Individualized Treatment Allocations with Distributional Welfare

Yifan Cui & Sukjin Han

Zhejiang University & University of Bristol

22 September 2023

Brown University

Policies for Heterogeneous Population

Individuals are heterogeneous

- so are their responses to treatments

When designing policies (i.e., treatment allocations), important to reflect this heterogeneity

$\Rightarrow$ individualized policies

Policies for Heterogeneous Population

Individuals are heterogeneous

- so are their responses to treatments

When designing policies (i.e., treatment allocations), important to reflect this heterogeneity

$\Rightarrow$ individualized policies

Policy design depends on policymaker's specific objective

- utilitarian (i.e., sum or mean) (Manski 04)

  vs.

- non-utilitarian (e.g., prioritarian, maximin)

Exploring Non-Utilitarian Paradigm

There may be settings where utilitarian goal is less sensible

- especially when target population exhibits skewed heterogeneity (e.g., outliers)
- possibility of non-utilitarian welfare (Manski 04)

The purpose of this paper: To explore objectives of (non-utilitarian) policymaker who concerns...

- distribution (e.g., tails) of treatment effects
- vote shares

Preliminaries

Observables:

- Y: outcome; D: binary treatment; X: covariates

Unobservables:

- Y_d (for $d = 1, 0$): potential outcomes

Policy:

- $\delta : \mathcal{X} \to \mathcal{A} \subseteq [0, 1]$ is a treatment allocation rule based on X
  - e.g., $\mathcal{A} = \{0, 1\}$ corresponds to the deterministic rule
  - e.g., $\mathcal{A} = [0, 1]$ corresponds to the stochastic rule
- $\mathcal{D}$: (potentially constrained) space of δ

Preliminaries

Observables:

- Y: outcome; D: binary treatment; X: covariates

Unobservables:

- Y_d (for $d = 1, 0$): potential outcomes

Policy:

- $\delta : \mathcal{X} \to \mathcal{A} \subseteq [0, 1]$ is a treatment allocation rule based on X
  - e.g., $\mathcal{A} = \{0, 1\}$ corresponds to the deterministic rule
  - e.g., $\mathcal{A} = [0, 1]$ corresponds to the stochastic rule
- $\mathcal{D}$: (potentially constrained) space of δ

A policymaker (PM) wants to choose $\delta \in \mathcal{D}$ that optimizes a certain welfare criterion

Review: Mean (Utilitarian) Welfare

Utilitarian PM is interested in optimal policy δ^*_{ATE} that satisfies

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X)Y_1 + (1 - \delta(X))Y_0]$$

- with deterministic rule, the criterion can be written as $E[Y_{\delta(X)}]$

Review: Mean (Utilitarian) Welfare

Utilitarian PM is interested in optimal policy δ^*_{ATE} that satisfies

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X)Y_1 + (1-\delta(X))Y_0]$$

- with deterministic rule, the criterion can be written as $E[Y_{\delta(X)}]$

Because

$$\begin{aligned} E[\delta(X)Y_1 + (1-\delta(X))Y_0] &= E[Y_0 + \delta(X)(Y_1 - Y_0)] \\ &= E[Y_0] + E[\delta(X)E[Y_1 - Y_0|X]], \end{aligned}$$

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X)E[Y_1 - Y_0|X]]$$

- conditional ATE as "welfare gain"
- subject to constraints, δ^*_{ATE} maximizes the average of conditional ATE selected (or weighted) by δ

Review: Mean (Utilitarian) Welfare

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) E[Y_1 - Y_0 | X]]$$

- with no constraint, $\delta^*_{ATE}(x) = 1\{E[Y_1 - Y_0 | X = x] \geq 0\}$

Review: Mean (Utilitarian) Welfare

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) E[Y_1 - Y_0 | X]]$$

- with no constraint, $\delta^*_{ATE}(x) = 1\{E[Y_1 - Y_0 | X = x] \geq 0\}$
- can be sensitive to outliers

Review: Mean (Utilitarian) Welfare

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) E[Y_1 - Y_0 | X]]$$

- with no constraint, $\delta^*_{ATE}(x) = 1\{E[Y_1 - Y_0 | X = x] \geq 0\}$
- can be sensitive to outliers
- e.g., given $X = x$, few individuals with high $Y_1 - Y_0$ can make $E[Y_1 - Y_0 | X = x] > 0$

$Y_1 - Y_0 \mid X = x$
7
5
-1
-1
-1
-1
-1
-2
-2
-2

$$E[Y_1 - Y_0 | X = x] = 1/10$$

Review: Mean (Utilitarian) Welfare

$$\delta^*_{ATE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) E[Y_1 - Y_0 | X]]$$

- with no constraint, $\delta^*_{ATE}(x) = 1\{E[Y_1 - Y_0 | X = x] \geq 0\}$
- can be sensitive to outliers
- e.g., given $X = x$, few individuals with high $Y_1 - Y_0$ can make $E[Y_1 - Y_0 | X = x] > 0$
- suggests to treat *all* individuals with $X = x$ even though the treatment harms the majority

$Y_1 - Y_0 \| X = x$
7
5
-1
-1
-1
-1
-1
-2
-2
-2

$$E[Y_1 - Y_0 | X = x] = 1/10$$

This Paper: Quantile of Treatment Effects as Welfare Gain

We propose

$$\delta^* \equiv \delta^*_{QoTE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0 | X)]$$

- $Q_\tau(Y_1 - Y_0|X)$ is τ-quantile of $Y_1 - Y_0$ (QoTE) given X

This Paper: Quantile of Treatment Effects as Welfare Gain

We propose

$$\delta^* \equiv \delta^*_{QoTE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- $Q_\tau(Y_1 - Y_0|X)$ is τ-quantile of $Y_1 - Y_0$ (QoTE) given X
- δ^* maximizes the average of conditional QoTE selected (or weighted) by δ
- with no constraint, $\delta^*(x) = 1\{Q_\tau(Y_1 - Y_0|X = x) \geq 0\}$

This Paper: Quantile of Treatment Effects as Welfare Gain

We propose

$$\delta^* \equiv \delta^*_{QoTE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- $Q_\tau(Y_1 - Y_0|X)$ is τ-quantile of $Y_1 - Y_0$ (QoTE) given X
- δ^* maximizes the average of conditional QoTE selected (or weighted) by δ
- with no constraint, $\delta^*(x) = 1\{Q_\tau(Y_1 - Y_0|X = x) \geq 0\}$
- τ (i.e., rank in individual TEs) represents a reference group chosen by the PM

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- decision less sensitive to outliers
  - "within-group fairness"
    (Leqi & Kennedy 21)

$Y_1 - Y_0|X = x$

7
5
-1
-1
-1
-1
-1
-2
-2
-2

$$Q_{0.5}(Y_1 - Y_0|X = x) = -1$$

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0 | X)]$$

- decision less sensitive to outliers
  - "within-group fairness" (Leqi & Kennedy 21)
- τ is chosen by the PM to set a reference group

$Y_1 - Y_0 | X = x$

7
5
-1
-1
-1
-1
-1
-2
-2
-2

$$Q_{0.5}(Y_1 - Y_0 | X = x) = -1$$

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0 | X)]$$

- decision less sensitive to outliers
  - "within-group fairness" (Leqi & Kennedy 21)
- τ is chosen by the PM to set a reference group
  - large τ: negligent PM

$Y_1 - Y_0 | X = x$

7
5
-1
-1
-1
-1
-1
-2
-2
-2

$$Q_{0.9}(Y_1 - Y_0 | X = x) = 5$$

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- decision less sensitive to outliers
  - "within-group fairness" (Leqi & Kennedy 21)
- τ is chosen by the PM to set a reference group
  - large τ: negligent PM
  - small τ: prudent PM

$Y_1 - Y_0|X = x$

7
5
−1
−1
−1
−1
−1
−2
−2
−2

$Q_{0.2}(Y_1 - Y_0|X = x) = -2$

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0 | X)]$$

- decision less sensitive to outliers
  - "within-group fairness" (Leqi & Kennedy 21)
- τ is chosen by the PM to set a reference group
  - large τ: negligent PM
  - small τ: prudent PM

$Y_1 - Y_0 | X = x$

7
5
2
2
2
1
1
1
1
-2

$$Q_{0.2}(Y_1 - Y_0 | X = x) = 1$$

Alternatives in Literature: Quantile Welfare

Wang, Zhou, Song & Sherwood 18: Quantile of $Y_{\delta(X)}$, focusing on deterministic regime

$$\delta^*_? \in \arg\max_{\delta \in \mathcal{D}} Q_\tau(Y_{\delta(X)})$$

- no closed-form solution for optimal policy $\delta^*_?$
  - interpretation of welfare gain is elusive
- lack of "across-group fairness" (Leqi & Kennedy 21):
  - decision for one group is influenced by TEs of other groups

Alternatives in Literature: Quantile Welfare

Leqi & Kennedy 21: Average of conditional quantile, focusing on deterministic regime

$$\delta^*_{QTE} \in \arg\max_{\delta \in \mathcal{D}} E[Q_\tau(Y_{\delta(X)}|X)]$$

Alternatives in Literature: Quantile Welfare

Leqi & Kennedy 21: Average of conditional quantile, focusing on deterministic regime

$$\delta^*_{QTE} \in \arg\max_{\delta \in \mathcal{D}} E[Q_\tau(Y_{\delta(X)}|X)]$$

But because

$$\begin{aligned} E[Q_\tau(Y_{\delta(X)}|X)] &= E[\delta(X)Q_\tau(Y_1|X) + (1-\delta(X))Q_\tau(Y_0|X)] \\ &= E[Q_\tau(Y_0|X)] + E[\delta(X)\{Q_\tau(Y_1|X) - Q_\tau(Y_0|X)\}], \end{aligned}$$

$$\delta^*_{QTE} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X)\{Q_\tau(Y_1|X) - Q_\tau(Y_0|X)\}]$$

- δ^*_{QTE} maximizes the average of conditional QTE selected by δ
- with no constraint, $\delta^*_{QTE}(x) = 1\{Q_\tau(Y_1|X=x) - Q_\tau(Y_0|X=x) \geq 0\}$

Alternatives in Literature: Quantile Welfare

$$\delta^*_{QTE} \in \arg\max_{\delta \in \mathcal{D}} E[Q_\tau(Y_{\delta(X)}|X)]$$

- with no constraint, $\delta^*_{QTE}(x)$
  $= 1\{Q_\tau(Y_1|X = x)$
  $-Q_\tau(Y_0|X = x) \geq 0\}$

Alternatives in Literature: Quantile Welfare

$$\delta^*_{QTE} \in \arg\max_{\delta \in \mathcal{D}} E[Q_\tau(Y_{\delta(X)}|X)]$$

- with no constraint, $\delta^*_{QTE}(x) = 1\{Q_\tau(Y_1|X = x) - Q_\tau(Y_0|X = x) \geq 0\}$
- QTE is difference of Q_τ's of potentially different individuals
  - hard to justify esp. in making individualized decision
  - hard to define prudence or negligence

$Y_1 \mid X = x$	$Y_0 \mid X = x$
2	0
7	2
2	1
1	0
5	3
2	1
1	0
0	0
1	1
1	0

$$Q_{0.5}(Y_1) - Q_{0.5}(Y_0) = 1$$

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- still, the notion of welfare *level* is unclear

This Paper: Quantile of Treatment Effects as Welfare Gain

$$\delta^* \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- still, the notion of welfare *level* is unclear

Another interpretation of $\delta^*(x) = 1\{Q_\tau(Y_1 - Y_0|X = x) \geq 0\}$:

- suppose individuals who benefit from treatment vote for it
- with $\tau = 0.5$, δ^* is a policy that obeys *majority vote*:

$$\begin{aligned} & Q_{0.5}(Y_1 - Y_0|X) \geq 0 \\ \Leftrightarrow \quad & F_{Y_1 - Y_0|X}(0) \leq 1/2 \\ \Leftrightarrow \quad & P[Y_1 \geq Y_0|X] \geq 1/2 \\ \Leftrightarrow \quad & P[Y_1 \geq Y_0|X] \geq P[Y_1 < Y_0|X] \end{aligned}$$

- consistent with a PM who has political incentive and whose decision is influenced by vote shares

This Paper: Quantile of Treatment Effects as Welfare Gain

- can be generalized by considering $Q_{0.5-\alpha/2}(Y_1 - Y_0|X) \geq 0$, which is equivalent to

$$P[Y_1 \geq Y_0|X] \geq P[Y_1 < Y_0|X] + \alpha$$

where $\alpha \geq 0$ is vote share margin

Optimal Policies Robust to Model Ambiguity

QoTE is generally not point-identified even under unconfoundedness

Optimal Policies Robust to Model Ambiguity

QoTE is generally not point-identified even under unconfoundedness

We propose an optimal policy robust to model ambiguity:

$$\delta^*_{mmw} \in \arg\max_{\delta \in \mathcal{D}} \min_{F_{Y_1,Y_0|X} \in \mathcal{F}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- $\mathcal{F} \equiv \mathcal{F}(P)$ is the identified set of $F_{Y_1,Y_0|X}$ given data P

Optimal Policies Robust to Model Ambiguity

QoTE is generally not point-identified even under unconfoundedness

We propose an optimal policy robust to model ambiguity:

$$\delta^*_{mmw} \in \arg\max_{\delta \in \mathcal{D}} \min_{F_{Y_1,Y_0|X} \in \mathcal{F}} E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$$

- $\mathcal{F} \equiv \mathcal{F}(P)$ is the identified set of $F_{Y_1,Y_0|X}$ given data P

Alternatively,

$$\delta^*_{mmr} \in \arg\min_{\delta \in \mathcal{D}} \max_{F_{Y_1,Y_0|X} \in \mathcal{F}} E[\{\delta^\dagger(X) - \delta(X)\} Q_\tau(Y_1 - Y_0|X)]$$

- $\delta^\dagger = 1\{Q_\tau(Y_1 - Y_0|\cdot) \geq 0\} \in \arg\max_\delta E[\delta(X) Q_\tau(Y_1 - Y_0|X)]$ is the first-best policy

Optimal Policies Robust to Model Ambiguity

Define the identified interval for $Q_\tau(Y_1 - Y_0|X = x)$:

$$[Q_\tau^L(x), Q_\tau^U(x)] = \{Q_\tau(Y_1 - Y_0|X = x) : F_{Y_1,Y_0|X} \in \mathcal{F}\}$$

Assumption REC: The identified set $\mathcal{Q}(P)$ of $Q_\tau(Y_1 - Y_0|X)$ is rectangular, i.e.,

$$\mathcal{Q}(P) \equiv \{Q_\tau(Y_1 - Y_0|X = \cdot) : Q_\tau(Y_1 - Y_0|X = x) \in [Q_\tau^L(x), Q_\tau^U(x)]\}$$

- allows to interchange the max/min over $\mathcal{F}$ with the expectation over X (Kasy 16, D'Adamo 23)

▸ REC

Optimal Policies Robust to Model Ambiguity

Under REC, we can show

$$\delta^*_{mmw} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q^L_\tau(X)]$$

and

$$\delta^*_{mmr} \in \arg\max_{\delta \in \mathcal{D}} E\left[\delta(X)\Lambda(X)\right]$$

where

$$\begin{aligned}\Lambda(x) &= Q^U_\tau(x) \cdot 1\{Q^U_\tau(x) \geq 0\} + Q^L_\tau(x) \cdot 1\{Q^L_\tau(x) \leq 0\} \\ &= Q^U_\tau(x) \cdot 1\{Q^L_\tau(x) \geq 0\} + Q^L_\tau(x) \cdot 1\{Q^U_\tau(x) \leq 0\} \\ &\quad + \left(|Q^U_\tau(x)| - |Q^L_\tau(x)|\right) \cdot 1\{Q^L_\tau(x) < 0 < Q^U_\tau(x)\}\end{aligned}$$

Related Literature

Treatment choice and policy learning:

- Manski 04, 09, Hirano and Porter 09, Stoye 12, Kitagawa & Tetenov 18, Athey & Wager 21, Mbakop & Tabord-Meehan 21, Sakaguchi 21, Kitagawa, Sakaguchi, Tetenov 21, Ida, Ishihara, Ito, Kido, Kitagawa, Sakaguchi, Sasaki 22, among others
- Murphy, van der Laan & Robins 01, Murphy 03, Robins 04, Zhao, Zeng, Rush & Kosorok 12, Cui & Tchetgen Tchetgen 21, among others
- Wang et al. 18, Leqi & Kennedy 21, Kitagawa & Tetenov 21
- Kitagawa, Lee & Qiu 23

Treatment choice under ambiguity:

- Stoye 09, Kasy 16, Kallas & Zhou 21, Pu & Zhang 21, Cui 21, Yata 21, Han 23, D'Adamo 23

Possible Identifying Assumptions

Let $Q_\tau(x) \equiv Q_\tau(Y_1 - Y_0 | X = x)$ for simplicity

Possible Identifying Assumptions

Let $Q_\tau(x) \equiv Q_\tau(Y_1 - Y_0 | X = x)$ for simplicity

Informativeness of the bounds is useful

$$[Q_\tau^L(x), Q_\tau^U(x)] = \{Q_\tau(x) : F_{Y_1, Y_0|X} \in \mathcal{F}\}$$

We provide a range of identifying assumptions that the researcher may want to impose

- to shrink $\mathcal{F}$ and thus $[Q_\tau^L(x), Q_\tau^U(x)]$,
- sometimes to a singleton

Possible Identifying Assumptions

Let $Q_\tau(x) \equiv Q_\tau(Y_1 - Y_0|X = x)$ for simplicity

Informativeness of the bounds is useful

$$[Q_\tau^L(x), Q_\tau^U(x)] = \{Q_\tau(x) : F_{Y_1,Y_0|X} \in \mathcal{F}\}$$

We provide a range of identifying assumptions that the researcher may want to impose

- to shrink $\mathcal{F}$ and thus $[Q_\tau^L(x), Q_\tau^U(x)]$,
- sometimes to a singleton

First, to identify the marginal distribution of Y_d:

Assumption CI (Conditional Independence): $Y_d \perp D|X$ for $d \in \{0, 1\}$.

- Alternatively, panel quantile regression models can be used to identify $Q_\tau(Y_d|X)$ (Chernozhukov, Fernandez-Val, Hahn & Newey 13)

Possible Identifying Assumptions

No-assumption bounds on $Q_\tau(x)$ (besides CI):

- Makarov 81, Williamson & Downs 90
- may be uninformative

Possible Identifying Assumptions

No-assumption bounds on $Q_\tau(x)$ (besides CI):

- Makarov 81, Williamson & Downs 90
- may be uninformative

Assumption SI (Stochastic Increasing): For given $x \in \mathcal{X}$, $P[Y_1 \leq y_1 | Y_0 = \cdot, X = x]$ and $P[Y_0 \leq y_0 | Y_1 = \cdot, X = x]$ are nonincreasing.

- SI + CI produce informative bounds
- Frandsen & Lefgren 21

Possible Identifying Assumptions

No-assumption bounds on $Q_\tau(x)$ (besides CI):

- Makarov 81, Williamson & Downs 90
- may be uninformative

Assumption SI (Stochastic Increasing): For given $x \in \mathcal{X}$, $P[Y_1 \leq y_1 | Y_0 = \cdot, X = x]$ and $P[Y_0 \leq y_0 | Y_1 = \cdot, X = x]$ are nonincreasing.

- SI + CI produce informative bounds
- Frandsen & Lefgren 21

Assumption SD (Stochastic Dominance): For given $x \in \mathcal{X}$, (i) $P[Y_d \leq y | D = 1, X = x] \leq P[Y_d \leq y | D = 0, X = x]$; or (ii) $P[Y_1 \leq y | D = d, X = x] \leq P[Y_0 \leq y | D = d, X = x]$.

- SD(i) or SD(ii), without CI or with instruments
- Blundell, Gosling, Ichimura & Meghir 07, Lee 23

Possible Identifying Assumptions

Here are assumptions for point identification of $Q_\tau(x)$

Assumption CI2 (Joint Conditional Independence):
$(Y_1, Y_0) \perp D|X$.

Assumption DC (Deconvolution): $Y_1 - Y_0 \perp Y_0|X$.

- CI2 + DC point-identify $Q_\tau(x)$ (Heckman & Smith 95)

Possible Identifying Assumptions

Here are assumptions for point identification of $Q_\tau(x)$

Assumption CI2 (Joint Conditional Independence):
$(Y_1, Y_0) \perp D|X$.

Assumption DC (Deconvolution): $Y_1 - Y_0 \perp Y_0|X$.

- CI2 + DC point-identify $Q_\tau(x)$ (Heckman & Smith 95)

Assumption RY (Roy Model): (i) $D = 1\{Y_1 \geq Y_0\}$; (ii) large support of elements of X; (iii) additive errors in Y_d-equations.

Assumption RY2 (Extended Roy Model):
(i) $D = 1\{Y_1 \geq h(Y_0, X, Z)\}$; (ii) strict monotonicity of h;
(iii) $(Y_0, Y_1) \perp Z|X$.

- RY or RY2 point-identifies $Q_\tau(x)$ (Heckman & Smith 95, Lee & Park 22)

Possible Identifying Assumptions

Assumption RI (Rank Invariance): (i) $Y_d = m_d(X, U_d)$; (ii) $m_d(x, \cdot)$ is strictly increasing; (iii) $U_1|_{X=x} = U_0|_{X=x}$.

- Heckman, Smith & Clements 97, Chernozhukov & Hansen 05
- generalized version in Heckman, Smith & Clements 97

Possible Identifying Assumptions

Assumption RI (Rank Invariance): (i) $Y_d = m_d(X, U_d)$; (ii) $m_d(x, \cdot)$ is strictly increasing; (iii) $U_1|_{X=x} = U_0|_{X=x}$.

- Heckman, Smith & Clements 97, Chernozhukov & Hansen 05
- generalized version in Heckman, Smith & Clements 97

Assumption RY (Mutual Independence): $Y_1 \perp Y_0 | X, C$ for some variable C.

- RY + CI point-identify $Q_\tau(x)$
- relates to factor models (Abbring & Heckman 07)

Possible Identifying Assumptions

Assumption RI (Rank Invariance): (i) $Y_d = m_d(X, U_d)$; (ii) $m_d(x, \cdot)$ is strictly increasing; (iii) $U_1|_{X=x} = U_0|_{X=x}$.

- Heckman, Smith & Clements 97, Chernozhukov & Hansen 05
- generalized version in Heckman, Smith & Clements 97

Assumption RY (Mutual Independence): $Y_1 \perp Y_0 | X, C$ for some variable C.

- RY + CI point-identify $Q_\tau(x)$
- relates to factor models (Abbring & Heckman 07)

Assumption SYM (Symmetric Distribution): The distribution of $Y_1 - Y_0 | X$ is symmetric.

- Then $Q_{0.5}(Y_1 - Y_0 | X) = E[Y_1 - Y_0 | X]$, which is point-identified under CI

Calculating Bounds on $Q_\tau(x)$

Let $C(u,v|X)$ be the copula for $(U,V) \equiv (F_{Y_1}(Y_1), F_{Y_0}(Y_0))$ conditional on X

Then, by Sklar's Theorem,

$$\begin{aligned} P[Y_1 - Y_0 \le t|X] &= P[F^{-1}_{Y_1|X}(U) - F^{-1}_{Y_0|X}(V) \le t|X] \\ &= \int 1\{F^{-1}_{Y_1|X}(u) - F^{-1}_{Y_0|X}(v) \le t\} dC(u,v|X) \end{aligned}$$

Calculating Bounds on $Q_\tau(x)$

Let $C(u, v|X)$ be the copula for $(U, V) \equiv (F_{Y_1}(Y_1), F_{Y_0}(Y_0))$ conditional on X

Then, by Sklar's Theorem,

$$\begin{aligned} P[Y_1 - Y_0 \leq t|X] &= P[F^{-1}_{Y_1|X}(U) - F^{-1}_{Y_0|X}(V) \leq t|X] \\ &= \int 1\{F^{-1}_{Y_1|X}(u) - F^{-1}_{Y_0|X}(v) \leq t\} dC(u, v|X) \end{aligned}$$

Therefore, with $\Delta \equiv Y_1 - Y_0$,

$$F^L_{\Delta|X}(t) = \inf_{C(\cdot,\cdot|X)\in\mathcal{C}} \int 1\{F^{-1}_{Y_1|X}(u) - F^{-1}_{Y_0|X}(v) \leq t\} dC(u, v|X)$$

$$F^U_{\Delta|X}(t) = \sup_{C(\cdot,\cdot|X)\in\mathcal{C}} \int 1\{F^{-1}_{Y_1|X}(u) - F^{-1}_{Y_0|X}(v) \leq t\} dC(u, v|X)$$

where $\mathcal{C}$ is the class of copulas restricted by identifying assumptions

Calculating Bounds on $Q_\tau(x)$

For τ-quantile $Q_\tau(x)$ for $\Delta|X = x$, we can obtain its lower and upper bounds as

$$Q^L_\tau(X) = F^{U,-1}_{\Delta|X}(\tau)$$
$$Q^U_\tau(X) = F^{L,-1}_{\Delta|X}(\tau)$$

In practice, we need to approximate $C(u, v|x)$ to transform above optimization into linear programs

- two approaches

Calculating Bounds on $Q_\tau(x)$: Approach I

For Makarov bounds, consider (suppressing X)

$$F^L_\Delta(t) = \min_{c(\cdot,\cdot)} \sum_{j=1}^{k} \sum_{i=1}^{k} 1\{F^{-1}_{Y_1}(r(i)) - F^{-1}_{Y_0}(r(j)) \leq t\} c(i,j)$$

$$F^U_\Delta(t) = \max_{c(\cdot,\cdot)} \sum_{j=1}^{k} \sum_{i=1}^{k} 1\{F^{-1}_{Y_1}(r(i)) - F^{-1}_{Y_0}(r(j)) \leq t\} c(i,j)$$

where

$$r(i) = \frac{2i-1}{2k}$$

and

$$\sum_{s=1}^{k} c(s,j) = 1/k, \text{for } j = 1 \ldots k$$

$$\sum_{s=1}^{k} c(i,s) = 1/k, \text{for } i = 1 \ldots k$$

Calculating Bounds on $Q_\tau(x)$: Approach I

Additionally, e.g., Assumption SI imposes

$$\left\{ \left\{ \sum_{s=1}^{i} c(s,j) \geq \sum_{s=1}^{i} c(s,j+1) \right\}_{j=1}^{k-1} \right\}_{i=1}^{k}$$

$$\left\{ \left\{ \sum_{s=1}^{i} c(i,s) \geq \sum_{s=1}^{j} c(i+1,s) \right\}_{i=j}^{k-1} \right\}_{j=1}^{k}$$

Calculating Bounds on $Q_\tau(x)$: Approach II

Alternatively, we can approximate $C(u, v|x)$ using Bernstein copula $C_B(u, v|x)$ (Sancetta & Satchell 04)

Finally, $F_{Y_1|X}(y)$ and $F_{Y_0|X}(y)$ can be estimated using standard nonparametric or parametric methods

Theoretical Properties of Estimated Policy

Recall $Q_\tau(X) \equiv Q_\tau(Y_1 - Y_0|X)$ and our objective function is

$$V(\delta) \equiv E[\delta(X)Q_\tau(X)]$$

The regret of this "classification" is

$$R(\delta) \equiv V(\delta^\dagger) - V(\delta) = E[|Q_\tau(X)|1\{\delta(X) \neq sign(Q_\tau(X))\}]$$

- $sign(q) = 1$ when $q \geq 0$ and $sign(q) = 0$ when $q < 0$

Theoretical Properties of Estimated Policy

Recall $Q_\tau(X) \equiv Q_\tau(Y_1 - Y_0|X)$ and our objective function is

$$V(\delta) \equiv E[\delta(X) Q_\tau(X)]$$

The regret of this "classification" is

$$R(\delta) \equiv V(\delta^\dagger) - V(\delta) = E[|Q_\tau(X)| 1\{\delta(X) \neq sign(Q_\tau(X))\}]$$

- $sign(q) = 1$ when $q \geq 0$ and $sign(q) = 0$ when $q < 0$

$R(\delta)$ is not identified, thus we define maximum regret as

$$\bar{R}(\delta) \equiv \sup_{Q_\tau(\cdot) \in [Q_\tau^L(\cdot), Q_\tau^U(\cdot)]} E[|Q_\tau(X)| 1\{\delta(X) \neq sign(Q_\tau(X))\}]$$

Focus on the case where $\mathcal{D}$ is unrestricted

Theoretical Properties of Estimated Policy

Assumption EST: $F^{-1}_{\Delta|X}(\tau)$ is bounded a.s. and

$$\hat{Q}^L_\tau(X) - Q^L_\tau(X) = o_p(1),$$
$$\hat{Q}^U_\tau(X) - Q^U_\tau(X) = o_p(1).$$

- EST is implied by consistency of $\hat{F}_{Y_d|X}$ and consistency of the copula approximation

Theoretical Properties of Estimated Policy

Assumption EST: $F^{-1}_{\Delta|X}(\tau)$ is bounded a.s. and

$$\hat{Q}^L_\tau(X) - Q^L_\tau(X) = o_p(1),$$
$$\hat{Q}^U_\tau(X) - Q^U_\tau(X) = o_p(1).$$

- EST is implied by consistency of $\hat{F}_{Y_d|X}$ and consistency of the copula approximation

Theorem 1 (Regret Bounds)

Suppose EST holds. Then,

$$R(\hat{\delta}^{stoch}) \leq \bar{R}(\hat{\delta}^{stoch}) \leq E\left[\frac{Q^L_\tau(X)Q^U_\tau(X)}{Q^L_\tau(X)-Q^U_\tau(X)}1\{Q^L_\tau(X)<0<Q^U_\tau(X)\}\right]+o_p(1),$$

where the ratio $= 0$ when its denominator $= 0$, and

$$R(\hat{\delta}^{determ}) \leq \bar{R}(\hat{\delta}^{determ}) \leq E\left[\min(\max(Q^U_\tau(X),0),\max(-Q^L_\tau(X),0))\right]+o_p(1).$$

Theoretical Properties of Estimated Policy

Corollary 1 (Expected Regret Bounds)

Suppose EST holds. Then,

$$E_{P^n}\left[R(\hat{\delta}^{stoch})\right] \leq E\left[\frac{Q^L_\tau(X)Q^U_\tau(X)}{Q^L_\tau(X)-Q^U_\tau(X)}1\{Q^L_\tau(X)<0<Q^U_\tau(X)\}\right]+o(1),$$

where the ratio $=0$ when its denominator $=0$, and

$$E_{P^n}\left[R(\hat{\delta}^{determ})\right] \leq E\left[\min(\max(Q^U_\tau(X),0),\max(-Q^L_\tau(X),0))\right]+o(1).$$

Theoretical Properties of Estimated Policy

Corollary 1 (Expected Regret Bounds)

Suppose EST holds. Then,

$$E_{P^n}\left[R(\hat{\delta}^{stoch})\right] \leq E\left[\frac{Q_\tau^L(X)Q_\tau^U(X)}{Q_\tau^L(X)-Q_\tau^U(X)}1\{Q_\tau^L(X)<0<Q_\tau^U(X)\}\right]+o(1),$$

where the ratio $=0$ when its denominator $=0$, and

$$E_{P^n}\left[R(\hat{\delta}^{determ})\right] \leq E\left[\min(\max(Q_\tau^U(X),0),\max(-Q_\tau^L(X),0))\right]+o(1).$$

Leading term in each bound reduces to zero when either...

- bounds on $Q_\tau(X)$ excludes zero a.s.
- or $Q_\tau(X)$ is point-identified

Classification Method with Constrained Policy Class

Sometimes PM may be interested in parsimonious decision rules

- e.g., threshold policies with linear index

Classification Method with Constrained Policy Class

Sometimes PM may be interested in parsimonious decision rules

- e.g., threshold policies with linear index

Our proposed method readily extends to the case of constrained $\mathcal{D}$

Classification Method with Constrained Policy Class

Sometimes PM may be interested in parsimonious decision rules

- e.g., threshold policies with linear index

Our proposed method readily extends to the case of constrained $\mathcal{D}$

$$\delta^*_{mmw} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q^L_\tau(X)]$$

$$\delta^*_{mmr} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X)[Q^U_\tau(X) \cdot 1\{Q^U_\tau(X) \geq 0\} + Q^L_\tau(X) \cdot 1\{Q^L_\tau(X) \leq 0\}]]$$

We can consider convex relaxation by using hinge loss function $\phi(t) = \max(1 - t, 0)$ and adding regularization

- e.g., the outcome weighted learning framework (Zhao et al. 12)
- consistency with hinge loss is proved even when the classifier's prediction set is restricted (Kitagawa et al., 2021)

Classification Method with Constrained Policy Class

Sometimes PM may be interested in parsimonious decision rules

- e.g., threshold policies with linear index

Our proposed method readily extends to the case of constrained $\mathcal{D}$

$$\delta^*_{mmw} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X) Q^L_\tau(X)]$$

$$\delta^*_{mmr} \in \arg\max_{\delta \in \mathcal{D}} E[\delta(X)[Q^U_\tau(X) \cdot 1\{Q^U_\tau(X) \geq 0\} + Q^L_\tau(X) \cdot 1\{Q^L_\tau(X) \leq 0\}]]$$

We can consider convex relaxation by using hinge loss function $\phi(t) = \max(1 - t, 0)$ and adding regularization

- e.g., the outcome weighted learning framework (Zhao et al. 12)
- consistency with hinge loss is proved even when the classifier's prediction set is restricted (Kitagawa et al., 2021)

Then bound on $\bar{R}$ and thus bound on R can be obtained

Numerical Illustrations

Q: How policies differ across PM's criteria esp. when the QoTE is partially identified?

Numerical Illustrations

Q: How policies differ across PM's criteria esp. when the QoTE is partially identified?

Data-generating process:

- Draw (Y_1, Y_0) (or $(\log Y_1, \log Y_0)$) from $N(\mu, \Sigma)$ with $\mu = (\mu_1, \mu_0)'$ and

$$\Sigma = \begin{bmatrix} \sigma_1 & \rho_{10}\sqrt{\sigma_0\sigma_1} \\ \rho_{10}\sqrt{\sigma_0\sigma_1} & \sigma_0 \end{bmatrix}$$

- Draw D from $Bernoulli(0.5)$
- Generate $Y = DY_1 + (1 - D)Y_0$

Note: $Y_1|Y_0 \sim N(\mu_1 + \rho_{10}\sigma_1 Z_0, (1 - \rho^2\sigma_1))$ where $Z_0 = \frac{Y_0 - \mu_0}{\sigma_0}$

- thus SI holds when $\rho_{10} \geq 0$ (similarly for log-normal case)

Numerical Illustrations

When $\mathcal{D}$ is unrestricted, the true optimal policies are:

- $\delta^* = 1\{Q_\tau(Y_1 - Y_0) > 0\}$
  - e.g., $Q_\tau(Y_1 - Y_0) = \mu_1 - \mu_0 + \Phi^{-1}(\tau)\sqrt{\sigma_1^2 + \sigma_0^2 - 2\rho_{10}\sigma_1\sigma_0}$
- $\delta^*_{QTE} = 1\{Q_\tau(Y_1) - Q_\tau(Y_0) > 0\}$
  - e.g., $Q_\tau(Y_1) - Q_\tau(Y_0) = \mu_1 - \mu_0 + \Phi^{-1}(\tau)(\sigma_1 - \sigma_0)$
- $\delta^*_{ATE} = 1\{E[Y_1 - Y_0] > 0\}$
  - e.g., $E[Y_1 - Y_0] = \mu_1 - \mu_0$
- they are first best for both deterministic and stochastic policies
- under normality and SI, if $0 < \tau < 0.5$, we have...
  - $Q_\tau(Y_1 - Y_0) < Q_\tau(Y_1) - Q_\tau(Y_0)$ and
  - $Q_\tau(Y_1 - Y_0) < E(Y_1) - E(Y_0)$

For brevity, let $Q_\tau \equiv Q_\tau(Y_1 - Y_0)$

Numerical Illustrations

Unlike δ^*_{QTE} and δ^*_{ATE}, recall obtaining $\delta^* \equiv \delta^*_{QoTE}$ involves model uncertainty:

- for deterministic policy:

$$\delta^*_{mmr} = \begin{cases} 1\{|Q^U_\tau| \geq |Q^L_\tau|\} & \text{if } Q^L_\tau < 0 < Q^U_\tau \\ 1 & \text{if } Q^L_\tau \geq 0 \\ 0 & \text{if } Q^U_\tau \leq 0 \end{cases}$$

- for stochastic policy:

$$\delta^*_{mmr} \sim \begin{cases} Bernoulli\left(\frac{Q_U}{Q_U - Q_L}\right) & \text{if } Q^L_\tau < 0 < Q^U_\tau \\ 1 & \text{if } Q^L_\tau \geq 0 \\ 0 & \text{if } Q^U_\tau \leq 0 \end{cases}$$

In simulation, Q^L_τ and Q^U_τ are calculated under...

- no assumption (i.e., Makarov bounds) or Assumption SI

Numerical Illustrations

For δ^*_{mmr}, δ^*_{QTE} and δ^*_{ATE}, we estimate $\hat{\delta}^*$, $\hat{\delta}^*_{QTE}$ and $\hat{\delta}^*_{ATE}$

- by estimating Q^U_τ, Q^L_τ, $Q_\tau(Y_d)$, and $E[Y_d]$ $(d = 0, 1)$ using the generated experimental data

For $TE \in \{QoTE, QTE, ATE\}$, (recalling $\delta^* \equiv \delta^*_{QoTE}$)

- misclassification error: $E_{P^n}[1\{\hat{\delta}_{TE} \neq \delta^*_{TE}\}]$
- regret: $E_{P^n}[|TE| \cdot 1\{\hat{\delta}_{TE} \neq \delta^*_{TE}\}]$

We focus on $\tau = 0.25$

Numerical Results: Correct Classification Rate ($n = 50$)

	$\hat{\delta}^{SI,stoch}$	$\hat{\delta}^{stoch}$	$\hat{\delta}^{SI,determ}$	$\hat{\delta}^{determ}$	$\hat{\delta}_{QTE}$	$\hat{\delta}_{ATE}$
	Subgroup 1 (0 is not contained)					
δ^*	100%	100%	100%	100%	33.5%	93%
δ^*_{QTE}	0%	0%	0%	0%	66.5%	7%
δ^*_{ATE}	100%	100%	100%	100%	33.5%	93%
	Subgroup 2 (0 is not contained)					
δ^*	92%	90%	90.5%	94%	1%	17%
δ^*_{QTE}	8%	10%	9.5%	6%	99%	83%
δ^*_{ATE}	8%	10%	9.5%	6%	99%	83%

Numerical Results: Correct Classification Rate ($n = 50$)

	$\hat{\delta}^{SI,stoch}$	$\hat{\delta}^{stoch}$	$\hat{\delta}^{SI,determ}$	$\hat{\delta}^{determ}$	$\hat{\delta}_{QTE}$	$\hat{\delta}_{ATE}$
	Subgroup 3 (0 is contained)					
δ^*	79%	51%	84%	62%	99.5%	100%
δ^*_{QTE}	79%	51%	84%	62%	99.5%	100%
δ^*_{ATE}	79%	51%	84%	62%	99.5%	100%
	Subgroup 4 (0 is contained)					
δ^*	26%	49.5%	14.5%	44%	1.5%	0.5%
δ^*_{QTE}	74%	50.5%	85.5%	56%	98.5%	99.5%
δ^*_{ATE}	74%	50.5%	85.5%	56%	98.5%	99.5%
	Subgroup 5 (0 is contained, SI false)					
δ^*	82%	81%	83%	93.5%	66.5%	4%
δ^*_{QTE}	82%	81%	83%	93.5%	66.5%	4%
δ^*_{ATE}	18%	19%	17%	6.5%	33.5%	96%
	Subgroup 6 (0 is contained, SI false)					
δ^*	43%	61.5%	40%	61%	24%	0%
δ^*_{QTE}	57%	38.5%	60%	39%	76%	100%
δ^*_{ATE}	57%	38.5%	60%	39%	76%	100%

Numerical Results: Correct Classification Rate ($n = 50$)

	$\hat{\delta}^{SI,stoch}$	$\hat{\delta}^{stoch}$	$\hat{\delta}^{SI,determ}$	$\hat{\delta}^{determ}$	$\hat{\delta}_{QTE}$	$\hat{\delta}_{ATE}$
	Subgroup 8 (log normal, SI excludes 0)					
δ^*	95.5%	68%	95.5%	73%	100%	76.5%
δ^*_{QTE}	95.5%	68%	95.5%	73%	100%	76.5%
δ^*_{ATE}	4.5%	32%	4.5%	28%	0%	23.5%

Numerical Results: Correct Classification Rate ($n = 1000$)

	$\hat{\delta}^{SI,stoch}$	$\hat{\delta}^{stoch}$	$\hat{\delta}^{SI,determ}$	$\hat{\delta}^{determ}$	$\hat{\delta}_{QTE}$	$\hat{\delta}_{ATE}$
	Subgroup 1 (0 is not contained)					
δ^*	100%	100%	100%	100%	0%	100%
δ^*_{QTE}	0%	0%	0%	0%	100%	0%
δ^*_{ATE}	100%	100%	100%	100%	0%	100%
	Subgroup 2 (0 is not contained)					
δ^*	100%	99%	100%	100%	0%	0%
δ^*_{QTE}	0%	1%	0%	0%	100%	100%
δ^*_{ATE}	0%	1%	0%	0%	100%	100%

Numerical Results: Correct Classification Rate ($n = 1000$)

	$\hat{\delta}^{SI,stoch}$	$\hat{\delta}^{stoch}$	$\hat{\delta}^{SI,determ}$	$\hat{\delta}^{determ}$	$\hat{\delta}_{QTE}$	$\hat{\delta}_{ATE}$
	Subgroup 3 (0 is contained)					
δ^*	89%	59%	100%	95%	100%	100%
δ^*_{QTE}	89%	59%	100%	95%	100%	100%
δ^*_{ATE}	89%	59%	100%	95%	100%	100%
	Subgroup 4 (0 is contained)					
δ^*	21%	43%	0%	20%	0%	0%
δ^*_{QTE}	79%	57%	100%	80%	100%	100%
δ^*_{ATE}	79%	57%	100%	80%	100%	100%
	Subgroup 5 (0 is contained, SI false)					
δ^*	92%	74%	100%	100%	100%	0%
δ^*_{QTE}	92%	74%	100%	100%	100%	0%
δ^*_{ATE}	8%	26%	0%	0%	0%	100%
	Subgroup 6 (0 is contained, SI false)					
δ^*	34%	48%	6.5%	86%	0%	0%
δ^*_{QTE}	66%	52%	93.5%	14%	100%	100%
δ^*_{ATE}	66%	52%	93.5%	14%	100%	100%

Numerical Results: Correct Classification Rate ($n = 1000$)

	$\hat{\delta}^{SI,stoch}$	$\hat{\delta}^{stoch}$	$\hat{\delta}^{SI,determ}$	$\hat{\delta}^{determ}$	$\hat{\delta}_{QTE}$	$\hat{\delta}_{ATE}$
	Subgroup 8 (log normal, SI excludes 0)					
δ^*	100%	60%	100%	97%	100%	70.5%
δ^*_{QTE}	100%	60%	100%	97%	100%	70.5%
δ^*_{ATE}	0%	40%	0%	3%	0%	29.5%

Application I: Allocations of Right Heart Catheterization

We consider the right heart catheterization (RHC) (e.g., Hirano & Imbens 02)

- D: RHC (1 if received and 0 otherwise), a diagnostic procedure for critically ill patients
- Y: number of days from admission to death within 30 days

Studies like Connors et al. 21 found that patient survival is lower with RHC than without

- therefore, relevant policy question is to find patients for whom allocating (or avoiding!) RHC is life-saving

In the dataset, 5735 patients are divided into a treatment group (2184 patients) and a control group (3551 patients)

- X: age, sex, coma, DNR status, est'ed survival rate, ICU mortality score

Application I: Allocations of Right Heart Catheterization

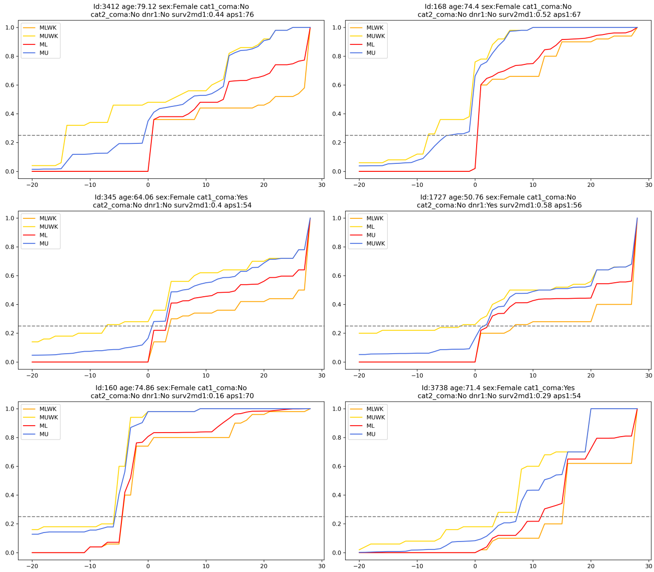


Figure: Bounds on the QoTE of Six Representative Patients ($\tau = 0.25$)

Application II: Allocations of Job Training

We consider the National Job Training Partnership Act (JTPA) Study (Bloom et al. 97, Abadie et al. 02, Kitagawa & Tetenov 18)

We use a subset that includes 9,223 adults; 6,133 received job training, while 3,090 did not

- D: job training
- Y: 30-month earnings after job training
- X: sex, years of education, high school diploma, previous earnings

Q: how do prudent/negligent policies look like, relative to e.g. utilitarian policy?

Application II: Allocations of Job Training

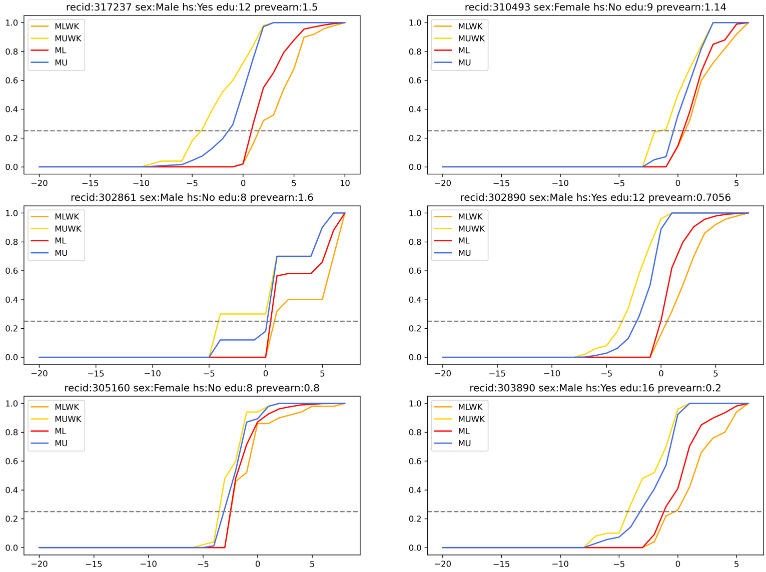
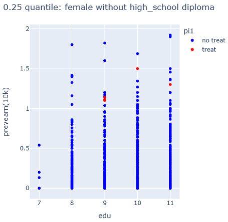
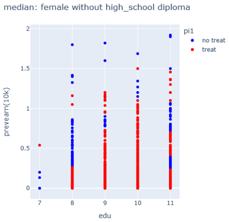


Figure: Bounds on the QoTE of Six Representative Workers ($\tau = 0.25$)

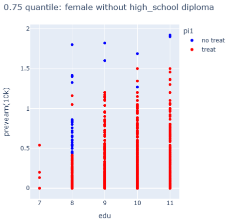
Application II: Allocations of Job Training



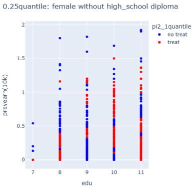
(a) $\tau = 0.25$



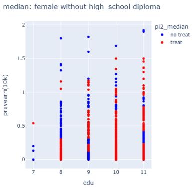
(b) $\tau = 0.5$



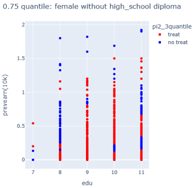
(c) $\tau = 0.75$



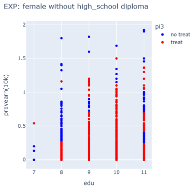
(d) 0.25-QTE



(e) 0.5-QTE



(f) 0.75-QTE



(g) ATE

Figure: Decisions for Female Workers Without High School Diploma

Concluding Remarks

The paper...

- proposes optimal allocation decisions
- under welfare gain defined by quantile of treatment effects
- that can also be motivated by vote shares;
- considers ambiguity-robust decisions;
- provides theoretical guarantee by calculating regret bounds;
- proposes a range of identifying assumptions

Extensions:

- interquartile range as equity target for an egalitarian PM
  - Kitagawa & Tetenov 21
- $E[Y_1 - Y_0 | Y_0 < c]$ as alternative prioritarian objective

Thank You! ☺

Rectangular Identified Set

Consider $X \in \{0, 1\}$

REC imposes that

$$\{(Q_\tau(0), Q_\tau(1)) : Q_\tau(x) \in [Q^L_\tau(x), Q^U_\tau(x)], x \in \{0, 1\}\}$$

is rectangular

Then

$$\begin{aligned}
\min_{F_{Y_1,Y_0|X}} E[\delta(X) Q_\tau(X)] &= \min\,[p_1 \delta(1) Q_\tau(1) + p_0 \delta(0) Q_\tau(0)] \\
&= p_1 \delta(1) \min Q_\tau(1) + p_0 \delta(0) \min Q_\tau(0) \\
&= p_1 \delta(1) Q^L_\tau(1) + p_0 \delta(0) Q^L_\tau(0) \\
&= E[\delta(X) \min Q_\tau(X)]
\end{aligned}$$

Return