -\frac{\alpha_1}{\alpha_0^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\alpha_0} \cdot \varphi_{1,1} - \frac{\alpha_1}{\alpha_0^2} \cdot \varphi_{0,1} \\ \vdots \\ \frac{1}{\alpha_0} \cdot \varphi_{1,J} - \frac{\alpha_1}{\alpha_0^2} \varphi_{0,J} \end{pmatrix}.
\end{aligned}
$$

Chapter 2

Inference on the Best Policies with Many Covariates

2.1 Introduction

Motivation and our contribution

Many empirical work requires an understanding of the impact of the most effective policies or treatments on a relevant response variable of interest. For instance, in randomized (factorial) experiments with multiple treatments, researchers may be interested in the most effective policies (combinations). In online platforms, decision makers may be interested in the top five advertising strategies. In financial portfolio management, managers might want to learn about the best-performing strategies among many alternatives. In practice, after different policy effect sizes are estimated from a random sample, researchers may naturally look into those policies with the largest effect sizes. Accurately measuring the performance of top policies allows policy makers to deliver better-informed decisions for forecasting the effects of future policy implementations.

Nevertheless, given the well-recognized "winner's curse" phenomenon, there can be considerable uncertainties concerning if the top policies with large estimated effect sizes are indeed effective in the population (see Section 2.1 for a literature review). In fact, due to the winner's curse phenomenon, literature documents that the estimated effect sizes of the best-performing policies without additional adjustments tend to be overly optimistic, rendering under-covered confidence intervals [108, 8]. In this manuscript, we refer to the optimistic bias introduced by the winner's curse phenomenon as the winner's curse bias. To mitigate this bias issue, we focus on the problem of constructing accurate point estimates and valid confidence intervals for the true effect sizes of the (observed) best policies. By the best policies, we refer to a user-supplied number of policies that have the largest (estimated) effects among a set of candidate policies (see Section 2.2 for a concrete problem setup), as we would expect that in practice researchers might want to focus on a few top policies of interest.

Other than the winner's curse phenomenon discussed above, an additional consideration

gains prominence in the evaluation of the most effective policies. Since policy (or intervention) variables are often not exogenous, researchers may adopt observational methods to estimate their effects. In recent years, given the increased availability of large datasets with rich covariate information, one commonly adopted approach in empirical works is to assume that the policy variables are exogenous after controlling for a sufficiently large set of factors or covariates. Such a consideration demandingly requires empirical researchers to estimate the policy effects in the presence of many covariates.

To simultaneously address the above-mentioned issues, in this article, we propose a procedure that not only is robust to the presence of many covariates, but also provides accurate point estimates and valid frequentist confidence intervals for multiple best policy effect sizes. By many covariates, we allow the number of covariates q_n to diverge with the sample size n as long as $\limsup_{n\to\infty} q_n/n < 1$. Note that this does not rule out the cases where q_n is fixed or $q_n = o(n)$. In other words, our inferential method remains valid when the dimension of the covariates q_n is fixed or $q_n = o(n)$. Our proposed confidence intervals are built upon resampling methods, and we demonstrate that they achieve exact nominal coverage as the sample size goes to infinity under fairly moderate assumptions. Our empirical evidence shows that conventional estimates ignoring the winner's curse issue are substantially upward biased, while our corrections reduce the winner's curse bias and increase coverage. As far as we know, valid statistical inferential tools on multiple best policies that lift the winner's curse while incorporating possibly many covariates have been lacking, and the contribution of our work is to bridge this gap and help policy makers deliver well-informed decisions in practice.

We illustrate our method with two empirical applications. In the first case study, we use the charitable giving data from [92] to evaluate the best pricing policies that motivate donors to give. Our results suggest that simple methods without adjusting for the winner's curse bias could be potentially overly optimistic in identifying the most effective polices. After accounting for the winner's curse bias, we do not find sufficient evidence to support that the second best pricing policy–asking the donor to give 25% more than his/her highest historical donation–is effective, implying that asking for a more "expensive" donation may not encourage donors to give. We nevertheless note that given our calibration only marginally reduces the effect size of the second best policy, the above conclusion might not warrant a different economic interpretation. In the second case study, we evaluate the effectiveness of the national supported work (NSW) program in different groups of workers. The NSW program is a job training program designed to prepare disadvantaged workers for employment, and it has been investigated in various studies [39, 107]. We apply the proposed approach to evaluate the performance of the NSW program on the most-affected subgroups of workers observed in the dataset. Our study results potentially suggest that married black workers might benefit from the NSW program with an average increase of $4,410 for their annual income.

Connection to the existing literature

One fundamental trend that drives the motivation of the methodology developed in this manuscript is the increasing availability of massive datasets and the associated increasing dimensionality. Such a trend brings scientists opportunities to deliver better-informed policies but, at the same time, presents challenges in developing econometric and statistical tools; see [54], [19], [52], [25] for example. A recent book [58] provides a thorough discussion of analytical methods that aim to address such challenges. Specifically, the increasing data availability brings challenges and also opportunities to better understand various policies whose effects can be inferred from data. Along this line, our manuscript aims at providing understating for policies that are estimated and selected to be the most effective from a pool of policies.

The winner's curse phenomenon and its related issues have been widely recognized in economics, statistics, and data science at large. Seminal works by [56, 55] point out that spurious discoveries can easily arise when target parameters are selected through data mining and statistical machine learning algorithms. Recent work by [8] considers performing conditional and unconditional inference on observed best policy and [9] extends the work to more general ranking problems, which is still different from our goal in conducting unconditional inference on multiple top policies. Moreover, while the conditional approaches in [8] and [9] produce optimal confidence intervals for the observed policy effects, their point estimates and confidence intervals can be conservative when they are applied unconditionally. [45] considers a method to handle the winner's curse bias with Tweedie's formula concerning the empirical Bayes theory. [108] consider a plug-in correction of the winner's curse bias and propose to construct confidence interval based on bootstrapping in the context of A/B testing, but the proposed method lacks theoretical justifications. In clinical trials for evaluating the largest observed treatment effect in multiple subpopulations, [64] propose a bootstrap-based confidence interval that achieves the exact nominal level as the sample size goes to infinity, though generalizing their method to make inference on several top policies might not be straightforward, especially in the presence of many covariates.

Our manuscript builds upon the literature on linear regression models with many or high dimensional covariates; see [84], [117], [115], [5], [50], [24], [23], [89] and the reference therein. In particular, [117] has established the asymptotic normality results for any contrasts of the ordinary least squares (OLS) coefficient vector estimator, when the dimension of the covariates divided by the sample size vanishes asymptotically. More recently, [25] have shown that a small subset of the OLS estimators for the regression coefficients are asymptotically normal without restricting the dimension of the covariates to be a vanishing fraction of the sample size. Moreover, [25] have proposed a robust covariance matrix estimator for the subset of the the OLS estimator under fairly general conditions. [89] has proposed an alternative covariance matrix estimator that can deal with designs with even large number of covariates under additional assumptions (Assumption 9 in the current manuscript).

Making inference on the best-performing policies is related to the literature on constructing confidence intervals for extrema parameters with bootstrap; see [7], [51], [178], [29], [35]

and the reference therein. Given the asymptotic distributions of extrema parameter estimators are often not normal, bootstrap-based methods can face serious difficulties when used to replicate the distribution of extrema of parameter estimators [117, 118]. While subsampling could overcome this issue faced by the classical bootstrap, it can exhibit very poor finite-sample performance because of the noise introduced by the vanishing subsample size. Different from our goal in constructing confidence intervals that achieve the exact nominal level, [72] and [29] propose to construct conservative bootstrap confidence intervals for extrema of parameters. In our current problem setup with many covariates, the problem becomes even more acute as [49] show through a mix of simulation and theoretical analyses that the bootstrap is fraught with problems in moderate high dimensions. In the context of meta-analyses, [35] propose an approach to make inference on ordered fixed study-specific parameters when different parameters are estimated independently from multiple studies.

Our method also contributes to the rapidly growing literature on program evaluations; see [59], [19], [97], [13], [2], [32], [57], [163] among many others. Under our asymptotic regime where the number of covariates q_n grows with the sample size n, the Neyman orthogonalization based approaches often need to work with models with sparse regression coefficients [19]. Rather than imposing such a sparsity assumption, our approach estimates the policy effects with regression adjustments without requiring the regression coefficients to be sparse. Because our approach only requires a consistent covariance matrix estimation for different policy effect estimators, we expect that the proposed framework on evaluating the best policies can be generalized when different policy effects are estimated with other off-shelf methods and we relegate such extensions for future work.

Notation. We work with triangular array data $\{\omega_{i,n} : i = 1, \ldots, n; n = 1, 2, \ldots\}$ where for each n, $\{\omega_{i,n} : i = 1, \ldots, n\}$ is defined on the probability space $(\Omega, \mathcal{S}, P_n)$. All parameters that characterize the distribution of $\{\omega_{i,n} : i = 1, \ldots, n\}$ are implicitly indexed by P_n and thus by n. We write vectors and matrices in bold font, and use regular font for univariate variables and constants.

2.2 Model setup and methodology

Problem setup and a revisit to the winner's curse phenomenon

Suppose we have a random sample $\{(y_{i,n}, \boldsymbol{x}'_{i,n}, \boldsymbol{w}'_{i,n})'\}_{i=1}^n$, we pose the problem in the framework of a linear regression model under heteroscedasticity

$$y_{i,n} = \boldsymbol{x}'_{i,n}\boldsymbol{\beta} + \boldsymbol{w}'_{i,n}\boldsymbol{\gamma}_n + u_{i,n}, \quad i = 1, \ldots, n, \tag{2.1}$$

where $y_{i,n}$ is the outcome variable, $\boldsymbol{x}_{i,n} \in \mathbb{R}^d$ are the treatment or policy variables of interest, $\boldsymbol{w}_{i,n} \in \mathbb{R}^{q_n}$ contains the confounding factors, $u_{i,n}$ is an unobserved error term, and the coefficient vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)'$ contains the treatment effect of $\boldsymbol{x}_{i,n}$ on the outcome $y_{i,n}$. We allow the linear model (2.1) to hold approximately by allowing $\mathbb{E}[u_{i,n}|\{\boldsymbol{x}_{i,n}\}_{i=1}^n, \{\boldsymbol{w}_{i,n}\}_{i=1}^n] \neq 0$.

We are also in a scenario where $\boldsymbol{w}_{i,n}$ is high-dimensional, in the sense that q_n can be a vanishing fraction of the sample size n as long as $\limsup_{n\to\infty} q_n/n < 1$. To simplify notations, we drop subscript n in univariate random variables in the rest of the manuscript. That is, for example, we denote $\gamma_{j,n}$ as γ_j.

We write the ordered values of $\beta_1, \ldots, \beta_d$ as $\beta_{(1)} \geq \ldots \geq \beta_{(d)}$. We adopt the ordinary least-squares (OLS) estimator $\hat{\boldsymbol{\beta}}$ (see Remark 4 for other possible estimates) to estimate $\boldsymbol{\beta}$ and write the order statistics of $\hat{\boldsymbol{\beta}}$ as $\hat{\beta}_{(1)} \geq \ldots \geq \hat{\beta}_{(d)}$. Because researchers in practice might hope to focus on a few top policies, given that d_0 is a user-supplied positive integer, our goal is to construct accurate point estimates and valid confidence intervals for two sets of quantities:

(1) the best policy effect sizes in the population: $\beta_{(1)}, \ldots, \beta_{(d_0)}$,

(2) the observed best policy effect sizes: $\beta_{\hat{j}}$, where $\hat{j} = \sum_{k=1}^{d} k \cdot \mathbb{1}(\hat{\beta}_k = \hat{\beta}_{(j)})$, for $j = 1, \ldots, d_0$.

The first set of quantities characterizes the effects of the top d_0 policies in the population and are thus fixed parameters. The second set of quantities describes the true effect sizes of the best performing policies observed in the random sample, and these quantities are thus "data-dependent parameters." Both sets of quantities can be of interest in different empirical applications [38, 30, 130], and our proposed procedure can be used to deliver valid statistical inference on both quantities (Theorem 2 and Corollary 1).

Remark 3 (Ties in the estimated policy effects)**.** *The second set of parameters is well defined if the observed policies do not have exact ties in the sense that $\hat{\beta}_{(1)} > \ldots > \hat{\beta}_{(d)}$. When the policies effect estimators solve to the interior points of the feasible parameter space, it is likely that no exact ties appear in the random sample. On the other hand, there can exist scenarios where, for example, d_0 is set as 2 but there are multiple policy effect sizes that tie at rank 2. In this case, one may choose instead a data-dependent $\hat{d}_0 = \max\{k : \hat{\beta}_{(k)} > \hat{\beta}_{(2)} - C_1 \cdot n^{-0.25}\}$. This new random d_0 will asymptotically be able to incorporate all the effect sizes that actually are equal to the true effect size associated with $\hat{\beta}_{(2)}$. In this way, the limiting value of $\hat{d}_0$ will not necessarily be 2, but can be a larger number than 2 to incorporate "very close" effect sizes with the rank-2 effect size. We also provide some related discussions in Remark 6.*

Remark 4 (Other possible estimators of $\boldsymbol{\beta}$)**.** *In the presence of many covariates when q_n is potentially large ($\limsup_{n\to\infty} q_n/n \to 1$ in our asymptotic regime) without assuming the coefficient γ_n to be sparse, we adopt the OLS estimator to estimate $\boldsymbol{\beta}$, because the OLS estimator has been thoroughly studied in the existing literature and enjoys favorable theoretical guarantees. In high dimensions when $q_n \gg n$, other estimators of $\boldsymbol{\beta}$ that incorporate model selection procedures can be adopted as well. Our procedure can produce valid statistical inference as long as the covariance matrix of $\hat{\boldsymbol{\beta}}$ can be consistently estimated. For example, under the sparsity assumption on $\boldsymbol{\gamma}_n$ documented in the literature [58], we may adopt the covariance matrix estimator from the de-sparsified Lasso procedure [160, 182].*

To fully realize the challenges on delivering valid statistical inference on these two sets of parameters in our current problem setup, we revisit the winner's curse phenomenon. When first discussed in common-value auctions, the winner's curse refers to the bidding behavior where bidders systematically overbid, resulting in an expected loss [26]. In our context of policy evaluations, the winner's curse refers to the issue that the observed best policies have the tendency to over-estimate the best policies in the population. We would thus often expect that neither $\mathbb{E}[\hat{\beta}_{(j)} - \beta_{(j)}]$ nor $\mathbb{E}[\hat{\beta}_{(j)} - \beta_{\hat{j}}]$ is close to zero, and the resulting confidence interval may fail to reach the nominal level. Such an issue becomes even more acute as we have many covariates $\boldsymbol{w}_{i,n}$ entering the inferential process.

We next illustrate the winner's curse issue through Example 1 with a simple simulation study, where we observe substantial winner's curse bias and under-covered confidence intervals for the top polices. In particular, Figure 1(b) demonstrates that coverage probabilities are worsened when a larger number of covariates are incorporated for estimating $\boldsymbol{\beta}$. It is worth pointing out that when $d = 3$, $\hat{\beta}_{(2)}$ is the median policy effect. Thus, the estimation bias is around 0, and the true standard deviation is much smaller than the estimated standard deviation, resulting in a confidence interval with close to 100% coverage. When d increases, the coverage probability gradually drops due to a larger estimation bias and inaccurately estimated standard deviation.

Example 1 (A simulation study demonstrating the winner's curse phenomenon with many covariates)**.** *We generate 1000 Monte Carlo samples following the setup in Model* (2.1)*. We generate* $\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ *with* $\Sigma_{jk} = 0.5^{|j-k|}$ *for* $j, k = 1, \ldots, d$*,* $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$ *with* $\tilde{\boldsymbol{w}}_{i,n} \sim \mathcal{N}(0, \boldsymbol{I}_{q_n})$*, where* $\boldsymbol{I}_{q_n}$ *is a* q_n*-dimensional identity matrix. We consider the case where no policy is effective (so that* $\boldsymbol{\beta} = 0$*,* $\beta_{\hat{1}} = \beta_{\hat{2}} = 0$*) and* $\gamma_j = 1/j$*, for* $j = 1, \ldots, q_n$*. We report the asymptotic bias of the conventional estimator (i.e.,* $\sqrt{n} \cdot \mathbb{E}[\hat{\beta}_{(j)} - \beta_{(j)}]$*) as well as the coverage probability of confidence intervals constructed based on normal approximation with the Eicker-White [47, 175] covariance matrix estimator defined in Eq* (2.6)*.*

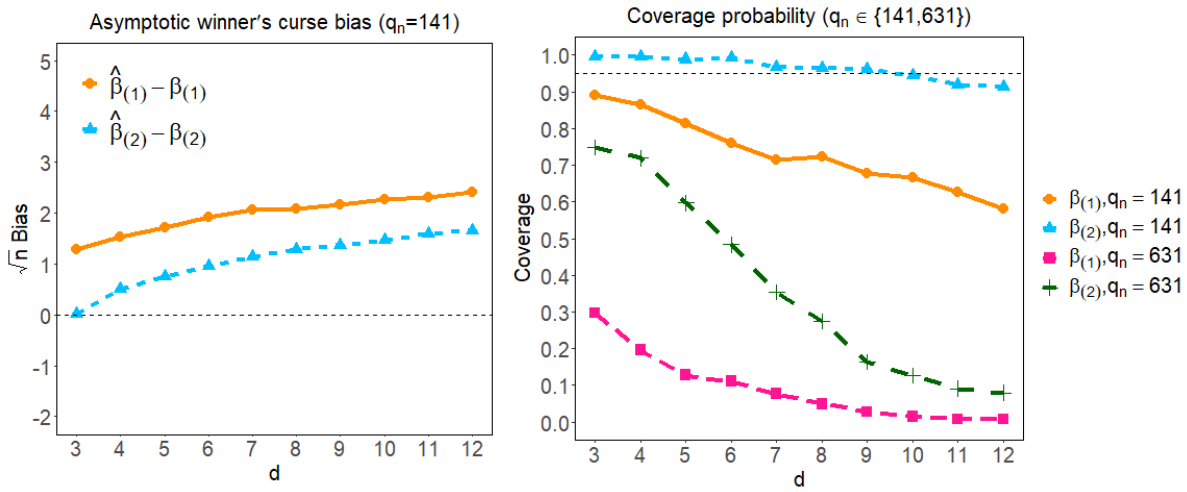
Methodology

Our method starts with the ordinary least-squares (OLS) estimator of $\boldsymbol{\beta}$, that is

$$\hat{\boldsymbol{\beta}} = \Big(\sum_{i=1}^{n} \hat{\boldsymbol{v}}_{i,n} \hat{\boldsymbol{v}}_{i,n}' \Big)^{-1} \Big(\sum_{i=1}^{n} \hat{\boldsymbol{v}}_{i,n} y_{i,n} \Big),$$

where $\hat{\boldsymbol{v}}_{i,n} = \sum_{j=1}^{n} (\boldsymbol{M}_n)_{i,j} \boldsymbol{x}_{j,n}$, and $(\boldsymbol{M}_n)_{i,j} \triangleq \mathbb{1}(i = j) - \boldsymbol{w}_{i,n}' \Big(\sum_{k=1}^{n} \boldsymbol{w}_{k,n} \boldsymbol{w}_{k,n}' \Big)^{-1} \boldsymbol{w}_{j,n}$. As we focus on the case when q_n can be a non-vanishing fraction of n, $n \to \infty$, we adopt the robust covariance matrix estimator proposed in [89]. We try to follow the author's notation as closely as possible:

$$\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}} \triangleq \hat{\boldsymbol{\Gamma}}_n^{-1} \hat{\boldsymbol{\Sigma}}_n^{\mathrm{KJ}} \hat{\boldsymbol{\Gamma}}_n^{-1},$$



(a) Winner's curse bias ($q_n = 141$)

(b) Coverage probability

Figure 2.1: Demonstration of the winner's curse phenomenon following the simulation setup in Example 1. The maximum Monte Carlo standard error for the asymptotic bias is 0.88. Panel (a) captures the asymptotic winner's curse bias when $q_n = 141$; Panel (b) captures the coverage probability when $q_n \in \{141, 631\}$ and the nominal level is 0.95.

where

$$\hat{\mathbf{\Gamma}}_n = \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{v}}_{i,n}\hat{\boldsymbol{v}}_{i,n}', \quad \hat{\Sigma}_n^{\mathrm{KJ}} \triangleq \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{v}}_{i,n}\hat{\boldsymbol{v}}_{i,n}' y_{i,n}\acute{u}_{i,n},$$

where $\acute{u}_{i,n} = \frac{\hat{u}_{i,n}}{(\boldsymbol{M}_n)_{i,i}}$, $\hat{u}_{i,n} = \sum_{j=1}^{n}(\boldsymbol{M}_n)_{i,j}(y_{j,n} - \boldsymbol{x}_{j,n}'\hat{\boldsymbol{\beta}})$, for $i = 1, \ldots, n$. Such an estimator is well-defined as long as $\min_i(\boldsymbol{M}_n)_{i,i} > 0$. If $\min_i(\boldsymbol{M}_n)_{i,i} = 0$, it means that the auxiliary regression produces a perfect prediction. So the observation does not carry information on $\boldsymbol{\beta}$ and can be ignored.

As shown in Example 1, the estimated top policy effect sizes with $\hat{\beta}_{(1)}, \ldots, \hat{\beta}_{(d_0)}$ are often biased upward for our target parameters due to the winner's curse phenomenon. Inspired by the procedure proposed by [35][1] for meta-analyses, we generate replicates of $\hat{\boldsymbol{\beta}}$ from a multivariate normal distribution

$$\hat{\boldsymbol{\beta}}^* \big| \{(y_{i,n}, \boldsymbol{x}_{i,n}', \boldsymbol{w}_{i,n}')'\}_{i=1}^n \sim \mathcal{N}(\hat{\boldsymbol{\beta}}, \hat{\mathbf{\Omega}}_n^{\mathrm{KJ}}/n), \quad \text{where } \hat{\boldsymbol{\beta}}^* = (\hat{\beta}_1^*, \ldots, \hat{\beta}_d^*)', \tag{2.2}$$

and we denote the ordered values of the vector $\hat{\boldsymbol{\beta}}^*$ as $\hat{\beta}_{(1)}^* \geq \ldots \geq \hat{\beta}_{(d)}^*$. Note that the above description of $\hat{\boldsymbol{\beta}}^*$ differs from some previous work on bootstrapping insofar we have

[1]Note that there is a typo in [35] for the definition of the near tie set. Although the near tie $\mathcal{H}_{(j)}$ in their manuscript was originally defined as $\mathcal{H}_{(j)} = \{k : |\beta_k - \beta_{(j)}| = O(n^{-\frac{1}{2}}),\ k = 1, \ldots, d\}$, their proof goes through when the near tie set is defined with $\mathcal{H}_{(j)} = \{k : |\beta_k - \beta_{(j)}| = o(n^{-\frac{1}{2}}),\ k = 1, \ldots, d\}$.

suppressed the role of "multiplier variables," and we have defined $\hat{\boldsymbol{\beta}}^*$ as a sample from $\mathcal{N}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Omega}}_n^{\text{KJ}}/n)$. Different from [35] that requires different estimators to be estimated from independent studies with non-overlapping random samples, our approach relaxes such an requirement and allows $\hat{\beta}_1, \ldots, \hat{\beta}_d$ to be correlated.

Next, given properly chosen b_{L} and b_{R} so that $b_{\text{L}} - b_{\text{R}} = O(n^{-\delta})$ with $\delta \in (0, \frac{1}{2})$ (see Supplementary Materials Section C.1 for their data-adaptive choices, and robustness to different choices of tuning parameters in Supplementary Materials, Section C.2), we estimate a "near tie" set that captures policies that have similar effect sizes to the j-th largest policy:

$$\widehat{\mathcal{H}}_{(j)} = \big\{k : \hat{\beta}^*_{(j)} - b_{\text{L}} \le \hat{\beta}^*_k \le \hat{\beta}^*_{(j)} + b_{\text{R}},\ k = 1, \ldots, d\big\}.$$

We then record the averages of $\hat{\beta}^*_1, \ldots, \hat{\beta}^*_d$ and of $\hat{\beta}_1, \ldots, \hat{\beta}_d$ in the estimated tie set $\widehat{\mathcal{H}}_{(j)}$ as

$$\tilde{\beta}^*_{(j)} = \frac{\sum_{k \in \widehat{\mathcal{H}}_{(j)}} \hat{\beta}^*_k}{|\widehat{\mathcal{H}}_{(j)}|}, \text{ and } \tilde{\beta}_{(j)} = \frac{\sum_{k \in \widehat{\mathcal{H}}_{(j)}} \hat{\beta}_k}{|\widehat{\mathcal{H}}_{(j)}|}, \tag{2.3}$$

where $|\widehat{\mathcal{H}}_{(j)}|$ denotes the cardinality of the set $\widehat{\mathcal{H}}_{(j)}$.

Finally, we apply the above resampling procedure to construct point estimates and confidence intervals for $\beta_{(j)}$ as well as $\beta_{\hat{j}}$ (as defined in Section 2.2 and in Eq (2.5)), $j = 1, \ldots, d_0$. Specifically, for confidence interval construction, we generate B independent samples of $\tilde{\beta}^*_{(j)}$ as in Eq (2.3), and then define $\hat{q}_{(j)}(\alpha/2)$ to be the empirical $\alpha/2$-quantile of the $B \ge 1$ samples (and similarly for $\hat{q}_{(j)}(1-\alpha/2)$), leading to a level-$\alpha$ confidence interval for $\beta_{(j)}$ with

$$\big[\hat{q}_{(j)}(\alpha/2),\ \hat{q}_{(j)}(1-\alpha/2)\big], \quad j = 1, \ldots, d_0.$$

Corollary 1 demonstrates that the above confidence interval also serves as an asymptotically exact level-α prediction interval for $\beta_{\hat{j}}$. For point estimates, we may either use $\tilde{\beta}_{(j)}$ or the averaged resampled statistics $\tilde{\beta}^*_{(j)}$ to estimate $\beta_{(j)}$ and $\beta_{\hat{j}}$.

2.3 Theoretical investigation

Notations and assumptions

Before discussing the theoretical results in detail, we revisit and introduce some notations and assumptions adopted in the manuscript. We denote the sample $\{(y_{i,n}, \boldsymbol{x}'_{i,n}, \boldsymbol{w}'_{i,n})'\}_{i=1}^n$ as $\{\boldsymbol{z}_{i,n}\}_{i=1}^n$. Recall $u_{i,n}$ is the random error in the considered linear model (2.1), we define

$$\varepsilon_{i,n} = u_{i,n} - \mathbb{E}[u_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n], \quad \boldsymbol{v}_{i,n} = \boldsymbol{x}_{i,n} - \mathbb{E}[\boldsymbol{x}_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n], \quad i = 1, \ldots, n. \tag{2.4}$$

Let $e_{i,n} = \mathbb{E}[u_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n]$, we further denote

$$\sigma_{i,n}^2 = \mathbb{E}[\varepsilon_{i,n}^2|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n], \quad \tilde{\boldsymbol{v}}_{i,n} = \sum_{j=1}^n (\boldsymbol{M}_n)_{i,j}\boldsymbol{v}_{j,n},$$

$$\rho_n^1 = \frac{\sum_{i=1}^n \mathbb{E}\left[e_{i,n}^2\right]}{n}, \quad \rho_n^2 = \frac{\sum_{i=1}^n \mathbb{E}\left[\mathbb{E}\left(e_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n\right)^2\right]}{n},$$

$$\boldsymbol{Q}_{i,n} = \mathbb{E}\Big[\boldsymbol{x}_{i,n} - \big(\sum_{j=1}^n \mathbb{E}[\boldsymbol{w}_{j,n}\boldsymbol{w}_{j,n}']\big)^{-1} \sum_{j=1}^n \mathbb{E}[\boldsymbol{w}_{j,n}\boldsymbol{x}_{j,n}']\Big|\{\boldsymbol{w}_{i,n}\}_{i=1}^n\Big],$$

$$\tilde{\boldsymbol{Q}}_{i,n} = \sum_{j=1}^n (\boldsymbol{M}_n)_{i,j}\boldsymbol{Q}_{i,n}.$$

For a policy j, we define the near tie set in the population as:

$$\mathcal{H}_{(j)} = \big\{k : |\beta_k - \beta_{(j)}| = o(n^{-\frac{1}{2}}),\ k = 1, \ldots, d\big\}.$$

Next, let $\hat{\boldsymbol{e}}_j$ denote a d-dimensional (sparse) vector with

$$\hat{\boldsymbol{e}}_j = (\hat{e}_{j,1}, \ldots, \hat{e}_{j,d}), \quad \hat{e}_{jk} = \frac{\mathbb{1}(k \in \widehat{\mathcal{H}}_{(j)})}{|\hat{\mathcal{H}}_{(j)}|}, \quad k = 1, \ldots, d.$$

We will use the notation $\mathbb{P}(\cdot|\{\boldsymbol{z}_{i,n}\}_{i=1}^n)$ to refer to the probability that is conditional on the random variables $\{\boldsymbol{z}_{i,n}\}_{i=1}^n$.

We make following assumptions throughout this section. Note that Assumptions 6-9 listed below largely follow the assumptions in [25] and [89], we list these assumptions along with their interpretations to present a full picture for our readers.

Assumption 6 (Sampling)**.** *The errors $\varepsilon_{i,n}$ are uncorrelated across i conditional on $\{\boldsymbol{x}_{i,n}\}_{i=1}^n$ and $\{\boldsymbol{w}_{i,n}\}_{i=1}^n$. Let $\{N_1, \ldots, N_{G_n}\}$ represents a partition of $\{1, \ldots, n\}$ with $\max_{1\le g\le G_n} |N_g| = O(1)$ such that $\{(\varepsilon_{i,n}, \boldsymbol{v}_{i,n}), i \in N_g\}$ (defined in (2.4)) are independent across g conditional on $\{\boldsymbol{w}_{i,n}\}_{i=1}^n$.*

Assumption 6 generalizes the classical independent and identically distributed (i.i.d.) setting to allow for repeated measurements or group structures in the observed data. For example, Assumption 6 allows the observed data to form clusters of finite sample sizes, and within-cluster dependency is allowed as long as the observations between clusters are independent.

Assumption 7 (Design)**.** *The dimension of the covariates $\boldsymbol{w}_{i,n}$ satisfies that*

$$\limsup_{n\to\infty} q_n/n < 1.$$

The minimum eigenvalue of the matrix $\sum_{i=1}^n \boldsymbol{w}_{i,n}\boldsymbol{w}_{i,n}'$ *is bounded away from 0 with probability approaching one, that is*

$$\lim_{n\to\infty} \mathbb{P}\Big(\lambda_{\min}\big(\sum_{i=1}^n \boldsymbol{w}_{i,n}\boldsymbol{w}_{i,n}'\big) > 0\Big) = 1.$$

Lastly,

$$\max_{1\le i\le n}\Big\{\mathbb{E}[\varepsilon_{i,n}^4|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n],\ \frac{1}{\sigma_{i,n}^2},\ \mathbb{E}[\boldsymbol{v}_{i,n}^4|\{\boldsymbol{w}_{i,n}\}_{i=1}^n],\ 1/\lambda_{\min}\Big(\frac{\sum_{i=1}^n \mathbb{E}\big[\tilde{\boldsymbol{v}}_{i,n}\tilde{\boldsymbol{v}}_{i,n}'|\{\boldsymbol{w}_{i,n}\}_{i=1}^n\big]}{n}\Big)\Big\} = O_p(1).$$

Assumption 7 contains three conditions. The first condition allows the dimension of the covariates $\boldsymbol{w}_{i,n}$ to grow at the sample rate as the sample size n. The second condition requires the matrix $\sum_{i=1}^n \boldsymbol{w}_{i,n}\boldsymbol{w}_{i,n}'$ to be full rank, which is necessary otherwise the OLS estimator would not be able to calculate the matrix $\boldsymbol{M}_n$. Furthermore, as noted in [25], such an assumption can be imposed by dropping any covariates in $\boldsymbol{w}_{i,n}$ that are collinear. The third condition contains conventional moment conditions for the covariates and heteroscedasticity.

Assumption 8 (Linear model approximation)**.** $\sum_{i=1}^n \mathbb{E}[||\boldsymbol{Q}_{i,n}||^2]/n = O(1)$, $\rho_n^1 + n(\rho_n^1 - \rho_n^2) + \rho_n^1 \cdot \sum_{i=1}^n \mathbb{E}[||\boldsymbol{Q}_{i,n}||^2] = o(1)$, *and* $\max_{1\le i\le n}||\hat{\boldsymbol{v}}_{i,n}||/\sqrt{n} = o_p(1)$, $n\rho_n^1 = O(1)$.

Assumption 8 mainly characterizes the difference between the mean squares of the conditional errors ρ_n^1 and the projection ρ_n^2 into the covariate space $\{\boldsymbol{w}_{i,n}\}$'s. The characterization of this difference involves $\sum_{i=1}^n \mathbb{E}[||\boldsymbol{Q}_{i,n}||^2]$ where $\boldsymbol{Q}_{i,n}$ describes the deviation of $\boldsymbol{x}_{i,n}$ from its population linear projection. Residuals of this linear projection, represented by $\hat{\boldsymbol{v}}_{i,n}$'s, are assumed to satisfy a negligibility condition after a maximization over all i's. This negligibility condition regularizes the distributional connection between $\boldsymbol{x}_{i,n}$'s and $\boldsymbol{w}_{i,n}$'s. We note that if the mean squares of $\boldsymbol{x}_{i,n}$'s are bounded and that an exogeneity condition $e_{i,n} = 0$ holds for all i and n, then the linear model approximation assumption naturally holds. Otherwise, if the exogeneity condition does not hold, Assumption 8 requires a small-bias condition $n\rho_n^1 = O(1)$.

Assumption 9 (Variance estimation)**.** $\lim_{n\to\infty} \mathbb{P}(\min_i \big(\boldsymbol{M}_n\big)_{i,i} > 0) = 1$,

$$\mathbb{P}\Big(\min_i \big(\boldsymbol{M}_n\big)_{i,i} > 0\Big) = O_p(1), \quad \frac{\sum_{i=1}^n ||\tilde{\boldsymbol{Q}}_{i,n}||^4}{n} = O_p(1),$$

and $\max_i ||\mu_{i,n}||/\sqrt{n} = o_p(1)$ *with* $\mu_{i,n} = \mathbb{E}\big[y_{i,n}|\{\boldsymbol{x}_{i,n}\}_{i=1}^n, \{\boldsymbol{w}_{i,n}\}_{i=1}^n\big]$.

Assumption 9 has two major parts. The first part regularizes the diagonal elements $\big(\boldsymbol{M}_n\big)_{i,i}$'s, essentially requiring the smallest diagonal element to be consistently bounded away from zero when n tends to infinity. Even though it is difficult to provide broadly

general primitives to validate this assumption, Assumption 2 of [25], Assumption 4 of [89], and the discussions therein provide sufficient conditions for this assumption to hold. The second part regularizes $\mu_{i,n}$'s and $\tilde{\mathbf{Q}}_{i,n}$'s in order to control the variance of $y_{i,n}$'s and the variance of $\mathbb{E}(\boldsymbol{v}_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n)$'s.

Assumption 10 (Policy effect sizes)**.** *For $\delta \in (0, \frac{1}{2})$, the asymptotic distance between the effects of policy $k \notin \mathcal{H}_{(j)}$ and $j \in \mathcal{H}_{(j)}$ diverges as $n \to \infty$:*

$$n^{\delta} \cdot \min_{k \notin \mathcal{H}_{(j)}} \left|\beta_{(j)} - \beta_k\right| \to \infty, \text{ as } n \to \infty, \quad j = 1, \ldots, d.$$

Assumption 10 requires that any policies outside the near tie set $\mathcal{H}_{(j)}$ have effect sizes sufficiently different from the ones in $\mathcal{H}_{(j)}$. In fixed dimensions when q_n does not grow with n, the underlying policy effect sizes $\beta_1, \ldots, \beta_d$ are constant with respect to the sample size n. The near tie set reduces to a "precise" tie set $\mathcal{H}_{(j)} = \{k : \beta_k = \beta_{(j)},\ k = 1, \ldots, d\}$, suggesting that $\min_{k \notin \mathcal{H}_{(j)}} \left|\beta_{(j)} - \beta_k\right|$ is a positive constant bounded away from zero. In such a case, Assumption 10 is automatically satisfied.

Properties of the proposed estimator

For the proposed estimator, we show that the following theorem holds:

Theorem 2. *Under Assumptions 6-10, for any $t \in \mathbb{R}$, for the resampled statistics, the following holds*

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\sqrt{n}\big(\tilde{\beta}^*_{(j)} - \tilde{\beta}_{(j)}\big)}{(\hat{\boldsymbol{e}}_j' \hat{\boldsymbol{\Omega}}_n^{KJ} \hat{\boldsymbol{e}}_j)^{\frac{1}{2}}} \leq t \Big| \{(y_{i,n}, \boldsymbol{x}_{i,n}', \boldsymbol{w}_{i,n}')'\}_{i=1}^n \right) = \Phi(t).$$

For the original statistics, it holds that

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\sqrt{n}\big(\tilde{\beta}_{(j)} - \beta_{(j)}\big)}{(\hat{\boldsymbol{e}}_j' \hat{\boldsymbol{\Omega}}_n^{KJ} \hat{\boldsymbol{e}}_j)^{\frac{1}{2}}} \leq t \right) = \Phi(t).$$

Furthermore, we have that $\lim_{n \to \infty} \mathbb{P}\Big(\mathbb{P}\big(\tilde{\beta}^*_{(j)} \leq \beta_{(j)} | \{\boldsymbol{z}_{i,n}\}_{i=1}^n\big) \leq s\Big) = s.$

Theorem 2 confirms that our proposed confidence interval for $\beta_{(j)}$ achieves exact $1 - \alpha$ coverage probability as the sample size goes to infinity when B is sufficiently large, which distinguishes the proposed inference procedure from simultaneous methods. Furthermore, Theorem 2 says that $\tilde{\beta}_{(j)}$ is a root-n consistent estimator of $\beta_{(j)}$, in the sense that $\forall \varepsilon > 0$, there exists $M > 0$ such that $\mathbb{P}\big(|\sqrt{n}(\tilde{\beta}_{(j)} - \beta_{(j)})| > M\big) \leq \varepsilon$, for $n \geq 1$.

As for the observed best policies, recall that we denote the observed j-th largest policy as

$$\hat{j} = \sum_{k=1}^{d} k \cdot \mathbb{1}(\hat{\beta}_k = \hat{\beta}_{(j)}). \tag{2.5}$$

The following corollary suggests that the proposed confidence interval for $\beta_{(j)}$ can also serve as an exact prediction interval for $\beta_{\hat{j}}$. Therefore, the proposed procedure in Section 2.2 can also be used to make inference on the observed top policies in a random sample:

Corollary 1. *Under Assumptions 6-10, we have that* $\lim_{n\to\infty}\mathbb{P}\Big(\mathbb{P}\big(\tilde{\beta}^*_{(j)} \leq \beta_{\hat{j}}|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\big) \leq s\Big) = s$. *Furthermore,* $\tilde{\beta}_{(j)}$ *is a "root-n consistent" estimator of the data-dependent parameter* $\beta_{\hat{j}}$ *in the sense that* $\forall\varepsilon > 0$, *there exists* $M > 0$ *such that* $\mathbb{P}\big(|\sqrt{n}(\tilde{\beta}_{(j)} - \beta_{\hat{j}})| > M\big) \leq \varepsilon$, *for* $n \geq 1$.

Remark 5 (Regression models with fixed effects)**.** *The proposed resampling-based approach can be used to calibrate multiple best policies when fixed effects are introduced in linear regression models (see [164] for comprehensive discussion). This suggests that our approach not only applies to independently sampled data, but also remains valid when there are repeated-measurements present in the data. These may include short panel data, and datasets in which, for example, two individuals have sampled from each household. To conserve space in the main manuscript, we have leave the detailed discussion in the Supplementary Materials (Section D).*

Remark 6 (Data dependent choice of d_0)**.** *In addition to a deterministic choice of* d_0, *another practically relevant scenario is a data dependent choice of* d_0. *An example of such a data dependent choice is* $\hat{d}_0 = \max\{k : \hat{\beta}_{(k)} > C\}$, *where* C *is a user-specified threshold for the effect size. A relatively complicated situation is that* C *coincides with some of the policy sizes in* $\beta_1, \beta_2, \cdots, \beta_d$. *In this situation, it is possible that no matter how large* n *is,* $\hat{d}_0$ *does not converge to a deterministic value but instead to a non-degenerate random variable. For the purpose of separation, we may adjust* $\hat{d}_0 = \max\{k : \hat{\beta}_{(k)} > C\}$ *to be* $\hat{d}'_0 = \max\{k : \hat{\beta}_{(k)} > C + C_1 \cdot n^{-0.25}\}$, *where* C_1 *is a constant that does not depend on* n. *The choice of* -0.25 *is tunable and may be of independent interest. By this new choice of* $\hat{d}'_0$, *the policy effects that exactly equal* C *will be eliminated almost surely when* n *tends to infinity. This elimination exactly matches the target to select all the policy sizes that are larger than* C. *In the limit of* n *tending to infinity,* $\max\{k : \hat{\beta}_{(k)} > C + C_1 \cdot n^{-0.25}\}$ *will converge almost surely to a set that contains all effect sizes larger than* C. *Therefore, the large-sample theory results for a pre-specified deterministic integer would still hold by plugging in* $\hat{d}'_0$.

2.4 Simulation studies

Simulation design

We generate i.i.d. Monte Carlo samples of $\{(y_{i,n}, \boldsymbol{x}'_{i,n}, \boldsymbol{w}'_{i,n})\}_{i=1}^n$ from the model

$$y_{i,n} = \boldsymbol{x}'_{i,n}\boldsymbol{\beta} + \boldsymbol{w}'_{i,n}\gamma_n + \varepsilon_{i,n}, \quad i = 1, \ldots, n.$$

We consider various data generating processes (DGP) for different choices of the policy variable $\boldsymbol{x}_{i,n}$, the covariates $\boldsymbol{w}_{i,n}$ and the random noise $\varepsilon_{i,n}$. The first DGP follows a similar setup taking from [89] and [25], where we generate many (sparse) dummy variables entering the estimation of $\boldsymbol{\beta}$. We generate $\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$ with $\Sigma_{jk} = 0.5^{|j-k|}$ for $j,k = 1,\ldots,d$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$ with $\tilde{\boldsymbol{w}}_{i,n} \sim \mathcal{N}(0,\boldsymbol{I}_{q_n})$ and $\boldsymbol{I}_{q_n}$ is a q_n-dimensional identity matrix, and $\varepsilon_{i,n} \sim \mathcal{N}(0,1)$. The second DGP considers a case with dummy policy random variables and normal covariates, where we generate $\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0)$ with $\tilde{\boldsymbol{x}}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0,\boldsymbol{I}_{q_n})$ and $\varepsilon_{i,n} \sim \mathcal{N}(0,1)$. In the Supplementary Materials, we have further included DGPs with more realistic error terms beyond normal distribution, including error terms with asymmetric and bimodal distributions. For most of the DGPs, we investigate both homoscedastic as well as heteroscedastic models. See Supplementary Materials Section C for detailed description and simulation results.

As for the coefficients, we consider three DGPs that vary in $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}_n$. The first DGP considers the case in which no policy is effective (meaning that $\boldsymbol{\beta} = 0$), and the coefficient $\gamma_j = 1/j$, for $j = 1,\ldots,q_n$. We refer to this case as the "homogeneity" case since β_j's take the same value zero. The second and the third DGPs consider cases where policy effects are generated from $\beta_j = \Phi^{-1}\big(\frac{j}{d+1}\big)$ for $j = 1,\ldots,d$, and the coefficients are either $\boldsymbol{\gamma}_n = 0$ or $\gamma_j = 1/j$, for $j = 1,\ldots,q_n$. We refer to this case as the "heterogeneity(1)" case and "heterogeneity(2)" case, respectively, since different policies have heterogeneous effects.

We set the sample size $n \in \{700, 2000\}$ to mimic the sample size in our case studies, the number of policies $d \in \{5, 10\}$, and the dimension of the covariates q_n from $q_n \in \{1, 141, 281, 421, 561, 631\}$. All statistics reported below are computed based on over 1,000 Monte Carlo replications. To avoid redundancy, we present the results for $n = 700$ and $d = 5$ in the main manuscript, and rests are provided in the Supplementary Materials (Section C).

To demonstrate the robustness of the adopted covariance matrix estimator, we compare our proposal with three alternative covariance matrix estimators. The first one we compare with is the covariance matrix estimator proposed by [25]: $\hat{\boldsymbol{\Omega}}_n^{\texttt{HCK}} = \hat{\boldsymbol{\Gamma}}_n^{-1}\hat{\Sigma}_n^{\texttt{HCK}}\hat{\boldsymbol{\Gamma}}_n^{-1}$, where $\hat{\Sigma}_n^{\texttt{HCK}} \triangleq \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n \kappa_{ij,n}^{\texttt{HCK}}\hat{\boldsymbol{v}}_{i,n}\hat{\boldsymbol{v}}_{i,n}'\hat{u}_{j,n}^2$, $\hat{u}_{j,n} = \sum_{k=1}^n(\boldsymbol{M}_n)_{j,k}(y_{k,n} - \boldsymbol{x}_{k,n}'\hat{\boldsymbol{\beta}})$, and

$$\kappa_n^{\texttt{HCK}} = \begin{pmatrix} M_{11,n}^2 & \cdots & M_{1n,n}^2 \\ \vdots & \ddots & \vdots \\ M_{n1,n}^2 & \cdots & M_{nn,n}^2 \end{pmatrix}^{-1} = (\boldsymbol{M}_n \odot \boldsymbol{M}_n)^{-1},$$

with $\odot$ denoting the Hadamard product. The estimator $\hat{\Sigma}_n^{\texttt{HCK}}$ is well-defined whenever $(\boldsymbol{M}_n \odot \boldsymbol{M}_n)$ is invertible. We use the acronym "HCK" to denote this estimator in the following parts. The second one we compare with is the classical Eicker-White covariance matrix estimator [47, 175] of the form:

$$\hat{\boldsymbol{\Omega}}_n^{\texttt{EW}} = \hat{\boldsymbol{\Gamma}}_n^{-1}\hat{\Sigma}_n^{\texttt{EW}}\hat{\boldsymbol{\Gamma}}_n^{-1}, \tag{2.6}$$

where $\hat{\Sigma}_n^{\texttt{EW}} \triangleq \frac{1}{n}\sum_{i=1}^n \hat{\boldsymbol{v}}_{i,n}\hat{\boldsymbol{v}}_{i,n}'\hat{u}_{i,n}^2$ and $\hat{u}_{i,n} = \sum_{j=1}^n(\boldsymbol{M}_n)_{i,j}(y_{j,n} - \boldsymbol{x}_{j,n}'\hat{\boldsymbol{\beta}})$. We use the acronym "EW" to denote this estimator in our simulation results section. Huber-Eicker-White standard error is also known as the HC0 standard error, where HC stands for "heteroskedasticity

robust." The last covariance matrix estimator we adopted is a variant of the HC0 estimator:

$$\hat{\boldsymbol{\Omega}}_n^{\texttt{HC3}} = \hat{\boldsymbol{\Gamma}}_n^{-1}\hat{\Sigma}_n^{\texttt{HC3}}\hat{\boldsymbol{\Gamma}}_n^{-1}, \quad \text{where } \hat{\Sigma}_n^{\texttt{HC3}} \triangleq \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{v}}_{i,n}\hat{\boldsymbol{v}}_{i,n}' \frac{\hat{u}_{i,n}^2}{(\boldsymbol{M}_n)_{i,j}^2}. \tag{2.7}$$

The above estimator upward reweights regression residuals, and we use the acronym "HC3" to denote this estimator in our simulation results section.

Simulation results

We summarize our main takeaways from the simulation results presented in Table 2.1-2.3, where we have compared our proposed approach ("Proposed + KJ") in Section 2.2 with four other methods. "Proposed + EW", "Proposed + HC3", and "Proposed + HCK" refer to methods adjusting for the winner's curse bias but use $\hat{\boldsymbol{\Omega}}_n^{\texttt{EW}}$, $\hat{\boldsymbol{\Omega}}_n^{\texttt{HC3}}$, and $\hat{\boldsymbol{\Omega}}_n^{\texttt{HCK}}$, respectively, to estimate the covariance matrix of $\boldsymbol{\beta}$. "No adjustment+KJ" refers to the approach with no adjustment for the winner's curse bias and adopts the robust covariance matrix estimator proposed by [89] to make inference on the best policies. We present the coverage probabilities and $\sqrt{n}$-scaled biases for the top two policies in the population, i.e., $\beta_{(1)}$ and $\beta_{(2)}$. As the simulation results are rather similar for the observed top two policies in the random sample, i.e., $\beta_{\hat{1}}$, $\beta_{\hat{2}}$, we present these results in the Supplementary Materials (Section C.4).

Our simulation results confirm our theoretical results presented in Theorem 2. When no policy is effective, our proposed method not only successfully suppresses the winner's curse bias for the top two policies but also attains near nominal coverage (Table 2.2). Similar pattern can also be observed when top policies are effective (i.e., β_j's are heterogeneous, and Table 2.1 in particular). In nearly all designs and for a range of considered values of q_n, our proposal yields close to nominal confidence interval, though some under coverage is observed for large values of q_n. The method with no adjustment is obviously biased upward due to the winner's curse phenomenon, thus it provides under-covered confidence intervals and point estimates with rather large biases. In all considered cases, both the EW-based method and the HC3-based method tend to lose coverage when $q_n \geq 141$, and the HCK-based method tends to produce under-covered confidence interval whenever $q_n \geq 561$. In moderately high dimensions so that q_n/n is approximately one half, the proposed method with the HCK variance estimator has comparable performances with our approach.

2.5 Case studies

Case study I: Charitable giving

In the past half century, charitable giving by individuals in the United States has grown and it has contributed to more than two percent of the annual GDP since 1998 [114]. Charitable giving is often driven by altruism, while as suggested by many field experiments, improper policies adopted by the demand side–fundraisers–may impair the supply side's (individual

Table 2.1: Simulation results ($d = 5$, heterogeneity, $\beta_{(1)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1})$, $\gamma_n = 0$, j = 1,..., d				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.97(0.01)	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.97(0.01)
	$\sqrt{n}$Bias	-0.04(0.06)	-0.03(0.04)	-0.03(0.04)	-0.04(0.05)	0.05(0.06)
$q_n = 141$	Cover	0.96(0.01)	0.96(0.01)	0.94(0.01)	0.95(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.04(0.05)	-0.04(0.04)	-0.04(0.04)	0.06(0.06)	0.06(0.07)
$q_n = 281$	Cover	0.96(0.01)	0.95(0.01)	0.82(0.02)	0.80(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.05(0.06)	-0.06(0.05)	-0.06(0.03)	-0.10(0.07)	-0.08(0.08)
$q_n = 421$	Cover	0.95(0.02)	0.94(0.01)	0.79(0.01)	0.76(0.01)	0.78(0.01)
	$\sqrt{n}$Bias	-0.05(0.05)	-0.06(0.05)	-0.07(0.05)	-0.12(0.09)	0.11(0.09)
$q_n = 561$	Cover	0.95(0.01)	0.92(0.01)	0.65(0.02)	0.63(0.01)	0.68(0.01)
	$\sqrt{n}$Bias	-0.07(0.07)	-0.09(0.07)	-0.17(0.10)	-0.20(0.12)	0.15(0.13)
$q_n = 631^*$	Cover	0.93(0.01)	0.91(0.01)	0.51(0.02)	0.48(0.01)	0.55(0.01)
	$\sqrt{n}$Bias	-0.17(0.08)	-0.19(0.10)	-0.28(0.11)	-0.35(0.22)	-0.26(0.13)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0)$, $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed + EW	No adjustment+KJ
$q_n = 1$	Cover	0.97(0.01)	0.97(0.01)	0.95(0.01)	0.96(0.01)	0.97(0.01)
	$\sqrt{n}$Bias	-0.02(0.07)	-0.02(0.04)	-0.01(0.03)	-0.07(0.11)	-0.05(0.09)
$q_n = 141$	Cover	0.96(0.01)	0.95(0.01)	0.94(0.01)	0.94(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	-0.02(0.03)	-0.02(0.02)	-0.03(0.02)	0.11(0.12)	-0.06(0.12)
$q_n = 281$	Cover	0.95(0.01)	0.94(0.01)	0.87(0.01)	0.85(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.03(0.04)	-0.03(0.03)	-0.04(0.02)	0.14(0.12)	-0.08(0.13)
$q_n = 421$	Cover	0.95(0.01)	0.94(0.01)	0.78(0.01)	0.76(0.01)	0.75(0.01)
	$\sqrt{n}$Bias	-0.03(0.03)	-0.05(0.04)	-0.08(0.05)	-0.19(0.17)	0.19(0.14)
$q_n = 561$	Cover	0.95(0.01)	0.92(0.01)	0.63(0.02)	0.61(0.01)	0.63(0.01)
	$\sqrt{n}$Bias	-0.04(0.04)	-0.08(0.06)	-0.19(0.10)	-0.30(0.26)	-0.24(0.22)
$q_n = 631$	Cover	0.94(0.01)	0.91(0.01)	0.49(0.02)	0.45(0.01)	0.68(0.01)
	$\sqrt{n}$Bias	-0.06(0.07)	-0.11(0.09)	-0.23(0.13)	-0.42(0.20)	0.39(0.29)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(1)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(1)}$. " * " indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

donors) motivation of giving [6]. Therefore, to effectively attract resources from individual donors, fundraisers need to properly design donation incentives. One of the donation incentives is matching grant which means that a matching donor pledges to match any donation from other donors with certain ratio and up to some threshold. As the price elasticity of matching donation may differ from other donation incentives, we hope to carefully investigate different pricing policies in a matching donation and study if the observed top two performing policies are indeed effective.

We work with the charitable giving data in [92]. [92] conduct a field experiment that explores the price elasticity in a matching donation. The field experiment involves 50,083 previous donors to a political charity. Individuals are randomly assigned to two groups: treatment ($n = 33,396$) and control ($n = 16,687$). In the control group, individuals receive a standard letter with no matching details. In the treatment group, each potential donor receives a letter with three strategies: (1) match ratio, (2) match size, and (3) ask amount.

Table 2.2: Simulation results ($d = 5$, homogeneity, $\beta_{(1)}$)

		$\boldsymbol{\beta} = 0, \quad \boldsymbol{\gamma}_j = 1/j$				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma), \quad \boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.97(0.01)	0.96(0.01)	0.96(0.01)	0.93(0.02)	0.90(0.01)
	$\sqrt{n}$Bias	0.02(0.03)	0.02(0.03)	0.03(0.03)	0.03(0.03)	1.64(0.04)
$q_n = 141$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.89(0.02)	0.88(0.01)
	$\sqrt{n}$Bias	0.03(0.04)	0.04(0.04)	0.03(0.04)	0.12(0.04)	1.78(0.04)
$q_n = 281$	Cover	0.96(0.01)	0.94(0.01)	0.90(0.02)	0.85(0.01)	0.83(0.01)
	$\sqrt{n}$Bias	0.03(0.04)	0.04(0.04)	0.05(0.04)	0.22(0.03)	2.03(0.05)
$q_n = 421$	Cover	0.95(0.01)	0.93(0.01)	0.82(0.02)	0.79(0.01)	0.74(0.02)
	$\sqrt{n}$Bias	0.05(0.05)	0.18(0.05)	0.24(0.06)	0.36(0.03)	2.63(0.06)
$q_n = 561$	Cover	0.95(0.01)	0.93(0.01)	0.67(0.02)	0.73(0.01)	0.63(0.02)
	$\sqrt{n}$Bias	0.08(0.09)	0.51(0.05)	0.74(0.06)	0.44(0.04)	3.74(0.09)
$q_n = 631^*$	Cover	0.93(0.01)	0.89(0.01)	0.53(0.02)	0.50(0.01)	0.45(0.02)
	$\sqrt{n}$Bias	0.18(0.09)	1.21(0.09)	1.84(0.11)	2.42(0.06)	5.10(0.12)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0), \quad \boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.94(0.01)	0.90(0.01)
	$\sqrt{n}$Bias	0.03(0.04)	0.04(0.05)	0.05(0.05)	0.07(0.07)	2.75(0.06)
$q_n = 141$	Cover	0.96(0.01)	0.96(0.01)	0.93(0.01)	0.90(0.01)	0.83(0.01)
	$\sqrt{n}$Bias	0.05(0.05)	0.05(0.06)	0.17(0.07)	0.31(0.07)	3.29(0.08)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.90(0.01)	0.88(0.01)	0.75(0.02)
	$\sqrt{n}$Bias	0.07(0.08)	0.07(0.07)	0.31(0.08)	0.54(0.05)	3.59(0.08)
$q_n = 421$	Cover	0.95(0.01)	0.95(0.01)	0.85(0.01)	0.86(0.01)	0.65(0.02)
	$\sqrt{n}$Bias	0.04(0.04)	0.10(0.11)	0.73(0.13)	0.64(0.06)	4.58(0.11)
$q_n = 561$	Cover	0.93(0.01)	0.90(0.02)	0.59(0.02)	0.61(0.01)	0.53(0.02)
	$\sqrt{n}$Bias	0.13(0.07)	0.19(0.13)	2.00(0.13)	1.73(0.08)	5.90(0.13)
$q_n = 631$	Cover	0.90(0.01)	0.78(0.02)	0.38(0.02)	0.30(0.01)	0.33(0.02)
	$\sqrt{n}$Bias	0.50(0.12)	2.47(0.16)	5.16(0.19)	7.68(0.18)	6.51(0.19)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(1)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(1)}$. " * " indicates that $\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Within each strategy, individuals are randomly assigned to a sub-policy detailed below.

For the match ratio strategy, there are three sub-policies: (1) 1:1 (the matching donor contributes the same amount as the individual donor), (2) 2:1 (the matching donor contributes twice as many as the individual donor), (3) 3:1 (the matching donor contributes three times as many as the individual donor). For the match size strategy, there are four sub-policies with different pledge amounts: (1) \$25,000, (2) \$50,000, (3) \$100,000, and (4) unstated amount. For the ask amount strategy, individual donors are asked to give same amount, 25% more or 50% more than their largest past donation.

In our study, we focus on the treatment "ask amount" with three pricing policies, and we study the subpopulation ($n = 7,938$) of unmarried males living in red counties or red states. Red county (state) refers to a county (state) in which residents predominantly vote for the Republican Party. The outcome of interest is the donation amount. We have adjusted $q_n = 1,049$ covariates including the donors' demographic information (26 variables), census

Table 2.3: Simulation results ($d = 5$, heterogeneity, $\beta_{(2)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1})$, $\gamma_n = 0$				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.97(0.01)	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	0.02(0.06)	0.02(0.06)	0.01(0.06)	-0.05(0.07)	-0.04(0.07)
$q_n = 141$	Cover	0.97(0.01)	0.96(0.01)	0.93(0.01)	0.94(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.02(0.04)	-0.02(0.04)	-0.04(0.03)	-0.06(0.06)	-0.07(0.08)
$q_n = 281$	Cover	0.95(0.01)	0.94(0.01)	0.88(0.02)	0.89(0.01)	0.85(0.01)
	$\sqrt{n}$Bias	-0.03(0.04)	-0.03(0.03)	-0.07(0.03)	-0.11(0.10)	0.19(0.11)
$q_n = 421$	Cover	0.95(0.01)	0.94(0.02)	0.81(0.02)	0.80(0.01)	0.77(0.01)
	$\sqrt{n}$Bias	-0.03(0.03)	-0.03(0.04)	-0.10(0.06)	-0.15(0.12)	-0.23(0.15)
$q_n = 561$	Cover	0.95(0.01)	0.94(0.02)	0.67(0.02)	0.65(0.01)	0.63(0.01)
	$\sqrt{n}$Bias	0.03(0.03)	0.05(0.06)	0.12(0.08)	-0.17(0.13)	-0.26(0.18)
$q_n = 631^*$	Cover	0.94(0.01)	0.93(0.01)	0.56(0.02)	0.53(0.01)	0.50(0.01)
	$\sqrt{n}$Bias	-0.07(0.07)	-0.18(0.06)	-0.23(0.08)	-0.26(0.17)	0.47(0.22)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0)$, $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.98(0.01)	0.98(0.01)	0.98(0.01)	0.96(0.01)	0.98(0.01)
	$\sqrt{n}$Bias	-0.08(0.12)	-0.09(0.12)	-0.10(0.13)	0.09(0.12)	-0.07(0.10)
$q_n = 141$	Cover	0.97(0.01)	0.97(0.01)	0.95(0.01)	0.95(0.01)	0.97(0.01)
	$\sqrt{n}$Bias	-0.09(0.12)	-0.10(0.13)	-0.12(0.13)	0.09(0.13)	-0.08(0.10)
$q_n = 281$	Cover	0.97(0.01)	0.97(0.01)	0.90(0.01)	0.87(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	-0.11(0.14)	-0.10(0.14)	-0.15(0.14)	-0.18(0.14)	-0.10(0.11)
$q_n = 421$	Cover	0.96(0.01)	0.95(0.02)	0.80(0.02)	0.75(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.14(0.16)	-0.16(0.18)	-0.20(0.18)	-0.22(0.17)	-0.15(0.15)
$q_n = 561$	Cover	0.96(0.01)	0.94(0.02)	0.63(0.02)	0.60(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	0.14(0.18)	0.20(0.23)	-0.24(0.22)	-0.30(0.23)	0.19(0.20)
$q_n = 631$	Cover	0.94(0.01)	0.93(0.02)	0.58(0.02)	0.55(0.01)	0.52(0.01)
	$\sqrt{n}$Bias	0.15(0.20)	0.24(0.26)	0.28(0.25)	-0.35(0.13)	0.56(0.24)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(2)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(2)}$. " * " indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

information (27 variables), and their two-way interaction terms. Our results are summarized in Table 2.4.

Results in Table 2.4 suggest that, without any calibration, asking the donor either to give the same amount or to give 25% more than their highest past donation seems to be the best policies that significantly increase the donation amount. Specifically, our results from running a simple linear regression model suggest that asking the individual donor to give the same amount of their largest past donation appears to be the most effective pricing policy, and it on average raises \$0.67 (95% CI = (0.09, 1.25), p-value = 0.023) per donor. Asking the individual donor to give 25% more than their largest past donation is the second most effective policy, with an increased donation by \$0.66 (95% CI = (0.01, 1.31), p-value = 0.046) per donor.

Because we pick the most effective policies from a random sample, these estimates are potentially subject to the winner's curse bias. We thus apply the proposed method to

Method	Policies(Ask amount)	Est (95% CI)	p-value
Uncalibrated	Same	0.67 (0.09, 1.25)	0.023*
	25% more	0.66 (0.01, 1.31)	0.046*
	50% more	0.33 (-0.21, 0.86)	0.235
Calibrated	Same	0.63 (0.10, 1.20)	0.025*
	25% more	0.56(-0.01, 1.07)	0.052

Table 2.4: Estimated treatment effects (Est), 95% confidence intervals (95% CI), and two-sided p-values for the three "ask amount" policies. "Uncalibrated" refers to the study results obtained without any adjustment, and the confidence intervals are constructed based on normal approximation with the estimated covariance matrix $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$. "Calibrated" refers to our proposed methodology. The computational time is 741 seconds on a Lenovo NeXtScale nx360m5 node (24 cores per node) equipped with Intel Xeon Haswell processor.

carefully examine these seemly effective policies. After calibrating for the winner's curse bias, we confirm that the asking for the same amount policy remains as the most effective policy, though with a slightly smaller estimated effect size (Est = \$0.63, 95% CI = (0.10, 1.20), p-value = 0.025). This result is moderately aligned with the analysis in [92], whose results suggest that donors from red states or red counties are more willing to contribute, partially because the collaborating charity is politically oriented. However, for the effect of the second best policy–asking to donate 25% more than past donation–is shifted downward, and it no longer has significant impact in promoting the donation amount (Est = \$0.56, 95% CI = (-0.01, 1.07), p-value = 0.052). This result might be partially explained by the observation that donors are more motivated by a lower "price" of donation [173]. In sum, our analyses suggest that the best pricing policy of charitable giving for unmarried males living in the Republican Party dominated voting regions could be asking for the same amount as their highest previous donation, and asking for more donations may not incentivize the donors to give. Though given the obtained p-value before and after calibration for the second best policy is rather close to the 5 percent threshold, we note that such a conclusion should also be viewed with caution.

Case study II: National supported work (NSW) program

In this case study, we revisit a dataset from the National Supported Work (NSW) program. The NSW program is a labor training program implemented in 1970's that provides work experience to disadvantaged workers. Our proposed method can also be used to evaluate if the job training program is indeed beneficial for certain groups of workers. To do so, the structural component $\boldsymbol{x}_{i,n}$ in the model (2.1) would include variables representing the inter-

actions between the treatment variable (the job training program) and different subgroup indicator variables of interest.

We use the field experiment dataset adopted in [39] ($n = 455$), in which 185 workers are in the treatment group and 260 workers are in the control group. This dataset consists of a treatment indicator variable, an outcome variable measured by the participant post-treatment earnings in 1978, and eight baseline variables (including age, years of education, an indicator for high school degree, indicators for Black and Hispanic, marital status, and pre-treatment earnings in 1974 and 1975). We further add three sets of additional covariates following the setup in [59]: (1) $\mathbb{1}$(1974 earnings =0) and $\mathbb{1}$(1975 earnings =0); (2) all first-order interactions; (3) all polynomials up to the 2nd-order. The final dataset includes 51 covariates. We aim to investigate the effectiveness of the NSW program in four groups of workers: (1) married Black workers, (2) unmarried Black workers, (3) married Non-Black workers, and (4) unmarried Non-Black workers. The summarized results are shown in Table 2.5.

Method	Subgroups	Est (95% CI) ($\$10^3$)	p-value
Uncalibrated	Black, married	4.35 (0.89, 7.81)	0.014*
	Black, unmarried	1.10(-0.55, 2.75)	0.190
	Non-Black, married	1.33(-6.63, 9.29)	0.743
	Non-Black, unmarried	1.40(-2.61, 5.40)	0.494
Calibrated	Black, married	4.41(1.74, 8.50)	0.009*

Table 2.5: Estimated treatment effects (Est), 95% confidence intervals (CI), in units $\$10^3$/year, and two-sided p-values for the four subgroups in the NSW study ($n = 445$, $q_n = 51$). "Uncalibrated" refers to the study results obtained without any adjustment, and the confidence intervals are constructed based on normal approximation with the estimated covariance matrix $\hat{\mathbf{\Omega}}_n^{\mathrm{KJ}}$. "Calibrated" refers to our proposed methodology. The computational time is 122 seconds on a Lenovo NeXtScale nx360m5 node (24 cores per node) equipped with Intel Xeon Haswell processor.

Table 2.5 demonstrates that without adjusting for the winner's curse bias, married Black workers (estimated treatment effect = 4.35, 95% CI = (0.89, 7.81), p-value = 0.014, in units $\$10^3$) seem to benefit from the program the most. After accounting for the winner's curse bias issue, our approach potentially confirms that the treatment effect of the NSW program for the married Black workers is still significant, and the calibrated treatment effect remains roughly the same (Est= 4.41, 95% CI = (1.74, 8.50), p-value = 0.009, in units $\$10^3$/year).

The dataset collected from the NSW program has been frequently analyzed in the past decade, and our results are largely in-line with current understandings gathered in past studies. For example, although not focusing on the same groups of workers, [86] suggest that married and unemployed Black workers with some college education have increased their

post-treatment earnings for about 38%. [39] show that the job training program yields positive treatment effect on the overall Black participants. In this case study, our approach may help to confirm the seemly effective subgroup observed in a random sample while providing a statistically justified estimate accounting for the winner's curse bias.

2.6 Discussion

In this article, we have introduced an approach to evaluate multiple best policies based on resampling in the context of a linear model with many covariates. While our approach is numerically reliable and theoretically grounded, it is worthwhile to generalize our framework so that the policy effects can be estimated with other off-shelf methods that are, for example, robust to the high-dimensional confounders or to the presence of interference and noncompliance. Our current theoretical analysis suggests that our proposed approach can be readily extended as long as the covariance matrix between different policies can be consistently estimated. It is thus desirable for us to provide a general framework to broaden future applications for other disciplines in general.

2.7 Supplementary Materials

Theorem 1

Review of notations and assumptions

We denote the sample $\{(y_{i,n}, \boldsymbol{x}'_{i,n}, \boldsymbol{w}'_{i,n})'\}_{i=1}^n$ as $\{\boldsymbol{z}_{i,n}\}_{i=1}^n$. Recall $u_{i,n}$ is the random error in the considered linear model:

$$y_{i,n} = \boldsymbol{x}'_{i,n}\boldsymbol{\beta} + \boldsymbol{w}'_{i,n}\boldsymbol{\gamma}_n + u_{i,n}, \quad i = 1, \ldots, n, \tag{2.8}$$

we define

$$\varepsilon_{i,n} = u_{i,n} - \mathbb{E}[u_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n], \quad \boldsymbol{v}_{i,n} = \boldsymbol{x}_{i,n} - \mathbb{E}[\boldsymbol{x}_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n], \quad i = 1, \ldots, n. \tag{2.9}$$

Let $e_{i,n} = \mathbb{E}[u_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n]$, we further denote

$$\hat{u}_i = \sum_{j=1}^n (\boldsymbol{M}_n)_{i,j}(y_{j,n} - \boldsymbol{x}'_{j,n}\hat{\boldsymbol{\beta}}), \quad \hat{\boldsymbol{v}}_{i,n} = \sum_{i=1}^n (\boldsymbol{M}_n)_{i,j}\boldsymbol{x}_{j,n},$$

$$(\boldsymbol{M}_n)_{i,j} = \mathbb{1}(i=j) - \boldsymbol{w}'_{i,n}\Big(\sum_{k=1}^n \boldsymbol{w}_{k,n}\boldsymbol{w}'_{k,n}\Big)^{-1}\boldsymbol{w}_{j,n},$$

$$\sigma^2_{i,n} = \mathbb{E}[\varepsilon^2_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n, \{\boldsymbol{x}_{i,n}\}_{i=1}^n], \quad \tilde{\boldsymbol{v}}_{i,n} = \sum_{j=1}^n (\boldsymbol{M}_n)_{i,j}\boldsymbol{v}_{j,n},$$

$$\rho^1_n = \frac{\sum_{i=1}^n \mathbb{E}\big[e^2_{i,n}\big]}{n}, \quad \rho^2_n = \frac{\sum_{i=1}^n \mathbb{E}\Big[\mathbb{E}\big(e_{i,n}|\{\boldsymbol{w}_{i,n}\}_{i=1}^n\big)^2\Big]}{n},$$

$$\boldsymbol{Q}_{i,n} = \mathbb{E}\Big[\boldsymbol{x}_{i,n} - \big(\sum_{j=1}^n \mathbb{E}[\boldsymbol{w}_{j,n}\boldsymbol{w}'_{j,n}]\big)^{-1}\sum_{j=1}^n \mathbb{E}[\boldsymbol{w}_{j,n}\boldsymbol{x}'_{j,n}]\Big|\{\boldsymbol{w}_{i,n}\}_{i=1}^n\Big].$$

We will use the notation $\mathbb{P}(\cdot|\{\boldsymbol{z}_{i,n}\}_{i=1}^n)$ to refer to the probability that is conditional on the random variables $\{\boldsymbol{z}_{i,n}\}_{i=1}^n$.

For a policy j, recall our definition of the near tie set:

$$\mathcal{H}_{(j)} = \big\{k : |\beta_k - \beta_{(j)}| = o(n^{-\frac{1}{2}}),\ k = 1, \ldots, d\big\}.$$

This suggests that $\forall k \in \mathcal{H}_{(j)}$, there exist a sequence $\delta_n \to 0$ as $n \to 0$, such that

$$\beta_k = \beta_{(j)} + n^{-\frac{1}{2}} \cdot \delta_n, \quad \forall k \in \mathcal{H}_{(j)}.$$

Next, let $\hat{\boldsymbol{e}}_j$ denote a d-dimensional (sparse) vector based on the estimated tie set $\widehat{\mathcal{H}}_{(j)}$ with

$$\hat{\boldsymbol{e}}_j = (\hat{e}_{j,1}, \ldots, \hat{e}_{j,d}), \quad \hat{e}_{jk} = \frac{\mathbb{1}(k \in \widehat{\mathcal{H}}_{(j)})}{|\widehat{\mathcal{H}}_{(j)}|}, \quad k = 1, \ldots, d,$$

and define a $d-$dimensional sparse index vector based on the true near-tie set $\mathcal{H}_{(j)}$ as

$$\boldsymbol{e}_j = (e_{j,1}, \ldots, e_{j,d}), \quad e_{jk} = \frac{\mathbb{1}(k \in \mathcal{H}_{(j)})}{|\mathcal{H}_{(j)}|}, \quad k = 1, \ldots, d. \tag{2.10}$$

We make following assumptions throughout this section:

Assumption 11 (Sampling)**.** *The errors $\varepsilon_{i,n}$ are uncorrelated across i conditional on*

$$\{\boldsymbol{x}_{i,n}\}_{i=1}^n$$

and

$$\{\boldsymbol{w}_{i,n}\}_{i=1}^n.$$

Let $\{N_1, \ldots, N_{G_n}\}$ represents a partition of $\{1, \ldots, n\}$ with $\max_{1 \le g \le G_n} |N_g| = O(1)$ such that $\{(\varepsilon_{i,n}, \boldsymbol{v}_{i,n}), i \in N_g\}$ (defined in (2.9)) are independent across g conditional on $\{\boldsymbol{w}_{i,n}\}_{i=1}^n$.

Assumption 12 (Design). *The dimension of the covariates $\boldsymbol{w}_{i,n}$ satisfies that*

$$\limsup_{n\to\infty} q_n/n < 1.$$

The minimum eigenvalue of of the matrix $\sum_{i=1}^n \boldsymbol{w}_{i,n}\boldsymbol{w}_{i,n}'$ is bounded away from 0 with probability approaching one, that is

$$\lim_{n\to\infty} \mathbb{P}\Big(\lambda_{\min}\big(\sum_{i=1}^n \boldsymbol{w}_{i,n}\boldsymbol{w}_{i,n}'\big) > 0\Big)\Big) = 1.$$

Lastly,

$$\max_{1\le i\le n}\Big\{\mathbb{E}[\varepsilon_{i,n}^4|\{\boldsymbol{w}_{i,n}\}_{i=1}^n,\{\boldsymbol{x}_{i,n}\}_{i=1}^n],\ \frac{1}{\sigma_{i,n}^2},\\ \mathbb{E}[\boldsymbol{v}_{i,n}^4|\{\boldsymbol{w}_{i,n}\}_{i=1}^n],\ 1/\lambda_{\min}\big(\frac{\sum_{i=1}^n \mathbb{E}\big[\tilde{\boldsymbol{v}}_{i,n}\tilde{\boldsymbol{v}}_{i,n}'|\{\boldsymbol{w}_{i,n}\}_{i=1}^n\big]}{n}\big)\Big\} = O_p(1).$$

Assumption 13 (Linear model approximation). $\sum_{i=1}^n \mathbb{E}[||Q_{i,n}||^2]/n = O(1)$, $\rho_n^1 + n(\rho_n^1 - \rho_n^2) + \rho_n^2 \cdot \sum_{i=1}^n \mathbb{E}[||Q_{i,n}||^2] = o(1)$, *and* $\max_{1\le i\le n}||\hat{\boldsymbol{v}}_{i,n}||/\sqrt{n} = o_p(1)$, $n\rho_n^1 = O(1)$.

Assumption 14 (Variance estimation). $\lim_{n\to\infty}\mathbb{P}(\min_i \big(\boldsymbol{M}_n\big)_{i,i} > 0) = 1$,

$$\mathbb{P}\Big(\min_i \big(\boldsymbol{M}_n\big)_{i,i} > 0\Big) = O_p(1), \quad \frac{\sum_{i=1}^n ||\tilde{\boldsymbol{Q}}_{i,n}||^4}{n} = O_p(1),$$

and $\max_i ||\mu_{i,n}||/\sqrt{n} = o_p(1)$ *with* $\mu_{i,n} = \mathbb{E}\big[y_{i,n}|\{\mathbf{x}_{i,n}\}_{i=1}^n,\{\mathbf{w}_{i,n}\}_{i=1}^n\big]$.

Assumption 15 (Policy effect sizes). *For $\delta \in (0,\frac{1}{2})$, the asymptotic distance between the effects of policy k and j diverges as $n\to\infty$:*

$$n^\delta \cdot \min_{k\notin\mathcal{H}_{(j)}} \big|\beta_{(j)} - \beta_k\big| \to \infty, \ \text{as } n\to\infty, \quad j = 1,\dots,d.$$

Proof of Theorem 2

In this section, we show the following theorem holds:

Theorem 3. *Under Assumptions 11-15, for any $t\in\mathbb{R}$, for the resampled statistics, the following holds precisely*

$$\mathbb{P}\left(\frac{\sqrt{n}\big(\tilde{\beta}_{(j)}^* - \tilde{\beta}_{(j)}\big)}{(\hat{\boldsymbol{e}}_j'\hat{\boldsymbol{\Omega}}_n^{KJ}\hat{\boldsymbol{e}}_j)^{\frac{1}{2}}} \le t\Big|\{(y_{i,n},\boldsymbol{x}_{i,n}',\boldsymbol{w}_{i,n}')'\}_{i=1}^n\right) = \Phi(t).$$

For the original statistics, it holds that

$$\lim_{n\to\infty}\mathbb{P}\left(\frac{\sqrt{n}\big(\tilde{\beta}_{(j)}-\beta_{(j)}\big)}{(\hat{\boldsymbol{e}}_j'\hat{\boldsymbol{\Omega}}_n^{KJ}\hat{\boldsymbol{e}}_j)^{\frac{1}{2}}}\leq t\right)=\Phi(t).$$

In addition, we show that

$$\lim_{n\to\infty}\mathbb{P}\Big(\mathbb{P}\big(\tilde{\beta}^*_{(j)}\leq\beta_{(j)}|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\big)\leq s\Big)=s.$$

Proof. Our proof of Theorem 1 entails the following steps:

Step 1. Under Assumptions 11-13, [89] has shown the following holds

$$(\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}})^{-\frac{1}{2}}\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\rightsquigarrow N(0,\boldsymbol{I}_d),$$

where $\boldsymbol{I}_d$ is a d-dimensional identity matrix. Therefore, following the definition of $\boldsymbol{e}_j$ in Eq (2.10), we have

$$(\boldsymbol{e}_j'\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\Big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}\Big)\rightsquigarrow\mathcal{N}(0,1).\tag{2.11}$$

Step 2. Because of Lemma 2, we have

$$\lim_{n\to\infty}\mathbb{P}(\hat{\mathcal{H}}_{(j)}\neq\mathcal{H}_{(j)})=0.$$

Combing this with (2.11), we have

$$\begin{aligned}\Phi(t)&=\lim_{n\to\infty}\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\Big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}\Big)\leq t\Big)\\&=\lim_{n\to\infty}\Bigg[\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\Big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}\Big)\leq t\Big|\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big)\\&\mathbb{P}(\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)})+o_{\mathbb{P}}(1)\Bigg]\\&=\lim_{n\to\infty}\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\Big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}\Big)\leq t\Big|\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big),\end{aligned}$$

in which $o_{\mathbb{P}}(1)$ is lower bounded by zero and upper bounded by $\mathbb{P}(\hat{\mathcal{H}}_{(j)}\neq\mathcal{H}_{(j)})$, which tends to zero when n tends to infinity. We use this same $o_{\mathbb{P}}(1)$ notion throughout this proof. Now we have

$$\lim_{n\to\infty}\mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\Big(\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\hat{\mathcal{H}}_{(j)}|}-\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\beta_k}{|\hat{\mathcal{H}}_{(j)}|}\Big)\leq t\Big|\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big)=\Phi(t).$$

Next, we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\hat{\mathcal{H}}_{(j)}|}-\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\beta_k}{|\hat{\mathcal{H}}_{(j)}|}\big)\le t\Big)\\
&=\lim_{n\to\infty}\Bigg[\mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\hat{\mathcal{H}}_{(j)}|}-\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\beta_k}{|\hat{\mathcal{H}}_{(j)}|}\big)\le t\Big|\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big)\Big)\\
&\quad\mathbb{P}(\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)})+o_{\mathbb{P}}(1)\Bigg]\\
&=\lim_{n\to\infty}\mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\hat{\mathcal{H}}_{(j)}|}-\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\beta_k}{|\hat{\mathcal{H}}_{(j)}|}\big)\le t\Big|\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big)\\
&=\Phi(t)
\end{aligned}$$

The following holds precisely following the definition of the resampling procedure:

$$\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k^*}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)\le t\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\Big)=\Phi(t). \tag{2.12}$$

We now show that

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k^*}{|\hat{\mathcal{H}}_{(j)}|}-\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\hat{\mathcal{H}}_{(j)}|}\big)\le t\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\Big)\\
&=\lim_{n\to\infty}\Bigg[\mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k^*}{|\hat{\mathcal{H}}_{(j)}|}-\frac{\sum_{k\in\hat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\hat{\mathcal{H}}_{(j)}|}\big)\le t\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n,\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big)\\
&\quad\mathbb{P}(\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)})+o_{\mathbb{P}}(1)\Bigg]\\
&=\lim_{n\to\infty}\Bigg[\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k^*}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)\le t\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n,\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)}\Big)\\
&\quad\mathbb{P}(\hat{\mathcal{H}}_{(j)}=\mathcal{H}_{(j)})+o_{\mathbb{P}}(1)\Bigg]\\
&=\lim_{n\to\infty}\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k^*}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)\le t\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\Big)\\
&=\Phi(t).
\end{aligned}$$

Recall our definition in the main manuscript

$$\tilde{\beta}_{(j)}^*=\frac{\sum_{k\in\widehat{\mathcal{H}}_{(j)}}\hat{\beta}_k^*}{|\widehat{\mathcal{H}}_{(j)}|},\ \text{and}\ \tilde{\beta}_{(j)}=\frac{\sum_{k\in\widehat{\mathcal{H}}_{(j)}}\hat{\beta}_k}{|\widehat{\mathcal{H}}_{(j)}|}, \tag{2.13}$$

we thus have reached the conclusion presented in the theorem:

$$\lim_{n\to\infty}\mathbb{P}\Big((\hat{\boldsymbol{e}}_j'\hat{\Sigma}_n^{\mathrm{KJ}}\hat{\boldsymbol{e}}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\tilde{\beta}^*_{(j)}-\tilde{\beta}_{(j)}\big)\le t\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\Big)=\Phi(t).$$

Step 3. Lastly, to prove the bootstrap consistency, we show that

$$\lim_{n\to\infty}\mathbb{P}\Big(\mathbb{P}\big(\tilde{\beta}^*_{(j)}\le\beta_{(j)}|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\big)\le s\Big)=s.$$

Note that

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}\big(\tilde{\beta}^*_{(j)}\le\beta_{(j)}|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\big)\\
=&\lim_{n\to\infty}\mathbb{P}\Big((\boldsymbol{e}_j'\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}^*_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)\\
\le&(\boldsymbol{e}_j'\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)\Big|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\Big)\\
=&\lim_{n\to\infty}\Phi((\boldsymbol{e}_j'\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)).
\end{aligned}$$

Therefore, by the bounded convergence theorem, we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}\Big(\mathbb{P}\big(\tilde{\beta}^*_{(j)}\le\beta_{(j)}|\{\boldsymbol{z}_{i,n}\}_{i=1}^n\big)\le s\Big)\\
=&\mathbb{P}(\lim_{n\to\infty}\Phi((\boldsymbol{e}_j'\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big))\le s)\\
=&\mathbb{P}(\lim_{n\to\infty}(\boldsymbol{e}_j'\hat{\boldsymbol{\Omega}}_n^{\mathrm{KJ}}\boldsymbol{e}_j)^{-\frac{1}{2}}\cdot\sqrt{n}\big(\frac{\sum_{k\in\mathcal{H}_{(j)}}\beta_k}{|\mathcal{H}_{(j)}|}-\frac{\sum_{k\in\mathcal{H}_{(j)}}\hat{\beta}_k}{|\mathcal{H}_{(j)}|}\big)\le\Phi^{-1}(s))\\
=&\mathbb{P}(N(0,1)\le\Phi^{-1}(s))\\
=&s.
\end{aligned}$$

□

Lemmas and corollary

Lemma 2

Lemma 2. *Suppose $w_{k,(j)}=\mathbb{1}(k\in\hat{\mathcal{H}}_{(j)})$, for $j,k=1,\dots,d$, under Assumptions 11-15, we have the following argument holds $\forall\varepsilon>0$,*

$$\lim_{n\to\infty}\mathbb{P}\big(|w_{k,(j)}-\mathbb{1}(k\in\mathcal{H}_{(j)})|>\varepsilon\big)=0.$$

Proof. We start with reviewing and introducing some notations to pave the way for a clear proof. Recall that $\tilde{\beta}^*_{(j)}$ is the j-th largest effect size for the resampled statistics $\hat{\beta}_{(1)}, \ldots, \hat{\beta}_{(d)}$, suppose $\tilde{\beta}^*_{(j)}$ is resampled statistics from the normal distribution centered at $\hat{\beta}_{\check{j}}$, that is

$$\tilde{\beta}^*_{(j)} \big| \{\boldsymbol{z}_{i,n}\}_{i=1}^n \sim \mathcal{N}(\hat{\beta}_{\check{j}}, (\hat{\Sigma}_n^{\mathrm{KJ}})_{\check{j}\check{j}}), \quad \check{j} = \sum_{k=1}^{d} k \cdot \mathbb{1}(\hat{\beta}^*_k = \tilde{\beta}^*_{(j)}),$$

where $(\hat{\Sigma}_n^{\mathrm{KJ}})_{\check{j}\check{j}}$ is the $\check{j}$th component in the diagonal of the matrix $\hat{\Sigma}_n^{\mathrm{KJ}}$.

Recall we define the near tie $\mathcal{H}_{(j)}$ set as

$$\mathcal{H}_{(j)} = \big\{k : |\beta_k - \beta_{(j)}| = o(n^{-\frac{1}{2}}),\ k = 1, \ldots, d\big\}.$$

We further define two sets of policies that have effect sizes lower/larger than the policies in the set $\mathcal{H}_{(j)}$:

$$\mathcal{H}^L_{(j)} = \big\{k :\ \beta_k < \min_{m \in \mathcal{H}_{(j)}} \{\beta_m\}\ k = 1, \ldots, d\big\}, \quad \mathcal{H}^U_{(j)} = \big\{k :\ \beta_k > \max_{m \in \mathcal{H}_{(j)}} \{\beta_m\}\ k = 1, \ldots, d\big\}.$$

As for the estimated near tie set, we have for any $j \in \hat{\mathcal{H}}_{(j)}$ that

$$-b_L \le \hat{\beta}^*_k - \tilde{\beta}^*_{(j)} \le b_R, \quad \text{with } |b_R - b_L| = O(n^{-\delta}),$$

where $\delta \in (0, 0.5)$. Thus, there exists a positive constant C such that

$$\frac{|\hat{\beta}^*_k - \tilde{\beta}^*_{(j)}|}{n^{-\delta}} < C, \quad \forall j \in \hat{\mathcal{H}}_{(j)}.$$

Our proof is composed of the following three steps:

Step 1. We first show that the policy with the jth largest policy effect size in the resampled statistics falls into the set $\mathcal{H}_{(j)}$ with high probability, that is

$$\lim_{n \to \infty} \mathbb{P}\Big(\check{j} \in \mathcal{H}_{(j)}\Big) = 1. \tag{2.14}$$

Because $\hat{\beta}^*_{\check{j}} \in [\min_{j \in \mathcal{H}_{(j)}} \hat{\beta}^*_j, \max_{j \in \mathcal{H}_{(j)}} \hat{\beta}^*_j]$ by definition, coupled with the fact that

$$\Big\{ \max_{k \in \mathcal{H}^L_{(j)}} \hat{\beta}^*_k < \min_{j \in \mathcal{H}_{(j)}} \hat{\beta}^*_j \le \max_{j \in \mathcal{H}_{(j)}} \hat{\beta}^*_j < \min_{k \in \mathcal{H}^U_{(j)}} \hat{\beta}^*_k \Big\} \subset \Big(\check{j} \in \mathcal{H}_{(j)}\Big),$$

it is suffice to show

$$\lim_{n \to \infty} \mathbb{P}\left(\max_{k \in \mathcal{H}^L_{(j)}} \hat{\beta}^*_k < \min_{j \in \mathcal{H}_{(j)}} \hat{\beta}^*_j \le \max_{j \in \mathcal{H}_{(j)}} \hat{\beta}^*_j < \min_{k \in \mathcal{H}^U_{(j)}} \hat{\beta}^*_k \right) = 1.$$

Under Assumption 15, by Lemma 3, for any $k \in \mathcal{H}^L_{(j)}$ and $m \in \mathcal{H}_{(j)}$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{\beta}^*_k < \hat{\beta}^*_m\right) = 1.$$

Similarly, for any $k \in \mathcal{H}^U_{(j)}$ and $m \in \mathcal{H}_{(j)}$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{\beta}^*_m < \hat{\beta}^*_k\right) = 1.$$

Step 2. We then show, for $k \notin \mathcal{H}_{(j)}$

$$\lim_{n\to\infty} \mathbb{P}\big(w_{k,(j)} > \varepsilon, k \notin \mathcal{H}_{(j)}\big) = 0.$$

For any $\varepsilon > 0$ and $k \notin \mathcal{H}_{(j)}$, we have the following holds

$$\begin{aligned}
\mathbb{P}\big(w_{k,(j)} > \varepsilon\big) &= \mathbb{P}\big(\mathbb{1}(k \in \hat{\mathcal{H}}_{(j)}) > \varepsilon\big) \\
&= \mathbb{P}\big(\mathbb{1}(k \in \hat{\mathcal{H}}_{(j)}) > \varepsilon | k \in \hat{\mathcal{H}}_{(j)}\big) \cdot \mathbb{P}\big(k \in \hat{\mathcal{H}}_{(j)}\big) \\
&\quad + \mathbb{P}\big(\mathbb{1}(k \in \hat{\mathcal{H}}_{(j)}) > \varepsilon | k \notin \hat{\mathcal{H}}_{(j)}\big) \cdot \mathbb{P}\big(k \notin \hat{\mathcal{H}}_{(j)}\big) \\
&\leq \mathbb{P}\big(k \in \hat{\mathcal{H}}_{(j)}\big) \\
&\overset{\text{Def}}{=} \mathbb{P}\Big(\frac{|\hat{\beta}^*_k - \hat{\beta}^*_{(j)}|}{n^{-\delta}} < C\Big) \\
&= \mathbb{P}\Big(\frac{|\hat{\beta}^*_k - \hat{\beta}^*_{\check{j}}|}{n^{-\delta}} < C\Big) \\
&= \mathbb{P}\Big(|\hat{\beta}^*_k - \hat{\beta}^*_{\check{j}}| < n^{-\delta} \cdot C\Big) \\
&= \mathbb{P}\Big(|(\hat{\beta}^*_k - \beta_k) - (\hat{\beta}^*_{\check{j}} - \beta_{\check{j}}) + (\beta_k - \beta_{\check{j}})| < n^{-\delta} \cdot C\Big) \\
&\leq \mathbb{P}\Big(n^\delta |\beta_k - \beta_{\check{j}}| - n^\delta |\hat{\beta}^*_k - \beta_k| - n^\delta |\hat{\beta}^*_{\check{j}} - \beta_{\check{j}}| < C\Big) \\
&= \mathbb{P}\Big(n^\delta |\hat{\beta}^*_k - \beta_k| + n^\delta |\hat{\beta}^*_{\check{j}} - \beta_{\check{j}}| > n^\delta |\beta_k - \beta_{\check{j}}| + C\Big)
\end{aligned}$$

By definition, for $k \notin \mathcal{H}_{(j)}$

$$\begin{aligned}
\mathbb{P}\big(n^\delta |\beta_k - \beta_{\check{j}}| < C\big) &\leq \mathbb{P}\big(n^\delta |\beta_k - \beta_{\check{j}}| < C, \check{j} \in \mathcal{H}_{(j)}\big) + \mathbb{P}\big(\check{j} \notin \mathcal{H}_{(j)}\big) \\
&\leq \max_{j \in \mathcal{H}_j} \mathbb{P}\big(n^\delta |\beta_k - \beta_j| < C, j \in \mathcal{H}_{(j)}\big) + \mathbb{P}\big(\check{j} \notin \mathcal{H}_{(j)}\big).
\end{aligned}$$

Under Assumption 15, Lemma 3 and the conclusion in Eq (2.14) in Step 1 suggest that by letting $n \to \infty$ on both side, we have the above probability converges to zero. Based on above derivation, we have shown that $\lim_{n\to\infty} \mathbb{P}\big(w_{k,(j)} > \varepsilon, k \notin \mathcal{H}_{(j)}\big) = 0$, for $k \notin \mathcal{H}_{(j)}$.

Step 3. We are left to prove that for all $k \in \mathcal{H}_{(j)}$, the following holds $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} \mathbb{P}\big(|w_{k,(j)} - 1| > \varepsilon\big) = 0.$$

Following similar arguments, for a positive constant C, we have for $k, j \in \mathcal{H}_{(j)}$ the following statement holds

$$\begin{aligned}
\mathbb{P}\big(|w_{k,(j)} - 1| > \varepsilon\big) &= \mathbb{P}\big(|\mathbb{1}(k \in \hat{\mathcal{H}}_{(j)}) - 1| > \varepsilon\big) \\
&= \mathbb{P}\big(|\mathbb{1}(k \in \hat{\mathcal{H}}_{(j)}) - 1| > \varepsilon | k \in \hat{\mathcal{H}}_{(j)}\big) \cdot \mathbb{P}\big(k \in \hat{\mathcal{H}}_{(j)}\big) \\
&\quad + \mathbb{P}\big(|\mathbb{1}(k \in \hat{\mathcal{H}}_{(j)}) - 1| > \varepsilon | k \notin \hat{\mathcal{H}}_{(j)}\big) \cdot \mathbb{P}\big(k \notin \hat{\mathcal{H}}_{(j)}\big) \\
&\leq \mathbb{P}\big(k \notin \hat{\mathcal{H}}_{(j)}\big) \\
&\overset{\text{Def}}{=} \mathbb{P}\Big(\frac{|\hat{\beta}^*_k - \hat{\beta}^*_{(j)}|}{n^{-\delta}} \geq C\Big) \\
&= \mathbb{P}\Big(\frac{|\hat{\beta}^*_k - \hat{\beta}^*_{\check{j}}|}{n^{-\delta}} \geq C\Big) \\
&= \mathbb{P}\Big(|(\hat{\beta}^*_k - \beta_k) - (\hat{\beta}^*_{\check{j}} - \beta_{\check{j}}) + (\beta_k - \beta_{\check{j}})| \geq n^{-\delta} \cdot C\Big) \\
&\leq \mathbb{P}\Big(|\hat{\beta}^*_k - \beta_k| + |\hat{\beta}^*_{\check{j}} - \beta_{\check{j}}| + |\beta_k - \beta_{\check{j}}| \geq n^{-\delta} \cdot C\Big) \\
&\leq \mathbb{P}\Big(n^{\frac{1}{2}}|\hat{\beta}^*_k - \beta_k| + n^{\frac{1}{2}}|\hat{\beta}^*_{\check{j}} - \beta_{\check{j}}| + n^{\frac{1}{2}}|\beta_k - \beta_{\check{j}}| \geq n^{\frac{1}{2}-\delta} \cdot C\Big).
\end{aligned}$$

By definition of the near-tie set, for $k \in \mathcal{H}_{(j)}$, we have

$$\begin{aligned}
\mathbb{P}\big(n^{\frac{1}{2}}|\beta_k - \beta_{\check{j}}| < C\big) &\leq \mathbb{P}\big(n^{\delta}|\beta_k - \beta_{\check{j}}| < C, \check{j} \in \mathcal{H}_{(j)}\big) + \mathbb{P}\big(\check{j} \notin \mathcal{H}_{(j)}\big) \\
&\leq \max_{j \in \mathcal{H}_j} \mathbb{P}\big(n^{\delta}|\beta_k - \beta_j| < C, j \in \mathcal{H}_{(j)}\big) + \mathbb{P}\big(\check{j} \notin \mathcal{H}_{(j)}\big).
\end{aligned}$$

Again, under Assumption 15, Lemma 3 and the conclusion in Eq (2.14) we have derived in Step 1, by letting $n \to \infty$ on both side, we have the above probability converges to 1.

□

Lemma 3

Lemma 3. *Under Assumption 15, we show that for all $k \in \{1, \ldots, d\}$, any positive constant C and $\delta \in (0, \frac{1}{2})$, the following statement holds*

$$\lim_{n\to\infty} \mathbb{P}\big(|\hat{\beta}^*_k - \beta_k| \geq n^{-\delta} \cdot C\big) = 0.$$

Proof. Note that

$$\sqrt{n}(\hat{\beta}_k^* - \beta_k) = \sqrt{n}(\hat{\beta}_k - \beta_k) + \mathcal{N}\big(0, (\hat{\mathbf{\Omega}}_n^{\text{KJ}})_{k,k}\big). \tag{2.15}$$

Because $\sqrt{n}(\hat{\beta}_k - \beta_k)$ converges in distribution to a finite-value random variable and $\hat{\mathbf{\Omega}}_n^{\text{KJ}}$ converges in probability to a finite-value matrix when n tends to infinity [89], for any given $\epsilon > 0$, there exists an M, such that

$$\mathbb{P}(\sqrt{n}|\hat{\beta}_k^* - \beta_k| > M) < \epsilon. \tag{2.16}$$

Then, for any n such that

$$n > \big(\frac{M}{C}\big)^{\frac{1}{\frac{1}{2}-\delta}}$$

we have that

$$\mathbb{P}\big(|\hat{\beta}_k^* - \beta_k| \geq n^{-\delta} \cdot C\big) < \epsilon$$

and therefore

$$\limsup_{n\to\infty} \mathbb{P}\big(|\hat{\beta}_k^* - \beta_k| \geq n^{-\delta} \cdot C\big) < \epsilon.$$

Note that the above inequality holds for arbitrary $\epsilon > 0$. Therefore, we have

$$\limsup_{n\to\infty} \mathbb{P}\big(|\hat{\beta}_k^* - \beta_k| \geq n^{-\delta} \cdot C\big) = 0,$$

completing the proof. □

Lemma 4. *Denote the selected policy as*

$$\hat{j} = \sum_{k=1}^{d} k \cdot \mathbb{1}(\hat{\beta}_k = \hat{\beta}_{(j)}),$$

we show that

$$\lim_{n\to\infty} \mathbb{P}\big(\hat{j} \in \mathcal{H}_{(j)}\big) = 1.$$

Proof. This is a direct result from Step 1 and Step 2 in the proof for Theorem 2. □

Corollary 1

Corollary 2. *Under Assumptions 11-15, we have that* $\lim_{n\to\infty} \mathbb{P}\Big(\mathbb{P}\big(\tilde{\beta}_{(j)}^* \leq \beta_{\hat{j}} | \{\boldsymbol{z}_{i,n}\}_{i=1}^n\big) \leq s\Big) = s.$

Proof. Because of the consistency in Lemma 3, we have

$$\lim_{n\to\infty} \mathbb{P}(\beta_{\hat{j}} = \beta_{(j)}) = 1.$$

Therefore,

$$\lim_{n\to\infty} |\mathbb{P}\big(\tilde{\beta}_{(j)}^* \leq \beta_{(j)} | \{\boldsymbol{z}_{i,n}\}_{i=1}^n\big) - \mathbb{P}\big(\tilde{\beta}_{(j)}^* \leq \beta_{(j)} | \{\boldsymbol{z}_{i,n}\}_{i=1}^n\big)| \leq \lim_{n\to\infty} \mathbb{P}(\beta_{(j)} \neq \hat{\beta}_{(j)} | \{\boldsymbol{z}_{i,n}\}_{i=1}^n) = 0.$$

The result then follows by applying Step 3 in the Proof of Theorem 2. □

Additional simulation and empirical results

Practical implementation

In this section, we discuss the choice of tuning parameters (including B, δ, b_{L} and b_{R}) of the proposed method in Section 2.2. For the number of repetitions for our resampling procedure, we recommend using $B = 2,000$ as a good balance between computational load and statistical inference accuracy.

For the tuning pair $(b_{\text{R}}, b_{\text{L}})$, from our theoretical analysis, we need to ensure that the distance between b_{R} and b_{L} is of the order n^{δ} with $\delta \in (0, 0.5)$ to guarantee the statistical validity of our proposed procedure. To achieve this goal, for the policy $\beta_{(j)}$, we adopt the tuning pair of the form

$$b_{\text{L}}^{j} = n^{-\delta} \cdot s_{\hat{j}}^{\delta} \cdot c_{\text{L}}^{j}, \quad b_{\text{R}}^{j} = n^{-\delta} \cdot s_{\hat{j}}^{\delta} \cdot c_{\text{R}}^{j},$$

where $s_{\hat{j}}$ is the $\hat{j}$th element in the diagonal of the estimated covariance matrix $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$, c_{L}^{j} and c_{R}^{j} are positive constants.

The constants c_{L}^{j} and c_{R}^{j} can significantly impact the performance of the proposed approach in finite samples. In the extreme cases, on the one hand, if both c_{L}^{j} and c_{R}^{j} are overly large, the estimated near tie set might include more policies than necessary and our approach is only valid if all true policy effects are closely ties. On the other hand, if c_{L}^{j} and c_{R}^{j} are both closer to zero, our approach reduces to a standard parametric bootstrap approach, which is problematic in the presence of tied policy effects. To present a robust algorithm in finite samples, we thus adopt the following "double-bootstrap" method as discussed in [35] (note that double-bootstrapped statistics are labelled with double-star superscripts):

1. For $j = 1, ..., d$, set $\beta_j^* = \Delta \cdot \frac{\sum_{j=1}^{d} \hat{\beta}_j}{d} + (1 - \Delta) \cdot \hat{\beta}_j$, where

$$\Delta = \min\Big\{1, \frac{\sum_{j=1}^{d} s_{\hat{j}}}{n \sum_{j=1}^{d} (\hat{\beta}_j - \bar{\hat{\beta}})^2} \times n^{0.05}\Big\}.$$

2. For every candidate pair $(c_{\text{L}}, c_{\text{R}})$ such that $c_{\text{L}} \in \mathcal{C}_{\text{L}}$ and $c_{\text{R}} \in \mathcal{C}_{\text{R}}$, do

   For $t \leftarrow 1$ to T, do

   i. Generate $\hat{\boldsymbol{\beta}}^* = (\hat{\beta}_1^*, ..., \hat{\beta}_d^*)$ from $\mathcal{N}(\boldsymbol{\beta}^*, \hat{\Omega}_n^{\text{KJ}}/n)$, where $\boldsymbol{\beta}^* = (\beta_1^*, \ldots, \beta_d^*)'$, and denote the ordered values in $\boldsymbol{\beta}^*$ as $\beta_{(1)}^* \geq \ldots \geq \beta_{(d)}^*$.

   ii. For $r \leftarrow 1$ to R, do

   A. Generate double bootstrap statistics $\hat{\boldsymbol{\beta}}^{**} \triangleq (\hat{\beta}_1^{**}, ..., \hat{\beta}_d^{**})'$ from $\mathcal{N}(\hat{\boldsymbol{\beta}}^*, \hat{\Omega}_n^{\text{KJ}}/n)$, and denote the ordered values of $\hat{\boldsymbol{\beta}}^{**}$ as $\hat{\beta}_{(1)}^{**} \geq ... \geq \hat{\beta}_{(d)}^{**}$.

   B. Record $w_{k,(j)}^{**} = \mathbb{1}\{-c_{\text{L}} \cdot n^{-\delta} \cdot s_{\hat{j}}^{\delta} \leq (\hat{\beta}_k^{**} - \beta_{(j)}^{**}) \leq c_{\text{R}} \cdot n^{-\delta} \cdot s_{\hat{j}}^{\delta}\}$ and $\tilde{\beta}_{(j)}^{**} = \sum_{k=1}^{d} w_{k,(j)}^{**} \hat{\beta}_j^{**} / \sum_{k=1}^{d} w_{k,(j)}^{**}$, for $j = 1, \ldots, d$.

iii. Calculate $B_{j,t}(c_{\mathrm{L}}, c_{\mathrm{R}}) = \frac{1}{R}\sum_{r=1}^{R} \mathbb{1}\big(\tilde{\beta}^{**,r}_{(j)} \leq \beta^{*,r}_{(j)}\big)$, for $j = 1, \ldots, d$.

3. Record the loss function

$$L_j(c_{\mathrm{L}}, c_{\mathrm{R}}) = \frac{1}{T}\sum_{t=1}^{T}\Big(B_{j,(t)}(c_{\mathrm{L}}, c_{\mathrm{R}}) - \frac{t}{T+1}\Big)^2, \tag{2.17}$$

where $B_{j,(t)}(c_{\mathrm{L}}, c_{\mathrm{R}})$ is the t-th smallest statistics in $B_{j,1}(c_{\mathrm{L}}, c_{\mathrm{R}}), \ldots, B_{j,T}(c_{\mathrm{L}}, c_{\mathrm{R}})$.

4. Choose the pair $(c^j_{\mathrm{L}}, c^j_{\mathrm{R}})$ for inferring $\beta_{(j)}$ and $\beta_{\hat{j}}$ that minimizes $L_j(c_{\mathrm{L}}, c_{\mathrm{R}})$, that is

$$(c^j_{\mathrm{L}}, c^j_{\mathrm{R}}) = \min_{(c_{\mathrm{L}}, c_{\mathrm{R}})\in\mathcal{C}} L_j(c_{\mathrm{L}}, c_{\mathrm{R}}), \quad j = 1, \ldots, d_0.$$

Note that we only use the above procedure to choose the tuning parameters c^j_{L} and c^j_{R}, meaning that we do not use the resampled statistics in Step 1 to carry out inference on $\beta_{(j)}$. In Step 1, Δ is adopted to stabilize the performance of the tuning parameter selection in finite samples, and Δ only takes a close-to-zero value whenever limited variation is found between policy effect estimates.

Following Theorem 2, we know that $\mathbb{P}\big(\tilde{\beta}^*_{(j)} \leq \beta_{(j)} | \{\boldsymbol{z}_{i,n}\}_{i=1}^n\big)$ roughly follows Unif$(0,1)$ when the sample size n is large. Given a desirable tuning pair $(c_{\mathrm{L}}, c_{\mathrm{R}})$, we would thus expect that $B_{j,(1)}(c_{\mathrm{L}}, c_{\mathrm{R}}), ..., B_{j,(T)}(c_{\mathrm{L}}, c_{\mathrm{R}})$ share a similar distribution with the ordered statistics of i.i.d. Unif$(0,1)$ random variables. The loss function defined in Eq (2.17) measures the average of squared differences between $B_{j,(t)}(c_{\mathrm{L}}, c_{\mathrm{R}})$ and the expected value of the order statistics of the Unif$(0,1)$ random variables. Given the rational above, we would expect that the optimal tuning pair $(c^j_{\mathrm{L}}, c^j_{\mathrm{R}})$ minimize such a loss.

We further comment on several implementation details. Our numerical results suggest using $R = 200$ and $T = 40$ can provide reasonable choice of the tuning parameters in finite samples. In addition, when the loss function $L_j(c_{\mathrm{L}}, c_{\mathrm{R}})$ do not fluctuate substantially over all considered pairs $(c_{\mathrm{L}}, c_{\mathrm{R}})$. In this case, let γ denote the 97.5th percentile of $L_U = \frac{1}{T}\sum_{t=1}^{T}(U_{(b)} - \frac{t}{T+1})^2$ and $U_{(1)}, ..., U_{(T)}$ are ordered observations from Unif$(0,1)$ distribution, we choose $(\bar{c}^j_L, \bar{c}^j_R)$ which is the mean of all plausible pairs such that $L_j(c_{\mathrm{L}}, c_{\mathrm{R}}) < \gamma$. Lastly, as for the candidate region of c_{L} and c_{R}, we first consider selecting c_L from 0 to $\frac{2(\hat{\beta}_{(1)} - \hat{\beta}_{(j)})n^\delta}{s^\delta_{\hat{j}}}$ and c_R from 0 to $\frac{2(\hat{\beta}_{(j)} - \hat{\beta}_{(d)})n^\delta}{s_{\hat{j}}\delta}$. Then based on the values of $L_j(c_{\mathrm{L}}, c_{\mathrm{R}})$ for different tuning pairs, we may choose to expand or shrink the candidate region to make our algorithm more efficient.

Robustness to different tuning choice

We summarizing our simulation results with different choices of the tuning parameter $\delta \in \{0.05, 0.15, 0.25\}$, $R \in \{200, 500\}$ and $T \in \{40, 100, 200\}$. To avoid redundancy, we showcase the results with $\boldsymbol{\beta} = 0$ and $\beta_j = \Phi^{-1}\big(\frac{j}{d+1}\big)$, $j = 1, \ldots, d$ while q_n takes value 141 or 561. We

report the coverage probabilities and asymptotic biases for estimating the top two policies (i.e., $d_0 = 2$) $\beta_{(1)}$ and $\beta_{(2)}$. Supplementary Materials Table 2.6 summarizes the simulation results with different choice of δ and fixed $R = 200$ and $T = 40$. There, we observe that the performance of our method is overall robust to the choice of different δ in a variety of settings. Though when q_n is large and no policy is effective, smaller δ likely leads to under-covered confidence intervals. Supplementary Materials Table 2.7 summarizes the simulation results under $R \in \{200, 500\}$ and $T \in \{40, 100, 200\}$, while fixing $\delta = 0.25$. Our results demonstrate that when R or T increases, the coverage probabilities are slightly increased and biases are marginally reduced. Overall, we observe that the proposed method is not very sensitive to the choice of various tuning parameters δ, T, and R. To guide readers for the selection of tuning parameters to reach an optimal accuracy and computational efficiency trade-off, we further provide the computational time under various choices of T and R in the Supplementary Materials Table 2.8. In practice, to reduce computational cost while maintaining valid statistical inference, we adopt the following tuning set in the rest of the numerical studies: $R = 200$, $T = 100$, and $\delta = 0.25$.

Table 2.6: Coverage probability and asymptotic bias with different choices of δ

		No policy is effective, $\beta_{(1)} = \beta_{(2)} = 0$					
		$q_n = 141$			$q_n = 561$		
		δ=0.05	δ=0.15	δ=0.25	$\delta = 0.05$	δ=0.15	δ=0.25
$\beta_{(1)}$	Cover	0.97(0.00)	0.95(0.01)	0.96(0.01)	0.93(0.01)	0.92(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	0.01(0.02)	0.03(0.04)	-0.02(0.02)	0.03(0.01)	0.03(0.01)	-0.01(0.01)
$\beta_{(2)}$	Cover	0.98(0.00)	0.94(0.01)	0.97(0.01)	0.93 (0.01)	0.93(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	0.00(0.01)	0.02(0.01)	0.01(0.01)	0.03(0.01)	0.02(0.01)	0.01(0.01)
		Top two policies are effective, $\beta_{(1)} = 0.97, \beta_{(2)} = 0.43$					
		$q_n = 141$			$q_n = 561$		
		δ=0.05	δ=0.15	δ=0.25	$\delta = 0.05$	δ=0.15	δ=0.25
$\beta_{(1)}$	Cover	0.94(0.01)	0.98(0.00)	0.95(0.01)	0.93(0.01)	0.95(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.05(0.05)	-0.03(0.04)	-0.04(0.05)	-0.07(0.06)	-0.06(0.07)	-0.07(0.07)
$\beta_{(2)}$	Cover	0.94(0.01)	0.97(0.01)	0.95(0.01)	0.95(0.01)	0.96(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.08(0.08)	0.05(0.06)	0.06(0.06)	0.08(0.08)	0.07(0.08)	0.08(0.09)

Note: We fix $R = 200$ and $T = 40$. "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(j)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(j)}$. Monte Carlo standard errors are provided in the parenthesis.

Computational time with different tuning parameters

In this section, we summarize computational time with respect to various choices of R, T, and n to make the computational costs transparent for readers. Here, we fix δ at 0.25 and select 20 candidate tuning parameters. Table 2.8 demonstrates that the computational

costs are largely determined by T and R. When both T and R reach 500, running our method once takes approximately one hour. For simulation study with multiple iterations, we recommend setting $R = 200$ and $T \leq 200$ to achieve a reasonable trade-off between accuracy and computational efficiency.

Table 2.7: Coverage probability and asymptotic bias with different choices of T and R

			No policy is effective, $\beta_{(1)} = \beta_{(2)} = 0$					
			$q_n = 141$			$q_n = 561$		
			$T = 40$	$T = 100$	$T = 200$	$T = 40$	$T = 100$	$T = 200$
$\beta_{(1)}$	$R = 200$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.96(0.02)	0.96(0.02)
		$\sqrt{n}$Bias	-0.02(0.02)	0.03(0.06)	0.02(0.06)	-0.01(0.01)	-0.01(0.12)	0.01(0.12)
	$R = 500$	Cover	0.96(0.01)	0.95(0.01)	0.96(0.01)	0.97(0.01)	0.96(0.01)	0.96(0.01)
		$\sqrt{n}$Bias	0.01(0.02)	-0.02(0.04)	0.01(0.02)	0.01(0.01)	0.01(0.01)	0.01(0.01)
$\beta_{(2)}$	$R = 200$	Cover	0.96(0.01)	0.96(0.01)	0.97(0.01)	0.96(0.01)	0.96(0.01)	0.97(0.01)
		$\sqrt{n}$Bias	0.01(0.01)	-0.01(0.01)	0.01(0.01)	0.01(0.01)	0.01(0.01)	-0.01(0.01)
	$R = 500$	Cover	0.95(0.01)	0.95(0.01)	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.96(0.01)
		$\sqrt{n}$Bias	0.01(0.01)	0.01(0.01)	0.01(0.01)	0.01(0.01)	0.01(0.01)	0.00(0.01)
			Top two policies are effective, $\beta_{(1)} = 0.97, \beta_{(2)} = 0.43$					
			$q_n = 141$			$q_n = 561$		
			$T = 40$	$T = 100$	$T = 200$	$T = 40$	$T = 100$	$T = 200$
$\beta_{(1)}$	$R = 200$	Cover	0.95(0.01)	0.95(0.01)	0.96(0.01)	0.94(0.01)	0.95(0.01)	0.95(0.01)
		$\sqrt{n}$Bias	-0.04(0.05)	0.02(0.08)	0.02(0.08)	-0.07(0.07)	0.07(0.16)	0.06(0.18)
	$R = 500$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.94(0.01)	0.95(0.01)	0.95(0.01)
		$\sqrt{n}$Bias	0.02(0.04)	0.01(0.04)	0.01(0.05)	0.05(0.07)	0.05(0.07)	0.04(0.07)
$\beta_{(2)}$	$R = 200$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.96(0.01)
		$\sqrt{n}$Bias	0.06(0.06)	0.05(0.06)	0.05(0.06)	0.08(0.09)	0.07(0.08)	0.07(0.09)
	$R = 500$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.96(0.01)	0.96(0.01)
		$\sqrt{n}$Bias	0.06(0.08)	0.05(0.06)	0.05(0.06)	0.07(0.08)	0.07(0.08)	0.06(0.09)

Note: We fix $\delta = 0.25$. "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(j)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(j)}$. Monte Carlo standard errors are provided in the parenthesis.

Simulation results: $\beta_{\hat{j}}$

The simulation results presented in Table 2.9 and 2.10 help us confirm our theoretical analyses in Corollary 1, and we observe similar trends compared to the results in the main manuscript.

Simulation results: $d = 10$

This section provides an additional set of simulation results when $d = 10$, which is larger than the setting ($d = 5$) adopted in the main manuscript. We investigate the performance

Computational time with respect to various n, q_n, T, and R

		$q_n = 141$ ($s \times 10^3$)			$q_n = 561$ ($s \times 10^3$)		
		$T = 40$	$T = 200$	$T = 500$	$T = 40$	$T = 200$	$T = 500$
$n = 500$	$R = 200$	0.10	0.48	1.27	0.11	0.48	1.35
	$R = 500$	0.23	1.20	3.31	0.24	1.38	3.50
$n = 2000$	$R = 200$	0.10	0.50	1.29	0.12	0.51	1.40
	$R = 500$	0.26	1.23	3.44	0.27	1.39	3.98
$n = 5000$	$R = 200$	0.11	0.52	1.32	0.14	0.55	1.51
	$R = 500$	0.27	1.25	3.50	0.30	1.40	4.05

Table 2.8: The unit: 1,000 seconds. We fix $\delta = 0.25$ and set 20 candidate tuning pairs for $(c_{\text{L}}, c_{\text{R}})$. The simulations are performed on a Lenovo NeXtScale nx360m5 node (24 cores per node) equipped with Intel Xeon Haswell processor. The core frequency is 2.3 Ghz and supports 16 floating-point operations per clock period.

of the five methods for estimating $\beta_{(2)}$, $\beta_{(5)}$, and $\beta_{(10)}$. Table 2.11 - 2.14 demonstrate that without adjustment, the coverage probabilities for $\beta_{(5)}$ fall below 80% when $q_n \geq 281$, while our proposed method reaches nominal level coverage regardless the ranking of β_j. "Proposed + EW", "Proposed + HCK", and "Proposed + HC3" show similar trends compared to those when $d = 5$.

Simulation results: realistic error terms

In this section, we consider two DGPs of generating more practical errors beyond simple i.i.d. Gaussian noises. For the first DGP, we generate covariates from $\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma)$ and $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0, \boldsymbol{I}_{q_n})$, and then generate random noise from (1) an asymmetric distribution with the density function $0.5\phi(\varepsilon|-0.5, 0.25) + 0.5\phi(\varepsilon|0.5, 1)$; (2) a bimodal distribution with the density function $0.5\phi(\varepsilon|-1.5, 0.25) + 0.5\phi(\varepsilon|1.5, 1)$, where $\phi(\varepsilon|\mu, \sigma^2)$ denotes the density function of a normal random variable with mean μ and variance σ^2. The simulation results are summarized in Supplementary Materials Table 2.16. We further study this setting with a larger sample size, $n = 2,000$. This sample size is closer to the sample size adopted in our case study I. The simulation results under $n = 2,000$ are summarized in Supplementary Materials Table 2.18. We also consider the design with both $\beta \neq 0$ and $\gamma \neq 0$. The simulation results are summarized in Supplementary Materials Table 2.15–2.18.

For the second DGP, we consider heteroscedastic errors following the setup in [25] with: $\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$ with $\tilde{\boldsymbol{w}}_{i,n} \sim \mathcal{N}(0, \boldsymbol{I}_{q_n})$, and $\varepsilon_{i,n} \sim \mathcal{N}(0, 1)$ with $\mathbb{V}[\varepsilon_{i,n}|\boldsymbol{x}_{i,n}, \boldsymbol{w}_{i,n}] = c_\varepsilon(1+(t(x_{1,i,n})+\boldsymbol{l}'\boldsymbol{w}_{i,n})^2/4)$ and $\mathbb{V}[x_{k,i,n}|\boldsymbol{w}_{i,n}] = c_{x_k}(1+(\boldsymbol{l}'\boldsymbol{w}_{i,n})^2/4)$, where $x_{k,i,n}$ denotes the kth component of the vector $\boldsymbol{x}_{i,n}$. The constants c_ε and c_{x_k} are chosen so that $\mathbb{V}[\varepsilon_{i,n}] = \mathbb{V}[x_{k,i,n}] = 1$ and $t(a) = a\mathbb{1}(-1 \leq a \leq 1) + \text{sgn}(a)(1 - \mathbb{1}(-1 \leq a \leq 1))$. $\boldsymbol{l}$ is the conformable vector of ones. The simulation results are summarized in Supplementary Materials Table 2.15.

Table 2.9: Simulation results ($d = 5$, heterogeneity, $\beta_{\hat{1}}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1}), \quad \gamma_n = 0, \quad j = 1, \ldots, d$				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma), \quad \boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.96(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.04(0.06)	0.04(0.05)	0.04(0.05)	-0.04(0.05)	0.04(0.06)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.93(0.01)	0.95(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.07(0.07)	0.06(0.07)	0.07(0.06)	0.06(0.06)	0.05(0.07)
$q_n = 281$	Cover	0.94(0.01)	0.95(0.01)	0.84(0.01)	0.82(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.09(0.08)	-0.07(0.07)	-0.10(0.08)	-0.11(0.07)	-0.07(0.08)
$q_n = 421$	Cover	0.94(0.01)	0.91(0.01)	0.76(0.02)	0.75(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	-0.09(0.10)	-0.10(0.09)	-0.15(0.09)	-0.16(0.09)	-0.10(0.09)
$q_n = 561$	Cover	0.94(0.01)	0.90(0.01)	0.67(0.02)	0.65(0.01)	0.78(0.01)
	$\sqrt{n}$Bias	-0.15(0.14)	-0.12(0.10)	-0.17(0.23)	-0.25(0.12)	0.15(0.11)
$q_n = 631^*$	Cover	0.92(0.01)	0.89(0.02)	0.45(0.02)	0.42(0.01)	0.68(0.01)
	$\sqrt{n}$Bias	-0.19(0.18)	-0.22(0.13)	-0.35(0.29)	-0.54(0.22)	0.28(0.18)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0), \quad \boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.97(0.01)	0.97(0.01)	0.95(0.01)	0.95(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	-0.05(0.10)	-0.06(0.09)	-0.06(0.10)	-0.10(0.11)	-0.04(0.09)
$q_n = 141$	Cover	0.97(0.01)	0.95(0.01)	0.94(0.01)	0.93(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.06(0.11)	-0.08(0.12)	-0.07(0.11)	0.13(0.12)	0.09(0.12)
$q_n = 281$	Cover	0.96(0.01)	0.94(0.01)	0.86(0.02)	0.85(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.09(0.13)	-0.10(0.13)	-0.10(0.13)	-0.15(0.12)	-0.09(0.13)
$q_n = 421$	Cover	0.94(0.01)	0.93(0.01)	0.75(0.02)	0.72(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	0.11(0.17)	-0.12(0.13)	0.18(0.17)	-0.20(0.17)	0.14(0.14)
$q_n = 561$	Cover	0.94(0.01)	0.90(0.01)	0.51(0.02)	0.48(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	-0.15(0.22)	-0.21(0.20)	-0.25(0.23)	-0.46(0.26)	-0.21(0.20)
$q_n = 631$	Cover	0.91(0.01)	0.90(0.01)	0.48(0.02)	0.45(0.01)	0.80(0.01)
	$\sqrt{n}$Bias	-0.21(0.20)	-0.23(0.22)	0.41(0.30)	-0.53(0.20)	0.35(0.22)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{\hat{1}}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{\hat{1}}$. " * " indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Table 2.15 shows that, under the second DGP, our proposed method has slightly compromised performance, but still reaches nominal level coverage when $q_n \leq 421$. Table 2.16 and 2.17 demonstrate that the performance of our method is robust even when both $\beta \neq 0$ and $\gamma \neq 0$, and $q_n \leq 561$. Table 2.18 suggests that when sample size increases, our proposed method has smaller bias and improved coverage probabilities when $q_n = 631$. The other considered methods show similar trends to the settings under homoscedastic errors.

Additional analysis for case study I

In this section, we revisit case study I with a much smaller model that only includes the main effects. The results are summarized in Table 2.19. Table 2.19 shows that, overall, the results under a smaller model do not change substantively. But "asking 25% more" no longer has a significant impact on donation amount even without calibration.

Table 2.10: Simulation results ($d = 5$, heterogeneity, $\beta_{\hat{2}}$)

		$\boldsymbol{\beta}_j = \Phi^{-1}\left(\frac{j}{d+1}\right), \quad \boldsymbol{\gamma}_n = 0, \quad j = 1, \ldots, d$				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma), \quad \boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.97(0.01)	0.97(0.01)	0.96(0.01)	0.97(0.01)
	$\sqrt{n}$Bias	-0.04(0.06)	-0.03(0.06)	-0.03(0.06)	-0.04(0.07)	-0.04(0.06)
$q_n = 141$	Cover	0.96(0.01)	0.96(0.02)	0.88(0.02)	0.90(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.05(0.08)	-0.05(0.08)	-0.10(0.08)	-0.09(0.06)	-0.07(0.07)
$q_n = 281$	Cover	0.95(0.01)	0.94(0.02)	0.86(0.02)	0.84(0.01)	0.91(0.02)
	$\sqrt{n}$Bias	0.07(0.09)	0.07(0.09)	0.13(0.08)	-0.15(0.10)	0.12(0.10)
$q_n = 421$	Cover	0.94(0.01)	0.93(0.02)	0.77(0.02)	0.72(0.02)	0.71(0.02)
	$\sqrt{n}$Bias	-0.10(0.13)	-0.12(0.13)	-0.15(0.11)	-0.17(0.13)	-0.19(0.17)
$q_n = 561$	Cover	0.94(0.01)	0.92(0.02)	0.65(0.02)	0.60(0.01)	0.69(0.02)
	$\sqrt{n}$Bias	-0.16(0.17)	-0.18(0.16)	-0.18(0.15)	0.20(0.13)	0.35(0.22)
$q_n = 631^*$	Cover	0.93(0.01)	0.92(0.02)	0.44(0.02)	0.42(0.01)	0.50(0.02)
	$\sqrt{n}$Bias	-0.18(0.17)	-0.23(0.21)	-0.45(0.19)	-0.49(0.17)	0.48(0.30)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0), \quad \boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.98(0.01)	0.96(0.01)	0.96(0.01)	0.97(0.01)
	$\sqrt{n}$Bias	-0.08(0.13)	-0.07(0.13)	-0.10(0.12)	0.09(0.12)	-0.10(0.10)
$q_n = 141$	Cover	0.96(0.01)	0.96(0.01)	0.94(0.01)	0.95(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.09(0.14)	0.10(0.14)	0.13(0.14)	0.12(0.13)	0.10(0.14)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.92(0.01)	0.90(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.11(0.16)	-0.11(0.15)	0.17(0.15)	0.19(0.14)	0.13(0.17)
$q_n = 421$	Cover	0.95(0.01)	0.92(0.02)	0.82(0.02)	0.78(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.18(0.20)	-0.20(0.20)	-0.25(0.20)	-0.32(0.17)	-0.20(0.20)
$q_n = 561$	Cover	0.94(0.01)	0.92(0.02)	0.67(0.02)	0.64(0.01)	0.73(0.01)
	$\sqrt{n}$Bias	-0.26(0.24)	0.22(0.24)	-0.41(0.23)	-0.48(0.23)	0.27(0.25)
$q_n = 631$	Cover	0.92(0.01)	0.87(0.02)	0.52(0.02)	0.48(0.01)	0.50(0.01)
	$\sqrt{n}$Bias	0.30(0.27)	0.41(0.29)	0.50(0.28)	-0.65(0.13)	0.55(0.28)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{\hat{2}}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{\hat{2}}$. " * " indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Extension to regression models with fixed effects

As stated in the main manuscript, our approach extends to linear panel data models with fixed effects. We shall briefly discuss this connection below. Because it is a common practice to include the subscript t to denote time in panel data analyses, to avoid using triple subscript, we drop subscript n in all considered random variables in the discussion below.

Suppose we have access to one panel data with cross-sectional observations denoted by $i \in \mathcal{N} = \{1, \ldots, N\}$ and time periods $t \in \mathcal{T} = \{1, \ldots, T\}$. Consider the following fixed effects panel data model

$$y_{it} = \boldsymbol{\beta}' \boldsymbol{x}_{it} + c_i + e_{d_{it}} + u_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T,$$

where c_i is an unobserved effect that varies across sections but is assumed to be constant over time, $y_{it} \in \mathbb{R}$ is the observed outcome, $\boldsymbol{x}_{it} \in \mathbb{R}^{d \times 1}$ contains the policy variables of interest,

Table 2.11: Simulation results ($d = 10$, heterogeneity, $\beta_{(2)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1})$, $\gamma_n = 0$, j = 1, ..., d				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.97(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.05(0.07)	-0.04(0.07)	-0.05(0.06)	-0.06(0.07)	0.06(0.06)
$q_n = 141$	Cover	0.96(0.01)	0.95(0.01)	0.96(0.01)	0.96(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.06(0.07)	-0.07(0.07)	0.07(0.08)	-0.07(0.07)	-0.06(0.06)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.92(0.01)	0.89(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	-0.07(0.09)	-0.08(0.09)	0.15(0.10)	-0.18(0.08)	0.20(0.08)
$q_n = 421$	Cover	0.96(0.01)	0.94(0.01)	0.88(0.01)	0.80(0.02)	0.90(0.01)
	$\sqrt{n}$Bias	-0.08(0.11)	-0.10(0.11)	0.20(0.11)	-0.25(0.10)	-0.25(0.09)
$q_n = 561$	Cover	0.95(0.01)	0.94(0.01)	0.75(0.01)	0.67(0.02)	0.89(0.01)
	$\sqrt{n}$Bias	0.14(0.15)	0.15(0.15)	-0.27(0.16)	-0.31(0.13)	-0.28(0.12)
$q_n = 631^*$	Cover	0.93(0.01)	0.91(0.01)	0.65(0.01)	0.57(0.02)	0.88(0.01)
	$\sqrt{n}$Bias	0.22(0.20)	0.27(0.19)	-0.33(0.20)	0.35(0.17)	0.38(0.16)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0)$, $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.97(0.01)	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.05(0.11)	-0.07(0.11)	0.06(0.12)	-0.08(0.11)	-0.09(0.09)
$q_n = 141$	Cover	0.96(0.01)	0.95(0.01)	0.94(0.01)	0.93(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.07(0.12)	0.10(0.13)	0.13(0.13)	0.15(0.13)	0.10(0.11)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.91(0.01)	0.88(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	0.14(0.15)	0.16(0.16)	0.20(0.13)	-0.28(0.15)	-0.29(0.13)
$q_n = 421$	Cover	0.94(0.01)	0.93 (0.01)	0.85(0.01)	0.76(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	0.20(0.19)	0.22(0.18)	0.25(0.17)	0.37(0.18)	0.32(0.16)
$q_n = 561$	Cover	0.93(0.01)	0.93(0.01)	0.77(0.01)	0.67(0.02)	0.90(0.01)
	$\sqrt{n}$Bias	0.22(0.23)	0.25(0.23)	-0.30(0.19)	0.41(0.22)	0.44(0.19)
$q_n = 631$	Cover	0.90(0.01)	0.88(0.01)	0.67(0.01)	0.60(0.02)	0.88(0.01)
	$\sqrt{n}$Bias	0.31(0.25)	0.45(0.26)	-0.48(0.13)	0.55(0.25)	0.61(0.23)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(2)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(2)}$. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

and error terms u_{it}'s are uncorrelated conditional on $\boldsymbol{x}_{it}$ and d_{it}. $e_{d_{it}}$ is an unobserved effect indexed by an observed indexing variable $d_{it} \in \{1, \ldots, G\}$, and is assumed to be constant across all observations that share the same value of d_{it}. When $e_{d_{it}} = 0$, this model reduces to the one-way fixed effects model studied in [153], otherwise the above model coincides with the one studied in [164].

To concretely introduce the connection of the above model and our model setup, consider the case when $e_{d_{it}} \neq 0$, we stack the data over cross-sectional observations and time periods.

Table 2.12: Simulation results ($d = 10$, homogeneity, $\beta_{(5)}$)

		$\boldsymbol{\beta} = 0, \quad \boldsymbol{\gamma}_j = 1/j, \quad j = 1, \ldots, q_n$				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma), \quad \boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.93(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	0.01(0.01)	0.01(0.01)	0.01(0.01)	0.01(0.01)	0.04(0.02)
$q_n = 141$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.93(0.01)	0.82(0.01)
	$\sqrt{n}$Bias	-0.01(0.01)	-0.01(0.01)	0.02(0.02)	0.02(0.01)	-0.17(0.02)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.92(0.01)	0.80(0.01)
	$\sqrt{n}$Bias	0.02(0.02)	0.02(0.02)	0.02(0.02)	0.05(0.02)	-0.18(0.02)
$q_n = 421$	Cover	0.93(0.02)	0.93(0.02)	0.93(0.01)	0.80(0.02)	0.78(0.01)
	$\sqrt{n}$Bias	-0.02(0.02)	-0.02(0.02)	0.05(0.04)	0.17(0.03)	0.21(0.03)
$q_n = 561$	Cover	0.93(0.02)	0.92(0.02)	0.93(0.01)	0.75(0.02)	0.76(0.01)
	$\sqrt{n}$Bias	0.02(0.02)	-0.03(0.02)	0.05(0.04)	0.32(0.04)	-0.28(0.04)
$q_n = 631^*$	Cover	0.91(0.02)	0.90(0.01)	0.90(0.01)	0.73(0.02)	0.75(0.01)
	$\sqrt{n}$Bias	-0.12(0.05)	-0.30(0.05)	0.14(0.09)	0.36(0.06)	-0.34(0.05)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0), \quad \boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed + EW	No adjustment+KJ
$q_n = 1$	Cover	0.95(0.01)	0.95(0.02)	0.94(0.01)	0.93(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	0.01(0.01)	0.01(0.01)	-0.02(0.02)	0.01(0.01)	0.06(0.03)
$q_n = 141$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.93(0.01)	0.81(0.01)
	$\sqrt{n}$Bias	-0.01(0.01)	-0.01(0.01)	-0.02(0.02)	0.01(0.00)	0.19(0.03)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.92(0.01)	0.78(0.01)
	$\sqrt{n}$Bias	0.02(0.02)	0.02(0.02)	-0.02(0.03)	0.06(0.02)	-0.21(0.04)
$q_n = 421$	Cover	0.93(0.01)	0.93 (0.01)	0.92(0.02)	0.80(0.01)	0.75(0.01)
	$\sqrt{n}$Bias	-0.01(0.00)	-0.01(0.00)	-0.04(0.03)	0.17(0.03)	0.37(0.05)
$q_n = 561$	Cover	0.93(0.01)	0.93(0.01)	0.92(0.01)	0.75(0.02)	0.70(0.01)
	$\sqrt{n}$Bias	0.03(0.02)	0.03(0.02)	-0.08(0.06)	0.18(0.04)	-0.48(0.05)
$q_n = 631$	Cover	0.92(0.01)	0.92(0.01)	0.91(0.01)	0.75(0.02)	0.66(0.01)
	$\sqrt{n}$Bias	-0.12(0.05)	-0.30(0.05)	-0.32(0.11)	0.36(0.06)	-0.65(0.09)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(2)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(2)}$. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Define

$$
\begin{aligned}
\mathbf{y} =& (y_{11}, \ldots, y_{1T}, y_{21}, \ldots, y_{2T}, \ldots, y_{N1}, \ldots, y_{NT})' \in \mathbb{R}^{NT \times 1}, \\
\boldsymbol{x} =& (\boldsymbol{x}_{11}, \ldots, \boldsymbol{x}_{1T}, \boldsymbol{x}_{21}, \ldots, \boldsymbol{x}_{2T}, \ldots, \boldsymbol{x}_{N1}, \ldots, \boldsymbol{x}_{NT})' \in \mathbb{R}^{NT \times d}, \\
\boldsymbol{w} =& \Big(\mathbf{g}_1, \mathbf{g}_2\Big) \in \mathbb{R}^{NT \times (N+G)}, \\
\mathbf{g}_1 =& \big(\mathbf{1}_{(i=j)}\big)_{(i,t) \in \mathcal{N} \times \mathcal{T}}^{j \in \mathcal{N}}, \quad \mathbf{g}_2 = \big(\mathbf{1}_{(d_{it}=d)}\big)_{(i,t) \in \mathcal{N} \times \mathcal{T}}^{d \in \{1, \ldots, G\}}, \\
\boldsymbol{\gamma}_n =& (c_1, \ldots, c_N, e_1, \ldots, e_G)' \in \mathbb{R}^{(N+G) \times 1}, \\
\mathbf{u} =& (u_{11}, \ldots, u_{1T}, u_{21}, \ldots, u_{2T}, \ldots, u_{N1}, \ldots, u_{NT})' \in \mathbb{R}^{NT \times 1}.
\end{aligned}
$$

With the above notations, the fixed effects panel data model can be written as the following

$$
\mathbf{y} = \boldsymbol{x}\boldsymbol{\beta} + \boldsymbol{w}\boldsymbol{\gamma}_n + \mathbf{u}. \tag{2.18}
$$

Table 2.13: Simulation results ($d = 10$, heterogeneity, $\beta_{(5)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1})$, $\gamma_n = 0$, j = 1, ..., d				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.95(0.01)	0.92(0.02)	0.96(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.06(0.07)	0.07(0.07)	-0.12(0.09)	0.07(0.07)	0.06(0.06)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.92(0.01)	0.92(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.08(0.08)	0.08(0.08)	-0.12(0.08)	0.12(0.08)	-0.07(0.07)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.91(0.01)	0.91(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	0.08(0.08)	0.08(0.08)	-0.13(0.09)	0.12(0.08)	0.10(0.07)
$q_n = 421$	Cover	0.95(0.01)	0.95(0.01)	0.84(0.02)	0.82(0.02)	0.93(0.01)
	$\sqrt{n}$Bias	0.10(0.11)	0.11(0.11)	-0.13(0.07)	0.14(0.11)	-0.12(0.09)
$q_n = 561$	Cover	0.94(0.01)	0.94(0.01)	0.80(0.01)	0.79(0.02)	0.92(0.01)
	$\sqrt{n}$Bias	0.12(0.13)	0.13(0.13)	-0.18(0.06)	0.20(0.12)	-0.15(0.10)
$q_n = 631^*$	Cover	0.93(0.01)	0.92(0.01)	0.78(0.01)	0.76(0.01)	0.90(0.01)
	$\sqrt{n}$Bias	0.16(0.14)	0.18(0.14)	-0.28(0.26)	0.24(0.13)	0.18(0.12)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0)$, $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.96(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	0.05(0.11)	0.06(0.11)	-0.08(0.11)	0.08(0.11)	-0.06(0.09)
$q_n = 141$	Cover	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.95(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.06(0.11)	0.08(0.11)	-0.11(0.12)	0.10(0.11)	0.09(0.10)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.93(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.12(0.13)	0.12(0.13)	0.10(0.10)	0.14(0.12)	-0.10(0.11)
$q_n = 421$	Cover	0.95(0.01)	0.95(0.01)	0.91(0.01)	0.90(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.11(0.14)	0.15(0.15)	-0.20(0.08)	0.25(0.14)	-0.12(0.12)
$q_n = 561$	Cover	0.94(0.01)	0.94(0.01)	0.85(0.01)	0.83(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	-0.15(0.18)	-0.16(0.17)	-0.26(0.20)	0.28(0.18)	-0.15(0.15)
$q_n = 631$	Cover	0.93(0.01)	0.92(0.01)	0.75(0.01)	0.72(0.02)	0.93(0.01)
	$\sqrt{n}$Bias	0.25(0.20)	0.30(0.19)	-0.70(0.34)	0.68(0.20)	0.28(0.18)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(2)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(2)}$. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

This indicates that our approach also goes through in linear panel data models, as long as we can construct an estimator of $\boldsymbol{\beta}$ that converges to a Gaussian distribution with its covariance matrix being consistently estimated.

[89] has shown that the covariance matrix estimator $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ remains consistent in one-way fixed effect panel data regression models when $e_{d_{it}} = 0$. This suggests that our approach can be naturally extended to make inference on multiple best policies in one-way fixed effect models. In addition, [25] have shown that the covariance matrix estimator $\hat{\boldsymbol{\Omega}}_n^{\text{HCK}}$ is consistent in both one-way and two-way fixed effect panel data regression models. Since our resampling based approach only requires a consistent covariance matrix estimator to calibrate multiple best policy effects, this suggests that in two-way fixed effect models, what our approach can be adopted when using $\hat{\boldsymbol{\Omega}}_n^{\text{HCK}}$ to estimate the covariance matrix of $\hat{\boldsymbol{\beta}}$. Lastly, we note that our Assumption 1 in the main manuscript requires error terms to be conditionally uncorrelated within each observation i. This condition does rule out dynamic models as those discussed

Table 2.14: Simulation results ($d = 10$, heterogeneity, $\beta_{(10)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1})$, $\gamma_n = 0$, j = 1, ..., d				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.94(0.01)	0.95(0.01)	0.94(0.01)	0.94(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.06(0.07)	0.03(0.06)	0.06(0.06)	0.07(0.07)	-0.04(0.05)
$q_n = 141$	Cover	0.93(0.02)	0.93(0.02)	0.92(0.02)	0.92(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	0.06(0.06)	0.06(0.06)	0.08(0.07)	0.08(0.06)	0.06(0.06)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.90(0.01)	0.89(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.07(0.08)	0.08(0.08)	0.11(0.08)	0.10(0.07)	0.07(0.07)
$q_n = 421$	Cover	0.94(0.01)	0.94(0.01)	0.84(0.01)	0.82(0.02)	0.93(0.01)
	$\sqrt{n}$Bias	0.08(0.09)	0.09(0.09)	0.16(0.10)	0.17(0.09)	-0.11(0.08)
$q_n = 561$	Cover	0.93(0.01)	0.91(0.01)	0.68(0.02)	0.61(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	0.16(0.14)	0.18(0.14)	0.20(0.10)	0.24(0.13)	0.15(0.12)
$q_n = 631^*$	Cover	0.92(0.01)	0.90(0.01)	0.53(0.02)	0.50(0.02)	0.82(0.01)
	$\sqrt{n}$Bias	0.20(0.18)	0.23(0.18)	-0.50(0.12)	0.55(0.18)	-0.30(0.17)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0)$, $\boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.96(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	0.08(0.11)	0.08(0.11)	0.11(0.12)	0.09(0.11)	-0.08(0.09)
$q_n = 141$	Cover	0.94(0.01)	0.94(0.01)	0.93(0.01)	0.92(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.10(0.12)	0.11(0.12)	0.14(0.13)	0.15(0.12)	0.11(0.11)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.88(0.01)	0.84(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.11(0.15)	0.12(0.15)	0.16(0.14)	0.19(0.15)	0.12(0.13)
$q_n = 421$	Cover	0.94(0.01)	0.94(0.01)	0.85(0.01)	0.82(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.14(0.16)	0.16(0.16)	0.20(0.16)	0.25(0.16)	0.15(0.15)
$q_n = 561$	Cover	0.93(0.02)	0.93(0.01)	0.65(0.01)	0.62(0.02)	0.92(0.01)
	$\sqrt{n}$Bias	0.19(0.20)	0.23(0.21)	-0.24(0.15)	0.30(0.21)	-0.24(0.20)
$q_n = 631$	Cover	0.94(0.01)	0.92(0.01)	0.57(0.01)	0.50(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	0.26(0.29)	0.30(0.28)	-0.85(0.23)	0.88(0.31)	-0.33(0.29)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(10)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(10)}$. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

in [164].

Table 2.15: Simulation results ($d = 5,$ heteroscedasticity, heterogeneity, $\beta_{(1)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1}), \quad \boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma), \quad \boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98)), \quad \text{j} = 1, \ldots, \text{d}$				
		$\boldsymbol{\gamma}_n = 0$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.06(0.07)	0.07(0.07)	0.07(0.07)	0.07(0.07)	0.07(0.07)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.88(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	-0.12(0.12)	-0.12(0.12)	-0.12(0.12)	-0.16(0.07)	-0.12(0.12)
$q_n = 281$	Cover	0.95(0.01)	0.93(0.02)	0.92(0.01)	0.79(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.11(0.13)	-0.13(0.13)	-0.33(0.13)	-0.35(0.12)	-0.12(0.12)
$q_n = 421$	Cover	0.94(0.01)	0.93(0.01)	0.89(0.01)	0.73(0.02)	0.93(0.01)
	$\sqrt{n}$Bias	-0.14(0.15)	-0.18(0.15)	-0.44(0.15)	-0.52(0.15)	-0.19(0.15)
$q_n = 561$	Cover	0.93(0.01)	0.91(0.01)	0.82(0.02)	0.66(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	-0.22(0.19)	-0.32(0.19)	-0.67(0.18)	-0.71(0.19)	-0.33(0.18)
$q_n = 631^*$	Cover	0.92(0.01)	0.90(0.01)	0.76(0.02)	0.56(0.02)	0.90(0.01)
	$\sqrt{n}$Bias	-0.34(0.24)	-0.50(0.26)	-0.73(0.30)	-0.80(0.24)	-0.46(0.24)
		$\gamma_k = 1/k, \quad k = 1, \ldots, q_n$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.08(0.09)	0.08(0.09)	0.09(0.09)	0.09(0.09)	0.09(0.09)
$q_n = 141$	Cover	0.95(0.01)	0.94(0.01)	0.92(0.01)	0.87(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.12(0.16)	0.14(0.16)	0.17(0.14)	0.36(0.14)	0.14(0.14)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.88(0.01)	0.82(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	-0.13(0.14)	0.14(0.14)	-0.33(0.11)	-0.45(0.13)	0.15(0.12)
$q_n = 421$	Cover	0.93(0.01)	0.92(0.01)	0.80(0.01)	0.73(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	-0.21(0.18)	-0.24(0.18)	-0.40(0.11)	-0.56(0.11)	-0.24(0.13)
$q_n = 561^*$	Cover	0.92(0.01)	0.91(0.01)	0.67(0.01)	0.53(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	-0.28(0.22)	-0.35(0.21)	-0.47(0.17)	-0.51(0.19)	-0.37(0.20)
$q_n = 631^*$	Cover	0.91(0.01)	0.89(0.01)	0.59(0.01)	0.51(0.02)	0.88(0.01)
	$\sqrt{n}$Bias	-0.29(0.25)	-0.33(0.26)	-0.55(0.21)	-0.61(0.18)	-0.42 (0.22)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(1)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(1)}$. In this heteroscedastic design, $\tilde{\boldsymbol{w}}_{i,n} \sim \mathcal{N}(0, \boldsymbol{I}_{q_n})$, $\varepsilon_{i,n} \sim \mathcal{N}(0,1)$, $\mathbb{V}[\varepsilon_{i,n}|\boldsymbol{x}_{i,n}, \boldsymbol{w}_{i,n}] = c_\varepsilon(1+(t(x_{1,i,n})+\boldsymbol{l}'\boldsymbol{w}_{i,n})^2/4)$, and $\mathbb{V}[x_{k,i,n}|\boldsymbol{w}_{i,n}] = c_{x_k}(1+(\boldsymbol{l}'\boldsymbol{w}_{i,n})^2/4)$, where $x_{k,i,n}$ denotes the kth component of the vector $\boldsymbol{x}_{i,n}$. The constants c_ε and c_{x_k} are chosen so that $\mathbb{V}[\varepsilon_{i,n}] = \mathbb{V}[x_{k,i,n}] = 1$ and $t(a) = a\mathbb{1}(-1 \leq a \leq 1) + sgn(a)(1 - \mathbb{1}(-1 \leq a \leq 1))$. $\boldsymbol{l}$ is the conformable vector of ones. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Table 2.16: Simulation results ($d = 5$, heteroscedasticity, heterogeneity, $\beta_{(1)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1}), \quad j = 1, \ldots, d, \quad \boldsymbol{\gamma}_k = 1/k, \quad k = 1, \ldots, q_n$				
		$\boldsymbol{x}_{i,n} \sim N(0, \Sigma), \quad \boldsymbol{w}_{i,n} \sim N(0, I_{q_n}), \quad \varepsilon_i \sim f(\varepsilon) = 0.5\phi(\varepsilon \mid -0.5, 0.25) + 0.5\phi(\varepsilon \mid 0.5, 1)$				
		Proposed+KJ	Proposed+HCK	Proposed + HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.02(0.05)	-0.02(0.05)	-0.02(0.05)	-0.02(0.05)	0.05(0.05)
$q_n = 141$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.92(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.05(0.06)	-0.05(0.06)	-0.05(0.06)	-0.07(0.06)	0.06(0.06)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.93(0.01)	0.89(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	0.06(0.06)	0.06(0.06)	-0.10(0.07)	0.14(0.06)	0.07(0.07)
$q_n = 421$	Cover	0.94(0.01)	0.93(0.01)	0.92(0.01)	0.73(0.02)	0.92(0.01)
	$\sqrt{n}$Bias	-0.09(0.09)	-0.12(0.09)	-0.14(0.09)	0.17(0.09)	-0.10(0.08)
$q_n = 561^*$	Cover	0.93(0.02)	0.92(0.01)	0.92(0.01)	0.58(0.02)	0.92(0.01)
	$\sqrt{n}$Bias	-0.12(0.13)	-0.15(0.13)	-0.15(0.12)	-0.20(0.13)	0.14(0.12)
$q_n = 631^*$	Cover	0.94(0.01)	0.91(0.01)	0.90(0.01)	0.44(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	0.17(0.17)	0.18(0.16)	0.17(0.13)	0.25(0.17)	0.19(0.17)
		$\boldsymbol{x}_{i,n} \sim N(0, \Sigma), \quad \boldsymbol{w}_{i,n} \sim N(0, I_{q_n}), \quad \varepsilon_i \sim f(\varepsilon) = 0.5\phi(\varepsilon \mid -1.5, 0.25) + 0.5\phi(\varepsilon \mid 1.5, 1)$				
		Proposed+KJ	Proposed+HCK	Proposed + HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.07(0.10)	-0.08(0.10)	-0.09(0.10)	-0.08(0.10)	-0.08(0.09)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.93(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.10(0.12)	-0.10(0.12)	-0.12(0.12)	-0.13(0.12)	0.12(0.10)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.92(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.11(0.12)	-0.11(0.12)	-0.12(0.12)	-0.14(0.12)	-0.13(0.13)
$q_n = 421$	Cover	0.94(0.01)	0.94(0.01)	0.92(0.02)	0.78(0.02)	0.93(0.01)
	$\sqrt{n}$Bias	-0.14(0.15)	-0.14(0.15)	-0.16(0.15)	-0.20(0.15)	-0.17(0.16)
$q_n = 561^*$	Cover	0.94(0.01)	0.92(0.01)	0.91(0.01)	0.64(0.02)	0.92(0.01)
	$\sqrt{n}$Bias	-0.20(0.21)	-0.23(0.21)	-0.26(0.24)	0.42(0.22)	-0.24(0.19)
$q_n = 631^*$	Cover	0.93(0.01)	0.92(0.01)	0.90(0.01)	0.47(0.02)	0.87(0.01)
	$\sqrt{n}$Bias	-0.29(0.28)	-0.31(0.28)	-0.33(0.17)	0.85(0.29)	0.41(0.29)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(1)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(1)}$. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Table 2.17: Simulation results ($d = 5$, heterogeneity, $\beta_{(1)}$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1}), \quad j = 1, \ldots, d, \quad \boldsymbol{\gamma}_k = 1/k, \quad k = 1, \ldots, q_n$				
		$\boldsymbol{x}_{i,n} \sim \mathcal{N}(0, \Sigma), \quad \boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$				
		Proposed+KJ	Proposed+HCK	Proposed + HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.95(0.01)	0.96(0.01)	0.98(0.01)	0.96(0.01)	0.96(0.01)
	$\sqrt{n}$Bias	0.05(0.05)	0.05(0.05)	0.04(0.05)	0.05(0.05)	0.05(0.06)
$q_n = 141$	Cover	0.94(0.01)	0.93(0.01)	0.95(0.01)	0.91(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	-0.06(0.06)	-0.10(0.06)	-0.06(0.06)	-0.12(0.06)	0.06(0.06)
$q_n = 281$	Cover	0.94(0.01)	0.93(0.01)	0.94(0.01)	0.88(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	0.07(0.07)	0.10(0.07)	0.07(0.07)	0.13(0.07)	0.07(0.08)
$q_n = 421$	Cover	0.94(0.01)	0.93(0.01)	0.94(0.01)	0.78(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	0.09(0.09)	0.12(0.09)	0.09(0.09)	0.14(0.09)	0.10(0.09)
$q_n = 561^*$	Cover	0.93(0.02)	0.89(0.01)	0.92(0.01)	0.59(0.02)	0.90(0.01)
	$\sqrt{n}$Bias	-0.10(0.13)	-0.15(0.13)	0.13(0.12)	-0.19(0.13)	0.14(0.12)
$q_n = 631^*$	Cover	0.93(0.01)	0.92(0.01)	0.92(0.01)	0.43(0.02)	0.82(0.01)
	$\sqrt{n}$Bias	0.19(0.17)	0.17(0.16)	0.16(0.13)	-0.30(0.18)	0.21(0.16)
		$\boldsymbol{x}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{x}}_{i,n} > 0), \quad \boldsymbol{w}_{i,n} \sim \mathcal{N}(0, I)$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.95(0.01)	0.95(0.01)	0.96(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.08(0.09)	0.09(0.09)	0.09(0.09)	0.08(0.09)	-0.09(0.09)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	-0.10(0.11)	-0.10(0.11)	-0.11(0.11)	-0.11(0.11)	0.11(0.11)
$q_n = 281$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.92(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.11(0.12)	0.11(0.12)	0.12(0.11)	0.14(0.12)	0.12(0.12)
$q_n = 421$	Cover	0.94(0.01)	0.94(0.01)	0.93(0.02)	0.73(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	-0.14(0.15)	-0.14(0.15)	-0.14(0.14)	-0.22(0.15)	-0.15(0.15)
$q_n = 561$	Cover	0.94(0.01)	0.93(0.01)	0.92(0.01)	0.61(0.02)	0.93(0.01)
	$\sqrt{n}$Bias	0.19(0.19)	0.21(0.19)	0.20(0.16)	0.44(0.20)	-0.20(0.19)
$q_n = 631^*$	Cover	0.94(0.01)	0.93(0.01)	0.90(0.01)	0.50(0.02)	0.91(0.01)
	$\sqrt{n}$Bias	0.24(0.26)	0.30(0.28)	0.35(0.21)	0.72(0.28)	0.32(0.28)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(1)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(1)}$. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Table 2.18: Simulation results ($d = 5$, heteroscedasticity, heterogeneity, $\beta_{(1)}$, $n = 2000$)

		$\beta_j = \Phi^{-1}(\frac{j}{d+1})$, $\boldsymbol{x}_{i,n} \sim \mathcal{N}(0,\Sigma)$, $\boldsymbol{w}_{i,n} = \mathbb{1}(\tilde{\boldsymbol{w}}_{i,n} \geq \Phi^{-1}(0.98))$, j = 1,..., d				
		$\boldsymbol{\gamma}_n = 0$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.95(0.01)	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.02(0.03)	0.03(0.03)	0.03(0.03)	0.03(0.03)	0.03(0.03)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.91(0.02)	0.94(0.01)
	$\sqrt{n}$Bias	-0.04(0.04)	-0.04(0.04)	-0.04(0.04)	-0.06(0.04)	-0.04(0.04)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.93(0.01)	0.88(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	0.04(0.05)	0.05(0.05)	0.07(0.05)	-0.12(0.05)	0.07(0.05)
$q_n = 421$	Cover	0.94(0.01)	0.94(0.01)	0.91(0.01)	0.85(0.01)	0.93(0.01)
	$\sqrt{n}$Bias	-0.05(0.05)	-0.05(0.05)	-0.09(0.05)	-0.15(0.05)	-0.07(0.05)
$q_n = 561$	Cover	0.94(0.01)	0.93(0.01)	0.91(0.01)	0.82(0.01)	0.91(0.01)
	$\sqrt{n}$Bias	-0.05(0.05)	-0.06(0.05)	-0.10(0.05)	-0.18(0.05)	-0.11(0.05)
$q_n = 631^*$	Cover	0.93(0.01)	0.91(0.01)	0.90(0.01)	0.78(0.01)	0.90(0.01)
	$\sqrt{n}$Bias	-0.07(0.05)	-0.09(0.05)	-0.12(0.05)	-0.22(0.05)	-0.14(0.05)
		$\gamma_k = 1/k$, $k = 1,\ldots,q_n$				
		Proposed+KJ	Proposed+HCK	Proposed+HC3	Proposed+EW	No adjustment+KJ
$q_n = 1$	Cover	0.96(0.01)	0.96(0.01)	0.95(0.01)	0.94(0.01)	0.95(0.01)
	$\sqrt{n}$Bias	0.04(0.05)	0.05(0.05)	0.05(0.05)	0.05(0.05)	0.05(0.05)
$q_n = 141$	Cover	0.95(0.01)	0.95(0.01)	0.94(0.01)	0.93(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.07(0.08)	0.08(0.08)	0.08(0.08)	0.11(0.08)	0.08(0.08)
$q_n = 281$	Cover	0.94(0.01)	0.94(0.01)	0.94(0.01)	0.92(0.01)	0.94(0.01)
	$\sqrt{n}$Bias	0.08(0.08)	0.10(0.08)	0.13(0.08)	0.14(0.08)	0.08(0.08)
$q_n = 421$	Cover	0.94(0.01)	0.93(0.01)	0.92(0.01)	0.88(0.01)	0.92(0.01)
	$\sqrt{n}$Bias	0.08(0.08)	0.12(0.08)	0.14(0.08)	0.18(0.08)	0.15(0.08)
$q_n = 561^*$	Cover	0.94(0.01)	0.91(0.01)	0.90(0.01)	0.83(0.02)	0.90(0.01)
	$\sqrt{n}$Bias	0.09(0.08)	0.15(0.08)	0.17(0.07)	0.23(0.08)	0.18(0.08)
$q_n = 631^*$	Cover	0.92(0.01)	0.90(0.01)	0.86(0.01)	0.76(0.01)	0.85(0.01)
	$\sqrt{n}$Bias	0.12(0.08)	0.17(0.08)	0.20(0.05)	0.28(0.08)	0.23(0.08)

Note: "Cover" is the empirical coverage of the 95% confidence interval for $\beta_{(1)}$ and " $\sqrt{n}$Bias " captures the root-n scaled Monte Carlo bias for estimating $\beta_{(1)}$. In this heteroscedastic design, $\tilde{\boldsymbol{w}}_{i,n} \sim \mathcal{N}(0, \boldsymbol{I}_{q_n})$, $\varepsilon_{i,n} \sim \mathcal{N}(0,1)$, $\mathbb{V}[\varepsilon_{i,n}|\boldsymbol{x}_{i,n}, \boldsymbol{w}_{i,n}] = c_\varepsilon(1+(t(x_{1,i,n})+\boldsymbol{l}'\boldsymbol{w}_{i,n})^2/4)$, and $\mathbb{V}[x_{k,i,n}|\boldsymbol{w}_{i,n}] = c_{x_k}(1+(\boldsymbol{l}'\boldsymbol{w}_{i,n})^2/4)$, where $x_{k,i,n}$ denotes the kth component of the vector $\boldsymbol{x}_{i,n}$. The constants c_ε and c_{x_k} are chosen so that $\mathbb{V}[\varepsilon_{i,n}] = \mathbb{V}[x_{k,i,n}] = 1$ and $t(a) = a\mathbb{1}(-1 \leq a \leq 1) + sgn(a)(1 - \mathbb{1}(-1 \leq a \leq 1))$. $\boldsymbol{l}$ is the conformable vector of ones. "*" indicates that $\hat{\boldsymbol{\Omega}}_n^{\text{KJ}}$ is not positive semi-definite in some Monte Carlo samples.

Method	Policies(Ask amount)	Est (95% CI)	p-value
Uncalibrated	Same	0.70(0.10, 1.29)	0.023*
	25% more	0.67(-0.04, 1.37)	0.065
	50% more	0.38(-0.21, 0.96)	0.205
Calibrated	Same	0.66(0.07, 1.24)	0.026*

Table 2.19: Uncalibrated and calibrated results under a smaller model with main effects only ($n = 7,938, p = 53$). Estimated treatment effects (Est), 95% confidence intervals (95% CI), and two-sided p-values for the three "ask amount" policies. "Uncalibrated" refers to the study results obtained without any adjustment, and the confidence intervals are constructed based on normal approximation with the estimated covariance matrix $\hat{\mathbf{\Omega}}_n^{\text{KJ}}$. "Calibrated" refers to our proposed methodology. The computational time is 533 seconds on a Lenovo NeXtScale nx360m5 node (24 cores per node) equipped with Intel Xeon Haswell processor.

Chapter 3

Adaptive Experiments Toward Learning Treatment Effect Heterogeneity

3.1 Introduction

Motivation

Understanding and characterizing treatment effect heterogeneity have become increasingly important in many scientific fields. For example, in precision health, identifying differential treatment effects serves as an important step towards materializing the benefits of precision health because it provides evidence regarding how groups of patients with specific characteristics respond to a given treatment either in efficacy or in adverse effects [133]. As another example, in large internet companies and social science research, studying the impact of marketing offers on consumer purchases and evaluations of the effectiveness of government programs or public policies across different subgroups inform more effective policy-making [93, 95].

Existing literature in this research area has mostly focused on conducting retrospective post hoc analyses in observational or randomized experiment data. Even with large-scale observational data or carefully collected randomized experiment data, statistical bias in these analyses can not only arise due to the violation of untestable causal assumptions and the presence of unmeasured confounders in observational data [14, 12, 74, 131, 41, 83], but also may arise due to the widespread winner's curse phenomenon when ignoring seemingly promising heterogeneous treatment effects are selected from the data in an ad hoc fashion [64, 66, 169, 119, 36].

More concretely, in observational studies, causal conclusions are commonly established under a set of causal assumptions. For example, one of the commonly imposed causal assumptions in practice is the unconfoundedness assumption, which states that conditional on measured confounders, the treatment assignment is as good as random [see 126, for

example]. Given the unconfoundedness assumption is untestable, there is no guarantee as to whether established causal conclusions under this assumption are valid [87]. In classical randomized experiment data, although carrying out valid causal conclusion does not require imposing untestable causal assumptions, exploring treatment effect heterogeneity could still be susceptible to the winner's curse bias when researchers iteratively search for subgroups with high treatment levels and then only report the results for those subgroupns with large effects [64, 8, 152].

Our contribution

In this chapter, rather than conducting retrospective analysis in existing datasets, we look into the problem from a different perspective and propose an adaptive data collection mechanism–*adaptive randomized experiments*–to gather reliable causal evidence specifically targeted toward learning treatment effect heterogeneity. By adaptive, we mean that experimenters are allowed to sequentially allocate and modify experimenter measurement efforts (such as the treatment allocation probability and the proportions of sequentially enrolled subgroups) and adapt to the accrued evidence during the experiment [141, 79, 149, 166, 165, 143, 179, 60, 141]; see Section 3.1 for literature review and Section 3.2 for a detailed introduction. To collect robust evidence towards learning treatment effect heterogeneity, we formalize our experimental design goal as maximizing the probability of correctly selecting subpopulations (or subgroups) who respond favorably to the treatment under the language of large deviation theory (Section 3.3). Without loss of generality, we refer to the best subpopulation or the best subgroup as the subpopulation with the highest treatment level in this chapter. In what follows, we further break down our contributions from two perspectives:

On the methodology side, our proposed adaptive experiment strategy offers two potential benefits compared with classical retrospective analysis. On the one hand, compared with conducting post hoc analysis in observational data, because treatments are still randomly assigned in adaptive experiments and the treatment variables are independent of any potential unmeasured confounding variables, the proposed adaptive experiment design strategy generates samples enabling valid causal conclusions without imposing any untestable causal assumptions. On the other hand, compared with conducting post hoc analysis from randomized experiment data, our design is equipped with the flexibility to sequentially revise the experimental strategy. Thus, the proposed adaptive experimental design strategy possesses the ability to periodically detect individuals who respond favorably to the treatment and then optimize experimental effort spending based on the inferred context. As a part of this endeavor, this design feature offers advantages in improving the statistical efficiency of detecting treatment heterogeneity compared to completely randomized experiments (Proposition 2).

On the theoretical side, we first leverage the large deviation principle to find target "oracle" allocation strategies to spend the experimental efforts (Section 3.3), and demonstrate these oracle allocations are attainable using our proposed design strategies (Lemma 5). In particular, compared with completely randomized experiments, the derived oracle allocations

demonstrate the benefit of our design in terms of statistical efficiency gain (Proposition 2) and the ability to maintain a high correct selection probability of the best subgroup while allowing the top few subgroups to have closely tied treatment effects (Proposition 3). Next, unlike classical response adaptive designs, throughout this chapter, we do not restrict the outcome to following any parametric forms, alleviating the burden of choosing what type of parametric assumptions should be used in practice. Under mild moment restrictions on the outcome variables, we show that the proposed design delivers an asymptotic normally distributed estimator for the subgroup with the highest treatment level, and its variance can be consistently estimated (Theorem 4 and Theorem 6). Third, because the adopted allocations during the experiment and the subgroup treatment effect estimator rely on the sequentially accumulated historical data, characterizing the statistical properties of our proposed design and making inferences on the proposed estimator can be rather challenging, and some of the theoretical results might be of independent interest (see Supplementary Materials for details).

Existing literature

Adaptive experiments have been commonly adopted in clinical trials where patients are enrolled sequentially based on certain eligibility criteria, and in recent years they have been quickly picked up by online platforms to complete A/B tests or digital randomized experiments. Existing adaptive experiment design strategies can be roughly divided into covariate adaptive design, response adaptive design, and (adaptive) enrichment design. Covariate adaptive design has been widely used in modern clinical trials to balance treatment assignments across important prognostic factors. This design refers to a randomized treatment allocation scheme that depends solely on participant covariate information but is independent of the observed outcomes. Response adaptive design often refers to the design strategy in which the treatment assignment probabilities are adapted during the experiment based on the accrued evidence in the outcomes, with the goal of simultaneously achieving the experimental objectives and preserving statistical inference validity [79, 69, 60, 139]. Adaptive enrichment designs are often adopted in clinical trials and use interim data to identify treatment-sensitive patient subgroups by changing patient enrollment criteria. In these designs, experimenters often partition the population into pre-defined subgroups based on biomarkers measured at baseline and enroll patients in multiple stages [143]. In what follows, we provide a more detailed literature review of the above-mentioned adaptive experimental design strategies and their connections and differences to our design strategy.

In the response adaptive design literature, the early design can be traced back to Pólya's urn model [46]. Based on a randomized urn model and Zelen's play-the-winner-rule [180], Wei and Durham later developed the randomized play-the-winner rule in clinical trial settings [174]. More discussions on the urn-based designs can be found in [141] and [139]. The asymptotic properties of urn models are discussed in [88] and [17]. Besides the urn type of designs, another conventional class of response adaptive designs is the doubly adaptive biased coin (DBCD) design. The early DBCD design can be found in [48], which has its root

in Efron's biased coin design [44]. The asymptotic properties of DBCD designs are studied in various works [80, 159, 81]. However, much of the existing work on response adaptive designs aims to optimize the estimation efficiency of the overall treatment effect but is not tailored to study treatment effect heterogeneity [140, 78]. Although many response adaptive designs are carried out in fully adaptive settings, Hahn et al. propose a design in a multi-stage setting (Table 3.1). More specifically, this work proposes to revise the treatment assignment probability by minimizing the asymptotic variance of the average treatment effect estimator. Nevertheless, this design is carried out in two stages and is not designed to identify treatment effect heterogeneity.

Instead of relying solely on the outcome variable to optimize for the experimental goals, one may further incorporate covariate information. RAR that further incorporate covariate information is known as covariate-adjusted response adaptive (CARA) designs [18, 142, 184]. Early work in [128] and [181] proposes to balance covariates based on the biased coin design. Hu et al. propose a family of CARA designs that could account for both efficiency and ethics [82]. Aletti et al. generalize the CARA design framework to incorporate nonparametric estimates of the conditional response function [4]. Some other CARA designs are discussed in [113], [167], and [185]. In addition, the theoretical framework for the rerandomization design, that is, to balance covariates in different treatment arms before conducting experiments, is established in [122], and [121]. Later, Zhou et al. extend from the prior work and propose to conduct sequential rerandomization where subjects are sequentially enrolled in groups [187].

In the adaptive enrichment design literature, the early work in [60] considers revising the enrollment proportions of two discrete patient subgroups defined by a single biomarker and provides conditions under which the type I error rate is preserved. Some later work with the similar setup includes [172], [144], and [145]. Simon and Simon develop a more general framework of adaptive enrichment designs that can handle multiple biomarkers at the interim analysis. Stallard considers overlapping subgroups defined by a continuous biomarker [151]. However, much of the existing work on adaptive enrichment designs aims to preserve the type I error rate instead of identifying the best subgroup with high probability. In addition, many adaptive enrichment designs are carried out across multiple stages but are not conducted in a fully adaptive manner.

Our work also leverages the mathematical framework provided by the large deviation theory [77, 40, 157]. The general large deviation theory shares a natural connection with concentration inequality. An early theorem in large deviation framework–Sanov's theorem–lays out the foundation for various concentration inequality results [75, 76, 43].

3.2 A synthesized adaptive experiment framework

In this section, we introduce a unified design framework in a two-arm (a treatment arm and a control arm) experiment along with notation. Our design framework integrates both classical response adaptive randomization (RAR) design and enrichment design frequently adopted in practice.

Suppose the experiment participants are sequentially enrolled in T stages. The total number of enrolled subjects is $N = \sum_{t=1}^{T} n_t$, where n_t denotes the number of subjects in Stage t, for $t = 1, \ldots, T$. For the data collected in Stage t, we denote subject i's treatment assignment status as $D_{it} \in \{0, 1\}$, $i = 1, \ldots, n_t$, with $D_{it} = 1$ being the treatment arm and $D_{it} = 0$ being the control arm. Denote subject i's covariate information as $X_{it} \in \mathbb{R}^p$ and the observed outcome as $Y_{it} \in \mathbb{R}$. To formally introduce treatment (or causal) effects, we follow the Neyman-Rubin causal model [125, 148] in this chapter. Define $Y_{it}(d)$ as the potential outcome we would have observed if subject i receives treatment d at Stage t, for $d \in \{0, 1\}$. The observed outcome can then be written as

$$Y_{it} = D_{it}Y_{it}(1) + (1 - D_{it})Y_{it}(0), \quad i = 1, \ldots, n_t, \ t = 1, \ldots, T.$$

In line with existing literature in adaptive experiments, we assume that the outcomes are observed without delay, and their underlying distributions do not shift over time [79]. Furthermore, we define the history, i.e., the collected data, up to Stage t as

$$\boldsymbol{\mathcal{H}}_t := \{\mathcal{H}_s\}_{s=1}^{t} \triangleq \{(Y_{is}, D_{is}, X_{is}), i = 1, \ldots, n_s\}_{s=1}^{t}.$$

To investigate treatment effect heterogeneity, we divide the sample space $\mathcal{X}$ of the covariates X_{it} into m non-overlapping regions, denoted as $\{\mathcal{S}_j\}_{j=1}^{m}$ (an extension of an overlapping division shall be discussed in Section 3.9). In clinical settings, each division of the sample space is frequently referred to as a subgroup [11, 96, 179]; each subgroup of subjects, by definition, has different characteristics. Here, we assume subgroups are pre-specified. We measure the effectiveness of the treatment in each subgroup by taking the mean difference between the potential outcomes in the treated and controlled arms:

$$\tau_j = \mathbb{E}[Y_{it}(1) - Y_{it}(0) | X_{it} \in \mathcal{S}_j], \ j = 1, \ldots, m.$$

We further denote the total number of subjects enrolled in subgroup j as $N_j = \sum_{t=1}^{T} n_{tj}$.

In adaptive experiments, practitioners aim to sequentially allocate experimental efforts to reach certain pre-specified design goals. Such efforts often include actively recruiting subjects of different characteristics in multiple stages and revising treatment assignment (or allocation) probabilities based on accrued evidence during the experiment. In the current literature, there are two widely adopted design strategies focusing on dispensing these experimental efforts differently, which we shall discuss below:

The first strategy is called response adaptive randomization (RAR) design or covariate-adjusted response adaptive (CARA) design. In these designs, experiments can sequentially revise the treatment assignment strategies based on responses accrued during the experiment, but, unlike enrichment designs, often do not change the enrollment criteria across multiple stages. RAR designs incorporating additional covariate information are more frequently referred to as covariate-adjusted response adaptive (CARA) designs. The design goals of response adaptive designs tend to vary in different application areas, and we refer interested

readers for [134] for a comprehensive review. Formally, by defining the treatment assignment probability (or propensity scores) for subjects in subject j as

$$e_{t,j} := \mathbb{P}(D_{it} = 1 | X_{it} \in \mathcal{S}_j),\ t = 1, \ldots, T,\ j = 1, \ldots, m,$$

RAR and CARA design aim to dynamically revise $e_{t,j}$ to reach desired design goals.

The second strategy is called (adaptive) enrichment design, which has been frequently carried out in clinical settings to identify patient subgroups that benefit the most from a given treatment [60, 150, 105, 143]. In these designs, experimenters often fix the treatment allocation probability during the entire experiment, but they sequentially enroll different subgroups of participants over different stages. Here, the word "enrichment" spells out the action of actively recruiting a new batch of subjects who may have characteristics different from the previous stage, and the word "adaptive" indicates that the enrollment proportions of subjects with different characteristics can be adaptively revised based on the current understanding of treatment effect heterogeneity. Formally, by defining an auxiliary variable $Z_{it} \in \{1, 0\}$ that indicates if subject i is enrolled at Stage t, we introduce the enrichment proportion of subjects falling into region $\mathcal{S}_j$ in Stage t as

$$p_{t,j} := \mathbb{P}(X_{it} \in \mathcal{S}_j | Z_{it} = 1),\ t = 1, \ldots, T,\ j = 1, \ldots, m.$$

Enrichment designs sequentially revise $p_{t,j}$ across multiple stages to reach their design objectives.

Our proposed adaptive experimental design framework unifies response adaptive designs and enrichment designs by formalizing them as a sequential policy learning problem (see Table 3.1 for a summary). We hope that this unified framework broadens the practicability of the proposed design framework under various practical constraints. In particular, we define a sequential policy $\boldsymbol{\pi}$ consisting of a sequence of policies $\pi_1, \ldots, \pi_T$, and each π_t is a mapping from the historical data $\boldsymbol{\mathcal{H}}_t := \{\mathcal{H}_s\}_{s=1}^t$ accumulated up to Stage t to either the subgroup enrichment proportions $\boldsymbol{p}_{t+1} \triangleq (p_{t+1,1}, \ldots, p_{t+1,m})$, or to the treatment assignment probabilities $\boldsymbol{e}_{t+1} \triangleq (e_{t+1,1}, \ldots, e_{t+1,m})$, that is:

$$\begin{aligned} &\pi_t : \{\mathcal{H}_s\}_{s=1}^{t-1} \to \boldsymbol{e}_t \triangleq (e_{t,1}, \ldots, e_{t,m}). && \text{Response adaptive design} \\ &\pi_t : \{\mathcal{H}_s\}_{s=1}^{t-1} \to \boldsymbol{p}_t \triangleq (p_{t,1}, \ldots, p_{t,m}), && \text{Adaptive enrichment design.} \end{aligned}$$

Other than dispending different experimental strategies, practitioners can also flexibly choose the number of stages T and the number of participants n_t in each stage of the experiment. We refer to experimental design strategies with large n_t and finite T as multi-stage designs, and we refer to designs with small n_t and large T as fully adaptive designs. While both designs tend to share similar large sample properties, they have different strengths and can often be applied in scenarios with different practical constraints. On the one hand, multi-stage designs can be preferable in clinical settings or social experiments where experimenters often have a limited number of opportunities to revise the experimental effort allocation during the experiment (see [93, 62] for example). Fully adaptive designs are more

Table 3.1: Examples of frequentist data collection mechanisms in adaptive experiments.

π_t	$n_t = 1$ with large T Fully adaptive	Large n_t with finite T Multi-stage
Response adaptive design $\{\mathcal{H}_s\}_{s=1}^t \to \{\boldsymbol{e}_s\}_{s=1}^{t-1}$	Response adaptive [141, 149, 79, 166] Covariate-adjusted response adaptive [167, 82, 113, 185, 4]	Adaptive propensity score [69] Sequential rerandomization [122, 121, 187]
Enrichment design $\{\mathcal{H}_s\}_{s=1}^t \to \{\boldsymbol{p}_s\}_{s=1}^{t-1}$	Not available	Frequentist enrichment design [105, 150, 60, 151, 144]

readily to be integrated into digital experiments such as online A/B testing or digital clinical trials in which sequentially allocating experimental efforts in a large number of stages is more practical and less costly (see [95, 134] for example). On the other hand, shall be seen in our simulation studies in Section 3.7, benefiting from frequently updated experimental strategy, fully adaptive designs tend to have superior finite sample performance compared to multi-stage designs when the sample size N is rather small.

Benefiting from the above framework, while existing adaptive experiments normally target one of the experimental schemes listed in Table 3.1, the design strategies we shall propose can be applied in all four settings. This demonstrates that the proposed design strategy is flexible and completes existing frequentist adaptive design strategies, suggesting our designs can be potentially applied to online experiments conducted in e-commerce platforms, clinical trials conducted in health industries, and policy evaluation experiments conducted for social science research. In what follows, we introduce the general goal of our design strategy.

3.3 Design objectives and oracle allocation strategies: A large deviation perspective

When participants are enrolled sequentially in an adaptive experiment, the adoption of multi-stage sequential policy learning allows experiment designers to sequentially modify their policies to reach a particular design goal, either by adapting the treatment allocation of participants or by enrolling fewer participants in non-effective treatment groups. In the hope of collecting robust evidence towards learning treatment effect heterogeneity, our design goal aims to inform practitioners that the identified subgroups who mostly benefited (or were harmed) from the treatment in the experiment are indeed the best-performing (or worst-performing) subgroups in the population.

To illustrate the importance of correctly identifying the best-performing subgroup from experiments, we shall provide an example. MONET1 study is a randomized controlled

trial aiming to evaluate the treatment efficacy of a new drug (motesanib combined with carboplatin/paclitaxel) on advanced nonsquamous nonsmall-cell lung cancer (NSCLC). Researchers identify that the East Asian subgroup exhibits the largest treatment effect [96]. In light of the identified best-performing subgroup, Amgen, the drug company, invests a large amount of budget in moving forward with this drug. Unfortunately, the follow-up trial concludes that the identified best-performing subgroup is, in fact, not significant. The failure of MONET1 suggests that correctly identifying the best-performing subgroup from the experiment is crucial for further generalizing scientific findings.

To formulate our design objective from a statistical standpoint, when the experiment ends, we aim to construct reliable estimators of the subgroup average treatment effect so that the probability of correctly identifying the subgroups with the most beneficial (or harmful) effects is maximized. More formally, without loss of generality, we assume that the population subgroup average treatment effects satisfy $\tau_1 > \tau_2 > \ldots > \tau_m$ (generalizations to other possible effect orders are provided in Section 3.9), and we have constructed consistent estimators $\hat{\tau}_1, \ldots, \hat{\tau}_m$ of $\tau_1, \ldots, \tau_m$ based on the collected data at the end of the experiment. Because the joint distribution of $\hat{\tau}_1, \ldots, \hat{\tau}_m$ not only depends on the underlying data distribution of the potential outcome and covariates but also crucially relies on the treatment assignment mechanism and subgroup enrollment proportions, these estimators can also be viewed as a function of the historical data and the corresponding policy adopted in the adaptive experiment. Then, in a simple case where we aim to find the best subgroup with the largest treatment effect in the population (i.e., the first subgroup $\mathcal{S}_1$), our design objective is to find a sequential policy $\boldsymbol{\pi}$ belonging to a set of feasible policies $\boldsymbol{\Pi}$, so that the probability of the estimated first subgroup treatment effect margins out the others is maximized. As in this simple case, the first subgroup has the largest treatment effect in the population, the correction selection probability can be written as $\mathbb{P}\big(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j\big)$.

Unfortunately, without imposing additional parametric distributional assumptions on the historical data, directly searching for a policy that maximizes the correct selection probability results in an intractable optimization problem, as deriving a general analytic form of the correct selection probability is nearly impossible. One seemingly natural alternative is to consider solving this optimization problem in an asymptotic sense. Because by letting the total sample size N go to infinity, we can approximate the distribution of $\hat{\tau}_j$ with a Gaussian distribution under mild conditions. Nevertheless, given $\tau_1 > \tau_2$ and for any policy π, the correct selection probability $\mathbb{P}(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j)$ grows exponentially fast to one as $N \to \infty$ and thus is no longer a function of $\boldsymbol{\pi}$, suggesting that directly searching for a sequential policy maximizing the correct selection probability in an asymptotic sense is infeasible.

To avoid the above-mentioned issues, for now, we abandon the study of a sequential policy maximizing correct selection probability and instead study in an ideal but unrealistic situation. For us, the ideal situation is that, before the experiment starts, we acquire perfect knowledge about the joint distribution of the potential outcomes and the covariates. Such an oracle then allows us to study the best strategy to allocate experimental efforts so that the correct selection probability is maximized. Nevertheless, depending on the underlying distribution, because the correct selection probability can still take a complicated form with

finite N or tend to 1 with large N, searching for the oracle allocation strategy remains a challenging task. This motivates us to further magnify the correction selection probability through the lens of the large deviation principle [77, 40], which allows us to precisely characterize the correction selection probability with the minimum of a set of *rate functions* $\{G(\mathcal{S}_1, \mathcal{S}_j)\}_{j=2}^m$, that is

$$\lim_{N\to\infty} \frac{1}{N} \log\Big(1 - \mathbb{P}\big(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j\big)\Big) = -\min_{2\leq j\leq m} G(\mathcal{S}_1, \mathcal{S}_j; e_1, p_1, e_j, p_j).$$

and we defer the mathematical details of the above derivation to Section 3.6. The rate function $G(\mathcal{S}_1, \mathcal{S}_j; e_1, p_1, e_j, p_j)$, which often has a closed-form expression relying on treatment allocations and subgroup enrichment proportions in the best subgroup and subgroup j, measures the exponential decay rate of the probability of the rare event that the estimated best subgroup treatment effect $\hat{\tau}_1$ is smaller than the estimated subgroup treatment effect $\hat{\tau}_j$ as the sample size $N \to \infty$.

Borrowing the language similar to [42] and [53], we are ready to define *oracle allocation strategies* in the response adaptive designs and the enrichment designs as the following. In response adaptive designs, when the enrollment criteria are fixed, and the subgroup proportions cannot be modified, we define the oracle treatment allocation probabilities $\boldsymbol{e}^* \triangleq (e_1^*, \ldots, e_m^*)$ as the solution to the following constraint optimization problem:

$$\max_{\boldsymbol{e}\in(0,1)^m} \Big\{ \min_{2\leq j\leq m} G(\mathcal{S}_1, \mathcal{S}_j; e_1, e_j) : \sum_{j=1}^m p_j e_j = c_1,\ c_2 \leq e_j \leq 1 - c_2 \Big\},$$

where c_1 and c_2 are pre-specified positive constants between zero and one. Similarly, in enrichment designs, when the treatment assignment probabilities in different subgroups are fixed, and the propensity scores $\boldsymbol{e} = (e_1, \ldots, e_m)$ cannot be modified, we define the oracle subgroup enrichment proportions $\boldsymbol{p}^* \triangleq (p_1^*, \ldots, p_m^*)$ as the solution to the following constraint optimization problem

$$\max_{\boldsymbol{p}\in(0,1)^m} \Big\{ \min_{2\leq j\leq m} G(\mathcal{S}_1, \mathcal{S}_j; p_1, p_j) : \sum_{j=1}^m p_j = 1,\ p_j \geq 0 \Big\}.$$

The closed-form solution of the above two optimization problems critically relies on the specific choice of the subgroup treatment effect estimators, and we thus leave more detailed discussions of the oracle allocation strategies for the response adaptive design and the enrichment design in Sections 3.4 and 3.5. As shall be made clear in later sections, the oracle allocation strategies offer considerable advantages over traditional adaptive experimental designs, including improving the efficiency in estimating the best subgroup treatment effect (Proposition 2) and allowing the population treatment effect of the second-best subgroup to stay closer to that of the best subgroup (Proposition 3).

In practice, when experimenters have no prior knowledge about the joint distribution of the subgroup treatment effect estimators, adaptive experiments offer a natural environment

to sequentially learn the unknown parameters in each subgroup and adjust the allocation of experimental efforts during the experiment. In the following sections, we aim to answer the following two research questions: When we have no prior information about the data-generating process, is it possible to carry out adaptive experimental design strategies that sequentially study the joint distribution of the underlying data and meanwhile use learned information to better allocate experimental efforts as the experiment progresses? When the experiment is finished, can such designs produce subgroup treatment effect estimators that have competing performances with the ones under the oracle allocation strategies?

3.4 Response adaptive design with adaptive treatment allocation

In this section, we start with introducing the oracle treatment allocation strategy in response adaptive designs. We propose two design strategies for fully adaptive and multi-stage settings (Table 3.1).

Oracle treatment allocation in response adaptive designs

As the rate function depends on the choice of the subgroup treatment effect estimators, in this section, we adopt the classical inverse propensity score weighting (IPW) [138] to estimate the subgroup treatment effects

$$\hat{\tau}_j = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{\mathbb{1}_{(X_i\in\mathcal{S}_j)}}{p_j}\frac{D_iY_i}{e_j}\right\} - \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{\mathbb{1}_{(X_i\in\mathcal{S}_j)}}{p_j}\frac{(1-D_i)Y_i}{1-e_j}\right\}, \quad j=1,\ldots,m. \tag{3.1}$$

We leave an extension with the augmented IPW estimator [136] to Section 3.9.

Benefiting from the simple form of the above IPW estimator and based on the Gartner-Ellis Theorem [40, ch.2.3] in the large deviation theory, we are able to derive a closed form expression of the rate function $G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e})$ with

$$G(\mathcal{S}_1,\mathcal{S}_j;e_1,e_j) = \frac{(\tau_j-\tau_1)^2}{2\big(\sigma_1^2(e_1)+\sigma_j^2(e_j)\big)}, \quad \sigma_j^2(e_j) = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_je_j} + \frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_j]}{p_j(1-e_j)}, \tag{3.2}$$

where $\sigma_j^2(e_j)$ measures the variance of $\hat{\tau}_j$.

With the closed-form expression of the rate function in hand, we are now ready to explore the oracle treatment allocation $\boldsymbol{e}^* \triangleq (e_1^*,\ldots,e_m^*)$, which solves the following optimization

problem:

Problem *A*

$$\begin{aligned}
&\max_{\boldsymbol{e}} \min_{2\le j\le m} \frac{(\tau_j-\tau_1)^2}{2\big(\sigma_1^2(e_1)+\sigma_j^2(e_j)\big)}, && \leftarrow \text{Maximize correct selection probability}\\
&\text{s.t.} \sum_{j=1}^{m} p_j e_j = c_1, && \leftarrow \text{“Cost”/practical constraint}\\
&\quad c_2 \le e_j \le 1-c_2,\ j=1,\dots,m, && \leftarrow \text{Feasibility constraints}
\end{aligned}$$

where c_1 and c_2 are positive constants between 0 and 1. Here, the “cost” or practical constraint restricts the proportion of subjects receiving the treatment, and the feasibility constraint restricts the treatment assignment probability in each subgroup to be bounded away from zero and one. Because the objective function is the minimum of $m-1$ rate function, the above optimization problem is nonlinear. We instead work with its equivalent epigraph representation:

Problem *B*

$$\begin{aligned}
&\max_{\boldsymbol{e}} z, && \leftarrow \text{Linear objective function}\\
&\text{s.t.} \sum_{j=1}^{m} p_j e_j = c_1, && \leftarrow \text{“Cost”/practical constraint}\\
&\quad c_2 \le e_j \le 1-c_2,\ j=1,\dots,m, && \leftarrow \text{Feasibility constraints}\\
&\quad \frac{(\tau_j-\tau_1)^2}{2\big(\sigma_1^2(e_1)+\sigma_j^2(e_j)\big)} - z \ge 0,\ j=2,\dots,m. && \leftarrow \text{Equivalent to maximize}\\
& && \quad\ \text{correct selection probability}
\end{aligned}$$

The above epigraph representation, on the one hand, yields a concave optimization problem that can be efficiently solved by open-source software such as `IPOPT` [168] and `GUROBI` [67]; on the other hand, it allows us to explore the Lagrangian dual problem and derive a neat form of the oracle treatment allocations in some simplified cases (see Remark 7 below), providing statistical insights into the proposed design strategy.

Remark 7 (Oracle treatment allocation in a simplified case)**.** *Let $c_1 = 1/2$, that is half of the subjects can be assigned to treatment, and assume (1) the expectation of the conditional variance of the potential outcome is the same in the treatment and control arms for each subgroup, that is $\mathbb{V}[Y(1)|X\in\mathcal{S}_j] = \mathbb{V}[Y(0)|X\in\mathcal{S}_j]$, for $j=1,\dots,m$, and (2) each subgroup has equal proportion with $p_j = \frac{1}{m}$. The oracle treatment allocation can be found by solving the following equations:*

$$\begin{cases}
\frac{(\tau_j-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{e_1(1-e_1)}+\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{e_j(1-e_j)}} = \frac{(\tau_k-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{e_1(1-e_1)}+\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{e_k(1-e_k)}},\\
\sum_{l=1}^{m} p_l e_l = \frac{1}{2}.
\end{cases}$$

Because e_j does not have closed-form expression, we shall provide some intuitions on e_j in Figure 3.1. We consider three subgroups and two difference scenarios: (I) Let $\tau_1 = 3$, $\tau_3 = 0.5$, $\sigma_1 = \sigma_2 = \sigma_3 = 2$. We show the change of e_2 with respect to τ_2. (II) Let $\tau_1 = 3$, $\tau_2 = 2$, $\tau_3 = 1$, and $\sigma_1 = \sigma_3 = 2$. We show the change of e_2 with respect to σ_2.

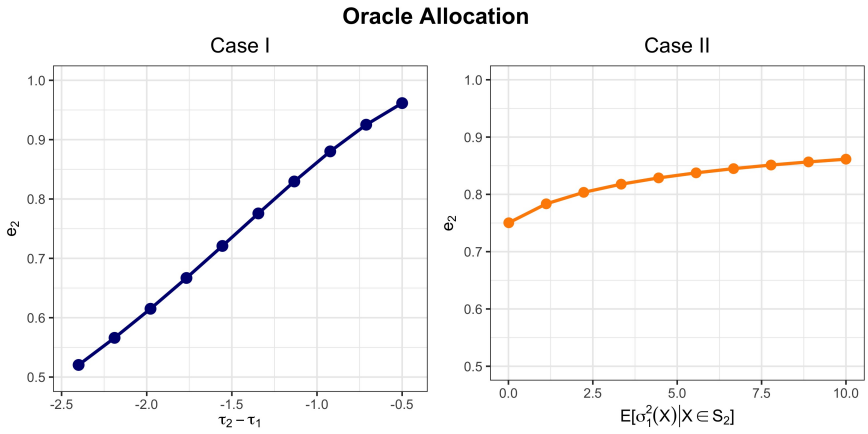


Figure 3.1: The change of oracle treatment allocation in subgroup 2.

If we further assume $\sigma_1^2(e_1) \ll \sigma_j^2(e_j)$, we can derive the closed-form expression of the oracle treatment allocation as

$$e_j^* = \frac{\alpha_j}{\frac{2}{m}\sum_{j=1}^m \alpha_j}, \quad \left(\alpha_j - \frac{1}{2}\right)^2 = \left|\frac{1}{4} - C \times \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]}{(\tau_j - \tau_1)^2}\right| \in \left(0, \frac{1}{4}\right), \quad j = 2, \ldots, m, \quad (3.3)$$

where C is a known constant ensures $\alpha_j \in (0, 1)$ and its specific form can be found in the Supplementary Material Section B, and e_1^* is the solution to the following equation:

$$\frac{e_1^2(1-e_1)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_1](2e_1 - 1)} = \sum_{j=2}^{d} \frac{e_j^{*2}(1-e_j^*)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_j](2e_j^* - 1)}.$$

Figure 3.1 provides some intuitions on the oracle treatment allocation strategy. Intuitively, when τ_j is closer to τ_1 or when the variance in subgroup j increases, subgroup j will require more experimental efforts to explore, and thus the oracle treatment allocation design will assign a larger portion of subjects to the treatment arm in subgroup j. The term $\frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]}{(\tau_j - \tau_1)^2}$ in Eq (3.3) shares the similar interpretation.

After obtaining the oracle treatment allocation, we want to approximate the oracle treatment allocation with accrued data in an adaptive experiment environment. In what follows, we shall discuss our proposed adaptive treatment allocation strategy in fully adaptive settings and multi-stage settings.

Fully adaptive case with large T and small n_t

In this section, we provide our proposed design strategy in the fully adaptive setting with large T and small n_t (Table 3.1). The derived sequential policy $\boldsymbol{\pi}_{\texttt{FA-RA}} = (\pi_1, \ldots, \pi_{t-1})$ enables us to dynamically allocate the experimental efforts so that the derived subgroup treatment effect estimator shares the same performance as the one delivered by the oracle allocation strategies.

Stage 1. *Enroll a portion of subjects from all the considered subgroups. Then randomly assign subjects in each subgroup to the treatment arm with a pre-specified propensity score, such as* $e_{1j} = \frac{1}{2}$, $j = 1, \ldots, m$.

As we have no prior information about enrolling participants, Stage 1 of our design serves as an explorations stage, in which we obtain initial estimates of the subgroup treatment effects $\hat{\tau}_{1j}$ and their standard errors $\hat{\sigma}^2_{1j}$:

$$\begin{aligned}
\hat{\tau}_{1j} &= \frac{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} D_{i1} Y_{i1}}{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} D_{i1}} - \frac{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} (1-D_{i1}) Y_{i1}}{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} (1-D_{i1})}, \\
\hat{\sigma}^2_{1j} &= \frac{\sum_{i=1}^{n_1} \left(Y_{i1} D_{is} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} - \frac{1}{n_1}\sum_{i=1}^{n_1}\{Y_{i1} D_{i1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)}\}\right)^2}{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} D_{i1}} \\
&\quad - \frac{\sum_{i=1}^{n_1} \left(Y_{i1}(1-D_{i1}) \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} - \sum_{i=1}^{n_1}\{Y_{i1}(1-D_{i1}) \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)}\}\right)^2}{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1}\in\mathcal{S}_j)} (1-D_{i1})}.
\end{aligned}$$

Stage t, for $t = 2, \ldots, T-1$. *Obtain* $\hat{\boldsymbol{e}}^*_t$ *by solving the sample analogue of* ***Problem B****, that is*

$$\hat{\boldsymbol{e}}^*_t = \arg\max_{\boldsymbol{e}} \left\{ z : \sum_{j=1}^{m} \hat{p}_{t,j} e_j = c_1,\ c_2 \le e_j \le 1-c_2,\ \min_{2\le j\le m} \frac{(\hat{\tau}_{t-1,j} - \hat{\tau}_{t-1,1})^2}{2\big(\hat{\sigma}^2_{t-1,1}(e_1) + \hat{\sigma}^2_{t-1,j}(e_j)\big)} - z \ge 0 \right\}, \tag{3.4}$$

where

$$\begin{aligned}
\hat{\tau}_{t-1,j} &= \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is} Y_{is}}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is}} - \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is}) Y_{is}}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is})}, \\
\hat{\sigma}^2_{t-1,j}(e_j) &= \sum_{s=1}^{t-1} \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\left(Y_{is} D_{is} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} - \frac{1}{n_s}\sum_{i=1}^{n_s}\{Y_{is} D_{is} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)}\}\right)^2}{\hat{p}_{t-1,j} e_j} \\
&\quad - \sum_{s=1}^{t-1} \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\left(Y_{is}(1-D_{is}) \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} - \frac{1}{n_s}\sum_{i=1}^{n_s}\{Y_{is}(1-D_{is}) \mathbb{1}_{(X_{is}\in\mathcal{S}_j)}\}\right)^2}{\hat{p}_{t-1,j}(1-e_j)}.
\end{aligned}$$

Assign treatment according to the learned policy $\pi_t : \{\mathcal{H}_s\}_{s=1}^{t-1} \to \hat{\boldsymbol{e}}^*_t = (\hat{e}^*_{t,1}, \ldots, \hat{e}^*_{t,m})$.

In each Stage t, based on the newly collected data from the previous stage $\{\mathcal{H}_{t-1}\}$, we renew our understanding of the underlying data distribution and obtain a pair of updated estimates $(\hat{\tau}_{t-1,j}, \hat{\sigma}^2_{t-1,j})$ for each subgroup. These updated estimates thus enable us to better mimic the behavior of the oracle treatment allocation strategy by solving a refined optimization problem defined in Eq (3.4) and revise the treatment assignment accordingly. In Stage t, we assign treatments based on $\hat{\boldsymbol{e}}^*_T$ using the historical data collected up to Stage $T-1$.

Statistical inference after Stage T. *Construct the final subgroup treatment effect estimator along with its standard error:*

$$\hat{\tau}_j = \frac{n_1}{N} \cdot \Big(\frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} D_{i1} Y_{i1}}{\hat{e}_1 \cdot \hat{p}_j} - \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} (1 - D_{i1}) Y_{i1}}{(1 - \hat{e}_1) \cdot \hat{p}_j} \Big) \tag{3.5}$$

$$+ \frac{N - n_1}{N} \Big(\frac{1}{N - n_1} \sum_{s=2}^{T} \sum_{i=1}^{n_s} \frac{\mathbb{1}_{(X_{is} \in \mathcal{S}_j)} D_{is} Y_{is}}{\hat{e}^*_s \cdot \hat{p}_j} - \frac{1}{N - n_1} \sum_{s=2}^{T} \sum_{i=1}^{n_s} \frac{\mathbb{1}_{(X_{is} \in \mathcal{S}_j)} (1 - D_{is}) Y_{is}}{(1 - \hat{e}^*_s) \cdot \hat{p}_j} \Big),$$

$$\hat{\sigma}^2_j = \frac{n_1}{N} \Big(\frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\big(Y_{i1} D_{i1} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} - \bar{Y}_{1j}\big)^2}{\hat{e}_1 \cdot \hat{p}_j} + \frac{1}{n_1} \sum_{s=1}^{n_1} \frac{\big(Y_{i1}(1 - D_{i1}) \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} - \bar{Y}_{0j}\big)^2}{(1 - \hat{e}_1) \cdot \hat{p}_j} \Big) \tag{3.6}$$

$$+ \frac{N - n_1}{N} \Big(\frac{1}{N - n_1} \sum_{s=2}^{T} \sum_{i=1}^{n_s} \frac{\big(Y_{is} D_{is} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} - \bar{Y}_{1j}\big)^2}{\hat{e}^*_s \cdot \hat{p}_j} \tag{3.7}$$

$$+ \frac{1}{N - n_1} \sum_{s=2}^{T} \sum_{i=1}^{n_s} \frac{\big(Y_{is}(1 - D_{is}) \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} - \bar{Y}_{0j}\big)^2}{(1 - \hat{e}^*_s) \cdot \hat{p}_j} \Big).$$

where $\hat{p}_j = \frac{\sum_{s=1}^{T} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)}}{N}$,

$$\bar{Y}_{1j} = \frac{1}{N} \sum_{s=1}^{T} \sum_{i=1}^{n_s} Y_{is} D_{is} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)},$$

and

$$\bar{Y}_{0j} = \frac{1}{N} \sum_{s=1}^{T} \sum_{i=1}^{n_s} Y_{is} (1 - D_{is}) \mathbb{1}_{(X_{is} \in \mathcal{S}_j)}.$$

Then, identify the best subgroup as the one exhibiting the maximal treatment effect size:

$$j^* = \underset{1 \le j \le m}{argmax}\ \hat{\tau}_j. \tag{3.8}$$

Lastly, construct a two-sided level-α confidence interval for the selected best subgroup as

$$\Big[\hat{\tau}_{j^*} \pm \Phi^{-1}(1 - \alpha/2) \cdot \hat{\sigma}_{j^*} / \sqrt{N} \Big]. \tag{3.9}$$

Multi-stage case with small T and large n_t

In this section, we provide an alternative multi-stage design strategy with small T and large n_t, when experimenters can not revise the treatment assignment strategy too frequently. Stage 1 and the statistical inference after Stage t are the same in fully adaptive and multi-stage settings. In Stage t, however, the multi-stage setting requires an additional step, as shown below:

Stage t, for $t = 2, \ldots, T-1$. *(a) Solve for $\hat{\boldsymbol{e}}_t^*$ as in the fully adaptive setting. (b) In each subgroup, assign subjects to the treatment arm with probability $\tilde{e}_{tj}^*$, where*

$$\tilde{e}_{tj}^* = \frac{1}{n_{tj}}\Big((\hat{e}_{tj}^* \sum_{s=1}^{t} n_{sj}) - \sum_{s=1}^{t-1} n_{sj}\tilde{e}_{sj}^* \Big),\ j = 1, \ldots, m,\ t = 2, \ldots, T.$$

To understand the extra step in the multi-stage setting, we can consider a scenario with $T = 2$. Recall that our experimental goal is to maximize the correct selection probability of identifying the best subgroup, and the correct selection probability is maximized under the proposed oracle treatment allocation. Therefore, we want the actual treatment allocation after Stage 2 to approximate the oracle treatment allocation. Considering that the initial stage treatment allocation $\hat{e}_{1j}$ might be far away from the oracle treatment allocation, we want to account for the number of subjects already assigned to the treatment arm in Stage 1 when allocating subjects in Stage 2 following the step shown above.

3.5 Adaptive enrichment design

In this section, we start by introducing the oracle enrichment design. We then propose two design strategies for multi-stage and fully adaptive settings (Table 3.1).

Oracle subgroup enrichment proportions

Based on the Gartner-Ellis Theorem, we are able to derive the closed-form expression of the rate function in the subgroup enrichment design:

$$G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{p}) = \frac{(\tau_j - \tau_1)^2}{2\big(\sigma_1^2(p_1) + \sigma_j^2(p_j)\big)}, \quad \sigma_j^2(p_j) = \frac{\mathbb{E}[\sigma_1^2(X)|X \in \mathcal{S}_j]}{p_j e_j} + \frac{\mathbb{E}[\sigma_0^2(X)|X \in \mathcal{S}_j]}{p_j(1-e_j)}. \tag{3.10}$$

With the closed form of the rate function, we can find the oracle subgroup enrichment proportion $\boldsymbol{p}^* \triangleq (p_1^*, \ldots, p_m^*)$ that solves the following optimization problem:

Problem *C*

$$\begin{aligned} &\max_{\boldsymbol{e}} z, && \leftarrow \text{Linear objective function} \\ &\text{s.t. } \sum_{j=1}^{m} p_j = 1,\ p_j > 0,\ j = 1, \ldots, m, && \leftarrow \text{Feasibility constraints} \\ &\frac{(\tau_j - \tau_1)^2}{2\big(\sigma_1^2(p_1) + \sigma_j^2(p_j)\big)} - z \geq 0,\ j = 2, \ldots, m. && \leftarrow \text{Equivalent to maximize} \\ & && \quad \text{correct selection probability} \end{aligned}$$

The feasibility constraints suggest that the subgroup enrichment proportions should sum up to 1, and each subgroup enrichment proportion is non-negative. To provide some intuition on the oracle enrichment proportion $\boldsymbol{p}^*$, we present the closed-form expression of $\boldsymbol{p}^*$ in a simplified setting (Remark 8).

Remark 8. *In a simplified setting, we assume $\sigma_1^2(p_1) \ll \sigma_j^2(p_j)$. Let $\sigma_j^2 := \mathbb{E}[\sigma_1^2(X)|X \in \mathcal{S}_j]/e_j + \mathbb{E}[\sigma_0^2(X)|X \in \mathcal{S}_j]/(1-e_j)$. The closed-form expression of the oracle enrichment proportion is*

$$p_j^* = \frac{\beta_j}{\sum_{j=1}^m \beta_j},\ \beta_j = \begin{cases} \sigma_j^2/(\tau_j - \tau_1)^2 & j \neq 1, \\ \sigma_j \sqrt{\sum_{l \neq 1} \beta_l^2/\sigma_l^2} & j = 1. \end{cases} \tag{3.11}$$

From Eq (3.11), we observe that for subgroup j, the enrichment proportion depends on both the subgroup variance and the squared distance between subgroup j's treatment effect and the best subgroup's treatment effect. Intuitively, when τ_j is closer to τ_1 or when subgroup j has a larger variance σ_j^2, p_j^* increases, and experimenters will enroll a larger proportion of subjects from subgroup j.

Proposed adaptive enrichment designs

In this section, we illustrate our proposed enrichment design in the multi-stage setting with large n_t and small T (Table 3.1). The derived sequential policy $\boldsymbol{\pi}_{\text{MS-AE}} = (\pi_1, \ldots, \pi_{t-1})$ enables us to adaptively revise subgroup enrichment proportions so that the estimated subgroup treatment effect estimator achieves the similar performance as the one obtained by the oracle enrichment strategy.

Stage 1. *Enroll subjects from all the considered subgroups and set the subgroup enrichment proportion as $p_{1j} = \frac{1}{m}$. In each subgroup, randomly assign the subjects to the treatment arm*

with $e_{1j} = \frac{1}{2}$. *We estimate* $\hat{\tau}_{1j}$ *and* $\hat{\sigma}^2_{1j}$ *as*

$$\hat{\tau}_{1j} = \frac{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} D_{i1} Y_{i1}}{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} D_{i1}} - \frac{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} (1 - D_{i1}) Y_{i1}}{\sum_{i=1}^{n_1} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} (1 - D_{i1})},$$

$$\begin{aligned}\hat{\sigma}^2_{1j} = &\frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\left(Y_{i1} D_{is} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} - \frac{1}{n_1} \sum_{i=1}^{n_1} \{Y_{i1} D_{i1} \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)}\}\right)^2}{p_{1j} \cdot \sum_{i=1}^{n_1} D_{i1}} \\ &+ \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{\left(Y_{i1} (1 - D_{i1}) \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)} - \frac{1}{n_1} \sum_{i=1}^{n_1} \{Y_{i1} (1 - D_{i1}) \mathbb{1}_{(X_{i1} \in \mathcal{S}_j)}\}\right)^2}{p_{1j} \cdot \sum_{i=1}^{n_1} (1 - D_{i1})}.\end{aligned}$$

Here, we consider $p_{1j} = \frac{1}{m}$, $j = 1, \ldots, m$, meaning that an equal number of subjects are enrolled from each subgroup. Choosing equal enrichment proportions at the initial stage is sensible because we want to start exploring each subgroup with the same amount of experimental effort.

Stage t, for $t = 2, \ldots, T-1$. *(a) Obtain* $\hat{\boldsymbol{p}}^*_t$ *by solving for the sample analogue of* ***Problem C****, that is*

$$\hat{\boldsymbol{p}}^*_t = \arg\max_{\boldsymbol{p}} \Big\{ z : \sum_{j=1}^{m} p_j = 1,\ p_j > 0,\ j = 1, \ldots, m,\ \min_{2 \leq j \leq m} \frac{(\hat{\tau}_{t-1,j} - \hat{\tau}_{t-1,1})^2}{2\big(\hat{\sigma}^2_{t-1,1}(p_1) + \hat{\sigma}^2_{t-1,j}(p_j)\big)} - z \geq 0 \Big\},$$

where

$$\hat{\tau}_{t-1,j} = \frac{\sum_{s=1}^{t-1} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} D_{is} Y_{is}}{\sum_{s=1}^{t-1} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} D_{is}} - \frac{\sum_{s=1}^{t-1} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} (1 - D_{is}) Y_{is}}{\sum_{s=1}^{t-1} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} (1 - D_{is})},$$

$$\begin{aligned}\hat{\sigma}^2_{t-1,j}(p_j) = &\sum_{s=1}^{t-1} \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\left(Y_{is} D_{is} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} - \frac{1}{n_s} \sum_{i=1}^{n_s} \{Y_{is} D_{is} \mathbb{1}_{(X_{is} \in \mathcal{S}_j)}\}\right)^2}{p_j \hat{e}_{t-1,j}} \\ &+ \sum_{s=1}^{t-1} \frac{1}{n_s} \sum_{i=1}^{n_s} \frac{\left(Y_{is} (1 - D_{is}) \mathbb{1}_{(X_{is} \in \mathcal{S}_j)} - \frac{1}{n_s} \sum_{i=1}^{n_s} \{Y_{is} (1 - D_{is}) \mathbb{1}_{(X_{is} \in \mathcal{S}_j)}\}\right)^2}{p_j (1 - \hat{e}_{t-1,j})}.\end{aligned}$$

(b) Sample new subjects from each subgroup with probability $\tilde{p}^*_{tj}$*, where*

$$\tilde{p}^*_{tj} = \frac{1}{n_t} \Big(\big(\hat{p}^*_{tj} \sum_{s=1}^{t} n_s\big) - \sum_{j=1}^{t-1} n_{tj} \Big),\ j = 1, \ldots, m.$$

Following the similar reasoning in Section 3.4, we re-scale the estimated optimal subgroup enrichment proportion $\hat{p}^*_{tj}$ by the number of enrolled subjects in the previous stages. In Stage T, we enrich subjects from each subgroup based on $\hat{\boldsymbol{p}}^*_T$.

Statistical inference after Stage T. *Construct the final subgroup treatment effect estimator:*

$$\hat{\tau}_j = \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is} Y_{is}}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} D_{is}\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}} - \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is}) Y_{is}}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} (1-D_{is})\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}}, \tag{3.12}$$

$$\begin{aligned}\hat{\sigma}_j^2 = &\frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \left(Y_{is} D_{is}\mathbb{1}_{(X_{is}\in\mathcal{S}_j)} - \frac{1}{n_s}\sum_{i=1}^{n_s}\{Y_{is} D_{is}\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}\}\right)^2}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} D_{is}\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}} \\ &+ \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \left(Y_{is}(1-D_{is})\mathbb{1}_{(X_{is}\in\mathcal{S}_j)} - \frac{1}{n_s}\sum_{i=1}^{n_s}\{Y_{is}(1-D_{is})\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}\}\right)^2}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} (1-D_{is})\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}}.\end{aligned} \tag{3.13}$$

Then, identify the best subgroup as the one exhibiting the maximal treatment effect size:

$$j^* = \underset{1\le j\le m}{argmax}\ \hat{\tau}_j.$$

Lastly, construct a two-sided level-α confidence interval for the identified best subgroup j^* as

$$\left[\hat{\tau}_{j^*} \pm \Phi^{-1}(1-\alpha/2)\cdot \hat{\sigma}_{j^*}/\sqrt{N}\right],\ j = 1\ldots, m.$$

In a fully adaptive setting, after solving for $\hat{p}^*_{tj}$ in Stage t, we sample subjects from each subgroup with probability $\hat{p}^*_{tj}$. The initial stage and the inferential step after Stage T follow the same procedures as the multi-stage setting.

3.6 Theoretical investigation

In this section, we investigate the theoretical properties of our proposed adaptive experiment strategies. We start with listing assumptions followed by their interpretations. Then we introduce the theoretical properties of our proposed adaptive experiment strategies. Because the statistical properties of our proposed adaptive experiment in the fully adaptive case are similar to the multi-stage setting, we focus on the fully adaptive setting in this section. We work with the following assumptions in this section:

Assumption 16. *For $t = 1,\ldots,T$, $i = 1,\ldots,n_t$ and $d \in \{0,1\}$, the potential outcome $Y_{it}(d)$ has bounded first and $2+\delta$ moments for some $\delta > 0$, that is $\mathbb{E}|Y_{it}(d)| < \infty$ and $\mathbb{E}|Y_{it}(d)|^{2+\delta} < \infty$.*

Assumption 16 puts a mild moment condition on the potential outcomes over different stages.

Assumption 17. *There are total $m \ge 2$ subgroups, and the subgroup treatment effects can be monotonically ordered with $\tau_1 > \tau_2 > \ldots > \tau_m$.*

For simplicity, here we assume there are no exact ties among the population subgroup treatment effects.

Assumption 18. *We assume subgroup proportions p_j are bounded away from 0 and 1 by a positive constant, $j = 1, \ldots, m$.*

Assumption 18 is required for the response adaptive design, which assumes that the proportion of each considered subgroup be bounded away from 0 and 1.

In what follows, we first establish the strong consistency result of the treatment allocation probabilities and the enrichment proportions adopted in our designs:

Lemma 5. *For $j = 1, \ldots, m$, as $T \to \infty$, for the response adaptive design, under Assumptions 16-18, we have*

$$\frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} D_{is}\mathbb{1}_{(X_{is}\in\mathcal{S}_j)}}{N_j} \to e_j^*, \quad \textit{almost surely},$$

where $N_j = N \cdot p_j$. For the adaptive enrichment design, under Assumptions 16-17,

$$\frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}(X_{is} \in \mathcal{S}_j)}{N} \to p_j^* \quad \textit{almost surely.}$$

Lemma 5 suggests that under our proposed response adaptive design and adaptive enrichment design, the total number of subjects allocated to the treatment arm in subgroup j or enrolled from subgroup j converges to the oracle treatment allocation or the oracle enrollment proportions almost surely as the number of stages (or sample size, equivalently) goes to infinity. In other words, although our proposed designs in Section 3.4 and 3.5 utilize no prior knowledge about the underlying data distribution before the experiment starts, they can allocate experimental efforts in a similar fashion to the oracle strategies when the sample size is sufficiently large. Given the stochastic nature of our proposed designs, we prove the above result leveraging the concept of downcrossing and stopping times [73]. The proof is provided in Supplementary Materials Section B.

Theorem 4. *Suppose $\frac{n_1}{N} \to c$, with c being a positive constant between 0 and 1, as $N \to \infty$, the subgroup treatment effects and the standard deviations of the estimated subgroup treatment effects can be consistently estimated almost surely. For the response adaptive design, under Assumptions 17 to 18,*

$$\hat{\tau}_j^{RA} - \tau_j = O\Big(\sqrt{\frac{\log\log N}{N}}\Big), \textit{ almost surely}, \quad \hat{\sigma}_j^{2^{RA}} - \sigma_j^2(e_j^*) = O\Big(\sqrt{\frac{\log\log N}{N}}\Big), \textit{ almost surely},$$
$$\textit{where } \sigma_j^2(e_j^*) = \frac{\mathbb{E}[\sigma_1^2(X)|X \in \mathcal{S}_j]}{p_j e_j^*} + \frac{\mathbb{E}[\sigma_0^2(X)|X \in \mathcal{S}_j]}{p_j(1-e_j^*)}.$$

For the adaptive enrichment design, under Assumptions 17 to 16,

$$\hat{\tau}_j^{AE} - \tau_j = O\Big(\sqrt{\frac{\log\log N}{N}}\Big), \text{ almost surely}, \quad \hat{\sigma}_j^{2^{AE}} - \sigma_j^2(p_j^*) = O\Big(\sqrt{\frac{\log\log N}{N}}\Big), \text{ almost surely},$$
$$\text{where } \sigma_j^2(p_j^*) = \frac{\mathbb{E}[\sigma_1^2(X)|X \in \mathcal{S}_j]}{p_j^* e_j} + \frac{\mathbb{E}[\sigma_0^2(X)|X \in \mathcal{S}_j]}{p_j^*(1-e_j)}.$$

Lastly, the estimated subgroup treatment effects converge to Gaussian distributions when N tends to infinity, that is

$$\sqrt{N}\big(\hat{\tau}_j^{RA} - \tau_j\big) \to \mathcal{N}\big(0,\ \sigma_j^2(e_j^*)\big), \quad \sqrt{N}\big(\hat{\tau}_j^{AE} - \tau_j\big) \to \mathcal{N}\big(0,\ \sigma_j^2(p_j^*)\big).$$

where $j = 1, \ldots, m$.

The first part of the consistency results provided above leverages the law of the iterated logarithm and martingale methods [79, 71], and demonstrates that the standard error for each subgroup converges to the one reached by the oracle strategy. The second part of the asymptotic normality result leverages the martingale central limit theorem [71].

The theoretical results established in Theorem 4 suggest that the selected best subgroup treatment effect $\hat{\tau}_{j^*}$ is a strongly consistent estimate of τ_1 and its variance can be well estimated by $\hat{\sigma}_{j^*}^2$. This suggests that the constructed confidence interval for the best subgroup in Eq (3.9) reaches the nominal coverage when the sample size goes to infinity.

In this section, we have demonstrated the statistical validity of our design. In particular, our designs provide competing performances with the ones given by the oracle strategies. To further demonstrate the benefits of our design, in the next section, we compare the asymptotic property of our proposed response adaptive design with the one reached by the completely randomized design.

Comparison with completely randomized experiments

In this section, we shall compare our proposed response adaptive design with completely randomized experiments in which the treatment is randomly assigned, often with 1/2 probability throughout the entire experiment. The comparisons shall be illustrated from three aspects: (1) the large deviation rate, (2) the semiparametric efficiency, (3) and the distance between τ_1 and τ_2.

First, we shall compare the large deviation rates between the two designs. We start with establishing a finite sample lower bound on the correct selection probability in Theorem 6 followed by the large deviation rate comparion in Lemma 5.

Lemma 6. *Under Assumptions 17-18, suppose $\frac{n_1}{N} \to c$ with c being a positive constant between 0 and 1, there exist some positive constants $C_1, \ldots, C_m$ such that the correct selection probability is lower bounded by:*

$$\mathbb{P}(\hat{\tau}_1(\hat{e}_1^*) \geq \max_{2\leq j\leq m} \hat{\tau}_j(\hat{e}_j^*)) \geq 1 - C_1 \exp\Big(-\frac{(\tau_1-\tau_2)^2 cN}{8\sigma_1^2(e_1^*)m}\Big) - \sum_{j=2}^{m} C_j \exp\Big(-\frac{(\tau_2 - \tau_j + \frac{\tau_1-\tau_2}{2})^2}{2\sigma_j^2(e_j^*)m}\Big).$$

The above theoretical result suggests that the identified best subgroup j^* after Stage T (see Eq (3.8) in Section 3.4) is the best subgroup (i.e., the subgroup with the largest treatment effect) with high probability when the number of stages goes to infinity. Note that $\frac{n_1}{N} \to c$ is commonly assumed in response adaptive design literature [79]. Following Theorem 6, we compare the large deviation rates under the proposed response adaptive design and the complete randomization design in Lemma 5.

Theorem 5 (Large deviation rate comparison)**.** *Under Assumptions 17-18 and suppose* $\frac{n_1}{N} \to c$ *with* c *being a positive constant between 0 and 1,*

$$\lim_{N\to\infty} \frac{1}{N} \log\left(1 - \mathbb{P}(\hat{\tau}_1(\hat{e}_1^*) \geq \max_{2\leq j\leq m} \hat{\tau}_j(\hat{e}_j^*))\right) \leq \lim_{N\to\infty} \frac{1}{N} \log\left(1 - \mathbb{P}(\hat{\tau}_1(1/2) \geq \max_{2\leq j\leq m} \hat{\tau}_j(1/2))\right).$$

Lemma 5 suggests that the large deviation rate under our proposed response adaptive design converges to zero exponentially faster than the complete randomization as the sample size goes to inifinity. In other words, our proposed response adaptive design is able to correctly identify the best subgroup with an efficient use of the samples.

Second, we shall compare the asymptotic efficiency gain of the proposed design for estimating the best subgroup treatment effect with the complete randomization design. To provide a fair comparison, we again do not restrict the potential outcome to follow any parameter models. In this case, as the treatment is randomly assigned independent of the potential outcomes, any subgroup treatment effect estimators obtained from response adaptive designs and complete randomization share the same form of a variance lower bound derived by [69], that is

$$\mathbb{V}_j(e_j) = \frac{\mathbb{E}[\sigma_1^2(X)|X \in \mathcal{S}_j]}{p_j e_j} + \frac{\mathbb{E}[\sigma_0^2(X)|X \in \mathcal{S}_j]}{p_j(1-e_j)} + \frac{\mathbb{V}[\mathbb{E}(Y(1)-Y(0)|X \in \mathcal{S}_j)]}{p_j}. \tag{3.14}$$

Based on the treatment assignment probability adopted in different designs, the achievable variance lower bound for complete randomization is obtained by replacing e_j with $1/2$, that is $\mathbb{V}_j(1/2)$, and the achievable variance lower bound for the response adaptive design is obtained by replacing e_j with e_j^*, that is $\mathbb{V}_j(e_j^*)$. Note that the variance lower bound $\mathbb{V}_j(e_j^*)$ is achievable under our design when the augmented IPW estimator, instead of IPW, is adopted to estimate the subgroup treatment effects. These two variance lower bounds thus allow us to compare the performance of our design with complete randomization when the most efficient estimator is adopted to estimate the subgroup treatment effects. To provide further insights into this comparison, we consider a simple case formalized in Proposition 2.

Proposition 2 (Efficiency comparison)**.** *In a simplified setting, we assume (1)* $\mathbb{V}[Y(1)|X \in \mathcal{S}_j] = \mathbb{V}[Y(0)|X \in \mathcal{S}_j] = \sigma_j^2$*,* $j = 1, \ldots, m$*, and (2)* $\mathbb{V}[Y(1)|X \in \mathcal{S}_1] = 0$*. Then,*

$$\begin{cases} \mathbb{V}_j(e_j^*) \leq \mathbb{V}_j(1/2), & \text{if } \frac{(\tau_j-\tau_1)^2}{\sigma_j^2} \leq 4, \\ \mathbb{V}_j(e_j^*) > \mathbb{V}_j(1/2), & \text{if } \frac{(\tau_j-\tau_1)^2}{\sigma_j^2} > 4. \end{cases}$$

Proposition 2 shows the efficiency comparison between our proposed response adaptive design and the complete randomization design. When estimating the best subgroup treatment effect, the asymptotic variance under our proposed response adaptive design is always smaller than the complete randomization design. However, when τ_j is far away from τ_1 or when the expected variance of the outcome in subgroup j is small, our proposed response adaptive design is less likely to have efficiency gain. Proposition 2 entails the efficiency trade-off between our proposed design and the complete randomization design.

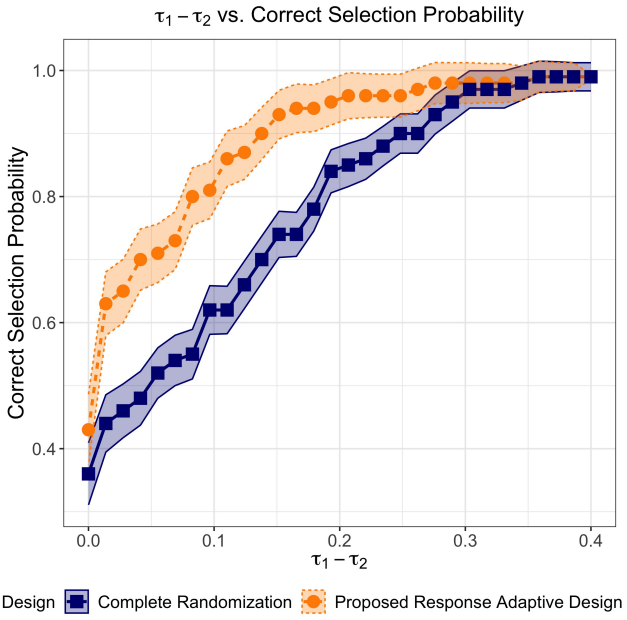


Figure 3.2: Comparison between the proposed response adaptive design and the complete randomization design with respect to various distances between τ_1 and τ_2.

Next, we shall compare the allowed distance between τ_1 and τ_2 in order to reach a fixed correct selection probability level under the two designs.

Proposition 3 (Distance comparison)**.** *Assume* $\mathbb{V}[Y(1)|X \in \mathcal{S}_j] = \mathbb{V}[Y(0)|X \in \mathcal{S}_j] = \sigma_j^2$. *Suppose we aim to reach a correct selection probability of at least* $1 - \varepsilon$. *For some positive constants* $C < \infty$ *and* $C' < \infty$, *under the complete randomization design, the distance between* τ_1 *and* τ_2 *is characterized as*

$$\tau_1 - \tau_2 \geq N^{-\frac{1}{2}} \cdot \left| C \cdot \log(\varepsilon) \cdot \sigma^2 \right|^{\frac{1}{2}}.$$

Under our proposed response adaptive design, the distance between τ_1 *and* τ_2 *is characterized as*

$$\tau_1 - \tau_2 \geq N^{-\frac{2}{3}} \cdot \left| C' \cdot \log(\varepsilon) \cdot \sigma^2 \right|^{\frac{1}{2}}.$$

Proposition 3 says that to reach a correct selection probability level of $1-\varepsilon$, while the completely randomizated design requires τ_1 to be well-separated from τ_2 by a distance of $N^{-\frac{1}{2}}$ order, our proposed adaptive design strategy allows τ_2 to stay within the $N^{\frac{1}{2}}$-neighborhood of τ_1. This implies that our proposed design allows τ_2 to be closer to τ_1 compared to the completely randomized experiment. Furthermore, we note that methods performing valid post hoc analysis on τ_1 require the distance between τ_1 and τ_2 to be a positive constant, and our approach relaxes this assumption and allows the second best subgroup treatment effect τ_2 to stay in a vanishing neighborhood of the best subgroup.

To verify Proposition 3, we provide a simple simulation in Figure 3.2. We set $\tau_1 = 1.6$, $\tau_3 = 1.4$, and $\tau_2 = \tau_1 - \delta$, where $\delta = c(0.03, 0.06, 0.09, 0.12, 0.15)$. We compare the correct selection probability under the proposed design and the complete randomization design with respect to various distances between τ_1 and τ_2. Figure 3.2 demontrates that to reach a certain correct selection probability level, our proposed design allows τ_2 to stay closer to τ_1. In other words, when τ_1 is close to τ_2, our proposed response adaptive design correctly distinghuish the best subgroup from the second best subgroup with a higher probability.

The implication of Figure 3.2 also provides some insights regarding the winner's curse bias mitigation under our proposed design, which shall be illustrated in Remark 9.

Remark 9. *The winner's curse bias of the estimated best subgroup treatment effect can be expressed as*

$$\mathbb{E}[\hat{\tau}_1] - \tau_1 = \mathbb{E}\big[\mathbb{E}[\hat{\tau}_1 | \mathbb{1}(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j)]\big] - \tau_1 = \mathbb{E}[\hat{\tau}_1] \cdot \mathbb{P}(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j) - \tau_1.$$

The winner's curse bias not only depends on the consistent estimation of τ_1 but also the correct identification of the best subgroup membership. Given that our adaptive experiment strategy is specifically designed for best subgroup identification, the winner's curse bias is mitigated by maxizing the probability of correctly selecting subgroup 1 as the best subgroup.

3.7 Simulation studies

In this section, we investigate the performance of our proposed response adaptive design and adaptive enrichment design. We summarize the takeaways from our simulation studies as follows. First, our proposed adaptive experiment strategies are able to maximize the correct selection probability. Compared to other conventional experimental strategies, our proposal takes a smaller sample size to reach the same level of correct selection probability. Second, under the proposed adaptive experiment strategies, the winner's curse bias on estimating the best subgroup treatment effect is mitigated. Third, compared to complete randomization, our proposed response adaptive design has a smaller variance when estimating the best subgroup treatment effect, which further confirms the efficiency comparison in Section 3.6.

Synthetic case studies

Our simulation design mimics adaptive experiments in our first case study. We have four non-overlapping subgroups representing clothing category, and we denote the subgroup membership for each subject i as $\boldsymbol{\mathcal{S}} = \big(\mathbb{1}(X_i \in \mathcal{S}_1), \ldots, \mathbb{1}(X_i \in \mathcal{S}_4)\big)^\intercal$. We generate the potential outcome as

$$Y_i(d)|X_i \in \mathcal{S}_j \sim \mathcal{N}(\mu_{d,j},\ \sigma_{d,j}).$$

We obtain the following parameters from the real data: $\boldsymbol{\mu}_1 = (4.14, 4.12, 4.43, 4.48)^\intercal$, $\boldsymbol{\mu}_0 = (4.83, 3.72, 4.02, 4.31)^\intercal$, $\boldsymbol{\sigma}_1 = (1.17, 1.06, 0.80, 0.90)^\intercal$, and $\boldsymbol{\sigma}_0 = (0.39, 1.57, 1.23, 1.10)^\intercal$. The subgroup proportions are $\boldsymbol{p} = (0.20, 0.16, 0.56, 0.08)^\intercal$. We denote

$$\boldsymbol{\tau} = (-0.69, 0.38, 0.41, 0.18)^\intercal$$

as the true subgroup treatment effects. The treatment assignment D_i is decided based on our adaptive experiment strategies, which shall be discussed later in the section. We investigate two adaptive experiment strategies: (1) response adaptive design and (2) adaptive enrichment design. For each adaptive experiment strategy, we consider the *multi-stage* setting and the *fully adaptive* setting.

For the response adaptive design in the multi-stage setting, we consider various sample sizes for the first stage, that is, $n_1 \in \{150, 222, 294, \ldots, 800\}$. For the number of experiment stages, we consider $T = 2$ and $T = 4$. When $T = 2$, we set $n_2 = 200$; when $T = 4$, we set $n_2 = n_3 = n_4 = 100$. In the multi-stage setting, we compare two methods: (1) our proposed response adaptive design and (2) the complete randomization design. The complete randomization design is a non-adaptive design commonly adopted in traditional randomized controlled trials. In the complete randomization design, one sets $e_{tj} = \frac{1}{2}$ across all the stages, $t = 1, \ldots, T,\ j = 1, \ldots, m$. Our proposed response adaptive design in multi-stage settings follows the procedures in Section 3.4.

For the response adaptive design in the fully adaptive setting, we fix stage 1 sample size as $n_1 = 250$ and $n_t = 1$ for $t = 2, \ldots, T$, where $T \in \{30, 90, 150, 210, 270, 330, 400\}$. The fully adaptive setting is commonly seen in conventional response adaptive designs. Therefore, besides the complete randomization design, we will also compare our proposed response adaptive design with an existing response adaptive design. In this setting, we compare three designs: (1) our proposed response adaptive design, (2) the complete randomization design, and (3) the doubly adaptive biased coin design [183].

For the adaptive enrichment design, we follow the same simulation setup as in the response adaptive design. In the fully-adaptive setting, we compare two designs: (1) our proposed adaptive enrichment design and (2) the equal enrichment design. The equal enrichment design is a non-adaptive design, where one sets $p_{tj} = \frac{1}{m}$, $t = 1, \ldots, T$. For our proposed enrichment design, we follow the procedures in Section 3.5. In the multi-stage setting, we compare three designs: (1) our proposed adaptive enrichment design, (2) equal enrichment design, and (3) adaptive enrichment with combination testing. The combination testing approach distributes the type I error rate across experimental stages. Based on the

computed type I error rate each stage aims to reach, one can estimate the corresponding enrichment proportions. We implement the combination testing approach using `R` package `rpact` [106].

We evaluate the performance of each adaptive experiment strategy from two aspects. First, we compare the experimental efforts (i.e., sample size) needed to reach various correct selection probability levels: $\{0.75, 0.8, 0.85, 0.9, 0.95, 0.99\}$. Second, we compare the $\sqrt{N}$-scaled bias of the estimated best subgroup treatment effect.

Simulation results

Figure 3.3 and Figure 3.4 show the comparison of our proposed response adaptive design with the other conventional designs in the fully adaptive setting and the multi-stage setting, respectively. From the two figures, we observe that our proposed response adaptive design has efficiency gain in two aspects. *First*, our proposed design makes efficient use of the experimental efforts. More concretely, to reach the same correct selection probability level, our proposed response adaptive design takes a much smaller sample size compared to the other two designs. We conjecture that the reason is two-fold. On the one hand, the complete randomization design does not revise treatment allocation adaptively. Thus the design may continue assigning subjects to the subgroup that has already been well-explored. On the other hand, although the doubly adaptive biased coin design is a response adaptive design, its experimental goal is to minimize the variance of the estimated treatment effect instead of identifying the best subgroup. While our proposed response adaptive design is tailored to the goal of maximizing the probability of correctly identifying the best subgroup.

Second, our proposed response adaptive design is efficient in estimating the best subgroup treatment effect. The bias of the estimated best subgroup treatment effect has a smaller variance compared to the other conventional designs, which confirms our theoretical results in Section 3.6. In Section 3.6, we show that our proposed response adaptive design has an efficiency gain in estimating the best subgroup treatment effect. We observe a similar efficiency gain in both fully adaptive and multi-stage settings.

Besides the above-mentioned efficiency gain, Figure 3.4 also shows that when estimating the best subgroup treatment effect, the complete randomization design and the doubly adaptive biased coin design yield upward biases, especially when the sample size is small. The upward bias is the winner's curse bias [45, 8]. Our proposed method can mitigate the winner's curse bias when estimating the best subgroup treatment effect.

Figure 3.5 and Figure 3.6 show the comparison of our proposed adaptive enrichment design with the conventional enrichment designs. These two figures demonstrate that our proposed adaptive enrichment design takes the smallest sample size to reach the same correct selection probability as the conventional designs. As for the bias comparison, our proposed adaptive enrichment design yields a smaller bias in estimating the best subgroup treatment effect when the sample size is small.

The simulation results confirm the theoretical investigations in Section 3.6. In practice, when an experimenter can only enroll a limited number of subjects, our proposed adaptive

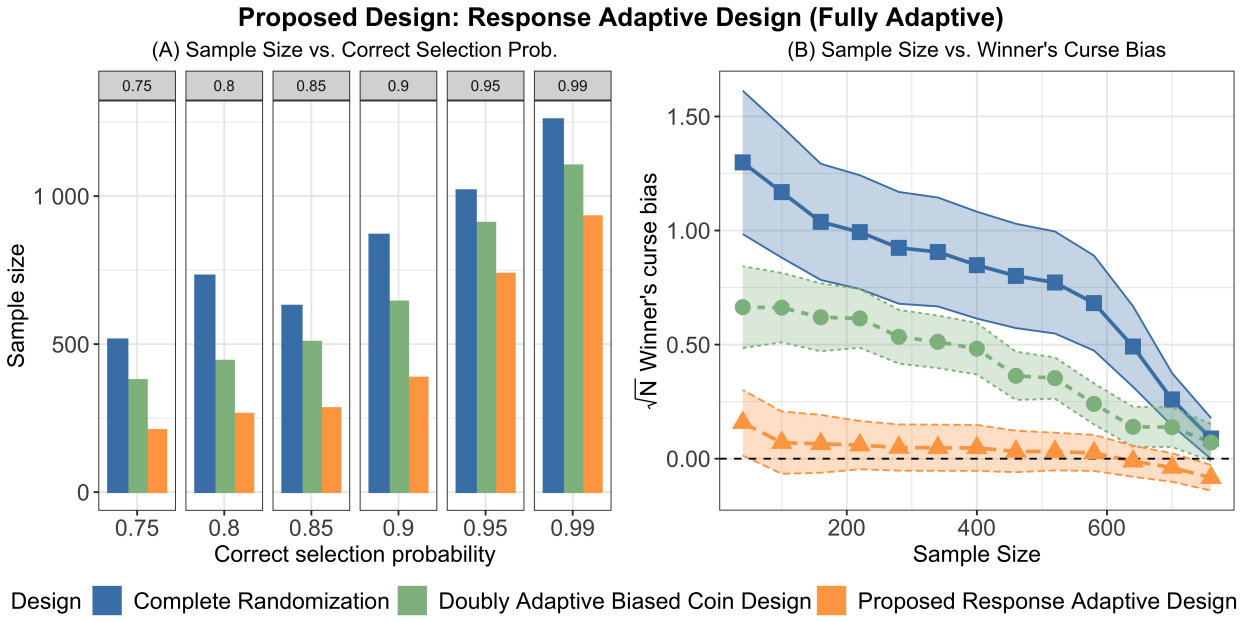


Figure 3.3: Comparison of the proposed response adaptive design, the complete randomization design, and the doubly adaptive biased coin design under the fully adaptive setting. (A) shows the sample size comparison under various correct selection probability levels. (B) shows the $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

experiment strategies demonstrate more efficient use of the samples to identify the best subgroup with a higher probability and can reduce the winner's curse bias on estimating the best subgroup treatment effect.

3.8 Case studies

In this section, we investigate the application of our proposed adaptive experiment strategy in two case studies. In the first case study, we investigate which clothing category would benefit the most from an inclusive advertising strategy. In the second study, we investigate which patient subgroup would benefit the most from genetically-guided therapy.

Case study I

For online clothing websites, an essential advertising strategy is to display images of human models wearing clothing products. Some studies suggest that there exists a "pro-thin" bias in fashion advertising; that is, clothing websites tend to display idealized human models wearing size small clothes [109, 1]. In light of the recent social campaigns calling for fashion marketing inclusiveness, some fashion companies have revised their advertising strategies by

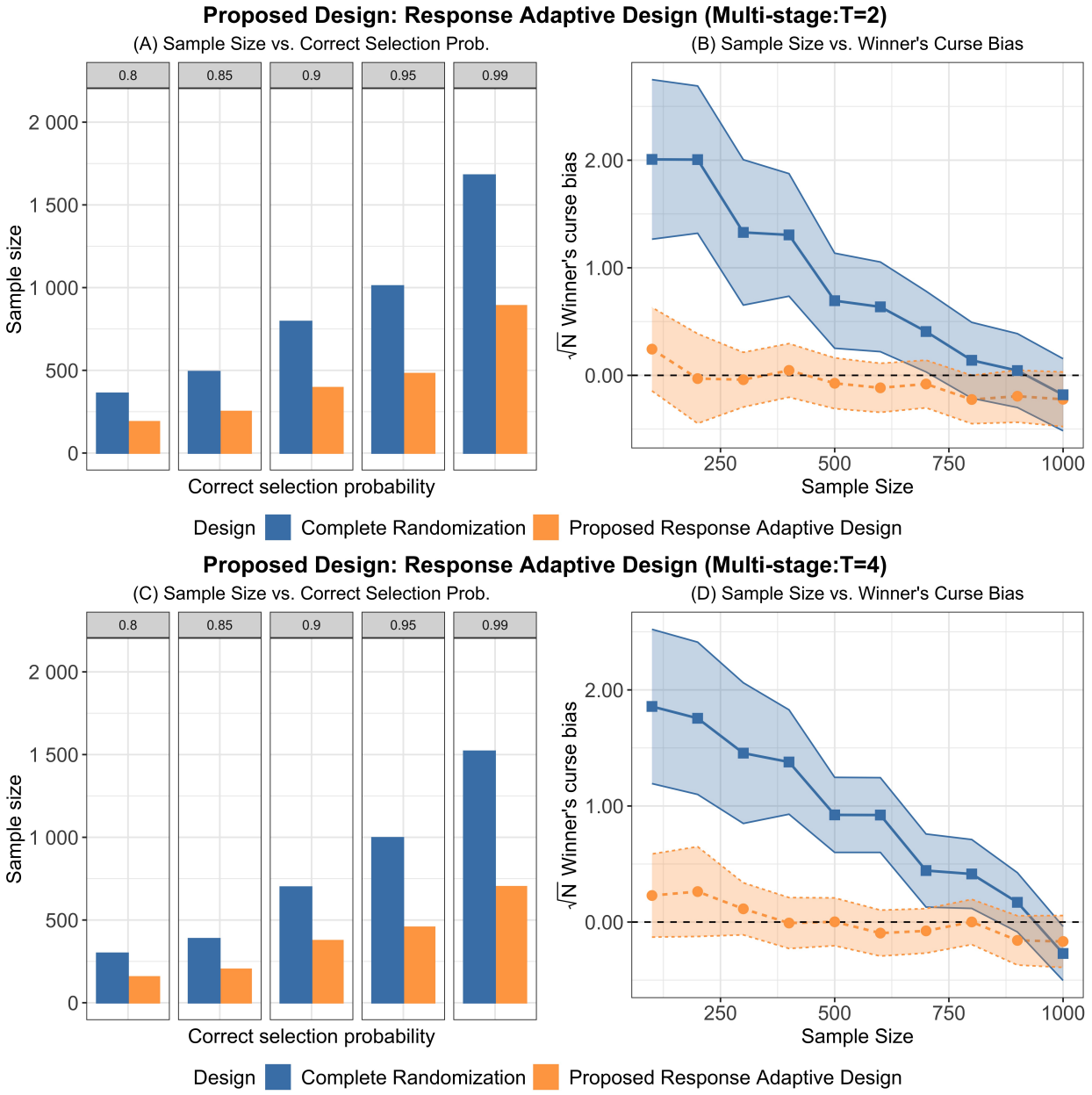


Figure 3.4: Comparison of the proposed response adaptive design and the complete randomization design under the multi-stage setting ($T = 2$ and $T = 4$). (A) shows the sample size comparison under various correct selection probability levels. (B) shows the $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

incorporating images of human models in a broader range of body shapes [34]. Although it is hypothesized that the inclusive advertising strategy could improve customer satisfaction,

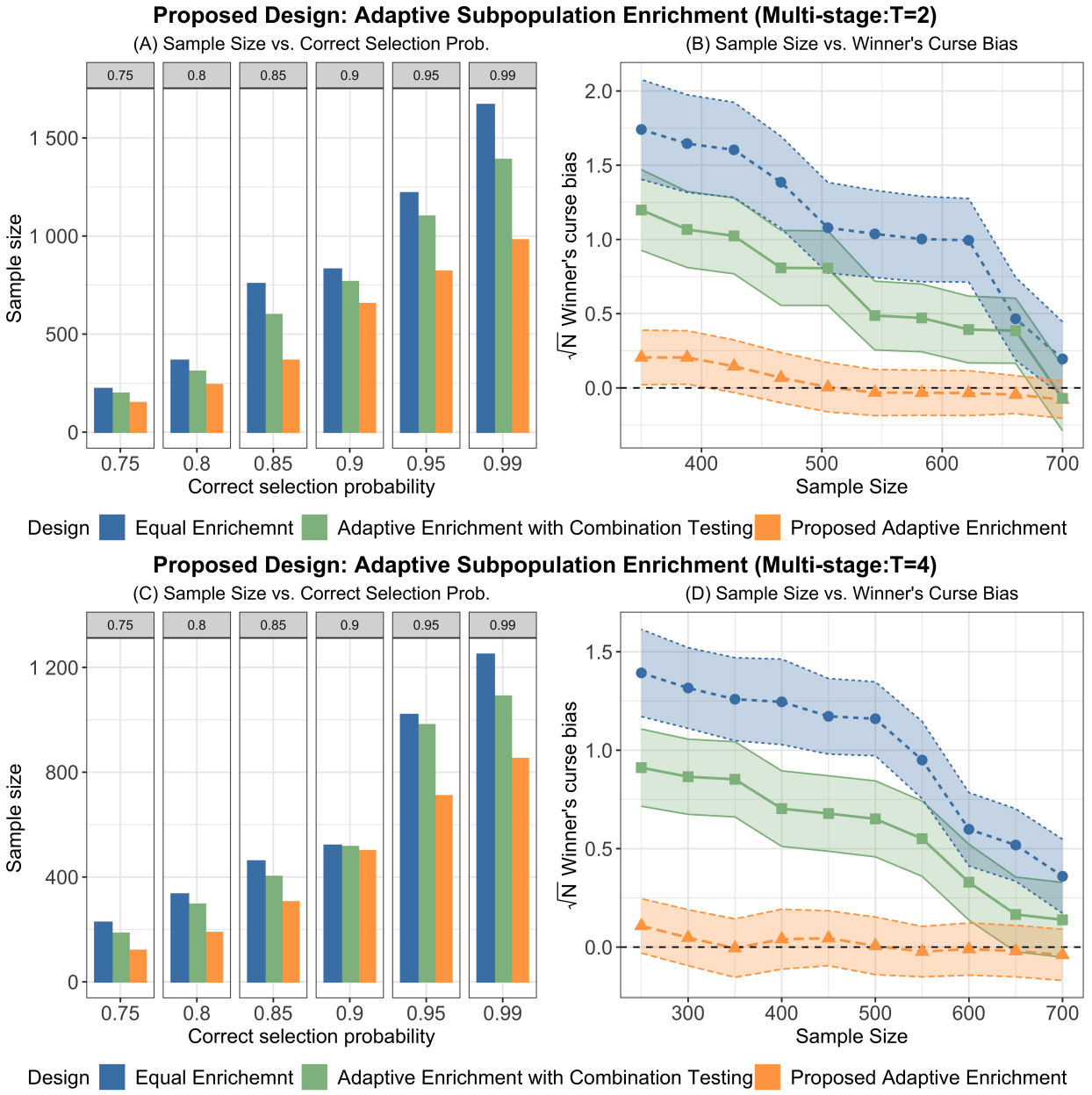


Figure 3.5: Comparison of the proposed adaptive subgroup enrichment design and the equal enrichment design under the multi-stage setting ($T = 2$ and $T = 4$). (A) shows the sample size comparison under various correct selection probability levels. (B) shows the $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

it remains unknown which clothing category benefits the most from the inclusive advertising strategy [90]. In this case study, our goal is to identify which clothing category benefits the

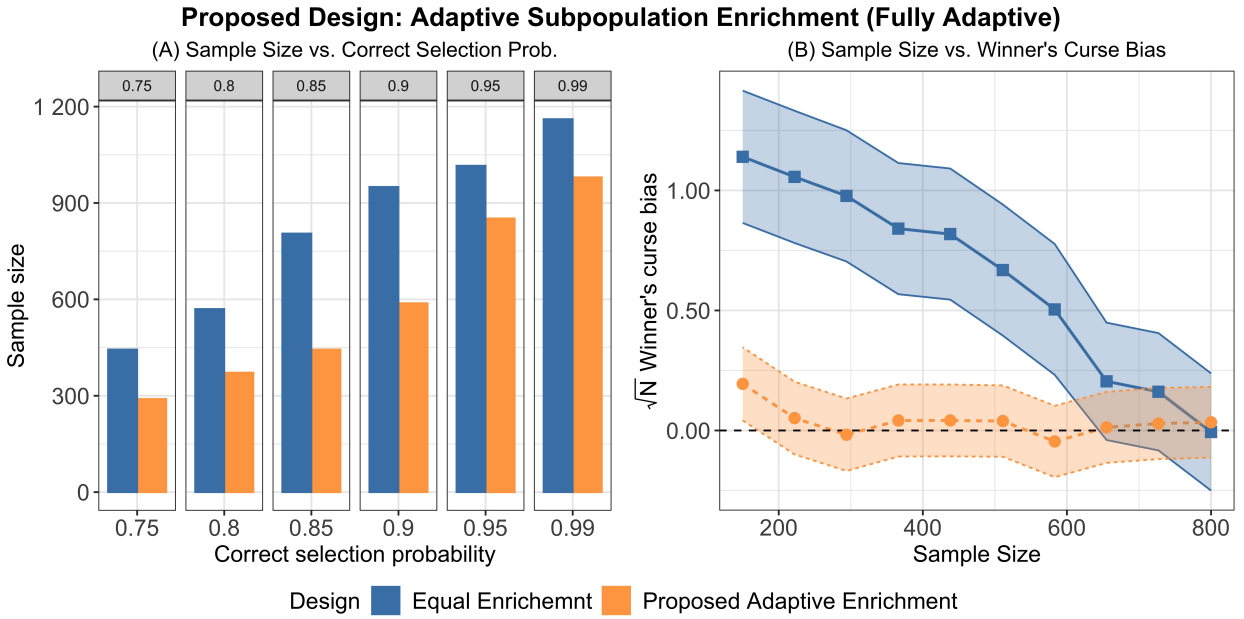


Figure 3.6: Comparison of the proposed adaptive subgroup enrichment design and the equal enrichment design under the fully adaptive setting. (A) shows the sample size comparison under various correct selection probability levels. (B) shows the $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

most from displaying inclusive body shape images.

In this case study, we use e-commerce data from ModCloth, a women's clothing website. The data are collected and processed by [171]. The original dataset contains 99, 893 observations collected from 2010 to 2019. For each clothing product, the website may display one of the two types of human model images: (1) a model wearing size small; (2) a model wearing size small, and a model wearing size large (Figure 3.7). We define the treatment variable as $D = 1$ if displaying "small/large" images and $D = 0$ if only displaying "small" images. We consider four clothing categories: (1) bottoms, (2) tops, (3) dresses, and (4) outwear. To quantify customer satisfaction, we use customer ratings on a scale from 0 to 5.

We assume that customers visit the website sequentially. At each time point t, one customer views only one product. We adaptively revise the probability of displaying inclusive body shape images. To initialize the experiment, we enroll $n_1 = 300$ customers in the first stage. We randomly sample $n_t = 1$ customer in the following stages, where $t = 2, \ldots, 100$. We compare three adaptive experiment strategies: (1) our proposed response adaptive design, (2) the complete randomization design, and (3) the doubly adaptive biased coin design. We then evaluate the treatment effect of the inclusive advertising strategy (i.e., display "small/large" images) in each clothing category. The results are summarized in Figure 3.8.

Figure 3.8 shows that, under our proposed response adaptive design, the advertising

Figure 3.7: An example of two different advertising strategies taken from ModCloth website. The left panel shows an inclusive advertising strategy of displaying both small and plus-size human models. The right panel shows a conventional advertising strategy of only displaying human models wearing size small.

strategy of displaying inclusive body shape images has the largest improvement in customer ratings in the outwear clothing category. Compared to the other two conventional designs, our proposed response adaptive design estimates the treatment effect in the outwear clothing category with high efficiency. In the other clothing categories, our response adaptive design does not have significant efficiency gains, which supports our theoretical results in Section 3.6 (Proposition 2). The case study results suggest that, on the one hand, under our proposed response adaptive design, one can identify the clothing category that benefits the most from the inclusive advertising strategy with smaller experimental efforts. On the other hand, the inclusive advertising strategy could have practical marketing benefits and positive social impacts because such an advertising strategy could not only improve customer satisfaction but also boost customer self-esteem and reduce body-focused anxiety [34, 90].

Case study II

CYP2D6 (cytochrome P450 2D6)-substrate drugs are commonly prescribed for treating major depressive disorder (MDD), yet MDD patients may respond differently to CYP2D6-substrate drugs due to patients' heterogeneous metabolic behavior [15]. The heterogeneous metabolic behavior is related to the CYP2D6 enzyme, which catalyzes the metabolism of CYP2D6-substrate drugs, and this enzyme is encoded by the *CYP2D6* gene [146]. With the presence of polymorphisms in the *CYP2D6* gene, the side effects of taking CYP2D6-substrate drugs might be prominent [129, 10]. In this case, a genetically-guided therapy that

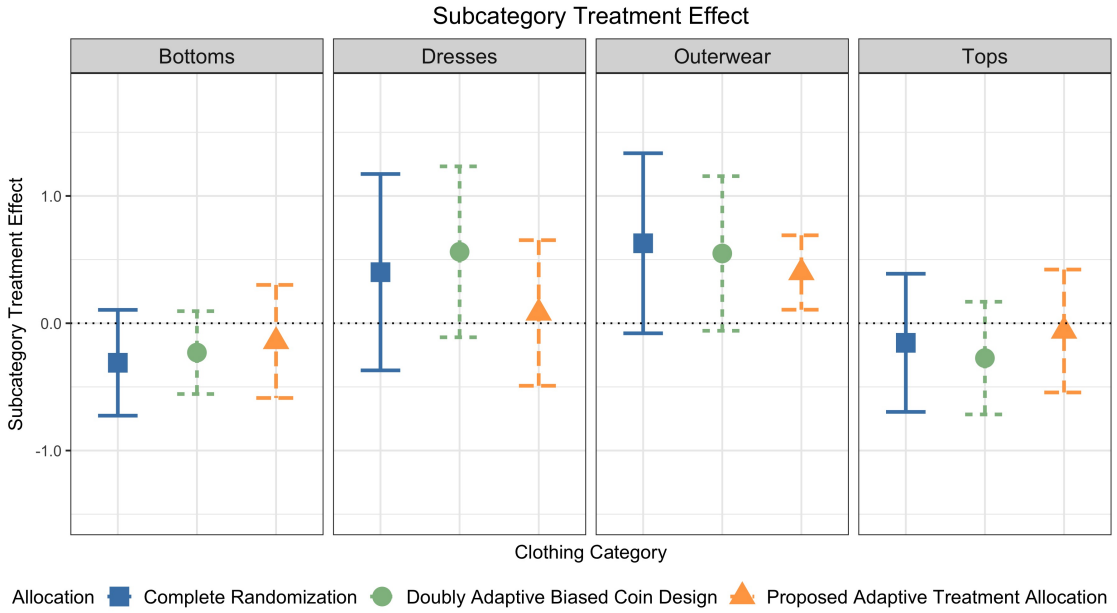


Figure 3.8: The estimated treatment effects in the four clothing categories under the complete randomization design, the doubly adaptive biased coin design, and our proposed response adaptive design. The initialization stage includes $n_1 = 300$ customers. In the following stages, we have $n_t = 1$, where $t = 2, \ldots, 100$.

prescribes CYP2D6-substrate drugs based on the *CYP2D6* gene could potentially provide a more precise and beneficial treatment for MDD patients. In this study, we use synthetic data from CYP-GUIDES (Cytochrome Psychotropic Genotyping Under Investigation for Decision Support) trial [147] to investigate which MDD patient subgroup would benefit the most from the genetically-guided therapy.

The original trial is conducted at the Institute of Living at Hartford Hospital, consisting of 1459 patients that are genotyped for *CYP2D6* [156]. In this study, there are two considered therapies: (1) the standard therapy ($D = 0$), which does not rely on CYP2D6 functional status; (2) the genetically-guided therapy ($D = 1$), in which the drugs are prescribed based on CYP2D6 functional status. More specifically, for the genetically-guided therapy, patients who have abnormal CYP2D6 function are proscribed medications metabolized by the CYP2D6 enzyme. The outcome of interest is the length of stay in the hospital, measured by hours. The shorter the length of stay, the more beneficial the therapy is. In addition, we consider six patient age subgroups: (1) 18-20; (2) 21-30; (3) 31-40; (4) 41-50; (5) 51-60; (6) > 60.

We assume that patients are administered sequentially, and physicians treat one patient at a time. We adaptively revise the probability of applying genetically-guided therapy. To initialize the experiment, we enroll $n_1 = 100$ patients in the first stage. We randomly sample

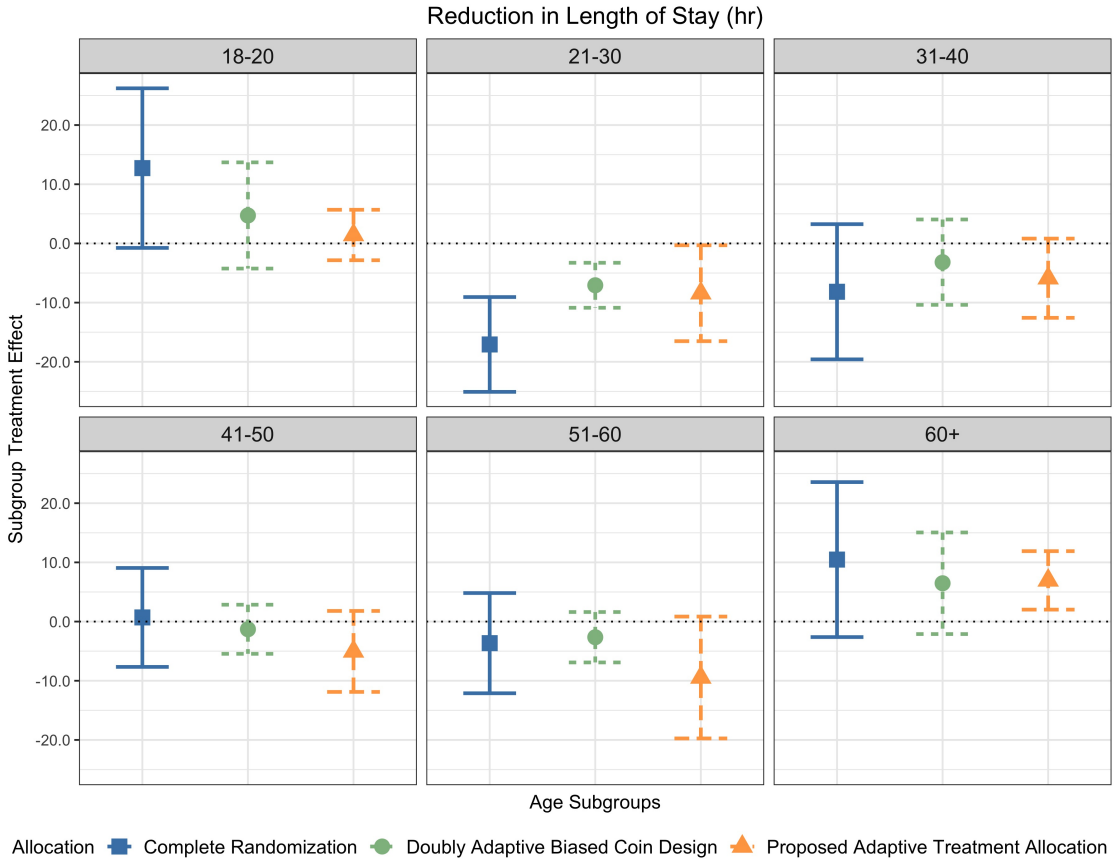


Figure 3.9: The estimated treatment effects in the six age subgroups under the complete randomization design, the doubly adaptive biased coin design, and our proposed response adaptive design. The initialization stage includes $n_1 = 100$ patients. In the following stages, we have $n_t = 1$, where $t = 2, \ldots, 200$.

$n_t = 1$ patient in the following stages, where $t = 2, \ldots, 200$. To generate synthetic outcomes, for example, if patient A's outcome under the genetically-guided therapy is not observed in the original data, we would match patient A with patient B who shares similar baseline characteristics but is treated by the genetically-guided therapy. Then we use the matched patient's outcome as the outcome we would have observed for patient A under the genetically-guided therapy. We compare three adaptive experiment strategies: (1) our proposed response adaptive design, (2) the complete randomization design, and (3) the doubly adaptive biased coin design. We then evaluate the treatment effect of the genetically-guided therapy in each patient age subgroup. The results are summarized in Figure 3.9.

Figure 3.9 demonstrates that the genetically-guided therapy reduces the length of stay the most in patients who are older than 60 under our proposed response adaptive design. Compared to the other two designs, our proposed response adaptive design estimates the

treatment effects in the age group older than 60 and the age group 18-20 with high efficiency. Our response adaptive design does not have significant efficiency gains in the other age subgroups, which verifies the theoretical results in Proposition 2. This case study's results suggest that prescribing drugs for MDD patients based on individual CYP2D6 functional status could potentially deliver more informed medical treatments and thus better improve patients' welfare, especially for patients who are older than 60.

3.9 Discussion

In this manuscript, we propose an adaptive experimental framework specifically designed to study treatment effect heterogeneity. It encompasses both response adaptive design and adaptive enrichment design in either fully adaptive or multi-stage experimental settings. Our proposed adaptive experimental framework allows various future extensions. First, instead of only identifying the best subgroup, decision-makers may also want to look into the top-performing subgroups. The current framework could be naturally extended to identify the top-performing subgroups with high probability. Second, in the current setting, we restrict subgroups to be non-overlapped. In practice, overlapping subgroups are also frequently seen in various settings; therefore, extending the adaptive experimental framework to handle overlapping subgroups could be of practical interest. Third, in the current framework, we assume there are no near or exact ties among the population subgroup treatment effects. In future work, we would like to relax this assumption by allowing subgroups to have near ties or exact ties.

3.10 Supplementary Materials

Proof of Remark 1

To prove Remark 1 in the main manuscript, we rely on the following lemma:

Lemma 7. *(a) The rate function under the adaptive treatment allocation design is*

$$G(\mathcal{S}_1, \mathcal{S}_j, \boldsymbol{e}) = \frac{(\tau_j - \tau_1)^2}{\sigma_1^2(e_1) + \sigma_j^2(e_j)},$$
$$\sigma_j^2(e_j) = \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]}{p_j e_j} + \frac{\mathbb{V}[Y(0)|X \in \mathcal{S}_j]}{p_j(1-e_j)},$$

(b) and the following equality holds:

$$G(\mathcal{S}_1, \mathcal{S}_j, \boldsymbol{e}) = G(\mathcal{S}_1, \mathcal{S}_k, \boldsymbol{e}).$$

Lemma 8. *By assuming* $\mathbb{V}[Y(1)|X \in \mathcal{S}_j] = \mathbb{V}[Y(0)|X \in \mathcal{S}_j]$*, the following holds:*

$$\frac{e_j(1-e_j)}{e_k(1-e_k)} = \frac{p_k(\tau_k - \tau_1)^2/\mathbb{V}[Y(1)|X \in \mathcal{S}_k]}{p_j(\tau_j - \tau_1)^2/\mathbb{V}[Y(1)|X \in \mathcal{S}_j]} + R_n,$$

where $R_n = \frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]} \cdot \frac{\frac{1}{A_j}-\frac{1}{A_k}}{\frac{1}{A_1}+\frac{1}{A_k}}$, $\frac{1}{A_j} - \frac{1}{A_k} = \frac{p_j e_j(1-e_j)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]} - \frac{p_k e_k(1-e_k)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}$, *and* $\frac{1}{A_1} = \frac{p_1 e_1(1-e_1)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}$. *We assume* $R_n = 0$ *by assuming* $\mathbb{V}[Y(1)|X \in \mathcal{S}_1] = 0$.

Proof. By Lemma 7, we have

$$G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e}) = G(\mathcal{S}_1, \mathcal{S}_k; \boldsymbol{e}),$$
$$\frac{(\tau_j - \tau_1)^2}{\sigma_1^2(e_1) + \sigma_j^2(e_j)} = \frac{(\tau_k - \tau_1)^2}{\sigma_1^2(e_1) + \sigma_k^2(e_k)}.$$

Therefore,

$$\frac{(\tau_j - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{p_1 e_1} + \frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_1]}{p_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_j e_j} + \frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_j]}{p_j(1-e_j)}}$$
$$= \frac{(\tau_k - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{p_1 e_1} + \frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_1]}{p_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_k e_k} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_k(1-e_k)}}.$$

In Remark 1, we assume $\mathbb{V}[Y(1)|X \in \mathcal{S}_j] = \mathbb{V}[Y(0)|X \in \mathcal{S}_j]$:

$$\frac{(\tau_j - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{p_1 e_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_j e_j(1-e_j)}} = \frac{(\tau_k - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{e_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{e_k(1-e_k)}}.$$

By Lemma 8, we have

$$\frac{(\tau_j - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_j e_j(1-e_j)}} = \frac{(\tau_k - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_k e_k(1-e_k)}},$$
$$\frac{e_j(1-e_j)}{e_k(1-e_k)} = \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]/p_j(\tau_j - \tau_1)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_k]/p_k(\tau_k - \tau_1)^2}.$$

One can compute e_j as

$$e_j^* = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - 4\frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]}{p_j(\tau_j - \tau_1)^2}}, \ j \neq 1.$$

□

Proof of Lemma 7

Proof. (a) The large derivation theory states that for some random variable x,

$$I(x) = \sup_{\lambda}(\lambda x - \Lambda_x(\lambda)),$$

where $\Lambda_x(\lambda)$ is the log moment generating function (MGF) of x. In our problem, we have two random quantities $\hat{\tau}_1$ and $\hat{\tau}_j$, which are the estimates of τ_1 and τ_j, respectively. We want to derive the rate function $I(\hat{\tau}_1, \hat{\tau}_j)$ starting with deriving the log MGF.

Step 1. Derive log MGF

The log MGF of $\hat{\tau}_1, \hat{\tau}_j$ is $\Lambda_{\hat{\tau}_1,\hat{\tau}_j}(\lambda_1, \lambda_j) = \log\Big(\mathbb{E}\big(e^{\lambda_1\hat{\tau}_1}e^{\lambda_j\hat{\tau}_j}\big)\Big)$. We want to incorporate p_1 and p_j into the log MGF.

$$\begin{aligned}
\log\Big(\mathbb{E}\big(e^{\frac{\lambda_1}{p_1}p_1\hat{\tau}_1+\frac{\lambda_j}{p_j}p_j\hat{\tau}_j}\big)\Big) &= p_1\log\Big(\mathbb{E}\big(e^{\frac{\lambda_1}{p_1}\hat{\tau}_1}\big)\Big) + p_j\log\Big(\mathbb{E}\big(e^{\frac{\lambda_j}{p_j}\hat{\tau}_j}\big)\Big),\\
I_{p_1,p_j}(\hat{\tau}_1,\hat{\tau}_j) &= \sup_{\lambda_1,\lambda_j}(\lambda_1\hat{\tau}_1 - p_1\Lambda_1(\lambda_1/p_1) + \lambda_j\hat{\tau}_j - \Lambda_j(\lambda_j/p_j)),\\
&= \sup_{\lambda_1}(\lambda_1\hat{\tau}_1 - p_1\Lambda_1(\lambda_1/p_1)) + \sup_{\lambda_j}(\lambda_j\hat{\tau}_j - p_j\Lambda_j(\lambda_1/p_j)),\\
&= p_1\sup_{\lambda_1/p_1}(\frac{\lambda_1}{p_1}\hat{\tau}_1 - p_1\Lambda_1(\lambda_1/p_1)) + p_j\sup_{\lambda_j/p_j}(\frac{\lambda_j}{p_j}\hat{\tau}_j - p_j\Lambda_j(\lambda_1/p_j)),\\
&= p_1I(\hat{\tau}_1) + p_jI(\hat{\tau}_j).
\end{aligned}$$

Therefore, $I_{p_1,p_j}(\hat{\tau}_1,\hat{\tau}_j) = p_1I(\hat{\tau}_1) + p_jI(\hat{\tau}_j)$.

Step 2. Derive the general form of $G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{\pi})$

According to the large deviation principle, for a set B such that

$$\inf_{x\in B^o} I(x) = \inf_{x\in\bar{B}} I(x),$$

where B^o and $\bar{B}$ denote the interior and the closure of set B respectively,

$$G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{\pi}) = \inf_{\hat{\tau}_1\le\hat{\tau}_j}\{p_1I(\hat{\tau}_1) + p_jI(\hat{\tau}_j)\}.$$

The form of $G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{\pi})$ can be further simplified. First, we inspect the functional form of $I_1(\hat{\tau}_1)$ and $I_j(\hat{\tau}_j)$. Assume $\hat{\tau}_1$ is asymptotically normally distributed with mean τ_1 and variance σ_1^2, then $I_1(\hat{\tau}_1) = \frac{(\hat{\tau}_1-\tau_1)^2}{2\sigma_1^2}$. Similarly, $I_j(\hat{\tau}_j) = \frac{(\hat{\tau}_j-\tau_j)^2}{2\sigma_j^2}$. The functional form of $I_1(\hat{\tau}_1)$ suggests that $I_1(\hat{\tau}_1)$ is a function which decreases in $\hat{\tau}_1$ when $\hat{\tau}_1 < \tau_1$, and increases in $\hat{\tau}_1$ when $\hat{\tau}_1 > \tau_1$. $I_j(\hat{\tau}_j)$ shares the similar functional behavior. Therefore, we can search for the region $\tau_j \le \hat{\tau}_1 \le \hat{\tau}_j \le \tau_1$, that is

$$\begin{aligned}
G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{\pi}) &= \inf_{\hat{\tau}\in[\tau_j,\tau_1]}\{p_1I_1(\hat{\tau}) + p_jI_j(\hat{\tau})\},\\
&= \inf_{\hat{\tau}}\{p_1I_1(\hat{\tau}) + p_jI_j(\hat{\tau})\}.
\end{aligned}$$

Step 3. Derive the explicit form of $G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{\pi})$
First, we derive the rate function under adaptive allocation design.

$$\begin{aligned}
\frac{\partial\big(p_1 I_1(\hat{\tau}) + p_j I_j(\hat{\tau})\big)}{\partial \hat{\tau}} &= \frac{p_1(\hat{\tau} - \tau_1)}{\sigma_1^2} + \frac{p_j(\hat{\tau} - \tau_j)}{\sigma_j^2} = 0, \\
\hat{\tau}^* &= \frac{p_1/\sigma_1^2}{p_1/\sigma_1^2 + p_j/\sigma_j^2} \cdot \tau_1 + \frac{p_j/\sigma_j^2}{p_1/\sigma_1^2 + p_j/\sigma_j^2} \cdot \tau_j, \\
\text{Plug in } \hat{\tau}_j^*,\ G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{p}) &= \frac{(\tau_j - \tau_1)^2}{2(\sigma_1^2/p_1 + \sigma_j^2/p_j)}, \\
G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e}) &= \frac{(\tau_j - \tau_1)^2}{2(\sigma_1^2(e_1) + \sigma_j^2(e_j))}.
\end{aligned}$$

(b) Given $G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e}) = \frac{(\tau_j - \tau_1)^2}{\sigma_1^2(e_1) + \sigma_j^2(e_j)}$, the optimization problem is

$$\begin{aligned}
&\max_{\boldsymbol{e}} \min_{2 \le j \le m} G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e}), \\
&\text{s.t. } \sum_{l=1}^{m} p_l e_l \le c_1, \\
&\quad c_2 \le e_l \le 1 - c_2,\ l = 1, \ldots, m.
\end{aligned}$$

The original problem can be transformed via the epigraph representation

$$\begin{aligned}
&\max_{\boldsymbol{e}}\ z, \\
&\text{s.t. } G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e}) \ge z,\ j = 2, \ldots, m, \\
&\sum_{l=1}^{m} p_l e_l \le c_1,\ 0 < c_1 < 1, \\
&\quad c_2 \le e_l \le 1 - c_2,\ l = 1, \ldots, m,\ c > 0.
\end{aligned}$$

The Lagrangian associated with the above problem is

$$\begin{aligned}
L(\boldsymbol{e}, \boldsymbol{\lambda}, \gamma, \boldsymbol{\alpha}, \boldsymbol{\beta}, z) = z &+ \sum_{j=2}^{m} \lambda_j \big(G(\mathcal{S}_1, \mathcal{S}_j; \boldsymbol{e}) - z\big) \\
&+ \gamma\big(c_1 - \sum_{l=1}^{m} p_l e_l\big) + \sum_{l=1}^{m} \alpha_l (e_l - c_2) + \sum_{l=1}^{m} \beta_l (1 - c_2 - e_l).
\end{aligned}$$

The Lagrangian primal problem is

$$\max_{\boldsymbol{e}} L_{\text{primal}}(\boldsymbol{e}, \boldsymbol{\lambda}, \gamma, \boldsymbol{\alpha}, \boldsymbol{\beta}, z) = \max_{\boldsymbol{e}} \min_{\boldsymbol{\lambda}, \gamma, \boldsymbol{\alpha}, \boldsymbol{\beta}, z} L(\boldsymbol{e}, \boldsymbol{\lambda}, \gamma, \boldsymbol{\alpha}, \boldsymbol{\beta}, z).$$

The Lagrangian dual problem is

$$\min_{\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z} \max_{\boldsymbol{e}} L(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z) = \min_{\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z} L_{\text{dual}}(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z).$$

By taking derivative of $L(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z)$, we can obtain the following:

$$\frac{\partial L(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z)}{\partial e_j} = \lambda_j \frac{\partial G(\mathcal{S}_1,\mathcal{S}_j;\boldsymbol{e})}{\partial e_j} - \gamma p_j + \alpha_j - \beta_j = 0, \tag{3.15}$$

$$\frac{\partial L(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z)}{\partial e_1} = \sum_{j=2}^{m} \lambda_j \frac{\partial G(\mathcal{S}_1,\mathcal{S}_j;\boldsymbol{e})}{\partial e_1} - \gamma p_1 + \alpha_1 - \beta_1 = 0, \tag{3.16}$$

$$\frac{\partial L(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z)}{\partial z} = 1 - \sum_{j=2}^{m} \lambda_j = 0 \implies \sum_{j=2}^{m} \lambda_j = 1, \tag{3.17}$$

$$\frac{\partial L(\boldsymbol{e},\boldsymbol{\lambda},\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},z)}{\partial \lambda_j} = G(\mathcal{S}_1,\mathcal{S}_j;\boldsymbol{e}) - z = 0, \ j = 2,\ldots,m. \tag{3.18}$$

By Condition (4), we have

$$G(\mathcal{S}_1,\mathcal{S}_j;\boldsymbol{e}) = G(\mathcal{S}_1,\mathcal{S}_k;\boldsymbol{e}).$$

□

Proof of Lemma 8

Proof. By Lemma 7,

$$\frac{(\tau_j-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{p_1e_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_je_j(1-e_j)}} = \frac{(\tau_k-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{p_1e_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_ke_k(1-e_k)}}.$$

Let $A_1 = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}{p_1e_1(1-e_1)}$, $A_j = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_je_j(1-e_j)}$, $A_k = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_ke_k(1-e_k)}$,

$$
\begin{aligned}
\frac{(\tau_j-\tau_1)^2}{A_1+A_j} &= \frac{(\tau_k-\tau_1)^2}{A_1+A_k},\\
(\tau_j-\tau_1)^2\frac{1}{(A_1+A_j)} &= (\tau_k-\tau_1)^2\frac{1}{(A_1+A_k)},\\
(\tau_j-\tau_1)^2\frac{1}{(\frac{1}{A_1}+\frac{1}{A_j})A_1A_j} &= (\tau_k-\tau_1)^2\frac{1}{(\frac{1}{A_1}+\frac{1}{A_k})A_1A_k},\\
(\frac{1}{A_1}+\frac{1}{A_j})A_1A_j(\tau_k-\tau_1)^2 &= (\frac{1}{A_1}+\frac{1}{A_k})A_1A_k(\tau_j-\tau_1)^2,\\
\frac{A_j(\tau_k-\tau_1)^2}{A_k(\tau_j-\tau_1)^2} &= \frac{\frac{1}{A_1}+\frac{1}{A_k}}{\frac{1}{A_1}+\frac{1}{A_j}},\\
\frac{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}{p_je_j(1-e_j)}\cdot(\tau_k-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_ke_k(1-e_k)}\cdot(\tau_j-\tau_1)^2} &= \frac{\frac{1}{A_1}+\frac{1}{A_k}}{\frac{1}{A_1}+\frac{1}{A_j}},\\
\frac{\frac{1}{A_1}+\frac{1}{A_k}}{\frac{1}{A_1}+\frac{1}{A_j}}\cdot\frac{e_j(1-e_j)}{e_k(1-e_k)} &= \frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]},\\
\frac{e_j(1-e_j)}{e_k(1-e_k)} &= \frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}\cdot\frac{\frac{1}{A_1}+\frac{1}{A_j}}{\frac{1}{A_1}+\frac{1}{A_k}},\\
\frac{e_j(1-e_j)}{e_k(1-e_k)} &= \frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}\cdot\Big(1+\frac{\frac{1}{A_j}-\frac{1}{A_k}}{\frac{1}{A_1}+\frac{1}{A_k}}\Big),\\
\frac{e_j(1-e_j)}{e_k(1-e_k)} &= \frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}\\
&\quad+\underbrace{\frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}\cdot\frac{\frac{1}{A_j}-\frac{1}{A_k}}{\frac{1}{A_1}+\frac{1}{A_k}}}_{R_n},\\
\text{where } \frac{1}{A_j}-\frac{1}{A_k} &= \frac{p_je_j(1-e_j)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]}-\frac{p_ke_k(1-e_k)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]},\\
\frac{1}{A_1} &= \frac{p_1e_1(1-e_1)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]}.
\end{aligned}
$$

To set $R_n = 0$, we can assume $\frac{1}{A_1}\to\infty$, thus $\mathbb{V}[Y(1)|X\in\mathcal{S}_1] = 0$. □

Proof of Lemma 1

To prove Lemma 1 in the main manuscript, we rely on the following assumptions and lemma:

Assumption 19. *For $t = 1, \ldots, T$, $i = 1, \ldots, n_t$ and $d \in \{0, 1\}$, the potential outcome $Y_{it}(d)$ has bounded first and $2 + \delta$ moments for some $\delta > 0$, that is $\mathbb{E}|Y_{it}(d)| < \infty$ and $\mathbb{E}|Y_{it}(d)|^{2+\delta} < \infty$.*

Assumption 20. *There are total $m \geq 2$ subgroups, and the subgroup treatment effects can be monotonically ordered with $\tau_1 > \tau_2 > \ldots > \tau_m$.*

Assumption 21. *We assume subgroup proportions p_j are bounded away from 0 and 1 by a positive constant, $j = 1, \ldots, m$.*

Lemma 9. *Under Assumptions 19-21, $\hat{e}^*_{tj} \to e^*_j$, for $t = 1, \ldots, T$, $j = 1, \ldots, m$, as $T \to \infty$.*

Lemma 10. *Under Assumptions 19-21, $\hat{\tau}_{tj} \to \tau_j$, $t = 1, \ldots, T$, $T \to \infty$.*

Proof. We want to show $\frac{\sum_{i:i\in\mathcal{S}_j} \mathbb{1}(D_i=1)}{N} \to e^*_j$ as $N \to \infty$. In the fully sequential setting, in the following proof, we will use T and N interchangeably.

Step 1. Show the number of downcrossing is finite.

Definition 1 (Downcrossing)**.** *Let $\varphi(x; \boldsymbol{\theta}) : [0, 1] \times \mathbb{R}^d \to [0, 1]$. The function $t(\boldsymbol{\theta}) : \mathbb{R}^d \to [0, 1]$ is a generalized downcrossing of φ if for any $\boldsymbol{\theta} \in \mathbb{R}^d$,*

$$\forall x < t(\boldsymbol{\theta}),\ \varphi(x; \boldsymbol{\theta}) \geq t(\boldsymbol{\theta}),\ \text{and } \forall x > t(\boldsymbol{\theta}),\ \varphi(x; \boldsymbol{\theta}) \leq t(\boldsymbol{\theta}).$$

In our setting, $t(\boldsymbol{\theta})$ is equivalent to e^*_j, which denotes the optimal treatment allocation. More concretely, we let $e^*_j := t(\tau_j, \sigma_j)$ and $\hat{e}_{tj} := t(\hat{\tau}_{t-1,j}, \hat{\sigma}_{t-1,j})$, $t = 1, 2, \ldots, T$, $T \to \infty$. $\varphi(.)$ is an allocation rule, $\varphi(g_t; \hat{\tau}_{t,j}, \hat{\sigma}_{t,j})$, where $g_t = \frac{\sum_{i:i\in\mathcal{S}_j} \mathbb{1}(D_i=1)}{t}$ denotes the actual treatment allocation up to stage t. We can see $\varphi(g_t; \hat{\tau}_{t,j}, \hat{\sigma}_{t,j})$ is decreasing in g_t because when g_t is larger than the optimal allocation proportion, we tend to allocate fewer subjects to the treatment arm; when g_t is smaller than the optimal allocation proportion, we tend to allocate more subjects to the treatment arm.

Given that $\varphi(g_t; \hat{\tau}_{t,j}, \hat{\sigma}_{t,j})$ is decreasing in g_t, there exists a unique downcrossing. Next, we will show $\frac{\sum_{i:i\in\mathcal{S}_j} \mathbb{1}(D_i=1)}{N} \to e^*_j$ as $N_j \to \infty$ leveraging the unique downcrossing.

Step 2. Convergence of the actual allocation

Denote the martingale process at each stage t as $\{M_t; \mathcal{F}_t\}$, where $M_t = \sum_{i:i\in\mathcal{S}_j, i\leq t} \Delta M_i = \sum_{i:i\in\mathcal{S}_j, i\leq t} \{D_i - \mathbb{E}[D_i|\mathcal{F}_{t-1}]\}$. We denote $\mathcal{F}_t$ as the σ-algebra up to stage t ($\mathcal{F}_1$ denotes the trivial σ-field). $\mathcal{F}_t = \sigma(D_1, \ldots, D_t; Y_1, \ldots, Y_t)$.

Assume in the first stage, two treatment arms of subgroup j are assigned with equal number of subjects. We denote the total number of subjects in stage 1 as n_1. Let $\lambda_t = \max\{s : n_1 + 1 \leq s \leq t,\ g_s \leq t(\hat{\tau}_{sj}, \hat{\sigma}_{sj})\}$. λ_t represents the maximum number of subject before the downcrossing. By Definition 1, at each step $i \leq \lambda_t$, $\varphi(g_i; \hat{\tau}_{ij}, \hat{\sigma}_{ij}) \geq t(\hat{\tau}_{sj}, \hat{\sigma}_{sj})$; at

each step $i > \lambda_t$, $\varphi(g_i; \hat{\tau}_{ij}, \hat{\sigma}_{ij}) \leq t(\hat{\tau}_{sj}, \hat{\sigma}_{sj})$. Let $D_t := \sum_{i:i\in\mathcal{S}_j, i\leq t} D_i$,

$$\begin{aligned}
D_t &= D_{\lambda_t+1} + \sum_{i:i\in\mathcal{S}_j, \lambda_t+2\leq i\leq t} \{D_i - \mathbb{E}[D_i|\mathcal{F}_{i-1}] + \mathbb{E}[D_i|\mathcal{F}_{i-1}]\}, \\
&= D_{\lambda_t+1} + \sum_{i:i\in\mathcal{S}_j, \lambda_t+2\leq i\leq t} \Delta M_i + \sum_{i:i\in\mathcal{S}_j, \lambda_t+2\leq i\leq t} \varphi(g_{i-1}; \hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j}), \\
&\leq D_{\lambda_t} + 1 + M_t - M_{\lambda_t+1} + \sum_{i:i\in\mathcal{S}_j, \lambda_t+2\leq i\leq t} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j}).
\end{aligned}$$

Since $D_{\lambda_t} \leq \lambda_t \cdot t(\hat{\tau}_{\lambda_t}, \hat{\sigma}_{\lambda_t})$, we have

$$\begin{aligned}
D_t &\leq \lambda_t \cdot t(\hat{\tau}_{\lambda_t}, \hat{\sigma}_{\lambda_t}) + 1 + M_t - M_{\lambda_t+1} + \sum_{i:i\in\mathcal{S}_j, \lambda_t+2\leq i\leq t} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j}), \\
&= \lambda_t \cdot t(\hat{\tau}_{\lambda_t}, \hat{\sigma}_{\lambda_t}) + 1 + M_t - M_{\lambda_t+1} + \\
&\quad + \sum_{i:i\in\mathcal{S}_j, 1\leq i\leq t} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j}) - \sum_{i:i\in\mathcal{S}_j, 2\leq i\leq \lambda_t+1} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j}) - t(\hat{\tau}_{1,j}, \hat{\sigma}_{1,j}), \\
&= \Big(\lambda_t \cdot t(\hat{\tau}_{\lambda_t}, \hat{\sigma}_{\lambda_t}) - \sum_{i:i\in\mathcal{S}_j, 2\leq i\leq \lambda_t+1} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j})\Big) \\
&\quad + M_t - M_{\lambda_t+1} + 1 - t(\hat{\tau}_{1,j}, \hat{\sigma}_{1,j}) \\
&\quad + \sum_{i:i\in\mathcal{S}_j, 1\leq i\leq t} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j}) + nt(\hat{\tau}_{\lambda_n}, \hat{\sigma}_{\lambda_n}) - nt(\hat{\tau}_{\lambda_n}, \hat{\sigma}_{\lambda_n}), \\
D_t - nt(\hat{\tau}_n, \hat{\sigma}_n) &\leq \Big(\lambda_t \cdot t(\hat{\tau}_{\lambda_t}, \hat{\sigma}_{\lambda_t}) - \sum_{i:i\in\mathcal{S}_j, 2\leq i\leq \lambda_t+1} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j})\Big) \\
&\quad + M_t - M_{\lambda_t+1} + 1 - t(\hat{\tau}_{1,j}, \hat{\sigma}_{1,j}) \\
&\quad - \Big(n \cdot t(\hat{\tau}_n, \hat{\sigma}_n) - \sum_{i:i\in\mathcal{S}_j, 1\leq i\leq n} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j})\Big).
\end{aligned}$$

By Lemmas 9 and 10, we have

$$\begin{aligned}
t(\hat{\tau}_{n,j}, \hat{\sigma}_{n,j}) \to t(\tau_j, \sigma_j), \text{ that is } t(\hat{\tau}_{n,j}, \hat{\sigma}_{n,j}) \to e_j^*, \ a.s. \\
\therefore \lim_{n\to\infty} \frac{1}{n}\Big(n \cdot t(\hat{\tau}_n, \hat{\sigma}_n) - \sum_{i:i\in\mathcal{S}_j, 1\leq i\leq n} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j})\Big) = 0, \ a.s. \\
\text{and } \lim_{n\to\infty} \frac{1}{n}\Big(\lambda_t \cdot t(\hat{\tau}_{\lambda_t}, \hat{\sigma}_{\lambda_t}) - \sum_{i:i\in\mathcal{S}_j, 2\leq i\leq \lambda_t+1} t(\hat{\tau}_{i-1,j}, \hat{\sigma}_{i-1,j})\Big) = 0, \ a.s.,
\end{aligned}$$

since $\lambda_t \to \infty$ as $t \to \infty$ (i.e., $N \to \infty$). Therefore,

$$\frac{D_t - N \cdot t(\hat{\tau}_{N,j}, \hat{\sigma}_{N,j})}{N} \to 0, \ a.s.$$

Given that $t(\hat{\tau}_{N,j}, \hat{\sigma}_{N,j}) \to e_j^*$,

$$\lim_{N\to\infty} \frac{\sum_{i:i\in\mathcal{S}_j} D_i}{N} = e_j^*, \ a.s.$$

□

Proof of Lemma 9

From the proof of Lemma 7, we denote the Lagrangian dual function under the adaptive experiment setting as G_t and under the oracle setting as g. Supppose $||G_t - g||_\infty \to 0$. Let K be a bounded and compact set and fix ε. We let $K_\varepsilon = K - B_\varepsilon(e_j^*)$. K_ε is a compact, bounded, and closed set. For some $e_{j,\varepsilon}^* \in K_\varepsilon$, $g(e_{j,\varepsilon}^*) < g(e_j^*)$ by definition. Define $\delta = \sup_K g - \sup_{K_\varepsilon} g > 0$. Suppose $||G_t - g||_\infty < \delta/2$, we have

$$\sup_K G_t < \sup_{K_\varepsilon} g + \frac{\delta}{2} = g(e_j^*) - \frac{\delta}{2},$$
$$\sup_K G_t \geq G_t(e_j^*) > g(e_j^*) - \frac{\delta}{2} = g(e_j^*) - \frac{\delta}{2}.$$

Because $\hat{e}_{tj}^* \in B_\varepsilon(e_j^*)$,

$$\mathbb{P}(||G_t - g||_\infty < \delta/2) \leq \mathbb{P}(||\hat{e}_{tj}^* - e_j^*|| < \varepsilon),$$
$$\mathbb{P}(||\hat{e}_{tj}^* - e_j^*|| \geq \varepsilon) \leq \mathbb{P}(||G_t - g||_\infty \geq \delta/2) \to 0.$$

Therefore, $\hat{e}_{tj}^* \to e_j^*$ as $T \to \infty$.

Proof of Lemma 10

Proof. **Step 1. Show Z_t is a martingale sequence.**

Let $Z_t = \frac{D_t Y_t \mathbb{1}(X_t \in \mathcal{S}_j)}{\hat{e}_{tj}} - \frac{(1-D_t) Y_t \mathbb{1}(X_t \in \mathcal{S}_j)}{1-\hat{e}_{tj}}$, $t = 1, \ldots, T$. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Ω denotes a set, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and P is a probability measured defined on $\mathcal{F}$. Let $\mathcal{F}_t$, $t = 1, 2, \ldots$ be an increasing sequence of σ-algebras of $\mathcal{F}$ sets. We want to show Z_t is a martingale sequence by verifying the following conditions:

(i) Z_t is measurable with respect to $\mathcal{F}_t$;

(ii) $\mathbb{E}(|Z_t|) < \infty$;

(iii) $\mathbb{E}[Z_t|\mathcal{F}_{t-1}] = Z_{t-1}$ almost surely.

Condition (i) and (ii) hold by Assumption 19. Next, we check condition (iii).

$$\begin{aligned}\mathbb{E}[Z_t|\mathcal{F}_{t-1}] &= \mathbb{E}\Big[\frac{D_tY_t\mathbb{1}(X_t\in\mathcal{S}_j)}{\hat{e}_{tj}} - \frac{(1-D_t)Y_t\mathbb{1}(X_t\in\mathcal{S}_j)}{1-\hat{e}_{tj}}\Big|\mathcal{F}_{t-1}\Big],\\ &= \frac{\mathbb{E}[D_t|\mathcal{F}_{t-1}]\mathbb{E}[Y_t\mathbb{1}(X_t\in\mathcal{S}_j)|\mathcal{F}_{t-1}]}{\hat{e}_{tj}} - \frac{\mathbb{E}[(1-D_t)|\mathcal{F}_{t-1}]\mathbb{E}[Y_t\mathbb{1}(X_t\in\mathcal{S}_j)|\mathcal{F}_{t-1}]}{1-\hat{e}_{tj}},\\ &= \frac{D_{t-1}Y_{t-1}\mathbb{1}(X_{t-1}\in\mathcal{S}_j)}{\hat{e}_{t-1,j}} - \frac{(1-D_{t-1})Y_{t-1}\mathbb{1}(X_{t-1}\in\mathcal{S}_j)}{1-\hat{e}_{t-1,j}} = Z_{t-1}.\end{aligned}$$

Because conditions (i)–(iii) hold for Z_t, Z_t is a martingale sequence.

Step 2. Show optional stopping theorem holds for $\boldsymbol{Z}_t$.

Theorem 6 (Doob's Optional Stopping Theorem)**.** *Let $\boldsymbol{Z} = \{Z_t\}_{t=1}^T$ be a martingale and S a stopping time, both with respect to a filtration $\{\mathcal{F}_t\}_{t=1}^T$. Suppose that any one of the following conditions holds,*

(a) There exists a positive integer c such that $S \le c$, a.s.;

(b) $\mathbb{E}[S] < \infty$, and there exists a positive real number C such that $\mathbb{E}[|Z_{t+1} - Z_t|\big|\mathcal{F}_t] \le C$, a.s. $\forall t$;

(c) There exists a constant C', such that $|Z_{t\wedge S}| \le C'$ a.s. $\forall t$.

Then $\mathbb{E}[Z_S] = \mathbb{E}[Z_1]$.

Denote α_i as the minimum total sample size required to have i subjects in subpopulation j assigned to the treatment arm, i.e., $\alpha_i = \min\{N : N_j = i\}$, $i = 1, 2, \ldots$. $\boldsymbol{Z}_i$ denotes the sequence of random variables where $t = 1, \ldots, i$, and $\boldsymbol{\eta}_i$ is an independent copy of $\boldsymbol{Z}_i$. Define $\boldsymbol{W}_i = \boldsymbol{Z}_{\alpha_i}\mathbb{1}(\alpha_i < \infty) + \boldsymbol{\eta}_i\mathbb{1}(\alpha_i = +\infty)$.

Let S be a stopping time. we want to show $\mathbb{E}[Z_S] = \mathbb{E}[Z_1]$ using Doob's optional stopping theorem. We first write the stopped process as

$$\begin{aligned}W_S &= W_1 + \sum_{t=1}^{S-1\wedge t-1}(W_{t+1} - W_t),\\ |W_S| &\le |Z_1| + \sum_{t=1}^{S-1\wedge t-1}|W_{t+1} - W_t|,\\ &\le |Z_1| + \sum_{t=1}^{\infty}|W_{t+1} - W_t|\cdot\mathbb{1}(S > t),\end{aligned}$$

by the monotone convergence theorem,

$$\begin{aligned}&\mathbb{E}\Big[|Z_1| + \sum_{t=1}^{\infty} |W_{t+1} - W_t| \cdot \mathbb{1}(S > t)\Big] \\ &= \mathbb{E}[|Z_1|] + \sum_{t=1}^{\infty} \mathbb{E}[|W_{t+1} - W_t| \cdot \mathbb{1}(S > t)].\end{aligned}$$

Under Assumption 19, condition (c) in Theorem 6 is satisfied. If $\alpha_i < \infty$, condition (a) is satisfied. If condition (b) holds,

$$\begin{aligned}&\mathbb{E}[|Z_1|] + \sum_{t=1}^{\infty} \mathbb{E}[|W_{t+1} - W_t| \cdot \mathbb{1}(S > t)], \\ &= \mathbb{E}[|Z_1|] + \sum_{t=1}^{\infty} \mathbb{E}\big[\mathbb{E}[|W_{t+1} - W_t| \big| \mathcal{F}_t] \cdot \mathbb{1}(S > t)\big], \\ &\leq \mathbb{E}[|Z_1|] + c \sum_{t=1}^{\infty} \mathbb{P}(S > t), \\ &= \mathbb{E}[|Z_1|] + c\mathbb{E}[\tau] < \infty.\end{aligned}$$

Therefore, we can use Theorem 6 to conclude $\mathbb{E}[W_i] = \mathbb{E}[W_1] = \mathbb{E}[Z_1]$, where $\boldsymbol{W}_i = \boldsymbol{Z}_{\alpha_i} \mathbb{1}(\alpha_i < \infty) + \boldsymbol{\eta}_i \mathbb{1}(\alpha_i = +\infty)$.

Step 3. Consistency of $\hat{\tau}_j$.

$\mathbb{E}[W_S] = \mathbb{E}[W_1] = \mathbb{E}[Z_1]$ implies that $\boldsymbol{W}_i = \boldsymbol{Z}_{\alpha_i} \mathbb{1}(\alpha_i < \infty) + \boldsymbol{\eta}_i \mathbb{1}(\alpha_i = +\infty)$ has a common distribution the same as that of Z_1.

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} W_t = \frac{1}{T}\sum_{t=1}^{T} Z_t = \frac{1}{T}\sum_{t=1}^{T} Z_1 = Z_1, \\ = \frac{\sum_{i=1}^{n_1} D_i Y_i \mathbb{1}(X_i \in \mathcal{S}_j)}{\sum_{i=1}^{n_1} (D_i \mathbb{1}(X_i \in \mathcal{S}_j))/n_{1j}} - \frac{\sum_{i=1}^{n_1} (1 - D_i) Y_i \mathbb{1}(X_i \in \mathcal{S}_j)}{\sum_{i=1}^{n_1} ((1 - D_i) \mathbb{1}(X_i \in \mathcal{S}_j))/n_{1j}}.\end{aligned}$$

Since we assume that stage 1 generates consistent estimate of τ_j, that is $\hat{\tau}_{1j} \to \tau_j$, we can conclude that $\hat{\tau}_{tj} \to \tau_j$, a.s. by the strong law of large numbers for martingale sequence, and $\hat{\tau}_{tj} - \tau_j = O(\sqrt{\frac{\log\log T}{T}})$ a.s. □

Proof of Theorem 1

Proof. In this proof, we want to show the asymptotic normality of $\hat{\tau}_j$ by checking the conditions for using martingale central limit theorem (CLT).

$$\hat{\tau}_j = \frac{1}{N}\sum_{t=2}^{N}\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big) - \Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big),$$
$$\text{Let } Z_t = \Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big) - \Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big) - \tau_j$$

$\{Z_t\}_{t=1}^N$ is a martingale difference sequence because

$$\mathbb{E}[Z_t|\mathcal{F}_{t-1}] = \mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big) - \Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big) - \tau_j\Big|\mathcal{F}_{t-1}\Big] = 0.$$

To apply martingale CLT, the martingale difference sequence needs to satisfy the following conditions

(a) $\mathbb{E}[Z_t^2] = \sigma_{tj}^2(e_j^*)$, $\frac{1}{N}\sum_{t=1}^N \sigma_{tj}^2(e_j^*) \to \sigma_j^2(e_j^*)$ as $N\to\infty$, where $\sigma_j^2(e_j^*) = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j e_j^*} + \frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_j]]}{p_j(1-e_j^*)}$, $\sigma_{tj}^2(e_j^*) = \frac{\mathbb{V}[Y(1)|X_t\in\mathcal{S}_j]}{p_j e_j^*} + \frac{\mathbb{V}[Y(0)|X_t\in\mathcal{S}_j]}{p_j(1-e_j^*)}$.

(b) $\mathbb{E}[|Z_t|^r] < \infty$, $r>2$;

(c) $\frac{1}{N}\sum_{t=1}^N Z_t^2 \to \sigma_j^2(e_j^*)$ as $N\to\infty$.

Step 1. Check condition (a) Let

$$\sigma_j^2(e_j^*) = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j e_j^*} + \frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_j]]}{p_j(1-e_j^*)},$$
$$\sigma_{tj}^2(e_j^*) = \frac{\mathbb{V}[Y(1)|X_t\in\mathcal{S}_j]}{p_j e_j^*} + \frac{\mathbb{V}[Y(0)|X_t\in\mathcal{S}_j]}{p_j(1-e_j^*)},$$
$$\mathbb{E}[Z_t^2] - \sigma_{tj}^2(e_j^*) = \underbrace{\mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})} - \frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})} - \tau_j\Big)^2\Big]}_{A}$$
$$- \frac{\mathbb{V}[Y(1)|X_t\in\mathcal{S}_j]}{p_j e_j^*} - \frac{\mathbb{V}[Y(0)|X_t\in\mathcal{S}_j]}{p_j(1-e_j^*)}.$$

$$
\begin{aligned}
A &= \mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})} - \frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})} - \tau_j\Big)^2\Big], \\
&= \mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] + \mathbb{E}\Big[\Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] + \mathbb{E}[\tau_j^2] \\
&- 2\mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)\Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)\Big] \\
&- 2\mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})} - \frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)\cdot\tau_j\Big], \\
&= \mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] + \mathbb{E}\Big[\Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] + \tau_j^2 \\
&\quad - 2\mathbb{E}\Big[\mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})} - \frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)\cdot\tau_j\Big]\Big|\mathcal{F}_{t-1}\Big], \\
&= \mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] + \mathbb{E}\Big[\Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] - \tau_j^2.
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[Z_t^2] - \sigma_{tj}^2(e_j^*) &= \mathbb{E}\Big[\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] + \mathbb{E}\Big[\Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2\Big] - \tau_j^2 \\
&\quad - \frac{\mathbb{V}[Y(1)|X_t\in\mathcal{S}_j]}{p_j e_j^*} - \frac{\mathbb{V}[Y(0)|X_t\in\mathcal{S}_j]}{p_j(1-e_j^*)}, \\
&\le \mathbb{E}\Big[\Big|\Big(\frac{D_t Y_t}{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2 - \Big(\frac{D_t Y_t}{e_j^*}\Big)^2\Big|\Big] + \mathbb{E}\Big[\Big|\Big(\frac{(1-D_t)Y_t}{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big)^2 \\
&\quad - \Big(\frac{(1-D_t)Y_t}{1-e_j^*}\Big)^2\Big|\Big] - \tau_j^2, \\
&\le C\cdot\mathbb{E}\Big[\Big|\Big(\frac{D_t Y_t}{\sqrt{\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}}\Big) - \Big(\frac{D_t Y_t}{\sqrt{e_j^*}}\Big)\Big|\Big] + \mathbb{E}\Big[\Big|\Big(\frac{(1-D_t)Y_t}{\sqrt{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}}\Big) \\
&\quad - \Big(\frac{(1-D_t)Y_t}{\sqrt{1-e_j^*}}\Big)\Big|\Big] - \tau_j^2, \\
&\le C_1\cdot\mathbb{E}\Big[\Big|\Big(\sqrt{\mathbb{P}(D_t=1|\mathcal{F}_{t-1}}D_t Y_t\Big) - \Big(D_t Y_t\sqrt{e_j^*}\Big)\Big|\Big] \\
&\quad + C_2\cdot\mathbb{E}\Big[\Big|\Big((1-D_t)Y_t\sqrt{1-\mathbb{P}(D_t=1|\mathcal{F}_{t-1})}\Big) \\
&\quad - \Big((1-D_t)Y_t\sqrt{1-e_j^*}\Big)\Big|\Big] - \tau_j^2.
\end{aligned}
$$

By Lemma 9, $\sqrt{\mathbb{P}(D_t=1|\mathcal{F}_{t-1}) - e_j^*} \to 0$,

$$
\mathbb{E}[Z_t^2] - \sigma_{tj}^2(e_j^*) \to 0, \text{ and } \frac{1}{N}\sum_{t=1}^{N}\sigma_{tj}(e_j^*) \to \sigma_j^2(e_j^*).
$$

Step 2. Check condition (b)
Condition (b) is satisfied because $\mathbb{E}[Y_t] < \infty$ and $\frac{1}{\hat{e}_{tj}} < \infty$.
Step 3. Check condition (c)
To check condition (c), we bound

$$Z_t^2 - \mathbb{E}[Z_t^2|\mathcal{F}_{t-1}] + \mathbb{E}[Z_t^2|\mathcal{F}_{t-1}] - \sigma_j^2(e_j^*),$$

where $\frac{1}{N}\sum_{t=1}^{N} Z_t^2 - \mathbb{E}[Z_t^2|\mathcal{F}_{t-1}] \to 0$, by the weak law of large numbers for martingale difference sequence. □

Proof of Lemma 2

Proof. Denote $\hat{\tau}_j^r$ as the estimated causal effect of subpopulation j under sample size r. The correct selection event can be expressed as

$$\begin{aligned}
\boldsymbol{\mathcal{E}} &:= \bigcap_{r=n_1}^{N} \left\{ \left\{ \hat{\tau}_1^r \geq \frac{\tau_1 + \tau_2}{2} \right\} \bigcap \left\{ \bigcap_{j \neq 1} \left\{ \hat{\tau}_j^r \leq \frac{\tau_1 + \tau_2}{2} \right\} \right\} \right\}, \\
&:= \bigcap_{r=n_1}^{N} \left\{ \left\{ \hat{\tau}_1^r \geq \tau_1 - \frac{\tau_1 - \tau_2}{2} \right\} \bigcap \left\{ \bigcap_{j \neq 1} \left\{ \hat{\tau}_j^r \leq \tau_j + \tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2} \right\} \right\} \right\}.
\end{aligned}$$

The correct selection probability can be characterized as

$$\begin{aligned}
&1 - \mathbb{P}(\hat{\tau}_1 \leq \max_{j=2,\ldots,m} \hat{\tau}_j) \\
&\geq 1 - \mathbb{P}(\boldsymbol{\mathcal{E}}^c) \\
&\geq 1 - \sum_{r=n_1}^{N} \left[\mathbb{P}\Big(\hat{\tau}_1^r < \tau_1 - \frac{\tau_1 - \tau_2}{2} \Big) + \sum_{j=2}^{m} \mathbb{P}\Big(\hat{\tau}_1^r > \tau_j + \tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2} \Big) \right], \\
&\geq 1 - \left[\sum_{r=n_1}^{\infty} \mathbb{P}\Big(\hat{\tau}_1^r < \tau_1 - \frac{\tau_1 - \tau_2}{2} \Big) + \sum_{j=2}^{m} \sum_{r=n_1}^{\infty} \mathbb{P}\Big(\hat{\tau}_1^r > \tau_j + \tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2} \Big) \right], \\
&\geq 1 - \left[\sum_{r=n_1}^{\infty} \exp\Big(- \frac{(\tau_1 - \tau_2)^2 r}{8\sigma_1^2(p_1^*)} \Big) + \sum_{j=2}^{m} \sum_{r=n_1}^{\infty} \exp\Big(- \frac{(\tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2})^2 r}{2\sigma_j^2(p_j^*)} \Big) \right], \\
&\geq 1 - \left[\frac{\exp\Big(- \frac{(\tau_1 - \tau_2)^2 n_1}{8\sigma_1^2(p_1^*)m} \Big)}{1 - \exp\Big(- \frac{(\tau_1 - \tau_2)^2 (n_2 - n_1)}{8\sigma_1^2(p_1^*)} \Big)} + \sum_{j=2}^{m} \frac{\exp\Big(- \frac{(\tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2})^2 n_1}{2\sigma_j^2(p_j^*)m} \Big)}{1 - \exp\Big(- \frac{(\tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2})^2 (n_2 - n_1)}{2\sigma_j^2(p_j^*)} \Big)} \right], \\
&= 1 - \left[C_1' \exp\Big(- \frac{(\tau_1 - \tau_2)^2 n_1}{8\sigma_1^2(p_1^*)m} \Big) + \sum_{j=2}^{m} C_j' \exp\Big(- \frac{(\tau_2 - \tau_j + \frac{\tau_1 - \tau_2}{2})^2 n_1}{2\sigma_j^2(p_j^*)m} \Big) \right],
\end{aligned}$$

where

$$C_1 = \frac{1}{1-\exp\Big(-\frac{(\tau_1-\tau_2)^2(n_2-n_1)}{8\sigma_1^2(p_1^*)}\Big)} = \frac{1}{1-\exp\Big(-\frac{(\tau_1-\tau_2)^2}{8\sigma_1^2(p_1^*)}\Big)},$$
$$C_j = \frac{1}{1-\exp\Big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2(n_2-n_1)}{2\sigma_j^2(p_j^*)}\Big)} = \frac{1}{1-\exp\Big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2}{2\sigma_j^2(p_j^*)}\Big)},$$

given that $n_2 - n_1 = 1$ in fully sequential experiments. Similarly, the rate of correct selection probability is bounded by

$$1-\mathbb{P}(\hat{\tau}_1 \leq \max_{j=2,\ldots,m} \hat{\tau}_j)$$
$$\geq 1-\left[C_1' \exp\Big(-\frac{(\tau_1-\tau_2)^2 n_1}{8\sigma_1^2(p_1^*)m}\Big) + \sum_{j=2}^{m} C_j' \exp\Big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2 n_1}{2\sigma_j^2(e_j^*)m}\Big)\right],$$
$$C_1' = \frac{1}{1-\exp\Big(-\frac{(\tau_1-\tau_2)^2(n_2-n_1)}{8\sigma_1^2(e_1^*)}\Big)} = \frac{1}{1-\exp\Big(-\frac{(\tau_1-\tau_2)^2}{8\sigma_1^2(e_1^*)}\Big)},$$
$$C_j' = \frac{1}{1-\exp\Big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2(n_2-n_1)}{2\sigma_j^2(e_j^*)}\Big)} = \frac{1}{1-\exp\Big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2}{2\sigma_j^2(e_j^*)}\Big)}.$$

□

Proof of Theorem 2

Proof. Following Lemma 2, we can show that under our proposed design:

$$\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j)\right) \leq -\frac{(\tau_1-\tau_2)^2 c}{8\sigma_1^2(e_1^*)m}$$
$$+\frac{1}{N}\log\Big(1+\exp\big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2}{2\sigma_j^2(e_j^*)m}+\frac{(\tau_1-\tau_2)^2 n_1}{8\sigma_1^2(e_1^*)m}\big)\Big).$$

Under the complete randomization design:

$$\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j)\right) \leq -\frac{(\tau_1-\tau_2)^2 c}{8\sigma_1^2(1/2)m}$$
$$+\frac{1}{N}\log\Big(1+\exp\big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2}{2\sigma_j^2(1/2)m}+\frac{(\tau_1-\tau_2)^2 n_1}{8\sigma_1^2(1/2)m}\big)\Big).$$

When $N \to \infty$, we have

$$\lim_{N\to\infty}\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1(e_1^*) \geq \max_{2\leq j\leq m} \hat{\tau}_j(e_j^*))\right) = -\frac{(\tau_1-\tau_2)^2 c}{8\sigma_1^2(e_1^*)m},$$

and

$$\lim_{N\to\infty}\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1(1/2)\geq\max_{2\leq j\leq m}\hat{\tau}_j(1/2))\right)=-\frac{(\tau_1-\tau_2)^2c}{8\sigma_1^2(1/2)m}.$$

By our optimization problem setup, e_1^* maximizes the rate function. Therefore,

$$\lim_{N\to\infty}\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1(e_1^*)\geq\max_{2\leq j\leq m}\hat{\tau}_j(e_j^*))\right)\leq\lim_{N\to\infty}\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1(1/2)\geq\max_{2\leq j\leq m}\hat{\tau}_j(1/2))\right).$$

□

Proof of Proposition 1

Proof.

$$\mathbb{V}^{\texttt{Proposed}}=\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j}+\frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_j]]}{p_j(1-e_j)},$$
$$\mathbb{V}^{\texttt{CR}}=\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j\cdot 1/2}+\frac{\mathbb{V}[Y(0)|X\in\mathcal{S}_j]]}{p_j\cdot 1/2}.$$

Assume $\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]=\mathbb{V}[Y(0)|X\in\mathcal{S}_j]]$,

$$\mathbb{V}^{\texttt{Proposed}}=\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)},\quad \mathbb{V}^{\texttt{CR}}=\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j\cdot 1/2\cdot 1/2},$$
$$\text{When}\quad\frac{\mathbb{V}^{\texttt{CR}}}{\mathbb{V}^{\texttt{Proposed}}}=\frac{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j\cdot 1/2\cdot 1/2}}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)}}\geq 1,$$
$$\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j\cdot 1/2\cdot 1/2}\geq\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)},$$
$$4\geq\frac{1}{e_j(1-e_j)}$$

Given $\frac{(\tau_j-\tau_1)^2}{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}=\frac{1}{e_j(1-e_j)}$,

$$\mathbb{V}^{\texttt{Proposed}}\leq\mathbb{V}^{\texttt{CR}},\ \text{when}\ \frac{(\tau_j-\tau_1)^2}{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}\leq 4.$$

□

Proof of Proposition 2

To reach a correct selection probability level of $1-\delta$, we have, we have

$$1-\delta\geq 1-C_1\exp\Big(-\frac{(\tau_1-\tau_2)^2cN}{8\sigma_1^2(e_1^*)m}\Big)-\sum_{j=2}^{m}C_j\exp\Big(-\frac{(\tau_2-\tau_j+\frac{\tau_1-\tau_2}{2})^2}{2\sigma_j^2(e_j^*)m}\Big).$$

As $N \to \infty$,

$$\delta = C_1 \exp\Big(- \frac{(\tau_1 - \tau_2)^2 cN}{8\sigma_1^2(e_1^*)m}\Big),$$
$$\frac{1}{N}\log\delta \cdot \sigma_1^2(e_1) = (\tau_1 - \tau_2)^2,$$
$$\frac{1}{N}\log\delta \cdot \Big(\frac{\mathbb{V}(Y(1)|X \in \mathcal{S}_1)}{p_1 e_1} + \frac{\mathbb{V}(Y(0)|X \in \mathcal{S}_1)}{p_1(1-e_1)}\Big) = (\tau_1 - \tau_2)^2.$$

We assume $\mathbb{V}(Y(1)|X \in \mathcal{S}_j) = \mathbb{V}(Y(0)|X \in \mathcal{S}_j)$. Under the complete randomization design,

$$\frac{1}{N}\log\delta \cdot \Big(\frac{\mathbb{V}(Y(1)|X \in \mathcal{S}_1)}{p_1 \cdot \frac{1}{2}} + \frac{\mathbb{V}(Y(0)|X \in \mathcal{S}_1)}{p_1 \cdot \frac{1}{2}}\Big) = (\tau_1 - \tau_2)^2,$$
$$\frac{1}{N}\log\delta \cdot \Big(\frac{2\mathbb{V}(Y(1)|X \in \mathcal{S}_1)}{p_1 \cdot \frac{1}{2}}\Big) = (\tau_1 - \tau_2)^2,$$
$$\frac{1}{N^{\frac{1}{2}}} \cdot \sqrt{\log\delta \cdot C} = \tau_1 - \tau_2.$$

Under our proposed design,

$$\frac{1}{N}\log\delta \cdot \Big(\frac{\mathbb{V}(Y(1)|X \in \mathcal{S}_1)}{p_1 \cdot e_1^*} + \frac{\mathbb{V}(Y(1)|X \in \mathcal{S}_1)}{p_1 \cdot (1-e_1^*)}\Big) = (\tau_1 - \tau_2)^2,$$
$$\frac{1}{N}\log\delta \cdot C' \cdot (\tau_1 - \tau_2)^{\frac{1}{2}} = (\tau_1 - \tau_2)^2,$$
$$\frac{1}{N^{\frac{2}{3}}}\log\delta \cdot C' = \tau_1 - \tau_2.$$

In the following proof, we derive the closed-form expression of e_1^* under a simplified setting.

Proof. **Assume** $\mathbb{E}[Y(1)|X \in \mathcal{S}_j] = \mathbb{E}[Y(0)|X \in \mathcal{S}_j]$,

$$\sigma_j^2(e_j) = \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]]}{p_j e_j} + \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]]}{p_j(1-e_j)} = \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]]}{p_j e_j(1-e_j)}.$$

Step 1. Derive Optimal treatment allocation for subgroup $j \neq 1$.
Using the first-order optimality condition $G(\mathcal{S}_1, \mathcal{S}_j, \boldsymbol{e}) = G(\mathcal{S}_1, \mathcal{S}_k, \boldsymbol{e})$:

$$\frac{(\tau_j - \tau_1)^2}{\sigma_1^2(e_1) + \sigma_j^2(e_j)} = \frac{(\tau_k - \tau_1)^2}{\sigma_1^2(e_1) + \sigma_k^2(e_k)},$$
$$\frac{(\tau_j - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_1]]}{p_1 e_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]]}{p_j e_j(1-e_j)}} = \frac{(\tau_k - \tau_1)^2}{\frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_1]]}{p_1 e_1(1-e_1)} + \frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_k]]}{p_k e_k(1-e_k)}}.$$

Assume $\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]]}{p_1e_1(1-e_1)} \ll \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_1e_j(1-e_j)}$, $j=2,\ldots,d$,

$$\begin{aligned}
&\frac{(\tau_j-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)}} = \frac{(\tau_k-\tau_1)^2}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]]}{p_ke_k(1-e_k)}},\\
&\frac{(\tau_j-\tau_1)^2}{(\tau_k-\tau_1)^2} = \frac{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)}}{\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]]}{p_ke_k(1-e_k)}} = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)}\cdot\frac{p_ke_k(1-e_k)}{\mathbb{V}[Y(1)|X\in\mathcal{S}_k]]},\\
&\frac{e_j(1-e_j)}{e_k(1-e_k)} = \frac{p_k(\tau_k-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_k]]}{p_j(\tau_j-\tau_1)^2/\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}.
\end{aligned} \tag{3.19}$$

From Eq (1), $(\tau_j-\tau_1)^2 = \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j(1-e_j)}$, one can compute e_j^* as

$$\begin{aligned}
&e_j^2 - e_j + \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j(\tau_j-\tau_1)^2} = 0,\\
&e_j^* = \frac{1}{2} \pm \frac{1}{2}\sqrt{1-4\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j(\tau_j-\tau_1)^2}},\ j\neq 1.
\end{aligned}$$

Step 2. Derive Optimal treatment allocation for subgroup 1.
Using the first-order optimality condition $\sum_{j=2}^m \frac{\partial G(\mathcal{S}_1,\mathcal{S}_j,\boldsymbol{e})/\partial e_1}{\partial G(\mathcal{S}_1,\mathcal{S}_j,\boldsymbol{e})/\partial e_j} = 1$. Assume $\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]]}{p_1e_1(1-e_1)} \ll \frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_1e_j(1-e_j)}$, $j=2,\ldots,d$.

$$\begin{aligned}
\frac{\partial G(\mathcal{S}_1,\mathcal{S}_j,\boldsymbol{e})}{\partial e_1} &= -(\tau_j-\tau_1)^2\Big(\sigma_1^2(e_1)+\sigma_j^2(e_j)\Big)^{-2}\cdot\Big(\frac{-\mathbb{V}[Y(1)|X\in\mathcal{S}_1]]}{p_1e_1^2}+\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]]}{p_1(1-e_1)^2}\Big).\\
\frac{\partial G(\mathcal{S}_1,\mathcal{S}_j,\boldsymbol{e})}{\partial e_j} &= -(\tau_j-\tau_1)^2\Big(\sigma_1^2(e_1)+\sigma_j^2(e_j)\Big)^{-2}\cdot\Big(\frac{-\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j^2}+\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j(1-e_j)^2}\Big).
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=2}^m\Big\{\frac{\partial G(\mathcal{S}_1,\mathcal{S}_j,\boldsymbol{e})/\partial e_1}{\partial G(\mathcal{S}_1,\mathcal{S}_j,\boldsymbol{e})/\partial e_j}\Big\}\\
&=\sum_{j=2}^m\Big\{\Big(\frac{-\mathbb{V}[Y(1)|X\in\mathcal{S}_1]]}{p_1e_1^2}+\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]]}{p_1(1-e_1)^2}\Big)\Big/\Big(\frac{-\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_je_j^2}\\
&+\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]]}{p_j(1-e_j)^2}\Big)\Big\},\\
&=\sum_{j=2}^m\Big\{\Big(\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_1]](2e_1-1)}{p_1e_1^2(1-e_1)^2}\Big)\Big/\Big(\frac{\mathbb{V}[Y(1)|X\in\mathcal{S}_j]](2e_j-1)}{p_je_j^2(1-e_j)^2}\Big)\Big\}=1,
\end{aligned}$$

$$1 = \Big(\frac{\mathbb{V}[Y(1)|X \in \mathcal{S}_1]](2e_1-1)}{p_1 e_1^2(1-e_1)^2}\Big) \cdot \sum_{j=2}^{d} \frac{p_j e_j^2(1-e_j)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]](2e_j-1)},$$
$$\frac{p_1 e_1^2(1-e_1)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_1]](2e_1-1)} = \sum_{j=2}^{m} \frac{p_j e_j^2(1-e_j)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]](2e_j-1)}.$$

Let $a := p_1$, $b := \mathbb{V}[Y(1)|X \in \mathcal{S}_1]]$, $d := \sum_{j=2}^{m} \frac{p_j e_j^2(1-e_j)^2}{\mathbb{V}[Y(1)|X \in \mathcal{S}_j]](2e_j-1)}$,

$$\frac{a e_1^2(1-e_1)^2}{b(2e_1-1)} = d,$$
$$\text{Let } c = \frac{bd}{a},$$
$$e_1^4 - 2e_1^3 + e_1^2 - 2c \cdot e_1 + c = 0,$$
$$e_1^* = -\frac{-2}{4} \pm S \pm \frac{1}{2}\sqrt{-4S^2 - 2p \pm \frac{q}{S}},$$
$$\text{where } p = \frac{-1}{2},\ q = -2c,$$
$$S = \frac{1}{2}\sqrt{-\frac{2}{3}(-\frac{1}{2}) + \frac{1}{3}(Q + \frac{\Delta_0}{Q})},$$
$$Q = (\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2})^{1/3},$$
$$\Delta_0 = 1,\ \Delta_1 = 2 + 108c.$$
$$e_1^* = \frac{1}{2} \pm S \pm \frac{1}{2}\sqrt{-4S^2 + 1 \pm \frac{-2c}{S}},$$
$$S = \frac{1}{2}\sqrt{\frac{1}{3} + \frac{1}{3}(Q + \frac{1}{Q})},\ Q = \Big(\frac{(2+108c) + \sqrt{(2+108c)^2 - 4}}{2}\Big)^{1/3}$$

Therefore, $e_1^* = C'(\tau_2 - \tau_1)^{\frac{1}{2}}$. □

Proof for the multi-stage setting

Policy consistency proof

Proof.

$$\hat{e}_j = \frac{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j) D_i}{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)},$$
$$= \frac{\sum_{i=1}^{n_1} \mathbb{1}(X_{i1} \in \mathcal{S}_j) D_i + \sum_{t=2}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j) D_i}{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)},$$

Let U_i be an i.i.d uniform random variable in $[0,1]$, $e_{1j} = \frac{1}{2}$,

$$\begin{aligned}
\hat{e}_j &= \frac{\sum_{i=1}^{n_1} \mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq e_{1j}) + \sum_{t=2}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq \hat{e}_{tj})}{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)}, \\
&= \frac{n_1}{N}\frac{\frac{1}{n_1}\sum_{i=1}^{n_1} \mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq e_{1j})}{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)} + \sum_{t=2}^{T}\frac{n_t}{N}\frac{\frac{1}{n_t}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq \hat{e}_{tj})}{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}(X_{it} \in \mathcal{S}_j)}, \\
\text{where } &= \frac{1}{n_1}\sum_{i=1}^{n_1} \mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq e_{1j}) - \mathbb{E}[\mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq e_{1j})] + \mathbb{E}[\mathbb{1}(U_i \leq e_{1j})] \\
&= \mathbb{E}[\mathbb{1}(X_{it} \in \mathcal{S}_j)\mathbb{1}(U_i \leq e_{1j})] + O_p(\frac{1}{\sqrt{n_1}}), \\
&= e_{1j}\mathbb{P}(\mathcal{S}_j) + O_p(\frac{1}{\sqrt{n_1}}).
\end{aligned}$$

Next, we work on $\sqrt{n_t}\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(U_i \leq \hat{e}_{tj})$:

$$
\begin{aligned}
&\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq\hat{e}_{tj}) - \frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})\\
&+\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj}),\\
&+\sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})] - \sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq\hat{e}_{tj})]\\
&-\sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})] + \sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq\hat{e}_{tj})]\\
&=\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj}) + \sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq\hat{e}_{tj})]\\
&-\sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})]\\
&-\sqrt{n_t}\Big(\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj}) - \mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})]\Big)\\
&+\sqrt{n_t}\Big(\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq \hat{e}_{tj}) - \mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq \hat{e}_{tj})]\Big),\\
&=\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj}) + \sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq\hat{e}_{tj})]\\
&-\sqrt{n_t}\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})] + o_P(1),\\
&=\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})\\
&+\frac{\partial}{\partial e^*_{tj}}\Big(\mathbb{E}[\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})]\Big)\sqrt{n_t}(\hat{e}_{tj}-e^*_{tj}) + o_P(1),\\
&=\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq e^*_{tj})\\
&+\frac{\partial}{\partial e^*_{tj}}\Big(e^*_{tj}\Big)\sqrt{n_t}(\hat{e}_{tj}-e^*_{tj}) + o_P(1),\\
&=e^*_{tj}\mathbb{P}(\mathcal{S}_j) + O_p(\frac{1}{\sqrt{n_t}}) + \sqrt{n_t}(\hat{e}_{tj}-e^*_{tj}) + o_P(1),
\end{aligned}
$$

In sum, $\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\leq\hat{e}_{tj}) = e^*_{tj}\mathbb{P}(\mathcal{S}_j) + O_p(\frac{1}{\sqrt{n_t}})$.

Therefore,

$$\begin{aligned}
\hat{e}_j &= \frac{n_1}{N}\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\le e_{1j})}{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)} + \sum_{t=2}^{T}\frac{n_t}{N}\frac{\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)\mathbb{1}(U_i\le \hat{e}_{tj})}{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)}, \\
&= \frac{n_1}{N}\frac{e_{1j}\mathbb{P}(\mathcal{S}_j)+O_p(\frac{1}{\sqrt{n_1}})}{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)} + \sum_{t=2}^{T}\frac{n_t}{N}\frac{e^*_{tj}\mathbb{P}(\mathcal{S}_j)+O_p(\frac{1}{\sqrt{n_t}})}{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\mathbb{1}(X_{it}\in\mathcal{S}_j)}, \\
&= \frac{n_1}{N}(e_{1j}+O_p(\frac{1}{\sqrt{n_1}}) + \sum_{t=2}^{T}\frac{n_t}{N}\big(e^*_{tj}+O_p(\frac{1}{\sqrt{n_t}})\big), \\
&= e^*_j + O_p(\frac{1}{\sqrt{N}})
\end{aligned}$$

□

Proof of Theorem 1

Proof. Let $\hat{\tau}_j = \hat{\tau}_j^{\mathrm{T}} - \hat{\tau}_j^{\mathrm{C}}$. We want to show $\hat{\tau}_j^{\mathrm{T}} - \tau_j^{\mathrm{T}} = O_p(n^{-1/2})$ and $\hat{\tau}_j^{\mathrm{C}} - \tau_j^{\mathrm{C}} = O_p(n^{-1/2})$.

$$\begin{aligned}
\hat{\tau}_j^{\mathrm{T}} &= \frac{1}{\sum_{t=1}^{T} n_{tj}}\Big(\sum_{i=1}^{n_1}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{e_1} + \sum_{t=2}^{T}\sum_{i=1}^{n_t}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\hat{e}_{tj}}\Big), \\
&= \frac{1}{\sum_{t=1}^{T} n_{tj}}\Big(\sum_{i=1}^{n_1}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\mathbb{P}(D_i=1,X_i\in\mathcal{S}_j)/\mathbb{P}(\mathcal{S}_j)} + \sum_{t=2}^{T}\sum_{i=1}^{n_t}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\mathbb{P}(D_i=1,X_i\in\mathcal{S}_j)/\mathbb{P}(\mathcal{S}_j)}\Big), \\
&= \frac{1}{\sum_{t=1}^{T} n_{tj}}\Big(\frac{\sum_{i=1}^{n_1}\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\sum_{i=1}^{n_1}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)D_i}{n_{1j}}} + \sum_{t=2}^{T}\frac{\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\sum_{i=1}^{n_t}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)D_i}{n_{tj}}}\Big), \\
&= \frac{1}{\sum_{t=1}^{T} n_{tj}}\Big(n_{1j}\cdot\frac{\sum_{i=1}^{n_1}\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\sum_{i=1}^{n_1}\mathbb{1}(X_i\in\mathcal{S}_j)D_i} + \sum_{t=2}^{T} n_{tj}\cdot\frac{\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)D_iY_i}{\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)D_i}\Big).
\end{aligned}$$

Since the second stage data depend on the first stage data, we introduce a uniform random variable U_i, such that the first stage treatment is defined as $D_i = \mathbb{1}(U_i \le e_{1j})$, and the following stage treatment is defined as $D_i = \mathbb{1}(U_i \le \hat{e}_{tj})$, Therefore,

$$\begin{aligned}
\hat{\tau}_j^{\mathrm{T}} = \frac{1}{\sum_{t=1}^{T} n_{tj}}\Big(&n_{1j}\frac{n_1\cdot\frac{1}{n_1}\sum_{i=1}^{n_1}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_1)Y_i}{n_1\cdot\frac{1}{n_1}\sum_{i=1}^{m}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_1)} \\
&+ \sum_{t=2}^{T} n_{tj}\frac{n_t\cdot\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le \hat{e}_{tj})Y_i}{n_t\cdot\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le \hat{e}_{tj})}\Big).
\end{aligned}$$

For the numerators,

$$\begin{aligned}
\frac{1}{n_1}\sum_{i=1}^{n_1}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_1)Y_i &= \mathbb{E}[\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_1)Y_i] + O_p(n^{-1/2}),\\
&= e_1\mathbb{E}[Y_i(1)|X_i\in\mathcal{S}_j]\mathbb{P}(\mathcal{S}_j) + O_p(n^{-1/2}),\\
&= e_1\cdot\tau_j^{\mathrm{T}}\cdot\mathbb{P}(\mathcal{S}_j) + O_p(n^{-1/2}),\\
\frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le \hat{e}_{tj})Y_i &= \frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e^*_{tj})Y_i\\
&+ \frac{\partial}{\partial e^*_{tj}}\mathbb{E}[\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e^*_{tj})Y_i]\\
&\cdot\sqrt{n}(\hat{e}_{tj}(X_i) - e^*_{tj}) + O_p(1),\\
&= \frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e^*_{tj})Y_i\\
&+ \frac{\partial}{\partial e^*_{tj}}\Big(e^*_{tj}\tau_j^{\mathrm{T}}(X)\mathbb{P}(\mathcal{S}_j)\Big)\\
&\cdot\sqrt{n}(\hat{e}_{tj}(X) - e^*_{tj}(X)) + O_p(1),\\
&= \frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_{tj})Y_i + \Big(\tau_j^{\mathrm{T}}\mathbb{P}(\mathcal{S}_j)\Big)\\
&\cdot O_p(1) + O_p(1),\\
&= \sqrt{n_t}(e_{tj}\cdot\tau_j^{\mathrm{T}}\cdot\mathbb{P}(\mathcal{S}_j) + O_p(1)),\\
\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le \hat{e}_{tj})Y_i &= e_{tj}\cdot\tau_j^{\mathrm{T}}\cdot\mathbb{P}(\mathcal{S}_j) + O_p(n^{-1/2}).
\end{aligned}$$

For the denominators,

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_{1j})\\
&= \mathbb{E}[\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le e_{1j})] + O_p(n^{-1/2})),\\
&= e_{1j}\mathbb{P}(\mathcal{S}_j) + O_p(n^{-1/2})),\\
&\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(X_i\in\mathcal{S}_j)\mathbb{1}(U_i\le \hat{e}_{tj}(X)),\\
&= e^*_{tj}(X)\cdot\mathbb{P}(\mathcal{S}_j) + O_p(n^{-1/2})).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\hat{\tau}_j^{\mathrm{T}} &= \frac{1}{\sum_{t=1}^T n_t}\Bigg(n_{1j}\cdot\frac{n_1\cdot\Big(e_{1j}\tau_j^{\mathrm{T}}\mathbb{P}(\mathcal{S}_j)+O_p(n^{-1/2})\Big)}{m\cdot\Big(e_{1j}\mathbb{P}(\mathcal{S}_j)+O_p(n^{-1/2})\Big)} \\
&\quad+\sum_{t=2}^T n_{tj}\cdot\frac{n_t\cdot\Big(e_{tj}\tau_j^{\mathrm{T}}\mathbb{P}(\mathcal{S}_j)+O_p(n^{-1/2})\Big)}{n_t\cdot\Big(e_{tj}\mathbb{P}(\mathcal{S}_j)+O_p(n^{-1/2})\Big)}\Bigg), \\
&= \frac{n_{1j}\cdot\tau_j^{\mathrm{T}}+\sum_{t=2}^T n_{tj}\cdot\tau_j^{\mathrm{T}}}{\sum_{t=1}^T n_{tj}}+O_p(n^{-1/2}), \\
&= \tau_j^{\mathrm{T}}+O_p(n^{-1/2}).
\end{aligned}$$

Similarly, $\hat{\tau}_j^{\mathrm{C}} = \tau_j^{\mathrm{C}}+O_p(n^{-1/2})$.

Next, we want to prove the asymptotic normality of the proposed estimator in the multi-stage setting. Denote $\tau^{\mathrm{T}}(X_i) := \mathbb{E}[Y_i(1)|X_i]$, $\tau^{\mathrm{C}}(X_i) := \mathbb{E}[Y_i(0)|X_i]$.

$$\begin{aligned}
\sqrt{N}(\hat{\tau}_j-\tau_j) &= \frac{1}{\sqrt{N}}\Bigg(\sum_{i=1}^{n_1}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{D_i^*}{e_{1j}}\big(Y_i-\tau^{\mathrm{T}}(X_i)\big)+\tau^{\mathrm{T}}(X_i)\Big) \\
&\quad+\sum_{t=2}^T\sum_{i=1}^{n_t}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{D_i^*}{e_{tj}}\big(Y_i-\tau^{\mathrm{T}}(X_i)\big)+\tau^{\mathrm{T}}(X_i)\Big)\Bigg) \\
&\quad-\frac{1}{\sqrt{N}}\Bigg(\sum_{i=1}^{n_1}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{1-D_i^*}{1-e_{1j}}\big(Y_i-\tau^{\mathrm{C}}(X_i)\big)+\tau^{\mathrm{C}}(X_i)\Big) \\
&\quad+\sum_{t=2}^T\sum_{i=1}^{n_t}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{1-D_i^*}{1-e_{tj}}\big(Y_i-\tau^{\mathrm{C}}(X_i)\big)+\tau^{\mathrm{C}}(X_i)\Big)\Bigg)-\tau_j+o_p(1) \\
&= \frac{\sqrt{n_1}}{\sqrt{N}}D+\sum_{t=2}^T\frac{\sqrt{n_t}}{\sqrt{N}}E+o_p(1),
\end{aligned}$$

where

$$\begin{aligned}
D &:= \frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{D_i^*}{e_{1j}}\big(Y_i-\mu_1(X_i)\big)+\mu_1(X_i)\Big) \\
&\quad-\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{1-D_i^*}{1-e_{1j}}\big(Y_i-\mu_0(X_i)\big)+\mu_0(X_i)\Big)-\tau_j, \\
E &:= \frac{1}{\sqrt{n_t}}\sum_{i=1}^{n_t}\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{D_i^*}{e_{tj}}\big(Y_i-\mu_1(X_i)\big)+\mu_1(X_i)\Big) \\
&\quad-\frac{\mathbb{1}(X_i\in\mathcal{S}_j)}{\mathbb{P}(\mathcal{S}_j)}\Big(\frac{1-D_i^*}{1-e_{tj}}\big(Y_i-\mu_0(X_i)\big)+\mu_0(X_i)\Big)-\tau_j.
\end{aligned}$$

$$
\begin{aligned}
D \xrightarrow{d} & N\Bigg(0, \mathbb{E}\Big[\frac{\mathbb{1}(X_i \in \mathcal{S}_j)}{\mathbb{P}^2(\mathcal{S}_j)}\Big(\frac{1}{e_{1j}}(Y_i - \mu_1(X_i))^2 \\
& + \frac{1}{1-e_{1j}}(Y_i - \mu_0(X_i))^2 + \big(\mu_1(X_i) - \mu_0(X_i) - \tau_j\big)^2\Big)\Big]\Bigg), \\
E \xrightarrow{d} & N\Bigg(0, \mathbb{E}\Big[\frac{\mathbb{1}(X_i \in \mathcal{S}_j)}{\mathbb{P}^2(\mathcal{S}_j)}\Big(\frac{1}{e_{tj}}(Y_i - \mu_1(X_i))^2 \\
& + \frac{1}{1-e_{tj}}(Y_i - \mu_0(X_i))^2 + \big(\mu_1(X_i) - \mu_0(X_i) - \tau_j\big)^2\Big)\Big]\Bigg), \\
\frac{\sqrt{n_1}}{\sqrt{N}} D \xrightarrow{d} & N\Bigg(0, \mathbb{E}\Big[\frac{\mathbb{1}(X_i \in \mathcal{S}_j)}{\mathbb{P}^2(\mathcal{S}_j)}\Big(\frac{n_1/N}{e_{1j}}\sigma_1^2(X_i) \\
& + \frac{n_1/N}{1-e_{1j}}\sigma_0^2(X_i) + \frac{n_1}{N}\big(\tau^{\mathrm{T}}(X_i) - \tau^{\mathrm{C}}(X_i) - \tau_j\big)^2\Big)\Big]\Bigg), \\
\frac{\sqrt{n_t}}{\sqrt{N}} E \xrightarrow{d} & N\Bigg(0, \mathbb{E}\Big[\frac{\mathbb{1}(X_i \in \mathcal{S}_j)}{\mathbb{P}^2(\mathcal{S}_j)}\Big(\frac{n_t/N}{e_{tj}}\sigma_1^2(X_i) \\
& + \frac{n_t/N}{1-e_{tj}}\sigma_0^2(X_i) + \frac{n_t}{N}\big(\tau^{\mathrm{T}}(X_i) - \tau^{\mathrm{C}}(X_i) - \tau_j\big)^2\Big)\Big]\Bigg), \\
\sqrt{N}(\hat{\tau}_j - \tau_j) = & \frac{\sqrt{n_1}}{\sqrt{N}} D + \sum_{t=2}^{T} \frac{\sqrt{n_t}}{\sqrt{N}} E + o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
\sqrt{N}(\hat{\tau}_j - \tau_j) \xrightarrow{d} N\Bigg(0, & \mathbb{E}\Big[\frac{\mathbb{1}(X_i \in \mathcal{S}_j)}{\mathbb{P}^2(\mathcal{S}_j)}\Big(\big(\frac{n_1/N}{e_{1j}} + \sum_{t=2}^{T}\frac{n_t/N}{e_{tj}}\big)\sigma_1^2(X_i) \\
& + \big(\frac{n_1/N}{1-e_{1j}} + \sum_{t=2}^{T}\frac{n_t/N}{1-e_{tj}}\big)\sigma_0^2(X_i) \\
& + \big(\tau^{\mathrm{T}}(X_i) - \tau^{\mathrm{C}}(X_i) - \tau_j\big)^2\Big)\Big]\Bigg),
\end{aligned}
$$

where $\mu_1(X_i) = \mathbb{E}[Y_i(1)|X_i]$, $\mu_0(X_i) = \mathbb{E}[Y_i(0)|X_i]$.

□

Bibliography

[1] Ulf Aagerup. "The influence of real women in advertising on mass market fashion brand perception". In: *Journal of Fashion Marketing and Management: An International Journal* 15.4 (2011), pp. 486–502.

[2] Alberto Abadie and Matias D Cattaneo. "Econometric methods for program evaluation". In: *Annual Review of Economics* 10 (2018), pp. 465–503.

[3] Jason Abrevaya, Yu-Chin Hsu, and Robert P Lieli. "Estimating conditional average treatment effects". In: *Journal of Business & Economic Statistics* 33.4 (2015), pp. 485–505.

[4] Giacomo Aletti, Andrea Ghiglietti, and William F Rosenberger. "Nonparametric covariate-adjusted response-adaptive design based on a functional urn model". In: *The Annals of Statistics* 46.6B (2018), pp. 3838–3866.

[5] Stanislav Anatolyev. "Inference in regression models with many regressors". In: *Journal of Econometrics* 170.2 (2012), pp. 368–382.

[6] James Andreoni and John Miller. "Giving according to GARP: An experimental test of the consistency of preferences for altruism". In: *Econometrica* 70.2 (2002), pp. 737–753.

[7] Donald WK Andrews. "Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space". In: *Econometrica* (2000), pp. 399–405.

[8] Isaiah Andrews, Toru Kitagawa, and Adam McCloskey. *Inference on winners*. Tech. rep. National Bureau of Economic Research, 2019.

[9] Isaiah Andrews et al. "Inference for Losers". In: *American Economic Association Papers and Proceedings* (2022+).

[10] Manar Hamed Arafa and Hebatallah Husseini Atteia. "Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6) are associated with long term tramadol treatment-induced oxidative damage and hepatotoxicity". In: *Toxicology and Applied Pharmacology* 346 (2018), pp. 37–44.

[11] Susan F Assmann et al. "Subgroup analysis and other (mis) uses of baseline data in clinical trials". In: *The Lancet* 355.9209 (2000), pp. 1064–1069.

[12] Susan Athey and Guido Imbens. "Recursive partitioning for heterogeneous causal effects." In: *Proceedings of the National Academy of the Sciences* 113.27 (2016), pp. 7353–7360.

[13] Susan Athey and Guido W Imbens. "The state of applied econometrics: Causality and policy evaluation". In: *Journal of Economic Perspectives* 31.2 (2017), pp. 3–32.

[14] Susan Athey, Guido W Imbens, and Stefan Wager. "Approximate residual balancing: debiased inference of average treatment effects in high dimensions". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 80.4 (2018), pp. 597–623.

[15] G Ayano. "Psychotropic medications metabolized by cytochromes P450 (CYP1A2) enzyme and relevant drug interactions: Review of articles". In: *Austin Journal of Pharmacology and Therapeutics* 4.2 (2016), pp. 2–5.

[16] Philipp Bach et al. "DoubleML–An Object-Oriented Implementation of Double Machine Learning in R". In: *arXiv preprint arXiv:2103.09603* (2021).

[17] Zhi-Dong Bai and Feifang Hu. "Asymptotics in randomized urn models". In: *Advances In Statistics*. World Scientific, 2008, pp. 334–360.

[18] Uttam Bandyopadhyay and Atanu Biswas. "Allocation by randomized play-the-winner rule in the presence of prognostic factors". In: *Sankhyā: The Indian Journal of Statistics, Series B* (1999), pp. 397–412.

[19] Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. "Inference on treatment effects after selection among high-dimensional controls". In: *The Review of Economic Studies* 81.2 (2014), pp. 608–650.

[20] David Benkeser and Mark van der Laan. "The highly adaptive lasso estimator". In: *2016 IEEE international conference on data science and advanced analytics (DSAA)*. IEEE. 2016, pp. 689–696.

[21] Peter J Bickel et al. *Efficient and adaptive estimation for semiparametric models*. Vol. 4. Johns Hopkins University Press Baltimore, 1993.

[22] Leo Breiman. "Random forests". In: *Machine learning* 45.1 (2001), pp. 5–32.

[23] Matias D Cattaneo, Michael Jansson, and Xinwei Ma. "Two-step estimation and inference with possibly many included covariates". In: *The Review of Economic Studies* 86.3 (2019), pp. 1095–1122.

[24] Matias D Cattaneo, Michael Jansson, and Whitney K Newey. "Alternative asymptotics and the partially linear model with many regressors". In: *Econometric Theory* 34.2 (2018), pp. 277–301.

[25] Matias D Cattaneo, Michael Jansson, and Whitney K Newey. "Inference in linear regression models with many covariates and heteroscedasticity". In: *Journal of the American Statistical Association* 113.523 (2018), pp. 1350–1361.

[26] Gary Charness and Dan Levin. "The origin of the winner's curse: a laboratory study". In: *American Economic Journal: Microeconomics* 1.1 (2009), pp. 207–36.

[27] Shreyasi Chatterjee and Amritpal Mudher. "Alzheimer's disease and type 2 diabetes: a critical assessment of the shared pathological traits". In: *Frontiers in neuroscience* 12 (2018), p. 383.

[28] Tianqi Chen and Carlos Guestrin. "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794.

[29] Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. "Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors". In: *The Annals of Statistics* 41.6 (2013), pp. 2786–2819.

[30] Victor Chernozhukov, Sokbae Lee, and Adam M Rosen. "Intersection bounds: estimation and inference". In: *Econometrica* 81.2 (2013), pp. 667–737.

[31] Victor Chernozhukov and Vira Semenova. *Simultaneous inference for best linear predictor of the conditional average treatment effect and other structural functions*. Tech. rep. cemmap working paper, 2018.

[32] Victor Chernozhukov et al. *Double/debiased machine learning for treatment and structural parameters*. 2018.

[33] Victor Chernozhukov et al. "Double/debiased/neyman machine learning of treatment effects". In: *American Economic Review* 107.5 (2017), pp. 261–65.

[34] Melissa D Cinelli and Lifeng Yang. "The role of implicit theories in evaluations of "plus-size" advertising". In: *Journal of Advertising* 45.4 (2016), pp. 472–481.

[35] Brian Claggett, Minge Xie, and Lu Tian. "Meta-analysis with fixed, unknown, study-specific parameters". In: *Journal of the American Statistical Association* 109.508 (2014), pp. 1660–1671.

[36] David I Cook, Val J Gebski, and Anthony C Keech. "Subgroup analysis in clinical trials". In: *Medical Journal of Australia* 180.6 (2004), p. 289.

[37] Richard K Crump et al. "Dealing with limited overlap in estimation of average treatment effects". In: *Biometrika* 96.1 (2009), pp. 187–199.

[38] AP Dawid. "Selection paradoxes of Bayesian inference". In: *Multivariate analysis and its applications*. Institute of Mathematical Statistics, 1994, pp. 211–220.

[39] Rajeev H Dehejia and Sadek Wahba. "Propensity score-matching methods for non-experimental causal studies". In: *Review of Economics and statistics* 84.1 (2002), pp. 151–161.

[40] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*. Vol. 38. Springer Science & Business Media, 2009.

[41] Habiba Djebbari and Jeffrey Smith. "Heterogeneous impacts in PROGRESA". In: *Journal of Econometrics* 145.1-2 (2008), pp. 64–80.

[42] David L Donoho and Jain M Johnstone. "Ideal spatial adaptation by wavelet shrinkage". In: *biometrika* 81.3 (1994), pp. 425–455.

[43] MD Donsker and SR Srinivasa Varadhan. "On the principal eigenvalue of second-order elliptic differential operators". In: *Communications on Pure and Applied Mathematics* 29.6 (1976), pp. 595–621.

[44] Bradley Efron. "Forcing a sequential experiment to be balanced". In: *Biometrika* 58.3 (1971), pp. 403–417.

[45] Bradley Efron. "Tweedie's formula and selection bias". In: *Journal of the American Statistical Association* 106.496 (2011), pp. 1602–1614.

[46] Florian Eggenberger and George Pólya. "Über die statistik verketteter vorgänge". In: *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 3.4 (1923), pp. 279–289.

[47] Friedhelm Eicker. "Asymptotic normality and consistency of the least squares estimators for families of linear regressions". In: *The annals of mathematical statistics* (1963), pp. 447–456.

[48] Jeffrey R Eisele. "The doubly adaptive biased coin design for sequential clinical trials". In: *Journal of Statistical Planning and Inference* 38.2 (1994), pp. 249–261.

[49] Noureddine El Karoui and Elizabeth Purdom. "Can we trust the bootstrap in high-dimensions? The case of linear models". In: *The Journal of Machine Learning Research* 19.1 (2018), pp. 170–235.

[50] Noureddine El Karoui et al. "On robust regression with high-dimensional predictors". In: *Proceedings of the National Academy of Sciences* 110.36 (2013), pp. 14557–14562.

[51] Jianqing Fan, Peter Hall, and Qiwei Yao. "To how many simultaneous hypothesis tests can normal, student's t or bootstrap calibration be applied?" In: *Journal of the American Statistical Association* 102.480 (2007), pp. 1282–1288.

[52] Jianqing Fan, Fang Han, and Han Liu. "Challenges of big data analysis". In: *National science review* 1.2 (2014), pp. 293–314.

[53] Jianqing Fan and Runze Li. "Variable selection via nonconcave penalized likelihood and its oracle properties". In: *Journal of the American statistical Association* 96.456 (2001), pp. 1348–1360.

[54] Jianqing Fan, Jinchi Lv, and Lei Qi. "Sparse high-dimensional models in economics". In: *Annu. Rev. Econ.* 3.1 (2011), pp. 291–317.

[55] Jianqing Fan, Qi-Man Shao, and Wen-Xin Zhou. "Are discoveries spurious? Distributions of maximum spurious correlations and their applications". In: *Annals of statistics* 46.3 (2018), p. 989.

[56] Jianqing Fan and Wen-Xin Zhou. "Guarding against spurious discoveries in high dimensions". In: *The Journal of Machine Learning Research* 17.1 (2016), pp. 7123–7156.

[57] Jianqing Fan et al. "Optimal covariate balancing conditions in propensity score estimation". In: *Journal of Business & Economic Statistics* (2021), pp. 1–14.

[58] Jianqing Fan et al. *Statistical foundations of data science*. Chapman and Hall/CRC, 2020.

[59] Max H Farrell. "Robust inference on average treatment effects with possibly more covariates than observations". In: *Journal of Econometrics* 189.1 (2015), pp. 1–23.

[60] Dean Follmann. "Adaptively changing subgroup proportions in clinical trials". In: *Statistica Sinica* (1997), pp. 1085–1102.

[61] Lei Gao et al. "Shared genetic etiology between type 2 diabetes and Alzheimer's disease identified by bioinformatics analysis". In: *Journal of Alzheimer's Disease* 50.1 (2016), pp. 13–17.

[62] Paul J Gertler, Sebastian W Martinez, and Marta Rubio-Codina. "Investing cash transfers to raise long-term living standards". In: *American Economic Journal: Applied Economics* 4.1 (2012), pp. 164–92.

[63] Susan Gruber and Mark J van der Laan. "A targeted maximum likelihood estimator of a causal effect on a bounded continuous outcome". In: *The International Journal of Biostatistics* 6.1 (2010).

[64] Xinzhou Guo and Xuming He. "Inference on selected subgroups in clinical trials". In: *Journal of the American Statistical Association* (2020), pp. 1–19.

[65] Xinzhou Guo et al. "Assessing Heterogeneous Risk of Type II Diabetes Associated with Statin Usage: Evidence from Electronic Health Record Data". In: *arXiv preprint arXiv:2205.06960* (2022).

[66] Xinzhou Guo et al. "Sharp Inference on Selected Subgroups in Observational Studies". In: *arXiv preprint arXiv:2102.11338* (2021).

[67] LLC Gurobi Optimization. *Gurobi optimizer reference manual*. 2018.

[68] Jinyong Hahn. "On the role of the propensity score in efficient semiparametric estimation of average treatment effects". In: *Econometrica* (1998), pp. 315–331.

[69] Jinyong Hahn, Keisuke Hirano, and Dean Karlan. "Adaptive experimental design using the propensity score". In: *Journal of Business & Economic Statistics* 29.1 (2011), pp. 96–108.

[70] Jaroslav Hájek. "Comment on An essay on the logical foundations of survey sampling by Basu, D". In: *Foundations of Statistical Inference* 236 (1971).

[71] Peter Hall and Christopher C Heyde. *Martingale limit theory and its application*. Academic press, 2014.

[72] Peter Hall and Hugh Miller. "Bootstrap confidence intervals and hypothesis tests for extrema of parameters". In: *Biometrika* 97.4 (2010), pp. 881–892.

[73] Bruce M Hill, David Lane, and William Sudderth. "A strong law for some generalized urn processes". In: *The Annals of Probability* (1980), pp. 214–226.

[74] Jennifer L Hill. "Bayesian nonparametric modeling for causal inference". In: *Journal of Computational and Graphical Statistics* 20.1 (2011), pp. 217–240.

[75] W Hoeffding. "On a theorem of VM Zolotarev". In: *Theory of Probability & Its Applications* 9.1 (1964), pp. 89–91.

[76] Wassily Hoeffding. "Asymptotically optimal tests for multinomial distributions". In: *The Annals of Mathematical Statistics* (1965), pp. 369–401.

[77] Frank Hollander. *Large deviations*. Vol. 14. American Mathematical Soc., 2000.

[78] Feifang Hu and William F Rosenberger. "Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons". In: *Journal of the American Statistical Association* 98.463 (2003), pp. 671–678.

[79] Feifang Hu and William F Rosenberger. *The theory of response-adaptive randomization in clinical trials*. John Wiley & Sons, 2006.

[80] Feifang Hu and Li-Xin Zhang. "Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials". In: *The Annals of Statistics* 32.1 (2004), pp. 268–301.

[81] Feifang Hu, Li-Xin Zhang, and Xuming He. "Efficient randomized-adaptive designs". In: *The Annals of Statistics* (2009), pp. 2543–2560.

[82] Jianhua Hu, Hongjian Zhu, and Feifang Hu. "A unified family of covariate-adjusted response-adaptive designs based on efficiency and ethics". In: *Journal of the American Statistical Association* 110.509 (2015), pp. 357–367.

[83] Ying Huang, Peter B Gilbert, and Holly Janes. "Assessing treatment-selection markers using a potential outcomes framework". In: *Biometrics* 68.3 (2012), pp. 687–696.

[84] Peter J Huber. "Robust regression: asymptotics, conjectures and Monte Carlo". In: *The annals of statistics* (1973), pp. 799–821.

[85] Kosuke Imai and Marc Ratkovic. "Covariate balancing propensity score". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76.1 (2014), pp. 243–263.

[86] Kosuke Imai and Marc Ratkovic. "Estimating treatment effect heterogeneity in randomized program evaluation". In: *The Annals of Applied Statistics* 7.1 (2013), pp. 443–470.

[87] Kosuke Imai et al. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies". In: *American Political Science Review* 105.4 (2011), pp. 765–789.

[88] Svante Janson. "Functional limit theorems for multitype branching processes and generalized Pólya urns". In: *Stochastic Processes and their Applications* 110.2 (2004), pp. 177–245.

[89] Koen Jochmans. "Heteroscedasticity-robust inference in linear regression models with many covariates". In: *Journal of the American Statistical Association* (2020), pp. 1–10.

[90] Bo Ra Joo and Juanjuan Wu. "The impact of inclusive fashion advertising with plus-size models on female consumers: The mediating role of brand warmth". In: *Journal of Global Fashion Marketing* 12.3 (2021), pp. 260–273.

[91] Joseph DY Kang and Joseph L Schafer. "Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data". In: *Statistical science* 22.4 (2007), pp. 523–539.

[92] Dean Karlan and John A List. "Does price matter in charitable giving? Evidence from a large-scale natural field experiment". In: *American Economic Review* 97.5 (2007), pp. 1774–1793.

[93] Dean S Karlan and Jonathan Zinman. "Credit elasticities in less-developed economies: Implications for microfinance". In: *American Economic Review* 98.3 (2008), pp. 1040–68.

[94] Shakeeb Khan and Elie Tamer. "Irregular identification, support conditions, and inverse weight estimation". In: *Econometrica* 78.6 (2010), pp. 2021–2042.

[95] Eugene Kharitonov et al. "Sequential testing for early stopping of online experiments". In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2015, pp. 473–482.

[96] K Kubota et al. "Phase III study (MONET1) of motesanib plus carboplatin/paclitaxel in patients with advanced nonsquamous nonsmall-cell lung cancer (NSCLC): Asian subgroup analysis". In: *Annals of oncology* 25.2 (2014), pp. 529–536.

[97] Jannis Kueck et al. "Estimation and Inference of Treatment Effects with L_2-Boosting in High-Dimensional Settings". In: *arXiv preprint arXiv:1801.00364* (2017).

[98] Sören R Künzel et al. "Metalearners for estimating heterogeneous treatment effects using machine learning". In: *Proceedings of the national academy of sciences* 116.10 (2019), pp. 4156–4165.

[99] Mark van der Laan and Susan Gruber. "One-step targeted minimum loss-based estimation based on universal least favorable one-dimensional submodels". In: *The international journal of biostatistics* 12.1 (2016), pp. 351–378.

[100] Mark J van der Laan and Susan Gruber. "Collaborative Double Robust Targeted Maximum Likelihood Estimation". In: *The International Journal of Biostatistics* 6.1 (2010), p. 17.

[101] Mark J van der Laan and James M Robins. *Unified methods for censored longitudinal data and causality*. Springer Science & Business Media, 2003.

[102] Mark J van der Laan and Sherri Rose. *Targeted learning in data science*. Springer, 2018.

[103] Mark J van der Laan and Sherri Rose. *Targeted learning: causal inference for observational and experimental data*. Springer Science & Business Media, 2011.

[104] Mark J van der Laan and Donald Rubin. "Targeted Maximum Likelihood Learning". In: *The International Journal of Biostatistics* 2.1 (2006).

[105] Tze Leung Lai, Philip W Lavori, and Ka Wai Tsang. "Adaptive enrichment designs for confirmatory trials". In: *Statistics in Medicine* 38.4 (2019), pp. 613–624.

[106] Daniel Lakens, Friedrich Pahlke, and Gernot Wassmer. "Group sequential designs: A tutorial". In: (2021).

[107] Robert J LaLonde. "Evaluating the econometric evaluations of training programs with experimental data". In: *The American economic review* (1986), pp. 604–620.

[108] Minyong R Lee and Milan Shen. "Winner's curse: Bias estimation for total effects of features in online controlled experiments". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018, pp. 491–499.

[109] Emma E Levine and Maurice E Schweitzer. "The affective and interpersonal consequences of obesity". In: *Organizational Behavior and Human Decision Processes* 127 (2015), pp. 66–84.

[110] Jonathan Levy et al. "A fundamental measure of treatment effect heterogeneity". In: *Journal of Causal Inference* 9.1 (2021), pp. 83–108.

[111] John A Lewis. "In defence of the dichotomy". In: *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry* 3.2 (2004), pp. 77–79.

[112] Lin Li and Christian Hölscher. "Common pathological processes in Alzheimer disease and type 2 diabetes: a review". In: *Brain research reviews* 56.2 (2007), pp. 384–402.

[113] Yunzhi Lin, Ming Zhu, and Zheng Su. "The pursuit of balance: an overview of covariate-adaptive randomization techniques in clinical trials". In: *Contemporary clinical trials* 45 (2015), pp. 21–25.

[114] John A List. "The market for charitable giving". In: *Journal of Economic Perspectives* 25.2 (2011), pp. 157–80.

[115] J Scott Long and Laurie H Ervin. "Using heteroscedasticity consistent standard errors in the linear regression model". In: *The American Statistician* 54.3 (2000), pp. 217–224.

[116] Jacqueline MacArthur et al. "The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog)". In: *Nucleic acids research* 45.D1 (2017), pp. D896–D901.

[117] Enno Mammen. "Bootstrap and wild bootstrap for high dimensional linear models". In: *The annals of statistics* (1993), pp. 255–285.

[118] Enno Mammen. *When does bootstrap work?: asymptotic results and simulations*. Vol. 77. Springer Science & Business Media, 2012.

[119] Roland A Matsouaka, Junlong Li, and Tianxi Cai. "Evaluating marker-guided treatment selection strategies". In: *Biometrics* 70.3 (2014), pp. 489–499.

[120] Samia Mora et al. "Statins for the primary prevention of cardiovascular events in women with elevated high-sensitivity C-reactive protein or dyslipidemia: results from the Justification for the Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER) and meta-analysis of women from primary prevention trials". In: *Circulation* 121.9 (2010), pp. 1069–1077.

[121] Kari Lock Morgan and Donald B Rubin. "Rerandomization to balance tiers of covariates". In: *Journal of the American Statistical Association* 110.512 (2015), pp. 1412–1421.

[122] Kari Lock Morgan and Donald B Rubin. "Rerandomization to improve covariate balance in experiments". In: *The Annals of Statistics* 40.2 (2012), pp. 1263–1282.

[123] Whitney K Newey. "Semiparametric efficiency bounds". In: *Journal of applied econometrics* 5.2 (1990), pp. 99–135.

[124] Whitney K Newey. "The asymptotic variance of semiparametric estimators". In: *Econometrica: Journal of the Econometric Society* (1994), pp. 1349–1382.

[125] Jerzy S. Neyman. "On the application of probability theory to agricultural experiments. essay on principles. section 9.(statistical science (1990), 5, 465-480)". In: *Annals of Agricultural Sciences* 10 (1923), pp. 1–51.

[126] Judea Pearl, Madelyn Glymour, and Nicholas P Jewell. *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.

[127] Maya L Petersen et al. "Diagnosing and responding to violations in the positivity assumption". In: *Statistical methods in medical research* 21.1 (2012), pp. 31–54.

[128] Stuart J Pocock and Richard Simon. "Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial". In: *Biometrics* (1975), pp. 103–115.

[129] C Rafaniello et al. "The predictive value of ABCB1, ABCG2, CYP3A4/5 and CYP2D6 polymorphisms for risperidone and aripiprazole plasma concentrations and the occurrence of adverse drug reactions". In: *The pharmacogenomics journal* 18.3 (2018), pp. 422–430.

[130] Yoshiyasu Rai. "Statistical inference for treatment assignment policies". In: *Unpublished Manuscript* (2018).

[131] Stephen W Raudenbush and Howard S Bloom. "Learning about and from a distribution of program impacts using multisite trials". In: *American Journal of Evaluation* 36.4 (2015), pp. 475–499.

[132] Bruce Reed et al. "Associations between serum cholesterol levels and cerebral amyloidosis". In: *JAMA neurology* 71.2 (2014), pp. 195–200.

[133] National Science Foundation reports. "Statistics at a crossroad: who is for the challenge?" In: (2019).

[134] David S Robertson et al. "Response-adaptive randomization in clinical trials: from myths to practical considerations". In: *arXiv preprint arXiv:2005.00564* (2020).

[135] James Robins and Andrea Rotnitzky. "Recovery of information and adjustment for dependent censoring using surrogate markers." In: *AIDS Epidemiology, Methodological issues Bikhäuser* (1992), pp. 297–331.

[136] James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. "Estimation of regression coefficients when some regressors are not always observed". In: *Journal of the American statistical Association* 89.427 (1994), pp. 846–866.

[137] Paul R Rosenbaum and Donald B Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects". In: *Biometrika* 70 (1983), pp. 41–55.

[138] Paul R Rosenbaum and Donald B Rubin. "The central role of the propensity score in observational studies for causal effects". In: *Biometrika* 70.1 (1983), pp. 41–55.

[139] William F Rosenberger. "Randomized urn models and sequential design". In: (2002).

[140] William F Rosenberger and Feifang Hu. "Maximizing power and minimizing treatment failures in clinical trials". In: *Clinical Trials* 1.2 (2004), pp. 141–147.

[141] William F Rosenberger and John M Lachin. "The use of response-adaptive designs in clinical trials". In: *Controlled clinical trials* 14.6 (1993), pp. 471–484.

[142] William F Rosenberger, AN Vidyashankar, and Deepak K Agarwal. "Covariate-adjusted response-adaptive designs for binary response". In: *Journal of biopharmaceutical statistics* 11.4 (2001), pp. 227–236.

[143] Michael Rosenblum, Ethan X Fang, and Han Liu. "Optimal, two-stage, adaptive enrichment designs for randomized trials, using sparse linear programming". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82.3 (2020), pp. 749–772.

[144] Michael Rosenblum and Mark J van der Laan. "Optimizing randomized trial designs to distinguish which subpopulations benefit from treatment". In: *Biometrika* 98.4 (2011), pp. 845–860.

[145] Michael Rosenblum, Han Liu, and En-Hsu Yen. "Optimal tests of treatment effects for the overall population and two subpopulations in randomized trials, using sparse linear programming". In: *Journal of the American Statistical Association* 109.507 (2014), pp. 1216–1228.

[146] Gualberto Ruaño et al. "Length of psychiatric hospitalization is correlated with CYP2D6 functional status in inpatients with major depressive disorder". In: *Biomarkers in medicine* 7.3 (2013), pp. 429–439.

[147] Gualberto Ruaño et al. "Subanalysis of the CYP-GUIDES Trial: CYP2D6 functional stratification and operational timeline selection". In: *Psychiatry Research* 297 (2021), p. 113571.

[148] Donald B Rubin. "Estimating causal effects of treatments in randomized and nonrandomized studies." In: *Journal of educational Psychology* 66.5 (1974), p. 688.

[149] Daniel J Russo et al. "A tutorial on thompson sampling". In: *Foundations and Trends® in Machine Learning* 11.1 (2018), pp. 1–96.

[150] Noah Simon and Richard Simon. "Adaptive enrichment designs for clinical trials". In: *Biostatistics* 14.4 (2013), pp. 613–625.

[151] Nigel Stallard. "Adaptive enrichment designs with a continuous biomarker". In: *Biometrics* (2022).

[152] Nigel Stallard, Susan Todd, and John Whitehead. "Estimation following selection of the largest of two normal means". In: *Journal of Statistical Planning and Inference* 138.6 (2008), pp. 1629–1638.

[153] James H Stock and Mark W Watson. "Heteroskedasticity-robust standard errors for fixed effects panel data regression". In: *Econometrica* 76.1 (2008), pp. 155–174.

[154] Xiaogang Su et al. "Subgroup analysis via recursive partitioning". In: *Journal of Machine Learning Research* 10.Feb (2009), pp. 141–158.

[155] Daniel I Swerdlow et al. "HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials". In: *The Lancet* 385.9965 (2015), pp. 351–361.

[156] Joseph Tortora et al. "Clinical database of the CYP-guides trial: An open data resource on psychiatric hospitalization for severe depression". In: *Data in Brief* 30 (2020), p. 105457.

[157] Hugo Touchette. "The large deviation approach to statistical mechanics". In: *Physics Reports* 478.1-3 (2009), pp. 1–69.

[158] Anastasios A Tsiatis. *Semiparametric theory and missing data*. Springer, 2006.

[159] Yevgen Tymofyeyev, William F Rosenberger, and Feifang Hu. "Implementing optimal allocation in sequential binary response experiments". In: *Journal of the American Statistical Association* 102.477 (2007), pp. 224–234.

[160] Sara Van de Geer et al. "On asymptotically optimal confidence regions and tests for high-dimensional models". In: *The Annals of Statistics* 42.3 (2014), pp. 1166–1202.

[161] Aad W Van der Vaart. *Asymptotic statistics*. Vol. 3. Cambridge university press, 2000.

[162] Tyler J VanderWeele et al. "Selecting optimal subgroups for treatment using many covariates". In: *Epidemiology* 30.3 (2019), p. 334.

[163] Gonzalo Vazquez-Bare. "Identification and estimation of spillover effects in randomized experiments". In: *Journal of Econometrics* Forthcoming (2021).

[164] Valentin Verdier. "Estimation and inference for linear models with two-way fixed effects and sparsely matched data". In: *Review of Economics and Statistics* 102.1 (2020), pp. 1–16.

[165] Sofia S Villar, Jack Bowden, and James Wason. "Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges". In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 30.2 (2015), p. 199.

[166] Sofia S Villar, Jack Bowden, and James Wason. "Response-adaptive designs for binary responses: how to offer patient benefit while being robust to time trends?" In: *Pharmaceutical statistics* 17.2 (2018), pp. 182–197.

[167] Sofia S Villar and William F Rosenberger. "Covariate-adjusted response-adaptive randomization for multi-arm clinical trials using a modified forward looking Gittins index rule". In: *Biometrics* 74.1 (2018), pp. 49–57.

[168] Andreas Wächter and Lorenz T Biegler. "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming". In: *Mathematical programming* 106.1 (2006), pp. 25–57.

[169] Stefan Wager and Susan Athey. "Estimation and inference of heterogeneous treatment effects using random forests". In: *Journal of the American Statistical Association* (2017). DOI: doi.org/10.1080/01621459.2017.1319839.

[170] Willem M van der Wal and Ronald B Geskus. "ipw: an R package for inverse probability weighting". In: *Journal of Statistical Software* 43 (2011), pp. 1–23.

[171] Mengting Wan et al. "Addressing marketing bias in product recommendations". In: *Proceedings of the 13th international conference on web search and data mining*. 2020, pp. 618–626.

[172] Sue-Jane Wang, Robert T O'Neill, and HM James Hung. "Approaches to evaluation of treatment effect in randomized clinical trials with genomic subset". In: *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry* 6.3 (2007), pp. 227–244.

[173] Mal Warwick. *Testing, Testing 1, 2, 3: Raise More Money with Direct Mail Tests*. John Wiley & Sons, 2003.

[174] LJ Wei and S Durham. "The randomized play-the-winner rule in medical trials". In: *Journal of the American Statistical Association* 73.364 (1978), pp. 840–843.

[175] Halbert White. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". In: *Econometrica: journal of the Econometric Society* (1980), pp. 817–838.

[176] Marvin N Wright, S Wager, and P Probst. "Ranger: A fast implementation of random forests". In: *R package version 0.12* 1 (2020).

[177] Peter Würtz et al. "Metabolomic profiling of statin use and genetic inhibition of HMG-CoA reductase". In: *Journal of the American college of cardiology* 67.10 (2016), pp. 1200–1210.

[178] Minge Xie, Kesar Singh, and Cun-Hui Zhang. "Confidence intervals for population ranks in the presence of ties and near ties". In: *Journal of the American Statistical Association* 104.486 (2009), pp. 775–788.

[179] Yanxun Xu et al. "Subgroup-based adaptive (SUBA) designs for multi-arm biomarker trials". In: *Statistics in Biosciences* 8.1 (2016), pp. 159–180.

[180] Marvin Zelen. "Play the winner rule and the controlled clinical trial". In: *Journal of the American Statistical Association* 64.325 (1969), pp. 131–146.

[181] Marvin Zelen. "The randomization and stratification of patients to clinical trials". In: *Journal of Chronic Diseases* (1994), pp. 365–375.

[182] Cun-Hui Zhang and Stephanie S Zhang. "Confidence intervals for low dimensional parameters in high dimensional linear models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76.1 (2014), pp. 217–242.

[183] Lanju Zhang and William F Rosenberger. "Response-adaptive randomization for clinical trials with continuous outcomes". In: *Biometrics* 62.2 (2006), pp. 562–569.

[184] Li-Xin Zhang et al. "Asymptotic properties of covariate-adjusted response-adaptive designs". In: *The Annals of Statistics* 35.3 (2007), pp. 1166–1182.

[185] Wanying Zhao et al. "Incorporating covariates information in adaptive clinical trials for precision medicine". In: *Pharmaceutical Statistics* 21.1 (2022), pp. 176–195.

[186] Wenjing Zheng and Mark J van der Laan. "Asymptotic theory for cross-validated targeted maximum likelihood estimation". In: *U.C. Berkeley Division of Biostatistics Working Paper Series* (2010).

[187] Quan Zhou et al. "Sequential rerandomization". In: *Biometrika* 105.3 (2018), pp. 745–752.

[188] Julie M Zissimopoulos et al. "Sex and race differences in the association between statin use and the incidence of Alzheimer disease". In: *JAMA neurology* 74.2 (2017), pp. 225–232.