Locally Robust Policy Learning: Inequality, Inequality of Opportunity and Intergenerational Mobility*

Joël Terschuur

Universidad Carlos III de Madrid

July 22, 2024

Click here for the latest version of the paper.

Abstract

Policy makers need to decide whether to treat or not to treat heterogeneous individuals. The optimal choice depends on the welfare function that the policy maker has in mind. I study a general setting for policy learning with semiparametric Social Welfare Functions (SWFs) that can be estimated by locally robust/orthogonal moments estimable by U-statistics. This rich class of SWFs substantially expands the setting in Athey and Wager (2021) and accommodates a wider range of distributional preferences. Three main applications of the general theory motivate the paper: (i) Inequality aware SWFs, (ii) Inequality of Opportunity aware SWFs and (iii) Intergenerational Mobility SWFs. I use the Panel Study of Income Dynamics (PSID) to assess the effect of attending preschool on adult earnings and estimate optimal policy rules based on parental years of education and parental income.

JEL Classification: C13; C14; C21; D31; D63; I24

Keywords: local robustness, U-statistics, Inequality, Intergenerational mobility, empirical welfare maximization.

R package (forthcoming): `https://joelters.github.io/home/code/`

*Research funded by Ministerio de Ciencia e Innovación, grant ECO2017-86675-P, MCI/AEI/FEDER/UE, grant PGC 2018-096732-B-100, grant PID2021-127794NB-I00, grant PRE2020-092794 and Comunidad de Madrid, grants EPUC3M11 (VPRICIT) and H2019/HUM-589. I am grateful to Juan Carlos Escanciano and Toru Kitagawa for their guidance and support, and to Miguel Ángel Delgado, Juan José Dolado, Ignacio Ortuño, Nazarii Salish, Jan Stuhler and Carlos Velasco. *Email address: jrobert@eco.uc3m.es*

1 Introduction

Whenever a policy or treatment has heterogeneous effects it is important to decide carefully who should be treated. In the simplest case in which we care only about the average outcome, there are no budgetary limits and the treatment effect is positive for everyone, it follows that the best policy is to treat everyone. However, in most cases, we do not have such a luxury. We might have a limited budget, distributional concerns or negative treatment effects for some individuals. In these cases, it is important to decide whether *to treat or not to treat* different individuals. This is the problem of policy learning.

Such a problem is omnipresent not only in economics but in business, law, education, medicine and many other fields of inquiry. While in economics we might want to know whether to provide training or not to the unemployed, decide different rules to assign conditional cash transfers or even whether a transfer should be given unconditionally, in business we might want to know whether to provide a discount to a customer or not, or whether we should send price recommendations to some stores and not to others. Judges might have to decide whether to release someone on parole or not based on the recidivism probability of the individual. In education, we might want to know whether to provide a scholarship to a student or not or whether we should provide additional extra-curricular lessons. Certain medicines might be beneficial for some but detrimental for others or there might not be enough vaccines to cover the whole population as seen in the COVID-19 pandemic.

The inherent ethical and distributional considerations in all these examples are quite different. Hence, it is important to have a general framework that can accommodate different welfare functions. While the framework has to be as general as possible, it also needs to allow for the estimation of optimal policy rules with certain statistical guarantees. In this paper, I provide a framework to compute such optimal rules for a rich class of semiparametric welfare functions, estimable by U-statistics. This includes, among many others, the standard additive welfare function but also: (i) Inequality aware SWFs, (ii) Inequality of Opportunity (IOp) aware SWFs and (iii) Intergenerational Mobility SWFs. These three SWFs are of great interest to policy makers and motivate this paper.

To my knowledge, there is no prior work on IOp and Intergenerational mobility aware social welfare functions in the policy learning literature. IOp is the part of inequality that is explained by circumstances X outside the control of the individual, e.g. sex, race, parental education or parental income. Hence, IOp SWFs are useful whenever we do not want to penalize all inequality but just unfair inequality (i.e. inequality explained by circumstances). Based on the seminal contributions in Van De Gaer (1993), Fleurbaey (1995) and Roemer (1998) the IOp literature has grown and focused on how to measure IOp. A popular measure of IOp is the Gini of the best predictions (in mean squared error sense) of the outcome Y given the circumstances X, i.e. $G(\gamma(X))$ where $\gamma(X) = \mathbb{E}[Y|X]$ and where henceforth $G(Z)$ denotes the Gini inequality index of the generic random variable Z. To accommodate a possibly high-dimensional set of circumstances, IOp literature has started using machine learners to estimate the predictions (e.g Brunori et al. (2019a), Brunori et al. (2019b), Brunori et al. (2021), Brunori and Neidhöfer (2021), Rodríguez et al. (2021), Carranza (2022) or Hufe et al. (2022) among others). As usual in such two-step procedures, the bias-variance trade-off in the prediction might allow for some bias which can creep into the second stage. Escanciano and Terschuur (2023) provide locally robust IOp estimators that are robust to such biases. I use these results to construct IOp aware SWFs.

Inequality aware SWFs have been studied before in Kasy (2016) and Kitagawa and Tetenov (2021). A popular welfare function for a random outcome Y is $W = \mathbb{E}[Y](1 - G(Y))$. This welfare function values the average outcome but penalizes high inequality (the Gini is between 0 and 1, where 0 is complete equality and 1 complete inequality). This framework allows for a simpler way to deal with the case in which the population eligible for treatment and the population whose welfare we care about differ. For instance, we might want to study how to allocate some transfer among the poor but care about the inequality in the whole population. I do so by using the fact that the Gini coefficient can be written as a U-statistic which prevents me from having to use the cumulative distribution function of the outcome as in Kitagawa and Tetenov (2021).

Leqi and Kennedy (2021) study optimal treatment regimes to maximize average conditional quantiles. Wang et al. (2018) study quantile-optimal treatment regimes and adapt their theory to study the minimization of Gini's mean difference. They also employ the second-order U-statistics nature of the Gini mean difference and obtain asymptotic theory for a particular class of policy rules by using empirical U-process methods. I avoid the use of U-processes by using an alternative representation of U-statistics as sums-of-i.i.d. blocks, as explained in Clémençon et al. (2008).

While inequality aware SWFs look at the distribution of Y, IOp aware SWFs focus on the distribution of the predictions $\gamma(X) = \mathbb{E}[Y|X]$. A natural IOp aware SWF would be $W = \mathbb{E}[\gamma(X)](1 - G(\gamma(X))) = \mathbb{E}[Y](1 - G(\gamma(X)))$, which only penalizes inequalities explained by circumstances. This example adds an extra unknown nuisance parameter, $\gamma(X)$, on top of the conditional expectations/propensity scores needed to identify treatment effects. There is no previous work on policy learning with general semiparametric welfare functions which depend on additional unknown functions aside from those needed to identify treatment effects. There is an example of a specific welfare function that depends on conditional quantiles (average conditional median) in Leqi and Kennedy (2021).

Intergenerational mobility is the study of the relationship between the outcomes of parents and the outcomes of their children. The Kendall-τ is a popular measure of intergenerational mobility in the literature (see Chetty et al. (2014) or Kitagawa et al. (2018)). It looks at whether it is the case that if the parents of individual i are richer than those of j, i is also richer than j. A natural intergenerational mobility aware SWF would be $W = -|\tau - t|$ for some target Kendall-τ $t \in [-1, 1]$. This example is of interest in, for instance, deciding the allocation of higher education scholarships to students based on individual characteristics whenever the treatment effect of education on long-term income can be identified and there is a policy interest in reducing the association between parental income and child's income.

The technical goal in the policy learning literature, sometimes called empirical welfare maximization or offline policy learning in the computer science literature, is to find an optimal allocation rule π which maps individual characteristics to a binary decision $\{0, 1\}$ of treatment or no treatment. This optimal rule is searched within a class Π of plausible treatment rules to maximize some welfare function. Following the seminal work in Manski (2004), I search for an optimal policy in the plausible class so as to minimize regret, i.e. the expected difference between the best possible welfare and the welfare evaluated at the estimated policy. Other relevant work on treatment rules in econometrics includes Dehejia (2005), Hirano and Porter (2009), Stoye (2009, 2012), Chamberlain (2011), Bhattacharya and Dupas (2012), Tetenov (2012), Kasy (2016), Kitagawa and Tetenov (2018, 2021), Athey and Wager (2021), or Zhou et al. (2023).

Non-parametric estimation of unknown functions in semiparametric welfare functions poses a challenge to the statistical guarantees of estimated policy rules. This is due to the slow convergence rate of non-parametric estimators such as kernels or machine learners. The semiparametric literature has developed methods to overcome this problem by using locally robust/orthogonal scores. These are alternative moment conditions that identify the quantity of interest and allow for its estimation at a parametric ($\sqrt{n}$) rate. I expand previous work by considering any semiparametric welfare function, possibly defined as a U-statistic, which can be estimated by locally robust/orthogonal scores. The main theoretical result is to provide an asymptotic upper bound to the regret of the estimated policy rule.

This paper is close to Kitagawa and Tetenov (2021) in taking into account other distributional aspects aside from the mean and it is also closely related to Athey and Wager (2021), Leqi and Kennedy (2021) and Zhou et al. (2023) in making use of the latest semiparametric literature on locally robust/orthogonal scores (e.g. Chernozhukov et al. (2022)) to obtain parametric rates of convergence even with slow nonparametric first steps. I also build upon Escanciano and Terschuur (2023) to expand policy learning results to welfare functions defined by U-statistics. The key result in Athey and Wager (2021) is to find rates of the regret that optimally depend on the complexity of the policy class, Π, and the sample size in observational settings where the propensity score is unknown. They do so for average-treatment-like welfare functions. I substantially generalize this setting by allowing arbitrary semiparametric welfare functions, possibly defined as U-statistics, as long as they can be identified with locally robust/orthogonal scores.

Empirically, treatment allocation with inequality, IOp and IGM welfare functions poses many challenges. First, we need rich information on circumstances and parental income which are absent in many modern datasets. Second, we need a credible identification strategy to identify treatment effects. In this paper, I tackle these challenges by looking at the effect of attending preschool on adult earnings using the Panel Study of Income Dynamics (PSID) dataset. This application has many advantages. First, any variable that induces preschool attendance can be considered a circumstance under the (very reasonable) assumption that we cannot hold the kid responsible for these variables. Second, PSID has been following families for nearly 50 years meaning we have rich information on family background. Third, PSID allows us to look at long-term outcomes such as adult earnings.

The empirical application is not free of problems. The treatment is not randomly assigned. Hence, the identification of treatment effects relies on the assumption of selection on observable circumstances. Also, attending preschool is not a binary treatment since some preschools might be better than others. Furthermore, I have no information on the cost of treatment. Finally, allocating children to preschool based on their circumstances might not be enforceable or ethical. In fact, I show that the preschool choices observed in the sample differ greatly from the estimated optimal rules even when we maximize the average outcome. This points out that parents have different considerations when sending their kids to preschool aside from future earnings. Hence, I interpret the empirical exercise as a thought experiment to illustrate the main contributions of the paper.

The effect of preschool on short/medium/long-term outcomes has been extensively studied and there is a public interest in expanding public preschool programs in the US. The share of 4-year-olds in public preschool has grown from 14% in 2002 to 34% in 2019 and many states and large cities in the US now operate large-scale public preschool programs (Gray-Lobe et al. (2023)). The first popular small-scale randomized preschool experiments in the US were the High/Scope Perry Preschool project and Carolina Abecedarian project whose participants have been followed for decades leading to many studies showing positive effects (Campbell and Ramey (1994), Campbell et al. (2012), Heckman et al. (2013), García et al. (2020)). Gray-Lobe et al. (2023) use admission lotteries to study the impact of the large-scale public preschool in Boston on a range of outcomes. They find positive effects and varying treatment effects based on gender. Heckman and Raut (2016) study the impact of preschool using a structural model and find that a tax-financed public preschool program targeted at children with poor socioeconomic status increases average earnings and increases intergenerational mobility.

In this paper, I find that the effect of preschool attendance is heterogeneous. While on average preschool has a positive effect on adult earnings, children with highly educated mothers and high parental income are negatively affected by preschool. This is in line with results in the psychology and economics literature which document that in early educational institutions, there is less interaction with adults and hence there can be a negative effect of attending such institutions if the interactions with adults in the household are of "higher-quality" (see Fort et al. (2020)).

These heterogeneous effects have different implications when estimating optimal treatment rules for different welfare functions. I compute optimal treatment rules based on parental income and mother's education in a class of decision trees of depth two. Estimated optimal rules that maximize the average try to treat anyone who has a positive treatment effect. Inequality aware SWFs include individuals to treatment who have a negative treatment effect since the decrease in average earnings is compensated by a decrease in inequality. This possibility could be ruled out by restricting ourselves to trees that do not treat groups with negative estimated treatment effects. The same happens with the IGM welfare which has no average motive at all and only cares about decreasing the association between parental income and child's income to zero. I find that the additive and IOp estimated optimal policy rules coincide while the inequality and IGM estimated rules treat individuals who do not benefit from treatment or do not treat individuals who do benefit from treatment. The coincidence of the additive and IOp rules is specific to the structure of the heterogeneous treatment effects in the data and is not a general result. In this empirical application, we can see that this happens mostly because maximizing the average already decreases IOp drastically.

I start by introducing the main welfare objects which are going serve as guiding examples of the general theory in Section 2. Section 3 elaborates on the general theory for general welfare functions identified by locally robust/orthogonal scores which are linear on the distribution of the data (i.e. not defined as U-statistics) and Section 4 expands the results to general welfare functions, possibly defined as U-statistics. Section 5 provides upper bounds on the regret of estimated policies and Section 6 deals with the empirical application. All proofs are in the Appendix.

2 Welfare economics for inequality, IOp and rank correlations

The policy learning literature is at the intersection of welfare economics and econometrics. Before we delve into the econometric problem of estimating optimal rules and evaluating their statistical performance, I present in this section the main welfare objects we are going to be interested in. Suppose we have some continuous random outcome $Y_i \in \mathbb{R}^+$. The most basic welfare function is the additive welfare based on the average outcome

$$W = \mathbb{E}[Y_i].$$

The above welfare does not care about other distributional aspects apart from the average outcome. A first approach to include distributional concerns in our analysis is to follow Dalton

(1920) and Atkinson et al. (1970) and consider increasing and concave transformations $u(\cdot)$ of the outcome[1]

$$W = \mathbb{E}[u(Y_i)].$$

This welfare function will already rank two outcome distributions in the same way for all increasing and concave $u(\cdot)$ if the Lorenz curve of one of the distributions is everywhere above the Lorenz curve of the other distribution and has equal or higher mean; equivalently if one distribution second-order stochastically dominates the other. However, if we want to obtain a complete ordering we need to specify $u(\cdot)$ further. One popular choice is

$$u(y) = \begin{cases} \frac{y^{1-\theta}}{1-\theta} & \text{if } \theta \in (0,1) \\ \log(y) & \text{if } \theta = 1, \end{cases}$$

where θ captures the concavity of $u(\cdot)$ and can therefore be interpreted as an inequality aversion parameter. This paper also focuses on welfare which is aware of Inequality of Opportunity (IOp). IOp is the part of total inequality which can be explained by circumstances, i.e. by variables that are outside the control of the individual such as parental education or parental income. Let $X_i \in \mathbb{R}^k$ be such a random vector of circumstances. Let also $\gamma(X_i) = \mathbb{E}[Y_i|X_i]$, i.e. the best predictor (in mean squared error sense) of the outcome Y_i given the circumstances X_i. By looking at the distribution of $\gamma(X_i)$ instead of that of the outcome Y_i we get IOp averse welfare functions. For instance,

$$W = \mathbb{E}[u(\gamma(X_i))].$$

If there is no IOp, circumstances are unable to predict the outcome and we have that the best predictor is the unconditional mean: $\gamma(X_i) = \mathbb{E}[Y_i]$. In this case, we have that $W = u(\mathbb{E}[Y_i])$ so we only care about the average income (with a different scale due to $u(\cdot)$). If we have maximum IOp, the outcome is a deterministic function of the circumstances and $\gamma(X_i) = Y_i$. Then, $W = \mathbb{E}[u(Y_i)]$. Since all inequality is IOp, we are back to the inequality averse welfare function.

Another option to take into account distributional concerns is to weigh differently different parts of the distribution. Let F_Y be the distribution of the outcome and F_Y^{-1} be the quantiles. Then, for some weights $w(\cdot)$ a planner might have the following welfare in mind

$$W = \int_0^1 F_Y^{-1}(\tau) w(\tau) d\tau.$$

This welfare has been used in Mehran (1976), Donaldson and Weymark (1980), Weymark (1981), Donaldson and Weymark (1983) or Aaberge et al. (2021). If we let $w_k(\tau) = (k-1)(1-\tau)^{k-2}$ we get what is known as the extended Gini family of social welfare functions. In this paper, I

[1] With abuse of notation I call W to all welfare functions as they appear.

focus on $k = 3$ which is known as the standard Gini social welfare function and can be shown to be

$$\begin{aligned} W &= \mathbb{E}[Y_i](1 - G(Y_i)) \\ &= (1/2)\mathbb{E}[Y_i + Y_j - |Y_i - Y_j|], \end{aligned}$$

where the second equality follows from the fact that we can write the Gini of Y_i as $G(Y_i) = \mathbb{E}[|Y_i - Y_j|]/\mathbb{E}[Y_i + Y_j]$ where Y_j is a copy of Y_i (i.e. the Gini can be interpreted as a normalized expected absolute distance between the outcomes of two individuals taken at random). The welfare above is additive as long as there is no inequality ($G(Y_i) = 0$) and penalizes positive values of the Gini coefficient. Again, if we do not care about inequality but only about IOp we can look at the distribution of $\gamma(X_i)$ instead of the distribution of Y_i. In that case, we have

$$\begin{aligned} W &= \mathbb{E}[\gamma(X_i)](1 - G(\gamma(X_i))) \\ &= (1/2)\mathbb{E}[\gamma(X_i) + \gamma(X_j) - |\gamma(X_i) - \gamma(X_j)|]. \end{aligned}$$

If there is no IOp, then $G(\gamma(X_i)) = 0$ and we are back in the additive case. If there is full IOp, then $G(\gamma(X_i)) = G(Y_i)$ and we are back to the standard Gini social welfare function of outcome Y_i. Finally, I also consider the problem of intergenerational mobility. Let $X_{1i} \in \mathbb{R}$ be the parental outcome. A non-linear measure of association between Y_i and X_{1i} is the Kendall-τ

$$\tau = \mathbb{E}[sgn(Y_i - Y_j)sgn(X_{1i} - X_{1j})],$$

where $sgn(a) = \mathbb{1}(a > 0) - \mathbb{1}(a < 0)$. This parameter is popular in the intergenerational mobility literature (see Chetty et al. (2014) or Kitagawa et al. (2018)) where X_{1i} is parental income and Y_i is the child's income. It takes values between 1 and -1. $\tau = 1$ means that whenever an individual has a higher income than another, she also has a higher parental income and vice versa. $\tau = -1$ is the opposite, whenever someone has a higher income, she has a lower parental income. For some target Kendall-τ $t \in [-1, 1]$ an intergenerational mobility aware welfare function is

$$W = -\left|\mathbb{E}\left[sgn(Y_i - Y_j)sgn(X_{1i} - X_{1j})\right] - t\right|.$$

To my knowledge, this welfare function has not been used before in the literature. Note that it allows us to treat problems much more general than intergenerational mobility. Setting $t = 0$, maximizing this welfare function corresponds to allocating a treatment to minimize the dependence between two variables Y_i and X_{1i}. For instance, we could study how to allocate scholarships to make academic attainment less dependent on parental income.

3 Policy learning with general orthogonal scores

Consider random variables $(Y_i(1), Y_i(0), D_i, X_i) \sim F_0$ where $(Y_i(1), Y_i(0)) \in \mathcal{Y} \times \mathcal{Y}$ are real-valued potential outcomes, i.e. $Y_i(1)$ is the outcome of individual i under treatment and $Y_i(0)$ is the outcome of individual i in the absence of treatment. D_i is a binary treatment and $X_i \in \mathcal{X}$ is now a vector of pre-treatment covariates. Let $\gamma^{(j)}(X_i) = \mathbb{E}[Y_i(j)|X_i] \in \Gamma$ for $j = 0, 1$ be potential predictions, i.e. the predictions of the potential outcomes given X_i. We observe an i.i.d. sample $(Z_1, ..., Z_n)$ with $Z_i = (Y_i, D_i, X_i) \in \mathcal{Z}$ and $Y_i = Y_i(1)D_i + Y_i(0)(1 - D_i) \in \mathcal{Y}$. Let $\pi : \mathcal{X} \mapsto \{0, 1\}$ be a treatment rule which indicates who receives treatment and Π be a collection of such treatment rules. We are interested in choosing a policy $\pi \in \Pi$ so as to maximize the following welfare function

$$W(\pi) = \mathbb{E}[g(Y_i(1), X_i, \gamma^{(1)})\pi(X_i) + g(Y_i(0), X_i, \gamma^{(0)})(1 - \pi(X_i))]. \tag{3.1}$$

In the simplest example of additive welfare, $g(Y_i(j), X_i, \gamma^{(j)}) = Y_i(j)$ for $j = 0, 1$. Importantly, in (3.1), g can depend on possibly infinite-dimensional unknown nuisance parameters γ. While throughout the paper I consider γ to be a conditional expectation of the outcome given X, this framework can be extended to allow for much more general first steps such as high-dimensional quantile regressions (see Ichimura and Newey (2022)).

Example 1 (IOp Atkinson) *If we are interested in an inequality averse SWF we can use Atkinson SWF,* $W(\pi) = \mathbb{E}[u(Y_i(1))\pi(X_i) + u(Y_i(0))(1 - \pi(X_i))]$ *with* $u(\cdot)$ *a concave function and* X_i *a vector of circumstances. In this case, the optimal policy can be estimated using the methods in Kitagawa and Tetenov (2018) and Athey and Wager (2021). If we want an IOp averse SWF we can look at the distribution of* $\gamma(X_i)$ *instead of at the distribution of* Y_i*:*

$$W(\pi) = \mathbb{E}[u(\gamma^{(1)}(X_i))\pi(X_i) + u(\gamma^{(0)}(X_i))(1 - \pi(X_i))].$$

■

(3.1) is not observable since for a given individual we do not observe both potential outcomes. To identify (3.1) we first need our sample to come from an experimental or observational experiment where the policy has already been implemented. Let $e(X_i) = \mathbb{P}(D_i = 1|X_i)$ be the propensity score. I assume that the following holds.

Assumption 1 *i)* $(Y_i(1), Y_i(0)) \perp D_i | X_i$,

ii) There exists $\kappa \in (0, 1/2]$ *such that* $e(x) \in [\kappa, 1 - \kappa]$.

The next proposition states the first identification result. There are two ways of identifying welfare, either using what is usually called the direct method (DM) based on conditional expectations or using Inverse Propensity Score Weighting (IPW). I focus on the DM approach since it leads to simpler expressions. I derive all the results in the paper for the IPW approach in Appendix A. Let $\gamma(D_i, X_i) = \mathbb{E}[Y_i|D_i, X_i]$, $\gamma_j(X_i) = \gamma(j, X_i)$ for $j = 0, 1$ and $\varphi(D_i, X_i, \gamma) = \mathbb{E}[g(Y_i, X_i, \gamma)|D_i, X_i]$.

Proposition 3.1 *Under Assumption 1, $W(\pi)$ is identified as*

$$W(\pi) = \mathbb{E}[\varphi(1, X_i, \gamma_1)\pi(X_i) + \varphi(0, X_i, \gamma_0)(1 - \pi(X_i))].$$

Note that if g only depends on potential nuisance parameters but not on actual potential outcomes directly, i.e. $g(u, X_i, \gamma^{(j)}) = g(t, X_i, \gamma^{(j)}) \equiv g(X_i, \gamma^{(j)})$ for all $u, t \in \mathcal{Y}$ then φ is known since $\varphi(D_i, X_i, \gamma) = g(X_i, \gamma)$. This is the case in all IOp examples such as Example 1 and Example 3 below. This is not the case in the rest of the examples. Hence, depending on which case we are we will have either γ or (γ, φ) as nuisance parameters. To enjoy local robustness to high dimensional and ML first steps, I provide orthogonal scores in the next result. First, I need the following assumption to take care of the nuisance parameter γ.

Assumption 2 *There exist (α_1, α_0) such that for any $\tilde{\gamma} \in L_2$ and $j = 0, 1$ and $\tau \geq 0$*

$$\frac{d}{d\tau}\mathbb{E}[\varphi(j, X_i, \bar{\gamma}_\tau)]\bigg|_{\tau=0} = \frac{d}{d\tau}\mathbb{E}[\alpha_j(D_i, X_i)\bar{\gamma}_\tau(D_i, X_i)]\bigg|_{\tau=0},$$

where $\bar{\gamma}_\tau = \gamma + \tau\tilde{\gamma}$ and $\mathbb{E}[\alpha_j(D_i, X_i)^2] < \infty$.

This is a common assumption in the semiparametric and orthogonal moments literature (e.g. (4.1) in Newey (1994)) and allows for φ to depend non-linearly on γ generalizing Assumption 1 in Athey and Wager (2021). Since γ enters φ only through g, a sufficient condition is to assume a similar result for the function g instead of φ and then it will be straightforward to find α. Orthogonal scores usually take the form of the original identifying score plus mean zero correction terms based on residuals which make the score locally robust to first steps.

Proposition 3.2 *The orthogonal score is given by*

$$\Gamma_i(\pi) = \Gamma_{1i}\pi(X_i) + \Gamma_{0i}(1 - \pi(X_i)),$$

where

$$\Gamma_{1i} = \varphi(1, X_i, \gamma) + \frac{D_i}{e(X_i)}(g(Y_i, X_i, \gamma_1) - \varphi(1, X_i, \gamma)) + \alpha_1(D_i, X_i)(Y_i - \gamma(D_i, X_i)),$$
$$\Gamma_{0i} = \varphi(0, X_i, \gamma) + \frac{1 - D_i}{1 - e(X_i)}(g(Y_i, X_i, \gamma_0) - \varphi(0, X_i, \gamma)) + \alpha_0(D_i, X_i)(Y_i - \gamma(D_i, X_i)).$$

As expected, orthogonal scores are formed by identifying scores (φ already identifies the welfare) and correction terms for nuisance parameters φ and γ. Note that whenever g does not depend on the potential outcomes directly (e.g. Examples 1 and 3) we have that $g(Y_i, X_i, \gamma_j) - \varphi(j, X_i, \gamma) = 0$ for $j = 0, 1$ so we have $\Gamma_{1i} = \varphi(1, X_i, \gamma) + \alpha_1(D_i, X_i, g)(Y_i - \gamma(D_i, X_i))$ and $\Gamma_{0i} = \varphi(0, X_i, \gamma) + \alpha_0(D_i, X_i, g)(Y_i - \gamma(D_i, X_i))$. To estimate the welfare for a given $\pi \in \Pi$ we employ cross-fitting as in Chernozhukov et al. (2022). Let the data be split in L groups $I_1, ..., I_l$, then

$$\hat{W}_n(\pi) = \frac{1}{n} \sum_{l=1}^{L} \sum_{i \in I_l} \hat{\Gamma}_{1i,l} \pi(X_i) + \hat{\Gamma}_{0i,l}(1 - \pi(X_i)),$$

where

$$\hat{\Gamma}_{1i,l} = \hat{\varphi}_l(1, X_i, \hat{\gamma}_l) + \frac{D_i}{\hat{e}_l(X_i)}(Y_i - \hat{\varphi}_l(1, X_i, \hat{\gamma}_l)) + \hat{\alpha}_{1,l}(D_i, X_i, \nu)(Y_i - \hat{\gamma}_l(D_i, X_i)),$$
$$\hat{\Gamma}_{0i,l} = \hat{\varphi}_l(0, X_i, \hat{\gamma}_l) + \frac{1 - D_i}{1 - \hat{e}_l(X_i)}(Y_i - \hat{\varphi}_l(0, X_i, \hat{\gamma}_l)) + \hat{\alpha}_{0,l}(D_i, X_i, \nu)(Y_i - \hat{\gamma}_l(D_i, X_i)),$$

and $(\hat{\varphi}_l, \hat{e}_l, \hat{\gamma}_l, \hat{\alpha}_{j,l})$, $j = 0, 1$, are estimators of the nuisance functions which do not use observations in I_l. Again, whenever g does not depend on the potential outcomes, the middle term in both expressions is zero. This is the case in the example of Atkinson welfare IOp.

Example 1 (IOp Atkinson (cont.)) *For $\theta \in (0, 1]$, let*

$$U(\gamma(x)) = \begin{cases} \frac{\gamma(x)^{1-\theta}}{1-\theta} & \text{if } \theta \in (0,1) \\ \log(\gamma(x)) & \text{if } \theta = 1. \end{cases}$$

In this case, $g = U$ which only depends on the nuisance parameters. The orthogonal score for $\theta \in (0, 1]$ is

$$\begin{aligned} \Gamma_i(\pi) = {} & U(\gamma(1, X_i)) + \frac{\gamma(D_i, X_i)^{-\theta} D_i}{e(X_i)}(Y_i - \gamma(D_i, X_i))\pi(X_i) \\ & + U(\gamma(0, X_i)) + \frac{\gamma(D_i, X_i)^{-\theta}(1 - D_i)}{1 - e(X_i)}(Y_i - \gamma(D_i, X_i))(1 - \pi(X_i)), \end{aligned}$$

i.e. $\alpha_1(D_i, X_i, g) = e(X_i)^{-1}\gamma(D_i, X_i)^{-\theta} D_i$ and $\alpha_0(D_i, X_i, g) = (1 - e(X_i))^{-1}\gamma(D_i, X_i)^{-\theta}(1 - D_i)$. ■

The estimator of the optimal treatment rule among a class of rules Π is

$$\hat{\pi} = \arg\max_{\pi \in \Pi} \hat{W}_n(\pi).$$

Before analyzing the statistical performance of such a rule let us first extend the results in this section to welfare functions estimable by U-statistics. This will allow us to consider inequality, IOp and intergenerational mobility aware SWFs based on the Gini coefficient and the Kendall-τ.

4 Policy learning with U-statistics

Let now $\pi_{ab}(X_i, X_j) = \mathbb{1}(\pi(X_i) = a) \times \mathbb{1}(\pi(X_j) = b)$ with $a, b \in \{0, 1\}$. Now we consider the following SWFs

$$W(\pi) = \mathbb{E}\Bigg[\sum_{(a,b)\in\{0,1\}^2} g(Y_i(a), X_i, Y_j(b), X_j, \gamma^{(a)}, \gamma^{(b)})\pi_{ab}(X_i, X_j)\Bigg]. \tag{4.1}$$

There are several differences compared to the previous setting. First, $W(\pi)$ now depends on expected pairwise comparisons. For instance, in the Gini, we look at the expected absolute distance between two individuals taken at random. Second, we are summing across $\{0, 1\}^2$. This is because we have to take into account when both members of the pair are under treatment, or just one of them or none of them. Finally, we have $\pi_{ab}(X_i, X_j)$ instead of $\pi(X_i)$ since we need to account for when both member of the pair are allocated to treatment, just one of them or none of them.

Example 2 (Inequality) *We can accommodate the standard Gini welfare function with*

$$g(Y_i(a), Y_j(b)) = (1/2)(Y_i(a) + Y_j(b) - |Y_i(a) - Y_j(b)|).$$

■

Example 3 (Inequality of Opportunity IOp) *We can apply the standard Gini welfare function to the distribution of the predictions to get $\mathbb{E}[\gamma(X_i)](1 - G(\gamma(X_i)))$. This fits our setting by letting*

$$g(X_i, X_j, \gamma^{(a)}, \gamma^{(b)}) = (1/2)(\gamma^{(a)}(X_i) + \gamma^{(b)}(X_j) - |\gamma^{(a)}(X_i) - \gamma^{(b)}(X_j)|).$$

■

Example 4 (Kendal-τ) *If we want to allocate a treatment targeting a specific Kendall-τ, say $t \in \mathbb{R}$, we have to extend our setting to transformations of the right-hand side of 4.1. We can define*

$$g(Y_i(a), X_{1i}, Y_j(b), X_{1j}) = sgn(Y_i(a) - Y_j(b))sgn(X_{1i} - X_{1j}),$$

and let

$$W(\pi) = -\Bigg|\mathbb{E}\Bigg[\sum_{(a,b)\in\{0,1\}^2} g(Y_i(a), X_{1i}, Y_j(b), X_{1j})\pi_{ab}(X_i, X_j)\Bigg] - t\Bigg|.$$

■

For $a, b \in \{0, 1\}$ let now $\varphi(a, X_i, b, X_j, \gamma_a, \gamma_b) = \mathbb{E}[g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)|D_i = a, X_i, D_j = b, X_j]$ and $e_{ab}(X_i, X_j) = e_a(X_i)e_b(X_j)$ where for $c \in \{0, 1\}$, $e_c(X_i) = \mathbb{P}(D_i = c|X_i)$.

Proposition 4.1 *Under Assumption 1, $W(\pi)$ in (4.1) is identified in the following way*

$$W(\pi) = \mathbb{E}\Big[\sum_{(a,b)\in\{0,1\}^2} \varphi(a, X_i, b, X_j, \gamma_a, \gamma_b)\pi_{ab}(X_i, X_j).\Big],$$

For an IPW version of this result see Appendix A. Again, note that whenever g does not depend directly on the potential outcomes we have that $\varphi(a, X_i, b, X_j, \gamma_a, \gamma_b) = g(X_i, X_j, \gamma_a, \gamma_b)$ as we can see in Example 3 below. Whenever g does depend on the potential outcomes, $\varphi(a, X_i, b, X_j, \gamma_a, \gamma_b)$ must be estimated using dyadic regressions. Now we apply Proposition 4.1 to identify the welfare in each of our three main examples.

Example 2 (Inequality (cont.)) *In this example, welfare is identified by*

$$W(\pi) = \mathbb{E}\Bigg[\frac{1}{2}\sum_{(a,b)\in\{0,1\}^2} \mathbb{E}(Y_i + Y_j - |Y_i - Y_j| \mid D_i = a, X_i, D_j = b, X_j)\pi_{ab}(X_i, X_j)\Bigg].$$

■

Example 3 (IOp (cont.)) *In this example, welfare is identified by*

$$W(\pi) = \frac{1}{2}\mathbb{E}\Bigg[\sum_{(a,b)\in\{0,1\}^2} \Big(\gamma_a(X_i) + \gamma_b(X_j) - |\gamma_a(X_i) - \gamma_b(X_j)|\Big)\pi_{ab}(X_i, X_j)\Bigg].$$

■

Example 4 (Intergenerational mobility (cont.)) *In this example, welfare is identified by*

$$W(\pi) = -\Bigg|\mathbb{E}\Bigg[\frac{1}{2}\sum_{(a,b)\in\{0,1\}^2} \mathbb{E}(sgn(X_{1i} - X_{1j})sgn(Y_i - Y_j) \mid D_i = a, X_i, D_j = b, X_j)\pi_{ab}(X_i, X_j)\Bigg] - t\Bigg|.$$

■

Example 3 does not depend on the potential outcomes which makes the expression simpler. To compute the orthogonal scores we need to assume a similar linearization property as the one in Assumption 2 and to the linearization assumed in Escanciano and Terschuur (2023).

Assumption 3 *There exist $\alpha_{ab,p}$, $P < \infty$, and (c_{1p}, c_{2p}) for $p = 1, ..., P$, such that for all $(a, b) \in \{0, 1\}^2$ the following linearization holds*

$$\frac{d}{d\tau}\mathbb{E}[\varphi(a, X_i, b, X_j, \bar{\gamma}_\tau)] = \frac{d}{d\tau}\mathbb{E}\Bigg[\sum_{p=1}^{P} \alpha^{\gamma}_{ab,p}(D_i, X_i, D_j, X_j)(c_{1p}\bar{\gamma}_\tau(D_i, X_i) + c_{2p}\bar{\gamma}_\tau(D_j, X_j))\Bigg],$$

where $\bar{\gamma}_\tau$ is defined as in Assumption 2 and $\mathbb{E}[\alpha^{\gamma}_{ab,p}(D_i, X_i, D_j, X_j)^2] < \infty$.

Again, γ enters φ only through g so a sufficient condition that would allow to compute α_{ab} is to assume a linearization like the above for g instead of for φ. P is usually not greater than two. Now we are ready to present the result of the orthogonal scores for welfare functions estimable by U-statistics.

Proposition 4.2 *The orthogonal scores are given by*

$$\Gamma_{ij}(\pi) = \sum_{(a,b)\in\{0,1\}^2} \Gamma_{ij}^{ab}\pi_{ab}(X_i, X_j),$$

where

$$\Gamma_{ij}^{ab} = \varphi(a, X_i, b, X_j, \gamma_a, \gamma_b) + \phi_{ab}^{\varphi}(D_i, X_i, D_j, X_j, \varphi, \alpha^e) + \phi_{ab}^{\gamma}(D_i, X_i, D_j, X_j, \gamma, \alpha^{\gamma}),$$

where

$$\phi_{ab}^{\gamma}(D_i, X_i, D_j, X_j, e, \alpha^{\gamma}) = \sum_{p=1}^{P} \alpha_{ab,p}^{\gamma}(D_i, X_i, D_j, X_j, e)(c_{1p}Y_i + c_{2p}Y_j - c_{1p}\gamma(D_i, X_i) - c_{2p}\gamma(D_j, X_j)),$$

$$\phi_{ab}^{\varphi}(D_i, X_i, D_j, X_j, \varphi, \alpha^m) = \alpha_{ab}^{\varphi}(D_i, X_i, D_j, X_j)(g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b) - \varphi(D_i, X_i, D_j, X_j, \gamma_a, \gamma_b)),$$

and

$$\alpha_{ab}^{\varphi}(D_i, X_i, D_j, X_j) = \frac{D_{ij}^{ab}}{e_{ab}(X_i, X_j)}, \quad D_{ij}^{ab} = \mathbb{1}(D_i = a)\mathbb{1}(D_j = b).$$

Once again, note that whenever g does not directly depend on the potential outcomes then $g = \varphi$ and we have that $\phi_{ab}^{\varphi} = 0$. For a version of this result using IPW see Appendix A. Now we can see how Proposition 4.2 applies to our examples.

Example 2 (Inequality (cont.)) *In this example, we have that*

$$\begin{aligned}\Gamma_{ij}^{ab} &= \frac{1}{2}\mathbb{E}(Y_i + Y_j - |Y_i - Y_j| \mid D_i = a, X_i, D_j = b, X_j) \\ &+ \frac{D_{ij}^{ab}}{2e_{ab}(X_i, X_j)}(Y_i + Y_j - |Y_i - Y_j| - \mathbb{E}(Y_i + Y_j - |Y_i - Y_j| \mid D_i = a, X_i, D_j = b, X_j)).\end{aligned}$$

■

Example 3 (IOp (cont.)) *I introduce the orthogonal score of the IOp example as a Proposition with its proof in the Appendix.*

Proposition 4.3 *Assume for all $(a, b) \in \{0, 1\}^2$ that either (i) $\mathbb{P}(\gamma_a(X_i) - \gamma_b(X_j) = 0) = 0$ or that (ii) $x_i \neq x_j \implies \gamma_a(X_i) - \gamma_b(X_j) \neq 0$ and let $\delta_{ij}^{ab} = sgn(\gamma_a(X_i) - \gamma_b(X_j))$, then*

$$\begin{aligned}\Gamma_{ij}^{ab} = \frac{1}{2}\bigg(&\gamma_a(X_i) + \gamma_b(X_j) - |\gamma_a(X_i) - \gamma_b(X_j)| \\ &+ \frac{\mathbb{1}(D_i = a)}{e_a(X_i)}(1 - \delta_{ij}^{ab})(Y_i - \gamma(D_i, X_i)) + \frac{\mathbb{1}(D_j = b)}{e_b(X_j)}(1 + \delta_{ij}^{ab})(Y_j - \gamma(D_j, X_j))\bigg).\end{aligned}$$

These assumptions deal with the point of non-differentiability of the absolute value. They hold if $\gamma_c(X_i)$ for $c \in \{a, b\}$ are continuous random variables (e.g. γ_c is strictly monotonic on a continuous random variable). When all circumstances are discrete (ii) can be a credible assumption. For a thorough discussion see Escanciano and Terschuur (2023). ■

To estimate the welfare in these examples for a given $\pi \in \Pi$ I use an adaptation to U-statistics of the cross-fitting used before (see Escanciano and Terschuur (2023)). I split the pairs $\{(i,j) \in \{1,...,n\}^2 : i < j\}$ in L groups $I_1, ..., I_l$, then

$$\hat{W}_n(\pi) = \binom{n}{2}^{-1} \sum_{l=1}^{L} \sum_{(i,j)\in I_l} \hat{\Gamma}_{ij,l}(\pi), \tag{4.2}$$

where $\hat{\Gamma}_{ij,l}$ is the same as Γ_{ij} but with all nuisance parameters replaced by estimators which do not use observations in the pairs in I_l. As before, the estimator of the optimal treatment rule among a class of rules Π is

$$\hat{\pi} = \arg\max_{\pi\in\Pi} \hat{W}_n(\pi).$$

For the Intergenerational mobility example, the estimation is slightly different.

Example 4 (Intergenerational mobility (cont.)) *The orthogonal score is given by*

$$\begin{aligned}\Gamma_{ij}^{ab} &= \mathbb{E}(sgn(X_{1i} - X_{1j})sgn(Y_i - Y_j) \mid D_i = a, X_i, D_j = b, X_j) \\ &+ \frac{D_{ij}^{ab}}{e_{ab}(X_i, X_j)}(sgn(X_{1i} - X_{1j})sgn(Y_i - Y_j) - \mathbb{E}(sgn(X_{1i} - X_{1j})sgn(Y_i - Y_j) \mid D_i = a, X_i, D_j = b, X_j)).\end{aligned}$$

The estimator of the welfare for a given $\pi \in \Pi$ and target t is

$$\hat{W}_n(\pi) = -\left|\binom{n}{2}^{-1} \sum_{l=1}^{L} \sum_{(i,j)\in I_l} \sum_{(a,b)\in\{0,1\}^2} \hat{\Gamma}_{ij,l}^{ab} \pi_{ab}(X_i, X_j) - t\right|. \tag{4.3}$$

■

5 Asymptotic statistical guarantees

Now it is useful to make clear the dependence of the scores Γ_{ij}^{ab} on the data and the nuisance parameters. Hence, I let now $\Gamma_{ij}^{ab} = \psi_{ab}(Z_i, Z_j, \gamma, \varphi, \alpha)$, where

$$\psi_{ab}(Z_i, Z_j, \gamma, \nu, \alpha) = \varphi(a, X_i, b, X_j, \gamma_a, \gamma_b) + \phi_{ab}^{\gamma}(Z_i, Z_j, \gamma, \alpha^{\gamma}) + \phi_{ab}^{\varphi}(Z_i, Z_j, \varphi, \alpha^{\nu}).$$

ψ_{ab} is the sum of an identifying function (m_{ab}) plus other functions (($\phi_{ab}^{\gamma}, \phi_{ab}^{\nu}$) which are correction terms needed to achieve orthogonality to the nuisance parameters γ and φ. In general, we have that for a given treatment rule π, orthogonal scores are given by

$$\Gamma_{ij}(\pi) = \sum_{(a,b)\in\{0,1\}^2} \psi_{ab}(Z_i, Z_j, \gamma, \varphi, \alpha)\pi_{ab}(X_i, X_j).$$

This framework accommodates also the welfare functions which are not defined as U-statistics if $\psi_{ab}(Z_i, Z_j, \gamma, \varphi, \alpha)$ does not depend on Z_j and only depends on a so that we could rewrite it as $\psi_a(Z_i, \gamma, \varphi, \alpha)$ for $a \in \{0, 1\}$. For this reason, I stick to this notation and do not state all conditions and results for welfare functions that are not U-statistics and those that are. The intergenerational mobility example does not fit in this general setting, however, the results extend easily to this example by Corollary 1 at the end of this section. In the next subsections I give conditions on the convergence of the nuisance parameters and on the complexity of the policy class Π which will allow me to prove asymptotical statistical guarantees for the estimated treatment rules.

5.1 Conditions on the nuisance parameter estimators

I give high-level conditions for the estimators of all nuisance parameters that have to be used to estimate the welfare. These conditions have been shown to hold for a variety of non-parametric estimators such as kernels or sieve estimators. The assumptions below are analogous to those in Escanciano and Terschuur (2023).

Assumption 4 $\mathbb{E}[|\psi(Z_i, Z_j, \gamma, \varphi, \alpha)|^2] < \infty$, $\omega \in \{\gamma, \varphi\}$ *and for* $(a, b) \in \{0, 1\}^2$

(i) $n^{\lambda_\gamma}\sqrt{\mathbb{E}(|\varphi(a, x_i, b, x_j, \hat{\gamma}_l) - \varphi(a, x_i, b, x_j, \gamma)|^2)} = o(1)$;

(ii) $n^{\lambda_\varphi}\sqrt{\mathbb{E}(|\hat{\varphi}_l(a, x_i, b, x_j, \gamma) - \varphi(a, x_i, b, x_j, \gamma)|^2)} = o(1)$;

(iii) $n^{\lambda_\gamma}\sqrt{\mathbb{E}(|\phi^{\gamma}_{ab}(z_i, z_j, \hat{\gamma}_l, \alpha^\gamma) - \phi^{\gamma}_{ab}(z_i, z_j, \gamma, \alpha^\gamma)|^2)} = o(1)$;

(iv) $n^{\lambda_\varphi}\sqrt{\mathbb{E}(|\phi^{\varphi}_{ab}(z_i, z_j, \hat{\varphi}_l, \alpha^\varphi) - \phi^{\varphi}_{ab}(z_i, z_j, \varphi, \alpha^\varphi)|^2)} = o(1)$;

(v) $n^{\lambda_\alpha}\sqrt{\mathbb{E}(|\phi^{\omega}_{ab}(z_i, z_j, \omega, \hat{\alpha}^\omega_l) - \phi^{\omega}_{ab}(z_i, z_j, \omega, \alpha^\omega)|^2)} = o(1)$,

where $1/4 < \lambda_\gamma, \lambda_\varphi, \lambda_\alpha$.

These are mild mean-square consistency conditions for $\hat{\gamma}_l$, $\hat{\varphi}_l$ and $\hat{\alpha}_l$ separately. Assumption 4 often follows from the L2 convergence rates of the nuisance estimators. There is a large literature checking L_2-convergence rates for different machine learners under low-level sparsity or smoothness conditions on the nuisance parameters. The traditional non-parametric literature gives rates for kernel regression and sieves/series (e.g. Chen (2007)). For L_1-penalty estimators such as Lasso see, e.g., Belloni and Chernozhukov (2011) and Belloni and Chernozhukov (2013). Also for low-level conditions for shrinkage and kernel estimators see Appendix B in Sasaki and Ura (2021). Rates for L_2-boosting in low dimensions are found in Zhang and Yu (2005), and more recently Kueck et al. (2023) find rates for L_2-boosting with high dimensional data. For results on versions of random forests see Wager and Walther (2015) and Athey et al. (2019). Finally, for single-layer, sigmoid-based neural networks see Chen and White (1999) and for a modern setting of deep neural networks with rectified linear (ReLU) activation function see Farrell et al. (2021). Note that $\hat{\varphi}_l$ estimates conditional expectations where both the dependent variables and the conditioning ones are indexed by both i and j. Stute (1991) calls such objects conditional U-statistics and studies the asymptotic properties of Nadaraya-Watson nonparametric estimators of these quantities. Graham et al. (2021) study the nonparametric estimation of such nuisance parameters when the dependent variable is dyadic. They provide asymptotic and supremum norm rate results and also propose to use a Nadaraya-Watson estimator. In this paper, I run the machine learning algorithms on the stacked pairs. Unfortunately, not much is known about rates for such dyadic machine learning regressions which are also very computationally demanding. Another option that avoids dyadic regression is to use the IPW approach in Appendix A. Define now the following interaction terms for $\omega \in \{\gamma, \varphi\}$ and let $||\cdot||$ denote the L2 norm.

$$
\begin{aligned}
\hat{\xi}_{ij,ab,l} &= \hat{\varphi}_l(a, X_i, b, X_j, \hat{\gamma}_l) - \varphi(a, X_i, b, X_j, \hat{\gamma}_l) - \hat{\varphi}_l(a, X_i, b, X_j, \gamma) + \varphi(a, X_i, b, X_j, \gamma), \\
\hat{\xi}^{\omega}_{ij,ab,l} &= \phi_{ab}(z_i, z_j, \hat{\omega}_l, \hat{\alpha}^{\omega}_l) - \phi_{ab}(z_i, z_j, \omega, \hat{\alpha}^{\omega}_l) - \phi_{ab}(z_i, z_j, \hat{\omega}_l, \alpha^{\omega}) + \phi_{ab}(z_i, z_j, \omega, \alpha^{\omega}).
\end{aligned}
$$

Assumption 5 *For each $l = 1, ..., L$*

(i) $\int\int \phi^{\gamma}_{ab}(z_i, z_j, \gamma, \hat{\alpha}^{\gamma}_l) F(dz_i) F(dz_j) = 0$ *and* $\int\int \phi^{\varphi}_{ab}(Z_i, Z_j, \varphi, \hat{\alpha}^{\varphi}_l) F(dz_i) F(dz_j) = 0$.

(ii) $\mathbb{E}(||\hat{\gamma}_l - \gamma||^2) = o(n^{-2\lambda_\gamma})$, $\mathbb{E}(||\hat{\varphi}_l - \varphi||^2) = o(n^{-2\lambda_\varphi})$ *and*

$$
\begin{aligned}
|\mathbb{E}[(\varphi(a, X_i, b, X_j, \tilde{\gamma}) + \phi^{\gamma}_{ab}(Z_i, Z_j, \tilde{\gamma}, \alpha^{\gamma}))\pi_{ab}(X_i, X_j)]| &\leq C||\tilde{\gamma} - \gamma||^2 \\
|\mathbb{E}[(\tilde{\varphi}(a, X_i, b, X_j, \gamma) + \phi^{\varphi}_{ab}(Z_i, Z_j, \tilde{\varphi}, \alpha^{\varphi}))\pi_{ab}(X_i, X_j)]| &\leq C||\tilde{\varphi} - \varphi||^2.
\end{aligned}
$$

Assumption 5 (i) is usually easy to verify from visual inspection and (ii) requires L2 convergence rates and some smoothness. Note that C is uniform over $\pi \in \Pi$.

Assumption 6 *For each $l = 1, ..., L$*

$$
\sqrt{n}\mathbb{E}(\hat{\xi}^{\omega}_{ij,ab,l}\pi_{ab}(X_i, X_j)) = o(1).
$$

These are rate conditions on the remainder terms $\hat{\xi}^{\omega}_l(w_i, w_j)$. Often, Assumption 6 follows if $\sqrt{n}||\hat{\alpha}^{\omega}_l - \alpha||||\hat{\omega}_l - \omega|| = o(1)$. For example Athey and Wager (2021), in the proof of their Lemma 4, use Cauchy-Schwarz inequality to get a bound on their interaction term which does not depend on π and only on the product of L2 norms. This is the precise assumption that allows for parametric rates even with slow nonparametric estimators. In essence, it is enough for the product of the nonparametric estimators to go to zero at a parametric rate.

5.2 Conditions on the complexity of the policy class

The complexity of the policy class must also be restricted. If all sorts of subsets of $\mathcal{X}$ are allowed to decide who should be treated then we get overfitted policy rules. As in Athey and Wager (2021) I measure the policy class complexity with its VC dimension (see for instance Wainwright (2019)) which I allow to grow with the sample size. Hence, from now on I subscript the policy class by n, Π_n.

Assumption 7 *There are constants $0 < \beta < 1/2$ and $n^* \geq 1$ such that for all $n \geq n^*$, $VC(\Pi_n) < n^\beta$.*

Examples of finite VC-dimension classes are linear eligibility scores or generalized eligibility scores (see Kitagawa and Tetenov (2018)). Policy classes whose VC-dimension can increase with the sample size are for example decision trees which get deeper with sample size (see Athey and Wager (2021)).

5.3 Upper bounds

Let now

$$
\begin{aligned}
W(\pi) &= \mathbb{E}\Bigg[\sum_{(a,b)\in\{0,1\}^2} \psi_{ab}(Z_i, Z_j, \gamma, \varphi, \alpha)\pi_{ab}(X_i, X_j)\Bigg],\\
\widetilde{W}_n(\pi) &= \binom{n}{2}^{-1} \sum_{i<j}\Bigg[\sum_{(a,b)\in\{0,1\}^2} \psi_{ab}(Z_i, Z_j, \gamma, \varphi, \alpha)\pi_{ab}(X_i, X_j)\Bigg],\\
\hat{W}_n(\pi) &= \binom{n}{2}^{-1} \sum_{l=1}^{L}\sum_{(i,j)\in I_l}\Bigg[\sum_{(a,b)\in\pi} \psi_{ab}(Z_i, Z_j, \hat{\gamma}_l, \hat{\varphi}_l, \hat{\alpha}_l)\pi_{ab}(X_i, X_j).\Bigg],
\end{aligned}
$$

$W(\pi)$ and $\widetilde{W}_n(\pi)$ are the welfare and the infeasible estimator of the welfare at policy rule π when all nuisance parameters are known. $\hat{W}_n(\pi)$ is the feasible estimator which we already introduced in (4.2). Let $W^*_{\Pi_n} = \sup_{\pi\in\Pi_n} W(\pi)$ be the best possible welfare. I want to give upper bounds to the regret: $\mathbb{E}[W^*_{\Pi_n} - W(\hat{\pi})]$, i.e. the expected difference between the best possible welfare and the welfare evaluated at the estimated policy. As usual, I start bounding the regret as follows

$$
\begin{aligned}
\mathbb{E}[W^*_{\Pi_n} - W(\hat{\pi})] &\leq 2\mathbb{E}\Bigg[\sup_{\pi\in\Pi_n} |\hat{W}_n(\pi) - W(\pi)|\Bigg]\\
&\leq 2\mathbb{E}\Bigg[\sup_{\pi\in\Pi_n} |\hat{W}_n(\pi) - \widetilde{W}_n(\pi)|\Bigg] + 2\mathbb{E}\Bigg[\sup_{\pi\in\Pi_n} |\widetilde{W}_n(\pi) - W(\pi)|\Bigg], \qquad (5.1)
\end{aligned}
$$

where in the second inequality I have added and subtracted $\widetilde{W}_n(\pi)$ and used the triangle inequality. The second term above is just a standard centered U-process indexed by $\pi \in \Pi_n$. I start as in Athey and Wager (2021) by showing the rate of convergence of this second term. I work for some fixed $(a,b) \in \{0,1\}^2$ and define the following set

$$\Pi_{ab,n} = \{\pi_{ab} : \pi \in \Pi_n\}.$$

The first step is to bound it by the Rademacher complexity which I define as

$$\mathcal{R}_n(\Pi_{ab,n}) = \mathbb{E}_\varepsilon\left(\sup_{\pi\in\Pi_n}\left|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})\right|\right).$$

Lemma 1

$$\mathbb{E}\left[\sup_{\pi\in\Pi_n}|\widetilde{W}_n(\pi) - W(\pi)|\right] \leq \mathbb{E}[2\mathcal{R}_n(\Pi_{ab,n})].$$

Now we want an asymptotic upper bound for $\mathbb{E}[\mathcal{R}_n(\Pi_{ab})]$. Importantly, we want the bound to depend on the following variance

$$S_{ab} = \mathbb{E}[\Gamma^{2\,ab}_{i,j}].$$

While Kitagawa and Tetenov (2018) and others provide bounds in terms of the max of the scores, Athey and Wager (2021) provide bounds based on the variance and the efficient variance. The next result provides a bound on the Rademacher complexity based on S_{ab}.

Lemma 2 *Assume that Γ^{ab}_{ij} has bounded support for $(a,b) \in \{0,1\}^2$. Then, under Assumptions 4 and 6*

$$\mathbb{E}[\mathcal{R}_n(\Pi_{ab,n})] = \mathcal{O}\left(\sqrt{\frac{S_{ab}\cdot VC(\Pi_{ab,n})}{\lfloor n/2\rfloor}}\right).$$

The boundedness assumption can be generalized to sub-gaussianity. However, this generalization comes at the cost of making the (already involved) proofs substantially less tractable. Now we want to provide asymptotic upper bounds for the first term in (5.1). Escanciano and Terschuur (2023) show that for given $\pi \in \Pi_n$

$$\sqrt{n}(\hat{W}_n(\pi) - \widetilde{W}_n(\pi)) \to_p 0.$$

The next result makes the above uniform in $\pi \in \Pi_n$.

Lemma 3 (Uniform coupling) *Under Assumptions 4 and 6*

$$\sqrt{n}\mathbb{E}[\sup_{\pi\in\Pi_n}|\hat{W}_n(\pi) - \widetilde{W}_n(\pi)|] = \mathcal{O}\left(1 + \frac{VC(\Pi_{ab,n})}{\lfloor n/2\rfloor^{\min(\lambda_\gamma,\lambda_\nu,\lambda_\alpha)}}\right).$$

Finally, using Lemmas 2 and 3 the following holds.

Theorem 1 *Suppose Assumptions 4 and 6 hold, that Assumption 7 holds with $\beta < \min(\lambda_\gamma, \lambda_\nu, \lambda_\alpha)$. Then*

$$\mathbb{E}[W^*_{\Pi_n} - W(\hat{\pi})] = \mathcal{O}\left(\sqrt{\frac{S_{ab}\cdot(2VC(\Pi_n) - 1)}{\lfloor n/2\rfloor}}\right).$$

Corollary 1 *The bound in Theorem 1 applies to the Intergenerational mobility example.*

6 Empirical application

In the empirical application, I study the optimal allocation of children to preschool. I make use of the Panel Study of Income Dynamics (PSID) database which has been following families for nearly 50 years. The nature of this survey allows us to observe a rich set of circumstances and long-term outcomes. In 1995, PSID asked adults between 18-30 years old about their participation in preschool. Hence, we can track the long-term outcomes of these individuals. I take as an outcome the average earnings from 25 to 35 years old. I assume selection on observables holds. In particular, I condition on sex, birthyear, average parental income in the 5 years before birth, mother's education, father's education, father's occupation and whether the individual is black. In Table 1 we see the results of estimating the Average Treatment Effect (ATE), Gini, IOp and Intergenerational mobility as captured by the rank correlation of parents and child income.

Outcome	ATE	se	p-value	Gini	IOp	IGM	n
Earnings 25-35	4622.063	1083.865	0	0.392	0.172	0.168	2971

Table 1: ATE, Gini, IOp and Kendal-τ

To estimate the ATE, I use the doubly robust Augmented Inverse Propensity weighted scores from Robins et al. (1994) using Conditional Inference Forests to estimate the regression functions and propensity scores. I chose CIF by cross-validation among a pool of different machine learners. Under the assumption of no selection on observables, we observe a sizeable and significant positive effect of attending preschool of 4,622$ of added annual earnings. Dollars have been adjusted by the CPI to 2010 dollars. We see that the Gini coefficient is 0.39 and that IOp is 0.17, meaning that almost 44% of total inequality can be explained by the circumstances we observe. The Kendall-τ is around 0.17 which indicates a positive association between parental and child incomes.

I compute optimal treatment rules based on parental income and the mother's years of education. I set the target in the Kendall-τ welfare to zero, meaning that the aim is to completely erase intergenerational persistence. As the policy class, I use 2-depth decision trees. Unfortunately, the U-statistic nature of the welfare function prevents me from using the computational shortcuts in Athey and Wager (2021) since the sub-trees are not independent optimization problems. To ease the computational problem I use the deciles of parental income as cutting points instead of all the observed values of parental income. I do an exhaustive search meaning that I consider all possible 2-depth decision trees.

The first result is that the optimal treatment allocation is the same for the additive and the IOp welfare. Although this might seem surprising, it is perfectly possible if decreasing inequality of opportunity is not compensated by increases in the average. In fact, as reported in Table 2, the estimated rule maximizing the average already drastically decreases IOp. I show the optimal rule under these two welfares in Figure 1. At the terminal nodes, I report the number of observations, the conditional average treatment effect (CATE) in the node and the proportion of observations in the terminal node that are treated in the data ($\hat{p}$). For additive/IOp welfare, we see that the first cutting point is whether parental income is below or above the 40th percentile (51,515$). If an observation is below this cut-off the tree splits according to the education of the mother. If parental income is below the 40th percentile and the mother's education is less than college (below 13 years) the tree allocates the observation to treatment. We see that the CATE in this node is positive so, as we would expect, an additive policy maker treats these observations. If parental income is below the 40th percentile but the mother is highly educated we see that the CATE is actually negative and hence the additive policy maker does not allocate the individual to treatment. For high parental income, we also split on the mother's education but at a higher level of education. If your parental income is higher than the 40th percentile and your mother attended college or less (16 years of education or less) you are allocated to treatment and in this node, we have very large positive effects. However, we do not allocate kids with high parental income and high maternal education to preschool since the CATE in this group is negative.

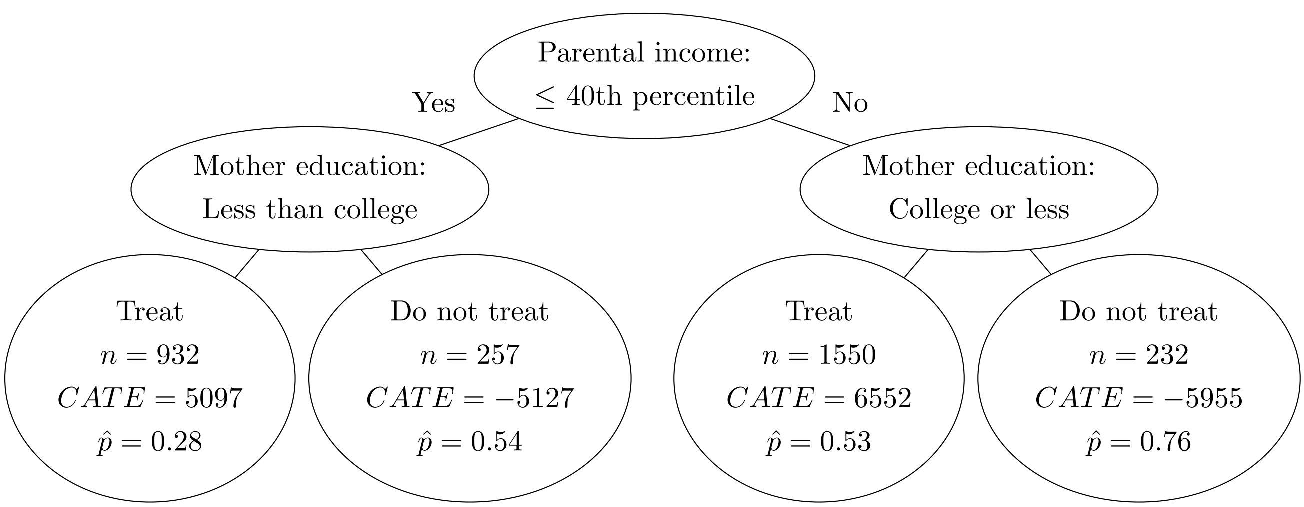

Figure 1: Estimated optimal policy rule under additive and IOp welfare.

We can conjecture that there is a substitution effect between the quality of the preschool institution and mother's education. If we take parental income to be a proxy for the quality of preschool, we see that it is enough for the mother to have more than 13 years of education for the child to be better off without preschool. However, for children who attend better preschools (have higher parental income), the mother has to have more than 16 years of education for the child to be better off without preschool. This is in line with results in the psychology and economics literature which document that in early educational institutions, there is less interaction with adults and hence there can be a negative effect of attending such institutions if the interactions with adults in the household are of "high-quality" (see Fort et al. (2020)). To decrease IOp further, we would need to treat advantaged kids who do not benefit from preschool. The fact that the optimal policy rule is the same for the additive and the IOp welfare indicates that the penalization of inequality of opportunity is not severe enough to treat advantaged kids who do not benefit from preschool or to not treat kids who do not benefit from preschool. Interestingly, the treatment allocation observed in the data is quite far from the optimal rule. This suggests that future earnings are not the only consideration when parents decide whether to send their kids to preschool.

In figure 2, we see the optimal policy rule for the inequality welfare function, i.e. now we penalize all inequalities and not just the ones explained by circumstances. We see that the tree is the same except for the first cutting point on parental income. Now we first divide individuals into those with parental income lower and higher than the 20th percentile (37,699$). Then, the division based on the mother's education is the same. Hence, compared to the previous tree, we shift 20% of the population to the right side subtree. For instance, a kid who has a parental income of 40,000$ and whose mother has 16 years of education would not be treated under the additive/IOp welfare but is treated under the inequality based optimal rule. Although masked by other observations in the node, this 20% of the population who is switched to treatment has an estimated negative CATE. When we penalize all sorts of inequalities, it starts becoming optimal to decrease the average outcome to decrease inequality.

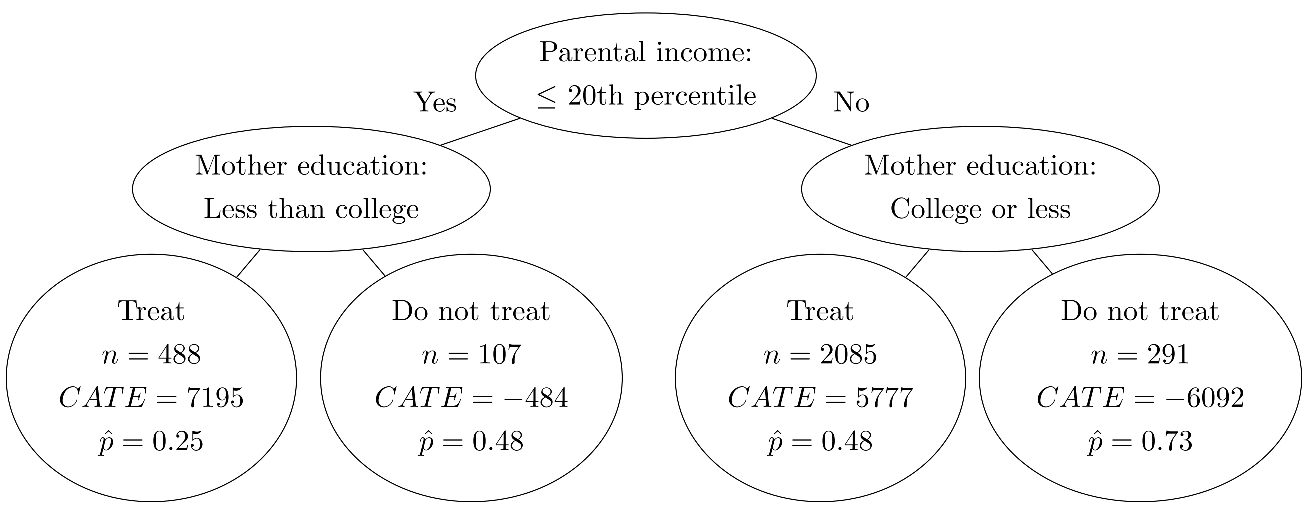

Figure 2: Estimated optimal policy rule under inequality welfare.

In Figure 3 we see the results for an intergenerational mobility aware welfare function. Notice that in the intergenerational mobility welfare there is no efficiency motive and we target a zero Kendall-τ. Hence, the optimal policy is even more controversial since individuals with positive treatment effects are not treated and individuals with negative treatment effects are treated. In this case, if the mother has more than 13 years of education but less than 16 years of education (college education), the optimal policy does not treat even though there are positive treatment effects. Even more extreme, if the mother is very highly educated, the optimal policy treats even though there are negative treatment effects. For kids with maternal education below college, the optimal policy treats only if you are in the lowest 50% of the distribution of parental income (below 58,270$). If your mother has less than college but your parents are in the richest half of the income distribution you are not treated even though there are positive treatment effects.

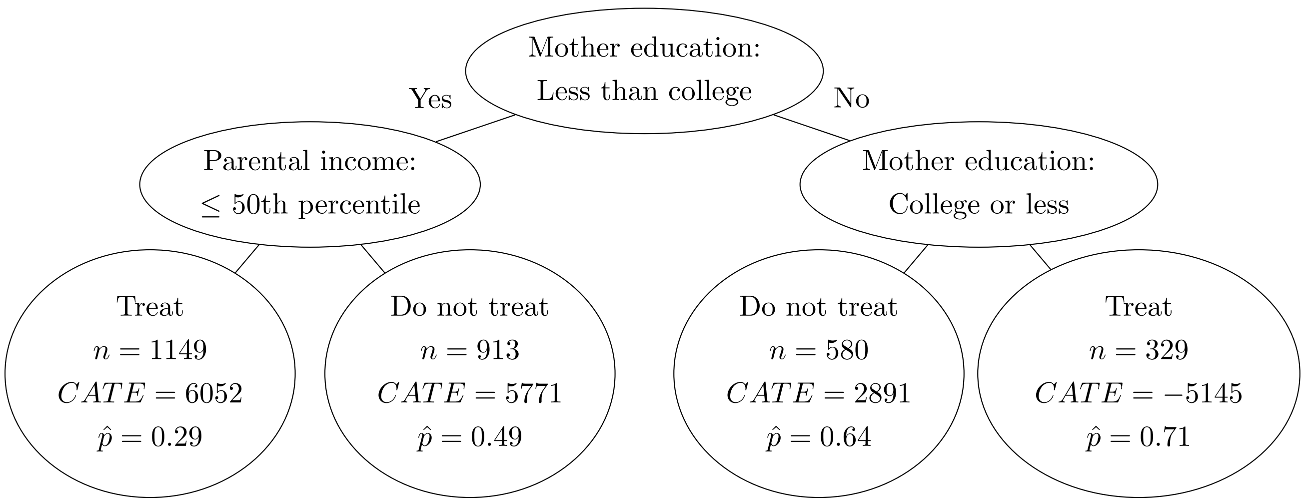

Figure 3: Estimated optimal policy rule under intergenerational mobility welfare.

Finally, in Table 2 we see a summary of the results and compare the estimated optimal treatments with situations in which no one or everyone is treated. If we focus first on the welfare column, we see that for additive, IOp and inequality welfares, treating no one gives the worst welfare. In the IGM case, we see that treating no one and treating everyone give basically the same welfare (note that maximal welfare in the IGM case is 0). Since the optimal policy rule is the same for the additive and the IOp welfare, I leave blank the cells for the mean, Gini, IOp, IGM and share treated for IOp since the values are the same as for the additive welfare. As expected, the additive and IOp welfares have an estimated optimal policy rule that attains the highest average outcome and the lowest IOp. The decrease in IOp under this rule is quite drastic. While in the sample we can explain 44% of total inequality with circumstances (IOp/Gini), at the optimal additive/IOp rule we explain 35%. In part, this explains why both rule coincide. In other settings, maximizing the average might go against IOp which does not happen here.

The estimated optimal treatment rule for IGM gives basically the same IOp as the the estimated optimal policy rule based on the IOp welfare, but it does so at a much larger cost in terms of the average outcome. Again, as expected, the estimated optimal policy rule under the inequality welfare gives the lowest Gini compared to the additive and IOp welfares. It is interesting to note that it gives the highest IOp across all welfares. The IGM welfare estimated rule gives virtually the same Gini as the inequality welfare but again, at a much higher cost in terms of the average outcome. As expected, the IGM rule gives the lowest Kendall-τ.

		Welfare	Mean	Gini	IOp	IGM	Share treated
Additive	**Optimal rule**	39727	39727	0.392	0.138	0.15	0.84
	Treat no one	34169	34169	0.4	0.162	0.148	0
	Treat everyone	38778	38778	0.383	0.142	0.15	1
IOp	**Optimal rule**	34231	·	·	·	·	·
	Treat no one	28640	·	·	·	·	·
	Treat everyone	33282	·	·	·	·	·
Inequality	**Optimal rule**	24165	39383	0.386	0.141	0.153	0.87
	Treat no one	20518	34169	0.4	0.162	0.148	0
	Treat everyone	23942	38778	0.383	0.142	0.15	1
IGM	**Optimal rule**	-0.086	35951	0.383	0.139	0.086	0.5
	Treat no one	-0.148	34169	0.4	0.162	0.148	0
	Treat everyone	-0.15	38778	0.383	0.142	0.15	1
Sample	·	·	36197	0.392	0.172	0.168	0.47

Table 2: Welfare, mean, Gini, IOp, IGM and share treated for different optimal policy rules compared with policies which treat no one and everyone. I also show the actual values observed in the sample. The dots in the IOp rows indicate that the optimal policy rule is the same as in the additive case.

Finally, comparing the results with what we observe with the treatment allocation in the sample, we see that we achieve a higher mean with all other welfares except with the IGM welfare. The Gini in the sample is the same as the one under the estimated optimal additive and IOp rule. The observed IOp in the sample is higher than the one achieved under the estimated rules of all other welfares. IGM observed in the sample is the lowest (highest Kendall-τ) compared to all welfares. Finally, the share of treated in the sample is also lower than the one achieved under the estimated optimal rule of all other welfares.

7 Conclusion

This paper extends previous work on policy learning to accommodate general semiparametric welfare functions estimable by U-statistics. This opens the analysis to highly policy-relevant welfare functions such as inequality, inequality of opportunity and intergenerational mobility aware SWFs. The inequality of opportunity SWF is especially useful when we do not want to penalize all sorts of inequality but just unfair sources of inequality. In the empirical application, inequality and IGM aware SWFs assign groups with negative treatment effects to treatment. However, the additive and IOp rules coincide. Further work is needed, particularly to ease the computational burden of the method. The application of convex surrogates in Kitagawa et al. (2021) is a promising avenue to achieve this. Another interesting extension is to allow for multiple treatments as in Zhou et al. (2023) in a U-statistics setting. In our application, this could be useful to study the optimal allocation of children to different types of preschools. Finally, it would be interesting to extend the results to the case of continuous treatments as in Athey and Wager (2021).

8 Appendix

A Inverse Propensity Weighting (IPW) results

In this section, I show the identification and local robustness results in general for both (DM) and (IPW). The next proposition states the first identification result.

Proposition 8.1 *Under Assumption 1, $W(\pi)$ is identified as*

$$W(\pi) = \mathbb{E}[m_1(Z_i, \gamma, \nu)\pi(X_i) + m_0(Z_i, \gamma, \nu)(1 - \pi(X_i))],$$

with $\nu \in \{\varphi, e\}$ and where m_1 and m_0 can be any of the following

$$\begin{aligned} &\textit{(DM)} \quad m_1(Z_i, \gamma, \varphi) = \varphi(1, X_i, \gamma_1), \quad m_0(Z_i, \gamma, \varphi) = \varphi(0, X_i, \gamma_0) \\ &\textit{(IPW)} \quad m_1(Z_i, \gamma, e) = \frac{g(Y_i, X_i, \gamma_1)D_i}{e(X_i)}, \quad m_0(Z_i, \gamma, e) = \frac{g(Y_i, X_i, \gamma_0)(1 - D_i)}{1 - e(X_i)}. \end{aligned}$$

Now I show the identification result for U-statistics estimable quantities.

Proposition 8.2 *Under Assumption 1, $W(\pi)$ in (4.1) is identified in the following ways*

$$W(\pi) = \mathbb{E}\Big[\sum_{(a,b)\in\{0,1\}^2} m_{ab}(Z_i, Z_j, \gamma, \nu)\pi_{ab}(X_i, X_j)\Big],$$

where $\nu \in \{\varphi, e\}$ and m_{ab} can be any of the following

$$\begin{aligned} &\textit{(DM)} \quad m_{ab}(Z_i, Z_j, \gamma, \varphi) = \varphi(a, X_i, b, X_j, \gamma_a, \gamma_b) \\ &\textit{(IPW)} \quad m_{ab}(Z_i, Z_j, \gamma, e) = g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D_{ij}^{ab}/e_{ab}(X_i, X_j). \end{aligned}$$

Next, I introduce the Assumption necessary for computing locally robust scores and the results of local robustness with U-statistics estimable quantities.

Assumption 8 *There exist α_{ab}, $P < \infty$, and (c_{1p}, c_{2p}) for $p = 1, ..., P$, such that for all $(a, b) \in \{0, 1\}^2$ the following linearization holds*

$$\frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i, Z_j, \bar{\gamma}_\tau, \nu)] = \mathbb{E}\Big[\sum_{p=1}^{P} \alpha_{ab,p}^{\gamma}(D_i, X_i, D_j, X_j)(c_{1p}\bar{\gamma}_\tau(D_i, X_i) + c_{2p}\bar{\gamma}_\tau(D_j, X_j))\Big],$$

where $\bar{\gamma}_\tau$ is defined as in Assumption 2.

Proposition 8.3 *The orthogonal scores are given by*

$$\Gamma_{ij}(\pi) = \sum_{(a,b)\in\{0,1\}^2} \Gamma_{ij}^{ab}\pi_{ab}(X_i, X_j),$$

where depending on whether we identify with DM or IPW we have

$$(DM) \quad \Gamma_{ij}^{ab} = \varphi(a, X_i, b, X_j, \gamma_a, \gamma_b) + \phi_{ab}^{\varphi}(D_i, X_i, D_j, X_j, \varphi, \alpha^e) + \phi_{ab}^{\gamma}(D_i, X_i, D_j, X_j, \gamma, \alpha^{\gamma})$$

$$(IPW) \quad \Gamma_{ij}^{ab} = \frac{g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b) D_{ij}^{ab}}{e_{ab}(X_i, X_j)} + \phi_{ab}^{e}(D_i, X_i, D_j, X_j, e, \alpha^e) + \phi_{ab}^{\gamma}(D_i, X_i, D_j, X_j, \gamma, \alpha^{\gamma}),$$

where

$$\phi_{ab}^{\gamma}(D_i, X_i, D_j, X_j, e, \alpha^{\gamma}) = \sum_{p=1}^{P} \alpha_{ab,p}^{\gamma}(D_i, X_i, D_j, X_j, e)(c_{1p} Y_i + c_{2p} Y_j - c_{1p}\gamma(D_i, X_i) - c_{2p}\gamma(D_j, X_j)),$$

$$\phi_{ab}^{\varphi}(D_i, X_i, D_j, X_j, \varphi, \alpha^m) = \alpha_{ab}^{\varphi}(D_i, X_i, D_j, X_j)(g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b) - \varphi(D_i, X_i, D_j, X_j, \gamma_a, \gamma_b)),$$

$$\phi_{ab}^{e}(D_i, X_i, D_j, X_j, e, \alpha^e) = \alpha_{ab,1}^{e}(X_i)(\mathbb{1}(D_i = a) - e_a(X_i)) + \alpha_{ab,2}^{e}(X_j)(\mathbb{1}(D_j = b) - e_b(X_j)),$$

and

$$\alpha_{ab}^{\varphi}(D_i, X_i, D_j, X_j) = \frac{D_{ij}^{ab}}{e_{ab}(X_i, X_j)},$$

$$\alpha_{ab,1}^{e}(X_i) = -\mathbb{E}\left[\frac{g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b) D_{ij}^{ab}}{e_a(X_i)^2 e_b(X_j)} \middle| X_i\right],$$

$$\alpha_{ab,2}^{e}(X_j) = -\mathbb{E}\left[\frac{g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b) D_{ij}^{ab}}{e_a(X_i) e_b(X_j)^2} \middle| X_j\right].$$

B Auxiliary lemmas

In this section, we prove some lemmas which will be needed to prove the main results. Let us first define some important objects. For a fixed sample $\{X_i, \Gamma_{i,\lfloor n/2 \rfloor + i}\}_{i=1}^{\lfloor n/2 \rfloor}$ we have

$$\widetilde{\Pi}_{ab} = \{\pi_{ab}(X_1, X_{\lfloor n/2 \rfloor + 1}), ..., \pi_{ab}(X_{\lfloor n/2 \rfloor}, X_n) : \pi \in \Pi\}.$$

For $\pi, \pi' \in \widetilde{\Pi}_{ab}$ define the following distances

$$D_n^2(\pi, \pi') = \frac{\sum_{i=1}^{\lfloor n/2 \rfloor} \Gamma_{i,\lfloor n/2 \rfloor + i}^{2\,ab} (\pi_{ab}(X_i, X_{\lfloor n/2 \rfloor + i}) - \pi'_{ab}(X_i, X_{\lfloor n/2 \rfloor + i}))^2}{\sum_{i=1}^{\lfloor n/2 \rfloor} \Gamma_{i,\lfloor n/2 \rfloor + i}^{2\,ab}},$$

$$H(\pi, \pi') = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\pi_{ab}(X_i, X_{\lfloor n/2 \rfloor + i}) \neq \pi'_{ab}(X_i, X_{\lfloor n/2 \rfloor + i})).$$

Let $N_{D_n}(\varepsilon, \widetilde{\Pi}_{ab}, \{X_i, \Gamma_{i,\lfloor n/2 \rfloor + i}\}_{i=1}^{\lfloor n/2 \rfloor})$ be the number of balls of radius ε needed to cover $\widetilde{\Pi}_{ab}$ under distance D_n. Define the same object for the Hamming distance H and let

$$N_H(\varepsilon, \widetilde{\Pi}_{ab}) = \sup\{N_H(\varepsilon, \widetilde{\Pi}_{ab}, \{X_i\}_{i=1}^{m}) : X_1, ..., X_m \in \mathcal{X}, m \geq 1\}.$$

Note $N_H(\varepsilon, \widetilde{\Pi}_{ab})$ does not depend on m. It will be useful to bound N_{D_n} with N_H which is what we do in the next lemma.

Lemma 4 *For fixed* $\{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}$ *we have that*

$$N_{D_n}(\varepsilon, \widetilde{\Pi}_{ab}, \{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}) \le N_H(\varepsilon^2, \widetilde{\Pi}_{ab}).$$

Proof: Take an auxiliary sample $\{X'_j\}_{j=1}^m$ contained in $\{X_i\}_{i=1}^n$ such that

$$\left| |B_i| - \frac{m\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}{\sum_{i=1}^{\lfloor n\rfloor/2} \Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}} \right| \le 1,$$

where $B_i = \{j \in \{1, ..., m\} : X'_j = X_i\}$. Then, for $\pi, \pi' \in \widetilde{\Pi}_{ab}$

$$D_n^2(\pi, \pi') = \frac{1}{m} \sum_{i=1}^{\lfloor n/2\rfloor} \underbrace{\frac{m\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}{\sum_{k=1}^{\lfloor n/2\rfloor} \Gamma^{2\,ab}_{k,\lfloor n/2\rfloor+k}}}_{\ge |B_i|-1} \mathbb{1}(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i}) \ne \pi'_{ab}(X_i, X_{\lfloor n/2\rfloor+i})).$$

So

$$\begin{aligned}
D_n^2(\pi, \pi') &\ge \sum_{i=1}^{\lfloor n/2\rfloor} \frac{|B_i|}{m} \mathbb{1}(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i}) \ne \pi'_{ab}(X_i, X_{\lfloor n/2\rfloor+i})) - O(1/m) \\
&= \sum_{i=1}^{\lfloor n/2\rfloor} \frac{|B_i|}{m} \frac{1}{|B_i|} \sum_{j\in B_i} \mathbb{1}(\pi_{ab}(X'_j, X'_{\lfloor n/2\rfloor+j}) \ne \pi'_{ab}(X'_j, X'_{\lfloor n/2\rfloor+j})) - O(1/m) \\
&= \frac{1}{m} \sum_{i=1}^{\lfloor n/2\rfloor} \sum_{j\in B_i} \mathbb{1}(\pi_{ab}(X'_j, X'_{\lfloor n/2\rfloor+j}) \ne \pi'_{ab}(X'_j, X'_{\lfloor n/2\rfloor+j})) - O(1/m).
\end{aligned}$$

In the first equality above we have used the fact that all summands in the inner sum are the same since for all $j \in B_i$ we know that $(X_i, X_{\lfloor n/2\rfloor+i}) = (X'_j, X'_{\lfloor n/2\rfloor+j})$. Now we notice that the sum $\sum_{i=1}^{\lfloor n/2\rfloor} \sum_{j\in B_i}$ might sum some pairs more than once (e.g. if $(X_1, X_{\lfloor n/2\rfloor+1}) = (X_2, X_{\lfloor n/2\rfloor+2})$ then $B_1 = B_2$). Using this fact and that $\{X'_j\}_{j=1}^m$ is contained in $\{X'_i\}_{i=1}^m$ we have that

$$\begin{aligned}
D_n^2(\pi, \pi') &\ge \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}(\pi_{ab}(X'_j, X'_{\lfloor n/2\rfloor+j}) \ne \pi'_{ab}(X'_j, X'_{\lfloor n/2\rfloor+j})) - O(1/m) \\
&= H(\pi, \pi') - O(1/m).
\end{aligned}$$

Hence, $H(\pi, \pi') \le D_n^2(\pi, \pi') + O(1/m)$. Since N_H does not depend on m, we can make m arbitrarily large and conclude that

$$N_{D_n}(\varepsilon, \widetilde{\Pi}_{ab}, \{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}) \le N_H(\varepsilon^2, \widetilde{\Pi}_{ab}).$$

■

Now we prove that the sequence of covers we use in the proof of Lemma 2 exists.

Lemma 5 *There exists a sequence of covers* $\{B_k\}_{k=0}^K$ *with* $K < \infty$ *of* $\widetilde{\Pi}_{ab}$ *with* $B_k \subset \widetilde{\Pi}_{ab}$ *such that for* $k = 0, ..., K$

- *For all* $\pi \in \widetilde{\Pi}_{ab}$*, there exists* $b \in B_k$ *such that* $D_n(\pi, b) \leq 2^{-k}$*,*
- $|B_k| = N_{D_n}(2^{-k}, \widetilde{\Pi}_{ab}, \{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}) \leq |\widetilde{\Pi}_{ab}|$.

Proof: First note that $|\widetilde{\Pi}_{ab}| < 2^{\lfloor n/2\rfloor} < \infty$ since X_i's are fixed. Since $\widetilde{\pi}_{ab}$ is finite and $B_k \subset \widetilde{\Pi}_{ab}$ for all k, there exists finite K such that we can set $B_K = \widetilde{\Pi}_{ab}$. This is because for any B_k which is a strict subset of $\widetilde{\Pi}_{ab}$ there exist $\pi \in \widetilde{\Pi}_{ab}$ such that for all $b \in B_k$, $D_n(b, \pi) > a > 0$ and there exists $K > 0$ such that $2^{-K} < a$. K is finite since there are only finitely many subsets of $\widetilde{\Pi}_{ab}$. For B_{K-1} we can look through all possible strict subsets for one which satisfies our conditions, if we do not find any we know that $B_{K-1} = \widetilde{\Pi}_{ab}$ does satisfy them. In this way, we can go backwards and build the sequence of covers. ■

The next Lemma relates the VC dimension of $\widetilde{\Pi}_{ab}$ to that of Π.

Lemma 6 $VC(\widetilde{\Pi}_{ab}) \leq 2VC(\Pi) - 1$.

Proof: Let $\pi_t(X_i) = \mathbb{1}(\pi(X_i) = t)$ for $t \in \{0, 1\}$. Define $\Pi_t = \{\mathbb{1}(\pi(X_i) = t) : \pi \in \Pi\}$. Note that $\Pi_1 = \Pi$ and that $VC(\Pi_0) = VC(\Pi_1)$ by Lemma 9.7 in Kosorok (2008). Now note that for any $(a, b) \in \{0, 1\}^2$

$$\widetilde{\Pi}_{ab} = \{\pi_a \cdot \pi_b : (\pi_a, \pi_b) \in \Pi_a \times \Pi_b\},$$

so Lemma 9.9 (ii) in Kosorok (2008) yields the desired result. ■

C Proofs of main results

Proof of Proposition 3.1: See Proof of Proposition 8.1. ■

Proof of Proposition 4.1: See Proof of Proposition 8.2. ■

Proof of Proposition 4.2: See Proof of Proposition 8.3. ■

Proof of Proposition 8.1: I proof only the identification of the first term of the welfare since the second one follows in the same manner.

$$\begin{aligned}
\mathbb{E}[g(Y_i(1), X_i, \gamma^{(1)})\pi(X_i)] &= \mathbb{E}[\mathbb{E}(g(Y_i(1), X_i, \gamma_1)|X_i)\pi(X_i)] \\
&= \mathbb{E}[\mathbb{E}(g(Y_i(1), X_i, \gamma_1)|D_i = 1, X_i)\pi(X_i)] \\
&= \mathbb{E}[\mathbb{E}(g(Y_i, X_i, \gamma_1)|D_i = 1, X_i)\pi(X_i)] \\
&= \mathbb{E}\left[\mathbb{E}\left(\frac{g(Y_i, X_i, \gamma_1)D_i}{e(X_i)}|X_i\right)\pi(X_i)\right] \\
&= \mathbb{E}\left[\frac{g(Y_i, X_i, \gamma_1)D_i}{e(X_i)}\pi(X_i)\right],
\end{aligned}$$

the first equality follows by LIE and the fact that by selection on observables and definition of Y_i, we have that $\mathbb{E}[Y_i(1)|X_i] = \mathbb{E}[Y_i(1)|D_i = 1, X_i] = \mathbb{E}[Y_i|D_i = 1, X_i]$. The second equality follows from selection on observables, and the third equality from the definition of Y_i and already establishes the identification by the direct method. $\blacksquare$

Proof of Proposition 3.2: Let $d/d\tau$ be the derivative with respect to τ evaluated at $\tau = 0$, let $\varphi_\tau = \varphi + \tau\tilde{\varphi}$ for some $\tilde{\varphi}$ in the space where φ lives and $\mathbb{E}_\tau$ be the expectation with respect to $F + \tau(H - F)$ for some alternative distribution H. Then

$$\frac{d}{d\tau}\mathbb{E}[\varphi_\tau(1, X_i, \bar{\gamma}_\tau(1, X_i))\pi(X_i)] = \frac{d}{d\tau}\mathbb{E}[\varphi_\tau(1, X_i, \gamma(1, X_i))\pi(X_i)] + \frac{d}{d\tau}\mathbb{E}[\varphi(1, X_i, \bar{\gamma}_\tau(1, X_i))\pi(X_i)].$$

For the first term note that

$$\begin{aligned}
\frac{d}{d\tau}\mathbb{E}[\varphi_\tau(1, X_i, \gamma(1, X_i))\pi(X_i)] &= \frac{d}{d\tau}\mathbb{E}\left[\frac{D_i}{e(X_i)}\varphi_\tau(1, X_i, \gamma(1, X_i))\pi(X_i)\right] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau\left[\frac{D_i}{e(X_i)}\varphi_\tau(1, X_i, \gamma(1, X_i))\pi(X_i)\right] \\
&- \frac{d}{d\tau}\mathbb{E}_\tau\left[\frac{D_i}{e(X_i)}\varphi(1, X_i, \gamma(1, X_i))\pi(X_i)\right] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau\left[\frac{D_i}{e(X_i)}(g(Y_i, X_i, \gamma(1, X_i)) - \varphi(1, X_i, \gamma(1, X_i)))\pi(X_i)\right],
\end{aligned}$$

where we use LIE in the first equality, then we use the chain rule and finally that $\varphi_\tau(1, X_i, \gamma(1, X_i))$ is a projection of $g(Y_i, X_i, \gamma(1, X_i))$. For the second term, we have

$$\begin{aligned}
\frac{d}{d\tau}\mathbb{E}[\varphi(1, X_i, \bar{\gamma}_\tau(1, X_i))\pi(X_i)] &= \frac{d}{d\tau}\mathbb{E}[\alpha_1(D_i, X_i)\bar{\gamma}_\tau(1, X_i)\pi(X_i)] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau[\alpha_1(D_i, X_i)\bar{\gamma}_\tau(1, X_i)\pi(X_i)] \\
&- \frac{d}{d\tau}\mathbb{E}_\tau[\alpha_1(D_i, X_i)\gamma_\tau(1, X_i)\pi(X_i)] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau[\alpha_1(D_i, X_i)(Y_i - \gamma(1, X_i))\pi(X_i)],
\end{aligned}$$

where we use Assumption 2 in the first equality, then the chain rule and then the fact that γ is a projection. Then, following Chernozhukov et al. (2022) we have that

$$\Gamma_{1i} = \varphi(1, X_i, \gamma) + \frac{D_i}{e(X_i)}(g(Y_i, X_i, \gamma_1) - \varphi(1, X_i, \gamma)) + \alpha_1(D_i, X_i)(Y_i - \gamma(D_i, X_i)).$$

The arguments for Γ_{0i} are the analogous. $\blacksquare$

Proof of Proposition 8.2:

$$
\begin{aligned}
W(\pi) &= \mathbb{E}\Big[\sum_{(a,b)\in\{0,1\}^2} g(Y_i(a),X_i,Y_j(b),X_j,\gamma^{(a)},\gamma^{(b)})\pi_{ab}(X_i,X_j)\Big]\\
&= \mathbb{E}\Big[\mathbb{E}\Big(\sum_{(a,b)\in\{0,1\}^2} g(Y_i(a),X_i,Y_j(b),X_j,\gamma^{(a)},\gamma^{(b)})\Big|X_i,X_j\Big)\pi_{ab}(X_i,X_j)\Big]\\
&= \mathbb{E}\Big[\mathbb{E}\Big(\sum_{(a,b)\in\{0,1\}^2} g(Y_i(a),X_i,Y_j(b),X_j,\gamma^{(a)},\gamma^{(b)})\Big|X_i,D_i=a,X_j,D_j=b\Big)\pi_{ab}(X_i,X_j)\Big]\\
&= \mathbb{E}\Big[\mathbb{E}\Big(\sum_{(a,b)\in\{0,1\}^2} g(Y_i,X_i,Y_j,X_j,\gamma^{(a)},\gamma^{(b)})\Big|X_i,D_i=a,X_j,D_j=b\Big)\pi_{ab}(X_i,X_j)\Big]\\
&= \mathbb{E}\Big[\mathbb{E}\Big(\sum_{(a,b)\in\{0,1\}^2} \frac{g(Y_i,X_i,Y_j,X_j,\gamma^{(a)},\gamma^{(b)})D_{ij}^{ab}}{e_{ab}(X_i,X_j)}\Big|X_i,X_j\Big)\pi_{ab}(X_i,X_j)\Big]\\
&= \mathbb{E}\Big[\sum_{(a,b)\in\{0,1\}^2} \frac{g(Y_i,X_i,Y_j,X_j,\gamma^{(a)},\gamma^{(b)})D_{ij}^{ab}}{e_{ab}(X_i,X_j)}\pi_{ab}(X_i,X_j)\Big],
\end{aligned}
$$

where in the second equality I use LIE, in the third I use selection on observables, in the fourth I use the definition of Y_i. The identification by the direct method is in the fourth equality while the IPW is the last equality. ■

Proof of Proposition 8.3: Let us start with the DM identification. As usual, let $d/d\tau$ be the derivative at $\tau=0$. Let me also make the dependence on φ explicit: $m_{ab}(Z_i,Z_j,\gamma,\varphi)=\varphi(a,X_i,b,X_j,\gamma_a,\gamma_b)$, let also $\varphi_\tau=\varphi+\tau\tilde{\varphi}$ for some $\tilde{\varphi}\in L_2$. By the chain rule

$$
\frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i,Z_j,\bar{\gamma}_\tau,\varphi_\tau)]=\frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i,Z_j,\bar{\gamma}_\tau,\varphi)]+\frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i,Z_j,\gamma,\varphi_\tau)].
$$

By Assumption 8 we have that the first term is

$$
\frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i,Z_j,\bar{\gamma}_\tau,\varphi)]=\mathbb{E}\Big[\sum_{p=1}^{P}\alpha^{\gamma}_{ab,p}(D_i,X_i,D_j,X_j)(c_{1p}\bar{\gamma}_\tau(D_i,X_i)+c_{2p}\bar{\gamma}_\tau(D_j,X_j))\Big],
$$

so by Lemma 1 and equation (2.16) in Escanciano and Terschuur (2023) we have that

$$
\phi^{\gamma}_{ab}(D_i,X_i,D_j,X_j,e,\alpha^{\gamma})=\sum_{p=1}^{P}\alpha^{\gamma}_{ab,p}(D_i,X_i,D_j,X_j,e)(c_{1p}Y_i+c_{2p}Y_j-c_{1p}\gamma(D_i,X_i)-c_{2p}\gamma(D_j,X_j)).
$$

For the second term notice that

$$
\begin{aligned}
\mathbb{E}[\varphi(a,X_i,b,X_j,\gamma_a,\gamma_b)] &= \mathbb{E}\Big[\varphi(a,X_i,b,X_j,\gamma_a,\gamma_b)\frac{D_{ij}^{ab}}{e_{ab}(X_i,X_j)}\Big]\\
&= \mathbb{E}\Big[\varphi(a,X_i,b,X_j,\gamma_a,\gamma_b)\frac{1}{e_{ab}(X_i,X_j)}\Big|D_{ij}^{ab}=1\Big]\mathbb{P}(D_i=a,D_j=b)\\
&= \mathbb{E}\Big[\varphi(D_i,X_i,D_j,X_j,\gamma_a,\gamma_b)\frac{D_{ij}^{ab}}{e_{ab}(X_i,X_j)}\Big].
\end{aligned}
$$

So by the same arguments

$$\phi^{\varphi}_{ab}(D_i, X_i, D_j, X_j, \varphi, \alpha^m) = \alpha^{\varphi}_{ab}(D_i, X_i, D_j, X_j)(g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b) - \varphi(D_i, X_i, D_j, X_j, \gamma_a, \gamma_b)),$$

with $\alpha^{\varphi}_{ab}(D_i, X_i, D_j, X_j) = D^{ab}_{ij}/e_{ab}(X_i, X_j)$. For the IPW identification let me make the dependence on the propensity score explicit: $m_{ab}(Z_i, Z_j, \gamma, \varphi, e) = g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}/e_{ab}(X_i, X_j)$. For $c \in \{0, 1\}$ let $e_{c,\tau} = e_c + \tau\tilde{e}_c$ for some $\tilde{e}_c \in L_2$ and $e_\tau = (e_{a,\tau}, e_{b,\tau})$. Then

$$\frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i, Z_j, \bar{\gamma}_\tau, e_\tau)] = \frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i, Z_j, \bar{\gamma}_\tau, e)] + \frac{d}{d\tau}\mathbb{E}[m_{ab}(Z_i, Z_j, \gamma, e_\tau)].$$

For the first term, we have the same result as above by using Assumption 8. For the second term note

$$\begin{aligned}
\frac{d}{d\tau}\mathbb{E}\left[\frac{g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_{a,\tau}(X_i)e_{b,\tau}(X_j)}\right] &= \frac{d}{d\tau}\mathbb{E}\left[\frac{-g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_a(X_i)^2 e_b(X_j)}e_{a,\tau}(X_i)\right] \\
&+ \frac{d}{d\tau}\mathbb{E}\left[\frac{-g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_a(X_i)e_b(X_j)^2}e_{b,\tau}(X_j)\right] \\
&= \frac{d}{d\tau}\mathbb{E}\left[\mathbb{E}\left(\frac{-g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_a(X_i)^2 e_b(X_j)}\middle| X_i\right)e_{a,\tau}(X_i)\right] \\
&+ \frac{d}{d\tau}\mathbb{E}\left[\mathbb{E}\left(\frac{-g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_a(X_i)e_b(X_j)^2}\middle| X_j\right)e_{b,\tau}(X_j)\right].
\end{aligned}$$

So by the same arguments as before

$$\phi^{e}_{ab}(D_i, X_i, D_j, X_j, e, \alpha^e) = \alpha^e_{ab,1}(X_i)(\mathbb{1}(D_i = a) - e_a(X_i)) + \alpha^e_{ab,2}(X_j)(\mathbb{1}(D_j = b) - e_b(X_j)),$$

where

$$\begin{aligned}
\alpha^e_{ab,1}(X_i) &= -\mathbb{E}\left[\frac{g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_a(X_i)^2 e_b(X_j)}\middle| X_i\right], \\
\alpha^e_{ab,2}(X_j) &= -\mathbb{E}\left[\frac{g(Y_i, X_i, Y_j, X_j, \gamma_a, \gamma_b)D^{ab}_{ij}}{e_a(X_i)e_b(X_j)^2}\middle| X_j\right].
\end{aligned}$$

■

Proof of Proposition 3: Let $\gamma_{c,\tau} = \gamma_c + \tau\tilde{\gamma}_c$ for some $\tilde{\gamma}_c \in L_2$. We have that for $(a, b) \in \{0, 1\}^2$

$$\begin{aligned}
\frac{d}{d\tau}\mathbb{E}[\gamma_{a,\tau}(X_i) + \gamma_{b,\tau}(X_j)] &= \frac{d}{d\tau}\mathbb{E}\left[\gamma_{a,\tau}(X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} + \gamma_{b,\tau}(X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right] \\
&= \frac{d}{d\tau}\mathbb{E}\left[\bar{\gamma}_\tau(D_i, X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} + \bar{\gamma}_\tau(D_j, X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau\left[\bar{\gamma}_\tau(D_i, X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} + \bar{\gamma}_\tau(D_j, X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right] \\
&- \frac{d}{d\tau}\mathbb{E}_\tau\left[\gamma(D_i, X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} + \gamma(D_j, X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau\left[\frac{\mathbb{1}(D_i = a)}{e_a(X_i)}(Y_i - \gamma(D_i, X_i)) + \frac{\mathbb{1}(D_j = b)}{e_b(X_j)}(Y_j - \gamma(D_j, X_j))\right].
\end{aligned}$$

Also, let $\Delta_{a,b} = \gamma_a(X_i) - \gamma_b(X_j)$, then

$$\frac{d}{d\tau}\mathbb{E}[|\gamma_{a,\tau}(X_i) - \gamma_{b,\tau}(X_j)|] = \frac{d}{d\tau}\mathbb{E}[|\Delta_{ab} + \tau(\tilde{\gamma}_a(X_i) - \tilde{\gamma}_b(X_j))|].$$

As shown in Escanciano and Terschuur (2023), the Gateaux derivative of the mapping $\Delta \mapsto \mathbb{E}(|\Delta|)$ is some direction ν (assuming no point mass at zero, which follows from the assumptions in the Proposition) is $\mathbb{E}[sgn(\Delta)\nu]$. Hence, by the chain rule

$$\begin{aligned}
\frac{d}{d\tau}\mathbb{E}[\gamma_{a,\tau}(X_i) + \gamma_{b,\tau}(X_j)] &= \frac{d}{d\tau}\mathbb{E}[sgn(\gamma_a(X_i) - \gamma_b(X_j))(\gamma_{a,\tau}(X_i) - \gamma_{b,\tau}(X_j))] \\
&= \frac{d}{d\tau}\mathbb{E}\left[sgn(\gamma_a(X_i) - \gamma_b(X_j))\left(\gamma_{a,\tau}(X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} - \gamma_{b,\tau}(X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right)\right] \\
&= \frac{d}{d\tau}\mathbb{E}\left[sgn(\gamma_a(X_i) - \gamma_b(X_j))\left(\gamma_\tau(D_i, X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} - \gamma_\tau(D_j, X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right)\right] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau\left[sgn(\gamma_a(X_i) - \gamma_b(X_j))\left(\gamma_\tau(D_i, X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} - \gamma_\tau(D_j, X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right)\right] \\
&- \frac{d}{d\tau}\mathbb{E}_\tau\left[sgn(\gamma_a(X_i) - \gamma_b(X_j))\left(\gamma(D_i, X_i)\frac{\mathbb{1}(D_i = a)}{e_a(X_i)} - \gamma(D_j, X_j)\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}\right)\right] \\
&= \frac{d}{d\tau}\mathbb{E}_\tau\left[sgn(\gamma_a(X_i) - \gamma_b(X_j))\left(\frac{\mathbb{1}(D_i = a)}{e_a(X_i)}(Y_i - \gamma(D_i, X_i)) - \frac{\mathbb{1}(D_j = b)}{e_b(X_j)}(Y_j - \gamma(D_j, X_j))\right)\right].
\end{aligned}$$

So by the results in Escanciano and Terschuur (2023), the locally robust score is given by

$$\begin{aligned}
2\Gamma_{ij}^{ab} &= \gamma_a(X_i) + \gamma_b(X_j) - |\gamma_a(X_i) - \gamma_b(X_j)| \\
&+ \frac{\mathbb{1}(D_i = a)}{e_a(X_i)}(Y_i - \gamma(D_i, X_i)) + \frac{\mathbb{1}(D_j = b)}{e_b(X_j)}(Y_j - \gamma(D_j, X_j)) \\
&- sgn(\gamma_a(X_i) - \gamma_b(X_j))\left(\frac{\mathbb{1}(D_i = a)}{e_a(X_i)}(Y_i - \gamma(D_i, X_i)) - \frac{\mathbb{1}(D_j = b)}{e_b(X_j)}(Y_j - \gamma(D_j, X_j))\right) \\
&= \gamma_a(X_i) + \gamma_b(X_j) - |\gamma_a(X_i) - \gamma_b(X_j)| \\
&+ (1 - sgn(\gamma_a(X_i) - \gamma_b(X_j)))\frac{\mathbb{1}(D_i = a)}{e_a(X_i)}(Y_i - \gamma(D_i, X_i)) \\
&+ (1 + sgn(\gamma_a(X_i) - \gamma_b(X_j)))\frac{\mathbb{1}(D_j = b)}{e_b(X_j)}(Y_j - \gamma(D_j, X_j)).
\end{aligned}$$

■

Before proving the rest of the main results I introduce a representation of U-statistics which will be very useful for the coming proofs. For any function $f : \mathcal{Z}^2 \to \mathbb{R}$ let $\mathbb{U}_n f(X_i, X_j) = \binom{n}{2}^{-1} \sum_{i<j} f(X_i, X_j)$. Let κ be the permutations of $\{1, ..., n\}$, then, as in Clémençon et al. (2008), we can rewrite

$$\mathbb{U}_n f(Z_i, Z_j) = \frac{1}{n!}\sum_{\kappa} \lfloor n/2 \rfloor^{-1} \sum_{i=1}^{\lfloor n/2 \rfloor} f(Z_{\kappa(i)}, Z_{\kappa(\lfloor n/2 \rfloor + i)}). \tag{8.1}$$

This expresses $\mathbb{U}_n f(Z_i, Z_j)$ as a (dependent) sum of averages of i.i.d. random variables (i.e. $f(Z_{\kappa(i)}, Z_{\kappa(\lfloor n/2\rfloor+i)})$ are i.i.d. for $i = 1, ..., \lfloor n/2\rfloor$).

Proof of Lemma 1: Using the definition of $W(\pi)$ and $\widetilde{W}_n(\pi)$ and the triangle inequality we know that

$$\begin{aligned}\mathbb{E}\left[\sup_{\pi\in\Pi}|\widetilde{W}_n(\pi) - W(\pi)|\right] &= \mathbb{E}\left[\sup_{\pi\in\Pi}\left|\mathbb{U}_n \sum_{(a,b)\in\{0,1\}^2}\left(\Gamma_{ij}^{ab}\pi_{ab}(X_i, X_j) - \mathbb{E}[\Gamma_{ij}^{ab}\pi_{ab}(X_i, X_j)]\right)\right|\right] \\ &\leq \sum_{(a,b)\in\{0,1\}^2}\mathbb{E}\left[\sup_{\pi\in\Pi}\left|\mathbb{U}_n\left(\Gamma_{ij}^{ab}\pi_{ab}(X_i, X_j) - \mathbb{E}[\Gamma_{ij}^{ab}\pi_{ab}(X_i, X_j)]\right)\right|\right].\end{aligned}$$

By the representation used in (8.1) we can rewrite the above as

$$\begin{aligned}\sum_{(a,b)\in\{0,1\}^2}\mathbb{E}\Bigg[\sup_{\pi\in\Pi}\Bigg|\frac{1}{n!}\sum_{\kappa}\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Big(&\Gamma^{ab}_{\kappa(i)\kappa(\lfloor n/2\rfloor+i)}\pi_{ab}(X_{\kappa(i)}, X_{\kappa(\lfloor n/2\rfloor+i)}) \\ &- \mathbb{E}[\Gamma^{ab}_{\kappa(i)\kappa(\lfloor n/2\rfloor+i)}\pi_{ab}(X_{\kappa(i)}, X_{\kappa(\lfloor n/2\rfloor+i)})]\Big)\Bigg|\Bigg]. \end{aligned} \tag{8.2}$$

Introduce an independent ghost sample $(Z_1', ..., Z_n')$ which is distributed as $(Z_1, ..., Z_n)$, Rademacher random variables ε_i, $i = 1, ..., n$, such that $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = 1/2$ and construct ghost scores $\Gamma_{ij}'^{ab}$ using the ghost sample. Let $\mathbb{E}_Z$ be the expectation with respect to the distribution of the sample $(Z_1, ..., Z_n)$ and define $\mathbb{E}_{Z'}$ and $\mathbb{E}_\varepsilon$ similarly. Define the Rademacher complexity as

$$\mathcal{R}_n(\Pi) = \mathbb{E}_\varepsilon\left(\sup_{\pi\in\Pi}\left|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})\right|\right).$$

Again the key here is that the summands of the sum inside the expectation in $\mathcal{R}_n(\Pi)$ are independent. We are now ready to use a classical symmetrization argument, since Z_i' has the same distribution as Z_i we have that (8.2) is equal to

$$\begin{aligned}
&\sum_{(a,b)\in\{0,1\}^2} \mathbb{E}_Z\left[\sup_{\pi\in\Pi}\left|\frac{1}{n!}\sum_{\kappa}\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Bigg(\Gamma^{ab}_{\kappa(i)\kappa(\lfloor n/2\rfloor+i)}\pi_{ab}(X_{\kappa(i)},X_{\kappa(\lfloor n/2\rfloor+i)})\right.\right.\\
&\qquad\left.\left. -\mathbb{E}_{Z'}[\Gamma'^{ab}_{\kappa(i)\kappa(\lfloor n/2\rfloor+i)}\pi_{ab}(X'_{\kappa(i)},X'_{\kappa(\lfloor n/2\rfloor+i)})]\Bigg)\right|\right]\\
&\qquad\le \frac{1}{n!}\sum_{\kappa}\sum_{(a,b)\in\{0,1\}^2}\mathbb{E}_{Z,Z'}\left[\sup_{\pi\in\Pi}\left|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Bigg(\Gamma^{ab}_{\kappa(i)\kappa(\lfloor n/2\rfloor+i)}\pi_{ab}(X_{\kappa(i)},X_{\kappa(\lfloor n/2\rfloor+i)})\right.\right.\\
&\qquad\left.\left. -\Gamma'^{ab}_{\kappa(i)\kappa(\lfloor n/2\rfloor+i)}\pi_{ab}(X'_{\kappa(i)},X'_{\kappa(\lfloor n/2\rfloor+i)})\Bigg)\right|\right]\\
&\qquad= \sum_{(a,b)\in\{0,1\}^2}\mathbb{E}_{Z,Z',\varepsilon}\left[\sup_{\pi\in\Pi}\left|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Bigg(\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})\right.\right.\\
&\qquad\left.\left. -\Gamma'^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X'_i,X'_{\lfloor n/2\rfloor+i})\Bigg)\right|\right]\\
&\qquad\le \sum_{(a,b)\in\{0,1\}^2}\mathbb{E}_{Z,Z',\varepsilon}\left[\sup_{\pi\in\Pi}\left|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})\right|\right.\\
&\qquad\left. +\left|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma'^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X'_i,X'_{\lfloor n/2\rfloor+i})\right|\right]\\
&\qquad= \sum_{(a,b)\in\{0,1\}^2}\mathbb{E}[2\mathcal{R}_n(\Pi)].
\end{aligned}$$

The first inequality follows from Jensen's and triangle inequalities, the second equality uses the fact that the vector $(Z_{\pi(i)}, Z_{\pi(\lfloor n/2\rfloor+i)}, Z'_{\pi(i)}, Z'_{\pi(\lfloor n/2\rfloor+i)})$ is identically distributed across $i = 1, ..., \lfloor n/2\rfloor$ for all permutations in κ (so we can just take the permutation $\kappa(i) = i$) and the fact that $\varepsilon_i(\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i}) - \Gamma'^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X'_i, X'_{\lfloor n/2\rfloor+i}))$ and $\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i}) - \Gamma'^{ab}_{i,\lfloor n/2\rfloor+i}\pi_{ab}(X'_i, X'_{\lfloor n/2\rfloor+i})$ have the same distribution, the third inequality uses the triangle inequality and the last equality uses that $Z_i \sim Z'_i$ and the definition of the Rademacher complexity.
■

Proof of Lemma 2: Note that Lemma 5 gives us a sequence of covers B_k for $k = 0, ..., K$ of $\tilde{\Pi}_{ab}$ of radius less than 2^{-k} for some K. For any $j = 1, ..., J$ with $J = \lceil \log_2(\lfloor n/2\rfloor)(1-\beta)\rceil$ and $\pi \in \tilde{\Pi}_{ab}$ let $b_j : \tilde{\Pi}_{ab} \mapsto \tilde{\Pi}_{ab}$ be an operator such that $b_j(\pi)$ is an approximating policy from the cover B_j such that $D_n(\pi, b_j(\pi)) \le 2^{-j}$, such an approximation exists by Lemma 5. By the same Lemma we also know that $|\{b_j(\pi) : \pi \in \tilde{\Pi}_{ab}\}| \le N_{D_n}(2^{-j}, \tilde{\Pi}_{ab}, \{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor})$. Let $\underline{J} = \lceil 1/2\log_2(\lfloor n/2\rfloor)(1-\beta)\rceil$. By using a telescope sum and the approximations $b_0, ..., b_J$ we can decompose the Rademacher complexity as

$$\mathcal{R}_n(\Pi) = \mathbb{E}_\varepsilon\Bigg\{\sup_{\pi\in\Pi}\Bigg|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\Bigg[b_0(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))$$
$$+\sum_{j=1}^{\underline{J}}\Bigg(b_j(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{j-1}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))\Bigg)$$
$$+\left(b_J(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{\underline{J}}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))\right)$$
$$+\left(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})-b_J(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))\right)\Bigg]\Bigg|\Bigg\}.$$

Note that since the distance D_n is bounded by 1, by the second property in Lemma 5 we have that b_0 can be any policy in $\tilde{\Pi}_{ab}$. Hence, we can set $b_0(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))=0$ for all $i=1,...,\lfloor n/2\rfloor$. We approach each of the terms above in turn. Note that $b_0,...,b_J$ is a sequence of increasingly accurate approximations. The first step is to notice that the last term above is negligible, i.e. the term involving the closest approximation vanishes at a $\sqrt{n}$ rate. By using Cauchy-Schwarz and multiplying and dividing we get

$$\sqrt{\lfloor n/2\rfloor}\sup_{\pi\in\Pi}\Bigg|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})-b_J(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\Bigg|$$
$$\leq\sqrt{\lfloor n/2\rfloor}\sup_{\pi\in\Pi}\frac{\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Big|\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})-b_J(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\Big|^2}}{\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}}$$
$$\times\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}$$
$$=\sqrt{\lfloor n/2\rfloor}\sup_{\pi\in\Pi}D_n(\pi_{ab},b_J(\pi_{ab}))\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}$$
$$\leq\sqrt{\lfloor n/2\rfloor}2^{-J}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}$$
$$=\frac{M}{\lfloor n/2\rfloor^{1/2-\beta}}\to 0,$$

where in the last inequality we use Lemma 5 and in the last inequality we use the fact that $J=\lceil\log_2(\lfloor n/2\rfloor)(1-\beta)\rceil$ and the boundedness assumption. Now we show that the second to last term of the Rademacher decomposition is also negligible. Notice that $\{\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_J(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{\underline{J}}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\}_{i=1}^{\lfloor n/2\rfloor}$ are zero mean (conditional on $\{X_i,\Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}$) i.i.d. random variables. They are also bounded below by $a_i=-|\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_J(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{\underline{J}}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))|$

and above by $b_i = -a_i$. Hence, by Hoeffding's inequality

$$\begin{aligned}
&\mathbb{P}_\varepsilon\left(\left|\sum_{i=1}^{\lfloor n/2\rfloor} \varepsilon_i \Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_J(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})) - b_{\underline{J}}(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})))\right| \geq t\right)\\
&\leq 2\exp\left(-\frac{2t^2}{\sum_{i=1}^{\lfloor n/2\rfloor}(b_i - a_i)^2}\right)\\
&= 2\exp\left(-\frac{t^2}{D_n^2(b_J(\pi_{ab}), b_{\underline{J}}(\pi_{ab}))\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\right).
\end{aligned}$$

Hence, for any $a > 0$ we have that

$$\begin{aligned}
&\mathbb{P}_\varepsilon\Bigg(\left|\sqrt{\lfloor n/2\rfloor}\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor} \varepsilon_i \Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_J(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})) - b_{\underline{J}}(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})))\right|\\
&\geq a2^{2-\underline{J}}\sqrt{\frac{\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}{\lfloor n/2\rfloor}}\Bigg)\\
&\leq 2\exp\left(-\frac{a^2 4^{2-\underline{J}}}{D_n^2(b_J(\pi_{ab}), b_{\underline{J}}(\pi_{ab}))}\right)\\
&\leq 2\exp\left(-\frac{a^2 4^{2-\underline{J}}}{\sum_{j=\underline{J}}^{J-1} D_n^2(b_j(\pi_{ab}), b_{j+1}(\pi_{ab}))}\right)\\
&\leq 2\exp\left(-\frac{a^2 4^{2-\underline{J}}}{\left(\sum_{j=\underline{J}}^{J-1} 2^{-(j-1)}\right)^2}\right)\\
&\leq 2\exp(-a^2),
\end{aligned}$$

where we have used triangle inequality in the second inequality and the fact that $\sum_{j=\underline{J}}^{J-1} 2^{-(j-1)} = 2^{2-\underline{J}} - 2^{2-J} \leq 2^{2-\underline{J}}$ in the last inequality. This holds for any policy, hence

$$\begin{aligned}
&\mathbb{P}_\varepsilon\Bigg(\sup_{\pi\in\Pi}\left|\lfloor n/2\rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2\rfloor} \varepsilon_i \Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_J(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})) - b_{\underline{J}}(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})))\right|\\
&\geq a2^{2-\underline{J}}\sqrt{\frac{\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}{\lfloor n/2\rfloor}}\Bigg)\\
&\leq 2|\{b_J(\pi_{ab}), b_{\underline{J}}(\pi_{ab})\}|\exp(-a^2)\\
&\leq 2N_{D_n}(2^{-J}, \widetilde{\Pi}_{ab}, \{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor})\exp(-a^2)\\
&\leq 2N_H(2^{-2J}, \widetilde{\Pi}_{ab})\exp(-a^2)\\
&= 2\exp(\log(N_H(2^{-2J}, \widetilde{\Pi}_{ab})))\exp(-a^2)\\
&\leq 2\exp(5VC(\widetilde{\Pi}_{ab})\log(2^{2J}) - a^2)\\
&\leq 2\exp(5VC(\widetilde{\Pi}_{ab})\log(2^{-2(1-\beta)\log_2(\lfloor n/2\rfloor)}) - a^2),
\end{aligned}$$

where in the first inequality I use the union bound, in the second inequality I use properties of the approximations (see Zhou et al. (2023)), in the third I use Lemma 4 and in the fourth inequality I bound the log of the Hamming covering number by the VC dimension using a result in Haussler (1995). Let now

$$a = \frac{2^{\underline{J}}}{\sqrt{\log(\lfloor n/2 \rfloor)\lfloor n/2 \rfloor^{-1}\sum_{i=1}^{\lfloor n/2 \rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2 \rfloor+i}}},$$

so that

$$a2^{2-\underline{J}}\sqrt{\frac{\sum_{i=1}^{\lfloor n/2 \rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2 \rfloor+i}}{\lfloor n/2 \rfloor}} = \frac{4}{\sqrt{\log(\lfloor n/2 \rfloor)}}.$$

Finally,

$$\begin{aligned}
&\mathbb{P}_{\varepsilon}\left(\sup_{\pi\in\Pi}\left|\lfloor n/2 \rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2 \rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2 \rfloor+i}(b_J(\pi_{ab}(X_i,X_{\lfloor n/2 \rfloor+i})) - b_{\underline{J}}(\pi_{ab}(X_i,X_{\lfloor n/2 \rfloor+i})))\right|\right. \\
&\qquad\left.\geq \frac{4}{\sqrt{\log(\lfloor n/2 \rfloor)}}\right) \\
&\leq 2\exp\left(5VC(\widetilde{\Pi}_{ab})\log(\lfloor n/2 \rfloor^{-2(1-\beta)}) - \frac{\lfloor n/2 \rfloor^{-\beta}}{\log(\lfloor n/2 \rfloor)\sum_{i=1}^{\lfloor n/2 \rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2 \rfloor+i}}\right) \\
&\quad\leq 2\exp\left\{-5\lfloor n/2 \rfloor^{\beta}\log\left(\lfloor n/2 \rfloor^{2(1-\beta)}\right) - \frac{1}{\lfloor n/2 \rfloor^{\beta}\log(\lfloor n/2 \rfloor)M^2}\right\} \to 0,
\end{aligned}$$

where I have used Assumption 7 and the boundedness assumption.

$$\mathbb{E}\left(\sup_{\pi\in\Pi}\left|\lfloor n/2 \rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2 \rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2 \rfloor+i}(b_J(\pi_{ab}(X_i,X_{\lfloor n/2 \rfloor+i})) - b_{\underline{J}}(\pi_{ab}(X_i,X_{\lfloor n/2 \rfloor+i})))\right|\right) \to 0,$$

since for any sequence of random variables X_n and sequence of real numbers a_n if $\lim_{n\to\infty}\mathbb{P}(X_n \leq a_n) = 1$ and $\lim_{n\to\infty} a_n = 0$, then $\lim_{n\to\infty}\mathbb{E}(X_n) = 0$ (proof of this fact uses $\mathbb{E}(X_n) = \int_0^\infty \mathbb{P}(X_n > u)\,du$). Hence, we have proven that

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_n(\Pi)] = \mathbb{E}&\left\{\sup_{\pi\in\Pi}\left|\lfloor n/2 \rfloor^{-1}\sum_{i=1}^{\lfloor n/2 \rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2 \rfloor+i}\left[\sum_{j=1}^{\underline{J}}\left(b_j(\pi_{ab}(X_i,X_{\lfloor n/2 \rfloor+i})) - b_{j-1}(\pi_{ab}(X_i,X_{\lfloor n/2 \rfloor+i}))\right)\right]\right|\right\} \\
&+ o\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Hence I have left what Zhou et al. (2023) call the effective regime. Let $j \in \{1,...,\underline{J}\}$ and a_j be some constant depending on j. As before, conditional on $\{X_i, \Gamma_{i,\lfloor n/2 \rfloor+i}\}_{i=1}^{\lfloor n/2 \rfloor}$ we can apply

Hoeffding inequality and then use the definition of D_n to get

$$
\begin{aligned}
&\mathbb{P}_\varepsilon\Bigg(\Bigg|\lfloor n/2\rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_j(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{j-1}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\Bigg| \\
&\qquad\geq a_j 2^{2-j}\sqrt{\frac{\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}{\lfloor n/2\rfloor}}\Bigg) \\
&\qquad\leq 2\exp\Bigg(-\frac{a_j^2 4^{2-j}}{D_n^2(b_j(\pi_{ab}),b_{j-1}(\pi_{ab}))}\Bigg) \\
&\qquad\leq 2\exp\Bigg(\frac{-a_j^2 4^{2-j}}{4^{-(j-1)}}\Bigg) \\
&\qquad= 2\exp\Bigg(-4a_j^2\Bigg),
\end{aligned}
$$

where in the last inequality we have used the fact that $D_n(b_j(\pi_{ab}),b_{j-1}(\pi_{ab}))\leq 2^{-(j-1)}$ by Lemma 5. Now we let

$$
a_j^2(k)=2\log\Bigg(\frac{2j^2}{\delta_k}N_H(4^{-j},\widetilde{\Pi}_{ab})\Bigg),
$$

where δ_k is some sequence of real numbers indexed by $k\in\mathbb{N}$. For notational convenience define

$$
R_j=\sup_{\pi\in\Pi}\Bigg|\lfloor n/2\rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}(b_j(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{j-1}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\Bigg|.
$$

Then we have that

$$
\begin{aligned}
\mathbb{P}\Bigg(R_j\geq a_j(k)2^{-j}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\Bigg) &\leq 2|\{b_j(\pi_{ab}),b_{j-1}(\pi_{ab})\}|\exp(-a_j^2(k)/2) \\
&\leq 2N_{D_n}(2^{-j},\widetilde{\Pi}_{ab},\{X_i,\Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor})\exp(-a_j^2(k)/2) \\
&\leq 2N_H(2^{-2j},\widetilde{\Pi}_{ab})\exp(-a_j^2(k)/2) \\
&= 2N_H(4^{-j},\widetilde{\Pi}_{ab})\exp(-\log(N_H(4^{-j},\widetilde{\Pi}_{ab})2j^2/\delta_k)) \\
&= \frac{\delta_k}{j^2}.
\end{aligned}
$$

Sum across j and apply this bound with $\delta_k=1/2^k$ to note that

$$
\begin{aligned}
\sum_{j=1}^{J}\mathbb{P}\Bigg(R_j\geq a_j(k)2^{-j}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\Bigg) &\leq \sum_{j=1}^{J}\frac{\delta_k}{j^2} \\
&\leq \sum_{j=1}^{\infty}\frac{\delta_k}{j^2} \\
&\leq \frac{1.7}{2^k}.
\end{aligned}
$$

Let F_{R_j} be the cumulative distribution function of R_j (conditional on $\{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}$). We can bound the following object of interest in the following way

$$
\begin{aligned}
&\mathbb{E}_\varepsilon\left[\sup_{\pi\in\Pi}\left|\lfloor n/2\rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\sum_{j=1}^{J}(b_j(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})) - b_{j-1}(\pi_{ab}(X_i, X_{\lfloor n/2\rfloor+i})))\right|\right] \\
&\quad\le \sum_{j=1}^{J}\mathbb{E}_\varepsilon[R_j] \\
&\quad= \int_0^\infty \sum_{j=1}^{J}(1 - F_{R_j}(r))\,dr \\
&\quad\le \int_0^\infty \sum_{j=1}^{J}\mathbb{P}(R_j \ge r)\,dr \\
&\quad\le \sum_{k=0}^{\infty}\sum_{j=1}^{J}\frac{1.7}{2^k}a_j(k)2^{-j}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}} \\
&\quad\le \sum_{k=0}^{\infty}\sum_{j=1}^{J}\frac{1.7}{2^k}\sqrt{2}\sqrt{\log(2^{k+1}j^2N_H(4^{-j},\widetilde{\Pi}_{ab}))}2^{-j}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}} \\
&\quad\le 1.7\sqrt{2}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\sum_{k=0}^{\infty}2^{-k}\sum_{j=1}^{J}2^{-j}\sqrt{(k+1)\log 2 + 2\log j + \log N_H(4^{-j},\widetilde{\Pi}_{ab})} \\
&\quad\le 1.7\sqrt{2}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\sum_{k=0}^{\infty}2^{-k}\sum_{j=1}^{J}2^{-j}\left(\sqrt{k+1} + \sqrt{2\log j} + \sqrt{5VC(\widetilde{\Pi}_{ab})\log(4^j)}\right) \\
&\quad\le 1.7\sqrt{2}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\sum_{k=0}^{\infty}2^{-k}\Bigg(\sqrt{k+1}\sum_{j=1}^{\infty}2^{-j} + \sqrt{2}\sum_{j=1}^{\infty}2^{-j}\sqrt{\log j} \\
&\quad+ \sqrt{5VC(\widetilde{\Pi}_{ab})}\sum_{j=1}^{\infty}2^{-j}\sqrt{\log 4^j}\Bigg) \\
&\quad\le 1.7\sqrt{2}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\left(\sum_{k=0}^{\infty}2^{-k}\sqrt{k+1} + \frac{\sqrt{2}}{2}\sum_{k=0}^{\infty}2^{-k} + \sqrt{5VC(\widetilde{\Pi}_{ab})}1.6\sum_{k=0}^{\infty}2^{-k}\right) \\
&\quad\le 1.7\sqrt{2}\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\left(5 + 3.2\sqrt{5VC(\widetilde{\Pi}_{ab})}\right).
\end{aligned}
$$

So taking expectations over $\{X_i, \Gamma_{i,\lfloor n/2\rfloor+i}\}_{i=1}^{\lfloor n/2\rfloor}$, using this bound and the Jensen's inequality we get

$$
\begin{aligned}
&\mathbb{E}\Bigg[\sup_{\pi\in\Pi}\Bigg|\lfloor n/2\rfloor^{-1/2}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\sum_{j=1}^{J}(b_j(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{j-1}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\Bigg|\Bigg]\\
&\le 1.7\sqrt{2}\Big(5+8\sqrt{5VC(\widetilde{\Pi}_{ab})}\Big)\mathbb{E}\Bigg[\sqrt{\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}}\Bigg]\\
&\le 1.7\sqrt{2}\Big(5+8\sqrt{5VC(\widetilde{\Pi}_{ab})}\Big)\sqrt{\mathbb{E}\Big[\Gamma^{2\,ab}_{i,\lfloor n/2\rfloor+i}\Big]}\\
&= 1.7\sqrt{2}\Big(5+8\sqrt{5VC(\widetilde{\Pi}_{ab})}\Big)\sqrt{S_{ab}}\\
&\le C\sqrt{VC(\widetilde{\Pi}_{ab})S_{ab}},
\end{aligned}
$$

for some constant $C>0$. Dividing both sides by $\sqrt{\lfloor n/2\rfloor}$ we get

$$
\begin{aligned}
&\mathbb{E}\Bigg[\sup_{\pi\in\Pi}\Bigg|\lfloor n/2\rfloor^{-1}\sum_{i=1}^{\lfloor n/2\rfloor}\varepsilon_i\Gamma^{ab}_{i,\lfloor n/2\rfloor+i}\sum_{j=1}^{J}(b_j(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i}))-b_{j-1}(\pi_{ab}(X_i,X_{\lfloor n/2\rfloor+i})))\Bigg|\Bigg]\\
&\le C\sqrt{\frac{VC(\widetilde{\Pi}_{ab})S_{ab}}{\lfloor n/2\rfloor}},
\end{aligned}
$$

and hence

$$
\begin{aligned}
\mathbb{E}[\mathcal{R}_n(\Pi)] &\le C\sqrt{\frac{VC(\widetilde{\Pi}_{ab})S_{ab}}{\lfloor n/2\rfloor}}+o\Big(\frac{1}{\sqrt{n}}\Big)\\
&=\mathcal{O}\Bigg(\sqrt{\frac{VC(\widetilde{\Pi}_{ab})S_{ab}}{\lfloor n/2\rfloor}}\Bigg).
\end{aligned}
$$

■

Proof of Lemma 3: Define the following random variables

$$
\begin{aligned}
\hat{R}^{(1)}_{ij,ab,l} &= m_{ab}(Z_i,Z_j,\hat{\gamma}_l,\nu)-m_{ab}(Z_i,Z_j,\gamma,\nu)\\
\hat{R}^{(2)}_{ij,ab,l} &= m_{ab}(Z_i,Z_j,\gamma,\hat{\nu}_l)-m_{ab}(Z_i,Z_j,\gamma,\nu)\\
\hat{R}^{(3)}_{ij,ab,l} &= \phi^{\gamma}_{ab}(Z_i,Z_j,\hat{\gamma}_l,\alpha^{\gamma})-\phi^{\gamma}_{ab}(Z_i,Z_j,\gamma,\alpha^{\gamma})\\
\hat{R}^{(4)}_{ij,ab,l} &= \phi^{\gamma}_{ab}(Z_i,Z_j,\gamma,\hat{\alpha}^{\gamma}_l)-\phi^{\gamma}_{ab}(Z_i,Z_j,\gamma,\alpha^{\gamma})\\
\hat{R}^{(5)}_{ij,ab,l} &= \phi^{\nu}_{ab}(Z_i,Z_j,\hat{\nu}_l,\alpha^{\nu})-\phi^{\nu}_{ab}(Z_i,Z_j,\nu,\alpha^{\nu})\\
\hat{R}^{(6)}_{ij,ab,l} &= \phi^{\nu}_{ab}(Z_i,Z_j,\nu,\hat{\alpha}^{\nu}_l)-\phi^{\nu}_{ab}(Z_i,Z_j,\nu,\alpha^{\nu}).
\end{aligned}
$$

Then,

$$
\mathbb{E}\Big[\sup_{\pi\in\Pi_n}|\hat{W}_n(\pi)-\widetilde{W}_n(\pi)|\Big]=\mathbb{E}\Bigg(\sup_{\pi\in\Pi_n}\Bigg|\binom{n}{2}^{-1}\sum_{l=1}^{L}\sum_{(i,j)\in I_l}\sum_{(a,b)\in\pi}\sum_{k=1}^{6}(\hat{R}^{(k)}_{ij,ab,l}+\hat{\xi}_{ij,ab,l}+\hat{\xi}^{\gamma}_{ij,ab,l}+\hat{\xi}^{\nu}_{ij,ab,l})\pi_{ab}(X_i,X_j)\Bigg|\Bigg).
$$

By repeated use of the triangle inequality

$$(\dagger)\quad \mathbb{E}\Big[\sup_{\pi\in\Pi_n}|\hat{W}_n(\pi)-\widetilde{W}_n(\pi)|\Big] \leq \sum_{l=1}^{L}\sum_{(a,b)\in\pi}\mathbb{E}\bigg(\sup_{\pi\in\Pi_n}\bigg|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}(\hat{R}^{(1)}_{ij,l}+\hat{R}^{(3)}_{ij,l})\pi_{ab}(X_i,X_j)\bigg|\bigg)$$
$$+\sum_{l=1}^{L}\sum_{(a,b)\in\pi}\mathbb{E}\bigg(\sup_{\pi\in\Pi_n}\bigg|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}(\hat{R}^{(2)}_{ij,l}+\hat{R}^{(5)}_{ij,l})\pi_{ab}(X_i,X_j)\bigg|\bigg)$$
$$+\sum_{l=1}^{L}\sum_{(a,b)\in\pi}\mathbb{E}\bigg(\sup_{\pi\in\Pi_n}\bigg|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}(\hat{R}^{(4)}_{ij,l}+\hat{R}^{(6)}_{ij,l})\pi_{ab}(X_i,X_j)\bigg|\bigg)$$
$$+\sum_{l=1}^{L}\sum_{(a,b)\in\pi}\mathbb{E}\bigg(\sup_{\pi\in\Pi_n}\bigg|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}(\hat{\xi}_{ij,l}+\hat{\xi}^{\gamma}_{ij,l}+\hat{\xi}^{\nu}_{ij,l})\pi_{ab}(X_i,X_j)\bigg|\bigg).$$

I will bound each of the terms separately. The same arguments apply for all $l=1,...,L$ and $(a,b)\in\pi$, hence we focus on some fixed (a,b) and l. Let N_l^c be the observations not in I_l. By adding and subtracting $\mathbb{E}[(\hat{R}^{(1)}_{ij,ab,l}+\hat{R}^{(3)}_{ij,ab,l})\pi_{ab}(X_i,X_j)|N_l^c]$ and applying the triangle inequality we get that the summands of the first term are bounded by

$$\mathbb{E}\bigg(\sup_{\pi\in\Pi_n}\bigg|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}(\hat{R}^{(1)}_{ij,l}+\hat{R}^{(3)}_{ij,l})\pi_{ab}(X_i,X_j)-\mathbb{E}[(\hat{R}^{(1)}_{ij,ab,l}+\hat{R}^{(3)}_{ij,ab,l})\pi_{ab}(X_i,X_j)|N_l^c])\bigg|\bigg)\quad(\star)$$
$$+\,\mathbb{E}\bigg(\sup_{\pi\in\Pi_n}\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}|\mathbb{E}[(\hat{R}^{(1)}_{ij,ab,l}+\hat{R}^{(3)}_{ij,ab,l})\pi_{ab}(X_i,X_j)|\hat{\gamma}_l]|\bigg).\quad(\star\star)$$

By Assumption 5 we know that

$$\begin{aligned}|\mathbb{E}[\hat{R}^{(1)}_{ij,ab,l}+\hat{R}^{(3)}_{ij,ab,l}|N_l^c]| &= |\mathbb{E}[m_{ab}(Z_i,Z_j,\hat{\gamma}_l,\nu)+\phi^{\gamma}_{ab}(Z_i,Z_j,\hat{\gamma}_l,\alpha^{\gamma})|\hat{\gamma}_l]|\\ &\leq C||\hat{\gamma}_l-\gamma||^2.\end{aligned}$$

Applying the conditional Jensen's inequality (on the absolute value) in $(\star\star)$ and noting that the resulting expression is maximized by treating everybody we get that

$$(\star\star)\leq C\mathbb{E}[||\hat{\gamma}_l-\gamma||^2]\underbrace{\binom{n}{2}^{-1}|I_l|}_{\leq 1}=o(n^{-2\lambda_\gamma})=o(1/\sqrt{n}),$$

where the last equality follows since $2\lambda_\gamma \geq 1/2$. For $(\star)$, note that

$$
\begin{aligned}
(\star) \leq & \binom{n}{2}^{-1} |I_l| \mathbb{E}\left(\sup_{\pi \in \Pi_n} \left| |I_l|^{-1} \sum_{(i,j) \in I_l} \hat{R}^{(1)}_{ij,l} \pi_{ab}(X_i, X_j) - \mathbb{E}[\hat{R}^{(1)}_{ij,ab,l} \pi_{ab}(X_i, X_j) | N_l^c] \right| \right) \\
& + \binom{n}{2}^{-1} |I_l| \mathbb{E}\left(\sup_{\pi \in \Pi_n} \left| |I_l|^{-1} \sum_{(i,j) \in I_l} \hat{R}^{(3)}_{ij,l} \pi_{ab}(X_i, X_j) - \mathbb{E}[\hat{R}^{(3)}_{ij,ab,l} \pi_{ab}(X_i, X_j) | N_l^c]) \right| \right) \\
= & \binom{n}{2}^{-1} |I_l| \mathbb{E}\left[\mathbb{E}\left(\sup_{\pi \in \Pi_n} \left| |I_l|^{-1} \sum_{(i,j) \in I_l} \hat{R}^{(1)}_{ij,l} \pi_{ab}(X_i, X_j) - \mathbb{E}[\hat{R}^{(1)}_{ij,ab,l} \pi_{ab}(X_i, X_j) | N_l^c] \right| \middle| N_l^c \right) \right] \\
& + \binom{n}{2}^{-1} |I_l| \mathbb{E}\left[\mathbb{E}\left(\sup_{\pi \in \Pi_n} \left| |I_l|^{-1} \sum_{(i,j) \in I_l} \hat{R}^{(3)}_{ij,l} \pi_{ab}(X_i, X_j) - \mathbb{E}[\hat{R}^{(3)}_{ij,ab,l} \pi_{ab}(X_i, X_j) | N_l^c] \right| \middle| N_l^c \right) \right].
\end{aligned}
$$

The inner expectations are the expected supremum of centered U-processes. Using Lemma 1 We can bound these inner expectations with Rademacher complexities. However, in the same way we used the construction in Equation (8.1) in Lemma 1 to be able to bound the U-process with a Rademacher complexity which involves a sum of independent terms, we need to use such a construction for each fold I_l. Take the cross-fitting technique in Escanciano and Terschuur (2023) where we split $\{1, ..., n\}$ into sets $\mathcal{C} = \{C_1, ..., C_K\}$ and take the intersection between $\mathcal{C}^2$ and the set $\{(i,j) \in \{1, ..., n\}^2 : i < j\}$. I_l can be either a triangle ($I_l \in T$, where $T = \{I_l : i \in C_f, j \in C_g, f < g, (i,j) \in I_l\}$) or a square ($I_l \in S$, where $S = \{I_l : i \in C_f, j \in C_g, f = g, (i,j) \in I_l\}$) and that in each case we can bound the U-process with the following Rademacher complexities

$$
\mathcal{R}_{n,l}(\Pi_{ab}) = \begin{cases} \mathbb{E}_\varepsilon\left(\sup_{\pi \in \Pi} \left| |C_k|^{-1} \sum_{i=1}^{|C_k|} \varepsilon_i \hat{R}^{(q)}_{\rho(i,k),|C_k|+i} \pi_{ab}(X_{\rho(i,k)}, X_{|C_k|+i}) \right| \right) & \text{if } I_l \in S \\ \mathbb{E}_\varepsilon\left(\sup_{\pi \in \Pi} \left| \lfloor |C_k|/2 \rfloor^{-1} \sum_{i=1}^{\lfloor |C_k|/2 \rfloor} \varepsilon_i \hat{R}^{(q)}_{i,\lfloor |C_k|/2 \rfloor + i,l} \pi_{ab}(X_i, X_{\lfloor |C_k|/2 \rfloor + i}) \right| \right) & \text{if } I_l \in T, \end{cases}
$$

for $q = 1, 3$. Hence, by Lemmas 1 and 2 we have that

$$
\binom{n}{2}^{-1} |I_l| \mathbb{E}\left(\sup_{\pi \in \Pi_n} \left| |I_l|^{-1} \sum_{(i,j) \in I_l} \hat{R}^{(1)}_{ij,l} \pi_{ab}(X_i, X_j) - \mathbb{E}[\hat{R}^{(1)}_{ij,ab,l} \pi_{ab}(X_i, X_j) | N_l^c] \right| \middle| N_l^c \right) = \mathcal{O}\left(\sqrt{\frac{S^{(1)}_{ab,l} VC(\Pi_{ab,n})}{\lfloor C_k/2 \rfloor}} \right),
$$

where $S^{(1)}_{ab,l} = \mathbb{E}[\hat{R}^{(1)^2}_{ij,l} | N_l^c]$. Noting that $\mathbb{E}[S^{(1)}_{ab,l}] = \mathbb{E}[(m_{ab}(Z_i, Z_j, \hat{\gamma}_l, \nu) - m_{ab}(Z_i, Z_j, \gamma, \nu))^2]$ and using Assumption 4, Jensen's inequality, the fact that $|I_l| = |C_k| \times |C_m|$ if $I_l = I(C_k, C_m)$ and $|I_l| = |C_k| \times |C_k - 1|/2$ if $I_l = I(C_k, C_k)$ and that for evenly sized folds $|C_k|/(n-1) \leq 1$ for all $k = 1, ..., K$ we have that

$$
\begin{aligned}
& \mathbb{E}\left(\sup_{\pi \in \Pi_n} \left| \binom{n}{2}^{-1} \sum_{(i,j) \in I_l} \hat{R}^{(1)}_{ij,l} \pi_{ab}(X_i, X_j) - \mathbb{E}[\hat{R}^{(1)}_{ij,ab,l} \pi_{ab}(X_i, X_j) | N_l^c]) \right| \right) \\
& = \mathcal{O}\left(\sqrt{VC(\Pi_{ab,n}) \frac{a((1 - K^{-1})n)^2}{n^{1+2\lambda_\gamma}}} \right).
\end{aligned}
$$

The same bound applies by using the same arguments when we replace $\hat{R}^{(1)}_{ij,l}$ by $\hat{R}^{(3)}_{ij,l}$. Also, this bound applies to all folds I_l, hence, summing across all folds gives us the same asymptotic bound. As a result, we have bounded the first term on the right-hand side in (†). For the second term, we can follow exactly the same steps as with the first term to get the same bounds with λ_γ replaced by λ_ν. For the third term in (†) note that by Assumption 5 (i) (global robustness of α), we have that $\mathbb{E}[\hat{R}^{(4)}_{ij,ab,l}|N_l^c] = \mathbb{E}[\hat{R}^{(6)}_{ij,ab,l}|N_l^c] = 0$. Hence, we do not need to add and subtract anything and we can apply Lemmas 1 and 2 directly to get that for $q = 4, 6$

$$\mathbb{E}\left(\sup_{\pi\in\Pi_n}\left|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}\hat{R}^{(q)}_{ij,l}\pi_{ab}(X_i,X_j)\right|\right) = \mathcal{O}\left(\sqrt{VC(\Pi_{ab,n})\frac{a((1-K^{-1})n)^2}{n^{1+2\lambda_\alpha}}}\right).$$

Finally, the bound for the last term in (†) follows directly from Assumption 6

$$\mathbb{E}\left(\sup_{\pi\in\Pi_n}\left|\binom{n}{2}^{-1}\sum_{(i,j)\in I_l}(\hat{\xi}_{ij,l}+\hat{\xi}^{\gamma}_{ij,l}+\hat{\xi}^{\nu}_{ij,l})\pi_{ab}(X_i,X_j)\right|\right) = \mathcal{O}\left(\frac{a(1-K^{-1})}{\sqrt{n}}\right).$$

Putting everything together we know that

$$\begin{aligned}
\sqrt{n}\mathbb{E}\left[\sup_{\pi\in\Pi_n}|\hat{W}_n(\pi)-\widetilde{W}_n(\pi)|\right] &= \mathcal{O}\left(\sqrt{VC(\Pi_{ab,n})\frac{a((1-K^{-1})n)^2}{n^{2\lambda_\gamma}}}\right)\\
&\quad+\mathcal{O}\left(\sqrt{VC(\Pi_{ab,n})\frac{a((1-K^{-1})n)^2}{n^{2\lambda_\nu}}}\right)\\
&\quad+\mathcal{O}\left(\sqrt{VC(\Pi_{ab,n})\frac{a((1-K^{-1})n)^2}{n^{2\lambda_\alpha}}}\right)\\
&\quad+\mathcal{O}\left(a(1-K^{-1})\right)+o(1)\\
&= \mathcal{O}\left(a((1-K^{-1})n)\left(1+\sqrt{\frac{VC(\Pi_{ab,n})}{n^{2\min(\lambda_\gamma,\lambda_\nu,\lambda_\alpha)}}}\right)\right).
\end{aligned}$$

■

Proof of Theorem 1: Follows from Lemmas 2, 3 and 6. ■

Proof of Corollary 1: Let Γ^{ab}_{ij} and $\hat{\Gamma}^{ab}_{ij,l}$ be defined as in the Intergenerational mobility example and let

$$\begin{aligned}
K(\pi) &= \mathbb{E}\left[\sum_{(a,b)\in\{0,1\}^2}\Gamma^{ab}_{ij}\pi_{ab}(X_i,X_j)\right],\\
\widetilde{K}_n(\pi) &= \binom{n}{2}^{-1}\sum_{i<j}\left[\sum_{(a,b)\in\{0,1\}^2}\Gamma^{ab}_{ij}\pi_{ab}(X_i,X_j)\right],\\
\hat{K}_n(\pi) &= \binom{n}{2}^{-1}\sum_{l=1}^{L}\sum_{(i,j)\in I_l}\left[\sum_{(a,b)\in\pi}\hat{\Gamma}^{ab}_{ij,l}(Z_i,Z_j,\hat{\gamma}_l,\hat{\nu}_l,\hat{\alpha}_l)\pi_{ab}(X_i,X_j).\right].
\end{aligned}$$

Note also that $W(\pi) = -|K(\pi) - t|$. Hence, we can write the regret as

$$\mathbb{E}\left[sup_{\pi \in \Pi_n} - |K(\pi) - t| + |K(\hat{\pi}) - t|\right] \leq \mathbb{E}\left[\sup_{\pi \in \Pi_n} |K(\pi) - K(\hat{\pi})|\right].$$

The result follows from applying Theorem 1 with W replaced by K. $\blacksquare$

References

Aaberge, R., T. Havnes, and M. Mogstad (2021): "Ranking intersecting distribution functions," *Journal of Applied Econometrics*, 36, 639–662.

Athey, S., J. Tibshirani, and S. Wager (2019): "Generalized random forests," *The Annals of Statistics*, 47, 1148–1178.

Athey, S. and S. Wager (2021): "Policy learning with observational data," *Econometrica*, 89, 133–161.

Atkinson, A. B. et al. (1970): "On the measurement of inequality," *Journal of economic theory*, 2, 244–263.

Belloni, A. and V. Chernozhukov (2011): "l1-penalized quantile regression in high-dimensional sparse models," *MIT Department of Economics Working Paper No. 09-11.*

——— (2013): "Least squares after model selection in high-dimensional sparse models," *Bernoulli*, 19, 521 – 547.

Bhattacharya, D. and P. Dupas (2012): "Inferring welfare maximizing treatment assignment under budget constraints," *Journal of Econometrics*, 167, 168–196.

Brunori, P., P. Hufe, and D. Mahler (2021): "The roots of inequality: Estimating inequality of opportunity from regression trees and forests," *IZA Discussion Paper No. 14689.*

Brunori, P. and G. Neidhöfer (2021): "The evolution of inequality of opportunity in Germany: A machine learning approach," *Review of Income and Wealth*, 67, 900–927.

Brunori, P., F. Palmisano, and V. Peragine (2019a): "Inequality of opportunity in sub-Saharan Africa," *Applied Economics*, 51, 6428–6458.

Brunori, P., V. Peragine, and L. Serlenga (2019b): "Upward and downward bias when measuring inequality of opportunity," *Social Choice and Welfare*, 52, 635–661.

Campbell, F. A., E. P. Pungello, M. Burchinal, K. Kainz, Y. Pan, B. H. Wasik, O. A. Barbarin, J. J. Sparling, and C. T. Ramey (2012): "Adult outcomes as a function of an early childhood educational program: an Abecedarian Project follow-up." *Developmental psychology*, 48, 1033.

Campbell, F. A. and C. T. Ramey (1994): "Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families," *Child development*, 65, 684–698.

Carranza, R. (2022): "Upper and Lower Bound Estimates of Inequality of Opportunity: A Cross-National Comparison for Europe," *Review of Income and Wealth.*

Chamberlain, G. (2011): "1011 Bayesian Aspects of Treatment Choice," in *The Oxford Handbook of Bayesian Econometrics*, Oxford University Press.

Chen, X. (2007): "Large sample sieve estimation of semi-nonparametric models," *Handbook of econometrics*, 6, 5549–5632.

Chen, X. and H. White (1999): "Improved rates and asymptotic normality for nonparametric neural network estimators," *IEEE Transactions on Information Theory*, 45, 682–691.

Chernozhukov, V., J. C. Escanciano, H. Ichimura, W. K. Newey, and J. M. Robins (2022): "Locally robust semiparametric estimation," *Econometrica*, 90, 1501–1535.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014): "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 129, 1553–1623.

Clémençon, S., G. Lugosi, and N. Vayatis (2008): "Ranking and empirical minimization of U-statistics," *The Annals of Statistics*, 36, 844–874.

Dalton, H. (1920): "The measurement of the inequality of incomes," *The Economic Journal*, 30, 348–361.

Dehejia, R. H. (2005): "Program evaluation as a decision problem," *Journal of Econometrics*, 125, 141–173.

Donaldson, D. and J. A. Weymark (1980): "A single-parameter generalization of the Gini indices of inequality," *Journal of economic Theory*, 22, 67–86.

——— (1983): "Ethically flexible Gini indices for income distributions in the continuum," *Journal of Economic Theory*, 29, 353–358.

Escanciano, J. C. and J. R. Terschuur (2023): "Machine Learning Inference on Inequality of Opportunity," *arXiv preprint arXiv:2206.05235.*

Farrell, M. H., T. Liang, and S. Misra (2021): "Deep neural networks for estimation and inference," *Econometrica*, 89, 181–213.

Fleurbaey, M. (1995): "Equal opportunity or equal social outcome?" *Economics & Philosophy*, 11, 25–55.

Fort, M., A. Ichino, and G. Zanella (2020): "Cognitive and noncognitive costs of day care at age 0–2 for children in advantaged families," *Journal of Political Economy*, 128, 158–205.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2020): "Quantifying the life-cycle benefits of an influential early-childhood program," *Journal of Political Economy*, 128, 2502–2541.

Graham, B. S., F. Niu, and J. L. Powell (2021): "Minimax risk and uniform convergence rates for nonparametric dyadic regression," Tech. rep., National Bureau of Economic Research.

Gray-Lobe, G., P. A. Pathak, and C. R. Walters (2023): "The long-term effects of universal preschool in Boston," *The Quarterly Journal of Economics*, 138, 363–411.

Haussler, D. (1995): "Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension," *Journal of Combinatorial Theory, Series A*, 69, 217–232.

Heckman, J., R. Pinto, and P. Savelyev (2013): "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American Economic Review*, 103, 2052–2086.

Heckman, J. J. and L. K. Raut (2016): "Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model," *Journal of econometrics*, 191, 164–175.

Hirano, K. and J. R. Porter (2009): "Asymptotics for statistical treatment rules," *Econometrica*, 77, 1683–1701.

Hufe, P., A. Peichl, P. Schüle, and J. Todorović (2022): "Fairness in Europe: A Multidimensional Comparison," in *CESifo Forum*, München: ifo Institut-Leibniz-Institut für Wirtschaftsforschung an der ..., vol. 23, 45–51.

Ichimura, H. and W. K. Newey (2022): "The influence function of semiparametric estimators," *Quantitative Economics*, 13, 29–61.

Kasy, M. (2016): "Partial identification, distributional preferences, and the welfare ranking of policies," *Review of Economics and Statistics*, 98, 111–131.

Kitagawa, T., M. Nybom, and J. Stuhler (2018): "Measurement error and rank correlations," Tech. rep., cemmap working paper.

Kitagawa, T., S. Sakaguchi, and A. Tetenov (2021): "Constrained classification and policy learning," *arXiv preprint arXiv:2106.12886*.

Kitagawa, T. and A. Tetenov (2018): "Who should be treated? empirical welfare maximization methods for treatment choice," *Econometrica*, 86, 591–616.

——— (2021): "Equality-minded treatment choice," *Journal of Business & Economic Statistics*, 39, 561–574.

Kosorok, M. R. (2008): *Introduction to empirical processes and semiparametric inference.*, Springer.

Kueck, J., Y. Luo, M. Spindler, and Z. Wang (2023): "Estimation and inference of treatment effects with L2-boosting in high-dimensional settings," *Journal of Econometrics*, 234, 714–731.

Leqi, L. and E. H. Kennedy (2021): "Median optimal treatment regimes," *arXiv preprint arXiv:2103.01802*.

Manski, C. F. (2004): "Statistical treatment rules for heterogeneous populations," *Econometrica*, 72, 1221–1246.

Mehran, F. (1976): "Linear measures of income inequality," *Econometrica: Journal of the Econometric Society*, 805–809.

Newey, W. K. (1994): "The asymptotic variance of semiparametric estimators," *Econometrica: Journal of the Econometric Society*, 1349–1382.

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, 846–866.

Rodríguez, L. J. C., G. A. Marrero, J. G. R. Hernández, and P. S. Rojo (2021): "Inequality of Opportunity in Spain," *Hacienda Pública Española/Review of Public Economics*, 237, 153–185.

Roemer, J. E. (1998): *Equality of opportunity*, Harvard University Press.

Sasaki, Y. and T. Ura (2021): "Estimation and inference for policy relevant treatment effects," *Journal of Econometrics*.

Stoye, J. (2009): "Minimax regret treatment choice with finite samples," *Journal of Econometrics*, 151, 70–81.

——— (2012): "Minimax regret treatment choice with covariates or with limited validity of experiments," *Journal of Econometrics*, 166, 138–156.

Stute, W. (1991): "Conditional U-statistics," *The Annals of Probability*, 812–825.

Tetenov, A. (2012): "Statistical treatment choice based on asymmetric minimax regret criteria," *Journal of Econometrics*, 166, 157–165.

Van De Gaer (1993): "Equality of opportunity and investment in human capital," Phd thesis, Katholieke Universiteit Leuven.

Wager, S. and G. Walther (2015): "Adaptive concentration of regression trees, with application to random forests," *arXiv preprint arXiv:1503.06388*.

Wainwright, M. J. (2019): *High-dimensional statistics: A non-asymptotic viewpoint*, vol. 48, Cambridge university press.

Wang, L., Y. Zhou, R. Song, and B. Sherwood (2018): "Quantile-optimal treatment regimes," *Journal of the American Statistical Association*, 113, 1243–1254.

Weymark, J. A. (1981): "Generalized Gini inequality indices," *Mathematical Social Sciences*, 1, 409–430.

Zhang, T. and B. Yu (2005): "Boosting with early stopping: Convergence and consistency," *The Annals of Statistics*, 33, 1538 – 1579.

Zhou, Z., S. Athey, and S. Wager (2023): "Offline multi-action policy learning: Generalization and optimization," *Operations Research*, 71, 148–183.