

Individualized decision making under partial identification: three perspectives, two optimality results, and one paradox

Yifan Cui

National University of Singapore

ABSTRACT. Unmeasured confounding is a threat to causal inference and gives rise to biased estimates. In this paper, we consider the problem of individualized decision making under partial identification. Firstly, we argue that when faced with unmeasured confounding, one should pursue individualized decision making using partial identification in a comprehensive manner. We establish a formal link between individualized decision making under partial identification and classical decision theory by considering a lower bound perspective of value/utility function. Secondly, building on this unified framework, we provide a novel minimax solution (i.e., a rule that minimizes the maximum regret for so-called opportunists) for individualized decision making/policy assignment. Lastly, we provide an interesting paradox drawing on novel connections between two challenging domains, i.e., individualized decision making and unmeasured confounding. Although motivated by instrumental variable bounds, we emphasize that the general framework proposed in this paper would in principle apply for a rich set of bounds that might be available under partial identification.

Keywords: Decision making strategies, Individualized preferences, Partial identification, Mixed strategy, Optimality, Sharpness

MEDIA SUMMARY

In the era of big data, observational studies are a treasure for both association analysis and causal inference, with the potential to improve decision making. Depending on the set of assumptions one is willing to make, one might achieve either point, sign, or partial identification of causal effects. In particular, under partial identification, it might be inevitable to make suboptimal decisions. Policymakers caring about decision making would face the following important question: what are optimal strategies corresponding to different risk preferences?

In this article, the author offers a unified framework which generalizes several decision making strategies in the literature. Building on this unified framework, the author also provides a novel minimax solution (i.e., a rule that minimizes the maximum regret for so-called opportunists) for individualized decision making/policy assignment.

1. THE POWER OF STORY-TELLING: DIFFERENT VIEWS MIGHT LEAD TO DIFFERENT DECISIONS

Suppose one is playing a two-armed slot machine. The rewards R_{-1} and R_1 are the payoffs for hitting the jackpot of each arm, respectively. For simplicity, let us assume that both arms always give positive rewards ($R_{-1}, R_1 > 0$), i.e., one is guaranteed not to lose and therefore would not refrain from playing this game. However, due to some uncertainty, one does not have prior knowledge of the exact value of R_{-1} and R_1 . Fortunately, suppose there is a magic instrument, which can help one to identify the range of rewards.

By only providing one with the left panel of Figure 1, i.e., the range of $R_1 - R_{-1}$, most people might opt to the arm -1 . But wait a minute... Where am I and why am I looking at the left panel without knowing the real payoffs? After looking at the right panel, the decision might be changed depending on a person's risk preference.

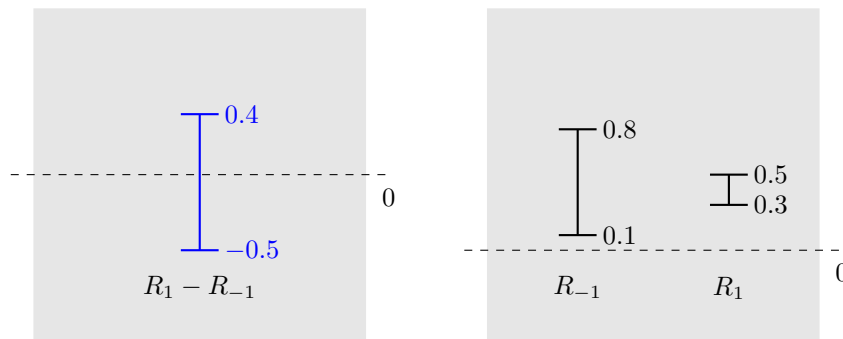


Figure 1. The left panel: the possible range of $R_1 - R_{-1}$; The right panel: the possible ranges of R_{-1} and R_1 respectively.

Is there such an instrument in real-life? The answer is in the affirmative. One such instruments is a so-called instrumental variable (IV). In statistics and related disciplines, IV method is used to estimate causal relationships when randomized experiments are not feasible or when there is non-compliance in a randomized experiment. Intuitively, a valid IV induces changes in the explanatory variable but otherwise has no direct effect on the dependent variable, allowing one to uncover the causal effect of the explanatory variable on the dependent variable. Under certain IV models, one can obtain bounds for counterfactual means. So how would one pursue decision making when faced with partial identification? The rest of the paper offers a comprehensive view of individualized decision making under partial identification as well as several novel solutions to various decision/policy making strategies.

1.1. Introduction. An optimal decision rule provides a personalized action/treatment strategy for each participant in the population based on one's individual characteristics. A prevailing strand of work has been devoted to estimating optimal decision rules (Athey and Wager, 2021; Murphy, 2003; Murphy et al., 2001; Qian and Murphy, 2011; J. M. Robins, 2004; Zhang, Tsiatis, Davidian, et al., 2012; Zhao et al., 2012 and many others), we refer to Chakraborty and Moodie (2013), Kosorok and Laber (2019), and Tsiatis et al. (2019) for an up-to-date literature review on this topic.

Recently, there has been a fast-growing literature on estimating individualized decision rules based on observational studies subject to potential unmeasured confounding (Cui & Tchetgen Tchetgen, 2020, 2021a, 2021b; Han, 2019, 2020, 2021; Kallus et al., 2019; Kallus & Zhou, 2018; Pu &

Zhang, 2021; Qiu et al., 2021a, 2021b; Yaddowsky et al., 2018; Zhang & Pu, 2021). In particular, Cui and Tchetgen Tchetgen (2021b) pointed out that one could identify treatment regimes that maximize lower bounds of the value function when one has only partial identification through an IV. Pu and Zhang (2021) further proposed an IV-optimality criterion to learn an optimal treatment regime, which essentially recommends the treatment for patients for whom the estimated conditional average treatment effect bound covers zero based on the length of the bounds, i.e., based on the left panel of Figure 1. See more details in Cui and Tchetgen Tchetgen (2021a, 2021b) and Zhang and Pu (2021).

In this paper, we provide a comprehensive view of individualized decision making under partial identification through maximizing the lower bounds of the value function. This new perspective unifies various classical decision making strategies in classical decision theory. Building on this unified framework, we also provide a novel minimax solution (for so-called opportunists who fail to lose) for individualized decision making/policy assignment. In addition, we point out that there is a mismatch between different optimality results, i.e., an “optimal” rule that attains one criterion does not necessarily attain the other. Such mismatch is a distinctive feature of individualized decision making under partial identification, and therefore makes the concept of universal optimality for decision making under uncertainty ill-defined. Lastly, we provide a paradox to illustrate that a non-individualized decision can conceivably lead to a superior outcome than an individualized decision under partial identification. The provided paradox also sheds light on that using IV bounds as sanity check or policy improvement might also be of interest.

To conclude this section, we briefly introduce notation used throughout the paper. Let Y denote the outcome of interest and $A \in \{-1, 1\}$ be a binary action/treatment indicator. Throughout it is assumed that larger values of Y are more desirable. Suppose that U is an unmeasured confounder of the effect of A on Y . Suppose also that one has observed a pre-treatment binary IV $Z \in \{-1, 1\}$. Let X denote a set of fully observed pre-IV covariates. Throughout we assume the complete data are independent and identically distributed realizations of (Y, X, A, Z, U) ; thus the observed data are (Y, X, A, Z) .

2. A BRIEF REVIEW OF OPTIMAL DECISION RULES WITH NO UNMEASURED CONFOUNDING

An individualized decision rule is a mapping from the covariate space to the action space $\{-1, 1\}$. Suppose Y_a is a person’s potential outcome under an intervention that sets A to value a , $Y_{\mathcal{D}(X)}$ is the potential outcome under a hypothetical intervention that assigns A according to the rule \mathcal{D} , i.e., $Y_{\mathcal{D}(X)} \equiv Y_1 I\{\mathcal{D}(X) = 1\} + Y_{-1} I\{\mathcal{D}(X) = -1\}$, $E[Y_{\mathcal{D}(X)}]$ is the value function (Qian & Murphy, 2011), and $I\{\cdot\}$ is the indicator function. Throughout the paper, we make the following standard consistency and positivity assumptions: (1) For a given regime \mathcal{D} , $Y = Y_{\mathcal{D}(X)}$ when $A = \mathcal{D}(X)$ almost surely. That is, a person’s observed outcome matches his/her potential outcome under a given decision rule when the realized action matches his/her potential assignment under the rule; (2) We assume that $\Pr(A = a|X) > 0$ for $a = \pm 1$ almost surely. That is, for any observed covariates X , a person has an opportunity to take either action.

We wish to identify an optimal decision rule \mathcal{D}^* that admits the following representation, i.e.,

$$(2.1) \quad \mathcal{D}^*(X) = \text{sign}\{E(Y_1 - Y_{-1}|X) > 0\} \text{ or } \mathcal{D}^* = \arg \max_{\mathcal{D}} E[Y_{\mathcal{D}(X)}].$$

A significant amount of work has been devoted to estimating optimal decision rules relying on the following unconfoundedness assumption:

Assumption 1. (Unconfoundedness) $Y_a \perp\!\!\!\perp A|X$ for $a = \pm 1$.

The assumption essentially rules out the existence of an unmeasured factor U that confounds the effect of A on Y upon conditioning on X . It is straightforward to verify that under Assumption 1, one can identify the value function $E[Y_{\mathcal{D}(X)}]$ for a given decision rule \mathcal{D} . Furthermore, the optimal decision rule in Equation (2.1) is identified from the observed data

$$\mathcal{D}^*(X) = \text{sign}\{\mathcal{C}(X) > 0\},$$

where $\mathcal{C}(X) = E(Y|X, A = 1) - E(Y|X, A = -1) = E(Y_1 - Y_{-1}|X)$ denotes the conditional average treatment effect (CATE). As established by Qian and Murphy (2011), learning optimal decision rules under Assumption 1 can be formulated as

$$\mathcal{D}^* = \arg \max_{\mathcal{D}} E \left[\frac{I\{\mathcal{D}(X) = A\}Y}{\Pr(A|X)} \right],$$

where $\Pr(A|X)$ is the probability of taking A given X . Zhang, Tsiatis, Laber, et al. (2012) proposed to directly maximize the value function over a parametrized set of functions. Rather than maximizing the above value function, Rubin and van der Laan (2012), Zhang, Tsiatis, Davidian, et al. (2012), and Zhao et al. (2012) transformed the above problem into a weighted classification problem,

$$\arg \min_{\mathcal{D}} E\{|\mathcal{C}(X)|I[\text{sign}\{\mathcal{C}(X) > 0\} \neq \mathcal{D}(X)]\}.$$

The ensuing classification approach was shown to have appealing robustness properties, particularly in a randomized study where no model assumption on Y is needed.

3. INSTRUMENTAL VARIABLE WITH PARTIAL IDENTIFICATION

In this section, instead of relying on Assumption 1, we allow for unmeasured confounding, which might cause biased estimates of optimal decision rules. Let $Y_{z,a}$ denote the potential outcome had, possibly contrary to fact, a person's IV and treatment value been set to z and a , respectively. Suppose that the following assumption holds.

Assumption 2. (Latent unconfoundedness) $Y_{z,a} \perp\!\!\!\perp (Z, A)|X, U$ for $z, a = \pm 1$.

This assumption essentially states that together U and X would in principle suffice to account for any confounding bias. Because U is not observed, we propose to account for it when a valid IV Z is available that satisfies the following standard IV assumptions (Cui & Tchetgen Tchetgen, 2021b):

Assumption 3. (IV relevance) $Z \not\perp\!\!\!\perp A|X$.

Assumption 4. (Exclusion restriction) $Y_{z,a} = Y_a$ for $z, a = \pm 1$ *almost surely*.

Assumption 5. (IV independence) $Z \perp\!\!\!\perp U|X$.

Assumption 6. (IV positivity) $0 < \Pr(Z = 1|X) < 1$ *almost surely*.

Assumptions 3-5 are well-known IV conditions, while Assumption 6 is needed for nonparametric identification (Greenland, 2000; Hernan & Robins, 2006). Assumption 3 requires that the IV is associated with the treatment conditional on X . Note that Assumption 3 does not rule out confounding of the Z - A association by an unmeasured factor, however, if present, such factor must be independent of U . Assumption 4 states that there can be no direct causal effect of Z on Y not

mediated by A . Assumption 5 states that the direct causal effect of Z on Y would be identified conditional on X if one were to intervene on $A = a$. Figure 2 provides a graphical representation of Assumptions 4 and 5.

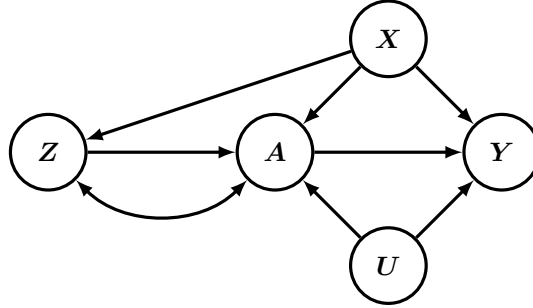


Figure 2. Causal directed acyclic graph with unmeasured confounding. The bi-directed arrow between Z and A indicates the possibility that there may be unmeasured common causes confounding their association.

While Assumptions 3-6 together do not suffice for point identification of the counterfactual mean and average treatment effect, a valid IV, even under minimal four assumptions, can partially identify the counterfactual mean and average treatment effect, i.e., lower and upper bounds might be formed. Let $\mathcal{L}_{-1}(X)$, $\mathcal{U}_{-1}(X)$, $\mathcal{L}_1(X)$, $\mathcal{U}_1(X)$ denote lower and upper bounds for $E(Y_{-1}|X)$ and $E(Y_1|X)$; hereafter, we consider lower and upper bounds for $E(Y_1 - Y_{-1}|X)$ of form $\mathcal{L}(X) = \mathcal{L}_1(X) - \mathcal{U}_{-1}(X)$ and $\mathcal{U}(X) = \mathcal{U}_1(X) - \mathcal{L}_{-1}(X)$, respectively; sharp bounds for $E(Y_1 - Y_{-1}|X)$ in certain prominent IV models have been shown to take such a form, see for instance Robins-Manski bound (Manski, 1990; J. Robins, 1989), Balke-Pearl bound (Balke & Pearl, 1997), Manski-Pepper bound under a monotone IV assumption (Manski & Pepper, 2000) and many others. Here, we consider the following Balke-Pearl bounds (Balke & Pearl, 1997) for a binary outcome as our running example. Let $p_{y,a,z,x}$ denote $\Pr(Y = y, A = a|Z = z, X = x)$, and

$$\begin{aligned}
 \mathcal{L}_{-1}(x) &= \max \left\{ \begin{array}{c} p_{1,-1,1,x} \\ p_{1,-1,-1,x} \\ p_{1,-1,-1,x} + p_{1,1,-1,x} - p_{-1,-1,1,x} - p_{1,1,1,x} \\ p_{-1,1,-1,x} + p_{1,-1,-1,x} - p_{-1,-1,1,x} - p_{-1,1,1,x} \end{array} \right\}, \\
 \mathcal{U}_{-1}(x) &= \min \left\{ \begin{array}{c} 1 - p_{-1,-1,1,x} \\ 1 - p_{-1,-1,-1,x} \\ p_{-1,1,-1,x} + p_{1,-1,-1,x} + p_{1,-1,1,x} + p_{1,1,1,x} \\ p_{1,-1,-1,x} + p_{1,1,-1,x} + p_{-1,1,1,x} + p_{1,-1,1,x} \end{array} \right\}, \\
 \mathcal{L}_1(x) &= \max \left\{ \begin{array}{c} p_{1,1,-1,x} \\ p_{1,1,1,x} \\ -p_{-1,-1,-1,x} - p_{-1,1,-1,x} + p_{-1,-1,1,x} + p_{1,1,1,x} \\ -p_{-1,1,-1,x} - p_{1,-1,-1,x} + p_{1,-1,1,x} + p_{1,1,1,x} \end{array} \right\}, \\
 \mathcal{U}_1(x) &= \min \left\{ \begin{array}{c} 1 - p_{-1,1,1,x} \\ 1 - p_{-1,1,-1,x} \\ p_{-1,-1,-1,x} + p_{1,1,-1,x} + p_{1,-1,1,x} + p_{1,1,1,x} \\ p_{1,-1,-1,x} + p_{1,1,-1,x} + p_{-1,-1,1,x} + p_{1,1,1,x} \end{array} \right\}.
 \end{aligned}$$

Additionally, one could proceed with other partial identification assumptions and corresponding bounds. We refer to references cited in Balke and Pearl (1997) and a review paper by Swanson et al. (2018) for alternative bounds.

We conclude this section by providing multiple settings in real-life where an IV is available but Assumption 1 is not likely to hold: 1) In a double-blind placebo-randomized trial in which participants are subject to noncompliance, the treatment assignment is a valid IV; 2) Another classical example is that in sequential, multiple assignment, randomized trials (SMARTs) in which patients are subject to noncompliance, the adaptive intervention is a valid IV. We note that the later proposed randomized minimax solution in Section 5.3 offers a promising strategy for this setting; 3) In social studies, a classical example is estimating the causal effect of education on earnings. Residential proximity to a college is a valid IV. We will further elaborate the third example in the next section.

4. A REAL-WORLD EXAMPLE

In this section, we first consider a real-world application on the effect of education on earnings using data from the National Longitudinal Study of Young Men (Card, 1993; Okui et al., 2012; Tan, 2006; Wang et al., 2017; Wang & Tchetgen Tchetgen, 2018), which consist of participants on 5525 men aged between 14 and 24 in 1966. Among them, 3010 provided valid education and wage responses in the 1976 follow-up. Following Tan (2006) and Wang and Tchetgen Tchetgen (2018), we consider education beyond high school as a binary action/treatment (i.e., A). A practically relevant question is the following: which students would be better off starting college to maximize their earnings?

In this study, there might be unmeasured confounders even after adjusting for observed covariates, e.g., unobserved preferences for education levels might be an unmeasured confounder that is likely to be associated with both education and wage. We follow Card (1993), Wang et al. (2017), and Wang and Tchetgen Tchetgen (2018) and use presence of a nearby four-year college as an instrument (i.e., Z). In this data set, 2053 (68.2%) lived close to a four-year college, and 1521 (50.5%) had education beyond high school. To illustrate the IV bounds with binary outcomes, we follow Wang et al. (2017) and Wang and Tchetgen Tchetgen (2018) to dichotomize the outcome wage (i.e., Y) at its median, that is 5.375 dollars per hour. While we only use this as an illustrating example, we note that dichotomizing earnings might affect decision making, and therefore in practice one might conduct a sensitivity analysis around the choice of cut-off. Following Wang and Tchetgen Tchetgen (2018), we adjust for age, race, father and mother’s education levels, indicators for residence in the south and a metropolitan area and IQ scores (i.e., X), all measured in 1966. Among them, race, parents’ education levels and residence are included as they may affect both the IV and outcome; age is included as it is likely to modify the effect of education on earnings; and IQ scores, as a measure of underlying ability, are included as they may modify both the effect of proximity to college on education, and the effect of education on earnings.

We use random forests to estimate the probability of $p_{y,a,z,x}$ (with default tuning parameters in Liaw and Wiener (2002)) and then construct estimates of Balke-Pearl bounds $\mathcal{L}_{-1}(X)$, $\mathcal{U}_{-1}(X)$, $\mathcal{L}_1(X)$, $\mathcal{U}_1(X)$, $\mathcal{L}(X)$, $\mathcal{U}(X)$. To streamline our presentation, we consider the subset of individuals of age 15, parents’ education level 11 years, non-black, and residence in a non-south and metropolitan area. Their IV CATE and counterfactual mean bounds $\mathcal{L}(X)$, $\mathcal{U}(X)$, $\mathcal{L}_{-1}(X)$, $\mathcal{U}_{-1}(X)$, $\mathcal{L}_1(X)$, $\mathcal{U}_1(X)$ are presented in Figure 3.

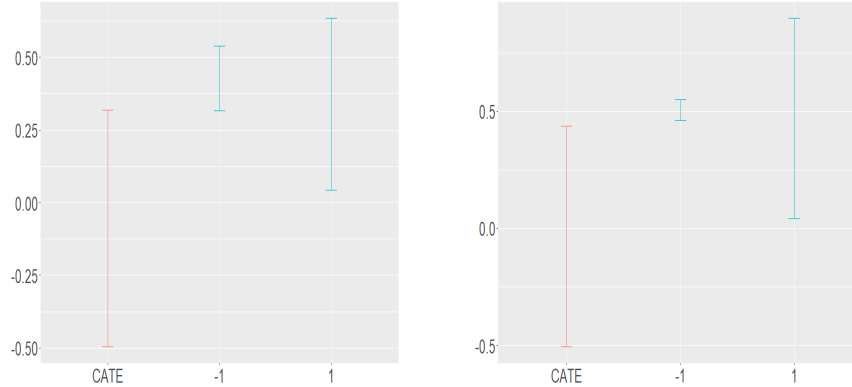


Figure 3. IV CATE and counterfactual mean bounds for two subjects with IQ scores 84.00 and 102.45, where $A = 1$ and -1 refer to education beyond high school or not, respectively.

The shape of IV bounds looks similar to the slot machine example of Figure 1 given at the beginning of the paper. When faced with uncertainty, what are different decision making strategies? In the next section, we provide a new perspective of viewing optimal decision making under partial identification beyond just looking at contrast or value function. Except for the real-world example, for pedagogical purposes, we focus on the population level of IV bounds instead of their empirical analogs throughout.

5. THE LOWER BOUND PERSPECTIVE: A UNIFIED CRITERION

In Section 5.1, we link the lower bound framework to well established decision theory from an investigator’s perspective. In Section 5.2, we extend our framework to take into account individual preferences of participants. In Section 5.3, we provide a formal solution to achieve a minimax regret goal by leveraging a randomization scheme. In Section 5.4, we reveal a mismatch between deterministic/randomized minimax regret and maximin utility, and conclude that there is no universal concept of optimality for decision making under partial identification.

5.1. A generalization of classical decision theory. In this section, we establish a formal link between individualized decision making under partial identification and classical decision theory. The set of rules $\mathcal{D}(w(x), x)$ which maximize the following lower bounds of $E[Y_{\mathcal{D}(X)}]$,

$$\left\{ E_X \{ [1 - w(X)] [\mathcal{L}(X) I \{ \mathcal{D}(X) = 1 \} + \mathcal{L}_{-1}(X)] + w(X) [-\mathcal{U}(X) I \{ \mathcal{D}(x) = -1 \} + \mathcal{L}_1(X)] \} : \right. \\ \left. \text{where } w(x) \text{ can depend on } \mathcal{D}(x), 0 \leq w(x) \leq 1, \text{ for any } x \right\},$$

is denoted by \mathcal{D}^{opt} . The derivation of lower bounds of $E[Y_{\mathcal{D}(X)}]$ is provided in the Appendix. Hereinafter, we refer to reasoning decision making strategy from \mathcal{D}^{opt} as the lower bound criterion, where, as can be seen later, $w(x)$ reflects the investigator’s preferences.

In Table 1, we provide examples of decision making criteria that have previously appeared in classical decision theory and we connect each such criterion to a corresponding $w(x)$. Hereafter, for a rule \mathcal{D} , we formally define utility as value function $E[Y_{\mathcal{D}(X)}]$ and regret as $E[Y_{\mathcal{D}^*(X)}] - E[Y_{\mathcal{D}(X)}]$. We give the formal definition of each rule in Table 1 except that the mixed strategy is deferred to Section 5.3. In the following definitions, min or max without an argument is taken with respective to $E[Y_{\mathcal{D}(X)}]$ (recall that $E[Y_{\mathcal{D}(X)}] = E[E[Y_1|X]I\{\mathcal{D}(X) = 1\} + E[Y_{-1}|X]I\{\mathcal{D}(X) = -1\}]$),

and $E(Y_{-1}|X)$, $E(Y_1|X)$ satisfy $\mathcal{L}_{-1}(X) \leq E(Y_{-1}|X) \leq \mathcal{U}_{-1}(X)$, $\mathcal{L}_1(X) \leq E(Y_1|X) \leq \mathcal{U}_1(X)$, respectively), and \mathcal{D} belongs to the set of all deterministic rules.

- Maximax utility (optimist): $\max_{\mathcal{D}} \max E[Y_{\mathcal{D}(X)}]$;
- Maximin utility (pessimist): $\max_{\mathcal{D}} \min E[Y_{\mathcal{D}(X)}]$;
- Minimax regret (opportunist): $\min_{\mathcal{D}} \max(E[Y_{\mathcal{D}^*(X)}] - E[Y_{\mathcal{D}(X)}])$;
- Hurwicz criterion: $\max_{\mathcal{D}} (\alpha \max E[Y_{\mathcal{D}(X)}] + (1 - \alpha) \min E[Y_{\mathcal{D}(X)}])$;
- Healthcare decision making: $\max_{\mathcal{D}} E[E(Y_{-1}|X) + \mathcal{L}(X)I\{\mathcal{D}(X) = 1\}]$.

For example, for the left panel of Figure 3, maximax utility criterion recommends $A = 1$; maximin utility criterion recommends $A = -1$; minimax regret criterion recommends $A = -1$.

Notably, all criteria in Table 1 reduce to \mathcal{D}^* under point identification. For a more complete treatment of decision making strategies and formal axioms of rational choice, we refer to Arrow and Hurwicz (1972). Interestingly, we note that a (deterministic) minimax regret criterion coincides with Hurwicz criterion with $\alpha = 1/2$ as $\mathcal{L}(X) = \mathcal{L}_1(X) - \mathcal{U}_{-1}(X)$ and $\mathcal{U}(X) = \mathcal{U}_1(X) - \mathcal{L}_{-1}(X)$.

Decision theory criterion	$w(x)$	The corresponding rule
Maximax utility (optimist)	$I(P > Q)$	$\mathcal{D}(w(x), x) = \begin{cases} 1 & \mathcal{U}_1(x) > \mathcal{U}_{-1}(x), \\ -1 & \mathcal{U}_1(x) < \mathcal{U}_{-1}(x). \end{cases}$
Maximin utility (pessimist)	$I(P < Q)$	$\mathcal{D}(w(x), x) = \begin{cases} 1 & \mathcal{L}_1(x) > \mathcal{L}_{-1}(x), \\ -1 & \mathcal{L}_1(x) < \mathcal{L}_{-1}(x). \end{cases}$
Minimax regret (opportunist)	$1/2$	$\mathcal{D}(w(x), x) = \begin{cases} 1 & \mathcal{L}(x) > 0 \text{ or } \mathcal{L}(x) < 0 < \mathcal{U}(x), \mathcal{U}(x) > \mathcal{L}(x) , \\ -1 & \mathcal{U}(x) < 0 \text{ or } \mathcal{L}(x) < 0 < \mathcal{U}(x), \mathcal{U}(x) < \mathcal{L}(x) . \end{cases}$
Hurwicz criterion ($0 \leq \alpha \leq 1$: Hurwicz index)	$\alpha I(P > Q) + (1 - \alpha)I(P < Q)$	$\mathcal{D}(w(x), x) = \begin{cases} 1 & (1 - \alpha)\mathcal{L}_1(x) + \alpha\mathcal{U}_1(x) > (1 - \alpha)\mathcal{L}_{-1}(x) + \alpha\mathcal{U}_{-1}(x), \\ -1 & (1 - \alpha)\mathcal{L}_1(x) + \alpha\mathcal{U}_1(x) < (1 - \alpha)\mathcal{L}_{-1}(x) + \alpha\mathcal{U}_{-1}(x). \end{cases}$
Healthcare decision making ($A = -1$: standard of care)	0	$\mathcal{D}(w(x), x) = \begin{cases} 1 & \mathcal{L}(x) > 0, \\ -1 & \mathcal{L}(x) < 0. \end{cases}$
...
Mixed strategy	a Bernoulli $w(x)$	a stochastic $\mathcal{D}(w(x), x)$

Table 1. Different representations of $w(x)$ for various decision making strategies defined in the Appendix, where $P \equiv \mathcal{L}(x)I\{\mathcal{D}(x) = 1\} + \mathcal{L}_{-1}(x)$ and $Q \equiv -\mathcal{U}(x)I\{\mathcal{D}(x) = -1\} + \mathcal{L}_1(x)$. The arguments of x and \mathcal{D} in P and Q are omitted for simplicity. To streamline the presentation, we omit the case of tiebreaking in $\mathcal{D}(w(x), x)$.

Remark 1. While both lower bound criterion and Hurwicz criterion have an index, they are conceptually and technically different. The index $w(x)$ being a number between 0 and 1 refers to the preference of actions; with $w(x)$ being a weighted average of $I(P < Q)$ and $I(P > Q)$, the lower bound criterion balances pessimism and optimism; however, it may not be straightforward for Hurwicz’s criterion to balance preferences on treatments/actions.

5.2. Incorporating individualized preferences: numeric/symbolic/stochastic inputs. We note that the lower bound criterion also sheds light on the process of data collection for individualized decision making. As individuals in the population of interest may ultimately exhibit different preferences for selecting optimal decisions, it may be unreasonable to assume that all participants share a common preference for evaluating optimality of an individualized decision rule under partial identification. An investigator might collect participants’ risk preferences over the space of rational choices to construct an individualized decision rule. Therefore, we use the subscript r (a participant’s observed preference) to remind ourselves that $w_r(x)$ depends not only on x but also on an

individual's risk preference, i.e., $r \in \mathcal{R}$ determines a specific form of $w_r(x)$ (see Table 1), where \mathcal{R} is a collection of different risk preferences. Such $w_r(x)$ results in a decision rule $\mathcal{D}(w_r(x), x)$ depending on both x (standard individualization, e.g., in the sense of subgroup identification) and r (individualized risk preferences when faced uncertainty), where r can be collected from each individual.

Remark 2. *We note that part of the elegance of this lower bound framework is that the risk preference does not come into play if there is no uncertainty about optimal decision, i.e., if $0 \notin (\mathcal{L}(x), \mathcal{U}(x))$, regardless what $w_r(x)$ being chosen, $\mathcal{D}(w_r(x), x) = \mathcal{D}^*(x)$.*

Remarkably, the recorded index $w_r(x)$ for each x could be numeric/symbolic/stochastic, i.e., fall into either one of the following three categories, while the participants only need to specify a category and input a number between 0 and 1 if the first two categories are chosen:

- Treatment/action preferences: Input a number β between 0 and 1 which indicates preference on treatments/actions with larger β in favor of $A = 1$. Here, $w_r(x) = \beta$. In observational studies, most applied researchers upon observing $0 \in (\mathcal{L}(x), \mathcal{U}(x))$ would rely on standard of care ($A = -1$) and opt to wait for more conclusive studies, which corresponds to $\beta = 0$. In a placebo-controlled study with $A = -1$ denoting placebo, $\beta = 0$ represents concerns about safeness/aversion of treatment.
- Utility/risk preferences: Input a number β between 0 and 1 and let symbolic input $w_r(x) = \beta I(P > Q) + [1 - \beta]I(P < Q)$, where β refers to the coefficient of optimism. For instance, $\beta = 0$ puts the emphasis on the worst possible outcome, and refers to risk aversion; and likewise $\beta = 1/2, 1$ refer to risk neutral and risk taker, respectively.
- An option for opportunists who are unwilling to lose: rendering $w_r(x)$ random as a Bernoulli random variable, see Section 5.3 for details.

We highlight that the proposed index $w_r(x)$ unifies various concepts in artificial intelligence, economics, and statistics, which holds promising to provide a satisfactory regime for each individual through a machine intelligent way.

5.3. A randomized minimax regret solution for opportunists. In this section, we consider whether an investigator/participant who happens to be an opportunist, can do better in terms of protecting the worst case regret than the minimax regret approach in Table 1?

An opportunist might not put all of his or her eggs in one basket. This mixed strategy is also known as mixed portfolio in portfolio optimization. Let $p(x)$ denote the probability of taking $A = 1$ given $X = x$, by the definition of the minimax regret criterion, one essentially needs to solve the following for $p(x)$,

$$\min_{p(x)} \max([1 - p(x)] \max\{\mathcal{U}(x), 0\}, p(x) \max\{-\mathcal{L}(x), 0\}),$$

which leads to the following solution

$$p(x) = \begin{cases} 1 & \mathcal{L}(x) > 0, \\ 0 & \mathcal{U}(x) < 0, \\ \frac{\mathcal{U}(x)}{\mathcal{U}(x) - \mathcal{L}(x)} & \mathcal{L}(x) < 0 < \mathcal{U}(x). \end{cases}$$

Such a choice of $p(x)$ guarantees the worst case regret no more than

$$\begin{cases} 0 & \mathcal{U}(x) < 0 \text{ or } \mathcal{L}(x) > 0, \\ -\frac{\mathcal{L}(x)\mathcal{U}(x)}{\mathcal{U}(x)-\mathcal{L}(x)} & \mathcal{L}(x) < 0 < \mathcal{U}(x). \end{cases}$$

We formalize the above result in the following theorem.

Theorem 5.1. *Define the stochastic policy $\tilde{\mathcal{D}}$ as $\tilde{\mathcal{D}}(x) = 1$ with probability $p(x)$, the corresponding regret is bounded by*

$$E[Y_{\mathcal{D}^*(X)}] - E[Y_{\tilde{\mathcal{D}}(X)}] \leq E \left[-\frac{\mathcal{L}(X)\mathcal{U}(X)}{\mathcal{U}(X) - \mathcal{L}(X)} I\{\mathcal{L}(X) < 0 < \mathcal{U}(X)\} \right],$$

where $E[Y_{\tilde{\mathcal{D}}(X)}] = E_X[E_{\tilde{\mathcal{D}}}[E_{Y_{\tilde{\mathcal{D}}}}[Y_{\tilde{\mathcal{D}}(X)}|\tilde{\mathcal{D}}, X]|X]]$.

In contrast, by only considering deterministic rules, a minimax regret approach guarantees the worst case regret for $X = x$ which is no more than

$$\min(\max\{\mathcal{U}(x), 0\}, \max\{-\mathcal{L}(x), 0\}).$$

It is clear that

$$-\frac{\mathcal{L}(x)\mathcal{U}(x)}{\mathcal{U}(x) - \mathcal{L}(x)} < \min\{-\mathcal{L}(x), \mathcal{U}(x)\} \text{ if } \mathcal{L}(x) < 0 < \mathcal{U}(x).$$

Therefore, the proposed mixed strategy gives a sharper minimax regret bound than Zhang and Pu (2021) and Pu and Zhang (2021), and therefore is sharper than any deterministic rules.

Remark 3. *The result in this section does not necessarily rely on $\mathcal{L}(x)$ being defined as $\mathcal{L}_1(x) - \mathcal{U}_{-1}(x)$ and $\mathcal{U}(x)$ being defined as $\mathcal{U}_1(x) - \mathcal{L}_{-1}(x)$.*

Remark 4. *The proposed mixed strategy leads to $w(x)$ or $w_r(x)$ a Bernoulli random variable with probability $p(x)$, and therefore a stochastic rule $\mathcal{D}(w(x), x)$ or $\mathcal{D}(w_r(x), x)$ assigning 1 with probability $p(x)$. Note that $w_r(x)$ being a Bernoulli random variable with parameter $p(x)$, and $w_r(x)$ being a scalar $p(x)$ are fundamentally different: the former one provides a stochastic decision rule. In other words, participants with the same x can receive different recommendations; while the latter one leads to a deterministic rule. That is, all participants with the same x receive the same recommendation.*

5.4. No universal optimality for decision making under partial identification. As can be easily seen from Table 1 as well as Section 5.3, there is a mismatch between deterministic/randomized minimax regret and maximin utility. In fact, each of the three rules corresponds to a different decision strategy. Such mismatch is a distinctive feature of partial identification.

On the one hand, it is notable that $\{\mathcal{L}(x), \mathcal{U}(x)\}$ provides complementary information to the analyst as it might inform the analyst as to when he/she might refrain from making a decision; mainly, if such an interval includes zero so that there is no evidence in the data as to whether the action/treatment is on average beneficial or harmful for individuals with that value of x . One might need to conduct randomized experiments in order to draw a causal conclusion if $0 \in (\mathcal{L}(x), \mathcal{U}(x))$. On the other hand, the decision making must in general be considered a game of four numbers $\{\mathcal{L}_1(x), \mathcal{L}_{-1}(x), \mathcal{L}(x), \mathcal{U}(x)\}$ rather than two, e.g., $\{\mathcal{L}_1(x), \mathcal{L}_{-1}(x)\}$ or $\{\mathcal{L}(x), \mathcal{U}(x)\}$.

From the above point of view, the concept of optimality of a decision rule under partial identification cannot be absolute, rather, it is relative to a particular choice of decision making criterion, whether it is minimax, maximax, maximin, etc. Furthermore, an individualized decision rule might

incorporate participants’ risk preference as it might be unreasonable to assume everyone sharing a common preference. In the Appendix, we provide expressions for the minimum utility, maximum regret, and maximum mis-classification rate of certain “optimal” rules in Table 1 (including maximin utility and deterministic/randomized minimax regret rule) for practical uses.

6. A PARADOX: $1+1<2$

In this section, we provide an interesting paradox regarding the use of partial identification to conduct individualized decision making. To streamline our presentation, we use (deterministic) minimax regret rule as a running example, however, any rule $\mathcal{D} \in \mathcal{D}^{\text{opt}}$ can suffer the same paradox. To simplify exposition, we consider the case with no U , that unbeknownst to the analyst unmeasured confounding is absent. We consider the following model with covariate X (e.g., female/male) distributed on $\{0, 1\}$ with equal probabilities,

$$\begin{aligned} \Pr(Y = 1|X, A) &= X/16 + 1/5A + 1/15, \\ \Pr(A = 1|X, Z) &= X/16 + 2/5Z + 1/2, \\ Z &\sim \text{Bernoulli}(1/2), \end{aligned}$$

with a slight abuse of notation, we use 0, 1 coding for Z, A here. It is easy to see that the optimal rule is $\mathcal{D}^* = 1$ for the entire population. After a simple calculation, the Balke-Pearl conditional average treatment effect bounds for $X = 0, 1$ both contain zero with $|\mathcal{L}(0)| < |\mathcal{U}(0)|$ and $|\mathcal{L}(1)| > |\mathcal{U}(1)|$. The Balke-Pearl average treatment effect bounds marginalizing over X also contain zero and $|\mathcal{L}| < |\mathcal{U}|$.

As it is unbeknownst to the analyst whether unmeasured confounding is present or whether if X is an effect modifier, there are several possible strategies for the analyst to analyze the data.

- (1) If one is concerned about individualized decision making but does not worry about unmeasured confounding, one runs a standard regression type analysis and gets the right answer.
- (2) If one is concerned about unmeasured confounding but is only interested in decision making based on the population level (i.e., based on average treatment effect analysis), one can obtain IV bounds on the average treatment effect and also get the right answer.
- (3) If one is concerned about individualized decision making and also worries about unmeasured confounding, one gets the wrong answer for a subgroup.

We summarize results of the above strategies of analyses in Table 2.

	$X = 0$	$X = 1$
(1)	✓	✓
(2)	✓	✓
(3)	✓	×

Table 2. Correct/incorrect decisions using three types of data analyses.

As can be seen from the table, mixing up two very difficult domains (individualized recommendation + unmeasured confounding) might make life harder ($1+1<2$). There are several lessons one can learn from this paradox:

- a) A comparison between (1) and (3): it would be a good idea to first conduct a standard analysis (e.g., assume Assumption 1) or other point identification approaches (e.g., assume Assumption 7

of Cui and Tchetgen Tchetgen, 2021b) and then use IV bounds as a sanity check or say policy improvement;

b) A comparison between (2) and (3): the paradox sheds light on the clear need for carefully distinguish variables used to make individualized decisions from variables used to address confounding concerns; similar to but different from Simpson’s paradox, the aggregated and disaggregated answers can be opposite for a substantial subgroup.

c) (3) by itself: it might be a rather risky undertaking to narrow down an interval estimate to a definite decision given the overwhelming uncertainty; overly accounting for unmeasured confounding might erroneously recommend a sub-optimal decision to a subgroup.

As motivated by the comparison between (1) and (3), we formalize the policy improvement idea following Kallus and Zhou (2018). Note that minimizing the worst-case possible regret against a baseline policy \mathcal{D}_0 would improve upon those individuals for whom $\mathcal{D}_0(X) = -1, \mathcal{L}(X) > 0$ and $\mathcal{D}_0(X) = 1, \mathcal{U}(X) < 0$. We revisit the real data example in Section 4. We first run a standard analysis (random forest: Y on X, A) and obtain $\mathcal{D}_0(X) = \text{sign}\{\Pr(Y|X, A = 1) - \Pr(Y|X, A = -1)\}$; among 3010 subjects, 2106 have $\mathcal{D}_0(X) = 1$ and 904 have $\mathcal{D}_0(X) = -1$. Then we calculate IV conditional average treatment effect bounds, and there are 323 subjects with $\mathcal{L}(X) > 0$ and 45 subjects with $\mathcal{U}(X) < 0$. Then we use IV bounds as a sanity check/improvement: only 4 subjects with $\mathcal{D}_0(X) = -1$ switch to 1, and 8 subjects with $\mathcal{D}_0(X) = 1$ switch to -1 . Therefore, for most subjects in this application, the IV bounds do not necessarily invalidate the standard regression analysis, while IV bounds are still helpful to validate/invalidate decisions for a subgroup.

7. DISCUSSION

In this paper, we illustrated how one might pursue individualized decision making using partial identification in a comprehensive manner. We established a formal link between individualized decision making under partial identification and classical decision theory by considering a lower bound perspective of value/utility function. Building on this unified framework, we provided a novel minimax solution for opportunists who fail to lose. We also pointed out that there is a mismatch between maximin utility and minimax regret. Moreover, we provided an interesting paradox to ground several interesting ideas on individualized decision making and unmeasured confounding. To conclude, we list the following points which might be worth considering in future research.

- As the proper use of multiple IVs is of growing interest in a lot of applications including statistical genetics studies, one could possibly construct multiple IVs and then try to find multiple bounds to conduct a better sanity check or improvement. Another possibility is to strengthen multiple IVs (Ertefaie et al., 2018; Zubizarreta et al., 2013). A stronger IV might provide a tighter bound, and therefore a sign identification may be achieved (Cui & Tchetgen Tchetgen, 2020).
- The proposed minimax regret method by leveraging a randomization scheme might be of interest in optimal control settings such as reinforcement learning and contextual bandit where exploitation and exploration are under consideration. In addition, given observational data in which a potential IV is available, one can use different strategies to construct an initial randomized policy for use in a reinforcement learning and bandit algorithm.
- One important difference between decision making with IV partial identification and classical decision theory is the source of uncertainty. For the former one, unmeasured confounding creates uncertainty, and overthinking confounding might create overwhelming uncertainty.

Therefore, to better assess the uncertainty, it would also be of great interest to formalize a sensitivity analysis procedure for point identification such as under assumptions of no unmeasured confounding or no unmeasured common effect modifiers (Cui & Tchetgen Tchetgen, 2021b). A similar question has also been raised by Han (2021).

Disclosure Statement. The author has no conflicts of interest to declare.

Acknowledgments. The author is thankful to three referees, action editor, and editor-in-chief for useful comments which led to an improved manuscript.

APPENDIX A. DERIVATION OF LOWER BOUNDS OF VALUE FUNCTION

The following was originally derived in Cui and Tchetgen Tchetgen (2021b). It is helpful to provide it here.

Proof. Note that

$$\begin{aligned} E[Y_{\mathcal{D}(X)}|X] &= E(Y_1|X)I\{\mathcal{D}(X) = 1\} + E(Y_{-1}|X)I\{\mathcal{D}(X) = -1\}, \\ E[Y_{\mathcal{D}(X)}|X] &= E(Y_1 - Y_{-1}|X)I\{\mathcal{D}(X) = 1\} + E(Y_{-1}|X), \\ E[Y_{\mathcal{D}(X)}|X] &= E(Y_{-1} - Y_1|X)I\{\mathcal{D}(X) = -1\} + E(Y_1|X). \end{aligned}$$

By $\mathcal{L}_{-1}(X) \leq E(Y_{-1}|X) \leq \mathcal{U}_{-1}(X)$ and $\mathcal{L}_1(X) \leq E(Y_1|X) \leq \mathcal{U}_1(X)$, one has the following bounds,

$$\begin{aligned} &(1 - w(X))[\mathcal{L}(X)I\{\mathcal{D}(X) = 1\} + \mathcal{L}_{-1}(X)] + w(X)[-U(X)I\{\mathcal{D}(X) = -1\} + \mathcal{L}_1(X)] \\ \text{(A.1)} \quad &\leq \mathcal{L}_1(X)I\{\mathcal{D}(X) = 1\} + \mathcal{L}_{-1}(X)I\{\mathcal{D}(X) = -1\} \leq E[Y_{\mathcal{D}(X)}|X], \end{aligned}$$

where $0 \leq w(x) \leq 1$ for any x . Therefore, we complete the proof by taking expectations on both sides of Equation (A.1). \square

APPENDIX B. MINIMUM UTILITY, MAXIMUM REGRET, AND MAXIMUM MIS-CLASSIFICATION RATE OF SEVERAL ‘‘OPTIMAL’’ RULES

We give the minimum value function, maximum regret, and maximum mis-classification rate over $\mathcal{D} \in \mathcal{D}^{opt}$ expressed in terms of the observed data:

$$\begin{aligned} &E[\max(\mathcal{L}_{-1}(X), \mathcal{L}_1(X))I\{0 \notin (\mathcal{L}(X), \mathcal{U}(X))\} + \min(\mathcal{L}_{-1}(X), \mathcal{L}_1(X))I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \\ &E[\max(|\mathcal{L}(X)|, |\mathcal{U}(X)|)I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \\ &E[I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \end{aligned}$$

respectively. While the maximum mis-classification rate remains the same, the minimum value function and maximum regret for a given \mathcal{D} can be different. For instance, the minimum value function and maximum regret of the maximin rule in Table 1 are:

$$\begin{aligned} &E[\max(\mathcal{L}_{-1}(X), \mathcal{L}_1(X))], \\ &E\left[|\mathcal{L}(X)|I\{\mathcal{L}_{-1}(X) < \mathcal{L}_1(X)\} + |\mathcal{U}(X)|I\{\mathcal{L}_{-1}(X) > \mathcal{L}_1(X)\}I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \end{aligned}$$

respectively. The minimum value function and maximum regret of the minimax rule in Table 1 are:

$$\begin{aligned} &E\left[\max(\mathcal{L}_{-1}(X), \mathcal{L}_1(X))I\{0 \notin (\mathcal{L}(X), \mathcal{U}(X))\} \right. \\ &\quad \left. + [\mathcal{L}_1(X)I\{|\mathcal{L}(X)| < |\mathcal{U}(X)|\} + \mathcal{L}_{-1}(X)I\{|\mathcal{L}(X)| > |\mathcal{U}(X)|\}]I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \right. \\ &E[\min(|\mathcal{L}(X)|, |\mathcal{U}(X)|)I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \end{aligned}$$

respectively. The minimum value function and maximum regret of the randomized minimax rule in Section 5.3 are:

$$\begin{aligned} &E\left[\max(\mathcal{L}_{-1}(X), \mathcal{L}_1(X))I\{0 \notin (\mathcal{L}(X), \mathcal{U}(X))\} \right. \\ &\quad \left. + \left[\mathcal{L}_1(X)\frac{\mathcal{U}(X)}{\mathcal{U}(X) - \mathcal{L}(X)} + \mathcal{L}_{-1}(X)\frac{-\mathcal{L}(X)}{\mathcal{U}(X) - \mathcal{L}(X)}\right]I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}], \right. \\ &E\left[-\frac{\mathcal{L}(X)\mathcal{U}(X)}{\mathcal{U}(X) - \mathcal{L}(X)}I\{0 \in (\mathcal{L}(X), \mathcal{U}(X))\}\right], \end{aligned}$$

respectively.

REFERENCES

- Arrow, K. J., & Hurwicz, L. (1972). An optimality criterion for decision-making under ignorance. *Uncertainty and Expectations in Economics (Oxford)*.
- Athey, S., & Wager, S. (2021). Policy learning with observational data. *Econometrica*, 89(1), 133–161. <https://doi.org/10.3982/ECTA15732>
- Balke, A., & Pearl, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92(439), 1171–1176. <https://doi.org/10.1080/01621459.1997.10474074>
- Card, D. (1993). *Using geographic variation in college proximity to estimate the return to schooling* (tech. rep.). National Bureau of Economic Research.
- Chakraborty, B., & Moodie, E. (2013). *Statistical methods for dynamic treatment regimes*. Springer. <https://doi.org/10.1007/978-1-4614-7428-9>
- Cui, Y., & Tchetgen Tchetgen, E. (2020). On a necessary and sufficient identification condition of optimal treatment regimes with an instrumental variable. *arXiv preprint arXiv:2010.03390*.
- Cui, Y., & Tchetgen Tchetgen, E. (2021a). Machine intelligence for individualized decision making under a counterfactual world: A rejoinder. *Journal of the American Statistical Association*, 116(533), 200–206. <https://doi.org/10.1080/01621459.2021.1872580>
- Cui, Y., & Tchetgen Tchetgen, E. (2021b). A semiparametric instrumental variable approach to optimal treatment regimes under endogeneity (with discussion). *Journal of the American Statistical Association*, 116(533), 162–173. <https://doi.org/10.1080/01621459.2020.1783272>
- Ertefaie, A., Small, D. S., & Rosenbaum, P. R. (2018). Quantitative evaluation of the trade-off of strengthened instruments and sample size in observational studies. *Journal of the American Statistical Association*, 113(523), 1122–1134. <https://doi.org/10.1080/01621459.2017.1305275>
- Greenland, S. (2000). An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology*, 29(4), 722–729. <https://doi.org/10.1093/ije/29.4.722>
- Han, S. (2019). Optimal dynamic treatment regimes and partial welfare ordering. *arXiv preprint arXiv:1912.10014*.
- Han, S. (2020). Identification in nonparametric models for dynamic treatment effects. *Journal of Econometrics*. <https://doi.org/10.1016/j.jeconom.2019.08.014>
- Han, S. (2021). Comment: Individualized treatment rules under endogeneity. *Journal of the American Statistical Association*, 116(533), 192–195. <https://doi.org/10.1080/01621459.2020.1831923>
- Hernan, M., & Robins, J. (2006). Instruments for causal inference: An epidemiologist’s dream? *Epidemiology (Cambridge, Mass.)*, 17, 360–72. <https://doi.org/10.1097/01.ede.0000222409.00878.37>
- Kallus, N., Mao, X., & Zhou, A. (2019). Interval estimation of individual-level causal effects under unobserved confounding. In K. Chaudhuri & M. Sugiyama (Eds.), *Proceedings of machine learning research* (pp. 2281–2290). PMLR. <http://proceedings.mlr.press/v89/kallus19a.html>
- Kallus, N., & Zhou, A. (2018). Confounding-robust policy improvement. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), *Advances in neural information processing systems*. Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2018/file/3a09a524440d44d7f19870070a5ad42f-Paper.pdf>
- Kosorok, M. R., & Laber, E. B. (2019). Precision medicine. *Annual Review of Statistics and Its Application*, 6(1), 263–286. <https://doi.org/10.1146/annurev-statistics-030718-105251>
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomforest. *R News*, 2(3), 18–22. <https://CRAN.R-project.org/doc/Rnews/>
- Manski, C. F. (1990). Nonparametric bounds on treatment effects. *The American Economic Review*, 80(2), 319–323.

- Manski, C. F., & Pepper, J. V. (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica*, *68*, 997–1010. <https://doi.org/10.3386/t0224>
- Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(2), 331–355. <https://doi.org/10.1111/1467-9868.00389>
- Murphy, S. A., van der Laan, M. J., Robins, J. M., & Group, C. P. P. R. (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association*, *96*(456), 1410–1423. <https://doi.org/10.1198/016214501753382327>
- Okui, R., Small, D. S., Tan, Z., & Robins, J. M. (2012). Doubly robust instrumental variable regression. *Statistica Sinica*, 173–205. <https://doi.org/10.5705/ss.2009.265>
- Pu, H., & Zhang, B. (2021). Estimating optimal treatment rules with an instrumental variable: A partial approach. *Journal of the Royal Statistical Society: Series B*. <https://doi.org/10.1111/rssb.12413>
- Qian, M., & Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of statistics*, *39*(2), 1180. <https://doi.org/10.1214/10-AOS864>
- Qiu, H., Carone, M., Sadikova, E., Petukhova, M., Kessler, R. C., & Luedtke, A. (2021a). Optimal individualized decision rules using instrumental variable methods (with discussion). *Journal of the American Statistical Association*, *116*(533), 174–191. <https://doi.org/10.1080/01621459.2020.1745814>
- Qiu, H., Carone, M., Sadikova, E., Petukhova, M., Kessler, R. C., & Luedtke, A. (2021b). Rejoinder: Optimal individualized decision rules using instrumental variable methods. *Journal of the American Statistical Association*, *116*(533), 207–209. <https://doi.org/10.1080/01621459.2020.1865166>
- Robins, J. (1989). The analysis of randomized and non-randomized aids treatment trials using a new approach to causal inference in longitudinal studies. *US Public Health Service*.
- Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. *Proceedings of the second seattle Symposium in Biostatistics*, 189–326. https://doi.org/10.1007/978-1-4419-9076-1_11
- Rubin, D. B., & van der Laan, M. J. (2012). Statistical issues and limitations in personalized medicine research with clinical trials. *The international journal of biostatistics*, *8*(1), 18. <https://doi.org/10.1515/1557-4679.1423>
- Swanson, S. A., Hernán, M. A., Miller, M., Robins, J. M., & Richardson, T. S. (2018). Partial identification of the average treatment effect using instrumental variables: Review of methods for binary instruments, treatments, and outcomes. *Journal of the American Statistical Association*, *113*(522), 933–947. <https://doi.org/10.1080/01621459.2018.1434530>
- Tan, Z. (2006). Regression and weighting methods for causal inference using instrumental variables. *Journal of the American Statistical Association*, *101*(476), 1607–1618. <https://doi.org/10.1198/016214505000001366>
- Tsiatis, A. A., Davidian, M., Holloway, S. T., & Laber, E. B. (2019). *Dynamic treatment regimes: Statistical methods for precision medicine*. CRC Press. <https://doi.org/10.1201/9780429192692>
- Wang, L., Robins, J. M., & Richardson, T. S. (2017). On falsification of the binary instrumental variable model. *Biometrika*, *104*(1), 229–236. <https://doi.org/10.1093/biomet/asw064>
- Wang, L., & Tchetgen Tchetgen, E. (2018). Bounded, efficient and multiply robust estimation of average treatment effects using instrumental variables. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, *80*(3), 531. <https://doi.org/10.1111/rssb.12262>
- Yadlowsky, S., Namkoong, H., Basu, S., Duchi, J., & Tian, L. (2018). Bounds on the conditional and average treatment effect with unobserved confounding factors. *arXiv preprint arXiv:1808.09521*.

- Zhang, B., Tsiatis, A. A., Davidian, M., Zhang, M., & Laber, E. (2012). Estimating optimal treatment regimes from a classification perspective. *Stat*, *1*(1), 103–114. <https://doi.org/10.1002/sta.411>
- Zhang, B., Tsiatis, A. A., Laber, E. B., & Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics*, *68*(4), 1010–1018. <https://doi.org/10.1111/j.1541-0420.2012.01763.x>
- Zhang, B., & Pu, H. (2021). Discussion of Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020). *Journal of the American Statistical Association*, *116*(533), 196–199. <https://doi.org/10.1080/01621459.2020.1832500>
- Zhao, Y., Zeng, D., Rush, A. J., & Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, *107*(499), 1106–1118. <https://doi.org/10.1080/01621459.2012.695674>
- Zubizarreta, J. R., Small, D. S., Goyal, N. K., Lorch, S., Rosenbaum, P. R., et al. (2013). Stronger instruments via integer programming in an observational study of late preterm birth outcomes. *The Annals of Applied Statistics*, *7*(1), 25–50. <https://doi.org/10.1214/12-AOAS582>