

# Equality-Minded Treatment Choice

Toru Kitagawa & Aleksey Tetenov

To cite this article: Toru Kitagawa & Aleksey Tetenov (2021) Equality-Minded Treatment Choice, Journal of Business & Economic Statistics, 39:2, 561-574, DOI: [10.1080/07350015.2019.1688664](https://doi.org/10.1080/07350015.2019.1688664)

To link to this article: <https://doi.org/10.1080/07350015.2019.1688664>



© 2019 The Authors. Published with license  
by Taylor & Francis Group, LLC



View supplementary material [↗](#)



Published online: 09 Dec 2019.



Submit your article to this journal [↗](#)



Article views: 1548



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)

## Equality-Minded Treatment Choice

Toru Kitagawa<sup>a</sup> and Aleksey Tetenov<sup>b</sup>

<sup>a</sup>Department of Economics, University College London, London UK, and Cemmap, Institute for Fiscal Studies, London, UK; <sup>b</sup>Institute of Economics and Econometrics, Geneva School of Economics and Management, University of Geneva, Geneva, Switzerland

### ABSTRACT

The goal of many randomized experiments and quasi-experimental studies in economics is to inform policies that aim to raise incomes and reduce economic inequality. A policy maximizing the sum of individual incomes may not be desirable if it magnifies economic inequality and post-treatment redistribution of income is infeasible. This article develops a method to estimate the optimal treatment assignment policy based on observable individual covariates when the policy objective is to maximize an *equality-minded rank-dependent social welfare function*, which puts higher weight on individuals with lower-ranked outcomes. We estimate the optimal policy by maximizing a sample analog of the rank-dependent welfare over a properly constrained set of policies. We show that the average social welfare attained by our estimated policy converges to the maximal attainable welfare at  $n^{-1/2}$  rate uniformly over a large class of data distributions when the propensity score is known. We also show that this rate is minimax optimal. We provide an application of our method using the data from the National JTPA Study. Supplementary materials for this article are available online.

### ARTICLE HISTORY

Received February 2017  
Accepted September 2019

### KEYWORDS

Gini coefficient; Inequality index; Program evaluation; Social welfare; Treatment choice

## 1. Introduction

In causal inference studies analyzing experimental or quasi-experimental data, treatment response generally varies with individual observable characteristics. Learning about such heterogeneity from the data is essential for designing *individualized treatment rules* that assign treatments on the basis of individual observable characteristics. The optimal individualized treatment rule maximizes a social welfare criterion representing the policy maker's preferences over population distributions of post-treatment outcomes. The literature on statistical treatment choice initiated by Manski (2004) emphasizes this perspective of welfare-based empirical policy design and pursues statistically sound ways to estimate the optimal treatment assignment rule.

Research on statistical treatment rules typically focuses on the *additive social welfare* criterion (sometimes called “utilitarian”) defined as the mean of the outcomes in the population, even though welfare economics offers a variety of alternative criteria. The additive social welfare criterion offers analytical and computational convenience because the optimal treatment rule then depends only on the conditional average treatment effect. Empirical researchers studying causal impacts of social programs have stressed the importance of evaluating distributional impacts, which are overlooked when only mean outcomes are considered (e.g., Bitler, Gelbach, and Hoynes 2006). The distributional impact of a policy is especially important when the policy maker is concerned about *economic inequality* in the population.

We study the problem of treatment assignment that aims to maximize a rank-dependent *social welfare function* (SWF), which has the form

$$W \equiv \int Y_i \cdot \omega(\text{Rank}(Y_i)) di, \quad (1)$$

where  $Y_i$  is individual  $i$ 's outcome,  $\text{Rank}(Y_i)$  is the outcome rank of  $i$  from the bottom of the outcome distribution, and  $\omega(\cdot)$  is a nonnegative weight assigned to each rank. The additive SWF is a special case of (1) with constant  $\omega(\cdot)$ . The class of *generalized Gini* SWFs proposed by Mehran (1976) and Weymark (1981) consists of SWFs of the form (1) with *nonincreasing*  $\omega(\cdot)$ . It closely relates to *income inequality indices*, including the widely used Gini index. Blackorby and Donaldson (1978) showed that, given a specification of  $\omega(\cdot)$ , the rank-dependent SWF can be written as a product of the average outcome and one minus a generalized relative index of inequality, for example, Gini. This implies that these SWFs generate a ranking of outcome distributions that is increasing in the average outcome and decreasing in the chosen index of inequality. While inequality measures are predominantly applied to net income, our analysis allows  $Y_i$  to denote other outcome variables of interest, including functions of income, consumption, wealth, or human capital. We will therefore refer to  $Y_i$  simply as “the outcome” in this article.

The goal is to choose a treatment rule  $\delta$  that assigns individuals to one of two treatments  $d \in \{0, 1\}$  depending on their observable pretreatment covariates  $X \in \mathcal{X}$ . This choice is made after experimental data have been collected and analyzed.

**CONTACT** Toru Kitagawa  [t.kitagawa@ucl.ac.uk](mailto:t.kitagawa@ucl.ac.uk) 

 Supplementary materials for this article are available online. Please go to [www.tandfonline.com/UBES](http://www.tandfonline.com/UBES).

© 2019 The Authors. Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We do not consider the problem of optimal experimental design in this article, taking the design as given. We assume that an individual's treatment outcome does not depend on treatments received by others. The policy-maker in our setup can only impact the distribution of outcomes through the choice of a treatment assignment rule and cannot combine it with other redistributive policies.

Finding a policy that maximizes a rank-dependent SWF is a nontrivial problem without a closed-form solution even if the conditional distributions of potential outcomes ( $P(Y_0|X)$  and  $P(Y_1|X)$ ) are known (in a slight abuse of notation typical in the literature,  $Y_i$  will denote individual  $i$ 's realization of random variable  $Y \equiv (1 - D)Y_0 + DY_1$ , whereas  $Y_0$ ,  $Y_1$ , and  $Y_d$  will denote random variables for potential outcomes of treatments 0, 1, and  $d \in \{0, 1\}$ ). Under an additive SWF ( $\int u(Y_i)di$ ) (averaging either outcomes  $Y_i$  or, more generally, their nonlinear transformations  $u(Y_i)$ ), it is optimal to assign for each subgroup the treatment with the highest conditional mean  $E(u(Y_d)|X)$ . In contrast, a rank-dependent SWF is non-decomposable across subgroups, as the ranking of treatment assignment rules for a given covariate value may change depending on the treatment assignment of other subgroups (see Section B in the online supplement).

We show in [Theorem 2.1](#) that an equality-minded rank-dependent SWF is always maximized by a non-randomized treatment rule (assigning the same treatment to all individuals with identical covariates). This result greatly simplifies the space of treatment rules that need to be considered. It also allows us to index treatment rules by their *decision sets*  $G \subset \mathcal{X}$ , denoting all values of the covariates  $\{X \in G\}$  for which treatment 1 is assigned.

Our aim is to estimate from the sample data a treatment assignment rule  $\widehat{G}$  belonging to a constrained (but generally large) *set of feasible policies*  $\mathcal{G}$ , which is a collection of non-randomized treatment rules indexed by their decision sets. Policy makers often face legal, ethical, or political constraints that restrict how individual characteristics can be used to determine treatment assignment. One of the advantages of our framework is that it easily incorporates such exogenous restrictions. Our analytical results also require  $\mathcal{G}$  to satisfy a certain complexity restriction (a finite *VC-dimension*) to prevent overfitting. Kitagawa and Tetenov (2018a) argue that this is not restrictive for many public policy applications and provide rich examples of treatment rule classes that satisfy this complexity restriction.

We propose estimating the treatment rule  $\widehat{G}$  by maximizing a sample analog of  $W(G)$ , the SWF evaluated at the population distribution of  $Y$  that is realized if treatment assignment rule  $G$  is implemented. The general idea of estimating a policy by maximizing an empirical welfare criterion is in line with the method developed by Kitagawa and Tetenov (2018a) for the additive welfare case, hence, we refer to the method proposed in the current article for rank-dependent SWFs as *equality-minded empirical welfare maximization* (EWM). However, the results of Kitagawa and Tetenov (2018a) cannot be directly applied in this setting. We had to construct a suitable sample analog of  $W(G)$  using the whole distribution function and derive new methods to show that this distributional welfare analog (which is not a sample average, as in Kitagawa and Tetenov (2018a)) converges

uniformly over a VC-class of policies. Additionally, we develop a novel proof technique to establish this convergence for unbounded outcome distributions using a weak tail condition, whereas the results in Kitagawa and Tetenov (2018a) for additive welfare require outcomes to be bounded.

We evaluate the statistical performance of  $\widehat{G}$  in terms of *regret*  $E_{P^n} \left[ \sup_{G \in \mathcal{G}} W(G) - W(\widehat{G}) \right]$ , which is the average welfare loss relative to the maximum welfare achievable in  $\mathcal{G}$  with respect to the sampling distribution  $P^n$  of a size  $n$  sample. We derive a non-asymptotic and distribution-free upper bound on regret in terms of the sample size  $n$  and a measure of complexity of  $\mathcal{G}$ , and show that it converges to zero at  $n^{-1/2}$  rate. We also show that this rate is minimax optimal over a minimally constrained class of population distributions, ensuring that no other data-driven treatment rule can lead to a faster welfare loss convergence rate uniformly over the class of distributions.

The remainder of this article is organized as follows. [Section 1.1](#) provides an overview of related literature. [Section 2](#) discusses the properties of equality-minded rank-dependent social welfare functions and their application to the analysis of treatment choice. In [Section 3](#), we introduce the general analytical framework and show the convergence rate properties of the EWM rule for rank-dependent welfare. [Section 4](#) provides extensions of the model that incorporate cost or capacity constraints and that allow the sampled population to be only a subset of the full population on which social welfare is defined. In [Section 5](#), we apply our method to the experimental data from the National JTPA Study. Main proofs are collected in Appendix A. An online supplement contains additional proofs, examples, and extensions.

### 1.1. Related Literature

The analysis of statistical treatment rules was pioneered by Manski (2004), and is a growing area of research in econometrics. Important recent developments can be found in Dehejia (2005), Hirano and Porter (2009), Stoye (2009, 2012), Chamberlain (2011), Bhattacharya and Dupas (2012), Tetenov (2012), Kasy (2016, 2018), Kitagawa and Tetenov (2018a), Mbakop and Tabord-Meehan (2018), and Athey and Wager (2018), among others. All the existing works on treatment choice except for Kasy (2016) posit an additive welfare criterion as the objective function of the policy maker. Motivated by policy concerns about economic inequality, the current article instead analyzes the treatment choice problem for a class of rank-dependent social welfare functions that embody inequality aversion.

The main feature distinguishing the current analysis from the EWM approach for the additive welfare case considered in Kitagawa and Tetenov (2018a) is that the rank-dependent welfare criterion is nondecomposable. Computing the empirical welfare criterion then requires that the whole distribution of outcomes that would be generated by each policy is estimated first, before a nonlinear transformation is applied to this distribution estimate. This problem has not been previously considered in econometrics nor in the machine learning and statistics literatures on empirical risk minimization problems (Vapnik 1998), where the empirical risk criterion always takes

the form of a sample average (with the exception of Wang et al. (2018), who maximized one specific quantile of the outcome distribution). Another novel technical contribution of this article is that we allow outcomes to be unbounded (which is important for analysis of economic outcomes like earnings) with only a weak restriction on the tail of the potential outcome distribution.

Kasy (2016) analyzed treatment choice for a class of social welfare functions including rank-dependent social welfare. Our approach differs from his in several aspects. First, Kasy (2016) considered a linear approximation of the rank-dependent welfare function around a status-quo policy to discuss (partial) identification of a welfare-improving local policy change. We instead focus on a globally optimal policy without invoking approximations. Second, we assume that the welfare criterion is point-identified by the sampling process, while Kasy (2016) focused on partial identification of the welfare criterion and construction of the social planner’s incomplete preference ordering over policies. Third, we study estimation of an optimal policy and examine optimality of the estimator in terms of welfare regret convergence rate, while Kasy (2016) studied asymptotically valid inference on the welfare rankings.

Aaberge, Havnes, and Mogstad (2013) estimated a rank-dependent social welfare function of two policy alternatives: with and without uniform implementation of the treatment. Firpo and Pinto (2016) estimated the impact of uniform implementation of the treatment on measures of inequality, including the Gini coefficient. In contrast, the focus of the current article is on estimating the optimal treatment rule from a large class of individualized assignment rules.

We consider social welfare functions that satisfy the axiom of *anonymity*, that is, social welfare functions that are functionals of the distribution of outcomes after treatment assignment and that are indifferent to reshuffling of the outcomes between individuals. Thus, our objective does not depend on the distribution of individual treatment effects  $P(Y_1 - Y_0)$ , which has also received attention recently in the program evaluation literature (Heckman, Smith, and Clements 1997; Firpo and Ridder 2019; Fan and Park 2010).

Building social welfare functions satisfying the Pigou–Dalton principle of transfers (that a transfer of income from a higher ranked individual to a lower ranked individual that does not change their ranks is always desirable) is one of the central themes in the literature of inequality measurement and welfare economics (see Cowell 1995, 2000; Lambert 2001). Currently, there are two widely used classes of social welfare functions that meet the Pigou–Dalton principle. The first is the class of Atkinson-type SWFs (Atkinson 1970),  $W(F) = \int_0^\infty U(y)dF(y)$ , where  $F(y)$  is the cumulative distribution function (cdf) of the outcome and  $U(\cdot)$  is a concave nondecreasing function. Since the Atkinson-type social welfare function is linear in  $F$ , the EWM approach of Kitagawa and Tetenov (2018a) can be readily applied by defining the outcome observations as  $U(Y)$ . The second class, which is this article’s main focus, is the class of rank-dependent social welfare functions introduced by Mehran (1976), Blackorby and Donaldson (1978), and Weymark (1981) and axiomatized by Yaari (1988). The key axiom of Yaari (1988) that distinguishes rank-dependent social welfare from Atkinson-type social welfare is *invariance under*

*a rank-preserving lump-sum change of incomes at the upper tail*, which means that the preference ordering between two income distributions  $F$  and  $F'$  is invariant to any fixed lump-sum increase (decrease) in income of all those above (below) the  $\tau$ th quantile of  $F$  and  $F'$  for any  $\tau \in (0, 1)$ . On the other hand, the key axiom that characterizes the Atkinson-type social welfare is the independence axiom: the preference ordering between  $F$  and  $F'$  is invariant to any mixing with another common income distribution. As noted in Machina (1982), the rank-dependent social welfare generalizes the Atkinson-type social welfare exactly as rank-dependent expected utility generalizes classical von Neumann–Morgenstern expected utility (Machina 1982; Quiggin 1982) by relaxing the controversial independence axiom. These rich and insightful works in welfare economics have not yet been well linked to econometrics and empirical analysis for policy design. One of the main aims of the current article is to establish a link between these two important literatures.

## 2. Treatment Choice With Equality-Minded Social Welfare Functions

We call a SWF *equality-minded* if it satisfies the *Pigou–Dalton principle of transfers*: a transfer of income from a higher ranked individual to a lower ranked individual is always desirable when it does not change their ranks in the income distribution. The equality-minded SWFs analyzed in this article are the rank-dependent SWFs with decreasing welfare weights (also called generalized Gini SWFs), introduced by Mehran (1976) and Weymark (1981) and axiomatized by Yaari (1988). An equality-minded rank-dependent SWF admits the following representation:

$$W_\Lambda(F) \equiv \int_0^\infty \Lambda(F(y))dy, \tag{2}$$

where  $\Lambda(\cdot) : [0, 1] \rightarrow [0, 1]$  is a nonincreasing, nonnegative function with  $\Lambda(0) = 1$  and  $\Lambda(1) = 0$ . A rank-dependent SWF satisfies the Pigou–Dalton principle of transfers if and only if  $\Lambda(\cdot)$  is convex.

The term *rank-dependent* is due to an equivalent representation of (2) as a weighted sum of outcomes. Given that a convex  $\Lambda(\cdot)$  is almost everywhere differentiable, we can apply integration by parts to equivalently express  $W_\Lambda(F)$  as

$$W_\Lambda(F) = \int_0^1 F^{-1}(\tau)\omega(\tau)d\tau, \tag{3}$$

where  $F^{-1}(\tau) \equiv \inf\{y : F(y) \geq \tau\}$  is the  $\tau$ th quantile of the outcome distribution and  $\omega(\tau) \equiv \frac{d[1-\Lambda(\tau)]}{d\tau}$ . Thus,  $W_\Lambda(F)$  is a weighted average of outcomes  $Y$ , where  $\omega(\tau)$  specifies the *rank-specific* welfare weight assigned to individuals with outcomes at the  $\tau$ th quantile. If the SWF is equality-minded then  $\Lambda(\cdot)$  is convex, hence,  $\omega(\cdot)$  is nonincreasing and assigns larger welfare weights to individuals with lower outcomes.

Throughout the article, we consider equality-minded SWFs satisfying the following assumption:

**Assumption 2.1 (SWF).** The policy-maker’s SWF has representation (2), where  $\Lambda(\cdot) : [0, 1] \rightarrow [0, 1]$  is a nonincreasing, convex function with  $\Lambda(0) = 1$ ,  $\Lambda(1) = 0$ , and its right derivative at the origin is finite:  $|\Lambda'(0)| < \infty$ .

An important family of social welfare functions satisfying Assumption 2.1 is the *extended Gini* family considered in Donaldson and Weymark (1980, 1983) and Aaberge, Havnes, and Mogstad (2013):

$$W_k(F) \equiv \int_0^\infty (1 - F(y))^{k-1} dy = \int_0^\infty \Lambda_k(F(y)) dy = \int_0^1 F^{-1}(\tau) \omega_k(\tau) d\tau, \tag{4}$$

where  $\Lambda_k(\tau) \equiv (1 - \tau)^{k-1}$  and the welfare weight function is  $\omega_k(\tau) \equiv (k - 1)(1 - \tau)^{k-2}$ . Extended Gini social welfare functions are equality-minded for  $k > 2$ . Setting  $k = 2$  yields the additive welfare function  $W_2(F) = \int_0^\infty (1 - F(y)) dy = E(Y)$ , which is not equality-minded.

The standard *Gini social welfare function* (Blackorby and Donaldson 1978; Weymark 1981) corresponds to  $k = 3$  in the extended Gini family, with  $\Lambda_3(\tau) = (1 - \tau)^2$  and welfare weights  $\omega_3(\tau) = 2(1 - \tau)$ . It could also be written as

$$W_{\text{Gini}}(F) = E(Y)(1 - I_{\text{Gini}}(F)), \tag{5}$$

where  $I_{\text{Gini}}(F) = 1 - \frac{\int_0^1 F^{-1}(\tau) \cdot 2(1-\tau) d\tau}{E(Y)}$  is the widely used Gini inequality index.

Assumption 2.1 implies that the rank-specific weight function  $\omega(\cdot)$  defined in (3) does not asymptote at the origin, that is, the welfare weight assigned to the lowest rank is bounded. This restriction rules out SWFs that closely approximate the Rawlsian social welfare function.

We consider the problem of choosing a policy that assigns individuals to one of two treatments  $d \in \{0, 1\}$  to maximize the chosen SWF. A treatment assignment rule  $\delta : \mathcal{X} \rightarrow [0, 1]$  specifies the proportion of individuals with observable pretreatment covariates  $X \in \mathcal{X} \subset \mathbb{R}^{d_x}$  who will be assigned to treatment 1 by the policy-maker. The policy randomly assigns individuals with covariates  $X$  to the two treatments with probabilities  $1 - \delta(X)$  and  $\delta(X)$ . The population distribution of outcomes induced by treatment rule  $\delta$  has cdf

$$F_\delta(y) \equiv \int_{\mathcal{X}} [(1 - \delta(x))F_{Y_0|X=x}(y) + \delta(x)F_{Y_1|X=x}(y)] dP_X(x), \tag{6}$$

where  $Y_0$  and  $Y_1$  denote the potential outcomes of the two treatments with conditional distributions  $F_{Y_0|X}$  and  $F_{Y_1|X}$  given  $X$  and  $P_X$  is the marginal distribution of  $X$ .

If the population distribution of  $(Y_0, Y_1, X)$  were known, the optimal policy maximizing the social welfare function (2) would be

$$\delta^* \in \arg \max_{\delta} W_\Lambda(F_\delta). \tag{7}$$

For the additive welfare function (the mean of  $Y$ ), the welfare maximization problem simplifies to

$$\delta_{\text{util}}^* \in \arg \max_{\delta} \int_{\mathcal{X}} [(1 - \delta(x))E(Y_0|X = x) + \delta(x)E(Y_1|X = x)] dP_X(x). \tag{8}$$

This social welfare function is additive across covariates and depends on the outcome distributions only through their conditional means  $E(Y_d|X)$ . Then the optimal policy is

$$\delta_{\text{util}}^* = 1 \{x \in \mathcal{X} : E(Y_1|X = x) > E(Y_0|X = x)\}.$$

In contrast, the optimal rule for a rank-dependent welfare function (2) depends on the whole conditional distributions of potential outcomes  $F_{Y_0|X}$  and  $F_{Y_1|X}$ , not only on their means. The optimal rule can differ from the one maximizing an additive welfare if there is no first-order stochastic dominance relationship between  $F_{Y_0|X}$  and  $F_{Y_1|X}$  for some covariate values.

Even with the knowledge of the distribution of  $(Y_0, Y_1, X)$ , a simple characterization of the optimal rule does not seem available for rank-dependent SWFs. The following theorem mitigates this complication by substantially reducing the set of candidate treatment rules that need to be considered.

**Theorem 2.1.** If  $W_\Lambda(\cdot)$  satisfies Assumption 2.1, then for every measurable treatment rule  $\delta : \mathcal{X} \rightarrow [0, 1]$ , there exists a non-randomized treatment rule  $\delta_G(x) \equiv 1\{x \in G\}$  for some Borel set  $G \subset \mathcal{X}$ , such that  $W_\Lambda(F_{\delta_G}) \geq W_\Lambda(F_\delta)$ .

If all upper level sets of  $\delta$  belong to a collection  $\mathcal{G}$  of Borel subsets of  $\mathcal{X}$ :

$$\{x : \delta(x) \geq t\} \in \mathcal{G}, \forall t \in \mathbb{R},$$

then there exists  $\delta_G(x)$ ,  $G \in \mathcal{G}$ , such that  $W_\Lambda(F_{\delta_G}) \geq W_\Lambda(F_\delta)$ .

*Proof.* See Appendix A. □

This theorem shows that a treatment assignment rule maximizing equality-minded rank-dependent welfare is non-randomized (assigns all individuals with the same covariates to the same treatment). We can therefore restrict our search for an optimal policy to the set of non-randomized rules that can be succinctly characterized by their *decision sets*  $G \subset \mathcal{X}$ . Decision set  $G$  determines the group of individuals  $\{X \in G\}$  to whom treatment 1 is assigned. With abuse of notation, we denote the welfare of a non-randomized treatment rule with decision set  $G$  by  $W_\Lambda(G)$ , suppressing the cumulative distribution function in its argument,

$$W_\Lambda(G) \equiv W_\Lambda(F_G),$$

$$F_G(y) \equiv \int_{\mathcal{X}} [F_{Y_0|X=x}(y)1\{x \notin G\} + F_{Y_1|X=x}(y)1\{x \in G\}] dP_X(x). \tag{9}$$

Our goal is to estimate from the sample data a treatment assignment rule that attains the maximum level of social welfare  $\sup_{G \in \mathcal{G}} W_\Lambda(G)$  over the *set of feasible policies*  $\mathcal{G} \equiv \{G \subset \mathcal{X}\}$ , which

is a collection of nonrandomized treatment rules (subsets of the covariate space  $\mathcal{X}$ ). An important feature of our empirical welfare maximization approach is that the complexity of  $\mathcal{G}$  is constrained by a finite *Vapnik-Cervonenkis (VC) dimension* (defined in Appendix A).

**Assumption 2.2 (VC).** The class of decision sets  $\mathcal{G}$  has a finite VC-dimension  $\nu < \infty$ .

The VC-dimension is a restriction on the complexity of the set of feasible policies. Without it, maximizing a sample analog of  $W_\Lambda(G)$  over  $G$  can lead to arbitrarily complicated policies (overfitting) and prevent us from learning the optimal policy on the basis of a finite number of observations. It does not require  $\mathcal{G}$  to be finite and allows

for very large classes of treatment rules. For example, a class of treatment rules defined by a linear equation in functions of  $x$ ,  $\mathcal{G} \equiv \left\{ G = \{x : \sum_{j=1}^m \beta_j f_j(x) \geq 0\}, \beta \in \mathbb{R}^m \right\}$  has a finite VC-dimension. See Kitagawa and Tetenov (2018a) for other examples of classes  $\mathcal{G}$  that satisfy Assumption 2.2. An example of  $\mathcal{G}$  that does not satisfy Assumption 2.2 is the class of all monotone treatment rules  $\mathcal{G} \equiv \left\{ G = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq f(x_1)\}, f : \mathbb{R} \rightarrow \mathbb{R} \text{ increasing} \right\}$  considered by Mbakop and Tabord-Meehan (2018) in the additive welfare case.

### 3. EWM for Equality-Minded Welfare

We proceed to propose our method of estimating the treatment rule in finite samples and analyze its properties.

The population from which the sample is drawn is characterized by  $P$ , a joint distribution of  $(Y_0, Y_1, D_i, X_i)$ , where  $X_i \in \mathcal{X} \subset \mathbb{R}^d$  refers to observable pretreatment covariates of individual  $i$ ,  $Y_0, Y_1 \in \mathbb{R}_+$  are the potential outcomes of treatments 0 and 1, and  $D_i \in \{0, 1\}$  is a binary indicator of the individual's experimental treatment. The observed experimental outcome is  $Y_i = (1 - D_i)Y_0 + D_iY_1$ .

The data are a size  $n$  random sample from  $P$  of observations  $Z_i = (Y_i, D_i, X_i)$ . Based on this data, the policy-maker has to choose a conditional treatment rule  $G \in \mathcal{G}$  that determines whether individuals with covariates  $X$  in the target population will be assigned to treatment 0 or to treatment 1. The following are our maintained assumptions about the class  $\mathcal{P}$  of population distributions of  $(Y_0, Y_1, D, X)$ :

*Assumption 3.1.*

(UCF) *Unconfoundedness*:  $(Y_0, Y_1) \perp D|X$ .

(TC) *Tail condition*: There exists  $\Upsilon < \infty$  such that for all  $P \in \mathcal{P}$

$$\int_0^\infty \sqrt{P(Y_d > y)} dy \leq \Upsilon. \tag{10}$$

(SO) *Strict overlap*: There exist  $\kappa \in (0, 1/2]$  such that the propensity score satisfies  $e(x) \in [\kappa, 1 - \kappa]$  for all  $x \in \mathcal{X}$ .

These assumptions generally hold if the data come from an experiment with randomized treatment assignment. In observational studies, on the other hand, *unconfoundedness* rules out situations in which the observed treatment assignments depend on subjects' unobserved characteristics that can be associated with their potential outcomes. *Strict overlap* can also be violated in an observational study if only one of the treatments is assigned in the sampling process for some covariate values. We do not constrain any feature of the joint distribution of  $(Y_0, Y_1, X)$  except that the distributions of  $Y_0$  and  $Y_1$  satisfy the tail condition (TC). A sufficient condition for (TC) is that

$$\sup_{P \in \mathcal{P}, d \in \{0,1\}} E[Y_d^{2+\Delta}] < \infty \tag{11}$$

for some  $\Delta > 0$ . The outcome variable and the covariates can be discrete, continuous, or their combination, and the support of  $X$  does not have to be bounded.

Throughout the main text, we maintain the assumption that the propensity score  $e(X) \equiv P(D = 1|X)$  is known, as is

usually the case in experimental data. Section C in the online supplement extends the analysis to observational data for which the propensity score is unknown and needs to be estimated.

We estimate the treatment rule by maximizing a sample analog of the population SWF. The *equality-minded EWM treatment rule*  $\widehat{G}$  maximizes a sample analog  $\widehat{W}_\Lambda(G)$  of the welfare criterion over the set of feasible rules  $G \in \mathcal{G}$ . The unknown outcome distribution  $F_G$  induced by treatment rule  $G$  in (9) could be estimated by

$$\widehat{F}_G(y) \equiv 1 - \frac{1}{n} \sum_{i=1}^n \left[ \frac{D_i}{e(X_i)} \cdot 1\{X_i \in G\} + \frac{1 - D_i}{1 - e(X_i)} \cdot 1\{X_i \notin G\} \right] \cdot 1\{Y_i > y\}. \tag{12}$$

Under Assumption 3.1 (UCF),  $\widehat{F}_G(y)$  is an unbiased estimator of  $F_G(y)$ .

The sample analog of welfare (Equation (2)) is defined as

$$\widehat{W}_\Lambda(G) \equiv \int_0^\infty \Lambda(\widehat{F}_G(y) \vee 0) dy \tag{13}$$

and the equality-minded EWM treatment rule is then

$$\widehat{G} \in \arg \max_{G \in \mathcal{G}} \widehat{W}_\Lambda(G). \tag{14}$$

The maximum ( $\vee$ ) of  $\widehat{F}_G(y)$  and 0 is taken in (13) because  $\widehat{F}_G(y)$  may take values smaller than 0, for which  $\Lambda(\cdot)$  is not defined. The summands in (12) are nonnegative, so  $\widehat{F}_G(y) \leq 1$  for all  $y$ .  $\widehat{F}_G(y) = 1$  for all  $y \geq \max_{1 \leq i \leq n} Y_i$ .  $\widehat{F}_G$  may not be a proper cdf because  $\lim_{y \rightarrow -\infty} \widehat{F}_G(y) = 1 - \frac{1}{n} \sum_{i=1}^n \left[ \frac{D_i}{e(X_i)} \cdot 1\{X_i \in G\} + \frac{1 - D_i}{1 - e(X_i)} \cdot 1\{X_i \notin G\} \right]$  may be either below or above zero in finite samples.

We also consider properties of the *normalized equality-minded EWM rule*

$$\widehat{G}^R \in \arg \max_{G \in \mathcal{G}} \widehat{W}_\Lambda^R(G), \text{ where } \widehat{W}_\Lambda^R(G) \equiv \int_0^\infty \Lambda(\widehat{F}_G^R(y)) dy, \tag{15}$$

using a normalized cdf sample analog

$$\widehat{F}_G^R(y) \equiv \begin{cases} 1\{y \geq \min_{1 \leq i \leq n} Y_i\} & \text{if } \widehat{F}_G(-\infty) = 1, \\ 1 - \frac{1 - \widehat{F}_G(y)}{1 - \widehat{F}_G(-\infty)} & \text{if } \widehat{F}_G(-\infty) < 1, \end{cases} \tag{16}$$

which always yields a proper cdf. The ranking of treatment rules by the normalized criterion  $\widehat{W}_\Lambda^R(G)$  is invariant to positive affine transformations of outcomes  $Y$ , whereas the ranking by  $\widehat{W}_\Lambda(G)$  is not.

Note that

$$1 - \widehat{F}_G^R(y) = \frac{1}{n} \sum_{i=1}^n \left[ \frac{\frac{D_i}{e(X_i)} \cdot 1\{X_i \in G\} + \frac{1 - D_i}{1 - e(X_i)} \cdot 1\{X_i \notin G\}}{\frac{1}{n} \sum_{i=1}^n \left( \frac{D_i}{e(X_i)} \cdot 1\{X_i \in G\} + \frac{1 - D_i}{1 - e(X_i)} \cdot 1\{X_i \notin G\} \right)} \right] \cdot 1\{Y_i > y\}.$$

This is similar to the idea of normalizing propensity score weights recommended for the estimation of the average treatment effect (Imbens 2004).

### 3.1. Rate Optimality of EWM

The next theorem derives a uniform upper bound of the average welfare loss of the EWM rule.

**Theorem 3.1.** Under Assumptions 2.1 and 2.2, the average welfare loss of treatment rules  $\widehat{G}$  and  $\widehat{G}^R$  satisfies

$$\sup_{P \in \mathcal{P}} E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_\Lambda(G) - W_\Lambda(\widehat{G}) \right] \leq C |\Lambda'(0)| \frac{\Upsilon}{\kappa} \sqrt{\frac{\nu}{n}} \quad (17)$$

for all  $n > 1$ , and

$$\begin{aligned} \sup_{P \in \mathcal{P}} E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_\Lambda(G) - W_\Lambda(\widehat{G}^R) \right] \\ \leq |\Lambda'(0)| \frac{\Upsilon}{\kappa} \left( C_1^R \sqrt{\frac{\nu}{n}} + 4ne^{-C_2^R \kappa^2 n} \right) \end{aligned} \quad (18)$$

for all  $n > C_3^R \left(\frac{1-\kappa}{\kappa}\right)^2 \nu$ , where  $\mathcal{P}$  is the class of all distributions satisfying Assumption 3.1 and  $C, C_1^R, C_2^R, C_3^R > 0$  are universal constants.

*Proof.* The proof of (17) is in Appendix A. The proof of (18) is in the online supplement.  $\square$

This theorem shows that for a large class of data generating processes characterized by Assumption 3.1, the welfare of the equality-minded EWM rule is guaranteed to converge to the maximal attainable welfare no slower than at  $n^{-1/2}$  rate (the second term in bound (18) is of a lower order). The uniform convergence rate of  $n^{-1/2}$  coincides with that of the EWM rule for the additive welfare shown in Theorem 2.1 of Kitagawa and Tetenov (2018a). This is a nontrivial result, given that the rank-dependent welfare function depends on the whole conditional distributions of the potential outcomes given covariates, rather than only on their conditional means, as is the case for the additive welfare criterion.

The next theorem provides a universal lower bound for the worst-case average welfare loss of any treatment rule.

**Theorem 3.2.** Suppose that Assumptions 2.1 and 2.2 hold with  $\nu \geq 2$ , then for any non-randomized treatment choice rule  $\widehat{G}$  that is a function of the sample, and for any  $\tau^* \in (0, 1]$  at which  $\Lambda(\cdot)$  is differentiable, it holds

$$\begin{aligned} \sup_{P \in \mathcal{P}} E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_\Lambda(G) - W_\Lambda(\widehat{G}) \right] \\ \geq \frac{e^{-4}}{2} \Upsilon |\Lambda'(\tau^*)| \sqrt{\tau^*} \sqrt{\frac{\nu-1}{n}} \text{ for all } n \geq 4 \frac{\nu-1}{\tau^*}, \end{aligned} \quad (19)$$

where  $\mathcal{P}$  is the class of all distributions satisfying Assumption 3.1.

*Proof.* See the online supplement.  $\square$

Since  $\Lambda(\cdot)$  is convex and  $|\Lambda'(0)| > 0$ , there also exists some  $\tau^* > 0$  for which  $|\Lambda'(\tau^*)| > 0$ . Hence, the bound (19) is always positive for some  $\tau^* > 0$ . A comparison of the lower bound of this theorem with the welfare loss upper bound of the EWM rule obtained in Theorem 3.1 shows that the EWM rule is *minimax*

*rate optimal* over the class of data generating processes satisfying Assumption 3.1. We can therefore claim that in the absence of any additional restrictions other than Assumption 3.1, no other data-driven procedure for obtaining a nonrandomized rule can outperform the EWM rule in terms of the uniform convergence rate over  $\mathcal{P}$ . This optimality claim is analogous to that of the EWM rule for the additive welfare case (Kitagawa and Tetenov 2018a, Theorems 2.1 and 2.2), and the minimax optimal rate attained by the equality-minded EWM rule is the same as the optimal rate in the additive welfare case. It is remarkable to see that even in the absence of any analytical characterization of the true optimal assignment rule in terms of the population distribution of  $(Y_0, Y_1, X)$ , maximizing the empirical welfare leads to a policy that, if implemented, is guaranteed to reach the maximum attainable social welfare at the minimax optimal rate.

It is also worth noting that the VC-dimension of  $\mathcal{G}$  appears in the same order both in the upper and lower bound expressions of Theorems 3.1 and 3.2. Since these bounds are non-asymptotic, we can let  $\nu$  increase with the sample size, and we can still conclude the minimax rate optimality of the equality-minded EWM rule. This insight is similar to the EWM rule for the additive welfare case (Kitagawa and Tetenov 2018a, Remark 2.6).

## 4. Extensions

### 4.1. Social Welfare is Defined on a Population Larger than the Sampled Population

One of the distinguishing features of rank-dependent social welfare is that it is not additive over subpopulations (see Section B in the online supplement for an illustration). If the subpopulation for which the policy intervention takes place (e.g., unemployed workers) is only a subset of the whole population on which the rank-dependent SWF is defined (e.g., the population of a country), it is important to explicitly take into account the outcome distribution for the rest of the population in estimating the optimal assignment rule.

Suppose that the social welfare function is defined on a population with distribution  $J$  that is a mixture of two subpopulations with distributions  $F$  and  $H$ :

$$J = \eta F + (1 - \eta)H, \quad \eta \in (0, 1). \quad (20)$$

Let  $F$  be the outcome distribution on the targeted subpopulation from which the experimental data are sampled and on which the estimated treatment policy is to be implemented. Let  $H$  be the outcome distribution for the rest of the population (excluded from the sampling process and unaffected by the chosen treatment assignment rule). The mixture weight  $\eta$  represents the size of subpopulation  $F$ . For simplicity, we assume that  $\eta$  and  $H$  are known to the social planner. We also assume that the outcome distribution  $H$  is invariant to the treatment assignment policy applied to subpopulation  $F$ , for example, there are no spillover or general equilibrium effect across  $F$  and  $H$ .

Implementing treatment assignment rule  $\{X \in G\}$  on subpopulation  $F$  leads to full population welfare equal to

$$W_\Lambda(J_G) \equiv \int_0^\infty \Lambda(\eta F_G(y) + (1 - \eta)H(y)) dy,$$

where  $F_G(\cdot)$  is the cdf defined in (9). The empirical welfare maximization method in this case consists of maximizing a sample analog of  $W_\Lambda(J_G)$ ,

$$\widehat{G} \in \arg \max_{G \in \mathcal{G}} W_\Lambda(\eta \widehat{F}_G + (1 - \eta)H),$$

where  $\widehat{F}_G$  is defined in (12).

The uniform convergence proof of [Theorem 3.1](#) can be easily extended to this case, the only change being the proportionality of the bound to  $\eta$ .

**Corollary 4.1.** Under [Assumptions 2.1, 2.2, and 3.1](#),

$$\sup_{P \in \mathcal{P}} E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_\Lambda(J_G) - W_\Lambda(J_{\widehat{G}}) \right] \leq \eta C |\Lambda'(0)| \frac{\Upsilon}{\kappa} \sqrt{\frac{v}{n}}, \quad (21)$$

where  $C > 0$  is a universal constant defined in [Theorem 3.1](#).

### 4.2. Cost of Treatment

In the preceding sections, we did not take into account the cost of treatment even though cost differences among treatments are often an important consideration in practice. In this section, we discuss how to take the cost of treatment into account in the estimation of welfare maximizing treatment assignment policies.

Let  $0 \leq c(x) < \infty$ ,  $x \in \mathcal{X}$ , be the cost of treatment 1 for a subject whose observable characteristics are  $x$ . We assume that treatment 0 is cost-free and  $c(\cdot)$  is known. For the additive social welfare function, we can easily incorporate treatment costs into the EWM criterion by subtracting the per-capita cost of treatment  $C(G) \equiv \int_G c(x) dP_X(x)$  from  $\widehat{W}(G)$ . The additive social welfare criterion depends only on the average treatment cost, it is invariant to assumptions about *who pays the cost*. For rank-dependent social welfare this invariance does not hold, hence, we have to be explicit about who bears the cost in the construction of the social welfare criterion. We illustrate this using two cost allocation scenarios.

In the first scenario, assume that the outcome variable is income and the cost of treatment is *self-financed* by each recipient of the treatment. Specifically, the income of individuals assigned to treatment 1 (individuals with  $X \in G$ ) will be reduced by the full cost of their own treatment  $c(X)$ . The transformed potential outcomes in this scenario are  $\tilde{Y}_{1i} \equiv Y_{1i} + \bar{c} - c(X_i)$  and  $\tilde{Y}_{0i} \equiv Y_{0i} + \bar{c}$ . We add the constant  $\bar{c} \equiv \sup_{x \in \mathcal{X}} c(x) < \infty$  to all outcomes to keep them nonnegative in line with [Assumption 3.1](#) (TC). The welfare ranking of policies is unchanged when a constant is added uniformly to all outcomes.

The rank-dependent SWF of policy  $G$  with self-financed treatment cost is

$$\begin{aligned} W_\Lambda^{sf}(G) &\equiv \int_0^\infty \Lambda(F_G^{sf}(y)) dy, \\ F_G^{sf}(y) &\equiv \int_{\mathcal{X}} \left[ F_{\tilde{Y}_{0|X=x}}(y) 1\{x \notin G\} + F_{\tilde{Y}_{1|X=x}}(y) 1\{x \in G\} \right] dP_X(x), \end{aligned} \quad (22)$$

where  $F_{\tilde{Y}_{0|X=x}(\cdot)}$  and  $F_{\tilde{Y}_{1|X=x}(\cdot)}$  are the cdfs of the transformed potential outcomes. An empirical analog for  $W_\Lambda^{sf}(G)$  can be obtained by replacing  $\widehat{F}_G(y)$  in (13) by

$$\widehat{F}_G(y) \equiv 1 - \frac{1}{n} \sum_{i=1}^n \left[ \frac{D_i}{e(X_i)} \cdot 1\{X_i \in G\} + \frac{1 - D_i}{1 - e(X_i)} \cdot 1\{X_i \notin G\} \right] \cdot 1\{\tilde{Y}_i > y\}, \quad (23)$$

where  $\tilde{Y}_i \equiv Y_i + \bar{c} - D_i \cdot c(X_i)$ . Since this modification does not affect the validity of [Assumption 3.1](#), the EWM rule with self-financed treatment cost attains the uniform welfare loss upper bounds of [Theorem 3.1](#) with  $\Upsilon + \bar{c}$  in place of  $\Upsilon$ .

In the second scenario, suppose that the treatment cost is financed by all of the population members equally via a lump-sum transfer. The average per-capita treatment cost  $C(G)$  is subtracted from every individual's income regardless of their covariates and assigned treatment. Using representation (3), the rank-dependent SWF with equal lump-sum treatment costs can be expressed as

$$W_\Lambda^{ls}(G) \equiv \int_0^1 [F_G^{-1}(\tau) + \bar{c} - C(G)] \omega(\tau) d\tau = W(G) + \bar{c} - C(G), \quad (24)$$

using the fact that  $\int_0^1 \omega(\tau) d\tau = \Lambda(0) - \Lambda(1) = 1$  and adding  $\bar{c}$  to ensure nonnegative outcomes. Per-capita treatment cost of policy  $G$  could be estimated using its sample analog  $\widehat{C}(G) \equiv \frac{1}{n} \sum_{i=1}^n c(X_i) \cdot 1\{X_i \in G\}$  and the EWM rule is obtained by maximizing  $\widehat{W}_\Lambda^{ls}(G) \equiv \widehat{W}_\Lambda(G) + \bar{c} - \widehat{C}(G)$  over  $G \in \mathcal{G}$ .

In this article, we do not consider the joint optimization of the treatment assignment and cost allocation. However, the comparison of  $W_\Lambda^{sf}(G)$  and  $W_\Lambda^{ls}(G)$  shows that the allocation of treatment costs across the population can be used as an additional vehicle of policy intervention to increase a rank-dependent SWF.

### 4.3. Capacity-Constrained Treatment

Another practical concern ruled out in the preceding sections is a capacity constraint limiting the proportion of population that can be assigned to treatment. Suppose that the proportion of the target population that could receive treatment 1 cannot exceed  $K \in (0, 1)$ . If  $P_X$  is unknown, then policies that seem to satisfy the capacity constraint based on the sample estimates of  $P_X(G)$  may not actually satisfy it in the population. The analysis of the welfare loss needs to take into account what happens if the proposed policy is infeasible. For tractability, we continue to restrict attention only to non-randomized treatment rules (the result in [Theorem 2.1](#) need not hold with a capacity constraint).

For the additive welfare case, Kitagawa and Tetenov (2018a) proposed a capacity-constrained EWM procedure assuming that if a proposed treatment rule  $G$  violates the capacity constraint ( $P_X(G) > K$ ) then the scarce treatment is randomly rationed to a fraction  $\frac{K}{P_X(G)}$  of individuals with  $X \in G$  independently of  $(Y_0, Y_1, X)$ . This random rationing approach can be straightforwardly extended to the EWM for the rank-dependent social welfare.

With the capacity constraint and random rationing, the cdf of outcomes generated by policy  $G$  can be written as

$$F_G^K(y) = \int_{\mathcal{X}} \left[ F_{Y_0|X=x}(y) + \min \left\{ 1, \frac{K}{P_X(G)} \right\} \times (F_{Y_1|X=x}(y) - F_{Y_0|X=x}(y)) \cdot 1\{x \in G\} \right] dP_X(x). \quad (25)$$



Hence, the social welfare under the capacity constraint and random rationing is  $W_{\Lambda}^K(G) \equiv \int_0^{\infty} \Lambda(F_G^K(y))dy$ . Its sample analog can be constructed by replacing  $\widehat{F}_G(y)$  in (13) with

$$\widehat{F}_G^K(y) \equiv 1 - \frac{1}{n} \sum_{i=1}^n \left[ \frac{1-D_i}{1-e(X_i)} + \min \left\{ 1, \frac{K}{P_{X,n}(G)} \right\} \right] \cdot 1\{Y_i > y\}, \tag{26}$$

where  $P_{X,n}(G) \equiv \frac{1}{n} \sum_{i=1}^n 1\{X_i \in G\}$  is a sample analog of  $P_X(G)$ .

**Theorem 4.1** establishes a finite sample bound for the capacity-constrained equality-minded EWM rule similar to the bound in **Theorem 3.1**.

*Theorem 4.1.* Under **Assumptions 2.1, 2.2, and 3.1**,

$$\begin{aligned} & \sup_{P \in \mathcal{P}} E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_{\Lambda}^K(G) - W_{\Lambda}^K(\widehat{G}^K) \right] \\ & \leq \left( \frac{C_{K1}}{K} + C_{K2} \right) |\Lambda'(0)| \frac{\Upsilon}{\kappa} \sqrt{\frac{v}{n}}, \end{aligned} \tag{27}$$

where  $C_{K1}, C_{K2} > 0$  are universal constants.

*Proof.* See the online supplement. □

### 5. Empirical Illustration

To illustrate equality-minded empirical treatment choice, we apply our method to the experimental data from the National Job Training Partnership Act (JTPA) Study. A detailed description of the study and an assessment of average program effects for five large subgroups of the target population is found in Bloom et al. (1997). The study randomized whether applicants were eligible to receive a mix of training, job-search assistance, and other services for a period of 18 months. It collected background information about the applicants prior to random assignment, as well as administrative and survey data on applicants' earnings in the 30-month period following assignment. Our sample consists of 9223 observations with available data on years of education and preprogram earnings from the sample of adults (22 years and older) used in the original evaluation of the program and in subsequent studies (Bloom et al. 1997; Heckman, Ichimura, and Todd 1997; Abadie, Angrist, and Imbens 2002). The probability of being assigned to the treatment was two-thirds in this sample.

For this illustration, total individual earnings in the 30-month period following program assignment serve as the measure of income. We use three social welfare functions from the extended Gini family (4) with parameters  $k \in \{2, 3, 6\}$ .  $k = 2$  corresponds to the additive social welfare, which is not equality-minded.  $k = 3$  corresponds to the standard Gini SWF with welfare weights  $\omega_3(\tau) = 2(1 - \tau)$  and  $k = 6$  corresponds to an extended Gini SWF with welfare weights  $\omega_6(\tau) = 5(1 - \tau)^4$ , which places even greater weight on low-ranked outcomes.

For simplicity, we consider only the distribution of earnings in the population sampled for the experiment in the social welfare function. This embodies concerns about inequality *within* the study population (JTPA-eligible economically disadvantaged adults). In practice, policy makers are more

likely to be concerned with inequality in the overall population, which also includes individuals outside of the experiment's sampling frame. Then the social welfare function should be evaluated on the income distribution of the whole population of interest.

Pretreatment variables on which we consider conditioning treatment assignment are the individual's years of education and earnings in the year prior to assignment. We do not use race, sex, or age to condition treatment assignment. Though treatment effects may vary with these characteristics, using them to condition treatment assignment is often socially unacceptable and illegal. Education and earnings are verifiable characteristics, which is also important for conditioning treatment assignment. The performance of treatment rules that condition on unverifiable characteristics is hard to evaluate if individuals change their self-reported characteristics to obtain their desired treatment assignment (either in the experiment or when the policy is implemented).

**Table 1** compares empirical estimates of social welfare measures (representative income) for a few alternative treatment rules. First, we consider simple treatment rules that either assign no one or everyone to treatment. Second, we consider empirically optimal rules from the class of *quadrant treatment rules*:

$$\begin{aligned} \mathcal{G}_Q \equiv & \{ \{x : s_1(\text{education} - t_1) > 0 \\ & \& s_2(\text{prior earnings} - t_2) > 0\}, s_1, s_2 \in \{-1, 0, 1\}, \\ & t_1, t_2 \in \mathbb{R} \}. \end{aligned} \tag{28}$$

This class of treatment eligibility rules is easily implementable and is often used in practice. To be assigned to treatment according to such rules, an individual's education and preprogram earnings both have to be above (or below) some specific thresholds. Third, we consider empirically optimal rules from the class of *linear treatment rules*:

$$\begin{aligned} \mathcal{G}_{LES} \equiv & \{ \{x : \beta_0 + \beta_1 \cdot \text{education} + \beta_2 \cdot \text{prior earnings} > 0\}, \\ & \beta_0, \beta_1, \beta_2 \in \mathbb{R} \}. \end{aligned} \tag{29}$$

**Table 1.** Estimated representative income under alternative treatment rules that condition on education and preprogram earnings.

| Treatment rule   | Representative income |              |              | Proportion to be treated |
|--|-----------------------|--------------|--------------|--------------------------|
|  | Average (k = 2)       | Gini (k = 3) | Gini (k = 6) |                          |
| Treat no one   | \$15,311              | \$6769       | \$1561       | 0                        |
| Treat everyone   | \$16,487              | \$7423       | \$1786       | 1                        |
| <i>Quadrant class conditioning on years of education and preprogram earnings</i> |                       |              |              |                          |
| Maximize average income  | \$16,646              | \$7490       | \$1807       | 95%                      |
| Maximize standard Gini SWF (k = 3)   | \$16,462              | \$7522       | \$1828       | 80%                      |
| Maximize extended Gini SWF (k = 6)   | \$16,153              | \$7388       | \$1835       | 52%                      |
| <i>Linear class conditioning on years of education and preprogram earnings</i>   |                       |              |              |                          |
| Maximize average income  | \$16,670              | \$7528       | \$1820       | 96%                      |
| Maximize standard Gini SWF (k = 3)   | \$16,489              | \$7537       | \$1839       | 78%                      |
| Maximize extended Gini SWF (k = 6)   | \$16,154              | \$7405       | \$1852       | 51%                      |

The first column in Table 1 displays the estimated average income under each treatment rule. The second column shows the standard Gini social welfare, expressed in terms of the representative income of the policy (the income distribution generated by the policy is valued as much as an equal income distribution with the representative income). The third column shows the representative income under an extended Gini SWF with  $k = 6$ . The fourth column lists the proportion of the target population assigned to treatment by each policy.

Figure 1 compares the quadrant treatment rules maximizing the average income, the standard Gini SWF, and an extended Gini SWF ( $k = 6$ ). Figure 2 compares the linear treatment rules maximizing the same three criteria. The size of black dots shows the number of individuals with different covariate

values. Many individuals would be assigned to treatment by treatment rules maximizing any of the considered welfare functions, but there are also notable differences. Treatment rules maximizing the standard Gini SWF target a smaller proportion of the population, focusing on individuals with lowest preprogram earnings. Treatment rules maximizing the more equality-minded extended Gini SWF assign even fewer individuals to treatment. The estimated treatment rules change discontinuously with the Gini parameter  $k$ . For example, the linear treatment rule maximizing the additive SWF ( $k = 2$ ) is also optimal for a range of other Gini parameters ( $k = 2.25, 2.5, 2.75$ ), whereas the linear rule maximizing the extended Gini SWF ( $k = 6$ ) is also optimal for larger parameters ( $k = 7, 8, 9$ ).

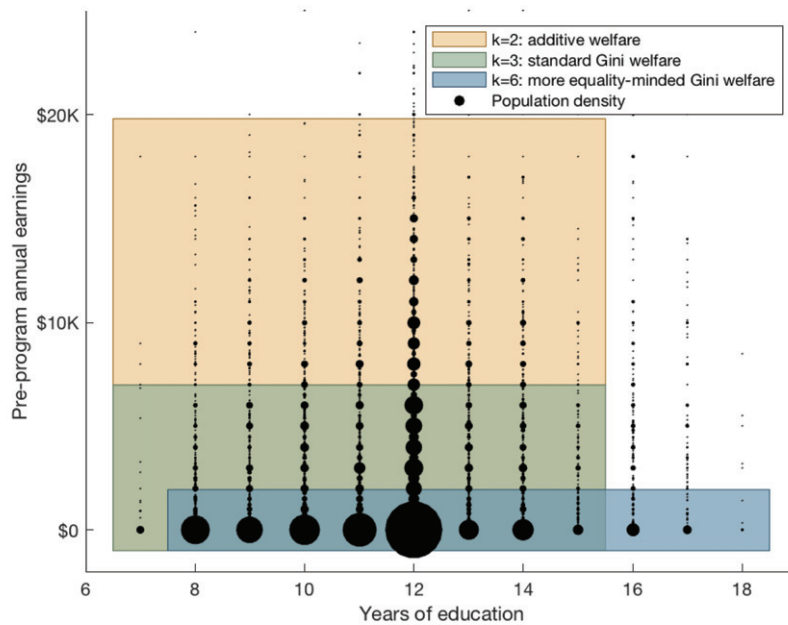


Figure 1. Treatment rules from the quadrant class that maximize welfare functions from the extended Gini family (including the additive welfare).

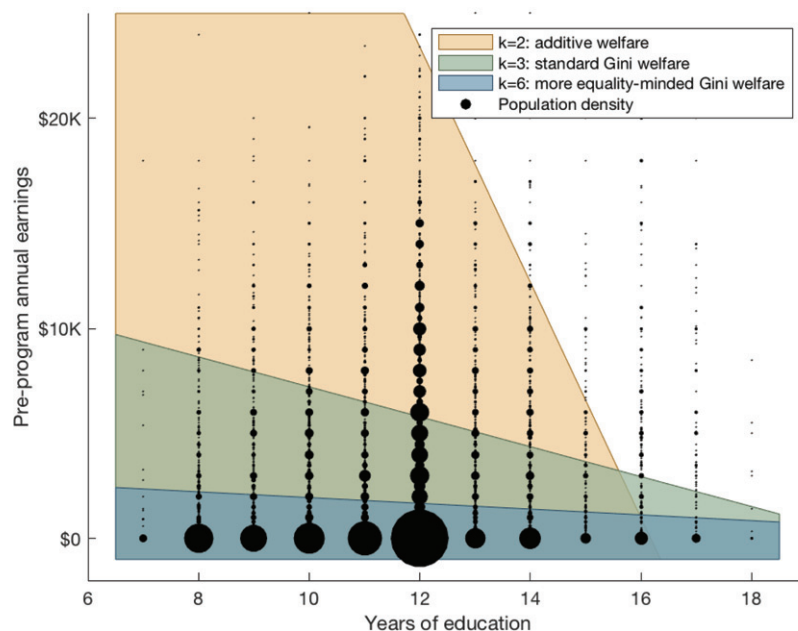


Figure 2. Treatment rules from the linear class that maximize welfare functions from the extended Gini family (including the additive welfare).

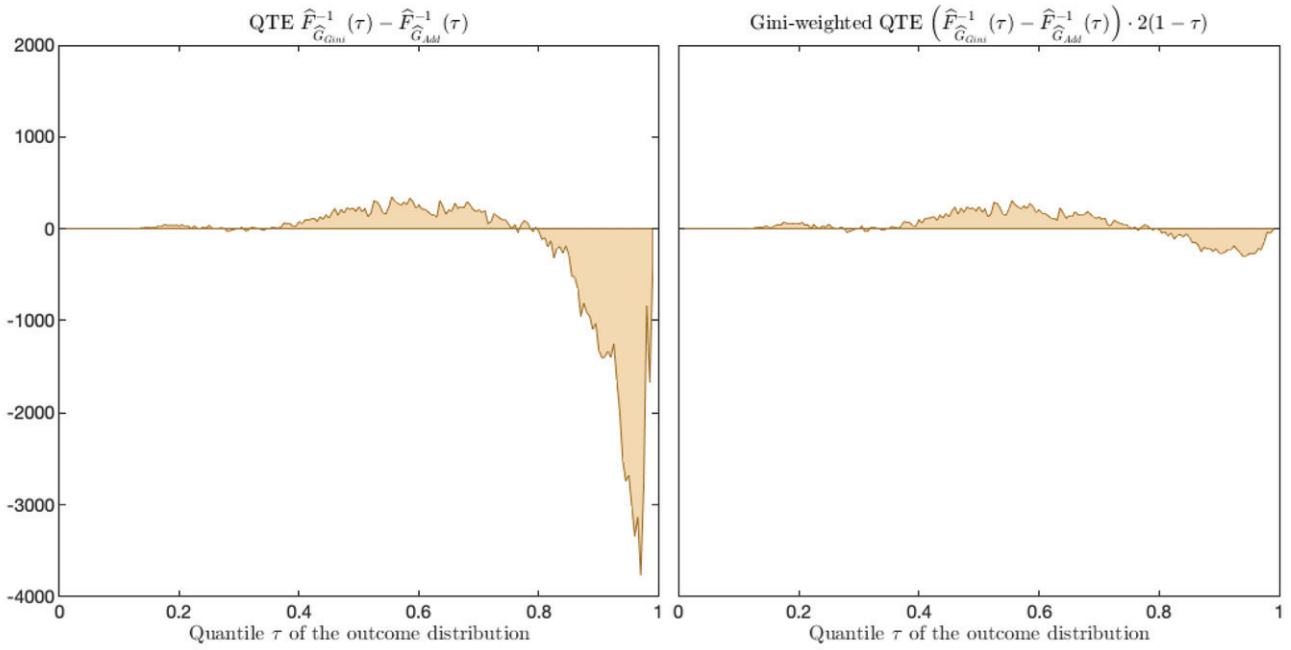


Figure 3. Quantile treatment effects of the Gini welfare maximizing rule compared to the additive welfare maximizing rule. Treatment rules from the quadrant class.

Figure 3 explores the trade-off between treatment rules maximizing different social welfare functions. We compute the income distributions generated by  $\hat{G}_{Add}$ , the quadrant treatment rule maximizing average income, and by  $\hat{G}_{Gini}$ , the quadrant rule maximizing the Gini SWF. The left panel displays the difference between the income distributions generated by these treatment rules at each quantile:  $\hat{F}_{\hat{G}_{Gini}}^{-1}(\tau) - \hat{F}_{\hat{G}_{Add}}^{-1}(\tau)$ . The average-maximizing treatment rule  $\hat{G}_{Add}$  generates an income distribution in which top quantiles (0.8 and higher) are substantially higher than in the income distribution generated by the Gini treatment rule. However, the distribution produced by the Gini treatment rule is better at midrange quantiles (0.4–0.8). The additive welfare criterion equally weights changes of all quantiles, hence, it favors  $\hat{G}_{Add}$ .

The standard Gini welfare criterion, in contrast, uses decreasing welfare weights  $\omega_3(\tau) = 2(1 - \tau)$ . The right panel of Figure 3 displays the same quantile differences between the two income distributions weighted by  $\omega_3(\tau)$ . With these equality-minded welfare weights, the gains offered by treatment rule  $\hat{G}_{Add}$  at top quantiles get a lower welfare weight than the gains offered by  $\hat{G}_{Gini}$  in the middle of the income distribution, hence,  $\hat{G}_{Gini}$  is preferred under the Gini SWF.

## 6. Conclusion

This article develops the first method for individualized treatment choice when the policy maker's objective is to maximize an equality-minded rank-dependent SWF. We showed that the average social welfare obtained by the estimated policy converges at the minimax-optimal  $n^{-1/2}$  rate. The key restriction underlying these rate results is the complexity restriction (Assumption 2.2 (VC)) imposed on the set of feasible policies. This complexity restriction still allows for rich classes of individualized treatment rules and offers a flexible and convenient way to incorporate exogenous constraints

that policy makers face in realistic settings of policy design. Our analytical results cover a general class of equality-minded rank-dependent SWFs, and the shown regret bounds are valid even when the complexity of policies grows with the sample size. Efficient computation for the equality-minded EWM and a data-driven choice of complexity for policies (e.g., as proposed by Mbakop and Tabord-Meehan (2018) for the additive welfare case) remain open questions.

## Appendix A: Lemmas and Proofs

*Proof of Theorem 2.1.* Denote an upper level set of  $\delta(x)$  at level  $u \in [0, 1]$  by  $G(u) \equiv \{x \in \mathcal{X} : \delta(x) \geq u\}$ . By noting that

$$\delta(x) = \int_0^1 1\{x \in G(u)\} du,$$

we can rewrite  $F_\delta(y)$  defined in (6) as

$$\begin{aligned} F_\delta(y) &= \int_{\mathcal{X}} \left[ \int_0^1 1\{x \notin G(u)\} du \cdot F_{Y_0|X=x}(y) \right. \\ &\quad \left. + \int_0^1 1\{x \in G(u)\} du \cdot F_{Y_1|X=x}(y) \right] dP_X(x) \\ &= \int_0^1 \left[ \int_{\mathcal{X}} (1\{x \notin G(u)\} \cdot F_{Y_0|X=x}(y) \right. \\ &\quad \left. + 1\{x \in G(u)\} \cdot F_{Y_1|X=x}(y)) dP_X(x) \right] du \\ &= \int_0^1 F_{G(u)}(y) du, \end{aligned}$$

where  $F_{G(u)}(y)$  is the distribution of outcomes induced by treatment rule  $\delta_{G(u)} \equiv 1\{x \in G(u)\}$ . By convexity of  $\Lambda(\cdot)$ , we obtain

$$\Lambda(F_\delta(y)) \leq \int_0^1 \Lambda(F_{G(u)}(y)) du,$$

and this leads to

$$W_\Lambda(F_\delta) \leq \int_0^1 W_\Lambda(F_{G(u)}) du \equiv \bar{W}_\Lambda. \quad (\text{A.1})$$

Suppose that  $\bar{W}_\Lambda - W_\Lambda(F_{G(u)}) > 0$  for all  $u \in [0, 1]$ . Then the integral of this function over the set  $u \in [0, 1]$  of positive measure must also be strictly positive,

$$0 < \int_0^1 (\bar{W}_\Lambda - W_\Lambda(F_{G(u)})) du = \bar{W}_\Lambda - \bar{W}_\Lambda,$$

which is a contradiction. Therefore, there exists  $u^* \in [0, 1]$  for which  $W_\Lambda(F_{G(u^*)}) \geq \bar{W}_\Lambda$ , hence,  $W_\Lambda(F_{G(u^*)}) \geq W_\Lambda(F_\delta)$ . If all upper level sets  $G(u)$  of  $\delta$  belong to  $\mathcal{G}$ , then also  $G(u^*) \in \mathcal{G}$ . □

The following five lemmas will be used in the proof of [Theorem 3.1](#). The first lemma establishes a quadratic upper bound for the function  $t^{-1/2}$  for  $t \geq 1$ .

**Lemma A.1.** Let  $t_0 > 1$ , define

$$g(t) \equiv \begin{cases} 0 & \text{for } t = 0, \\ t^{-1/2} & \text{for } t \geq 1, \end{cases} \quad (\text{A.2})$$

$$h(t) \equiv t_0^{-1/2} - \frac{1}{2}t_0^{-3/2}(t - t_0) + t_0^{-2}(t - t_0)^2. \quad (\text{A.3})$$

Then  $g(t) \leq h(t)$  for  $t = 0$  and for all  $t \geq 1$ .

*Proof of Lemma A.1.* For  $t = 0$ ,

$$h(0) = t_0^{-1/2} + \frac{1}{2}t_0^{-3/2}t_0 + t_0^{-2}t_0^2 = \frac{3}{2}t_0^{-1/2} + 1 > 0 = g(0).$$

Now consider the function  $(h - g)(t)$  and its derivatives for  $t \geq 1$ :

$$(h - g)(t) = t_0^{-1/2} - \frac{1}{2}t_0^{-3/2}(t - t_0) + t_0^{-2}(t - t_0)^2 - t^{-1/2},$$

$$(h - g)'(t) = -\frac{1}{2}t_0^{-3/2} + 2t_0^{-2}(t - t_0) + \frac{1}{2}t^{-3/2},$$

$$(h - g)''(t) = 2t_0^{-2} - \frac{3}{4}t^{-5/2}, \text{ and}$$

$$(h - g)'''(t) = \frac{15}{8}t^{-7/2}.$$

First, we will show that  $(h - g)(t) \geq 0$  for  $t \in [1, t_0]$ . The function is positive at  $t = 1$ :

$$\begin{aligned} (h - g)(1) &= t_0^{-1/2} - \frac{1}{2}t_0^{-3/2}(1 - t_0) + t_0^{-2}(1 - t_0)^2 - 1 \\ &= t_0^{-1/2} - \frac{1}{2}t_0^{-3/2} + \frac{1}{2}t_0^{-1/2} + t_0^{-2} - 2t_0^{-1} + 1 - 1 \\ &= \frac{1}{2}t_0^{-1/2} \left( 3 - 4t_0^{-1/2} - t_0^{-1} + 2t_0^{-3/2} \right) \\ &= \frac{1}{2}t_0^{-1/2} \left( 1 - t_0^{-1/2} \right)^2 \left( 3 + 2t_0^{-1/2} \right) > 0 \end{aligned}$$

because  $t_0^{-1/2} > 0$ . At  $t = t_0$ ,  $(h - g)(t_0) = 0$ . We will next show that  $(h - g)(t) \geq 0$  between  $t = 1$  and  $t = t_0$ .

The second derivative of  $(h - g)$  is positive at  $t = t_0$ ,

$$(h - g)''(t_0) = 2t_0^{-2} - \frac{3}{4}t_0^{-5/2} = t_0^{-2} \left( 2 - \frac{3}{4}t_0^{-1/2} \right) > 0 \quad (\text{A.4})$$

because  $t_0 > 1$  by assumption, hence,  $t_0^{-1/2} < 1$ . Since the third derivative is positive on  $[1, t_0]$ , it follows that the second derivative is either positive everywhere on  $[1, t_0]$ , or it is first negative on some interval  $[1, t_2)$  and then positive on  $(t_2, t_0]$ .

The first derivative of  $(h - g)$  equals zero at  $t = t_0$ :

$$(h - g)'(t_0) = -\frac{1}{2}t_0^{-3/2} + 2t_0^{-2}(t_0 - t_0) + \frac{1}{2}t_0^{-3/2} = 0.$$

If the second derivative is positive everywhere on  $[1, t_0]$ , then the first derivative must be negative everywhere on  $[1, t_0)$ . If the second derivative changes sign from negative to positive, then the first derivative must either be negative on  $[1, t_0)$  or it could switch sign from positive on some interval  $[1, t_1)$  to negative on  $(t_1, t_0)$ .

Since  $(h - g)(1) > 0$ ,  $(h - g)(t_0) = 0$ , and  $(h - g)$  is either decreasing on  $[1, t_0)$  or increasing on  $[1, t_1)$  and then decreasing on  $(t_1, t_0)$ , it follows that  $(h - g)(t) \geq 0$  for all  $t \in [1, t_0]$ .

Second, consider  $t > t_0$ . At  $t = t_0$ ,  $(h - g)(t_0) = 0$ ,  $(h - g)'(t_0) = 0$ , and the second derivative is positive for all  $t > t_0$  because it is positive at  $t = t_0$  (A.4) and the third derivative is positive for all  $t \geq t_0$ . It follows that  $(h - g)(t) > 0$  for all  $t > t_0$ . □

The second lemma applies the bound in [Lemma A.1](#) to the expectation of the function  $g(\cdot)$  of a binomial variable.

**Lemma A.2.** Suppose that random variable  $B \sim \text{Binomial}(n, p)$  with  $np > 1$  and  $g(\cdot)$  is the function defined in (A.2). Then

$$E[g(B)] < 2(np)^{-1/2}.$$

*Proof of Lemma A.2.* Let  $h(\cdot)$  be the function defined in (A.3) with  $t_0 = np = E[B]$ . [Lemma A.1](#) shows that  $g(t) \leq h(t)$  for all values in the support of  $B$  ( $t \in \{0, 1, 2, \dots\}$ ), therefore

$$\begin{aligned} E[g(B)] &\leq E[h(B)] = E \left[ (np)^{-1/2} - \frac{1}{2}(np)^{-3/2}(B - E[B]) \right. \\ &\quad \left. + (np)^{-2}(B - E[B])^2 \right] \\ &= (np)^{-1/2} - \frac{1}{2}(np)^{-3/2} \cdot 0 + (np)^{-2} \text{var}[B] \\ &= (np)^{-1/2} + (np)^{-2}(np)(1 - p) \\ &< (np)^{-1/2} + (np)^{-1} < 2(np)^{-1/2} \end{aligned}$$

because  $np > 1$  implies  $p > 0$  and  $(np)^{-1} < (np)^{-1/2}$ . □

Let  $\mathbf{x}^l \equiv \{x_1, \dots, x_l\}$  be a finite set with  $l \geq 1$  points in  $\mathcal{X}$ . Given a class of subsets  $\mathcal{G}$  in  $\mathcal{X}$ , define  $N(\mathbf{x}^l) = |\{\mathbf{x}^l \cap G : G \in \mathcal{G}\}|$  as the number of different subsets of  $\mathbf{x}^l$  picked out by  $G \in \mathcal{G}$ . The VC-dimension  $v \geq 1$  of  $\mathcal{G}$  is defined by the largest  $l$  such that  $\sup_{\mathbf{x}^l} N(\mathbf{x}^l) = 2^l$  holds (Vapnik 1998). See Vapnik (1998), Dudley (1999, chap. 4), and van der Vaart and Wellner (1996) for extensive discussions. Note that the VC-dimension is smaller by one compared to the VC-index used to measure the complexity of a class of sets in empirical process theory, for example, van der Vaart and Wellner (1996).

Let  $Z_i = (Y_i, D_i, X_i) \in \mathcal{Z}$ , where  $\mathcal{Z} \equiv \mathbb{R}_+ \times \{0, 1\} \times \mathcal{X}$ . The subgraph of a real-valued function  $f : \mathcal{Z} \mapsto \mathbb{R}$  is the set

$$SG(f) \equiv \{(z, t) \in \mathcal{Z} \times \mathbb{R} : 0 \leq t \leq f(z) \text{ or } f(z) \leq t \leq 0\}.$$

The third lemma is reproduced from Kitagawa and Tetenov (2018b, Lemma A.1). It establishes a link between the VC-dimension of a class of subsets in the covariate space  $\mathcal{X}$  and the VC-dimension of a class of subgraphs of functions on  $\mathcal{Z}$ .

**Lemma A.3.** Let  $\mathcal{G}$  be a VC-class of subsets of  $\mathcal{X}$  with VC-dimension  $v < \infty$ . Let  $g$  and  $h$  be two given functions from  $\mathcal{Z}$  to  $\mathbb{R}$ . Then the set of functions from  $\mathcal{Z}$  to  $\mathbb{R}$

$$\mathcal{F} = \{f_G(z) = g(z) \cdot 1\{x \in G\} + h(z) \cdot 1\{x \notin G\} : G \in \mathcal{G}\}$$

is a VC-subgraph class of functions with VC-dimension less than or equal to  $v$ .

The fourth lemma, reproduced from Kitagawa and Tetenov (2018b, Lemma A.4), is a maximal inequality that bounds the mean of a supremum of a centered empirical process indexed by a VC-subgraph class of functions.

**Lemma A.4.** Let  $\mathcal{F}$  be a class of uniformly bounded functions, that is, there exists  $\bar{F} < \infty$  such that  $\|f\|_\infty \leq \bar{F}$  for all  $f \in \mathcal{F}$ . Assume that  $\mathcal{F}$  is a VC-subgraph class with VC-dimension  $\nu < \infty$ . Then, there is a universal constant  $C_1$  such that

$$E p^n \left[ \sup_{f \in \mathcal{F}} |E_n(f) - E_P(f)| \right] \leq C_1 \bar{F} \sqrt{\frac{\nu}{n}}$$

holds for all  $n \geq 1$ .

The last novel lemma allows us to prove Theorem 3.1 for unbounded outcomes.

**Lemma A.5.** Let  $\mathcal{F}$  be a class of uniformly bounded functions, that is, there exists  $\bar{F} < \infty$  such that  $\|f\|_\infty \leq \bar{F}$  for all  $f \in \mathcal{F}$ . Assume that  $\mathcal{F}$  is a VC-subgraph class with VC-dimension  $\nu < \infty$ . Let  $(Y, Z) \sim P$ , where  $Y \geq 0$  is a scalar ( $Y$  and  $Z$  may be dependent). Let  $\{(Y_i, Z_i)\}_{i=1}^n \sim P^n$  be an iid sample from  $P$ . Assume that

$$\int_0^\infty \sqrt{P(Y > y)} dy \leq M. \tag{A.5}$$

Then, there is a universal constant  $C_T$  such that

$$\begin{aligned} & \int_0^\infty E p^n \left[ \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} - E_P[f(Z) 1\{Y > y\}] \right| \right] dy \\ & \leq C_T \bar{F} M \sqrt{\frac{\nu}{n}} \end{aligned} \tag{A.6}$$

holds for all  $n \geq 1$ .

**Proof of Lemma A.5.** We start by deriving upper bounds for each value of  $y, y > 0$ , on

$$\begin{aligned} & E p^n [\xi_n(y)], \\ \xi_n(y) & \equiv \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} - E_P[f(Z) 1\{Y > y\}] \right|. \end{aligned}$$

First, consider values of  $y$  for which  $nP(Y > y) \leq 1$ . Due to the envelope condition,  $\left| \frac{1}{n} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \right| \leq \bar{F} \cdot \frac{1}{n} \sum_{i=1}^n 1\{Y_i > y\}$  for any  $f \in \mathcal{F}$ , hence,

$$\begin{aligned} E p^n \left[ \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \right| \right] & \leq \bar{F} \cdot E p^n \left[ \frac{1}{n} \sum_{i=1}^n 1\{Y_i > y\} \right] \\ & = \bar{F} P(Y > y). \end{aligned}$$

Also,  $E_P[f(Z) 1\{Y > y\}] = P(Y > y) E_P[f(Z)|Y > y] \leq \bar{F} P(Y > y)$ . It follows that

$$\begin{aligned} E p^n [\xi_n(y)] & \leq E p^n \left[ \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \right| \right] \\ & \quad + \sup_{f \in \mathcal{F}} E_P[f(Z) 1\{Y > y\}] \\ & \leq 2\bar{F} P(Y > y) \leq 2\bar{F} \sqrt{\frac{P(Y > y)}{n}}, \end{aligned} \tag{A.7}$$

where the last inequality holds because  $\sqrt{nP(Y > y)} \leq 1$ .

Second, we consider values of  $y$  for which  $nP(Y > y) > 1$ . Define random variables  $N_y \equiv \sum_{i=1}^n 1\{Y_i > y\}$  for the number of observations in the data with  $Y_i > y$ , then

$$\frac{1}{n} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} = \begin{cases} 0 & \text{if } N_y = 0, \\ \frac{N_y}{n} \cdot \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} & \text{if } N_y \geq 1. \end{cases}$$

If  $N_y \geq 1$  then

$$\begin{aligned} \xi_n(y) & = \sup_{f \in \mathcal{F}} \left| \frac{N_y}{n} \cdot \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \right. \\ & \quad \left. - P(Y > y) E_P[f(Z)|Y > y] \right| \\ & = \sup_{f \in \mathcal{F}} \left| \begin{aligned} & \frac{N_y}{n} \cdot \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \\ & - P(Y > y) \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \\ & + P(Y > y) \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \\ & - P(Y > y) E_P[f(Z)|Y > y] \end{aligned} \right| \\ & \leq \left| \frac{N_y}{n} - P(Y > y) \right| \sup_{f \in \mathcal{F}} \left| \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} \right| \\ & \quad + P(Y > y) \sup_{f \in \mathcal{F}} \left| \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} - E_P[f(Z)|Y > y] \right| \\ & \leq \bar{F} \left| \frac{N_y}{n} - P(Y > y) \right| \text{ (the sum has } N_y \text{ nonzero terms, each} \\ & \quad \text{bounded by } \bar{F} \text{)} \\ & \quad + P(Y > y) \sup_{f \in \mathcal{F}} \left| \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} - E_P[f(Z)|Y > y] \right|. \end{aligned} \tag{A.8}$$

Note that  $\{Z_i\}_{i:Y_i>y}$  is an iid sample of size  $N_y$  from the conditional distribution  $P(Z|Y > y)$ . We next apply the bound in Lemma A.4 for each value of  $N_y \geq 1$ :

$$\begin{aligned} E p^n \left[ \sup_{f \in \mathcal{F}} \left| \frac{1}{N_y} \sum_{i=1}^n f(Z_i) 1\{Y_i > y\} - E_P[f(Z)|Y > y] \right| \middle| N_y \right] \\ \leq C_1 \bar{F} \sqrt{\frac{\nu}{N_y}}. \end{aligned} \tag{A.9}$$

Combining inequality (A.8) with bound (A.9), and using the definition  $g(t) = t^{-1/2}$  for  $t \geq 1$  from (A.2), we obtain a bound on the conditional expectation of  $\xi_n(y)$  for  $N_y \geq 1$ :

$$E p^n [\xi_n(y)|N_y] \leq \bar{F} \left| \frac{N_y}{n} - P(Y > y) \right| + P(Y > y) C_1 \bar{F} \sqrt{\nu} g(N_y). \tag{A.10}$$

If  $N_y = 0$  then  $1\{Y_i > y\} = 0$  for all  $i$ , hence,

$$\begin{aligned} \xi_n(y) & = \sup_{f \in \mathcal{F}} E_P[f(Z) 1\{Y > y\}] = P(Y > y) \sup_{f \in \mathcal{F}} E_P[f(Z)|Y > y] \\ & \leq \bar{F} P(Y > y) = \bar{F} \left| \frac{N_y}{n} - P(Y > y) \right| + P(Y > y) C_1 \bar{F} \sqrt{\nu} g(N_y), \end{aligned}$$

where the last equality uses the definition  $g(0) = 0$  from (A.2). Therefore, the conditional expectation bound (A.10) also holds for  $N_y = 0$ .

The unconditional expectation of  $\xi_n(y)$  is then bounded by

$$E_{P^n} [\xi_n(y)] \leq \bar{F} E_{P^n} \left[ \frac{N_y}{n} - P(Y > y) \right] + P(Y > y) C_1 \bar{F} \sqrt{v} E_{P^n} [g(N_y)]. \tag{A.11}$$

The random variable  $N_y$  has a Binomial( $n, P(Y > y)$ ) distribution, hence,

$$E_{P^n} \left| \frac{N_y}{n} - P(Y > y) \right| \leq \sqrt{\text{var} \left( \frac{N_y}{n} \right)} = \sqrt{\frac{P(Y > y)(1 - P(Y > y))}{n}} \leq \sqrt{\frac{P(Y > y)}{n}}.$$

Since  $nP(Y > y) > 1$ , applying Lemma A.2 yields  $E_{P^n} [g(N_y)] \leq \frac{2}{\sqrt{n} \sqrt{P(Y > y)}}$ . Combining this inequality with (A.11) and  $v \geq 1$  we obtain

$$\begin{aligned} E_{P^n} [\xi_n(y)] &\leq \bar{F} \sqrt{\frac{P(Y > y)}{n}} + P(Y > y) C_1 \bar{F} \sqrt{v} \frac{2}{\sqrt{n} \sqrt{P(Y > y)}} \\ &\leq 2(1 + C_1) \bar{F} \sqrt{\frac{v}{n}} \sqrt{P(Y > y)} = C_T \bar{F} \sqrt{\frac{v}{n}} \sqrt{P(Y > y)}, \end{aligned} \tag{A.12}$$

where  $C_T \equiv 2(1 + C_1)$ . This bound is higher than the bound (A.7) derived for  $y$  such that  $nP(Y > y) \leq 1$ , hence, bound (A.12) holds for all  $y \geq 0$ .

The last step is to integrate the bound (A.12) over  $y$  and apply (A.5):

$$\int_0^\infty E_{P^n} [\xi_n(y)] dy \leq \int_0^\infty C_T \bar{F} \sqrt{\frac{v}{n}} \sqrt{P(Y > y)} dy \leq C_T \bar{F} M \sqrt{\frac{v}{n}}.$$

□

**Proof of Theorem 3.1.** Take an arbitrary set  $G^* \in \mathcal{G}$ , then

$$\begin{aligned} W_\Lambda(G^*) - W_\Lambda(\hat{G}) &= W_\Lambda(G^*) - \hat{W}_\Lambda(\hat{G}) + \hat{W}_\Lambda(\hat{G}) - W_\Lambda(\hat{G}) \\ &\leq W_\Lambda(G^*) - \hat{W}_\Lambda(G^*) + \hat{W}_\Lambda(\hat{G}) - W_\Lambda(\hat{G}) \\ &\leq 2 \sup_{G \in \mathcal{G}} |\hat{W}_\Lambda(G) - W_\Lambda(G)|, \end{aligned}$$

where the second line follows since  $\hat{W}_\Lambda(\hat{G})$  maximizes  $\hat{W}_\Lambda(G)$  over  $G \in \mathcal{G}$ . It follows that

$$\sup_{G \in \mathcal{G}} W_\Lambda(G) - W_\Lambda(\hat{G}) \leq 2 \sup_{G \in \mathcal{G}} |\hat{W}_\Lambda(G) - W_\Lambda(G)|. \tag{A.13}$$

Since  $\Lambda(\cdot)$  is convex and nonincreasing,

$$\begin{aligned} \sup_{G \in \mathcal{G}} |\hat{W}_\Lambda(G) - W_\Lambda(G)| &= \sup_{G \in \mathcal{G}} \left| \int_0^\infty \Lambda(\hat{F}_G(y) \vee 0) dy \right. \\ &\quad \left. - \int_0^\infty \Lambda(F_G(y)) dy \right| \\ &\leq \sup_{G \in \mathcal{G}} \int_0^\infty |\Lambda(\hat{F}_G(y) \vee 0) - \Lambda(F_G(y))| dy \\ &\leq \int_0^\infty \sup_{G \in \mathcal{G}} |\Lambda(\hat{F}_G(y) \vee 0) - \Lambda(F_G(y))| dy \\ &\leq |\Lambda'(0)| \int_0^\infty \sup_{G \in \mathcal{G}} |\hat{F}_G(y) - F_G(y)| dy. \end{aligned} \tag{A.14}$$

Combining (A.13) and (A.14), the average welfare loss of  $\hat{G}$  can be bounded by

$$\begin{aligned} E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_\Lambda(G) - W_\Lambda(\hat{G}) \right] \\ \leq 2|\Lambda'(0)| \int_0^\infty E_{P^n} \left[ \sup_{G \in \mathcal{G}} |\hat{F}_G(y) - F_G(y)| \right] dy. \end{aligned} \tag{A.15}$$

By Lemma A.3, the class of functions  $\mathcal{W} = \{w_G(\cdot) : G \in \mathcal{G}\}$ , where

$$w_G(Z_i) \equiv \left[ \frac{D_i}{e(X_i)} \cdot 1\{X_i \in G\} + \frac{1 - D_i}{1 - e(X_i)} \cdot 1\{X_i \notin G\} \right] \tag{A.16}$$

is a VC-subgraph class with VC-dimension of at most  $v$ . Assumption 3.1 (SO) implies that  $w_G(Z_i) \in [0, \frac{1}{\kappa}]$ , hence, functions in  $\mathcal{W}$  are uniformly bounded by  $\frac{1}{\kappa}$ .

Since  $F_G(y) = 1 - E_P [w_G(Z) \cdot 1\{Y > y\}]$  and  $\hat{F}_G(y) = 1 - \frac{1}{n} \sum_{i=1}^n w_G(Z_i) \cdot 1\{Y_i > y\}$ ,

$$\begin{aligned} |\hat{F}_G(y) - F_G(y)| \\ = \left| \frac{1}{n} \sum_{i=1}^n w_G(Z_i) \cdot 1\{Y_i > y\} - E_P [w_G(Z) \cdot 1\{Y > y\}] \right|. \end{aligned}$$

It follows from Assumption 3.1 (TC) and  $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$  that

$$\begin{aligned} \int_0^\infty \sqrt{P(Y > y)} dy &= \int_0^\infty \sqrt{P(Y_1 > y, D = 1) + P(Y_0 > y, D = 0)} dy \\ &\leq \int_0^\infty [\sqrt{P(Y_1 > y)} + \sqrt{P(Y_0 > y)}] dy \leq 2\Upsilon. \end{aligned} \tag{A.17}$$

Applying Lemma A.5 to (A.15) yields

$$E_{P^n} \left[ \sup_{G \in \mathcal{G}} W_\Lambda(G) - W_\Lambda(\hat{G}) \right] \leq 4C_T |\Lambda'(0)| \frac{\Upsilon}{\kappa} \sqrt{\frac{v}{n}}.$$

Setting  $C = 4C_T$  completes the proof of (17).

The proof of (18) is found in the online supplement. □

### Supplementary Materials

Supplementary Materials, available on the journal's website, include the dataset and replication code for the empirical application, an online appendix containing an illustrative example of the properties of rank-dependent SWFs, proofs of Theorems 3.1, 3.2 and Theorem 4.1, and an extension of the method with estimated propensity score.

### Acknowledgments

We would like to thank Ignacio Monzón, John Weymark, Max Tabord-Meehan, and Jeff Rowley for useful comments. We have also benefitted from presenting this work at Academia Sinica, Bristol ESG, RES, Cowles Conference on Econometrics, IAAE, SCW, AMES, Cemmap Advances in Econometrics, CEU, EEA-ESEM, Oxford, and Northwestern Conference of Former Econometrics Students.

### Funding

Financial support from the ESRC through the ESRC Centre for Microdata Methods and Practice (CeMMAP) (grant number RES-589-28-0001) and the European Research Council (Starting grant number 715940) is gratefully acknowledged. Toru Kitagawa thanks the Visiting Researcher Fellowship from Collegio Carlo Alberto.

### References

Aaberge, R., Havnes, T., and Mogstad, M. (2013), "A Theory for Ranking Distribution Functions," IZA Discussion Paper, No. 7738. [563,564]  
 Abadie, A., Angrist, J., and Imbens, G. (2002), "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica*, 70, 91–117. [568]

- Athey, S., and Wager, S. (2018), "Efficient Policy Learning," arXiv no. 1702.0289. [562]
- Atkinson, A. B. (1970), "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–263. [563]
- Bhattacharya, D., and Dupas, P. (2012), "Inferring Welfare Maximizing Treatment Assignment Under Budget Constraints," *Journal of Econometrics*, 167, 168–196. [562]
- Bitler, M. P., Gelbach, J. B., and Hoynes, H. W. (2006), "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments," *American Economic Review*, 96, 988–1012. [561]
- Blackorby, C., and Donaldson, D. (1978), "Measures of Relative Equality and Their Meaning in Terms of Social Welfare," *Journal of Economic Theory*, 18, 59–80. [561,563,564]
- Bloom, H. S., Orr, L. L., Bell, S. H., Cave, G., Doolittle, F., Lin, W., and Bos, J. M. (1997), "The Benefits and Costs of JTPA Title II-A Programs: Key Findings From the National Job Training Partnership Act Study," *Journal of Human Resources*, 32, 549–576. [568]
- Chamberlain, G. (2011), "Bayesian Aspects of Treatment Choice," in *The Oxford Handbook of Bayesian Econometrics*, eds. J. Geweke, G. Koop, and H. van Dijk, Oxford: Oxford University Press, pp. 11–39. [562]
- Cowell, F. (1995), *Measuring Inequality* (2nd ed.), Hemel Hempstead: Harvester Wheatsheaf. [563]
- (2000), "Measurement of Inequality," in *Handbook of Income Distribution* (Vol. 1), eds. A. Atkinson and F. Bourguignon, Amsterdam: Elsevier. [563]
- Dehejia, R. H. (2005), "Program Evaluation as a Decision Problem," *Journal of Econometrics*, 125, 141–173. [562]
- Donaldson, D., and Weymark, J. A. (1980), "A Single-Parameter Generalization of the Gini Indices of Inequality," *Journal of Economic Theory*, 22, 67–86. [564]
- (1983), "Ethically Flexible Gini Indices for Income Distributions in the Continuum," *Journal of Economic Theory*, 29, 353–358. [564]
- Dudley, R. (1999), *Uniform Central Limit Theorems*, Cambridge, UK: Cambridge University Press. [571]
- Fan, Y., and Park, S. S. (2010), "Sharp Bounds on the Distribution of Treatment Effects and Their Statistical Inference," *Econometric Theory*, 26, 931–951. [563]
- Firpo, S., and Pinto, C. (2016), "Identification and Estimation of Distributional Impacts of Interventions Using Changes in Inequality Measures," *Journal of Applied Econometrics*, 31, 457–486. [563]
- Firpo, S., and Ridder, G. (2019), "Partial Identification of the Treatment Effect Distribution and Its Functionals," *Journal of Econometrics*, 213, 210–234. [563]
- Heckman, J. J., Ichimura, H., and Todd, P. E. (1997), "Matching as an Econometric Evaluation Estimator: Evidence From Evaluating a Job Training Programme," *The Review of Economic Studies*, 64, 605–654. [568]
- Heckman, J. J., Smith, J., and Clements, N. (1997), "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts," *The Review of Economic Studies*, 64, 487–535. [563]
- Hirano, K., and Porter, J. R. (2009), "Asymptotics for Statistical Treatment Rules," *Econometrica*, 77, 1683–1701. [562]
- Imbens, G. W. (2004), "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics*, 86, 4–29. [565]
- Kasy, M. (2016), "Partial Identification, Distributional Preferences, and the Welfare Ranking of Policies," *Review of Economics and Statistics*, 98, 111–131. [562,563]
- (2018), "Optimal Taxation and Insurance Using Machine Learning—Sufficient Statistics and Beyond," *Journal of Public Economics*, 167, 205–219. [562]
- Kitagawa, T., and Tetenov, A. (2018a), "Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice," *Econometrica*, 86, 591–616. [562,563,565,566,567]
- (2018b), "Supplement to 'Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice,'" *Econometrica*. [571,572]
- Lambert, P. J. (2001), *The Distribution and Redistribution of Income* (3rd ed.), Manchester: Manchester University Press. [563]
- Machina, M. (1982), "Expected Utility' Analysis Without the Independence Axiom," *Econometrica*, 50, 277–323. [563]
- Manski, C. F. (2004), "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72, 1221–1246. [561,562]
- Mbakop, E., and Tabord-Meehan, M. (2018), "Model Selection for Treatment Choice: Penalized Welfare Maximization," arXiv no. 1609.03167. [562,565,570]
- Mehran, F. (1976), "Linear Measures of Income Inequality," *Econometrica*, 44, 805–809. [561,563]
- Quiggin, J. (1982), "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343. [563]
- Stoye, J. (2009), "Minimax Regret Treatment Choice With Finite Samples," *Journal of Econometrics*, 151, 70–81. [562]
- (2012), "Minimax Regret Treatment Choice With Covariates or With Limited Validity of Experiments," *Journal of Econometrics*, 166, 138–156. [562]
- Tetenov, A. (2012), "Statistical Treatment Choice Based on Asymmetric Minimax Regret Criteria," *Journal of Econometrics*, 166, 157–165. [562]
- van der Vaart, A. W., and Wellner, J. A. (1996), *Weak Convergence and Empirical Processes*, New York: Springer. [571]
- Vapnik, V. N. (1998), *Statistical Learning Theory*, New York: Wiley. [562,571]
- Wang, L., Zhou, Y., Song, R., and Sherwood, B. (2018), "Quantile-Optimal Treatment Regimes," *Journal of the American Statistical Association*, 113, 1243–1254. [563]
- Weymark, J. A. (1981), "Generalized Gini Inequality Indices," *Mathematical Social Sciences*, 1, 409–430. [561,563,564]
- Yaari, M. E. (1988), "A Controversial Proposal Concerning Inequality Measurement," *Journal of Economic Theory*, 44, 381–397. [563]