UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Policy Design and Statistical Inference

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Davide Viviano

Committee in charge:

Professor Graham Elliott, Chair
Professor Jelena Bradic
Professor James Fowler
Professor Paul Niehaus
Professor Yixiao Sun
Professor Kaspar Wüthrich

2022

The Dissertation of Davide Viviano is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

To my parents, brother and grandparents. To Holly and her love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I acknowledge Graham Elliott as my committee chair for his constant support. Since my first year, Graham has always been a tremendous mentor and source of professional and personal advice. It has been an honor to become a scholar under his guidance.

Yixiao Sun took an early interest in my research and has been an invaluable source of support since then. Our discussions were always exciting and illuminating. I am very grateful for his help throughout this Ph.D.

Kaspar Wüthrich has provided thoughtful guidance throughout this Ph.D. journey. He has been an incredible mentor and collaborator and has always been very generous with his time.

Jelena Bradic taught me how to be a tenacious and analytical researcher from the beginning of this Ph.D. Our collaborations have led to exciting contributions to economics and statistics.

James Fowler's joy and excitement have continuously inspired and guided me since the second year of my Ph.D. His way of thinking has had a tremendous impact on me as a scholar.

Paul Niehaus provided extensive support through mentorship and collaborations during the last years of this Ph.D. His guidance has been invaluable to interpret economic problems from an interdisciplinary and critical perspective.

I want to thank my professors and mentors at UCSD, Barcelona GSE, and Luiss University. I am particularly thankful to Giuseppe Ragusa, who introduced me to econometrics during my bachelor's at Luiss and has been an incredible source of support during my first steps as an undergrad. I am thankful to Omiros Papaspiliopoulos, David Rossell, Christian Brownleess, and Gabor Lugosi for teaching me statistical inference and machine learning at Barcelona GSE.

I am also indebted to many people who provided professional support throughout this Ph.D. with helpful feedback and comments. A non-comprehensive list includes Brendan Beare, Tim Christensen, Itzik Fadlon, James Hamilton, Brian Karrer, Toru Kitagawa, Xinwei Ma, Craig Mcintosh, Karthik Muralidharan, Fredrik Sävje, Joel Sobel, Alex Tetenov, Ye Wang, Yinchu Zhu, Ying Zhu.

On a personal level, I would like to thank my parents, brother, grandparents, and my grandfather Luigi in particular. I could not live without their unconditional love. I am very thankful to Holly for her love and support through these years of our Ph.D. It was the biggest blessing to live this adventure together.

I am grateful to my friends in Rome, who, thousands of miles away, never made me feel far away. I am particularly thankful for the support of my best friend, Diletta. She has always guided me through a happy life. I am grateful for the love and support from Rome of the two Federico, Andrea, Cristiano, Fabrizio, Maurizio, and Mary. Our friendship is one of the best gifts one can hope for.

My friends Denise, Pietro, Edoardo, Tjeerd, Victor, Alec, Giampaolo made these years in San Diego joyful.

Chapters 1 and 2, in part, are being prepared for submission or resubmission for publication. Davide Viviano is the sole author.

Chapter 3, in part, is prepared for resubmission for publication Viviano, Davide; Bradic, Jelena. The Chapter is coauthored with Jelena Bradic. The dissertation author is the primary investigator.

Chapter 4, in part, is prepared for submission for publication Viviano, Davide; Wüthrich Kaspar; Niehaus, Paul. The Chapter is coauthored with Paul Niehaus and Kaspar Wüthrich. The dissertation author is one of the primary investigators of this material.

VITA

2016	Bachelor of Science in Economics, Luiss University
2017	Master of Science in Data Science, Barcelona Graduate School of Economics
2022	Doctor of Philosophy in Economics, University of California at San Diego

FIELDS OF STUDY

Primary Field: Econometrics; Secondary Field: Development Economics

ABSTRACT OF THE DISSERTATION

Essays in Policy Design and Statistical Inference

by

Davide Viviano

Doctor of Philosophy in Economics

University of California San Diego, 2022

Professor Graham Elliott, Chair

This thesis is devoted to designing and analyzing statistical decision rules to improve public policy. The first three chapters study the estimation and inference of welfare-maximizing treatment rules in the context of experimental design, spillover effects, and algorithmic fairness. The last chapter focuses on statistical inference in experiments with multiple policy effects (either of different interventions or on different outcomes).

Specifically, Chapter 1 proposes an experimental design for estimation and inference of welfare-maximizing policies with spillover effects. As a first contribution, we introduce a single-wave experiment to estimate the marginal effect of a change in treatment probabilities, taking spillover effects into account. Using the marginal effect, we propose a practical test for

welfare-maximizing policy. As a second contribution, we design a multiple-wave experiment to estimate treatment rules and maximize welfare. We provide asymptotic and small-sample guarantees of the procedure and study its numerical properties in simulations calibrated to existing experiments.

Chapter 2 studies how to allocate treatments on a network, using information from an existing (quasi)-experiment. We introduce a method that maximizes the sample analog of social welfare when spillovers occur. We construct semi-parametric welfare estimators and cast the optimization problem into a mixed-integer linear program. We derive guarantees on the regret and illustrate the method's advantage when targeting information on social networks.

Chapter 3 studies the problem of allocating treatments when policymakers have preferences for non-discriminatory actions. We adopt the non-maleficence perspective of "first do no harm": we select the fairest allocation within the Pareto frontier. We derive off-the-shelf optimization procedures and regret bounds on the unfairness of the estimated policy function. We illustrate our method using an application from education economics.

Chapter 4 focuses on inference in the presence of multiple hypothesis testing. The use of multiple hypothesis testing adjustments varies widely in applied economic research, without consensus on when and how it should be done. We provide a game-theoretic foundation for this practice. We show that adjustments with many interventions must depend on the nature of scale economies in the research production function and on economic interactions between policy decisions. When research examines multiple outcomes, our framework motivates aggregating outcomes into sufficient statistics for policy-making.

Chapter 1

Policy Design in Experiments with Unknown Interference

Introduction

One of the goals of a government or NGO is to estimate the welfare-maximizing policy. Network interference is often a challenge: treating an individual may also generate spillovers and affect the design of the policy. For instance, approximately 40% of experimental papers published in the "top-five" economic journals in 2020 mention spillover effects as a possible threat when estimating the effect of the program.[1] Researchers have become increasingly interested in experimental designs for choosing the treatment rule (policy) which maximizes welfare.[2] But when it comes to experiments on networks, standard approaches are geared toward the estimation of treatment effects. Estimation of treatment effects, on its own, is not sufficient for welfare-maximization.[3] For example, when assigning cash transfers, these may have the largest direct effect when given to people living in remote areas but generate the smallest spillovers. This trade-off has significant policy implications when treating each individual is costly or infeasible.

This chapter studies experimental designs in the presence of network interference when the goal is welfare maximization. The main difficulty in these settings is that interactions can be

[1] This is based on the author's calculation. Top-five economic journals are *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, *Review of Economic Studies*.

[2] See Bubeck and Cesa-Bianchi (2012) and Kasy and Sautmann (2019) for a discussion.

[3] Examples of treatment effects are the direct effects of the treatment and the overall effect, i.e., the effect if we treat all individuals, compared to treating *none*. For welfare maximization, none of these estimands is sufficient.

challenging to measure, and collecting network information can be very costly as it may require enumerating all individuals and their connections in the population.[4] We, therefore, focus on a setting with limited information on the network. This is formalized by assuming that units are organized into a *finite* number of large clusters, such as schools, districts, or regions, and interact through an unobserved network (and in unknown ways) within each cluster. In the cash-transfer program, we may expect that treatments generate spillovers to those living in the same or nearby villages[5], but spillovers are negligible between individuals in different regions. We propose the first experimental design to analyze and estimate *welfare-maximizing* treatment rules in the presence of unobserved spillovers on networks.

We make two main contributions. As a first contribution, we introduce a design where researchers randomize treatments and collect outcomes once (*single-wave experiment*) to (i) test whether one or more treatment allocation rules, such as the one currently implemented by the policymaker, maximize welfare; (ii) estimate how we can improve welfare with a (small) change to allocation rules. The experiment is based on a simple idea. With a small number of clusters, we do not have enough information to precisely estimate the welfare-maximizing treatment rule. However, if we take *two* clusters and assign treatments in each cluster independently with slightly different (locally perturbated) probabilities, we can estimate the marginal effect of a change in the treatment assignment rule (which we will refer to as marginal *policy* effect, MPE). For instance, in the cash-transfer example, the MPE defines the marginal effect of treating more people in remote areas, taking spillover effects into account.[6] Using the MPE, we introduce a practical test for whether there exists a welfare-improving treatment allocation rule. As this chapter suggests, researchers should report estimates of the MPE and test for welfare-maximizing policies. The MPE indicates the *direction* for a welfare improvement, and the test provides evidence on whether it is worth conducting additional experiments to estimate a welfare-improving policy.

[4] See Breza et al. (2020) for a discussion on the cost associated with collecting network information.

[5] For instance, Egger et al. (2019) document spillovers from cash transfers between nearby villages.

[6] This is the derivative of welfare with respect to the policy's parameters, taking spillovers into account, different from what is known in observational studies as marginal treatment effect (Carneiro et al., 2010), which instead depends on the individual selection into treatment mechanism.

Specifically, the experiment *pairs* clusters and randomizes treatments independently within clusters, with local perturbations to treatment probabilities within each pair. The difference in treatment probabilities balances the bias and variance of a difference-in-differences estimator. We show that the estimator for each pair converges to the marginal effect and derive properties for inference with finitely many clusters. The experiment separately estimates the direct and spillover effects, which are of independent interest. These are the effect on the recipients and the marginal effect of increasing the neighbors' treatment probability.[7]

As a second contribution, we offer an adaptive (i.e., *multiple-wave*) experiment to estimate welfare-maximizing allocation rules. The goal here is to adaptively randomize treatments to estimate the welfare-maximizing policy while also improving participants' welfare.[8] We propose an experiment that guarantees tight small-sample upper bounds for *both* the (i) out-of-sample regret, i.e., the difference between the maximum attainable welfare and the welfare evaluated at the estimated policy deployed on a new population, and the (ii) in-sample regret, i.e., the regret of the experimental participants. The experiment groups clusters into pairs, using as many pairs as the number of iterations (or more); every iteration, it randomizes treatments in a cluster and perturbs the treatment probability within each pair; finally, it updates policies sequentially, using the information on the marginal effects from a different pair via gradient descent. We illustrate the existence of a bias in adaptive experiments with repeated sampling and develop a novel algorithm with circular updates to avoid the bias.

From a theoretical perspective, a corollary of the sequential experiment's small-sample guarantees is that the out-of-sample regret converges at a faster-than-parametric rate in the number of clusters and iterations, and similarly, the in-sample regret.[9] We note that there are no regret guarantees tailored to unobserved interference. Existing results for treatment choice with

[7] Also, between-clusters local perturbations accommodate settings where policymakers cannot allow much variation in how treatments are assigned between different clusters because of exogenous constraints.

[8] Improving participants' welfare is desirable for large-scale experiments, common on online platforms (Karrer et al., 2021), and of increasing interest in development studies (Muralidharan and Niehaus, 2017).

[9] The average out-of-sample regret converges at a rate $1/T$, where T is the number of iterations and proportional to the number of clusters, and $\log(T)/T$ the in-sample regret. For the out-of-sample regret, we derive an exponential rate $\exp(-c_0 T)$, for a positive constant c_0 under additional restrictions (see Sec 1.4.1).

$i.i.d.$ data, treating clusters as sampled observations, would instead imply a slower convergence in the number of clusters.[10] We achieve a faster rate by (a) exploiting *within*-cluster variation in assignments and *between* clusters' local perturbations; (b) deriving concentration within each cluster as the cluster's size increases; (c) assuming and leveraging decreasing marginal effects of increasing neighbors' treatment probability.

We illustrate the numerical properties of the method with calibrated experiments that use data from an information diffusion experiment (Cai et al., 2015) and a cash-transfer program (Alatas et al., 2012, 2016). We show that our test can, in expectation, lead to welfare improvements up to fifty percentage points if, upon rejection of the null hypothesis that increasing treatment probabilities does not improve welfare, policy-makers increase treatment probabilities by five percent. When designing an adaptive experiment, the proposed method substantially improves both out-of-sample and in-sample regret, even with few iterations.

Throughout the text, we assume that the maximum degree grows at an appropriate slower rate than the cluster size; covariates and potential outcomes are identically distributed between clusters; treatment effects do not carry over in time. We discuss how such assumptions can be relaxed in the Appendix.

This chapter adds to both the literature on single-wave and multiple-wave experiments. In the context of *single-wave* (or two-wave) experiments, existing network experiments include clustered experiments (Eckles et al., 2017; Ugander et al., 2013; Karrer et al., 2021) and saturation designs (Baird et al., 2018; Pouget-Abadie, 2018). References with observed networks include Basse and Airoldi (2018b), Jagadeesan et al. (2020), Viviano (2020) among others.[11] See also

[10] For treatment choice, Kitagawa and Tetenov (2018) establish distribution-free lower bounds of order $1/\sqrt{n}$. In the literature on bandit feedback Shamir (2013) provides lower bounds of order $1/\sqrt{n}$ for continuous stochastic optimization procedure with strongly-convex functions (see also Bubeck et al. (2011)); optimization connects to bandits of Flaxman et al. (2004); Agarwal et al. (2010) which, however, provide slower rates for high-probability bounds (see also Section 1.4.1). Wager and Xu (2021) provide rates of order $1/T$ for in-sample regret but leverage an explicit model for market interactions with infinite asymptotics. Here, I do not impose assumptions on the interference mechanism and consider finitely many clusters. Kasy and Sautmann (2019) provides bounds for either (but not both) notions of regret in finite sample.

[11] For the analysis on the bias of average treatment effect estimators with interference see also Basse and Feller (2018), Johari et al. (2020), Basse and Airoldi (2018a), and Imai et al. (2009) when matching with different-sized clusters for overall average treatment effects.

Bai (2019); Tabord-Meehan (2018) with $i.i.d.$ data. These authors study experimental designs for inference on treatment effects only, but not the problem of inference on welfare-maximizing policies. Different from all the above references, we propose a design that allows us to identify the marginal effect under interference (and treatment effects). We introduce the first test and design for inference on policy optimality under unobserved interference. This difference motivates us to introduce a novel design consisting of local perturbations between clusters. The idea of using the information on marginal effects connects to the literature on optimal taxation (Saez, 2001; Chetty, 2009), which considers observational studies with independent units.

With *multiple-wave* experiments, we introduce a framework for adaptive experimentation with unknown interference. We connect to the literature on adaptive exploration (Bubeck and Cesa-Bianchi, 2012; Russo et al., 2017; Kasy and Sautmann, 2019, among others), and the one on derivative free stochastic optimization (Flaxman et al., 2004; Kleinberg, 2005; Shamir, 2013; Agarwal et al., 2010, among others). These references do not study the problem of network interference. Here, we leverage between-clusters perturbations and within-cluster concentration to obtain high-probability bounds on the regret with fast rates. In related work, Wager and Xu (2021) have studied prices estimation via local experimentation in the different contexts of a single market with asymptotically independent agents. They assume infinitely many individuals and an explicit model for market prices. As noted by the authors, the structural assumptions imposed in the above reference do not allow for spillovers on a network (i.e., individuals may depend arbitrarily on neighbors' assignments). Our setting differs due to network spillovers and the fact that individuals are organized into finitely many independent components (clusters), where such spillovers are unobserved. These differences motivate (i) the proposed design mechanism, which exploits two-level randomization at the cluster and individual level instead of individual-level randomization, and (ii) pairing and perturbations between clusters. From a theoretical perspective, network dependence and repeated sampling induce novel challenges for an adaptive experiment which we address here.

We relate to the literature on inference on treatment effects under interference and

draw from Hudgens and Halloran (2008) for definitions of potential outcomes. However, this literature considers existing (quasi)experiments and does not study experimental design and welfare-maximization. Aronow and Samii (2017); Manski (2013); Leung (2020); Ogburn et al. (2017); Goldsmith-Pinkham and Imbens (2013); Li and Wager (2020) assume an observed network, while Vazquez-Bare (2017), Hudgens and Halloran (2008), Ibragimov and Müller (2010) consider clusters among others. Sävje et al. (2021) study inference of the direct effect of treatment only. Our focus on policy optimality and experimental design differs from the above references. We show that inference on policy-optimality requires information on the MPE, which, we demonstrate, can be estimated with a clusters pair.

More broadly, we connect to the literature on statistical treatment rules, on estimation Manski (2004); Kitagawa and Tetenov (2018); Athey and Wager (2021); Bhattacharya and Dupas (2012); Stoye (2009); Mbakop and Tabord-Meehan (2021); Kitagawa and Wang (2021); Sasaki and Ura (2020); Viviano (2019), among others, and inference Andrews et al. (2019); Rai (2018); Armstrong and Shen (2015); Kasy (2016);[12] see also the literature on classification (Elliott and Lieli, 2013). This literature considers an existing experiment and does not study experimental design. Here, instead, we leverage an adaptive procedure to maximize out-of-sample and participants' welfare. Also, this literature has not studied policy design with unobserved interference. We broadly relate also to the literature on centrality measures (see for a review Bloch et al., 2019), which mainly focuses on model-based approaches with a single observed network – different from here where we leverage clusters' variations; the one on peer groups composition (Graham et al., 2010), and the one on inference with externalities (e.g., Bhattacharya et al., 2013, which, however, does not study inference on policy optimality). These references also do not study experimental designs.

[12] See also Kato and Kaneko (2020); Hadad et al. (2019); Imai and Li (2019); Hirano and Porter (2020a), which do not allow for testing for policy optimality, but construct confidence bands for the welfare.

Setup and Method's Overview

This section introduces conditions, estimands, and a brief overview of the method. We consider a setting with K clusters, where K is an even number. We assume that each cluster has N individuals, while our framework directly extends to clusters of different but proportional sizes. Observables and unobservables are jointly independent between clusters but not necessarily within clusters. Each cluster k is associated with a vector of outcomes, covariates, treatment assignments, and an adjacency matrix which is different for each cluster. These are $Y_{i,t}^{(k)} \in \mathscr{Y}, D_{i,t}^{(k)} \in \{0,1\}, X_i^{(k)} \in \mathscr{X} \subseteq \mathbb{R}^p, A^{(k)} \in \mathscr{A}$, respectively. Here, $(Y_{i,t}^{(k)}, D_{i,t}^{(k)})$ denote the outcome and treatment assignment of individual i at time t in cluster k, respectively, $X_i^{(k)}$ are time-invariant (baseline) covariates, and $A^{(k)}$ is a cluster-specific adjacency matrix. Each period t, researchers only observe a random subsample

$$\left(Y_{i,t}^{(k)}, X_i^{(k)}, D_{i,t}^{(k)}\right)_{i=1}^{n}, \quad n = \lambda N, \quad \lambda \in (0,1],$$

where n defines the sample size of observations from each cluster and is proportional to the cluster size.[13] There are T periods in total. While units sampled each period may or may not be the same, with abuse of notation, we index sampled units $i \in \{1, \cdots, n\}$. Whenever we provide asymptotic analyses, we let N grow, and K be fixed.

Setup: Covariates, Network and Potential Outcomes

Next, we introduce conditions on the covariates, network, and potential outcomes. These conditions guarantee that Lemma 1.2.1 (in Section 1.2.2) holds; practitioners may skip this subsection and directly refer to Section 1.2.2 for their implications, keeping in mind the covariates' distribution in Equation (1.1). We now discuss the network and covariates. We assume that individuals can form a link with a subset of individuals in each cluster. Formally, in each cluster,

[13] Our framework directly extends when $n = g(N)$ for a generic monotonic function $g(\cdot)$, and when we sample a different but proportional number of individuals from each cluster.

nodes are spaced under some latent space (Lubold et al., 2020) and can interact with at most the $\gamma_N^{1/2}$ closest nodes under the latent space. We say $1\{i_k \leftrightarrow j_k\} = 1$ if individual i can interact with j in cluster k and zero otherwise. Conditional on the indicators $1\{i_k \leftrightarrow j_k\}$,

$$(X_i^{(k)}, U_i^{(k)}) \sim_{i.i.d.} F_{U|X}F_X, \quad A_{i,j}^{(k)} = l\Big(X_i^{(k)}, X_j^{(k)}, U_i^{(k)}, U_j^{(k)}\Big)1\{i_k \leftrightarrow j_k\} \tag{1.1}$$

for an arbitrary and unknown function $l(\cdot)$ and unobservables $U_i^{(k)}$. Equation (1.1) states that whether two individuals interact depends on: (i) whether they are close enough within a certain latent space (captured by $1\{i_k \leftrightarrow j_k\}$); (ii) their covariates and unobserved individual heterogeneity (i.e., X_i, U_i) which capture homophily. Equation (1.1) also states that covariates are $i.i.d.$ unconditionally on $A^{(k)}$, but not necessarily conditionally.[14] Figure 2.2 provides an illustration. Here, we condition on the indicators $1\{i_k \leftrightarrow j_k\}$ (which can differ across clusters) to control the network's maximum degree but not on the network $A^{(k)}$. Equivalently, we can interpret such indicators as exogenously drawn from some arbitrary distribution.[15]

Assumption 1.2.1 (Network)**.** For $i \in \{1, \cdots, N\}, k \in \{1, \cdots, K\}$, let (i) Equation (1.1) hold given the indicators $1\{i_k \leftrightarrow j_k\}$, for some unknown $l(\cdot)$; (ii) $\sum_{j=1}^{n} 1\{i_k \leftrightarrow j_k\} = \gamma_N^{1/2}$.

Assumption 1.2.1 states the following: before being born, each individual may interact with $\gamma_N^{1/2}$ many other individuals. After the birth, the individual's gender, income, and parental status determine her type and the distribution of her and her potential connections' edges.[16] Here $\gamma_N^{1/2}$ captures the degree of dependence. Whenever $\gamma_N^{1/2}$ equals N, we impose no restriction on the individual's number of connections. In Theorem 1.3.1 the maximum degree can be equal to the cluster's size, while subsequent results require more restrictive restrictions.[17]

[14] We might also augment $l(\cdot)$ with additional $i.i.d.$ exogenous η_{ij} without affecting our results.

[15] Formally, $\mathscr{I}_k \sim \mathscr{P}_k, \quad (X_i^{(k)}, U_i^{(k)})|\mathscr{I}_k \sim_{i.i.d.} F_{U|X}F_X, \quad A_{i,j}^{(k)}|\mathscr{I}_k = l\Big(X_i^{(k)}, X_j^{(k)}, U_i^{(k)}, U_j^{(k)}\Big)1\{i_k \leftrightarrow j_k\}$, where $\mathscr{I}_k$ is the matrix of such indicators in cluster k and $\mathscr{P}_k$ is a cluster-specific distribution left unspecified.

[16] Networks formed from pairwise interactions have also been discussed in Jackson and Wolinsky (1996), Li and Wager (2020), Leung (2019). Here, we impose such restrictions to obtain easy-to-interpret conditions on the degree. Assumption 1.2.1 is sufficient but not necessary.

[17] Assumption 1.2.1 would not be required if we were to observe neighbors' assignments as in Viviano (2019).

We now discuss potential outcomes. Under interference, outcomes depend on individuals' covariates, unobservables and neighbors' treatments. That is, $Y_{i,t}^{(k)}(\mathbf{d}_1^{(k)},\cdots,\mathbf{d}_t^{(k)})$ denotes the potential outcome of individual i at time t. Here $\mathbf{d}_s^{(k)} \in \{0,1\}^N$ denotes the treatment assignments at time s of all individuals in cluster k. The following condition is imposed.

Assumption 1.2.2 (Potential outcomes)**.** Suppose that for any $i,t,k,\mathbf{d}_s^{(k)} \in \{0,1\}^N, s \leq t$

$$Y_{i,t}^{(k)}(\mathbf{d}_1^{(k)},\cdots,\mathbf{d}_t^{(k)}) = r\Big(\mathbf{d}_{i,t}^{(k)},\mathbf{d}_{j:A_{i,j}^{(k)}>0,t}^{(k)},X_i^{(k)},X_{j:A_{i,j}^{(k)}>0}^{(k)},U_i,U_{j:A_{i,j}^{(k)}>0},A_{i,\cdot}^{(k)},\nu_{i,t}^{(k)}\Big)+\tau_k+\alpha_t$$

for some unknown function $r(\cdot)$, stationary (but possibly serially dependent) unobservables $\nu_{i,\cdot}^{(k)}|X^{(k)},U^{(k)} \sim_{i.i.d.} P_\nu$, fixed effects τ_k,α_t.

Assumption 1.2.2 imposes three conditions. First, treatment effects do not carry-over in time. Second, potential outcomes are stationary up to separable fixed effects (unobservables can depend over time).[18] Third, potential outcomes can depend on neighbors' assignments, neighbors' covariates, and neighbors' identities, with arbitrary and unknown heterogeneity in spillover effects. Note that $A_{i,\cdot}^{(k)}$ denotes the set of connections of individual i in cluster k, with $\{j : A_{i,j}^{(k)} > 0\}$ denoting those individuals with some connection to i. Assumption 1.2.2 imposes no condition on the dependence on neighbors' assignments but assumes that cluster fixed effects do not depend on treatment assignments, which is important for identification.

Remark 1 (Extensions)**.** In Appendix A.1 and Appendix A.2 we present extensions. In Appendix A.1.1 individuals depend on general equilibrium effects, and in Appendix A.1.3 on the past assignments. Appendix A.2.2 presents non-separable fixed effects. Appendix A.2.4 studies staggered adoption. Appendix A.1.2, A.1.5 present effects heterogeneous on clusters' types. □

[18] Such condition is often implicitly imposed in studies on experimental design (Kasy and Sautmann, 2019). For a discussion on the no-carry-over assumption, see Athey and Imbens (2018). We relax it in Section A.1.3.

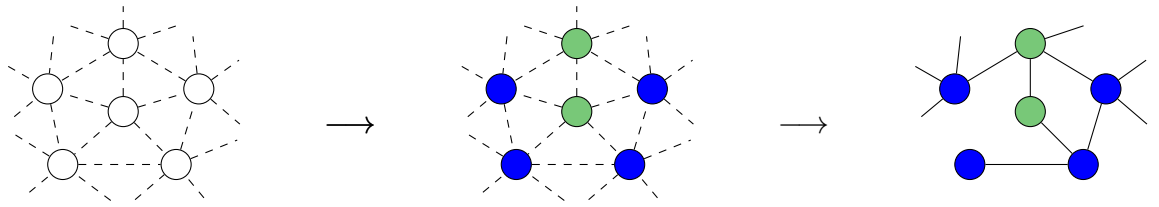


Figure 1.1. Example of the network formation model, with $\gamma_N = 5$. Individuals' are assigned different types which may or may not be observed by the researcher (corresponding to different colors). Individuals interact based on their types and form links among the possible connections. The possible connections and the realized adjacency matrix remain unobserved.

Policy and Welfare Maximization

The goal of this paper is to estimate a policy (treatment assignment rule) that maximizes welfare. We focus on a parametric class of policies, indexed by some parameter β. Formally, a policy

$$\pi(\cdot;\beta) : \mathscr{X} \mapsto [0,1], \quad \beta \in \mathscr{B},$$

is a map that prescribes the individual treatment probability based on covariates. Here $\mathscr{B}$ is a compact parameter space, and $\pi(x,\beta)$ is twice differentiable in β. The experiment assigns treatments independently based on $\pi(\cdot)$, and time/cluster-specific parameters $\beta_{k,t}$.

Assumption 1.2.3 (Treatment assignments in the experiment)**.** For given parameters $\beta_{k,t}$

$$D_{i,t}^{(k)} | X^{(k)}, \beta_{k,t} \sim_{i.i.d.} \text{Bern}\Big(\pi(X_i^{(k)};\beta_{k,t})\Big),$$

which, for short of notation, we refer to as $D_{i,t}^{(k)} | X^{(k)}, \beta_{k,t} \sim \pi(X_i^{(k)}, \beta_{k,t})$.

Assumption 1.2.3 defines a treatment rule in experiments. Treatments are assigned independently based on covariates and time and cluster-specific parameters $\beta_{k,t}$ (whose choice is discussed in the next sections). The assignment in Assumption 1.2.3 is easy to implement: it can be implemented in an online fashion and does not require network information, which justifies its choice; also, it generalizes assignments in saturation designs studied for inference

on treatment effects (Baird et al., 2018). Our goal is to estimate the welfare-maximizing β (see Remark 2).[19] Our framework extends to continuous treatments, omitted for brevity.[20]

An example of assignment rule is treating individuals with equal probability (Akbarpour et al., 2018), i.e., $\pi(\cdot;\beta) = \beta \in [0,1]$. We can also *target* treatments, i.e., $\pi(x;\beta) = \beta_x$, indicating the treatment probability for $X_i^{(k)} = x$ (with $\mathscr{X}$ discrete).

Throughout our discussion, whenever we write $\pi(\cdot;\beta)$, omitting the subscripts (k,t), we refer to a generic exogenous (i.e., not data dependent) vector of parameters β. We define $\mathbb{E}_\beta[\cdot]$ the expectation taken over the distribution of treatments assigned according to $\pi(\cdot;\beta)$.

Lemma 1.2.1 (Outcomes)**.** *Under Assumption 1.2.1, 1.2.2, under an assignment in Assumption 1.2.3 with exogenous (i.e., not data-dependent) $\beta_{k,t}$ the following holds:*

$$Y_{i,t}^{(k)} = y\left(X_i^{(k)}, \beta_{k,t}\right) + \varepsilon_{i,t}^{(k)} + \alpha_t + \tau_k, \quad \mathbb{E}_{\beta_{k,t}}\left[\varepsilon_{i,t}^{(k)} | X_i^{(k)}\right] = 0, \tag{1.2}$$

for some function $y(\cdot)$ unknown to the researcher. In addition, for some unknown $m(\cdot)$, $\mathbb{E}_{\beta_{k,t}}\left[Y_{i,t}^{(k)} | D_{i,t}^{(k)} = d, X_i^{(k)} = x\right] = m(d,x,\beta_{k,t}) + \alpha_t + \tau_k$.

The proof is in Appendix A.3.2. Equation (1.2) states that the outcome depends on two components. The first is the conditional expectation given the individual covariates and the parameter $\beta_{k,t}$, *unconditional* on covariates, adjacency matrix, individual, and neighbors' assignments. We can interpret the functions $y(\cdot)$ and $m(\cdot)$ as functions which depend on observables only. The dependence with $\beta_{k,t}$ captures spillover effects since treatments' distribution depends on $\beta_{k,t}$, while we average over neighbors' treatments and covariates. The second component $\varepsilon_{i,t}$ are unobservables that also depend on the neighbors' assignments and covariates. As shown in Appendix A.3.2, under the above conditions, such unobservables only depend on γ_N many others, where $\gamma_N^{1/2}$ is the maximum degree of the network (see Assumption 1.2.1). Also, note

[19] In Theorem 1.4.5 we show that the optimum obtained under Assumption 1.2.3 is asymptotically equivalent to the one with arbitrary dependent assignments under additional conditions on spillovers and costs.

[20] All our results hold for $D_{i,t}^{(k)} | X_i^{(k)}, \beta_{k,t} = \pi(X_i^{(k)}, \beta_{k,t})$, where $\pi(\cdot;\beta)$ is smooth in β.

that Lemma 1.2.1 assumes that $\beta_{k,t}$ is exogenous. We guarantee exogeneity with a careful choice of the experimental design discussed in subsequent sections.

Example 1.2.1. Let $D_{i,t}^{(k)} \sim_{i.i.d.} \text{Bern}(\beta)$ be exogenous, $\mathscr{N}_i = \{j : A_{i,j} = 1\}, A_{i,j} \in \{0,1\}$, and

$$Y_{i,t} = \alpha_t + D_{i,t}\phi_1 + \frac{\sum_{j\in\mathscr{N}_i} D_{j,t}^{(k)}}{|\mathscr{N}_i|}\phi_2 - \Big(\frac{\sum_{j\in\mathscr{N}_i} D_{j,t}}{|\mathscr{N}_i|}\Big)^2\phi_3 + \nu_{i,t}, \quad \mathbb{E}[\nu_{i,t}] = 0. \tag{1.3}$$

Equation (1.3) states that outcomes depend on the individual treatment, and the percentage of treated neighbors. With some algebra, taking expectations, $Y_{i,t} = \alpha_t + \beta\phi_1 + \beta\phi_2 - \beta\phi_3\kappa - \beta^2\phi_3(1-\kappa) + \varepsilon_{i,t}$ where $\varepsilon_{i,t}$ also depends on neighbors' assignments and $\kappa = \mathbb{E}[1/|\mathscr{N}_i|]$. □

Lemma 1.2.1 provides the basis for the definition of (utilitarian) welfare.

Definition 1.2.1 (Welfare)**.** For treatments as assigned in Assumption 1.2.3 with exogenous β parameter, let welfare be

$$W(\beta) = \int y(x,\beta)dF_X(x). \tag{1.4}$$

We define welfare as the expected outcome, had treatments been assigned with policy $\pi(\cdot,\beta)$. The expectation is taken over treatment assignments, covariates, and adjacency matrix. We interpret $y(x,\beta)$ the outcome *net of costs*[21], and incorporate the costs in the outcome function.[22] We define the welfare-maximizing policy and the marginal effect[23] as

$$\beta^* \in \arg\sup_{\beta\in\mathscr{B}} W(\beta), \quad V(\beta) = \frac{\partial W(\beta)}{\partial\beta}. \tag{1.5}$$

The marginal effect defines the derivative of the welfare with respect to the vector of

[21] This is standard in the literature (Kitagawa and Tetenov, 2018). However, some applications may not have explicit definitions of costs. For these cases, one possible choice of the cost is the opportunity cost, had the treatment being assigned to a population with no externalities. See Section 1.4.2 for a discussion.

[22] Namely, we can parametrize $Y_{i,t} = Y_{i,t} - c(X_{i,t};\beta)$, for a cost function $c(\cdot;\beta)$.

[23] Here, we are assuming differentiability. See Assumption 1.3.1 for explicit conditions.

parameters β. A useful insight is that, under mild regularity conditions, we can write

$$V(\beta) = \int \Big[\underbrace{\pi(x;\beta)\frac{\partial m(1,x,\beta)}{\partial \beta} + (1-\pi(x;\beta))\frac{\partial m(0,x,\beta)}{\partial \beta}}_{(S)} + \underbrace{\frac{\partial \pi(x;\beta)}{\partial \beta}\Delta(x,\beta)}_{(D)} \Big] dF_X(x), \tag{1.6}$$

where $\Delta(x,\beta) = m(1,x,\beta) - m(0,x,\beta)$. The marginal effect depends on the direct effect, weighted by the probability of treatment (D); the marginal spillover effect, i.e., the marginal effect of increasing neighbors' treatment probabilities. Equation (1.6) follows in the spirit of the direct and indirect effects decomposition in Hudgens and Halloran (2008).[24]

Remark 2 (Assumption 1.2.3 and unconstrained optimum)**.** In Theorem 1.4.5, we compare $W(\beta^*)$ to the unconstrained optimum W_N^*, where treatments can be assigned arbitrary. We show, under additional conditions, that $W(\beta^*) - W_N^* \to 0$, as $N \to \infty$, whenever the treatment costs are the opportunity costs of an intervention with no spillovers. In a cash-transfer program these are the opportunity costs had the cash transfers be given to individuals spread out on a large area instead of individuals in nearby villages. □

Method's Overview and Example

We now give an overview and example. Consider a policy-maker who must allocate cash-transfers to *half* of the population. Let $X_i \in \{0,1\}$, equal to one for households farer from the district's center than the median household and zero otherwise, with $P(X_i = 1) = 1/2$. Due to the constraint, the policy-maker assigns treatments $D_{i,t}|X_i = x \sim \text{Bern}(\pi(x,\beta))$, where $\pi(x,\beta) = x\beta + (1-x)(1-\beta)$ is the treatment probability for $x \in \{0,1\}$.[25] Different treatment probabilities for people in remote areas produce different welfare effects, and assigning all treatments to individuals in remote areas is sub-optimal. This is illustrated in Figure 1.2, where we report $W(\beta)$, calibrated to data from Alatas et al. (2012, 2016).[26]

[24] We also note that in recent work, Hu et al. (2021) have proposed targeting as causal estimand the average indirect effect, which is different from (S) for heterogenous assignments. Also, Graham et al. (2010) present peer effects' decompositions in the different contexts of peer groups' formation.

[25] This follows from the budget constraint where $\beta_1 P(X_i = 1) + \beta_2 P(X_i = 0) = 1/2$, where $\beta_1 = \beta, \beta_2 = 1-\beta$ are the treatment probabilities for people in remote and closer areas from the center, respectively.

[26] See Appendix A.4.3 for details.

Single-wave experiment: hypothesis testing

First, we would like to test whether a certain (baseline) intervention β, such as the one currently implemented by the government or NGO, is welfare-optimal. That is, we test $H_0 : W(\beta) = W(\beta^*)$. Its rejection is informative on whether a (small) change to β improves welfare. We use the marginal effect for:

(a) *Hypothesis testing*: assuming that β^* is an interior point, if $\frac{\partial W(\beta)}{\partial \beta} \neq 0$, then H_0 is false;

(b) *Policy update*: estimate the welfare-improving *direction* (increase or decrease β).

Here, (a) tests whether the line's slope in Figure 1.2 is zero. Similarly, a rejection of one-sided hypotheses $\frac{\partial W(\beta)}{\partial \beta} \leq 0$ suggests to increase β (treat more people in remote areas).

We proceed to construct estimators of the marginal effect. We start from Equation (1.6). The direct effect (D) can be identified from a single network, taking the difference between treated and untreated outcomes. However, the spillover effect (S) cannot be identified from a single network when unobserved. We instead exploit *two* clusters' variation.

We take two clusters such as two regions. We collect *baseline* ($t = 0$) outcomes and covariates; we then randomize treatments with slightly different probabilities between the regions. In the first region, we treat individuals in remote areas ($X_i = 1$) with probability $\beta + \eta_n$. Here, η_n is a small deterministic number (local perturbation). The remaining individuals are treated with probability $1 - \beta - \eta_n$. In the second region, we treat individuals in remote area with probability $\beta - \eta_n$, and the remaining ones with probability $1 - \beta + \eta_n$.

As shown in Figure 1.2, we can estimate welfare for two different but similar treatment probabilities; the line's slope between the points is approximately equal to the marginal effect. For a suitable choice of η_n (see Theorem 1.3.1), a consistent marginal effect's estimator is

$$\widehat{V}_{(k,k+1)}(\beta) = \frac{1}{2\eta_n}\left[\bar{Y}_1^{(k)} - \bar{Y}_0^{(k)}\right] - \frac{1}{2\eta_n}\left[\bar{Y}_1^{(k+1)} - \bar{Y}_0^{(k+1)}\right], \tag{1.7}$$

where $\bar{Y}_t^{(h)}$ is the outcomes' sample average in cluster h at time t, $Y_{i,0}$ is the baseline outcome

with no experiment in place yet, and $(k, k+1)$ index the two clusters. The above estimator is a difference-in-differences; we subtract baseline outcomes due to fixed effects.

We can then leverage the marginal effect for hypothesis testing. We take $K = 2G, G > 1$ clusters and 1) we match clusters in G pairs; 2) within each pair g, we obtain an estimate $\widehat{V}_g(\beta)$ for pair g; 3) we construct a (scale-invariant) test statistics using G pairs, which does not depend on the estimator's variance.[27] As we discuss in Remark 5, pairing clusters guarantees finite clusters asymptotics. Formal details of the algorithm are in Section 1.3.

Algorithm 1 illustrates the experiment in the cash-transfer example with two clusters. With more than two clusters, we pair clusters and implement Algorithm 1 in each pair.

Algorithm 1. Local perturbation with *two* clusters, β is a scalar

Require: Value β, $K = 2$ with $h \in \{k, k+1\}$, constant $\bar{C}$.

1: $t = 0$ (baseline): either nobody receives treatments or treatments are assigned with $\pi(\cdot;\beta)$ (either case is allowed).
   a: Experimenters collect baseline outcomes: for n units in each cluster observe $Y_{i,0}^{(h)}, X_i^{(h)}, h \in \{k, k+1\}$.
2: $t = 1$: experiment starts
   a: Based on the target parameter β, assign treatments for $X_i = 1$ as

$$D_{i,1}^{(k)} | \beta, X_i^{(k)} = x \sim \begin{cases} \text{Bern}(\pi(x, \beta + \eta_n)) \text{ if } h = k \\ \text{Bern}(\pi(x, \beta - \eta_n)) \text{ if } h = k+1 \end{cases}, \quad \bar{C} n^{-1/2} < \eta_n < \bar{C} n^{-1/4}.$$

   b: For n units in each cluster $h \in \{k, k+1\}$ observe $Y_{i,1}^{(h)}$.
3: Estimate the marginal effect as in Equation (1.7).

Remark 3 (Inference on treatment effects)**.** Our design permits us to estimate separately the direct treatment effect and the spillover effect, which may be of independent interest. We estimate the direct effect with a means' difference between treated and controls, *pooled* across clusters. In Theorem 1.3.3, we show that the estimated direct effect has a negligible bias of order $o(n^{-1/2})$. We estimate the spillover effect by taking the outcomes' difference between clusters (see Section 2.3).[28] □

[27] See Equation (1.10). Scale-invariance follows from Ibragimov and Müller (2010).

[28] Also, if the experimenter's goal is precise inference on direct effects only, β can be chosen based on variance

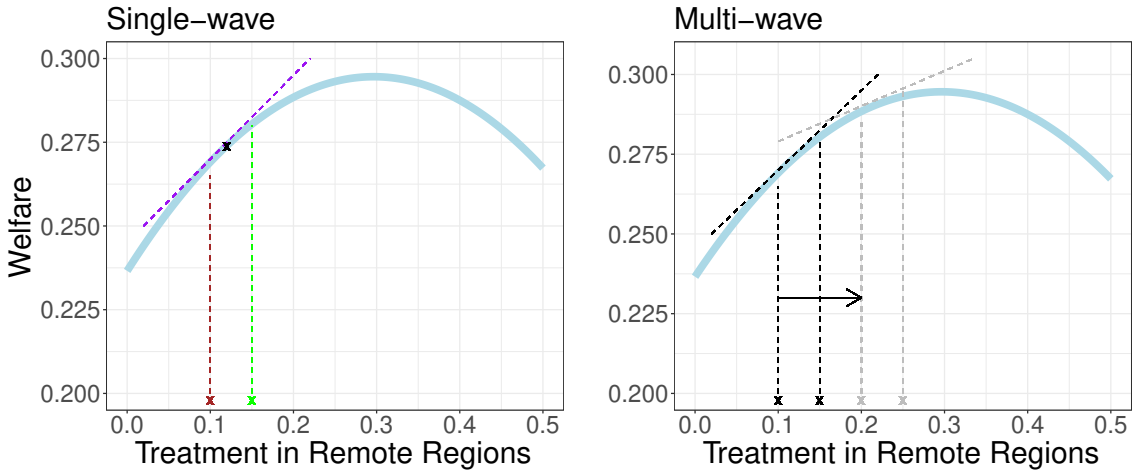


Figure 1.2. Example of experimental design. Half of the population is treated. The left panel is a single-wave experiment with two clusters. In the first cluster, we assign the policy colored in green and the second cluster colored in brown. The right panel is a two-wave experiment. In the first period, we use a pair of clusters to estimate the marginal effect (black color), and we update the policy for a different pair (gray color).

Multiple-wave experiment: estimation of β^*

We now show how we can estimate β^* using T experimentation periods and $K \geq 2(T+1)$ many clusters. The policy-maker assigns cash transfers and collects outcomes sequentially. Every period t, she assigns treatments in cluster k based on parameters $\beta_{k,t}$, with two goals. First, at time T, she wants to obtain an estimate $\hat{\beta}^*$, which well approximates β^*. Second, she wants to improve the experimental participants' welfare.

We maximize welfare with the following algorithm: 1) we pair clusters and organize pairs in a circle as in Figure 1.3; 2) every step t, we estimate the marginal effect within each pair; 3) we then update the policy in a given clusters' pair using the information on the marginal effect from the subsequent pair on the circle. We refer to Step 3 as circular cross-fitting.[29] Step 3 guarantees that the experiment is unconfounded. See Section 1.4 for details.[30]

considerations, but we can employ our method – which only requires small perturbations to β – to identify spillover effects.

[29]Cross-fitting algorithms were used by Chernozhukov et al. (2018) in the different context of double-machine learning with $i.i.d.$ data. Our procedure differs due to the adaptive sampling and the circular structure.

[30]Also, note that our method also extends when treatments can only be assigned once. See Appendix A.2.4.

The right panel in Figure 1.2 illustrates the procedure for two waves: we use one pair of clusters to estimate the marginal effect, which we then use for the second pair (and vice-versa, see Algorithm 3). The experiment assumes and leverages the concavity of welfare, generally attained under decreasing marginal effects of neighbors' treatments. Under lack of concavity, the experiment returns a *local* optimum. In Section 1.4 we measure the method's performance based on the in-sample and out-of-sample regret and show that with high-probability

$$W(\beta^*) - W(\hat{\beta}) = \mathscr{O}(1/T), \quad \max_k \frac{1}{T}\sum_{t=1}^{T}\Big[W(\beta^*) - W(\beta_{k,t})\Big] = \mathscr{O}(\log(T)/T).$$

The first equation indicates the out-of-sample, and the second the in-sample regret. The regret scales at rate $O(1/K)$, as we choose $2T+2=K$. Under additional restrictions, we also derive exponential rates $O(\exp(-c_0 K))$ for the out-of-sample regret, for some $c_0 > 0$.

Remark 4. An alternative approach is to estimate $y(\cdot)$ by assigning different policies to clusters and extrapolating the overall effect. We do not consider this alternative for two reasons. For a generic p-dimensional vector β, the out-of-sample regret is either sensitive to the model used for extrapolation or suffers a curse of dimensionality (e.g., when a grid search is employed). Second, this method does not control the *in-sample* regret, i.e., it must incur significant in-sample welfare loss to estimate the response function $y(\cdot)$. Appendix A.1.4 presents details. □

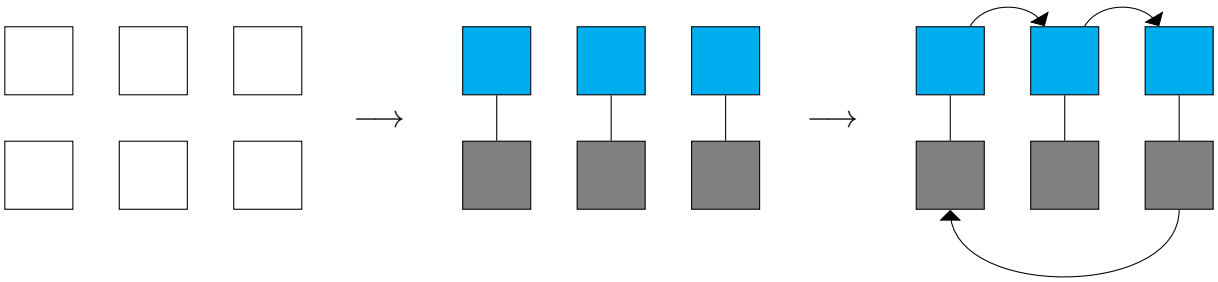


Figure 1.3. Circular cross-fitting method. Clusters (rectangles) are paired. Within each pair, researchers assign different treatment probabilities to clusters with different colors. Finally, the policy in each pair is updated using information from the consecutive pair.

Single-wave Experiment

In this section, we turn to the design and analysis of a single-wave experiment. We consider an experiment to test the following hypothesis.

Definition 1.3.1 (Testable implication)**.** Let $\beta^* \in \mathscr{B}$ be an interior point. $W(\beta) = W(\beta^*)$ implies

$$H_0 : V^{(j)}(\beta) = 0, \quad \forall j \in \{1, \cdots, l\}, l \leq p. \tag{1.8}$$

The above implication is at the core of our approach. We can test whether l arbitrary entries of the marginal effect are equal to zero. Rejection implies a lack of global optimality. For expositional convenience, we consider $l = 1$ only (test the first entry being zero). In Appendix A.2.1 we show how our method generalizes to $l > 1$. We may also test $V^{(j)}(\beta) \leq 0$; for example, for $\pi(x, \beta) = \beta_x$ ($\mathscr{X}$ discrete), one-sided test is informative for whether we should increase treatment probabilities for individuals with $x = j$ (without assuming that β^* is in the interior). Finally, it is useful to define the vector

$$\underline{e}_j = \left[0, \cdots, 0, 1, 0, \cdots, 0\right], \text{ where } \underline{e}_j \in \{0,1\}^p, \text{ and } \underline{e}_j^{(j)} = 1. \tag{1.9}$$

Algorithm 2 presents the design. The algorithm pairs clusters. Within each pair, it estimates the first entry of the marginal effect (since here we test $V^{(1)}(\beta) = 0$) using local perturbations – as discussed in Section 1.2.3. It then constructs a scale-invariant test statistics. Without loss of generality, we index clusters such that each pair contains two consecutive clusters $\{k, k+1\}$ with k being an odd number.

In the following lines, we discuss the estimation of marginal and treatment effects and guarantees for inference on H_0.

Algorithm 2. One wave experiment for inference with $l = 1$

Require: Value $\beta \in \mathbb{R}^p$ (exogenous), K clusters, constant $\bar{C}$, size α;

1: Organize clusters into $G = K/2$ pairs with consecutive indexes $\{k, k+1\}$;

2: For each pair $g = \{k, k+1\}$, k is odd, run Algorithm 1, with at $t = 1$,

$$D_{i,1}^{(k)} | \beta, X_i^{(k)} = x \sim \begin{cases} \text{Bern}(\pi(x, \beta + \eta_n \underline{e}_1)) \text{ if } h = k \\ \text{Bern}(\pi(x, \beta - \eta_n \underline{e}_1)) \text{ if } h = k+1 \end{cases}, \quad \bar{C} n^{-1/2} < \eta_n < \bar{C} n^{-1/4},$$

and estimate the marginal effect as in Equation (1.6).

3: Construct the test statistics

$$\mathscr{T}_n = \frac{\sqrt{G} \bar{V}_n(\beta)}{\sqrt{(G-1)^{-1} \sum_g (\widehat{V}_g(\beta) - \bar{V}_n(\beta))^2}}, \quad \bar{V}_n(\beta) = \frac{1}{G} \sum_g \widehat{V}_g(\beta); \tag{1.10}$$

here $\widehat{V}_g$ is the marginal effect estimated in pair g.

4: Construct the test as $1\Big\{ |\mathscr{T}_n| > \text{cv}_{G-1}(\alpha) \Big\}$ with size α; $\text{cv}_{G-1}(\alpha)$ is the size α t-test's critical quantile with $G-1$ degrees of freedom.

Estimation of Marginal and Treatment Effects

The experiment we just described permits us to estimate three quantities of independent interest: the marginal effect, the direct effect, and the spillover effect. These should be reported by researchers once the experiment is concluded. We describe the estimators below.

Equation (1.6) provides the marginal effect estimator $\widehat{V}_g(\beta)$ for each pair of clusters g. Researchers may report $\bar{V}_n(\beta)$ (Equation 1.10) in their results – the average across clusters' pairs. We show below that both $\bar{V}_n(\beta)$ and $\widehat{V}_g(\beta)$ provide a consistent estimate of $V^{(1)}(\beta)$.[31]

The experiment also allows us to estimate the direct effect of the treatment and the (marginal) spillover effect separately, respectively defined as:[32]

$$\Delta(\beta) = \int \Big[m(1, x, \beta) - m(0, x, \beta) \Big] dF_X(x), \quad S_1(d, \beta) = \int \frac{\partial m(d, x, \beta)}{\partial \beta^{(1)}} dF_X(x).$$

The direct effect is the treatment effect, keeping fixed the neighbors' treatment probability. $S_1(\cdot)$,

[31] Note that our discussion directly extends to estimating each entry of $V(\beta)$ as shown in Appendix A.2.1.

[32] Here, we are implicitly assuming that $m(\cdot)$ is uniformly bounded to invoke the dominated convergence theorem. See the next subsection for formal assumptions.

the spillover effect, is the marginal effect of a small change in the first entry of β, keeping fixed individual treatment status.[33] For a given pair of clusters $(k,k+1)$, we estimate

$$\widehat{\Delta}_k(\beta) = \frac{1}{2n} \sum_{h\in\{k,k+1\}} \sum_{i=1}^{n} \Big[\frac{D_{i,1}^{(h)} Y_{i,1}^{(h)}}{\pi(X_i^{(h)}, \beta + \eta_n v_h \underline{e}_1)} - \frac{(1-D_{i,1}^{(h)}) Y_{i,1}^{(h)}}{1-\pi(X_i^{(h)}, \beta + \eta_n v_h \underline{e}_1)} \Big],$$

where $v_h = 1\{h=k\} - 1\{h=k+1\}$.

The estimator pools observations between the two clusters and takes a difference between treated and control units within each cluster, divided by the probability of treatments. This follows similarly to classical Horvitz-Thompson estimators (Horvitz and Thompson, 1952). Importantly, we divide by the probability of treatments, taking into account the perturbation η_n. We average direct effects across clusters' pairs to obtain a single measure $\bar{\Delta}_n = \frac{1}{G}\sum_g \widehat{\Delta}_g(\beta)$. The indirect effect is estimated as follows:

$$\widehat{S}_{(k,k+1)}(0,\beta) = \frac{1}{2n} \sum_{h\in\{k,k+1\}} \frac{v_h}{\eta_n} \sum_{i=1}^{n} \Big[\frac{Y_{i,1}^{(h)}(1-D_{i,1}^{(h)})}{1-\pi(X_i^{(h)}, \beta + v_h \eta_n \underline{e}_1)} - \bar{Y}_0^{(h)} \Big].$$

The estimator takes a difference between the control units between the two clusters, divided by their corresponding treatment probabilities. Researchers may report the between pairs average $\bar{S}_n(0,\beta) = \frac{1}{G}\sum_g \widehat{S}_g(0,\beta)$,[34] which captures spillovers on the control units.

Theoretical Analysis

Next, we study theoretical guarantees. The following regularity condition is imposed.

Assumption 1.3.1 (Regularity 1)**.** Suppose that for all $x \in \mathscr{X}, d \in \{0,1\}$, $\pi(x,\beta), m(d,x,\beta)$ are uniformly bounded and twice differentiable with bounded derivatives.[35]

[33] Similarly to the marginal effect, our setting also extends to estimating $S_j(\cdot)$ for arbitrary entries of β as in Appendix A.2.1.

[34] Here, $\hat{S}(1,\beta)$ follows similarly and omitted for brevity.

[35] Assumption 1.3.1 imposes smoothness and boundedness restrictions. These restrictions hold for a large set of linear and non-linear functions, assuming that $\mathscr{X}$ is compact. Boundedness is often imposed in the literature, see, e.g., Kitagawa and Tetenov (2018).

Theorem 1.3.1 (Consistency of the marginal effect)**.** *Suppose that* $\varepsilon_{i,t}^{(k)}$ *is sub-gaussian. Let Assumptions 1.2.1, 1.2.2, 1.3.1 hold. Let* $\text{Var}(\sqrt{n}\hat{V}_{(k,k+1)}(\beta)) = \mathcal{O}(\rho_n)$*. Then with probability at least* $1-\delta$*, for any* $\delta \in (0,1)$*,*

$$\left|\hat{V}_{(k,k+1)}(\beta) - V^{(1)}(\beta)\right| = \mathcal{O}\Big(\eta_n + \min\Big\{\sqrt{\frac{\gamma_N \log(\gamma_N/\delta)}{n\eta_n^2}}, \sqrt{\frac{\rho_n}{n\eta_n^2\delta}}\Big\}\Big),$$

where $\hat{V}_{(k,k+1)}$ *is estimated as in Algorithm 2.*

For $\gamma_N \log(\gamma_N)/N^{1/3} = o(1), \eta_n = n^{-1/3}, \hat{V}_{(k,k+1)}(\beta) \to_p V(\beta)$.

The proof is in Appendix A.3.3. Theorem 1.3.1 shows that we can consistently estimate the marginal effects with two large clusters. Consistency depends on the degree of dependence among unobservables $\varepsilon_{i,t}^{(k)}$ (which also depends on neighbors' assignments). The convergence rate depends on the *minimum* between the maximum degree of the network, which is proportional to $\gamma_N^{1/2}$, and the covariances among unobservables, captured by ρ_n. If either the network has a degree that grows at a slower rate than N (recall that $n/N = O(1)$) or a degree equal to N but vanishing covariances, we can consistently estimate the marginal effects. The theorem also illustrates the trade-off in the choice of the deviation parameter η_n: a larger deviation parameter η_n decreases the variance, but it increases the bias. Appendix A.2.5 provides an explicit rule of thumb for choosing η_n.

Corollary 1. *Under the conditions in Theorem 1.3.1, letting* $\gamma_N \log(\gamma_N)/n^{1/3} = o(1), \eta_n = n^{-1/3}$*, for any K,* $\bar{V}_n \to_p V^{(1)}(\beta)$.

The above corollary illustrates consistency once we pool information from different clusters (with K being finite). Next, we study inference assuming the following condition.

Assumption 1.3.2 (Regularity 2)**.** Assume that for treatments as assigned in Algorithm 2, for all $k \in \{1, \cdots, K\}$, $\varepsilon_{i,t}^{(k)}$ has bounded fourth moment and for some $\bar{C}_k > 0$, $\rho_n \geq 1$,

$$\text{Var}\Big(\frac{1}{\sqrt{n}}\Big[\bar{Y}_1^{(k)} - \bar{Y}_0^{(k)}\Big]\Big) = \bar{C}_k \rho_n. \tag{1.11}$$

Assumption 1.3.2 imposes standard moment bounds and a lower bound on the variance of the estimator, attained under independence and positive dependence.

Theorem 1.3.2. *Let Assumptions 1.2.1, 1.2.2, 1.3.1, 1.3.2 hold. Let* $n^{1/4}\eta_n = o(1), \gamma_N/N^{1/4} = o(1)$*,* $K < \infty$*. Then for each pair* $(k,k+1)$*, for* $\widehat{V}_{(k,k+1)}$ *estimated as in Algorithm 2,*

$$\text{Var}\Big(\widehat{V}_{(k,k+1)}\Big)^{-1/2}\Big(\widehat{V}_{(k,k+1)} - V^{(1)}(\beta)\Big) \to_d \mathcal{N}(0,1).$$

The proof is in Appendix A.3.3. Theorem 1.3.2 guarantees asymptotic normality. The theorem assumes that the maximum degree $\gamma_N^{1/2}$ grows at a slower rate than the sample size of order $N^{1/8}$ (and hence $n^{1/8}$ since n is proportional to N). This condition is stronger than what is required for consistency only.[36] Given Theorem 1.3.2, we conduct inference with scale-invariant test statistics without necessitating estimation of the (unknown) variance.

Corollary 2. *Let the conditions in Theorem 1.3.2 hold. For* $4 \leq K < \infty$*,* $\alpha \leq 0.08$*,*

$$\lim_{n\to\infty} P\Big(|\mathscr{T}_n| \leq \text{cv}_{K/2-1}(\alpha)\Big|H_0\Big) \geq 1-\alpha, \tag{1.12}$$

where $\text{cv}_{K/2-1}(h)$ *is the size-h critical value of a t-test with* $K/2-1$ *degrees of freedom.*

The proof is in Appendix A.3.5. The theorem guarantees asymptotically valid inference on H_0 as $n \to \infty$ and K is finite. With $l = 1$ the proof is a direct consequence of Theorem 1.3.2, combined with properties of pivotal statistics in Ibragimov and Müller (2010).[37] In Appendix A.2.1 we provide expressions for the test statistics and derivations for $l > 1$.

To our knowledge, this is the first set of results for inference on welfare-maximizing policies with unknown interference.

We conclude this section with a study on the estimated direct and spillover effect.

[36] We conjecture that weaker restrictions on the degree are possible, as for consistency. We leave their study to future research.

[37] See also Chernozhukov et al. (2018) for a discussion on pivotal inference in the different context of synthetic controls.

Theorem 1.3.3 (Asymptotically neglegible bias of the direct effect)**.** *Let Assumptions 1.2.1, 1.2.2, 1.3.1 hold, and* $\eta_n = o(n^{-1/4})$*. Then for all pairs* $(k,k+1)$*,* $\mathbb{E}\left[\widehat{\Delta}_{(k,k+1)}(\beta)\right] = \Delta(\beta) + o(n^{-1/2})$*. Similarly,* $\mathbb{E}\left[\bar{\Delta}_n(\beta)\right] = \Delta(\beta) + o(n^{-1/2})$*, where the second term does not depend on K.*

The proof is in Appendix A.3.3. Theorem 1.3.3 shows that the bias of the estimated direct effect is asymptotically negligible at a rate faster than the parametric rate $n^{-1/2}$ when *pooling* observations from different clusters. Our insight here is that, with pairing and perturbations of opposite signs, the first-order bias cancels out. This result implies that our experimental design induces a bias that can be ignored for estimation and inference.[38] Given that the bias is asymptotically negligible, we can use existing results for inference on direct effects with a single network (e.g., Sävje et al. 2021). For completeness, we show consistency below.

Corollary 3. *Suppose that* $\varepsilon_{i,t}^{(k)}$ *is sub-gaussian. Let Assumptions 1.2.1, 1.2.2, 1.3.1 hold, and* $\pi(x,\beta) \in (\kappa, 1-\kappa), \kappa \in (0,1)$ *for all* $x \in \mathscr{X}$*. Let* $\eta_n = o(n^{-1/4})$*. Then with probability at least* $1-\delta$*, for any* $\delta \in (0,1)$*, for any K*

$$\left|\bar{\Delta}_n - \Delta(\beta)\right| = \mathscr{O}\Big(\sqrt{\frac{\gamma_N \log(\gamma_N/\delta)}{Kn}}\Big) + o(n^{-1/2}),$$

The proof is in Appendix A.3.5. The corollary requires strict overlap (standard in the literature on causal inference) and shows that we can attain consistency for $K < \infty$, $n \to \infty$.

The following result is on the bias of the marginal spillover effects estimators.

Theorem 1.3.4 (Marginal Spillover effects)**.** *Let Assumptions 1.2.1, 1.2.2, 1.3.1 hold. Then for all pairs* $(k,k+1)$*,* $\mathbb{E}\left[\widehat{S}_{(k,k+1)}(0,\beta)\right] = S_1(0,\beta) + \mathscr{O}(\eta_n)$*.*

The proof is in Appendix A.3.3. Theorem 1.3.4 shows that the bias converges to zero as $\eta_n \to 0$. The rate is slower than the rate of the direct effect's bias (since $\eta_n > n^{-1/2}$). We obtain a slower rate because the marginal spillover effect depends on *between* clusters variations. Consistency and inference follow verbatim as discussed for the marginal effect.

[38] Note that $\eta_n = o(n^{-1/4})$ is consistent with requirements in previous theorems.

Remark 5 (Pairing clusters permits finite-clusters asymptotics)**.** Pairing plays a fundamental role in our design with finite K. In the absence of pairing, the bias of the marginal and treatment effects would not converge to zero for finite K. To gain further intuition, consider a uni-dimensional setting ($X = 1$ almost surely), and $\pi(1,\beta) = \beta$. Then

$$\mathbb{E}\Big[\widehat{V}_{(k,k+1)}(\beta)\Big] = \frac{1}{2}\sum_{h\in\{k,k+1\}} \frac{v_h}{\eta_n} y(1;\beta + v_h\eta_n)$$

$$= \underbrace{\frac{1}{2}\sum_{h\in\{k,k+1\}} \frac{v_h}{\eta_n} y(1;\beta)}_{(i)} + \underbrace{\frac{\partial y(1;\beta)}{\partial \beta}}_{(ii)} + \mathscr{O}(\eta_n), \quad v_h = \begin{cases} 1 \text{ if } h = k \\ -1 \text{ if } h = k+1. \end{cases}$$

Component (i) induces a bias, while (ii) is the target estimand. Observe that (i) equals zero because of the paired design: v_h is one for one cluster and minus one for the other cluster. Suppose instead that a paired design was not implemented, and instead we have $v_h \in \{\pm 1\}$ with equal probability.[39] Then (i) would scale to zero at a slow rate $1/\sqrt{K\eta_n^2}$ after averaging across all the clusters, requiring infinite clusters asymptotics. □

Multi-wave Experiment

In this section, we design the adaptive experiment and derive its theoretical properties. For illustrative purposes, we provide the algorithm for the one-dimensional case $p = 1$, in Algorithm 3, that is when $\beta \in \mathscr{B} = [\mathscr{B}_1, \mathscr{B}_2]$ is a scalar. In Remark 6 and formally in Appendix A.5 we provide the complete algorithm for the p-dimensional case. Theoretical results are for the general p-dimensional case (p is finite). A description is below the algorithmic box.

The algorithm pairs clusters, and initializes clusters at the same starting value β_0, $\check{\beta}_1^1 =$

[39] Random probabilities assignments are common when estimating treatment effects with saturation design (Baird et al., 2018). Pairing is common in applications in the different context of estimating overall average treatment effects with (different-sized) cluster experiments (Imai et al., 2009).

Algorithm 3. Multiple-wave experiment with β scalar

Require: Starting value β_0, K clusters, $T+1$ periods, constant $\bar{C}$.

1: Create pairs of clusters $\{k,k+1\}, k \in \{1,3,\cdots,K-1\}$;

2: $t=0$ (initialization):

a: Assign treatments as $D_{i,0}^{(h)}|X_i^{(h)} = x \sim \text{Bern}(\pi(x,\beta_0))$ for all $h \in \{1,\cdots,K\}$.

b: For n units in each cluster observe $Y_{i,0}^{(h)}, h \in \{1,\cdots,K\}$; initalize $\widehat{V}_{k,t}=0$, $\check{\beta}_k^0 = \beta_0$.

3: **while** $1 \le t \le T$ **do**

a: Define

$$\check{\beta}_h^t = \begin{cases} P_{\mathscr{B}_1,\mathscr{B}_2-\eta_n}\left[\check{\beta}_h^{t-1} + \alpha_{h+2,t}\widehat{V}_{h+2,t-1}\right], & h \in \{1,\cdots,K-2\}, \\ P_{\mathscr{B}_1,\mathscr{B}_2-\eta_n}\left[\check{\beta}_h^{t-1} + \alpha_{1,t}\widehat{V}_{1,t-1}\right], & h \in \{K-1,K\}; \end{cases}$$

where $\alpha_{k,t}$ is the learning rate (see Remark 7), and $P_{\mathscr{B}_1,\mathscr{B}_2-\eta_n}$ is the projection operator.

b: Assign treatments as (for $\bar{C}n^{-1/2} < \eta_n < \bar{C}n^{-1/4}$)

$$D_{i,t}^{(h)}|X_i^{(h)} = x \sim \text{Bern}(\pi(x,\beta_{h,t})), \quad \beta_{h,t} = \begin{cases} \check{\beta}_h^t + \eta_n \text{ if } h \text{ is odd} \\ \check{\beta}_h^t - \eta_n \text{ if } h \text{ is even} \end{cases} \tag{1.13}$$

c: For n units in each cluster $h \in \{1,\cdots,K\}$ observe $Y_{i,t}^{(h)}$;

d: For each pair $\{k,k+1\}$, estimate

$$\hat{V}_{k,t} = \hat{V}_{k+1,t} = \frac{1}{2\eta_n}\left[\bar{Y}_t^{(k)} - \bar{Y}_0^{(k)}\right] - \frac{1}{2\eta_n}\left[\bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k+1)}\right].$$

4: **end while**

5: Return $\hat{\beta}^* = \frac{1}{K}\sum_{k=1}^K \check{\beta}_k^T$

$\cdots = \check{\beta}_K^1 = \beta_0$. At $t=0$, it randomizes treatments independently as

$$t=0: \quad D_{i,t}^{(k)}|X_i^{(k)} = x \sim \pi(x;\beta_0), \quad \text{for all } (i,k).$$

Here, β_0 is chosen exogenously, e.g., it is the current policy in place. Over each iteration t, we assign treatments based on $\beta_{k,t}$ for cluster k at time t which equals the parameter $\check{\beta}_k^t$ obtained from a previous iteration plus a positive (negative) perturbation η_n in the first (second) cluster in a pair. The local perturbation follows similarly to what discussed in the previous section; also, by construction, $\check{\beta}_k^t$ is the same for a given pair $(k,k+1)$, where k is odd. We choose $\check{\beta}_k^{t+1}$ via

circular cross-fitting: we wrap clusters in a circle and update the parameter in a pair of clusters $(k, k+1)$ using information from the subsequent pair (see Figure 1.3). The algorithm runs over T periods and returns $\hat{\beta}^* = \frac{1}{K}\sum_{k=1}^{K} \check{\beta}_k^{T+1}$.

In our experiment, we update the policy in each clusters' pair with information from a subsequent pair. This approach guarantees that the estimated policy used to randomize treatments in cluster k does not depend on observables and unobservables in that same cluster, assuming that the number of clusters is twice as large as the number of iterations.

Lemma 1.4.1 (Unconfoundedness)**.** *Let $T/p + 1 \leq K/2$. Consider the experimental design in Algorithm A.5.1 for generic p-dimensions (and Algorithm 3 for $p = 1$). Then for any k,*

$$\left(\beta_{k,1}, \cdots, \beta_{k,T}\right) \perp \left\{Y_{i,t}^{(k)}(\mathbf{d}), X_i^{(k)}, \mathbf{d} \in \{0,1\}^N\right\}_{i \in \{1,\cdots,N\}, t \leq T}.$$

The proof is in Appendix A.3.2. Lemma 1.4.1 shows that the parameters used in the experiment are independent of potential outcomes and covariates in the same cluster. Namely, the circular cross-fitting breaks the dependence due to repeated sampling, which would otherwise confound the experiment.[40] See Remark 8 for a discussion.

Remark 6 (p-dimensional case: Algorithm A.5.1)**.** The algorithm for the p-dimensional case follows similarly to the uni-dimensional case with a minor change: we consider T/p many *waves*/iterations, each consisting of p periods. Within each wave w, every period, we perturb a single coordinate of $\check{\beta}_k^w$, compute the marginal effect for that coordinate, and repeat over all coordinates $j \in \{1, \cdots p\}$ before making the next policy update to select $\check{\beta}_k^{w+1}$. □

Remark 7 (Learning rate)**.** We are now left to discuss how "large" should be the step size, i.e., if we know that the marginal effect is positive, by how much should we increase the treatment probability? Assuming strong concavity of the objective function, the learning rate $\alpha_{k,t}$ should

[40] This setting is different from previous literature on adaptive experimentation, which focuses on settings where repeated sampling does not occur. See, for example, Kasy and Sautmann (2019); Wager and Xu (2021).

be of order $1/t$. A more robust choice is

$$\alpha_{k,t} = \begin{cases} \frac{J}{T^{1/2-v/2}||\hat{V}_{k,t}||} \text{ if } ||\hat{V}_{k,t}||_2^2 > \frac{c}{\breve{T}^{1-v}} - \varepsilon_n, \\ 0 \text{ otherwise} \end{cases}, \tag{1.14}$$

for a positive ε_n, $\varepsilon_n \to 0$, and small constants $1 \geq v, c > 0$.[41] Here, the learning rate divides the estimated marginal effect by its norm (known as gradient norm rescaling, Hazan et al. 2015) and guarantees control of the out-of-sample regret under strict quasi-concavity.[42] This choice is appealing since it guarantees comparable step sizes between different clusters. □

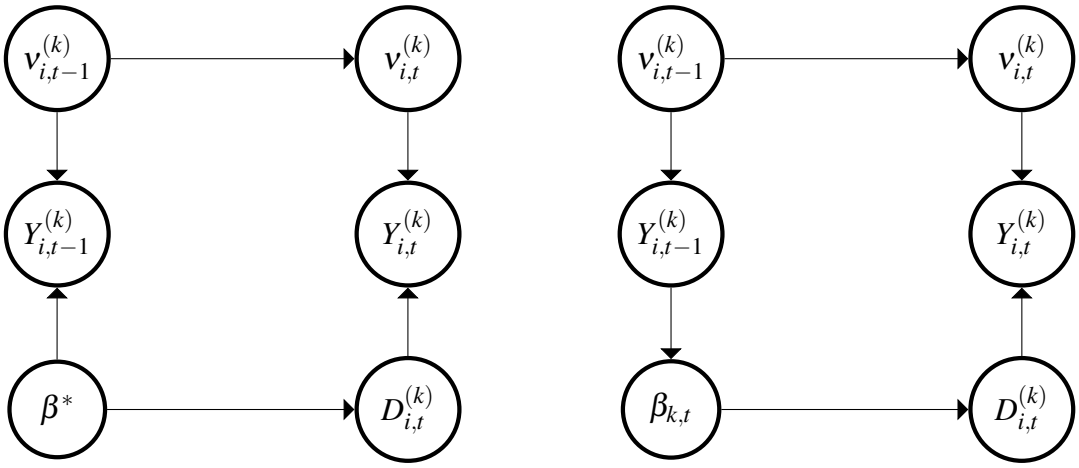
Remark 8 (Chapter 1: Why circular cross-fitting? Bias with repeated sampling)**.** Here, we illustrate the source of bias if the circular cross fitting was not employed. Every period, the researcher can only identify the expected outcome of $Y_{i,t}^{(k)}$ conditional on the parameter $\beta_{k,t}$, namely $\widetilde{W}(\beta_{k,t}) = \mathbb{E}_{\beta_{k,t}}[Y_{i,t}^{(k)}|\beta_{k,t}]$. If $\beta_{k,t}$ was chosen exogenously, based on information from a different cluster, then $\mathbb{E}_{\beta_{k,t}}[Y_{i,t}^{(k)}|\beta_{k,t}] = \mathbb{E}_{\beta_{k,t}}[Y_{i,t}^{(k)}] = W(\beta_{k,t})$, where $W(\beta_{k,t})$ defines the expected welfare once we deploy the policy $\beta_{k,t}$ on a new population. However, this is not the case if $\beta_{k,t}$ is estimated using information on $Y_{i,t-1}^{(k)}$. Consider the example where the outcome depends on some auto-correlated unobservables $\nu_{i,t}$ and treatment assignments in Figure 1.4. The *dependence* structure of Figure 1.4 implies that: $W(\beta_{k,t}) = \mathbb{E}_{\beta_{k,t}}[Y_{i,t}^{(k)}] \neq \mathbb{E}_{\beta_{k,t}}[Y_{i,t}^{(k)}|\beta_{k,t}] = \widetilde{W}(\beta_{k,t})$, if $\beta_{k,t}$ depends on covariates and unobservables previous outcomes (and so on unobservables $\nu_{i,t}^{(k)}$) in cluster k. Here $W(\beta_{k,t})$ captures the estimand of interest. Instead, $\widetilde{W}(\beta_{k,t})$ denotes what we can identify. Our algorithm breaks such dependence and guarantees unconfounded experimentation as shown in Lemma 1.4.1. □

Theoretical Guarantees

Next, we derive theoretical properties. Let $\breve{T} = T/p$. We assume the following.

[41] Formally, we let $\varepsilon_n \propto \sqrt{\frac{\gamma_N}{\eta_n^2 n}} + \eta_n$.

[42] More discussion is included in Appendix A.2.3.



Policy on a new population Experiment with repeated sampling

Figure 1.4. Why circular cross-fitting? The left-panel shows the dependence structure when a static policy is implemented on a new population (we omit $D_{i,t-1}^{(k)}$ for expositional convenience). The right panel shows the dependence structure of a sequential experiment that uses the same units for policy updates over subsequent periods in the presence of *repeated* sampling.

Assumption 1.4.1. Let : (A) $\varepsilon_{i,t}^{(k)}$ be sub-gaussian; (B) $K \geq 2(T/p+1)$.

Condition (A) states that unobservables have sub-gaussian tails (attained, for example, by bounded random variables); (B) assumes that we have at least twice as many clusters as the number of waves, which guarantees that Lemma 1.4.1 (unconfoundedness) holds.

In the following results, we impose the following restriction.

Assumption 1.4.2 (Strong concavity)**.** Assume that $W(\beta)$ is σ-strongly concave over $\mathscr{B}$, for some arbitrary $\sigma > 0$.

Strong concavity is a common feature of objective functions (Bottou et al., 2018) and guarantees uniqueness of the optimum. We relax Assumption 1.4.2 in Appendix A.2.3.

Theorem 1.4.2. *Let Assumptions 1.2.1, 1.2.2, 1.3.1, 1.4.1, 1.4.2 hold. Take a small* $1/4 > \xi > 0$*,* $\alpha_{k,w} = J/w$ *for a finite* $J \geq 1/\sigma$*. Let* $n^{1/4-\xi} \geq C\sqrt{p\log(n)\gamma_N T^{B_p}\log(KT)}$*,* $\eta_n = 1/n^{1/4+\xi}$*, for finite constants* $B_p, C > 0$*. Then with probability at least* $1-1/n$*, for a constant* $\bar{C}' < \infty$*, independent of* (p,n,N,K,T)*,*

$$||\beta^* - \hat{\beta}^*||^2 \leq \frac{p\bar{C}'}{\check{T}}.$$

The proof is in Appendix A.3.3. Theorem 1.4.2 provides a *small sample* upper bound on the distance between the estimated policy and the optimal one. The bound only depends on T (and not n) since n is assumed to be sufficiently larger than T.

Corollary 4. *Let the conditions in Theorem 1.4.2 hold. Let* $K = 2(T/p+1)$. *Then with probability at least* $1-1/n$, *for a constant* $C' < \infty$ *independent of* (p,n,N,K,T),

$$W(\beta^*) - W(\hat{\beta}^*) \leq \frac{pC'}{K}.$$

The proof is in Appendix A.3.5. The above corollary formalizes the out-of-sample regret bound as we choose $K = 2(T/p+1)$.

Researchers may wonder whether the procedure is "harmless" also on the in-sample units. We provide in-sample guarantees in the following theorem.

Theorem 1.4.3 (In-sample regret)**.** *Let the conditions in Theorem 1.4.2 hold. Then with probability at least* $1-1/n$, *for a constant* $c < \infty$ *independent of* (p,n,N,K,T),

$$\max_{k\in\{1,\cdots,K\}} \frac{1}{\check{T}} \sum_{w=1}^{\check{T}} \left[W(\beta^*) - W(\check{\beta}_k^w) \right] \leq c\frac{p\log(\check{T})}{\check{T}}.$$

The proof is in Appendix A.3.3. Theorem 1.4.3 guarantees that the cumulative welfare in *each* cluster k, incurred by deploying the current policy $\check{\beta}_k^w$ at wave w (recall that in the general p-dimensional case we have $\check{T}$ many waves), converges to the largest achievable welfare at a rate $\log(T)/T$, also for those units participating in the experiment.[43] This result guarantees that the proposed design is not *harmful* to experimental participants.

We conclude this sub-section by deriving a faster (exponential) convergence rate of the out-of-sample regret (but not in-sample regret) with a different choice of the learning rate.

[43] By a first-order Taylor expansion, a corollary is that the bound also holds for $\check{\beta}_k^w \pm \eta_n$ up to an additional factor which scales to zero at rate η_n (and therefore negligible under the conditions imposed on n).

Theorem 1.4.4 (Out-of-sample regret with larger sample size)**.** *Let Assumptions 1.2.1, 1.2.2, 1.3.1, 1.4.1, 1.4.2 hold, with $W(\beta)$ being τ-smooth, and $K = 2T + 2$. Take a small $1/4 > \xi > 0$, $\alpha_{k,w} = 1/\tau$. Let $n^{1/4-\xi} \geq C\sqrt{p\log(n)\gamma_N e^{TB_p}\log(KT)}$, $\eta_n = 1/n^{1/4+\xi}$, for finite constants $B_p, C > 0$. Then with probability at least $1 - 1/n$, for constants $0 < c_0, c_0' < \infty$, independent of (n, N, K, T),*

$$W(\beta^*) - W(\hat{\beta}^*) \leq c_0 \exp(-c_0' K).$$

The proof is in Section A.3.3. The main additional restriction in Theorem 1.4.4 is that the sample size grows *exponentially* in the number of iterations (instead of polynomially). The theorem leverages properties of the gradient descent under strong concavity and smoothness (Bubeck, 2014), combined with derivations discussed below.

To our knowledge, these are the first regret guarantees under unknown (and partial) interference.

We now contrast with past literature. In the online optimization literature, the rate $1/T$ is common for convex optimization (Bottou et al., 2018), assuming independent units (see Duchi et al., 2018, for out-of-sample regret rates only). Differently, here, because of interference, we leverage between-clusters perturbations. Also, we do not have direct access to the gradient, and related optimization procedures are those of Flaxman et al. (2004); Agarwal et al. (2010), where regret can converge at rate $O(1/T)$ in expectation only, while high-probability bounds are $1/\sqrt{T}$.[44] Here, we exploit within-cluster concentration and between clusters' variation to control large deviations of the estimated gradients and obtain faster rates for high-probability bounds. In our derivations, the perturbation parameter depends on the sample size, which is different from the references above, and the idea of circular estimation is novel due to repeated sampling. Wager and Xu (2021) derive $1/T$ regret guarantees in the different settings of market pricing, as $n \to \infty$, with independent units and samples for each wave. Our results are in a finite sample and do not impose independence or modeling assumptions other than partial interference.

[44] See Theorem 6 in Agarwal et al. (2010) and discussion below.

These differences require a different set of techniques for derivations. The proof of the theorem (i) uses concentration arguments for locally dependent graphs (Janson, 2004); (ii) it uses the within-cluster and between-clusters variation for consistent estimation of the marginal effect, together with the matching design to guarantee that there is a vanishing bias when estimating marginal spillover effects for $K < \infty$; (iii) it uses a recursive argument to bound the cumulative error obtained through the estimation and circular cross-fitting.

Remark 9 (Non-adaptive experiment)**.** In Appendix A.1.4 we introduce an experiment to estimate β^* without adaptive randomization, in the spirit of Section 1.3. Appendix A.1.4 shows how we can leverage perturbations and marginal effects as in Algorithm 1 to control the out-of-sample regret without an adaptive experiment, but also formally shows the drawbacks of a non-adaptive experiment, since it does not control the in-sample regret. □

A Comparison with the Unconstrained Optimum

Next, we study the following question: how does β^* compare to the policy that assigns treatments without restrictions on the policy function? We study this under more restrictive conditions. We omit the super-script k since our argument applies to any cluster. Consider

$$W_N^* - W(\beta^*), \quad W_N^* = \sup_{\mathscr{P}_N(\cdot)\in\mathscr{F}} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\Big[\mathbb{E}_{D\sim\mathscr{P}_N(A,X)}[Y_{i,t}|A,X]\Big] \tag{1.15}$$

where $\mathscr{F}$ denotes the set of conditional distribution of the vector $D \in \{0,1\}^N$, given the network A and the covariates of all observations X. Equation (1.15) denotes the difference between the expected potential outcomes, evaluated at the global optimum over all possible assignments (conditional on A, X), and the welfare evaluated at the optimal policy β^*.

Assumption 1.4.3 (Discrete parameter space, assignment and minimum degree)**.** Assume that $X_i \in \mathscr{X}$, $\mathscr{X} = \{1, \cdots, |\mathscr{X}|\}$, $|\mathscr{X}| < \infty$. Let $\pi(x,\beta) = \beta_x$, and $\mathscr{B} = [0,1]^{|\mathscr{X}|}$. Assume in addition that $\inf_{x,x',u'} \int l(x,u,x',u')dF_{U|X=x}(u) \geq \kappa$, for some $\kappa \in (0,1]$.

Assumption 1.4.3 states that we assign treatments based on finitely many observable types.[45] Each type $x \in \mathscr{X}$ is assigned a different probability β_x, which can take any value between zero and one. Assumption 1.4.3 also states that conditional on individual's type (X_i, U_i), any other unobserved type U_j can form a connection with individual i with some positive probability, conditional on observables X_j, provided that i and j are connected under the latent space representation (recall Equation 1.1).[46] The second restriction is on the potential outcomes. Let

$$\begin{aligned} Y_{i,t}(\mathbf{d}_t) &= \Big[\Delta(X_i) - c(X_i)\Big]\mathbf{d}_{i,t} + \mathscr{S}_{i,t}(\mathbf{d}_t) + \nu_{i,t}, \quad \mathbb{E}[\nu_{i,t}|X,A] = 0 \\ \mathscr{S}_{i,t}(\mathbf{d}_t) &= s\Big(\frac{\sum_{j\neq i} A_{i,j}\mathbf{d}_{j,t}1\{X_j = 1\}}{\max\{\sum_{j\neq i} A_{i,j}1\{X_j = 1\}, 1\}}, \cdots, \frac{\sum_{j\neq i} A_{i,j}\mathbf{d}_{j,t}1\{X_j = |\mathscr{X}|\}}{\max\{\sum_{j\neq i} A_{i,j}1\{X_j = |\mathscr{X}|\}, 1\}}\Big). \end{aligned} \tag{1.16}$$

Here $\Delta(\cdot)$ is the conditional direct treatment effect, and $c(\cdot)$ is the cost of the treatment. The function $s(\cdot)$ captures the spillover effects. The spillover effects depend on the fraction of treated neighbors. Spillovers are heterogeneous in the neighbors' types.

Assumption 1.4.4 (Opportunity costs of an equal-impact intervention with no spillovers)**.** Assume that $\Delta(x) = c(x)$ for all $x \in \mathscr{X}$.

Assumption 1.4.4 states that the cost of the treatment is the opportunity cost had the treatment been assigned to the same individuals who are disconnected. In a cash-transfer program, we may assign treatments to individuals in the same or nearby villages or to individuals spread out on an entire state or continent without creating spillover effects in the latter case. The cost of the treatment to assign treatments in the same villages is their opportunity cost, i.e., the direct treatment effects are assumed to be the same.[47]

[45] This is often imposed, see e.g., Manski (2004), Graham et al. (2010).

[46] This condition is consistent with Assumption 1.2.1, since the assumption states that the expected minimum degree is bounded from below by $\kappa\gamma_N^{1/2}$, which is smaller than the maximum degree $\gamma_N^{1/2}$.

[47] Note that, under such restriction, treating each individual is not optimal only if, for *some* treatment configurations, there are negative spillovers. For instance, giving cash transfers to richer individuals may decrease the average satisfaction with the program.

Theorem 1.4.5. *Let Equation* (1.16) *holds, with $s(\cdot)$ twice differentiable with bounded derivatives. Suppose that Assumption 1.2.1, 1.2.2, 1.4.3, 1.4.4 hold. Then as $N \to \infty, \gamma_N \to \infty$,*

$$\sup_{\mathscr{P}_N(\cdot)\in\mathscr{F}} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\Big[\mathbb{E}_{D\sim\mathscr{P}_N(A,X)}[Y_{i,t}|A,X]\Big] - W(\beta^*) \to 0.$$

The proof is in Appendix A.3.3. Theorem 1.4.5 shows that the assignment mechanism that maximizes welfare, when treatments can be assigned arbitrarily, converges to β^* with individualized treatments. It implies that collecting network information is, on average, not useful to improve welfare in this context. The theorem assumes that the maximum degree converges to infinity, but it may converge at a slower rate than N, consistently with our conditions in previous theorems. This is a novel result in the context of the literature on targeting networked individuals, which depends on the cost assumption.[48]

Corollary 5. *Let the conditions in Theorem 1.4.2 and Theorem 1.4.5 hold. Then for a constant $C < \infty$ independent of (N,n,T,K), $\lim_{N\to\infty,\gamma_N\to\infty} P\Big(W_N^* - W(\hat{\beta}) \leq \frac{C}{T}\Big) = 1$.*

Calibrated Experiments

In this section, we study the properties of the methodology in numerical studies. We calibrate simulations to data from Cai et al. (2015) and Alatas et al. (2012, 2016), while making simplifying assumptions whenever necessary. In the first calibration, the outcome is insurance adoption, and the treatment is whether an individual received an intensive information session in the experiment. In the second calibration, the treatment is whether a household received a cash transfer, and the outcome is program satisfaction.[49]

[48] In a different context, Akbarpour et al. (2018) show that for a class of diffusion mechanisms, random seeding is approximately optimal as we choose a few more seeds. Here, we do not study the problem from the perspective of network diffusion but instead focus on an exogenous interference mechanism with heterogeneity and provide a different set of results. Our result hinges on Assumption 1.4.4.

[49] The experiment of Cai et al. (2015) contains multiple arms assigned at the household and village level. Here, we only focus on the treatment effects of intensive information sessions, pooling the remaining arms together for simplicity. The experiment of Alatas et al. (2012) contains different arms assigned at the village level, as well as

Throughout these simulations, we study the problem of choosing a univariate parameter β, which denotes the unconditional treatment probability. In each cluster k, we generate data as

$$Y_{i,t} = \phi_0 + \phi_1 D_{i,t} + \phi_2 S_{i,t} + \phi_3 S_{i,t}^2 - cD_{i,t} + \eta_{i,t}, \quad S_{i,t} = \frac{\sum_{j \neq i} A_{i,j} D_{i,t}}{\max\{1, \sum_{j \neq i} A_{i,j}\}}, \tag{1.17}$$

where c is the cost of the treatment and $\eta_{i,t} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2)$. We consider two sets of parameters $\left(\phi_0, \phi_1, \phi_2, \phi_3, \sigma^2\right)$ calibrated to data from Cai et al. (2015) and Alatas et al. (2012, 2016) respectively. We obtain information on neighbors' treatment directly from data from Cai et al. (2015). For the second application, we merge data from Alatas et al. (2012), and Alatas et al. (2016) and use information from approximately one hundred observations whose neighbors' treatments are all observable to estimate the parameters.[50] For either application, we estimate a linear model as in Equation (1.17) also controlling for additional covariates to guarantee unconfoundedness of the treatment.[51] We consider as cost of treatment $c = \phi_1$, i.e., the opportunity cost of allocating the treatment to a population of disconnected individuals.

We generate clusters with $N = 600$ units, and sample $n \in \{200, 400, 600\}$. We generate a geometric network of the form $A_{i,j} = 1\left\{||U_i - U_j||_1 \leq 2\rho/\sqrt{N}\right\}, U_i \sim_{i.i.d.} \mathcal{N}(0, I_2)$, where the parameter ρ governs the density of the network. The geometric formation process and the $1/\sqrt{N}$ follows similarly to simulations in Leung (2020). We consider two networks, a "sparse network" with $\rho = 2$, reported in the main text, and a "dense network", with $\rho = 6$, studied in the Appendix. Throughout our analysis, without loss of generality, we report welfare divided

information on cash transfers assigned at the household level. Here, we study the effect of cash transfers only and control for village-level treatments when estimating the parameters of interest.

[50] This is different from Figure 1.2 where we use information from individuals whose 80% or more neighbors are observable. We make this choice in Section 1.2 to increase precision to estimate heterogeneous effects. This approach introduces a sampling bias in the estimation procedure, which we ignore for simplicity, given that our goal is not the analysis of the original experiment but only calibrating numerical studies.

[51] For Cai et al. (2015) the covariates are gender, age, rice area, literacy level, a coefficient that captures the risk aversion, the baseline disaster probability, education, and a dummy containing information on whether the individual has one to five friends. For Alatas et al. (2012) we control for the education level, village-level treatments, i.e., how individuals have been targeted in a village (i.e., via a proxy variable for income, a community-based method, or a hybrid), the size of the village, the consumption level, the ranking of the individual poverty level, the gender, marital status, household size, the quality of the roof and top (which are indicators of poverty).

by its maximum $W(\beta^*)$ (i.e., $W(\beta^*) = 1$), and we subtract the intercept ϕ_0 since ϕ_0 does not depend on β.

We conclude with details on estimation. We fix the perturbation parameter $\eta_n = 10\%$;[52] similarly, in the adaptive experiment, we choose the learning rate $10\%/\sqrt{t}$ with gradient norm rescaling as Remark 7. This choice guarantees that for each iteration, we only vary treatment probabilities by at most 10%, and the size of the variation decreases over each iteration, in the same spirit of learning rate under strong concavity without norm rescaling.[53] Since the model does not allow for time-varying fixed effects, we estimate marginal effects without baseline outcomes. For the multi-wave experiment, we initialize parameters at a small treatment probability $\beta = 0.2$.[54]

One-wave Experiment

First, we study the properties of the one wave experiment as we vary the number of clusters K and the sample size from each cluster n. We are interested in testing the one-sided null of whether we should increase the number of treated individuals to increase welfare, i.e.,

$$H_0 : \frac{\partial W(\beta)}{\partial \beta} \leq 0, \quad H_1 = \frac{\partial W(\beta)}{\partial \beta} > 0 \quad \beta \in [0.1, \cdots, \beta^*]. \tag{1.18}$$

In Figure A.4.1, in the Appendix, we report the power of the test as a function of the regret for $\rho = 2$. Power is increasing in the regret, the number of clusters, and sample size. However, the marginal improvement in the power from twenty to thirty clusters is small. This result is suggestive of the benefit of the method even with few clusters and small sample size.

Figure 1.5 illustrates the benefit of the method as, upon rejection of H_0, we recommend increasing the treatment probability by 5% as a function of the baseline treatment probability.

[52] Here 10% is consistent with the rule of thumb in Section A.2.5 which would prescribe values between 7% and 12% as we vary n. In Figure A.4.7 we report results as we vary η_n.

[53] This choice is preferable to $10\%/\sqrt{T}$ because it allows for larger steps in the initial iterations. A valid alternative choice is also $10\%/t$, corresponding to the one under strong concavity. The latter case has a practical drawback: updates become very small after a few iterations. For a comparison, see Figure A.4.5.

[54] This choice guarantees that no less than 10% of individuals are treated once we impose the negative perturbation.

For example, for $\beta = 0.2$, it indicates the relative welfare increase if we were to increase the treatment probability to 0.25, relative to the status quo where $\beta = 0.2$. The figure shows that the relative improvement can be as large as fifty percentage points compared to the status quo. In addition, while the gain is increasing in the sample size, substantial gains can be obtained even if the sample size from each cluster is as small as $n = 200$. This is particularly relevant for targeting information: in such a case, differences in power across different sample sizes only occur when the regret is small (see Figure A.4.1) or, equivalently, when we are already close to the optimum. As a result, once we take welfare effects into account, a larger sample size may lead to negligible improvements in welfare. As this example suggests, when the goal is welfare maximization, power analysis may be complemented by the welfare analysis discussed here.

Results are robust as we increase the density of the network with $\rho = 6$ (see Appendix A.4). In Table 1.1 we report the size of the test, which provides supportive evidence to the theoretical findings.

Multiple-wave Experiment

Next, we study the performance of the adaptive experiment. We let $T \in \{5, 10, 15, 20\}$. In Table 1.2 we report the welfare improvement of the proposed method with respect to a grid search method that samples observations from an equally spaced grid between $[0.1, 0.9]$ with a size equal to the number of clusters (i.e., $2T$). We consider the best competitor between the one that maximizes the estimated welfare obtained from a correctly specified quadratic function and the one that chooses the treatment with the largest value within the grid. The panel at the top of Table 1.2 reports the out-of-sample welfare improvement. The improvement is positive and up to three percentage points for targeting information and up to sixty percentage points for targeting cash transfers. Improvements are generally larger for larger T. In one instance only, for $T = 5$ and a small sample size $n = 200$, we observe a negative effect for targeting information of two percentage points. The panel at the bottom of Table 1.2 reports positive and large improvements for the in-sample welfare across all the designs, worst-case across clusters.

For the worst-case regret, we fix the number of clusters to $K = 40$ for our method and study the properties as a function of the number of iterations.[55] The improvements are twice as large for targeting information and thirty percentage points larger for targeting cash transfers. These are often increasing in T with a few exceptions.[56]

These results illustrate large benefits for *both* estimating policies to be implemented on a new population and to maximize participants' welfare. Figure 1.6 reports the out-of-sample regret of the method (and Figure A.4.2 the in-sample regret). The regret converges to zero as we increase the number of iterations. The larger sample size guarantees a smaller regret. In Appendix A.4.1, A.4.2, we report results for $\rho = 6$, consistent with findings in the main text. In Appendix A.4.3 we provide simulations with covariates using data from Alatas et al. (2012).

Table 1.1. Simulations: coverage. One wave experiment. 200 replications. Coverage for testing H_0 (size is 5%). First panel corresponds to $\rho = 2$, and second panel to $\rho = 6$.

	Information				Cash Transfer			
$K =$	10	20	30	40	10	20	30	40
$n = 200$	0.905	0.950	0.905	0.900	0.920	0.940	0.915	0.895
$n = 400$	0.980	0.960	0.900	0.925	0.980	0.960	0.895	0.930
$n = 600$	0.975	0.970	0.955	0.945	0.970	0.995	0.960	0.935
$n = 200$	0.925	0.880	0.880	0.900	0.925	0.940	0.905	0.905
$n = 400$	0.980	0.940	0.920	0.920	0.980	0.960	0.900	0.925
$n = 600$	0.975	0.890	0.930	0.995	0.975	0.995	0.950	0.915

Calibrated Simulations to M-Turk Experiment

Finally, in this subsection we briefly describe calibrated simulations using information from an M-Turk experiment. The experiment's goal was to increase information about the

[55] Fixing $K = 40$ allows us to change the number of iterations up to $T = 20$ while keeping fixed the number of clusters.

[56] The reason improvements are not always increasing in T is that uniform concentration may deteriorate for large T and small n as we consider the worst-case welfare across clusters.

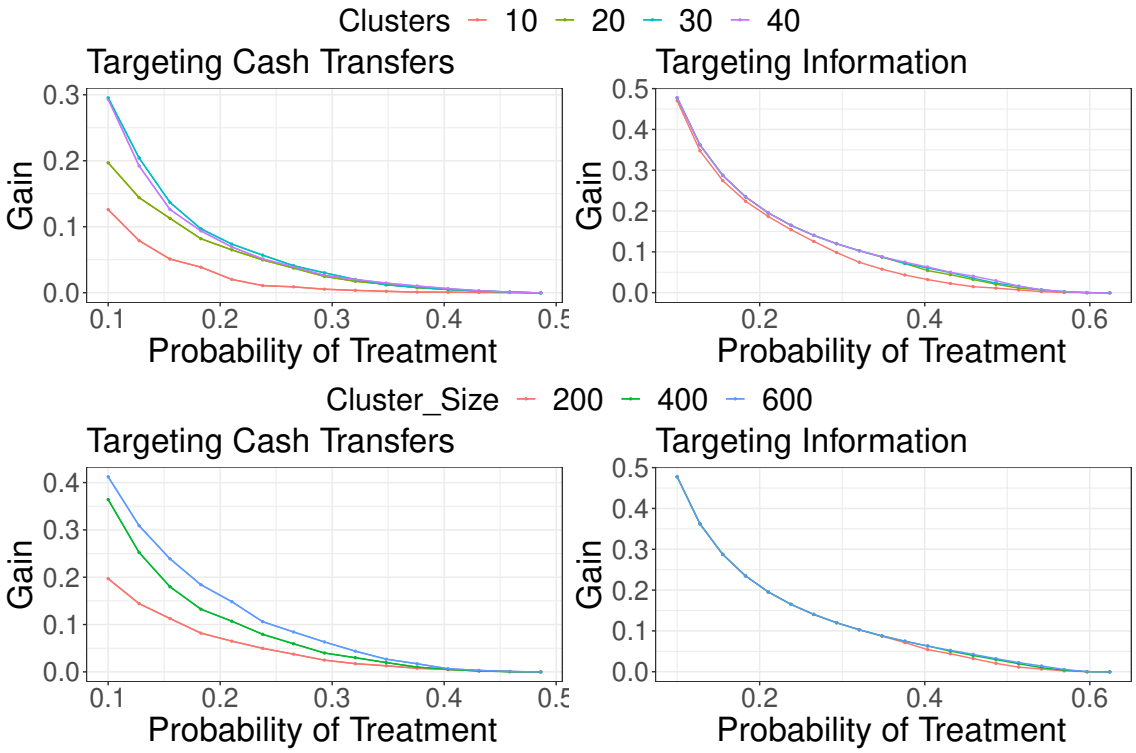


Figure 1.5. Simulations: one-wave experiment. $\rho = 2$. Expected percentage increase in welfare from increasing the probability of treatment β by 5% upon rejection of H_0. Here, the x-axis reports $\beta \in [0.1, \cdots, \beta^* - 0.05]$. The panels at the top fix $n = 400$ and vary the number of clusters. The panels at the bottom fix $K = 20$ and vary n.

Table 1.2. Simulations: multi-wave experiment. 200 replications. Relative improvement in welfare with respect to best competitor for $\rho = 2$. The panel at the top reports the out-of-sample regret and the one at the bottom the worst case in-sample regret across clusters.

	Information				Cash Transfer			
$T =$	5	10	15	20	5	10	15	20
$n = 200$	-0.026	0.014	0.043	0.033	0.295	0.390	0.528	0.322
$n = 400$	0.0003	0.026	0.026	0.035	0.462	0.444	0.589	0.563
$n = 600$	0.002	0.035	0.021	0.022	0.485	0.489	0.622	0.644
$n = 200$	1.103	1.370	1.451	1.447	0.254	0.276	0.305	0.323
$n = 400$	1.400	1.616	1.667	1.626	0.282	0.329	0.367	0.379
$n = 600$	1.546	1.771	1.828	1.751	0.279	0.335	0.364	0.368

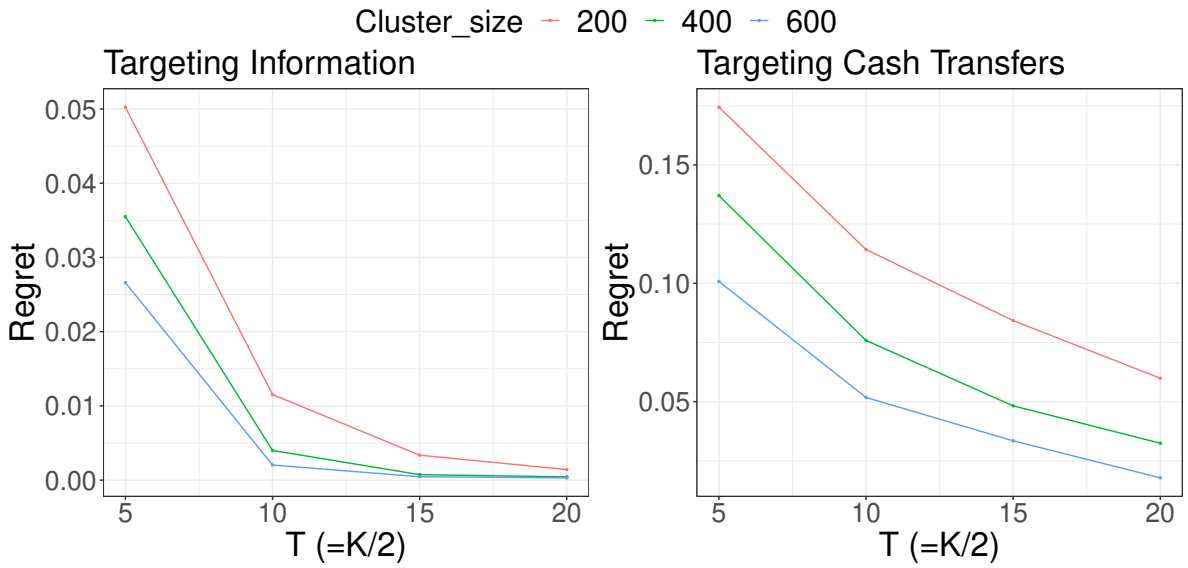


Figure 1.6. Simulations: out-of-sample regret. Adaptive experiment $\rho = 2$. 200 replications. The panel reports the out-of-sample regret of the method as a function of the number of iterations.

severity of COVID-19 to increase vaccination adoption. Details are contained in Appendix A.4.4.[57]

Specifically, at the beginning of March 2020 (before the vaccination campaign was extensively implemented), we ran an M-Turk experiment where each individual was assigned either of two arms. A control arm, which consisted of a survey asking basic questions on characteristics of the participants, and a treatment arm. The treatment arm was first assigned simple survey questions. Then, individuals under treatment were assigned three questions about COVID, whose correct answer was rewarded with a small economic incentive. The scope of the treatment was to increase awareness of the severity of the disease by asking questions and showing the correct answers to facilitate information transmission. At the end of the survey, both control and treated units were asked when they would have done the vaccine. Our outcome of interest is binary and equals whether individuals would have done the vaccine either as soon as possible or during the spring. We estimate the model with 1035 participants.[58] We estimate

[57] The experiment was certified an IRB exempt by UCSD, Human Research Protections Program.

[58] We collected information from 2411 participants. We removed 158 observations that had already received the vaccine in March 2020 and 203 observations that took less than thirty seconds and more than five minutes to take the survey. We also removed all those observations who were not living in the US.

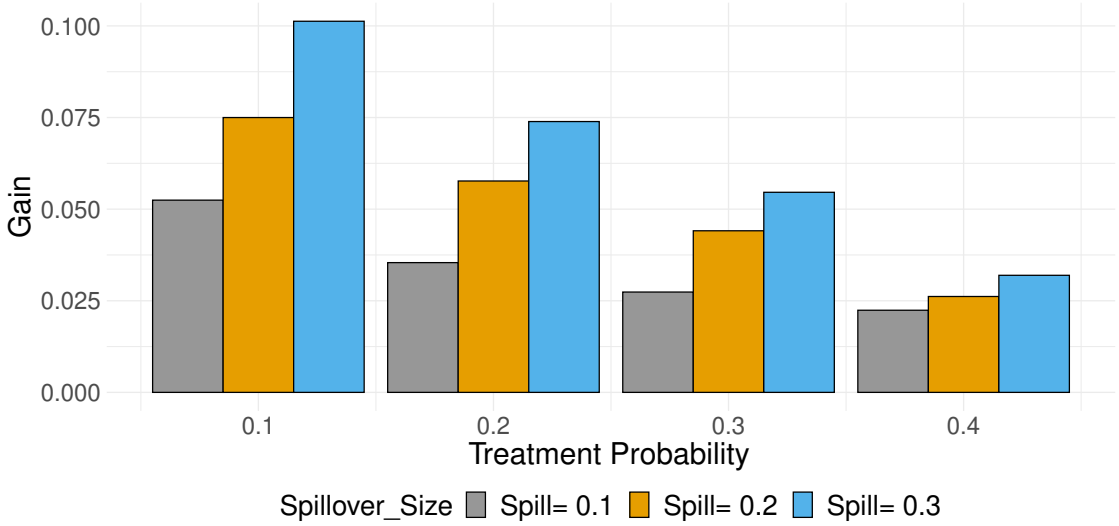


Figure 1.7. Calibrated simulations with Amazon M-Turk. One wave experiment calibrated to M-Turk experiment. Relative welfare improvement for increasing the probability of treatment by ten percent, conditional on rejecting the null hypothesis. 200 replications. Different columns correspond to different levels of spillover effects (captured by α). Here $K = 20$.

treatment effects by running a simple linear regression, where the treatment dummy interacts with the dummy, indicating whether the individual classifies herself as liberal, conservative, or "prefer not to say". We find heterogeneity of treatment effects, with positive effects of the treatment on liberals only. The model and policy function are described in Appendix A.4.4.

We simulate a one-wave experiment calibrated to the results from the M-Turk experiment and report results in Figure 1.7. We simulate spillover effects as those are not observable in the experiment and study the method's properties for different levels of spillover effects. In the figure, we report the relative welfare improvement of increasing treatment probabilities on the group receiving positive effects of ten percent upon rejection of the null hypothesis H_0 in Equation (1.18). As the size of the spillovers increase, welfare improvements upon rejection increase and, similarly power, with improvements up to twelve percentage points.

Dynamic Treatment Effects: an Overview

This section briefly discusses an extension with dynamic treatments. We omit covariates and assume that $X_i = 1$, and defer to Appendix A.1.3 formal details.

In the presence of carry-over effects, outcomes also depend on past treatments. For treatments assigned with exogenous parameters $(\beta_{k,1}, \cdots, \beta_{k,t})$ as in Definition 1.2.3, we let $Y_{i,t}^{(k)} = \Gamma(\beta_{k,t}, \beta_{k,t-1}) + \varepsilon_{i,t}^{(k)}, \mathbb{E}_{\beta_{k,1:t}}\left[\varepsilon_{i,t}^{(k)}\right] = 0$, for some unknown $\Gamma(\cdot)$, $\varepsilon_{i,t}^{(k)}$. The components $\beta_{k,t}, \beta_{k,t-1}$ capture present and carry-over effects that result from individual and neighbors' treatments in the past two periods. We study the problem of estimating a path of treatment probabilities $(0, \beta_1, \cdots, \beta_T)$ from an experiment, where, in the first period, we assume for simplicity that none of the individuals is treated. We then implement this path on a new population without having access to the outcomes of such a new population. We maximize long-run welfare, defined as follows:

$$\mathscr{W}(\{\beta_s\}_{s=1}^{T^*}) = \sum_{t=1}^{T^*} q^t \Gamma(\beta_t, \beta_{t-1}), \quad \text{for a given horizon } T^*, \text{ and discounting factor } q < 1.$$

The long-run welfare defines the cumulative (discounted) welfare effect obtained from a certain sequence of decisions $(\beta_1, \beta_2, \cdots)$. Our goal is to find the sequence that maximizes the long-run welfare. We start from the following observation: by the first-order conditions

$$\frac{\partial \Gamma(\beta_t, \beta_{t-1})}{\partial \beta_t} + q \frac{\partial \Gamma(\beta_{t+1}, \beta_t)}{\partial \beta_t} = 0, \quad \text{for all } t. \tag{1.19}$$

Equation (1.19) shows that the choice of the welfare-maximizing parameter β_{t+1} may depend on the previous two decisions. Using ideas from reinforcement learning (Sutton and Barto, 2018), we parametrize future treatment probabilities based on past treatment probabilities as $\beta_{t+1} = h_\theta(\beta_t, \beta_{t-1}), \theta \in \Theta$, for some given function $h_\theta(\cdot)$, and find the parameter $\theta \in \Theta$, which maximizes welfare. The algorithm estimates the function $\Gamma(\cdot)$ using a single wave experiment

and then maximizes

$$\widehat{\theta} \in \arg\max_{\theta\in\Theta} \sum_{t=1}^{T^*} q^t \widehat{\Gamma}(\beta_t, \beta_{t-1}), \quad \beta_t = h_\theta(\beta_{t-1}, \beta_{t-2}) \quad \forall t \geq 1, \quad \beta_0 = \beta_{-1} = 0.$$

Our main insight here is on how to design the experiment to estimate $\widehat{\Gamma}(\cdot)$. If $\widehat{\Gamma}$ is estimated with randomization based on a simple grid-search procedure, the rate of convergence of the regret would be $1/\sqrt{K}$ (see Appendix A.1.3). However, in Appendix A.1.3 we show that by using local perturbations – similarly to what was discussed in previous sections – and leveraging information from the estimated gradient, we can achieve a convergence rate of the out-of-sample regret of order $1/K$ (but not faster in n).[59] The single-wave procedure comes at a cost: the rate is specific to the one-dimensional setting and carry-overs over two consecutive periods; in p dimensions, the rate would be of much slower order due to the curse of dimensionality. This is different from the adaptive design discussed in previous sections, where the dimension does not affect the rate in K, and it opens new research questions in the presence of carry-over effects.

Conclusions

This chapter makes two main contributions. First, it introduces a single-wave experimental design to estimate the marginal effect of the policy and test for policy optimality. The experiment also allows us to identify and estimate treatment effects, which can be of independent interest. Second, it introduces an adaptive experiment to maximize welfare. We derive asymptotic properties for inference and provide a set of guarantees on the in-sample and out-of-sample regret. To our knowledge, this is the first paper to study inference on marginal effects and adaptive experimentation with unobserved interference.

In a single-wave experiment, we encourage researchers to identify and report estimates

[59] The idea is as follows: we randomize probabilities $(\beta_1, \beta_2) \in [0,1]^2$ from a coarse grid and use small groups (three) clusters to estimate the partial derivatives at each point. We extrapolate the value of $\Gamma(\cdot)$ throughout the set $[0,1]^2$ with a first-order Taylor approximation around the closest point in the grid, using the information on the estimated marginal effect.

of the marginal effects. We show that we can use the information on the marginal effects to conduct hypothesis testing on policy optimality and, ultimately, incorporate uncertainty in decision-making. Future research may explore notions of efficiency in this setting.

Our work opens new questions also from a theoretical perspective. The main assumption is that clusters are observable before the experiment starts. We leave it to future research to study properties when (i) such clusters are not fully disconnected, in the same spirit of Leung (2021); (ii) such clusters need to be estimated, similarly to graph-clustering procedures as in Ugander et al. (2013); (iii) clusters present different distributions (see Appendix A.1.2). Similarly, it may be interesting to study the properties of our method, as the degree of interference is proportional to the sample size. This is theoretically possible, as illustrated in Theorem 1.3.1, and we leave its analysis to future research. An open question is how to combine policy learning with procedures that impute the network (e.g., Alidaee et al., 2020; Breza et al., 2020; Manresa, 2013).

Finally, an important assumption of the literature on adaptive experimentations is that welfare is a function of observable characteristics. We leave it to future research to study whether (and how) we may allow for unobserved utilities.

Chapter 1 is currently being prepared for submission for publication of the material. Davide Viviano is the sole author of this material.

Chapter 2

Policy Targeting under Network Interference

Introduction

This chapter studies the problem of optimally allocating treatments in the presence of spillover effects. Unlike the previous chapter, we consider a setting where researchers collect information from an existing experiment or observational study. Researchers observe the neighbors' identities of the participants, but no independent clusters are necessarily available. Using such information, we introduce a method that maximizes the average social welfare sample analog when spillovers occur.[1] Relevant applications include public policy programs, cash transfer programs, educational programs, information campaigns, to cite some.[2]

Consider a setting where units are connected in a network. Individuals are assigned treatments whose effects are assumed to propagate locally in the network (i.e., to their neighbors). Researchers sample n units from an experiment or a quasi-experiment, collecting information on their covariates, treatment assignments, outcomes, covariates, and assignments of their neighbors. We do not require knowledge of the entire population network. Researchers' goal is to estimate a treatment allocation rule for new applications. For example, consider the problem discussed in

[1] While Chapter 1 mostly studies the problem of online policy design, this chapter focuses on offline policy design.

[2] Some examples of relevant applications are Barrera-Osorio et al. (2011); Egger et al. (2019); Opper (2016); Zubcsek and Sarvary (2011); Bond et al. (2012). Spillover effects have been documented in development economics (Banerjee et al., 2013), social economics (Sobel, 2006), medicine (Christakis and Fowler, 2010) among many others.

Cai et al. (2015) on studying the effect of information sessions on insurance take-up in villages subject to environmental disasters. Then our method uses the experimental data collected by Cai et al. (2015) to estimate which individuals the insurance company should target in new villages.

The first challenge for policy targeting is that the network on the target units may be unobserved due to the cost associated with collecting it on large populations.[3] For instance, in the example from Cai et al. (2015), the insurance company has only access to the network information of the experimental participants. It is costly or infeasible to collect network information from all target individuals in a region or country. Considering this, we develop a method that allows for arbitrary constraints on the policy space, with the network not necessarily observed on the target sample. We leverage the dependence between covariates and network information on the in-sample units to target individuals.

A second challenge is heterogeneity in treatment effects: an individual may respond differently to her and her neighbors' treatment depending on her type (e.g., her age or education). We impose and leverage the exogenous and anonymous interference assumption[4], often documented in applications[5], to estimate welfare without imposing conditions on the individual heterogeneity in treatment effects.

Our method, entitled Network Empirical Welfare Maximization (NEWM), estimates the welfare of the network as a function of the policy using arbitrary (machine-learning) estimators; it then solves an optimization procedure over the policy space using an exact optimization

[3]The reader may refer to Breza et al. (2020) for a discussion on the cost associated with collecting network information.

[4]In particular, we assume that potential outcomes are functions of individual and neighbors' treatment assignments (but not neighbors' identity), the number of neighbors, arbitrary individual characteristics which are observable in the experiment but may not be observed on the target sample (e.g., covariates, centrality measures, or summary statistics of neighbors covariates), and exogenous unobservables. See Manski (2013) for a discussion on exogenous and anonymous interference.

[5]Cai et al. (2015) show that inviting individuals to information sessions generates large effects on their friends' insurance take-up in subsequent sessions. However, an individual's take-up decision does not depend on her friends' take-up decision, consistently with an exogenous interference model. Examples of studies with models consistent with the assumption of anonymous and exogenous interference include Sinclair et al. (2012); Duflo et al. (2011); Muralidharan et al. (2017), where for the second reference, networks can be considered groups of classrooms with units within each classroom being fully connected. Athey et al. (2018) provide a general framework for testing the exogenous and anonymous assumption.

algorithm. We interpret the policy targeting as a treatment choice problem (Manski, 2004; Kitagawa and Tetenov, 2018; Athey and Wager, 2021), which we extend to the case of network interference. We evaluate the method's performance based on its maximum regret, i.e., on the difference between the largest achievable welfare and the welfare from deploying the estimated policy function. Unlike in the previous chapter, here, the regret also depends on the topology of the network, as no independent clusters are available.

From a theoretical perspective, we make three contributions: (i) we derive the first set of guarantees on the regret for treatment rules with spillover effects; (ii) we introduce an estimation procedure to achieve the $n^{-1/2}$-convergence rate for estimation with machine learning estimators and networked units; (iii) we show that for a large class of policy functions, the optimization problem can be written as a mixed-integer linear program, which can be solved using off-the-shelf optimization routines.

First, we discuss the identification of social welfare under network interference for a generic network formation model. Identification relies on the unconfoundedness of treatment assignments and the conditional network exogeneity, i.e., potential outcomes independent of the neighbors' identity, conditional on individual observable characteristics.[6] We then study semi-parametric (machine-learning) estimators for the welfare and analyze the performance of the estimated policy. We show that under regularity conditions, the regret of the estimated policy scales at the minimax rate $1/\sqrt{n}$, under bounded degree. In the presence of an unbounded degree, we characterize the rate as a function of the maximum degree.

New challenges in our derivations are: (i) individuals' dependence on neighbors' assignments that we control through contraction inequalities; (ii) statistical dependence, which invalidates standard symmetrization arguments, and which is arbitrary unconditionally on the network. Also, with networked units, semi-parametric (machine-learning) estimators may present

[6]Network exogeneity is often and explicitly stated when inference is performed unconditionally on the network structure (Leung, 2020). The reader may refer to Goldsmith-Pinkham and Imbens (2013) for a discussion. Network exogeneity is attained if, for example, two individuals form a link based on observable individual-level characteristics and unobservables, which are independent of potential outcomes.

non-vanishing bias even when cross-fitting methods are employed (Chernozhukov et al., 2018; Athey and Wager, 2021). We introduce a novel cross-fitting algorithm for networked observations to control the bias and achieve the regret's minimax rate.[7]

We guarantee that the optimization procedure achieves the in-sample optimum by casting the problem in a mixed-integer linear program, which can be solved using off-the-shelf algorithms. We show that we can achieve a linear representation of spillover effects in the objective function by introducing an additional set of linear constraints and binary decision variables.

Data from Cai et al. (2015) illustrate the advantages of the method. The NEWM method leads to (out-of-sample) improvements up to twelve percentage points compared to methods that ignore network effect (Akbarpour et al., 2018; Kitagawa and Tetenov, 2018; Athey and Wager, 2021), despite network information *not* being accessible on the target villages.

This chapter builds on the growing literature on statistical treatment choice. Examples include Kitagawa and Tetenov (2018), Kitagawa and Tetenov (2019), Athey and Wager (2021), Mbakop and Tabord-Meehan (2021). A list of additional references on optimal treatment allocation includes Armstrong and Shen (2015), Bhattacharya and Dupas (2012), Hirano and Porter (2009), Stoye (2009), Stoye (2012), Tetenov (2012) among others. Further connections are more broadly related to the literature on classification (Elliott and Lieli, 2013). Our focus is quite different from previous references: we estimate an optimal policy when treatments generate spillovers. This paper is the first to study the properties of targeting under network interference in the context of the empirical welfare maximization literature.

It is important to distinguish our framework from the $i.i.d.$ setting of multi-valued treatments of Kitagawa and Tetenov (2018); Zhou et al. (2018). A first conceptual difference is that here individuals depend on neighbors' assignments, while treatments are individual-specific. This structure permits the network on the target units not to be observable. The second difference is that individuals exhibit dependence, and hence standard theoretical arguments based on $i.i.d.$

[7]The algorithm consists of coloring individuals, assigning the same colors to units far enough in the network, and estimating different nuisance functions using units assigned to the same color.

sampling fail, such as symmetrization. This motivates (a) a novel estimation algorithm[8]; (b) theoretical arguments that exploit properties of the network in our derivations. The optimization program differs due to the dependence of individuals on neighbors' assignments.

A contribution of this chapter is to bridge the literature on causal inference and statistical treatment choice with the literature on network interference and influence maximization. This latter strand of literature mostly focuses on detecting the most influential "seeds" based on particular notions of centrality (Bloch et al., 2017) but often ignores heterogeneity in treatment effects and constraints on the policy space. Recent advances include Jackson and Storms (2018), Akbarpour et al. (2018), Banerjee et al. (2017), Banerjee et al. (2014), Galeotti et al. (2017) among others. Examples in computer science and marketing are Kempe et al. (2003), Eckles et al. (2019) and references therein. A disadvantage of using centrality measures for targeting is that an empirical comparison is possible only between two or a few decision rules through cluster experiments (Banerjee et al., 2017; Kim et al., 2015; Chin et al., 2018). Our contribution is complementary to this literature. The main differences are: we estimate policies from an infinite but constrained set of candidates, hence allowing the network on the target sample not to be observable; we leverage local interference while allowing for heterogeneity.

Spillover effects have been studied in the context of policy intervention from different angles. Bhattacharya et al. (2019) and Wager and Xu (2021) study the effect of a *global* policy intervention on social welfare, using partial identification and a sequential experiment respectively. This chapter considers individualized policy interventions instead and point identification due to the local interference assumption. In the presence of spillover effects, Li et al. (2019), Graham et al. (2010), Bhattacharya (2009) consider the problem of optimal *allocation of individuals* across independent groups. Differences are: (i) policy functions denote group assignment mechanisms instead of binary treatment allocations, inducing a different definition and identification of the welfare function; (ii) the allocation does not allow for constrained environments, and

[8]Namely when using standard cross-fitting for policy learning (Athey and Wager, 2021), the predicted propensity score and conditional mean of individual i is dependent on unobservables and observables of i since it depends on observables of i's friend, j, who is statistically dependent on i.

(iii) the authors assume a clustered network structures with small independent clusters. Su et al. (2019) discuss instead a closed-form expression for treatment allocations under interference in an *unconstrained* environment, assuming linearity. In this chapter, instead, we do not impose such structural assumptions. The constraints lead to a different estimation strategy and justify the regret analysis discussed here. Laber et al. (2018) consider a Bayesian structural model whose estimation relies on computational intensive Monte Carlo methods.

We build a connection to the econometric and statistical literature on inference on networks, including literature on social interaction (Manski, 2013; Manresa, 2013; Auerbach, 2019), and more generally causal inference under interference (Liu et al., 2019; Li et al., 2019; Hudgens and Halloran, 2008; Goldsmith-Pinkham and Imbens, 2013; Sobel, 2006; Sävje et al., 2021). The exogenous and anonymous interference condition is most closely related to Leung (2020). However, knowledge of treatment effects is not sufficient to construct optimal treatment rules in the presence of either (or both) constraints on the policy functions or treatment effects heterogeneity. We use the notion of non-parametric estimators from the literature on causal inference (Aronow and Samii, 2017) for the construction of welfare. Here, we devise a novel algorithm to achieve fast convergence rates with interference and semi-parametric estimators.

More recent papers have followed our work, studying targeting on networks in new directions. In a more recent paper, Kitagawa and Wang (2021) discuss the problem of targeting in a parametric setting in the presence of a SIR network. The advantage of a model-based approach is the lack of a pilot experiment, but its validity depends on the specific application in mind. Following our work, Ananth (2020) introduces more general treatment configurations arbitrarily dependent on the network information of the target units. However, such assignments are infeasible in applications when the network of the target units is costly or impossible to collect. When feasible, these assignments may be prone to overfitting.[9] Here, individuals can be targeted based on observable covariates, which may also include individual-specific network statistics when available from the target population.

[9]Overfitting can be the result of a violation of the bounded VC-dimension.

Setup and Identification

This section discusses the main identification conditions, the causal estimands of interest, and defines utilitarian welfare under interference.

Experiment and Sampling

The policy targeting exercise considered in this paper consists of two steps. In the first step, researchers collect data from an experiment or quasi-experiment. In the second step, they estimate a policy recommendation to be implemented on an independent sample $j \in \mathscr{J}$.

We start introducing the notation. Each individual is associated with an arbitrary vector of covariates $Z_i \in \mathscr{Z}$ This vector is random, and it may contain individual, neighbors' covariates statistics, and network statistics. Therefore, the entries of Z_i may exhibit dependence. We impose restrictions in Assumption 2.3.2.

There are E (finitely) many individuals who participate in the experiment and who constitute an independent network. These individuals are connected under an adjacency matrix A. We define

$$A \in \mathscr{A} \subseteq \{0,1\}^{E\times E}, \quad \mathscr{N}_i = \Big\{k \in \{1,\cdots,E\} \setminus i : A_{i,j} = 1\Big\}, \quad Z = (Z_i)_{i=1}^{E},$$

where $\mathscr{A}$ denotes the set of symmetric and unweighted adjacency matrices, $\mathscr{N}_i$ denotes all neighbors of individual i, and Z the matrix of covariates of all participants.[10] For example, in Cai et al. (2015) $\{1,\cdots,E\}$ denotes all individuals in those villages in the experiment. In full generality, we let edges be drawn as $A_{i,j} \sim \mathscr{P}_{i,j}$, with $\mathscr{P}_{i,j}$ being left unspecified, and arbitrary dependent. Define $|\mathscr{N}_i|$ the cardinality of the set $\mathscr{N}_i$.

For some function $f_D(\cdot)$ known in an experiment, and to be estimated in a quasi-

[10]Namely, we say that two individuals (i,j) are connected if $A_{i,j} = 1$, where $A_{i,j}$ denote the edge between individual i and individual j, with $A_{i,j} = A_{j,i} \in \{0,1\}$ and $A_{i,i} = 0$.

experiment, a binary treatment is assigned to each experimental participant as follows:

$$D_i = f_D(Z_i, \varepsilon_{D_i}), \quad \varepsilon_{D_i} \sim_{iid} \mathscr{D}, \quad i \in \{1, \cdots, E\}. \tag{2.1}$$

We discuss explicit conditions on ε_{D_i} in Section 2.2.2. Researchers sample $n \leq E$ experimental participants at random, and observe the vector

$$\left(Y_i, Z_i, Z_{j \in \mathscr{N}_i}, D_i, D_{j \in \mathscr{N}_i}, \mathscr{N}_i\right)_{i=1}^{n}, \quad \text{for some } n \leq E$$

where Y_i denotes the post-treatment outcome.[11]

Interference

In the presence of interference, the outcome of individual i depends on its treatment assignment and all other units' treatment assignments. For a given vector $D^E = (D_i)_{i=1}^E$ denoting the vector of treatment assignments of experimental participants, we let

$$Y_i = \tilde{r}_E(i, D^E, A, Z_i, \varepsilon_i), \quad i \in \{1, \cdots, E\} \tag{2.2}$$

for some (unobserved) random variables ε_i. The above equation states that experimental participants depend on the assignment of all units in the experiment. In the following condition, we impose that the outcomes only depend on the treatment assigned to their first-degree neighbors, and we make assumptions about interactions being anonymous.

Assumption 2.2.1 (Interference)**.** For all $(v,u) \in \{1, \cdots, E\}^2$, for any $D^E = (D_k)_{k=1}^E \in \{0,1\}^E$, $\tilde{D}^E = (\tilde{D}_k)_{k=1}^E \in \{0,1\}^E$ and $A, \tilde{A} \in \mathscr{A}$,

$$\tilde{r}_E(v, D^E, A, z, e) = \tilde{r}_E(u, \tilde{D}^E, \tilde{A}, z, e)$$

[11] Such a sample can be constructed as follows: researchers sample n random vertices; for each of these vertices, they observe the individual outcome, covariates, treatment assignments, and neighbors' identities; for each individual, they then collect information on neighbors' treatment assignments and covariates.

for all $z \in \mathscr{Z}$, $e \in \Big\{\text{supp}(\varepsilon_v) \cup \text{supp}(\varepsilon_u)\Big\}$, if all the following three conditions hold: (i) $\sum_k A_{v,k} = \sum_k \tilde{A}_{u,k}$; (ii) $\sum_k A_{v,k} D_k^E = \sum_k \tilde{A}_{u,k} \tilde{D}_k^E$; (iii) $D_v^E = \tilde{D}_u^E$.

Assumption 2.2.1 postulates that outcomes only depend on (i) the number of first-degree neighbors, (ii) the number of first-degree treated neighbors (iii) individual's treatment status as well as covariates of interest. Under Assumption 2.2.1 we can write for each individual $i \in \{1, \cdots, E\}$

$$Y_i = r\Big(D_i, \sum_{k \in \mathscr{N}_i} D_k, Z_i, |\mathscr{N}_i|, \varepsilon_i\Big), \tag{2.3}$$

for some possibly unknown function $r(\cdot)$.

Under the above condition, the outcome can depend on the number of treated neighbors, the share of treated neighbors, or whether at least one neighbor is treated. Heterogeneity is allowed and captured by the dependence of $r(\cdot)$ with Z_i and $|\mathscr{N}_i|$. The assumption is consistent with findings in Cai et al. (2015), where the authors document exogenous interference and lack of endogenous spillover effects. The model is closely related to Leung (2020), and Athey et al. (2018) provide a general framework for testing anonymous and local interference. We complement Equation (2.3) with the following condition.

Assumption 2.2.2 (Unconfoundedness)**.** For some arbitrary function $f_D(.)$, let the following hold: for all $i \in \{1, \cdots, E\}$,

(A) $D_i = f_D(Z_i, \varepsilon_{D_i})$ and ε_{D_i} are *i.i.d.*, and $\varepsilon_{D_i} \perp \Big(Z, A, (\varepsilon_j)_{j=1}^E\Big)$;

(B) For each i, $\varepsilon_i \perp \Big(A, Z\Big)\Big|Z_i, |\mathscr{N}_i|$;

Condition (A) states that the treatment is randomized in the experiment based on observable Z_i. Since Z_i may contain network information of a given individual, the assumption also accommodates randomization schemes where treatment assignment is based on network information. Condition (B) imposes that the network and others' covariates are independent of unobservables. Network exogeneity is attained if, for example, two individuals form a link

based on observable individual-level characteristics and unobservables, which are independent of potential outcomes.[12] We illustrate an example of a network formation model satisfying Assumption 2.2.2.

Example 2.2.1 (Network Formation)**.** *Let* $A_{i,j} = g(Z_i, Z_j, \alpha_{i,j}), (Z_i, \varepsilon_i) \sim_{i.i.d.} \Gamma$ *for some unknown function* $g(\cdot)$ *and* $\alpha_{i,j}$ *denoting exogenous unobservables jointly independent of observables and* $(\varepsilon_j)_{j=1}^{E}$ *(but possibly dependent with each other). Then Condition (B) in Assumption 2.2.2 holds.*

Policy Targeting: Problem Description

The planner's goal is to design a treatment mechanism that maximizes social welfare on units $j \in \mathscr{I}$, where $\mathscr{I}$ denotes the target population. For instance, in Cai et al. (2015), $\mathscr{I}$ defines the units in new villages where the experiment has not been conducted. The target population is separated from the experimental participants, and connected under a matrix $G \in \mathscr{G} \subset \{0,1\}^{|\mathscr{I}|\times|\mathscr{I}|}$ possibly unknown to the researcher.

With an abuse of notation, we will denote for all $j \in \mathscr{I}$, $\mathscr{N}_j = \{k \in \mathscr{I} \setminus j : G_{j,k} = 1\}$, the set of neighbors of individual j. The matrix G and covariates $(Z_j)_{j\in\mathscr{I}}$ are realized before treatments are assigned. The policy-maker has limited information on the target sample. Namely, she has only access to the variables

$$X_j \subseteq Z_j, \quad X_j \in \mathscr{X} \subseteq \mathscr{Z}, \quad j \in \mathscr{I},$$

denoting a *subset* of individuals' characteristics, which may also include arbitrary network information which is included in Z_j. She *does not* necessarily observe the neighbors' identity of the target participants. Her goal is to design a policy as follows:

(A) Individuals may be treated differently, depending on observables characteristics;

[12] Network exogeneity is often stated when inference is performed unconditionally on the network structure (Leung, 2020). The reader may refer to Goldsmith-Pinkham and Imbens (2013) for a discussion.

(B) The assignment mechanism must be easy to implement without requiring knowledge of the population network and covariates of all other individuals;

(C) The assignment mechanism can be subject to arbitrary constraints (e.g., ethical or economic constraints).

We formalize the above conditions by defining an *individualized* treatment assignment

$$\pi : \mathscr{X} \mapsto \{0,1\}, \quad \pi \in \Pi, \tag{2.4}$$

where Π denotes the set of constraints imposed on the decision function.[13] The map amounts to a partition of $\mathscr{X}$, the support of X_j. The policy function may depend on arbitrary observables X_j, such as measures of centrality, or covariates available to the policy-maker. This policy function (i) can be implemented in an online fashion, individual by the individual; (ii) it is feasible since it does not require knowledge of the population network and covariates of all units.[14]

The decision rule $\pi(\cdot)$, states that each unit $j \in \mathscr{I}$ is treated according to $D_j = \pi(X_j)$ for all $j \in \mathscr{I}$. The planner's goal is to choose the policy π that maximizes welfare. The utilitarian welfare is defined as the expected outcome once we fix $D_j = \pi(X_j)$ for all $j \in \mathscr{I}$. Namely,

$$W(\pi) = \frac{1}{|\mathscr{I}|} \sum_{j \in \mathscr{I}} \mathbb{E}\Big[Y_j \Big| \Big\{ (D_j)_{j \in \mathscr{I}} = (\pi(X_j))_{j \in \mathscr{I}} \Big\}\Big].$$

Importantly, the expectation is not conditional on a particular value of the treatment assignments $(\pi(X_j))_{j \in \mathscr{I}}$, but instead, it is conditional on the event that the treatment assignment mechanism is such that $D_j = \pi(X_j)$ for each unit in the target sample.

Welfare effects depend on two main components: (i) the direct effect of the treatment since each treatment is chosen based on $\pi(X_j)$, and (ii) the spillover effect since each neighbors' exposures also depend on the policy π. For example, welfare on the *experimental* participants,

[13] An example is the capacity constraint of the form $\int_{x \in \mathscr{X}} \pi(x) dF_X(x) \leq K$ for a constant K (Kitagawa and Tetenov, 2018).

[14] It also avoids overfitting with a domain constant in the size of the population.

reads as follows

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[r\Big(\pi(X_i),\sum_{k\in\mathscr{N}_i}\pi(X_k),Z_i,|\mathscr{N}_i|,\varepsilon_i\Big)\Big], \tag{2.5}$$

that denotes the effect of π also mediated through neighbors' assignments. Its definition follows from the deterministic nature of the policy $\pi \in \Pi$.

Unfortunately, we can only identify the welfare of the experimental participants. Discrepancies between the expected welfare of sample units and the targeted welfare $W(\pi)$ will reflect in the quality of the estimated policy function. We introduce the following definition.

Definition 2.2.1 (Sample-Target Discrepancy)**.** Define

$$\mathscr{K}_{\Pi}(n) = \sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[r\Big(\pi(X_i),\sum_{k\in\mathscr{N}_i}\pi(X_k),Z_i,|\mathscr{N}_i|,\varepsilon_i\Big)\Big] - W(\pi)\Big|,$$

for some function $\mathscr{K}_{\Pi}(.) : \mathbb{Z} \mapsto \mathbb{R}_+$.

Here, $\mathscr{K}_{\Pi}(.)$ denotes the difference between the expected welfare of in-sample units and the welfare of the target population. Such difference captures the bias induced by estimating the policy function on units that are possibly drawn from a different population than target units.

While we derive our results also as a function of the discrepancy, we can guarantee that $\mathscr{K}_{\Pi}(\cdot) = 0$ under the following conditions.

Assumption 2.2.3 (Local interference on the target population)**.** For a given target population $\mathscr{I}$, and for each (deterministic) policy function $\pi : \mathscr{X} \mapsto \{0,1\}, \pi \in \Pi$ let

$$W(\pi) = \frac{1}{|\mathscr{I}|}\sum_{j\in\mathscr{I}}\mathbb{E}\Big[r\Big(\pi(X_j),\sum_{k\in\mathscr{N}_j}\pi(X_k),Z_j,|\mathscr{N}_j|,\varepsilon_j\Big)\Big] \tag{2.6}$$

for some unobservables $(\varepsilon_j)_{j\in\mathscr{I}}$.

Assumption 2.2.3 states that the local interference model in Assumption 2.2.1 also holds on the target sample. In the following lemma, we provide conditions for the discrepancy to be equal to zero.

Lemma 2.2.1. *Let Assumption 2.2.1, 2.2.2, 2.2.3 hold. Suppose in addition that* $(Z_i, Z_{k\in\mathcal{N}_i}, |\mathcal{N}_i|, \varepsilon_i) \sim \mathscr{P}$ *for all* $i \in \{1,\cdots,E\} \cup \mathscr{I}$*. Then* $\mathscr{K}_\Pi(n) = 0$*, i.e., for all* $\pi \in \Pi$*,*

$$W(\pi) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\Big[r\Big(\pi(X_i), \sum_{k\in\mathcal{N}_i} \pi(X_k), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big].$$

Lemma 2.2.1 shows that $\mathscr{K}_\Pi(n)$ is equal to zero whenever the vector of unobservables ε_i, individual covariates Z_i, number of neighbors, and neighbors' covariates are drawn jointly from the same distribution. To gain further intuition, note that the condition in the above example is satisfied if the conditional distribution of $(Z_i, \varepsilon_i, Z_{N_i})\Big||\mathcal{N}_i| = l \sim \mathscr{P}_1(l)$, for some distribution $\mathscr{P}_1(l)$ which depends on the number of connections of an individual, and the *unconditional* distribution of the degree is the same across units. In Appendix B.1.2 we allow sampled and target units to have covariates and degrees drawn from different distributions.

Remark 10 (Spillovers on non-compliance)**.** Consider spillovers which also occur over individuals' compliance. Namely, let $D_i \in \{0,1\}$ denote the assigned treatment and $T_i \in \{0,1\}$ denotes the selected treatment from individual i. We model non-compliance as follows:

$$Y_i = r\Big(T_i, \sum_{k\in\mathcal{N}_i} T_k, Z_i, |\mathcal{N}_i|, \varepsilon_i\Big), \quad T_i = h_\theta\Big(D_i, \sum_{k\in\mathcal{N}_i} D_k, Z_i, |\mathcal{N}_i|, \nu_i\Big), \quad \nu_i \sim_{i.i.d.} \mathscr{V} \tag{2.7}$$

where ν_i denote an exogenous unobservables, independent from ε_i, and $(r(\cdot), \theta)$ are unknown, with θ denoting the set of parameters indexing h. Intuitively, treatment effects and compliance depend on neighbors' selected and assigned treatments. In Appendix B.1.1 we discuss conditions for identification, assuming that researchers collect information also on second-degree neighbors.

□

Identification

We discuss the problem of identifying the welfare of the experimental participants. We introduce two additional conditions.

Assumption 2.2.4. For all (i,j), $P(\varepsilon_i \leq t|Z_i = z, |\mathscr{N}_i| = l) = P(\varepsilon_j \leq t|Z_j = z, |\mathscr{N}_j| = l)$ for all $z \in \mathscr{Z}$, $l \in \mathbb{Z}$, $t \in \text{supp}(\varepsilon_i) \cup \text{supp}(\varepsilon_j)$.

The condition states that unobservables are identically distributed.

Assumption 2.2.5 (M-Local Dependence)**.** Assume that for any $A \in \mathscr{A}, i \in \{1, \cdots, E\}$ and some possibly unknown $M \geq 2$,

$$\varepsilon_i \perp (\varepsilon_j)_{j \notin \mathscr{N}_{i,M-1}} | A, Z,$$

where $\mathscr{N}_{i,M-1}$ denote the set of nodes connected to i by at most $M-1$ edges.[15]

Assumption 2.2.5 states that unobservables are conditionally independent for individuals that are not in a neighborhood up with a radius $M-1$.[16] Next, we discuss the causal estimands.

Definition 2.2.2 (Conditional Mean)**.** Under Assumption 2.2.4 the conditional mean function for each $i \in \{1, \cdots, E\}$, is defined as

$$m(d,s,z,l) = \mathbb{E}\Big[r(d,s,Z_i,|\mathscr{N}_i|,\varepsilon_i)\Big|Z_i = z, |\mathscr{N}_i| = l\Big]. \tag{2.8}$$

Assumption 2.2.4 is necessary in order for the conditional mean function to be identical across units. The second causal estimand of interest is the propensity score.

Definition 2.2.3. Under Condition (i) in Assumption 2.2.2, for each $i \in \{1, \cdots, E\}$ we define[17]

$$\begin{aligned} e(d,s,\mathbf{x},z,l) &= P\Big(D_i = d, \sum_{k \in \mathscr{N}_i} D_k = s \Big| Z_{k \in \mathscr{N}_i} = \mathbf{x}, Z_i = z, |\mathscr{N}_i| = l\Big) \\ &= P\Big(D_i = d | Z_i = z\Big) \sum_{u_1, \cdots, u_l : \sum_v u_v = s} \prod_{k=1}^{l} P\Big(D_{\mathscr{N}_i^{(k)}} = u_k \Big| Z_{\mathscr{N}_i^{(k)}} = \mathbf{x}^{k,\cdot}\Big). \end{aligned} \tag{2.9}$$

for $d \in \{0,1\}, s \in \mathbb{Z}, s \leq l$.

[15] Formally, such set is defined as the set $\{k : A_{i,k}^{M-1} \neq 0\}$.

[16] Local dependency graphs are often assumed in the presence of network data, see, e.g., Leung (2020).

[17] The second equaition decomposes the probability of the event into sums of probabilities of disjoint events, each corresponding to a given combination of treatment assignments $(u_1, \cdots, u_l)$, whose sum equals to s.

Definition 2.2.3 defines the probability of treatment given individual and neighbors covariates. Condition (i) in Assumption 2.2.2 guarantees that the propensity score does not depend on the index of unit i or of its neighbors. The definition of the propensity score follows from the literature on multi-valued treatments (Imbens, 2000), while here, the individuals' exposures also depend on the random number of neighbors and neighbors' covariates. Here, the number of treated neighbors $\sum_{k\in\mathcal{N}_i}\pi(X_k)$ depends on the array of covariates $X_{k\in\mathcal{N}_i}$, via the policy function π. If we were to construct the propensity score, which did not also depend on such covariates, we would not identify welfare effects.

Lemma 2.2.2 (Identification)**.** *Let Assumption 2.2.1, 2.2.2, 2.2.4 hold. Let* $S_i(\pi)=\sum_{k\in\mathcal{N}_i}\pi(X_k)$. *Then for any function* $\pi\in\Pi$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[r(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k),Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[m\Big(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k),Z_i,|\mathcal{N}_i|\Big)\Big]\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,\pi(X_i)=D_i\}}{e\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}Y_i\Big].\end{aligned}$$

The proof is in Appendix B.4.2. The above result guarantees the identification of the welfare of the experimental participants. The result does *not* require conditions on the distribution of covariates and on the dependence across unobservables.

Network Empirical Welfare Maximization

We now discuss the procedure and its theoretical properties. We defer a discussion on the optimization method to Section 2.4.

Empirical Welfare

We start introducing some notation. We let $S_i(\pi)=\sum_{k\in\mathcal{N}_i}\pi(X_k)$ be the assigned treatment to neighbors of i under policy π.

A nonparametric estimator of the welfare function is denoted as

$$W_n^{ipw}(\pi, e) = \frac{1}{n}\sum_{i=1}^{n} \frac{1\{S_i(\pi) = \sum_{k\in\mathcal{N}_i} D_k, \pi(X_i) = D_i\}}{e\Big(\pi(X_i), S_i(\pi), Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|\Big)} Y_i, \tag{2.10}$$

which also depends on the propensity score e. One possible disadvantage of the above estimator is the large variance. Therefore, we also consider the following double robust estimator (AIPW) of the welfare function of the following form:

$$\begin{aligned} W_n^{aipw}(\pi, m^c, e^c) =& \frac{1}{n}\sum_{i=1}^{n} \frac{1\{S_i(\pi) = \sum_{k\in\mathcal{N}_i} D_k, \pi(X_i) = D_i\}}{e^c\Big(\pi(X_i), S_i(\pi), Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|\Big)} \Big(Y_i - m^c\Big(\pi(X_i), S_i(\pi), Z_i, |\mathcal{N}_i|\Big)\Big) \\ &+ \frac{1}{n}\sum_{i=1}^{n} m^c\Big(\pi(X_i), S_i(\pi), Z_i, |\mathcal{N}_i|\Big), \end{aligned} \tag{2.11}$$

for some arbitrary functions (m^c, e^c) which we will refer to as the pseudo-true conditional mean and propensity score. The welfare inherits double robust properties similarly to what discussed in Robins et al. (1994), i.e. if either $m^c = m$ or $e^c = e$, the welfare is equal in expectation to Equation (2.5) (see Appendix B.4.2). Throughout our discussion, the estimated conditional mean and propensity score will be denoted as $\hat{m}$ and $\hat{e}$, respectively. The propensity score can be easily estimated as a function of conditional marginal probabilities (see Definition 2.2.3).

Known Propensity Score

We start discussing theoretical guarantees of the policy that maximizes $W_n^{aipw}(\pi, m^c, e^c)$, for some *arbitrary* functions m^c and e^c, which, only throughout this subsection, are assumed not to be data-dependent unless otherwise specified. Formally,

$$\hat{\pi}^{aipw}_{m^c,e^c} \in \arg\max_{\pi\in\Pi} W_n^{aipw}(\pi, m^c, e^c), \tag{2.12}$$

defines the "oracle" policy, which has access to the pseudo-true functions (m^c, e^c).

Assumption 2.3.1. Let the following hold:

(LP) $m^c(d,s,z,l)$ is L-Lipschitz a.e. in its second argument, for some $\infty > L \geq 0$;

(OV) For all $d \in \{0,1\}$, there exist some $\delta \in (0,1)$ such that $e^c\Big(d,s,\mathbf{x},z,l\Big) \in (\delta, 1-\delta)$ for all $s \leq l$, for all $z \in \mathscr{Z}, l \in \mathbb{Z}, \mathbf{x} \in \mathscr{Z}^l \times \mathbb{Z}^l$.

(TC) For $\Gamma_1^2, \Gamma_2^2 < \infty$, $\forall i \in \{1,...,n\}$, $\mathbb{E}\Big[\sup_{d\in\{0,1\},s\leq|\mathscr{N}_i|} r(d,s,Z_i,|\mathscr{N}_i|,\varepsilon_i)^3 \Big| A,Z\Big] < \Gamma_1^2$, and $\mathbb{E}\Big[\sup_{d\in\{0,1\}} m^c(d,0,Z_i,|\mathscr{N}_i|)^3 \Big| A,Z\Big] < \Gamma_2^2$, almost surely for each $i \in \{1,\cdots,E\}$.

(VC) π belongs to a function class of point-wise measurable functions[18] Π, where Π has finite VC dimension.[19]

Condition (LP) is satisfied for *any* bounded $m^c(.)$[20], but it also accomodates for unbounded $m^c(.)$. Remarkably, the condition is agnostic on the *true* conditional mean function $m(.)$. Condition (OV) is the usual overlap condition, often imposed in the causal inference literature. We discuss trimming at the end of this subsection. Condition (TC) imposes moment conditions on the outcome. Condition (VC) imposes restrictions on the geometric complexity of the function class of interest. It is commonly assumed in the literature on empirical welfare maximization, and examples include Kitagawa and Tetenov (2018) and Athey and Wager (2021), among others. For instance, if individuals are assigned based on a threshold crossing rule, the VC dimension equals the number of variables used in the assignment. Under condition (OV) in Assumption 2.3.1 we define δ_0 a constant such that $\max_{d\in\{0,1\}} e^c(d,0,Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|) \in (\delta_0, 1-\delta_0)$ almost surely, where $\delta_0 \geq \delta$ by definition.[21]

Covariates can be endogenous to the network and exhibit arbitrary dependence. In the following lines, we impose conditions on the distribution (and dependence) of covariates.

[18] Point-wise measurability can be replaced by measurability of each function $\pi \in \Pi$, but, in this latter case, since the pointwise supremum may not be necessarily measurable the supremum function must be interpreted as the *lattice* supremum. See Appendix B.7.

[19] The VC dimension denotes the cardinality of the largest set of points that the function π can shatter. The VC dimension is commonly used to measure the complexity of a class. See for example Devroye et al. (2013).

[20] To see why the claim hold, let $|m^c(d,S,Z_i)| < B$ a.e. Since $S \in \mathbb{Z}$, $|m^c(d,S,Z_i) - m^c(d,S',Z_i)| \leq 2B|S-S'|$.

[21] To observe why, note that under (OV) δ defines the lowest propensity score over all possible configurations of number of treated neighbors, also including the case where none of the neighbors is treated.

Assumption 2.3.2 (Covariates' distribution)**.** Assume that we can write $Z_i = h(Q_i, L_i)$, for some unknown function $h(\cdot)$, where $Q_i \in \mathscr{Q}, L_i \in \mathscr{L}$, where

$$(A) \quad (Q_j)_{j=1}^E \perp (\varepsilon_j)_{j=1}^E \Big| A, (L_j)_{j=1}^E, \quad Q_j \perp Q_{k \notin \mathscr{N}_{j,M}} | A, (L_j)_{j=1}^E, \quad (Q_j, Q_{k \in \mathscr{N}_j}) | A, (L_k)_{k=1}^E \sim \mathscr{T}(|\mathscr{N}_j|)$$

$$(B) \quad |\mathscr{L}| < \infty$$

for some possibly unknown distribution $\mathscr{T}$.

Assumption 2.3.2 decomposes covariates into two main components which are possibly *unknown* to the researcher. The first component Q_i has arbitrary support, but it is exogenous to the network structure and is locally dependent. The second component L_i can instead exhibit arbitrary dependence, but it has finite support. Intuitively, Assumption 2.3.2 states that either we have finitely many "types" of individuals, or we have infinitely many, but continuities are not predictive of the network formation. In the presence of continuous variables arbitrarily dependent on the network, the assumption holds after binning.

Example 2.3.1 (Finitely many types)**.** Suppose that $|\mathscr{Z}| < \infty$, i.e., the support of Z_i is finite dimensional. Then Assumption 2.3.2 holds. □

Let $\mathscr{N}_E = \max_{i \in \{1,\cdots,E\}} |\mathscr{N}_i| + 1$ denote the maximum degree of the network of experimental participants[22], and $\mathscr{N}_n = \max_{i \in \{1,\cdots,n\}} |\mathscr{N}_i| + 1$ the maximum degree of the sampled units. We can now state the first theorem.

Theorem 2.3.1 (Oracle Regret)**.** *Let Assumptions 2.2.1, 2.2.2, 2.2.4, 2.2.5, 2.3.1, 2.3.2 hold. Assume that either (or both)* (i) $m^c(.) = m(.)$ *and/or* (ii) *both* $e^c(.) = e(.)$*. Then,*

$$\mathbb{E}\Big[\sup_{\pi \in \Pi} W(\pi) - W(\hat{\pi}^{aipw}_{m^c,e^c})\Big] \leq \mathbb{E}[\mathscr{P}_{M,|\mathscr{L}|}(\mathscr{N}_E)] \frac{\bar{C}(L+1)(\Gamma_1+\Gamma_2)}{\delta \delta_0^{\sqrt{3}}} \sqrt{\frac{\mathrm{VC}(\Pi)}{n}} + 2\mathscr{K}_\Pi(n),$$

for a finite constant $\bar{C} < \infty$ *independent of* $(L, \Gamma_1, \Gamma_2, \delta, \delta_0, E, n)$ *and* $\mathscr{P}_{M,|\mathscr{L}|}$ *being a polynomial function with finite degree.*

[22] In our proofs, we only use information from the maximum degree of the n sampled units and their neighbors up to the *Mth* degree. Here we instead consider the maximum degree $\mathscr{N}_E$ of all experimental participants, which is larger than what is considered in our proofs, for expositional convenience only.

Corollary 6 (Known propensity score)**.** *Let* $\hat{\pi}^{ipw} \in \arg\max_{\pi\in\Pi} W_n^{ipw}(\pi, e)$ *for known propensity score and let the conditions in Theorem 2.3.1 and Lemma 2.2.1 hold. Suppose that* $|\mathcal{N}_i| < J < \infty$ *almost surely. Then*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{ipw})\Big] \leq \bar{C}'\sqrt{\mathrm{VC}(\Pi)/n}$$

for a constant $\bar{C}' < \infty$ *indendent of* (n, E).

The proof of the theorem is in Appendix B.5. Theorem 2.3.1 provides a non-asymptotic upper bound on the regret, and it is the first result of this type under network interference. The theorem is double robust to the misspecification of m^c and $1/e^c$. The corollary provides regret guarantees when researchers maximize the welfare with a known propensity score. The bound in the above corollary is distribution-free (i.e., it holds for all data-generating processes satisfying the conditions in the corollary). For bounded degree, which in our application using data from Cai et al. (2015) equals five, the regret scales at rate $1/\sqrt{n}$, which has been shown to match the maximin lower-bound in the $i.i.d.$ with no-interference (Kitagawa and Tetenov, 2018). We observe that assuming that sampled units are representative of the target population (Assumption 2.2.3) is not invoked in Theorem 2.3.1 since the regret is expressed as a function of the discrepancy between the target and sample units $\mathscr{K}_\Pi(n)$. However, it is assumed in Lemma 2.2.1 and the corollary. The proof consists of (i) controlling the Rademacher complexity in the presence of spillover effects, using extensions of contraction inequalities to accommodate the lack of sub-exponential tail decay; (ii) dealing with dependence for symmetrization, and bounding the groups of dependent units using the chromatic number.

The regret also depends on the network topology through a polynomial function of the maximum degree (see Theorem B.5.1). The polynomial function depends on the degree of dependence and on the number of endogenous types captured by $|\mathscr{L}|$ (see Assumption 2.3.2). In Section 2.3.4 we illustrate how the rate can improve in terms of the maximum degree under weaker dependence conditions.

Finally, note that a direct corollary of Theorem 2.3.1 is that the bound holds if the

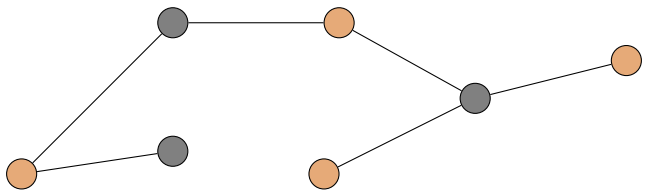


Figure 2.1. Network cross-fitting: simple illustration with $M = 1$ (note that in practice $M \geq 2$). For each unit, estimators are constructed using information from the other units assigned to the same color.

propensity score is known and the conditional mean estimated on an independent sample.

Estimated Nuisances and Network Cross-fitting

Next, we discuss regret guarantees in the presence of unknown propensity score, with the estimated policy being

$$\hat{\pi}^{aipw}_{\hat{m},\hat{e}} \in \arg\max_{\pi\in\Pi} W^{aipw}_n(\pi,\hat{m},\hat{e}). \tag{2.13}$$

A key challenge is represented by the dependence structure.

We propose a modification of the *cross-fitting* algorithm (Chernozhukov et al., 2018) to account for the local dependence of observations and a possibly fully connected network. The algorithm assumes that the researcher knows M and observes the network of the experimental participants (but not necessarily of the target units). Extensions, when the dependence structure is unknown, are discussed in Section 2.3.4. The algorithm groups individuals such that each group contains conditionally independent observations. It then estimates the nuisance function for a given observation using information from all other units in the same group.

Algorithm 4. Network Cross-Fitting

Require: Adjacency matrix A of sampled individuals; degree of dependence M.

1: Create K folds with K such that you assign to each fold individuals who are not neighbors and do not share common neighbors up to the M^{th} degree conditional on A;
2: For each unit j in fold k, estimate the conditional mean, and the propensity score used to predict j's values, using all observations *within* the fold k *only*, with the exception of j.

In the following assumption, we discuss conditions on the convergence rate of the proposed procedure.

Assumption 2.3.3. Assume that for each $d \in \{0,1\}, s \in \mathbb{Z}$, $\mathscr{A}_n \times \mathscr{B}_n = \mathscr{O}(n^{-v})$ for some $v \geq 1/2$, where

$$\begin{aligned}\mathscr{A}_n &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[\sup_{d,s}\Big(\hat{m}(d,s,Z_i,|\mathscr{N}_i|) - m(d,s,Z_i,|\mathscr{N}_i|)\Big)^2\Big]} \\ \mathscr{B}_n &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[\sup_{d,s}\Big(\frac{1}{\hat{e}(d,s,Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|)} - \frac{1}{e(d,s,Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|)}\Big)^2\Big]}.\end{aligned} \tag{2.14}$$

Assume in addition that $\hat{m}$ and $1/\hat{e}$ are uniformly bounded by a finite constant.

Under the above conditions, we can state the following theorem. The proof is contained in Appendix B.5.

Theorem 2.3.2. *Let Assumption 2.2.1, 2.2.2, 2.2.4, 2.2.5, 2.3.1, 2.3.2, 3.4.3 hold, with $m^c = m, e^c = e$. Let estimation being performed as in Algorithm 1. Then*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}_{\hat{m},\hat{e}})\Big] \leq \bar{C}'\mathbb{E}[\mathscr{P}_{M,|\mathscr{L}|}(\mathscr{N}_E)]\sqrt{\frac{\mathrm{VC}(\Pi)}{n}} + 2\mathscr{K}_\Pi(n)$$

for a constant $\bar{C}' < \infty$ independent of (n,E), and $\mathscr{P}_{M,|\mathscr{L}|}$ being a polynomial function with finite degree.

Theorem 2.3.2 shows that the rate of convergence of the estimator does not affect the convergence rate of the regret under Assumption 3.4.3. Under the conditions in Lemma 2.2.1, $\mathscr{K}_\Pi(n) = 0$, and hence the regret converges to zero as the degree grows at an appropriate slower rate than the sample size.

Remark 11 (Assumption 3.4.3 and increasing degree)**.** Assumption 3.4.3 imposes conditions on the *product* of the convergence rate of the estimator to the *true* conditional mean and propensity score function, in the same spirit of standard conditions in the i.i.d. setting (e.g., Farrell 2015). Observe that the condition in Assumption 3.4.3 is satisfied for general machine-learning estimators under bounded degree, since, in the presence of bounded degree, by Brooks (1941)'s

theorem, we can construct finitely many partitions of independent observations. Assumption 3.4.3 instead imposes a faster rate of convergence than $n^{-1/4}$ for both the propensity score and the conditional mean function, as the maximum degree is increasing. For instance, letting $M = 2$ (i.e., individuals are dependent on friends and friends of friends), the number of partitions is of order $\mathscr{N}_E^2$. In this cases, each nuisance function is estimated on a sample with $n/\mathscr{N}_E^2$ many observations, and hence Assumption 3.4.3 is satisfied, if each nuisance convergences at a rate $n^{1/4}\mathscr{N}_E^2$. This equals the parametric rate if $\mathscr{N}_E = E^{1/8}$ and $n = E$ (all participants are sampled), and can be slower than the parametric rate if $\mathscr{N}_E < E^{1/8}$. On the other hand, if the degree grows at a faster rate than $E^{1/8}$ Assumption 3.4.3 can be relaxed to hold for an arbitrary rate n^{-v} and the regret also depends on such a rate. We omit this case for the sake of brevity only. □

Three Additional Results

Next, we discuss three additional results: targeting with poor overlap, faster rates with clustered networks, estimation with a partially observed network.

Trimming to Control Overlap

Strict overlap can be violated in the presence of a few nodes with an unbounded degree. We address this challenge by proposing a trimming estimator. To guarantee overlap, we introduce the following trimming estimator:

$$W_n^{tr}(\pi;\kappa_n) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1\{S_i(\pi) = \sum_{k\in\mathscr{N}_i} D_k, \pi(X_i) = D_i\}}{e\Big(\pi(X_i), S_i(\pi), Z_{k\in\mathscr{N}_i}, Z_i, |\mathscr{N}_i|\Big)} Y_i \times 1\Big\{|\mathscr{N}_i| \leq \kappa_n\Big\}\right\}, \tag{2.15}$$

for a particular choice of κ_n which we discuss below. The above estimator accounts for the spillover effect of treating nodes with a large set of neighbors to their connections, but it excludes from estimation the direct effect on the largely connected nodes.

Theorem 2.3.3. *Let* $\hat{\pi}^{tr}_{\kappa_n} \in \arg\max_{\pi\in\Pi} W_n^{tr}(\pi;\kappa_n)$ *and suppose that* $e(d,s,\mathbf{x},z,l) \in (\delta_n, 1-\delta_n)$ *for all* $d \in \{0,1\}, \mathbf{x} \in \mathscr{X}^l, z \in \mathscr{Z}, s \leq \kappa_n, l \leq \kappa_n$, *and assume that* $P(|\mathscr{N}_i| \leq \kappa_n) > c > 0$, *for some*

$c \in (0,1)$. *Suppose that conditions in Theorem 2.3.1 hold. Suppose in addition* $Y_i \in [-B,B]$ *for a universal constant B. Let the conditions in Lemma 2.2.1 hold. Then the following holds:*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{tr}_{\kappa_n})\Big] \leq \frac{\bar{C}\mathbb{E}[\mathscr{P}_{M,|\mathscr{L}|}(\mathscr{N}_E)]B}{\delta_n}\sqrt{\frac{\mathrm{VC}(\Pi)}{n}} + 2BP(|\mathscr{N}_i| > \kappa_n)$$

for a finite constant $\bar{C} < \infty$ *independent of* (E,n).

The proof of the theorem is in Appendix B.5. Theorem 2.3.3 shows that the rate of convergence still depends polynomially on the maximum degree, but the overlap constant δ_n is potentially much larger than the worst-case constant δ. The additional price to pay is that the regret also depends on the probability that the degree exceeds a certain threshold. Such probability can be small whenever the *number* of nodes with a large degree grows at a slower rate than the sample size. A simple example is $\sqrt{n}$-many individuals (assuming $n = E$) having a growing degree greater than κ, in which case $P(|\mathscr{N}_i| > \kappa) = O(1/\sqrt{n})$.

Remark 12 (Increasing overlap by restricting the number of treatment exposures)**.** Additional modeling restrictions may be imposed to reduce the dimensionality of the problem and guarantee strict overlap (also without trimming). For example, suppose that the researcher imposes the following restriction, for some ordered τ_1, τ_2, τ_3:

$$r(d,s,z,l,e) = \begin{cases} \bar{r}_1(d,z,l,e) \text{ if } s/l \leq \tau_1 \\ \bar{r}_2(d,z,l,e) \text{ if } \tau_1 < s/l \leq \tau_2 \\ \bar{r}_3(d,z,l,e) \text{ if } \tau_2 < s/l \leq \tau_3 \end{cases} \tag{2.16}$$

for some possibly unknown functions $\bar{r}_1, \bar{r}_2, \bar{r}_3$. Then the number of possible exposures reduces to six different exposures. In this case, the balancing score defines the probability of each of this exposure, having a larger probability compared to the propensity score in Definition 2.2.3. □

Faster Rates with Growing Degree under Weaker Dependence

A natural question is whether we can achieve faster rates as a function of the degree. We formalize this in the following theorem, where we show that weaker dependence conditions guarantee a faster rate in the degree.

Theorem 2.3.4. *Let Assumption 2.2.1, 2.2.2, 2.2.4, 2.2.5, 2.3.1 hold. Suppose that* $K_i = \Big(Y_i, D_i, D_{k\in\mathscr{N}_i}, Z_i, Z_{k\in\mathscr{N}_i}, |\mathscr{N}_i|\Big) \sim \mathscr{P}$ *and that we can partition* $(K_i)_{i=1}^n$ *into K groups each containing mutually independent observations. Let either* $m^c = m$ *or* $e^c = e$*. Then*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}_{m^c,e^c})\Big] \leq \bar{C}K\sqrt{\frac{\mathbb{E}[\log(\mathscr{N}_n)\mathscr{N}_n]}{n}}$$

for a finite constant $\bar{C}$ *independent of* (n, E, K).

The proof is contained in Appendix B.5. Such a stronger result requires the additional restriction: we can construct finitely many groups of vectors K_i with units independent within each group (as in the case of many small clusters). The result shows that the degree affects the function class complexity at a rate $\sqrt{\mathbb{E}[\mathscr{N}_n^{1/2}]}$ only, while the remaining polynomial terms in Theorem 2.3.1 capture the dependence structure of the data. From an inspection of the proof, the reader may observe that the result holds regardless of whether Z_i is discrete or continuous.

Given a simpler dependence structure, for Theorem 2.3.4, the proof follows similarly to the $i.i.d.$ case of Kitagawa and Tetenov (2018) in its symmetrization argument, but with an important modification that uses contraction arguments, extended to deal under lack of sub-exponential tails' decay. To gain further intuition on its main motivation, consider first as a "loose" definition of the Rademacher complexity

$$\mathscr{R}_n(\Pi) = \mathbb{E}_\sigma\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i f_i\Big(\sum_{k\in\mathscr{N}_i}\pi(X_i), \pi(X_i)\Big)\Big|\Big] \tag{2.17}$$

for some data-dependent functions $f_i(\cdot)$ Lipschitz in their second argument, and $(\sigma_i)_{i=1}^n$ denoting

$i.i.d.$ Rademacher random variables. For example, with a bounded outcome f_i may be defined as the outcome multiplied by its corresponding inverse probability weight.

In standard $i.i.d.$ settings Equation (2.17) bounds (up-to a constant term) the regret for any realization of the data, with $f_i(\cdot, \pi(X_i))$ not depending on the first entry due to lack of interference. However, under interference, the conditional Rademacher complexity as in Equation (2.17) is not a valid upper bound of the regret due to dependence. It is, however, a valid upper bound under lack of statistical dependence among observations, used in Theorem 2.3.4.

This result illustrates the trade-off among different assumptions: when individuals exhibit arbitrary network dependence as in Theorem 2.3.1, and independence only conditional on the network, standard symmetrization arguments cannot directly be used, and the regret also depends on higher moments of the network of all experimental participants. This, instead does not occur if stronger independence conditions are imposed.

Regret with Partially Observed Networks

We conclude this section by characterizing the regret rate with M unknown, and information about the neighbors (but not of the entire network of the sampled participants) is accessible. We assume that the researcher estimates nuisance functions using the entire sample, and we characterize the rate as a function of the convergence rate of each estimator. We first impose the following condition.

Assumption 2.3.4. For some $\xi_1, \xi_2 > 0$,

$$\begin{aligned}
&\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n} \sup_{d\in\{0,1\}, s\leq|\mathcal{N}_i|} \Big|\hat{m}(d,s,Z_i,|\mathcal{N}_i|) - m^c(d,s,Z_i,|\mathcal{N}_i|)\Big|\Big] = \mathcal{O}(1/n^{\xi_1}).\\
&\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n} \sup_{d\in\{0,1\}, s\leq|\mathcal{N}_i|} \Big|\Big(Y_i - m^c(d,s,Z_i,|\mathcal{N}_i|)\Big)\Big(e^c(d,s,O_i) - \hat{e}(d,s,O_i)\Big)\Big|\Big] = \mathcal{O}(1/n^{\xi_2}).
\end{aligned} \tag{2.18}$$

where $O_i = (Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|)$. In addition, assume that $\hat{e}(\cdot) \in (\delta, 1-\delta)$ almost surely.

Assumption 2.3.4 imposes conditions on the convergence rate of $\hat{m}$ to its *pseudo*-true

value m^c and the convergence rate of $\hat{e}$ to the pseudo true e^c.[23] Here, we do not require the network cross-fitting algorithm for our results to hold, at the expense of a slower convergence rate for semi-parametric estimators.

Theorem 2.3.5. *Let Assumption 2.2.1, 2.2.2, 2.2.4, 2.2.5, 2.3.1, 2.3.2, 2.3.4 hold. Assume that either $m^c(.) = m(.)$ or $e^c(.) = e(.)$. Then,*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}_{\hat{m},\hat{e}})\Big] \leq \mathscr{O}\Big(1/n^{\xi} + \mathbb{E}[\mathscr{P}_{M,|\mathscr{L}|}(\mathscr{N}_E)]\sqrt{\frac{\mathrm{VC}(\Pi)}{n}}\Big) + 2\mathscr{K}_{\Pi}(n),$$

where $\xi = \min\{\xi_1, \xi_2\}$, for $\mathscr{P}_{M,|\mathscr{L}|}$ being a polynomial function with finite degree.

Theorem 2.3.5 provides a uniform bound on the regret, and it is double robust to correct specification of the conditional mean and the propensity score. The theorem's result depends on the convergence rate of $\hat{e}$ and $\hat{m}$ to their *pseudo*-true value. For parametric estimators of the conditional mean and the propensity score and bounded degree, the regret bounds scale at rate $1/\sqrt{n}$. However, for the general machine-learning estimator, the rate can be slower than the parametric one, reflecting the "cost" of the lack of knowledge of the degree of dependence M.

Mixed Integer Linear Program Formulation

In this section, we discuss the optimization procedure. Firstly, we define the estimated effect of assigning to unit i treatment d, after treating s of its neighbors. For the double robust estimator, this quantity is defined as

$$g_i(d,s) = \frac{1\{\sum_{k\in\mathscr{N}_i} D_k = s, D_i = d\}}{e^c\Big(d,s,Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|\Big)}\Big(Y_i - m^c\Big(d,s,Z_i,|\mathscr{N}_i|\Big)\Big) + m^c\Big(d,s,Z_i,|\mathscr{N}_i|\Big), \tag{2.19}$$

where we omit the dependence of $g_i(\cdot)$ with m^c and e^c for sake of brevity. Secondly, we define $I_i(\pi,h) = 1\Big\{\sum_{k\in\mathscr{N}_i}\pi(X_k) = h\Big\}$ the indicator of whether h neighbors of individual i have been

[23] Convergence rate for penalized regression on networks is found in He and Song (2018) among others. Estimation via the method of moments that satisfy the high-level conditions in Hansen (1982) also guarantees parametric convergence rates $1/\sqrt{n}$ of the estimators of interest.

treated under policy π. We observe that the following holds:

$$\sum_{h=0}^{|\mathcal{N}_i|}\Big\{\Big(g_i(1,h)-g_i(0,h)\Big)\pi(X_i)I_i(\pi,h)+I_i(\pi,h)g_i(0,h)\Big\}=g_i\Big(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k)\Big). \quad (2.20)$$

Namely, the estimated treatment effect on unit i, obtained after implementing policy π, is the sum of effects obtained by treating zero to all neighbors of i. Each element in the sum is weighted by the indicator $I_i(\pi,h)$, and only one of these indicators is equal to one. We can then define n variables p_i that denote the treatment assignment of each unit after restricting the policy function in the function class of interest. Namely, we let $p_i=\pi(X_i),\pi\in\Pi$. For example, for policy functions of the form $\pi(X_i)=1\{X_i^\top\beta\geq 0\},\beta\in\mathscr{B}$, similarly to Kitagawa and Tetenov (2018) we write

$$\frac{X_i^\top\beta}{|C_i|}<p_i\leq\frac{X_i^\top\beta}{|C_i|}+1,\quad C_i>\sup_{\beta\in\mathscr{B}}|X_i^\top\beta|,\quad p_i\in\{0,1\},$$

where p_i is equal to one if $X_i^\top\beta$ is positive and zero otherwise. The key intuition is to introduce additional decision variables to represent $I_i(\pi,h)$ through linear constraints. We define the following variables:

$$t_{i,h,1}=1\Big\{\sum_{k\in\mathcal{N}_i}p_k\geq h\Big\},\quad t_{i,h,2}=1\Big\{\sum_{k\in\mathcal{N}_i}p_k\leq h\Big\},\quad h\in\{0,\cdots,|\mathcal{N}_i|\}.$$

The first variable is one if at least h neighbors are treated, and the second variable is one if at most h neighbors are treated. The goal is to achieve a linear representation of such variables. Observe that

$$t_{i,h,1}+t_{i,h,2}=\begin{cases}1 \text{ if and only if } \sum_{k\in\mathcal{N}_i}p_k\neq h\\ 2 \text{ otherwise}\end{cases}\quad\Rightarrow\quad t_{i,h,1}+t_{i,h,2}-1=I_i(\pi,h). \quad (2.21)$$

Therefore, we can define $t_{i,h,1},t_{i,h,2}$ using mixed-integer linear constraints. Namely, the variable

$t_{i,h,1}$ can equivalently be defined as[24]

$$\frac{(\sum_k A_{i,k}p_k - h)}{|\mathscr{N}_i|+1} < t_{i,h,1} \leq \frac{(\sum_k A_{i,k}p_k - h)}{|\mathscr{N}_i|+1} + 1, \quad t_{i,h,1} \in \{0,1\}. \tag{2.22}$$

Similar reasoning follows for $t_{i,h,2}$. We now can write the objective function as

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{h=0}^{|\mathscr{N}_i|}\Big\{\Big(g_i(1,h)-g_i(0,h)\Big)p_i(t_{i,h,1}+t_{i,h,2}-1)+(t_{i,h,1}+t_{i,h,2}-1)g_i(0,h)\Big\}. \tag{2.23}$$

The above objective function leads to a quadratic mixed integer program. On the other hand, quadratic programs can be computationally expensive to solve. We write the problem as a mixed-integer linear program introducing one additional set variables, that we call $u_{i,h}$ for $h \in \{0,\cdots,|\mathscr{N}_i|\}$, with $u_{i,h} = p_i(t_{i,h,1}+t_{i,h,2}-1)$. We provide the complete formulation below.

$$\max_{\{u_{i,h}\},\{p_i\},\{t_{i,1,h},t_{i,2,h}\},\beta\in\mathscr{B}} \frac{1}{n}\sum_{i=1}^{n}\sum_{h=0}^{|\mathscr{N}_i|}\Big\{\Big(g_i(1,h)-g_i(0,h)\Big)u_{i,h}+g_i(0,h)(t_{i,h,1}+t_{i,h,2}-1)\Big\} \tag{2.24}$$

under the constraints:[25]

$$\begin{aligned}
&(A) \quad p_i = \pi(X_i), \quad \pi \in \Pi \\
&(B) \quad \frac{p_i+t_{i,h,1}+t_{i,h,2}}{3} - 1 < u_{i,h} \leq \frac{p_i+t_{i,h,1}+t_{i,h,2}}{3}, \quad u_{i,h} \in \{0,1\} \quad \forall h \in \{0,\cdots,|\mathscr{N}_i|\}, \\
&(C) \quad \frac{(\sum_k A_{i,k}p_k - h)}{|\mathscr{N}_i|+1} < t_{i,h,1} \leq \frac{(\sum_k A_{i,k}p_k - h)}{|\mathscr{N}_i|+1} + 1, \quad t_{i,h,1} \in \{0,1\}, \quad \forall h \in \{0,\cdots,|\mathscr{N}_i|\} \\
&(D) \quad \frac{(h - \sum_k A_{i,k}p_k)}{|\mathscr{N}_i|+1} < t_{i,h,2} \leq \frac{(h - \sum_k A_{i,k}p_k)}{|\mathscr{N}_i|+1} + 1, \quad t_{i,h,2} \in \{0,1\}, \quad \forall h \in \{0,\cdots,|\mathscr{N}_i|\}.
\end{aligned} \tag{2.25}$$

[24]The equation holds for the following reason. Suppose that $h < \sum_k A_{i,k}p_k$. Since $\frac{(\sum_k A_{i,k}p_k - h)}{|\mathscr{N}_i|+1} < 0$, the left-hand side of the inequality is negative and the right hand side is positive and strictly smaller than one. Since $t_{i,h,1}$ is constrained to be either zero or one, in the latter case, it equals zero. Suppose now that $h \geq \sum_k A_{i,k}p_k$. Then the left-hand side is bounded from below by zero, and the right-hand side is bounded from below by one. Therefore $t_{i,h,1}$ is set to be one.

[25]To motivate the constraint for $u_{i,h}$, notice first that we can write $p_i(t_{i,h,1}+t_{i,h,2}-1) = p_i \times t_{i,h,1} \times t_{i,h,2}$ since $(t_{i,h,1}+t_{i,h,2}-1)$ is equal to one if both variables are ones and zero if either of the two variables are ones and the other is zero. The case where both variables are zero never occurs by construction. Therefore we can write $\frac{p_i+t_{i,h,1}+t_{i,h,2}}{3} - 1 < u_{i,h} \leq \frac{p_i+t_{i,h,1}+t_{i,h,2}}{3}, \quad u_{i,h} \in \{0,1\}$.

The first constraint can be replaced by methods discussed in previous literature such as maximum scores (Florios and Skouras, 2008), while the additional constraints are justified by the presence of interference.[26] In the presence of capacity constraints, the problem can be formulated as above after adding additional linear constraints on the maximum number of treated units. Whenever units have no neighbors, the objective function is proportional to the one discussed in Kitagawa and Tetenov (2018) under no interference.[27] Therefore, the formulation provided generalizes the MILP formulation to the case of interference.[28]

Theorem 2.4.1. *The π^* solves the optimization problem in Equation* (2.24) *under the constraints in Equation* (2.25) *if and only if $\pi^* \in argmax_{\pi\in\Pi} W_n^{aipw}(\pi, m^c, e^c)$.*

Theorem 2.4.1 is a direct consequence of the argument in the current section, and it permits solving the optimization problem over the function class Π using off-the-shelf methods.

Empirical Application

We now illustrate the proposed method using data originated from Cai et al. (2015). The authors study the effect of an information session on insurance adoption in 47 villages in China, documenting (i) positive spillover effects resulting from direct treatments to neighbors; (ii) absence of endogenous spillovers.[29] To evaluate our procedure's performance, we "simulate" the following environment: researchers collect information on the first 25 villages. We estimate the policy to target the individuals in the remaining villages, which in total are 22. In the remaining villages, we assume that the policy-maker does not have access to the network information. Throughout our discussion, we consider the population of individuals having at

[26] In practice, we observe that including additional (superfluous) constraints stabilizes the optimization problem. These are $\sum_h (t_{i,h,1} + t_{i,h,2} - 1) = 1$ for each i and $\sum_i \sum_h u_{i,h} = \sum_i p_i$.

[27] This follows from the fact that under no interference the second component in the objective function is constant and the first component only depends on the individual treatment allocation.

[28] Also, observe that the formulation differs from those provided for allocation of an individual into small peer groups (Li et al., 2019) since the latter case does not account for the individualized treatment assignments, encoded in the constraints (A)-(D), and in the variables in the objective function $t_{i,h}$.

[29] Endogenous spillovers define the effect of increasing insurance take-up as a function of the purchase decision of direct neighbors.

least one neighbor.[30] Individuals are connected under a "strong" adjacency matrix, whose edges are equal to one if both individuals of a given pair indicated the other as a connection. The training set contains $n = 1315$ observations, and the maximum degree is bounded by five. The test set has 1401 units. For simplicity, we assume full compliance with the treatment.[31]

The outcome of interest is insurance adoption, and it is binary. While the experiment has more than two arms, we only focus on the effects of assigning individuals to intensive information sessions for simplicity. The intensive information sessions were randomized at the household level. The experiment of Cai et al. (2015) consists of two rounds: two consecutive sets of information sessions were performed within a few days. The authors assume that spillovers occur only to individuals participating in the second information session. To capture these effects, we estimate the policy function using an asymmetric adjacency matrix, where individuals participating in the first round of information sessions have no incoming edges. We evaluate the performance of the remaining villages using the true population adjacency matrix.

Estimation of the conditional mean function is performed non-parametrically using Random Forest (Breiman, 2001).[32] We maximize welfare using the double robust estimator. We estimate the individual probability of treatment as in Equation (2.9) using logistic regression.[33] We compare the methods using the estimated doubly-robust estimator on observations from the remaining villages.[34]

[30] This follows from the fact that the optimization problem over individuals having no connections can be treated as a separate problem.

[31] In the experiment, more than 90% of farmers attended the sessions.

[32] The function depends on the percentage of treated neighbors, the individual treatment assignments, and their interaction. We also condition on age, rice area, risk aversion, their interactions with the treatment assignments as well as gender, age, literacy level, the index of risk aversion, the probability of a climate disaster, education, and the number of friends of the participant.

[33] The same covariates used for the conditional mean function are also used for the propensity score.

[34] To guarantee overlap, we trim the propensity score whenever the joint probability of individual and neighbors' treatment is below 5% when evaluating the methods out-of-sample. Results are robust if we choose 2% trimming. For trimmed units, extrapolation using the conditional mean function is performed, which guarantees that we evaluate the performance on all the 1401 units.

Table 2.1. Application: degree distribution. Two individuals are friends if both indicate the other as a friend.

1	2	3	4	5
53.3 %	30.3 %	12.2 %	2.99 %	0.92 %

We consider a linear policy rule of the following form:

$$\pi(X_i) = 1\Big\{\beta_0 + \text{age}\beta_1 + \text{rice\_area}\beta_2 + \text{risk\_adversion}\beta_3 \geq 0\Big\}. \tag{2.26}$$

We study the effect with four different levels of capacity constraints, namely $20\%, 30\%, 40\%$ of individuals are treated.[35] We compare the proposed method to four competitors: (i) the EWM rule discussed in Kitagawa and Tetenov (2018), with estimated propensity score of individual treatment (i.e., it ignores network effects), and a policy function class as in Equation (2.26); (ii) the doubly robust method of Athey and Wager (2021) where, however, the conditional mean function also controls for the network information; (iii) the method that targets at random the same number of individuals as NEWM.[36] In Table 2.1 we report the percentage of individuals with a certain number of neighbors. The strong network presents a small degree that facilitates computations but may *under*-estimate spillover effects.

We collect results in Table 2.2. In the table, we report the Feasible NEWM with policy as in Equation (2.26) and the best competitor among a random allocation, EWM with propensity score and EWM with also a regression adjustment. Table 2.2 also collects results for the "oracle" NEWM method. This method consists of maximizing the double-robust welfare over each

[35] We also compute results for capacities 50% whose results are comparable to the results with 40% constraint. We omit those since the optimization of feasible NEWM does not achieve a zero dual gap within the time constraint for 50% constraint. Also, whenever the number of treated units on the target sample exceeds the capacity constraints, we treat the same number of units as the capacity constraint, treating those with the largest estimated score $X_i^\top \widehat{\beta}$.

[36] We estimate the model using a MILP program as in Kitagawa and Tetenov (2018) for the EWM methods and as in the main text for the NEWM method. The dual gap of each estimated method is zero, with the exception of the oracle method. For the oracle method, since we estimate different policies for different villages, we incur a dual gap over a few villages. However, as noted in Table 2.2, the dual gap of the oracle method does not affect its performance relative to the other methods.

village on the *target* sample, estimating a village-specific linear decision rule which also depends on the *number of neighbors* of each individual. The method is named "oracle" since it has access to the outcomes and the network information from the target villages. Welfare is measured as 12 RMB per mu per season.[37]

We observe that the proposed method uniformly outperforms those methods that do not account for spillover effects. The improvement is up to twelve percentage points. This can be economically relevant once the policy is implemented at scale. The method underperforms relative to the oracle method since the feasible NEWM estimator does not directly use information from the target villages other than the age, rice area, and risk aversion of each individual. In Figure 2.2, we compare the oracle method to the feasible method in terms of treatment probabilities. Two facts are worth noticing. First, (i) the assignments under the oracle method positively correlate with the degree of individuals. However, the method does *not* treat all the units with the largest degree; instead, it balances treatments across different sub-populations to exploit heterogeneity in treatment effects. Second, (ii) the feasible method treats more individuals with the largest degree, although network information is not used by the policy function. The figure shows how the NEWM method exploits information on the dependence between the degree and observable covariates from in-sample units for best targeting individuals when network information is not directly accessible by the policy-maker. In Table 2.3, we report the estimated policy function's coefficients. We observe a positive dependence of the optimal treatment rule with the risk aversion and the rice area of the individual and a negative dependence with age.

In Appendix B.2 we include a numerical study that shows that our procedure uniformly outperforms those methods that ignore network effects.

[37] 1 RMB is 0.15$ and one mu is 0.067 hectare

Table 2.2. Application: out-of-sample welfare comparisons. Welfare is measured as the probability of insurance adoption times insurance premium. Feas NEWM stands for feasible NEWM. Best competitor reports the welfare among the random assignment EWM with propensity score and double-robust EWM. Oracle is the NEWM that estimates different policies in the target villages, having access to outcomes and network information in those villages. Third column reports the welfare improvement between the feasible NEWM and the best competitor. Different rows correspond to the case where capacity constraints. 90% lower bound denotes the lower bound constructed using cluster robust standard errors.

	Feas NEWM	Best Comp	Welfare Imp	Oracle NEWM
Welf: Capacity 0.2	2.168	2.118	2%	4.219
90% *Lower bound*	1.18	1.13		
Welf: Capacity 0.3	2.574	2.299	12%	4.842
90% *Lower bound*	2.23	1.92		
Welf: Capacity 0.4	2.781	2.492	12%	5.363
90% *Lower bound*	1.96	1.87		

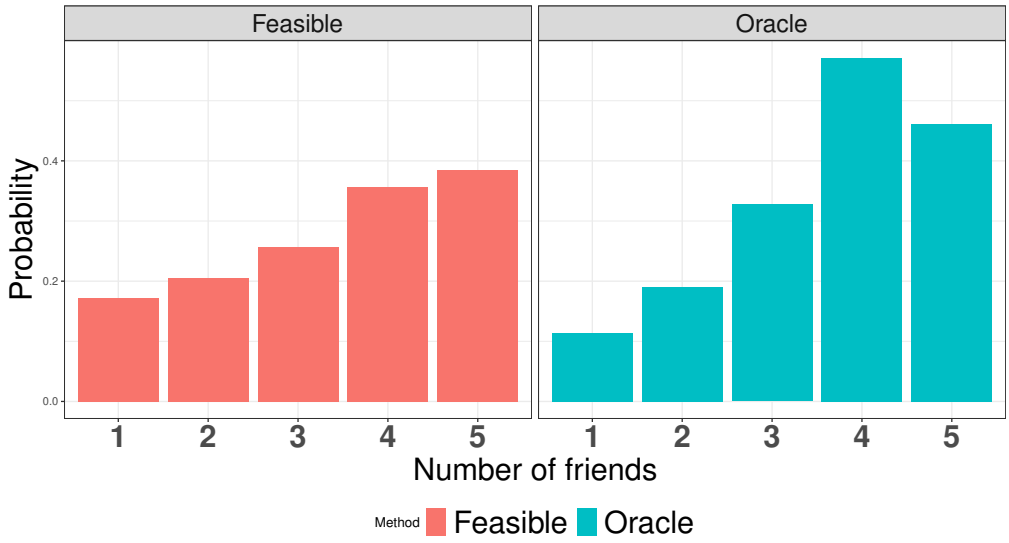


Figure 2.2. Treatment probability as a function of the number of neighbors. The plot reports the probability of being treated as a function of the number of neighbors when 20% of individuals are treated. In red the feasible NEWM and in blue the oracle.

Table 2.3. Application: estimated coefficients of the policy function. Feasible NEWM, with coefficients rescaled by the size of the risk-adversion coefficient.

	Intercept	Age	Rice Area	Risk Adversion
Capacity 0.2	6.944	-1	2.204	1
Capacity 0.3	-0.919	-0.426	1.512	1
Capacity 0.4	0.594	-0.132	0.330	1

Conclusions

In this chapter, we have introduced a method for estimating treatment rules under network interference. We consider constrained environments, and we accommodate policy functions that do not necessarily depend on network information. The proposed methodology is valid for a generic class of network formation models, and it relies on semi-parametric estimators. We cast the problem into a mixed-integer linear program and derive guarantees on the regret.

Our method assumes anonymous and exogenous interactions. Future research can address the case of endogenous interactions by explicitly modeling the endogenous component. Similarly, an avenue for future research is to replace the exogeneity of the network formation with assumptions on the (endogenous) network formation process.

Finally, the literature on influence maximization has often relied on structural models, while this chapter has focused on semiparametric estimation procedures. Exploring the trade-offs of these different approaches remains an open research question.

Chapter 2 is currently being prepared for submission for publication of the material. Davide Viviano is the sole author of this material.

Chapter 3

Fair Policy Targeting

Introduction

In previous chapters, we have shown how researchers can estimate treatment allocation rules from experiments or quasi-experiments. However, one of the major concerns of targeting interventions on individuals is discrimination: individualized treatments may induce disparities in sensitive attributes such as age, gender, or race. Motivated by evidence of policymakers' preferences towards non-discriminatory actions (Cowgill and Tucker, 2019), this chapter designs fair and efficient targeting rules for applications in social welfare and health programs. We construct treatment allocation rules using data from experiments or quasi-experiments, where, here we assume $i.i.d.$ sampling, and we develop policies that trade-off efficiency and fairness.

Fair targeting is a controversial task due to the lack of consensus on the formulation of the decision problem. Conventional approaches mostly developed in computer science consist in designing algorithmic decisions that maximize the *expected* utility across all individuals by imposing fairness constraints on the decision space of the policymaker (Nabi et al., 2019).[1] In contrast, the economic literature has outlined the importance of taking into account the welfare effects of such policies (Kleinberg et al., 2018). Fairness constraints on the policymaker's decision space may ultimately lead to sub-optimal welfare for both sensitive groups. This is a significant limitation when the policymakers are concerned with the effects of their decisions

[1] For a review, the reader may refer to Corbett-Davies and Goel (2018).

on each individual's utilities: absent legal constraints, we may not want to impose unnecessary constraints on the policy if such constraints are *harmful* for some or all individuals.

This chapter studies fair and Pareto optimal treatment rules. We discuss targeting in a setting where decision-makers prefer allocations for which we cannot find any other policy that strictly improves welfare for one of the two sensitive groups without decreasing welfare on the opposite group. Within such a set, she then chooses the fairest allocation. The decision problem is conceived for applications in social welfare and health programs and motivated by the Hippocratic notion of "first do no harm" ("*primum non-nocere*") (Rotblat, 1999): instead of imposing possibly harmful fairness constraints on the decision space, we restrict the set of admissible solutions to the Pareto optimal set, and among such, we choose the fairest one. For example, during a health-program campaign, the policymakers may not be willing to decrease all individuals' health status to gain fairness. Instead, they may be willing to *trade-off* health status of different groups (e.g., young and old individuals) when considering fairness. The framework that we propose has three desirable properties: (i) it applies to general notions of fairness which may reflect different decision makers' preferences; (ii) it guarantees Pareto efficiency of the policy function, with the relative importance of each group solely chosen based on the notion of fairness adopted by the decision-maker; (iii) it also allows for arbitrary legal or ethical constraints, incorporating as a special case the presence of fairness constraints whenever such constraints are *binding* on policymakers' decisions due to ethical or legal considerations.[2] We name our method Fair Policy Targeting.

We contribute to the statistical treatment choice literature by introducing the notion, estimation procedures, and studying properties of Pareto optimal and fair treatment allocation rules. We allow for general notions of fairness, and as a contribution of independent interest, we define envy-freeness fairness (Varian, 1976) within the context of policy targeting.

The decision problem consists of lexicographic preferences of the policymaker of the

[2]In the presence of binding fairness constraints, our proposed policy achieves a lower unfairness compared to the policy that maximizes welfare under fairness constraints while being Pareto optimal. See Section 3.2.4 for details.

following form: (i) Pareto dominant allocations are preferred over dominated ones; (ii) Pareto optimal allocations are ranked based on fairness considerations. We identify the Pareto frontier as the set of maximizers over any weighted average of each group's welfares. Therefore, such an approach embeds as a special case maximizing a weighted combination of welfares of each sensitive group such as in Athey and Wager (2021), Kitagawa and Tetenov (2018)[3], and in Rambachan et al. (2020). The above references take a specific weighted combination of welfares with weights as given, while in our case, weights are part of the decision problem and are directly selected to maximize fairness. This has important practical implications: our procedure is solely based on the notion of fairness adopted by the social planner, and it does not require specific importance weights assigned to each sensitive group, which would be hard to justify to the general public.

Estimating the set of Pareto optimal allocations represents a fundamental challenge since (i) the set consists of maximizers over a continuum of weights between zero and one; (ii) each maximizer of the welfare (or a weighted combination of welfares) is often not unique (Elliott and Lieli, 2013). To overcome these issues, we show that the Pareto frontier can be approximated using simple linear constraints. We use a discretization argument, and we evaluate weighted combinations of the objective functions separately to construct a polyhedron that contains Pareto allocations. Our approach drastically simplifies the optimization algorithm: instead of estimating the entire set of Pareto allocations, we maximize fairness under easy-to-implement linear constraints. We provide theoretical guarantees on our approach, and we show that the distance between the Pareto frontier obtained via linear constraints and its population counterpart converges uniformly to zero at a rate $1/\sqrt{n}$.

We study regret guarantees, i.e., the difference between the estimated policy function's expected unfairness against the minimal possible unfairness achieved by Pareto optimal allocations. We characterize the rate under high-level conditions for general notions of unfairness and

[3]Under the utilitarian perspective considered in Kitagawa and Tetenov (2018), Athey and Wager (2021), the welfare maximization problem is equivalent to maximizing a weighted combination of the welfare of different groups with weights equal to corresponding probabilities. See Section 3.2 for more discussion.

derive upper bounds that scale at rate $1/\sqrt{n}$, in several examples, and a lower bound that matches the same rate. We conclude with an application and a calibrated numerical study on targeting student awards and discuss the advantages of the proposed method compared to alternatives that ignore Pareto optimality.

This chapter relates to a growing literature on statistical treatment rules (Sun, 2020; Manski, 2004; Athey and Wager, 2021; Armstrong and Shen, 2015; Bhattacharya and Dupas, 2012; Hirano and Porter, 2009; Kitagawa and Tetenov, 2018, 2019; Mbakop and Tabord-Meehan, 2021; Stoye, 2012; Tetenov, 2012; Viviano, 2019; Zhou et al., 2018). Further connections are also related to the literature on classification (Elliott and Lieli, 2013). However, none of these discuss the design of fair and Pareto optimal decisions.

Fairness is a rising concern in economics, see Cowgill and Tucker (2019), Kleinberg et al. (2018), Rambachan et al. (2020). The authors provide economic insights on the characteristics of optimal decision rules when discrimination bias occurs. Here, we answer the different questions about the design and estimation of the optimal targeting rule within a statistical framework and derive the method's properties. A further difference is the decision problem with a multi-objective, instead of a single-objective utility function, as in previous references.

Additional references include Kasy and Abebe (2020) that provide comparative statics on the impact of fairness on the individuals' welfare, focusing on the analysis of algorithms, and Narita (2021) who motivates fairness based on incentive compatibility in the different contexts of the design of experiments.

In computer science, Pareto optimality has been considered in the context of binary predictions by Balashankar et al. (2019) and Martinez et al. (2019). The authors propose semi-heuristic and computationally intensive procedures for estimating Pareto efficient classifiers. Xiao et al. (2017) discuss the different problems of estimation of a Pareto allocation that trade-offs fairness and individual utilities for recommender systems, where the relative importance weights of the different objectives are selected a-priori. These references do not address the treatment choice problem discussed in the current paper.

References in computer science include Chouldechova (2017), Dwork et al. (2012), Hardt et al. (2016) among others. Corbett-Davies and Goel (2018) contain a review. Additional work also includes Liu et al. (2017) who discuss fair bandits, and Ustun et al. (2019) who propose decoupled estimation of tree classifiers without allowing for exogenous (legal or economic) constraints on the policy space. While the above references address the decision problem as a prediction problem, several papers discuss algorithmic fairness within a causal framework (Coston et al., 2020; Kilbertus et al., 2017; Nabi et al., 2019; Kusner et al., 2019). All such papers estimate decision rules under fairness constraints without discussing Pareto optimality. The different decision problem considered here is motivated by applications in social welfare and health programs. We show that when not binding on policy-makers decisions, fairness constraints may lead to Pareto-dominated allocations and possibly harmful policies for advantaged and disadvantaged individuals. When fairness constraints are binding, instead, the decision problem proposed in this paper leads to fairer allocations compared to a constrained welfare maximization problem while not being Pareto dominated.

Decision Making and Fairness

We start by introducing some notation. For each unit, we denote with $S \in \mathscr{S}$ a sensitive or protected attribute. For expositional convenience, we let $\mathscr{S} = \{0,1\}$, with $S = 1$ denoting the disadvantaged group, and $X \in \mathscr{X} \subseteq \mathbb{R}^p$ individual characteristics. We define the post-treatment outcome with $Y \in \mathscr{Y} \subseteq \mathbb{R}$ realized only once the sensitive attribute, covariates, and the treatment assignment are realized. We define $Y(d)$, $d \in \{0,1\}$ the potential outcomes under treatment d. The observed Y satisfies the Single Unit Treatment Value Assumption (SUTVA) (Rubin, 1990). The vector of potential outcomes, covariates, sensitive attributes and treatment assignments, (Y_i, D_i, S_i, X_i) are $i.i.d.$. Let

$$e(x,s) = P(D = 1|X = x, S = s), \quad p_1 = P(S = 1) \tag{3.1}$$

be the propensity score and the probability of being assigned to the disadvantaged group. Finally, we assume that treatments are randomized independently of potential outcomes (Imbens and Rubin, 2015).

Assumption 3.2.1 (Treatment Unconfoundedness)**.** For $d \in \{0,1\}$, $Y(d) \perp D|X,S$.

Social Welfare

Given observables, (Y_i, X_i, D_i, S_i) we seek to design a treatment assignment rule (i.e. policy function) $\pi : \mathscr{X} \times \mathscr{S} \mapsto \mathscr{T} \subseteq [0,1], \pi \in \Pi$ that depends on the individual characteristics and protected attributes, and which can be either probabilistic or deterministic.[4] Here, Π incorporates given and binding legal or economic constraints that restrict the policymaker's decision space. The welfare generated by a policy π on those individuals with sensitive attribute $S = s$ is defined as[5]

$$W_s(\pi) = \mathbb{E}\Big[(Y(1) - Y(0))\pi(X,S)\Big|S = s\Big]. \tag{3.2}$$

Under the utilitarian perspective (Manski, 2004), the welfare maximization problem, i.e., the population counterpart of the empirical welfare maximization (EWM) (Kitagawa and Tetenov, 2018), solves

$$\max_{\pi \in \Pi} \Big\{ p_1 W_1(\pi) + (1 - p_1) W_0(\pi) \Big\}$$

where p_1 is defined as in Equation (3.1). However, whenever the sensitive group is a *minority* group, welfare maximization assigns a small weight to the welfare of the minority, disproportionally favoring the majority group. An alternative approach is to maximize the welfare separately for each possible sensitive group by designing different policies for different groups (Ustun et al., 2019). This approach may violate discriminatory laws, i.e., the resulting policy function violates the constraint in Π. A simple example is when, due to legal reasons, the policy $\pi(x,s)$ must be constant in the sensitive attribute s. Instead, we consider a framework where the policymaker

[4] It is deterministic if $\mathscr{T} = \{0,1\}$ and probabilistic if $\mathscr{T} = [0,1]$.

[5] Welfare is interpreted from an *intention-to-treat* perspective similarly to Kitagawa and Tetenov (2018), Athey and Wager (2021).

simultaneously maximizes each group's welfare, imposing Pareto efficiency on the estimated policy under arbitrary legal or economic constraints encoded in Π. Given the set of efficient policies, the planner then selects the least unfair one. Our approach is designed for social and welfare programs where legal constraints naturally occur and where, given such constraints, the policymaker's preferences align with classical notions of "first do no harm".

Pareto Principle for Treatment Rules

The set of Pareto optimal choices is defined as Π_o, and it contains all such allocations $\pi \in \Pi$ for which the welfare for one of the two groups cannot be improved without reducing the welfare for the opposite group. We characterize Π_o in the following lemma.

Lemma 3.2.1 (Pareto Frontier)**.** *The set* $\Pi_o \subseteq \Pi$ *is such that*

$$\Pi_o = \Big\{ \pi_\alpha : \pi_\alpha \in \arg\sup_{\pi \in \Pi} \alpha W_1(\pi) + (1-\alpha) W_0(\pi), \quad \alpha \in (0,1) \Big\}. \tag{3.3}$$

The lemma follows directly from Negishi (1960), whose proof is included in Appendix C.3. It will be convenient to define

$$\bar{W}_\alpha = \sup_{\pi \in \Pi} \alpha W_1(\pi) + (1-\alpha) W_0(\pi), \tag{3.4}$$

the largest value of the objective in Equation (3.3) for a fixed α. In the following examples, we show that Pareto allocations *generalize* notions of optimal treatment rules from previous literature.

Example 3.2.1 (Welfare Maximization)**.** The population equivalent of the EWM problem belongs to the Pareto frontier. Namely,

$$\arg\max_{\pi \in \Pi} \Big\{ p_1 W_1(\pi) + (1-p_1) W_0(\pi) \Big\} \subseteq \Pi_o.$$

An alternative approach consists in maximizing weighted combinations of the welfare with the weights for each group as given. For instance the allocation (Rambachan et al., 2020)

$$\check{\pi}_{\omega} \in \arg\max_{\pi \in \Pi} \Big\{ \omega W_1(\pi) + (1-\omega) W_0(\pi) \Big\} \subseteq \Pi_{\circ}, \tag{3.5}$$

for some *specific* weight ω belongs to the Pareto frontier. □

Pareto optimal allocations are often non-unique, allowing for flexibility in the choice of efficient policies. The policy-maker must appeal to some preferential ranking principle based on her preferences. We discuss those in the following lines.

Decision Problem

We start by defining $\mathscr{C}(\Pi)$ the *choice set* of the policy maker (Mas-Colell et al., 1995), where $\mathscr{C}$ is a choice function with $\mathscr{C}(\{\pi_1, \pi_2\}) = \pi_1$ if π_1 is strictly preferred to π_2. We let

$$\text{UnFairness} : \Pi \mapsto \mathbb{R} \tag{3.6}$$

an operator which quantifies the unfairness of a policy. We leave unspecified UnFairness and provide examples in Section 3.4.3 and Section 3.5. We now state the planner's preferences.

Assumption 3.2.2 (Policy-maker's Preferences)**.** Preferences are rational[6] with:
(i) $\mathscr{C}(\{\pi_1, \pi_2\}) = \pi_1$ if $W_1(\pi_1) \geq W_1(\pi_2)$ and $W_0(\pi_1) \geq W_0(\pi_2)$ and either (or both) of the two inequalities hold strictly; (ii) for each π_1, π_2, where neither π_1 Pareto dominates π_2 nor π_2 Pareto dominates π_1, $\mathscr{C}(\{\pi_1, \pi_2\}) = \pi_1$ if UnFairness(π_1) $<$ UnFairness(π_2); (iii) for each π_1, π_2 , where neither Pareto dominates the other and with equal UnFairness, $\mathscr{C}(\{\pi_1, \pi_2\}) = \{\pi_1, \pi_2\}$.

Assumption 3.2.2 postulates lexicographic preferences of the following form: (i) an allocation is strictly preferred to another if it weakly improves welfare for both groups and

[6]Rational preferences imply transitivity and completeness (Mas-Colell et al., 1995).

strictly improves welfare for at least one group; (ii) given two allocations where none of the two Pareto dominates the other, allocations are ranked based on fairness.

Proposition 3.2.2 (Decision Problem)**.** *Under Assumption 3.2.2,* $\pi^\star \in \mathscr{C}(\Pi)$ *if and only if*

$$\begin{aligned} &\pi^\star \in \arg\inf_{\pi\in\Pi} \text{UnFairness}(\pi) \\ &\textit{subject to } \alpha W_1(\pi)+(1-\alpha)W_0(\pi) \geq \bar{W}_\alpha, \textit{ for some } \alpha \in (0,1). \end{aligned} \tag{3.7}$$

The proof is contained in Appendix C.3. Proposition 3.2.2 formally characterizes the policy-makers decision problem, which consists of minimizing the policy's unfairness criterion under the condition that the policy is Pareto optimal. The policy-maker does not maximize a weighted combination of welfares, with some *pre-specified* and hard-to-justify weights. Instead, each group's importance (i.e., α) is implicitly chosen within the optimization problem to maximize fairness. This approach allows for a transparent choice of the policy based on the policymakers' definition of fairness.

Example 3.2.2 (Why Pareto Efficiency? A simple example)**.** Let $X = 1$ for simplicity, take $\tau_s, \phi \in (0,1), s \in \{0,1\}$ and let $Y(d) = \tau_S d + \varepsilon(d)$, with $\mathbb{E}[\varepsilon(d)|S] = 0$. Consider a class of probabilistic decision rules

$$\Pi = \Big\{\pi(x,s) = \beta_s, \quad \beta_1, \beta_0 \in (0,1), \quad \beta_0 p_0 + \beta_1 p_1 \leq \phi\Big\},$$

with the share of treated units being at most ϕ. Let UnFairness be the difference in the groups' welfares, namely $|\tau_1\beta_1 - \tau_0\beta_0|$. The smallest possible unfairness is zero, since we can choose $\beta_1 = \beta_0 = 0$ with one of the fairest allocation selecting none of the individuals to treatment. Consider now the Pareto frontier, defined as:

$$\Pi_\circ = \Big\{\pi(x,s) = \beta_s^*, \quad \beta_0^* = \frac{\phi - p_1\beta_1^*}{p_0}, \beta_1^* \in [0,1]\Big\} \subset \Pi. \tag{3.8}$$

The set of Pareto allocation rules out all those allocation for which the capacity constraint is attained with strict inequality, also excluding $\beta_1 = \beta_0 = 0$. The proposed policy assigns all benefits to individuals, and it trade-offs *who* to treat to minimize $|\tau_1\beta_1 - \tau_0\beta_0|$.[7] □

Three properties of $\pi^\star$

This section compares the properties of the policy in Proposition 3.2.2 with existing alternatives. Such properties are stated as corollaries of Proposition 3.2.2.

We first contrast with $\check{\pi}_\omega$ in Equation (3.5), consisting of maximizing welfare for some *pre-specified* importance weights assigned to individuals of the two groups (Rambachan et al., 2020). We show that $\pi^\star$ leads to a weakly smaller UnFairness than $\check{\pi}_\omega$.

Corollary 7 (Maximization with importance weights)**.** *Let $\pi^\star$ as in Equation* (3.7)*. Then* $\text{UnFairness}(\pi^\star) \leq \text{UnFairness}(\check{\pi}_\omega), \forall \omega \in (0,1)$*, where $\check{\pi}_\omega$ is defined in Equation* (3.5)*. In addition, $\check{\pi}_\omega$ does not Pareto dominate $\pi^\star$.*

Corollary 7 shows that UnFairness of $\pi^\star$ is Pareto optimal and *uniformly* smaller than UnFairness of the policy that maximizes a weighted combination of the welfares of the two groups. This follows from the importance weights being chosen to minimize UnFairness in our case. Next, we compare maximization with fairness constraints. Define

$$\Pi(\kappa) = \left\{ \pi \in \Pi : \text{UnFairness}(\pi) \leq \kappa \right\} \subseteq \Pi,$$

the set of policies with constraints on the largest unfairness. We constrast $\pi^\star$ to

$$\widetilde{\pi} \in \arg \max_{\pi \in \Pi(\kappa)} p_1 W_1(\pi) + (1 - p_1) W_0(\pi) \tag{3.9}$$

which maximizes the welfare imposing fairness constraints (Nabi et al., 2019).

[7] Observe that the level of unfairness with the frontier may or may not be potentially strictly larger than the unfairness obtained in an unconstrained scenario. Namely, to achieve zero unfairness for every $\pi \in \Pi_o$, we need that $\tau_1\beta_1^* = \tau_0\beta_0^*$. Substituting $\beta_0^* = \phi/p_0 - p_1\beta_1^*/p_0$ this would require $\beta_1^* = \frac{\phi}{p_0}(\tau_1/\tau_0 + p_1/p_0)^{-1}$ which is not necessarily feasible (i.e., the expression is larger than one).

Corollary 8 (Maximization with fairness constraints)**.** *Let* $\pi^\star$ *as in Equation* (3.7) *and* $\widetilde{\pi}$ *as in Equation* (3.9)*. Suppose that* $\widetilde{\pi} \in \Pi_o$ *(i.e., it belongs to the Pareto frontier). Then* $\text{UnFairness}(\pi^\star) \leq \text{UnFairness}(\widetilde{\pi})$*. Suppose instead that* $\widetilde{\pi} \notin \Pi_o$*. Then* $\text{UnFairness}(\pi^\star) \leq \text{UnFairness}(\pi_o)$ *for all* $\pi_o \in \Pi_o$ *that Pareto dominate* $\widetilde{\pi}$*.*

Corollary 8 shows that if $\widetilde{\pi}$ *is* Pareto optimal, then its UnFairness is larger than UnFairness of $\pi^\star$. When instead $\widetilde{\pi}$ is *not* Pareto optimal, its Pareto dominant allocations have larger UnFairness than $\pi^\star$. Further intuition can be gained under strong duality, which we discuss in Appendix C.1.1. Intuitively, the constraint in Proposition 3.2.2 holding for *some* weighted combinations of welfares (instead of a particular choice of the weights) is key to achieve lower unfairness of $\pi^\star$ relative to $\widetilde{\pi}$, when $\widetilde{\pi}$ is Pareto efficient.[8]

Finally, we compare $\pi^\star$ to $\widetilde{\pi}$ in the presence of *binding* fairness constraints. This corresponds to $\Pi(\kappa) = \Pi$, i.e. the fairness constraints are incorporated in Π.

Corollary 9 (Binding capacity constraints)**.** *Consider* $\pi^\star$ *as in Equation* (3.7) *and* $\widetilde{\pi}$ *as in Equation* (3.9)*. Suppose that fairness constraints are binding to the policy-maker, i.e.* $\Pi(\kappa) = \Pi$*. Then* $\text{UnFairness}(\pi^\star) \leq \text{UnFairness}(\widetilde{\pi})$*. In addition,* $\tilde{\pi}$ *does not Pareto dominate* $\pi^\star$*.*

Corollary 9 shows that with binding fairness constraints, incorporated in the function class Π, $\pi^\star$ achieves a lower UnFairness than the allocation that maximizes the utilitarian welfare under such constraints. Appendix C.3 contains the proofs of the above corollaries.

[8] Under strong duality, the dual of $\widetilde{\pi}$ corresponds to minimize UnFairness for *one particular* weighted combination of welfare exceeding a certain threshold. In contrast, our decision problem imposes the constraint that *some* weighted combination of welfares exceeds a certain threshold. This difference reflects the difference between the lexicographic preferences that we propose as opposed to an additive social planner's utility. It guarantees that whenever $\widetilde{\pi}$ is Pareto optimal, its fairness is dominated by the one under $\pi^\star$.

Fair Targeting: Estimation

We now construct an estimator of $\pi^\star$. We introduce some notation, and we define

$$m_{d,s}(x) = \mathbb{E}\Big[Y_i(d)\Big|X_i = x, S_i = s\Big], \quad \Gamma_{d,s,i} = \frac{1\{S_i = s\}}{p_s}\Big[\frac{1\{D_i = d\}}{e(X_i,S_i)}\Big(Y_i - m_{d,s}(X_i)\Big) + m_{d,s}(X_i)\Big] \tag{3.10}$$

the conditional mean of the group s under treatment d, and the doubly robust score (Robins et al., 1994), respectively. We let $\hat{\Gamma}_{d,s,i}$ the estimated counterpart of $\Gamma_{d,s,i}$. Define

$$\hat{W}_s(\pi) = \frac{1}{n}\sum_{i=1}^{n}\Big(\hat{\Gamma}_{1,s,i} - \hat{\Gamma}_{0,s,i}\Big)\pi(X_i,s). \tag{3.11}$$

the estimated welfare built upon semi-parametric literature (Newey, 1990; Robins and Rotnitzky, 1995), with $\hat{m}_{d,s}(.), \hat{e}(.), \hat{p}_s$, constructed via cross-fitting (Chernozhukov et al., 2018). Details of the cross-fitting procedure are contained in Appendix C.1.2. We consider first general notions of fairness, and introduce the corresponding estimator below.

Definition 3.3.1 (Empirical UnFairness)**.** We define $\mathscr{V}_n(\pi, p_s, e, m)$ an unbiased estimate of UnFairness(π) which depends on observables and the population propensity score and conditional mean. We write $\widehat{\mathscr{V}_n}(\pi) = \mathscr{V}_n(\pi, \hat{p}_s, \hat{e}, \hat{m})$, the empirical counterpart with estimated nuisances functions.

We defer to Section 3.4.3 and Section 3.5 explicit examples of $\widehat{\mathscr{V}_n}(\pi)$.

(Approximate) Pareto Optimality

Next, we characterize the Pareto frontier using linear inequalities. In a first step, we discretize the Pareto frontier, and construct a grid of equally spaced values $\alpha_j \in (0,1)$, $j \in \{1,...,N\}$, with $N = \sqrt{n}$. We approximate the Pareto frontier using the set

$$\widehat{\Pi}_\circ = \Big\{\pi_\alpha \in \Pi : \pi_\alpha \in \arg\sup_{\pi\in\Pi}\Big\{\alpha\hat{W}_0(\pi) + (1-\alpha)\hat{W}_1(\pi)\Big\}, \text{ s.t. } \alpha \in \{\alpha_1,...,\alpha_N\}\Big\}. \tag{3.12}$$

The grid's choice is arbitrary, as long as values are *equally spaced*.

The set $\widehat{\Pi}_{\text{o}}$ may be hard, if not impossible, to directly estimate, since we may have uncountably many solutions (Manski and Thompson, 1989; Elliott and Lieli, 2013). Instead of directly estimating $\widehat{\Pi}_{\text{o}}$, we characterize it through linear constraints. First, we find the largest empirical welfare achieved on the discretized Pareto Frontier defined as

$$\bar{W}_{j,n} = \sup_{\pi\in\Pi}\Big\{\alpha_j\hat{W}_0(\pi) + (1-\alpha_j)\hat{W}_1(\pi)\Big\}, \text{ for each } j \in \{1,...,N\}, \tag{3.13}$$

which can be obtained through standard optimization routines (Kitagawa and Tetenov, 2018; Zhou et al., 2018). We then construct an approximate Pareto frontier as follows:

$$\widehat{\Pi}_{\text{o}}(\lambda) = \Big\{\pi \in \Pi : \exists j \in \{1,...,N\} \text{ such that } \alpha_j\hat{W}_{0,n}(\pi) + (1-\alpha_j)\hat{W}_{1,n}(\pi) \geq \bar{W}_{j,n} - \frac{\lambda}{\sqrt{n}}\Big\}, \tag{3.14}$$

where $\widehat{\Pi}_{\text{o}}(0) = \widehat{\Pi}_{\text{o}}$, and $\widehat{\Pi}_{\text{o}} \subseteq \widehat{\Pi}_{\text{o}}(\lambda)$ for any $\lambda \geq 0$. Notice that $-\frac{\lambda}{\sqrt{n}}$ imposes that the resulting policy is "approximately" Pareto optimal. Here, λ guarantees that $\widehat{\Pi}_{\text{o}}(\lambda)$ contains all Pareto optimal policies with high-probability, while controlling the distance between the estimated frontier and its population counterpart (see Section 3.4.1). The estimated policy is defined as

$$\hat{\pi}_\lambda \in \arg\min_{\pi\in\widehat{\Pi}_{\text{o}}(\lambda)} \widehat{\mathscr{V}}_n(\pi). \tag{3.15}$$

Remark 13 (The choice of the grid and λ)**.** The choice of λ depends on the function class Π. In Section 3.4 we discuss guarantees by imposing that $\lambda/\sqrt{n} \geq M\sqrt{v}/N$, for some finite constant M with λ increasing in the geometric complexity v of Π, and where we choose $N = \sqrt{n}$. In contrast, the function class's complexity does not affect the choice of the grid (i.e., N). This is because the welfare loss due to the grid's approximation error is uniformly bounded by a constant independent of Π.[9] □

[9]Namely, take a grid of $N+1$ equally spaced α_j. Then the approximation error reads as $\sup_{\pi\in\Pi}|\alpha W_1(\pi) + (1-\alpha)W_0(\pi) - \max_{\alpha_j\in\{\alpha_1,\cdots,\alpha_N\}}\alpha_j W_1(\pi) - (1-\alpha)W_0(\pi)| \leq 2M/N$, which is uniformly bounded by M where M

Optimization: Mixed Integer Quadratic Program

We can now provide a mixed-integer quadratic program (MIQP) formulation for optimization. To do so, we represent the constraints on the policy space with the variables $\mathbf{z}_s = (z_{s,1}, \cdots, z_{s,n}), z_{s,i} = \pi(X_i, s), \pi \in \Pi$. Here $\mathbf{z}_s$ have simple representation for general classes of policy functions, such as either probabilistic rules which we derive in Appendix C.1.2 or deterministic linear decision rules (Florios and Skouras, 2008).

We introduce an additional set of decision variables that guarantee the constraints in Equation (3.14) hold. The vector $\mathbf{u} = (u_1, ..., u_N) \in \{0,1\}^N$ encodes the locations on the grid of α for which the supremum in (3.14) is reached at; here, $u_j = 1$ whenever the constraint in Equation (3.14) holds for α_j. The chosen policy must be Pareto optimal, i.e., u_j must be equal to one for at least one j. To ensure this, we introduce a simple constraint $\sum_{j=1}^{N} u_j \geq 1$.

Combining such constraints, it directly follows that $\hat{\pi}_\lambda$ satisfies Equation (3.15) if and only if

$$\hat{\pi}_\lambda \in \arg\min_{\pi} \min_{\mathbf{z_0}, \mathbf{z_1}, \mathbf{u}} \qquad \widehat{\mathscr{V}}_n(\pi) \tag{3.16}$$

$$\begin{aligned}
\text{subject to} \quad & z_{s,i} = \pi(X_i, s), \quad 1 \leq i \leq n && \text{(A)} \\
& u_j \alpha_j \langle \hat{\Gamma}_{1,0} - \hat{\Gamma}_{0,0}, z_0 \rangle + u_j (1 - \alpha_j) \langle \hat{\Gamma}_{1,1} - \hat{\Gamma}_{0,1}, z_1 \rangle \geq u_j n \bar{W}_{j,n} - \sqrt{n}\lambda && \text{(B)} \\
& \langle \mathbf{1}, \mathbf{u} \rangle \geq 1 && \text{(C)} \\
& \pi \in \Pi && \text{(D)} \\
& u_j \in \{0,1\}, \quad 1 \leq j \leq N. && \text{(E)}
\end{aligned}$$

Constraints (A), (C), and (E) are mixed-integer linear constraints, while Constraint (B) is quadratic. However, notice that we can further simplify (B) as a linear constraint at the expense of introducing additional Nn binary variables and $2Nn$ additional constraints (e.g., see Wolsey and Nemhauser 1999; Viviano 2019). Finally, (D) is either linear or quadratic for deterministic

bounds the first moment of the potential outcomes independent of Π.

assignments and linear probability models. Hence the objective admits a MIQP representation whenever $\widehat{\mathscr{V}_n}(\pi)$ admits linear representation in π, which we discuss for several examples in the following section.

Theoretical Guarantees

In this section, we discuss the theoretical properties of the solution in Equation (3.16). We first introduce three conditions.

Assumption 3.4.1. Suppose that the following conditions hold: (A) Π has finite VC-dimension, denoted as v; (B) Π is pointwise measurable.

Condition (A) restricts the function class of interest of the policy function. Simple examples where the finite VC-dimension condition holds are linear decision rules (Manski, 1975), and decision trees (Zhou et al., 2018).[10] Condition (B) ensures the measurability of the supremum of the empirical process of interest (Rai, 2018).

Assumption 3.4.2. Let: (i) $e(X_i,s), p_s \in (\delta, 1-\delta)$, almost surely, for $\delta \in (0,1)$, for all $s \in \{0,1\}$; (ii) $Y_i(d) \in [-M,M]$, for some $M < \infty$, for all $d \in \{0,1\}$ almost surely.

Condition (i) imposes the standard overlap assumption; Condition (ii) assumes uniformly bounded outcomes. See e.g. Mbakop and Tabord-Meehan (2021) for related conditions. The following assumptions are imposed on the estimators.

Assumption 3.4.3 (Nuisances' regularities)**.** There exist some $\xi_1 \geq 1/4, \xi_2 \geq 1/4$, such that:

$$\mathbb{E}\Big[\Big(\hat{m}_{d,s}(X_i) - m_{d,s}(X_i)\Big)^2\Big] = \mathscr{O}(n^{-2\xi_1}), \quad \mathbb{E}\Big[\Big(1\Big/\hat{p}_s\hat{e}(X_i,s) - 1\Big/p_s e(X_i,s)\Big)^2\Big] = \mathscr{O}(n^{-2\xi_2}). \tag{3.17}$$

for all $s,d \in \{0,1\}$, where X_i is out-of-sample. In addition, for a finite constant M and $\delta \in (0,1)$, $\sup_{d\in\{0,1\},s\in\{0,1\},x\in\mathscr{X}} |\hat{m}_{d,s}(x)| < M$, and $\hat{e}(X,S), \hat{p} \in (\delta, 1-\delta)$ almost surely.

[10] In the former case, the VC-dimension is bounded by the number of covariates, whereas in the latter case is bounded by the exponential of the number of layers in the tree (Athey and Wager, 2021; Zhou et al., 2018).

Assumption 3.4.3 states that the *product* of the mean-squared error of the estimated propensity score and conditional mean converges at the parametric. This condition is standard in the doubly-robust literature (Chernozhukov et al., 2018; Farrell, 2015). Assumption 3.4.3 also states that the conditional mean function and the propensity score functions are uniformly bounded. The conditions can be stated asymptotically, in which case results should be interpreted in the asymptotic sense only (Athey and Wager, 2021).[11]

Guarantees on the Pareto Frontier

It is interesting to study the behavior of the estimated frontier relative to its population counterpart. We do so in the following theorems.

Theorem 3.4.1. *Under Assumptions 3.2.1, 3.4.1-3.4.3, for any* $\gamma \in (0,1), \lambda \geq 0$*, a universal constant* $c_0 < \infty$*, with probability larger than* $1-\gamma$*,*

$$\sup_{\alpha\in(0,1),\pi\in\Pi} \left| \alpha W_0(\pi) + (1-\alpha) W_1(\pi) - \max_{\alpha_j \in \{\alpha_1,\cdots,\alpha_N\}} \left\{ \alpha_j \widehat{W}_0(\pi) + (1-\alpha_j)\widehat{W}_1(\pi) - \frac{\lambda}{\sqrt{n}} \right\} \right| \leq c_0 \sqrt{\frac{v}{n}} + c_0 \sqrt{\frac{\log(2/\gamma)}{n}} + \frac{\lambda}{\sqrt{n}}. \tag{3.18}$$

Theorem 3.4.1 shows that the distance between the estimated Pareto frontier and its population counterpart converges to zero at rate $1/\sqrt{n}$ for a choice of $\lambda = \mathcal{O}(1)$. A natural question is whether the estimated Pareto frontier also contains all Pareto optimal allocations for a finite λ. We complement Theorem 3.4.1 showing that with high probability the set of estimated allocations $\widehat{\Pi}_o(\lambda)$ contains the Pareto frontier for a suitable (finite) choice of λ.

Theorem 3.4.2. *Let Assumptions 3.2.1, 3.4.1-3.4.3 hold. Then for any* $\gamma \in (0,1), \lambda \geq \underline{b}(\sqrt{v} + \sqrt{\log(2/\gamma)} + 1)$*, for a constant* $\underline{b} > 0$*, independent of the sample size,* $N = \sqrt{n}$*, it follows that* $\mathbb{P}\Big(\Pi_o \subseteq \widehat{\Pi}_o(\lambda)\Big) \geq 1-\gamma$.

Theorem 3.4.2 complements Theorem 3.4.1 showing that it suffices $\lambda = \mathcal{O}(1)$ (and hence a slackness of order $\mathcal{O}(1/\sqrt{n})$) for the set of estimated allocations to contain the Pareto frontier.

[11] We also observe that uniformly boundedness can be replaced by uniform consistency as in Athey and Wager (2021), in which case, however, regret results are derived in an asymptotic sense.

The proofs of Theorems 3.4.1, 3.4.2 are contained in Appendix C.3. Theorem 3.4.2 uses *finite sample* properties of the estimated (discretized) frontier showing that it concentrates around its population counterpart, uniformly over Π at rate $\sqrt{v/n}$. The choice of $\lambda/\sqrt{n}$ matches the upper-bound on the maximal deviations.

General Fairness Bounds

Given the guarantees on the frontier, we next analyze guarantees on the (un)fairness of the policy. We start our discussion by introducing regret bounds for generic notions of unfairness under high-level assumptions and then provide examples of upper and lower bounds.

Assumption 3.4.4 (High-level conditions on UnFairness)**.** For some $\eta > 0, \gamma > 0$,

$$\mathbb{P}\Big(\sup_{\pi \in \Pi} \Big| \widehat{\mathscr{V}}_n(\pi) - \text{UnFairness}(\pi) \Big| \leq \mathscr{K}(\Pi, \gamma) n^{-\eta} \Big) \geq 1 - \gamma$$

for some constant $\mathscr{K}(\Pi, \gamma) < \infty$. Also assume that UnFairness(π) is uniformly bounded.

Assumption 3.4.4 states that the estimated unfairness converges with probability $1-\gamma$ to population unfairness uniformly over Π at rate $n^{-\eta}$ for some arbitrary η. The constant $\mathscr{K}(\Pi, \gamma)$ depends on the function class' complexity and the probability γ. We characterize the constant and the rate η in examples in Section 3.4.3 and Appendix C.1.4.

Theorem 3.4.3. *Let Assumptions 3.2.1, 3.4.1-3.4.4 hold. Then for some constants* $0 < c_0, \underline{b} < \infty$*, independent of n,* $\lambda \geq \underline{b}(\sqrt{v} + \sqrt{\log(2/\gamma)} + 1), N = \sqrt{n}$*, with probability at least* $1 - 2\gamma$*,*

$$\text{UnFairness}(\hat{\pi}_\lambda) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) \leq \frac{c_0}{\sqrt{n}} + \frac{c_0 \mathscr{K}(\Pi, \gamma)}{n^\eta}. \tag{3.19}$$

The proof is contained in Appendix C.3.[12] Theorem 3.4.3 characterizes the convergence

[12]The proof consists in (i) first using Theorem 3.4.2 to show that the set of Pareto allocations is contained with high probability within the estimated set of allocations; (ii) second, bounding the regret with twice the difference between the estimated and population UnFairness, taking the supremum over $\pi \in \widehat{\Pi}_o$ (which contains Π_o with high probability), and using the fact that the VC-dimension of Π bounds the VC-dimension of the set of Pareto allocations.

rate of the UnFairness of the estimated policy relative to the lowest unfairness within the class of Pareto allocations. To our knowledge, this is the first result of this type of fair policy. The rate depends on the convergence rate of the estimated UnFairness. In the following paragraphs, we discuss examples and sufficient conditions for Assumption 3.4.4 to hold and formally characterize the rate of convergence η and the constant $\mathscr{K}(\cdot)$.

Regret: Examples and Rate Characterization

Here we discuss three examples, one based on policy predictions, a second based on the welfare effect, and a third based on incentive compatibility.

Definition 3.4.1 (Prediction disparity)**.** Prediction disparity and its empirical counterpart take the following form

$$C(\pi) = \mathbb{E}\Big[\pi(X,S)|S=0\Big] - \mathbb{E}\Big[\pi(X,S)|S=1\Big], \quad \hat{C}(\pi) = \frac{\sum_{i=1}^{n}\pi(X_i)(1-S_i)}{n(1-\hat{p}_1)} - \frac{\sum_{i=1}^{n}\pi(X_i)S_i}{n\hat{p}_1},$$

Prediction disparity captures fairness based on the probability of treatment of different groups. The second notion of UnFairness measures welfare effects disparities *between* the two groups.

Definition 3.4.2 (Welfare disparity)**.** Define the welfare disparity and its empirical counterpart as follows.

$$D(\pi) = W_0(\pi) - W_1(\pi), \quad \widehat{D}(\pi) = \widehat{W}_0(\pi) - \widehat{W}_1(\pi).$$

Between-groups disparity captures the difference in *welfare* between the advantaged group $(S=0)$ and the disadvantaged group $(S=1)$, relative to the baseline.[13]

The policymaker may also consider $|D(\pi)|$ or $|C(\pi)|$ as measures of UnFairness, in which case the policymaker treats the two groups symmetrically, whose regret bounds are discussed in

[13]Recall the definition of welfare in Equation (3.2) where we only consider the effect under treatment the effect under control.

Appendix C.3.13. One last example is based on the notion of incentive compatibility, motivated by discussion in Narita (2021).

Definition 3.4.3 (Incentive compatibility)**.** Incentive compatibility is defined as

$$\mathscr{I}(\pi) = I_1(\pi) + I_0(\pi), \quad I_s(\pi) = \mathbb{E}\Big[\pi(X, 1-s)(Y(1) - Y(0))|S = s\Big] - \mathbb{E}\Big[\pi(X, s)(Y(1) - Y(0))|S = s\Big]$$

with estimator $\widehat{\mathscr{I}}(\pi) = \hat{I}_1(\pi) + \hat{I}_0(\pi)$, $\hat{I}_s(\pi) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\Gamma}_{1,s,i} - \hat{\Gamma}_{0,s,i})\pi(X_i, 1-s) - \widehat{W}_s(\pi)$.

Here $I_s(\pi)$ captures fairness based on the incentive of an individual in revealing her sensitive attribute: $I_s(\pi)$ is positive if the welfare of an individual generated from reporting her sensitive attribute incorrectly is larger than the welfare obtained if she reported it correctly. Additional notions, such as predictive parity, can also be considered and omitted for the sake of brevity, see Appendix C.1.4 for details. Finally, observe that each of the three definitions above considers UnFairness linear in the policy π, and hence optimization can be performed via MIQP.

Upper and Lower Bounds: Rate Characterization

In the following theorem, we discuss the rate of the regret-bound.

Theorem 3.4.4 (Regret bound)**.** *Let Assumptions 3.2.1, 3.4.1-3.4.3 hold. Let either (i)* $\text{UnFairness}(\pi) = D(\pi)$*, and* $\widehat{\mathscr{V}}_n(\pi) = \widehat{D}(\pi)$*, (ii) or* $\text{UnFairness}(\pi) = C(\pi)$*, and* $\widehat{\mathscr{V}}_n(\pi) = \widehat{C}(\pi)$*, (iii) or* $\text{UnFairness}(\pi) = \mathscr{I}(\pi)$*, and* $\widehat{\mathscr{V}}_n(\pi) = \widehat{\mathscr{I}}(\pi)$*. Then for some constants* $0 < \underline{b}, c_0 < \infty$ *independent of the sample size, for any* $\gamma \in (0,1), \lambda \geq \underline{b}(\sqrt{v} + \sqrt{\log(2/\gamma)} + 1), N = \sqrt{n}$*, with probability at least* $1 - 2\gamma$*,*

$$\text{UnFairness}(\hat{\pi}_\lambda) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) \leq c_0\sqrt{\frac{v}{n}} + c_0\sqrt{\frac{\log(2/\gamma)}{n}}.$$

The proof is included in Appendix C.3. Theorem 3.4.4 characterizes the regret bound for three different notions of UnFairness. The bound scales at rate $1/\sqrt{n}$. The lower bound depends, however, on the notion of unfairness. In the following lines, we derive a lower bound for any data-dependent policy which achieves the same rate for the predictive disparity.

Theorem 3.4.5 (Lower bound)**.** *Let* Π *be such that* $\pi(x,s)$ *is constant in its last argument* s *for all* $x \in \mathscr{X}, \pi \in \Pi$*, and with finite VC-dimension* $v \geq 3$*. Let* $\text{UnFairness}(\pi) = C(\pi)$*, and* $\widehat{\mathscr{V}_n}(\pi) = \widehat{C}(\pi)$*. Let* $\mathscr{U}$ *be the set of distributions of* (X,S) *and* $\mathscr{P}(X,S) = \{P_{Y,D|(X,S)} :$ *such that* $|Y| < M$ *a.s., and* $P(D=1|X,S) \in (\delta, 1-\delta)\}$*. Then, there exists a distribution* $P_{X,S,Y,D} = P_{X,S}P_{Y,D|X,S}$ *with* $P_{X,S} \in \mathscr{U}, P_{Y,D|X,S} \in \mathscr{P}(X,S)$*, such that for every rule* $\pi_n \in \Pi_o$ *based upon* $(X_1,S_1,Y_1,D_1),\cdots,(X_n,S_n,Y_n,D_n)$*, for finite constants constant* $0 < c_0, \bar{C} < \infty$ *independent of n, and any* $\gamma \in (0,1/4)$*,* $n \geq \max\{\bar{C}\log(1/(4\gamma)), v-1\}$*, with probability at least* γ

$$\text{UnFairness}(\pi_n) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) \geq \sqrt{\frac{c_0 \log(\frac{1}{4\gamma})}{n}}.$$

The proof is contained in Appendix C.3, and, to our knowledge, it is the first result of this type for fair and Pareto optimal policies. The lower bound states that we can find distribution and some positive (non-vanishing) probability γ such that any data-dependent policy π_n achieves a regret which scales to zero at a rate no faster than $1/\sqrt{n}$. Observe that a direct corollary of such a result is that the rate of the lower bound is also achieved in expectation. The condition imposes a restriction on the set of policies Π: Π does not contain policies that use the sensitive attribute as a covariate. This class of policies occurs if anti-discriminatory laws are enforced and incorporated over the set Π. The lower bound applies to prediction disparity, and we leave to future research a more comprehensive study of lower bounds under generic notions of fairness. The derivation modifies arguments in the empirical risk minimization literature (Devroye et al., 2013) due to the dependence of the objective function with the *conditional* probability of treatment.

Throughout this section, we have considered some examples *distributional* notions of fairness, i.e., notions of fairness that depend on distributional statements relative to the sensitive attribute, but not necessarily counterfactual ones. Distributional notions of fairness are often used in the literature, see e.g. Kasy and Abebe (2020), Donini et al. (2018), Narita (2021). *Counterfactual* notions depend instead on counterfactual statements relative to the *sensitive* attribute (Kilbertus et al., 2017). We discuss one counterfactual notion in Section 3.5.

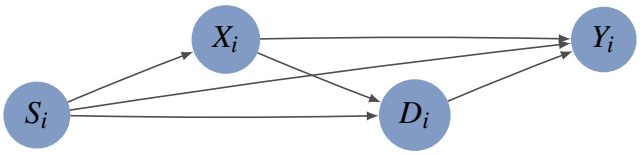


Figure 3.1. Example of a Directed Acyclical Graph.

Counterfactual UnFairness

This section is of independent interest, and it discusses a novel notion of UnFairness which connects the literature on causal fairness (Kilbertus et al., 2017) and the economic literature on envy-freeness (Varian, 1976). The notion is based on counterfactual statements relative to the *sensitive* attribute. This section defines $Y(d,s), X(s)$ the potential outcome and covariates as functions of the sensitive attribute s. The following causal model is considered.

Assumption 3.5.1. Let (A) $Y(d,s) \perp (D,S)|X(s)$, (B) $X(s) \perp S$.

Assumption 3.5.1 is only required for estimation with a counterfactual notion of fairness discussed in the following lines and not required for notions of fairness discussed in the previous section. Condition (A) and (B) in Assumption 3.5.1 state that the sensitive attribute is independent of potential outcomes and covariates, while it allows for the dependence of *observed* covariates and outcomes with the sensitive attribute. This dependence is captured by indexing potential outcomes and covariates by the sensitive attribute. Figure 3.1 displays a directed acyclical graph under which Assumption 3.5.1 holds. We observe that the dependence between the sensitive attribute and outcomes and covariates can occur through *unobserved* characteristics, which are dependent on both outcomes and sensitive attributes as long as observables do not *causally* affect the sensitive attribute. Assumption 3.5.1 is satisfied in the DAGs discussed in Kilbertus et al. (2017), where sensitive attributes do not have causal parents.

Let the conditional welfare, for the policy function being assigned to the opposite attribute, i.e., the effect of $\pi(x,s_1)$, on the group s_2, conditional on covariates, be

$$V_{\pi(x,s_1)}(x,s_2) = \mathbb{E}\Big[\pi(x,s_1)Y_i(1,s_2) + (1-\pi(x,s_1))Y_i(0,s_2)\Big|X_i(s_2) = x\Big]. \tag{3.20}$$

Envy refers to the concept that "an allocation is equitable if and only if no agent prefers another agent's bundle to his own" (Varian, 1976). We say that the agent with attribute s_2 *envies* the agent with attribute s_1, if her welfare (on the right-hand side of Equation 3.21) exceeds the welfare she would have received had her covariate and policy been assigned the opposite attribute (left-hand side of Equation 3.21), namely

$$\mathbb{E}_{X(s_1)}\Big[V_{\pi(X(s_1),s_1)}\Big(X(s_1),s_2\Big)\Big] > \mathbb{E}_{X(s_2)}\Big[V_{\pi(X(s_2),s_2)}\Big(X(s_2),s_2\Big)\Big]. \tag{3.21}$$

We then measure the unfairness towards an individual with attribute s_2 as

$$\mathscr{A}(s_1,s_2;\pi) = \mathbb{E}_{X(s_1)}\Big[V_{\pi(X(s_1),s_1)}(X(s_1),s_2)\Big] - \mathbb{E}_{X(s_2)}\Big[V_{\pi(X(s_2),s_2)}\Big(X(s_2),s_2\Big)\Big]. \tag{3.22}$$

Equation (3.22) makes a connection to previous notions of *counterfactual fairness* (Kilbertus et al., 2017), while, differently from previous references, (i) we provide formal justification to fairness using an envy-freeness argument; (ii) we construct the definition of fairness based on *distributional* impact of the treatment allocation rule on the welfare. It is complementary to Kusner et al. (2019), who compare the effect of the same policy on individuals with the opposite sensitive attribute, lacking an envy-based justification.

We can then estimate $\mathscr{A}(\cdot)$ as follows:

$$\begin{aligned}\mathscr{A}_n(s,s';\pi) =& \frac{1}{n\hat{p}_s}\sum_{i:S_i=s}\Big\{\hat{m}_{1,s'}(X_i)\pi(X_i,s) + \hat{m}_{0,s'}(X_i)(1-\pi(X_i,s))\Big\} \\ & - \frac{1}{n}\sum_{i=1}^{n}\Big\{\hat{\Gamma}_{1,s,i}\pi(X_i,s) - \hat{\Gamma}_{0,s,i}(1-\pi(X_i,s))\Big\}.\end{aligned} \tag{3.23}$$

Whenever we aim not to discriminate in either direction (women with respect to men and vice-

versa), we define unfairness by taking the sum of the effects $\mathscr{A}(s_1, s_2; \pi)$ and $\mathscr{A}(s_2, s_1; \pi)$,[14] and define counterfactual envy-freeness and its empirical counterpart as

$$\mathrm{E}(\pi) = \mathscr{A}(1,0;\pi) + \mathscr{A}(0,1;\pi), \quad \hat{\mathrm{E}}(\pi) = \mathscr{A}_n(1,0;\pi) + \mathscr{A}_n(0,1;\pi). \tag{3.24}$$

Different from prediction notions of fairness, theoretical guarantees also require the following condition.

Assumption 3.5.2. Assume that for some $\zeta > 0$,

$$\mathbb{E}\Big[\Big(\hat{m}_{d,s_1}(X_i(s_2)) - m_{d,s_1}(X_i(s_2))\Big)^2\Big] = \mathscr{O}(n^{-2\zeta}), \quad \forall s_1, s_2 \in \{0,1\}, d \in \{0,1\}.$$

Assumption 3.5.2 states that the estimator of the conditional mean function for each sensitive attribute and treatment status $s, d \in \{0,1\}$, must converge to the true conditional mean function in mean-squared error at some arbitrary rate $2\zeta > 0$. Here, we require convergence in l_2 for a given sensitive attribute conditional on the *opposite* sensitive attribute, due to the particular notion of fairness considered.[15]

Theorem 3.5.1 (Counterfactual envy-freeness)**.** *Let Assumptions 3.4.1-3.4.3, 3.5.1 and 3.5.2 hold. Let* $\mathrm{UnFairness}(\cdot) = \mathrm{E}(\cdot)$ *and* $\hat{\mathscr{V}}_n(\cdot) = \hat{E}(\cdot)$*. Then for some constants* $0 < \underline{b}, c_0 < \infty$ *independent of the sample size, for any* $\gamma \in (0,1), \lambda \geq \underline{b}(\sqrt{v} + \sqrt{\log(2/\gamma)} + 1), N = \sqrt{n}$*, with probability at least* $1 - 2\gamma$*,*

$$\mathrm{UnFairness}(\hat{\pi}) - \inf_{\pi \in \Pi_o} \mathrm{UnFairness}(\pi) \leq c_0 \sqrt{\frac{v}{n^{2\zeta}}} + c_0 \sqrt{\frac{\log(2/\gamma)}{n}}.$$

A corollary of Theorem 3.5.1 is that under the parametric rate of convergence of the conditional mean function, the regret bound scales at rate $n^{-1/2}$. Interestingly, the convergence rate is of order slower than $n^{-1/2}$ for non-parametric estimators compared to the notions of

[14] Such an approach builds on the notion of "social envy" discussed in Feldman and Kirman (1974).

[15] Namely, to estimate fairness, we need to extrapolate relative to the *opposite* group.

UnFairness discussed in Section 3.4. The slower convergence rate is because counterfactual envy-freeness requires estimating the conditional mean function on the population with attribute $S = s_1$ while averaging over the covariates' distribution with the opposite attribute, therefore requiring extrapolation. This result showcases the *trade-off* in the choice of a counterfactual notion of unfairness relative to predictive ones. It opens new questions on trade-offs across different notions of UnFairness which we leave to future research.

Empirical Application and Numerical Study

We now discuss the empirical application. This section designs a policy that assigns students to entrepreneurial programs, imposing fairness on gender. We use data that originated from Lyons and Zhang (2017). The paper studies the effect of an entrepreneurship training and incubation program for undergraduate students in North America on subsequent entrepreneurial activity. We have in total 335 observations, of which 53% treated and the remaining under control, and 26% of applicants are women.[16] The population of interest is the pool of final applicants. We construct a targeting rule that assigns the award to the finalist based on the applicant's observable characteristics. We maximize subsequent entrepreneurial activity, which is captured using a dummy variable, indicating whether the participant worked in the startup once the program ended. The study is a quasi-experiment, and, as noted in Lyons and Zhang (2018), the focus on the pool of final applicants mitigates the selection on unobservables. Similarly to Lyons and Zhang (2018) we control for residual confounding of the treatment assignment among final applicants through individual level observable characteristics and an observable quality score of the final applicant. Estimation of the nuisance functions is through penalized regression and discussed in Appendix C.2.1.

We consider three notions of UnFairness: (i) *counterfactual envy*; (ii) (ii) *predictive disparity*, which minimizes the probability of treatment between the two groups as in Definition 3.4.1; (iii) *predictive disparity* with absolute value (i.e. it denotes the absolute difference between

[16]Data is available at https://www.openicpsr.org/openicpsr/project/113492/version/V1/view.

the probability of treatment between the two groups). While (ii) and (iii) do not impose conditions on the distribution of the sensitive attribute, counterfactual envy ((i)) assumes unconfoundedness also of the sensitive attribute. Such a condition is equivalent to assuming that the decision to change gender is exogenous.

We consider *linear* decision rules, given their large use in economics (Manski, 1975)[17]

$$\Pi = \Big\{ \pi(x, \text{fem}) = 1\Big\{ \beta_0 + \beta_1 \text{fem} + x^\top \phi \geq 0 \Big\}, \quad (\beta_0, \beta_1, \phi) \in \mathscr{B} \Big\}. \tag{3.25}$$

We allow covariates x to be either (1) the years to graduation, years of entrepreneurship, the region of the start-up, the major, the school rank, or (2) the score assigned to the candidate by the interviewer and the school rank. We refer to these two cases respectively as *Case 1* and *Case 2*. We consider in-sample capacity constraints imposed on the function class with at most 150 individuals selected for the treatment.[18]

We compare the proposed methodology to the method that maximizes the empirical welfare with the double robust score (Athey and Wager, 2021). We consider three nested function classes for the welfare maximization method. The first does not impose any restriction except for the functional form in Equation (3.25). The second, imposes that $\beta_1 = 0$. The third class imposes that $\beta_1 = 0$ *and* that the average effect of the policy on females is at least as large as the one on males. The function classes are

$$\begin{aligned}
\Pi_1 =& \Pi, \quad \Pi_2 = \Big\{ \pi(x) = 1\Big\{ \beta_0 + x^\top \phi \geq 0 \Big\} \Big\}, \\
\Pi_3 =& \Big\{ \pi(x) = 1\Big\{ \beta_0 + x^\top \phi \geq 0 \Big\}, \mathbb{E}_n\Big[(Y_i(1) - Y_i(0)) \pi(X_i) \Big| S = 1 \Big] \geq \mathbb{E}_n\Big[(Y_i(1) - Y_i(0)) \pi(X_i) \Big| S = 0 \Big] \Big\},
\end{aligned}$$

where $\mathbb{E}_n[\cdot]$ denote the empirical expectation, estimated using the doubly-robust method.

[17] This is estimated solving Equation (3.16) with a small slackness parameter of order 10^{-6}. The reader may refer to Appendix C.1.3 for details.

[18] The validity of the in-sample capacity constraints follows from a uniform concentration argument of the capacity constraint around its expectation. Formally, we can bound $\mathbb{E}\Big[\sup_{\pi \in \Pi} \Big| \frac{1}{n} \sum_{i=1}^n \pi(X_i) - \mathbb{E}[\pi(X_i)] \Big| \Big]$ using the Dudley's entropy integral bound (Wainwright, 2019), that guarantees that the estimated policy controls with high probability the percentage of treated individuals by $150/335$ up to a small tolerance of order $\sqrt{v/n}$. We can then show using the bounded difference inequality that $\sup_{\pi \in \Pi} \Big| \frac{1}{n} \sum_{i=1}^n \pi(X_i) - \mathbb{E}[\pi(X_i)] \Big|$ concentrates at rate $1/\sqrt{n}$ around its expectation with high probability.

Table 3.1. Application: results from fair targeting. The first two columns report the welfare improvement plus the baseline value. The last column reports the importance weights assigned by the method to the welfare of female students. FTP Envy refers to the Fair Targeting rule that minimizes envy-freeness unfairness; FTP Predictive Disp (Definition 3.4.1) refers to the Pareto allocation that minimizes the difference in probability of treatment (Abs indicate in absolute value); Welfare Max. 1 denotes the method that maximizes the empirical welfare considering Π_1, and similarly Welfare Max. 2, 3 for the function classes, respectively Π_2, Π_3.

	Welfare Female		Welfare Male		Importance Weight	
	Case 1	Case 2	Case 1	Case 2	Case 1	Case 2
Fair Envy	0.376	0.372	0.272	0.195	0.384	0.487
FTP Pred	0.432	0.374	0.224	0.180	0.847	0.924
FTP Pred Abs	0.433	0.351	0.208	0.235	0.924	0.487
Welfare Max. 1	0.376	0.351	0.272	0.235	0.266	0.266
Welfare Max. 2	0.288	0.307	0.285	0.238	0.266	0.266
Welfare Max. 3	0.331	0.307	0.265	0.238	0.266	0.266

In Figure 3.2 we plot the Pareto frontier over each function class.[19] The figure shows that restricting the function class of interest leads to Pareto-dominated allocations. This outlines the limitations of maximizing welfare under fairness constraints: such constraints can result in harm for both types of individuals. Instead, the proposed method enforces Pareto optimality in the least constrained environment (red line) and then selects the policy based on fairness considerations.

In Table B.2.1 we collect results[20] of the welfare of female and male students, as well as the relative importance weight assigned to each group for methods that maximize different UnFairness measures. In the table, we observe that minimizing Envy and Predictive Disparity leads to (weakly) larger welfare effects on the minority group. Envy leads to comparable results to welfare maximization for *Case* 1 due to the discreteness of the frontier.[21] We observe an

[19] The value functions over the Pareto frontier can be exactly recovered as follows: we solve 2 optimization problems for each α_j, $j \in \{1, ..., N\}$. For each of these problems, we impose constraints on the welfare of one of the two groups being larger than the other and vice-versa; we then select the subset of solutions that are not Pareto dominated by the other, and we plot the corresponding welfares in the figure.

[20] In computations, the competitors (Welfare Maximization) achieves the global optimum (dual gap equal to zero). For the proposed method, we impose a maximum time limit on the MIQP.

[21] Observe that even if the weight α is larger for FTP Envy and FTP Parity Abs in *Case* 1 and 2 respectively, this does not lead to a different result than Welfare Max. 1 due to the discreteness of the Pareto frontier.

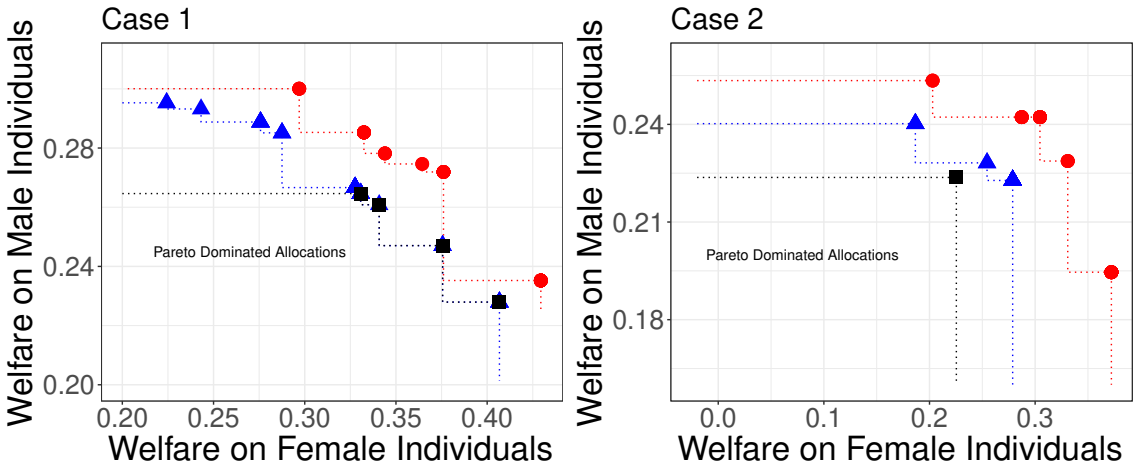


Figure 3.2. (Discretized) Pareto frontier under deterministic linear policy rule (estimated through MIQP). Dots denote Pareto optimal allocations. Red dots (circle) correspond to Π_1, blue dots (triangle) to Π_2 and black dots (square) to Π_3.

increase in the welfare of female students when minimizing the *absolute* difference between probabilities of treatments for *Case* 1 and comparable results to the welfare maximization method for *Case* 2. The table shows that the proposed method finds importance weights assigned to each group solely based on the notion of fairness provided, without requiring any prior specification of relative weights assigned to each group. The method that maximizes the empirical welfare instead assigns to the sensitive group the importance weight equal to its corresponding probability, small for minorities. In two settings only, the results coincide with the proposed method due to the discreteness of the frontier.

Figure 3.3 reports the unfairness level for different sets of covariates, with unfairness measured as the difference in the probability of treatments between the two groups. Overall, Figure 3.3 shows that the level of the unfairness of the proposed method is uniformly smaller than the unfairness achieved by maximizing welfare, consistently with results in Section 3.2.

Finally, it is interesting to compare also with probabilistic decision rules, which are allowed in our framework. Figure 3.3 also collects result also for a probabilistic policy function (in green) which is a super-set of Π in Equation (3.25) and assigns different probabilities of

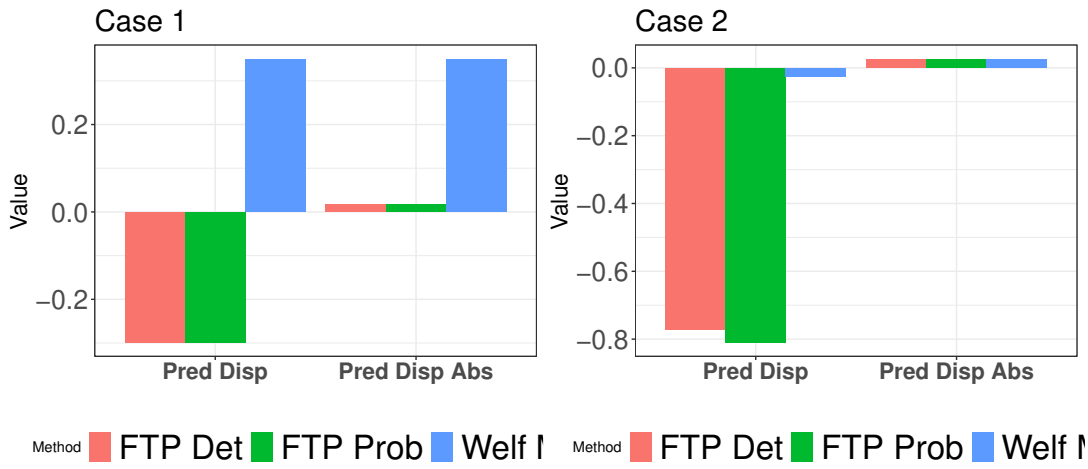


Figure 3.3. Application: UnFairness comparisons. Unfairness level of the Fair Policy Targeting method with a deterministic allocation rule (in red), with a probabilistic decision rule (in green), and of the welfare maximization method (in blue). Pred disp refers to Definition 3.4.1 and Pred disp abs to Definition 3.4.1 in absolute value. Smaller values indicate smaller UnFairness.

treatments to groups below and above the hyperplane in Equation (3.25) (see Appendix C.1.3).[22] Results are mostly comparable across probabilistic or deterministic decisions. However, we find that a probabilistic decision enlarges the set of Pareto efficient allocations relative to the maximum score function, which we discuss in Appendix C.2.1.

A Calibrated Experiment

Next, we conducted a calibrated experiment to test the method's performance. We estimate the model using a penalized regression using data from Lyons and Zhang (2017) as we do in the empirical application, and we run the simulations calibrated to the estimated model. Covariates and sensitive attributes are drawn with replacement from the empirical distribution, and the policy is estimated using all covariates also used for estimation. Appendix C.2.2 contains details. We consider 500 replications with $n = 400$, while in Appendix C.2.2, we show that results are robust for $n = 600$. For computational reasons, we consider linear probabilistic treatment assignments $X^\top \beta$ constrained between zero and one, with β computed using a linear program

[22] Formally, the function class is $\left\{\pi_\beta(X,S) = p_1 1\{X_i^\top \beta + S\beta_0 > 0\} + p_0 1\{X^\top \beta + S\beta_0 \leq 0\}, p_1, p_0 \in [0,1], \beta \in \mathscr{B}\right\}$.

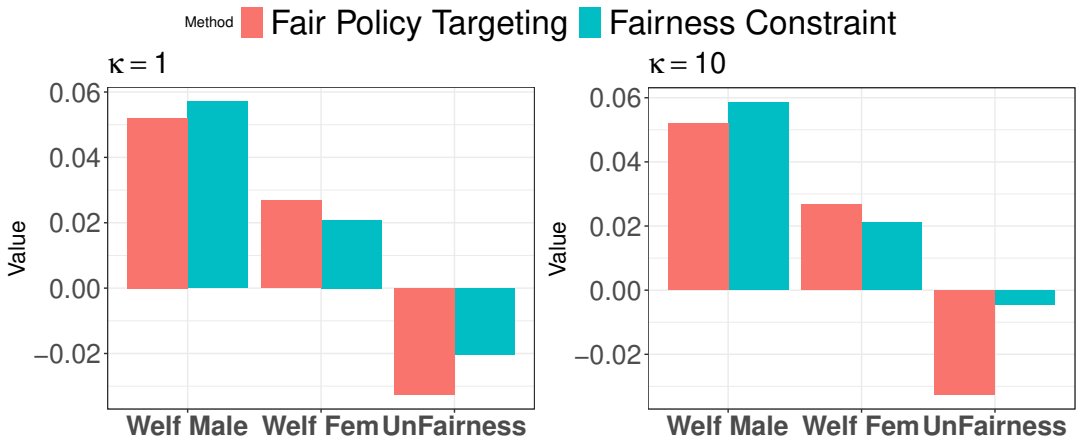


Figure 3.4. Numerical study calibrated to Lyons and Zhang (2017). Results averaged over 500 replications using a linear probabilistic decision rule constrained between zero and one. $n = 400$. The left panel collects results for $\kappa = 1$ (stricted fairness constaints) and the right panel for $\kappa = 10$ (slacker fairness constraints). Colored in red is the proposed method, and in blue, the method that maximizes the empirical welfare imposing fairness constraints. In each panel, the first two columns report the welfare of each group, and the third column the difference in probability of being treated between females and males students.

with linear constraints.[23] We contrast our procedure (estimated without fairness constraints) to the procedure that considers the same function class, but it maximizes the empirical welfare under the additional constraint that $\text{UnFairness}_n(\pi) \leq \kappa/n$, where $\text{UnFairness}_n(\pi)$ is the empirical counterpart of the notion of predictive parity as in Definition 3.4.1.[24] We let $\kappa \in \{1, 10\}$. We impose the same capacity constraint as in the application, with the maximum number of treated students equal to one hundred and fifty. We collect results in Figure 3.4, where we report the population welfare and unfairness of each method (based on the simulation design), averaged across the simulations. The figure corroborates our theoretical findings. First, Fair Policy Targeting is not Pareto dominated by the existing competitor: it leads to strictly *larger* welfare on female students, but to smaller welfare on the male students. Second, our proposed procedure also leads to *smaller* unfairness. This property follows from the fact that our procedure either (or

[23] Such assignments are chosen for computation reasons since they do not involve integer variables. See Appendix C.1.3 for a detailed description.

[24] We observe that also other measures may be considered. We consider predictive parity for computational convenience and the sake of brevity.

both) is strictly Pareto dominant or leads to smaller UnFairness while not being Pareto dominated. The key intuition is the following: even when fairness constraints are imposed, the competitor imposes a small importance weight on the minority group. This results in UnFairness that, while satisfying the constraint, can still be large. Fair Policy Targeting instead chooses the importance weight assigned to each group to minimize UnFairness, and, as a result, either (or both) leads to lower unfairness without being Pareto dominated as in this case, or it is strictly Pareto dominant.

Conclusion

In this chapter, we have introduced a novel method for estimating fair and optimal treatment allocation rules. We proposed a multi-objective decision problem, where the policymaker aims to select the least unfair policy in the set of Pareto optimal allocations. We discuss a set of theoretical guarantees on the estimated policy and provide an application. From a theoretical perspective, we open new questions on the trade-offs between predictive and causal notions of fairness and its corresponding regret bound. Counterfactual notions require extrapolation, hence possibly leading to a slower convergence rate. We leave to future research a comprehensive study of properties of different notions of fairness in terms of their implied regret. From a practical perspective, an interesting new direction is estimation with non-utilitarian within-group welfare measures.

Chapter 3 has been prepared for submission for publication. Viviano, Davide; Bradic, Jelena. Chapter 3 is co-authored with Jelena Bradic. Davide Viviano is the principal investigator.

Chapter 4

(When) Should you adjust inference for multiple hypothesis testing?

Introduction

This chapter departs from the problem of choosing whom to treat discussed in previous chapters and focuses on the problem of statistical inference in the presence of multiple policy effects.

Empirical papers in economics usually test more than one hypothesis. Historically researchers have treated these tests as independent when conducting inference. But recently, some have begun to approach "multiple hypothesis testing" (MHT) scenarios differently, using procedures that adjust one test for the presence of others in order to control an aggregate error rate (e.g., the family-wise error rate (FWER) or the false discovery rate (FDR)) or testing for effects on indices that aggregate multiple outcomes. Considering experimental papers published in "top 5" journals, for example, the share that adjusted inference in one of these ways increased from 0% to 39% over the past decade (left-hand panel of Figure 4.1), so that there is now wide variation in *whether* they do so, as well as *how* they do so (right-hand panel). This raises several simple but important questions: when (if at all), how, and why should researchers be required to adjust for multiplicity? While there is a general sense that simply ignoring MHT issues could create problematic incentives, there is little consensus on how exactly they should be addressed. This leaves referees and editors to rely on intuition or taste when evaluating submitted papers.

In this chapter, we seek to understand whether and why MHT procedures might arise as desirable solutions within the process of scientific communication, and if they do, what observable features of the economic environment determine the right procedures to use. To this end, we study a model of the academic publication process in which a benevolent social planner chooses professional norms with respect to MHT adjustments, taking into account the researchers' incentives. The framework embeds three core ideas. First, policy decisions are influenced by the summary recommendations (in particular, hypothesis tests) contained in research papers. Second, while this makes research results a public good, the costs of producing them are born privately by the researcher. She decides whether or not to incur these costs, conducting a pre-specified experiment with multiple hypotheses based on its chances of subsequent publication.[1] Hypothesis testing protocols must therefore balance the twin goals of (i) motivating the production of research and (ii) producing good policy guidance from it. Finally, we examine how to balance these objectives conservatively, selecting protocols that maximize worst-case social welfare, which we refer to as maximin protocols.

In the case of a single hypothesis, these assumptions can be used to rationalize standard testing procedures, as shown in an insightful paper by Tetenov (2016).[2] To study the multiple hypothesis case, however, we require notions of optimality that are more nuanced than those in the single-hypothesis case. We show that with multiple hypotheses, no maximin rule dominates all others (in contrast to the single-hypothesis case). Motivated by this non-existence result, we develop an appropriate notion of *local power* with the property that any maximin rule that is locally most powerful is also admissible, allowing us to select among admissible maximin rules.

A defining feature of our approach is to identify hypothesis tests with corresponding policy decisions. This is what ultimately allows us to select between alternative testing procedures based on their welfare consequences. The test/decision correspondence is particularly clear in

[1] As this suggests, our focus is on experimental research. We discuss relationships to non-experimental work in Section 4.4.

[2] While we focus on the research publication process, our framework can also be used to study MHT in other contexts such as regulatory approval as analyzed in Tetenov (2016).

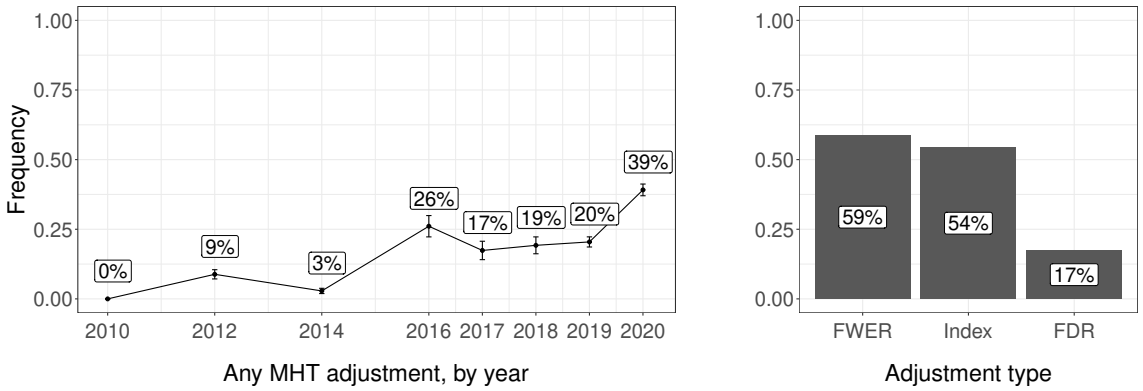


Figure 4.1. Multiple testing adjustments in "top-5" experimental publications. The left-hand panel reports the share of experimental papers that report at least one multiple hypothesis testing adjustment, including both indexing and control of compound error rates, by year of publication. The right-hand panel reports the frequency of each adjustment type, pooling across years. Adjustment types are not mutually exclusive. Authors' calculations based on a review of papers published in the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, the *Quarterly Journal of Economics*, and the *Review of Economic Studies*. We exclude articles from the American Economic Review Papers and Proceedings.

situations where researchers study multiple *interventions* or effects on multiple *sub-populations*.[3] We therefore begin in Section 4.2 by studying these cases. We first establish that (under certain conditions) separate hypothesis tests based on threshold-crossing protocols—in particular, the t-tests that are ubiquitous in applied work—are optimal. We then draw two broad conclusions about the role of multiplicity.

First, it is often optimal to adjust testing thresholds (i.e., critical values) for the number of hypotheses being tested. A loose intuition for this result is as follows. The worst states of the world are those in which the status quo of no intervention is best, as, in these states, a research study has only downside. Here it is desirable to keep publication probabilities low enough that the researcher chooses not to experiment. If the hypothesis testing rule were invariant to the number of hypotheses being tested, then for sufficiently many hypotheses, this condition

[3]This is common in practice. For example, 27 of 124 field experiments published in "top-5" journals between 2007 and 2017 feature factorial designs with more than one treatment (Muralidharan et al., 2020). Moreover, the average number of subgroups analyzed in the 34 field experiments published in top economics journals between 2005 and 2009 is 6.4 (Fink et al., 2014).

would be violated: the chance of getting a study published due solely to false positives would be high enough that the researcher would choose to conduct one. Some adjustment for hypothesis count is thus optimal. We believe that this logic aligns fairly closely with the lay intuition that researchers should not be allowed to test many hypotheses and then "get credit" for false discoveries.

Second (and as this suggests), the research cost function determines whether and how much adjustment is required. We can, in fact, pick a research cost function such that *no* further adjustment is required, as the costs of doing research scale with the number of hypotheses tested in just such a way as to "build in" the needed correction. More generally, optimal testing rules compensate for residual imbalances in researcher incentives with respect to the number of hypotheses, taking the researcher's costs into account. As a result, the framework can explain when common criteria emerge as appropriate solutions and when they do not, as a function of the economic environment. When research costs are fixed, for example, it is optimal to control the average size of tests (e.g., via a Bonferroni correction), while when costs scale in exact proportion to the number of tests *no* MHT adjustment is required.

In our base model, the interaction between hypotheses is driven by interdependence in the costs of conducting research on them. We also consider a series of extensions in which there is successively more scope for additional interactions of other kinds. Specifically, we introduce interactions in the *publication process* in the form of a threshold rule where papers are published if they find enough results, and then in the *economic effects* of the treatments being evaluated—allowing, for example, for complementary interventions.[4] These extensions yield some interesting results, including (in one case) FWER control at the level of *groups* of hypotheses sufficient for publication (as opposed to the individual hypotheses).

In Section 4.3 we turn to the case of multiple *outcomes*. Here a prerequisite step is to specify whether and how these outcomes inform multiple *decisions*. One interpretation is that

[4]In these more general settings, maximin optimality can be very conservative; we, therefore, focus on a weaker notion of maximin optimality, corresponding to weak size control in the literature.

the research informs an audience of multiple policy-makers with heterogeneous preferences over outcomes, in the spirit of Andrews and Shapiro (2021). One education minister may care more about literacy gains, for example, while another cares more about numeracy. If it is unknown *which* policy-maker will act on the results of the paper and (as before) we maximize worst-case welfare, then optimal hypothesis testing rules again adjust for the number of outcomes. But in this case, maximin rules can be very conservative: we show that in our leading case, separate t-testing is maximin only if it has *zero* power, for example, and that maximin tests with non-zero power include testing based on the minimum of a group of t-statistics.

A reasonable alternative interpretation of the multiple outcomes case is that several outcomes usually matter even for a single decision. An education minister might reasonably wish to examine effects on both literacy and numeracy, for example, when deciding whether to scale up an intervention. Under this interpretation, optimal recommendations are based on a test for effects on a weighted average of the outcomes—that is, on indexing. The optimal indexing procedure depends, in turn, on the meaning of the outcomes. If (as in the literacy/numeracy case) they capture distinct arguments in the welfare function, then optimal weights should represent the *economic* preferences of the decision-maker. If alternatively, they are best understood as alternative measures of or proxies for the *same* underlying concept, then optimal weights minimize the *statistical* variance of the resulting index (in the spirit of the approach proposed by Anderson (2008)). A natural example of the latter case would be observing the effects on two different noisy tests of literacy.

Our paper draws inspiration from other work using economic models to inform the choice of statistical procedures. The audience for research results plays a central (albeit passive) role, as in recent work on scientific communication (e.g., Frankel and Kasy, 2018; Andrews and Shapiro, 2021). More generally, the preferences and incentives of researchers drive the analysis (e.g., Chassang et al., 2012; Tetenov, 2016; Spiess, 2018; Henry and Ottaviani, 2019; Yoder, 2019; Banerjee et al., 2020; McCloskey and Michaillat, 2020; Williams, 2021). Our contribution is to adapt and apply this perspective to the analysis of multiple hypothesis testing specifically.

In doing so, we aim to provide some practical guidance to applied researchers navigating the extensive statistical literature on MHT. This literature focuses on the design of algorithmic procedures for controlling particular notions of compound error.[5] We refer to Efron (2008a) and Romano et al. (2010) for overviews. We draw in particular on List et al. (2019)'s helpful distinction between different types of multiplicity and show how these distinctions lead to meaningful differences in optimal testing procedures.[6] Overall, while the literature provides many different MHT procedures, few statistical optimality results exist (e.g., Lehmann et al., 2005; Romano et al., 2011). We, therefore, discuss in Section 4.4 what takeaways for practice one might reasonably draw from our theoretical results—recognizing that this inevitably requires judgment, taking into account important considerations which are (necessarily) outside of our model. Practitioners may wish to skip directly to this section, which is intended to be relatively self-contained.

Our paper also relates to an extensive literature at the intersection between decision theory and hypothesis testing, dating back to Wald (1950) and Robbins (1951) and, more recently, Manski (2004) and Tetenov (2012) in the context of statistical treatment choice.[7] Previous work has motivated notions of compound error control in single-agent non-strategic environments; see, for example, Storey (2003) and Efron (2008b) for a Bayesian interpretation of the FDR and Lehmann and Romano (2005b) for a discussion of the FWER. We complement this literature by developing a model of the publication process that explicitly incorporates the incentives and constraints of the researchers. Relative to the decision-theoretic approach, this has two main advantages. First, it lets us characterize *when* MHT adjustments are appropriate—and also

[5]See, e.g., Holm (1979); Westfall and Young (1993); Benjamini and Hochberg (1995); Benjamini and Liu (1999); Storey (2002); Storey et al. (2004); Lehmann and Romano (2005a); Lee and Shaikh (2014); Romano and Wolf (2016); List et al. (2019); **?** among many others.

[6]There are other useful taxonomies in the literature. For example, Rubin (2021) distinguishes between settings where the joint intersection null is rejected if one hypothesis is rejected (disjunction testing), settings where all individual nulls need to be rejected to reject the joint union null (conjunction testing), and individual testing. We will briefly mention the connection of our results to disjunction testing in Section 4.3.4.

[7]The literature on statistical treatment choice mostly focuses on non-strategic planners' problems. See also Hirano and Porter (2009); Kitagawa and Tetenov (2018); Hirano and Porter (2020b); Athey and Wager (2021) for recent contributions.

when they are *not*—as a function of measurable features of the research and publication process. Second, it allows us to justify and discriminate between different notions of compound error (e.g., FWER and FDR) in the same framework based on these same economic fundamentals.

Multiple Interventions and Multiple Subgroups

Our goal is to understand what professional norms with respect to MHT adjustments lead to desirable welfare outcomes. To this end, we consider a model of the academic publication process and ask what statistical procedures a benevolent social planner would choose, taking into account a representative experimental researcher's incentives. The planner's problem is most analogous to that of the editorial board of an academic journal choosing editorial standards regarding the use of hypothesis testing procedures. We abstract from other factors that would be important for a positive description of the academic publication process, including the incentives of journal editors (e.g. maximizing the journal's impact factor) and competition between journals.[8]

In our model, multiple testing issues arise whenever research informs multiple policy decisions. We therefore start by discussing settings with multiple *interventions* or different *sub-populations*, as here there is a clear one-to-one mapping between multiple hypothesis tests and multiple policy decisions. The case of multiple *outcomes* is more subtle. In particular, there may not be a one-to-one mapping between hypothesis tests and policy decisions since multiple tests may only inform one decision. We will turn to this case in Section 4.3.

Setup and Model

The social planner prescribes and commits to a hypothesis testing protocol, restricting how the researcher can report discoveries and make recommendations. Given this protocol, the researcher decides whether to run an experiment with $T \geq 1$ different non-exclusive treatments.

[8]We can also incorporate some of these (e.g. the role of space constraints in journals in Section 4.2.6 and Appendix D.1.3).

Treatments may represent either different interventions or different sub-populations to whom a given intervention might be applied. What is important is that there are $J = T$ distinct and non-exclusive policy decisions to be made. Throughout this section, to simplify the exposition, we will not explicitly distinguish between T and J and write J everywhere.

If the researcher experiments, she draws a vector of statistics $X \sim F_\theta$, and incurs a cost $C(J) > 0$, which may depend on the number of treatments. Here, $X \in \mathscr{X} \subseteq \mathbb{R}^J$ and $\theta \in \Theta$ is the parameter of interest. The costs $C(J)$ are sunk after the experiment is conducted and do not depend on θ. We assume that every experiment is written up and submitted to the journal. Research designs (defined by J and F_θ) arise exogenously. Our main results continue to hold when we relax this assumption, with the most-powerful characterization requiring a standard symmetry assumption; see Appendix D.1.2. This captures situations where, for example, researchers collaborate with implementation partners such as NGOs who present them with the opportunity to work on an evaluation whose parameters are largely fixed by the partner's capacity or the size of the population it serves. Research designs are endogenous in the broad sense that the social planner's choice of a hypothesis testing protocol determines which designs are implemented and published.

The researcher reports results in the form of a vector of *discoveries* or *recommendations*

$$r(X;J) = (r_1(X;J), \ldots, r_J(X;J))^\top, \tag{4.1}$$

where $r_j : \mathscr{X} \mapsto \{0,1\}$ and $r_j(X;J) = 1$ if and only if treatment j is recommended. We will refer to r as *recommendation function* or *hypothesis testing protocol*. In Equation (4.1) we assume that researchers report all the J tests that they conduct, abstracting from p-hacking and selective reporting.[9] This captures settings with pre-analysis plans, which are becoming more common in experimental research in economics (see, e.g., Miguel, 2021, Figure 2).

[9]To accommodate p-hacking and selective reporting, one could model researcher behavior as, for example, in McCloskey and Michaillat (2020). They propose critical values for single hypothesis tests that control size when researchers conduct studies until they find significant results or research becomes too costly.

The social planner chooses the types of statistical test(s) that the researcher may employ by selecting $r \in \mathscr{R}$, where $\mathscr{R}$ is a pre-specified and exogenous class of functions. Unless otherwise specified, we do not impose any restrictions on $\mathscr{R}$ other than pointwise measurability. Our focus will be on studying how optimal recommendation function(s) vary as a function of variation in the number J of treatments being tested.

Welfare depends on the researcher's recommendations. Specifically, we assume that the recommended (combination) of treatments is implemented in a target population by a policy-maker who is otherwise a passive player.[10] We assume that the target population is independent of the experimental sample but subject to the same data-generating process.[11] The policy-maker implements the researcher's recommendations irrespective of whether the paper gets published; we think of this as capturing the idea that the paper will eventually be published somewhere, and that policy-makers do not discriminate between papers based on the academic prestige of the outlet. If, on the other hand, the researcher does not experiment, the status quo is implemented.

To help simplify our specification of welfare, we introduce the *selector* function δ that indicates which of the $2^J - 1$ possible *combinations* of treatments is recommended,

$$\delta(r(X;J)) \in \{0,1\}^{2^J-1}, \quad \text{where } \sum_{k=1}^{2^J-1} \delta_k(r(X;J)) \in \{0,1\}. \tag{4.2}$$

If no treatment is implemented (i.e. $\sum_k \delta_k(r(X;J)) = 0$) then the status quo is maintained. For $k = 1, \ldots, 2^J - 1$, let $u_k(\theta)$ denote the social welfare generated by the combination of treatments $\delta_k(r(X;J)))$. We write the welfare effect of implementing the treatments recommended by the researcher concisely as $u(\theta)^\top \delta(r(X;J)))$, where $u(\theta) = \big(u_1(\theta), \ldots, u_{2^J-1}(\theta)\big)^\top$.

To derive our main results, we assume that welfare is additive.

Assumption 4.2.1 (Additive welfare)**.** Suppose that, for $k = 1, \ldots, J$, $u_k(\theta) = \theta_k$, and $u_{J+1}(\theta) =$

[10] To simplify notation we assume that the recommendation is always implemented, but our framework directly extends if (more realistically) the recommendation increases the probability of implementation.

[11] These are standard assumptions (e.g., Manski, 2004; Kitagawa and Tetenov, 2018; Athey and Wager, 2021).

$\theta_1 + \theta_2$, $u_{J+2}(\theta) = \theta_1 + \theta_3$, ..., $u_{2^J-1}(\theta) = \sum_{j=1}^{J} \theta_j$.[12]

Under Assumption 4.2.1, the welfare gains from implementing a combination of treatments is equal the sum of the welfare gains from implementing them individually. This is the case, for example, when the treatments are very different so that interaction effects are unlikely, or when each treatment corresponds to treating a different sub-population and there are no cross-group spillovers. In Section 4.2.6 and Appendix D.1.3, we extend our analysis to general welfare functions that allow for interaction effects.

We will often return to the following running example (or variants of it) in which the researcher studies $J = 2$ treatments using a linear regression model.

Example 4.2.1 (Running example)**.** Consider the problem of studying the effect of $J = 2$ non-exclusive experimental treatments D_1 and D_2 on an outcome of interest Y based on a sample with N observations. Suppose that

$$Y_i = \theta_1 D_{i,1} + \theta_2 D_{i,2} + \varepsilon_i, \;\; i = 1, \ldots, N, \;\; \varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \tag{4.3}$$

where σ^2 is known.[13] For simplicity, the baseline average outcome is normalized to zero, and θ_1 and θ_2 are the average treatment effects of D_1 and D_2 net of the costs of implementation. Under these assumptions, $\hat{\theta} \sim \mathcal{N}(\theta, \Sigma)$, where $\hat{\theta} = \left(\hat{\theta}_1, \hat{\theta}_2\right)^\top$ is the OLS estimator of $\theta = (\theta_1, \theta_2)^\top$, and the covariance matrix Σ is known. This fits into our framework by setting $X = \hat{\theta}$ in which case F_θ is the CDF of a $\mathcal{N}(\theta, \Sigma)$ distribution. Note that F_θ depends on the sample sizes in the experimental design, which we treat as exogenous.

The recommendation $r(X;2)$ can take four values: $r(X;2) = (0,0)$ (recommend baseline), $r(X;2) = (1,0)$ (recommend D_1), $r(X;2) = (0,1)$ (recommend D_2), and $r(X;2) = (1,1)$ (recommend D_1 and D_2). The selector δ is defined as $\delta((0,0)) = (0,0,0)$, $\delta((1,0)) = (1,0,0)$,

[12] Absent additional restrictions, assuming that the welfare effect of only implementing treatment j is θ_j is without loss of generality since we can always redefine the parameter of interest such that this assumption holds.

[13] This model is similar to the one studied in Section 5.5 of Elliott et al. (2015). The exact normality assumption is imposed for simplicity. Similar results follow from standard asymptotic approximations provided that N is large enough.

$\delta((0,1)) = (0,1,0)$, and $\delta((1,1)) = (0,0,1)$. A typical choice of the recommendation function (which will be shown to be optimal under some conditions) is

$$r(X;2) = \left(1\left\{\frac{\hat{\theta}_1}{\sqrt{\Sigma_{1,1}}} \geq t\right\}, 1\left\{\frac{\hat{\theta}_2}{\sqrt{\Sigma_{2,2}}} \geq t\right\}\right)^\top, \tag{4.4}$$

where the threshold t is chosen (optimally) by the social planner. That is, the recommendations are based on standard (one-sided) t-tests. □

We introduce two asymmetries in our model that are essential for justifying hypothesis testing, i.e., a protocol r that imposes a size control criterion. First, following Tetenov (2016) and Di Tillio et al. (2017), we impose asymmetry in the incentives. While the social planner maximizes welfare, the researcher's utility does not depend on welfare but instead depends on the expected publication prospects of a research project. If instead the incentives of the researcher and the social planner were aligned then hypothesis testing of any kind would be unnecessary.[14] The key assumption is that some (not necessarily all) researchers have missaligned preferences; we show in Appendix D.2.8 that all our results go through unchanged if some do not. Second, we assume that there is asymmetric information: the parameter θ is known to the researcher but unknown to the social planner, who chooses professional norms and commits to them ex ante. In Appendix D.1.1, we show that our main results continue to hold under imperfect information when the researcher has a prior about the parameter θ.

The (expected) researcher's utility of conducting the experiment depends on the costs of doing so and on the probability that this results in a publication, which in turn is a function of the number of the discoveries. We assume that the publication probability is exogenous. In doing so are abstracting from several features that would be important for a positive description of the academic publication process, including editor and referee behavior and incentives. The

[14] Namely, the optimal planner strategy would be not to impose any restriction on r. In this case the researcher would always report as true discoveries the ones corresponding to positive parameter's values.

researcher experiments if the expected utility, $\beta_r(\theta)$, is positive, where

$$\beta_r(\theta) = \int \rho\left(\sum_{j=1}^{J} r_j(x;J)\right) dF_\theta(x) - C(J). \tag{4.5}$$

Here $\rho : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is an exogenous and weakly increasing function, which captures the idea that the publication prospects of a paper depend on its expected influence on policy. This assumption is consistent with the notion that referees and researchers more generally treat rejections of the null as mattering.[15] That said, our framework can readily accommodate settings where the status quo view is that a policy is effective, and it is desirable to motivate researchers to find and report "null results" (i.e., evidence that it is not). We can capture such cases by flipping the interpretation of the treatment to be "no policy" instead of "the policy" and interpreting the treatment effect as being net of the cost of implementing the policy.

We first assume that $\beta_r(\theta)$ is linear in the number of discoveries ($\rho(x) = x$), and then examine a threshold-crossing publication rule ($\rho(x) = \gamma \cdot 1\{x \geq \kappa\}$) in Section 4.2.6.

Assumption 4.2.2 (Linear publication rule)**.** The researcher's utility conditional on experimenting is, up-to-rescaling by J,

$$\beta_r(\theta) = \int \sum_{j=1}^{J} r_j(x;J) dF_\theta(x) - C(J), \quad C(J) \leq J. \tag{4.6}$$

In Assumption 4.2.2, we interpret $\int \sum_{j=1}^{J} r_j(x;J) dF_\theta(x)$ as the publication probability, multiplied by a factor J. Thus, this assumption describes a setting where the publication probability is linear in the number of discoveries. The condition that $C(J) \leq J$ implies that there is at least one case (when all discoveries are reported with probability one) under which it is profitable for the researcher to experiment. Without such an upper bound, the researcher would never experiment. Assumption 4.2.2 allows the publication prospects to depend on factors

[15] For example, Andrews and Kasy (2019) estimate that rejections at the 5% level increase the chances of publication by over 30%.

beyond the number of discoveries, such as novelty, quality, and unobserved journal preferences, as long as these factors are independent of X.[16]

The following tie-breaking assumption simplifies our analysis by avoiding multiple equilibria.[17]

Assumption 4.2.3 (Tie-breaking rule)**.** Whenever the researcher is indifferent between experimenting or not, she makes the choice that yields the highest social welfare.

The social planner commits to a hypothesis testing protocol r, taking into account the best response of the researcher. Welfare depends on the protocol r and the parameter θ:

$$v_r(\theta) = \begin{cases} \int u(\theta)^\top \delta(r(x;J)) dF_\theta(x) & \text{if } \beta_r(\theta) > 0, \\ \max\left\{ \int u(\theta)^\top \delta(r(x;J)) dF_\theta(x), 0 \right\} & \text{if } \beta_r(\theta) = 0, \\ 0 & \text{if } \beta_r(\theta) < 0. \end{cases} \tag{4.7}$$

The second case ($\beta_r(\theta) = 0$) follows from Assumption 4.2.3.

Note that the structure of the model we study naturally justifies one-sided hypothesis testing, as researchers test whether or not a proposed treatment improves upon the status quo. In Appendix D.2.1 we outline ways to change the structure of the model in order to justify two-sided hypothesis testing.

We now study the optimal choice of hypothesis testing protocols. We first review the single hypothesis case ($J = 1$) previously analyzed by Tetenov (2016) and then consider the case with multiple hypotheses.

Review with a Single Hypothesis

Tetenov (2016) considers a game between an informed agent and a regulator, which is relevant, for instance, in the context of drug approvals. We explain his results using the

[16] This aspect of Assumption 4.2.2 is similar to the stylized model of the publication process with a single hypothesis in Brodeur et al. (2016).

[17] A similar assumption is imposed in Kamenica and Gentzkow (2011).

terminology of our framework. Without loss of generality, suppose that $u(\theta) \in [-1,1]$. Define the *null space* of parameter values as the set of parameters such that implementing the (single) treatment being studied would reduce welfare, $\Theta_0 := \{\theta : u(\theta) < 0\}$. Similarly, define the *alternative space* of parameter values as the set of parameters such that the treatment increases welfare $\Theta_1 := \{\theta : u(\theta) \geq 0\}$. Social welfare is $v_r(\theta) = u(\theta)$ if $\int r(x)dF_\theta(x) \geq C$, where we write $r(x) := r(x;1)$ and $C := C(1)$ for simplicity, and zero otherwise. That is, welfare is non-zero if the expected utility from experimenting, $\int r(x)dF_\theta(x)$, is larger than the cost of experimentation.

To justify single hypothesis testing, Tetenov (2016) focuses on maximin optimal recommendation functions, i.e., recommendation functions that maximize worst-case welfare,

$$r^* \in \arg\max_{r\in\mathscr{R}} \min_{\theta\in\Theta} v_r(\theta).$$

The focus on maximin recommendation functions is important for justifying standard hypothesis testing. Intuitively, the worst-case nature of the maximin criterion induces the social planner to treat size control and power asymmetrically. Standard hypothesis testing will not generally be optimal under alternative optimality criteria.

Proposition 1 in Tetenov (2016) demonstrates that a recommendation function is maximin optimal if and only if

$$\int r^*(x)dF_\theta(x) \leq C \quad \text{for all } \theta \in \Theta_0. \tag{4.8}$$

This result shows that maximin optimal recommendation functions are such that the researcher does not find it worthwhile experimenting whenever the treatment is welfare-reducing ($\theta \in \Theta_0$). For this to hold, the probability of publication in this state (given by the left-hand side of (4.8)) must be sufficiently low. The model thus rationalizes error control, i.e. control of the probability of falsely rejecting the null that the status quo of no treatment is best.

To select among the many alternative maximin recommendation functions, Tetenov

(2016) provides admissibility results under an additional monotone likelihood ratio property. He shows that admissible recommendation functions satisfy the following condition

$$\int r^*(x)dF_0(x) = C, \tag{4.9}$$

with the recommendation function taking the form of a threshold crossing protocols, $r(X) = 1\{X \geq t^*\}$. This result provides a formal justification for standard (one-sided) tests with conventional critical values.

Maximin Protocols and Size Control

In this section, we discuss the problem in the general case with $J > 1$ treatments. We have seen in Section 4.2.2 that the social planner maximizing worst-case welfare is important for justifying single-hypothesis testing. Since analyzing multiple testing requires a framework for justifying hypothesis testing in the first place, we will focus on maximin optimal recommendation functions; that is, recommendation functions that maximize worst-case welfare[18],

$$r^* \in \arg\max_{r \in \mathscr{R}} \min_{\theta \in \Theta} v_r(\theta).$$

We denote the set of maximin recommendation functions by $\mathscr{M}$.

Define the *null space*, the set of parameters such that welfare is weakly negative regardless of the choice of r, as follows.

Definition 4.2.1 (Null space)**.** The null space is $\Theta_0 := \Big\{\theta : u_j(\theta) < 0 \text{ for all } j\Big\}$. □

The following proposition provides a characterization of maximin hypothesis testing protocols r^*, generalizing Proposition 1 in Tetenov (2016) to the case of $J > 1$ hypotheses. It

[18]Maximin hypothesis testing protocols are attractive in settings where there are concerns about researchers experimenting with treatments that may hurt (groups of) individuals. To this end, we will show below that maximin protocols discourage experimentation when treatments have negative welfare impacts. Moreover, focusing on maximin rules allows us to obtain concrete recommendations that do not depend on the planner's prior about θ, which would be difficult to conceptualize in practice.

shows that our definition of the null space is directly connected to maximin optimality.

Proposition 4.2.1 (Maximin protocols)**.** *Let Assumption 4.2.3 hold and suppose that* $\Theta_0 \neq \emptyset$*. A recommendation function* r^* *is maximin-optimal, i.e.,*

$$r^* \in \arg\max_{r \in \mathscr{R}} \min_{\theta \in \Theta} v_r(\theta), \tag{4.10}$$

if and only if

$$\beta_{r^*}(\theta) \leq 0 \quad \forall \theta \in \Theta_0 \quad \text{and} \quad v_{r^*}(\tilde{\theta}) \geq 0 \quad \forall \tilde{\theta} \in \Theta \setminus \Theta_0. \tag{4.11}$$

Proof. See Appendix D.3.2. □

Proposition 4.2.1 shows that maximin optimality is equivalent to two conditions. First, as in the case with $J = 1$ hypotheses, maximin recommendation functions depend on the researcher's utility $\beta_r(\theta)$, and deter experimentation over Θ_0, where *all* treatments reduce welfare. Second, the welfare for $\theta \in \Theta \setminus \Theta_0$ must be non-negative. This second condition requires that if some treatments reduce the welfare, there must be other treatments that compensate them. The first condition captures a notion of size control. We show below that the second condition is non-binding for one-sided t-tests in the leading case where X is normally distributed. Note that Proposition 4.2.1 applies very generally, without relying on any particular functional form assumption on the researcher's utility.

We illustrate the definition of the null space and the characterization of maximin protocols in our running example.

Example 4.2.2 (Running example continued)**.** In our running example, the null space is $\Theta_0 = \{\theta \in \Theta : \theta_1 < 0 \text{ and } \theta_2 < 0\}$. Figure 4.2 provides a graphical illustration. By Proposition 4.2.1, a recommendation function $r^* = (r_1^*, r_2^*)^\top$ is maximin only if (but not necessarily if)

$$P(r_1^*(X;2) = 1|\theta_1, \theta_2) + P(r_2^*(X;2) = 1|\theta_1, \theta_2) \leq C(2), \quad \theta_1 < 0, \theta_2 < 0. \tag{4.12}$$

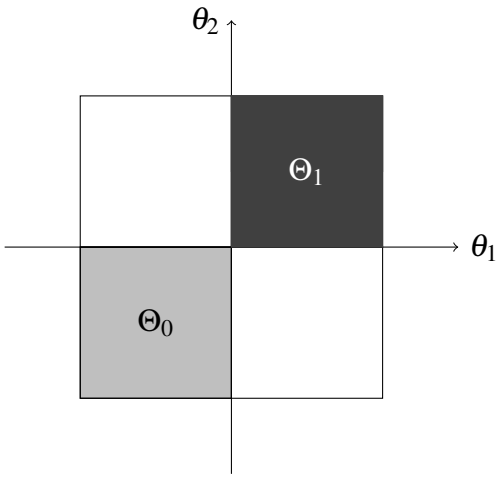


Figure 4.2. Graphical illustration with two hypotheses. Illustration of the null space $\Theta_0 = \{\theta \in \Theta : \theta_1 < 0 \text{ and } \theta_2 < 0\}$ and the alternative space $\Theta_1 = \{\theta \in \Theta : \theta_1 \geq 0 \text{ and } \theta_2 \geq 0\}$. See Remark 14 for a discussion of the two orthants where the coefficients have different signs.

Equation (4.12) shows that maxmin recommendation functions impose restrictions on size control (i.e., the probability of reporting a false discovery). □

Remark 14 (Null space)**.** The definition of the null space Θ_0 corresponds to the *global null hypothesis* in the literature. It is a subset of the *strong null space* $\tilde{\Theta}_0 = \{\theta : \theta_j < 0 \text{ for some } j\}$. We note that $\tilde{\Theta}_0$ plays an important role in the second condition of Proposition 4.2.1 ($v_r(\tilde{\theta}) \geq 0 \; \forall \tilde{\theta} \in \Theta \setminus \Theta_0$). Since $v_r(\tilde{\theta}) \geq 0$ for all $\tilde{\theta} \in \{\theta : \theta_j \geq 0 \text{ for all } j\}$ by definition, this condition is equivalent to assuming that welfare is positive for $\tilde{\theta} \in \tilde{\Theta}_0 \setminus \Theta_0$. □

Admissibility and Power

The set of maximin recommendation functions contains infinitely many elements, some of which may be very conservative. An example is the function that forces the researcher not to report any discoveries ($r_j(X;J) = 0$ for all j). This motivates the use of additional criteria for choosing among them. In other words, among recommendation functions that minimize the social planner's downside, how might she select those with large upside?

We proceed as follows. First, we show that no maximin recommendation function

dominates all other recommendation functions when $J > 1$. This is in sharp contrast to the single-hypothesis case in Section 4.2.2, where dominant recommendation functions exist in some cases.[19] Second, motivated by this result, we introduce a notion of local power and show that locally most powerful recommendation functions are admissible. Finally, we provide explicit characterizations of locally most powerful recommendation functions.

We start by defining a suitable notion of weak dominance.

Definition 4.2.2 (Weak dominance[20])**.** The recommendation function r *weakly* dominates r' if $v_r(\theta) \geq v_{r'}(\theta)$ for all $\theta \in \Theta$. A dominant recommendation function r over a set $\tilde{\mathscr{R}} \subseteq \mathscr{R}$ is a recommendation function that *weakly* dominates all other recommendation functions $r' \in \tilde{\mathscr{R}}$. □

Recall that maximin recommendation functions discourage experimentation, leading to zero welfare for $\theta \in \Theta_0$. Therefore, dominance among such rules is determined by their relative performance in $\Theta \setminus \Theta_0$ and, thus, captures a notion of power in our setting.

To derive the remainder of our results, we focus on the leading case where X is normally distributed.

Assumption 4.2.4 (Normality)**.** The statistic X is normally distributed, $X \sim \mathscr{N}(\theta, \Sigma)$, where Σ is positive definite.

The following proposition shows that there exist data-generating processes such that no maximin recommendation function weakly dominates all other recommendation functions. This result holds in particular when X is normally distributed.

Proposition 4.2.2 (No recommendation function dominates the others)**.** *Let $J > 1$. Let Assumptions 4.2.1, 4.2.2, and 4.2.3 hold. Then there exists a parameter space $\Theta \subseteq [-1,1]^J$ and a distribution $\{F_\theta, \theta \in \Theta\}$ such that no maximin recommendation function r (weakly) dominates*

[19] For example, under normality a uniformly most powerful test exist by classical results when interpreting power in terms of welfare effects (e.g., Van der Vaart, 2000, Chapter 15).

[20] This definition is expressed in terms of weak inequality in order to accommodate sets of different functions r having the same probabilities of discoveries.

all other maximin recommendation functions $r' \in \mathscr{M}$ for any cost $0 < C(J) < J$. Moreover, there exists such a distribution F_θ that satisfies Assumption 4.2.4.

Proof. The proof of Proposition 4.2.2 is based on the following observation. For any maximin recommendation function we can find a θ with one (or more) positive component such that the recommendation function is dominated by another function with the property that the positive component has the largest probability of discovery, and the remaining entries have zero probability of discovery. The assumption that $\theta_j \in [-1,1]$ is not restrictive, and the result continues to hold when $\theta_j \in [-M,M]$ for a finite constant $M > 0$. See Appendix D.3.2 for details.[21] □

In the terminology of classical hypothesis testing, Proposition 4.2.2 states that there are settings in which no uniformly most powerful test exists (where in our setting power is measured in terms of implied welfare). It implies that we cannot find recommendation functions that are dominant for all alternatives.

Motivated by this result, we focus instead on ranking maximin recommendation functions within a particular set of *local* (to the null space Θ_0) alternatives.

Definition 4.2.3 (ε-alternatives)**.** For $\varepsilon > 0$, define the local alternative space as

$$\Theta_1(\varepsilon) := \Big\{ \theta : u_j(\theta) \geq \varepsilon \text{ for some } j, u_j(\theta) \geq 0 \text{ for all } j \Big\}.$$

□

The set of ε-alternatives $\Theta_1(\varepsilon)$ is the set of parameters for which, for some policy decision, welfare is strictly positive by at least ε. Note that $\Theta_1(\varepsilon) \cap \Theta_0 = \emptyset$ for all $\varepsilon \geq 0$.

Based on Definition 4.2.3, we introduce the following notion of local power.

[21] Proposition 4.2.2 assumes that $C(J) < J$. If $C(J) = J$ the researcher will always be indifferent between experimenting or not, and we can assign probability one to each discovery without violating the constraint.

Definition 4.2.4 (Locally more powerful)**.** A recommendation function r is locally more powerful (or ε-more powerful) than r' if[22]

$$\liminf_{\varepsilon \downarrow 0} \left\{ \frac{1}{\varepsilon} \inf_{\theta \in \Theta_1(\varepsilon)} v_r(\theta) - \frac{1}{\varepsilon} \inf_{\theta' \in \Theta_1(\varepsilon)} v_{r'}(\theta') \right\} \geq 0. \tag{4.13}$$

□

Definition 4.2.4 introduces a partial ordering of recommendation functions based on their worst-case performance under ε-alternatives. It considers parameter values in an alternative space that contains the origin as $\varepsilon \to 0$. The difference of the utilities is rescaled by the location parameter ε to avoid trivial solutions.[23]

We say that a maximin recommendation function r is locally most powerful (or ε -most powerful) if it is (weakly) locally more powerful than any other maximin recommendation function r'. The following proposition confirms that any maximin and locally most powerful recommendation function is also *admissible* (i.e., not strictly dominated by any other recommendation function). Denote by $\mathscr{E}$ the set of protocols that are locally most powerful and maximin.

Proposition 4.2.3 (Admissibility)**.** *Suppose that* $[-1,1] \in \{u_j(\theta) : \theta \in \Theta\}$ *for all* j*. Then any maximin and locally most powerful recommendation function* r *is admissible with respect to any* $r' \notin \mathscr{E}$*.*

Proof. See Appendix D.3.2. □

The notion of admissibility in Proposition 4.2.3 treats as inadmissible any recommendation function that is dominated by another recommendation function, including recommendation functions that are not maximin. However, note that by Proposition 4.10 maximin recommendation

[22] In Appendix D.3.2, we show that the expression below is uniformly bounded for all (r, r').

[23] Rescaling by the location parameter is common in local asymptotic analyses and is standard practice when making optimality statements (e.g., Athey and Wager, 2021).

functions cannot be dominated over Θ_0 by recommendation functions that are not maximin.[24] Proposition 4.2.3 motivates our focus on maximin and locally most powerful recommendation functions in the following sections.

In Appendix D.2.3, we discuss two alternative notions of power: (i) a version of local power where we let *all* parameters be small and positive and (ii) a criterion inspired by the literature on weighted average power (WAP). We show that both of these arguably natural alternatives are not very useful in our context.

Remark 15 (Trivial maximin decision rules are not admissible)**.** Maximin recommendation functions that are not locally most powerful are not guaranteed to be admissible. For example, the recommendation function $r(X;J) = (0,\dots,0)^\top$ is maximin but not admissible if a non-trivial maximin recommendation function exists (see, for example, Proposition 4.2.5). □

Most Powerful Maximin Protocols

This section provides explicit characterizations and examples of locally most powerful and maximin recommendation functions with an additive welfare (Assumption 4.2.1) and a linear publication rule (Assumption 4.2.2). In Section 4.2.6 we discuss extensions to alternative publication rules and general welfare functions.

The following lemma provides a characterization of the locally most powerful recommendation functions.

Lemma 4.2.4 (Separate size control is locally most powerful)**.** *Let $J > 1$. Let Assumptions 4.2.1, 4.2.2, 4.2.3, and 4.2.4 hold, and let $\Theta = [-1,1]^J$. Then $r^* \in \mathscr{R}$ is maximin optimal and locally most powerful if and only if r^* satisfies Equation* (4.11) *and*

$$P(r_j^*(X;J) = 1 | \theta = 0) = \frac{C(J)}{J} \quad \forall j \in \{1,\dots,J\}, \tag{4.14}$$

[24] Since $\mathscr{E}$ might contain multiple protocols for which (4.13) holds with weak inequality, admissibility is stated with respect to those protocols that are not maximin and locally most powerful.

assuming that such r^* exists.

Proof. See Appendix D.3.1. We note that the proof does not rely on normality of X (Assumption 4.2.4). We only require that X is continuously distributed with CDF F_θ, which admits a PDF $f_\theta(x)$ that is continuous in θ for all $x \in \mathscr{X}$. □

Lemma 4.2.4 states that r^* is locally most powerful if and only if the rejection probability of each test at $\theta = 0$ (i.e. the boundary of the null space) is equal to $C(J)/J$. Condition (4.14) implies that the researcher is indifferent between experimenting and not experimenting and requires that the tests are similar at the boundary of the null space.

Lemma 4.2.4 has three important implications. First, it shows that a recommendation function is locally most powerful and maximin if and only if it imposes size control that is *separate*. That is, given restrictions on the marginal probabilities of rejecting individual hypotheses, no further restrictions are placed on the joint probabilities. Second, locally most powerful maximin recommendation functions impose symmetry across the different hypotheses. Symmetry is a desirable property absent additional restrictions on the relative importance of the different hypotheses.[25] Finally, Lemma 4.2.4 shows that whether and to what extent the level of these separate tests should depend on the number of hypotheses being tested—in other words, whether an adjustment for the presence of multiple hypothesis is required—depends on the structure of the research production function $C(J)$.

To illustrate the practical implications more concretely, it is useful to decompose (without loss of generality) the total research costs $C(J)$ into fixed and variable components c_f and $c_v(J)$:

$$C(J) = c_f + c_v(J). \tag{4.15}$$

Using this notation, the optimal level of size control in Lemma 4.2.4 is $C(J)/J = (c_f + c_v(J))/J$.

Consider first a case in which the research production function exhibits no economies

[25] Symmetry is a consequence of our construction of the local alternative space $\Theta_1(\varepsilon)$. In Appendix D.2.3, we consider an alternative definition of the local alternative space that does not imply symmetry.

of scale with respect to the number of hypotheses being tested. Specifically, suppose there are no fixed costs and that variable costs scale linearly in J, i.e. $c_f = 0$ and $c_v(J) = \alpha J$. Then r^* is locally most powerful if

$$P(r_j^*(X;J) = 1|\theta = 0) = \alpha \quad \forall j \in \{1,\ldots,J\}. \tag{4.16}$$

In other words, standard inference without adjustment for MHT is optimal in this case. While it is true that the researcher obtains a higher expected reward from taking on projects that test more hypotheses, the appropriate correction for this is already "built in" to the costs of conducting research, so that no further correction is required.

Now consider the case in which the research production function exhibits strong economies of scale. Specifically, suppose there are only fixed costs and that the marginal cost of testing an additional hypothesis is zero, i.e. $c_v(J) = 0$ and $c_f = \alpha$. Then r^* is locally most powerful if

$$P(r_j^*(X;J) = 1|\theta = 0) = \frac{\alpha}{J}, \quad \forall j \in \{1,\ldots,J\}. \tag{4.17}$$

In this case the optimal inferential procedure is to control average size, as for example via a Bonferroni correction. This correction is necessary to appropriately align the incentives of the researcher with those of the social planner, as without it she would have a disproportionate incentive to conduct projects with a large number of hypotheses.

An important question is what types of recommendation functions r^* are maximin and locally most powerful. The next proposition shows that under normality the widely-used *threshold-crossing* recommendation functions corresponding to standard one-sided tests are maximin optimal and locally most powerful.

Proposition 4.2.5 (Optimality of separate t-tests)**.** *Assume that the conditions in Lemma 4.2.4*

hold and $C(J) > 0$. Then the recommendation function

$$r_j^*(X;J) = 1\left\{X_j/\sqrt{\Sigma_{j,j}} \geq \Phi^{-1}\left(1 - C(J)/J\right)\right\}, \quad \forall j \in \{1,\ldots,J\}, \tag{4.18}$$

is maximin optimal and locally most powerful.

Proof. See Appendix D.3.2. □

Proposition 4.2.5 shows that standard one-sided t-tests with critical value $\Phi^{-1}\left(1 - C(J)/J\right)$ are optimal in our setting. The critical value depends on the number of hypotheses J and the structure of the cost function $C(J)$.

Remark 16 (Average size control is optimal)**.** Lemma 4.2.4 shows that most powerful maximin rules control average size, $\sum_{j=1}^{J} P(r_j^*(X;J) = 1|\theta = 0) = C(J)$. Many of the popular MHT corrections reviewed in the introduction do not directly target average size control and, thus, will generally not be optimal in our model. This explains why classical Bonferroni corrections are optimal in our model when $C(J)$ is constant, while common refinements of Bonferroni such as Holm (1979)'s method are not. By construction, Bonferroni satisfies average size control, whereas common refinements do not. The optimality of Bonferroni (and average size control) is driven by our choice of the publication rule. Bonferroni corrections may not be optimal with other publication rules. See Section 4.2.6 and Appendix D.1.3. □

Additional Forms of Interactions Between Treatments

So far, we have analyzed settings where treatments interact only via the research cost function. Here we briefly describe extensions to two settings with successively more scope for interaction between hypotheses in each configuration. In the first we continue to assume no economic interactions between interventions but allow for interactions in the publication process through a threshold publication rule. In the second we allow for arbitrary economic interactions between interventions—for example, complementary interventions—as well as a

threshold publication rule. We summarize our results here and refer to Appendix D.1.3 for a detailed discussion.

Linear Welfare and Threshold-Crossing Publication Rule

We introduce interactions in the publication rule by replacing the linear publication rule in Assumption 4.2.2 with a threshold rule in which papers that find sufficiently many results can be published:

$$\beta_r(\theta) = \gamma \int 1\left\{\sum_{j=1}^{J} r_j(x;J) \geq \kappa\right\} dF_\theta(x) - C(J) \tag{4.19}$$

Here the probability of publication is equal to γ if the number of discoveries exceeds κ, so that the paper has "done enough" to be considered for publication, and zero otherwise. Under the threshold-crossing publication rule (4.19), only papers with sufficient influence (i.e. more than κ discoveries) are published—for example, because there are capacity constraints at journals. With a threshold crossing publication rule the incremental value to the researcher of rejecting any given hypothesis may depend on the number of other hypotheses also rejected. The threshold crossing publication rule leads to more complicated optimal hypothesis testing protocols that depend on the joint distribution of X. Therefore, we restrict attention to independent recommendation functions for which $r_j(X;J) \perp r_{j'}(X;J)$ with $j \neq j'$.

For this class of recommendation functions, we show that one-sided t-tests, $r_j^*(X;J) = 1\left\{X/\sqrt{\Sigma_{j,j}} \geq \Phi^{-1}(1-p^*)\right\}$, are still optimal when $X \sim \mathcal{N}(\theta,\Sigma)$, where p^* depends on J, $C(J)$, and γ. When $C(J)/\gamma = \alpha$, which is equivalent to assuming a constant publication probability and constant costs in the number of discoveries, we can show that $p^* \asymp 1/J$ as $J \to \infty$. Thus, asymptotically, fixed-cost research production functions again rationalize Bonferroni corrections.

General Welfare and Threshold-Crossing Publication Rule

To introduce the possibility of economic interactions between the treatments being studied, in addition to interactions in the cost function $C(J)$ and the publication rule, we assume that $u_k(\theta) = \theta_k$ for $k = 1,\ldots,2^J-1$. Importantly, unlike Assumption 4.2.1, this allows for

interactions in the welfare impact of multiple treatments. With $J = 2$, for example, complementarities can be modeled as $u_1(\theta) = \theta_1$, $u_2(\theta) = \theta_2$, and $u_3(\theta) = \theta_3 = \theta_1 + \theta_2 + \zeta$ for some $\zeta > 0$. As a result, we have $X \in \mathscr{X} \subseteq \mathbb{R}^{2^J-1}$ since we are interested in all possible combinations of treatments.

To describe the theoretical results, let

$$\widetilde{\delta}_j(r(X;J)) = \delta_j(r(X;J))1\left\{\sum_{j=1}^{J} r_j(X;J) \geq \kappa\right\} \tag{4.20}$$

indicate the policy decisions taken *only* in the case that the experiment results in a paper that is publishable given the threshold publication rule.

We show that separate size control over each *group* of discoveries is locally most powerful,

$$P\Big(\widetilde{\delta}_j(r^*(X;J)) = 1|\theta = 0\Big) = \frac{C(J)}{\gamma|\mathscr{K}|} \quad \forall j \in \mathscr{K}, \tag{4.21}$$

where $\mathscr{K}$ denotes the set of indexes $k \in \{1,\ldots,2^J-1\}$ that corresponds to groups of κ or more hypotheses, and thus groups that are sufficient for publication. Equation (4.21) can be interpreted as a Bonferroni correction at the level of groups of tests. Any MHT procedure that satisfies Equation (4.21) controls the weak FWER at level $C(J)/\gamma$ at $\theta = 0$, namely

$$P\left(\widetilde{\delta}_j(r^*(X;J)) = 1 \text{ for at least one } j|\theta = 0\right) = C(J)/\gamma.$$

In other words, Equation (4.21) rationalizes (weak) FWER control between *groups* of hypotheses sufficient for publication. When a recommending a single treatment is sufficient for publication ($\kappa = 1$) this implies control of the probability of a single false rejection, i.e. of the standard notion of weak FWER control. More generally, however, FWER is applied not to each separate discovery $r_j(X;J)$ but instead to each group.

Finally, we provide an example of a maximin and locally most powerful recommendation

function. Given the structure of the publication process, it suffices to focus only on groups of treatments with more than κ elements. Formally, consider a vector $\widetilde{X} \in \mathbb{R}^{|\mathscr{K}|}$, where $\widetilde{X}$ is a sub-vector of $X \in \mathbb{R}^{2^J-1}$, with each entry corresponding to a statistic X_j, which itself corresponds to a certain *group* of treatments with more than κ elements. Assume that $\widetilde{X} \sim \mathscr{N}(\theta, I)$, then $\widetilde{\delta}_j(r(X;J)) = 1\left\{X_j > \max_{j' \neq j} \widetilde{X}_{j'} \text{ and } X_j > t\right\} 1\{j \in \mathscr{K}\}$ is maximin and locally most powerful if t is chosen such that $P\left(\max_j \widetilde{X}_j > t | \theta = 0\right) = C(J)/\gamma$. The independence between the entries of $\widetilde{X}$ is important here; without it the existence of an optimal recommendation function is not guaranteed. This example illustrates the complexity of optimal recommendation functions when there are potential interactions between interventions.

Multiple Outcomes

In the framework above we interpret a hypothesis test as a recommendation about a policy decision. Papers that examine multiple *interventions* or impacts within multiple *sub-groups* could clearly inform multiple policy decisions—which of the interventions to implement, or which of the sub-groups to treat. In some cases this gives rise as we have seen to a rationale for MHT adjustments.

With multiple *outcomes*, there may not be such a one-to-one mapping between hypothesis tests and policy decisions. This is because a paper that examines multiple outcomes may or may not inform multiple policy decisions. For example, a paper that measures the impact of a single reform on multiple measures of well-being might be used to guide the single decision whether or not to scale up that reform. In this case our framework interprets the paper as making a single recommendation. This raises the question whether that recommendation should be based on the aggregated results of tests of the individual outcomes, potentially adjusted for multiplicity, or on some other method of aggregating tests or outcomes.

An alternative interpretation of a paper that examines multiple outcomes is that it informs multiple decisions because its audience includes multiple, heterogeneous policy-makers, some

of whom care more about some outcomes than others. In this case we can interpret the paper as making multiple recommendations, and examine whether applying MHT adjustments to those recommendations is desirable.

Throughout this section we examine the case of a single treatment ($T = 1$) in order to focus attention on issues that are specific to multiple outcomes; the results can be extended to multiple treatments at the expense of additional notation.

Setup and Model

Consider an audience of $J \geq 1$ different policy-makers. Each policy-maker decides whether to implement a *single* treatment $(T = 1)$ based on its effects on G different outcomes $Y = (Y_1, \ldots, Y_G)^\top$. Since each policy-maker only makes one policy decision, we will use *policy-maker* and *policy decision* interchangeably. Below we will distinguish between settings with multiple policy-makers ($J > 1$) and a single policy-maker ($J = 1$).

The $G \geq J$ outcomes are associated with statistics $X = (X_1, \ldots, X_G)^\top \sim F_\theta$ which measure the effect of the treatment on those outcomes. For example, suppose that the researcher evaluates the effect of a treatment D on G outcomes, $Y_1, \ldots, Y_G$, using regression models $Y_{i,g} = \theta_g D_i + \varepsilon_{i,g}$ for all $g = 1, \ldots, G$. In this case, θ_g is the treatment effect on the gth outcome, the vector of statistics is $X = \hat{\theta}$, where $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_G)^\top$ is the OLS estimator of $\theta = (\theta_1, \ldots, \theta_G)^\top$, and F_θ is the CDF of a $\mathcal{N}(\theta, \Sigma)$ distribution.

We assume that policy-maker j's welfare is a weighted average of the effects on the different outcomes.

Assumption 4.3.1 (Policy-maker utility)**.** Policy-maker j's welfare is $u_j(\theta) = \theta^\top w_j^*$, for all $j \in \{1, \ldots, J\}$, where $\sum_g w_{g,j}^* = 1$. The weights $\{w_j^*\}$ are common knowledge.

In Assumption 4.3.1, we impose that the weights $\{w_j^*\}$ are common knowledge,[26] but also that they may vary across policy-makers. For example, the education ministry in one country

[26] In Appendix D.2.4 we consider the consequences of uncertainty about the policy-relevant weights and show that—if the social planner applies worst-case logic as above—this leads to extremely conservative testing procedures.

may be more concerned about literacy, while the education ministry in another is more concerned about numeracy.

The researcher makes J recommendations, one for each policy-maker, such that the recommendation function takes the following form

$$r(X;G,J) = (r_1(X;G,J),\ldots,r_J(X;G,J))^\top \in \{0,1\}^J,$$

where $r_j(X;G,J)$ is the recommendation to policy-maker j. Note that the hypothesis testing protocol r may depend on both the number of outcomes G and the number of policy-makers J. Note that this assumption (along with Assumption 4.3.1) may in some cases imply an additional aggregation step, where researchers first aggregate $X_1,\ldots,X_G$ into J indices and then perform tests based on these indices. We discuss this in more detail in Section 4.3.4.

The structure of the researcher utility $\beta_r(\theta)$ is the same as in Section 4.2:

$$\beta_r(\theta) = \int \rho\left(\sum_{j=1}^{J} r_j(x;G,J)\right) dF_\theta(x) - C(G), \tag{4.22}$$

where $\rho : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a weakly increasing function and $C(G)$ denotes the cost of experimentation. In Equation (4.22), the publication prospects depend on the number of policy recommendations, J, whereas the costs of the experiments are determined by the number of outcomes, G. This relaxes the assumption implicit in Equation (4.5) in Section 4.2 that these two quantities must be equal. We will consider different possible functions $\rho(\cdot)$ below.

Each policy-maker j is a passive player who implements the policy recommended by the researcher. Thus, policy-maker j's utility conditional on experimentation is $P(r_j(X)=1|\theta)u_j(\theta)$. The social planner is aware of the different policy-makers who may read and implement the paper's recommendations, but does not know for certain *which* will do so. He thus faces two sources of uncertainty, with respect to both the welfare effects of the treatment and the audience

for evidence on those effects.[27]

The social planner chooses a testing protocol to maximize worst-case welfare over both θ and j (the identity of the implementing policy-maker). Therefore, welfare can be written in compact form as

$$v_r(\theta) = \begin{cases} \min_{j\in\{1,\ldots,J\}} P(r_j(X;G,J)=1|\theta)u_j(\theta) & \text{if } \beta_r(\theta) > 0 \\ \max\left\{ \min_{j\in\{1,\ldots,J\}} P(r_j(X;G,J)=1|\theta)u_j(\theta), 0 \right\} & \text{if } \beta_r(\theta) = 0 \\ 0 & \text{if } \beta_r(\theta) < 0. \end{cases} \tag{4.23}$$

General Characterization of Maximin Protocols

Here we provide a characterization of maximin hypothesis testing protocols that is valid irrespective of the number of outcomes, G, and the number of policy-makers, J.

We begin by characterizing the null space. The null space takes the following form:

$$\Theta_0 := \left\{ \theta : u_j(\theta) < 0, \text{ for some } j \in \{1,\ldots,J\} \right\} \tag{4.24}$$

Intuitively, worst-case welfare is negative if any of the $u_j(\theta)$ is negative. To make explicit the connection to the null space with a single outcome in Section 4.2.2, note that

$$\left\{ \theta : u_j(\theta) < 0, \text{ for some } j \in \{1,\ldots,j\} \right\} = \cup_{j=1}^{J} \left\{ \theta : u_j(\theta) < 0 \right\}, \tag{4.25}$$

where each component in the right-hand side denotes the null space with a single outcome and a single hypothesis. This shows that the null space with multiple outcomes can be very large when there are many outcomes, which we will show leads to conservative testing protocols.

The following proposition provides a general characterization of maximin recommenda-

[27] An alternative interpretation is that every policy-maker decides to implement the treatment in independent populations based on the recommendations in the paper. In this case, these formalisms represent a social planner who desires non-harmful decisions in *each* population.

tion functions.

Proposition 4.3.1 (Maximin optimality)**.** *Let Assumption 4.2.3 hold and suppose that* $\Theta_0 \neq \emptyset$*. A recommendation function* r^* *is maximin-optimal with multiple outcomes, i.e.,*

$$r^* \in \arg\max_{r \in \mathscr{R}} \min_{\theta \in \Theta} v_r(\theta)$$

if and only if

$$\beta_{r^*}(\theta) \leq 0 \quad \forall \theta \in \Theta_0, \tag{4.26}$$

where Θ_0 *is defined in Equation* (4.24) *and* $v_r(\theta)$ *in Equation* (4.23)*.*

Proof. Note that result applies generally, and the proof does not rely on Assumptions 4.2.4 and 4.3.1. See Appendix D.3.2 for details. □

It is interesting to compare the result in Proposition 4.3.1 to that with multiple treatments in Proposition 4.2.1 because, in both settings, the researcher reports a vector of discoveries. The key difference is that the null space with multiple outcomes, Θ_0, contains not only parameter values for which all components are negative but also those for which *some* components are negative.[28] As a result we will see that this setup rationalizes stricter notions of size control than above, which can be conservative, as illustrated in Section 4.3.3.

Multiple Policy-Makers

Here we consider a setting with multiple policy-makers. Specifically, we assume that there are as many policy makers as outcomes, $J = G$. To simplify the exposition, we will not explicitly distinguish between G and J in this section and write J everywhere.

Proposition 4.3.1 can lead to conservative hypothesis testing protocols. To see why,

[28] Using standard hypothesis testing terminology, the null space corresponds to strong error control; see the discussion in Remark 14 and Lehmann and Romano (2005b).

consider a linear publication rule as in Assumption 4.2.2,

$$\beta_r(\theta) = \int \sum_{j=1}^{J} r_j(x;J) dF_\theta(x) - C(J), \tag{4.27}$$

where $r_j(x;J) := r_j(x;J,J)$ and $C(J) = \alpha J$ for some $\alpha \in (0,1)$. Suppose that each policy-maker cares about a different outcome, i.e., $w^*_{j,j} = 1$ for $j \in 1,\ldots,J$ and $w^*_{j,k} = 0$ for $k \neq j$. Proposition 4.3.1 implies that if we restrict the class of feasible recommendation functions to the (common) threshold crossing protocols, power needs to be zero for the rule to be maximin-optimal. Specifically, suppose that $X \sim \mathcal{N}(\theta, I)$, where $\theta_1,\ldots,\theta_J \in [-M,M]$, for some arbitrary large M. Consider the following threshold crossing protocol

$$r(X;J) = \left(1\left\{X_1 \geq t\right\},\ldots,1\left\{X_J \geq t\right\}\right)^\top.$$

Then for a large enough M and $\alpha \leq 1 - 1/J$, the threshold crossing rule is maximin optimal only in the trivial case where $t \to \infty$, which implies that the tests never reject and have zero power.[29]

(Finite) threshold crossing hypothesis testing protocols are not maximin optimal because they fail to discourage experimentation when worst-case welfare is negative. For example, suppose that the treatment has a large positive effect on all but the Jth outcome and a negative effect on the Jth outcome. In this case, the worst-off policy-maker is policy-maker J who cares only about the treatment effect on outcome J, which is negative. However, despite the worst-case welfare being negative, the researcher has an incentive to experiment because she will reject $J-1$ hypotheses with high probability.

This discussion suggests using "worst-case" protocols to ensure that the researcher experiments only if *all* effects are positive. Indeed, when the threshold t is chosen such that

[29] To see this, take $\theta_j \to \infty$ for all $j < J$ and $\theta_J < 0$. Then $\sum_{j=1}^{J} P(r_j(X;J) = 1|\theta_j) = P(r_j(X;J) = 1|\theta_J) + J - 1$ such that we need to impose that $P(r_j(X;J) = 1|\theta_J) \leq \alpha J - J + 1$, where $\alpha J - J + 1 \leq 0$ for $\alpha \leq 1 - 1/J$.

$P(X_j \geq t|\theta_j = 0) = C(J)/J$ for $j = 1,\ldots,J$, the following protocol is maximin optimal:[30]

$$r_j(X;J) = 1\left\{\min_{j\in\{1,\ldots,J\}} X_j \geq t\right\}, \quad \forall j \in \{1,\ldots,J\} \tag{4.28}$$

Due to the use of the minimum across all statistics and the choice of the threshold, the recommendation function (4.28) can be very conservative when there are many outcomes.

An important aspect of Proposition 4.3.1 is that it allows us to directly map the features of the publication process and researcher utility to different types of compound error rate control. Consider next the case of a threshold crossing publication rule, where papers are published if they "have done enough", i.e., if they report more than κ discoveries,

$$\beta_r(\theta) = \gamma \int 1\left\{\sum_{j=1}^{J} r_j(x;J) \geq \kappa\right\} dF_\theta(x) - C(J). \tag{4.29}$$

Here the probability of publication is equal to γ if the number of discoveries exceeds κ. We assume that $\gamma \in (C(J),1)$ such that there is a least one case where experimentation is profitable; otherwise the researcher would never experiment.

Proposition 4.3.1 implies that r^* is maximin if and only if

$$P\left(\sum_{j=1}^{J} r_j(X;J) \geq \kappa|\theta\right) = P\left(\text{at least } \kappa \text{ discoveries} \mid \theta\right) \leq \frac{C(G)}{\gamma} \quad \forall\theta \in \Theta_0$$

This criterion is quite restrictive; it is stronger than and implies strong control of the κ-FWER at level $C(G)/\gamma$. To illustrate, let $\kappa = 1$. Then we impose restrictions not only on the probability of at least one false discovery, but also on the probability of *any* discovery (true or false), whenever the treatment has a negative welfare effect on at least one outcome (i.e. $\theta \in \Theta_0$).

Finally, it is also possible to "invert" our research question and examine what researcher

[30]To see why this decision rule is maximin, note that $\sum_j P(r_j(X;J) = 1|\theta_j) = J\prod_{j=1}^{J} P(X_j \geq t|\theta_j)$. For some $j' \in \{1,\ldots,J\}$, let $\theta_j \to \infty$ for all $j \neq j'$, and $\theta_{j'} \leq 0$. It follows that the expression is bounded from above by $JP(X_{j'} \geq t|\theta_{j'} = 0) = C(J)$.

incentives and features of the publication process rationalize other popular criteria such as control of the FDR. Interestingly, it turns out that rationalizing FDR control in our model requires us to assume that the researcher is malevolent. Suppose that, as above, $w^*_{j,j} = 1$ for $j \in 1, \ldots, J$ and $w^*_{j,k} = 0$ for $k \neq j$. If

$$\beta_r(\theta) = \int \left[\sum_{j=1}^{J} \frac{1\left\{\theta_j < 0\right\} r_j(x;J)}{\sum_{j=1}^{J} r_j(x;J)} \cdot 1\left\{ \sum_{j=1}^{J} r_j(x;J) > 0 \right\} \right] dF_\theta(x) - C(J), \tag{4.30}$$

any decision rule that controls the FDR under the null hypothesis Θ_0 at level $C(J)$ is maximin optimal. Equation (4.30) imposes that the researcher is malevolent in the sense that her utility is increasing in the number of false discoveries. We interpret this result as suggesting that FDR control does not arise as a natural solution in our frequentist maximin framework. As we will discuss in Section 4.4, however, we see other cogent arguments for FDR control once one is willing to move outside such frameworks and consider decision-making from a Bayesian perspective.

Single Policy-Maker

In this section, we turn to the case of a single policy-maker ($J = 1$), indexed by $j = 1$. In this case, there is only one policy decision to be made such that $r_1(X;G) \in \{0,1\}$, where $r_1(X;G) := r_1(X;G,1)$, and the researcher utility becomes

$$\beta_r(\theta) = \int r_1(x;G) dF_\theta(x) - C(G).$$

Throughout this section, we will omit the policy-maker's index $j = 1$ to simplify the notation.

An immediate implication of the general result in Proposition 4.3.1 is that, if $X \sim \mathcal{N}(\theta, \Sigma)$, one-sided t-tests based on a weighted average of the outcome-specific statistics $X^\top w$

are maximin optimal for any choice of weights w summing to one,

$$r(X;G) = 1\left\{\frac{X^\top w}{\sqrt{w^\top \Sigma w}} > \Phi^{-1}(1-C(G))\right\}. \tag{4.31}$$

The choice of the optimal weights (i.e., the weights leading to the most powerful recommendation functions) depends on our interpretation of the outcomes Y, the corresponding statistics X, and the parameter space Θ. We consider two alternative interpretations. In the first, each entry of X measures impacts on distinct arguments in the policy-maker's utility function. For example, a study might report the effects of a single micro-credit intervention on livelihoods and women's empowerment, both of which matter independently to the policy-maker. Second, we assume that each entry of X is a distinct measure of the *same* underlying welfare-relevant outcome. This describes settings in which there are multiple ways of measuring the same construct, as for example the dual income- and expenditure-based measures of the transfer multiplier reported by Egger et al. (2020). It is also closely related to "latent variable" models in which multiple outcomes are interpreted as proxies for some deeper, underlying construct of interest, e.g. multiple measures of "health" in medical studies. See Appendix D.2.7.

We start by consider a setting where each entry of X measures impacts on economically distinct outcomes. We make the following assumption.

Assumption 4.3.2 (Economically distinct outcomes)**.** $\Theta = [-1,1]^G$

Assumption 4.3.2 encompasses two separate requirements. First, we normalize the parameter space of θ_g to $[-1,1]$.[31] Second, we assume that there are no cross-parameter restrictions, formalizing what we mean by "economically distinct" outcomes.

The next proposition shows that with economically distinct outcomes, one-sided t-tests based on the welfare-optimal index $X^\top w^*$ are optimal.

[31] This normalization is not restrictive, and we could instead use any interval $[-M,M]$ for $0 < M < \infty$.

Proposition 4.3.2. *Let Assumptions 4.2.3, 4.3.1, and 4.3.2 hold. Suppose further that Assumption 4.2.4 holds with* $w^{*\top}\Sigma w^* > 0$*. Then*

$$r^*(X;G) = 1\left\{\frac{X^\top w^*}{\sqrt{w^{*\top}\Sigma w^*}} > \Phi^{-1}(1-C(G))\right\} \tag{4.32}$$

is r^* *is locally most powerful. Namely, for all* $r' : \mathscr{X} \mapsto \{0,1\}$ *that are maximin, i.e.,* $r' \in \mathscr{M}$*,*

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\Big[\inf_{\theta\in\tilde{\Theta}(\varepsilon)}\theta^\top w^* P(r^*(X;G)=1|\theta) - \inf_{\theta\in\tilde{\Theta}(\varepsilon)}\theta^\top w^* P(r'(X;G)=1|\theta)\Big] \geq 0 \tag{4.33}$$

where $\tilde{\Theta}(\varepsilon) = \big\{\theta\in\Theta : \theta^\top w^* = \varepsilon\big\}$.

Proof. See Appendix D.3.2. □

The intuition for the optimality result in Proposition 4.3.2 is that choosing $X^\top w^*$ as the test statistic is optimal because it guarantees that

$$P\Big(r^*(X;G)=1|\theta^\top w_1^* = 0\Big) = C(G),$$

so that the test is similar at the boundary of the null space.

Unlike with multiple treatments in Section 4.2, the optimal critical value of the one-sided t-test (4.32), $\Phi^{-1}(1-C(G))$, is decreasing in G whenever $C(G)$ is increasing in G. This is because the researcher only tests one hypothesis such that she has no incentives to collect additional outcomes. Thus, if the social planner wants to incentivise the researcher to conduct experiments with many outcomes, she needs to lower the critical value to encourage them to do so.

The essence of Proposition 4.3.2 is that if a weighted average of the underlying parameters is what determines the welfare consequences of implementing the policy, then an analogous weighted average of the individual statistics is the appropriate test statistic. One important practical implication is that it is important to the publication process for the researcher to elicit

the optimal weights so that her paper can report a test of the appropriate index.

Conducting separate hypothesis tests on the outcomes individually (whether with or without MHT adjustment) is not optimal when there is a single policy-maker; this would discard valuable information about the relative magnitudes of the effects. See Appendix D.2.5 for a formal discussion and a simple example. We emphasize that this implication of model is fairly different from empirical practice, where researchers typically first run separate tests and then discuss significant effects. By contrast, our model suggests that aggregating information *before* testing is optimal from a welfare perspective.

Now consider the case in which each element of X is a distinct measure of a common underlying parameter. We make the following assumption.

Assumption 4.3.3 (Multiple measurements)**.** $\Theta = \{\theta_1 = \theta_2 = \cdots = \theta_G,\ \theta_1 \in [-1,1]\}$

Under Assumption 4.3.3, welfare conditional on the researcher recommending implementation of the treatment is $u(\theta) = \theta^\top w^* = \theta_1$ and the null space is

$$\Theta_0 = \Big\{\theta_1 : \theta_1 < 0\Big\}.$$

Since welfare and the null space are invariant to the choice of the weights when all components of θ are equal, it is optimal to aggregate $X_1, \ldots, X_G$ into a single statistic with variance-minimizing weights. This is in the spirit of disjunction testing (e.g., Rubin, 2021). The next proposition provides the formal result.

Proposition 4.3.3. *Let Assumptions 4.2.3, 4.2.4, 4.3.1, and 4.3.3 hold. Then*

$$r^*(X;G) = 1\left\{\frac{w^{s\top}X}{\sqrt{w^{s\top}\Sigma w^s}} > \Phi^{-1}(1-C(G))\right\},$$

where w^s minimizes $w^\top \Sigma w$ subject to $\sum_g w_g = 1$, is locally most powerful. That is

$$\lim_{\varepsilon \downarrow 0} P(r^*(X;G) = 1 | \theta = \varepsilon) - \lim_{\varepsilon \downarrow 0} P(r'(X;G) = 1 | \theta = \varepsilon) \geq 0$$

for all $r' : \mathscr{X} \mapsto \{0,1\}$ *that are maximin, i.e.,* $r' \in \mathscr{M}$.

Proof. See Appendix D.3.2. □

Proposition 4.3.3 implies that the most powerful threshold crossing rule r^* chooses the weights w^s to minimize $w^\top \Sigma w$, under the constraint that such weights sum to one. The above rule coincides with classical notions of uniformly most powerful tests in the statistical literature of single hypothesis testing (Van der Vaart, 2000; Wald, 1950).

Remark 17 (Aggregating statistics vs. aggregating outcomes)**.** The above arguments show that one-sided t-tests based on a weighted average of statistics X are optimal. In practice, researchers often first aggregate the outcomes Y and then run tests based on the aggregate outcome (e.g. Anderson, 2008). For linear estimators like OLS, these two approaches are equivalent; see Appendix D.2.6 for a further discussion. □

Discussion & Implications for Practice

In this section we discuss takeaways for practice that, in our view, one might reasonably draw from our analysis. Careful judgment is of course needed to map the results into new practices and norms, keeping in mind that professional norms should take into account many considerations that are necessarily "outside" of our model. For example, it may be beneficial to have consistent and coherent norms such that researchers know what to expect when they plan studies, even if these norms do not respect every detail of the problem we have studied here. We focus our discussion on broad and intuitive themes that seem likely to hold in a variety of plausible models of the research publication process, rather than on the functional forms that emerge as solutions to our specific model.

The first main idea we emphasize is that MHT adjustments are *potentially* appropriate when experimental research can reasonably be expected to guide multiple, distinct policy decisions. Studies that describe the effects of multiple interventions, or effects on multiple

sub-populations, clearly fit this description.[32] It is thus useful to distinguish these cases from the multiple outcomes case.

Whether MHT adjustments are *actually* appropriate in these cases, however, still depends on the costs of conducting research. This conclusion follows directly from the view of MHT as a way of "getting incentives right" for researchers. If incentives matter, then it must be *net* incentives, i.e. rewards net of costs, that determine behavior. And this means that the optimal structure of rewards depends on the structure of costs. Specifically, our results suggest that it is reasonable to request MHT adjustments (e.g. average size control, as via a Bonferroni correction) when it seems clear that the costs of conducting the research were largely invariant to the number of tests. If, however, costs scale with the number of hypotheses tested, then no correction at all may be more appropriate. Intuitively, to the extent that the appropriate correction is already "built in" to the costs of conducting research, no further correction is required.

In practice, the structure of the research cost function may often lie somewhere in-between these two extremes. Testing an additional intervention usually increases the costs of running an experiment, for example, but may well do so less than proportionately. Our framework would then motivate forms of multiple testing correction that are weaker than Bonferroni adjustments. There could even be *dis-economies* of scale in some research production processes, in which case our framework yields the interesting implication that hypothesis testing procedures should *reward* rather than penalize researchers for testing multiple hypotheses. Clearly, some actual data on the costs of doing program evaluation research and the way these scale with the number of interventions or sub-populations studied would be useful to sharpen guidance here.

In the case of multiple outcomes, the argument for MHT adjustment is not as clear. While it is quite common to measure impacts on many outcomes, what matters for decision-making is usually not whether we accept or reject hypotheses about these outcomes individually but rather whether in aggregate the effects we see are "worth more" than the costs of the intervention. The

[32] Indeed, the multiple-intervention case—genetic association testing in particular—has often been cited as the leading motivation for new MHT procedures proposed in the literature (see Dudoit et al., 2003; Efron, 2008a, for a review).

important task is therefore to conduct this aggregate assessment of worth.

The method of aggregation should reflect the meaning of the outcomes. If we interpret them as different statistical measures of some common underlying economic concept, then the right aggregation may be the one that is statistically most efficient. If we observe two distinct tests of literacy, for example, we might aggregate them based on how informative they are about actual reading and writing abilities, putting more weight on the more informative test (as for example in Anderson, 2008). But if multiple outcomes capture concepts with distinct economic value, then aggregation should instead reflect the relative economic value placed on each outcome. If we observe tests of literacy and of numeracy, for example, we might aggregate these using information about the relative market returns to each skill.

The clearest argument for introducing MHT adjustment in the case of multiple outcomes rests on the idea that some papers address heterogeneous audiences who value different things (e.g., Andrews and Shapiro, 2021). In this case our theory suggests that researchers could proceed by constructing multiple indices, each reflecting the preferences of a different member of this "policy audience," and then apply MHT adjustments across tests of these indices. This leads to very conservative procedures, however, and in practice seems an unreasonable burden to place on researchers. A more realistic approach when feasible may be for researchers to ask a policy-maker in their target audience to value outcomes, and report a test based on an index constructed using those valuations.

This approach would not preclude learning by other audience members. Journals already typically require the publication of enough data to replicate published results, which means that anyone with enough technical knowledge could re-do the analysis using alternative weights that reflect their own, distinct preferences. This process could be streamlined by publishing point estimates and covariance matrices, which are often all that is needed for audience members to draw their own conclusions (e.g. in the leading case of (asymptotically) normally distributed estimators). For non-technical audiences the process could be made more accessible using simple tools allowing them to "plug in" their own weights and obtain results customized to their

preferences. In this vision the publication process itself would become more interactive, enabling a dynamic exchange between researchers and their audience(s).

We have analyzed MHT within the context of a single paper examining results from a single experiment. In practice, however, researchers conduct many experiments and write many papers. This raises the question of how the mapping from experiments to papers should affect the use of MHT adjustments. The broad principle remains that MHT adjustments may be appropriate in our framework when research costs do not scale in proportion to the number of hypotheses being tested. For a single paper reporting results from multiple experiments, this suggests that the results should be treated as independent for hypothesis testing if the costs of conducting the underlying experiments do not interact. MHT adjustments might be appropriate, however, if the costs are interdependent. For multiple papers reporting results from the same experiment there will be cost complementarities if these exist in the underlying experiment, in which case MHT adjustments across those papers would be appropriate.

Our focus here has been on experimental research. In part this is pragmatic, as this is by far where MHT adjustments are most common in current practice.[33] It also reflects the fact, however, that in non-experimental studies it is not as clear how to require researchers to pre-commit to the tests they will report. Observational work is often iterative in complex ways, and issues such such as p-hacking may consequently loom larger than multiple testing. That said, the central role of the research cost function is a point of commonality across these problems.[34]

Finally, we emphasize that these takeaways presume that we choose to continue operating within a frequentist hypothesis-testing paradigm. There is of course a deeper question whether to move away from this paradigm altogether. In connection with this issue, we close with a remark on control of false discovery *rates*. While this approach does not emerge as a natural solution within our framework, several papers have pointed out a cogent rationale: controlling

[33] In 2020, fewer than 5% of non-experimental empirical papers even mention multiple testing as an potential issue.

[34] McCloskey and Michaillat (2020) show, for example, how research costs influence the appropriate critical values for single hypothesis testing in a setting where researchers can decide how long to continue gathering data before reporting results.

the (positive) False Discovery Rate can be interpreted as rejecting hypotheses with a sufficiently low Bayesian posterior probability (e.g., Storey, 2003; Gu and Koenker, 2020; Kline et al., 2021). In fact one can show that this argument applies even in the case of a *single* hypothesis, as its essence is simply that it makes sense to balance the costs of false positives and false negatives rather than to prioritize size control regardless of the power cost. In this sense we view the argument less as support for a particular solution to the MHT problem, and more as a reminder of the value of Bayesian reasoning more broadly.

Chapter 4 was prepared for publication Viviano, Davide; Wüthrich Kaspar; Niehaus, Paul. The Chapter is coauthored with Paul Niehaus and Kaspar Wüthrich. Davide Viviano is one of the primary investigators.

Appendix A

Appendix to Chapter 1

Main Extensions

Estimation with Global Interference

In this section, we relax the local dependency assumptions. The treatment affects each unit in a cluster k through a global interference mechanism mediated by a random variable $p_t^{(k)}$. For example, we can think of $p_t^{(k)}$ as the average number of treated units in a cluster or the adjusted price in a particular market or village due to the program. For simplicity, we consider the case where $X_i^{(k)} = 1$, i.e., the policy of interest is a global policy (e.g., the probability of treatment). We discuss assumptions on the outcome model below.

Assumption A.1.1 (Global interference)**.** Let treatments be assigned as in Definition 1.2.3 with exogenous vector of parameters $\beta_{k,1:t}$. Let

$$Y_{i,t}^{(k)} = \alpha_t + \tau_k + g\left(p_t^{(k)}, \beta_{k,t}\right) + \varepsilon_{i,t}^{(k)}, \quad \mathbb{E}_{\beta_{k,1:t}}\left[\varepsilon_{i,t}^{(k)} | p_t^{(k)}\right] = 0,$$

for some function $g(\cdot)$ unknown to the researcher, bounded and twice continuously differentiable with bounded derivatives, and unobservable $p_t^{(k)}$. Assume in addition that $\varepsilon_{i,t}^{(k)} \perp \varepsilon_{j \notin \mathscr{I}_i^{(k)},t}^{(k)} | \beta_{k,1:t}, p_t^{(k)}$ for some set $|\mathscr{I}_i^{(k)}| = \mathscr{O}(\gamma_N)$.

Assumption A.1.1 states that the outcome within each cluster is a function of a common factor, and treatment assignment rule $\beta_{k,t}$ plus unobservables centered around zero and locally

dependent. The factor $p_t^{(k)}$ also depends on the treatment assignments of all individuals. This is formalized below.

Assumption A.1.2 (Global interference component)**.** Let treatments be assigned as in Definition 1.2.3. Assume that

$$p_t^{(k)} = q(\beta_{k,t}) + o_p(\eta_n),$$

with $q(\beta)$ being unknown, bounded and twice continuously differentiable in β with uniformly bounded derivatives.

Assumption A.1.2 states that the factor can be expressed as the sum of two components. The first component $q(\cdot)$ depends on the policy parameter $\beta_{k,t}$ assigned at time t and on the distribution of covariates of all units in a cluster.[1] The second component is a stochastic component that depends on the realized treatment effects. We illustrate an example below.

Example A.1.1 (Within cluster average)**.** Suppose that

$$Y_{i,t}^{(k)} = t(\bar{D}_t^{(k)}, \nu_{i,t}), \quad \nu_{i,t}^{(k)} \sim_{i.i.d.} \mathscr{P}_\nu, \quad D_{i,t}^{(k)} \sim_{i.i.d.} \text{Bern}(\beta)$$

where $t(\cdot)$ is some arbitrary (smooth) function. Then $p_t^{(k)} = \bar{D}_t^{(k)}$ i.e., individuals depend on the average exposure in a cluster. We can write

$$Y_{i,t}^{(k)} = t(p_t^{(k)}, \nu_{i,t}^{(k)}) \quad \text{where} \quad p_t^{(k)} = \beta + \underbrace{(\bar{D}_t^{(k)} - \beta)}_{=O_p(n^{-1/2})},$$

which satisfies Assumption A.1.2 for $\eta_n = n^{-1/3}$ or larger. □

Example A.1.1 illustrates how we can accommodate a global interference mechanism whenever individuals depend on statistics of treatment assignments, such as their average.

[1] Observe that we can equivalently relax Assumption A.1.2 and assume that $q(\beta, \cdot)$ depends on the empirical distribution of covariates and use basic concentration arguments (Wainwright, 2019) to show that asymptotically the two definitions are equivalent.

Example A.1.1 does not allow for local spillovers between units but only for global interference.[2]

We are interested in the marginal effects defined below.

$$V_g(\beta) = \frac{\partial W_g(\beta)}{\partial \beta}, \quad W_g(\beta) = g(q(\beta), \beta).$$

Estimation of the marginal effect follows similarly to Equation (1.7). The following theorem guarantees consistency.

Theorem A.1.1. *Let Assumption A.1.1, A.1.2 hold with subgaussian* $\varepsilon_{i,t}^{(k)}$, $X = 1$. *Then for* $\hat{V}_{(k,k+1)}$ *estimated as in Algorithm 2, for k being odd holds:*

$$\left|\widehat{V}_{(k,k+1)} - V_g(\beta)\right| = \mathscr{O}_p\Big(\sqrt{\frac{\gamma_N \log(\gamma_N)}{\eta_n^2 n}} + \eta_n\Big) + o_p(1).$$

The proof is in Appendix A.3.4. Theorem A.1.1 guarantees consistency of the estimated gradient. The experiment can be conducted similarly to what discussed in Section 1.4 and omitted for the sake of brevity.

Matching Clusters with Distributional Embeddings

In this section, we turn to the problem of matching clusters, allowing for covariates having different distributions in different clusters. In particular, we assume that $X_i^{(k)} \sim_{i.i.d.} F_X^{(k)}$. The main distinction from previous sections is that $F_X^{(k)}$ is cluster-specific. The section works as follows: first, we characterize the bias of the difference in means estimators; second, we propose a matching algorithm that minimizes the worst-case bias.

We start from the simple setting with two clusters k, k' only, and two periods $t \in \{0,1\}$.

[2] Assumption A.1.2 builds on the model of demand as a function of individual prices in Wager and Xu (2021). The difference is that we do not rely on a specific modeling assumption of market interactions. Instead, we model outcomes as functions of exposures in a certain cluster and exploit the two-cluster variation for consistent estimation, as we discuss below.

Treatments are assigned as follows

$$\begin{aligned} t=0: &\quad D_{i,0}^{(h)} \sim \pi(X_i^{(h)};\beta_0), \quad h \in \{k,k'\} \\ t=1: &\quad D_{i,1}^{(k)} \sim \pi(X_i^{(k)};\beta), \quad D_{i,1}^{(k')} \sim \pi(X_i^{(k')};\beta'). \end{aligned} \tag{A.1.1}$$

Namely, at time $t=0$, treatments are assigned with a parameter β_0. At time $t=1$ treatments are assigned with parameter β in cluster k and β' in cluster k'.

The estimand of interest is the difference in the average effects in cluster k, formally

$$\omega_k = \int y(x;\beta) dF_X^{(k)}(x) - \int y(x;\beta') dF_X^{(k)}(x).$$

We study the properties of the difference in differences estimator

$$\widehat{\omega}_k(k') = \left[\bar{Y}_1^{(k)} - \bar{Y}_1^{(k')}\right] - \left[\bar{Y}_0^{(k)} - \bar{Y}_0^{(k')}\right],$$

which defines a difference in differences between the two clusters over two consecutive periods.

Our focus is to control the bias of the estimator. This is defined in the following lemma.

Lemma A.1.2. *Let Assumption 1.2.1, 1.2.2, and treatments assigned as in Equation* (A.1.1). *Then*

$$\mathbb{E}[\widehat{\omega}_k(k')] - \omega_k = \int \Big(y(x;\beta') - y(x;\beta_0)\Big) d\Big(F_X^{(k)}(x) - F_X^{(k')}(x)\Big).$$

Lemma A.1.2 shows that the bias depends on the difference between the expectations averaged over two different distributions. Unfortunately, the bias is unknown since it depends on the function $y(\cdot)$, which is not identifiable with finitely many clusters. We therefore bound the worst-case error over a class of functions $x \mapsto [y(x;\beta') - y(x;\beta_0)] \in \mathscr{M}$, with $\mathscr{M}$ defined below. The proof follows directly from Lemma 1.2.1 and rearrangement.

We start by defining $\mathscr{M}$ be a reproducing kernel Hilbert space (RKHS) equipped with a

norm $||\cdot||_{\mathscr{M}}$.[3] Without loss of generality, we study the worst-case functionals over the unit-ball. Formally, we focus on bounding the worst-case error of the form[4]

$$\begin{aligned}&\sup_{[y(\cdot;\beta')-y(\cdot;\beta_0)]\in\mathscr{M}:||y(\cdot;\beta')-y(\cdot;\beta_0)||_{\mathscr{M}}\leq 1}\left|\omega_k-\mathbb{E}[\widehat{\omega}_k(k')]\right|\\&=\sup_{f\in\mathscr{M}:||f||_{\mathscr{M}}\leq 1}\Big\{\int f(x)d(F_X^{(k)}-F_X^{(k')})\Big\}.\end{aligned}\tag{A.1.2}$$

The right-hand side is know as the maximum mean discrepancy (MMD), a measure of distances in RKHS (see Muandet et al., 2016, and references therein). It is known that the MMD can be consistently estimated using kernels. In particular, given a particular choice of a kernel $k(\cdot)$, which corresponds to a certain RKHS, we can estimate

$$\begin{aligned}\widehat{\text{MMD}}^2(k,k')&=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}h\Big(X_i^{(k)},X_i^{(k')},X_j^{(k)},X_j^{(k')}\Big),\\h(x_i,y_i,x_j,y_j)&=k(x_i,x_j)+k(y_i,y_j)-k(x_i,y_j)-k(x_j,y_i).\end{aligned}\tag{A.1.3}$$

The estimator estimates the squared MMD. It only depends on the kernel function $k(\cdot)$ and hence can be easily constructed in a finite sample without requiring an explicit characterization of RKHS. Here, $\widehat{\text{MMD}}^2(k,k')\to_p\left|\left|\mu_{F_X^{(k)}}-\mu_{F_X^{(k')}}\right|\right|^2_{\mathscr{M}}$ (Sriperumbudur et al., 2012).

We now turn to the problem of matching clusters. We do so using the estimated MMD in Equation (A.1.3). We first note that the estimator $\widehat{\text{MMD}}^2(k,k')$ only depends on *pre-treatment* variables, and hence can be computed *before* treatments are assigned. As a result, given cluster k, we can match k with the $k'\neq k$ having the smallest estimated MMD. Formally, the following matching algorithm is considered:

- construct

$$k'\in\arg\min_{k\neq k}\widehat{\text{MMD}}^2(k,k').\tag{A.1.4}$$

[3] A RKHS is an Hilbert space of functions where all the evaluations functionals are bounded, namely, where for each $f\in\mathscr{M}$, and $x\in\mathscr{X}$, $f(x)\leq C||f||_{\mathscr{M}}$ for a finite constant C. Intuitively, assuming that $[y(\cdot;\beta')-y(\cdot;\beta_0)]\in\mathscr{M}$ imposes smoothness conditions on the average effect as a function of x.

[4] Here Equation (A.1.2) follows directly from Lemma A.1.2 and the fact that the integral is a scalar.

based on the minimum estimated MMD in Equation (A.1.3).

- Randomize treatments as in Equation (A.1.1);
- Estimate $\widehat{\omega}_k(k')$.

When instead we want to match without replacement clusters, we can minimize some aggregate measures of error (e.g., the sum of estimated MMD across clusters).

Policy Choice with Dynamic Treatments

This section studies an experimental design when carry-overs occur. For simplicity, we omit covariates and assume that $X_i = 1$.

We start our discussion by introducing the dynamic model. For the sake of brevity, we directly impose a high-level condition on the outcome model.

Assumption A.1.3 (Dynamic model)**.** For treatments assigned with exogenous parameters $(\beta_{k,1}, \cdots, \beta_{k,t})$ as in Definition 1.2.3, let the followig hold

$$Y_{i,t}^{(k)} = \Gamma(\beta_t, \beta_{t-1}) + \varepsilon_{i,t}^{(k)}, \quad \mathbb{E}_{\beta_{k,1:t}}\left[\varepsilon_{i,t}^{(k)}\right] = 0,$$

for some unknown $\Gamma(\cdot)$, $\varepsilon_{i,t}^{(k)}$.

The components $\beta_{k,t}, \beta_{k,t-1}$ capture present and carry-over effects that result from individual and neighbors' treatments in the past two periods. Here, we allow for both panels and repeated cross-sections. We study the problem of estimating a path of treatment probabilities $(0, \beta_1, \cdots, \beta_T)$ from an experiment, where, in the first period, we assume for simplicity that none of the individuals is treated. This path is then implemented on a new population without having access to the outcomes of such a new population. We provide a simple example below.

Example A.1.2. Suppose that

$$Y_{i,t}^{(k)} = D_{i,t}^{(k)}\phi_1 + \frac{\sum_{j\neq i} A_{i,j}^{(k)} D_{i,t-1}}{\sum_{j\neq i} A_{i,j}^{(k)}}\phi_2 + \nu_{i,t}^{(k)}, \quad D_{i,t}^{(k)} \sim_{i.i.d.} \text{Bern}(\beta_t).$$

That is, individuals depend on their present treatment assignment and on the treatment assignments of the neighbors in the previous period. Let $\nu_{i,t}$ be a zero-mean random variable. The expression simplifies to

$$Y_{i,t}^{(k)} = \beta_t\phi_1 + \beta_{t-1}\phi_2 + \varepsilon_{i,t}^{(k)}$$

where $\varepsilon_{i,t}^{(k)}$ is zero mean, and depends on neighbors' and individual assignments. □

We now define the long-run welfare.

Definition A.1.1 (Long-run welfare)**.** Given an horizon T^*, define the long-run welfare as follows:

$$\mathscr{W}(\{\beta_s\}_{s=1}^{T^*}) = \sum_{t=1}^{T^*} q^t \Gamma(\beta_t, \beta_{t-1}),$$

for a known discounting factor $q < 1$, where $\beta_0 = 0$.

The long-run welfare deifines the cumulative (discounted) welfare obtained from a certain sequence of decisions $(\beta_1, \beta_2, \cdots)$.

Our goal is to maximize the long-run welfare. We parametrize future treatment probabilities based on past treatment probabilities as follows

$$\beta_{t+1} = h_\theta(\beta_t, \beta_{t-1}), \quad \theta \in \Theta.$$

The parametrization is imposed for computational convenience. For some arbitrary large T^*, the

objective function takes the following form

$$\begin{aligned} \widetilde{W}(\theta) &= \sum_{t=1}^{T^*} q^t \Gamma\Big(\beta_t, \beta_{t-1}\Big), \\ \beta_t &= h_\theta(\beta_{t-1}, \beta_{t-2}) \quad \text{for all } t \geq 1, \quad \beta_0 = \beta_{-1} = 0. \end{aligned} \tag{A.1.5}$$

Here $\widetilde{W}(\theta)$ denotes the long-run welfare indexed by a given policy's parameter θ. The objective function defines the discounted cumulative welfare induced by the policy h_θ.

Definition A.1.2 (Non-stationary policy decisions)**.** A non-stationary policy is defined as as a map $h_\theta : \mathscr{B} \times \mathscr{B}, \theta \in \Theta$. Define the non-stationary estimand as follows:

$$h_{\theta^*}(\cdot), \quad \theta^* \in \arg\max_{\theta \in \Theta} \widetilde{W}(\theta).$$

The algorithm estimates the function $\Gamma(\cdot)$ using a single wave experiment, i.e., we use a single period of experimentation. We then use the estimated function $\Gamma(\cdot)$ and its gradient for estimating the optimal policy.

The randomization and estimators are described in Algorithm A.5.2. We conduct the randomization using two periods of experimentation only. We partition the space $[0,1]^2$ into a grid $\mathscr{G}$ of equally spaced components (β_1^r, β_2^r) for each triad of clusters r. Within each triad, we induce small deviations to the parameters β. For each triad r, the algorithm returns

$$\widetilde{\Gamma}(\beta_2^r, \beta_1^r), \quad \widehat{g}_1(\beta_2^r, \beta_1^r), \quad \widehat{g}_2(\beta_2^r, \beta_1^r)$$

where the latter two components are the estimated partial derivatives of $\Gamma(\cdot)$, and $\widetilde{\Gamma}(\beta_2^r, \beta_1^r)$ is the within cluster average. For each pair of parameters (β_2, β_1), we estimate $\widehat{\Gamma}(\beta_2, \beta_1)$ as follows

$$\begin{aligned} &\widehat{\Gamma}(\beta_2, \beta_1) = \widetilde{\Gamma}(\beta_2^r, \beta_1^r) + \widehat{g}_2(\beta_2^r, \beta_1^r)(\beta_2 - \beta_2^r) + \widehat{g}_1(\beta_2^r, \beta_1^g)(\beta_1 - \beta_1^r), \\ &\text{where} \quad (\beta_1^r, \beta_2^r) = \arg \min_{(\tilde{\beta}_1, \tilde{\beta}_2) \in \mathscr{G}} \Big\{ ||\beta_1 - \tilde{\beta}_1||^2 + ||\beta_2 - \tilde{\beta}_2||^2 \Big\}. \end{aligned} \tag{A.1.6}$$

The idea is as follows: we estimate $\Gamma(\beta_2, \beta_1)$ at (β_2, β_1) using a a first-order Taylor approximation around the closest pairs of parameters in the grid $\mathscr{G}$. Given $\widehat{\Gamma}$, we estimate the welfare-maximizing parameter[5]

$$\widehat{\theta} \in \arg\max_{\theta \in \Theta} \sum_{t=1}^{T^*} q^t \widehat{\Gamma}(\beta_t, \beta_{t-1}), \quad \beta_t = h_\theta(\beta_{t-1}, \beta_{t-2}) \quad \forall t \geq 1, \quad \beta_0 = \beta_{-1} = 0.$$

In the following theorem, we study the behavior of $\widehat{\theta}$, in terms of out-of-sample regret.

Theorem A.1.3 (Out-of-sample regret)**.** *Let Assumption A.1.3 hold. Let $X = 1$, and suppose that $\Gamma(\beta_2, \beta_1)$ is twice differentiable with bounded derivatives. Let treatments be assigned as in Algorithm A.5.2. Suppose in addition that $\varepsilon_{i,t}^{(k)} \perp \varepsilon_{j \notin \mathscr{I}_i^{(k)}}^{(k)}$ where $|\mathscr{I}_i^{(k)}| \leq \gamma_N$, for some arbitrary γ_N and $\varepsilon_{i,t}^{(k)}$ is sub-gaussian. Let $\gamma_N \log(\gamma_N)/(\eta_n^2 n) = o(1)$. Then*

$$\lim_{n \to \infty} P\Big(\sup_{\theta \in \Theta} \widetilde{W}(\theta) - W(\widehat{\theta}) \leq \frac{\bar{C}}{K} \Big) = 1$$

for a constant $\bar{C}$ independent of K.

The proof is in Appendix A.3.4. To our knowledge, Algorithm A.5.2 is novel to the literature on experimental design.[6]

Theorem A.1.3 shows that with probability converging to one as the size of each cluster increases, the regret scales at a rate of $1/K$. To gain further intuition on the derivation of the

[5]Here, $\widehat{\theta}$ can be obtained using off-the-shelf algorithms. A simple example is running in-parallel multiple gradient descent algorithms initialized over different starting points and choosing the one which leads to the largest objective $\sum_{t=1}^{T^*} q^t \widehat{\Gamma}(\beta_t, \beta_{t-1})$.

[6]We note that optimal dynamic treatments have been studied in the literature on bio-statistics, see, e.g., Laber et al. (2014), while here we consider the different problems of the design of the experiment. Adusumilli et al. (2019) discuss off-line policy estimation in the presence of dynamic budget constraints with *i.i.d.* observations. The authors assume no carry-overs and do not discuss the problem of experimental design.

theorem, observe that we can bound

$$\sup_{\theta\in\Theta}\widetilde{W}(\theta)-\widetilde{W}(\widehat{\theta})\leq 2\sum_{t}q^{t}\times\underbrace{\sup_{(\beta_1,\beta_2)\in[0,1]^2}\left|\widehat{\Gamma}(\beta_2,\beta_1)-\Gamma(\beta_2,\beta_1)\right|}_{(A)}.$$

To bound (A), observe first that each element in the grid $\mathscr{G}$ has a distance of order $1/\sqrt{K}$ since the grid has two dimensions and $K/3$ components. As a result for any element (β_2,β_1), we can write

$$\begin{aligned}\Gamma(\beta_2,\beta_1)=&\underbrace{\Gamma(\beta_2^r,\beta_1^r)}_{(B)}+\underbrace{\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_1^r}(\beta_1-\beta_1^r)+\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_2^r}(\beta_2-\beta_2^r)}_{(C)}\\&+\underbrace{\mathscr{O}\left(||\beta_1-\beta_1^r||^2+||\beta_2-\beta_2^r||^2\right)}_{(D)}\end{aligned}$$

where $\beta^r\in\mathscr{G}$ is some value in the grid such that (D) is of order $1/K$. We can then show that $(B)+(C)-\widehat{\Gamma}(\beta_2^r,\beta_1^r)$ converges in probability to zero as n grows which is possible since we use the estimated gradient to construct $\widehat{\Gamma}$. If instead we had not used information on the estimated gradient, the rate would be dominated by (C) which is of order $1/\sqrt{K}$.

However, we note that different from previous sections, the rate $1/K$ is specific to the one-dimensional setting and carry-overs over two consecutive periods. In p dimensions, the rate would be of order $1/K^{2/(p+1)}$ due to the curse of dimensionality.

Welfare Maximization with a Non-Adaptive Experiment and Local Perturbations

Next, we revisit the non-adaptive experiment in Section 1.3 and introduce estimators of β^* *without* adaptivity. This sub-section serves two purposes. First, it sheds light on comparisons of the adaptive procedure with grid-search-type methods, showing drawbacks of the grid-search approach in terms of convergence of the regret. Second, it shows how, when an adaptive procedure is not available, we can still use information from the marginal effect estimated as we

propose in Algorithm 1 to improve rates of convergence in the number of clusters.

The algorithm that we propose is formally discussed in Algorithm A.5.3 in Appendix A.5 and works as follows. First, we construct a fine grid $\mathscr{G}$ of the parameter space $\mathscr{B}$ (with p dimensions), with equally spaced parameters under the $l2$-norm. Second, we pair clusters, and we assign a *different* parameter β^k for each pair $(k,k+1)$ from the grid $\mathscr{G}$. Third, in each pair, we estimate the gradient $\hat{V}_{(k,k+1)} \in \mathbb{R}^p$, by perturbing, sequentially for $T = p$ periods, one coordinate at a time of the parameter β^k.[7] We estimate welfare using a first-order Taylor expansion

$$\widehat{W}(\beta) = \bar{W}^{k^*(\beta)} + \hat{V}^{\top}_{(k^*(\beta),k^*(\beta)+1)}(\beta - \beta^{k^*(\beta)}), \quad \hat{\beta}^{ow} = \arg\max_{\beta\in\mathscr{B}} \widehat{W}(\beta), \tag{A.1.7}$$

$$\text{where } k^*(\beta) = \arg\min_{k\in\{1,3,\cdots\},\beta^k\in\mathscr{G}} ||\beta^k - \beta||^2, \quad \bar{W}^k = \frac{1}{2}\Big[\frac{1}{T}\sum_{t=1}^{T}\bar{Y}_t^k - \bar{Y}_0^k + \frac{1}{T}\sum_{t=1}^{T}\bar{Y}_t^{k+1} - \bar{Y}_0^{k+1}\Big].$$

Here, $\bar{Y}_t^k$ is the average outcome in cluster k at time t, and $\hat{V}_{(k^*,k^*+1)}$ is estimated as in Equation (A.5.3). Also, $\bar{W}^k$ denotes the average outcome, as we *pool* outcomes from two clusters in the same pair $(k,k+1)$. The estimator in Equation (A.1.7) uses a first-order Taylor expansion around β, using information from the closest element β^k.

We can now characterize guarantees of the estimator as $n \to \infty$, and $K, p < \infty$.

Theorem A.1.4. *Suppose that $\varepsilon_{i,t}^{(k)}$ is sub-gaussian. Let Assumptions 1.2.1, 1.2.2, 1.3.1 hold. Let $\eta_n = o(n^{-1/4})$. Let $\gamma_N \log(n\gamma_N K)/(\eta_n^2 n) = o(1)$. Consider $\hat{\beta}^{ow}$ as in Algorithm A.5.3, with $\mathscr{B} \subseteq [0,1]^p$. Then for a constant $\bar{C} < \infty$ independent of (n,T,K),*

$$\lim_{n\to\infty} P\Big(W(\beta^*) - W(\hat{\beta}^{ow}) \leq \frac{\bar{C}}{K^{2/p}}\Big) = 1.$$

The proof is in Appendix A.3.3. Theorem A.1.4 showcases two properties of the method. First, for $p = 1$, the rate of convergence is of order $1/K^2$, which is possible *because* we also estimate and leverage the gradient $\hat{V}$. Our insight here is to use local perturbations to recover

[7] Sequentiality here is for notational convenience only, and can be replaced by $T = 1$, but with $2p$ clusters allocated to each coordinate.

the gradient directly by choosing pairs of points on the grid that are close enough (but not too close, which we control through the perturbation η_n) so that we can recover $\hat{V}$ at a given point consistently as $n \to \infty$, fixing K. We then augment the estimator of the welfare with $\hat{V}$, since, otherwise, the rate would be slower in K.[8] One drawback of a grid search approach is that, as $p > 1$, the method suffers a curse of dimensionality and the rate in K decreases as p increases. This is different from the adaptive procedure (e.g., Corollary 4), where the rate in K does not depend on p. A second disadvantage of the grid search is that the method does not control the in-sample regret, formalized below.

Proposition A.1.5 (Non-vanishing in-sample regret)**.** *There exists a strongly concave $W(\cdot)$, such that, for $p = 1$, $W(\beta^*) - \frac{1}{K}\sum_{k=1}^{K} W(\beta^k) \geq c$, for a constant $c > 0$ independent of (n, K, T).*

Proposition A.1.5 shows that the grid search method performs poorly for the in-sample regret, which is of interest when optimizing participants' welfare (or costs of the experiment), differently from the adaptive procedure in the main text. Similar reasoning can be used for related procedures to the grid search approach.

Inference and Estimation with Observed Cluster Heterogeneity

In this subsection, we discuss an extension to allow for cluster heterogeneity. Consider $\theta_k \in \Theta$ to denote the cluster's type for cluster k, where Θ is a finite space (i.e., there are finitely many cluster types). Let θ_k be observable by the researcher and be non-random (i.e., conditions should be interpreted conditional on $\{\theta_k\}_{k=1}^{K}$). Consider the following assumptions.

Assumption A.1.4. Let the following holds:

(A) For each cluster k, let Assumption 1.2.1 holds, with $F_X, F_{U|X}$ replaced by $F_X(\theta_k), F_{U|X}(\theta_k)$ as a function of θ_k;

(B) Assumption 1.2.2 with $r(\cdot)$ which also depends on θ_k.

[8] By a second-order Taylor expansion, using information from the gradient guarantees that $\widehat{W}(\beta)$ converges to $W(\beta)$ up-to a second-order term of order $O(||\beta - \beta^k||^2)$, instead of a first-order term $O(||\beta - \beta^k||)$.

Assumption A.1.4 allows for both the distribution of covariates and unobservables and potential outcomes to also depend on the cluster's type θ_k. We can now state the following lemma.

Lemma A.1.6. *Under Assumption A.1.4, under an assignment in Assumption 1.2.3 with exogenous (i.e., not data-dependent) $\beta_{k,t}$ the following holds:*

$$Y_{i,t}^{(k)} = y\left(X_i^{(k)}, \beta_{k,t}, \theta_k\right) + \varepsilon_{i,t}^{(k)} + \alpha_t + \tau_k, \quad \mathbb{E}_{\beta_{k,t}}\left[\varepsilon_{i,t}^{(k)} | X_i^{(k)}\right] = 0, \tag{A.1.8}$$

for some function $y(\cdot)$ unknown to the researcher. In addition, for some unknown $m(\cdot)$, $\mathbb{E}_{\beta_{k,t}}\left[Y_{i,t}^{(k)} | D_{i,t}^{(k)} = d, X_i^{(k)} = x\right] = m(d,x,\beta_{k,t},\theta_k) + \alpha_t + \tau_k$.

Differently from Lemma 1.2.1, here the the functions also depend on the cluster's type θ_k. The proof of Lemma A.1.6 follows verbatim from the one of Lemma 1.2.1, taking here into account also the (deterministic) cluster's type. In the following lines, we discuss the single-wave for inference and multi-wave experiments for estimation under heterogeneity.

Single-wave experiment

In the context of a single-wave experiment, we are interested in testing the null hypothesis of whether a *class* of decisions $\beta(\theta), \theta \in \Theta$, which depends on the cluster's type, is optimal. Namely, let

$$W\left(\beta(\theta), \theta\right) = \int y(x, \beta(\theta), \theta) dF_X(\theta), \quad \beta : \Theta \mapsto \mathscr{B}, \quad \theta \in \Theta$$

be the welfare corresponding to cluster's type θ, for a decision rule $\beta(\theta)$. Also, let

$$V\left(\beta(\theta), \theta\right) = \frac{\partial W(b,\theta)}{\partial b}\Big|_{b=\beta(\theta)}$$

be the marginal effect with respect to changing $\beta(\theta)$ (for fixed θ). Our null hypothesis is

$$H_0 : V\Big(\beta(\theta), \theta\Big) = 0, \quad \forall \theta \in \Theta, \tag{A.1.9}$$

i.e., the (baseline) policy $\beta(\theta)$ is optimal for all clusters under consideration. The algorithm follows similarly to Algorithm 2 with the following modification: instead of matching arbitrary clusters, we construct pairs such that elements in the same pair $(k, k+1)$ are such that $\theta_k = \theta_{k+1}$. We can now state the following corollary.

Corollary 10. *Suppose that for all $x \in \mathscr{X}, d \in \{0,1\}, b \in \mathscr{B}, \theta \in \Theta, \pi(x,b), m(d,x,\beta,\theta)$ are uniformly bounded and twice differentiable with bounded derivatives. Let Assumption 1.3.2, A.1.4 hold. Consider Algorithm 2, with parameter $\beta(\theta)$ as a function of the cluster's type and for each pair of clusters $(k,k+1), k$ being odd, being such that $\theta_k = \theta_{k+1}$. Let $\mathscr{T}_n$ as in Algorithm 2. Then for $4 \leq K < \infty$, $\alpha \leq 0.08$,*

$$\lim_{n\to\infty} P\Big(|\mathscr{T}_n| \leq \mathrm{cv}_{K/2-1}(\alpha)\Big| H_0\Big) \geq 1-\alpha, \tag{A.1.10}$$

where $\mathrm{cv}_{K/2-1}(h)$ is the size-h critical value of a t-test with $K/2-1$ degrees of freedom, and H_0 is as defined in Equation (A.1.9)*.*

Corollary 10 states that the proposed algorithm guarantees asymptotic size control also in the presence of cluster heterogeneity, under the null hypothesis that the policy $\beta(\theta)$ as a function of the cluster's type is optimal for all types. Importantly, here we construct a test statistics using information from *all* clusters. Consistency of the marginal effect follows verbatim as in Theorem 1.3.1, assuming that clusters of the same type are matched.

Multi-wave experiment

For the multi-wave experiment, our goal is to find $\beta^*(\theta)$ such that

$$\beta^*(\theta) \in \arg\max_{b\in\mathscr{B}} W(b,\theta), \quad \forall \theta \in \Theta.$$

Similarly to the single-wave experiment, clusters (k,k') of the same type $\theta_k = \theta_{k'}$ are first matched together. The first extension consists of grouping clusters of the same type together, estimating separately $\beta^*(\theta)$ for each $\theta \in \Theta$. In this case the regret bound holds up-to a factor of order $\min_t P(\theta = t)$, with $P(\theta = t)$ denoting the (exact) share of clusters of type t. Intuitively, the same analysis carry-over to this case, as we restrict our attention to each subgroup separately. The second approach instead consists of updating the same policy from a given pair using information from that *same* pair.

Additional Extensions

Tests with a p-dimensional Vector of Marginal Effects

In the following lines we extend Algorithm 2 to testing the following null

$$H_0 : V^{(j)}(\beta) = 0, \text{ for some } p \geq l \geq 1,$$

where we consider a generic number of dimensions tested l. We introduce the algorithmic procedure in Algorithm A.2.1.

We define $\mathscr{K}_j$ the set of pairs in Algorithm A.2.1 used to estimate the j^{th} entry of $V(\beta)$. Define

$$\bar{V}_n^{(j)} = \frac{2l}{K} \sum_{k \in \mathscr{K}_j} \widehat{V}_k,$$

the average marginal effect for coordinate j estimated from those clusters is used to estimate the effect of the j^{th} coordinate. We construct

$$Q_{j,n} = \frac{\sqrt{K/(2l)}\bar{V}_n^{(j)}}{\sqrt{(K/(2l)-1)^{-1}\sum_{k \in \mathscr{K}_j}(\widehat{V}_k^{(j)} - \bar{V}_n^{(j)})^2}}, \quad \mathscr{T}_n = \max_{j \in \{1,\cdots,l\}} |Q_{j,n}|, \tag{A.2.3}$$

where $\mathscr{T}_n$ denotes the test statistics. The choice of the l-infinity norm is motivated by its theoretical properties: the statistics $Q_{j,n}$ follows an unknown distribution as a result of possibly

Algorithm A.2.1. One wave experiment for inference

Require: Value $\beta \in \mathbb{R}^p$, K clusters, 2 periods of experimentation, number of tests t.

1: Match clusters into pairs $K/2$ pairs with consecutive indexes $\{k, k+1\}$;

2: $t = 0$ (*baseline*):

a: Treatments are assigned at some baseline β_0 $D_{i,0}^{(h)} \sim \pi(X_i^{(h)}, \beta_0), h \in \{1, \cdots, K\}$ (e.g., none of the individuals is treated).

b: Collect baseline values: for n units in each cluster observe $Y_{i,0}^{(h)}, h \in \{1, \cdots, K\}$.

3: $t = 1$ (*experimentation-wave*)

4: Assign each pair of clusters $\{k, k+1\}$ to a coordinate $j \in \{1, \cdots, p\}$ (with the same number of pairs to each coordinate)

5: For each pair $\{k, k+1\}$, k is odd, assigned to coordinate j

a: Randomize

$$D_{i,1}^{(h)} \sim \begin{cases} \pi(X_i^{(h)}, \beta + \eta_n \underline{e}_j) \text{ if } h = k \\ \pi(X_i^{(h)}, \beta - \eta_n \underline{e}_j) \text{ if } h = k+1 \end{cases}, \quad n^{-1/2} < \eta_n \leq n^{-1/4}$$

b: For n units in each cluster $h \in \{k, k+1\}$ observe $Y_{i,1}^{(h)}$.

c: Estimate the marginal effect for coordinate j as

$$\widehat{V}_k = \frac{1}{2\eta_n}\left[\bar{Y}_1^{(k)} - \bar{Y}_0^{(k)}\right] - \frac{1}{2\eta_n}\left[\bar{Y}_1^{(k+1)} - \bar{Y}_0^{(k+1)}\right] \tag{A.2.1}$$

return

$$\widetilde{V}_n = \left[\widehat{V}_1, \widehat{V}_3, \cdots, \widehat{V}_{K-1}\right] \tag{A.2.2}$$

heteroskedastic variances of $\widehat{V}_k$ across different clusters. However, the upper-bound on the critical quantiles of the proposed test-statistic for unknown variance attains a simple expression under the proposed test-statistics. From a conceptual stand-point, the proposed test-statistic is particularly suited when a large deviation occurs over one dimension of the vector.[9] We now introduce the following theorem.

Theorem A.2.1 (Nominal coverage)**.** *Let Assumptions 1.2.1, 1.2.2, 1.3.2 hold. Let* $n^{1/4}\eta_n = o(1), \gamma_N^2/N^{1/4} = o(1)$, $K < \infty$. *Let* $K \geq 4l$, H_0 *be as defined in Equation* (1.8). *For any* $\alpha \leq 0.08$,

$$\lim_{n\to\infty} P\Big(\mathscr{T}_n \leq q_\alpha \Big| H_0\Big) \geq 1 - \alpha, \text{ where } q_\alpha = \mathrm{cv}_{K/(2l)-1}\Big(1 - (1-\alpha)^{1/l}\Big), \tag{A.2.4}$$

[9] Observe that alternatively, we may also consider randomization tests as discussed in Canay et al. (2017). This is omitted for the sake of brevity.

with $\text{cv}_{K/(2l)-1}(h)$ *denotes the critical value of a two-sided t-test with level h with test-statistic having* $K/(2l)-1$ *degrees of freedom.*

The proof is in Appendix A.3.4.

Non Separable Fixed Effects

In the following lines, we show how we can leverage direct and marginal spillover effects to identify (and then estimate) the marginal effects when fixed effects are non-separable in time and cluster identity.

Theorem A.2.2 (Marginal effects with non-separable fixed effects)**.** *Let* $X = 1$*, and suppose that* $m(d,1,\beta)$ *is bounded and twice differentiable with bounded derivatives for* $d \in \{0,1\}$*. Let Assumptions 1.2.1 hold. Suppose that fixed-effects are non-separable, with*

$$Y_{i,t}^{(k)} = m(D_{i,t}^{(k)},1,\beta) + \alpha_{k,t} + \varepsilon_{i,t}^{(k)}, \quad \mathbb{E}[\varepsilon_{i,t}^{(k)}] = 0, \quad D_{i,t}^{(k)} \sim_{i.i.d.} \text{Bern}(\beta),$$

and $m(1,1,\beta)$ *being a constant function in* β*. Then*

$$\mathbb{E}\Big[\hat{\Delta}_k(\beta) + \hat{S}(0,\beta)(1-\beta) - (1-\beta)\hat{S}(1,\beta)\Big] = V(\beta) + \mathscr{O}(\eta_n).$$

The proof is in Appendix A.3.4. Theorem A.2.2 shows that we can use the information on the spillover and direct treatment effects to identify the marginal effects in the presence of non-separable time and cluster fixed effects. The theorem leverages the assumption that spillovers only occur in the control individuals but not the treated. We note that in applications where treatment effects have disproportionately large effects on the outcome of an individual, and the treated outcomes do not depend on the neighbors' assignments, there might be some tension with the assumption of fixed effects assumption. Studying economic settings where the assumption does and does not hold is an interesting future direction.

Out-of-sample Regret with Strict Quasi-Concavity

In the following lines, we provide guarantees on the regret bounds for the adaptive algorithm in Section 1.4 under quasi-concavity. We replace Assumption 1.4.2 with the following condition.

Assumption A.2.1 (Local strong concavity and strict quasi-concavity)**.** Assume that the following conditions hold.

(A) For every $\beta, \beta' \in \mathscr{B}$, such that $W(\beta') - W(\beta) \geq 0$, then $V(\beta)^\top(\beta' - \beta) \geq 0$,

(B) For every $\beta \in \mathscr{B}$, $||V(\beta)||_2 \geq \mu||\beta - \beta^*||_2$, for a positive constant $\mu > 0$;

(C) $\frac{\partial^2 W(\beta)}{\partial \beta^2}\Big|_{\beta=\beta^*}$ has negative eigenvalues bounded away from zero at β^*, with $\beta^* \in \tilde{\mathscr{B}} \subset \mathscr{B}$ being in the interior of $\mathscr{B}$.

Condition (A) imposes a quasi-concavity of the objective function. The condition is *equivalent* to common definitions of quasi concavity (Hazan et al., 2015). Condition (A) holds when increasing the probability of treating more neighbors has decreasing marginal effects. Condition (B) assumes that the marginal effect only vanishes at the optimum, ruling out regions over which marginal effects remain constant at zero. A notion of strict quasi-concavity can be found in Hazan et al. (2015), where the authors assume (A) and that the gradient vanishes at the optimum only. Condition (C) also imposes that the function has negative definite Hessian at β^* only but not necessarily globally. The above restrictions guarantee strong concavity locally at the optimum, but not necessarily globally. We now introduce out-of-sample guarantees in this setting. In such a case, the choice of the learning rate consists of a gradient norm rescaling, as discussed in Remark 7.

Theorem A.2.3. *Let Assumptions 1.2.1, 1.2.2, 1.4.1, A.2.1 hold, and choose* $\alpha_{k,w}$ *as in Equation* (1.14)*, for arbitrary* $v \in (0,1)$*,* κ *as defined in Equation* (A.3.7)*, and* ε_n *as in Lemma A.3.12. Take a small* $1/4 > \xi > 0$*, and let* $n^{1/4-\xi} \geq \bar{C}\sqrt{\log(n) p \gamma_N T^2 e^{B_p T} \log(KT)}$*,* $\eta_n = 1/n^{1/4+\xi}$*, for*

finite constants $\infty > B_p, \bar{C} > 0$*. Then for* $T \geq \zeta^{1/v}$*, for a finite constant* $\zeta < \infty$*, with probability at least* $1 - 1/n$*,*

$$W(\beta^*) - W(\hat{\beta}^*) = \mathscr{O}(\check{T}^{-1+v}).$$

The proof is in Appendix A.3.4.

Staggered Adoption

In this section, we sketch the experimental design in the presence of staggered adoption, i.e., when treatments are assigned only once to individuals and post-treatment outcomes are collected once. The algorithm works similarly to what was discussed in Section 1.4 with one small difference: every period, we only collect information from a given clusters' pair and update the policy for the subsequent pair and proceed in an iterative fashion. See Algorithm A.5.4 for details.

For simplicity, we only discuss the case where $\beta \in \mathbb{R}$ (single coordinate), while for the case where $\beta \in \mathbb{R}^p$ the algorithm works similarly with the only difference that every perturbation to each coordinate requires a new pair of clusters (with in total $2p$ times the number of iterations many clusters).

Theorem A.2.4 (In-sample regret)**.** *Let the conditions in Theorem 1.4.3 hold and let* $\beta \in \mathbb{R}$*, with* $\check{\beta}^t$ *estimated as discussed in the current section. Then with probability at least* $1 - 1/n$*,*

$$\frac{1}{T}\sum_{t=1}^{T}\left[W(\beta^*) - W(\check{\beta}^t)\right] \leq \bar{C}\frac{p\log(T)}{T}$$

for a finite constant $\bar{C} < \infty$*.*

See Appendix A.3.4 for the proof.

Selection of η_n: Rule of Thumb

In this subsection we provide a rule of thumb for selecting η_n. Following Theorem 1.3.1 and following Lemma A.3.3 which provides exact constants, we can write with probability at least $1-1/n$

$$\left|\hat{V}_{(k,k+1)} - V(\beta)\right| \leq \sqrt{\frac{2\sigma^2\gamma_N \log(2\gamma_N n)}{n\eta_n^2}} + c\eta_n, \quad c = \left|\left|\frac{\partial^2 m(d,x,\beta)}{\partial \beta^2}\right|\right|_\infty,$$

where the l_∞ is taken with respect to each element of the Hessian, x, β.[10] Note that we cannot directly minimize the upper bound since otherwise, the bias and variance would converge to zero at the same rate, and we would violate the condition that $\eta_n = o(n^{-1/4})$ used for inference. Instead, we minimize

$$\min_{\eta_n} \sqrt{\frac{2\sigma^2\gamma_N \log(2\gamma_N n)}{n\eta_n^2}} + c\eta_n/s_n^2, \quad s_n = o(1),$$

where s_n penalizes the bias by an $o(1)$ component and chosen below. It follows that

$$\eta_n^2 = \sqrt{\frac{2s_n\sigma^2\gamma_N \log(2\gamma_N n)}{nc}}.$$

Note that under the conditions for inference, we assume that $\gamma_N/n^{1/4} = o(1)$. To also remove the logarithmic terms (which are asymptotically negligible), let $s_n = \gamma_N/(n^{1/4}\log(n\gamma_N))$, assumed to be $o(1)$ for inference. We can then write the solution to the optimization problem as

$$\gamma_N\sqrt{\frac{2\sigma^2}{c}}n^{-5/16} \approx \gamma_N\sqrt{\frac{2\sigma^2}{c}}n^{-1/3}.$$

[10] The constants for the upper bound for the variance follow from Lemma A.3.3, while the component $c\eta_n$ captures the bias obtained from a second-order Taylor expansion to $m(\cdot)$, noting that only one entry of β is perturbed by η_n in opposite directions for two clusters in a pair as discussed in Algorithm 1.

Here, we can replace σ^2 and c with some out-of-sample estimates of the outcomes' variance and curvature. Since, in practice, the researcher may impose a small sample upper bound on the bias, our proposal is

$$\eta_n = \begin{cases} \gamma_N \sqrt{\frac{2\sigma^2}{c}} n^{-1/3} & \text{if } \gamma_N \sqrt{\frac{2\sigma^2}{c}} n^{-1/3} \leq B \\ B \text{ otherwise} \end{cases}$$

where Bc denotes an upper bound on the bias of the estimator imposed by the researcher (e.g., $cB = 0.05$). The problem here is that γ_N is unknown. Therefore, whenever the researcher does not have a good guess for γ_N,we recommend choosing $\eta_n = \sqrt{\frac{2\sigma^2}{c}} n^{-1/3}$ (without the term γ_N) which also leads to valid inference, but slightly smaller small-sample bias (and larger small-sample variance) than the optimal choice.

Derivations

Notation and Definitions

First, we introduce conventions and notation. We define $x \lesssim y$ if $x \leq cy$ for a positive constant $c < \infty$. For K many clusters, we say that

$$\lfloor k \rceil = \begin{cases} k \text{ if } k \leq K \\ k - K \text{ otherwise.} \end{cases}$$

We will refer to $\hat{V}_{(k,k+1)}$ as $\hat{V}_k$ for k is odd for short of notation. Also, we define $\check{V}_{k,s} = \hat{V}_{\lfloor k+2 \rceil, s}$.

The following definition introduces the notion of a dependency graph (see also Janson 2004).

Definition A.3.1 (Dependency graph)**.** For given random variables $R_1, \cdots, R_n$, $W_n \in \{0,1\}^{n \times n}$ is a non-random matrix defined as dependecy graph of $(R_1, \cdots, R_n)$ if, for any i, $R_i \perp R_{j:W_n^{(i,j)}=0}$. We denote the dependency neighbors $\mathscr{N}_i = \{j : W_n^{(i,j)} = 1\}$. □

Intuitively, a dependency graph denotes a deterministic adjacency matrix with entry (i, j)

equal to one if (i, j) are statistical dependent.

Definition A.3.2 (Proper Cover)**.** Given an adjacency matrix A_n, with n rows and columns, a family $\mathscr{C}_n = \{\mathscr{C}_n(j)\}_j$ of disjoint subsets of $\{1,\cdots,n\}$ is a proper cover of A_n if $\cup_j \mathscr{C}_n(j) = \{1,\cdots,n\}$ and $\mathscr{C}_n(j)$ contains units such that for any pair of elements $\{i,k \in \mathscr{C}_n(j)\}$, then $A_n^{(i,k)} = 0$. □

Namely, a proper cover of A_n defines a set of disjoint sets, where each disjoint set contains some indexes of units that are not neighbors in A_n. Note that a proper cover always exists since, if A_n is fully connected, then the number of disjoint sets is just n, one for each element.

The size of the smallest proper cover is the chromatic number, defined as $\chi(A_n)$.

Definition A.3.3. (Chromatic Number) The chromatic number $\chi(A_n)$, denotes the size of the smallest proper cover of A_n. □

In the following lines we define the oracle descent procedure absent of sampling error. Let $\beta \in \mathscr{B} = [\mathscr{B}_1, \mathscr{B}_2]^p$, where $\mathscr{B}_1, \mathscr{B}_2$ are finite. Also, let $P_{\mathscr{B}_1,\mathscr{B}_2}$ be the projection operator onto $\mathscr{B}$.

Definition A.3.4 (Oracle gradient descent under strong concavity)**.** We define

$$\beta_w^{**} = P_{\mathscr{B}_1,\mathscr{B}_2}\left[\beta_{w-1}^{**} + \alpha_{w-1} V(\beta_{w-1}^{**})\right], \quad \beta_1^{**} = \beta_0, \tag{A.3.1}$$

with $\alpha_w = \frac{\eta}{w+1}$, equal for all clusters. □

Note that in our proofs, we will refer to the general p-dimensional case for the multi-wave experiment, which uses $\breve{T} = T/p$ waves. See Algorithm A.5.1.

Lemmas

Preliminary Lemmas

Lemma A.3.1. *(Ross, 2011) Let $X_1, ..., X_n$ be random variables such that $\mathbb{E}[X_i^4] < \infty$, $\mathbb{E}[X_i] = 0$, $\sigma^2 = \mathrm{Var}(\sum_{i=1}^n X_i)$ and define $W = \sum_{i=1}^n X_i/\sigma$. Let the collection $(X_1, ..., X_n)$ have dependency*

neighborhoods $\mathcal{N}_i$, $i = 1,...,n$ and also define $D = \max_{1\leq i\leq n}|\mathcal{N}_i|$. Then for Z a standard normal random variable, we obtain

$$d_W(W,Z) \leq \frac{D^2}{\sigma^3}\sum_{i=1}^{n}\mathbb{E}|X_i|^3 + \frac{\sqrt{28}D^{3/2}}{\sqrt{\pi}\sigma^2}\sqrt{\sum_{i=1}^{n}\mathbb{E}[X_i^4]}, \tag{A.3.2}$$

where d_W denotes the Wasserstein metric.

Lemma A.3.2. *(Brook's Theorem,Brooks (1941)) For any connected undirected graph G with maximum degree Δ, the chromatic number of G is at most Δ unless G is a complete graph or an odd cycle, in which case the chromatic number is $\Delta+1$.*

Concentration for Local Dependency Graphs

In the following lemma, we study the concentration of the average of locally dependent random variables (see also Janson 2004 for concentration with local dependency graphs) (we provide exact constants which are useful to derive the rule of thumb for η_n in Appendix A.2.5).

Lemma A.3.3 (Concentration for dependency graphs)**.** *Define $\{R_i\}_{i=1}^n$ sub-gaussian random variables with parameter $\sigma^2 < \infty$, forming a dependency graph with adjacency matrix A_n with maximum degree bounded by γ_N. Then with probability at least $1-\delta$, for any $\delta \in (0,1)$,*

$$\left|\frac{1}{n}\sum_{i=1}^{n}(R_i - \mathbb{E}[R_i])\right| \leq \sqrt{\frac{2\sigma^2\gamma_N\log(2\gamma_N/\delta)}{n}}.$$

Proof of Lemma A.3.3. First, we construct a proper cover $\mathscr{C}_n$ as in Definition B.3.1, with chromatic number $\chi(A_n)$. We can write

$$\left|\frac{1}{n}\sum_{i=1}^{n}(R_i - \mathbb{E}[R_i])\right| \leq \sum_{\mathscr{C}_n(j)\in\mathscr{C}_n}\underbrace{\left|\frac{1}{n}\sum_{i\in\mathscr{C}_n(j)}(R_i - \mathbb{E}[R_i])\right|}_{(A)}.$$

Here, we sum over each subset of index $\mathscr{C}_n(j) \in \mathscr{C}_n$ in the proper cover, and then we sum over

each element in the subset $\mathscr{C}_n(j)$. Observe now that by definition of the dependency graph, components in (A) are mutually independent. Using the Chernoff's bound (Wainwright, 2019), we have that with probability at least $1-\delta$, for any $\delta \in (0,1)$

$$\Big| \sum_{i \in \mathscr{C}_n(j)} (R_i - \mathbb{E}[R_i]) \Big| \leq \sqrt{2\sigma^2 |\mathscr{C}_n(j)| \log(2/\delta)},$$

where $|\mathscr{C}_n(j)|$ denotes the number of elements in $\mathscr{C}_n(j)$. As a result, using the union bound, we obtain that with probability at least $1-\delta$, for any $\delta \in (0,1)$

$$\Big| \frac{1}{n} \sum_{i=1}^{n} (R_i - \mathbb{E}[R_i]) \Big| \leq \underbrace{\frac{1}{n} \sum_{\mathscr{C}_n(j) \in \mathscr{C}_n} \sqrt{2\sigma^2 |\mathscr{C}_n(j)| \log(2\chi(A_n)/\delta)}}_{(B)}.$$

Using concavity of the square-root function, after multiplying and dividing (B) by $\chi(A_n)$, we have

$$\begin{aligned}(B) &\leq \frac{1}{n}\chi(A_n) \sqrt{2\sigma^2 \frac{1}{\chi(A_n)} \sum_{\mathscr{C}_n(j) \in \mathscr{C}_n} |\mathscr{C}_n(j)| \log(2\chi(A_n)/\delta)} \\ &= \frac{1}{n} \sqrt{2\sigma^2 \chi(A_n) n \log(\chi(A_n)/\delta)}.\end{aligned}$$

The last equality follows by the definition of proper cover. The final result follows by Lemma B.7.2. □

Proof of Lemma 1.2.1 and Local Dependence

Lemma 1.2.1 is stated as a corollary of Lemma A.3.4.

Lemma A.3.4. *Let Assumption 1.2.1, 1.2.2 hold. For treatment assigned as in Assumption 1.2.3 with exogenous parameter $\beta_{k,t}$ in cluster k at time t, Lemma 1.2.1 hold. Also, $\varepsilon_{i,t}^{(k)} \perp \{\varepsilon_{j,t}^{(k)}\}_{j \notin \mathscr{I}_i^{(k)}} | \beta_{k,t}$ for a set $|\mathscr{I}_i^{(k)}| = \mathscr{O}(\gamma_N)$.*

Proof of Lemma A.3.4. Under Assumption 1.2.2, we can write for some function g,

$$r\Big(D_{i,t}^{(k)}, D_{j:A_{i,j}^{(k)}>0,t}^{(k)}, X_i^{(k)}, X_{j:A_{i,j}^{(k)}>0}^{(k)}, A_i^{(k)}, U_i^{(k)}, U_{j:A_{i,j}^{(k)}>0}^{(k)}, \nu_{i,t}^{(k)}\Big) = g(Z_{i,t}^{(k)}).$$

Here, $Z_{i,t}^{(k)}$ depends on $A_i^{(k)}$, i.e., the edges of individual i, and on unobservables and observables of all those individuals such that $A_{i,j}^{(k)} > 0$, namely,

$$Z_{i,t}^{(k)} = \Big[D_{i,t}^{(k)}, X_i^{(k)}, \nu_{i,t}^{(k)}, A_i^{(k)} \otimes \Big(X^{(k)}, U^{(k)}, D_t^{(k)}\Big)\Big].$$

Importantly, under Assumption 1.2.1, $A_i^{(k)}$ is a function of $\Big\{\Big[X_j^{(k)}, U_j^{(k)}\Big], j : 1\{i_k \leftrightarrow j_k\} = 1\Big\}$, only, and each entry depends on (X_j, U_j, X_i, U_i) through the same function f for each individual. What is important, is that $\sum_j 1\{i_k \leftrightarrow j_k\} = \gamma_N^{1/2}$ for each unit i. Therefore, for some function $\tilde{g}$ (which depends on f in Assumption 1.2.1), we can equivalently write

$$Z_{i,t}^{(k)} = \tilde{g}(D_{i,t}^{(k)}, \nu_{i,t}^{(k)}, X_i^{(k)}, \tilde{Z}_{i,t}^{(k)}), \quad \tilde{Z}_{i,t}^{(k)} = \Big\{\Big[X_j^{(k)}, U_j^{(k)}, D_{j,t}^{(k)}\Big], j : 1\{i_k \leftrightarrow j_k\} = 1\Big\},$$

where $\tilde{Z}_{i,t}^{(k)}$ is the vector of $\Big[X_j^{(k)}, U_j^{(k)}, D_{j,t}^{(k)}\Big]$ of all individuals j with $1\{i_k \leftrightarrow j_k\} = 1$.

Now, observe that since $(U_i^{(k)}, X_i^{(k)}) \sim_{i.i.d.} F_{X|U}F_U$, and $\{\nu_{i,t}\}$ are $i.i.d.$ conditionally on $U^{(k)}, X^{(k)}$ (Assumption 1.2.2) and treatments are randomized independently (Assumption 1.2.3), we have

$$\Big[X_j^{(k)}, U_j^{(k)}, \nu_{j,t}^{(k)}, D_{j,t}^{(k)}\Big] \Big| \beta_{k,t} \sim_{i.i.d.} \mathscr{D}(\beta_{k,t})$$

is $i.i.d$ with some distribution $\mathscr{D}(\beta_{k,t})$ which only depends on the exogenous coefficient $\beta_{k,t}$ governing the distribution of $D_{i,t}^{(k)}$ under Definition 1.2.3. As a result for $\beta_{k,t}$ being exogenous, Lemma 1.2.1 holds since $\sum_j 1\{i_k \leftrightarrow j_k\} = \gamma_N^{1/2}$ for all i, hence $\tilde{Z}_{i,t}$ are identically distributed across units i. Similarly, also $\varepsilon_{i,t}^{(k)}|\beta_{k,t}$ is a measurable function of a vector $\Big[X_j^{(k)}, U_j^{(k)}, \nu_{j,t}^{(k)}, D_{j,t}^{(k)}\Big]_{j:1\{i_k \leftrightarrow j_k\}=1}$.[11] As a result since such vectors are independent conditional on $\beta_{k,t}$, $\varepsilon_{i,t}^{(k)}$ is mutally independent

[11] Here for notational convenience convenience only, we are letting $1\{i_k \leftrightarrow i_k\} = 1$.

with $\varepsilon_{v,t}^{(k)}$ for all v such that they do not share a common element $\left[X_j^{(k)}, U_j^{(k)}, \nu_{j,t}^{(k)}, D_{j,t}^{(k)}\right]$, that is, such that $\max_j 1\{i_k \leftrightarrow j_k\} 1\{v_k \leftrightarrow j_k\} = 0$.

There are at most $\gamma_N^{1/2} + \gamma_N$ many of $\varepsilon_{v,t}^{(k)}$ which can share a common neighbor with $\varepsilon_{i,t}^{(k)}$ ($\gamma_N^{1/2}$ many neighhbors and γ_N many neighbors of the neighbors), which concludes the proof. □

Concentration of the Average Outcomes

In this subsection, we provide three auxiliary lemmas.

Lemma A.3.5. *Suppose that treatments are assigned as in Assumption 1.2.3 with*

$$D_{i,0}^{(k)} \sim \pi(X_i^{(k)}, \beta_0), \quad D_{i,0}^{(k+1)} \sim \pi(X_i^{(k+1)}, \beta_0)$$
$$D_{i,t}^{(k)} \sim \pi(X_i^{(k)}, \beta), \quad D_{i,t}^{(k+1)} \sim \pi(X_i^{(k+1)}, \beta')$$

with exogenous parameters β_0, β, β'. Let Assumption 1.2.1, 1.2.2 hold. Then with probability at least $1-\delta$, for any $\delta \in (0,1)$

$$\left|\bar{Y}_t^{(k)} - \bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k)} + \bar{Y}_0^{(k+1)} - \int (y(x,\beta) - y(x,\beta'))dF_X(x)\right| = \mathcal{O}(\sqrt{\frac{\gamma_N \log(\gamma_N/\delta)}{n}}).$$

Proof of Lemma A.3.5. First, note that by Lemma A.3.4, we can write

$$\begin{aligned} \mathbb{E}\left[\bar{Y}_t^{(k)} - \bar{Y}_t^{(k+1)}\right] &= \int (y(x,\beta) - y(x,\beta'))dF_X(x) + \tau_k - \tau_{k+1} \\ \mathbb{E}\left[\bar{Y}_0^{(k)} - \bar{Y}_0^{(k+1)}\right] &= \tau_k - \tau_{k+1}. \end{aligned} \tag{A.3.3}$$

In addition, by Lemma A.3.4, $\varepsilon_{i,t}^{(k)}$ (and so $Y_{i,t}^{(k)}$) form a dependency graph with maximum degree bounded by γ_N. The proof completes by invoking Lemma A.3.3. □

Lemma A.3.6. *Let $y(x,\beta)$ be twice differentiable with uniformly bounded derivatives for all*

$x \in \mathscr{X}, \beta \in \mathscr{B}$. *Then for all* $\beta \in \mathscr{B}$, *where* $\mathscr{B}$ *is a compact space*

$$\int \Big[y(x, \beta + \eta_n \underline{e}_j) - y(x, \beta - \eta_n \underline{e}_j) \Big] dF_X(x) = 2\eta_n \int \frac{\partial y(x,\beta)}{\partial \beta^j} dF_X(x) + \mathscr{O}(\eta_n^2)$$
$$= 2\eta_n V^{(j)}(\beta) + \mathscr{O}(\eta_n^2).$$

Proof of Lemma A.3.6. We can write

$$y(x, \beta + \eta_n \underline{e}_j) = y(x,\beta) + \frac{\partial y(x,\beta)}{\partial \beta^j} \eta_n + \mathscr{O}(\eta_n^2)$$
$$y(x, \beta - \eta_n \underline{e}_j) = y(x,\beta) - \frac{\partial y(x,\beta)}{\partial \beta^j} \eta_n + \mathscr{O}(\eta_n^2)$$

from the mean-value theorem which guarantees that the first equality holds. The second equality holds by the dominated convergence theorem. □

Lemma A.3.7. *Let the conditions in Lemma A.3.5 hold. Let* $y(x,\beta)$ *be twice differentiable in* β *with uniformly bounded derivatives for all* $x \in \mathscr{X}, \beta \in \mathscr{B}$. *Suppose that* $\beta = \check{\beta} + \eta_n \underline{e}_j$ *and* $\beta' = \check{\beta} - \eta_n \underline{e}_j$. *Then with probability at least* $1 - \delta$, *for any* $\delta \in (0,1)$

$$\Big| \frac{\bar{Y}_t^{(k)} - \bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k)} + \bar{Y}_0^{(k+1)}}{2\eta_n} - V^{(j)}(\check{\beta}) \Big| = \mathscr{O}\Big(\sqrt{\frac{\gamma_N \log(\gamma_N/\delta)}{\eta_n^2 n}} + \eta_n \Big)$$

Proof of Lemma A.3.7. Using Lemma A.3.5 and the triangular inequality, with probability at least $1 - \delta$,

$$\Big| \frac{\bar{Y}_t^{(k)} - \bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k)} + \bar{Y}_0^{(k+1)}}{2\eta_n} - V^{(j)}(\check{\beta}) \Big|$$
$$\leq \Big| \frac{1}{2\eta_n} \int (y(x, \check{\beta} + \eta_n \underline{e}_j) - y(x, \check{\beta} - \eta_n \underline{e}_j)) dF_X(x) - V^{(j)}(\check{\beta}) \Big| + \mathscr{O}\Big(\sqrt{\frac{\gamma_N \log(\gamma_N/\delta)}{\eta_n^2 n}} \Big).$$

By Lemma A.3.6

$$\int (y(x, \check{\beta} + \eta_n \underline{e}_j) - y(x, \check{\beta} - \eta_n \underline{e}_j)) dF_X(x) = 2\eta_n V^{(j)}(\check{\beta}) + \mathscr{O}(\eta_n^2),.$$

□

Proof of Lemma 1.4.1

To prove the claim it suffices to show that $\check{\beta}_k^w$ is independent of potential outcomes and covariates in cluster k for all $w \in \{1, \cdots, \check{T}\}$, since $\beta_{k,t}$ is a deterministic function of some coefficient $\check{\beta}_k^w$ (see Algorithm A.5.1). Take k to be odd. To show that the claim holds it suffices to show that $\check{\beta}_k^w$ is a function of observables and unobservables only of those units in clusters $k' \notin \{k, k+1\}$. The recursive claim that we want to prove is the following: for all w, $\check{\beta}_k^w$ is exogenous with respect to potential outcomes and covariates in clusters with index $\{h > \lfloor k+2w+1 \rceil$ or $h \in \{k, k+1\}\}$. Clearly, for $\check{\beta}_k^1$ the lemma holds, since $\check{\beta}_k^1$ depends on the gradient in the pair $\{\lfloor k+2 \rceil, \lfloor k+3 \rceil\}$ only. Suppose that the lemma holds for all $w \leq \check{T} - 1$. Then consider $\check{\beta}_k^{\check{T}}$. Observe that $\check{\beta}_k^{\check{T}}$ is chosen based on the gradient $\hat{V}_{k+2,\check{T}-1}$ estimated in the previous wave in clusters $\{\lfloor k+2 \rceil, \lfloor k+3 \rceil\}$, and $\check{\beta}_k^{\check{T}-1}$. By the recursive algorithm, $\check{\beta}_k^{\check{T}-1}$ is exogenous with respect to covariates and potential outcomes in clusters with index $\{h > \lfloor k+2\check{T}-1 \rceil$ or $h \in \{k, k+1\}\}$, which is possible since $K \geq 2\check{T}$, hence $\lfloor k+2\check{T}-1 \rceil < k$. We only need to prove exogeneity of $\hat{V}_{k+2,\check{T}-1}$. The gradient estimated $\hat{V}_{k+2,\check{T}-1}$ is a function of the unobservables and observables at any time $t \leq T$ (where $T = \check{T}p$) in clusters $\{\lfloor k+2 \rceil, \lfloor k+3 \rceil\}$ and the policy $\check{\beta}_{k+2}^{\check{T}-1}$. Since $K \geq 2\check{T}$, again by the recursive algorithm $\check{\beta}_{k+2}^{\check{T}-1}$ is exogenous with respect to potential outcomes and covariates in clusters with index $\{h \geq \lfloor k+2\check{T} \rceil$ or $h \in \{k, k+1\}\}$ which completes the proof.

Lemmas for the Adaptive Experiment with Strong Concavity

In this section, we discuss theoretical guarantees of the algorithm, assuming the global strong concavity of the objective function $W(\beta)$. The following lemma follows by standard properties of the gradient descent algorithm (Bottou et al., 2018).

Lemma A.3.8. *For the learning rate as $\alpha_w = \eta/(w+1)$, and β_w^{**} as defined in Equation* (A.3.6)*, under Assumption 1.3.1, 1.4.1, 1.4.2, with σ-strong concavity, for $M \geq \eta \geq 1/\sigma$, for*

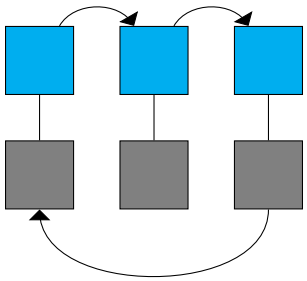


Figure A.3.1. Proof's strategy. Let $p = 1$. Since we have three clusters pairs (each pair of boxes), by assumption $T = 2$. Then the treatments at $T = 2$ in the first pair are assigned using information from the second pair at $T = 1$. Treatments in the second pair at $T = 1$, depend on information at $T = 0$ in the third pair. Hence, the parameter used at $T = 2$ in the first pair must be independent of covariates and potential outcomes in the first pair of clusters. The same reasoning applies to the other pairs of clusters.

any $M \in [1/\sigma, \infty)$*, and let* $L = \max\{2(\mathscr{B}_2 - \mathscr{B}_1)^2, G^2M^2\}$*, with* $G = \sup_\beta ||\frac{\partial W(\beta)}{\partial \beta}||_\infty$*. Then the following holds:*

$$||\beta_w^{**} - \beta^*||^2 \leq \frac{Lp}{w}$$

for a constant $L < \infty$*.*

Proof of Lemma A.3.8. The proof follows standard arguments of the gradient descent method (Bottou et al., 2018), where, here, we leverage strong concavity and the assumption that the gradient is uniformly bounded. Denote β^* the estimand of interest and recall the definition of β_w^{**} in Equation (A.3.6). We define ∇_{w-1} the gradient evaluated at β_{w-1}^{**}. From strong concavity, we can write

$$W(\beta^*) - W(\beta_w^{**}) \leq \frac{\partial W(\beta_w^{**})}{\partial \beta}(\beta^* - \beta_w^{**}) - \frac{\sigma}{2}||\beta^* - \beta_w^{**}||_2^2$$
$$W(\beta_w^{**}) - W(\beta^*) \leq \frac{\partial W(\beta^*)}{\partial \beta}(\beta_w^{**} - \beta^*) - \frac{\sigma}{2}||\beta^* - \beta_w^{**}||_2^2.$$

As a result, since $\frac{\partial W(\beta^*)}{\partial \beta} = 0$, we have

$$\Big(\frac{\partial W(\beta^*)}{\partial \beta} - \frac{\partial W(\beta_w^{**})}{\partial \beta}\Big)(\beta^* - \beta_w^{**}) = \frac{\partial W(\beta_w^{**})}{\partial \beta}(\beta^* - \beta_w^{**}) \geq \sigma||\beta_w^{**} - \beta^*||_2^2. \tag{A.3.4}$$

In addition, we can write:

$$||\beta_w^{**} - \beta^*||_2^2 = ||\beta^* - P_{\mathscr{B}_1,\mathscr{B}_2}(\beta_w^{**} + \alpha_{w-1}\nabla_{w-1})||_2^2 \leq ||\beta^* - \beta_w^{**} - \alpha_{w-1}\nabla_{w-1}||_2^2$$

where the last inequality follows from the fact that $\beta^* \in [\mathscr{B}_1, \mathscr{B}_2]^p$. Observe that we have

$$||\beta^* - \beta_w^{**}||_2^2 \leq ||\beta^* - \beta_{w-1}^{**}||_2^2 - 2\alpha_{w-1}\nabla_{w-1}(\beta^* - \beta_{w-1}^{**}) + \alpha_{w-1}^2||\nabla_{w-1}||_2^2.$$

Using Equation (A.3.4), we can write

$$||\beta_{w+1}^{**} - \beta^*||_2^2 \leq (1 - 2\sigma\alpha_w)||\beta_w^{**} - \beta^*||_2^2 + \alpha_w^2 G^2 p.$$

We now prove the statement by induction. Clearly, at time $w = 1$, the statement trivially holds. Consider a general time w. Then using the induction argument, we write

$$\begin{aligned}||\beta_{w+1}^{**} - \beta^*||_2^2 \leq (1 - 2\frac{1}{w+1})\frac{Lp}{w} + \frac{Lp}{(w+1)^2} &\leq (1 - 2\frac{1}{w+1})\frac{Lp}{w} + \frac{Lp}{w(w+1)} \\ &= (1 - \frac{1}{w+1})\frac{Lp}{w} = \frac{Lp}{w+1}.\end{aligned}$$

□

Lemma A.3.9. *Let Assumption 1.2.2, 1.2.1, 1.4.1 hold. Let α_w be as defined in Lemma A.3.8. Then with probability at least $1 - \delta$, for any $\delta \in (0,1)$, for all $w \geq 1$,*

$$\Big|\Big|P_{\mathscr{B}_1,\mathscr{B}_2 - \eta_n}\Big[\sum_{s=1}^{w}\alpha_s \check{V}_{k,s}\Big] - P_{\mathscr{B}_1,\mathscr{B}_2}\Big[\sum_{s=1}^{w}\alpha_s V(\beta_s^{**})\Big]\Big|\Big|_\infty = \mathscr{O}(P_{\check{T}}(\delta))$$

where $P_1(\delta) = \alpha_1 \times \mathrm{err}(\delta)$ and $P_w(\delta) = B_p\alpha_w P_{w-1}(\delta) + P_{w-1}(\delta) + \alpha_w \mathrm{err}_w(\delta)$, for a finite constant $B_p < \infty$, and $\mathrm{err}_w(\delta) = \mathscr{O}\Big(\sqrt{\gamma_N \frac{\log(p\check{T}K/\delta)}{\eta_n^2 n}} + p\eta_n\Big)$.

Proof of Lemma A.3.9. Recall that by Lemma A.3.7 we can write for every k and $w \in \{1, \cdots, \check{T}\}$

(here using the union bound),

$$\check{V}_{k,w}^{(j)} = V^{(j)}(\check{\beta}_{k+2}^{w}) + \mathscr{O}\Big(\sqrt{\gamma_N \frac{\log(K\check{T}/\delta)}{\eta_n^2 n}} + \eta_n\Big).$$

We now proceed by induction. We first prove the statement, assuming that the constraint is always attained. We then discuss the case of the constraint not being attained. Define (where we suppress the dependence with p for simplicity)

$$B = p \sup_{\beta} \Big|\Big| \frac{\partial^2 W(\beta)}{\partial \beta^2} \Big|\Big|_{\infty}.$$

Unconstrained case

Consider $w = 1$. Then since we initialize parameters at β_0 (recall that $\beta_0 = \beta_1^{**}$), for all clusters, we can write with probability $1-\delta$, for any $\delta \in (0,1)$,

$$\Big|\Big| \alpha_1 \check{V}_{k,1} - \alpha_1 V(\beta_0) \Big|\Big|_{\infty} = \alpha_1 \text{err}(\delta).$$

Consider $t = 2$, then we obtain for every $j \in \{1, \cdots, p\}$,

$$\alpha_2 \check{V}_{k,2}^{(j)} = \alpha_2 V^{(j)}(\check{\beta}_{k+2}^{2}) + \alpha_2 \text{err}(\delta) = \alpha_2 V^{(j)}(\beta_1^{**} + \alpha_1 V(\beta_1^{**}) + \alpha_1 \check{V}_{k,w} - \alpha_1 V(\beta_1^{**})) + \alpha_2 \text{err}(\delta).$$

Using the mean value theorem and Assumption 1.3.1, we obtain

$$\begin{aligned}
&\Big|\Big| \alpha_2 \check{V}_{k,2} - \alpha_2 V(\beta_2^{**}) \Big|\Big|_{\infty} \leq \alpha_2 \text{err}(\delta) + B\alpha_2\alpha_1 \text{err}(\delta) \\
&\Rightarrow \Big|\Big| \sum_{w=1}^{2} \alpha_w \check{V}_{k,w} - \sum_{w=1}^{2} \alpha_w V(\beta_w^{**}) \Big|\Big|_{\infty} \leq \alpha_2 \text{err}(\delta) + B\alpha_2\alpha_1 \text{err}(\delta) + \alpha_1 \text{err}(\delta).
\end{aligned}$$

Consider now a general w. Then we can write with probability $1-\delta$, for any $\delta \in (0,1)$,

$$\alpha_w \check{V}_{k,w} = \alpha_w V(\check{\beta}_{k+2}^{w-1}) + \alpha_w \text{err}(\delta).$$

Let $\tilde{P}_w^{(j)}(\delta) = \alpha_w \tilde{P}_{w-1}^{(j)}(\delta) + \tilde{P}_{w-1}^{(j)}(\delta) + \alpha_w \text{err}(\delta)$, with $\tilde{P}_1^{(j)}(\delta) = \alpha_1 \text{err}(\delta)$, the cumulative error for the jth coordinate. Then, recursively, we have (here, $\tilde{P}_{w-1}(\delta)$ is the vector of cumulative errors)

$$\alpha_w \check{V}_{k,w} = \alpha_w V(\beta_w^{**} + \tilde{P}_{w-1}(\delta)) + \alpha_w \text{err}(\delta).$$

Using the mean value theorem and Assumption 1.3.1, we obtain

$$\alpha_w \check{V}_{k,w} = \alpha_w V(\beta_w^{**}) + \alpha_w B \max_j \tilde{P}_{w-1}^{(j)}(\delta) + \alpha_w \text{err}(\delta).$$

Therefore, with probability $1 - w\tilde{\delta}$ (using the union bound)

$$\begin{aligned}\Big|\Big|\sum_{s=1}^{w} \alpha_s \check{V}_{k,s} - \sum_{s=1}^{w} \alpha_s V(\beta_s^{**})\Big|\Big|_\infty &\leq \Big|\Big|\alpha_w \check{V}_{k,w} - \alpha_w V(\beta_w^{**})\Big|\Big|_\infty + \Big|\Big|\sum_{s=1}^{w-1} \alpha_s \check{V}_{k,s} - \sum_{s=1}^{w-1} \alpha_s V(\beta_s^{**})\Big|\Big|_\infty \\ &\leq \alpha_w B P_{w-1}(\tilde{\delta}) + \alpha_w \text{err}(\tilde{\delta}) + P_{w-1}(\tilde{\delta}),\end{aligned}$$

where $P_{w-1}(\tilde{\delta})$ captures the largest cumulative error up-to iteration $w-1$ as defined in the statement of the lemma (the log-term as a function of p follows from the union bound). The proof completes once we write $\delta = \tilde{\delta}/w$.

Constrained case

Since the statement is true for $w = 1$, we can assume that it is true for all $s \leq w-1$ and prove the statement by induction. Since $\mathscr{B}$ is a compact space, we can write,

$$\begin{aligned}&\Big|\Big|P_{\mathscr{B}_1,\mathscr{B}_2-\eta_n}\Big[\sum_{s=1}^{w} \alpha_s \check{V}_{k,s}\Big] - P_{\mathscr{B}_1,\mathscr{B}_2}\Big[\sum_{s=1}^{w} \alpha_s V(\beta_s^{**})\Big]\Big|\Big|_\infty \\ &\leq \Big|\Big|P_{\mathscr{B}_1,\mathscr{B}_2-\eta_n}\Big[\sum_{s=1}^{w} \alpha_s \check{V}_{k,s}\Big] - P_{\mathscr{B}_1,\mathscr{B}_2-\eta_n}\Big[\sum_{s=1}^{w} \alpha_s V(\beta_s^{**})\Big]\Big|\Big|_\infty + \mathscr{O}(p\eta_n) \\ &\leq 2\Big|\Big|\sum_{s=1}^{w} \alpha_s \check{V}_{k,s} - \sum_{s=1}^{w} \alpha_s V(\beta_s^{**})\Big|\Big|_\infty + \mathscr{O}(p\eta_n)\end{aligned}$$

completing the proof. □

Lemma A.3.10. *Let the conditions in Lemma A.3.9 hold. Then with probability at least* $1-\delta$*, for any* $\delta \in (0,1)$*, for all* $w \geq 1, k \in \{1, \cdots, K\}$*,*

$$||\beta^* - \check{\beta}_k^w||_2^2 \leq \frac{Lp}{w} + pw^{2B_p} \times \mathscr{O}\Big(\gamma_N \frac{\log(p\check{T}K/\delta)}{\eta_n^2 n} + p^2\eta_n^2\Big),$$

for finite constants $B_p, L < \infty$*.*

Proof. Using the triangular inequality, we can write

$$||\beta^* - \check{\beta}_k^w||_2^2 \leq ||\beta^* - \beta_w^{**}||_2^2 + ||\check{\beta}_k^w - \beta_w^{**}||_2^2.$$

The first component on the right-hand side is bounded by Lemma A.3.8. Using Lemma A.3.9, we bound the second component with probability at least $1-\delta$, as follows

$$||\check{\beta}_k^w - \beta_w^{**}||_2^2 \leq p||\check{\beta}_k^w - \beta_w^{**}||_\infty^2 = p \times \mathscr{O}(P_w^2(\delta)).$$

We conclude the proof by explicitely characterizing the rate of $P_w(\delta)$ as defined in Lemma A.3.9. Following Lemma A.3.9, we can define recursively $P_w(\delta)$ for any $1 \leq w \leq \check{T}$ (recall that $\alpha_w \propto 1/w$) as

$$P_w(\delta) = (1 + \frac{B}{w})P_{w-1}(\delta) + \frac{1}{w}\text{err}_n(\delta), \quad P_1(\delta) = \text{err}_n(\delta).$$

where $\text{err}_n = \mathscr{O}(\sqrt{\gamma_N \frac{\log(p\check{T}K/\delta)}{\eta_n^2 n}} + p\eta_n)$. Take, without loss of generality, $B \geq 1$ (if $B < 1$), we can find an upper bound with a different $B = 1$. Substituting recursively each term, we can write[12]

$$P_w(\delta) \leq \text{err}_n(\delta) \sum_{s=1}^{w} \frac{1}{s} \prod_{j=s}^{w} (\frac{B}{j} + 1).$$

[12]The expression is $\leq$ instead of $=$ since the first term in the expression $\text{err}_n(\delta)$ multiplies by $(B/w+1)$ instead of just 1.

We now write

$$\sum_{s=1}^{w}\frac{1}{s}\prod_{j=s}^{w}(\frac{B}{j}+1)\leq\sum_{s=1}^{w}\frac{1}{s}\exp(\sum_{j=s}^{w}\frac{B}{j})\leq\sum_{s=1}^{w}\frac{1}{s}\exp\Big(1+B\log(w)-B\log(s)\Big)$$
$$\lesssim\sum_{s=1}^{w}\frac{1}{s^2}e^{B\log(w)+1}\lesssim w^B,$$

completing the proof.

□

Proof of the Theorems in the Main Text

Proof of Theorem 1.3.1

First observe that for any $\delta\in(0,1)$,

$$\mathbb{E}\Big[\hat{V}_k(\beta)\Big]=V^{(1)}(\beta)+\mathscr{O}(\eta_n),\quad P\Big(\Big|\hat{V}_k(\beta)-V^{(1)}(\beta)\Big|>\mathscr{O}\Big(\eta_n+\sqrt{\frac{\gamma_N\log(\gamma_N/\delta)}{n\eta_n^2}}\Big)\Big)\leq\delta,$$

with the proof of the first claim follows similarly as in the proof of Lemma A.3.6 and the second claim being a direct corollary of Lemma A.3.7. Finally observe that with probability at least $1-\delta$, for any $\delta\in(0,1)$, we also have

$$\Big|\widehat{V}_k(\beta)-V^{(1)}(\beta)\Big|=\mathscr{O}(\eta_n)+\mathscr{O}\Big(\sqrt{\frac{\rho_n}{\delta n\eta_n^2}}\Big),$$

by Chebishev inequality and the triangular inequality.

Proof of Theorem 1.3.2

Consider Algorithm 2 for a generic coordinate j. Let β be the target parameter as in Algorithm 2. By Lemma A.3.6, we have

$$\mathbb{E}[\widehat{V}_k^{(j)}]=V^{(j)}(\beta)+\mathscr{O}(\eta_n).$$

We have

$$\frac{\widehat{V}_k^{(j)} - \mathbb{E}[\widehat{V}_k^{(j)}]}{\sqrt{\mathrm{Var}(\widehat{V}_k^{(j)})}} = \frac{\widehat{V}_k^{(j)} - V^{(j)}(\beta)}{\sqrt{\mathrm{Var}(\widehat{V}_k^{(j)})}} + \mathscr{O}\Big(\frac{\eta_n}{\sqrt{\mathrm{Var}(\widehat{V}_k^{(j)})}}\Big).$$

Observe that under Assumption 1.3.2,

$$\mathscr{O}\Big(\frac{\eta_n}{\sqrt{\mathrm{Var}(\widehat{V}_k^{(j)})}}\Big) \leq \mathscr{O}(\eta_n^2 \times \sqrt{n}).$$

First, observe that by Lemma A.3.4, and the fact that covariates are independent, then $Y_{i,t}^{(k)} - Y_{i,0}^{(k)}$ form a locally dependent graph of maximum degree of order $\mathscr{O}(\gamma_N)$. We now invoke Lemma A.3.1. We can write

$$\begin{aligned}
&d_W\Big(\frac{1}{2\eta_n\sqrt{\mathrm{Var}(\widehat{V}_k^{(j)})}}\Big[\bar{Y}_t^{(k)} - \bar{Y}_0^{(k)}\Big] - \frac{1}{2\eta_n\sqrt{\mathrm{Var}(\widehat{V}_k^{(j)})}}\Big[\bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k+1)}\Big], \mathscr{G}\Big) \\
&\leq \underbrace{\frac{\gamma_N^2}{\sigma^3}\sum_{h\in\{k,k+1\}}\sum_{i=1}^{n}\Big[\mathbb{E}\Big|\frac{Y_{i,t}^{(k)} - Y_{i,0}^{(k)}}{\eta_n n}\Big|^3\Big]}_{(A)} + \underbrace{\frac{\sqrt{28}\gamma_N^{3/2}}{\sqrt{\pi}\sigma^2}\sqrt{\sum_{i=1}^{n}\Big[\mathbb{E}\Big|\frac{Y_{i,t}^{(k)} - Y_{i,0}^{(k)}}{\eta_n n}\Big|^4\Big]}}_{(B)}, \\
&\mathscr{G} \sim \mathscr{N}(0,1), \quad \sigma^2 = \mathrm{Var}\Big(\frac{1}{2\eta_n}\Big[\bar{Y}_t^{(k)} - \bar{Y}_0^{(k)}\Big] - \frac{1}{2\eta_n}\Big[\bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k+1)}\Big]\Big)
\end{aligned}$$

and d_W denotes the Wasserstein metric. We now inspect each argument on the right hand side. Under Assumption 1.3.2, $\sigma^2 \geq C_k \frac{1}{n\eta_n^2}$ for a constant $C_k > 0$, and the third and fourth moment are bounded. Hence, we have for a constant $C' < \infty$,

$$(A) \leq C' \frac{\gamma_N^2}{n^3\eta_n^3} \times n^{5/2}\eta_n^3 \asymp \frac{\gamma_N^2}{n^{1/2}} \to 0.$$

Similarly, for (B), we have

$$(B) \leq c' \frac{\gamma_N^{3/2} n\eta_n^2}{\eta_n^2 n^{3/2}} \asymp \frac{\gamma_N^{3/2}}{n^{1/2}} \to 0.$$

The proof completes.

Proof of Theorem 1.3.3

By Lemma 1.2.1, we can write (we omit the superscript k from $X^{(k)}$ for sake of brevity)

$$\mathbb{E}\Big\{\frac{1}{2n}\sum_{i=1}^{n}\Big[\frac{D_{i,1}^{(k+1)}Y_{i,1}^{(k+1)}}{\pi(X_i,\beta+\eta_n\underline{e}_1)}-\frac{(1-D_{i,1}^{(k+1)})Y_{i,1}^{(k+1)}}{1-\pi(X_i,\beta+\eta_n\underline{e}_1)}+\frac{D_{i,1}Y_{i,1}^{(k)}}{\pi(X_i,\beta-\eta_n\underline{e}_1)}-\frac{(1-D_{i,1}^{(k)})Y_{i,1}^{(k)}}{1-\pi(X_i,\beta-\eta_n\underline{e}_1)}\Big]\Big\}$$

$$=\frac{1}{2}\mathbb{E}\Big\{\Big[\frac{D_{i,1}^{(k+1)}Y_{i,1}^{(k+1)}}{\pi(X_i,\beta+\eta_n\underline{e}_1)}-\frac{(1-D_{i,1}^{(k+1)})Y_{i,1}^{(k+1)}}{1-\pi(X_i,\beta+\eta_n\underline{e}_1)}\Big]+\frac{1}{2n}\sum_{i=1}^{n}\Big[\frac{D_{i,1}Y_{i,1}^{(k)}}{\pi(X_i,\beta-\eta_n\underline{e}_1)}-\frac{(1-D_{i,1}^{(k)})Y_{i,1}^{(k)}}{1-\pi(X_i,\beta-\eta_n\underline{e}_1)}\Big]\Big\}$$

$$=\underbrace{\frac{1}{2}\int\Big[m(1,x,\beta+\eta_n\underline{e}_1)-m(0,x,\beta+\eta_n\underline{e}_1)+m(1,x,\beta-\eta_n\underline{e}_1)-m(0,x,\beta-\eta_n\underline{e}_1)\Big]dF_X(x)}_{(i)}.$$

The last equality follows from Lemma 1.2.1 and exogeneity of β. Doing a Taylor expansion to each component around β, we obtain that (i) equals

$$\int\Big[m(1,x,\beta)-m(0,x,\beta)+\frac{\partial m(1,x,\beta)}{2\partial\beta^1}\eta_n-\frac{\partial m(0,x,\beta)}{2\partial\beta^1}\eta_n-\frac{\partial m(1,x,\beta)}{2\partial\beta^1}\eta_n+\frac{\partial m(0,x,\beta)}{2\partial\beta^1}\eta_n\Big]dF_X(x)$$

$$+\mathscr{O}(\eta_n^2)=\int\Big[m(1,x,\beta)-m(0,x,\beta)\Big]dF_X(x)+\mathscr{O}(\eta_n^2),$$

which completes the proof, since $\eta_n=o(n^{-1/4})$.

Proof of Theorem 1.3.4

We are interested in studying

$$\mathbb{E}\Big\{\frac{1}{2n}\sum_{h\in\{k,k+1\}}\frac{v_h}{\eta_n}\sum_{i=1}^{n}\Big[\frac{Y_{i,1}^{(h)}(1-D_{i,1}^{(h)})}{1-\pi(X_i^{(h)},\beta+v_h\eta_n\underline{e}_1)}-\bar{Y}_0^{(h)}\Big]\Big\},\quad v_h=\begin{cases}1 \text{ if } h=k\\ -1 \text{ otherwise.}\end{cases}$$

Using Lemma 1.2.1, similarly to the derivation of Lemma A.3.6, we can write the above expression equal to

$$\frac{1}{2\eta_n}\int[m(0,x,\beta+\eta_n\underline{e}_1)-m(0,x,\beta-\eta_n\underline{e}_1)]dF_X(x).$$

Note that from the mean value theorem, and Assumption 1.3.1

$$m(0,x,\beta+\eta_n\underline{e}_1)-m(0,x,\beta-\eta_n\underline{e}_1)=m(0,x,\beta)-m(0,x,\beta)+2\frac{\partial m(0,x,\beta)}{\partial\beta^1}\eta_n+O(\eta_n^2)$$

which completes the proof.

Proof of Theorem 1.4.2

Consider Lemma A.3.10 where we choose $\delta = 1/n$. We can write for each k

$$||\beta^* - \check{\beta}_k^{\check{T}}||_2^2 \leq \frac{pL}{\check{T}} + \mathscr{O}(1/\check{T}),$$

for a finite constant $L < \infty$, since, under the conditions for n stated in the theorem, for finite B, the second component is $\mathscr{O}(1/\check{T})$. Note that

$$||\beta^* - \frac{1}{K}\sum_{k=1}^{K}\check{\beta}_k^{\check{T}}||_2^2 \leq \frac{1}{K}\sum_{k=1}^{K}||\beta^* - \check{\beta}_k^{\check{T}}||_2^2$$

by Jensen's inequality, which completes the proof.

Theorem 1.4.3

By the mean value theorem and Assumption 1.3.1, we have

$$\sum_{w=1}^{\check{T}} W(\beta^*) - W(\check{\beta}_k^w) \leq \bar{C}\sum_{w=1}^{\check{T}}||\beta^* - \check{\beta}_k^w||_2^2,$$

for a finite constant $\bar{C} < \infty$, since $\frac{\partial W(\beta^*)}{\partial \beta} = 0$, and the Hessian is uniformly bounded (Assumption 1.3.1). By Lemma A.3.10, choosing $\delta = 1/n$, and for n satisfying the conditions in Theorem 1.4.3, it follows that for all k,

$$\sum_{w=1}^{\check{T}} W(\beta^*) - W(\check{\beta}_k^w) \leq \sum_{w=1}^{\check{T}} \frac{p\kappa'}{w} \lesssim p\log(\check{T})$$

for $\kappa' < \infty$ being a finite constant. The proof completes.

Proof of Theorem 1.4.4

First, note that for a finite constant c_0, under Assumption 1.3.1 and Assumption 1.4.2

$$W(\beta^*) - W(\hat{\beta}^*) \leq c_0 ||\beta^* - \hat{\beta}||^2 \leq c_0 \frac{1}{K} \sum_{k=1}^{K} ||\beta^* - \check{\beta}_k^{\check{T}+1}||^2$$

where in the first inequality we used strong concavity (gradient equals zero), and in the second equality we used Jensen's inequality. Define β_w^{**} as in Equation (A.3.6), where, however, the learning rate is chosen so that $\alpha_w = 1/\tau$. Using the triangular inequality, we can write

$$||\beta^* - \check{\beta}_k^{\check{T}+1}||_2^2 \leq ||\beta^* - \beta_{\check{T}+1}^{**}||_2^2 + ||\check{\beta}_k^{\check{T}+1} - \beta_{\check{T}+1}^{**}||_2^2.$$

The first component is bounded by Theorem 3.10 in Bubeck (2014) (using the fact that $\mathscr{B}$ is compact) as follows:

$$||\beta^* - \beta_{\check{T}+1}^{**}||_2^2 \leq c_0 \exp(-c_0' 2(\check{T}+1)) = c_0 \exp(-K c_0')$$

for finite constants $0 < c_0, c_0' < \infty$, where we used the fact that $2(\check{T}+1) = K$. Using Lemma A.3.9, we bound the second component with probability at least $1-\delta$, as follows (for any $w \leq \check{T}+1$)

$$||\check{\beta}_k^w - \beta_w^{**}||_2^2 \leq p||\check{\beta}_k^w - \beta_w^{**}||_\infty^2 = p \times \mathscr{O}(P_w^2(\delta)).$$

We conclude the proof by explicitely characterizing the rate of $P_w(\delta)$ as defined in Lemma A.3.9. Following Lemma A.3.9, we can define recursively $P_w(\delta)$ for any $1 \leq w \leq \check{T}$ (recall that $\alpha_w \propto 1/w$) as

$$P_w(\delta) = (1+B)P_{w-1}(\delta) + \text{err}_n(\delta), \quad P_1(\delta) = \text{err}_n(\delta).$$

where $\text{err}_n = \mathcal{O}(\sqrt{\gamma_N \frac{\log(p\check{T}K/\delta)}{\eta_n^2 n}} + p\eta_n)$, and $B > 0$ is a finite constant with depends on p only. Using a recursive argument, we can write

$$P_w(\delta) \leq w(1+B)^w \text{err}_n(\delta).$$

The proof completes as we choose n sufficiently large as stated in the theorem.

Proof of Theorem 1.4.5

Recall that from Assumption 1.2.1, the maximum degree is $\gamma_N^{1/2}$. We break the proof into several steps. We will write the model dividing by $1/\sum_{j\neq i} A_{i,j} 1\{X_j = x\}$ instead of $1/\max\{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}, 1\}$ for notational convenience, but implicitely consider the expression $1/\sum_{j\neq i} A_{i,j} 1\{X_j = x\}$ equal to one if $\sum_{j\neq i} A_{i,j} 1\{X_j = x\} = 0$.

Upper bound on W_N^*

We first provide an upper bound on the largest achievable welfare. Recall that $\Delta = c$. Therefore, we can write

$$W_N^* \leq \frac{1}{N}\sum_{i=1}^N \sup_{\mathscr{P}} \mathbb{E}\Big[\mathbb{E}_{D\sim\mathscr{P}(A,X)}\Big[s\Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}\Big)\Big|A,X\Big]\Big].$$

Let

$$\beta^G = \arg \max_{\beta_1,\cdots,\beta_{|\mathscr{X}|}\in[0,1]^{|\mathscr{X}|}} s\Big(\beta_1,\cdots,\beta_{|\mathscr{X}|}\Big).$$

Note that since $D_j \in \{0,1\}$, we can write

$$\sup_{\mathscr{P}} \mathbb{E}\Big[\mathbb{E}_{D\sim\mathscr{P}(A,X)}\Big[s\Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}\Big)\Big|A,X\Big]\Big] \leq s\Big(\beta_1^G,\cdots,\beta_{|\mathscr{X}|}^G\Big).$$

Lower bound on $W(\beta^*)$

Using the fact that $\mathscr{B} = [0,1]^{|\mathscr{X}|}$, we can write[13]

$$\begin{aligned} W(\beta^*) &= \max_{\beta \in [0,1]^{|\mathscr{X}|}} \mathbb{E}_{\beta}\Big[s\Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}\Big)\Big] \\ &\geq \mathbb{E}_{\beta^G}\Big[s\Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}\Big)\Big], \end{aligned}$$

where we use the fact that $\beta^G = (\beta_1^G, \cdots, \beta_{|\mathscr{X}|}^G) \in [0,1]^{|\mathscr{X}|}$, and $\Delta(\cdot) = c(\cdot)$. From the mean value theorem

$$\begin{aligned} &\mathbb{E}_{\beta^G}\Big[s\Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}\Big)\Big] \\ &= s(\beta^G) + \mathbb{E}_{\beta^G}\Big\{\frac{\partial s(\beta)}{\partial \beta}\Big|_{\beta \in \Big[\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}, \beta^G\Big]} \times \\ &\times \Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}} - \beta^G\Big)\Big\}, \end{aligned}$$

where $\frac{\partial s(\cdot)}{\partial \beta}$ is evaluated at a point between the shared of treated neighbors and β^G.

Bound on the difference

Combining the two bounds, we can write

$$\begin{aligned} &W_N^* - W(\beta^*) \leq \\ &\underbrace{\Big|\mathbb{E}_{\beta^G}\Big\{\frac{\partial s(\beta)}{\partial \beta}\Big|_{\beta \in \Big[\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}, \beta^G\Big]} \times \Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}} - \beta^G\Big)\Big\}\Big|}_{(I)}, \end{aligned}$$

where we took the absolute value in the last equation.

Bound with Cauchy-Schwarz

We can now bound (I) as follows.

$$(I) \leq \sup_{\beta} \Big|\Big|\frac{\partial s(\beta)}{\partial \beta}\Big|\Big|_2 \times |\mathscr{X}| \max_{x\in\mathscr{X}} \underbrace{\sqrt{\mathbb{E}_{\beta^G}\Big[\Big(\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}} - \beta_x^G\Big)^2\Big]}}_{(II)},$$

[13] $\mathbb{E}_{\beta}\Big[s\Big(\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big]_{x\in\{1,\cdots,|\mathscr{X}|\}}\Big)\Big]$ does not depend on i similarly to Lemma 1.2.1.

where we first used Cauchy-Schwarz and then bound the first component by the supremum over β, x and the second component by the largest term over $x \in \mathscr{X}$ times the number of elements $|\mathscr{X}|$.

Bound for (II)

we now want to bound (II). We do so fixing $X_j = x$ and show that for all $x \in \mathscr{X}$ (and so also for the maximum) we can obtain a useful bound. Recall that here $\mathbb{E}_{\beta^G}$ indicates that $D_{i,t}|X_i^{(k)} = x \sim_{i.i.d.} \text{Bern}(\beta_x^G)$. Now, note that

$$\mathbb{E}_{\beta^G}\Big[\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big| X^{(k)}, A^{(k)}\Big] = \beta_x^G.$$

Also, note that since we take the maximum over $x \in \mathscr{X}$, here β_x is fixed, and therefore, $\text{Var}(\beta_x) = 0$. Therefore, we can write

$$\mathbb{E}_{\beta^G}\Big[\Big(\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}} - \beta_x^G\Big)^2\Big] = \mathbb{E}_{\beta^G}\Big[\text{Var}\Big(\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big| X^{(k)}, A^{(k)}\Big)\Big].$$

In addition,

$$\text{Var}_{\beta^G}\Big(\frac{\sum_{j\neq i} A_{i,j} D_j 1\{X_j = x\}}{\sum_{j\neq i} A_{i,j} 1\{X_j = x\}}\Big| X^{(k)}, A^{(k)}\Big) = \beta_x^G(1-\beta_x^G)\Big/\sum_{j\neq i} A_{i,j} 1\{X_j = x\},$$

since treatments are independent conditional on $X^{(k)}$, independent of $A^{(k)}$ conditional on $X^{(k)}$ by construction, and binary. Let $\kappa' = \kappa P(X = x)$, where, without loss of generality, $P(X = x) > 0$ since $|\mathscr{X}| < \infty$, and hence, if $P(X = x) = 0$, we can re-index $s(\cdot)$, for all types except $X = x$ and

conduct the same analysis as above, without the case $X = x$. Note that

$$\begin{aligned}
&\mathbb{E}\Big[\beta_x^G(1-\beta_x^G)\Big/\sum_{j\neq i}A_{i,j}1\{X_j=x\}\Big] = \beta_x^G(1-\beta_x^G)\mathbb{E}\Big[1\Big/\sum_{j\neq i}A_{i,j}1\{X_j=x\}\Big] \\
&\leq \beta_x^G(1-\beta_x^G)P\Big(\sum_{j\neq i}A_{i,j}1\{X_j=x\} < \kappa'\gamma_N^{1/4}\Big) + \beta_x^G(1-\beta_x^G)P\Big(\sum_{j\neq i}A_{i,j}1\{X_j=x\} \geq \kappa'\gamma_N^{1/4}\Big)\frac{1}{\kappa'\gamma_N^{1/4}} \\
&\leq \underbrace{\beta_x^G(1-\beta_x^G)P\Big(\sum_{j\neq i}A_{i,j}1\{X_j=x\} < \kappa'\gamma_N^{1/4}\Big)}_{(III)} + \frac{1}{\kappa'\gamma_N^{1/4}},
\end{aligned}$$

where in the second inequality we used the fact that we are implicitely using in our notation that $1/\sum_{j\neq i}A_{i,j}1\{X_j=x\}$ equals one if $\sum_{j\neq i}A_{i,j}1\{X_j=x\}=0$ for notational convenience.

Bound for (III)

We are left to derive a bound for (III), since the second term converges to zero as $\gamma_N \to \infty$. Define $h_x(X_i,U_i) = P(X=x)\int l(X_i,U_i,x,u)dF_{U|X=x}(u)$. Note that (recall that $1\{i\leftrightarrow j\}$ are fixed)

$$\mathbb{E}[A_{i,j}1\{X_j=x\}|X_i,U_i] = \mathbb{E}[l(X_i,U_i,X_j,U_j)1\{X_j=x\}1\{i\leftrightarrow j\}|X_i,U_i] = h_x(X_i,U_i)1\{i\leftrightarrow j\},$$

since, conditional on (X_i,U_i),the indicator $1\{i\leftrightarrow j\}$ is fixed (exogenous), and $(X_i,U_i) \sim_{i.i.d.} F_X F_{U|X}$. Also, recall that $\sum_j 1\{i\leftrightarrow j\} = \gamma_N^{1/2}$. Hence, only $\gamma_N^{1/2}$ many edges of i can at most be non-zero, while the remaining ones are zero almost surely. Therefore, using Hoeffding's inequality (Wainwright, 2019), and using independence conditional on X_i,U_i,

$$P\Big(\Big|\frac{1}{\gamma_N^{1/2}}\sum_{j\neq i}A_{i,j}1\{X_j=x\} - h_x(X_i,U_i)\Big| \leq \bar{C}\sqrt{\frac{\log(2\gamma_N)}{\gamma_N^{1/2}}}\Big|X_i,U_i\Big) \geq 1-1/\gamma_N, \tag{A.3.5}$$

for a finite constant $\bar{C} < \infty$. Observe that $h_x(X_i,U_i) \geq \kappa' > 0, \kappa' = P(X=x)\kappa$ almost surely by assumption. Define the event

$$\mathscr{E} = \Big\{|\sum_{j\neq i}A_{i,j}1\{X_j=x\} - \gamma_N^{1/2}h_x(X_i,U_i)| \leq \bar{C}\sqrt{\log(2\gamma_N)\gamma_N^{1/2}}\Big\}.$$

We can write

$$P\Big(\sum_{j\neq i}A_{i,j}1\{X_j=x\}<\kappa'\gamma_N^{1/4}\Big)=P\Big(\sum_{j\neq i}A_{i,j}1\{X_j=x\}-\gamma_N^{1/2}h_x(X_i,U_i)+\gamma_N^{1/2}h_x(X_i,U_i)<\kappa'\gamma_N^{1/4}\Big)$$
$$\leq P\Big(\gamma_N^{1/2}h_x(X_i,U_i)<\kappa'\gamma_N^{1/4}+|\sum_{j\neq i}A_{i,j}1\{X_j=x\}-\gamma_N^{1/2}h_x(X_i,U_i)|\Big)$$
$$\leq P\Big(\gamma_N^{1/2}h_x(X_i,U_i)<\kappa\gamma_N^{1/4}+|\sum_{j\neq i}A_{i,j}1\{X_j=x\}-\gamma_N^{1/2}h_x(X_i,U_i)|\Big|\mathscr{E}\Big)$$
$$+P\Big(\gamma_N^{1/2}h_x(X_i,U_i)<\kappa'\gamma_N^{1/4}+|\sum_{j\neq i}A_{i,j}1\{X_j=x\}-\gamma_N^{1/2}h(X_i,U_i)|\Big|\mathscr{E}^c\Big)\times P\Big(\mathscr{E}^c\Big).$$

Note that by Equation (A.3.5) (which holds conditionally and so also unconditionally)

$$P\Big(\gamma_N^{1/2}h_x(X_i,U_i)<\kappa'\gamma_N^{1/4}+|\sum_{j\neq i}A_{i,j}1\{X_j=x\}-\gamma_N^{1/2}h_x(X_i,U_i)|\Big|\mathscr{E}^c\Big)\times P\Big(\mathscr{E}^c\Big)\leq\frac{1}{\gamma_N}=o(1).$$

Finally, we can write for a finite constant $\bar{C}<\infty$,

$$P\Big(\gamma_N^{1/2}h_x(X_i,U_i)<\kappa'\gamma_N^{1/4}+|\sum_{j\neq i}A_{i,j}1\{X_j=x\}-\gamma_N^{1/2}h_x(X_i,U_i)|\Big|\mathscr{E}\Big)$$
$$\leq P\Big(\gamma_N^{1/2}h_x(X_i,U_i)<\kappa'\gamma_N^{1/4}+\bar{C}\sqrt{\log(2\gamma_N)\gamma_N^{1/2}}\Big|\mathscr{E}\Big)$$
$$\leq P\Big(\inf_{x,x',u'}\gamma_N^{1/2}h_x(x',u')<\kappa'\gamma_N^{1/4}+\bar{C}\sqrt{\log(2\gamma_N)\gamma_N^{1/2}}\Big|\mathscr{E}\Big)$$
$$=1\Big\{\inf_{x,x',u'}h_x(x',u')<\kappa'\gamma_N^{-1/4}+\bar{C}\sqrt{\log(2\gamma_N)}\gamma_N^{-1/4}\Big\}$$

which equals to zero for N,γ_N large enough, since $\inf_{x,x',u'}h_x(x',u')>0$.

Proof of Theorem A.1.4

Recall that $\mathscr{G}$ denotes a finite grid with $K/2$ elements. First, we bound

$$W(\beta^*)-W(\hat{\beta}^{ow})\leq 2\sup_{\beta\in[0,1]^p}\Big|W(\beta)-\hat{W}(\beta)\Big|.$$

By the mean value theorem, we can write for any $\beta^k\in\mathscr{G}$

$$W(\beta)=W(\beta^k)+V(\beta^k)^\top(\beta-\beta^k)+O\Big(||\beta^k-\beta||^2\Big),$$

Since we construct $\hat{W}(\beta)$ as in Equation (A.1.7), we can choose β^k closest to β. In such a case, by construction of the grid, $O\Big(||\beta^k - \beta||^2\Big) = O(1/K^{2/p})$ by construction of the grid. We can then write (using the fact that $p < \infty$)

$$\begin{aligned}
&\sup_{\beta\in[0,1]} \Big|W(\beta) - \hat{W}(\beta)\Big| \\
&\leq \sup_{\beta\in[0,1]^p, k\in\{1,\cdots,K\}} \Big|W(\beta^k) + V(\beta^k)^\top(\beta - \beta^k) - \bar{W}^k - \hat{V}_{(k,k+1)}^\top(\beta - \beta^k)\Big| + \mathscr{O}(1/K^{2/p}) \\
&\leq \sup_{k\in\{1,\cdots,K\}} \Big|W(\beta^k) - \bar{W}^k\Big| + ||V(\beta^k) - \hat{V}_{k,k+1}||_\infty \mathscr{O}(1) + \mathscr{O}(1/K^{2/p}).
\end{aligned}$$

Observe now that similarly to what discussed in Lemma A.3.6, where here the first order derivatives cancel out with a second order Taylor expansion, due to the opposite sign of $\pm\eta_n$ in each cluster,

$$2\mathbb{E}\Big[\bar{W}^k\Big] = \int y(x,\beta^k + \eta_n)dF_X(x) + \int y(x,\beta^k - \eta_n)dF_X(x) = 2\int y(x,\beta^k)dF_X(x) + O(\eta_n^2).$$

Using Lemma A.3.3, we can write for all $k \leq K$, with probability at least $1-\delta$,

$$\bar{W}^k = W(\beta^k) + O\Big(\sqrt{\gamma_N \log(pK\gamma_N/\delta)/n} + \eta_n^2\Big),$$

where we used the union bound over K, p in the expression. Similarly, from Lemma A.3.7, also using the union bound over K and p, with probability at least $1-\delta$,

$$||\hat{V}_{(k,k+1)} - V(\beta^k)||_\infty = O\Big(\sqrt{\gamma_N \log(Kp\gamma_N/\delta)/(n\eta_n^2)} + \eta_n\Big),$$

which concludes the proof as we choose $\delta = 1/n$, since $\eta_n = o(1)$, and p is finite.

Proofs for the Extensions

Proof of Theorem A.1.1

Observe that we can write

$$\mathbb{E}\Big[\bar{Y}_t^{(k)}\Big|p_t^{(k)}\Big] = \alpha_t + \tau_k + g\Big(q(\beta+\eta_n)+o_p(\eta_n),\beta+\eta_n\Big),$$
$$\mathbb{E}\Big[\bar{Y}_t^{(k)}\Big|p_t^{(k+1)}\Big] = \alpha_t + \tau_k + g\Big(q(\beta-\eta_n)+o_p(\eta_n),\beta-\eta_n\Big).$$

From a Taylor expansion in its first argument around $q(\beta+\eta_n)$, we obtain

$$g\Big(q(\beta+\eta_n)+o_p(\eta_n),\beta+\eta_n\Big) = g\Big(q(\beta+\eta_n),\beta+\eta_n\Big) + o_p(\eta_n)$$

and similarly once we subctract η_n. Therefore, we obtain

$$\mathbb{E}\Big[\bar{Y}_t^{(k)} - \bar{Y}_t^{(k+1)}\Big|p_t^{(k)},p_t^{(k+1)}\Big] = \tau_k - \tau_{k+1} + g\Big(q(\beta+\eta_n),\beta+\eta_n\Big) + o_p(\eta_n) - g\Big(q(\beta-\eta_n),\beta-\eta_n\Big).$$

We can now proceed with a Taylor expansion around of the functions $g(\cdot)$ around β to obtain (this follows similarly to Lemma A.3.6)

$$g\Big(q(\beta+\eta_n),\beta+\eta_n\Big) - g\Big(q(\beta-\eta_n),\beta-\eta_n\Big) = 2V_g(\beta)\eta_n + O(\eta_n^2).$$

In addition observe that since at the baseline β_0 is the same for both clusters,

$$\mathbb{E}[Y_0^{(k)} - Y_0^{(k+1)}|p_t^{(k)},p_t^{(k+1)}] = \tau_k - \tau_{k+1} + o_p(\eta_n).$$

The proof concludes from Lemma A.3.3 and the local dependence assumption in Assumption A.1.1.

Proof of Theorem A.2.3

In this subsection, we derive the theorem for the gradient descent method under Assumption A.2.1, for our extension where we relax global strong concavity. The derivation is split into

the following lemmas. First define the oracle descent as follows.

Definition A.3.5 (Oracle gradient descent)**.** We define for positive constants $\infty > \mu, \kappa > 0$, κ as defined in Lemma A.3.11, arbitrary $v \in (0,1)$

$$\beta_w^* = \begin{cases} P_{\mathscr{B}_1,\mathscr{B}_2}\left[\beta_{w-1}^* + \alpha_{w-1} V(\beta_{w-1}^*)\right] \text{ if } ||V(\beta_w^*)||_2 \geq \frac{\kappa}{\mu \check{T}^{1/2-v/2}} \\ \beta_{w-1}^* \text{ otherwise} \end{cases} \quad \beta_1^* = \beta_0, \qquad \text{(A.3.6)}$$

for $\alpha_w = \frac{J}{\check{T}^{1/2-v/2}||V(\beta_{w-1}^*)||}, J < 1$. □

Lemma A.3.11 (Adaptive gradient descent for quasi-concave functions and locally strong concave)**.** *Let $\mathscr{B}$ be compact. Define $G = \max\{\sup_{\beta\in\mathscr{B}} 2||\beta||^2, 1\}$. Let Assumption 1.3.1, 1.4.1, A.2.1 hold. Let κ be a positive finite constant, defined as in Equation* (A.3.7)*. Then for any $v \in (0,1)$, for $\check{T} \geq ((G+1)/J)^{1/v}$, the following holds:*

$$||\beta_{\check{T}}^* - \beta^*||^2 \leq \kappa \check{T}^{-1+v}.$$

Proof of Lemma A.3.11. To prove the statement, we use properties of gradient descent methods with gradient norm rescaling (Hazan et al., 2015), with modifications to the original arguments to explicitely obtain a rate T^{-1+v} for an arbitrary small v, and account for the formalization of local strong concavity based on the Hessian which we provide in our context.

Preliminaries

Clearly, if the algorithm terminates at w, under Assumption A.2.1 (B), this implies that

$$||\beta_w^* - \beta^*||_2^2 \leq \kappa \check{T}^{-1+v},$$

proving the claim. Therefore, assume that the algorithm did not terminate at time w. This implies that for any $\check{w} \geq 1$, $||\beta_{\check{w}}^* - \beta^*||_2^2 > \kappa \check{T}^{-1+v}$. Define $\varepsilon = \check{T}^{-1+v}$ and let ∇_w be the gradient evaluated at β_w^*. For every $\beta \in \mathscr{B}$, define $H(\beta)\Big|_{[\beta^*,\beta]}$ the Hessian evaluated at some point

$\tilde{\beta} \in [\beta^*, \beta]$, such that

$$W(\beta) = W(\beta^*) + \frac{1}{2}(\beta - \beta^*)^\top H(\beta)\Big|_{[\beta^*,\beta]}(\beta - \beta^*),$$

which always exists by the mean-value theorem and differentiability of the objective function. Define

$$\frac{1}{2}(\beta - \beta^*)^\top H(\beta)\Big|_{[\beta^*,\beta]}(\beta - \beta^*) = f(\beta) \leq 0,$$

where the inequality follows by definition of β^* (note that $f(\beta)$ also depends on $\tilde{\beta}$, whose dependence we implicitely suppressed).

Claim

We claim that

$$-|\lambda_{max}|||\beta - \beta^*||^2 \leq f(\beta) \leq -|\lambda_{\min}|||\beta - \beta^*||^2$$

for constants $\lambda_{\max} > \lambda_{\min} > 0$. The lower bound follows directly by Assumption 1.3.1, while the upper bound follows directly from Assumption A.2.1 (C), compactness of $\mathscr{B}$, and continuity of the Hessian. We provide details for the upper bound in the following paragraph.

Proof of the claim on the upper bound

We now use a contradiction argument. Suppose that the upper bound does not hold. Then since $\mathscr{B}$ is compact (and hence $||\beta - \beta^*||$ is bounded away from infinity for all $\beta \in \mathscr{B}$), and β^* is unique by (A, B) in Assumption A.2.1, there must exist a sequence $\beta_s \in \mathscr{B}, \beta_s \to \beta^*$ such that $f(\beta_s) \geq o(||\beta_s - \beta^*||^2)$. Recall that twice continuously differentiability of $W(\beta)$, we have that $H(\beta_s) \to H(\beta^*)$. As a result, we can find, for $s \geq S$, for S large enough, a point in the sequence such that (since p is finite)

$$2f(\beta_s) \leq (\beta_s - \beta^*)^\top H(\beta^*)(\beta_s - \beta^*) + \delta(s)||\beta_s - \beta^*||^2,$$

for $\delta(s) = |\tilde{\lambda}_{\max}(s)|$, where $\tilde{\lambda}_{\max}(s)$ is the maximum eigenvalue of $H(\beta)\Big|_{\beta\in[\beta_s,\beta^*]} - H(\beta^*)$. Note that such decomposition holds by symmetry of $H(\beta)\Big|_{\beta\in[\beta_s,\beta^*]} - H(\beta^*)$. Since $H(\beta^*)$ is negative definite, the above expression is bounded as follows

$$2f(\beta_s) \leq -(|\tilde{\lambda}_{min}| - \delta(s))||\beta_s - \beta^*||^2,$$

where $|\tilde{\lambda}_{min}| > 0$ is the minimum eigenvalue of $H(\beta^*)$ (in absolute value) bounded away from zero by Assumption A.2.1 (C). By continuity of the Hessian, $\delta(s) \to 0$, and we reach a contradiction. This result implies strong concavity *locally* at the optimum. Here $\lambda_{\max} > \lambda_{\min}$ since the lower bound holds for any finite and large enough $\lambda_{\max}$.

Cases

Define

$$\kappa = \frac{|\lambda_{max}|}{|\lambda_{min}|} \geq 1. \tag{A.3.7}$$

Observe now that if $||\beta_w^* - \beta^*||^2 \leq \varepsilon\kappa$, the claim trivially holds. Therefore, consider the case where

$$||\beta_w^* - \beta^*||^2 > \varepsilon\kappa.$$

Comparisons within the neighborhood

Take $\tilde{\beta} = \beta^* - \sqrt{\varepsilon}\frac{\nabla_w}{||\nabla_w||_2}$. Observe that

$$\begin{aligned} W(\tilde{\beta}) - W(\beta_w^*) &= \frac{1}{2}(\tilde{\beta} - \beta^*)^\top H(\tilde{\beta})\Big|_{[\beta^*,\tilde{\beta}]}(\tilde{\beta} - \beta^*) - \frac{1}{2}(\beta_w^* - \beta^*)^\top H(\beta_w^*)\Big|_{[\beta^*,\beta_w^*]}(\beta_w^* - \beta^*) \\ &\geq -|\lambda_{\max}|\varepsilon + |\lambda_{\min}|\varepsilon\kappa = 0. \end{aligned}$$

As a result, for all $\beta_w^* : ||\beta_w^* - \beta^*||^2 > \varepsilon\kappa$, using quasi-concavity

$$\nabla_w^\top(\tilde{\beta} - \beta_w^*) \geq 0 \Rightarrow \nabla_w^\top(\beta^* - \beta_w^*) \geq \sqrt{\varepsilon}||\nabla_w||_2 \tag{A.3.8}$$

Plugging in the above expression in the definition of β_w^*

By construction of the algorithm, we write

$$||\beta^* - \beta_{w+1}^*||^2 \leq ||\beta^* - \beta_w^*||^2 - 2\alpha_w J \nabla_w^\top (\beta^* - \beta_w^*) + J^2 \alpha_w^2 ||\nabla_w||^2.$$

By Equation (A.3.8), we can write

$$||\beta^* - \beta_{w+1}^*||^2 \leq ||\beta^* - \beta_w^*||^2 - 2J\alpha_w \sqrt{\varepsilon} ||\nabla_w||_2 + J^2 \alpha_w^2 ||\nabla_w||^2.$$

Plugging in the expression for α_w, and using the fact that $J \leq 1$, we have

$$||\beta^* - \beta_{w+1}^*||^2 \leq ||\beta^* - \beta_w^*||^2 - J\varepsilon.$$

Recursive argument

Recall that since the algorithm did not terminate, $||\beta^* - \beta_{\check{w}}^*||^2 > \varepsilon\kappa$, for all $\check{w} \leq w$. Using this argument recursively, we obtain

$$||\beta^* - \beta_{\check{T}}^*||^2 \leq ||\beta^* - \beta_0||^2 - J\sum_{s=1}^{\check{T}} \varepsilon = 2\max_{\beta \in \mathscr{B}} ||\beta||^2 - J\check{T}^v \leq G + 1 - J\check{T}^v.$$

Whenever $\check{T} > (G/J + 1/J)^{1/v}$, we have a contradiction. The proof completes. □

Lemma A.3.12. *Let Assumption 1.2.1, 1.2.2, 1.4.1, A.2.1 hold. Assume that*

$$\varepsilon_n \geq \sqrt{p}\Big[\bar{C}\sqrt{\gamma_N \frac{\log(\gamma_N \check{T} K/\delta)}{\eta_n^2 n}} + \eta_n\Big], \quad \frac{1}{4\mu \check{T}^{1/2 - v/2}} - \varepsilon_n \geq 0$$

for a finite constant $\bar{C} < 0$.

Then with probability at least $1-\delta$, for any $\delta \in (0,1)$ for any $w \leq \check{T}$,

$$\textit{either } (i) \left|\left|\check{\beta}_k^w - \beta_w^*\right|\right|_\infty = \mathscr{O}(P_w(\delta) + p\eta_n), \textit{ or } (ii) \left|\left|\check{\beta}_k^w - \beta^*\right|\right|_2^2 \leq \frac{p}{\check{T}^{1-v}}$$

where $P_1(\delta) = \text{err}(\delta)$ *and* $P_w(\delta) = \frac{2\sqrt{p}}{\nu_n} B_p \frac{1}{\check{T}^{1/2-\nu/2}} P_{w-1}(\delta) + P_{w-1}(\delta) + \frac{2\sqrt{p}}{\nu_n} \frac{1}{\check{T}^{1/2-\nu/2}} \text{err}(\delta)$*, for a finite constant* $B_p < \infty$*, and* $\text{err}(\delta) = \mathcal{O}\Big(\sqrt{\gamma_N \frac{\log(\gamma_N p \check{T} K/\delta)}{\eta_n^2 n}} + p\eta_n\Big)$*, with* $\nu_n = \frac{1}{\mu \check{T}^{1/2-\nu/2}} - 2\varepsilon_n$*.*

Proof of Lemma A.3.12. First, by Lemma 1.4.1, the estimated coefficients are exogenous. Hence, by invoking Lemma A.3.7 and the union bound, we can write for every k and t, $\delta \in (0,1)$,

$$\check{V}_{k,w}^{(j)} = V^{(j)}(\check{\beta}_{k+2}^w) + \mathcal{O}\Big(\sqrt{\gamma_N \frac{\log(\gamma_N K \check{T}/\delta)}{\eta_n^2 n}} + \eta_n\Big).$$

We now proceed by induction. We first prove the statement, assuming that the constraint is always attained. We then discuss the case of the constrained solution. Define

$$B = p \sup_{\beta} \Big|\Big| \frac{\partial^2 W(\beta)}{\partial \beta^2} \Big|\Big|_\infty.$$

Unconstrained case

Consider $w = 1$. Then since all clusters start from the same starting point β_0 recall that $(\beta_1^* = \beta_0)$, we can write with probability $1-\delta$, by the union bound over p (which hence enters in the $\log(p)$ component of err_n) and Lemma A.3.7

$$\Big|\Big| \check{V}_{k,1} - V(\beta_1^*) \Big|\Big|_\infty \leq \text{err}(\delta). \tag{A.3.9}$$

Consider now the case where the algorithm stops. This implies that it must be that $||\check{V}_{k,1}||_2 \leq \frac{1}{\mu \check{T}^{1/2-\nu/2}} - \varepsilon_n$. By Lemma A.3.7

$$||V(\beta_1^*)||_2 \leq ||\check{V}_{k,1}||_2 + \sqrt{p}\text{err}(\delta) \leq \frac{1}{\mu \check{T}^{1/2-\nu/2}} - \varepsilon_n + \sqrt{p}\text{err}(\delta) \leq \frac{1}{\mu \check{T}^{1/2-\nu/2}}. \tag{A.3.10}$$

since $\varepsilon_n \geq \sqrt{p}\text{err}(\delta)$. As a result, also the oracle algorithm stops at β_1^* by construction of ε_n. Suppose the algorithm does not stop. Then it must be that $||\check{V}_{k,1}|| \geq \frac{1}{\mu \check{T}^{1/2-v/2}} - \varepsilon_n$ and

$$||V_1(\beta_1^*)|| \geq \frac{1}{\mu \check{T}^{1/2-v/2}} - \varepsilon_n - \sqrt{p}\text{err}_1 \geq \frac{1}{\mu \check{T}^{1/2-v/2}} - 2\varepsilon_n := \nu_n > 0.$$

Observe now that

$$\begin{aligned} \Big|\Big|\frac{\check{V}_{k,1}}{||\check{V}_{k,1}||_2} - \frac{V(\beta_1^*)}{||V(\beta_1^*)||_2}\Big|\Big|_\infty &\leq \Big|\Big|\frac{\check{V}_{k,1} - V(\beta_1^*)}{||V(\beta_1^*)||_2}\Big|\Big|_\infty + \Big|\Big|\frac{\check{V}_{k,1}(||\check{V}_{k,1}||_2 - ||V(\beta_1^*)||_2)}{||V(\beta_1^*)||_2||\check{V}_{k,1}||_2}\Big|\Big|_\infty \\ &\leq \Big|\Big|\frac{\check{V}_{k,1} - V(\beta_1^*)}{||V(\beta_1^*)||_2}\Big|\Big|_\infty + \sqrt{p}\Big|\Big|\frac{\check{V}_{k,1} - V(\beta_1^*)}{||V(\beta_1^*)||_2}\Big|\Big|_\infty. \end{aligned} \tag{A.3.11}$$

The last inequality follows from the reverse triangular inequalities and standard properties of the norms. Then with probability at least $1-\delta$, for any $\delta \in (0,1)$

$$(\text{A.3.11}) \leq \frac{1}{\nu_n} \times 2\sqrt{p}\text{err}(\delta).$$

completing the claim for $w = 1$. Consider now a general w. Define the error until time $w-1$ as P_{w-1}.Then for every $j \in \{1, \cdots, p\}$, by Assumption 1.3.1, we have with probability at least $1 - w\delta$ (using the union bound),

$$\begin{aligned} &\check{V}_{k,w}^{(j)} = V^{(j)}(\check{\beta}_{k+2}^w) + \text{err}(\delta) = V^{(j)}(\beta_w^* + P_w(\delta)) + \text{err}(\delta) \\ &\Rightarrow \Big|\Big|\check{V}_{k,w} - V(\beta_w^*)\Big|\Big|_\infty \leq BP_w(\delta) + \text{err}(\delta), \end{aligned}$$

where the above inequality follows by the mean-value theorem and Assumption 1.3.1. Suppose now that $||\check{V}_{k,w}||_2 \leq \frac{1}{\mu \check{T}^{1/2-v/2}} - \varepsilon_n$. Then for the same argument as in Equation (A.3.10), we have

$$||V(\check{\beta}_k^w)||_2 \leq \frac{1}{\mu \check{T}^{1/2-v/2}}.$$

Under Assumption A.2.1 (B) this implies that

$$||\check{\beta}_k^w - \beta^*||_2^2 \leq \frac{1}{\check{T}^{1-v}},$$

which proves the statement. Suppose instead that the algorithm does not stop. Then we can write by the induction argument

$$\begin{aligned}
&\Big|\Big|\check{\beta}_k^w + \frac{1}{\check{T}^{1/2-v/2}} \frac{\check{V}_{k,w}}{||\check{V}_{k,w}||_2} - \beta_w^* - \frac{1}{\check{T}^{1/2-v/2}} \frac{V(\beta_w^*)}{||V(\beta_w^*)||_2}\Big|\Big|_\infty \\
&\leq P_w(\delta) + \frac{1}{\check{T}^{1/2-v/2}} \underbrace{\Big|\Big| \frac{\check{V}_{k,w}}{||\check{V}_{k,w}||_2} - \frac{V(\beta_w^*)}{||V(\beta_w^*)||_2}\Big|\Big|_\infty}_{(B)}.
\end{aligned}$$

Using the same argument in Equation (A.3.11), we have with probability at least $1-\delta$,

$$(B) \leq \frac{2\sqrt{p}}{\nu_n}\Big[\text{err}(\delta) + BP_w(\delta)\Big],$$

which completes the proof for the unconstrained case. The $\check{T}$ component in the error expression follows from the union bound across all $\check{T}$ events. The constrained case follows similarly and omitted for brevity. □

Lemma A.3.13. *Let the conditions in Lemma A.3.12 hold. Then with probability at least* $1-\delta$*, for any* $k \in \{1, \cdots, K\}$*, for any* $v \in (0,1), \delta \in (0,1)$*,* $\check{T} \geq \zeta^{1/v}$*, for* $\zeta < \infty$ *being a finite constant*

$$||\beta^* - \check{\beta}_k^{\check{T}}||_2^2 \leq \frac{\kappa}{\check{T}^{1-v}} + \check{T}e^{B_p\sqrt{p}\check{T}} \times \mathscr{O}\Big(\gamma_N \frac{\log(p\gamma_N \check{T}K/\delta)}{\eta_n^2 n} + p^2\eta_n^2\Big),$$

with $\kappa, B_p < \infty$ *being constants independent on* $(n, \check{T})$ *and* ε_n *as defined in Lemma A.3.12.*

Proof. We invoke Lemma A.3.12. Observe that we only have to check that the result holds for (i) in Lemma A.3.12, since otherwise the claim trivially holds. Using the triangular inequality,

we can write

$$||\beta^* - \check{\beta}_k^{\check{T}}||_2^2 \leq ||\beta^* - \beta_{\check{T}}^*||_2^2 + ||\check{\beta}_k^{\check{T}} - \beta_{\check{T}}^*||_2^2.$$

The first component on the right-hand side is bounded by Lemma A.3.11, with $\check{T} \geq \zeta^{1/v}$, ζ being a constant defined in Lemma A.3.11.

Using Lemma A.3.12, we bound with probability at least $1-\delta$, the second component as follows

$$||\check{\beta}_k^{\check{T}} - \beta_{\check{T}}^*||_2^2 \leq p||\check{\beta}_k^{\check{T}} - \beta_{\check{T}}^*||_\infty^2 = p \times \mathscr{O}(P_{\check{T}}^2(\delta)).$$

We conclude the proof by explicitely defining recursively, for all $1 < w \leq \check{T}$,

$$P_w = (1 + \frac{2B_p\sqrt{p}}{\nu_n \check{T}^{1/2-v/2}})P_{w-1} + \frac{1}{\check{T}^{1/2-v/2}}\text{err}_n(\delta), \quad P_1 = \text{err}_n(\delta).$$

where $\text{err}_n(\delta) = \frac{2\sqrt{p}}{\nu_n}\mathscr{O}(\sqrt{\gamma_N \frac{\log(pTK/\delta)}{\eta_n^2 n}} + p\eta_n)$, and $B < \infty$ denotes a finite constant. Using a recursive argument, we obtain

$$P_w = \text{err}_n(\delta)\sum_{s=1}^{w}\alpha_s \prod_{j=s}^{w}(\frac{2B_p\sqrt{p}}{\nu_n \check{T}^{1/2-v/2}} + 1).$$

Recall now that $\nu_n \geq \frac{1}{2\mu\check{T}^{1/2-v/2}}$ as in Lemma A.3.12. As a result we can bound the above expression as

$$\sum_{s=1}^{w}\alpha_s \prod_{j=s}^{w}(\frac{2B_p\sqrt{p}}{\nu_n \check{T}^{1/2-v/2}} + 1) \leq \sum_{s=1}^{w}\alpha_s \prod_{j=s}^{w}(\frac{8\mu^2\check{T}^{1/2-v/2}B\sqrt{p}}{\check{T}^{1/2-v/2}} + 1) \leq \sum_{s=1}^{w}\alpha_s \exp\Big(\sum_{j=s}^{w} 8\mu^2 B_p\sqrt{p}\Big).$$

Now we have

$$\exp\Big(\sum_{j=s}^{w} 8\mu^2 B_p\sqrt{p}\Big) \leq \exp\Big(8\mu^2(w-s)B_p\sqrt{p}\Big).$$

We now write

$$P_w(\delta) \leq \text{err}_n(\delta) \sum_{s=1}^{w} \alpha_s \exp\left(8\mu^2(w-s)B_p\sqrt{p}\right) \leq \text{err}_n(\delta)\check{T}^{1/2+v} \exp\left(8\mu^2\check{T}B_p\sqrt{p}\right),$$

where we replaced w with $\check{T}$. The proof completes.

□

Corollary 11. *Theorem A.2.3 holds.*

Proof. Consider Lemma A.3.12 where we choose $\delta = 1/n$. Observe that we choose $\varepsilon_n \leq \frac{1}{4\mu\check{T}^{1/2-v/2}}$, which is attained by the conditions in Lemma A.3.12 as long as n is small enough such that

$$\sqrt{p}\Big[\bar{C}\sqrt{\log(n)\gamma_N \frac{\log(p\gamma_N\check{T}K)}{\eta_n^2 n}} + \eta_n\Big] \leq \frac{1}{4\mu\check{T}^{1/2-v/2}}$$

attained under the assumptions stated in Lemma A.3.12. As a result, we have $\nu_n = \frac{1}{4\mu\check{T}^{1/2-v/2}}$. By Lemma A.3.13 for all k, with probability at least $1 - 1/n$,

$$||\check{\beta}_k^{\check{T}} - \beta^*||^2 \lesssim \frac{p}{\check{T}^{1-v}}.$$

Also, we have

$$||\beta^* - \frac{1}{K}\sum_k \check{\beta}_k^{\check{T}}||_2^2 \leq \frac{1}{K}\sum_k ||\check{\beta}_k^{\check{T}} - \beta^*||^2.$$

The proof concludes by Theorem A.3.13 and Assumption 1.3.1, after observing that

$$W(\beta^*) - W(\hat{\beta}^*) \lesssim ||\beta^* - \hat{\beta}^*||_2^2.$$

□

Proof of Theorem A.1.3

First, we bound

$$\sup_{\theta\in\Theta}\widetilde{W}(\theta)-W(\widehat{\theta})\leq 2\sum_t q^t\times\underbrace{\sup_{(\beta_1,\beta_2)\in[0,1]^2}\left|\widehat{\Gamma}(\beta_2,\beta_1)-\Gamma(\beta_2,\beta_1)\right|}_{(A)}\lesssim(A),$$

since $\sum_t q^t<\infty$. To bound (A) observe first that each element in the grid $\mathscr{G}$ has a distance of order $1/\sqrt{K}$, since the grid has two dimensions and $K/3$ components. As a result for any element (β_2,β_1), we can write

$$\Gamma(\beta_2,\beta_1)=\underbrace{\Gamma(\beta_2^r,\beta_1^r)}_{(B)}+\underbrace{\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_1^r}(\beta_1-\beta_1^r)+\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_2^r}(\beta_2-\beta_2^r)}_{(C)}+\underbrace{\mathscr{O}\Big(||\beta_1-\beta_1^r||^2+||\beta_2-\beta_2^r||^2\Big)}_{(D)}$$

where $\beta^r\in\mathscr{G}$ is some value in the grid such that (B) is of order $1/K$. We can now write

$$\begin{aligned}(A)\leq&\underbrace{\sup_{(\beta_1^r,\beta_2^r)\in\mathscr{G},||\beta_1^r-\beta_1||^2+||\beta_2^r-\beta_2||^2\lesssim 1/K}\left|\widetilde{\Gamma}(\beta_2^r,\beta_1^r)-\Gamma(\beta_2^r,\beta_1^r)\right|}_{(i)}\\&+\underbrace{\sup_{(\beta_1^r,\beta_2^r)\in\mathscr{G}}\left|\widehat{g}_2(\beta_2^r,\beta_1^r)-\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_2^r}\right|\Big(|\beta_2-\beta_2^r|+|\beta_1-\beta_1^r|\Big)}_{(ii)}\\&+\underbrace{\sup_{(\beta_1^r,\beta_2^r)\in\mathscr{G}}\left|\widehat{g}_1(\beta_2^r,\beta_1^r)-\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_1^r}\right|\Big(|\beta_2-\beta_2^r|+|\beta_1-\beta_1^r|\Big)}_{(iii)}+\mathscr{O}\Big(|\beta_2-\beta_2^r|^2+|\beta_1-\beta_1^r|^2\Big).\end{aligned}$$

We now study each component separately. We start from (i). We observe that under Assumption A.1.3, by doing a Taylor expansion around (β_1^r,β_2^r), it is easy to observe that we can write

$$\mathbb{E}[\bar{Y}_{t+1}^{(k)}]=\Gamma(\beta_2^r,\beta_1^r)+\mathscr{O}(\eta_n).$$

Therefore by Lemma A.3.3, and the union bound over K many elements in $\mathscr{G}$ as $\gamma_N\log(\gamma_N K)/n\to 0$, $(i)\to 0$. Consider now (ii). We observe that since $\mathscr{B}$ is compact, we have $\Big(|\beta_2-\beta_2^r|+|\beta_1-\beta_1^r|\Big)=\mathscr{O}(1)$. In addition, similarly to what discussed in Lemma A.3.7, it follows that with

probability at least $1-\delta$,

$$\left|\widehat{g}_1(\beta_2^r,\beta_1^r)-\frac{\partial\Gamma(\beta_2^r,\beta_1^r)}{\partial\beta_1^r}\right|=\mathscr{O}\Big(\sqrt{\frac{\gamma_N\log(\gamma_N/\delta)}{\eta_n^2 n}}+\eta_n\Big).$$

Hence, by the union bound as $\frac{\gamma_N\log(\gamma_N K)}{\eta_n^2 n}=o(1)$ $(ii)=o_p(1)$ and similarly (iii). The proof concludes after observing that $|\beta_1^r-\beta_1|^2+|\beta_2^r-\beta_2|^2\lesssim 1/K$ by construction of the grid.

Proof of Theorem A.2.1

Let $\tilde{K}=K/2l$. Take

$$t_z^j=\frac{\frac{1}{\sqrt{z}}\sum_{i=1}^z X_i^j}{\sqrt{(z-1)^{-1}\sum_{i=1}^z(X_i^j-\bar{X}^j)^2}},\quad X_i^j\sim\mathscr{N}(0,\sigma_i^j).$$

Recall that by Theorem 1 in Ibragimov and Müller (2010), we have that for $\alpha\leq 0.08$

$$\sup_{\sigma_1,\cdots,\sigma_q}P(|t_z|\geq \mathrm{cv}_\alpha)=P(|T_{q-1}|\geq \mathrm{cv}_\alpha),$$

where cv_α is the critical value of a t-test with level α, and T_{z-1} is a t-student random variable with $z-1$ degrees of freedom. The equality is attained under homoskedastic variances (Ibragimov and Müller, 2010). We now write

$$P\Big(\mathscr{T}_n\geq q|H_0\Big)=P\Big(\max_{j\in\{1,\cdots,l\}}|Q_{j,n}|\geq q|H_0\Big)=1-P\Big(|Q_{j,n}|\leq q\forall j|H_0\Big)=1-\prod_{j=1}^l P\Big(|Q_{j,n}|\leq q|H_0\Big),$$

where the last equality follows by between cluster independence. Observe now that by Theorem 1.3.2 and the fact that the rate of convergence is the same for all clusters (Assumption 1.3.2) [14], for all j, for some $(\sigma_1,\cdots,\sigma_z)$, $z=\tilde{K}$,

$$\sup_q\left|P\Big(|Q_{j,n}|\leq q|H_0\Big)-P\Big(|t_{\tilde{K}}^j|\leq q\Big)\right|=o(1).$$

[14] Here we use continuity of the Gaussian distribution, and the fact that l is finite.

As a result, we can write

$$\sup_{\sigma_1,\cdots,\sigma_K} \lim_{n\to\infty} 1 - \prod_{j=1}^{l} P\Big(|Q_{j,n}| \leq q | H_0\Big) = 1 - \prod_{j=1}^{l} \inf_{\sigma_1^j,\cdots,\sigma_{\tilde{K}}^j} P\Big(|t_{\tilde{K}}^j| \leq q\Big),$$

where we used the fact that we use different pairs of (independent) clusters for each entry j. Using the result in Bakirov and Szekely (2006), we have

$$\inf_{\sigma_1^j,\cdots,\sigma_{\tilde{K}}^j} P\Big(|t_{\tilde{K}}^j| \leq q\Big) = P\Big(|T_{\tilde{K}-1}| \leq q | H_0\Big).$$

Therefore,

$$1 - \prod_{j=1}^{l} \inf_{\sigma_1^j,\cdots,\sigma_{\tilde{K}}^j} P\Big(|t_{\tilde{K}}^j| \leq q | H_0\Big) = 1 - P^l\Big(|T_{\tilde{K}-1}| \leq q\Big).$$

Setting the expression equal to α, we obtain

$$1 - P^l\Big(|T_{\tilde{K}-1}| \leq q\Big) = \alpha \Rightarrow P\Big(|T_{\tilde{K}-1}| \geq q\Big) = 1 - (1-\alpha)^{1/l}.$$

The proof completes after solving for q.

Proof of Theorem A.2.2

By Lemma 1.2.1, we can write

$$\mathbb{E}[\hat{\Delta}_k(\beta)] = m(1,1,\beta) - m(0,1,\beta) + O(\eta_n^2).$$

Following the same strategy as in the proof of Theorem 1.3.4, it is easy to show that

$$\mathbb{E}[\hat{S}(0,\beta)] = \frac{\partial m(0,1,\beta)}{\partial \beta} + \frac{1}{2}\Big[\alpha_{t,k} - \alpha_{t-1,k} - \alpha_{t,k+1} + \alpha_{t-1,k+1}\Big] + \mathscr{O}(\eta_n).$$

Similarly, we can write

$$\mathbb{E}[\hat{S}(1,\beta)] = \frac{\partial m(1,1,\beta)}{\partial \beta} + \frac{1}{2}\Big[\alpha_{t,k} - \alpha_{t-1,k} - \alpha_{t,k+1} + \alpha_{t-1,k+1}\Big] + \mathcal{O}(\eta_n).$$

The proof completes after noticing that $\frac{\partial m(1,1,\beta)}{\partial \beta} = 0$.

Proof of Theorem A.2.4

The proof mimics the proof of Theorem 1.4.3.

Consider Lemma A.3.10 where we choose $\delta = 1/n$. Note that we can directly apply Lemma A.3.10 also to the gradient estimated with Algorithm A.5.4, since, by the circular-cross fitting argument, each parameter $\check{\beta}_k^w$ is estimated using sequentially pairs of different clusters as in Algorithm A.5.4. The rest of the proof follows verbatim from the one of Theorem 1.4.3 and omitted for brevity.

Proof of Proposition A.1.5

By concavity, we can write

$$W(\beta^*) - \frac{1}{K}\sum_{k=1}^{K} W(\beta^k) \geq W(\beta^*) - W(\frac{1}{K}\sum_{k=1}^{K}\beta^k) = W(\beta^*) - W(0.5),$$

which completes the proof, for a suitable choice of $W(\beta^*) - W(0.5)$ (e.g., a quadratic function with $\beta^* = 0.3$).

Proof of the Corollaries

In this subsection we provide proofs to the corollary which do not directly follow from the corresponding theorem.

Proof of Corollary 2

The result directly follows from Theorem A.2.1, here applied to $l = 1$. The reader may refer to the proof of Theorem A.2.1 for details.

Proof of Corollary 3

The corollary follows from Lemma A.3.3 and the triangular inequality. Note that the rate is Kn since, after pooling observations from clusters, we can equivalently interpret $\bar{\Delta}_n$ as the estimator from two clusters, each with $K/2 \times n$ observations.

Proof of Corollary 4

The corollary follows from a second-order Taylor expansion, using the assumption that the Hessian is uniformly bounded.

Proof of Corollary 10

From Theorem 1.3.2, we know that

$$\frac{(\hat{V}_{k,k+1} - V(\beta(\theta_k), \theta_k))}{\sqrt{\text{Var}(\hat{V}_{k,k+1})}} \rightarrow_d \mathcal{N}(0,1),$$

since $\theta_k = \theta_{k+1}$ by construction of the pairs (assumption in the corollary). Therefore, we can directly apply Theorem A.2.1, here applied to $l = 1$, since, under H_0, each estimated marginal effect (for each type) is centered around zero, and, the theorem does not require that clusters have the same variance. The reader may refer to the proof of Theorem A.2.1 for details.

Numerical Studies: Additional Results

One-wave Experiment

In Figure A.4.3 we report the power plot for $\rho = 6$. In Figure A.4.4 we report the welfare gain from increasing β by 5% upon rejection of H_0 for $\rho = 6$. In Figure A.4.7 we report

comparisons for different values of η_n.

Multiple-wave Experiment

In Table A.4.1 the comparison with competitors for $\rho = 6$. Results are robust as in the main text. In Figure A.4.5 we report a comparison among different learning rates, which are the one which rescales by $1/t$, the one that rescales by $1/\sqrt{T}$ and the one that rescales by $1/\sqrt{t}$. The best performing learning rate rescales the step size by a factor of order $1/\sqrt{t}$.

Calibrated Experiment with Covariates for Cash Transfers

In this subsection, we turn to a calibrated experiment where we also control for covariates, as discussed in Section 1.2.3. We use data from Alatas et al. (2012, 2016). We estimate a function heterogenous in the distance of the household's village from the district's center. We use information from approximately four hundred observations, whose eighty percent or more neighbors are observed. We let $X_i \in \{0,1\}, X_i = 1$ if the household is far from the district's center than the median household, and estimate

$$Y_i|X_i = x \quad = \quad \phi_0 + \tilde{X}_i\tau + D_i\phi_{1,x} + \frac{\sum_{j\neq i} A_{j,i}D_j}{\max\{\sum_{j\neq i} A_{j,i}, 1\}}\phi_{2,x} + \Big(\frac{\sum_{j\neq i} A_{j,i}D_j}{\max\{\sum_{j\neq i} A_{j,i}, 1\}}\Big)^2\phi_{3,x} + \eta_i, \tag{A.4.1}$$

where η_i are unobservables centered on zero conditional on $X_i = x$, and $\tilde{X}_i$ denotes controls which also include X_i.[15] Using the estimated parameter, we can then calibrate the simulations as follows. We let $\eta_{i,t}, \sim \mathcal{N}(0,\sigma^2)$, where σ^2 is the residual variance from the regression. We then generate the network and the covariate as follows:

$$A_{i,j} = 1\Big\{||U_i - U_j||_1 \leq 2\rho/\sqrt{N}\Big\}, \quad U_i \sim_{i.i.d.} \mathcal{N}(0, I_2), \quad X_i = 1\{U_i^{(1)} > 0\}.$$

[15] We also control for the education level, village-level treatments, i.e., how individuals have been targeted in a village (i.e., via a proxy variable for income, a community-based method, or a hybrid), the size of the village, the consumption level, the ranking of the individual poverty level, the gender, marital status, household size, the quality of the roof and top (which are indicators of poverty).

Here, $U_i^{(1)}$ is continuous and captures a measure of distances. Individuals are more likely to be friends if they have similar distances from the center, and X_i is equal to one if an individual is far from the district's center from the median household. We fix $\rho = 1.5$ to guarantee that the objective's function optimum is approximately equal to the optimum observed from the data (in calibration, the optimum is $\beta \approx 0.26$, while $\beta^* \approx 0.29$ on the data). We then generate data

$$Y_{i,t}|X_i = x \quad = \quad D_i\hat{\phi}_{1,x} + \frac{\sum_{j\neq i} A_{j,i}D_j}{\max\{\sum_{j\neq i} A_{j,i}, 1\}}\hat{\phi}_{2,x} + \Big(\frac{\sum_{j\neq i} A_{j,i}D_j}{\max\{\sum_{j\neq i} A_{j,i}, 1\}}\Big)^2\hat{\phi}_{3,x} + \eta_{i,t}. \quad \text{(A.4.2)}$$

where we removed covariates that did not interact with the treatment rule (i.e., do not affect welfare computations). Our policy function is $\pi(x;\beta) = x\beta + (1-x)(1-\beta)$ where β is the probability of treatment for individuals farer from the center. Here, we implicitely imposed a budget constraint $\beta P(X_i = 1) + (1-\beta)P(X_i = 0) = 1/2$, where, by construction $P(X_i = 1) = 1/2$.

We collect results for the one-wave experiment in Figure A.4.6, A.4.8 (left-panel), where we report power and the relative improvement from improving by 5% the treatment probability for people in remote areas as discussed in the main text. Welfare improvements (and power) are increasing in the cluster size and the number of clusters. However, such improvements are negligible as we increase clusters from twenty to forty, suggesting that twenty clusters are sufficient to achieve the largest welfare effects.[16] In the right-hand side panel of Figure A.4.8 we report the out-of-sample regret. The regret is generally decreasing in the number of iterations, especially as the regret is further away from zero. As the regret gets almost zero (0.06%), the regret oscillates around zero as the number of iterations increases due to sampling variation. This behavior is suggestive that for some applications, few iterations (in this case, ten) are sufficient to reach the optimum, up to a small error. In Table A.4.2, we observe perfect coverage for $n = 600$, and under-coverage by no more than five percentage points in the remaining cases.

[16]The order of magnitude of the welfare gain is smaller compared to simulations with the unconditional probability since, here, we always treat exactly half of the population. As a result, welfare oscillates between 0.24 and 0.29 only (as opposed to zero to one as in the unconditional case), as shown in Figure 1.2.

Calibrated Simulations to M-Turk Experiment: Additional Details

In this subsection, we provide details for estimation for the experiment in Section 1.5.3.

In the treatment arm there were multiple answers questions. The first question asked is Which of these events caused more deaths of Covid in the US? (more answers allowed), giving four options (World War I and II, 50 times more than 9/11, US Civil war); What is the percentage of people in the US who had Covid within the last year? (approximately); The number of people infected from Covid in the last year is comparable to ... (giving three options). Each correct answer rewarded a small bonus and was displayed right after the participant submitted her answer to the three questions (before the end of the survey). In addition, at the end of the survey, participants were asked again one of the three questions, whose correct answer rewards a bonus. Participants were made aware of the bonuses and the survey's structure. The scope of the treatment was to increase awareness of the severity of the disease by asking questions and showing the correct answers to facilitate information transmission. At the end of the survey, both controls and treated units were asked when they would have done the vaccine. Our outcome of interest is binary and equals whether individuals would have done the vaccine either as soon as possible or during the spring. We estimate the model with 1035 participants. We estimate treatment effects by running a simple linear regression, where the treatment dummy interacts with the dummy, indicating whether the individual classifies herself as liberal, conservative, or "prefer not to say". We consider a model and policy function as in Section A.4.3, where X, in this case, denotes whether an individual is liberal or conservative (drawn with the same DGP as in Section A.4.3 for simplicity), and $\rho = 2$ as in the main text. We calibrate $\hat{\phi}_{1,x}$ to the estimated direct treatment effect and fix the percentage of treatment units to fifty percent. A challenge here is that we do not know spillovers ϕ_{2x}, ϕ_{3x}. Therefore, we choose $\phi_{3,x} = r\phi_{2,x}$, where r is estimated from data on information diffusion from Cai et al. (2015) for simplicity.[17] We choose $\phi_{2,x} + r\phi_{2,x} = \max\{\alpha\phi_{1,x}, 0\}$, i.e., total spillovers equal direct effect $\phi_{1,x}$ times a constant

[17] Other choices are possible but omitted for brevity.

$\alpha \in \{0.1, 0.2, 0.3, 0.4\}$ if these are positive, and zero otherwise.

Additional Figures and Tables

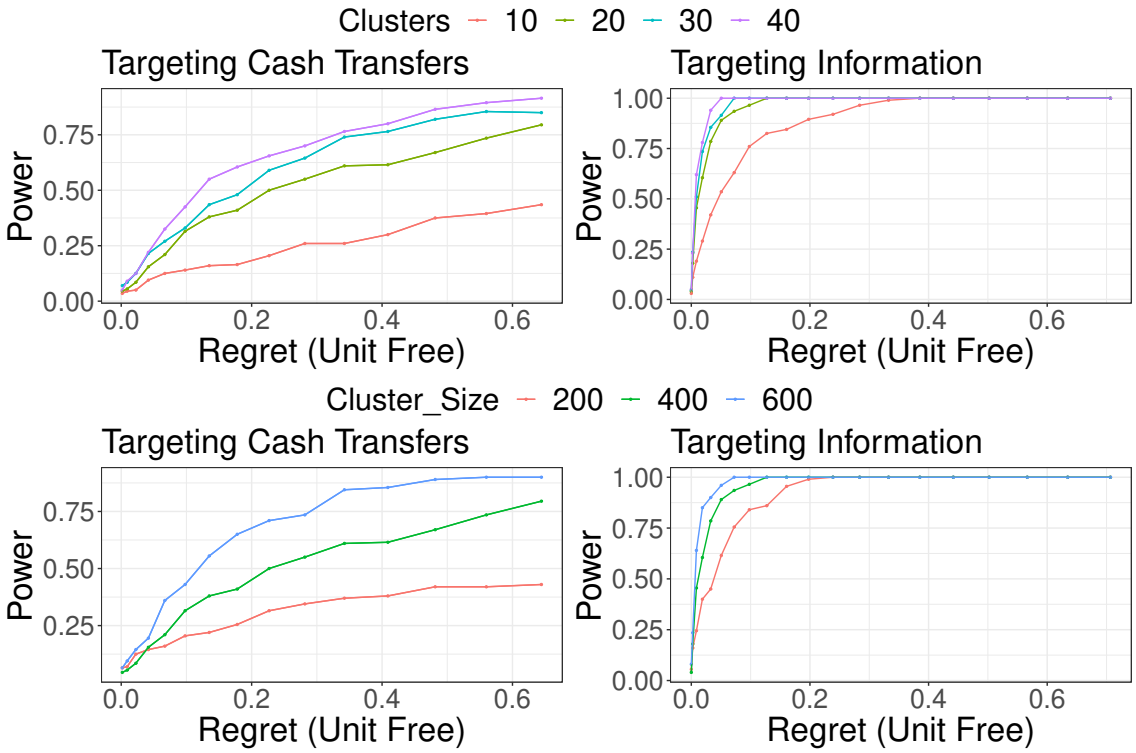


Figure A.4.1. Simulations: power in one-wave experiment in Section 1.5. 200 replications, $\rho = 2$. The panels at the top fix $n = 400$ and varies K. The panels at the bottom fix $K = 20$ and vary n.

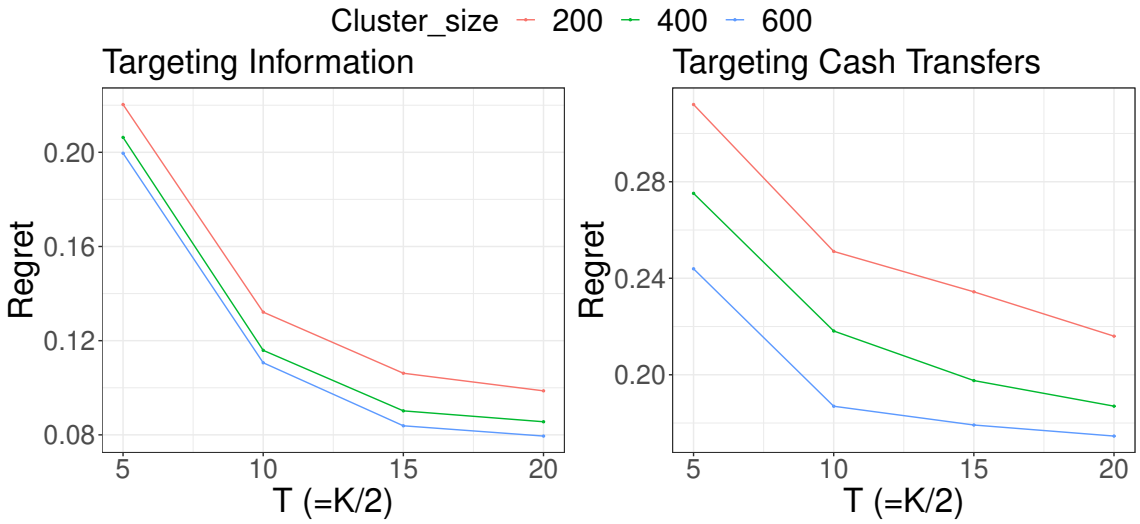


Figure A.4.2. Simulations: in-sample regret in Section 1.5. 200 replications. $\rho = 2$.

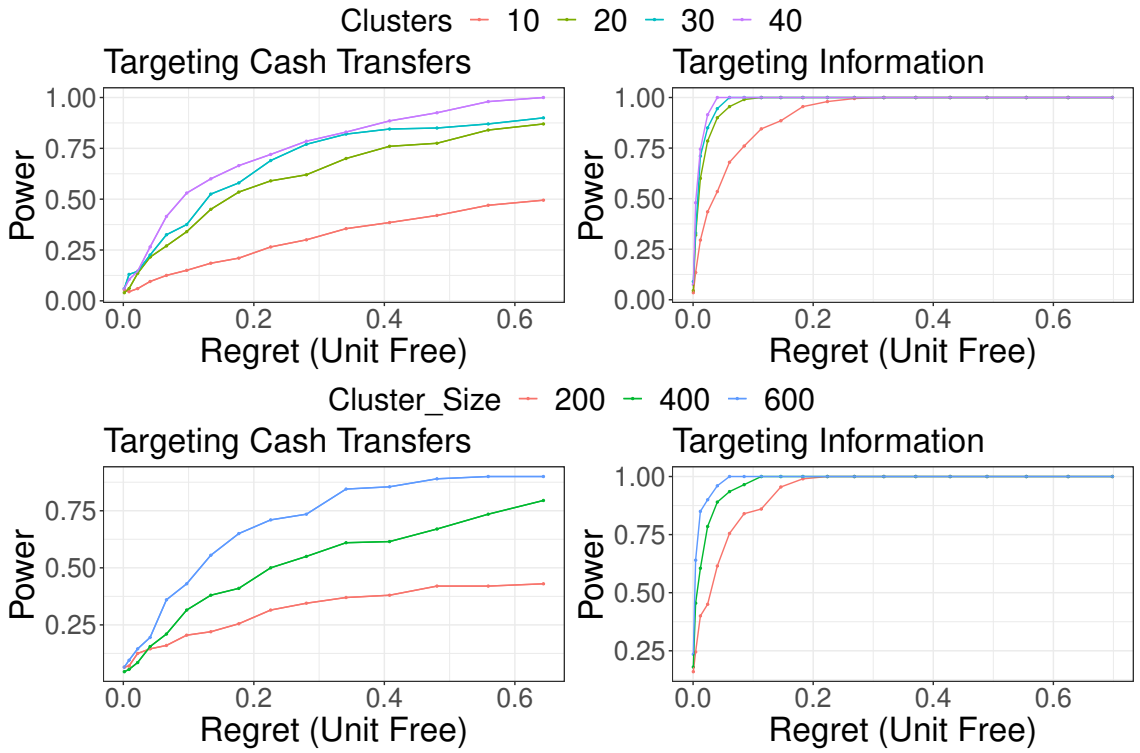


Figure A.4.3. Simulations: power plot with a dense network in Section 1.5, $\rho = 6$. The panels at the top fix $n = 400$ and varies K. The panels at the bottom fix $K = 20$ and vary n.

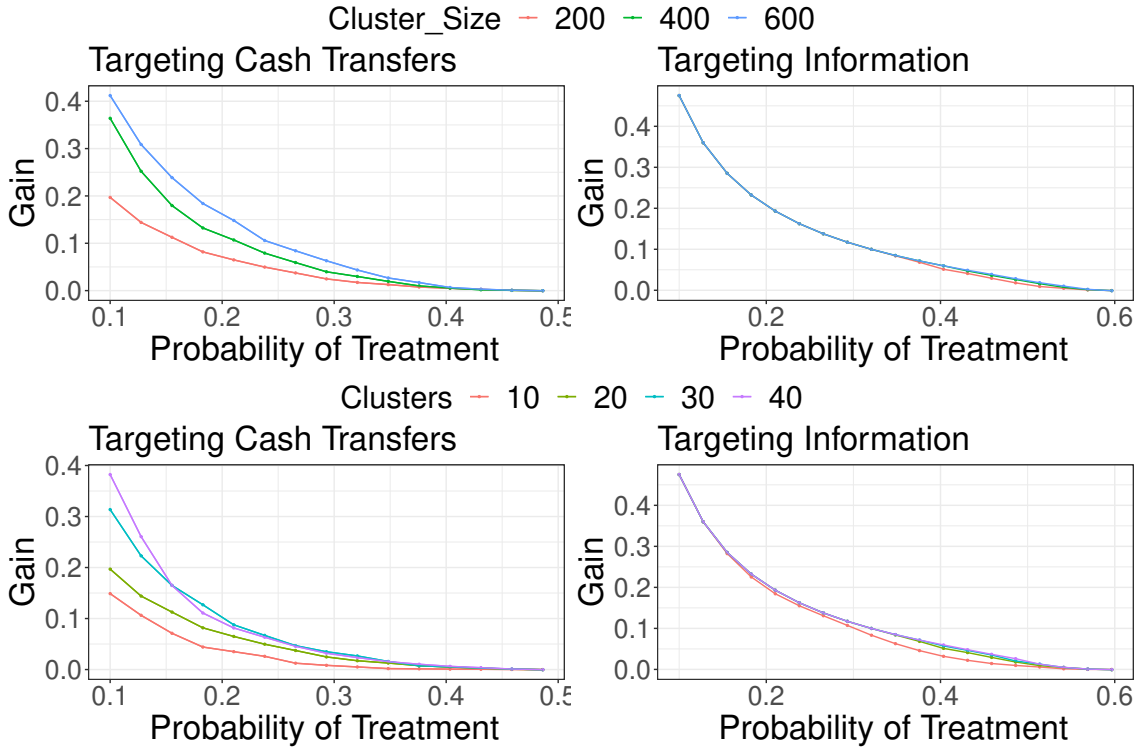


Figure A.4.4. Simulations: welfare analysis with a dense network. One-wave experiment in Section 1.5. $\rho = 6$. Expected percentage increase in welfare from increasing the probability of treatment β by 5% upon rejection of H_0. Here, the x-axis reports $\beta \in [0.1, \cdots, \beta^* - 0.05]$. The panels at the top fix $n = 400$ and varies the number of clusters. The panels at the bottom fix $K = 20$ and vary n.

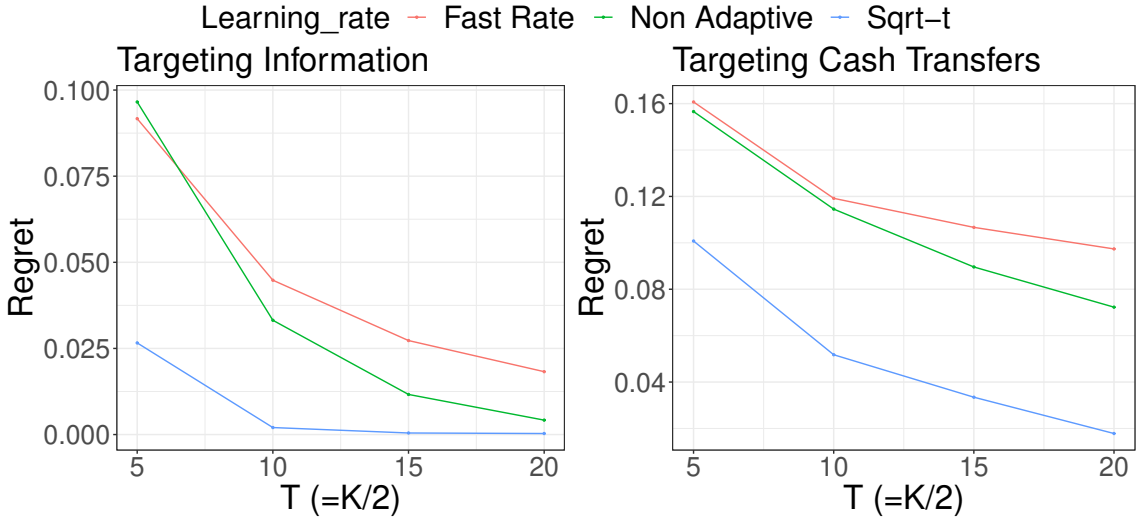


Figure A.4.5. Simulations: comparisons between different learning rates with experiment as in Section 1.5. 200 replications, $\rho = 2, n = 600, K = 2T$. Fast rate is of order $1/t$; non-adaptive dependsis of order $1/\sqrt{T}$; Sqrt-t is of order $1/\sqrt{t}$.

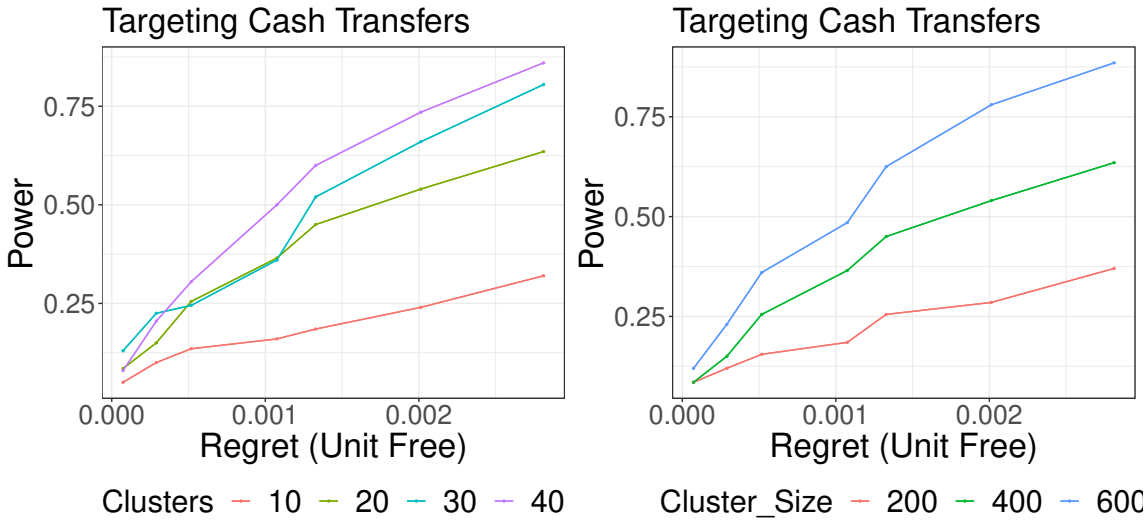


Figure A.4.6. Simulations: Single-wave experiment in Section A.4.3. Power, 200 replications.

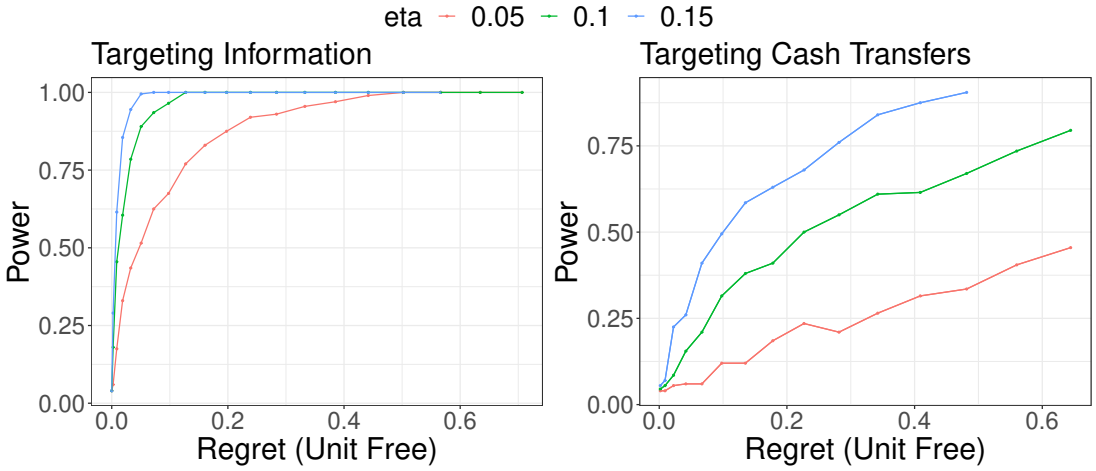


Figure A.4.7. Simulations: comparisons between different perturbations. One wave experiment calibrated to Alatas et al. (2012) and Cai et al. (2015). The plot reports power for different values of η_n varies, with $K = 200, n = 400$, with 200 replications.

Table A.4.1. Simulations: welfare with a dense network. Multiple-wave experiment in Section 1.5, $\rho = 6$. The panel at the top reports the out-of-sample regret and the one at the bottom the worst case in-sample regret across clusters.

	Information				Cash Transfer			
$T =$	5	10	15	20	5	10	15	20
$n = 200$	-0.023	0.012	0.039	0.029	0.405	0.475	0.542	0.358
$n = 400$	0.001	0.020	0.019	0.029	0.546	0.483	0.602	0.548
$n = 600$	0.0004	0.030	0.022	0.020	0.571	0.481	0.614	0.643
$n = 200$	1.238	1.488	1.546	1.516	0.347	0.415	0.429	0.416
$n = 400$	1.494	1.690	1.704	1.624	0.482	0.576	0.550	0.498
$n = 600$	1.579	1.791	1.809	1.689	0.606	0.689	0.664	0.586

Table A.4.2. Simulations: coverage with covariates. Single-wave experiment in Section A.4.3, 200 replications, size 5%.

$K =$	10	20	30	40
$n = 200$	0.940	0.950	0.895	0.900
$n = 400$	0.970	0.940	0.905	0.935
$n = 600$	0.950	0.970	0.950	0.940

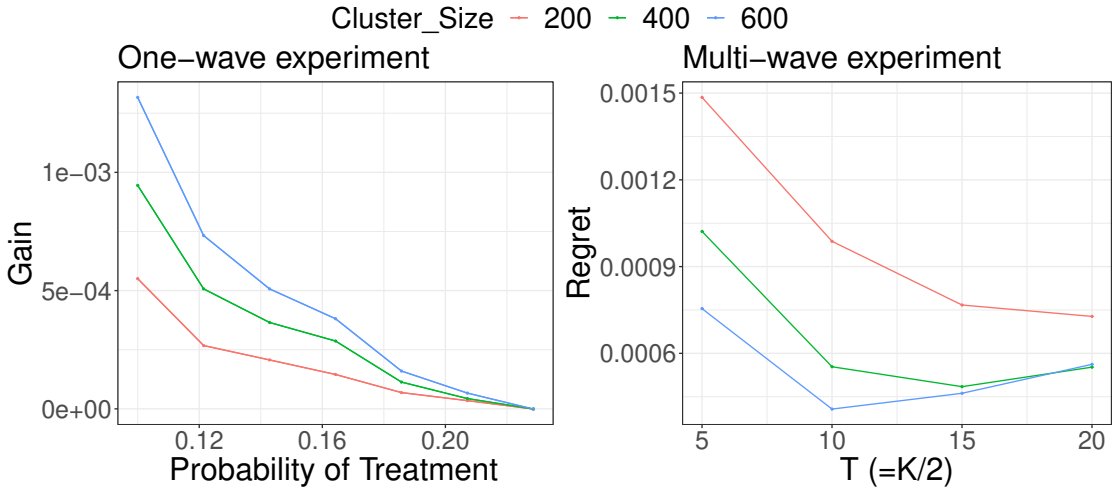


Figure A.4.8. Simulations: results with covariates (Section A.4.3). The panel on the left reports the relative welfare improvement from increasing the treatment probability in remote areas by 5% upon rejection of H_0. The panel on the right reports the in-sample regret. 400 replications.

Additional Algorithms

Algorithm A.5.1. Adaptive Experiment with Many Coordinates

Require: Starting value $\beta_0 \in \mathbb{R}$, K clusters, $T+1$ periods of experimentation, constant $\bar{C}$.

1: Create pairs of clusters $\{k, k+1\}, k \in \{1, 3, \cdots, K-1\}$;

2: $t = 0$(*baseline*):

a: Assign treatments as $D_{i,0}^{(h)} | X_i^{(h)} = x \sim \pi(x; \beta_0)$ for all $h \in \{1, \cdots, K\}$.

b: For n units in each cluster observe $Y_{i,0}^{(h)}, h \in \{1, \cdots, K\}$.

c: For cluster k initalize a gradient estimate $\widehat{V}_{k,t} = 0$ and initial parameters $\check{\beta}_k^o = \beta_0$.

3: **while** $1 \leq w \leq \check{T} = \frac{T}{p}$ **do**

4: **for each** $j \in \{1, \cdots, p\}$ **do**

a: Define

$$\check{\beta}_h^w = \begin{cases} P_{\mathscr{B}_1, \mathscr{B}_2 - \eta_n}\left[\check{\beta}_h^{w-1} + \alpha_{h+2,w-1}\widehat{V}_{h+2,w-1}\right], & h \in \{1, \cdots, K-2\}, \\ P_{\mathscr{B}_1, \mathscr{B}_2 - \eta_n}\left[\check{\beta}_h^{w-1} + \alpha_{h+2,w-1}\widehat{V}_{1,w-1}\right], & h \in \{K-1, K\}. \end{cases}$$

Here, $P_{\mathscr{B}_1, \mathscr{B}_2 - \eta_n}$ denotes the projection operator onto the set $[\mathscr{B}_1, \mathscr{B}_2 - \eta_n]^p$.

b: Assign treatments as (for a finite constant $\bar{C}$)

$$D_{i,t}^{(h)} | X_{i,t}^{(h)} = x \sim \pi(x, \beta_{h,w}), \quad \beta_{h,w} = \begin{cases} \check{\beta}_h^w + \eta_n \underline{e}_j \text{ if } h \text{ is odd} \\ \check{\beta}_h^t - \eta_n \underline{e}_j \text{ if } h \text{ is even} \end{cases}, \quad \bar{C} n^{-1/2} < \eta_n < \bar{C} n^{-1/4}$$

where $\underline{e}_j$ is the vector of zero, with entry j equal to one (see Equation (1.9)).

c: For n units in each cluster $h \in \{1, \cdots, K\}$ observe $Y_{i,t}^{(h)}$.

d: For each pair $\{k, k+1\}$, estimate

$$\hat{V}_{k,w}^{(j)} = \hat{V}_{k+1,w}^{(j)} = \frac{1}{2\eta_n}\left[\bar{Y}_t^{(k)} - \bar{Y}_0^{(k)}\right] - \frac{1}{2\eta_n}\left[\bar{Y}_t^{(k+1)} - \bar{Y}_0^{(k+1)}\right].$$

e: $t \leftarrow t+1$.

5: **end for**

f: $w \leftarrow w+1$.

6: **end while**

7: Return $\hat{\beta}^* = \frac{1}{K}\sum_{k=1}^K \check{\beta}_k^{\check{T}}$

Algorithm A.5.2. Dynamic Treatment Effects with $\beta \in \mathbb{R}$

Require: Parameter space $\mathscr{B}$, clusters $\{1,\cdots,K\}$, two periods $\{t,t+1\}$, perturbation η_n.

1: Group clusters into groups $r \in \{1,\cdots,K/3\}$ of $\{k,k+1,k+2\}$;

2: Construct a grid of parameters $\mathscr{G} \subset [0,1]^2$ equally spaced on $[0,1]^2$;

3: Assign each parameter $(\beta_1^r,\beta_2^r) \in \mathscr{G}$ to a different triad r.

4: **for** each $r \in \{1,\cdots,K/3\}$ **do**

$$\begin{aligned}
D_{i,t}^{(k)}|X_i^{(k)},\beta_1^r,\beta_2^r &\sim \pi(X_i^{(k)},\beta_2^r)\\
D_{i,t+1}^{(k)}|X_i^{(k)},\beta_1^r,\beta_2^r &\sim \pi(X_i^{(k)},\beta_1^r)\\
D_{i,t}^{(k+1)}|X_i^{(k)},\beta_1^r,\beta_2^r &\sim \pi(X_i^{(k)},\beta_2^r+\eta_n)\\
D_{i,t+1}^{(k+1)}|X_i^{(k)},\beta_1^r,\beta_2^r &\sim \pi(X_i^{(k)},\beta_1^r)\\
D_{i,t}^{(k+2)}|X_i^{(k)},\beta_1^r,\beta_2^r &\sim \pi(X_i^{(k)},\beta_2^r)\\
D_{i,t+1}^{(k+2)}|X_i^{(k)},\beta_1^r,\beta_2^r &\sim \pi(X_i^{(k)},\beta_1^r+\eta_n)
\end{aligned} \quad . \tag{A.5.1}$$

5: **end for**

6: For each $k \in \{1,4,\cdots,K-2\}$ estimate

$$\widehat{g_{1,k}} = \frac{\bar{Y}_{t+1}^{(k)} - \bar{Y}_{t+1}^{(k+2)}}{\eta_n}, \quad \widehat{g_{2,k}} = \frac{\bar{Y}_{t+1}^{(k)} - \bar{Y}_{t+1}^{(k+1)}}{\eta_n}, \quad \widetilde{\Gamma}_k = \frac{1}{3}\sum_{h\in\{k,k+1,k+2\}} \bar{Y}_{t+1}^{(h)} \tag{A.5.2}$$

Algorithm A.5.3. Welfare maximization with a "non-adaptive" experiment

Require: K clusters, $T = p$ periods; pairs of clusters $\{k,k+1\}, k \in \{1,3,\cdots,K-1\}$;

1: $t = 0$:

a: For n units in each cluster observe the baseline outcome $Y_{i,0}^{(h)}, h \in \{1,\cdots,K\}$.

b: Assign each pair $(k,k+1)$ to an element $\beta^k \in \mathscr{G}$, where $\mathscr{G}$ is an equally spaced grid of $\mathscr{B}$

2: **while** $1 \le t \le T$ **do**

a: Assign treatments as

$$D_{i,t}^{(h)} \sim \pi(1,\beta^h), \quad \beta^h = \begin{cases} \check{\beta}^h + \eta_n \underline{e}_t \text{ if } h \text{ is even} \\ \check{\beta}^{h-1} - \eta_n \underline{e}_t \text{ if } h \text{ is odd} \end{cases}, \quad n^{-1/2} < \eta_n \le n^{-1/4}$$

b: For each pair $(k,k+1)$, constructs the t entry

$$\hat{V}_{(k,k+1)}^{(t)}(\beta^k) = \frac{1}{2\eta_n}\left[\bar{Y}_t^k - \bar{Y}_0^k\right] - \frac{1}{2\eta_n}\left[\bar{Y}_t^{k+1} - \bar{Y}_0^{k+1}\right] \tag{A.5.3}$$

3: **end while**

4: Return $\hat{\beta}^{ow}$ as in Equation (A.1.7).

Algorithm A.5.4. Adaptive Experiment with staggered adoption

Require: Starting value $\beta \in \mathbb{R}$, K clusters, $T+1$ periods of experimentation.

1: Create pairs of clusters $\{k, k+1\}, k \in \{1, 3, \cdots, K-1\}$;

2: $t = 0$:

a: For n units in each cluster observe the baseline outcome $Y_{i,0}^{(h)}, h \in \{1, \cdots, K\}, \check{\beta}^0 = \beta$.

b: Initalize a gradient estimate $\widehat{V}_t = 0$

3: **while** $1 \leq t \leq T$ **do**

a: Sample without replaceent one pair of clusters $\{k, k+1\}$ not observed in previous iterations;

b: Define

$$\check{\beta}^t = \check{\beta}^{t-1} + \alpha_t \widehat{V}_t;$$

c: Assign treatments as

$$D_{i,t}^{(h)} \sim \pi(1, \beta_t), \quad \beta_t = \begin{cases} \check{\beta}^t + \eta_n \text{ if } h \text{ is even} \\ \check{\beta}^t - \eta_n \text{ if } h \text{ is odd} \end{cases}, \quad n^{-1/2} < \eta_n \leq n^{-1/4}$$

d: For n units in each cluster $h \in \{1, \cdots, K\}$ observe $Y_{i,t}^{(h)}$.

4: **end while**

5: Return $\hat{\beta}^* = \check{\beta}^T$

Appendix B

Appendix to Chapter 2

Extensions

In this section, we discuss two extensions: spillovers on non-compliance and target and sampled units having different distributions.

Spillovers on Non-Compliance: Identification of Welfare Effects

In the presence of non-compliance, we interpret the problem from an intention to treat perspective. A challenge is that the treatment assigned to an individual also influences the selection into the treatment of her friends. We use as a working model the one in Equation (2.7). The model is consistent with an exogenous interference model over the compliance: it states that the selection into the treatment of an individual depends on her and neighbors' treatment assignment. Denote

$$T_i(\pi) = h_\theta\Big(\pi(X_i), \sum_{k\in\mathcal{N}_i}\pi(X_k), Z_i, |\mathcal{N}_i|, \nu_i\Big)$$

the potential selected treatment, if treatments D_i were assigned under policy $\pi(\cdot)$ (i.e., $D_i = \pi(X_i)$)). In the following theorem, we show that the utilitarian welfare is the expected outcome, once treatments D_i are assigned according to the deterministic rule $\pi(X_i)$. Denote $\mathbb{E}_\pi$ the expectation conditional on the event that $D_i = \pi(X_i)$ for all $i \in \{1,\cdots,E\}$.

Theorem B.1.1. *Consider the model in Equation* (2.7)*. Assume that the following holds:*

$$\varepsilon_i \perp \Big(A, Z, (\nu_j)_{j=1}^E, (\varepsilon_{D_j})_{j=1}^E\Big)\Big|Z_i, |\mathcal{N}_i|, \quad \nu_i \perp \Big(A, Z, (\varepsilon_{D_j})_{j=1}^E\Big),$$

with ν_i *being i.i.d.. Then, for each* $i \in \{1, \cdots, E\}$*,*

$$\mathbb{E}_\pi[Y_i] = \mathbb{E}\Big[r\Big(T_i(\pi), \sum_{k\in\mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big] = \mathbb{E}\Big[\sum_{d\in\{0,1\}, s\in\{0,\cdots,|\mathcal{N}_i|\}} J_i(d,s) \times H_i(d,s,\pi)\Big], \tag{B.1.1}$$

where

$$\begin{aligned} J_i(d,s) &= \mathbb{E}\Big[Y_i\Big|Z_i, |\mathcal{N}_i|, T_i = d, \sum_{k\in\mathcal{N}_i} T_k = s\Big] \\ H_i(d,s,\pi) &= P_\theta\Big(T_i = d\Big|Z_i, |\mathcal{N}_i|, V_i(\pi)\Big) \times \\ &\qquad \sum_{u_1,\cdots,u_l:\sum_v u_v = s} \prod_{k=1}^{|\mathcal{N}_i|} P_\theta\Big(T_{\mathcal{N}_i^{(k)}} = u_k\Big|Z_{\mathcal{N}_i^{(k)}}, |\mathcal{N}_{\mathcal{N}_i^{(k)}}|, V_{\mathcal{N}_i^{(k)}}(\pi)\Big), \end{aligned} \tag{B.1.2}$$

where $V_i(\pi) = \Big\{D_i = \pi(X_i), \sum_{k\in\mathcal{N}_i} D_k = \sum_{k\in\mathcal{N}_i} \pi(X_k), \pi(X_i), \sum_{k\in\mathcal{N}_i} \pi(X_k)\Big\}$*, and* $P_\theta(T_i = 1|\cdot)$ *denotes the conditional probability of selection into treatment indexed by the parameters* θ*.*

The proof is contained in Appendix B.6. Theorem B.1.1 defines a causal estimand in terms of simple to estimate conditional expectations. This is commonly adopted in the literature on mediation analysis and g-estimation in bio-statistics (see, e.g., Naimi et al. 2017), while here we adopt such a framework in the context of interference.

The welfare effect of an incentive π depends on conditional mean functions that can be estimated from observed data. Here $J_i(\cdot)$ denotes the conditional mean of the outcome variable given the *selected* treatment, and $H_i(\cdot)$ denotes the conditional probability of selecting into treatment, conditional on the individual and neighbors' incentives. Observe that $H_i(\cdot)$ takes a *simple* expression which only depends on the individual probability of selected treatments $P_\theta(T_i = 1|\cdot)$, conditional on individual's and neighbors' treatment assignments.

Interestingly, $H_i(\cdot)$ also depends on the treatment assigned to the second-degree neighbors. This is intuitive: an incentive to Elena affects the selection into the treatment of her friend Matteo

whose decision affects the outcome of Matteo's friend Riccardo. Therefore identification requires information from first-degree and second-degree neighbors.

Contrasting with existing literature, a strand of literature on interference and non compliance includes Imai et al. (2018), Kang and Imbens (2016), Vazquez-Bare (2022), DiTraglia et al. (2020), and the recent work of Kim (2020). The above references discuss the identification of treatment effects under non-compliance without discussing the problem of welfare maximization. This is an important difference that motivates the exogenous interference model we propose in Equation (2.7) and the identification strategy that uses information from the second-degree neighbors.[1]

Different Target and Sample Units

Finally, we study the case where target and sampled units are drawn from different populations. The following condition is imposed.

Assumption B.1.1. Assume that target and sample units follow the following laws:

$$(Z_i, Z_{k\in\mathcal{N}_i})\Big||\mathcal{N}_i| \sim F_{s,|\mathcal{N}_i|}, \quad |\mathcal{N}_i| \sim G_s \quad \forall i \in \{1,\cdots,E\},$$
$$(Z_i, Z_{k\in\mathcal{N}_i})\Big||\mathcal{N}_i| \sim F_{t,|\mathcal{N}_i|}, \quad |\mathcal{N}_i| \sim G_t \quad \forall i \in \mathscr{I}.$$

with unobservables $\varepsilon_i|Z_i, Z_{k\in\mathcal{N}_i}, |\mathcal{N}_i|$ having the same conditional distribution in the target and sample population.

Assumption B.1.1 states that the distributions of individual covariates, neighbors' covariates, and the number of neighbors differ across the target and sampled units only. The assumption reads as follows: the joint distribution of $(Z_i, Z_{k\in\mathcal{N}_i})$, given the degree $|\mathcal{N}_i|$ is the same across units having the same degree and being in the same population (either target or

[1] We note in particular that the difference with Kim (2020) is that our model provides a simple closed-form expression of potential outcomes as functions of the *realized* treatments and exploits a local interference model also over the compliance, while Kim (2020) proposes potential outcomes as functions of expected treatments.

sampled population) and similarly the marginal distribution of $|\mathcal{N}_i|$, whereas such distributions may be different between target and sampled units.

Under the second condition in Assumption B.1.1 the conditional mean function is the same on target and sampled units, whereas the distribution of covariates and network may differ. We define $f_{s,|\mathcal{N}_i|}, f_{t,|\mathcal{N}_i|}$ the corresponding Radon-Nikodym derivatives of $F_{s,|\mathcal{N}_i|}, F_{t,|\mathcal{N}_i|}$, of sampled and target units with respect to a common dominating measure on $\mathcal{Z} \times \mathcal{Z}^{|\mathcal{N}_i|}$ and similarly g_s, g_t the corresponding Radon-Nikodym derivatives of G_s, G_t.[2] Then, we impose the following condition:

Assumption B.1.2. *Assume that*

$$f_{t,|\mathcal{N}_i|}(Z_i, Z_{k\in\mathcal{N}_i}) g_t(|\mathcal{N}_i|) = \rho(Z_i, Z_{k\in\mathcal{N}_i}, |\mathcal{N}_i|) f_{s,|\mathcal{N}_i|}(Z_i, Z_{k\in\mathcal{N}_i}) g_s(|\mathcal{N}_i|),$$

with $\rho(Z_i, Z_{k\in\mathcal{N}_i}, |\mathcal{N}_i|) \leq \bar{\rho} < \infty$ *almost surely.*

Assumption B.1.2 follows similarly to Kitagawa and Tetenov (2018), where the ratio of the two densities is assumed to be bounded almost surely. The empirical welfare criterion is constructed as follows:

$$\begin{aligned} &W_n^t(\pi, m^c, e^c) = \\ &\frac{1}{n}\sum_{i=1}^{n}\Big[\frac{1\{S_i(\pi) = \sum_{k\in\mathcal{N}_i} D_k, \pi(X_i) = D_i\}}{e^c\Big(\pi(X_i), S_i(\pi), Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|\Big)}\Big(Y_i - m^c(\pi(X_i), S_i(\pi), Z_i, |\mathcal{N}_i|)\Big)\rho(Z_i, Z_{k\in\mathcal{N}_i}, |\mathcal{N}_i|)\Big] \\ &+\frac{1}{n}\sum_{i=1}^{n}\Big[m^c(\pi(X_i), S_i(\pi), Z_i, |\mathcal{N}_i|)\rho(Z_i, Z_{k\in\mathcal{N}_i}, |\mathcal{N}_i|)\Big]. \end{aligned} \tag{B.1.3}$$

We can now state the following theorem whose proof is contained in Appendix B.5.

Theorem B.1.2. *Let Assumptions 2.2.1, 2.2.2, 2.2.3, 2.2.4, 2.2.5, 2.3.1, 2.3.2, B.1.1, B.1.2, and either* $m^c = m$ *or* $e^c = e$ *hold. Then we obtain that for* $\hat{\pi}^{aipw}_{m^c,e^c}$ *maximizing Equation* (B.1.3)*,*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}_{m^c,e^c})\Big] = \mathcal{O}\Big(\mathbb{E}[\mathscr{P}_{M,|\mathscr{L}|}(\mathcal{N}_E)]\sqrt{\frac{\mathrm{VC}(\Pi)}{n}}\Big),$$

[2]Since $|\mathcal{N}_i|$ is discrete, $g_s(l), g_t(l)$ corresponds to the probability of the number of neighbors being equal to l.

for a polynomial function $\mathscr{P}_{M,|\mathscr{L}|}$ with bounded degree.

A Numerical Study

In this section, we study the numerical performance of the proposed methodology in a small simulation study. We simulate data according to the following data generating process (DGP):

$$\begin{aligned} Y_i &= |\mathscr{N}_i|^{-1}\Big(X_i\beta_1 + X_i\beta_2 D_i + \mu\Big)\sum_{k\in\mathscr{N}_i} D_k + X_i\beta_3 D_i + \varepsilon_i \\ \varepsilon_i &= \eta_i/\sqrt{2} + \sum_{k\in\mathscr{N}_i}\eta_k/\sqrt{2|\mathscr{N}_i|}, \quad \eta_i \sim_{i.i.d.} \mathscr{N}(0,1), \end{aligned} \tag{B.2.1}$$

where $|\mathscr{N}_i|$ is set to be one for the elements with no neighbors. We simulate covariates as $X_{i,u} \sim_{i.i.d.} \mathscr{U}(-1,1)$ for $u \in \{1,2,3,4\}$. We draw $\beta_1, \beta_2, \mu \in \{-1,1\}^3$ independently and with equal probabilities. We draw $\beta_3 \in \{-1.5, 1.5\}$ with equal probabilities. We evaluate the NEWM method with and without balancing score adjustment, under correct specification of the conditional mean function. We impose trimming at 1% on the estimation of the balancing score. We compare the performance of the proposed methodology to two competing methods that ignore network effects. These are the empirical welfare maximization procedures with and without regression adjustment (Kitagawa and Tetenov, 2018; Athey and Wager, 2021).[3] For any of the method under consideration, we consider a policy function of the form

$$\pi(X_i) = 1\Big\{X_{i,1}\phi_1 + X_{i,2}\phi_2 + \phi_3 \geq 0\Big\}. \tag{B.2.2}$$

Optimization is performed using the MILP formulation.

[3]Namely, first, we consider the empirical welfare maximization method that does not account for network interference discussed in Kitagawa and Tetenov (2018) with a known balancing score of individual treatment. Second, we also compare the double robust formulation of the EWM method discussed in Athey and Wager (2021), with a linear model specification without network information.

In the first set of simulations, we consider a geometric network formation of the form

$$A_{i,j} = 1\Big\{ |X_{i,2} - X_{j,2}|/2 + |X_{i,4} - X_{j,4}|/2 \leq r_n \Big\} \tag{B.2.3}$$

where $r_n = \sqrt{4/2.75n}$ similarly to simulations in Leung (2020). In the second set of simulations, we generate Barabasi-Albert networks, where we first draw $n/5$ edges uniformly according to Erdős-Rényi graph with probabilities $p = 10/n$, and second, we draw sequentially connections of the new nodes to the existing ones with probability equal to the number of connection of each pre-existing node divided by the overall number of connections in the graph. We estimate the methods over 200 data sets, and we evaluate the performance of each estimate over 1000 networks, drawn from the same DGP. Results are collected in Table B.2.1. The table reports the welfare for different sample sizes. Observe that by construction, the sample size also corresponds to different data-generating processes of the network formation model, which, in turn, may shift by a constant the value of the regret.[4] The table shows that our procedure uniformly outperforms methods that ignore network effects.

Table B.2.1. Simulation results. Out-of-sample median *welfare* over 200 replications. DR is the method in Athey and Wager (2021) with estimated balancing score and EWM PS is the method in Kitagawa and Tetenov (2018) with known balancing score. NEWM1 is the proposed method with a correctly specified outcome model. NEWM2 is the double robust NEWM. G denotes the geometric network, and AB the Albert-Barabasi network. Different sample sizes correspond to different data-generating processes.

Welfare	$n = 50$		$n = 70$		$n = 100$		$n = 150$		$n = 200$	
	G	AB	G	AB	G	AB	G	AB	G	AB
DR	1.51	0.94	1.50	1.08	1.42	1.05	1.53	0.95	1.41	0.95
EWM PS	1.21	0.93	1.23	0.92	1.32	0.93	1.38	0.90	1.29	0.95
NEWM1	1.74	1.31	1.87	1.38	1.93	1.37	1.91	1.40	2.00	1.39
NEWM2	1.78	1.22	1.89	1.33	1.89	1.37	1.94	1.28	1.95	1.33

[4]Namely $\sup_{\pi \in \Pi} W(\pi)$ may differ across sample sizes because sample sizes correspond to different DGPs.

Proofs: Notation and Definitions

Before discussing the main results, we need to introduce the necessary notation. We define $S_i(\pi) = \sum_{k \in \mathcal{N}_i} \pi(X_k)$. We denote $\mathscr{A}_n$ the space of symmetric matrices in $\mathbb{R}^{n \times n}$ with entries being either zero or ones.

Definition B.3.1 (Proper Cover)**.** *Given an adjacency matrix $A_n \in \mathscr{A}_n$, with n rows and columns, a family $\mathscr{C}_n = \{\mathscr{C}_n(j)\}$ of disjoint subsets of $[n]$ is a proper cover of A_n if $\cup \mathscr{C}_n(j) = [n]$ and $\mathscr{C}_n(j)$ contains units such that for any pair of elements $\{(i,k) \in \mathscr{C}_n(j), k \neq i\}$, $A_n^{(i,k)} = 0$.*

The size of the smallest proper cover is the chromatic number, defined as $\chi(A_n)$.

Definition B.3.2 (Chromatic number)**.** *The chromatic number $\chi(A_n)$, denotes the size of the smallest proper cover of A_n.*

Definition B.3.3. *For a given matrix $A \in \mathscr{A}$, we define $A_n^2(A) \in \mathscr{A}_n$ the adjacency matrix where each row corresponds to a unit $i \in \{1,...,n\}$ and where two of such units are connected if they are neighbor or they share one first or second degree neighbor. Similarly $A_n^M(A)$ is the adjacency matrix obtained after connecting such units sharing common neighbors up to Mth degree. Here $N_{i,M}$ defines the set of neighbors of individual $i \in \{1, \cdots, n\}$ with adjacency matrix A_n^M.*

The proper cover of A_n^2 is defined as $\mathscr{C}_n^2 = \{\mathscr{C}_n^2(j)\}$ with chromatic number $\chi(A_n^2)$. Similarly $\mathscr{C}_n^M = \{\mathscr{C}_n^M(j)\}$ with chromatic number $\chi(A_n^M)$ is the proper cover of A_n^M.

Next, we discuss definitions on covering numbers.[5] For $z_1^n = (z_1,...,z_n)$ be arbitrary points in $\mathscr{Z}$, for a function class $\mathscr{F}$, with $f \in \mathscr{F}$, $f : \mathscr{Z} \mapsto \mathbb{R}$, we define,

$$\mathscr{F}(z_1^n) = \{f(z_1),...,f(z_n) : f \in \mathscr{F}\}. \tag{B.3.1}$$

Definition B.3.4. For a class of functions $\mathscr{F}$, with $f : \mathscr{Z} \mapsto \mathbb{R}$, $\forall f \in \mathscr{F}$ and n data points $z_1,...,z_n \in \mathscr{Z}$ define the l_q-covering number $\mathscr{N}_q\Big(\varepsilon, \mathscr{F}(z_1^n)\Big)$ to be the cardinality of the smallest

[5] Here we adopt similar notation to Chapter 28 and Chapter 29 of Devroye et al. (2013).

cover $\mathscr{S} := \{s_1, ..., s_N\}$, with $s_j \in \mathbb{R}^n$, such that for each $f \in \mathscr{F}$, there exist an $s_j \in \mathscr{S}$ such that $(\frac{1}{n}\sum_{i=1}^{n} |f(z_i) - s_j^{(i)}|^q)^{1/q} < \varepsilon$.

Throughout our analysis, for a random variable $X = (X_1, ..., X_n)$ we denote $\mathbb{E}_X[.]$ the expectation with respect to X. The Rademacher complexity is defined as follows.

Definition B.3.5. Let $X_1, ..., X_n$ be arbitrary random variables. Let $\sigma = \{\sigma_i\}_{i=1}^n$ be *i.i.d* Rademacher random variables (i.e., $P(\sigma_i = -1) = P(\sigma_i = 1) = 1/2$), independent of $X_1, ..., X_n$. We define the empirical Rademacher complexity as

$$\mathscr{R}_n(\mathscr{F}) = \mathbb{E}_\sigma \Big[\sup_{f \in \mathscr{F}} |\frac{1}{n} \sum_{i=1}^{n} \sigma_i f(X_i)| \Big| X_1, ..., X_n \Big]. \tag{B.3.2}$$

Main Lemmas

Lemma B.4.1. *The following holds:* $\chi(A_n) \leq \chi(A_n^M) \leq M \mathscr{N}_{n,M+1}^{M+1}$.

Proof of Lemma B.4.1. The first inequality follows by Definition B.3.3. The second inequality follows by Brook's Theorem (Brooks, 1941), since two individuals $(i, j) \in \{1, \cdots, n\}^2$ are connected if they share a common neighbor up to the *Mth* degree. The maximum degree is bounded by $\mathscr{N}_{n,1} + \mathscr{N}_{n,1} \times \mathscr{N}_{n,2} + \cdots + \prod_{s=1}^{M+1} \mathscr{N}_{n,s} \leq M \mathscr{N}_{n,M+1}^{M+1}$. □

Lemma B.4.2. *Let* $\mathscr{F}_1, \cdots, \mathscr{F}_k$ *be classes of bounded functions with VC-dimension* $v < \infty$ *and envelope* $\bar{F} < \infty$*, for* $k \geq 2$*. Let*

$$\mathscr{J} = \Big\{ f_1(f_2 + ... + f_k), \quad f_j \in \mathscr{F}_j, \quad j = 1, ..., k \Big\}, \quad \mathscr{J}_n(z_1^n) = \Big\{ h(z_1), ..., h(z_n); h \in \mathscr{J} \Big\}.$$

For arbitrary fixed points $z_1^n \in \mathbb{R}^d$, $\int_0^{2\bar{F}} \sqrt{\log\Big(\mathscr{N}_1\Big(u, \mathscr{J}(z_1^n)\Big)\Big)} du < c_{\bar{F}} \sqrt{k \log(k) v}$*. for a constant* $c_{\bar{F}} < \infty$ *that only depend on* $\bar{F}$*.*

Proof of Lemma B.4.2. Let $\mathscr{F}_{-1}(z_1^n) = \{f_2(z_1^n) + ... + f_k(z_1^n), f_j \in \mathscr{F}_j, j = 2, ..., k\}$. By Theorem

29.6 in Devroye et al. (2013),

$$\mathscr{N}_1\Big(\varepsilon, \mathscr{F}_{-1}(z_1^n)\Big) \leq \prod_{j=2}^{k} \mathscr{N}_1\Big(\varepsilon/(k-1), \mathscr{F}_j(z_1^n)\Big).$$

By Theorem 29.7 in Devroye et al. (2013),

$$\mathscr{N}_1\Big(\varepsilon, \mathscr{J}_n(z_1^n)\Big) \leq \prod_{j=2}^{k} \mathscr{N}_1\Big(\frac{\varepsilon}{2(k-1)\bar{F}}, \mathscr{F}_j(z_1^n)\Big) \mathscr{N}_1\Big(\frac{\varepsilon}{2\bar{F}}, \mathscr{F}_1(z_1^n)\Big). \tag{B.4.1}$$

By standard properties of covering numbers, for a generic set $\mathscr{H}$, $\mathscr{N}_1(\varepsilon, \mathscr{H}) \leq \mathscr{N}_2(\varepsilon, \mathscr{H})$. Therefore

$$\text{(B.4.1)} \leq \prod_{j=2}^{k} \mathscr{N}_2\Big(\frac{\varepsilon}{2(k-1)\bar{F}}, \mathscr{F}_j(z_1^n)\Big) \mathscr{N}_2\Big(\frac{\varepsilon}{2\bar{F}}, \mathscr{F}_1(z_1^n)\Big).$$

We apply now a uniform entropy bound for the covering number. By Theorem 2.6.7 of Van Der Vaart and Wellner (1996), we have that for a universal constant $C < \infty$,

$$\mathscr{N}_2\Big(\frac{\varepsilon}{2(k-1)\bar{F}}, \mathscr{F}_j(z_1^n)\Big) \leq C(v+1)(16e)^{(v+1)} \Big(\frac{2\bar{F}^2(k-1)}{\varepsilon}\Big)^{2v}$$

which implies that

$$\begin{aligned}
\log\Big(\mathscr{N}_1\Big(\varepsilon, \mathscr{J}_n(z_1^n)\Big)\Big) &\leq \sum_{j=1}^{k-1} \log\Big(\mathscr{N}_2\Big(\frac{\varepsilon}{2\bar{F}(k-1)}, \mathscr{F}_j(z_1^n)\Big)\Big) + \log\Big(\mathscr{N}_2\Big(\frac{\varepsilon}{2\bar{F}}, \mathscr{F}_1(z_1^n)\Big)\Big) \\
&\leq k \log\Big(C(v+1)(16e)^{v+1}\Big) + k2v \log(2C\bar{F}^2(k-1)/\varepsilon).
\end{aligned}$$

Since $\int_0^{2\bar{F}} \sqrt{k \log\Big(C(v+1)(16e)^{v+1}\Big) + k2v \log(2C\bar{F}^2(k-1)/\varepsilon)}\, d\varepsilon \leq c_{\bar{F}} \sqrt{k \log(k) v}$ for a constant $c_{\bar{F}} < \infty$, the proof completes. □

The following lemma controls the Rademacher complexity with individualized treatments.

Lemma B.4.3. *Let Π be a function class with $\pi : \mathbb{R}^d \mapsto \{0,1\}$ for any $\pi \in \Pi$, with finite VC-dimension, denoted as* $\mathrm{VC}(\Pi)$. *Take arbitrary random variables $(X_i^{(1)}, \cdots, X_i^{(h_i)}, V_i), X_i, V_i \in \mathbb{R}^d$, functions $f_i : \mathbb{Z} \mapsto \mathbb{R}$, with f_i L-lipschitz for all $i \in \{1, \cdots, n\}$. Let $\bar{h} = \max_i h_i$. Let $\sigma_1, \cdots, \sigma_n$ be*

independent Rademacher random variables (independent of $(X_i, V_i)_{i=1}^n$). Then for a constant $C < \infty$ independent of L, $\bar{h}$ and n

$$\mathbb{E}_\sigma\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n f_i\Big(\sum_{k\in\{1,\dots,h_i\}}\pi(X_i^{(k)})\Big)\pi(V_i)\sigma_i\Big|\Big] \le CL\sqrt{\frac{(\bar{h}+1)\log(\bar{h}+1)\mathrm{VC}(\Pi)}{n}}. \tag{B.4.2}$$

Proof of Lemma B.4.7. We first discuss the case where $h_i = h$ for all i. We then turn to the case where this condition fails at the end of the proof.

Lipschitz properties

First, we add and subtract the value of the function $f_i(0)$ at zero. Namely,

$$\begin{aligned}
&\mathbb{E}_\sigma\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\Big(f_i\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)})\Big) - f_i(0) + f_i(0)\Big)\pi(V_i)\Big|\Big] \\
&\le \underbrace{\mathbb{E}_\sigma\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\Big(f_i\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)})\Big) - f_i(0)\Big)\pi(V_i)\Big|\Big]}_{(1)} \\
&+ \underbrace{\mathbb{E}_\sigma\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i f_i(0)\pi(V_i)\Big|\Big]}_{(2)}
\end{aligned} \tag{B.4.3}$$

where the last inequality follows by the triangular inequality. We bound first (1). We decompose the supremum over $\pi \in \Pi$ as follows.

$$\begin{aligned}
&\mathbb{E}_\sigma\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\Big(f_i\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)})\Big) - f_i(0)\Big)\pi(V_i)\Big|\Big] \\
&\le \mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\Big(f_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)})\Big) - f_i(0)\Big)\pi_1(V_i)\Big|\Big].
\end{aligned}$$

Let $\phi_i(s) = \Big(f_i(s) - f_i(0)\Big)$. Conditional on the data, ϕ_i is not random. It is Lipschitz and $\phi_i(0) = 0$ by assumption and by the fact that Y_i is uniformly bounded. By Theorem 4.12 in Ledoux and Talagrand (2011) (see also Lemma B.4.5)

$$\begin{aligned}
&\mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\Big(f_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)})\Big) - f_i(0)\Big)\pi_1(V_i)\Big|\Big] \\
&\le L2\mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)})\Big)\pi_1(V_i)\Big|\Big].
\end{aligned} \tag{B.4.4}$$

Function class decomposition

We decompose the supremum as follows:

$$L\mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n\sigma_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)})\Big)\pi_1(V_i)\Big|\Big]$$
$$\leq \underbrace{L\mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi,\cdots,\pi_{h+1}\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n\sigma_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_{k+1}(X_i^{(k)})\Big)\pi_1(V_i)\Big|\Big]}_{(J)}.$$

We re-parametrize the class of functions as follows.

$$(J)=\mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi,\tilde\pi_2\in\Pi_2,\cdots,\tilde\pi_{h+1}\in\Pi_{h+1}}\Big|\frac{1}{n}\sum_{i=1}^n\sigma_i\Big(\sum_{k\in\{1,\dots,h\}}\tilde\pi_{k+1}(X_i^{(1)},\cdots,X_i^{(h)})\Big)\pi_1(V_i)\Big|\Big],$$

where $\Pi_2,\dots,\Pi_{h+1}$, are such that $\tilde\pi_j\in\Pi_j:\tilde\pi_j(X_i^{(1)},\cdots,X_i^{(k)},\cdots,X_i^{(h)})=\pi(X_i^{(j)})$. By Theorem 29.4 in Devroye et al. (2013), $\mathrm{VC}(\Pi_j)=\mathrm{VC}(\Pi)$ for all $j\in\{1,\dots,h\}$. Let

$$\tilde\Pi=\Big\{\pi_1\Big(\sum_{j=1}^h\pi_{j+1}\Big),\pi_k\in\Pi_k,k=1,\cdots,h+1\Big\}.$$

For any fixed point data point z_1^n, by Lemma B.4.2, the Dudley's integral of the function class $\tilde\Pi(z_1^n)$ is bounded by $C\sqrt{(h+1)\log(h+1)\mathrm{VC}(\Pi)}$, for a finite constant C. Therefore, by B.4.2 and the Dudley's entropy integral bound (Lemma B.7.8) we can bound the conditional Rademacher complexity

$$\mathbb{E}_\sigma\Big[\sup_{\pi_1\in\Pi_1,\pi_2\in\Pi_2,\dots,\pi_{h+1}\in\Pi_{h+1}}\Big|\frac{1}{n}\sum_{i=1}^n\sigma_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_{k+1}(X_i^{(k)})\Big)\pi_1(V_i)\Big|\Big]$$
$$\leq C\sqrt{(h+1)\frac{\mathrm{VC}(\Pi)\log(h+1)}{n}}$$

for a constant $C<\infty$.

Remaining components

For (2) in Equation (B.4.3) we can use the same argument used for (1). In particular, since Π has finite VC-dimension, and it contains functions mapping to $\{0,1\}$, by Theorem 2.6.7 of Van Der Vaart and Wellner (1996) [6], $\int_0^2\sqrt{\mathcal{N}_2(u,\Pi(z_1^n))}du<C\sqrt{\mathrm{VC}(\Pi)}$ for a constant C.

[6] The argument is the same used in the proof of Lemma B.4.2.

Conclusion

We are now left to discuss the case where h_i is different for each individual. In this case, we can assign to each unit covariates $\{X_{k\in\mathcal{N}_i}, \emptyset, \emptyset, \cdots, \emptyset\}$ with in total $\bar{h}$ and parametrizing $\pi(\emptyset) = 0$. The rest of the proof follows similarly as before. □

Lemma B.4.4. *$e^c(.) \in (\delta, 1-\delta)$ for $\delta \in (0,1)$. Then $\frac{1\{N=\sum_{k\in\mathcal{N}_i} D_k, d=D_i\}}{e^c\left(d,N,Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\right)}$ is $2/\delta$-Lipschitz in its second argument for all $d \in \{0,1\}$*

Proof of Lemma B.4.4. Let $O_i = (Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|)$. Then for any $N, N' \in \mathbb{Z}$

$$\Big|\frac{1\{N=\sum_{k\in\mathcal{N}_i} D_k, d=D_i\}}{e^c\Big(d,N,O_i\Big)} - \frac{1\{N'=\sum_{k\in\mathcal{N}_i} D_k, d=D_i\}}{e^c\Big(d,N',O_i\Big)}\Big| \leq \frac{2}{\delta}.$$

for $N \neq N'$. The last inequality follows from the fact that by the overlap condition and the triangular inequality. Since N is discrete, the right-hand side is at least $2/\delta|N-N'|$ which completes the proof. □

Auxiliary Lemmas for an Unbounded Outcome

The next lemma provides a bound on the Rademacher complexity in the presence of the composition of functions. We discuss the Ledoux-Talagrand contraction inequality (Ledoux and Talagrand, 2011; Chernozhukov et al., 2014) to the case of interest in this paper.

Lemma B.4.5. *Let $\phi_i : \mathbb{R} \mapsto \mathbb{R}$ be Lipschitz functions with parameter L, $\forall i \in \{1,...,n\}$, i.e., $|\phi_i(a) - \phi_i(b)| \leq L|a-b|$ for all $a,b \in \mathbb{R}$, with $\phi_i(0) = 0$. Then, for any $\mathcal{T} \subseteq \mathbb{R}^n$, with $t = (t_1,...,t_n) \in \mathbb{R}$, $\alpha = (\alpha_1,...,\alpha_n) \in \mathcal{A} \subseteq \{0,1\}^n$,*

$$\frac{1}{2}\mathbb{E}_\sigma\Big[sup_{t\in\mathcal{T},\alpha\in\mathcal{A}}\Big|\frac{1}{n}\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big|\Big] \leq L\mathbb{E}_\sigma\Big[sup_{t\in\mathcal{T},\alpha\in\mathcal{A}}\Big|\frac{1}{n}\sum_{i=1}^n \alpha_i\sigma_i t_i\Big|\Big].$$

Proof of Lemma B.4.5. The proof follows closely the one in Theorem 4.12 of Ledoux and Talagrand (2011) while dealing with the additional α vector. First note that if $\mathcal{T}$ is unbounded, there will be some setting so that the right hand side is infinity and the result trivially holds.

Therefore, we can focus to the case where $\mathscr{T}$ is bounded. First we aim to show that for $\mathscr{T} \subseteq \mathbb{R}^2$, for $\alpha \in \{0,1\}^2$,

$$\mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}} \alpha_1 t_1 + \sigma_2\phi(t_2)\alpha_2\Big] \leq \mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}} \alpha_1 t_1 + L\sigma_2 t_2\alpha_2\Big]. \tag{B.4.5}$$

If the claim above is true, than it follows that

$$\mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}} \alpha_1\phi_1(t_1)\sigma_1 + \sigma_2\phi(t_2)\alpha_2|\sigma_1\Big] \leq \mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}} \alpha_1\phi_1(t_1)\sigma_1 + L\sigma_2 t_2\alpha_2\Big|\sigma_1\Big].$$

as $\sigma_1\phi(t_1)$ simply transforms $\mathscr{T}$ (and it is still bounded, if not the claim would trivially holds), and we can iteretively apply this result. Hence, we first prove Equation (B.4.5). Define for $a,b \in \mathscr{A}$, $I(t,s,a,b) := \frac{1}{2}\Big(t_1a_1 + a_2\phi(t_2)\Big) + \frac{1}{2}\Big(s_1b_1 - b_2\phi(s_2)\Big)$. We want to show that the right hand side in Equation (B.4.5) is larger thant $I(t,s,a,b)$ for all $t,s \in \mathscr{T}$ and $a,b \in \mathscr{A}$. Since we are taking the supremum over t,s,a,b, we can assume without loss of generality that

$$t_1a_1 + a_2\phi(t_2) \geq s_1b_1 + b_2\phi(s_2), \quad s_1b_1 - b_2\phi(s_2) \geq t_1a_1 - a_2\phi(t_2). \tag{B.4.6}$$

We can now define four quantities of interest, being

$$m = b_1s_1 - b_2\phi(s_2), \quad n = b_1s_1 - Ls_2b_2, \quad m' = a_1t_1 + La_2t_2, \quad n' = a_1t_1 + a_2\phi(t_2).$$

We would like to show that $2I(t,s,a,b) = m + n' \leq m' + n$. We consider four different cases, similarly to the proof of Ledoux and Talagrand (2011) and argue that for any value of $(a_1,a_2,b_1,b_2) \in \{0,1\}^4$ the claim holds.

Case 1 Start from the case $a_2t_2, s_2b_2 \geq 0$. We know that $\phi(0) = 0$, so that $|b_2\phi(s_2)| \leq Lb_2s_2$. Now assume that $a_2t_2 \geq b_2s_2$. In this case

$$m - n = Lb_2s_2 - b_2\phi(s_2) \leq La_2t_2 - a_2\phi(t_2) = m' - n' \tag{B.4.7}$$

since $|a_2\phi(t_2) - b_2\phi(s_2)| \leq L|a_2t_2 - b_2s_2| = L(a_2t_2 - b_2s_2)$. To see why this last claim holds, note that for $a_2, b_2 = 1$, then the results hold by the condition $a_2t_2 \geq b_2s_2$ and Lipschitz continuity.

If instead $a_2 = 1, b_2 = 0$, the claim trivially holds. While the case $a_2 = 0, b_2 = 1$, then it must be that $s_2 = 0$ since we assumed that $a_2 t_2 \geq 0, b_2 s_2 \geq 0$ and $a_2 t_2 \geq b_2 s_2$. Thus $m - n \leq m' - n'$. If instead $b_2 s_2 \geq a_2 t_2$, then use $-\phi$ instead of ϕ and switch the roles of s, t giving a similar proof.

Case 2 Let $a_2 t_2 \leq 0, b_2 s_2 \leq 0$. Then the proof is the same as Case 1, switching the signs where necessary.

Case 3 Let $a_2 t_2 \geq 0, b_2 s_2 \leq 0$. Then we have $a_2 \phi(t_2) \leq L a_2 t_2$, since $a_2 \in \{0, 1\}$ and by Lipschitz properties of ϕ, $-b_2 \phi(s_2) \leq -b_2 L s_2$ so that $a_2 \phi(t_2) - b_2 \phi(s_2) \leq a_2 L t_2 - b_2 L s_2$ proving the claim.

Case 4 Let $a_2 t_2 \leq 0, b_2 s_2 \geq 0$. Then the claim follows simmetrically to Case 3.

We now conclude the proof. Denote $[x]_+ = \max\{0, x\}$ and $[x]_- = \max\{-x, 0\}$. Then we have

$$\begin{aligned}
\mathbb{E}\Big[\frac{1}{2}\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\Big|\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big|\Big] &\leq \mathbb{E}\Big[\frac{1}{2}\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\Big(\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big)_+\Big] \\
&\quad + \mathbb{E}\Big[\frac{1}{2}\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\Big(\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big)_-\Big] \\
&\leq \mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\Big(\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big)_+\Big]
\end{aligned}$$

where the last inequality follows by symmetry of σ_i and the fact that $(-x)_- = (x)_+$. note that $sup_x(x)_+ = (sup_x x)_+$. Therefore, using Equation (B.4.5)

$$\begin{aligned}
\mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\Big(\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big)_+\Big] &= \mathbb{E}\Big[\Big(\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big)_+\Big] \\
&\leq \mathbb{E}\Big[\Big(\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\sum_{i=1}^n L\sigma_i t_i\alpha_i\Big)_+\Big] \\
&\leq \mathbb{E}\Big[\Big|\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big|\Big] \leq \mathbb{E}\Big[\sup_{t\in\mathscr{T},\alpha\in\mathscr{A}}\Big|\sum_{i=1}^n \sigma_i\phi_i(t_i)\alpha_i\Big|\Big]
\end{aligned}$$

which completes the proof. □

Lemma B.4.6. *Let Z be an arbitrary random variable and $\mathscr{F}$ a class of uniformly bounded functions, i.e., there exist $\bar{F} < \infty$, such that $||f||_\infty \leq \bar{F}$ for all $f \in \mathscr{F}$. Let $Y_i \sim \mathscr{P}_i|Z$, where $Y \geq 0$ is a scalar. Let $(Y_i)_{i=1}^n|Z$ be pairwise independent across individuals i. Assume that for some $u > 0$, $\mathbb{E}[Y_i^{2+u}|Z] < B, \quad \forall i \in \{1, ..., n\}$. In addition assume that for any fixed points z_1^n, for some*

$V < \infty$, $\int_0^{2\bar{F}} \sqrt{\log\left(\mathscr{N}_1\left(u, \mathscr{F}(z_1^n)\right)\right)}du < \sqrt{V}$. *Let* σ_i *be i.i.d Rademacher random variables independent of* Y, Z. *Then there exist a constant* $0 < C_{\bar{F}} < \infty$ *that only depend on* $\bar{F}$ *and u, such that*

$$\int_0^{\infty} \mathbb{E}\Big[sup_{f\in\mathscr{F}}\Big|\frac{1}{n}\sigma_i \sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\Big|Z\Big]dy \leq C_{\bar{F}}\sqrt{\frac{BV}{n}}$$

for all $n \geq 1$.

Proof of Lemma B.4.6. The proof follows closely the one in Kitagawa and Tetenov (2019) (Lemma A.5) to express the bound as a function of the covering number (instead of the VC-dimension). First, define

$$\xi_n(y) = \sup_{f\in\mathscr{F}}\Big|\frac{1}{n}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i\Big|.$$

Denote $\bar{p}(y) = \frac{1}{n}\sum_{i=1}^{n} P(Y_i > y|Z)$.

Case 1

Consider first values of y for which $n\bar{p}(y) = \sum_{i=1}^{n} P(Y_i > y|Z) \leq 1$. Due to the envelope condition, and the definition of Rademacher random variables, we have

$$\Big|\frac{1}{n}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i\Big| \leq \bar{F}\frac{1}{n}\sum_{i=1}^{n} 1\{Y_i > y\}, \forall f \in \mathscr{F}.$$

Taking expectations we have $\mathbb{E}[\xi_n(y)] \leq \bar{F}\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n} 1\{Y_i > y\}\Big|Z\Big] = \bar{F}\bar{p}(y)$ and the right hand side is bounded by $\bar{F}\frac{1}{n}$ for this particular case.

Case 2

Consider now values of y such that $n\bar{p}(y) > 1$. Define the random variable $N_y = \sum_{i=1}^{n} 1\{Y_i > y\}$. Then we can write

$$\frac{1}{n}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i = \begin{cases} 0 \text{ if } N_y = 0 \\ \\ \frac{N_y}{n}\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i \text{ if } N_y \geq 1. \end{cases}$$

If $N_y \geq 1$, then

$$\begin{aligned}
\xi_n(y) &= \sup_{f\in\mathscr{F}}\Big|\frac{N_y}{n}\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)\sigma_i 1\{Y_i > y\}\Big| \\
&= \sup_{f\in\mathscr{F}}\Big|\frac{N_y}{n}\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)\sigma_i 1\{Y_i > y\} - \bar{p}(y)\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i + \\
&+ \bar{p}(y)\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i\Big| \\
&\leq \Big|\frac{N_y}{n} - \bar{p}(y)\Big|\sup_{f\in\mathscr{F}}\Big|\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i\Big| + \bar{p}(y)\sup_{f\in\mathscr{F}}\Big|\frac{1}{N_y}\sum_{i=1} f(Z_i)\sigma_i 1\{Y_i > y\}\Big|.
\end{aligned}$$

Denote $\mathbb{E}_\sigma$ the expectation only with respect to the Rademacher random variables σ. Conditional on N_y, $\xi_n(y)$ sums over N_y terms. Therefore, for a constant $0 < C_1 < \infty$ that only depend on $\bar{F}$,

$$\mathbb{E}_\sigma\Big[\bar{p}(y)\sup_{f\in\mathscr{F}}|\frac{1}{N_y}\sum_{i=1} f(Z_i)\sigma_i 1\{Y_i > y\}\Big|\Big] \leq C_1\bar{p}(y)\sqrt{V}g(N_y)$$

by Theorem 5.22 in Wainwright (2019) (Lemma B.7.8 in Appendix B.7) where $g(.)$ is defined in Lemma B.7.4. Similarly,

$$\mathbb{E}_\sigma\Big[\Big|\frac{N_y}{n} - \bar{p}(y)\Big|\sup_{f\in\mathscr{F}}\Big|\frac{1}{N_y}\sum_{i=1}^{n} f(Z_i)1\{Y_i > y\}\sigma_i\Big||Z\Big] \leq \Big|\frac{N_y}{n} - \bar{p}(y)\Big|C_1\sqrt{V}g(N_y).$$

For $N_y \geq 1$, it follows by the law of iterated expectations,

$$\mathbb{E}[\xi_n(y)|N_y, Z] \leq |\frac{N_y}{n} - \bar{p}(y)|C_1\sqrt{V}g(N_y) + C_1\bar{p}(y)\sqrt{V}g(N_y) \leq |\frac{N_y}{n} - \bar{p}(y)|C_1\sqrt{V} + C_1\bar{p}(y)\sqrt{V}g(N_y) \tag{B.4.8}$$

where the last inequality follows by the definition of the $g(.)$ function and the fact that $N_y \in \{1, 2, \cdots\}$. For $N_y = 0$ instead, we have

$$\mathbb{E}[\xi_n(y)|N_y, Z] \leq \Big|\frac{N_y}{n} - \bar{p}(y)\Big|C_1\sqrt{V}g(N_y) + C_1\bar{p}(y)\sqrt{V}g(N_y) = 0$$

by the definition of $g(.)$. Hence, the bound in Equation (B.4.8) always holds.

Unconditional expectation

We are left to bound the unconditional expectation with respect to N_y. Notice first that

$$\mathbb{E}\Big[\Big|\frac{N_y}{n} - \bar{p}(y)\Big||Z\Big] \leq C_1\sqrt{V}\sqrt{\mathbb{E}\Big[\Big|\frac{N_y}{n} - \bar{p}(y)\Big|^2\Big|Z\Big]} = C_1\sqrt{V}\sqrt{\text{Var}\Big(\frac{1}{n}\sum_{i=1}^{n} 1\{Y_i > y\}\Big|Z\Big)}.$$

By independence assumption,

$$C_1\sqrt{V}\sqrt{\mathrm{Var}\Big(\frac{1}{n}\sum_{i=1}^{n}1\{Y_i>y\}\Big|Z\Big)}\leq C_1\sqrt{V}\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}P(Y_i>y|Z)}.$$

In addition, since $n\bar{p}(y)>1$, by Lemma B.7.9 , $\mathbb{E}[g(N_y)|Z]\leq\frac{2}{\sqrt{n}\sqrt{\bar{p}(y)}}$. Combining the inequalities, it follows

$$\begin{aligned}\mathbb{E}[\xi_n(y)|Z]&\leq C_1\sqrt{V}\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}P(Y_i>y|Z)}+\bar{p}(y)C_1\sqrt{V}\frac{2}{\sqrt{n}\sqrt{\bar{p}(y)}}\\&\leq 2(1+C_1)\sqrt{\frac{V}{n}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}P(Y_i>y|Z)}.\end{aligned}$$

This bound is larger than the bound derived for $n\bar{p}(y)<1$, up to a constant factor $C_{\bar{F}}<\infty$. Therefore,

$$\mathbb{E}[\xi_n(y)|Z]\leq C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}P(Y_i>y|Z)}.$$

Integral bound

We can now write

$$\begin{aligned}\int_0^\infty\mathbb{E}[\xi_n(y)|Z]dy\leq\int_0^\infty C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^{n}\frac{P(Y_i>y|Z)}{n}}dy&=\int_0^1 C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^{n}\frac{P(Y_i>y|Z)}{n}}dy\\&\quad+\int_1^\infty C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^{n}\frac{P(Y_i>y|Z)}{n}}dy.\end{aligned}$$

The first term is bounded as follows. $\int_0^1 C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^{n}\frac{P(Y_i>y|Z)}{n}}dy\leq C_{\bar{F}}\sqrt{\frac{V}{n}}$. The second term is bounded instead as follows.

$$\begin{aligned}\int_1^\infty C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^{n}\frac{P(Y_i>y|Z)}{n}}dy&\leq C_{\bar{F}}\sqrt{\frac{V}{n}}\int_1^\infty\sqrt{\sum_{i=1}^{n}\frac{\mathbb{E}[Y_i^{2+u}|Z]}{ny^{2+u}}}dy\leq C_{\bar{F}}\sqrt{\frac{V}{n}}\int_1^\infty\sqrt{\frac{B}{y^{2+u}}}dy\\&\leq C'_{\bar{F}}\sqrt{\frac{VB}{n}}\end{aligned}$$

for a constant $C'_{\bar{F}}<\infty$ that only depend on $\bar{F}$ and u. □

Lemma B.4.7. *Let* Π *be a function class with* $\pi:\mathbb{R}^d\mapsto\{0,1\}$ *for any* $\pi\in\Pi$*, with finite VC-dimension, denoted as* $\mathrm{VC}(\Pi)$*. Let* $X^1,...,X^{h_i},V\in\mathbb{R}^d,O\in\mathbb{R}^k,Y\in\mathbb{R}$.

Let $K = (V_i, X_i^1, X_i^2, ..., X_i^{h_i}, O_i)_{i=1}^n$. Suppose that $\mathbb{E}[Y_i^{2+u}|K] < B_1 < \infty$ for all $i \in \{1, \cdots, n\}$ for some $u > 0$ and $(Y_i)_{i=1}^n$ are pairwise independent conditional on K. Let $\sigma_1, \cdots, \sigma_n$ be independent Rademacher random variables, independent of $(K_i, Y_i)_{i=1}^n$. Let $f : \mathbb{Z} \times \mathbb{R}^k \mapsto \mathbb{R}$ be L-Lipschitz in its first argument. Assume that $\mathbb{E}[f(0, O_i)^{2+u} Y_i^{2+u}|K] < B_2 < \infty$ for all $i \in \{1, \cdots, n\}$ for some $u > 0$. Let $\bar{h} = \max_i h_i$. Then for a constant $C < \infty$ that only depend on u,

$$\mathbb{E}_{Y,\sigma}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n} f\Big(\sum_{k\in\{1,\ldots,h\}} \pi(X_i^{(k)}), O_i\Big)\pi(V_i)\sigma_i Y_i\Big|\Big] \leq C(L+1)\sqrt{\frac{(\bar{h}+1)\log(\bar{h}+1)\mathrm{VC}(\Pi)(B_1+B_2)}{n}}, \tag{B.4.9}$$

where $\mathbb{E}_{Y,\sigma}$ denotes the conditional expectation taken only with respect to the variables $(Y_i, \sigma_i)_{i=1}^n$ conditional on K.

Proof of Lemma B.4.7. The proof follows similarly to Lemma B.4.3. We can consider the case where all units correspond to sums over h components, and discuss the case where this is violated at the end of the proof.

First decomposition

First, we add and subtract the value of the function $f(0, O_i)$ at zero. Namely, the left hand side in Equation (B.4.9) equals

$$\begin{aligned}
&\mathbb{E}_{Y,\sigma}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\Big(f\Big(\sum_{k\in\{1,\ldots,h\}}\pi(X_i^{(k)}), O_i\Big) - f\Big(0, O_i\Big) + f\Big(0, O_i\Big)\Big)Y_i\pi(V_i)\Big|\Big] \\
&\leq \underbrace{\mathbb{E}_{Y,\sigma}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\Big(f\Big(\sum_{k\in\{1,\ldots,h\}}\pi(X_i^{(k)}), O_i\Big) - f\Big(0, O_i\Big)\Big)Y_i\pi(V_i)\Big|\Big]}_{(1)} \\
&+ \underbrace{\mathbb{E}_{Y,\sigma}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i f\Big(0, O_i\Big)Y_i\pi(V_i)\Big|\Big]}_{(2)}
\end{aligned} \tag{B.4.10}$$

where the last inequality follows by the triangular inequality.

Decomposition with integral

We bound first (1). We write

$$\begin{aligned}
&\mathbb{E}_{Y,\sigma}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)Y_i\pi(V_i)\Big|\Big]\\
&=\mathbb{E}_{Y,\sigma}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)|Y_i|\mathrm{sign}(Y_i)\pi(V_i)\Big|\Big] \qquad \text{(B.4.11)}\\
&\leq 2\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)|Y_i|\pi(V_i)\Big|\Big]
\end{aligned}$$

where $\tilde{\sigma}_i$ are Rademacher random variables independent of $Y_1,\dots,Y_n$.[7] Since $|Y_i|>0$, we have

$$\begin{aligned}
\text{(D.3.7)} &= 2\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)\int_0^{\infty}1\{|Y_i|>y\}dy\pi(V_i)\Big|\Big]\\
&\leq 2\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\int_0^{\infty}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)1\{|Y_i|>y\}\pi(V_i)\Big|dy\Big]\\
&\leq 2\int_0^{\infty}\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)1\{|Y_i|>y\}\pi(V_i)\Big|\Big]dy
\end{aligned}$$

where the last inequality follows by the properties of the supremum function and Fubini theorem.

Lipschitz property

We decompose the supremum over $\pi\in\Pi$ as follows.

$$\begin{aligned}
&\int_0^{\infty}\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)1\{|Y_i|>y\}\pi(V_i)\Big|\Big]dy\\
&\leq\int_0^{\infty}\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy.
\end{aligned}$$

Let $\phi_i(N)=f\Big(N,O_i\Big)-f\Big(0,O_i\Big)$. Conditional on the data, ϕ_i is not random. By assumption it is Lipschitz and $\phi_i(0)=0$. By Lemma B.4.5,

$$\begin{aligned}
&\int_0^{\infty}\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(f\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)}),O_i\Big)-f\Big(0,O_i\Big)\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy\\
&\leq 2L\int_0^{\infty}\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\tilde{\sigma}_i\Big(\sum_{k\in\{1,\dots,h\}}\pi_2(X_i^{(k)})\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy. \qquad \text{(B.4.12)}
\end{aligned}$$

[7] To check the last claim the reader might consider that $P(\tilde{\sigma}_i=1|Y_i)=P(\sigma_i\mathrm{sign}(Y_i)=1|Y_i)=1/2$.

Supremum over Π

We decompose the supremum as follows:

$$L\int_0^\infty \mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i\Big(\sum_{k\in\{1,\cdots,h\}}\pi_2(X_i^{(k)})\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy$$
$$\leq \underbrace{L\int_0^\infty \mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi,\pi_2\in\Pi,\cdots,\pi_{h+1}\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i\Big(\sum_{k\in\{1,\ldots,h\}}\pi_{k+1}(X_i^{(k)})\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy}_{(J)}.$$

We re-parametrize the class of functions as follows.

$$(J)=$$
$$\int_0^\infty \mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi,\tilde{\pi}_2\in\Pi_2,\cdots,\tilde{\pi}_{h+1}\in\Pi_{h+1}}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i\Big(\sum_{k\in\{1,\ldots,h\}}\tilde{\pi}_{k+1}(X_i^{(1)},\cdots,X_i^{(h)})\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy$$

where $\Pi_2,...,\Pi_{h+1}$, are such that $\tilde{\pi}_j\in\Pi_j : \tilde{\pi}_j(X_i^{(1)},\cdots,X_i^{(k)},\cdots,X_i^{(h)}) = \pi(X_i^{(j)})$. By Theorem 29.4 in Devroye et al. (2013), $\mathrm{VC}(\Pi_j)=\mathrm{VC}(\Pi)$ for all $j\in\{1,...,h\}$.[8] Let

$$\tilde{\Pi}=\Big\{\pi_1\Big(\sum_{j=1}^h \pi_{j+1}\Big),\pi_k\in\Pi_k,k=1,\cdots,h+1\Big\}.$$

For any fixed point data point z_1^n, by Lemma B.4.2, the Dudley's integral of the function class $\tilde{\Pi}(z_1^n)$ is bounded by $C\sqrt{(h+1)\log(h+1)\mathrm{VC}(\Pi)}$, for a finite constant C. Therefore, by Lemma B.4.6 and B.4.2

$$\int_0^\infty \mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi_1\in\Pi_1,\pi_2\in\Pi_2,\ldots,\pi_{h+1}\in\Pi_{h+1}}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i\Big(\sum_{k\in\{1,\ldots,h\}}\pi_{k+1}(X_i^{(k)})\Big)1\{|Y_i|>y\}\pi_1(V_i)\Big|\Big]dy$$
$$\leq C\sqrt{B_1(h+1)\frac{\mathrm{VC}(\Pi)\log(h+1)}{n}}$$

for a constant $C<\infty$.

Remaining component

Next, we bound the term (2) in Equation (B.4.10). Following the same argument used for term (1), we have

$$\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i f\Big(0,O_i\Big)Y_i\pi(V_i)\Big|\Big]\leq 2\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i|f\Big(0,O_i\Big)Y_i|\pi(V_i)\Big|\Big]$$
$$\leq\int_0^\infty 2\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i 1\{|f\Big(0,O_i\Big)Y_i|>y\}\pi(V_i)\Big|\Big]dy.$$

[8] The reader might recognize that each $\pi_j\in\Pi_j$ can be written as the sum of π and functions constant at zero.

Since Π has finite VC-dimension, and it contains functions mapping to $\{0,1\}$, by Theorem 2.6.7 of Van Der Vaart and Wellner (1996) [9], $\int_0^2 \sqrt{\mathscr{N}_2(u,\Pi(z_1^n))}du < C\sqrt{\text{VC}(\Pi)}$ for a constant C. Hence, by Lemma B.4.6 and B.4.2,

$$\int_0^\infty 2\mathbb{E}_{Y,\tilde{\sigma}}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \tilde{\sigma}_i 1\{|f\Big(0,O_i\Big)Y_i| > y\}\pi(V_i)\Big|\Big]dy \leq c_0'\sqrt{\frac{B_2\text{VC}(\Pi)}{n}}$$

for a constant $c_0' < \infty$.

Conclusion

We are now left to discuss the case where h_i is different for each individual. In this case, we can assign to each unit covariates $\{X_i^{(1)},\cdots,X_i^{(h_i)},\emptyset,\emptyset,\cdots,\emptyset\}$ with in total $\bar{h}$ and parametrizing $\pi(\emptyset)=0$. The rest of the proof follows similarly as before. □

Lemmas for Identification

We conclude this section discussing the lemmas useful for identification.

Lemma B.4.8. *Let Assumption 2.2.1, 2.2.2, 2.2.4 hold. Then*

$$\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathscr{N}_i}D_k,\pi(X_i)=D_i\}}{e\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|\Big)}Y_i\Big|A,Z\Big] = m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|\Big).$$

Proof. Under Assumption 2.2.1, we can write

$$\begin{aligned}&\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathscr{N}_i}D_k,\pi(X_i)=D_i\}}{e\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|\Big)}Y_i\Big|A,Z\Big]\\&=\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathscr{N}_i}D_k,\pi(X_i)=D_i\}}{e\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|\Big)}r\Big(\pi(X_i),S_i(\pi),|\mathscr{N}_i|,Z_i,\varepsilon_i\Big)\Big|A,Z\Big].\end{aligned} \tag{B.4.13}$$

Under Assumption 2.2.2, we can then write

$$\begin{aligned}\text{(B.4.13)} = \mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathscr{N}_i}D_k,\pi(X_i)=D_i\}}{e\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|\Big)}|A,Z\Big]\times\\\times\mathbb{E}\Big[r\Big(\pi(X_i),S_i(\pi),|\mathscr{N}_i|,Z_i,\varepsilon_i\Big)\Big|A,Z\Big].\end{aligned}$$

[9] The argument is the same used in the proof of Lemma B.4.2.

By the first condition in Assumption 2.2.2 $\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,\pi(X_i)=D_i\}}{e\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}|A,Z\Big]=1$. Finally, by Assumption 2.2.2 and Assumption 2.2.4

$$\mathbb{E}\Big[r\Big(\pi(X_i),S_i(\pi),|\mathcal{N}_i|,Z_i,\varepsilon_i\Big)\Big|A,Z\Big]=m\Big(\pi(X_i),S_i(\pi),|\mathcal{N}_i|,Z_i,\Big),$$

where $m(\cdot)$ not being individual specific follows by Assumption 2.2.4. □

Lemma B.4.9. *Let $S_i(\pi)=\sum_{k\in\mathcal{N}_i}\pi(X_k)$. Let Assumption 2.2.1, 2.2.2, 2.2.4 hold. Then*

$$\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,d=D_i\}}{e^c\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}\Big(Y_i-m^c\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big)\Big)+m^c\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big)\Big|A,Z\Big]$$
$$=m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big),$$

if either $e^c=e$ or (and) $m^c=m$.

Proof. Whenever $e^c=e$, similarly to what discussed in Lemma B.4.8, we obtain that

$$\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,d=D_i\}}{e^c\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}Y_i\Big|A,Z\Big]=m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big).$$

The remaining component depending on m^c cancels out similarly as above. Let now $m^c=m$. Then

$$\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,d=D_i\}}{e^c\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}\Big(Y_i-m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big)\Big)\Big|A,Z\Big]$$
$$=\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,d=D_i\}}{e^c\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}\Big(r\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|,\varepsilon_i\Big)-m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big)\Big)\Big|A,Z\Big]$$
$$=\mathbb{E}\Big[\frac{1\{S_i(\pi)=\sum_{k\in\mathcal{N}_i}D_k,d=D_i\}}{e^c\Big(\pi(X_i),S_i(\pi),Z_{k\in\mathcal{N}_i},Z_i,|\mathcal{N}_i|\Big)}\Big|A,Z\Big]\times$$
$$\times\mathbb{E}\Big[\Big(r\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|,\varepsilon_i\Big)-m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big)\Big)\Big|A,Z\Big]=0.$$

□

Proofs of Lemmas in the Main Text

In this section we prove the lemmas in the main text.

Proof of Lemma 2.2.1. Under Assumption 2.2.3, the welfare on the target sample reads as follows

$$W(\pi) = \frac{1}{|\mathscr{I}|} \sum_{j \in \mathscr{I}} \mathbb{E}\Big[r\Big(\pi(X_j), \sum_{k \in \mathscr{N}_j} \pi(X_k), Z_j, |\mathscr{N}_j|, \varepsilon_j \Big) \Big].$$

Under the distributional condition in the statement of the lemma we have

$$W(\pi) = \mathbb{E}\Big[r\Big(\pi(X_j), \sum_{k \in \mathscr{N}_j} \pi(X_k), Z_j, |\mathscr{N}_j|, \varepsilon_j \Big) \Big]$$

is constant in j. In addition, by Assumption 2.2.1 and 2.2.2, welfare on the sampled units takes the expression

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\Big[r\Big(\pi(X_i), \sum_{k \in \mathscr{N}_i} \pi(X_k), Z_i, |\mathscr{N}_i|, \varepsilon_i \Big) \Big].$$

The proof completes again by the distributional assumption in the statement of the lemma. □

Proof of Lemma 2.2.2. By the law of iterated expectations, we write

$$\mathbb{E}\Big[r(\pi(X_i), \sum_{k \in \mathscr{N}_i} \pi(X_k), Z_i, |\mathscr{N}_i|, \varepsilon_i) \Big] = \mathbb{E}\Big[\mathbb{E}\Big[r(\pi(X_i), \sum_{k \in \mathscr{N}_i} \pi(X_k), Z_i, |\mathscr{N}_i|, \varepsilon_i) \Big| Z_i, \mathscr{N}_i, X_{k \in \mathscr{N}_i} \Big] \Big]$$

Under Assumption 2.2.2, 2.2.4 we obtain

$$\mathbb{E}\Big[\mathbb{E}\Big[r(\pi(X_i), \sum_{k \in \mathscr{N}_i} \pi(X_k), Z_i, |\mathscr{N}_i|, \varepsilon_i) \Big| Z_i, \mathscr{N}_i, X_{k \in \mathscr{N}_i} \Big] \Big] = \mathbb{E}\Big[m(\pi(X_i), \sum_{k \in \mathscr{N}_i} \pi(X_k), Z_i, |\mathscr{N}_i|) \Big].$$

The rest of the proof follows from Lemma B.4.8 and the law of iterated expectations. □

Regret Bounds

Theorem B.5.1. *Suppose the conditions in Theorem 2.3.1 holds. Then,*

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|W(\pi)-W_n(\hat{\pi}^{aipw}_{m^c,e^c})|\Big]$$
$$\leq |\mathscr{L}|^2(\Gamma_1+\Gamma_2)\frac{\bar{C}}{\delta}\Big(L+\frac{L}{\delta_0^{\sqrt{3}}}+\frac{1}{\delta_0^{\sqrt{3}}}\Big)\mathbb{E}[\mathscr{N}_{n,M}^{M}\mathscr{N}_n^{|\mathscr{L}|}\mathscr{N}_n^{3/2}\sqrt{\log(\mathscr{N}_{n,1})}]^2\sqrt{\frac{\mathrm{VC}(\Pi)}{n}}+\mathscr{K}_\Pi(n),$$

for a finite constant $\bar{C}<\infty$.

Proof of Theorem B.5.1. Let $S_i(\pi)=\sum_{k\in\mathscr{N}_i}\pi(X_k)$. We denote $L=(L_j)_{j=1}^E$ under Assumption 2.3.2. Recall under Assumption 2.3.2 that we can decompose covariates Z_i into two components, one L_i having arbitrary dependence and endogenous with respect to the network but with finite support, and the other, Q_i having local dependence but arbitrary support.

Preliminaries

Notice first that by Definition 2.2.1,

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|W(\pi)-W_n(\pi)|\Big]\leq\mathscr{K}_\Pi(n)+\mathbb{E}\Big[\sup_{\pi\in\Pi}|W^*(\pi)-W_n(\pi)|\Big]$$

where

$$W^*(\pi)=\frac{1}{n}\sum_{i=1}^n\mathbb{E}\Big[r\Big(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|,\varepsilon_i\Big)\Big]=\frac{1}{n}\sum_{i=1}^n\mathbb{E}\Big[m\Big(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|\Big)\Big],$$

and where the second equality follows from Assumption 2.2.1, 2.2.2 and 2.2.4 similarly to what discussed in the proof of Lemma B.4.9. By Lemma B.4.9, we have that $W^*(\pi)=\mathbb{E}[W_n(\pi,m^c,e^c)]$ under correct specification of either m^c or e^c. Therefore for $i\in\{1,\cdots,n\}$

$$W^*(\pi)=\frac{1}{n}\sum_{i=1}^n\mathbb{E}\Big[\frac{1\{\pi(X_i)=D_i,\sum_{k\in\mathscr{N}_i}\pi(X_k)=\sum_{k\in\mathscr{N}_i}D_k\}}{e^c\Big(\pi(X_i),S_i(\pi),O_i\Big)}\Big(Y_i-m^c\Big(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|\Big)\Big)\Big]+$$
$$+\mathbb{E}\Big[m^c\Big(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|\Big)\Big],$$

where $O_i=\Big(Z_{k\in\mathscr{N}_i},Z_i,|\mathscr{N}_i|\Big)$.

Decomposition

Using the triangular inequality, we decompose the supremum of the empirical process into three terms as follows.

$$\begin{aligned}\mathbb{E}\Big[\sup_{\pi\in\Pi}|W^*(\pi)-W_n(\pi,m^c,e^c)|\Big] \leq & \underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-\mathbb{E}[T_1(\pi,m^c,e^c)]\Big|\Big]}_{(A)} \\ & + \underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_2(\pi,m^c,e^c)-\mathbb{E}[T_2(\pi,m^c,e^c)]\Big|\Big]}_{(B)} \\ & + \underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_3(\pi,m^c)-\mathbb{E}[T_3(\pi,m^c)]\Big|\Big]}_{(C)},\end{aligned}$$

where

$$\begin{aligned}T_1(\pi,m^c,e^c) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1\{\pi(X_i)=D_i,\sum_{k\in\mathcal{N}_i}\pi(X_k)=\sum_{k\in\mathcal{N}_i}D_k\}}{e^c\Big(\pi(X_i),S_i(\pi),O_i\Big)}Y_i \\ T_2(\pi,m^c,e^c) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1\{\pi(X_i)=D_i,\sum_{k\in\mathcal{N}_i}\pi(X_k)=\sum_{k\in\mathcal{N}_i}D_k\}}{e^c\Big(\pi(X_i),S_i(\pi),O_i\Big)}m^c\Big(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k),Z_i,|\mathcal{N}_i|\Big) \\ T_3(\pi,m^c) &= \frac{1}{n}\sum_{i=1}^{n}m^c\Big(\pi(X_i),S_i(\pi),Z_i,|\mathcal{N}_i|\Big).\end{aligned} \tag{B.5.1}$$

The rest of the proof consists in providing upper bounds for (A), (B), and (C). We discuss (A) first.

Preparation for (A)

We create an independent copy of $(T_1(\cdot),A,L)$, defined as $(T_1'(\cdot),A',L')$ with the same distribution of $(T_1(\cdot),A,L)$. Using Jensen's inequality, we can write

$$\begin{aligned}\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-\mathbb{E}[T_1(\pi,m^c,e^c)]\Big|\Big] &= \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-\mathbb{E}[T_1'(\pi,m^c,e^c)]\Big|\Big] \\ &\leq \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-T_1'(\pi,m^c,e^c)\Big|\Big].\end{aligned}$$

We now use the law of iterated expectations to write

$$\begin{aligned}&\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-\mathbb{E}[T_1(\pi,m^c,e^c)]\Big|\Big]\\&\leq\mathbb{E}\Big[\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-T_1'(\pi,m^c,e^c)\Big|\Big|A,L,A',L'\Big]\Big].\end{aligned}$$

Denote each summand corresponding to T_1 as U_i and each summand corresponding to T_1' as U_i'. That is,

$$U_i=\frac{1\{\pi(X_i)=D_i,\sum_{k\in\mathscr{N}_i}\pi(X_k)=\sum_{k\in\mathscr{N}_i}D_k\}}{e^c\Big(\pi(X_i),S_i(\pi),O_i\Big)}Y_i.$$

Observe that $\Big[(U_i)_{i=1}^n,A,L\Big]$ and $\Big[(U_i')_{i'=1}^n,A',L'\Big]$ are independent by construction. Since we condition on (A,L), (A',L'), we can write

$$U_i\Big(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi\Big)$$

as an explicit function of $L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i}$ (and π) conditional on (A,L) and similarly $U_i'(L_i',|\mathscr{N}_i|',L_{k\in\mathscr{N}_i}',\pi)$. Conditional on $(A,L),(A',L')$, the marginal distribution of U_i,U_i' depends on (A,L,A',L') only through the individual, neighbors' covariates and number of neighbors under Assumption 2.2.2, Assumption 2.2.4.[10] Also, observe $U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)$ is exchangeable in $L_{k\in\mathscr{N}_i}$.[11] Define $\mathscr{N}_n'$ the maximum degree under A' and write $R_n=\max\{\mathscr{N}_n,\mathscr{N}_n'\}$. Define $\tilde{\mathscr{L}}_m\subset\mathscr{L}^m$ the space of vectors of dimension m with each component taking value in $\mathscr{L}$ after removing exchangable duplicates, where two vectors are duplicates if they are equal up-to

[10]To observe why, note that conditional on (L,A), U_i is just a function of $(L_i,Q_i,L_{k\in\mathscr{N}_i},Q_{k\in\mathscr{N}_i},|\mathscr{N}_i|,\varepsilon_i,\varepsilon_{D_i},\varepsilon_{D_{k\in\mathscr{N}_i}})$, where the distribution of $(\varepsilon_i,\varepsilon_{D_i},\varepsilon_{D_{k\in\mathscr{N}_i}})$ given (A,L) only depends on $(L_i,L_{k\in\mathscr{N}_i},|\mathscr{N}_i|)$ since ε_{D_i} are *i.i.d.* and exogenous.

[11]This can be observed from Definition 2.2.3 and Assumption 2.2.2 for the treatment assignments.

exchanging the entries of the vectors. We can then write

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|T_1(\pi,m^c,e^c)-T_1'(\pi,m^c,e^c)\Big|\Big|A,Z,A',Z'\Big]\\
&=\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)-\frac{1}{n}\sum_{i=1}^{n}U_i'(L_i',|\mathscr{N}_i|',L'_{k\in\mathscr{N}_i},\pi)\Big|\Big|A,L,A',L'\Big]\\
&\leq\sum_{m,m'\leq R_n}\sum_{l\in\mathscr{L}\times\tilde{\mathscr{L}}_m,l'\in\mathscr{L}\times\tilde{\mathscr{L}}_{m'}}\\
&\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big[U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)-U_i'(L_i',|\mathscr{N}_i|',L'_{k\in\mathscr{N}_i},\pi)\Big]I_i(l,l',m,m')\Big|\Big|A,L,A',L'\Big]
\end{aligned}
\tag{B.5.2}
$$

where $I_i(l,l',m,m')=1\{(L_i,L_{k\in\mathscr{N}_i})=l,(L_i',L'_{k\in\mathscr{N}_i})=l',m=|\mathscr{N}_i|,m'=|\mathscr{N}_i|'\}$. Finally, define $\mathscr{C}_n^M$ the proper cover of the adjacency matrix W_n where two elements are connected if they are neighbors of M degree either (or both) under A or A'. We can then write

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big[U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)-U_i'(L_i',|\mathscr{N}_i|',L'_{k\in\mathscr{N}_i},\pi)\Big]I_i(l,l',m,m')\Big|\Big|A,L,A',L'\Big]\\
&\leq\sum_{\mathscr{C}_n^M(j)\in\mathscr{C}_n^M}\\
&\underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{C}_n^M(j)}\Big[U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)-U_i'(L_i',|\mathscr{N}_i|',L'_{k\in\mathscr{N}_i},\pi)\Big]I_i(l,l',m,m')\Big|\Big|A,L,A',L'\Big]}_{(i)}.
\end{aligned}
$$

The goal is to bound (i).

Analysis of (i)

Observe first that each summand $U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)$ is a function of $\Big(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\varepsilon_i,Q_i,\varepsilon_{D_i},Q_{k\in\mathscr{N}_i},\varepsilon_{D_{k\in\mathscr{N}_i}},\pi\Big)$ only. In addition, by construction of $\mathscr{C}_n^M$, each summand is mutually independent with the others conditional on (A,L,A',L') by Assumption 2.2.5 and the first condition in Assumption 2.2.2 and Assumption 2.3.2. Finally, each non-zero summand is evaluated at the same values of individual, neighbors' covariates L_i and number of neighbors for both U_i,U_i'.

Symmetrization

Under the above observation, we note that each summand whose indicator I_i is non-zero in (i) is $i.i.d.$ conditional on A, L, A', L' under 2.2.5, 2.3.2 the fact that ε_{D_i} are $i.i.d.$ and exogenous (Assumption 2.2.2), since each summand is evaluated at the same values of $(L_i, L_{k\in\mathscr{N}_i}, |\mathscr{N}_i|)$. The remaining summands are instead equal to zero almost surely conditional on (A, L, A', L'). Hence, using a standard symmetrization argument, we bound (i) as follows

$$\begin{aligned}(i) &= \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{C}_n^M(j)}\sigma_i\Big[U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)-U_i'(L_i',|\mathscr{N}_i|',L_{k\in\mathscr{N}_i}',\pi)\Big]I_i(l,l',m,m')\Big|\Big|A,L,A',L'\Big]\\ &\leq \underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{C}_n^M(j)}\sigma_i U_i(L_i,|\mathscr{N}_i|,L_{k\in\mathscr{N}_i},\pi)I_i(l,l',m,m')\Big|\Big|A,L,A',L'\Big]}_{(j)}\\ &+\underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{C}_n^M(j)}\sigma_i U_i'(L_i',|\mathscr{N}_i|',L_{k\in\mathscr{N}_i}',\pi)I_i(l,l',m,m')\Big|\Big|A,L,A',L'\Big]}_{(jj)}\end{aligned} \tag{B.5.3}$$

where σ_i are $i.i.d.$ Rademacher random variables. Note here that a key step is to have $i.i.d.$ of all non-zero components. For those components that are zero, the value of σ_i is not relevant conditional on (A, L, A', L'), and hence we can just use σ_i to be all $i.i.d.$. We then bound (j) (and symmetrically (jj)) with Lemma B.4.7. To invoke Lemma B.4.7 we observe that under Assumption 2.3.1, by Lemma B.4.4

$$\phi_i(d,s) := \frac{1\{d = D_i, s = \sum_{k\in\mathscr{N}_i} D_k\}}{e^c\Big(d, N, O_i\Big)}$$

is $\frac{2}{\delta}$-Lipschitz in s conditional on the data. In addition, the function is bounded by Assumption 2.3.1, therefore

$$\mathbb{E}\Big[Y_i^3\frac{1\{\pi(X_i) = D_i, 0 = \sum_{k\in\mathscr{N}_i} D_k\}}{e^c\Big(\pi(X_i), 0, O_i\Big)^3}\Big|A,L\Big] < \frac{\Gamma_1^2}{\delta_0^3} < \infty,$$

since $\sigma(L)\subseteq\sigma(Z)$ Since Y_i is independent of (A', L') the moment bound also holds conditionally on (A', L'). Hence we can invoke Lemma B.4.7 for each term (j) and (jj) symmetrically.

Collecting the outer sum for (A)

Using a symmetric argument for each of the two terms in Equation (B.5.3), we can write by Lemma B.4.7 and collecting the outer sums in Equation (B.5.2)

$$(A) \leq 2|\mathscr{L}|^2 \bar{C}\mathbb{E}\Big[|\tilde{\mathscr{L}}_{\mathscr{N}_n}|^2 R_n^2 |\mathscr{C}_n^M| \frac{\Gamma_1 L}{\delta \delta_0^{3/2}} \sqrt{\frac{\mathscr{N}_n \log(\mathscr{N}_n)}{n}}\Big],$$

for a universal constant $\bar{C}$.

Bounding the constants to obtain a polynomial bound in the degree

We can bound $R_n \leq \mathscr{N}_n + \mathscr{N}_n'$, and $|\mathscr{C}_n^M| \leq \mathscr{N}_E^M \times \mathscr{N}_E^{M'}$ (see Lemma B.4.1). Note that $\mathscr{N}_E^M, \mathscr{N}_E^{M'}$ are independent. To characterize $|\tilde{\mathscr{L}}_{\mathscr{N}_n}|$, we can assume, since $|\mathscr{L}|$ is fixed with n that $\mathscr{N}_n > |\mathscr{L}|$. Observe that we can find at most $\mathscr{N}_n^{|\mathscr{L}|}$ many exchangeable combinations. To observe why, note that the number of exchangeable combinations is less or equal than the the number of distinct values that the vector $[V_1, \cdots, V_{|\mathscr{L}|}]$ can take with $V_1 \in \{1, \cdots, \mathscr{N}_n\}$.[12] The number of entries is of order $\mathscr{N}_n^{|\mathscr{L}|}$.

Conclusion

Finally, observe that the same reasoning directly applies also to $(B), (C)$, completing the proof.

□

Theorem 2.3.1 and Theorem 2.3.5

We state these two theorems as corollaries of Theorem B.5.1.

Corollary 12. *Theorem 2.3.1 holds.*

[12] Namely, we can represent the vector L_i as a vector of binary dummies. Due to exchangeability we are interested in the number of combinations that the sum of $\mathscr{N}_n$ of such vectors can take.

Proof. Following Kitagawa and Tetenov (2018),

$$\begin{aligned}\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}_{m^c,e^c})\Big] &= \mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W_n(\hat{\pi}_{m^c,e^c}, m^c, e^c) + W_n(\hat{\pi}_{m^c,e^c}, m^c, e^c) - W(\hat{\pi}_{m^c,e^c})\Big] \\ &\leq \mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W_n(\pi, m^c, e^c) + W_n(\hat{\pi}_{m^c,e^c}, m^c, e^c) - W(\hat{\pi}_{m^c,e^c})\Big] \\ &\leq \mathbb{E}\Big[2\sup_{\pi\in\Pi} |W(\pi) - W_n(\pi, m^c, e^c)|\Big]\end{aligned}$$

where the last inequality follows by the triangular inequality. The bound on the right hand side is given by Theorem B.5.1. □

Corollary 13. *Theorem 2.3.5 hold.*

Proof. Let $O_i = \Big(Z_{k\in\mathscr{N}_i}, Z_i, |\mathscr{N}_i|\Big)$ and $I_i(\pi) = 1\{D_i = \pi(X_i), \sum_{k\in\mathscr{N}_i} D_k = S_i(\pi)\}$. We have

$$\begin{aligned}\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}_{\hat{m},\hat{e}})\Big] \leq \underbrace{2\,\mathbb{E}\Big[\sup_{\pi\in\Pi} |W^{aipw}_n(\pi, m^c, e^c) - W(\pi)|\Big]}_{(I)} \\ + \underbrace{2\,\mathbb{E}\Big[\sup_{\pi\in\Pi} |W^{aipw}_n(\pi, \hat{m}, \hat{e}) - W_n(\pi, m^c, e^c)|\Big]}_{(II)}.\end{aligned}$$

Term (I) is bounded by Theorem B.5.1.

Bounding (II)

We now study (II). We separate two components of (II).

$$\begin{aligned}(II) = \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n \frac{I_i(\pi)}{\hat{e}(\pi(X_i), S_i(\pi), O_i)}\Big(Y_i - \hat{m}(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|\Big) \\ + \frac{1}{n}\sum_{i=1}^n \Big(\hat{m}(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|\Big) - m^c(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|\Big)\Big) \\ - \frac{1}{n}\sum_{i=1}^n \frac{I_i(\pi)}{e^c(\pi(X_i), S_i(\pi), O_i)}\Big(Y_i - m^c(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|\Big)\Big|\Big].\end{aligned} \tag{B.5.4}$$

We can now use the triangular inequality.

$$
\begin{aligned}
\text{(B.5.4)} \leq\ & \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{I_i(\pi)}{e^c(\pi(X_i),S_i(\pi),O_i)\hat{e}(\pi(X_i),S_i(\pi),O_i)}(Y_i-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|))\times \\
& \times\Big(e^c(\pi(X_i),S_i(\pi),O_i)-\hat{e}(\pi(X_i),S_i(\pi),O_i)\Big)\Big|\Big] \\
& +\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\hat{m}(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|\Big)-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)\Big)\Big|\Big] \\
& +\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{I_i(\pi)}{\hat{e}(\pi(X_i),S_i(\pi),O_i)}(\hat{m}(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|))\Big|\Big].
\end{aligned}
$$

By Assumption 2.3.1, and Holder's inequality we have that

$$
\begin{aligned}
& \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{I_i(\pi)}{e^c(\pi(X_i),S_i(\pi),O_i)\hat{e}(\pi(X_i),S_i(\pi),O_i)}(Y_i-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)) \\
& \times\Big(e^c(\pi(X_i),S_i(\pi),O_i)-\hat{e}(\pi(X_i),S_i(\pi),O_i)\Big)\Big|\Big] \\
& \leq \frac{1}{\delta^2}\mathbb{E}\Big[\sup_{\pi\in\Pi}\frac{1}{n}\sum_{i=1}^{n}\Big|(Y_i-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|))\times\Big(e^c(\pi(X_i),S_i(\pi),O_i)-\hat{e}(\pi(X_i),S_i(\pi),O_i)\Big|\Big)\Big].
\end{aligned}
\tag{B.5.5}
$$

Similarly,

$$
\begin{aligned}
& \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{I_i(\pi)}{\hat{e}(\pi(X_i),S_i(\pi),O_i)}\Big(\hat{m}(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)\Big)\Big|\Big] \\
& \leq \frac{1}{\delta}\mathbb{E}\Big[\sup_{\pi\in\Pi}\frac{1}{n}\sum_{i=1}^{n}\Big|\Big(\hat{m}(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)-m^c(\pi(X_i),S_i(\pi),Z_i,|\mathscr{N}_i|)\Big)\Big|\Big].
\end{aligned}
$$

Now note that for the conditional mean we have that the following expression

$$
\sup_{\pi\in\Pi}\Big|\hat{m}(\pi(X_i),\sum_{k\in\mathscr{N}_i}\pi(X_k),Z_i,|\mathscr{N}_i|)-m^c(\pi(X_i),\sum_{k\in\mathscr{N}_i}\pi(X_k),Z_i,|\mathscr{N}_i|)\Big|.
$$

is bounded by Holder's inequality by

$$
\begin{aligned}
& \Big|\sum_{h=1}^{|\mathscr{N}_i|}(m^c(\pi(X_i),h,Z_i,|\mathscr{N}_i|)-\hat{m}(\pi(X_i),h,Z_i,|\mathscr{N}_i|))1\{\sum_{k\in\mathscr{N}_i}\pi(X_k)=h\}\Big| \\
& \leq \max_{s\leq|\mathscr{N}_i|}|m^c(\pi(X_i),s,Z_i,|\mathscr{N}_i|)-\hat{m}(\pi(X_i),s,Z_i,|\mathscr{N}_i|)|,
\end{aligned}
$$

since $\sum_{h=1}^{|\mathscr{N}_i|}1\{\sum_{k\in\mathscr{N}_i}\pi(X_k)=h\}=1$. A similar bound applies to Equation (B.5.5). By Assumption 2.3.4 the proof completes. □

Proof of Theorem 2.3.3

The proof follows a similar strategy of Theorem B.5.1.

Studying the expectation

Define $I_i(\pi) = 1\Big\{D_i = \pi(X_i), \sum_{k\in\mathscr{N}_i} D_k = \sum_{k\in\mathscr{N}_i}\pi(X_k)\Big\}$. First, by Lemma 2.2.1 $\mathscr{K}_\Pi(n) = 0$, and by the distributional assumption in Lemma 2.2.1 we have

$$W(\pi) = \frac{1}{|\mathscr{I}|}\sum_{j\in\mathscr{I}}\mathbb{E}\Big[r\Big(\pi(X_j), \sum_{k\in\mathscr{N}_j}\pi(X_k), Z_j, |\mathscr{N}_j|, \varepsilon_j\Big)\Big] = \mathbb{E}\Big[m(\pi(X_i), \sum_{k\in\mathscr{N}_i}\pi(X_k), Z_i, |\mathscr{N}_i|)\Big]$$

for all $i \in \{1,\cdots,E\}$. Following the same proof as in Lemma B.4.8 and by the definition of conditional expectation

$$\begin{aligned}
&\mathbb{E}\Big[m\Big(\pi(X_i), \sum_{k\in\mathscr{N}_i}\pi(X_k), Z_i, |\mathscr{N}_i|\Big)\Big||\mathscr{N}_i| \le \kappa_n\Big]P(|\mathscr{N}_i| \le \kappa_n) \\
&= \mathbb{E}\Big[Y_i\frac{1\{D_i = \pi(X_i), \sum_{k\in\mathscr{N}_i} D_k = \sum_{k\in\mathscr{N}_i}\pi(X_k)\}}{e(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|)}1\{|\mathscr{N}_i| \le \kappa_n\}\Big].
\end{aligned} \tag{B.5.6}$$

Bounding the regret with two main terms

We can then write

$$\begin{aligned}
&\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{tr}) \le 2\sup_{\pi\in\Pi}\Big|W(\pi) - W_n^{tr}(\pi)\Big| \\
&= 2\sup_{\pi\in\Pi}\Big|W(\pi) - \frac{1}{n}\sum_{i=1}^n Y_i\frac{I_i(\pi)}{e(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|)}1\{|\mathscr{N}_i| \le \kappa_n\}\Big| \\
&\le \underbrace{2\sup_{\pi\in\Pi}\Big|\mathbb{E}\Big[m(\pi(X_i), \sum_{k\in\mathscr{N}_i}\pi(X_k), Z_i, |\mathscr{N}_i|)\Big||\mathscr{N}_i| \le \kappa_n\Big]P(|\mathscr{N}_i| \le \kappa_n) - \frac{1}{n}\sum_{i=1}^n Y_i\frac{I_i(\pi)1\{|\mathscr{N}_i| \le \kappa_n\}}{e(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|)}\Big|}_{(A)} \\
&+ 2BP(|\mathscr{N}_i| \ge \kappa_n),
\end{aligned}$$

where we used the fact that the outcome is uniformly bounded by B.

Bounding (A) and symmetrization

The rest of the proof follows similarly to the proof of Theorem B.5.1. Namely, we bound the above term using the symmetrization argument as in Equation (B.5.2) and below. Therefore,

we can show using the same argument in Theorem B.5.1 that

$$\mathbb{E}[(A)] \leq 2\bar{C}\mathbb{E}\Big[\sum_{m,m' \leq R_n} \sum_{l \in \mathscr{L} \times \tilde{\mathscr{L}}_m, l' \in \mathscr{L} \times \tilde{\mathscr{L}}_{m'}} \sum_{\mathscr{C}_n^M(j) \in \mathscr{C}_n^M} \underbrace{\mathbb{E}\Big[\sup_{\pi \in \Pi} \Big| \frac{1}{n} \sum_{i \in \mathscr{C}_n^M(j)} \sigma_i U_i(L_i, |\mathscr{N}_i|, L_{k \in \mathscr{N}_i}, \pi) I_i(l, l', m, m') \Big| | A, L, A', L' \Big]}_{(A)} \Big]$$

with the above quantities $\Big(I_i(l,l',m,m'), R_n, \mathscr{C}_n^M, A', L', (\sigma_i)_{i=1}^n, \tilde{\mathscr{L}}_m \Big)$ defined above and below Equation (B.5.2) and $U_i(L_i, |\mathscr{N}_i|, L_{k \in \mathscr{N}_i}, \pi)$ denoting random variables equal to the summands $Y_i \frac{I_i(\pi) 1\{|\mathscr{N}_i| \leq \kappa_n\}}{e(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|)}$. Such random variables are measurable with respect to $(\varepsilon_i, Q_i, Q_{k \in \mathscr{N}_i}, \varepsilon_{D_i}, \varepsilon_{D_{k \in \mathscr{N}_i}}, L_i, L_{k \in \mathscr{N}_i}, |\mathscr{N}_i|)$.

Rademacher complexity bound and conclusion

We then bound the Rademacher complexity obtained from the argument as in Theorem B.5.1 with Lemma B.4.3. In particular we observe that since Y_i is bounded we have that

$$Y_i \frac{I_i(\pi) 1\{|\mathscr{N}_i| \leq \kappa_n\}}{e(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|)}$$

is a Lipschitz function of $S_i(\pi)$ with constant $2B/\delta$, since $S_i(\pi)$ takes discretely many values. The bound on (A) follows by Lemma B.4.3. The outer sums are bounded as discussed at the end of the proof of Theorem B.5.1.

Proof of Theorem 2.3.2

Notation

Throughout the proof we define

$$
\begin{aligned}
&I_i(\pi) = 1\Big\{\pi(X_i) = D_i, \sum_{k\in\mathcal{N}_i}\pi(X_k) = \sum_{k\in\mathcal{N}_i} D_k\Big\}, \quad \tilde{I}_i(d,s) = 1\Big\{d = D_i, s = \sum_{k\in\mathcal{N}_i} D_k\Big\},\\
&e_i(\pi) = e\Big(\pi(X_i), \sum_{k\in\mathcal{N}_i}\pi(X_k), Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|\Big), \quad \tilde{e}_i(d,s) = e\Big(d, s, Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|\Big)\\
&m_i(\pi) = m\Big(\pi(X_i), \sum_{k\in\mathcal{N}_i}\pi(X_k), Z_i, |\mathcal{N}_i|\Big), \quad \tilde{m}_i(d,s) = m\Big(d, s, Z_i, |\mathcal{N}_i|\Big).
\end{aligned}
\tag{B.5.7}
$$

We also denote $\tilde{\varepsilon}_i = Y_i - m(\pi(X_i), \sum_{k\in\mathcal{N}_i}\pi(X_k), Z_i, |\mathcal{N}_i|)$. Whenever we use $\hat{e}, \hat{m}$ we refer to the estimated nuisances. Recall that Z denotes the matrix of covariates for all individuals $i \in \{1, \cdots, E\}$ in the network of participants.

Preliminaries

By Lemma B.7.3, we have

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}_{\hat{m},\hat{e}})\Big] \leq\ & \underbrace{2\,\mathbb{E}\Big[\sup_{\pi\in\Pi} |W^{aipw}_n(\pi, m^c, e^c) - W(\pi)|\Big]}_{(I)}\\
&+ \underbrace{2\,\mathbb{E}\Big[\sup_{\pi\in\Pi} |W^{aipw}_n(\pi, \hat{m}, \hat{e}) - W_n(\pi, m^c, e^c)|\Big]}_{(II)}.
\end{aligned}
$$

Term (I) is bounded by Theorem B.5.1. We now study (II).

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{\pi\in\Pi} |W^{aipw}_n(\pi, \hat{m}, \hat{e}) - W_n(\pi, m^c, e^c)|\Big]\\
&= \mathbb{E}\Big[\sup_{\pi\in\Pi} |\frac{1}{n}\sum_{i=1}^{n} \frac{I_i(\pi)}{\hat{e}_i(\pi)}(m_i(\pi) - \hat{m}_i(\pi)) + \varepsilon_i\frac{I_i(\pi)}{\hat{e}_i(\pi)} + \hat{m}_i(\pi) - m_i(\pi)|\Big]\\
&= \mathbb{E}\Big[\sup_{\pi\in\Pi} |\frac{1}{n}\sum_{i=1}^{n} \Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)} - \frac{I_i(\pi)}{e_i(\pi)}\Big)(m_i(\pi) - \hat{m}_i(\pi)) + \tilde{\varepsilon}_i\frac{I_i(\pi)}{\hat{e}_i(\pi)} + \Big(\frac{I_i(\pi)}{e_i(\pi)} - 1\Big)\hat{m}_i(\pi) - m_i(\pi)|\Big].
\end{aligned}
\tag{B.5.8}
$$

The last equality follows after adding and subctracting the component $\frac{I_i(\pi)}{e_i(\pi)}(m_i(\pi) -$

$\hat{m}_i(\pi)$). Observe now that we can bound the above component as follows.

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{\pi\in\Pi}|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big)(m_i(\pi)-\hat{m}_i(\pi))+\tilde{\varepsilon}_i\frac{I_i(\pi)}{\hat{e}_i(\pi)}+\Big(\frac{I_i(\pi)}{e_i(\pi)}-1\Big)\hat{m}_i(\pi)-m_i(\pi)|\Big] \\
\leq&\underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big)(m_i(\pi)-\hat{m}_i(\pi))\Big|\Big]}_{(i)}+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\underbrace{\tilde{\varepsilon}_i\Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big)}_{(ii)}\Big|\Big] \\
&+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\underbrace{\tilde{\varepsilon}_i\frac{I_i(\pi)}{e_i(\pi)}}_{(iii)}\Big|\Big]+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\underbrace{\Big(\frac{I_i(\pi)}{e_i(\pi)}-1\Big)(\hat{m}_i(\pi)-m_i(\pi))}_{(iv)}\Big|\Big].
\end{aligned} \tag{B.5.9}$$

The first step consists in showing that the terms (ii), (iii), (iv) are centered around zero. (iii) directly follows since none of the nuisance functions is estimated. We therefore discuss the expectation of (ii) and (iv).

Expectation of (ii)

We start from (ii). By Algorithm 1, we claim that

$$\mathbb{E}\Big[\tilde{\varepsilon}_i\Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big)\Big]=0. \tag{B.5.10}$$

This follows from the following argument: using the law of iterated expectations

$$E[(ii)]=\mathbb{E}\Big[\mathbb{E}\Big[\Big(r(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k),Z_i,|\mathcal{N}_i|,\varepsilon_i)-m_i(\pi)\Big)\Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big)\Big|\hat{e}_i,Z,A\Big]\Big]. \tag{B.5.11}$$

Observe now that $\hat{e}_i(\pi)$ is estimated on an independent sample from ε_i, conditional on (A,Z). Under Assumption 2.2.5 and the first condition in Assumption 2.2.2[13],

$$\begin{aligned}
&\mathbb{E}\Big[r(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k),Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big(\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big)\Big|\hat{e}_i,Z,A\Big]= \\
&m(\pi(X_i),\sum_{k\in\mathcal{N}_i}\pi(X_k),Z_i,|\mathcal{N}_i|)\mathbb{E}\Big[\frac{I_i(\pi)}{\hat{e}_i(\pi)}-\frac{I_i(\pi)}{e_i(\pi)}\Big|\hat{e}_i,Z,A\Big].
\end{aligned} \tag{B.5.12}$$

[13]Namely ε_{D_i} are $i.i.d.$ and exogenous.

Expectation of (iii)

Notice now that the same reasoning for (ii) directly applies to (iii) for the fact that

$$\mathbb{E}\Big[\frac{I_i(\pi)}{e_i(\pi)} - 1|\hat{m}_i, A, Z\Big] = \mathbb{E}\Big[\frac{I_i(\pi)}{e_i(\pi)} - 1|A, Z\Big] = 0$$

by the cross fitting argument.

Concentration bound for $(ii), (iii), (iv)$

We can therefore apply the same argument used for Theorem B.5.1 also to these three components directly (see the argument on Equation (B.5.2) and below) and obtain a bound of order $\mathscr{O}\Big(\mathbb{E}[\mathscr{P}_{M,|\mathscr{L}|}(\mathscr{N}_E)]\sqrt{\mathrm{VC}(\Pi)/n}\Big)$. Here, Lipschitz properties are guaranteed by the second condition of Assumption 3.4.3 for which the estimated $\hat{m}$ and $1/\hat{e}$ are bounded .

Bounding (i)

We are now left to bound (i). Observe that we have

$$\begin{aligned}(i) &\leq \mathbb{E}\Big[\sqrt{\frac{1}{n}\sum_{i=1}^{n}\sup_{d,s}\Big(\frac{1}{\tilde{e}_i(d,s)} - \frac{1}{\hat{\tilde{e}}_i(d,s)}\Big)^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\sup_{d,s}\Big(\tilde{m}_i(d,s) - \hat{\tilde{m}}_i(d,s)\Big)^2}\Big] \\ &\leq \sqrt{\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n}\sup_{d,s}\Big(\frac{1}{\tilde{e}_i(d,s)} - \frac{1}{\hat{\tilde{e}}_i(d,s)}\Big)^2\Big]}\sqrt{\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n}\sup_{d,s}\Big(\tilde{m}_i(d,s) - \hat{\tilde{m}}_i(d,s)\Big)^2\Big]} \leq \bar{C}n^{1/2},\end{aligned} \tag{B.5.13}$$

for a finite constant $\bar{C} < \infty$ by Assumption 3.4.3.

Proof of Theorem 2.3.4

As in the above theorem, it suffices to bound $\mathbb{E}\Big[\sup_{\pi\in\Pi}|W(\pi) - W_n(\pi)|\Big]$.

Expectation

We can write

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|W(\pi) - W_n(\pi)|\Big] \leq \mathscr{K}_\Pi(n) + \mathbb{E}\Big[\sup_{\pi\in\Pi}|W^*(\pi) - W_n(\pi)|\Big]$$

where

$$W^*(\pi) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[r\Big(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|, \varepsilon_i\Big)\Big] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[m\Big(\pi(X_i), S_i(\pi), Z_i, |\mathscr{N}_i|\Big)\Big],$$

and where the second equality follows from Assumption 2.2.1, 2.2.2 and 2.2.4 similarly to what discussed in the proof of Lemma B.4.9. By Lemma B.4.9, we have that $W^*(\pi) = \mathbb{E}[W_n(\pi, m^c, e^c)]$ under correct specification of either m^c or e^c. Therefore for $i \in \{1, \cdots, n\}$

$$W^*(\pi) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\Big[\frac{1\{\pi(X_i) = D_i, \sum_{k\in\mathcal{N}_i}\pi(X_k) = \sum_{k\in\mathcal{N}_i} D_k\}}{e^c\Big(\pi(X_i), S_i(\pi), O_i\Big)}\Big(Y_i - m^c\Big(\pi(X_i), S_i(\pi), Z_i, |\mathcal{N}_i|\Big)\Big)\Big] + \\ + \mathbb{E}\Big[m^c\Big(\pi(X_i), S_i(\pi), Z_i, |\mathcal{N}_i|\Big)\Big],$$

where $O_i = \Big(Z_{k\in\mathcal{N}_i}, Z_i, |\mathcal{N}_i|\Big)$.

Bound in terms of K

We can then decompose

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|W^*(\pi) - W_n(\pi)|\Big] \leq \sum_{k=1}^{K}\mathbb{E}\Big[\sup_{\pi\in\Pi}|W^*(\pi) - W_n^k(\pi)|\Big]$$

where $W_n^k(\pi)$ is the empirical welfare in a group k of independent observations.

Symmetrization and conclusion

Using Lemma B.7.1 we can use the standard symmetrization argument to $\mathbb{E}\Big[\sup_{\pi\in\Pi}|W^*(\pi) - W_n^k(\pi)|\Big]$, since each summand in $W_n^k(\pi)$ is independent and identically distributed. The proof concludes invoking Lemma B.4.7 and by Jensen's inequality to bound $\mathbb{E}[\sqrt{\mathcal{N}_n \log(\mathcal{N}_n)}] \leq \sqrt{\mathbb{E}[\mathcal{N}_n \log(\mathcal{N}_n)]}$.

Proof of Theorem B.1.2

By the law of iterated expectations we obtain that if either m^c or e^c are correctly specified, and under Assumption 2.2.3

$$\mathbb{E}[W_n^t(\pi, m^c, e^c)] = W(\pi) \tag{B.5.14}$$

similarly to what discussed in Lemma B.4.9. By trivial rearrengement, we obtain that

$$\sup_{\pi\in\Pi} W(\pi) - W(\hat{\pi}^{aipw}) \leq 2\sup_{\pi\in\Pi}\Big|W_n^t(\pi, m^c, e^c) - W(\pi)\Big|. \tag{B.5.15}$$

Notice now that the exact same argument for bounding the above term follows from the proof of Theorem B.5.1 holds, whereas in such a case the moment conditions also depend on $\bar{\rho}$. In particular the moment bounds are multiplied by $\bar{\rho}^3$. Similarly, the Lipschitz constant of the conditional mean function also must be multiplied by the term $\bar{\rho}$.

Proof of Theorem B.1.1

First equality

First we want to show that (note that $\Big\{D_i = \pi(X_i) \forall i \in \{1, \cdots, E\}\Big\}$ denotes the event that $D_i = \pi(X_i)$ but do not condition on the values of D_i)

$$\mathbb{E}\Big[r\Big(T_i, \sum_{k \in \mathcal{N}_i} T_k, Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big|\Big\{D_i = \pi(X_i) \forall i \in \{1, \cdots, E\}\Big\}\Big] = \mathbb{E}\Big[r\Big(T_i(\pi), \sum_{k \in \mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big].$$

The proof of the claim follows from the assumption that $(\varepsilon_j, \nu_j)_{j=1}^E \perp \Big(Z, A, (\varepsilon_{D_j})_{j=1}^E\Big)$. In particular, we can write

$$\begin{aligned}
&\mathbb{E}\Big[r\Big(T_i, \sum_{k \in \mathcal{N}_i} T_k, Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big|\Big\{D_i = \pi(X_i) \forall i \in \{1, \cdots, E\}\Big\}\Big] \\
&= \mathbb{E}\Big[r\Big(T_i(\pi), \sum_{k \in \mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big|\Big\{D_i = \pi(X_i) \forall i \in \{1, \cdots, E\}\Big\}\Big] \\
&= \mathbb{E}_{A,Z}\Big[\mathbb{E}\Big[r\Big(T_i(\pi), \sum_{k \in \mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big|\Big\{D_i = \pi(X_i) \forall i \in \{1, \cdots, E\}\Big\}, A, Z\Big]\Big] \\
&= \mathbb{E}_{A,Z}\Big[\mathbb{E}\Big[r\Big(T_i(\pi), \sum_{k \in \mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big|A, Z\Big]\Big] \\
&= \mathbb{E}\Big[r\Big(T_i(\pi), \sum_{k \in \mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i\Big)\Big].
\end{aligned}$$

Second equality

Second, we want to derive the main decomposition in the second equality. Let denote $\mathbb{E}_\pi$ the expectation conditional on the event $\Big\{D_i = \pi(X_i) \forall i \in \{1, \cdots, E\}\Big\}$. Using the law of iterated

expectations and the equality above, we have

$$\mathbb{E}_\pi\Big[r(T_i, \sum_{k\in\mathcal{N}_i} T_k, Z_i, |\mathcal{N}_i|, \varepsilon_i)\Big]$$
$$= \mathbb{E}\Big[\mathbb{E}\Big[r(T_i(\pi), \sum_{k\in\mathcal{N}_i} T_k(\pi), Z_i, |\mathcal{N}_i|, \varepsilon_i)\Big|Z, A\Big]\Big].$$

Using the definition of the expectation, we have that (I) equals

$$\underbrace{\sum_{s\in\{0,\cdots,|\mathcal{N}_i|\}} \mathbb{E}\Big[r(d,s,Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big|T_i(\pi)=d, \sum_{k\in\mathcal{N}_i} T_k(\pi)=s, Z, A\Big]}_{(i)} \times \underbrace{P\Big(T_i(\pi)=d, \sum_{k\in\mathcal{N}_i} T_k(\pi)=s\Big|A,Z\Big)}_{(ii)}.$$

We now discuss (i) and (ii) separately. First, we observe that since $(\varepsilon_j)_{j=1}^E \perp \Big(Z, A, (\varepsilon_{D_j}, \nu_j)_{j=1}^E\Big)$

$$(i) = \mathbb{E}\Big[r(d,s,Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big|Z_i,|\mathcal{N}_i|,T_i(\pi)=d,\sum_{k\in\mathcal{N}_i}T_k(\pi)=s,Z,A\Big]$$
$$= \mathbb{E}_\pi\Big[r(d,s,Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big|Z_i,|\mathcal{N}_i|,T_i=d,\sum_{k\in\mathcal{N}_i}T_k=s,Z,A\Big]$$
$$= \mathbb{E}\Big[r(d,s,Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big|Z_i,|\mathcal{N}_i|,T_i=d,\sum_{k\in\mathcal{N}_i}T_k=s,Z,A\Big]$$
$$= \mathbb{E}\Big[r(d,s,Z_i,|\mathcal{N}_i|,\varepsilon_i)\Big|Z_i,|\mathcal{N}_i|,T_i=d,\sum_{k\in\mathcal{N}_i}T_k=s\Big].$$

Consider now (ii). Observe that by indepedence and exogeneity of $(\nu_j)_{j=1}^E$, we can write

$$(ii) = P\Big(T_i(\pi)=d\Big|A,Z\Big) \times \sum_{u_1,\cdots,u_l:\sum_v u_v=s} \prod_{k=1}^{|\mathcal{N}_i|} P\Big(T_{\mathcal{N}_i^{(k)}}(\pi)=u_k\Big|A,Z\Big)$$

where the expression sums over all possible combinations of selected treatments such that the sum of the selected treatments of the neighbors is s. Using again exogeneity of ν_i, we have

$$P\Big(T_i(\pi)=d\Big|A,Z\Big) = P\Big(T_i=d\Big|Z_i,|\mathcal{N}_i|,D_i=\pi(X_i),\sum_{k\in\mathcal{N}_i}D_k=\sum_{k\in\mathcal{N}_i}\pi(X_k),\sum_{k\in\mathcal{N}_i}\pi(X_k),\pi(X_i)\Big).$$

Similar reasoning also applies to neighbors' selected treatments, which depend on the assigned treatments to the second-degree neighbors. The proof completes.

Preliminary Lemmas

This section collects a first set of lemmas from past literature that we invoke in our proofs.

Lemma B.7.1. *(Van Der Vaart and Wellner (1996)) Let $\sigma_1,...,\sigma_n$ be Rademacher sequence independent of $X_1,...,X_n$. Suppose that $X_1,\cdots,X_n$ are i.i.d.. Then*

$$\mathbb{E}\Big[sup_{f\in\mathscr{F}}\Big|\sum_{i=1}^{n} f(X_i)-\mathbb{E}[f(X_i)]\Big|\Big] \leq 2\mathbb{E}\Big[sup_{f\in\mathscr{F}}\Big|\sum_{i=1}^{n}\sigma_i f(X_i)\Big|\Big].$$

Lemma B.7.2. *(Brook's Theorem,Brooks (1941)) For any connected undirected graph G with maximum degree Δ, the chromatic number of G is at most Δ unless G is a complete graph or an odd cycle case the chromatic number is $\Delta+1$.*

The next lemma discuss the case of estimated conditional mean and balancing score function. The lemma follows from Kitagawa and Tetenov (2018).

Lemma B.7.3. *(From Kitagawa and Tetenov (2018)) The following holds.*

$$\sup_{\pi\in\Pi} W(\pi)-W(\hat{\pi}^{aipw}_{\hat{m},\hat{e}}) \leq 2\sup_{\pi\in\Pi}|W^{aipw}_n(\pi,m^c,e^c)-W(\pi)|+2\sup_{\pi\in\Pi}|\hat{W}^{aipw}_n(\pi,\hat{m},\hat{e})-W_n(\pi,m^c,e^c)|.$$

Lemma B.7.4. *(Lemma A.1, Kitagawa and Tetenov (2019)) Let $t_0>1$, define*

$$g(t):=\begin{cases} 0, & \text{for } t=0 \\ t^{-1/2}, & t\geq 1 \end{cases}, \quad h(t)=t_0^{-1/2}-\frac{1}{2}t_0^{-3/2}(t-t_0)+t_0^{-2}(t-t_0).$$

Then $g(t)\leq h(t)$, for $t=0$ and all $t\geq 1$.

In the next two lemmas we formalize two key properties of covering numbers.

Lemma B.7.5. *(Theorem 29.6, Devroye et al. (2013)) Let $\mathscr{F}_1,...,\mathscr{F}_k$ be classes of real functions on $\mathbb{R}^d$. For n arbitrary points $z_1^n=(z_1,...,z_n)$ in $\mathbb{R}^d$, define the sets $\mathscr{F}_1(z_1^n),...,\mathscr{F}_k(z_1^n)$ in $\mathbb{R}^n$ by*

$$\mathscr{F}_j(z_1^n)=\{f_j(z_1),...,f_j(z_n): f_j\in\mathscr{F}_j\}, \quad j=1,...,k.$$

Also, introduce

$$\mathscr{F}=\{f_1+...+f_k; f_j\in\mathscr{F}_j, \quad j=1,...,k\}.$$

Then for every $\varepsilon > 0$ *and* z_1^n,

$$\mathscr{N}_1\Big(\varepsilon, \mathscr{F}(z_1^n)\Big) \leq \prod_{j=1}^{k} \mathscr{N}_1\Big(\varepsilon/k, \mathscr{F}_j(z_1^n)\Big)$$

Lemma B.7.6. *(Pollard, 1990) Let* $\mathscr{F}$ *and* $\mathscr{G}$ *be classes of real valued functions on* $\mathbb{R}^d$ *bounded by* M_1 *and* M_2 *respectively. For arbitrary fixed points* z_1^n *in* $\mathbb{R}^d$*, let*

$$\mathscr{J}(z_1^n) = \{(h(z_1), ..., h(z_n); h \in \mathscr{J}\}, \quad \mathscr{J} = \{fg; f \in \mathscr{F}, g \in \mathscr{G}\}.$$

Then for every $\varepsilon > 0$ *and* z_1^n,

$$\mathscr{N}_1\Big(\varepsilon, \mathscr{J}(z_1^n)\Big) \leq \mathscr{N}_1\Big(\frac{\varepsilon}{2M_2}, \mathscr{F}(z_1^n)\Big)\mathscr{N}_1\Big(\frac{\varepsilon}{2M_1}, \mathscr{G}(z_1^n)\Big).$$

Lemma B.7.7. *(Wenocur and Dudley, 1981) Let* $g : \mathbb{R}^d \mapsto \mathbb{R}$ *be an arbitrary function and consider the class of functions* $\mathscr{G} = \{g + f, f \in \mathscr{F}\}$*. Then*

$$VC(\mathscr{G}) = VC(\mathscr{F})$$

where $VC(\mathscr{F})$*,* $VC(\mathscr{G})$ *denotes the VC dimension respectively of* $\mathscr{F}$ *and* $\mathscr{G}$*.*

An important relation between covering numbers and Rademacher complexity is given by Dudley's entropy integral bound.

Lemma B.7.8. *(From Theorem 5.22 in Wainwright (2019)) For a function class* $\mathscr{F}$ *of uniformly bounded functions and arbitrary fixed points* $z_1^n \in \mathbb{R}^d$*, and i.i.d. Rademacher random variables* $\sigma_1, ..., \sigma_n$ *in* $\mathbb{R}$*,*

$$\mathbb{E}_\sigma\Big[sup_{f\in\mathscr{F}}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i f(z_i)\Big|\Big] \leq \frac{32}{\sqrt{n}}\int_0^{D_q}\sqrt{\log\Big(\mathscr{N}_q\Big(u, \mathscr{F}(z_1^n)\Big)\Big)}du,$$

where D_q denotes the maximum diameter of $\mathscr{F}(z_1^n)$ according to the metric $d_q(f,g) = \left(\frac{1}{n}\sum_{i=1}^n (f(z_i) - g(z_i))^q\right)^{1/q}$.

Theorem 5.22 in Wainwright (2019) provides a general version of Lemma B.7.8. Equivalent versions of Lemma B.7.8 can be found also in Van Der Vaart and Wellner (1996).

Lemma B.7.9. *Let $X_1, ..., X_n$ be independent Bernoulli random variables with $b_i \sim Bern(p_i)$. Let $\bar{p} = \frac{1}{n}\sum_{i=1}^n p_i$ with $n\bar{p} > 1$ and $g(.)$ as defined in the Lemma B.7.4. Then*

$$\mathbb{E}\Big[g\Big(\sum_{i=1}^n X_i\Big)\Big] < 2(n\bar{p})^{-1/2}$$

Proof. The proof follows similarly as in Kitagawa and Tetenov (2019). Let $h(.)$ be the function defined in Lemma B.7.4 with $t_0 = n\bar{p}$. By Lemma B.7.4,

$$\begin{aligned}
\mathbb{E}\Big[g\Big(\sum_{i=1}^n X_i\Big)\Big] &\leq \mathbb{E}\Big[h\Big(\sum_{i=1}^n X_i\Big)\Big] \\
&= \mathbb{E}\Big[(n\bar{p})^{-1/2} - \frac{1}{2}\sqrt{(n\bar{p})^3}\Big(\sum_{i=1}^n X_i - \sum_{i=1}^n \mathbb{E}[X_i]\Big) + (n\bar{p})^{-2}\Big(\sum_{i=1}^n X_i - \sum_{i=1}^n \mathbb{E}[X_i]\Big)^2\Big] \\
&= (n\bar{p})^{-1/2} - \frac{1}{2}(n\bar{p})^{-3/2}0 + (n\bar{p})^{-2} Var\Big(\sum_{i=1}^n X_i\Big) \\
&= (n\bar{p})^{-1/2} + (n\bar{p})^{-2}\sum_{i=1}^n p_i(1-p_i) \\
&\leq (n\bar{p})^{-1/2} + (n\bar{p})^{-2}\sum_{i=1}^n p_i \\
&= (n\bar{p})^{-1/2} + (n\bar{p})^{-2} n\bar{p} \\
&= (n\bar{p})^{-1/2} + (n\bar{p})^{-1} < 2(n\bar{p})^{-1/2}
\end{aligned}$$

since $n\bar{p} > 1$ and $(n\bar{p})^{-1} < (n\bar{p})^{-1/2}$.

□

We conclude our discussion with an extension under lack of point-wise measurability.

Lemma B.7.10. *(Lemma 2.6, Hajłasz and Malỳ (2002)) Let $\mathscr{U}$ be a class of measurable functions defined on a measurable space $E \subset \mathbb{R}^n$. Then the lattice supremum defined as $\bigvee \mathscr{U}$ exists and*

there is a countable sub-family $\mathscr{V} \subset \mathscr{U}$ such that

$$\bigvee \mathscr{U} = \bigvee \mathscr{V} = \sup \mathscr{V}. \tag{B.7.1}$$

The above lemma has one important implication: whenever the function Π is not pointwise measurable, by taking the lattice supremum over a function class, we are guaranteed that such supremum corresponds to a supremum over a countable subset, for which the pointwise supremum is defined. By construction, such class $\mathscr{V}$ has VC-dimension bounded by the VC-dimension of the original function class $\mathscr{U}$.

Appendix C

Appendix to Chapter 3

Extensions and Mathematical Details

Comparison under Strong Duality

We now sketch the differences in the optimization problem with the one in Equation (3.9) assuming strong duality for expositional convenience and providing an intuition on the result in Proposition 8. Assuming strong-duality, the optimization problem of maximizing welfare under fairness constraint in Equation (3.9) can be equivalently rewritten as:

$$\tilde{\pi} \in \arg\min_{\pi \in \Pi} \text{UnFairness}(\pi), \quad \text{such that } p_1 W_1(\pi) + (1-p_1) W_0(\pi) \geq \lambda(\kappa) \tag{C.1.1}$$

for some constant $\lambda(\kappa) \leq \bar{W}_{p_1}$ which depends on κ. We now constrast Equation (C.1.1) to our proposed approach (Equation (3.7)). Suppose first that $\lambda(\kappa) = \bar{W}_{p_1}$, i.e., $\widetilde{\pi}$ *is* Pareto optimal. Then the constraint in Equation (C.1.1) is *stricter* than the constraint in Equation (3.7), since the latter case imposes that $\alpha W_1(\pi) + (1-\alpha) W_0(\pi) \geq \bar{W}_\alpha$, for *some* α, instead of for a particular chosen weight (e.g., p_1). As a result, $\pi^\star$ leads to a lower level of UnFairness whenever $\widetilde{\pi}$ *is* Pareto optimal, since $\pi^\star$ minimizes UnFairness under weaker constraints compared to $\widetilde{\pi}$. When instead $\widetilde{\pi}$ is *not* Pareto optimal, i.e., $\lambda(\kappa) < \bar{W}_{p_1}$, $\widetilde{\pi}$ is Pareto dominated by some other allocation $\widetilde{\widetilde{\pi}}$. However $\widetilde{\widetilde{\pi}}$ leads to a larger UnFairness than $\pi^\star$, while not Pareto dominating $\pi^\star$.

The key intuition is the following: under strong duality, the dual of $\widetilde{\pi}$ corresponds

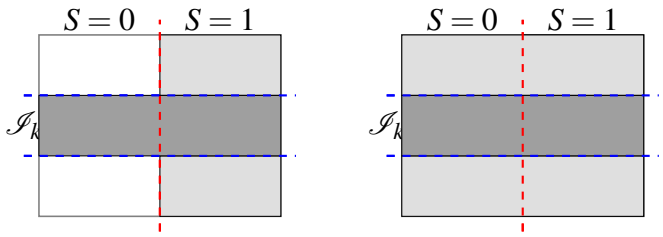


Figure C.1.1. Graphical representation of cross-fitting under two alternative models. The light gray area is the training set, used to construct an estimator of $\hat{m}_{d,s=1}$, whereas the darker gray area is an evaluation set, area in which a prediction of $\hat{m}_{d,s=1}$ is computed.

to minimizing UnFairness for *one particular* weighted combination of welfare exceeding a certain threshold. In contrast, our decision problem imposes the constraint that *some* weighted combination of welfares exceeds a certain threshold. This difference reflects the difference between the lexicographic preferences that we propose as opposed to an additive social planner's utility. It guarantees that whenever $\widetilde{\pi}$ is Pareto optimal, its fairness is dominated by the one under $\pi^\star$.

Cross-fitting with UnFairness

In this section we discuss cross-fitting with fairness. Two alternative cross-fitting procedures are available to the researcher. The first one, consists in dividing the sample into K folds and estimating the conditional mean $\hat{m}_{d,s}^{(-k(i))}(X_i)$ using observations for which $S=s$ only, after excluding the fold k corresponding to unit i (panel on the right in Figure C.1.1). Formally, let $i \in \mathscr{I}_k \cap \mathscr{S}_1$ where $\mathscr{I}_k$ is the k-th fold of the data and $\mathscr{S}_1 = \{i : S_i = s_1\}$. Let $\hat{m}_{d,s_2}^{(-k(i))}$ be an estimator obtained using samples not in the fold k, $\mathscr{I}_k^c \cap \mathscr{S}_1^c$ for which $\mathscr{S}_1^c = \{i : S_i = s_2\}$; for example by a random forest or linear regression of Y_j onto X_j for $S_j = s_2$, and $j \notin \mathscr{I}_k$. Such an approach does not impose parametric restrictions on the dependence of $m_{d,s}$ on the attribute s, at the expense of shrinking the effective sample size used for estimation. The second approach consists in further imposing additional parametric restrictions on the depends of $m_{d,s}$ on s and using all observations in all folds except k for estimating $\hat{m}_{d,s}^{(-k(i))}(X_i)$ (panel on the right in Figure C.1.1).

Linear or Quadratic Constraints for the Policy Function Space Representation

In this section we discuss mixed integer formualtions of probabilistic and deterministic decisions rules. Consider first a deterministic decision rule of the form

$$\Pi = \Big\{ \pi_\beta(X,S) = 1\{X^\top \beta + S\beta_0 > 0\}, \quad \beta \in \mathscr{B} \Big\}.$$

Then we can write the constraint (A) in Equation (3.16) as (Kitagawa and Tetenov, 2018)

$$\frac{X_i^\top \beta + s\beta_0}{|C_i|} < z_{s,i} \leq \frac{X_i^\top \beta + s\beta_0}{|C_i|} + 1, \quad C_i > \sup_{\beta \in \mathscr{B}} |X_i^\top \beta| + |\beta_0|, \quad z_{s,i} \in \{0,1\}.$$

Consider now the following probabilistic decision

$$\Pi = \Big\{ \pi_\beta(X,S) = p_1 1\{X_i^\top \beta + S\beta_0 > 0\} + p_0 1\{X^\top \beta + S\beta_0 \leq 0\}, p_1, p_0 \in [0,1], \beta \in \mathscr{B} \Big\}. \tag{C.1.2}$$

Then we can represent each decision variable as follows

$$\begin{aligned} &z_{s,i} = p_1 \xi_{s,i} + p_0(1 - \xi_{s,i}) \\ &\frac{X_i^\top \beta + s\beta_0}{|C_i|} < \xi_{s,i} \leq \frac{X_i^\top \beta + s\beta_0}{|C_i|} + 1, \quad C_i > \sup_{\beta \in \mathscr{B}} |X_i^\top \beta| + |\beta_0|, \quad \xi_{s,i} \in \{0,1\}. \end{aligned}$$

where we introduced the additional variables $\xi_{s,i}$. We use this probablistic rule in the empirical application.

One last type of function class of interest is a linear probability rule of the following form

$$z_{s,i} = X_i^\top \beta + \beta_0 S_i, \quad z_{s,i} \in [0,1]$$

which leads to fast computations due lack of integer variables in the program, and which we use for the simulations in Section 6.1 for computational feasibility.

Extension: Additional Notions of UnFairness

Predictive parity has been discussed in Kasy and Abebe (2020) among others. Here we consider its definition within the context of policy-targeting. Its notion requires additional assumption for its implementation, assuming *deterministic* treatment assignments $\pi(X_i) \in \{0,1\}$ (i.e., $\mathscr{T} = \{0,1\}$). The notion reads as follows:

$$P_s(\pi) = \Big|\mathbb{E}\Big[Y(1)\Big|\pi(X) = 1, S = s\Big] - \mathbb{E}\Big[Y(1)\Big|\pi(X) = 1\Big]\Big|.$$

Larger values of $P_s(\pi)$ increase UnFairness. Using the definition of the conditional expectation, and using consistency of potential outcomes, the following lemma holds.

Lemma C.1.1. *Let $\mathscr{T} = \{0,1\}$. Then following holds.*

$$P_s(\pi) = (1-p_s)\Big|\frac{\mathbb{E}\Big[Y(1)1\{S_i = s\}\pi(X)\Big]}{p_s\mathbb{P}(\pi(X) = 1|S = s)} - \frac{\mathbb{E}\Big[Y(1)\pi(X)1\{S = s'\}\Big]}{(1-p_s)\mathbb{P}(\pi(X) = 1|S = s')}\Big|.$$

Proof of Lemma C.1.1. Using the definition of conditional expectation:

$$\mathbb{E}\Big[Y\Big|\pi(X) = 1, S = s\Big] = \mathbb{E}\Big[\frac{Y(1)1\{S = s\}\pi(X)}{p_s P(\pi(X) = 1|S = s)}\Big]. \tag{C.1.3}$$

We also write

$$\mathbb{E}\Big[Y\Big|\pi(X) = 1\Big] = p_s\mathbb{E}\Big[Y\Big|\pi(X) = 1, S = s\Big] + (1-p_s)\mathbb{E}\Big[Y\Big|\pi(X) = 1, S = s'\Big].$$

Combining the expression with Equation (C.1.3) completes the proof. □

Given two sensitive groups $\mathscr{S} = \{0,1\}$, the corresponding notion of UnFairness we consider takes the following form:

$$P(\pi) \propto \frac{P_1(\pi)}{1-p_1} = \frac{P_0(\pi)}{p_1}. \tag{C.1.4}$$

We consider a double-robust estimator which takes the following form:

$$\widehat{\mathscr{V}_n}(\pi) = \Big| \sum_{i=1}^{n} \frac{\pi(X_i) S_i \Big\{ \frac{(Y_i - \hat{m}_1(X_i,S_i))D_i}{\hat{e}(X_i,S_i)} + \hat{m}_1(X_i,S_i) \Big\}}{n p_1 \mathbb{P}(\pi(X_i)=1|S_i=1)} - \frac{(1-S_i)\pi(X_i) \Big\{ \frac{(Y_i - \hat{m}_1(X_i,S_i))D_i}{\hat{e}(X_i,S_i)} + \hat{m}_1(X_i,S_i) \Big\}}{n(1-p_1)\mathbb{P}(\pi(X_i)=1|S_i=0)} \Big|. \tag{C.1.5}$$

Observe that the estimator depends on the estimated conditional mean function and propensity score, whereas p_s and $\mathbb{P}(\pi(X)=1|S=s)$ are assumed to be known. These two components can be obtained, for instance from census data, since p_s and $\mathbb{P}(\pi(X)=1|S=s)$ only depend on the distribution of covariates and sensitive attributes. Whenever $\mathbb{P}(\pi(X_i)=1|S_i=s)$ is replaced by its sampled analog $\mathbb{P}_n(\pi(X_i)=1|S_i=s) = \frac{1}{np_s}\sum_{i=1}^{n} \pi(X_i) 1\{S_i=s\}$, we require that $\mathbb{P}_n(\pi(X_i)=1|S_i=s)$ is bounded away from zero almost surely.

Theorem C.1.2 (Predictive parity)**.** *Let Assumptions 3.2.1, 3.4.1,3.4.2, 3.4.3 hold. Let either* UnFairness(π) *be defined using the notion of Predictive (dis)-parity. Assume that* $P(\pi(X,S)=1|S=1), P(\pi(X,S)=1|S=0) \in (\kappa, 1-\kappa)$ *for all* $\pi \in \Pi$, $\kappa \in (0,1)$. *Then for some constant* $c_0 < \infty$*, for any* $\gamma \in (0,1), \lambda \geq \underline{b}\sqrt{\frac{v\log(2/\gamma)}{n}}$*, for a constant* $\underline{b} > 0$*, independent of the sample size with probability at least* $1-2\gamma$*,*

$$\text{UnFairness}(\hat{\pi}) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) \leq c_0 \sqrt{\frac{\log(2/\gamma)}{n}} + c_0 \sqrt{\frac{v}{n}},$$

for a finite constant $c_0 < \infty$.

The proof is in Appendix C.3.14.

Remark 18 (Mixed-integer linear representation of Predictive Parity)**.** Let $\mathbb{P}(\pi(X_i)|S_i=1) = \frac{1}{p_s N}\sum_{i=1}^{N} \pi(X_i) S_i$ where N denotes the number of individuals whose census-information (i.e., baseline covariates and sensitive attributes) are observed. The optimization problem can be formulated as a mixed-integer *fractional* linear program for $\pi(X)$ satisfying a linear representation. This follows after the linearization of the constraint (B), which can be achieved by introducing $2N \times n$ many additional binary variables. Since fractional linear programs admit a mixed-integer

linear program representations (Charnes and Cooper, 1962), the optimization problem can be solved as a mixed integer linear program.

Numerical Studies and Application: Further Results and Details

Empirical Application

Estimation details

We control for confounding of the treatment assignment by estimating the probability of treatment using a penalized logistic regression, where we condition on the non-Caucasian attribute, gender, the average score, years to graduation, whether the individual had previously had entrepreneurship activities, the startup region (which a dummy since only two regions are considered), the degree (either engineer or business) and the school rank. We estimate the outcome using a penalized logistic regression, after conditioning on the above covariates, and any interaction term between gender, treatment assignment, and a vector of covariates, which include years to graduation, prior entrepreneurship, startup region, and the school rank. We estimate treatment effects using a doubly robust estimator. We use cross-fitting with five folds in our estimation.

Additional results for probabilistic treatment assignments

We also consider in our analysis the class of *probabilistic* assignment rules, which assign treatments with a probability decision as in Equation (C.1.2). Results are collected in Figure C.2.1, where we observe that the set of probabilistic decision Pareto dominates the determinitic ones up-to a small optimization error.

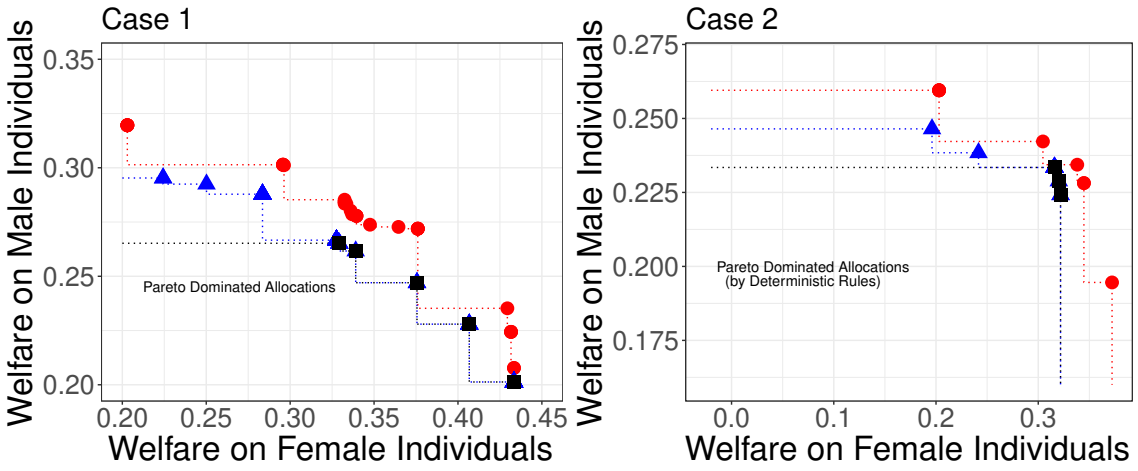


Figure C.2.1. Application: frontier of the probabilistic assignment rule. Dots denote Pareto optimal allocations. Red dots (circle) correspond to Π_1, blue dots (triangle) to Π_2 and black dots (square) to Π_3. The gray area denotes the set of allocations dominated by a *deterministic* decision rule.

Numerical Studies

Data genering process

We consider the following data generating process: we draw

$$S_i \sim_{i.i.d.} \text{Bern}(\hat{p}_1),$$

where $\hat{p}_1$ is the probability of being female in the experiment. We draw covariates $X|S=1$ from the empirical distribution of covariates of the female applicants and similarly we draw $X|S=0$ from the empirical distribution of male applicants. We draw $D|X,S \sim \text{Bern}(\hat{e}(X,S))$. Finally, we draw the potential outcome of interest as

$$Y(d)|X,S = \hat{m}_{d,S}(X) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0,1).$$

Both $\hat{e},\hat{m}$ are estimated using a penalized logistic regression from the empirical distribution.

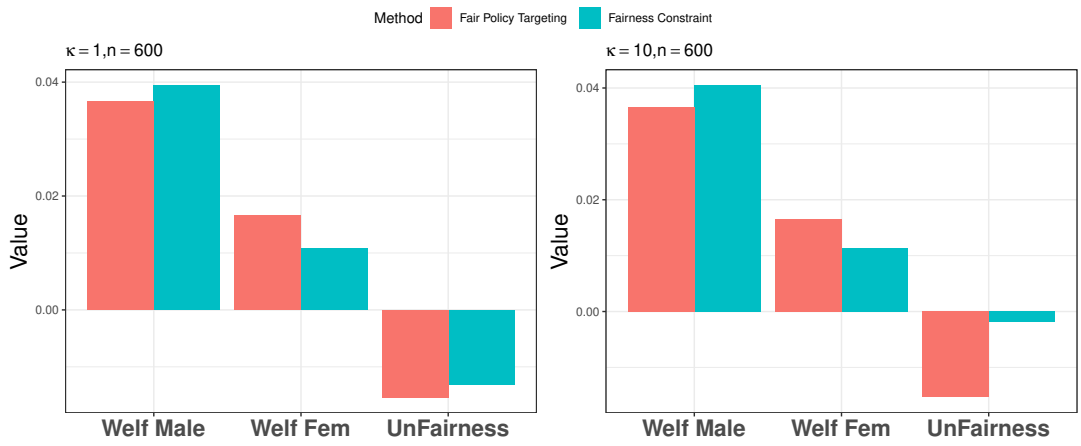


Figure C.2.2. Calibrated experiment to Lyons and Zhang (2017). Results averaged over 500 replications. $n = 600$. The left panel collects results for $\kappa = 1$ (stricted fairness constaint) and the right panel for $\kappa = 10$ (slacker fairness constraint). Colored in red is the proposed method, and in blue, the method that maximizes the empirical welfare imposing fairness constraints. In each panel, the first two columns report the welfare of each group, and the third column the difference in probability of being treated between females and males students.

Estimation

Over each iteration we estimate the propensity score and the conditional mean using a correctly model for the propensity score and a linear model for the conditional mean (estimated with penalized regression). We consider a class of probabilistic assignment rules, taking the following form:

$$\Pi = \Big\{ \pi(x) = \beta_0 + x^\top \beta, \quad \beta \in [-1,1]^{p+1} : \pi(x) \in [0,1], \quad \forall x \in \mathscr{X}_n \Big\},$$

where $\mathscr{X}_n$ denotes the empirical support of covariates. The above function class allows for faster computations compared to the maximum score function class since it does not require the use of any binary decision variable for the policy representation.

Additional results

In Figures C.2.2, we collect additional numerical results for $n = 600$ where we observe consistent results with those in the main text.

Main Proofs

Throughout the rest of our discussion we define

$$\Pi_{o,n} = \Big\{\pi_\alpha \in \Pi : \pi_\alpha \in \arg\sup_{\pi\in\Pi} \Big\{\alpha\hat{W}_0(\pi) + (1-\alpha)\hat{W}_1(\pi)\Big\},\ \text{s.t. } \alpha \in (0,1)\Big\}, \tag{C.3.1}$$

and let $N = \sqrt{n}$ as discussed in the main text. We denote $\alpha_1 - \alpha_2 = \varepsilon$, where, recall, the grid of $(\alpha_i)_{i=1}^N$ contains element equally spaced. We say that $x \lesssim y$ if $y \leq c_0 x$ for a finite constant c_0 independent of n.

Auxiliary Lemmas

Lemma C.3.1. *Under Assumption 3.2.1, 3.4.2 for any sensitive attribute $s \in \{0,1\}$*

$$W_s(\pi) = \mathbb{E}\Big[\frac{1\{S_i = s\}}{p_s}\Big(\frac{Y_i D_i}{e(X_i,s)} - \frac{Y_i(1-D_i)}{1-e(X_i,s)}\Big)\pi(X_i,s)\Big]. \tag{C.3.2}$$

Proof of Lemma C.3.1. Assumption 3.4.2 guarantees existence of the expectation. By definition of the conditional expectation

$$\text{(C.3.2)} = \mathbb{E}\Big[\Big(\frac{Y_i D_i}{e(X_i,s)} - \frac{Y_i(1-D_i)}{1-e(X_i,s)}\Big)\pi(X_i,s)\Big|S_i = s\Big].$$

Using the law of iterated expectations and Assumption 3.2.1 the result directly follows. □

Lemma C.3.2. *Let $W_{s,n} = \frac{1}{n}\sum_{i=1}^n(\Gamma_{1,s,i} - \Gamma_{0,s,i})\pi(X_i,s)$, where $\Gamma_{d,s,i}$ is defined as in Equation* (3.10)*. Let Assumptions 3.2.1, 3.4.1, and 3.4.2 hold. Then with probability at least $1-\gamma$,*

$$\begin{aligned}&\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi) + (1-\alpha)W_1(\pi) - \alpha W_{0,n}(\pi) + (1-\alpha)W_{1,n}(\pi)\Big| \\ &\leq \bar{C}\frac{M}{\delta^2}\sqrt{v/n} + \frac{\bar{C}M}{\delta^2}\sqrt{\log(2/\gamma)/n}\end{aligned} \tag{C.3.3}$$

for a universal constant $\bar{C} < \infty$. In addition,

$$\mathbb{E}\Big[\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big|\Big] \leq \bar{C}\frac{M}{\delta^2}\sqrt{v/n}. \quad \text{(C.3.4)}$$

Proof of Lemma C.3.2. Throughout the proof we refer to $\bar{C} < \infty$ as a universal constant. Observe first that under Assumption 3.4.2 and Assumption 3.4.1, we have

$$\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big|, \quad \text{(C.3.5)}$$

satisfies the bounded difference assumption (Boucheron et al., 2013) with constant $\frac{2M}{\delta^2 n}$. See for instance Boucheron et al. (2005). By the bounded difference inequality, with probability at least $1-\gamma$,

$$\begin{aligned}
&\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big| \\
&\leq \mathbb{E}\Big[\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big|\Big] \quad \text{(C.3.6)} \\
&+\bar{C}\frac{M}{\delta^2}\sqrt{\log(2/\gamma)/n}.
\end{aligned}$$

We now move to bound the expectation in the right-hand side of Equation (C.3.6). Under Assumption 3.2.1, we obtain by Lemma C.3.1 and trivial rearrangments, that

$$\mathbb{E}\Big[\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big] = 0. \quad \text{(C.3.7)}$$

Using the symmetrization argument (Van Der Vaart and Wellner, 1996), we can now bound the above supremum with the Radamacher complexity of the function class of interest (e.g., Athey and Wager (2021), Viviano (2019), Mbakop and Tabord-Meehan (2021)), which

combined with the triangle inequality reads as follows:

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big|\Big]\\
&\leq \mathbb{E}\Big[\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)-\alpha W_{0,n}(\pi)\Big|\Big]+\mathbb{E}\Big[\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|(1-\alpha)W_1(\pi)-(1-\alpha)W_{1,n}(\pi)\Big|\Big]\\
&\leq \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|W_0(\pi)-W_{0,n}(\pi)\Big|\Big]+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|W_1(\pi)-W_{1,n}(\pi)\Big|\Big]\\
&\leq \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\pi(X_i,1)\Gamma_{1,1,i}\Big|\Big]+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\pi(X_i,1)\Gamma_{0,1,i}\Big|\Big]\\
&+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\pi(X_i,0)\Gamma_{1,0,i}\Big|\Big]+\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\pi(X_i,0)\Gamma_{0,0,i}\Big|\Big],
\end{aligned} \tag{C.3.8}$$

where here σ_i are independent Radamacher random variables. We can study each component of the above expression separately. By the Dudley's entropy integral bound, since the VC-dimension of the function class Π is bounded by Assumption 3.4.1, and since each $\Gamma_{d,s,}$ is bounded, we obtain (see for instance Wainwright (2019)), under Assumption 3.4.1 (A) and (B), with trivial rearrangement

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\pi(X_i,s)\Gamma_{d,s,i}\Big|\Big]\leq\frac{M\bar{C}}{\delta^2}\sqrt{v/n}. \tag{C.3.9}$$

for each d,s. The remaining terms follow similarly. The proof is complete. □

Lemma C.3.3. *Let Assumptions 3.2.1, 3.4.1-3.4.3 hold. Then with probability at least* $1-\gamma$*,*

$$\begin{aligned}
&\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha\hat{W}_0(\pi)+(1-\alpha)\hat{W}_1(\pi)\Big|\\
&\leq\bar{C}\frac{M}{\delta^2}\sqrt{v/n}+\frac{\bar{C}M}{\delta^2}\sqrt{\log(2/\gamma)/n}
\end{aligned} \tag{C.3.10}$$

for a universal constant $\bar{C}<\infty$*.*

Proof of Lemma C.3.3. First observe that we can bound the above expression as

$$\begin{aligned}
&\sup_{\alpha\in(0,1)}\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha\hat{W}_0(\pi)+(1-\alpha)\hat{W}_1(\pi)\Big| \leq \\
&\sup_{\alpha\in(0,1)}\underbrace{\sup_{\pi\in\Pi}\Big|\alpha W_0(\pi)+(1-\alpha)W_1(\pi)-\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)\Big|}_{(I)} \\
&+\sup_{\alpha\in(0,1)}\underbrace{\sup_{\pi\in\Pi}\Big|\alpha W_{0,n}(\pi)+(1-\alpha)W_{1,n}(\pi)-\alpha\hat{W}_0(\pi)+(1-\alpha)\hat{W}_1(\pi)\Big|}_{(II)}.
\end{aligned} \tag{C.3.11}$$

Here $W_{s,n}$ is as defined in Lemma C.3.2. The term (I) is bounded as in Lemma C.3.2. Therefore, we are only left to discuss (II).

Using the triangular inequality, we only need to bound

$$\sup_{\pi\in\Pi}\Big|W_{0,n}(\pi)-\hat{W}_{0,n}(\pi)\Big|+\sup_{\pi\in\Pi}\Big|W_{1,n}(\pi)-\hat{W}_{1,n}(\pi)\Big|. \tag{C.3.12}$$

We bound the first term while the second term follows similarly. We write

$$\begin{aligned}
&\sup_{\pi\in\Pi}\Big|W_{s,n}(\pi)-\hat{W}_s(\pi)\Big| \\
&\leq\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}\frac{D_i(Y_i-m_{1,s}(X_i))}{e(X_i,s)}\pi(X_i,s)+\frac{1\{S_i=s\}}{p_s}m_{1,s}(X_i)\pi(X_i,s) \\
&-\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{\hat{p}_s}\frac{D_i(Y_i-\hat{m}_{1,s}(X_i))}{\hat{e}(X_i,s)}\pi(X_i,s)-\frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{1,s}(X_i)\pi(X_i,s)\Big| \\
&+\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}\frac{(1-D_i)(Y_i-m_{0,s}(X_i))}{1-e(X_i,s)}\pi(X_i,s)+\frac{1\{S_i=s\}}{p_s}m_{0,s}(X_i)\pi(X_i,s) \\
&-\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{\hat{p}_s}\frac{(1-D_i)(Y_i-\hat{m}_{s,0}(X_i))}{1-\hat{e}(X_i,s)}\pi(X_i,s)-\frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{0,s}(X_i)\pi(X_i,s)\Big|.
\end{aligned} \tag{C.3.13}$$

We discuss the first component while the second follows similarly.

With trivial re-arrengment, using the triangular inequality, we obtain that the following

holds

$$\begin{aligned}
&\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}\frac{D_i(Y_i-m_{1,s}(X_i))}{e(X_i,s)}\pi(X_i,s)+\frac{1\{S_i=s\}}{p_s}m_{1,s}(X_i)\pi(X_i,s)\\
&-\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{\hat{p}_s}\frac{D_i(Y_i-\hat{m}_{1,s}(X_i))}{\hat{e}(X_i,s)}\pi(X_i,s)-\frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{1,s}(X_i)\pi(X_i,s)\Big|\\
&\leq \underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s\hat{e}(X_i,s)}\Big)\pi(X_i,s)\Big|}_{(i)} \qquad \text{(C.3.14)}\\
&+\underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1\{S_i=s\}D_i}{\hat{e}(X_i,s)\hat{p}_s}-\frac{1\{S_i=s\}}{p_s}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))\pi(X_i,s)\Big|}_{(ii)}.
\end{aligned}$$

We study (i) and (ii) separately. We start from (i). Recall, that by cross fitting $\hat{e}(X_i,s)=\hat{e}^{-k(i)}(X_i,s)$, where $k(i)$ is the fold containing unit i. Therefore, observe that given the K folds for cross-fitting, we have

$$\begin{aligned}
&\Big|\frac{1}{n}\sum_{i=1}^{n}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s\hat{e}(X_i,s)}\Big)\pi(X_i,s)\Big|\\
&\leq \sum_{k\in\{1,\ldots,K\}}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(X_i,s)}\Big)\pi(X_i,s)\Big|.
\end{aligned} \qquad \text{(C.3.15)}$$

In addition, we have that

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{i\in\mathscr{I}_k}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(X_i,s)}\Big)\pi(X_i,s)\Big]\\
&=\mathbb{E}\Big[\mathbb{E}\Big[\sum_{i\in\mathscr{I}_k}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(X_i,s)}\Big)\pi(X_i,s)\Big|\hat{p}^{(-k(i))},\hat{e}^{(-k(i))}\Big]\Big]\\
&=0,
\end{aligned} \qquad \text{(C.3.16)}$$

by cross-fitting. By Assumption 3.4.3, we know that

$$\sup_{x\in\mathscr{X},s\in\mathscr{S}}\Big|\frac{1}{p_s e(x,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(x,s)}\Big|\leq 2/\delta^2 \qquad \text{(C.3.17)}$$

and therefore each summand in Equation (C.3.57) is bounded by a finite constant $2/\delta^2$. We now obtain, using the symmetrization argument (Van Der Vaart and Wellner, 1996), and the Dudley's entropy integral (Wainwright, 2019)

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|\frac{1}{n}\sum_{i\in\mathscr{I}_k}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(X_i,s)}\Big)\pi(X_i,s)|\Big|\hat{p}^{(-k(i))},\hat{e}^{(-k(i))}\Big]$$
$$\lesssim \frac{M}{\delta^2}\sqrt{v/n}. \tag{C.3.18}$$

In addition, by the bounded difference inequality (Boucheron et al., 2005), with probability at least $1-\gamma$, for a universial constant $c<\infty$

$$\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(X_i,s)}\Big)\pi(X_i,s)\Big|\leq$$
$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|\frac{1}{n}\sum_{i\in\mathscr{I}_k}1\{S_i=s\}D_i(Y_i-m_{1,s}(X_i))\Big(\frac{1}{p_s e(X_i,s)}-\frac{1}{\hat{p}_s^{(-k(i))}\hat{e}^{(-k(i))}(X_i,s)}\Big)\pi(X_i,s)|\Big|\hat{p}^{(-k(i))},\hat{e}^{(-k(i))}\Big]$$
$$+c\frac{M}{\delta^2}\sqrt{\frac{\log(2/\gamma)}{n}}. \tag{C.3.19}$$

We now consider the term (ii). Observe that we can write

$$\begin{aligned}(ii)\leq &\underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{D_i1\{S_i=s\}}{\hat{p}_s\hat{e}(X_i,s)}-\frac{D_i1\{S_i=s\}}{p_s e(X_i,s)}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))\pi(X_i,s)\Big|}_{(j)}\\ &+\underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{D_i1\{S_i=s\}}{p_s e(X_i,s)}-\frac{1\{S_i=s\}}{p_s}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))\pi(X_i,s)\Big|}_{(jj)}.\end{aligned} \tag{C.3.20}$$

We consider each term seperately. Consider (jj) first. Using the cross-fitting argument we obtain

$$\begin{aligned}&\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{D_i1\{S_i=s\}}{p_s e(X_i,s)}-\frac{1\{S_i=s\}}{p_s}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))\pi(X_i,s)\Big|\\ &\leq\sum_{k\in\{1,\ldots,K\}}\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}\Big(\frac{D_i1\{S_i=s\}}{p_s e(X_i,s)}-\frac{1\{S_i=s\}}{p_s}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}^{(-k(i))}(X_i))\pi(X_i,s)\Big|.\end{aligned} \tag{C.3.21}$$

Observe now that

$$\mathbb{E}\Big[\Big(\frac{D_i 1\{S_i=s\}}{p_s e(X_i,s)}-\frac{1\{S_i=s\}}{p_s}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}^{(-k(i))}(X_i))\pi(X_i,s)\Big|\hat{m}_{1,s}^{(-k(i))}\Big]=0. \tag{C.3.22}$$

Therefore, following the same argument discussed before, we obtain that with probability at least $1-\gamma$

$$\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}\Big(\frac{D_i 1\{S_i=s\}}{p_s e(X_i,s)}-\frac{1\{S_i=s\}}{p_s}\Big)(m_{1,s}(X_i)-\hat{m}_{1,s}^{(-k(i))}(X_i))\pi(X_i,s)\Big| \lesssim \frac{M}{\delta^2}\sqrt{\frac{v}{n}}+\frac{M}{\delta^2}\sqrt{\frac{\log(2/\gamma)}{n}}. \tag{C.3.23}$$

We are now left to bound (j). We obtain that

$$(j)\leq\sqrt{\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1}{\hat{p}_s\hat{e}(X_i,s)}-\frac{1}{p_s e(X_i,s)}\Big)^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))^2}. \tag{C.3.24}$$

Such a bound does not depend on π. Observe now that we can write by Assumption 3.4.3

$$\begin{aligned}&\sqrt{\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1}{\hat{p}_s\hat{e}(X_i,s)}-\frac{1}{p_s e(X_i,s)}\Big)^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))^2}\\&\leq\frac{1}{\delta}\sqrt{\sum_{k\in\{1,\dots,K\}}\frac{1}{n}\sum_{i\in\mathscr{I}_k}\Big(\frac{1}{\hat{e}^{(-k(i))}(X_i,s)\hat{p}^{-k(i)}}-\frac{1}{e(X_i,s)p_s}\Big)^2}\sqrt{\sum_{k\in\{1,\dots,K\}}\frac{1}{n}\sum_{i\in\mathscr{I}_k}(m_{1,s}(X_i)-\hat{m}_{1,s}^{(-k(i))}(X_i))^2}.\end{aligned} \tag{C.3.25}$$

By the bounded difference inequality, and the union bound we obtain that the following holds:

$$\begin{aligned}&\sqrt{\sum_{k\in\{1,\dots,K\}}\frac{1}{n}\sum_{i\in\mathscr{I}_k}\Big(\frac{1}{\hat{e}^{(-k(i))}(X_i,s)\hat{p}^{-k(i)}}-\frac{1}{e(X_i,s)p_s}\Big)^2}\sqrt{\sum_{k\in\{1,\dots,K\}}\frac{1}{n}\sum_{i\in\mathscr{I}_k}(m_{1,s}(X_i)-\hat{m}_{1,s}^{(-k(i))}(X_i))^2}\\&\leq K\sqrt{\mathbb{E}\Big[\Big(\frac{1}{\hat{e}(X_i,s)\hat{p}}-\frac{1}{e(X_i,s)p_s}\Big)^2\Big]}\sqrt{\mathbb{E}\Big[(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))^2\Big]}\\&+2\sqrt[4]{\log(2K/\gamma)/n}\sqrt{\mathbb{E}\Big[\Big(\frac{1}{\hat{e}(X_i,s)\hat{p}}-\frac{1}{e(X_i,s)p_s}\Big)^2\Big]}+2\sqrt[4]{\log(2K/\gamma)/n}\sqrt{\mathbb{E}\Big[(m_{1,s}(X_i)-\hat{m}_{1,s}(X_i))^2\Big]}\\&+2\sqrt{\log(2K/\gamma)/n},\end{aligned} \tag{C.3.26}$$

with probability at least $1-\gamma$. Under Assumption 3.4.3 and the union bound, the result completes since K is a finite number. □

Lemma C.3.4. *Let*

$$G(\alpha) = \sup_{\pi \in \Pi} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\} - \sup_{\pi \in \hat{\Pi}_o} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\}. \tag{C.3.27}$$

Define

$$\mathscr{G} = \{ G(\alpha), \alpha \in (0,1) \}.$$

Under Assumption 3.4.1, for any $\varepsilon > 0$, there exist a set $\{\alpha_1, ..., \alpha_{N(\varepsilon)}\}$, such that for all $\alpha \in (0,1)$,

$$|G(\alpha) - \max_{j \in \{1,...,N(\varepsilon)\}} G(\alpha_j)| \leq 4\varepsilon M, \tag{C.3.28}$$

and $N(\varepsilon) \leq 1 + 1/\varepsilon$.

Proof of Lemma C.3.4. We denote $\{\alpha_1, ..., \alpha_{N(\varepsilon)}\}$ an ε-cover of the interval $(0,1)$ with respect to the L1 norm. Namely, $\{\alpha_1, ..., \alpha_{N(\varepsilon)}\}$ are equally spaced numbers between $(0,1)$. Clearly, we have that the covering number $N(\varepsilon) \leq 1 + 1/\varepsilon$. We denote

$$G(\alpha) = \sup_{\pi \in \Pi} \alpha W_0(\pi) + (1-\alpha) W_1(\pi) - \sup_{\pi \in \hat{\Pi}_o} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\}. \tag{C.3.29}$$

To characterize the corresponding cover of the function class

$$\mathscr{G} = \{ G(\alpha), \alpha \in (0,1) \},$$

we claim that for any $\alpha \in (0,1)$, there exist an α_j in the ε cover such that

$$|G(\alpha) - G(\alpha_j)| \leq 4\varepsilon M. \tag{C.3.30}$$

Such a result follows by the argument outlined in the following lines.

Take α_j closest to α. Consider

$$
\begin{aligned}
&|G(\alpha) - G(\alpha_j)| \\
&= \Big| \sup_{\pi\in\Pi} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\} - \sup_{\pi\in\hat{\Pi}_o} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\} \\
&- \sup_{\pi\in\Pi} \Big\{ \alpha_j W_0(\pi) + (1-\alpha_j) W_1(\pi) \Big\} + \sup_{\pi\in\hat{\Pi}_o} \Big\{ \alpha_j W_0(\pi) + (1-\alpha_j) W_1(\pi) \Big\} \Big| \\
&\leq \underbrace{\Big| \sup_{\pi\in\Pi} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\} - \sup_{\pi\in\Pi} \Big\{ \alpha_j W_0(\pi) + (1-\alpha_j) W_1(\pi) \Big\} \Big|}_{(i)} \\
&+ \underbrace{\Big| \sup_{\pi\in\hat{\Pi}_o} \Big\{ \alpha W_0(\pi) + (1-\alpha) W_1(\pi) \Big\} - \sup_{\pi\in\hat{\Pi}_o} \Big\{ \alpha_j W_0(\pi) + (1-\alpha_j) W_1(\pi) \Big\} \Big|}_{(ii)}.
\end{aligned}
\tag{C.3.31}
$$

We study (i) and (ii) separately. Consider first (i). We observe the following fact: whenever

$$
\sup_{\pi\in\Pi} \alpha W_0(\pi) + (1-\alpha) W_1(\pi) - \sup_{\pi\in\Pi} \alpha_j W_0(\pi) + (1-\alpha_j) W_1(\pi) > 0 \tag{C.3.32}
$$

then we can bound

$$
(i) \leq \Big| \alpha W_0(\pi^*) + (1-\alpha) W_1(\pi^*) - \alpha_j W_0(\pi^*) + (1-\alpha_j) W_1(\pi^*) \Big|. \tag{C.3.33}
$$

Here $\pi^* \in \arg\sup_{\pi\in\Pi} \alpha W_0(\pi) + (1-\alpha) W_1(\pi)$. When instead

$$
\sup_{\pi\in\Pi} \alpha W_0(\pi) + (1-\alpha) W_1(\pi) - \sup_{\pi\in\Pi} \alpha_j W_0(\pi) + (1-\alpha_j) W_1(\pi) \leq 0 \tag{C.3.34}
$$

we can use the same argument by switching sign, which, with trivial rearrengment reads as

$$
(i) \leq \Big| \alpha W_0(\pi^{**}) + (1-\alpha) W_1(\pi^{**}) - \alpha_j W_0(\pi^{**}) + (1-\alpha_j) W_1(\pi^{**}) \Big|. \tag{C.3.35}
$$

Here $\pi^{**} \in \arg\sup_{\pi\in\Pi} \alpha_j W_0(\pi) + (1-\alpha_j)W_1(\pi)$. Therefore we obtain,

$$(i) \leq \sup_{\pi\in\Pi} \left| \alpha W_0(\pi) + (1-\alpha)W_1(\pi) - \alpha_j W_0(\pi) + (1-\alpha_j)W_1(\pi) \right| \leq 2|\alpha - \alpha_j| M \quad \text{(C.3.36)}$$

where the last inequality follows by Assumption 3.4.1 and the triangle inequality. Similar reasoning also applies to (ii). Since α_j was chosen to be the closest to α, we have $|\alpha_j - \alpha| \leq \varepsilon$. □

Proof of Lemma 3.2.1

The proof follows similarly to standard microeconomic textbook (Mas-Colell et al., 1995). Let

$$\tilde{\Pi} = \{\pi_\alpha : \pi_\alpha \in \arg\sup_{\pi\in\Pi} \alpha_1 W_0(\pi) + \alpha_2 W_1(\pi), \quad \alpha \in \mathbb{R}^2_+, \alpha_1 + \alpha_2 > 0\}. \quad \text{(C.3.37)}$$

Then we want to show that $\Pi_o = \tilde{\Pi}$. Trivially $\tilde{\Pi} \subseteq \Pi_o$, since otherwise the definition of Pareto optimality would be violated. Consider now some $\pi^* \in \Pi_o$. Then we show that there exist a vector $\alpha \in \mathbb{R}^2_+$, such that π^* maximizes the expression

$$\sup_{\pi\in\Pi} \alpha_1 W_0(\pi) + \alpha_2 W_1(\pi). \quad \text{(C.3.38)}$$

Denote the set

$$\mathscr{F} = \{(\tilde{W}_0, \tilde{W}_1) \in \mathbb{R}^2 : \exists \pi \in \Pi : \tilde{W}_0 \leq W_0(\pi) \text{ and } \tilde{W}_1 \leq W_1(\pi)\}. \quad \text{(C.3.39)}$$

Since $(0,0) \in \mathscr{F}$, such a set is non-empty. Notice now that $W_s(\pi)$ is linear is π for $s \in \{0,1\}$. Therefore, we obtain that the set $\mathscr{F}$ is a convex set, since it denotes the sub-graph of a concave functional. We denote $\bar{W} = (W_0(\pi^*), W_1(\pi^*))$ and $\mathscr{G} = \mathbb{R}^2_{++} + \bar{W}$ the set of welfares that strictly dominates π^*. Then $\mathscr{G}$ is non-empty and convex. Since $\pi^* \in \Pi_o$, we must have that

$\mathscr{F} \cap \mathscr{G} = \emptyset$. Therefore, by the separating hyperplane theorem, there exist an $\alpha \in \mathbb{R}^2$, with $\alpha \neq 0$, such that $\alpha^\top F \leq \alpha^\top (\bar{W} + d)$ for any $F \in \mathscr{F}$, $d \in \mathbb{R}^2_{++}$. Let $d_1 \to \infty$, it must be that $\alpha_1 \in \mathbb{R}_+$, and similarly for α_2. So $\alpha \in \mathbb{R}^2_+$. By letting $d \to 0$, we have that $\alpha^\top F \leq \alpha^\top \bar{W}$. This implies that

$$\alpha_1 W_0(\pi) + \alpha_2 W_1(\pi) \leq \alpha_1 W_0(\pi^*) + \alpha_2 W_1(\pi^*) \tag{C.3.40}$$

for any $\pi \in \Pi$ (since it is true for any $F \in \mathscr{F}$). Hence π^* maximizes welfare over all possible feasible allocations once reweighted by (α_1, α_2). Since the maximizer is invariant to multiplication of the objective function by constants, the result follows after dividing the objective function by the sums of the coefficients, which is non-zero by the separating hyperplane theorem. This completes the proof.

Proof of Proposition 3.2.2

First, observe that by rationality, preferences are complete and transitive. Observe also that the preference function equivalently correspond to lexico-graphic with $\pi \succ \pi'$ if π Pareto dominates π'. If instead neither π, π', Pareto dominates the other, then $\pi \succ \pi'$ is UnFairness (π) $<$ UnFairness (π'). Therefore, it must be that $\mathscr{C}(\Pi) \subseteq \Pi_o$, with $\pi^\star \in \mathscr{C}(\Pi)$ if and only if

$$\pi^\star \in \arg\min_{\pi \in \Pi_o} \text{UnFairness}(\pi).$$

By Lemma 3.2.1 the result directly follows.

Proof of Corollary 7

Define $\widetilde{\Pi} \subseteq \Pi$ the set of policies that satisfy the constraint in Equation (3.7) (i.e., feasible allocations). By Proposition 3.2.2 $\widetilde{\Pi} = \Pi_o$. Observe now that π_ω is a feasible allocation under the constraint in Equation (3.7). This directly implies the conclusion.

Proof of Corollary 8

If $\widetilde{\pi}$ is Pareto optimal, then it represents a feasible allocation (i.e. it satisfies the constraint in Equation (3.7)). If it is not, then any other allocation that *is* Pareto optimal and Pareto dominates $\widetilde{\pi}$ is feasible under the constraint in Equation (3.7) completing the proof.

Proof of Corollary 9

The proof follows directly from the fact that whenever fairness constraints are binding, the estimated policy contains as one possible solution the policy which maximizes the utilitarian welfare under fairness constraints. This follows from the fact that in such case

$$\widetilde{\pi} \in \left\{ \arg\max_{\pi \in \Pi} p_1 W_1(\pi) + (1-p_1) W_0(\pi) \right\} \subseteq \Pi_o,$$

since $\Pi = \Pi(\kappa)$.

Proof of Theorem 3.4.1

Throughout the proof we refer to $\bar{C} < \infty$ as a universal constant. We write

$$\begin{aligned}
&\sup_{\alpha \in (0,1)} \sup_{\pi \in \Pi} \left| \alpha W_0(\pi) + (1-\alpha) W_1(\pi) - \max_{\alpha_j \in \{\alpha_1, \ldots, \alpha_N\}} \alpha_j \hat{W}_0(\pi) - (1-\alpha_j) \hat{W}_1(\pi) - \lambda / \sqrt{n} \right| \\
&\leq \underbrace{\sup_{\alpha \in (0,1)} \sup_{\pi \in \Pi} \left| \alpha W_0(\pi) + (1-\alpha) W_1(\pi) - \alpha \hat{W}_0(\pi) - (1-\alpha) \hat{W}_1(\pi) \right|}_{(I)} + \frac{\lambda}{\sqrt{n}} \\
&+ \underbrace{\sup_{\alpha \in (0,1)} \sup_{\pi \in \Pi} \left| \alpha \hat{W}_0(\pi) + (1-\alpha) \hat{W}_1(\pi) - \max_{\alpha_j \in \{\alpha_1, \ldots, \alpha_N\}} \alpha_j \hat{W}_0(\pi) - (1-\alpha_j) \hat{W}_1(\pi) \right|}_{(II)}.
\end{aligned} \tag{C.3.41}$$

(I) is bounded as in Lemma C.3.3. (II) is bounded as follows.

$$(II) \leq \varepsilon \sup_{\pi \in \Pi} |\hat{W}_0(\pi)| + \varepsilon \sup_{\pi \in \Pi} |\hat{W}_1(\pi)|. \tag{C.3.42}$$

Under Assumption 3.4.3, the estimated conditional mean and propensity score are uniformly bounded. Therefore we obtain that

$$\varepsilon \sup_{\pi\in\Pi} |\hat{W}_0(\pi)| + \sup_{\pi\in\Pi} \varepsilon |\hat{W}_1(\pi)| \leq \bar{C}\varepsilon \frac{M}{\delta^2} \leq \bar{C}\frac{M}{N\delta^2}.$$

Proof of Theorem 3.4.2

Recall the definition of $\bar{W}_\alpha$ in Equation (3.4). The set of Pareto optimal policies reads as follows

$$\pi : \alpha W_1(\pi) + (1-\alpha)W_0(\pi) \geq \bar{W}_\alpha \text{ for some } \alpha \in (0,1).$$

Now it suffices to show for the claim to hold that

$$P\Big(\forall \alpha \in (0,1), \quad \max_{j\in\{1,\cdots,N\}} \bar{W}_\alpha - \bar{W}_{j,n} + \lambda(\gamma)/\sqrt{n} + \frac{\underline{b}}{N} \geq 0\Big) \leq \gamma,$$

where $\lambda(\gamma) = \underline{b}(\sqrt{v} + \sqrt{\log(2/\gamma)})$, whenever $N = \sqrt{n}$ (and hence $\lambda = \lambda(\gamma) + \underline{b}$). Observe that since $\{\alpha_1, \cdots, \alpha_N\}$ are equally spaced, we have that for all $\alpha \in (0,1)$

$$\sup_{\pi\in\Pi} \alpha W_1(\pi) + (1-\alpha)W_0(\pi) \geq \sup_{\pi\in\Pi} \alpha_j W_1(\pi) + (1-\alpha_j)W_0(\pi) + M\varepsilon$$

for some $j \in \{1, \cdots, N\}$ by Assumption 3.4.2 (ii). Taking $\underline{b} \geq M, \varepsilon = 1/N$, we have

$$\begin{aligned}
&P\Big(\forall \alpha \in (0,1), \quad \max_{j\in\{1,\cdots,N\}} \bar{W}_\alpha - \bar{W}_{j,n} + \lambda(\gamma)/\sqrt{n} + \frac{\underline{b}}{N} \geq 0\Big) \\
&\leq P\Big(\max_{j\in\{1,\cdots,N\}} \bar{W}_{\alpha_j} - \bar{W}_{j,n} + \lambda(\gamma)/\sqrt{n} \geq 0\Big).
\end{aligned}$$

We now observe that the following inequality holds:

$$\begin{aligned}
&\sup_{\pi\in\Pi} \alpha_j W_1(\pi) + (1-\alpha_j)W_0(\pi) - \bar{W}_{j,n} \\
&= \sup_{\pi\in\Pi} \Big\{ \alpha_j W_1(\pi) + (1-\alpha_j)W_0(\pi) \Big\} - \sup_{\pi\in\Pi} \Big\{ \alpha \hat{W}_1(\pi) + (1-\alpha)\hat{W}_0(\pi) \Big\} \\
&\leq 2 \sup_{\pi\in\Pi} \Big| \alpha_j W_1(\pi) + (1-\alpha_j)W_0(\pi) - \alpha_j \hat{W}_1(\pi) + (1-\alpha_j)\hat{W}_0(\pi) \Big|.
\end{aligned}$$

By Lemma C.3.3, with probability at least $1-\gamma$,

$$\sup_{\pi\in\Pi} \max_{\alpha_j, j\in\{1,\cdots,N\}} \Big| \alpha_j W_1(\pi) + (1-\alpha_j)W_0(\pi) - \alpha_j \hat{W}_1(\pi) + (1-\alpha_j)\hat{W}_0(\pi) \Big| \leq \bar{C}\sqrt{\frac{v}{n}} + \bar{C}\sqrt{\frac{\log(2/\gamma)}{n}}$$

for a finite constant $\bar{C}$ independent of n. By choosing $\underline{b} \geq 2\bar{C} + M$, the proof completes.

Proof of Theorem 3.4.3

By Theorem 3.4.2 with probability at least $1-\gamma$, $\Pi_o \subseteq \hat{\Pi}_o(\lambda)$ with $\hat{\Pi}_o(\lambda)$ in Equation (3.14). As a result, we can write with probability $1-\gamma$,

$$\text{UnFairness}(\hat{\pi}) - \inf_{\pi\in\Pi_o} \text{UnFairness}(\pi) \leq \text{UnFairness}(\hat{\pi}) - \inf_{\pi\in\hat{\Pi}_o(\lambda)} \text{UnFairness}(\pi).$$

We then write

$$\begin{aligned}
&\text{UnFairness}(\hat{\pi}) - \inf_{\pi\in\hat{\Pi}_o(\lambda)} \text{UnFairness}(\pi) \\
&= \text{UnFairness}(\hat{\pi}) - \hat{\mathscr{V}}_n(\hat{\pi}) + \hat{\mathscr{V}}_n(\hat{\pi}) - \inf_{\pi\in\hat{\Pi}_o(\lambda)} \text{UnFairness}(\pi).
\end{aligned}$$

Since $\hat{\pi}_\lambda \in \hat{\Pi}_o(\lambda)$, we have

$$\begin{aligned}
&\text{UnFairness}(\hat{\pi}) - \hat{\mathscr{V}}_n(\hat{\pi}) + \hat{\mathscr{V}}_n(\hat{\pi}) - \inf_{\pi\in\hat{\Pi}_o(\lambda)} \text{UnFairness}(\pi) \\
&\leq 2 \sup_{\pi\in\hat{\Pi}_o(\lambda)} \Big| \text{UnFairness}(\pi) - \hat{\mathscr{V}}_n(\pi) \Big| \leq 2 \sup_{\pi\in\Pi} \Big| \text{UnFairness}(\pi) - \hat{\mathscr{V}}_n(\pi) \Big|
\end{aligned}$$

where the last equality follows from the fact that $\hat{\Pi}_o(\lambda) \subseteq \Pi$. Assumption 3.4.4 bounds $\sup_{\pi\in\Pi} \Big| \text{UnFairness}(\pi) - \hat{\mathscr{V}}_n(\pi) \Big|$ completing the proof.

Proof of Theorem 3.4.4

For $\widehat{D}(\pi)$ it suffices to observe that

$$\sup_{\pi\in\Pi}\left|\widehat{W}(\pi_1)-\widehat{W}_0(\pi)-W_1(\pi)+W_0(\pi)\right| \leq \sup_{\pi\in\Pi}\left|\widehat{W}(\pi_1)-W_1(\pi)\right| + \sup_{\pi\in\Pi}\left|W_0(\pi)-\widehat{W}_0(\pi)\right|$$

with each term being bounded with probability at least $1-2\gamma$[1], by $\bar{C}\sqrt{v/n}+\bar{C}\sqrt{\log(2/\gamma)/n}$ for a finite constant $\bar{C}<\infty$, similarly to what discussed in the proof of Lemma C.3.3. The UnFairness bound follows as a corollary of Theorem 3.4.3.

For $\widehat{C}(\pi)$ the argument follows similarly, after noticing that we can bound

$$\begin{aligned}
&\sup_{\pi\in\Pi}\left|\frac{1}{n\hat{p}_1}\sum_{i=1}^{n}\pi(X_i)S_i-\mathbb{E}[\pi(X)|S=1]+\frac{1}{n(1-\hat{p}_1)}\sum_{i=1}^{n}\pi(X_i)(1-S_i)-\mathbb{E}[\pi(X)|S=0]\right| \\
&\leq \underbrace{\sup_{\pi\in\Pi}\left|\frac{1}{n\hat{p}_1}\sum_{i=1}^{n}\pi(X_i)S_i-\mathbb{E}[\pi(X)|S=1]\right|}_{(A)} + \underbrace{\sup_{\pi\in\Pi}\left|\frac{1}{(1-\hat{p}_1)n}\sum_{i=1}^{n}\pi(X_i)(1-S_i)-\mathbb{E}[\pi(X)|S=0]\right|}_{(B)}.
\end{aligned}$$

We proceed by bounding (A),while (B) follows similarly. We have

$$(A) \leq \underbrace{\sup_{\pi\in\Pi}\left|\frac{1}{p_1 n}\sum_{i=1}^{n}\pi(X_i)S_i-\mathbb{E}[\pi(X)|S=1]\right|}_{(i)} + \underbrace{\left|\frac{1}{p_1}-\frac{1}{\hat{p}_1}\right|}_{(ii)},$$

where the second component follows by the triangular inequality and the fact that $\pi(X_i)S_i\in\{0,1\}$. We now observe that each summand in (i) is centered around its expectation. Therefore, we can bound (i) using the Radamacher complexity of Π, with

$$\mathbb{E}[(i)] \leq \frac{2}{\delta}\mathbb{E}\Big[\sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\pi(X_i)S_i\right|\Big],$$

with $\sigma_1,\cdots,\sigma_n$ being independent Radamacher random variables. Using the Dudley's entropy bound (see Wainwright (2019)) it is easy to show that the right-hand side is bounded by $\bar{C}\sqrt{v/n}$ for a constant $\bar{C}<\infty$. Finally, using the bounded difference inequality (Boucheron et al., 2003),

[1] 2γ follows by the union bound.

with probability at least $1-\gamma$,

$$|(i)-\mathbb{E}[(i)]| \leq \bar{C}\sqrt{\frac{\log(2/\gamma)}{n}},$$

for a finite constant $\bar{C}$. The bound on the second component (ii) follows from standard property of the sample mean and the assumption that $\hat{p}_1 \geq \delta$. The final statement follows as a direct corollary of Theorem 3.4.3.

For $\mathscr{I}(\pi)$ the claim holds since

$$\begin{aligned}
&\sup_{\pi\in\Pi}\left|I_s(\pi)-\hat{I}_s(\pi)\right| \leq \\
&\underbrace{\sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{i=1}^{n}(\hat{\Gamma}_{1,s,i}-\hat{\Gamma}_{0,s,i})\pi(X_i,s')-\mathbb{E}\Big[(\Gamma_{1,s,i}-\Gamma_{0,s,i})\pi(X_i,s')\Big]\right|}_{(A)} \\
&+\underbrace{\sup_{\pi\in\Pi}\left|\frac{1}{n}\sum_{i=1}^{n}(\hat{\Gamma}_{1,s,i}-\hat{\Gamma}_{0,s,i})\pi(X_i,s)-\mathbb{E}[(\Gamma_{1,s,i}-\Gamma_{0,s,i})\pi(X_i,s)]\right|}_{(B)}.
\end{aligned} \tag{C.3.43}$$

Observe now that under Assumption 3.4.3, following the same argument in Lemma C.3.3, we can bound (A) and (B) as follows

$$(A) \lesssim \sqrt{\frac{v}{n}}+\sqrt{\frac{\log(2/\gamma)}{n}}, \quad (B) \lesssim \sqrt{\frac{v}{n}}+\sqrt{\frac{\log(2/\gamma)}{n}}.$$

with probability at least $1-\gamma$. The reader may refer to the proof of Lemma C.3.3 for details.

Proof of Theorem 3.4.5

First, since $\pi(x,s)$ is constant in s with an abuse of notation we can write $\pi(x)$ as a function of x only. We first observe that we can write

$$C(\pi)=\mathbb{E}\Big[(\frac{(1-S)}{1-p_1}-\frac{S}{p_1})\pi(X)\Big]=\mathbb{E}\Big[\frac{(p_1-S)}{(1-p_1)p_1}\pi(X)\Big]$$

For the lower bound it suffices to find one distribution which satisfies the condition. We choose $Y(1) = 0$, and $Y(0) = 0$ almost surely, which satisfies the bounded assumption on Y. This condition implies that any $\pi \in \Pi$ satisfies Pareto optimality, hence $\Pi_o = \Pi$.

Observe that the expression for $C(\pi)$ corresponds to the risk associated with a classifier $\pi(X)$ for classifying the sensitive attribute S with loss

$$l(S, \pi(X)) \propto (p_1 - S)\pi(X) = \begin{cases} p_1 - 1 \text{ if } S = 1, \pi(X) = 1 \\ p_1 \text{ if } S = 0, \pi(X) = 1 \\ 0 \text{ otherwise .} \end{cases} .$$

We now proceed following some of the steps in Theorem 14.5 and Theorem 14.6 in Devroye et al. (2013), but introducing modifications in the construction of the set of distributions under consideration and in the data-generating process due to the different loss function and its dependence with $P(S = 1)$ (which itself depends on the distribution of (X, S)).[2] We start by choosing D to be distributed as a Bernoulli random variable independent of (X, S). As a result, (Y, D) are independent of (X, S). Therefore, since (Y, D) is independent of (X, S) it suffices to focus on classifiers $\pi_n(X)$ constructed using information $(X_1, S_1), \cdots, (X_n, S_n)$ only. The rest of the proof consists in constructing a distribution of (X, S) such that the lower bound is attained. Recall that classifiers depend on X only and not on S by assumption.

Consider first the case where $(v - 1)/2$ is an integer. The case where it is not follows similarly to below and discussed at the end of the proof. We construct a family of distributions for (X, S), defined $\mathscr{F}$ as follows: first we find points $x_1, \cdots x_v$ that are shattered by Π_o. Each distribution in $\mathscr{F}$ is concentrated on the set of these points. A member in $\mathscr{F}$ is described by $v - 1$ bits $b_1, \cdots, b_{v-1}$. This is representated as a bit vector $b \subset \{0, 1\}^{v-1}$. Each bit vector that

[2] The lack of restriction on the error of the classifier represents a further difference.

we consider is assumed to sum to $(v-1)/2$, namely

$$\sum_{i=1}^{v-1} b_i = \frac{v-1}{2}.$$

Assume that $v-1 \leq n$. For each vector b, we let X put mass m at $x_i, i < v$, and mass $1-(v-1)m$ at x_v. This imposes the condition $(v-1)m \leq 1$, which will be satisfied. We choose for all b that we consider $P(S=1) = p_1 \in (\delta, 1-\delta)$ which we choose later in the proof. Next, introduce the constant $c \in (0, p_1)$. Let U a uniform random variable on $[0,1]$,

$$S = \begin{cases} 1 \text{ if } U \leq p_1 - c + 2cb_i, X = x_i, i < v \\ 1 \text{ if } U \leq p_1, X = x_v \\ 0 \text{ otherwise} \end{cases}.$$

Thus for $X = x_i, i < v$, S is one with probability $p_1 - c$ or $p_1 + c$, while for $X = x_v$ S is one with probability p_1. Now observe that the choice of S and the fact that $P(S=1) = p_1$ implies that

$$p_1 = \sum_{i=1}^{v-1} m(p_1 - c + 2cb_i) + p_1(1 - m(v-1)) = (v-1)mp_1 + p_1(1-m(v-1)), \quad \text{(C.3.44)}$$

since $c\sum_{i=1}^{v-1} b_i = c\frac{v-1}{2}$ by the restriction on $b \in \mathscr{B}$. The above expression is satisfied for any m, so no restrictions on m are implied by the Equation (D.3.18). With a simple argument, it is easy to show that one of the best rules[3] for b is the one which sets

$$f_b(x) = \begin{cases} 1 \text{ if } x = x_i, i < v, b_i = 1 \\ 0 \text{ otherwise.} \end{cases}$$

Such rule is feasible since it has VC-dimension v. Notice now that we can write for the decision

[3] A different which leads to the same objective is the one that classifies one also for $X = x_v$. This would be indifferent with respect to f_b since the loss function at $X = x_v$ is always zero in expectation for either prediction.

rule $f_b(x)$, $\mathbb{E}[l(S, f_b(X))|X = x_i] = -c$ for $i < v$, for fixed b. Observe now that we can write for any $\pi_n, X \in \{x_1, \cdots, x_{v-1}\}$, for fixed b,

$$\mathbb{E}[l(S, \pi_n(X))|X] - \mathbb{E}[l(S, f_b(X))|X] \geq 2c1\{\pi_n(X) \neq f_b(X)\},$$

since if $\pi_n(X) = 1 - f_b(X)$, then $\mathbb{E}[l(S, 1 - f_b(X))|X] = c$. Therefore we can bound for any π_n, and a fixed b

$$\begin{aligned} \text{UnFairness}(\pi_n) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) &\propto \mathbb{E}[l(S, \pi_n(X))] - \inf_{\pi \in \Pi} \mathbb{E}[l(S, \pi(X))] \\ &\geq \sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_b(x_j)\} \\ &\geq \sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_b(x_j)\}. \end{aligned} \tag{C.3.45}$$

Since we take the supremum over the class of distribution $P_b \in \mathscr{F}$ indexed by the bit-vector b, it suffices to provide upper bound with respect to b being a random variable and take expectations over b. We replace b by a uniformly distributed random variable B over $\mathscr{B} \subset \{0,1\}^{v-1}$, where $\mathscr{B}$ is the set of bit vectors which sum to $(v-1)/2$. We observe that for any $t \geq 0$,

$$\begin{aligned} &\sup_{(X,S) \in \mathscr{F}} P\Big(\text{UnFairness}(\pi_n) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) > t\Big) \\ &= \sup_b P\Big(\text{UnFairness}(\pi_n) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) > t\Big) \\ &\geq \mathbb{E}_b\Big[1\{\text{UnFairness}(\pi_n) - \inf_{\pi \in \Pi_o} \text{UnFairness}(\pi) > t\}\Big] \text{ (with random b)} \\ &\geq \mathbb{E}_b\Big[1\Big\{\sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_b(x_j)\} > t\Big\}\Big] \end{aligned}$$

where the last inequality uses Equation (C.3.45) and the monotonicity of the indicator function.

We can now write

$$\mathbb{E}_b\Big[1\Big\{\sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_b(x_j)\} > t\Big\}\Big]$$
$$= \frac{1}{|\mathscr{B}|} \sum_{(x_1,\cdots,x_n',s_1,\cdots,s_n)\in(\{x_1,\cdots,x_v\}\times\{0,1\})^2}$$
$$\sum_{b\in\mathscr{B}} 1\Big\{\sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_b(x_j)\} > t\Big\}\prod_{j=1}^{n} p_b(x_j', s_j)$$

with $p_b(x_j', s_j)$ denoting the joint probability of x_j', s_j. For a fixed b, define $b^c = (1-b_1,\cdots,1-b_{v-1})$. Observe that if $b \in \mathscr{B}$, then $b^c \in \mathscr{B}$ since we assumed that $(v-1)/2$ is an integer. Now observe that if

$$\frac{t}{2mc} \leq (v-1)/2, \tag{C.3.46}$$

then

$$1\Big\{\sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_b(x_j)\} > t\Big\} + 1\Big\{\sum_{j=1}^{v-1} 2mc1\{\pi_n(x_j) = 1 - f_{b^c}(x_j)\} > t\Big\} \geq 1$$

since it must be that either (or both) indicators are equal to one. Therefore for $t\Big/2mc \leq (v-1)/2$, the last expression in the lower bound above is bounded from below by

$$\frac{1}{|\mathscr{B}|} \sum_{(x_1,\cdots,x_n',s_1,\cdots,s_n)\in(\{x_1,\cdots,x_v\}\times\{0,1\})^2} \sum_{b\in\mathscr{B}} \frac{1}{2} \min\Big\{\prod_{j=1}^{n} p_b(x_j', s_j), \prod_{j=1}^{n} p_{b^c}(x_j', s_j)\Big\}.$$

By LeCam's inequality, we have that the above expression is bounded from below by (see Page 244 in Devroye et al. 2013)

$$\frac{1}{4|\mathscr{B}|} \sum_{b\in\mathscr{B}} \Big(\sum_{(x,s)} \sqrt{p_b(x,s)p_{b^c}(x,s)}\Big)^{2n}.$$

Observe that we have for $x = x_v$,

$$p_b(x,1) = p_{b^c}(x,1) = p_1(1-m(v-1)), \quad p_b(x,1) = p_{b^c}(x,1) = (1-p_1)(1-m(v-1)).$$

For $x = x_i, i < v$, we have

$$p_b(x,s)p_{b^c}(x,s) = m^2(p_1^2 - c^2), \quad s \in \{0,1\}.$$

Therefore, we obtain

$$\begin{aligned}\sum_{(x,s)} \sqrt{p_b(x,s)p_{b^c}(x,s)} &= (1-m(v-1)) + 2(v-1)m\sqrt{(p_1^2-c^2)} \\ &= (1-(v-1)m) + 2(v-1)m\sqrt{(p_1^2-c^2)}.\end{aligned}$$

Hence we can write

$$\frac{1}{4|\mathscr{B}|}\sum_{b\in\mathscr{B}}\sum_{(x,s)} \sqrt{p_b(x,s)p_{b^c}(x,s)} = \frac{1}{4}\Big\{(1-(v-1)m) + 2(v-1)m\sqrt{(p_1^2-c^2)}\Big\}.$$

Define $F = m(v-1)(p_1 - c)$. Then we can write

$$\begin{aligned}\frac{1}{4|\mathscr{B}|}\sum_{b\in\mathscr{B}}\Big(\sum_{(x,s)} \sqrt{p_b(x,s)p_{b^c}(x,s)}\Big)^{2n} &= \frac{1}{4}\Big\{(1-(v-1)m) + 2(v-1)m\sqrt{(p_1^2-c^2)}\Big\} \\ &= \frac{1}{4}\Big\{1 - \frac{F}{p_1-c}\Big(1-\sqrt{4p_1^2-4c^2}\Big)\Big\}^{2n}.\end{aligned}$$

We now choose $p_1 = 1/2$. We can now follow Devroye et al. (2013), end of Page 244 and write

$$\begin{aligned}\frac{1}{4}\Big\{1 - \frac{F}{p_1-c}\Big(1-\sqrt{4p_1^2-4c^2}\Big)\Big\}^{2n} &= \frac{1}{4}\Big\{1 - \frac{F}{p_1-c}\Big(1-\sqrt{1-4c^2}\Big)\Big\}^{2n} \\ &\geq \frac{1}{4}\Big\{1 - \frac{F}{p_1-c}4c^2\Big\}^{2n} \\ &\geq \frac{1}{4}\exp\Big(-\frac{16nFc^2}{1-2c}\Big/\Big(1-\frac{8Fc^2}{1-2c}\Big)\Big),\end{aligned}$$

where we used $1-x \geq e^{-x/(1-x)}$.

We now choose $c = \frac{t}{(v-1)m}$, which satisfies Equation (C.3.46), and where we need the condition that $0 < t \leq \frac{(v-1)m}{2}$ which we check later in the proof. We write

$$\frac{16nFc^2}{1-2c} \Big/ \Big(1 - \frac{8Fc^2}{1-2c}\Big) = \frac{16nFc^2}{1-2c-8Fc^2}.$$

Fix a constant $h \in (0,1)$ whose conditions will be discussed below together with the conditions for t. Take t, h such that $1-2c-8Fc^2 \geq h \in (0,1)$. Then it follows that (since $c = \frac{t}{(v-1)m}$)

$$\frac{16nFc^2}{1-2c-8Fc^2} \leq \frac{16nt^2F}{(v-1)^2m^2h}.$$

Hence, the lower bound reads as follows:

$$\sup_{(X,S)\in\mathscr{F}} P\Big(\text{UnFairness}(\pi_n) - \inf_{\pi\in\Pi_o} \text{UnFairness}(\pi) > t\Big) \geq \frac{1}{4}\exp\Big(-\frac{16nt^2F}{(v-1)^2m^2h}\Big).$$

Let $\frac{1}{4}\exp\Big(-\frac{16nt^2F}{(v-1)^2m^2h}\Big) = \kappa$. By re-arranging the expression, we write with probability at least κ, for some distribution in $\mathscr{F}$, for all π_n,

$$\text{UnFairness}(\pi_n) - \inf_{\pi\in\Pi_o} \text{UnFairness}(\pi) \geq \sqrt{\frac{F(v-1)^2m^2\log(\frac{1}{4\kappa})}{16nh}} \tag{C.3.47}$$

where we chose $t = \sqrt{\frac{F(v-1)^2m^2\log(\frac{1}{4\kappa})}{16nh}}$.

Next, we check the condition for t, h, and characterize the constants m, h, F. Recall that the conditions are the following:

$$0 < t \leq \frac{(v-1)m}{2}, \quad 1-2c-8Fc^2 \geq h, \quad c = \frac{t}{(v-1)m}, \quad F = m(v-1)(\frac{1}{2}-c), \quad 0 < m \leq \frac{1}{v-1},$$
$$t = \sqrt{\frac{F(v-1)^2m^2\log(\frac{1}{4\kappa})}{16nh}}, \quad h \in (0,1),$$

where the first condition on t follows from Equation (C.3.46). Take first $h = F/8$. Then the first condition on t implies that $n \geq \log(1/4\kappa)$. The second condition on h (with $h = F/8$) is satisfied

if the first inequality holds

$$1 - F/8 \geq c(2+4F) \geq c(2+8Fc)$$

since $c \in (0, 1/2)$. Now, observe that $F \leq 1/2$, hence it suffices to show that

$$c \leq \frac{1-1/16}{4} \Rightarrow \sqrt{\frac{\log(\frac{1}{4\kappa})}{2n}} \leq \frac{15}{64} \Rightarrow n \geq \bar{C}\log(1/4\kappa),$$

for a finite constant $\bar{C}$. The proof completes since the remaining conditions can be satisfied for an arbitrary choice of $0 < m < 1/(v-1)$.

We are left to show that the claim holds if $(v-1)/2$ is not an integer. For this case we follow the same steps of the proof where we construct a set of distributions $\mathscr{F}$ which puts mass m on $v-2$ $x_i, i < v-1$ and mass $\frac{1-(v-2)m}{2}$ on the remaining x_{v-1}, x_v. We construct a bit vector $b \in \mathscr{B} \subset \{0,1\}^{v-2}$ with $\sum_{i=1}^{v-2} b_i = \frac{v-2}{2}$ which must be equal to an integer since $\frac{v-1}{2}$ is not. We construct (since $v \geq 3$)

$$S = \begin{cases} 1 \text{ if } U \leq p_1 - c + 2cb_i, X = x_i, i < v-1 \\ 1 \text{ if } U \leq p_1, X = x_i, i \in \{v-1, v\} \\ 0 \text{ otherwise} \end{cases},$$

while the remaining part of the proof follows similarly to above.

Envy-Freeness UnFairness

Lemma C.3.5. *Under Assumption3.4.1, 3.4.2, 3.4.3, 3.5.1, 3.5.2, the following holds: with probability at least* $1-\gamma$*,*

$$\sup_{\pi \in \Pi} \left| \mathscr{A}(s, s'; \pi) - \mathscr{A}_n(s, s'; \pi) \right| \leq \frac{cM}{\delta^2} \sqrt{\frac{\log(2/\gamma)}{n}} + \frac{c}{\delta} n^{-\eta} + \sqrt{\frac{v}{n}} \tag{C.3.48}$$

for a universal constant $c < \infty$.

Proof of Lemma C.3.5. We consider the case where $s' \neq s$, whereas $s' = s$ follows trivially. Observe that we can write

$$\Big|\mathscr{A}(s,s';\pi) - \mathscr{A}_n(s,s';\pi)\Big| \leq$$
$$\underbrace{\Big|\mathbb{E}_{X(s)}\Big[V_{\pi(X(s),s)}(X(s),s')\Big] - \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{1,s'}(X_i)\pi(X_i,s) + \frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{0,s'}(X_i)(1-\pi(X_i,s))\Big)\Big|}_{(A(\pi))}$$
$$+ \underbrace{|\hat{W}_{s'}(\pi) - W_{s'}(\pi)|}_{(B(\pi))}. \tag{C.3.49}$$

The term $(B(\pi))$ is bounded uniformly as discussed in Lemma C.3.3. Therefore, we are only left to discuss bounds on $(A(\pi))$. To derive bounds in such a scenario, we first observe that we can write

$$\begin{aligned}
\sup_{\pi\in\Pi} A(\pi) \leq & \underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s) - \mathbb{E}\Big[\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)\Big]\Big|}_{(I)} \\
& + \underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}m_{0,s'}(X_i)(1-\pi(X_i,s)) - \mathbb{E}\Big[\frac{1\{S_i=s\}}{p_s}m_{0,s'}(X_i)(1-\pi(X_i,s))\Big]\Big|}_{(II)} \\
& + \underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{1,s'}(X_i) - \frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\Big)\pi(X_i,s)\Big|}_{(III)} \\
& + \underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1\{S_i=s\}}{\hat{p}_s}\hat{m}_{0,s'}(X_i) - \frac{1\{S_i=s\}}{p_s}m_{0,s'}(X_i)\Big)(1-\pi(X_i,s))\Big|}_{(IV)}.
\end{aligned}$$

We discuss (I) and (III), whereas (II) and (IV) follow similarly. Observe first that by Assumption 3.4.1 and the bounded difference inequality, with probability $1-\gamma$,

$$\begin{aligned}
&\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s) - \mathbb{E}\Big[\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)\Big]\Big| \\
&\leq \mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s) - \mathbb{E}\Big[\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)\Big]\Big|\Big] + \bar{C}\frac{M}{\delta}\sqrt{\log(2/\gamma)/n}
\end{aligned}$$

for a constant $\bar{C} < \infty$. Under Assumption 3.5.1 each summand is centered around zero. Using the symmetrization argument (Van Der Vaart and Wellner, 1996), we have

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)-\mathbb{E}\Big[\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)\Big]\Big|\Big] \leq$$
$$2\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)\Big|\Big]$$

where σ_i are $i.i.d.$ Radamacher random variables. Since $m_{1,s}$ is uniformly bounded and similarly p_s is bounded, and by Assumption 3.4.1, we obtain by the properties of the Dudley's entropy integral (Wainwright, 2019),

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\frac{1\{S_i=s\}}{p_s}m_{1,s'}(X_i)\pi(X_i,s)\Big|\Big] \leq \bar{C}\frac{M}{\delta}\sqrt{v/n}$$

for a universal constant $\bar{C} < \infty$. We now move to bound (III). Using the triangular inequality and Holder's inequality, we obtain

$$(III) \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}\Big|m_{1,s'}(X_i)-\hat{m}_{1,s'}(X_i)\Big| \tag{C.3.50}$$

The above bound is deterministic and it does not depend on π. Observe now that by consistency of potential outcomes and covariates

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}\Big|m_{1,s'}(X_i)-\hat{m}_{1,s'}(X_i)\Big| \\ =&\frac{1}{n}\sum_{i=1}^{n}\frac{1\{S_i=s\}}{p_s}\Big|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))\Big| \leq \frac{1}{n\delta}\sum_{i=1}^{n}\Big|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))\Big|.\end{aligned} \tag{C.3.51}$$

We now separate the contribution of each of the K folds using in the cross-fitting algorithm.

Namely, we define

$$\frac{1}{n\delta}\sum_{i=1}^{n}\left|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))\right| \leq \sum_{k\in\{1,\ldots,K\}}\frac{1}{n\delta}\sum_{i\in\mathscr{I}_k}\left|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}^{(-k(i))}(X_i(s))\right| \tag{C.3.52}$$

where $\mathscr{I}_k$ denotes the set of indexes in fold k, and $\hat{m}_{1,s'}^{(-k(i))}$ denotes the estimator obtained from all folds except k. Next, we bound the following term using Liaponuv inequality:

$$\frac{1}{n}\sum_{i\in\mathscr{I}_k}\mathbb{E}\Big[|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))|\Big] \lesssim \sqrt{\mathbb{E}\Big[|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))|^2\Big]} \leq cn^{-\eta}. \tag{C.3.53}$$

The last inequality follows by Assumption 3.5.2, for a universal constant $c<\infty$. Finally, we discuss exponential concentration of the empirical counterpart. By boundeness of $\hat{m}$ in Assumption 3.4.3, we have

$$\sup_{x\in\mathscr{X}}\left|m_{d,s'}(x)-\hat{m}_{d,s'}(x)\right| \leq 2M. \tag{C.3.54}$$

By the bounded difference inequality, with probability at least $1-\gamma$,

$$\frac{1}{n}\sum_{i\in\mathscr{I}_k}\left|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))\right| \leq \mathbb{E}\Big[\Big|m_{1,s'}(X_i(s))-\hat{m}_{1,s'}(X_i(s))\Big|\Big]+4M\sqrt{\log(2/\gamma)/n}. \tag{C.3.55}$$

Combining the above bounds, the proof completes. □

Corollary 14. *Theorem 3.5.1 holds.*

Proof. This follows from Theorem 3.5.1 and Lemma C.3.2. □

Regret Bounds for $|D(\pi)|$, and $|C(\pi)|$

To obtain UnFairness bounds for unfairness being defined as either $D(\pi)$ or $C(\pi)$ *in absolute value* it suffices to bound the following empirical processes

$$\sup_{\pi\in\Pi}\Big||\hat{C}(\pi)|-|C(\pi)|\Big|, \quad \sup_{\pi\in\Pi}\Big||\hat{D}(\pi)|-|D(\pi)|\Big|.$$

We bound the first on the left-hand side while the second follows similarly. We write by the reverse triangular inequality

$$\sup_{\pi\in\Pi}\Big||\hat{C}(\pi)|-|C(\pi)|\Big| \leq \sup_{\pi\in\Pi}\Big|\hat{C}(\pi)-C(\pi)\Big|.$$

The rest of the proof follows similarly to Theorem 3.5.1.

Proof of Theorem C.1.2

We write

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\widehat{P}_s(\pi)}{1-p_s}-\frac{P_s(\pi)}{1-p_s}\Big|\Big] \\
&\leq \underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\sum_{i=1}^n \pi(X_i)S_i\Big\{\frac{(Y_i-\hat{m}_1(X_i,S_i))D_i}{\hat{e}(X_i,S_i)}+\hat{m}_1(X_i,S_i)\Big\}}{np_1\mathbb{P}(\pi(X)=1|S=1)}-\mathbb{E}\Big[Y|\pi(X)=1,S=1\Big]\Big|\Big]}_{(A)} \\
&+\underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\sum_{i=1}^n \pi(X_i)(1-S_i)\Big\{\frac{(Y_i-\hat{m}_1(X_i,S_i))D_i}{\hat{e}(X_i,S_i)}+\hat{m}_1(X_i,S_i)\Big\}}{n\mathbb{P}(\pi(X)=1|S=0)(1-p_1)}-\mathbb{E}\Big[Y|\pi(X)=1,S=0\Big]\Big|\Big]}_{(B)}.
\end{aligned}$$

We study (A) while (B) follows similarly. First, we write

$$(A) \leq \underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\sum_{i=1}^n \pi(X_i)S_i\Big\{\frac{(Y_i-\hat{m}_1(X_i,S_i))D_i}{\hat{e}(X_i,S_i)}+\hat{m}_1(X_i,S_i)-\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}-m_1(X_i,S_i)\Big\}}{np_1\mathbb{P}(\pi(X)=1|S=1)}\Big|\Big]}_{(I)}$$
$$+\underbrace{\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\sum_{i=1}^n \pi(X_i)S_i\Big\{\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}+m_1(X_i,S_i)\Big\}}{np_1\mathbb{P}(\pi(X)=1|S=1)}-\mathbb{E}\Big[Y|\pi(X)=1,S=1\Big]\Big|\Big]}_{(II)}.$$

We study (I) first. Define

$$V_n(\pi)=\frac{1}{np_1}\sum_{i=1}^n \pi(X_i)S_i\Big\{\frac{(Y_i-\hat{m}_1(X_i,S_i))D_i}{\hat{e}(X_i,S_i)}+\hat{m}_1(X_i,S_i)-\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}-m_1(X_i,S_i)\Big\}.$$

We have

$$(I)\leq\frac{1}{\kappa}\mathbb{E}\Big[\underbrace{\sup_{\pi\in\Pi}|V_n(\pi)|}_{(a)}\Big].$$

We write

$$\begin{aligned}(a)\leq&\underbrace{\frac{1}{\delta}\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n S_iD_i(Y_i-m_{1,S_i}(X_i))\Big(\frac{1}{e(X_i,S_i)}-\frac{1}{\hat{e}(X_i,S_i)}\Big)\pi(X_i,S_i)\Big|}_{(j)}\\&+\underbrace{\frac{1}{\delta}\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^n\Big(\frac{D_i}{\hat{e}(X_i,S_i)}-1\Big)(m_{1,S_i}(X_i)-\hat{m}_{1,S_i}(X_i))\pi(X_i,S_i)S_i\Big|}_{(jj)}.\end{aligned}\tag{C.3.56}$$

We study (j) and (jj) separately. We start from (j). Recall, that by cross fitting $\hat{e}(X_i,S_i)=\hat{e}^{-k(i)}(X_i,S_i)$, where $k(i)$ is the fold containing unit i. Therefore, observe that given the K folds for cross-fitting, we have

$$\begin{aligned}&\Big|\frac{1}{n}\sum_{i=1}^n S_iD_i(Y_i-m_{1,S_i}(X_i))\Big(\frac{1}{e(X_i,S_i)}-\frac{1}{\hat{e}(X_i,S_i)}\Big)\pi(X_i,S_i)\Big|\\&\leq\sum_{k\in\{1,\ldots,K\}}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}S_iD_i(Y_i-m_{1,S_i}(X_i))\Big(\frac{1}{e(X_i,S_i)}-\frac{1}{\hat{e}^{(-k(i))}(X_i,S_i)}\Big)\pi(X_i,S_i)\Big|.\end{aligned}\tag{C.3.57}$$

In addition, we have that

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{i\in\mathscr{I}_k} S_i D_i (Y_i - m_{1,S_i}(X_i))\Big(\frac{1}{e(X_i,S_i)} - \frac{1}{\hat{e}^{(-k(i))}(X_i,S_i)}\Big)\pi(X_i,S_i)\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big[\sum_{i\in\mathscr{I}_k} S_i D_i (Y_i - m_{1,S_i}(X_i))\Big(\frac{1}{e(X_i,S_i)} - \frac{1}{\hat{e}^{(-k(i))}(X_i,S_i)}\Big)\pi(X_i,S_i)\Big|\hat{e}^{(-k(i))}\Big]\Big] = 0,
\end{aligned} \tag{C.3.58}$$

by cross-fitting. By Assumption 3.4.3, we know that

$$\sup_{x\in\mathscr{X},s\in\mathscr{S}}\Big|\frac{1}{e(x,s)} - \frac{1}{\hat{e}^{(-k(i))}(x,s)}\Big| \leq 2/\delta^2 \tag{C.3.59}$$

and therefore each summand in Equation (C.3.57) is bounded by a finite constant $2/\delta^2$. We now obtain, using the symmetrization argument (Van Der Vaart and Wellner, 1996), and the Dudley's entropy integral (Wainwright, 2019)

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}|\frac{1}{n}\sum_{i\in\mathscr{I}_k} S_i D_i (Y_i - m_{1,S_i}(X_i))\Big(\frac{1}{e(X_i,S_i)} - \frac{1}{\hat{e}^{(-k(i))}(X_i,S_i)}\Big)\pi(X_i,S_i)|\Big|\hat{e}^{(-k(i))}\Big] \lesssim \frac{M}{\delta^2}\sqrt{v/n}. \tag{C.3.60}$$

We now consider the term (jj). Observe that we can write

$$\begin{aligned}
(jj) \leq &\underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{D_i}{\hat{e}(X_i,S_i)} - \frac{D_i}{e(X_i,S_i)}\Big)(m_{1,S_i}(X_i) - \hat{m}_{1,S_i}(X_i))S_i\pi(X_i,S_i)\Big|}_{(v)} \\
&+ \underbrace{\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{D_i}{e(X_i,S_i)} - 1\Big)(m_{1,S_i}(X_i) - \hat{m}_{1,S_i}(X_i))\pi(X_i,S_i)S_i\Big|}_{(vv)}.
\end{aligned} \tag{C.3.61}$$

We consider each term seperately. Consider (vv) first. Using the cross-fitting argument we obtain

$$\begin{aligned}
&\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{D_i}{e(X_i,S_i)}-1\Big)(m_{1,S_i}(X_i)-\hat{m}_{1,S_i}(X_i))\pi(X_i,S_i)S_i\Big| \\
&\leq \sum_{k\in\{1,\dots,K\}}\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}\Big(\frac{D_i}{p_s e(X_i,S_i)}-1\Big)(m_{1,S_i}(X_i)-\hat{m}^{(-k(i))}_{1,S_i}(X_i))\pi(X_i,S_i)S_i\Big|.
\end{aligned} \tag{C.3.62}$$

Observe now that

$$\mathbb{E}\Big[\Big(\frac{D_i}{e(X_i,S_i)}-1\Big)(m_{1,S_i}(X_i)-\hat{m}^{(-k(i))}_{1,S_i}(X_i))\pi(X_i,S_i)S_i\Big|\hat{m}^{(-k(i))}_{1,S_i}\Big]=0. \tag{C.3.63}$$

Therefore, following the same argument discussed before,

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{n}\sum_{i\in\mathscr{I}_k}\Big(\frac{D_i}{e(X_i,S_i)}-1\Big)(m_{1,S_i}(X_i)-\hat{m}^{(-k(i))}_{1,S_i}(X_i))\pi(X_i,S_i)S_i\Big|\Big]\lesssim\frac{M}{\delta^2}\sqrt{\frac{v}{n}}. \tag{C.3.64}$$

We are now left to bound (v). We obtain that

$$(v)\leq\sqrt{\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1}{\hat{e}(X_i,S_i)}-\frac{1}{e(X_i,S_i)}\Big)^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_{1,S_i}(X_i)-\hat{m}_{1,S_i}(X_i))^2}. \tag{C.3.65}$$

Using Jensen inequality and Assumption 3.4.3 $\mathbb{E}[(v)]\lesssim n^{-1/2}$.

We now move to bound the expectation of (II). First, observe that by Lemma C.1.1, and standard properties of the double-robust estimator, we have that

$$\mathbb{E}\Big[\frac{\frac{1}{p_1 n}\sum_{i=1}^{n}\pi(X_i)S_i\Big\{\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}+m_1(X_i,S_i)\Big\}}{\mathbb{P}(\pi(X)=1|S=s)}\Big]=\mathbb{E}\Big[Y(1)|\pi(X)=1,S=1\Big].$$

Using the symmetrization argument (see Van Der Vaart and Wellner (1996)), we have

$$(II)\leq 2\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\frac{1}{p_1 n}\sum_{i=1}^{n}\pi(X_i)\sigma_i S_i\Big\{\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}+m_1(X_i,S_i)\Big\}}{\mathbb{P}(\pi(X)=1|S=s)}\Big|\Big],$$

where $\{\sigma_i\}$ are *i.i.d.* exogenous Radamacher random variables. Using the assumption that $P(\pi(X)=1|S=s)\in(\kappa,1-\kappa)$, we write

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{\frac{1}{p_1 n}\sum_{i=1}^{n}\pi(X_i)\sigma_i S_i\Big\{\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}+m_1(X_i,S_i)\Big\}}{\mathbb{P}(\pi(X)=1|S=s)}\Big|\Big]$$
$$\leq\frac{1}{\kappa}\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{p_1 n}\sum_{i=1}^{n}\pi(X_i)\sigma_i S_i\Big\{\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}+m_1(X_i,S_i)\Big\}\Big|\Big].$$

We now proceed using a standard argument. Using the fact that each summand in the above expression are uniformly bounded, and Π has finite VC-dimension, using the Dudley's entropy integral bound, it directly follows that

$$\mathbb{E}\Big[\sup_{\pi\in\Pi}\Big|\frac{1}{p_1 n}\sum_{i=1}^{n}\pi(X_i)\sigma_i S_i\Big\{\frac{(Y_i-m_1(X_i,S_i))D_i}{e(X_i,S_i)}+m_1(X_i,S_i)\Big\}\Big|\Big]\lesssim\sqrt{\frac{v}{n}}$$

which concludes the proof.

Appendix D

Appendix to Chapter 4

Main Extensions

Throughout the appendix, we will often suppress the dependence of the hypothesis testing protocol on J and G and simply write $r(X)$.

Imperfectly Informed Researchers

In Section 4.2 we assume that the researcher is perfectly informed and knows θ. Here we show that our main results continue to hold in settings where the researcher has imperfect information in the form of a prior about θ.[1] Denote this prior by $\pi \in \Pi$, where Π is the class of *all* distributions supported on Θ. Throughout this section, we assume that Π is unrestricted.[2] The prior π represents knowledge about θ that is available to both the researcher and the policy-maker, but not to the social planner.[3]

We assume that the vector of statistics X is drawn from a normal distribution conditional on θ, where θ itself is drawn from the prior π:

$$X \mid \theta \sim \mathcal{N}(\theta, \Sigma), \quad \theta \sim \pi, \quad \pi \in \Pi,$$

[1] In the single-hypothesis testing case, Tetenov (2016) gives results under imperfect information. However, these results rely on the Neyman-Pearson lemma, which is not applicable with multiple tests.

[2] This assumption is made for simplicity. For our theoretical results, we only need that the class of priors Π contains at least one element that is supported on the null space Θ_0, which holds by construction if Π is unrestricted.

[3] If the prior π was known to the social planner, she would act as a Bayesian decision-maker.

where Σ is positive definite.

The researcher acts as a Bayesian decision maker. She experiments whenever her ex-ante utility $\bar{\beta}_r(\pi)$ is positive, where

$$\bar{\beta}_r(\pi) = \int \beta_r(\theta) d\pi(\theta).$$

Welfare is

$$\bar{v}_r(\pi) = \begin{cases} \int\int u(\theta)^\top \delta(r(x)) dF_\theta(x) d\pi(\theta) & \text{if } \bar{\beta}_r(\pi) > 0 \\ \max\left\{ \int\int u(\theta)^\top \delta(r(x)) dF_\theta(x) d\pi(\theta), 0 \right\} & \text{if } \bar{\beta}_r(\pi) = 0 \\ 0 & \text{if } \bar{\beta}_r(\pi) < 0. \end{cases}$$

Under imperfect information, we define maximin rules with respect to the prior π, which is known to the researcher and policy-maker.

Definition D.1.1 (Π-maximin optimal)**.** We say that r^* is Π-maximin optimal if and only if

$$r^* \in \arg\max_{r \in \mathscr{R}} \inf_{\pi \in \Pi} \bar{v}_r(\pi).$$

Definition D.1.1 generalizes the notion of maximin optimality in Section 4.2.3, which is stated in terms of the parameter θ. When Π contains only point mass distributions, the two notions of maximin optimality are equivalent.

The following lemma provides a characterization of maximin recommendation functions.

Lemma D.1.1 (Conditions for maximin optimality)**.** *Let $\Theta_0 \neq \emptyset$ as in Definition 4.2.1. The recommendation function r^* is Π-maximin optimal (Definition D.1.1), if and only if* $\inf_{\pi \in \Pi} \bar{v}_{r^*}(\pi) \geq 0$.

Proof. See Appendix D.3.1. □

Lemma D.1.1 states that maximin optimality is equivalent to the worst-case welfare being non-negative. Based on this result, the next proposition shows that one-sided t-tests

with appropriately chosen critical values are maximin optimal and admissible under imperfect information.

Proposition D.1.2 (Maximin optimality and admissibility)**.** *Let $J > 1$. Let Assumption 4.2.1, 4.2.2, 4.2.3, and 4.2.4 hold. Suppose that $\Theta = [-1,1]^J$ and $C(J) > 0$. Then the recommendation function*

$$r_j^*(X) = 1\left\{X_j/\sqrt{\Sigma_{j,j}} \geq \Phi^{-1}\left(1 - C(J)/J\right)\right\}, \quad \forall j \in \{1,\ldots,J\}$$

is Π-maximin optimal. In addition, r^* is also admissible with respect to any $\pi \in \Pi$.

Proof. See Appendix D.3.2. The proof of the first result (maximin optimality) uses the duality properties of the linear program. The second result is a consequence of Proposition 4.2.5. The assumption that $\Theta = [-1,1]^J$ is made for simplicity and can be replaced with a compactness assumption on Θ. □

Proposition D.1.2 shows that the conclusions in Section 4.2.5 on the maximin optimality of t-tests remain valid under imperfect information. Proposition D.1.2 further states that r^* is admissible. While maximin optimality connects to size control in hypothesis testing, admissibility captures a notion of power. It implies that we cannot find any other decision rule that is more powerful than r^* for all prior distributions $\pi \in \Pi$.

Remark 19 (Locally most powerful rules under imperfect information)**.** Unlike the result under perfect information in Proposition 4.2.5, we do not characterize locally most powerful rules in Proposition D.1.2. This is because with imperfect information, there are several different notions of locally most powerful recommendation functions. For instance, we may consider local alternatives that assign ε probability to a certain parameter value in the alternative space, or alternatives that assign probability one to the parameter taking value ε. The first notion implicitly imposes a certain prior on a specific value of the parameter, which may be hard to justify in practice. The second notion coincides with the one discussed in the main text and is satisfied by the recommendation function in Proposition D.1.2. □

Our results for imperfectly informed researchers can be extended to the settings with threshold crossing publication rules and general welfare functions as in Section 4.2.6 and Appendix D.1.3. In these sections, maximin optimality is defined with respect to Θ_0 only. We can impose an equivalent condition under imperfect information with respect to

$$\Pi^{\text{weak}} = \left\{ \pi : \int_{\theta \in \Theta_0} \pi(\theta) d\theta = 1 \right\}.$$

The set Π^{weak} denotes the set of priors that impose mass one on the null space Θ_0. The notion of weak maximin optimality and the characterization of weakly maximin decisions is the same as the one in Equation (D.1.3). To see why, note that any decision rule that satisfies $\beta_r(\theta) \leq 0, \forall \theta \in \Theta_0$, also satisfies $\int \beta_r(\theta) d\pi(\theta) \leq 0$ for all $\pi \in \Pi^{\text{weak}}$.

Endogenous Number of Hypotheses

In Section 4.2, we assume that the number of hypotheses, J, is exogenous. Here we relax this assumption and consider a setting where the researcher can choose which and how many hypotheses to test.

The model is similar to that in Section 4.2 with small modifications. First, the social planner commits to a family of hypothesis testing protocols for any number of selected policies, $S \leq J$, $\{r_S\}_{S \in \{1,\ldots,J\}}$. Given $\{r_S\}_{S \in \{1,\ldots,J\}}$, the researcher chooses an arbitrary subset of treatments indexed by $s^* \in \{0,1\}^J$ out of the J possible treatments and decides whether to experiment. Conditional on experimentation, the researcher draws a vector $X_{s^*} \in \mathbb{R}^{S^*}$, where $S^* = \sum_{j=1}^{J} s_j^*$, containing information on the selected treatments and reports $r^*_{S^*}(X_{s^*})$. We assume that for the treatments that the researcher does not select, the status quo remains in place. This structure of the model reflects a setting where the social planner commits to a recommendation functions for any number of treatments, and the researcher then selects which treatment(s) to analyze.[4]

Formally, define an indicator $s_j \in \{0,1\}$ such that $s_j = 1$ if treatment j is selected by the

[4]Extensions to the case where r depends on the identity of selected treatments are possible but omitted for brevity.

researcher and $s_j = 0$ otherwise. If the researcher experiments, her utility is

$$\sum_{j=1}^{J} s_j r_{S,j}(X_s) - C(S), \quad S = \sum_{j=1}^{J} s_j.$$

Similar to Section 4.2, the utility is linear in the number of discoveries, and the cost depends on the number of tests S that the researcher conducts (for simplicity, we are assuming that the cost of each test is the same).

We assume that welfare is linear and additive $u_{r_S,s}(X_s) = \sum_{j=1}^{J} s_j \theta_j r_{S,j}(X_s)$. Under Assumption 4.2.3), we can write welfare as

$$v_r^E(\theta) = \begin{cases} \mathbb{E}_\theta\left[\sum_{j=1}^{J} s_j^* \theta_j r_{S^*,j}(X_{s^*})\right] & \text{if } \mathbb{E}_\theta\left[\sum_{j=1}^{J} s_j^* r_{S^*,j}(X_{s^*})\right] - C(S^*) > 0 \\ \max\left\{0, \mathbb{E}_\theta[u_{r_{S^*},s^*}(X_{s^*})]\right\} & \text{if } \mathbb{E}_\theta\left[\sum_{j=1}^{J} s_j^* r_{S^*,j}(X_{s^*})\right] - C(S^*) = 0 \\ 0 & \text{otherwise,} \end{cases} \tag{D.1.1}$$

where $S^* = \sum_{j=1}^{J} s_j^*$ and

$$s^* \in \arg\max_{s \in \{0,1\}^J} \sum_{j=1}^{J} s_j \mathbb{E}_\theta\left[r_{\sum_{j=1}^{J} s_j, j}(X_s)\right] - C\Big(\sum_{j=1}^{J} s_j\Big).$$

To alleviate the exposition, we omit the dependence of s^* and S^* on θ (that is, it should read $s^*(\theta)$ and $S^*(\theta)$).

The following lemma characterizes the class of maximin decisions.

Lemma D.1.3. *Let $\Theta = [-1,1]^J$ and consider the model described in the current section. Then r^* is maximin optimal, i.e.,*

$$r^* \in \arg\max_{r \in \mathscr{R}} \min_{\theta \in \Theta} v_r^E(\theta)$$

if and only if $v_{r^*}^E(\theta) \geq 0$ for all $\theta \in \Theta$.

Proof. See Appendix D.3.1. □

We conclude our discussion by showing that separate t-tests are also maximin optimal and locally most powerful when the number of hypotheses is selected endogenously. To do so, we introduce a slightly modified notion of local power.

Definition D.1.2 (Local power with endogenous treatments)**.** We say that r is locally more powerful than r' if

$$\left[\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\inf_{\theta\in\Theta_1^E(\varepsilon)} v_{r^*}^E(\theta)-\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\inf_{\theta\in\Theta_1(\varepsilon)} v_{r'}^E(\theta)\right]\geq 0,$$

where $\Theta_1^E(\varepsilon)=\big\{\theta:\theta_j\geq\varepsilon \text{ for some } j,\theta_{-j}=0\big\}$.

Definition D.1.2 modifies Definition 4.2.4 as follows. Instead of allowing $\theta_{-j}\geq 0$, it imposes that parameters are either ε (or larger) or zero. The main reason is technical: if the parameters can be smaller than ε but larger than zero, we can choose $\theta_{-j}\geq 0$ but smaller than ε (e.g., ε^2) so that the researcher selects all J treatments. However, each treatments leads to a welfare effect that converges to zero at a rate faster than ε (i.e., $o(\varepsilon)$).

The following proposition shows that, under some additional conditions, one-sided t-tests remain optimal even when the researcher selects the hypotheses she wants to test. Different from the case of exogenous J, here the threshold depends on the number of tests selected by the researcher. In what follows, we say that a hypothesis testing protocol exhibits symmetric size control if $P(r'_{S,j}(X_s)=1|\theta=0)=P(r'_{S,j'}(X_s)=1|\theta=0)$ for all $j,j'\in\{1,\ldots,S\}$ and $S\in\{1,\ldots,J\}$. As discussed in the main text, symmetry is a desirable property absent additional restrictions on the relative importance of the different hypotheses.

Proposition D.1.4 (Maximin optimality and local power)**.** *Let $J>1$. Suppose that Assumptions 4.2.1, 4.2.2, 4.2.3 hold and that $X_s\sim\mathcal{N}(\theta_s,\Sigma^s)$ for all $s\in\{0,1\}^J$. Suppose further that $\Theta=[-1,1]^J$ and $S\geq C(S)>0$ for all $S\in\{1,\ldots,J\}$. Then the protocol*

$$r^*_{S,j}(X_s)=1\Big\{X_{s,j}/\sqrt{\Sigma^s_{j,j}}\geq\Phi^{-1}\left(1-C(S)/S\right)\Big\},\quad\forall j\in\{1,\ldots,J\},S\in\{1,\ldots,J\},\quad\text{(D.1.2)}$$

is maximin optimal. If $C(S) = C(1)$ for all S, $\{r^*_{S,j}\}$ is locally most powerful among maximin protocols. If $C(S)$ is strictly increasing in S with $C(1) \geq C(S)/S$, $\{r^*_{S,j}\}$ is locally most powerful among maximin protocols with symmetric size control.

Proof. See Appendix D.3.2. □

Proposition D.1.4 states two results. First, one-sided t-tests are maximin optimal for any cost function $C(S)$. Second, they are also locally most powerful if either the costs are constant, or, if the costs are increasing in S, but we restrict attention to protocols with symmetric size control.

Additional Forms of Interactions Between Treatments

Section 4.2.6 provides a brief discussion of settings with additional forms of interactions. Here we present the detailed formal treatment of these settings. We focus on the case where the researcher knows θ. See Appendix D.1.1 for settings with imperfectly informed researchers.

Weakly Maximin Recommendation Functions

Maximin recommendation functions may be very conservative when we start considering additional forms of interactions between the treatments. Therefore, we introduce a weaker notion of maximin optimality that considers the worst case over $\Theta_0 := \{\theta : u_j(\theta) < 0 \text{ for all } j\}$ only, instead of Θ. It corresponds to the concept of *weak size control* in the MHT literature. Even under this weaker criterion we will often obtain conservative hypothesis testing protocols, which helps motivate attention to it.

Definition D.1.3 (Weak maximin optimality)**.** We say that r^* is weakly maximin if and only if

$$r^* \in \arg\max_{r \in \mathscr{R}} \min_{\theta \in \Theta_0} v_r(\theta).$$

Following the same argument as in the proof of Proposition 4.2.1, we can show that r^* is

weakly maximin if and only if[5]

$$\beta_{r^*}(\theta) \leq 0 \quad \text{for all } \theta \in \Theta_0. \tag{D.1.3}$$

Different from the notion of maximin optimality, Definition D.1.3 considers the worst-case allocation over the set Θ_0 instead of Θ. It is a weaker notion of optimality since it requires size control only over the subset of parameters that lead to negative treatment effects for each possible recommendation function. That is, it imposes size control only under the weak null.[6] By definition, all maximin protocols are also weakly maximin, while the converse is not necessarily true. A weakly maximin protocol is also maximin only if welfare is weakly positive over $\Theta \setminus \Theta_0$.

Linear Welfare and Threshold-Crossing Publication Rule

We now introduce interactions in the publication rule, replacing the linear rule in Assumption 4.2.2 with a threshold rule in which only papers that find sufficiently many results can be published.

Assumption D.1.1 (Threshold publication rule)**.** With a threshold publication rule, the researcher's utility conditional on experimenting is (up-to-rescaling)

$$\beta_r(\theta) = \gamma \int 1\left\{\sum_{j=1}^{J} r_j(x) \geq \kappa\right\} dF_\theta(x) - C(J) \tag{D.1.4}$$

for exogenous constants $\kappa \geq 0, \gamma > C(J)$.

The threshold crossing publication rule leads to optimal hypothesis testing protocols which depend on the joint distribution of X in a complicated way. To illustrate, consider the

[5]The formal argument is as follows: any recommendation function r yields weakly negative welfare for $\theta \in \Theta_0$. Therefore, the maximin recommendation function achieves zero utility over $\theta \in \Theta_0$, which holds if Equation (D.1.3) holds. This proves the "if" direction. If Equation (D.1.3) does not hold, then any r achieves negative utility, proving the "only if" direction.

[6]See Proschan and Brittain (2020) for related notions in the context of MHT.

leading case where $X \sim \mathcal{N}(\theta, \Sigma)$. One can show[7] that any weakly maximin and locally most powerful recommendation function must satisfy

$$P(r_j^*(X) = 1|\theta = 0) \geq p^*, \quad p^* = \min\{p_1^*, \ldots, p_J^*\}, \tag{D.1.5}$$

where $(p_1^*, \ldots, p_J^*)$ are the solutions to the following optimization problem

$$\begin{aligned}
&(p_1^*, \ldots, p_J^*) \in \arg \max_{p \in [0,1]^J} \min_{j \in \{1,\ldots,J\}} p_j \\
&\text{such that } \sum_{k \in \mathscr{K}} P(\delta_k(r(X)) = 1|\theta) \leq C(J)/\gamma \quad \forall \theta \in \Theta_0 \\
&\text{and } P(r_j(X) = 1|\theta = 0) = p_j \quad \forall j \in \{1, \cdots, J\}.
\end{aligned}$$

The above expression shows that most powerful recommendation functions impose that the probability of discovery of each separate treatment exceeds a certain (uniform) threshold p^*. The threshold depends on the joint distribution of the entries of X, which rules out separate size control if there is dependence between the entries of X.

Therefore, we restrict attention to the class of independent recommendation functions.

Assumption D.1.2 (Independent recommendation functions)**.** Consider a class of recommendation functions $r \in \mathscr{R}^{\text{ind}}$ with $r_j(X) \perp r_{j'}(X)$ with $j \neq j'$.

Assumption D.1.2 states that tests for distinct treatments are statistically independent. This holds under the normality Assumption 4.2.4 when Σ is a diagonal matrix and $r_j(X)$ is a function of X_j only. Assumption D.1.2 allows for separating the interactions arising from the threshold crossing publication rule from those occurring because of the statistical dependence between the components of X.[8]

[7] The proof follows similarly to the proof of Lemma 4.2.4 where the worst case alternative puts mass ε on the discovery with the smallest probability. As a result, the social planner wants to maximize the minimal probability across all discoveries.

[8] Independence assumptions have been commonly used as a starting point for developing approaches to multiple testing (e.g., Benjamini and Liu, 1999; Finner and Roters, 2001), and provide an interesting benchmark for contrasting our results against existing procedures and recommendations.

The following proposition characterizes the locally most powerful maximin recommendation functions under independence.

Proposition D.1.5 (Optimality of separate size control)**.** *Let $J > 1$. Let Assumptions 4.2.1, 4.2.3, 4.2.4, D.1.1, and D.1.2 hold, and let $\Theta = [-1,1]^J$. Then any $r^* \in \mathscr{R}^{\text{ind}}$ is weakly maximin optimal and locally most powerful if and only if r^* satisfies Equation* (D.1.3) *and*

$$\begin{aligned} &\lim_{\theta \downarrow 0} P\Big(r_j^*(X) = 1 | \theta\Big) = P(r_j^*(X) = 1 | \theta = 0) = p^* \quad \forall j \in \{1,\ldots,J\}, \\ &\text{where } p^* : \sum_{k \in \{\kappa,\cdots,J\}} \binom{J}{k} (p^*)^k = C(J)/\gamma, \end{aligned} \tag{D.1.6}$$

assuming such r^* exists.

Proof. See Appendix D.3.2. We note that the proof does not rely on normality of X (Assumption 4.2.4). We only require that X is continuously distributed with CDF F_θ, which admits a PDF $f_\theta(x)$ that is continuous in θ for all $x \in \mathscr{X}$. □

Proposition D.1.5 states that the optimal recommendation function involves separate size control and assigns to each false discovery the same probability p^*, which depends both on the number of hypotheses J and the threshold number of rejections κ needed for publication.

An immediate implication of Proposition D.1.5 is that under Assumption 4.2.4, the threshold crossing protocol of the form

$$r_j^*(X) = 1\Big\{X/\sqrt{\Sigma_{j,j}} \geq \Phi^{-1}(1-p^*)\Big\}, \quad \forall j \in \{1,\ldots,J\},$$

is (weakly) maximin optimal. The critical value depends on p^* defined in Proposition D.1.5.

As with a linear publication rule, a central implication of Proposition D.1.5 is that the way the size p^* of hypothesis tests should vary with the number J of hypotheses tested depends on the structure of the research cost function $C(J)$. To better understand the dependence of p^* on J, it is useful to analyze settings with very many treatments ($J \to \infty$). We consider two different

cases: one where $C(J)/\gamma$ is constant, and one where $C(J)/\gamma = 1/J$. The first case is essentially equivalent to assuming a constant publication probability and constant costs in the number of discoveries. The second case corresponds to assuming that the researcher's utility is increasing in J.

Corollary 15 (Asymptotic approximation)**.** *Assume that κ is fixed. Let p^* be as defined in Proposition D.1.5. Suppose that $C(J)/\gamma = \alpha$ for a constant $1 > \alpha > 0$ that does not depend on J. Then $p^* \asymp 1/J$ as $J \to \infty$. If, instead, $C(J)/\gamma = \alpha/J$, then $p^* \asymp 1/J^{(\kappa+1)/\kappa}$ as $J \to \infty$.*

Proof. See Appendix D.3.3. □

Corollary 15 shows that fixed-cost research production functions ($C(J)/\gamma = \alpha$) again rationalize Bonferroni-style corrections. Here for a threshold publication rule this holds asymptotically, as opposed to the linear publication rule case in which the result was exact. Interestingly, when $C(J)/\gamma = \alpha/J$, size control is of order $1/J^2$ for $\kappa = 1$ and approximately $1/J$ for $\kappa \gg 1$.

Figure D.1.1 plots the optimal level of size control for different values of J. It shows a comparison between the optimal level of size control under a linear publication rule (Lemma 4.2.4) and threshold crossing publication rule (Proposition D.1.5) with $\gamma = J$. We find that for any finite J the comparison of optimal test size under linear and threshold publication rules is ambiguous, depending on J and on the research cost function.

As Figure D.1.1 illustrates, the optimal level of size control with a threshold crossing publication rule also depends on the location of the publication threshold κ. One can show that p^* is increasing in the threshold κ. Intuitively, as the threshold increases, it becomes harder for the researcher to publish, and larger incentives are necessary to guarantee experimentation. As a result, for large-enough κ and fixed J, standard levels of size control such as 10% or 5% may be too stringent.

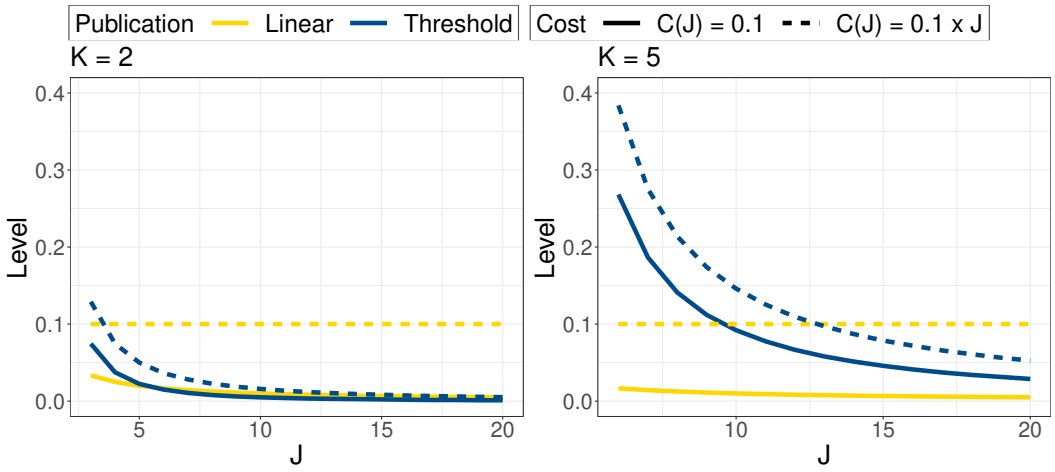


Figure D.1.1. Optimal size under linear and threshold publication rules. We set $C(J) \in \{0.1, 0.1 \times J\}$, where for the threshold crossing rule we fix $\gamma = J$. Different panels correspond to different values of κ for the threshold rule.

General Welfare and Threshold-crossing Publication Rule

We now introduce the possibility of economic interactions between the interventions being studied, in addition to interactions in the cost function $C(J)$ and the publication rule.

Assumption D.1.3 (General welfare)**.** $u_j(\theta) = \theta_j$ for all j, with $\Theta = [-1,1]^{2^J-1}$.

Importantly, unlike Assumption 4.2.1, Assumption D.1.3 allows for interactions in the welfare impact of multiple treatments. We illustrate this more general setup in the context of our running example.

Example D.1.1 (Fully saturated regression model)**.** Consider the fully saturated regression model

$$Y_i = D_{i,1}(1-D_{i,2})\theta_1 + D_{i,2}(1-D_{i,1})\theta_2 + D_{i,1}D_{i,2}\theta_3 + \varepsilon_i. \tag{D.1.7}$$

Unlike the "short" model (4.3), the "long" regression model (D.1.7) allows for interaction effects between the treatments, and θ_3 will differ from $\theta_1 + \theta_2$ in general. In this example, each entry of X corresponds to the OLS estimator of the effect of a *combination* of treatments, $X = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)^\top$. □

To state the results we define some additional notation. Let

$$\widetilde{\delta}_j(r(X)) = \delta_j(r(X))1\left\{\sum_{j=1}^{J} r_j(X) \geq \kappa\right\}, \quad \widetilde{\delta}(r(X)) \in \{0,1\}^{2^J-1} \tag{D.1.8}$$

indicate the policy decisions taken *only* in the case that the experiment results in a paper that is publishable given the threshold publication rule.

The next assumption is a generalization of Assumption 4.2.3.

Assumption D.1.4 (General tie-breaking assumption)**.** Assume that, conditional on X and for any $r \in \mathscr{R}$, the researcher does not write the paper (or equivalently, reports $r(X) = (0,\ldots,0)^\top$) if $\sum_{k=1}^{K} r_k(X) < \kappa$.

We can interpret Assumption D.1.4 as imposing some infinitesimally small cost on the writing and submission process. If the researcher knows that the paper will not be accepted after having seen X, she will not write up and submit the paper. We therefore consider the following modified version of the local alternative space.

Definition D.1.4 (General ε-alternatives)**.** For $\varepsilon > 0$, define the local alternative space as

$$\Theta_1(\varepsilon) := \Big\{\theta : u_j(\theta) \geq \varepsilon \text{ for some } j \in \mathscr{K}, u_j(\theta) \geq 0, \text{ for all } j\Big\},$$

where $\mathscr{K}$ denotes the set of indexes $k \in \{1,\ldots,2^J-1\}$ that corresponds to groups of κ or more hypotheses. □

Under Assumption D.1.4, the social planner only needs to consider groups of hypotheses for which it is profitable for the researcher to conduct research. Therefore, we can write the researcher utility and welfare as

$$\beta_r(\theta) = \gamma \int \widetilde{\delta}(r(x))^\top \mathbf{1} dF_\theta(x) - C(J) \quad \text{and} \quad v_r(\theta) = \int \widetilde{\delta}(r(x))^\top u(\theta) dF_\theta(x). \tag{D.1.9}$$

The expressions for $\beta_r(\theta)$ and $v_r(\theta)$ in Equation (D.1.9) incorporate that whenever the number discoveries does not exceed the threshold κ, the researcher does not submit the article such that the status-quo prevails.

The next proposition characterizes the optimal recommendation function for any given $\widetilde{\delta}_j$, which aggregates separate discoveries into a single recommendation. To state the proposition, we assume existence of an optimal solution. Such optimal solutions may not exist in general, and existence will depend on the distribution of X.

Proposition D.1.6 (Equal size control on compound error rates)**.** *Let $J > 1$. Suppose that Assumptions 4.2.3, 4.2.4, D.1.1, D.1.3, and D.1.4 hold. Suppose further that $\Theta = [-1,1]^{2^J-1}$. Let $\mathscr{K}$ be as in Definition 4.2.3. Then any $r^* \in \mathscr{R}$ is weakly maximin optimal and locally most powerful if and only if r^* satisfies Equation* (D.1.3) *and*

$$P\Big(\widetilde{\delta}_j(r^*(X)) = 1 | \theta = 0\Big) = \frac{C(J)}{\gamma|\mathscr{K}|} \quad \forall j \in \mathscr{K}, \tag{D.1.10}$$

assuming such r^* exists.

Proof. See Appendix D.3.2. We note that the proof does not rely on normality of X (Assumption 4.2.4). We only require that X is continuously distributed with CDF F_θ, which admits a PDF $f_\theta(x)$ that is continuous in θ for all $x \in \mathscr{X}$. □

Proposition D.1.6 shows that separate size control with a Bonferroni-type correction over each *group of discoveries* is maximin optimal and locally most powerful whenever the effect of each group of discoveries has equal weight on the researcher's utility. It rationalizes a specific form of FWER control.

Corollary 16 (Rationalization of the weak FWER)**.** *Let the conditions in Proposition D.1.6 hold. Then any MHT procedure that satisfies Equation* (D.1.10) *controls the weak FWER at level*

$C(J)/\gamma$ *at* $\theta = 0$*, namely*

$$P\left(\widetilde{\delta}_j(r^*(X)) = 1 \textit{ for at least one } j|\theta = 0\right) = C(J)/\gamma.$$

Proof. See Appendix D.3.3. □

In other words, Proposition D.1.6 rationalizes (weak) FWER control between *groups* of hypotheses sufficient for publication. Importantly, FWER is not applied to each separate discovery $r_j(X)$, but instead to each group. When a single result is sufficient for publication ($\kappa = 1$), however, this implies control of the probability of a single false rejection, i.e. of the standard notion of weak FWER control. This follows from the fact that rejecting any group of hypotheses implies rejecting its constituent members individually (for $\kappa = 1$, $\max_j \delta_j(r(X)) = 1$ if and only if $\max_j r_j(X) = 1$).

Next we provide an example of a hypothesis testing protocol that satisfies the conditions in Proposition D.1.6.

Example D.1.2 (Recommendation function with $|\mathscr{K}|$ and independence)**.** Consider a vector $\widetilde{X} \in \mathbb{R}^{|\mathscr{K}|}$, with each entry corresponding to a statistic X_j corresponding to a certain *group* of treatments with number of discoveries exceeding κ.[9] Note that $\widetilde{X}$ is a subvector of X. Assume that $X \sim \mathscr{N}(\theta, I)$, then the recommendation function[10]

$$\widetilde{\delta}_j(r(X)) = 1\left\{X_j > \max_{j' \neq j} \widetilde{X}_{j'} \text{ and } X_j > t\right\} 1\{j \in \mathscr{K}\} \tag{D.1.11}$$

is maximin and locally most powerful if t is chosen such that $P\left(\max_j \tilde{X}_j > t|\theta = 0\right) = C(J)/\gamma$.[11]

[9] Consider Example D.1.1 where $J = 2$: if $\kappa = 1$, then $|\mathscr{K}| = 2^2 - 1 = 3$; if instead, $\kappa = 2$, then $|\mathscr{K}| = 1$.

[10] This recommendation function bears some resemblance with step-down procedures in Lehmann and Romano (2005b, Chapter 9), where the maximum is considered a statistic of interest.

[11] We note that a simple threshold crossing protocol violates the constraint that $\sum_j \delta_j(r(X)) \leq 1$.

The recommendation function (D.1.11) is maximin optimal since[12]

$$P\left(\max_j \widetilde{X}_j > t|\theta\right) = P\left(\max_j(\widetilde{X}_j + \theta_j) > t|\theta = 0\right) \leq P\left(\max_j \widetilde{X}_j > t|\theta = 0\right)$$

for any $\theta \in \Theta_0 = \{\theta \in \Theta : \theta < 0\}$. It is locally most powerful because

$$\begin{aligned} &P\left(1\left\{\widetilde{X}_j > \max_{j' \neq j} \widetilde{X}_{j'} \text{ and } \widetilde{X}_j > t\right\}|\theta = 0\right) \\ &= P\left(\max_{j'} \widetilde{X}_{j'} > t|\theta = 0, \max_{j'} \widetilde{X}_{j'} \leq \tilde{X}_j\right) P\left(\max_{j'} \widetilde{X}_{j'} \leq \widetilde{X}_j|\theta = 0\right). \end{aligned}$$

□

Example D.1.2 provides an example of a maximin and locally most powerful recommendation function. The independence between the entries of X is important here; without it the existence of an optimal recommendation function is not guaranteed. Example D.1.2 also illustrates the complexity of optimal recommendation functions in the presence of potential interactions between interventions.

Additional Extensions and Details

One-sided and Two-sided Hypothesis Testing

The structure of the model we study naturally justifies one-sided hypothesis testing. Researchers test whether or not a proposed treatment strictly improves upon baseline. Treatments that lead to negative values of welfare are effectively excluded from consideration since the social planner discourages experimentation in this case. This structure is a direct consequence of the assumption that the policy-maker always implements the recommended treatment. Our framework can justify two-sided hypothesis testing under appropriate changes to the model. We outline these modifications here but leave the formal analysis of the resulting model for future

[12]Note that here we only need to consider $\widetilde{X}$, since for the remaining entries the recommendation is zero almost surely.

research.

Consider a policy-maker who randomly selects a treatment to implement using a uniform distribution when no recommendation is made. For each treatment, the researcher can make three recommendations: (i) implement the treatment; (ii) not implement the treatment; (iii) "do not know", which corresponds to the baseline, i.e., implement the treatment with some prior probability. Denote the recommendation for treatment j as $r_j(X) \in \{-1,0,1\}$, with $r_j(X) = -1$ corresponding to not implementing the treatment, $\tilde{r}_j(X) = 0$ corresponding to no recommendation, and $r_j(X) = 1$ corresponding to implementing the treatment. Consider a linear model without complementarities between treatments. The welfare generated by the recommendation $r_j(X)$ is

$$v_r(\theta) = \sum_{j=1}^{J} \int \left(1\Big\{ r_j(x) = 1 \Big\} \theta_{j,1} + 1\Big\{ r_j(x) = -1 \Big\} \theta_{j,-1} \right) dF_\theta(x).$$

The coefficients $\theta_{j,1}, \theta_{j,-1}$ capture the benefit net of the *opportunity* cost of implementing the treatment and not implementing the treatment respectively, relative to the random baseline. For the case where the policy-maker implements no treatment as baseline, then $\theta_{j,-1} = 0$, recovering our model formulation. By assuming that the baseline intervention is a random intervention, the model justifies *two-sided* hypothesis testing: not implementing a treatment has a benefit (or cost) relative to random implementation.

Σ-robust Recommendation Functions

In the main text, we assume that the experimental design and sample size (and thus the covariance matrix of X, Σ) are known to the social planner. In this setting, the planner chooses the recommendation function r to maximize worst-case welfare given Σ. Here we analyze a variant of our model in which the planner chooses r when Σ unknown and adversarially chosen by nature. We refer to recommendation functions that are maximin optimal in this setting as Σ-robust.

Definition D.2.1 (Σ-robust)**.** We say that r^* is Σ-robust if

$$r^* \in \arg\max_{r \in \mathscr{R}} \min_{\theta, \Sigma_{j,j} > \kappa \forall j} v_r(\theta), \text{ for some } \kappa > 0.$$

Definition D.2.1 states that the rule r^* is Σ-robust if it is maximin optimal not only with respect to θ but also with respect to Σ. The definition imposes a lower bound κ on the diagonal elements of the covariance matrix, ensuring that the signal-to-noise ratio $\theta_j/\sqrt{\Sigma_{j,j}}$ is uniformly bounded. This lower bound can be interpreted as a lower bound on the sample size necessary to publish a paper.

The next proposition shows that threshold crossing protocols are Σ-robust.

Proposition D.2.1. *Let $J > 1$. Let Assumption 4.2.1, 4.2.2, 4.2.3, and 4.2.4 hold, and let $\Theta = [-1,1]^J$. Then the recommendation function*

$$r_j^*(X) = 1\Big\{X_j/\sqrt{\Sigma_{j,j}} \geq \Phi^{-1}\left(1 - C(J)/J\right)\Big\}, \quad \forall j \in \{1,\ldots,J\}$$

is Σ-robust.

Proof. The proof mimics the proof of Proposition 4.2.5 (see Appendix D.3.2), since the optimization problem only depends on $\theta_j/\sqrt{\Sigma_{j,j}}$, and is omitted. □

Proposition D.2.1 shows that threshold crossing protocols and standard t-tests are also optimal in settings where the social planner seeks recommendation functions that are optimal irrespectively of the particular experimental design.

Alternative Notions of Power

In this section, we discuss two alternative notions of power: (i) a local notion of power, where the local alternatives are such that *every* parameter (instead of at least one) is small and positive; (ii) a global weighted average power (WAP) criterion that allows parameters in the alternative space to be large and positive.

Locally Most Powerful Recommendation Function in Every Direction

Consider the following power criterion.

Definition D.2.2 (Local power in every direction)**.** Let r be locally more powerful than r' in every direction if

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\Big(v_r(\theta^\varepsilon) - v_{r'}(\theta^\varepsilon)\Big) \geq 0, \quad \theta_j^\varepsilon = \varepsilon \quad \forall j \in \{1,\ldots,J\}.$$

In Definition D.2.2, the local alternatives are such that every parameter is equal to ε, instead of at least one as in Definition 4.2.4.

The following proposition provides a characterization of recommendation functions that are locally most powerful in every direction.

Proposition D.2.2 (Locally most powerful in every direction)**.** *Let $J > 1$. Let Assumptions 4.2.1, 4.2.2, 4.2.3, and 4.2.4 hold, and let $\Theta = [-1,1]^J$. Then, the recommendation function $r^*(X)$ is maximin and locally most powerful in every direction if it satisfies Equation* (4.10) *and*

$$\sum_{j=1}^{J} P(r_j^*(X) = 1 | \theta = 0) = C(J).$$

Proof. See Appendix D.3.2. □

Unlike Lemma 4.2.4, Proposition D.2.2 does not restrict the rejection probability of each individual hypothesis, but only imposes a condition on the aggregate rejection probability. As a consequence, Definition D.2.2 is less useful than Definition 4.2.4 for our purpose because it does not provide any guidance on how to specify the rejection probabilities for every test. For example, consider the following two recommendation functions

$$r(X) = \Big\{ r_1(X) = 1\Big\{ X_1/\sqrt{\Sigma_{1,1}} > \Phi^{-1}(1-C(J)) \Big\}, r_{j>1}(X) = 0 \Big\},$$
$$r'(X) = \Big\{ r_J(X) = 1\Big\{ X_J/\sqrt{\Sigma_{J,J}} > \Phi^{-1}(1-C(J)) \Big\}, r_{j<J}(X) = 0 \Big\}.$$

While both recommendation functions are maximin and locally most powerful in every direction, neither is locally most powerful according to Definition 4.2.4 since they violate the condition in Lemma 4.2.4 for every $j \in \{1,\ldots,J\}$.

A direct corollary of Proposition D.2.2 and Lemma 4.2.4 is that any locally most powerful recommendation function is also locally most powerful in every direction, while the converse is not necessarily true.

Corollary 17. *Under the conditions in Proposition D.2.2, any locally most powerful recommendation function is also locally most powerful in every direction.*

In view of Corollary 17, Definition 4.2.4 can be seen as a refinement of Definition D.2.2 that imposes additional symmetry, allowing us to overcome the indeterminacy implied by Proposition D.2.2.

Globally Most Powerful Rules

Consider the following WAP-style global notion of power.

Definition D.2.3 ($w(\theta)$-more powerful (WAP))**.** For a given weighting scheme $w(\theta)$: $\int_\Theta w(\theta)d\theta = 1$, r is $w(\theta)$-more powerful than r' if

$$\int_\Theta w(\theta)\Big(v_r(\theta) - v_{r'}(\theta)\Big)d\theta \geq 0.$$

Definition D.2.3 states that r is more powerful than r' if the difference in welfare, weighted by $w(\theta)$, is weakly positive.

The following proposition shows that no recommendation function is $w(\theta)$-most powerful uniformly in $w(\theta)$, even after restricting weights to be positive only over Θ_1.

Proposition D.2.3. *Let $J > 1, C(J) > 0$. Suppose that Assumptions 4.2.1, 4.2.2, 4.2.3, and 4.2.4 hold and that $\Theta = [-1,1]^J$. For any maximin rule $r \in \mathscr{M}$ there exists a set of weights*

$w(\theta) : \int_{\theta \in \Theta_1} w(\theta) = 1$ *such that*

$$\int_{\Theta} w(\theta)\Big(v_r(\theta) - v_{r'}(\theta)\Big)d\theta < 0.$$

for some r'.

Proof. See Appendix D.3.2. □

Proposition D.2.3 implies that the characterization of $w(\theta)$-most powerful recommendation functions will depend on the particular choice of $w(\theta)$, and ultimately on the properties of the recommendation function over the alternative space. In contrast, our local power criterion (Definition 4.2.4) has the advantage of yielding optimality results that do not depend on specifying weights over a potentially high-dimensional alternative space.

Multiple Outcomes Adversarial Weights

In the main text, we assume that the weights of the policy-makers w_j^* are known. Here we consider the case of unknown adversarial weights. For simplicity, we assume that there is only one policy-maker ($J = 1$) and suppress the indexing with j.

Suppose that the welfare weights w^* are unknown to the social planner (but not the policy-maker) and—in the spirit of the maximin criteria we studied above—chosen adversarially by nature. This assumption reflects information asymmetry between the researcher and policy-maker on the one hand and the planner on the other. As in the main text, we assume that the policy-maker passively implements the recommendation.

The adversarial welfare criterion is

$$v_r(\theta; w^{\text{adv}}), \quad w^{\text{adv}} \in \arg\min_{w \in \Delta} \sum_{g=1}^{G} w_g u_g(\theta), \tag{D.2.1}$$

where Δ denotes the $G-1$-simplex. The worst-case welfare conditional on the researcher

experimenting can be written as

$$\int \min_{w\in\Delta} r(x) \sum_{g=1}^{G} w_g u_g(\theta) dF_X(\theta) = \int r(x) dF_X(\theta) \sum_{g=1}^{G} w_g^{\text{adv}} u_g(\theta)$$
$$= \int r(x) dF_X(\theta) \min\Big\{u_1(\theta), \cdots, u_G(\theta)\Big\}.$$

The researcher's utility only depends on publication prospects and costs, and on the welfare and the weights. The null space takes the following form:

$$\Theta_0 := \big\{\theta : u_g(\theta) < 0, \text{ for some } g \in \{1,\ldots,G\}\big\} \tag{D.2.2}$$

where we use the superscript y to make explicit that this is the null space in the case of multiple outcomes. The definition of the null space is similar to Section 4.3.3.[13]

The next proposition characterizes the set of maximin recommendation functions with adversarial welfare weights.

Proposition D.2.4 (Adversarial weights)**.** *Let Assumption 4.2.3 hold and suppose that $\Theta_0 \neq \emptyset$. A recommendation function r^* is maximin-optimal with adversarial weights i.e.*

$$r^* \in \arg\max_{r\in\mathscr{R}} \min_{\theta\in\Theta} v_r(\theta; w^{\text{adv}})$$

if and only if

$$\beta_{r^*}(\theta) \leq 0 \quad \forall \theta \in \Theta_0, \tag{D.2.3}$$

where Θ_0 is defined in Equation (D.2.2)*.*

Proof. See Appendix D.3.2. □

Note that the structure of the result is similar to that in Proposition 1 in Tetenov (2016);

[13] It is interesting to note that, as a result, the problem in this subsection and the one in Section 4.3.3 lead to similar conclusions.

the key difference is that null set within which the researcher is deterred from experimenting is potentially much *larger*.

The next corollary provides an example of a maximin recommendation function in this context. As in Section 4.2.2 we write $C := C(1)$. Recall that here $r(X)$ is a scalar.[14]

Corollary 18 (Threshold crossing recommendation function with maximin protocols over outcomes)**.** *Suppose that*

$$X_g \sim \mathcal{N}\Big(\theta_g, 1\Big), \quad \theta_1, \ldots, \theta_G \in [-M, M], \quad X_g \perp X_{g' \neq g}.$$

Define $X := \min\{X_1, \ldots, X_G\}$. *Consider a linear publication rule as in Lemma 4.2.4, and a protocol* $r(X;t) = 1\{X \geq t\}$. *Then* $r^*(\cdot) := r(\cdot;t^*)$ *is maximin optimal, where* t^* *satisfies* $P(X_g \geq t^* | \theta_g = 0) = C$. *In addition, for some* M *large enough, any protocol* $r(X;t)$ *with* $P(X_g \geq t | \theta = 0) > C$ *is not maximin.*

Proof. See Appendix D.3.3. □

Corollary 18 characterizes a particular class of maximin protocols. Interestingly, the corollary shows that under independence, by choosing $X = \min\{X_1, \ldots, X_G\}$, a maximin threshold recommendation function imposes a very stringent size control of the form

$$P(r(X) = 1 | \theta = 0) = C^G, \tag{D.2.4}$$

Note that $C < 1$ if the researcher experiments, so that this implies the size of the test shrinks with respect to the number of outcomes at an exponential rate. This adjustment arises due to the adversarial nature of the model with unknown weights and the independence assumption.

Corollary 18 also states that any threshold crossing recommendation function that violates the above size control is not maximin optimal.[15] That said, there may exist an alternative class

[14] This is different from Section 4.3.3 where $r(X)$ is a $G \times 1$ vector.

[15] The idea of the proof is that by taking $X = \min_g X_g$, we violate maximin optimality whenever each parameter

of hypothesis testing protocols that lead to higher power as G increases; we leave this for future study.

Separate Test are not Optimal with a Single Policy Maker

In the main text, we show that with a single policy maker, testing based on a single index is optimal. Here we discuss why separate testing with or without MHT adjustments is not optimal in such settings. Specifically, we show that any most powerful hypothesis testing protocol based on X (such as the tests based on a single index in Section 4.3.4) is weakly more powerful than protocols based on separate testing with or without MHT adjustments.

Separate testing corresponds to a two-step protocol. The researcher reports a vector of tests $r_{\text{sep}}(X) \in \{0,1\}^G$. For example, $r_{\text{sep,g}}(X)$ may correspond to a one-sided tests about the effect of the treatment on the Y_g. Standard procedures for multiple hypothesis testing adjust the size of the tests r_{sep}. The size of each test is chosen by the social planner who optimizes over the (multivariate) function $r_{\text{sep}}(\cdot)$. The researcher chooses whether to experiment and, upon experimentation, reports G recommendations $r_{\text{sep}}(X)$. At this point, the policy-maker must aggregate such tests and make a single policy decision. Formally, for a given recommendation function $r_{\text{sep}}(X)$, the policy-maker chooses whether to implement the policy if $r_{\text{agg}}(r_{\text{sep}}(X)) = 1$, where r_{agg} is an exogenous "aggregator" function. The resulting overall protocol is simply the composition of r_{agg} and r_{sep}, $r_{\text{agg}} \circ r_{\text{sep}}$. Since we do not impose any restrictions on $r \in \mathscr{R}$, any two-step protocol $r_{\text{agg}} \circ r_{\text{sep}}$ can be replicated by an unrestricted protocol $r : \mathscr{X} \mapsto \{0,1\}$, while the converse is not true. Therefore, since $r_{\text{agg}} \circ r_{\text{sep}} \in \mathscr{R}$, it follows that if $r \in \mathscr{R}$ is more powerful than any other $r' \in \mathscr{R}$, r is more powerful than any two-step protocol. Intuitively, two-step protocols are more restrictive since they do not operate directly on X but instead reduce the information in X into binary indicators in the first step. This information reduction can result in a loss of power. Example D.2.1 provides an illustration.

except for one is large and positive, and for one g' only $\theta_{g'}$ is close to zero but negative. In such a case the probability of a discovery, which under independence is obtained by taking the product of each $P(X_g > t|\theta)$ exceeds the overall cost C and thus induces experimentation.

Example D.2.1 (Separate testing with or without MHT adjustments is not optimal)**.** Suppose that $G = 2$ and that $X \sim \mathcal{N}(\theta, I)$. Instead of reporting a single recommendation, conditional on experimentation, the researcher reports a vector of recommendations $r_{\text{sep}}(X;2) = (r_{\text{sep},1}(X;2), r_{\text{sep},2}(X;2)) \in \{0,1\}^2$, where $r_{\text{sep,g}}(X;2) = 1\{X_g > t_g\}$ for $g = 1,2$. The threshold t_g may embed a multiple testing adjustment, can depend on g, and is assumed to be positive and finite. The policy-maker then uses r_{agg} to aggregate $r_{\text{sep}}(X;2)$ into a single policy decision. Specifically, she implements the policy if $r_{\text{agg}}(r_{\text{sep}}(X;2)) = 1$.

This two-step protocol is less powerful than tests based on a single index. To see this, suppose that the policy-maker implements the treatment if there are "enough" rejections,

$$r_{\text{agg}}(r_{\text{sep}}(X;2)) = 1\{1\{X_1 > t_1\} + 1\{X_2 > t_2\} \geq \kappa\}.$$

It can be shown that if $\kappa \leq 1$, then $r_{\text{agg}} \circ r_{\text{sep}}$ is strictly dominated for $\theta^\top w^* < 0$ and thus not maximin. On the other hand, if $\kappa = 2$, then we can always find a configuration of parameters such that $r_{\text{agg}} \circ r_{\text{sep}}$ has approximately zero power. For example, suppose that researchers put equal weights on each outcome, formally $w^* = (1/2, 1/2)^\top$. Choose the first parameter θ_1 to be large and negative $\theta_1 = -M$, for $M > 0$, and the second parameter positive and larger in absolute value than θ_1, namely $\theta_2 = M + u$ for $u > 0$. In this case, the welfare is positive, $\theta^\top w^* = u/2 > 0$, while the two-step procedure has approximately zero power for large enough M, $P(r_{\text{agg}}(r_{\text{sep}}(X;2)) = 1 | \theta = (-M, M+u)) \approx 0$. The reason is that the researcher (almost) never rejects the first hypothesis, since the expectation of the first outcome can be arbitrary small ($\theta = -M$) such that $P(X_1 > t_1 | \theta_1 = -M) \approx 0$. □

Aggregating Statistics vs. Aggregating Outcomes

Here we show that (i) first estimating G separate OLS regressions and then aggregating the estimators and (ii) first aggregating the G outcomes and then running one OLS regression with the aggregate outcome are equivalent.

Suppose that we are interested in the effect of a binary treatment D on a vector of outcomes $Y = (Y_1, \ldots, Y_G)^\top$. Suppose further that, for $g = 1, \ldots, G$,

$$Y_{i,g} = \mu + \theta D_i + \varepsilon_{i,g}, \quad i = 1, \ldots, N, \quad \varepsilon_{i,g} \overset{iid}{\sim} \mathcal{N}(0, \sigma_g^2)$$

We choose X_g to be the OLS estimator of the effect of D on outcome Y_g,

$$X_g = \hat{\theta}_g = \frac{1}{N_1} \sum_{i:D_i=1} Y_{i,g} - \frac{1}{N_0} \sum_{i:D_i=0} Y_{i,g},$$

where N_1 is the number of treated units and N_0 is number of control units. This fits into our framework since $X = (\hat{\theta}_1, \ldots, \hat{\theta}_G)^\top \sim \mathcal{N}(\theta, \Sigma)$.

Now, instead of separately estimating G regressions and aggregating the corresponding OLS estimators, suppose that we first aggregate the outcomes into a single index using weights $w = (w_1, \ldots, w_G)^\top$, $Y^\top w$, and then run an OLS regression of $Y^\top w$ on D. The resulting OLS estimator can be written as

$$\frac{1}{N_1} \sum_{i:D_i=1} \sum_{g=1}^{G} w_g Y_{i,g} - \frac{1}{N_0} \sum_{i:D_i=0} \sum_{g=1}^{G} w_g Y_{i,g} = \sum_{g=1}^{G} w_g \left(\frac{1}{N_1} \sum_{i:D_i=1} Y_{i,g} - \frac{1}{N_0} \sum_{i:D_i=0} Y_{i,g} \right) = \sum_{g=1}^{G} w_g X_g.$$

This derivation shows that both approaches are equivalent.

A Factor Model for Multiple Outcomes

Suppose the researcher analyzes the effect of the treatment on some latent factor F (see, e.g., Ludwig et al., 2017). For example, for a health-related treatment, F may represent an underlying measure of overall health. The effect of the treatment on individual i is $F_i^1 - F_i^0 \overset{iid}{\sim} \mathcal{N}(\theta, 1)$, where F_i^1 and F_i^0 denote the potential factors with and without the treatment. The average treatment effect θ determines welfare upon experimentation.[16] We assume that the

[16] It is important to note that treatment effects are defined up-to re-scaling in this set-up.

outcome g of individual i is generated by a factor model

$$Y_{i,g} = \lambda_g \Big(F_i^1 D_i + F_i^0 (1 - D_i) \Big) + \varepsilon_{i,g}, \quad \big(\varepsilon_{i,1}, \ldots, \varepsilon_{i,G} \big) \mid F_i^1, F_i^0 \overset{iid}{\sim} \mathcal{N}(0, I_G), \tag{D.2.5}$$

where $\lambda_g \in (0, \infty)$ is an outcome-specific factor loading, and, for expositional convenience, we assumed that errors are independent and have the same variance.

If the researcher experiments, she samples N units and assigns a binary treatment $D_i \overset{iid}{\sim}$ Bern(p) to each unit i. Let X_g be the difference between the mean outcomes in the treated and control group,

$$X_g = \frac{1}{N} \sum_{i=1}^{n} Y_{i,g} \left(\frac{D_i}{p} - \frac{1 - D_i}{1 - p} \right), \tag{D.2.6}$$

Standard asymptotic results imply that $X \overset{a}{\sim} \mathcal{N}(\lambda\theta, \Sigma_N)$, where $\lambda = (\lambda_1, \ldots, \lambda_G)^\top$ and Σ_N is a covariance matrix. We are interested in constructing the recommendation function $r(X)$, which is maximin and (approximately) locally most powerful. Suppose that the researcher knows (λ, Σ), or (asymptotically) equivalently, that she has access to consistent estimators.[17] Then

$$r^*(X) = 1 \left\{ \sum_{g=1}^{G} \frac{X_g}{\lambda_g \sqrt{\lambda^{-1\top} \Sigma \lambda^{-1}}} > \Phi^{-1}(1 - C(G)) \right\},$$

is maximin optimal. In addition, for any maximin rule r', r^* achieves asymptotically the largest local power, namely,

$$\lim_{\varepsilon \downarrow 0} \Big[\inf_{\theta \in \tilde{\Theta}} \theta^\top \lambda P(r^*(X) = 1 | \theta) - \inf_{\theta \in \tilde{\Theta}} \theta^\top \lambda P(r'(X) = 1 | \theta) \Big] \geq 0,$$

[17] Asymptotic equivalence follows since

$$\sum_{g=1}^{G} \frac{X_g}{\hat{\lambda}_g \sqrt{\hat{\lambda}^{-1\top} \hat{\Sigma} \hat{\lambda}^{-1}}} = \sum_{g=1}^{G} \frac{X_g}{\lambda_g \sqrt{\lambda^{-1\top} \Sigma \lambda^{-1}}} + o_p(1), \quad \text{where } \sum_{g=1}^{G} \frac{X_g}{\lambda_g \sqrt{\lambda^{-1\top} \Sigma \lambda^{-1}}} \sim \mathcal{N}(\theta, 1).$$

Also, observe that consistent estimation can be achieved only up-to a rescaling constant (which is irrelevant for our testing procedure).

$\tilde{\Theta}(\varepsilon) = \{\theta \in \Theta : \lambda^\top \theta = \varepsilon\}$.[18]

(Some) Benevolent Researchers

One of the core assumptions in our baseline model is that the researcher and planner incentives are misaligned — the researcher's utility only depends on the benefits from publications and the costs of research. Here we show that all our main results continue to hold when some (but not all) researchers' incentives are aligned (i.e., care about social welfare). We focus on the model with multiple treatments (Section 4.2); results for the case of multiple outcomes (Section 4.3) can be derived using similar arguments.

Suppose that there are two types of researchers, indexed by $\omega \in \{0,1\}$. The utility of the researchers of type $\omega = 0$ is equal to welfare, $\delta(r(X))^\top u(\theta)$, whereas the the utility of researchers of type $\omega = 1$ is the same as in Section 4.2. We assume that both types occur with positive probability, $P(\omega = 1) \in (0,1)$.

Consider a setting where the researchers' type is unknown to the social planner, and the planner is maximin with respect to the researcher type:

$$\max_{r \in \mathscr{R}} \min_{\theta \in \Theta, \omega \in \{0,1\}} \widetilde{v}_r(\theta, \omega), \tag{D.2.7}$$

where (for $v_r(\theta)$ as defined in as Equation (4.7))

$$\widetilde{v}_r(\theta, \omega = 1) = v_r(\theta), \quad \widetilde{v}_r(\theta, \omega = 0) = \delta(r(X))^\top u(\theta) 1\left\{\delta(r(X))^\top u(\theta) \geq 0\right\}.$$

The first equality follows by definition of $v_r(\theta)$. The second equality follows from the fact that the benevolent researcher only experiments when welfare is positive.

We have the following equivalence result.

[18] Maximin optimality follows from the fact that $P(r(X) = 1|\theta) \leq C(G)$, for all $\theta \leq 0$. The power property follows by a first order Taylor expansion of $P(r^*(X) = 1|\theta)$ around $\theta = 0$, and continuity of the distribution of X in θ.

Proposition D.2.5. *Suppose that $u(\theta) = 0$, for some $\theta \in \Theta$. Then for any $\beta_r(\theta)$*

$$\min_{\theta\in\Theta,\omega\in\{0,1\}} \widetilde{v}_r(\theta;\omega) = \min_{\theta\in\Theta} v_r(\theta)$$

where $v_r(\theta)$ is given in Equation (4.7).

Proof. See Appendix D.3.2. □

Proposition D.2.5 shows that the social planner solves the same problem as she would have when there are only type $\omega = 1$ researchers (as in Section 4.2). It implies that all the results in Section 4.2 remain continue to hold when some researchers are benevolent.

Proofs

In our proofs we will sometimes write C instead of $C(J)$ and $P(r_j(X)|\theta)$ instead of $P(r_j(X) = 1|\theta)$ to lighten up the notation whenever it does not cause any confusion. For notational convenience, without loss of generality, whenever possible, we standardize the threshold utility with respect to γ, taking $\gamma = 1$. To distinguish maximin from weakly maximin, we refer to maximin as strongly maximin when is not clear from the context.

Lemmas

Here we collect the proofs of all lemmas.

Proof of Lemma 4.2.4

We prove the statement in the following steps. First observe that we can write the researcher's utility as

$$\sum_{j=1}^{J} P(r_j(X)|\theta) - C. \tag{D.3.1}$$

The proof proceeds as follows. We first find a lower bound on the worst-case power of r^*. We then argue that any weakly maximin recommendation function that does not satisfy maximin

optimality has a lower power using an upper bound on any maximin $r \neq r^*$.

Step 1: Maximin optimality.

This directly follows from Proposition 4.2.1.

Step 2: Locally more powerful calculations

We claim that

$$\liminf_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \inf_{\theta \in \Theta_1(\varepsilon)} v_{r^*}(\theta) \geq \frac{C}{J}.$$

We now show why. Define $\theta(\varepsilon) \in \Theta_1(\varepsilon)$ the vector of parameter under the local alternative. Observe that the utility under the local alternative reads as follows

$$(A) = \inf_{\theta(\varepsilon)} \sum_{j=1}^{J} \theta(\varepsilon) P(r_j^*(X) | \theta = \theta(\varepsilon)), \text{ such that } \theta_j(\varepsilon) \geq \varepsilon \text{ for some } j, \theta_j(\varepsilon) \geq 0 \quad \forall j.$$

We then write

$$\begin{aligned}(A) &\geq \inf_{w \in [0,1]^J : \sum_j w_j \geq \varepsilon, \theta(\varepsilon) \in \Theta_1(\varepsilon)} \sum_{j=1}^{J} w_j P(r_j^*(X) | \theta = \theta(\varepsilon)) \\ &\geq \inf_{w \in [0,1]^J : \sum_j w_j \geq \varepsilon, \theta' \in [0,1]^J : \sum_j \theta_j \geq \varepsilon} \sum_j w_j P(r_j^*(X) | \theta = \theta') := g(\varepsilon).\end{aligned}$$

Define $\mathscr{W}(\varepsilon_1, \varepsilon_2) = \left\{ (w, \theta) \in [0,1]^{2J} : \sum_j w_j \geq \varepsilon_1, \sum_j \theta_j \geq \varepsilon_2 \right\}$. Observe that we can write

$$\frac{1}{\varepsilon} g(\varepsilon) = \inf_{(w,\theta') \in \mathscr{W}(1,\varepsilon)} \sum_{j=1}^{J} w_j P(r_j^*(X) | \theta = \theta') = \inf_{(w,\theta') \in \mathscr{W}(1,1)} \sum_{j=1}^{J} w_j P(r_j^*(X) | \theta = \varepsilon \theta').$$

Observe that $\mathscr{W}(1,1)$ is a compact space. In addition $P(r_j^*(X) | \theta = \varepsilon \theta')$ is continuous in ε for any $\theta' \in \Theta$ by Assumption 4.2.4. As a result, $g(\varepsilon)/\varepsilon$ is a continuous function in ε. Therefore, we obtain that

$$\lim_{\varepsilon \to 0} \frac{g(\varepsilon)}{\varepsilon} = \inf_{(w,\theta') \in \mathscr{W}(1,1)} \sum_{j=1}^{J} w_j P(r_j^*(X) | \theta = \theta' \times 0) = \inf_{(w,\theta') \in \mathscr{W}(1,1)} \sum_{j=1}^{J} w_j \frac{C}{J} = \frac{C}{J}.$$

This completes the proof of our claim.

Step 3: Alternative set of maximin protocols.

We now claim that any maximin protocol which is not r^* must satisfy for *some* $j \in \{1,\cdots,J\}$,

$$P\Big(r_j(X)=1|\theta_k=0, \quad \forall k \in \{1,\cdots,J\}\Big) < \frac{C}{J}. \tag{D.3.2}$$

The claim holds for the following reasons. Consider a maximin recommendation function r' such that for all j Equation (D.3.17) does not hold. Then if r' is maximin optimal and satisfies Equation (D.3.2) with equality for all j, then there must be an r^* defined as in the proposition statement equal to r', which leads to a contradiction. Therefore it must be that r' is such that for some j Equation (D.3.2) is satisfies with reversed *strict* inequality and for all j is satisfied with reversed weak inequality. To observe why, by the above argument

$$\sum_j P\Big(r'_j(X)=1|\theta_k=0, \quad \forall k \in \{1,\cdots,J\}\Big) > C.$$

As a result, take $\theta = (-t,-t,\cdots,-t) \in \Theta_0$ for some small t. Then by Assumption 4.2.4 (namely, by continuity of F_θ), we have for t small enough

$$\sum_j P\Big(r'_j(X)=1|\theta_k=-t, \quad \forall k \in \{1,\cdots,J\}\Big) > C.$$

As a result, for t small enough, the policy r' contradicts Proposition 4.2.1.

Step 4: Power comparison.

Observe now that for the class of alternative treatments, we have

$$\inf_{\theta(\varepsilon)} \sum_{j=1}^{J} \theta(\varepsilon) P(r_j(X)|\theta=\theta(\varepsilon)) \leq \varepsilon P(r_j(X)=1|\theta_j=\varepsilon,\theta_{-j}=0),$$

since the allocation $(\theta_j = \varepsilon, \theta_{-j} = 0) \in \Theta_1(\varepsilon)$. Using Assumption 4.2.4 we have

$$\lim_{\varepsilon \to 0} P(r_j(X) = 1 | \theta_j = \varepsilon, \theta_{-j} = 0) = P(r_j(X) = 1 | \theta = 0) < \frac{C}{J}.$$

This completes the proof of the if statement.

Step 5: "Only if" statement

The only if statement follows from the fact that if r^* does not satisfy the condition in the proposition, then we can find a different function r' which leads to larger power than r^* following the same argument after Equation (D.3.2). As a result, in this case r^* violates the condition of local optimality.

Proof of Lemma D.1.1

First, recall that by construction $v_r(\theta) \leq 0$, for all $\theta \in \Theta_0$ and all r. As a result, for all r, $\int v_r(\theta) s\pi(\theta) \leq 0$, for all $\pi \in \Pi$, such that $\int_{\theta \in \Theta_0} \pi(\theta) d\theta = 1$. Since there exists at least one π which puts probability one on Θ_0, it follows that $\inf_{\pi \in \Pi} \widetilde{v}_r(\pi) \leq 0$ for all r. Therefore, r is maximin optimal if $\inf_{\pi \in \Pi} \widetilde{v}_r(\pi) \geq 0$.

Note now that by choosing $r^t(X) = (0, \cdots, 0)$ almost surely, we are guaranteed that $\inf_{\pi \in \Pi} \widetilde{v}_r(\pi) = 0$. Therefore, r^* is maximin optimal only if $\inf_{\pi \in \Delta} \widetilde{v}_r(\pi) \geq 0$, since otherwise dominated by r^t.

Proof of Lemma D.1.3

We first prove the only if statement. To achieve this goal, note that any decision rule r that does not satisfy $v_r(\theta) \geq 0$ is not maximin optimal, since otherwise the social planner can choose r' constant at zero for all S and guarantee that $v_{r'}(\theta) = 0, \quad \forall \theta \in \Theta$. Therefore any rule is maximin only if $v_{r^*}(\theta) \geq 0$ for all $\theta \in \Theta$. To prove the if statement, observe that for any r, $\min_{\theta \in \Theta} v_r(\theta) \leq 0$, since we can choose $\theta = (0, \cdots, 0)$. As a result, any rule is maximin if $\inf_{\theta \in \Theta} v_{r^*}(\theta) = 0$, that, under linearity and the assumption that $\Theta = [-1, 1]^J$, implies that

$v_{r^*}(\theta) \geq 0$ for all $\theta \in \Theta$, completing the proof.

Propositions

Here we collect the proofs of all propositions.

Proof of Proposition 4.2.1

Let $\Theta_0^s = \Theta \setminus \left\{\theta : u_j(\theta) \geq 0 \forall j\right\}$. We start with a general observation. Define $\theta^*(r) = \min_{\theta \in \Theta} v_r(\theta)$ the worst-case θ as a function of the recommendation function r. First, observe that since $\Theta_0 \neq \emptyset$, we have that $v_r(\theta) \leq 0$ for any $r \in \mathscr{R}, \theta \in \Theta_0$. Therefore, it must be that any recommendation function is maximin if and only if $v_{r^*}(\theta^*(r^*)) = 0$, since (i)if $v_{r^*}(\theta^*(r^*)) < 0$, then the social planner can choose $\tilde{r}(X) = (0,0,\ldots,0)$ and obtain $v_{\tilde{r}(X)}(\theta) = 0, \forall \theta \in \Theta$; (ii) if instead $v_{r^*}(\theta^*(r^*)) > 0$, then we reach a contradiction since we can find a $\theta \in \Theta_0$ which leads to non-positive utility. This shows that maximin rules can equivalently be characterized by $v_{r^*}(\theta^*(r^*)) = 0$.

Based on this observation, to prove the "if" direction, we only need to show that the worst-case utility under r^* equals zero, that is $v_{r^*}(\theta^*(r^*)) = 0$. Under Assumption 4.2.3, welfare is exactly zero for all $\theta \in \Theta_0$, as long as Equation (4.11) is satisfied. To complete the first direction of the claim, we are left to show that for any $\theta \in \Theta \setminus \Theta_0$ welfare is always non-negative. This is true since if $\theta \notin \Theta_0^s$ then the utility is trivially positive, since $u(\theta)$ has all entries weakly positive. If instead $\theta \in \Theta_0^s$, then either (i) $\theta \in \Theta_0$ or (ii) $\theta \in \Theta_0^s \setminus \Theta_0$. (i) was discussed before. Therefore consider (ii). Observe that since $v_{r^*}(\tilde{\theta}) \geq 0$ for all $\tilde{\theta} \in \Theta_0^s \setminus \Theta_0$ by assumption, it must be that $v_{r^*}(\theta^*(r^*)) = 0$. As a result if Equation (4.11) holds, r^* is maximin.

We now discuss the "only if" direction. Consider the case where $\beta_r(\theta) > 0$ for some $\theta \in \Theta_0$. Then we have $v_r(\theta^*(r)) \leq v_r(\theta) < 0$ for some $\theta \in \Theta_0$. Suppose instead that $v_{r^*}(\tilde{\theta}) < 0$, for some $\tilde{\theta} \in \Theta \setminus \Theta_0$. Then similarly $v_r(\theta^*(r)) \leq v_r(\tilde{\theta}) < 0$, completing the proof.

Proof of Proposition 4.2.2

We can write the researcher's payoffs *proportional to*

$$\sum_{j=1}^{J} P(r_j(X)|\theta) - C, \tag{D.3.3}$$

where we suppress the dependence of C with J for notational convenience. We take the parameter space[19]

$$\Theta = \left\{ \theta \in [-1,1]^J \text{ such that } \text{sign}(\theta_1) = \text{sign}(\theta_2) = \cdots = \text{sign}(\theta_{J-1}) \right\}$$

and the costs $C/J < 1$. corresponding to the positive and negative quadrants for all parameters except θ_J which can have an arbitrary sign irrespective of the sign of the other parameters. To prove the statement we show that there exists a maximin protocol that strictly dominates all others over an arbitrary set $\Theta' \subseteq \Theta$, and a different maximin recommendation function that it strictly dominates all others over some arbitrary set $\Theta'' \neq \Theta', \quad \Theta'' \subseteq \Theta$. We choose $\Theta' = (0, \cdots, 0, t)$ for a small t and $\Theta'' = (t, \cdots, t, 0)$ for a small t. Details are discussed in the following paragraphs. Since we can choose any distribution F_θ, we choose

$$X \sim \mathcal{N}(\theta, I).$$

Observe that we can define

$$\Theta_0 \subseteq \tilde{\Theta}_0 = \left\{ \theta : \theta_j \leq 0 \text{ for all } j \right\},$$

where $\tilde{\Theta}_0$ also contains those elements that lead to *weakly* negative welfare. We break the proof into several steps.

[19] This condition is not necessary by maximin optimality of threshold crossing protocols (Proposition 4.2.5) for arbitrary parameter spaces, but it simplifies the discussion.

Step 1: construction of the function class.

Define a class of recommendation functions of the form

$$\mathscr{R}^1 = \left\{ r \in \mathscr{R} : C \geq P\Big(r_J^1(X)|\theta = 0\Big) > \frac{C}{J} \text{ and } r \text{ is maximin } \right\}.$$

We claim that $\mathscr{R}^1 \neq \emptyset$ (note that $C/J < 1$ by assumption). To prove this claim it suffices to find a function $r \in \mathscr{R}^1$. An example is

$$\begin{aligned} P\Big(r_j(X)\Big|\forall\theta \in \Theta\Big) &= 0 \quad \forall j \leq J-1, \\ P\Big(r_J(X)|\forall\theta_J = 0, \forall\theta_{j<J} \leq 0\Big) &= \min\{C,1\}, \end{aligned} \tag{D.3.4}$$

and $P(r_J(X)|\theta_J, \theta_{-J})$ is constant in θ_{-J} and decreasing in θ_J, which exist since $\frac{C}{J} < 1$. r is maximin optimal since the researcher's utility is weakly non-negative for any $\theta \in \Theta$.[20] This follows by (i) monotonicity of $P(r_J(X) = 1|\theta_J, \theta_{-J})$ in θ_J and (ii) the fact that $P(r_J(X) = 1|\theta_J, \theta_{-J})$ is constant in θ_{-J}. A simple example satisfying such conditions is a threshold crossing recommendation function of the form

$$r_j(X) = 1\{X_j > t_j\}, t_{j<J} = \infty, t_J = \Phi^{-1}(1 - \min C, 1)$$

where $\Phi(\cdot)$ denotes the normal CDF. In addition, we observe that

$$\sup_{r \in \mathscr{R}^1} P\Big(r_J(X)|\theta_J = 0, \theta_{j<J} = 0\Big) \geq \min\{C,1\}$$

as a result of the above example. Define $\tau = \min\{C,1\}$ for the rest of the proof.

[20] See the proof of Proposition 4.2.1 for an explanation why a weakly non negative utility for all $\theta \in \Theta$ implies maximin optimality.

Step 2: comparisons with maximin protocols.

We now claim that for $\theta = (0,0,\cdots,0,t)$, for t approaching zero, there exists a maximin recommendation function $r^1 \in \mathscr{R}^1$ which leads to strictly larger welfare than any *maximin* decisions $r^2 \in \mathscr{R} \setminus \mathscr{R}^1$. To show our claim, it suffices to compare r^1, to any recommendation function $r^2 \notin \mathscr{R}^1$, such that

$$P\Big(r_J^2(X)|\theta = 0\Big) \leq \frac{C}{J}. \tag{D.3.5}$$

To see why, observe that whenever the above probability is between $(\frac{C}{J}, \tau]$, we contradict the statement that $r^2 \notin \mathscr{R}^1$. When instead

$$P\Big(r_J^2(X)|\theta = 0\Big) > \tau$$

the recommendation function r^2 is *not* maximin optimal, since this implies that $C \leq 1$, which in turn implies that by Assumption 4.2.4, the researcher would conduct experimentation under $(\theta_J = -t, \theta_{j<J} = -t)$, for some small positive t, leading to strictly negative welfare.

Step 3: Comparisons of welfare.

Observe now that for $\theta = (0,0,\cdots,0,t)$, the welfare is

$$\frac{1}{t} v_{r^1}(0,0,\cdots,0,t) = P\Big(r_J^1(X_1,X_2,\cdots,X_J)\Big|\theta_{-J} = 0, \theta_J = t\Big) = f_{r_J^1}(t).$$

Notice that $f_r(t)$ is a continuous function in t due to Assumption 4.2.4. By comparing the utilities under r^1 and r^2, and taking the difference we have

$$\sup_{r^1 \in \mathscr{R}^1} v_{r^1}(\theta_J = t, \theta_{j<J} = 0) - \sup_{r^2 : P(r_J^2(X)|\theta=0) \leq \frac{C}{J}} v_{r^2}(\theta_J = t, \theta_{j<J} = 0) =$$

$$t\Big[\sup_{r^1 \in \mathscr{R}^1} f_{r_J^1}(t) - \sup_{r^2 : P(r_J^2(X)|\theta=0) \leq \frac{C}{J}} f_{r_J^2}(t)\Big] =$$

$$t\Big[\sup_{r^1 \in \mathscr{R}^1} f_{r_J^1}(0) - \sup_{r^2 : P(r_J^2(X)|\theta=0) \leq \frac{C}{J}} f_{r_J^2}(0) + \varepsilon_{r_J^1}(t) - \varepsilon_{r_J^2}(t)\Big]$$

where $\lim_{t\to 0}\varepsilon_{r_J^1}(t) = \lim_{t\to 0}\varepsilon_{r_J^2}(t) = 0$, by Assumption 4.2.4. As a result, we can take some small $t > 0$, such that

$$\varepsilon_{r_J^1}(t) - \varepsilon_{r_J^2}(t) < \sup_{r^1\in\mathscr{R}^1} f_{r_J^1}(t) - \sup_{r^2:P(r_J^2(X)|\theta=0)\leq\frac{C}{J}} f_{r_J^2}(t),$$

since $\sup_{r^1\in\mathscr{R}^1} f_{r_J^1}(t) > \frac{C}{J}$, and therefore $\sup_{r^1\in\mathscr{R}^1} f_{r_J^1}(t) - \sup_{r^2:P(r_J^2(X)|\theta=0)\leq\frac{C}{J}} f_{r_J^2}(t) > 0$. For this case, we have

$$\sup_{r^1\in\mathscr{R}^1} v_{r^1}(\theta_J = t, \theta_{j<J} = 0) - \sup_{r^2:P(r_J^2(X)|\theta=0)\leq\frac{C}{J}} v_{r^2}(\theta_J = t, \theta_{j<J} = 0) > 0.$$

Therefore, for some t small enough, r^1 leads to strictly larger utility than any maximin recommendation function $r^2 \notin \mathscr{R}^1$.

Step 4: the recommendation function r^1 is not dominant

We are left to show that there exists a function $r^3 \notin \mathscr{R}^1$ which is maximin optimal and that leads to strictly larger utility than any $r^1 \in \mathscr{R}^1$ for some different combinations of θ. We choose $\theta = (t, t, \cdots, 0)$. We construct a set of decisions

$$\mathscr{R}^2 = \Big\{ r \notin \mathscr{R}^1 \text{ and } r \text{ is maximin optimal} \Big\}.$$

Observe that $\mathscr{R}^1 \cap \mathscr{R}^2 = \emptyset$ by definition.

Step 5: welfare computation for $\mathscr{R}^2$.

We claim that $\mathscr{R}^2$ is non-empty. An example is the treatment that assigns $P(r_J(X) = 1|\theta = 0) = 0$,

$$P(r_{j<J}(X) = 1|\theta_j = 0, \quad \forall\theta_{-j}) = \min\{\frac{C}{(J-1)}, 1\},$$

which $P(r_j(X) = 1|\theta_j, \theta_{-j})$ decreasing in θ_j and does not depend on θ_{-j}, for $j \neq J$. A threshold crossing recommendation function satisfies this condition. Such a recommendation function is

maximin by the assumption on Θ and the monotonicity of $P(r_j(X)=1|\theta_j)$ in θ_j. Consider the alternative $\breve{\theta}=(t,\cdots,t,0)$. Observe now that we have

$$\sup_{s^2\in\mathscr{R}^2} v_{s^2}(\breve{\theta})-\sup_{r^1\in\mathscr{R}^1} v_{r^1}(\breve{\theta})=t\times\Big[\sum_{j<J}P(s_j^2(X)=1|\theta=\breve{\theta})-\sum_{j<J}P(r_j^1(X)=1|\theta=\breve{\theta})\Big].$$

We write

$$\begin{aligned}&\Big[\sum_{j<J}P(s_j^2(X)=1|\theta=\breve{\theta})-\sum_{j<J}P(r_j^1(X)=1|\theta=\breve{\theta})\Big]\\&=\Big[\sum_{j<J}P(s_j^2(X)=1|\theta=0)-\sum_{j<J}P(r_j^1(X)=1|\theta=0)\Big]+\varepsilon(t)\end{aligned}$$

where $\varepsilon(t)\to 0$ as $t\to 0$ by continuity (Assumption 4.2.4).

Step 6: Upper bound on r^1.

We claim that

$$\sum_{j<J}P(r_j^1(X)=1|\theta=0)<(J-1)(\min\{C/(J-1),1\}).$$

We prove the claim by contradiction. Suppose that the above equation does not hold. Then it must be that (since $P(r_J^1(X)=1|\theta=0)>C/J$)

$$\sum_{j<J}P(r_j^1(X)=1|\theta=0)+P(r_J^1(X)=1|\theta=0)>(J-1)(\min\{C/(J-1),1\})+\frac{C}{J}. \quad \text{(D.3.6)}$$

Clearly if $C/(J-1)\leq 1$, the statement is true since otherwise we would contradict maximin optimality of r^1. Suppose that $C/(J-1)>1$. Then for r^1 to be maximin optimal we must have that $(J-1)+C/J\leq C$. However, it is easy to show that this implies that $C/J\geq 1$ which leads to a contradiction. This completes the claim.

Using continuity, we obtain that for t small enough any $r^1\in\mathscr{R}^1$ is dominated by r^2. The proof is complete.

Proof of Proposition 4.2.3

We structure the proof as follows: we first prove the first part of the claim. We show that any decision which is strongly maximin and locally more powerful than any other strongly maximin protocol is admissible. Second, we show that such decision is also admissible once compared to recommendation functions which are not maximin. Finally, we prove the second part of the claim (the only if statement).

Since admissibility is stated with respect to the rules $r' \notin \mathscr{E}$, by construction, for any $r \in \mathscr{E}$ and $r' \notin \mathscr{E}$, the inequality in Equation (4.13) holds with strict inequality if r' is maximin.

Step 1: Locally most powerful implies admissibility among strongly maximin protocols.

We first compare to maximin rules $r' \notin \mathscr{E}$. Observe first that we can take $u_j(\theta) = \varepsilon, u(\theta)_{i \neq j} = 0$ which belongs to $\Theta_1(\varepsilon)$ since by assumption $[-1,1] \in \{u_j(\theta), \theta \in \Theta\}$. Observe that by the definition of locally more powerful

$$0 \leq \inf_{\theta \in \Theta_1(\varepsilon)} v_r(\theta) = \inf_{\theta \in \Theta_1(\varepsilon)} \mathbb{E}[\delta(r(X))^\top u(\theta)]\varepsilon \leq \mathbb{E}[\delta(r(X))^\top (\varepsilon, 0, \cdots, 0)] \leq \varepsilon \qquad \text{(D.3.7)}$$

for some j, where the right-hand side follows by definition of $\Theta_1(\varepsilon)$. As a result we have that $0 \leq \inf_{\theta \in \Theta_1(\varepsilon)} v_r(\theta)/\varepsilon \leq 1$ and so the lim-inf and lim-sup are uniformly bounded. We now have that

$$(I) = \liminf_{\varepsilon \downarrow 0} \left\{ \frac{1}{\varepsilon} \inf_{\theta \in \Theta_1(\varepsilon)} v_r(\theta) - \frac{1}{\varepsilon} \inf_{\theta' \in \Theta_1(\varepsilon)} v_{r'}(\theta') \right\}$$

is finite since by Equation (D.3.7) the above expression for any $\varepsilon > 0$ is bounded from below by zero and from above by one. Now observe that by definition of lim-inf, there exists a *subsequence* $\varepsilon_n \downarrow 0$, such that $\frac{1}{\varepsilon_n} \inf_{\theta \in \Theta_1(\varepsilon_n)} v_r(\theta) - \frac{1}{\varepsilon_n} \inf_{\theta' \in \Theta_1(\varepsilon_n)} v_{r'}(\theta')$ converges to (I). Take some finite n over the subsequence but large enough such that

$$\frac{1}{\varepsilon_n} \inf_{\theta \in \Theta_1(\varepsilon_n)} v_r(\theta) - \frac{1}{\varepsilon_n} \inf_{\theta' \in \Theta_1(\varepsilon_n)} v_{r'}(\theta') > 0.$$

Such an ε_n exists since r' is maximin and $r' \notin \mathscr{E}$. Define $\theta_{\varepsilon_n} = \inf_{\theta' \in \Theta_1(\varepsilon_n)} v_{r'}(\theta')$. Observe that

$$\frac{1}{\varepsilon_n} v_r(\theta_{\varepsilon_n}) - \frac{1}{\varepsilon_n} v_{r'}(\theta_{\varepsilon_n}) \geq \frac{1}{\varepsilon_n} \inf_{\theta \in \Theta_1(\varepsilon_n)} v_r(\theta) - \frac{1}{\varepsilon_n} \inf_{\theta' \in \Theta_1(\varepsilon_n)} v_{r'}(\theta') > 0, \tag{D.3.8}$$

since[21]

$$\inf_{\theta \in \Theta_1(\varepsilon_n)} v_r(\theta) \leq v_r(\theta_{\varepsilon_n}).$$

Now observe that Equation (D.3.8) implies that r strictly dominates r' at some θ_{ε_n}. We now want to show that r is admissible within the class of maximin protocols. Since r is locally more powerful than all maximin protocols, we can apply the same reasoning to any other r'' for some (different) θ'_{ε_n}. Clearly, r is not dominated by any other maximin protocol recommendation function by the argument above.

Step 2: Comparison to non-maximin recommendation function.

Next we show that the strongly maximin locally most powerful r is also admissible within the larger class of strongly maximin and non-strongly maximin protocols. To show this we use a contradiction argument. Suppose that there exists a recommendation function $\tilde{r}$ which is *not* maximin. Then $\tilde{r}$ is dominated by r either over the parameter $\theta \in \Theta_0$ or over $\Theta \setminus \Theta_0$ since r is maximin. Therefore r is admissible with respect to all $r' \notin \mathscr{E}$.

Proof of Proposition 4.2.5

Observe that

$$P(r_j(X) = 1 | \theta = 0) = \frac{C}{J}, \quad \forall j \in \{1, \cdots, J\}.$$

[21] The inequality below follows from the fact that the constraint set $\Theta_1(\varepsilon)$ does not depend on r.

As a result, we only need to prove maximin optimality. Observe first that $r_j(X)$ is monotonically increasing in θ_j and constant in θ_{-j}. In addition, the null is defined as

$$\Theta_0 \subseteq \Big\{\theta : \theta_j \leq 0, \quad \forall j\Big\}.$$

By monotonicity, the first condition in Proposition 4.2.1 is satisfied. We show that also the second condition holds.

To show this it suffices to show that the worst-case objective function is weakly positive. With an abuse of notation we define θ_j the coefficient rescaled by $\sqrt{\Sigma_{j,j}}$ (which are finite by the assumption of Σ being positive-definite). We write the nature's adversarial game with a threshold crossing recommendation function as follows

$$\min_{\theta\in\Theta} \sum_{j=1}^{J} (1-\Phi(t-\theta_j))\theta_j, \qquad \text{s.t.} \sum_{j=1}^{J} (1-\Phi(t-\theta_j)) \geq J(1-\Phi(t)). \tag{D.3.9}$$

Observe that since $C(J) > 0$, t must be finite. We can write Θ as a compact space $[-M,M]^J$ for some finite M, by assumption:

$$\min_{\theta\in[-M,M]^J} \sum_{j=1}^{J} (1-\Phi(t-\theta_j))\theta_j, \qquad \text{s.t.} \sum_{j=1}^{J} (1-\Phi(t-\theta_j)) \geq J(1-\Phi(t)). \tag{D.3.10}$$

for some arbitrary large M whose choice is discussed below. Observe that Equation (D.3.10) is weakly smaller than Equation (D.3.9) hence a lower bound on Equation (D.3.10) suffices to prove the claim.

Claim.

We claim that welfare can only be negative if the minimizer θ^* is such that for some j, $\theta_j^* < 0$ and for some other $j' \neq j$ $\theta_{j'}^* > 0$. That there must exist some negative θ_j^* trivially follows from the objective function. That there must be a positive $\theta_{j'}^*$ follows directly from the constraint function: if such condition is not met and $\theta_j^* < 0$ for all j, then it follows that the

constraint is violated.

Focus on interior solution for M large enough.

We now argue that at least one component of the minimizer must satisfy $-\infty < \theta_j^* < 0$ for the resulting objective function to be strictly negative. To show this it suffices to observe that $z(1-\Phi(t-z)) \to 0$ as $z \to -\infty$. Therefore if for all $\theta_j^* < 0$ these are unbounded, then the objective function is trivially zero proving the claim. Second, following this same argument we also observe that it suffices to focus our analysis on solutions $\theta^* \in \Theta \subseteq [-M,M]^J$ for some arbitrary large but finite M. To see why observe that if *at least one* θ_j^* is unbounded its contribution to the objective function is zero, while it decreases the researcher's utility $\sum_{j=1}^J (1-\Phi(t-\theta_j^*))$, hence having a weakly positive effect on the objective function. For the $\theta_j^* > 0$ these must instead be finite since otherwise the objective function is strictly positive (hence, the claim trivially holds). Hence there must exist a minimizer θ^* which is in the interior of $[-M,M]^J$ for some arbitrary large and finite M.

Constraint qualification.

We now show that the KKT conditions are necessary for the optimality of $\theta^* \in \Theta \subset [-M,M]^J$. To show this we use the LICQ. In particular observe that the derivative of the constraint function is

$$-\sum_{j=1}^{J} \phi(t-\theta_j) \neq 0$$

for t being finite, for any point θ such that at least one θ_j is finite (in absolute value). If such condition is not met, then the objective function is (weakly) non negative as discussed in the previous paragraph.

Lagrangian.

We now study *necessary* conditions for the optimal solution of the problem in Equation (D.3.9). Consider the Lagrangian function

$$\sum_{j=1}^{J}(1-\Phi(t-\theta_j))\theta_j+\lambda\left[J(1-\Phi(t))-\sum_{j=1}^{J}(1-\Phi(t-\theta_j))\right]+\mu_{1,j}[\theta-M]+\mu_{2,j}[-\theta-M].$$

Now observe that by the argument in the second paragraph, and complementary slackness we can focus on the cases where $\mu_{1,j}=\mu_{2,j}=0$ for all j (i.e. θ^* is an interior of $[-M,M]^J$ for some finite M large enough). Taking first order conditions of the Lagrangian we obtain

$$\phi(t-\theta_j)\theta_j+(1-\Phi(t-\theta_j))=\lambda\phi(t-\theta_j)\Rightarrow\lambda^*=\frac{1}{\phi(t-\theta_j)}\left[\phi(t-\theta_j)\theta_j+(1-\Phi(t-\theta_j))\right].$$

Contradiction argument.

We conclude this proof using a contradiction argument using the claim we established at the beginning of the proof. Suppose that the objective function is strictly negative. The there must exists a j such that $\theta_j^*<0$ and $j'\neq j$ such that $\theta_{j'}^*>0$. In addition, observe that using the equation for the optimal λ, we can write

$$0>\theta_j^*=\lambda^*-\frac{(1-\Phi(t-\theta_j^*))}{\phi(t-\theta_j^*)}\quad 0<\theta_{j'}^*=\lambda^*-\frac{(1-\Phi(t-\theta_{j'}^*))}{\phi(t-\theta_{j'}^*)}.$$

Using the fact that t is finite, it follows that

$$\frac{1-\Phi(t-\theta_{j'}^*)}{\phi(t-\theta_{j'}^*)}<\lambda^*<\frac{1-\Phi(t-\theta_j^*)}{\phi(t-\theta_j^*)}.$$

Observe now that the expression implies

$$\frac{1-\Phi(z)}{\phi(z)}<\frac{1-\Phi(z')}{\phi(z')}\text{ for some }z<z'.$$

However, by standard properties of the normal CDF $(1-\Phi(z))/\phi(z)$ is monotonically *decreasing* in z hence leading to a contradiction.

Proof of Proposition 4.3.1

The proof mimics the proof of Proposition 4.2.1. Observe that since $\Theta_0 \neq \emptyset$, there exist a $\theta \in \Theta_0$ such that $v_r(\theta) \leq 0$ for any $r \in \mathscr{R}$. Then the maximin welfare equals zero, since otherwise the social planner can set $r(X) = 0$ and achieve zero welfare. Observe that the welfare is negative for all $\theta \in \Theta_0$. Under Assumption 4.2.3, welfare is exactly zero for all $\theta \in \Theta_0$, as long as Equation (D.2.3) is satisfied. In addition, welfare is always non-negative for $\theta \in \Theta \setminus \Theta_0$ and any r, for any $\theta \setminus \Theta_0$ each $u_g(\theta)$ is weakly positive. As a result if Equation (D.2.3) holds, r^* is maximin. Consider now the case where $\beta_r(\theta) > 0$ for some $\theta \in \Theta_0$. Then, we can find a $\theta \in \Theta_0$ which leads to negative utility, for any decision $r(\cdot)$. Assumption 4.2.3 is invoked for the equality in Equation (D.2.3) to be a weak inequality.

Proof of Proposition 4.3.2

First, we prove that the power function is continuous in ε. We then show that the local power of r^* is $C(G)$. Finally, we show that no maximin function r' can lead to a power larger than $C(G)$ in the limit as $\varepsilon \downarrow 0$. First, for any r, we can write

$$\begin{aligned}\frac{1}{\varepsilon}\inf_{\theta\in\tilde{\Theta}(\varepsilon)}\sum_{g=1}^{G} w_g^*\theta_g P(r(X)=1|\theta) &= \inf_{\theta\in\Theta\cap\{\theta:\sum_{g=1}^{G} w_g^*\theta_g=\varepsilon\}} P(r(X)=1|\theta) \\ &= \inf_{\tilde{\theta}\in\Theta\cap\{\tilde{\theta}:\sum_{g=1}^{G} w_g^*\tilde{\theta}_g=1\}} P(r(X)=1|\theta=\varepsilon\tilde{\theta}) = g_r(\varepsilon).\end{aligned}$$

Since the set $\Theta\cap\{\tilde{\theta}:\sum_{g=1}^{G} w_g^*\tilde{\theta}_g = 1\}$ is a compact set, and, by normality of X, $P(r(X)=1|\theta = \varepsilon\tilde{\theta})$ is continuous in ε for any $\tilde{\theta}$, we obtain that $g_r(\varepsilon)$ is a continuous function in ε for any r. This implies that

$$\lim_{\varepsilon\downarrow 0} g_{r^*}(\varepsilon) = C(G).$$

In addition, for any r' that is maximin, it must be that

$$\lim_{\varepsilon \downarrow 0} g_r(\varepsilon) \leq C(G), \quad r' \in \mathscr{M}$$

since otherwise we can take $\varepsilon < 0$ and small enough such $g_r(\varepsilon) > C(G)$, which contradicts maximin optimality.

Proof of Proposition 4.3.3

We now study local power, and constrast this to other maximim optimal decisions.[22] For any r, letting $\theta = \varepsilon$ local power equals

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \varepsilon P(r(X) = 1 | \theta = \varepsilon) = P(r(X) = 1 | \theta = 0).$$

By continuity of the density of X in θ, it suffices to show that r^* is such that for any maximin rule r',

$$P(r^*(X) = 1 | \theta = 0) - P(r'(X) = 1 | \theta = 0) \geq 0$$

to show that r^* is locally most powerful. This directly follows since for any maximin rule $P(r'(X) = 1 | \theta = 0) \leq C(G)$, since otherwise, by continuity we would violate maximin for $\theta = -\varepsilon$, for some small ε.

Proof of Proposition D.1.2

We first show maximin optimality. To show maximin optimality, it suffices to show that the worst-case objective function is weakly positive. With an abuse of notation, we define θ_j the coefficient rescaled by $\sqrt{\Sigma_{j,j}}$.

[22] Maximin optimality follows directly from the fac that at $\theta \leq 0$, the probability of discovery is smaller than the cost of experimentation.

Preliminaries for maximin optimality.

To show maximin optimality it suffices to show that the worst case utility, upon experimentation, is weakly positive. Therefore, we write the nature's adversarial game with a threshold crossing recommendation function as follows

$$\min_{\pi\in\Pi}\int\sum_{j=1}^{J}(1-\Phi(t-\theta_j))\theta_j d\pi(\theta),\quad \text{s.t.}\int\sum_{j=1}^{J}(1-\Phi(t-\theta_j))d\pi(\theta)\geq J(1-\Phi(t)). \tag{D.3.11}$$

Observe that since $C(J)>0$, t must be finite. It will be convenient to consider the equivalent optimization program

$$\max_{\pi\in\Pi}-\int\sum_{j=1}^{J}(1-\Phi(t-\theta_j))\theta_j d\pi(\theta),\quad \text{s.t.}\int\sum_{j=1}^{J}(1-\Phi(t-\theta_j))d\pi(\theta)\geq J(1-\Phi(t)). \tag{D.3.12}$$

Here, we inverted the sign of the objective and consequently inverted the maximum with the minimum. We will show that Equation (D.3.12) is bounded from *above* by zero, which suffices to show maximin optimality.

Finite dimensional optimization program.

The maximization over $\pi\in\Pi$ can be equivalently rewritten as a minimization over $\pi_1(\cdot),\cdots,\pi_J(\cdot),\theta_j\sim\pi_j$, i.e., with respect to marginal distributions. This follows directly by additivity. By Theorem 1, result 3 in Gaivoronski (1986), we can write the optimization problem in Equation (D.3.11) as an optimization over some finitely many $n\times J$ discrete points $\left(\theta_1^i,\cdots,\theta_J^i\right)_{i=1}^n$, each point θ_j^i having marginal probability $p_{i,j}$.[23] Hence, we write

$$\text{(D.3.11)}=\max_{\pi\in\Pi_n}\sum_{i=1}^{n}\int\sum_{j=1}^{J}(\Phi(t-\theta_j^i)-1)\theta_j p_{i,j},\quad \text{s.t.}\sum_{i=1}^{n}\sum_{j=1}^{J}(\Phi(t)-\Phi(t-\theta_j^i))p_{i,j}\geq 0, \tag{D.3.13}$$

[23]The conditions in the above reference are satisfied since $\Phi(t-\theta_j)$ is a continuous function in θ_j, Θ is a compact space, and we can find a distribution so that the constraint holds with strict inequality.

where

$$\Pi_n = \left\{ (p_{i,j})_{i=1,j=1}^{n,J}, p_{i,j} \geq 0, \sum_{i=1}^{n} p_{i,j} = 1 \text{ for all } j \in \{1, \cdots, J\} \right\},$$

for some $\left(\theta_1^i, \cdots, \theta_J^i\right)_{i=1}^n$.

Dual formulation.

The optimization in Equation (D.3.13) is a linear program with linear constraints. Therefore, the dual is directly defined as follows.

$$\begin{aligned} &\min_{y_1,\cdots,y_J,y_{J+1}} \sum_{j=1}^{J} y_j, \quad y_1, \cdots, y_J \in \mathbb{R}, \quad y_{J+1} \leq 0 \\ &\text{such that } \sum_{j=1}^{J} y_j \geq -y_{J+1}(\Phi(t) - \Phi(t-\theta_j^i)) - (1-\Phi(t-\theta_j^i))\theta_j^i \quad \forall (j,i). \end{aligned} \tag{D.3.14}$$

By weak duality, we have that

$$\text{(D.3.13)} \leq \text{(D.3.14)}.$$

Therefore, to prove that Equation (D.3.12) is bounded from above by zero, it is sufficient to prove that its dual is bounded from above by zero.

Upper bound on the dual's objective function.

To compute the upper bound on the dual's objective, it suffices to observe that the dual's objective is minimized over values $(y_1, \cdots, y_J)$, which can be arbitrary in $\mathbb{R}$ (and $y_{J+1} \leq 0$), but whose sum is constrained by the constraint in Equation (D.3.14). As a result, we have

$$\begin{aligned} \text{(D.3.14)} &\leq \min_{y_{J+1} \leq 0} \max_{\theta_j^i} -y_{J+1}(\Phi(t) - \Phi(t-\theta_j^i)) - (1-\Phi(t-\theta_j^i))\theta_j^i \\ &= \min_{y \geq 0} \max_{\theta_j^i} y(\Phi(t) - \Phi(t-\theta_j^i)) - (1-\Phi(t-\theta_j^i))\theta_j^i \\ &\leq \min_{y \geq 0} \max_{\check{\theta} \in [-1/\sqrt{\min_j \Sigma_{j,j}}, 1/\sqrt{\min_j \Sigma_{j,j}}]} y(\Phi(t) - \Phi(t-\check{\theta})) - (1-\Phi(t-\check{\theta}))\check{\theta}. \end{aligned}$$

The first equality follows directly by construction of the optimization program and the fact that we take the maximum over every possible value of θ_j^i. The second equality is a change of variable where we wrote $-y_{J+1} = y$, and the third inequality substitutes the maximum over the set of (unknown) $n \times J$ parameters θ_j^i over a minimization over a parameter $\breve{\theta}$ taking arbitrary values in the parameter space.

Claim of maximin optimality.

To show maximin optimality it sufficies to show that the function

$$\inf_{y\geq 0} \sup_{\breve{\theta}\in[-1/\sqrt{\min_j \Sigma_{j,j}},1/\sqrt{\min_j \Sigma_{j,j}}]} f(t,y,\breve{\theta}), \quad f(t,y,\breve{\theta}) := y(\Phi(t)-\Phi(t-\breve{\theta}))-(1-\Phi(t-\breve{\theta}))\breve{\theta}$$

is bounded from above from zero for all t. Positivity can be shown numerically. We provide an analytical argument below.

Check the function value.

Define

$$\breve{\theta}(y) \in \arg \max_{\theta\in[-1/\sqrt{\min_j \Sigma_{j,j}},1/\sqrt{\min_j \Sigma_{j,j}}]} y(\Phi(t)-\Phi(t-\breve{\theta}))-(1-\Phi(t-\breve{\theta}))\breve{\theta},$$

$$y^* \in \arg\min_{y\geq 0} y(\Phi(t)-\Phi(t-\breve{\theta}(y)))-(1-\Phi(t-\breve{\theta}(y)))\breve{\theta}(y).$$

Suppose first that $\breve{\theta}(\tilde{y}) = 0$ for some $\tilde{y} \geq 0$. Then it follows that

$$\min_{y\geq 0} f(t,y,\breve{\theta}(y)) \leq f(t,\tilde{y},\breve{\theta}(\tilde{y})) = 0.$$

We are left to discuss the case where $\breve{\theta}(y) \neq 0$. Notice that $\frac{(1-\Phi(t-\breve{\theta}(y))\breve{\theta}(y)}{\Phi(t)-\Phi(t-\breve{\theta}(y))} \geq 0$, for all $\breve{\theta}(y) \neq 0$ since if $\breve{\theta}(y) < 0$, the denominator and numerator are both negative and viceversa are both positive if $\breve{\theta}(y) > 0$. Therefore, (assuming $\breve{\theta}(y) \neq 0$, for all $y \geq 0$) we can always find a value $\tilde{y} \geq 0$ (since we minimize over $\tilde{y} \geq 0$), such that $\tilde{y} \leq \frac{(1-\Phi(t-\breve{\theta}(\tilde{y}))\breve{\theta}(\tilde{y})}{\Phi(t)-\Phi(t-\breve{\theta}(\tilde{y}))}$. In such a case, we obtain a

(weakly) negative valued objective function. As a result,

$$\min_{y\geq 0} y(\Phi(t)-\Phi(t-\check{\theta}(y)))-(1-\Phi(t-\check{\theta}(y)))\check{\theta}(y)\leq 0.$$

Admissibility.

Admissibility follows directly from most-powerful claimed in Proposition 4.2.5, and Proposition 4.2.3, where, in this case, admissibility is with respect to a point-mass distribution over some $\theta\in\Theta_1(\varepsilon)$, for some small enough $\varepsilon>0$.

Proof of Proposition D.1.4

We separately prove that the test is maximin and locally most powerful.

Maximin optimality

We start by proving maximin optimality. Following Lemma D.1.3 it suffices to show that $v_{r^*}(\theta)\geq 0$ for all $\theta\in\Theta$. Note that s_j^* depends on θ, so that we can explicitly write $s^*(\theta)$. To prove maximin optimality, it suffices to prove that

$$\min_{\theta\in\tilde{\Theta}}\sum_{j=1}^{J}(1-\Phi(t_{S^*(\theta)}-\theta_j))\theta_j s_j^*(\theta)\quad \text{s.t.}\sum_{j=1}^{J}(1-\Phi(t_{S^*(\theta)}-\theta_j))s_j^*(\theta)\geq S^*(\theta)(1-\Phi(t_{S^*(\theta)})).$$

is positive. This holds if we can prove that

$$\min_{\theta\in\tilde{\Theta}}\min_{s,S=\sum_j s_j}\sum_{j=1}^{J}(1-\Phi(t_S-\theta_j))\theta_j s_j\quad \text{s.t.}\sum_{j=1}^{J}(1-\Phi(t_S-\theta_j))s_j\geq S(1-\Phi(t_S)),$$

where in the second equation we allow s to be arbitrary. Switching the minimum over θ and over s, it suffices to show that

$$\min_{\theta\in\tilde{\Theta}}\sum_{j=1}^{J}(1-\Phi(t_S-\theta_j))\theta_j s_j\quad \text{s.t.}\sum_{j=1}^{J}(1-\Phi(t_S-\theta_j))s_j\geq S(1-\Phi(t_S)),$$

is positive for *all* $s \in \{0,1\}^J, S = \sum_{j=1}^J s_j$, where $\tilde{\Theta} = [-1/\min_{j,j}\sqrt{\Sigma_{j,j}}, 1/\min_{j,j}\sqrt{\Sigma_{j,j}}]^J$ after rescaling the coefficients by their corresponding variance. The argument is similar to that in the proof of Proposition 4.2.5. In particular, suppose first $s_j = 1, s_{-j} = 0$, i.e., only one treatment is selected. Then

$$(1 - \Phi(t_1 - \theta_j)) \geq (1 - \Phi(t_1)),$$

only if $\theta_j \geq 0$, guaranteeing that the objective function is positive. Suppose now that $S \geq 2$. Then since s_j is a binary indicator, without loss of generality, we can reshuffle coefficients' indexes so that $(s_1, \cdots, s_S) = (1, \cdots, 1)$, and write the objective function as

$$\min_{\theta \in \tilde{\Theta}_S} \sum_{j=1}^{S} (1 - \Phi(t_S - \theta_j))\theta_j, \quad \text{s.t.} \sum_{j=1}^{S} (1 - \Phi(t_S - \theta_j)) \geq S(1 - \Phi(t_S)),$$

where $\tilde{\Theta}_S = [-1/\min_{j,j}\sqrt{\Sigma_{j,j}}, 1/\min_{j,j}\sqrt{\Sigma_{j,j}}]^S$. The argument of Proposition 4.2.5 in Appendix D.3.2 follows verbatim and guarantees that the objective function is weakly positive. This completes the claim of maximin optimality.

We now discuss local power. The idea is to show that, in the worst case, any maximin rule cannot yield more local power than $C(1)$ (as defined in Equation (D.3.15)) for $\theta \in \Theta_1^E(\varepsilon)$.

Local power: Part 1

First, we claim that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \inf_{\theta \in \Theta_1(\varepsilon)} v_{r^*}^E(\theta) = C(1). \tag{D.3.15}$$

To prove the claim, observe first that for a single $\theta_j > 0$,

$$1 - \Phi(t_S - \theta_j) > (1 - \Phi(t_S)).$$

Hence, the researcher finds it profitable to select policy j. In such a case, if the researcher selects exactly one treatment j, she achieves a positive utility and conducts the experiment. In addition, since $\theta_{-j} = 0$, it will be profitable for the researcher to only select a single treatment if $C(S)$ is

strictly monotonic in S. If instead $C(S) = C(1)$ for all S, any combination of selected treatments, whenever $\theta_j = \varepsilon$ for a single j will leave welfare invariant and equal to $C(1)/S \leq C(1)$. By Assumption 4.2.3, the researcher, who is indifferent of whether select those treatments with $\theta_{-j} = 0$, will only select a single treatments corresponding to $\theta_j = \varepsilon$, maximizing the welfare (making it equal to $C(1)$). This proves the claim that the welfare is $C(1)$ whenever $\theta_j = \varepsilon$ for a single j. The case where $\theta_j > \varepsilon$ is not of interest since, once we look at the worst case scenario over $\theta_j \geq \varepsilon$, we will select $\theta_j = \varepsilon$.

Our goal is to show that for any other combination of $\theta_j \geq \varepsilon, \theta_{-j} = 0$ (e.g., more than one parameter is positive) welfare weakly improves. If this is the case, the claim is proven, since the worst case utility for $\theta \in \Theta_1^E(\varepsilon)$ must be $C(1)$. Clearly, if $S^* = 1$, we are back to the case discussed above regardless of the value of the other parameters. Here, the selected parameter must be the one equal to ε since it leads to the largest experimenter's utility.

Suppose now that there are $S^* \geq 2$ policies selected by the researcher. Then it must be that (by Assumption 4.2.3)

$$\sum_{j=1}^{S^*}(1-\Phi(t_{S^*}-\theta_j)) - C(S^*) > \sum_{j=1}^{S}(1-\Phi(t_S-\theta_j)) - C(S) \quad \forall S \leq S^*.$$

Since t_{S^*} and $C(S^*)$ are increasing in S^*, a *necessary* condition for this to hold for $\theta \in \Theta_1^E(\varepsilon)$ is that $\theta_j \geq \varepsilon$ for all $j \leq S^*$. This is true, since, otherwise, $\theta_j = 0$ by construction of $\Theta_1^E(\varepsilon)$, and, in such a case, the researcher's utility generated by selecting policy j is zero, but the threshold t_S increases in the number of selected policies, hence decreasing the researcher's utility if one more policy j with no effect is selected. Therefore it will be profitable to drop those tests with $\theta_j = 0$ to improve researcher's utility.

Therefore, for this case, the welfare divided by ε is

$$\frac{1}{\varepsilon}\sum_{j=1}^{S^*}\theta_j(1-\Phi(t_{S^*}-\theta_j)), \quad \theta_j \geq \varepsilon \quad \Rightarrow \quad \frac{1}{\varepsilon}\sum_{j=1}^{S^*}\theta_j(1-\Phi(t_{S^*}-\theta_j)) \geq C(S^*) \geq C(1),$$

since $\Phi(t_{S^*} - \theta_j)$ is monotonically descreasing in θ_j. This completes the claim and proves that for $\theta \in \Theta_1^E(\varepsilon)$, welfare is at least $C(1)$ and exactly $C(1)$ as we choose a single parameter to be $\geq \varepsilon$.

Local Power: Part 2

We show that in the worst case scenario, any maximin rule cannot improve local power more than $C(1)$ as in Equation (D.3.15) under the conditions stated. Welfare is (without loss of generality, let $r_{S,j}(X) = 0$ almost surely if $j > S$)

$$\begin{aligned}
&\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \inf_{\theta \in \tilde{\Theta}(\varepsilon)} \frac{1}{\varepsilon} \sum_{j=1}^{J} s_j^*(\theta) P(r_{S^*(\theta),j}(X_s) = 1 | \theta = \varepsilon) \theta_j \\
&\leq \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \inf_{\theta \in \tilde{\Theta}(\varepsilon)} \max_{s \in \{0,1\}^J} \sum_j s_j \theta_j P(r_{1^\top s,j}(X_s) = 1 | \theta) \\
&\leq \lim_{\varepsilon \downarrow 0} \max_s P(r_{1^\top s,1}(X_s) = 1 | \theta_1 = \varepsilon, \theta_{-1} = 0) \\
&= \max_s P(r_{1^\top s,1}(X_s) = 1 | \theta = 0)
\end{aligned}$$

where in the second inequality we chose $\theta_1 = \varepsilon, \theta_{-1} = 0$, and in the last inequality we used continuity of X. If $C(1) = C(S)$, it follows that $\max_s P(r_{1^\top s,1}(X_s) = 1 | \theta = 0) \leq C(1)$, since otherwise we can find a combination s and $\theta_1 = -t, \theta_{-1} = 0$, for a small enough t, which leads to positive researcher utility, but negative welfare, thus violating maximin optimality. If instead the protocol is symmetric, it must be that $P(r_{S,1}(X_s) = 1 | \theta = 0) \leq C(S)/S \leq C(1)$, for all s, S since, otherwise, we would violate maximin optimality.

Proof of Proposition D.1.5

For the sake of brevity we refer to maximin protocols as weakly maximin. The proof of the theorem follows similarly to Lemma 4.2.4. First, observe that the researcher's utility is linear in $\delta_k(r(X))$ for every $k \in \mathscr{K}$, since $\sum_{k \in \mathscr{K}} \delta_k(r(X)) \leq 1$ and $\delta_k(r(X)) \in \{0,1\}$. Namely,

the researcher's utility reads as follows

$$\sum_{k\in\mathscr{K}} P(\delta_k(r(X))=1|\theta)-C(J),$$

where $\mathscr{K}$ denotes the set of discoveries exceeding κ recommendations, where we rescaled $\gamma=1$. We first prove the first part of the statement.

Step 1: Maximin optimality.

Maximin optimality directly follows from Equation (D.1.3).

Step 2: Locally most powerful calculations.

We *claim* that

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\inf_{\theta\in\Theta_1(\varepsilon)} v_{r^*}(\theta)\geq p^*.$$

The claim holds following the same argument as in the proof of Lemma 4.2.4.

Step 3: Alternative set of maximin protocols

We now claim that any maximin protocol which is not r^* must satisfy for *some* $j\in\{1,\cdots,J\}$,

$$P\Big(r'_j(X)=1|\theta_k=0,\quad \forall k\Big)<p^*. \tag{D.3.16}$$

The claim follows directly from the incentive-compatibility constraint, following the same argument as in Lemma 4.2.4, with $P(\delta_j(r(X))|\theta)=p^{*k}$ for any group j having k many treatments (all treatments must be selected). There are $\binom{J}{k}$ many groups having k treatments. As a result, we can find a worst-case allocation for $r'_j(X)$ which leads to a utility bounded from above by εp^*. The rest of the proof follows similarly to Lemma 4.2.4 and Proposition D.1.6 below.

Proof of Proposition D.1.6

For the sake of brevity we refer to maximin protocols as weakly maximin. Observe that since the researcher never submits discoveries with less than κ treatments (Assumption D.1.4),

we only focus on discoveries having a positive effect on the researcher's utility. Therefore we refer to $\sum_k$ as $\sum_{k\in\mathscr{K}}$. Finally, observe that the researcher's utility is linear in $\delta_k(r(X))$ for every $k \in \mathscr{K}$, since $\sum_{k\in\mathscr{K}} \delta_k(r(X)) \leq 1$ and $\delta_k(r(X)) \in \{0,1\}$. As a result, we can prove the statement in the following steps following the proof of Lemma 4.2.4 with minor modifications. We first prove the first part of the statement.

Step 1: Maximin optimality.

This directly follows from Equation (D.1.3).

Step 2: Locally most powerful calculations.

We *claim* that

$$\liminf_{\varepsilon\downarrow 0} \frac{1}{\varepsilon} \inf_{\theta\in\Theta_1(\varepsilon)} v_{r^*}(\theta) \geq \frac{C}{|\mathscr{K}|}.$$

We now show why. Define $\theta(\varepsilon) \in \Theta_1(\varepsilon)$ the vector of parameter under the local alternative. Observe that the utility under the local alternative reads as follows

$$(A) = \inf_{\theta(\varepsilon)} \sum_{k\in\mathscr{K}} \theta(\varepsilon) P(\delta_k(r^*(X)) = 1 | \theta = \theta(\varepsilon)), \text{ s.t. } \theta_k(\varepsilon) \geq \varepsilon \text{ for some } k \in \mathscr{K}, \theta_k(\varepsilon) \geq 0, \forall k \in \mathscr{K}.$$

We then write

$$\begin{aligned}(A) &\geq \inf_{w\in[0,1]^{|\mathscr{K}|}:\sum_k w_k \geq \varepsilon, \theta(\varepsilon)\in\Theta_1(\varepsilon)} \sum_{k\in\mathscr{K}} w_k P(\delta_k(r^*(X)) = 1 | \theta = \theta(\varepsilon)) \\ &\geq \inf_{w\in[0,1]^{|\mathscr{K}|}:\sum_k w_k \geq \varepsilon, \theta\in[0,1]^{|\mathscr{K}|}:\sum_k \theta_k \geq \varepsilon} \sum_k w_k P(\delta_k(r^*(X)) | \theta) := g(\varepsilon).\end{aligned}$$

Define $\mathscr{W}(\varepsilon_1,\varepsilon_2) = \Big\{(w,\theta) \in [0,1]^{2\times|\mathscr{K}|} : \sum_j w_k \geq \varepsilon_1, \sum_k \theta_k \geq \varepsilon_2\Big\}$. Observe that we can write

$$\frac{1}{\varepsilon} g(\varepsilon) = \inf_{(w,\theta')\in\mathscr{W}(1,1)} \sum_{k\in\mathscr{K}} w_k P\Big(\delta_k(r^*(X)) | \theta = \theta'\varepsilon\Big).$$

Observe that $\mathscr{W}(1,1)$ is a compact space. In addition $P(\delta_k(r^*(X)) = 1|\theta)$ is continuous in θ by Assumption 4.2.4. As a result, $g(\varepsilon)/\varepsilon$ is a continuous function in ε. Therefore, we

obtain that

$$\lim_{\varepsilon \to 0} \frac{g(\varepsilon)}{\varepsilon} = \inf_{(w,\theta) \in \mathscr{W}(1,1)} \sum_{k \in \mathscr{K}} w_k P(\delta_k(r^*(X)) = 1 | \theta \times 0).$$

Observe now that by construction of the recommendation function, the above expression equals to

$$\inf_{(w,\theta) \in \mathscr{W}(1,1)} \sum_{k \in \mathscr{K}} w_k P(\delta_k(r^*(X)) = 1 | \theta \times 0) = \inf_{(w,\theta) \in \mathscr{W}(1,0)} \sum_{k \in \mathscr{K}} w_k \frac{C}{|\mathscr{K}|} = \frac{C}{|\mathscr{K}|}.$$

This completes the proof of our claim.

Step 3: Alternative set of maximin protocols.

We now claim that any maximin protocol which is not r^* must satisfy for *some* $j \in \mathscr{K}$,

$$P\Big(\delta_j(r(X)) = 1 | \theta_k = 0, \quad \forall k\Big) < \frac{C}{|\mathscr{K}|}. \tag{D.3.17}$$

The claim holds for the following reasons. Consider a maximin protocol r' such that for all k Equation (D.3.17) does not hold. Then if r' is maximin optimal and satisfies Equation (D.3.17) with equality for all k, then there must be an r^* defined as in the proposition statement equal to r. Therefore it must be that r' is such that for some k Equation (D.3.17) is satisfied with reversed *strict* inequality and for all k is satisfied with reversed weak inequality. To observe why, by the above argument

$$\sum_j P\Big(\delta_j(r'(X)) = 1 | \theta_k = 0, \quad \forall k\Big) > C.$$

As a result, take $\theta = (-t, -t, \cdots, -t) \in \Theta_0$ for some small t. Then by Assumption 4.2.4 (namely, by continuity of F_θ), we have for t small enough

$$\sum_j P\Big(\delta_j(r'(X)) = 1 | \theta_k = -t, \quad \forall k\Big) > C.$$

As a result, for t small enough, the protocol r' contradicts Proposition 4.2.1.

Step 4: Power comparison.

Observe now that for the class of alternative treatments, we have

$$\inf_{\theta(\varepsilon)} \sum_{k \in \mathscr{K}} \theta(\varepsilon) P(\delta_k(r(X)) | \theta = \theta(\varepsilon)) \leq \varepsilon P(\delta_k(r(X)) = 1 | \theta_k = \varepsilon, \theta_{-k} = 0),$$

since the allocation $(\theta_k = \varepsilon, \theta_{-k} = 0) \in \Theta_1(\varepsilon)$. Using Assumption 4.2.4 we have

$$\lim_{\varepsilon \to 0} P(\delta_k(r(X)) = 1 | \theta_k = \varepsilon, \theta_{-k} = 0) = P(\delta_k(r(X)) = 1 | \theta = 0) < \frac{C}{|\mathscr{K}|}.$$

This completes the "if" part of the statement. The "only if" part follows from the same argument used after Equation (D.3.17).

Proof of Proposition D.2.2

To prove the claim, it suffices to observe that under additivity

$$\frac{1}{\varepsilon} \sum_{j=1}^{J} \varepsilon P(r_j(X) = 1 | \theta = \varepsilon) = \sum_{j=1}^{J} P(r_j(X) = 1 | \theta = \varepsilon)$$

for any $\varepsilon \neq 0$. By taking the limit, and using continuity of the CDF of X in θ, we have

$$\lim_{\varepsilon \downarrow 0} \sum_{j=1}^{J} P(r_j(X) = 1 | \theta = \varepsilon) = \sum_{j=1}^{J} P(r_j(X) = 1 | \theta = 0).$$

By maximin optimality, we must have

$$\sum_{j=1}^{J} P(r_j(X) = 1 | \theta = 0) \leq C(J),$$

completing the proof.

Proof of Proposition D.2.3

Under Proposition 4.2.1 and continuity of X (Assumption 4.2.4) every maximin rule must satisfy

$$\sum_{j=1}^{J} P(r_j(X)=1|\theta=0) \leq C(J).$$

We can write the weighted power of r as

$$\int_{\theta\in\Theta_1} w(\theta) \sum_{j=1}^{J} P(r_j(X)=1|\theta)\theta_j d\theta.$$

We now discuss two cases. First, $P(r_J(X)=1|\theta=0)=C(J)$; second $P(r_J(X)=1|\theta=0)<C(J)$.

Case 1: $P(r_J(X)=1|\theta=0)=C(J)$

Suppose first that $P(r_J(X)=1|\theta=0)=C(J)$, which implies that $P(r_1(X)=1|\theta=0)=0$. Then choose $w(\theta)$ to be a point-mass distribution at $(\theta_1,\cdots,\theta_J)=(\varepsilon,0,\ldots,0)$ for some small $\varepsilon>0$. Take r' such that

$$r'_1(X)=1\Big\{X_1/\sqrt{\Sigma_{1,1}} \geq \Phi^{-1}(1-C(J))\Big\}, \quad r'_j(X)=0, \quad \forall j>1.$$

It is easy to show that $r'(X)$ is maximin under Assumption 4.2.2 and 4.2.4. Then it follows that

$$\int_{\theta\in\Theta} w(\theta)\Big(v_r(\theta)-v_{r'}(\theta)\Big)d\theta = \varepsilon\Big(P(r_1(X)=1|\theta=(\varepsilon,0,\ldots,0))-P(r'_1(X)=1|\theta=(\varepsilon,0,\ldots,0))\Big).$$

By continuity, it follows that as $\varepsilon\downarrow 0$

$$P(r_1(X)=1|\theta=(\varepsilon,0,\ldots,0))\to 0, \quad P(r'_1(X)=1|\theta=(\varepsilon,0,\ldots,0))\to C(J)>0.$$

Hence, by continuity, we can take $\varepsilon>0$ small enough such that

$$\int_{\theta\in\Theta} w(\theta)\Big(v_r(\theta)-v_{r'}(\theta)\Big)<0.$$

Case 2: $P(r_J(X)=1|\theta=0)<C(J)$

Suppose now that $P(r_J(X)=1|\theta=0)<C(J)$. Then we can take

$$r'_J(X)=1\Big\{X_1/\sqrt{\Sigma_{J,J}}\geq\Phi^{-1}(1-C(J))\Big\},\quad r'_j(X)=0,\quad\forall j<J,$$

and $w(\theta)$ being a point mass at $\theta=(0,\ldots,0,\varepsilon)$. The same argument follows verbatim as for the previous case with the first entry replacing the last entry.

Proof of Proposition D.2.4

The proof mimics the proof of Proposition 4.2.1, and uses the assumption of unconstrained $\mathscr{R}$. Observe that since $\Theta_0\neq\emptyset$, there exist a $\theta\in\Theta_0$ such that $v_r(\theta)\leq 0$ for any $r\in\mathscr{R}$. Then the maximin welfare equals zero, since otherwise the social planner can set $r(X)=0$ and achieve zero welfare. Observe that welfare is negative for all $\theta\in\Theta_0$. Under Assumption 4.2.3, welfare is exactly zero for all $\theta\in\Theta_0$, as long as Equation (D.2.3) is satisfied. In addition, the welfare is always non-negative for $\theta\in\Theta\setminus\Theta_0$ and any r, for any $\theta\setminus\Theta_0$ each $u_g(\theta)$ is weakly positive. As a result if Equation (4.11) holds, r^* is maximin. Consider now the case where $\beta_r(\theta)>0$ for some $\theta\in\Theta_0$. Then, we can find a $\theta\in\Theta_0$ which leads to negative utility, for any decision $r(\cdot)$. Assumption 4.2.3 is invoked for the equality in Equation (4.11) to be a weak inequality.

Proof of Proposition D.2.5

Let $\tilde{\theta}\in\Theta$ be such that $u(\tilde{\theta})=0$. Such a $\tilde{\theta}$ exists by assumption. We can then write

$$\min_{\theta\in\Theta}\widetilde{v}_r(\theta,\omega=1)=\min_{\theta}v_r(\theta)\leq v_r\tilde{\theta})\leq 0.$$

On the other hand, we can write

$$\min_{\theta\in\Theta}\widetilde{v}_r(\theta,\omega=0)\geq 0\Rightarrow\min_{\theta\in\Theta}\widetilde{v}_r(\theta,\omega=1)\leq\min_{\theta\in\Theta}\widetilde{v}_r(\theta,\omega=1).$$

It follows that

$$\min_{\theta\in\Theta,\omega}\widetilde{v}_r(\theta,\omega)=\min_{\omega}\min_{\theta\in\Theta}\widetilde{v}_r(\theta\in\Theta,\omega)=\min_{\theta\in\Theta}\widetilde{v}_r(\theta\in\Theta,\omega=1)=v_r(\theta).$$

The proof is complete.

Corollaries

Here we collect the proofs of the corollaries that are not immediate.

Proof of Corollary 15

We first prove the first statement (notice: we let $\gamma=1$ for notational convenience).

Case with $C(J)/\gamma=\alpha$.

We write

$$\sum_{k\in\{\kappa,\cdots,J\}}\binom{J}{k}(p^*)^k\leq\sum_{k\in\{0,\cdots,J\}}\binom{J}{k}(p^*)^k$$

By the Binomial theorem, we have

$$(1+p^*)^J=\sum_{k\in\{0,\cdots,J\}}\binom{J}{k}(p^*)^k.$$

We then observe that

$$(1+p^*)^J\leq\exp(p^*J).$$

Observe that for $p^*\leq 1/J$ the expression is $O(1)$. Therefore, any order $1/J$ or slower, guarantees that the publication probability is bounded from below by a finite constant. Faster order are instead not possible, since if p^* was of order faster than $1/J$ this would imply that the publication probability converges to zero. However, this would lead to a contradiction since the publication probability must equal $\alpha>0$. To rule out orders of convergence slower than $1/J$, we use a contradiction argument. The argument works as follows: any order of convergence slower than

$1/J$ for p^* implies that the publication probability converges to infinity. This would lead to a contradiction since the publication probability must equal $\alpha < 1$. First, take

$$\sum_{k\in\{\kappa,\cdots,J\}} \binom{J}{k}(p^*)^k \geq \binom{J}{\kappa}(p^*)^\kappa. \tag{D.3.18}$$

Suppose now that p^* is of order slower than $1/J$, e.g. $p^* \asymp h_J/(J)$, for some arbitrary $h_J \to \infty$ as $J \to \infty$. Then we have

$$\binom{J}{\kappa}(p^*)^\kappa \asymp (J-\kappa)^\kappa \frac{h_J^\kappa}{J^\kappa} \to \infty,$$

since $\kappa < \infty$, leading to a contradiction.

Case with $C(J)/\gamma = \alpha/J$.

We first start from the lower bound. We observe that we can write

$$J \times \sum_{k\in\{\kappa,\cdots,J\}} \binom{J}{k}(p^*)^k \geq J \times \binom{J}{\kappa}(p^*)^\kappa \geq J \times \frac{(J-\kappa)^\kappa}{\kappa^\kappa}\frac{1}{J^{\kappa+1}} = \frac{(J-\kappa)^\kappa}{\kappa^\kappa}\frac{1}{J^\kappa} = \frac{1}{\kappa^\kappa}(1-\kappa/J)^\kappa > 0$$

hence bounded away from zero. We now move to the upper bound. Consider first the case where $\kappa = 1$. Using the binomial theorem, we can write

$$\sum_{k\in\{\kappa,\cdots,J\}} \binom{J}{k}(p^*)^k = (1+p^*)^J - \sum_{k\in\{0,\cdots,\kappa-1\}} \binom{J}{k}(p^*)^k = (1+p^*)^J - 1 \leq \exp(1/J) - 1.$$

where the second equality follows from the fact that $\kappa = 1$ and the last equality by the fact that for $\kappa = 1, p^* = 1/J^2$. Using the mean value theorem, we have

$$\exp(1/J) - 1 = \frac{1}{J} + O(\frac{1}{J^2}) = O(1/J)$$

completing the claim for $\kappa = 1$. Let $\kappa > 1$. We have

$$J \times \sum_{k\in\{\kappa,\cdots,J\}} \binom{J}{k}(p^*)^k \leq J \times \sum_{k\in\{\kappa,\cdots,J\}} \frac{J^k}{k(k-1)}\frac{1}{J^{(1+1/\kappa)k}} = \sum_{k\in\{\kappa,\cdots,J\}} \frac{1}{k(k-1)}\frac{J}{J^{k/\kappa}}.$$

The inequality follows from the fact that $\binom{J}{k} \leq \frac{J^k}{k(k-1)}$ for $k > 1$ by definition of the binomial coefficient. Observe that since the sum starts from $\kappa > 1$ we have that for each summand $\frac{J}{J^{k/\kappa}} \leq 1$. Hence, we can write

$$\sum_{k \in \{\kappa, \cdots, J\}} \frac{1}{k(k-1)} \frac{J}{J^{k/\kappa}} \leq \sum_{k \in \{\kappa, \cdots, J\}} \frac{1}{k(k-1)} = O(1)$$

completing the proof since $\kappa > 1$.

Proof of Corollary 16

Observe that since

$$\sum_k \widetilde{\delta}_k(r(X)) \leq 1, \quad \widetilde{\delta}_k : \{0,1\}^J \mapsto \{0,1\}^{2^J - 1},$$

we have that

$$P\Big(\widetilde{\delta}_k(r^*(X)) = 1 \text{ for some } k | \theta = 0\Big) = \sum_k P\Big(\widetilde{\delta}_k(r^*(X)) = 1 | \theta = 0\Big),$$

since the events are disjoint. The result directly follows from Equation (D.1.10).

Proof of Corollary 18

First, observe that we can write:

$$P\Big(\min\{X_1, \cdots, X_G\} > t | \theta\Big) = \prod_{g=1}^{G} P(X_g > t | \theta). \tag{D.3.19}$$

Observe now that maximin optimality of r^* follows from the fact that the recommendation function is decreasing in each θ_j and that for any $\theta \in \Theta_0$ the probability of discovery cannot exceed the cost. To observe consider the extreme case where $\theta_g \to \infty$ for each $g \in \{1, \cdots, G-1\}$ and $\theta_G = 0$. Then $P(X_g > t | \theta = \infty) = 1$ while $P(X_G > t | \theta = 0) = C$. Hence $\prod_{g=1}^{G} P(X_g >$

$t|\theta) = C$ satisfying maximin optimality. The same argument applies if we shuffle the indexes. Finally, we discuss the second part of the claim. To show lack of maximin optimality we choose $\theta_{1:(G-1)}$ large enough so that $P(X_g > t|\theta) = 1$ for all $g < G$. As a result, any allocation with $P(X_G > t|\theta) > C$ is not maximin optimal. The same result applies to any other coordinate $g \neq G$.

Bibliography

Adusumilli, K., F. Geiecke, and C. Schilter (2019). Dynamically optimal treatment allocation using reinforcement learning. *arXiv preprint arXiv:1904.01047*.

Agarwal, A., O. Dekel, and L. Xiao (2010). Optimal algorithms for online convex optimization with multi-point bandit feedback. In *COLT*, pp. 28–40. Citeseer.

Akbarpour, M., S. Malladi, and A. Saberi (2018). Just a few seeds more: value of network information for diffusion. *Available at SSRN 3062830*.

Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2016). Network structure and the aggregation of information: Theory and evidence from indonesia. *American Economic Review 106*(7), 1663–1704.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, and J. Tobias (2012). Targeting the poor: evidence from a field experiment in indonesia. *American Economic Review 102*(4), 1206–40.

Alidaee, H., E. Auerbach, and M. P. Leung (2020). Recovering network structure from aggregated relational data using penalized regression. *arXiv preprint arXiv:2001.06052*.

Ananth, A. (2020). Optimal treatment assignment rules on networked populations. Technical report, working paper.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Andrews, I. and M. Kasy (2019). Identification of and correction for publication bias. *American Economic Review 109*(8), 2766–94.

Andrews, I., T. Kitagawa, and A. McCloskey (2019). Inference on winners. Technical report, National Bureau of Economic Research.

Andrews, I. and J. M. Shapiro (2021). A model of scientific communication. *Econometrica 89*(5), 2117–2142.

Armstrong, T. and S. Shen (2015). Inference on optimal treatment assignments.

Aronow, P. M. and C. Samii (2017). Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics 11*(4), 1912–1947.

Athey, S., D. Eckles, and G. W. Imbens (2018). Exact p-values for network interference. *Journal of the American Statistical Association 113*(521), 230–240.

Athey, S. and G. W. Imbens (2018). Design-based analysis in difference-in-differences settings with staggered adoption. Technical report, National Bureau of Economic Research.

Athey, S. and S. Wager (2021). Policy learning with observational data. *Econometrica 89*(1), 133–161.

Auerbach, E. (2019). Identification and estimation of a partially linear regression model using network data. *arXiv preprint arXiv:1903.09679*.

Bai, Y. (2019). Optimality of matched-pair designs in randomized controlled trials. *Available at SSRN 3483834*.

Baird, S., J. A. Bohren, C. McIntosh, and B. Özler (2018). Optimal design of experiments in the presence of interference. *Review of Economics and Statistics 100*(5), 844–860.

Bakirov, N. K. and G. J. Szekely (2006). Student's t-test for gaussian scale mixtures. *Journal of Mathematical Sciences 139*(3), 6497–6505.

Balashankar, A., A. Lees, C. Welty, and L. Subramanian (2019). What is fair? exploring pareto-efficiency for fairness constrained classifiers. *arXiv preprint arXiv:1910.14120*.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144), 1236498.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2014). Gossip: Identifying central individuals in a social network. Technical report, National Bureau of Economic Research.

Banerjee, A. V., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2017). Using gossips to spread information: theory and evidence from two randomized controlled trials.

Banerjee, A. V., S. Chassang, S. Montero, and E. Snowberg (2020). A theory of experimenters: Robustness, randomization, and balance. *American Economic Review 110*(4), 1206–30.

Barrera-Osorio, F., M. Bertrand, L. L. Linden, and F. Perez-Calle (2011). Improving the design of

conditional transfer programs: Evidence from a randomized education experiment in colombia. *American Economic Journal: Applied Economics 3*(2), 167–95.

Basse, G. and A. Feller (2018). Analyzing two-stage experiments in the presence of interference. *Journal of the American Statistical Association 113*(521), 41–55.

Basse, G. W. and E. M. Airoldi (2018a). Limitations of design-based causal inference and a/b testing under arbitrary and network interference. *Sociological Methodology 48*(1), 136–151.

Basse, G. W. and E. M. Airoldi (2018b). Model-assisted design of experiments in the presence of network-correlated outcomes. *Biometrika 105*(4), 849–858.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological) 57*(1), 289–300.

Benjamini, Y. and W. Liu (1999). A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference 82*(1-2), 163–170.

Bhattacharya, D. (2009). Inferring optimal peer assignment from experimental data. *Journal of the American Statistical Association 104*(486), 486–500.

Bhattacharya, D. and P. Dupas (2012). Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics 167*(1), 168–196.

Bhattacharya, D., P. Dupas, and S. Kanaya (2013). Estimating the impact of means-tested subsidies under treatment externalities with application to anti-malarial bednets. Technical report, National Bureau of Economic Research.

Bhattacharya, D., P. Dupas, and S. Kanaya (2019). Demand and welfare analysis in discrete choice models with social interactions. Technical report, National Bureau of Economic Research.

Bloch, F., M. O. Jackson, and P. Tebaldi (2017). Centrality measures in networks. *Available at SSRN 2749124*.

Bloch, F., M. O. Jackson, and P. Tebaldi (2019). Centrality measures in networks. *Available at SSRN 2749124*.

Bond, R. M., C. J. Fariss, J. J. Jones, A. D. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler (2012). A 61-million-person experiment in social influence and political mobilization. *Nature 489*(7415), 295.

Bottou, L., F. E. Curtis, and J. Nocedal (2018). Optimization methods for large-scale machine learning. *Siam Review 60*(2), 223–311.

Boucheron, S., O. Bousquet, and G. Lugosi (2005). Theory of classification: A survey of some recent advances. *ESAIM: probability and statistics 9*, 323–375.

Boucheron, S., G. Lugosi, and P. Massart (2003). Concentration inequalities using the entropy method. *The Annals of Probability 31*(3), 1583–1614.

Boucheron, S., G. Lugosi, and P. Massart (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Breza, E., A. G. Chandrasekhar, T. H. McCormick, and M. Pan (2020). Using aggregated relational data to feasibly identify network structure without network data. *American Economic Review 110*(8), 2454–84.

Brodeur, A., M. Lé, M. Sangnier, and Y. Zylberberg (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics 8*(1), 1–32.

Brooks, R. L. (1941). On colouring the nodes of a network. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 37, pp. 194–197. Cambridge University Press.

Bubeck, S. (2014). Convex optimization: Algorithms and complexity. *arXiv preprint arXiv:1405.4980*.

Bubeck, S. and N. Cesa-Bianchi (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning 5*(1), 1–122.

Bubeck, S., R. Munos, and G. Stoltz (2011). Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science 412*(19), 1832–1852.

Cai, J., A. De Janvry, and E. Sadoulet (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics 7*(2), 81–108.

Canay, I. A., J. P. Romano, and A. M. Shaikh (2017). Randomization tests under an approximate symmetry assumption. *Econometrica 85*(3), 1013–1030.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2010). Evaluating marginal policy changes and the average effect of treatment for individuals at the margin. *Econometrica 78*(1), 377–394.

Charnes, A. and W. W. Cooper (1962). Programming with linear fractional functionals. *Naval Research logistics quarterly 9*(3-4), 181–186.

Chassang, S., G. Padro I Miquel, and E. Snowberg (2012). Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review 102*(4), 1279–1309.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters.

Chernozhukov, V., D. Chetverikov, and K. Kato (2014). Gaussian approximation of suprema of empirical processes. *The Annals of Statistics 42*(4), 1564–1597.

Chernozhukov, V., K. Wuthrich, and Y. Zhu (2018). Practical and robust t-test based inference for synthetic control and related methods. *arXiv preprint arXiv:1812.10820*.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ. 1*(1), 451–488.

Chin, A., D. Eckles, and J. Ugander (2018). Evaluating stochastic seeding strategies in networks. *arXiv preprint arXiv:1809.09561*.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data 5*(2), 153–163.

Christakis, N. A. and J. H. Fowler (2010). Social network sensors for early detection of contagious outbreaks. *PloS one 5*(9).

Corbett-Davies, S. and S. Goel (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv preprint arXiv:1808.00023*.

Coston, A., A. Mishler, E. H. Kennedy, and A. Chouldechova (2020). Counterfactual risk assessments, evaluation, and fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 582–593.

Cowgill, B. and C. E. Tucker (2019). Economics, fairness and algorithmic bias. *preparation for: Journal of Economic Perspectives*.

Devroye, L., L. Györfi, and G. Lugosi (2013). *A probabilistic theory of pattern recognition*, Volume 31. Springer Science & Business Media.

Di Tillio, A., M. Ottaviani, and P. N. Sørensen (2017). Persuasion bias in science: Can economics help? *The Economic Journal 127*(605), F266–F304.

DiTraglia, F. J., C. García-Jimeno, R. O'Keeffe-O'Donovan, and A. Sanchez-Becerra (2020). Identifying causal effects in experiments with social interactions and non-compliance. *arXiv e-prints*, arXiv–2011.

Donini, M., L. Oneto, S. Ben-David, J. S. Shawe-Taylor, and M. Pontil (2018). Empirical risk minimization under fairness constraints. In *Advances in Neural Information Processing Systems*, pp. 2791–2801.

Duchi, J., F. Ruan, and C. Yun (2018). Minimax bounds on stochastic batched convex optimization. In *Conference On Learning Theory*, pp. 3065–3162. PMLR.

Dudoit, S., J. P. Shaffer, and J. C. Boldrick (2003). Multiple hypothesis testing in microarray experiments. *Statistical Science 18*(1), 71–103.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review 101*(5), 1739–74.

Dwork, C., M. Hardt, T. Pitassi, O. Reingold, and R. Zemel (2012). Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pp. 214–226.

Eckles, D., H. Esfandiari, E. Mossel, and M. A. Rahimian (2019). Seeding with costly network information. *Available at SSRN 3386417*.

Eckles, D., B. Karrer, and J. Ugander (2017). Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference 5*(1).

Efron, B. (2008a). Microarrays, empirical bayes and the two-groups model. *Statistical Science 23*(1), 1–22.

Efron, B. (2008b). Simultaneous inference: when should hypothesis testing problems be combined? *Annals of Applied Statistics 2*(1), 197–223.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. W. Walker (2019). General equilibrium effects of cash transfers: experimental evidence from kenya. Technical report, National Bureau of Economic Research.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. W. Walker (2020, December). General equilibrium effects of cash transfers: Experimental evidence from kenya. Working Paper 26600, National Bureau of Economic Research.

Elliott, G. and R. P. Lieli (2013). Predicting binary outcomes. *Journal of Econometrics 174*(1), 15–26.

Elliott, G., U. K. Müller, and M. W. Watson (2015). Nearly optimal tests when a nuisance parameter is present under the null hypothesis. *Econometrica 83*(2), 771–811.

Farrell, M. H. (2015). Robust inference on average treatment effects with possibly more

covariates than observations. *Journal of Econometrics 189*(1), 1–23.

Feldman, A. and A. Kirman (1974). Fairness and envy. *The American Economic Review*, 995–1005.

Fink, G., M. McConnell, and S. Vollmer (2014). Testing for heterogeneous treatment effects in experimental data: false discovery risks and correction procedures. *Journal of Development Effectiveness 6*(1), 44–57.

Finner, H. and M. Roters (2001). On the false discovery rate and expected type i errors. *Biometrical Journal 43*(8), 985–1005.

Flaxman, A. D., A. T. Kalai, and H. B. McMahan (2004). Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*.

Florios, K. and S. Skouras (2008). Exact computation of max weighted score estimators. *Journal of Econometrics 146*(1), 86–91.

Frankel, A. and M. Kasy (2018). Which findings should be published? Working Paper.

Gaivoronski, A. (1986). Linearization methods for optimization of functionals which depend on probability measures. In *Stochastic Programming 84 Part II*, pp. 157–181. Springer.

Galeotti, A., B. Golub, and S. Goyal (2017). Targeting interventions in networks. *arXiv preprint arXiv:1710.06026*.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

Graham, B. S., G. W. Imbens, and G. Ridder (2010). Measuring the effects of segregation in the presence of social spillovers: A nonparametric approach. Technical report, National Bureau of Economic Research.

Gu, J. and R. Koenker (2020). Invidious comparisons: Ranking and selection as compound decisions. arXiv:2012.12550.

Hadad, V., D. A. Hirshberg, R. Zhan, S. Wager, and S. Athey (2019). Confidence intervals for policy evaluation in adaptive experiments. *arXiv preprint arXiv:1911.02768*.

Hajłasz, P. and J. Malỳ (2002). Approximation in sobolev spaces of nonlinear expressions involving the gradient. *Arkiv för Matematik 40*(2), 245–274.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hardt, M., E. Price, and N. Srebro (2016). Equality of opportunity in supervised learning. In *Advances in neural information processing systems*, pp. 3315–3323.

Hazan, E., K. Levy, and S. Shalev-Shwartz (2015). Beyond convexity: Stochastic quasi-convex optimization. In *Advances in Neural Information Processing Systems*, pp. 1594–1602.

He, X. and K. Song (2018). Measuring diffusion over a large network. *arXiv preprint arXiv:1812.04195*.

Henry, E. and M. Ottaviani (2019). Research and the approval process: the organization of persuasion. *American Economic Review 109*(3), 911–55.

Hirano, K. and J. R. Porter (2009). Asymptotics for statistical treatment rules. *Econometrica 77*(5), 1683–1701.

Hirano, K. and J. R. Porter (2020a). Asymptotic analysis of statistical decision rules in econometrics. *Handbook of Econometrics*.

Hirano, K. and J. R. Porter (2020b). Chapter 4 - Asymptotic analysis of statistical decision rules in econometrics. In S. N. Durlauf, L. P. Hansen, J. J. Heckman, and R. L. Matzkin (Eds.), *Handbook of Econometrics, Volume 7A*, Volume 7 of *Handbook of Econometrics*, pp. 283–354. Elsevier.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics 6*, 65–70.

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association 47*(260), 663–685.

Hu, Y., S. Li, and S. Wager (2021). Average treatment effects in the presence of interference. *arXiv preprint arXiv:2104.03802*.

Hudgens, M. G. and M. E. Halloran (2008). Toward causal inference with interference. *Journal of the American Statistical Association 103*(482), 832–842.

Ibragimov, R. and U. K. Müller (2010). t-statistic based correlation and heterogeneity robust inference. *Journal of Business & Economic Statistics 28*(4), 453–468.

Imai, K., Z. Jiang, and A. Malai (2018). Causal inference with interference and noncompliance in two-stage randomized experiments. Technical report, Technical report Princeton University.

Imai, K., G. King, and C. Nall (2009). The essential role of pair matching in cluster-randomized experiments, with application to the mexican universal health insurance evaluation. *Statistical Science 24*(1), 29–53.

Imai, K. and M. L. Li (2019). Experimental evaluation of individualized treatment rules. *arXiv preprint arXiv:1905.05389*.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika 87*(3), 706–710.

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jackson, M. O. and E. Storms (2018). Behavioral communities and the atomic structure of networks. *Available at SSRN 3049748*.

Jackson, M. O. and A. Wolinsky (1996). A strategic model of social and economic networks. *Journal of economic theory 71*(1), 44–74.

Jagadeesan, R., N. S. Pillai, and A. Volfovsky (2020). Designs for estimating the treatment effect in networks with interference. *Annals of Statistics 48*(2), 679–712.

Janson, S. (2004). Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms 24*(3), 234–248.

Johari, R., H. Li, I. Liskovich, and G. Weintraub (2020). Experimental design in two-sided platforms: An analysis of bias. *arXiv preprint arXiv:2002.05670*.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Kang, H. and G. Imbens (2016). Peer encouragement designs in causal inference with partial interference and identification of local average network effects. *arXiv preprint arXiv:1609.04464*.

Karrer, B., L. Shi, M. Bhole, M. Goldman, T. Palmer, C. Gelman, M. Konutgan, and F. Sun (2021). Network experimentation at scale. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 3106–3116.

Kasy, M. (2016). Partial identification, distributional preferences, and the welfare ranking of policies. *Review of Economics and Statistics 98*(1), 111–131.

Kasy, M. and R. Abebe (2020). Fairness, equality, and power in algorithmic decision making. Technical report.

Kasy, M. and A. Sautmann (2019). Adaptive treatment assignment in experiments for policy choice. *Econometrica*.

Kato, M. and Y. Kaneko (2020). Off-policy evaluation of bandit algorithm from dependent

samples under batch update policy. *arXiv preprint arXiv:2010.13554*.

Kempe, D., J. Kleinberg, and É. Tardos (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 137–146. ACM.

Kilbertus, N., M. R. Carulla, G. Parascandolo, M. Hardt, D. Janzing, and B. Schölkopf (2017). Avoiding discrimination through causal reasoning. In *Advances in Neural Information Processing Systems*, pp. 656–666.

Kim, B. (2020). Analysis of randomized experiments with network interference and noncompliance. *arXiv preprint arXiv:2012.13710*.

Kim, D. A., A. R. Hwong, D. Stafford, D. A. Hughes, A. J. O'Malley, J. H. Fowler, and N. A. Christakis (2015). Social network targeting to maximise population behaviour change: a cluster randomised controlled trial. *The Lancet 386*(9989), 145–153.

Kitagawa, T. and A. Tetenov (2018). Who should be treated? Empirical welfare maximization methods for treatment choice. *Econometrica 86*(2), 591–616.

Kitagawa, T. and A. Tetenov (2019). Equality-minded treatment choice. *Journal of Business & Economic Statistics*, 1–14.

Kitagawa, T. and G. Wang (2021). Who should get vaccinated? individualized allocation of vaccines over sir network. *Journal of Econometrics*.

Kleinberg, J., J. Ludwig, S. Mullainathan, and A. Rambachan (2018). Algorithmic fairness. In *Aea papers and proceedings*, Volume 108, pp. 22–27.

Kleinberg, R. D. (2005). Nearly tight bounds for the continuum-armed bandit problem. In *Advances in Neural Information Processing Systems*, pp. 697–704.

Kline, P. M., E. K. Rose, and C. R. Walters (2021, July). Systemic discrimination among large u.s. employers. Working Paper 29053, National Bureau of Economic Research.

Kusner, M., C. Russell, J. Loftus, and R. Silva (2019). Making decisions that reduce discriminatory impacts. In *International Conference on Machine Learning*, pp. 3591–3600.

Laber, E. B., D. J. Lizotte, M. Qian, W. E. Pelham, and S. A. Murphy (2014). Dynamic treatment regimes: Technical challenges and applications. *Electronic journal of statistics 8*(1), 1225.

Laber, E. B., N. J. Meyer, B. J. Reich, K. Pacifici, J. A. Collazo, and J. M. Drake (2018). Optimal treatment allocations in space and time for on-line control of an emerging infectious disease. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 67*(4), 743–789.

Ledoux, M. and M. Talagrand (2011). Probability in banach spaces. classics in mathematics.

Lee, S. and A. M. Shaikh (2014). Multiple testing and heterogeneous treatment effects: re-evaluating the effect of PROGRESA on school enrollment. *Journal of Applied Econometrics 29*(4), 612–626.

Lehmann, E. L. and J. P. Romano (2005a). Generalizations of the familywise error rate. *The Annals of Statistics 33*(3), 1138–1154.

Lehmann, E. L. and J. P. Romano (2005b). *Testing statistical hypotheses*. Springer Science & Business Media.

Lehmann, E. L., J. P. Romano, and J. P. Shaffer (2005). On optimality of stepdown and stepup multiple test procedures. *The Annals of Statistics 33*(3), 1084 – 1108.

Leung, M. P. (2019). Inference in models of discrete choice with social interactions using network data. *Available at SSRN 3446926*.

Leung, M. P. (2020). Treatment and spillover effects under network interference. *Review of Economics and Statistics 102*(2), 368–380.

Leung, M. P. (2021). Network cluster-robust inference. *arXiv preprint arXiv:2103.01470*.

Li, S. and S. Wager (2020). Random graph asymptotics for treatment effect estimation under network interference. *arXiv preprint arXiv:2007.13302*.

Li, X., P. Ding, Q. Lin, D. Yang, and J. S. Liu (2019). Randomization inference for peer effects. *Journal of the American Statistical Association*, 1–31.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics 22*(4), 773–793.

Liu, L., M. G. Hudgens, B. Saul, J. D. Clemens, M. Ali, and M. E. Emch (2019). Doubly robust estimation in observational studies with partial interference. *Stat 8*(1), e214.

Liu, Y., G. Radanovic, C. Dimitrakakis, D. Mandal, and D. C. Parkes (2017). Calibrated fairness in bandits. *arXiv preprint arXiv:1707.01875*.

Lubold, S., A. G. Chandrasekhar, and T. H. McCormick (2020). Identifying the latent space geometry of network models through analysis of curvature. Technical report, National Bureau of Economic Research.

Ludwig, J., S. Mullainathan, and J. Spiess (2017). Machine-learning tests for effects on multiple outcomes. arXiv:1707.01473.

Lyons, E. and L. Zhang (2017). The impact of entrepreneurship programs on minorities. *American Economic Review 107*(5), 303–07.

Lyons, E. and L. Zhang (2018). Who does (not) benefit from entrepreneurship programs? *Strategic Management Journal 39*(1), 85–112.

Manresa, E. (2013). Estimating the structure of social interactions using panel data. *Unpublished Manuscript. CEMFI, Madrid*.

Manski (2004). Statistical treatment rules for heterogeneous populations. *Econometrica 72*(4), 1221–1246.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of econometrics 3*(3), 205–228.

Manski, C. F. (2013). Identification of treatment response with social interactions. *The Econometrics Journal 16*(1), S1–S23.

Manski, C. F. and T. S. Thompson (1989). Estimation of best predictors of binary response. *Journal of Econometrics 40*(1), 97–123.

Martinez, N., M. Bertran, and G. Sapiro (2019). Fairness with minimal harm: A pareto-optimal approach for healthcare. *arXiv preprint arXiv:1911.06935*.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

Mbakop, E. and M. Tabord-Meehan (2021). Model selection for treatment choice: Penalized welfare maximization. *Econometrica 89*(2), 825–848.

McCloskey, A. and P. Michaillat (2020). Incentive-compatible critical values. arXiv preprint arXiv:2005.04141.

Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives 35*(3), 193–214.

Muandet, K., K. Fukumizu, B. Sriperumbudur, and B. Schölkopf (2016). Kernel mean embedding of distributions: A review and beyond. *arXiv preprint arXiv:1605.09522*.

Muralidharan, K. and P. Niehaus (2017). Experimentation at scale. *Journal of Economic Perspectives 31*(4), 103–24.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2017). General equilibrium effects of (improving) public employment programs: Experimental evidence from india. Technical report,

National Bureau of Economic Research.

Muralidharan, K., M. Romero, and K. Wüthrich (2020). Factorial designs, model selection, and (incorrect) inference in randomized experiments. NBER Working Paper.

Nabi, R., D. Malinsky, and I. Shpitser (2019). Learning optimal fair policies. *Proceedings of machine learning research 97*, 4674.

Naimi, A. I., S. R. Cole, and E. H. Kennedy (2017). An introduction to g methods. *International journal of epidemiology 46*(2), 756–762.

Narita, Y. (2021). Incorporating ethics and welfare into randomized experiments. *Proceedings of the National Academy of Sciences 118*(1).

Negishi, T. (1960). Welfare economics and existence of an equilibrium for a competitive economy. *Metroeconomica 12*(2-3), 92–97.

Newey, W. K. (1990). Semiparametric efficiency bounds. *Journal of applied econometrics 5*(2), 99–135.

Ogburn, E. L., O. Sofrygin, I. Diaz, and M. J. van der Laan (2017). Causal inference for social network data. *arXiv preprint arXiv:1705.08527*.

Opper, I. M. (2016). Does helping john help sue? evidence of spillovers in education. *American Economic Review 109*(3), 1080–1115.

Pollard, D. (1990). Empirical processes: theory and applications. In *NSF-CBMS regional conference series in probability and statistics*, pp. i–86. JSTOR.

Pouget-Abadie, J. (2018). *Dealing with Interference on Experimentation Platforms*. Ph. D. thesis.

Proschan, M. A. and E. H. Brittain (2020). A primer on strong vs weak control of familywise error rate. *Statistics in medicine 39*(9), 1407–1413.

Rai, Y. (2018). Statistical inference for treatment assignment policies.

Rambachan, A., J. Kleinberg, J. Ludwig, and S. Mullainathan (2020). An economic approach to regulating algorithms.

Robbins, H. (1951). Asymptotically subminimax solutions of compound statistical decision problems. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California.

Robins, J. M. and A. Rotnitzky (1995). Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association 90*(429), 122–129.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association 89*(427), 846–866.

Romano, J. P., A. Shaikh, and M. Wolf (2011). Consonance and the closure method in multiple testing. *The International Journal of Biostatistics 7*(1).

Romano, J. P., A. M. Shaikh, and M. Wolf (2010). Hypothesis testing in econometrics. *Annu. Rev. Econ. 2*(1), 75–104.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters 113*, 38–40.

Ross, N. (2011). Fundamentals of stein's method. *Probability Surveys 8*, 210–293.

Rotblat, J. (1999). A hippocratic oath for scientists. *Science 286*(5444), 1475–1475.

Rubin, D. B. (1990). Formal mode of statistical inference for causal effects. *Journal of statistical planning and inference 25*(3), 279–292.

Rubin, M. (2021). When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing. *Synthese*, 1–32.

Russo, D., B. Van Roy, A. Kazerouni, I. Osband, and Z. Wen (2017). A tutorial on thompson sampling. *arXiv preprint arXiv:1707.02038*.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The review of economic studies 68*(1), 205–229.

Sasaki, Y. and T. Ura (2020). Welfare analysis via marginal treatment effects. *arXiv preprint arXiv:2012.07624*.

Sävje, F., P. M. Aronow, and M. G. Hudgens (2021). Average treatment effects in the presence of unknown interference. *The Annals of Statistics 49*(2), 673–701.

Shamir, O. (2013). On the complexity of bandit and derivative-free stochastic convex optimization. In *Conference on Learning Theory*, pp. 3–24. PMLR.

Sinclair, B., M. McConnell, and D. P. Green (2012). Detecting spillover effects: Design and analysis of multilevel experiments. *American Journal of Political Science 56*(4), 1055–1069.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? causal inference in the face of interference. *Journal of the American Statistical Association 101*(476), 1398–1407.

Spiess, J. (2018). Optimal estimation when researcher and social preferences are misaligned. Working Paper.

Sriperumbudur, B. K., K. Fukumizu, A. Gretton, B. Schölkopf, and G. R. Lanckriet (2012). On the empirical estimation of integral probability metrics. *Electronic Journal of Statistics 6*, 1550–1599.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(3), 479–498.

Storey, J. D. (2003). The positive false discovery rate: a bayesian interpretation and the q-value. *The Annals of Statistics 31*(6), 2013–2035.

Storey, J. D., J. E. Taylor, and D. Siegmund (2004). Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66*(1), 187–205.

Stoye, J. (2009). Minimax regret treatment choice with finite samples. *Journal of Econometrics 151*(1), 70–81.

Stoye, J. (2012). Minimax regret treatment choice with covariates or with limited validity of experiments. *Journal of Econometrics 166*(1), 138–156.

Su, L., W. Lu, and R. Song (2019). Modelling and estimation for optimal treatment decision with interference. *Stat 8*(1), e219.

Sun, L. (2020). Empirical welfare maximization with constraints. Technical report, Working paper.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement learning: An introduction*. MIT press.

Tabord-Meehan, M. (2018). Stratification trees for adaptive randomization in randomized controlled trials. *arXiv preprint arXiv:1806.05127*.

Tetenov, A. (2012). Statistical treatment choice based on asymmetric minimax regret criteria. *Journal of Econometrics 166*(1), 157–165.

Tetenov, A. (2016). An economic theory of statistical testing. Working Paper.

Ugander, J., B. Karrer, L. Backstrom, and J. Kleinberg (2013). Graph cluster randomization: Net-

work exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 329–337. ACM.

Ustun, B., Y. Liu, and D. Parkes (2019). Fairness without harm: Decoupled classifiers with preference guarantees. In *International Conference on Machine Learning*, pp. 6373–6382.

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

Van Der Vaart, A. W. and J. A. Wellner (1996). Weak convergence. In *Weak convergence and empirical processes*, pp. 16–28. Springer.

Varian, H. R. (1976). Two problems in the theory of fairness. *Journal of Public Economics 5*(3-4), 249–260.

Vazquez-Bare, G. (2017). Identification and estimation of spillover effects in randomized experiments. *arXiv preprint arXiv:1711.02745*.

Vazquez-Bare, G. (2022). Causal spillover effects using instrumental variables. *Journal of the American Statistical Association*, 1–12.

Viviano, D. (2019). Policy targeting under network interference. *arXiv preprint arXiv:1906.10258*.

Viviano, D. (2020). Experimental design under network interference. *arXiv preprint arXiv:2003.08421*.

Wager, S. and K. Xu (2021). Experimenting in equilibrium. *Management Science*.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*, Volume 48. Cambridge University Press.

Wald, A. (1950). *Statistical decision functions*. Wiley.

Wenocur, R. S. and R. M. Dudley (1981). Some special vapnik-chervonenkis classes. *Discrete Mathematics 33*(3), 313–318.

Westfall, P. H. and S. S. Young (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Volume 279. John Wiley & Sons.

Williams, C. (2021). Preregistration and incentives. SSRN 3796813.

Wolsey, L. A. and G. L. Nemhauser (1999). *Integer and combinatorial optimization*, Volume 55. John Wiley & Sons.

Xiao, L., Z. Min, Z. Yongfeng, G. Zhaoquan, L. Yiqun, and M. Shaoping (2017). Fairness-aware group recommendation with pareto-efficiency. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pp. 107–115.

Yoder, N. (2019). Designing incentives for heterogeneous researchers. SSRN 3154143.

Zhou, Z., S. Athey, and S. Wager (2018). Offline multi-action policy learning: Generalization and optimization. *arXiv preprint arXiv:1810.04778*.

Zubcsek, P. P. and M. Sarvary (2011). Advertising to a social network. *Quantitative Marketing and Economics 9*(1), 71–107.