

On Quantile Treatment Effects, Rank Similarity, and Multiple IVs

Sukjin Han & Haiqing Xu

U of Bristol & UT Austin

18 October 2022

UCL & CeMMAP Seminar

Observed vs. Counterfactual Distributions

Question:

How relationship between **observed** vs. **counterfactual distributions** plays a role in identification of treatment effects under endogeneity?

Observed vs. Counterfactual Distributions

Question:

How relationship between **observed** vs. **counterfactual distributions** plays a role in identification of treatment effects under endogeneity?

e.g., Condition:

For arbitrary A and \tilde{A} ,

if $F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}}$

then $F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}}$

Observed vs. Counterfactual Distributions

Condition 1.1 (preservation of stochastic dominance):

For arbitrary A and \tilde{A} ,

if $F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}}$

then $F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}}$

This paper...

- ▶ proposes this condition (and related ones) as possible source of identification...
- ▶ for quantile treatment effect (QTE) and average treatment effect (ATE) for treated and untreated populations, and
- ▶ proposes a simple procedure to calculate bounds using linear programming

Example

D : observed indicator of college degree (endogenous)

Y_1 : hypothetical earnings with college education

Y_0 : hypothetical earnings without college education

$Y = DY_1 + (1 - D)Y_0$: observed earnings

Z : instrument(s) for D (e.g., local earnings, distance to college)

Common parameters of interest are

$$QTE_{\tau} = Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$$

and

$$ATE = E[Y_1 - Y_0]$$

Example

D : observed indicator of college degree (endogenous)

Y_1 : hypothetical earnings with college education

Y_0 : hypothetical earnings without college education

$Y = DY_1 + (1 - D)Y_0$: observed earnings

Z : instrument(s) for D (e.g., local earnings, distance to college)

...or more fundamentally

$$QTE_{\tau}(d) = Q_{Y_1|D=d}(\tau) - Q_{Y_0|D=d}(\tau)$$

and

$$ATE(d) = E[Y_1 - Y_0|D = d]$$

Well-Known Approach: Rank Similarity

Assume

$$Y_d = q(d, U_d)$$

where $q(d, \cdot)$ is strictly increasing and $U_d \sim U[0, 1]$

Assume $D = h(Z, \eta)$ and $Z \perp U_d$

Rank similarity (Chernozhukov & Hansen 05):

$$F_{U_1|\eta, Z} = F_{U_0|\eta, Z}$$

- ▶ observationally equivalent to $U_1 = U_0 \equiv U$ (Chernozhukov & Hansen 13)
- ▶ reducing dimension of unobs'd heterogeneity

Under these assumptions, QTE_τ and ATE are point identified

Motivation: Relaxing Rank Similarity

Rank similarity (Chernozhukov & Hansen 05):

$$F_{U_1|\eta,Z} = F_{U_0|\eta,Z}$$

Strong assumption with fragile empirical justification

- ▶ e.g., wages in US CPS for past decades (Massoumi & Wang 19)

Motivation: Relaxing Rank Similarity

Rank similarity (Chernozhukov & Hansen 05):

$$F_{U_1|\eta} = F_{U_0|\eta}$$

Strong assumption with fragile empirical justification

- ▶ e.g., wages in US CPS for past decades (Massoumi & Wang 19)

Recall **Condition 1.1**:

$$\begin{array}{l} \text{For arbitrary } A \text{ and } \tilde{A}, \\ F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} \\ \implies \\ F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} \end{array}$$

Motivation: Relaxing Rank Similarity

Rank similarity (Chernozhukov & Hansen 05):

$$F_{U_1|\eta} = F_{U_0|\eta}$$

Strong assumption with fragile empirical justification

- ▶ e.g., wages in US CPS for past decades (Massoumi & Wang 19)

Stronger than Condition 1.1 is **Condition 2**:

For arbitrary A and \tilde{A} ,

$$\begin{aligned} F_{Y_1|\eta \in A} &\leq F_{Y_1|\eta \in \tilde{A}} \\ &\iff \\ F_{Y_0|\eta \in A} &\leq F_{Y_0|\eta \in \tilde{A}} \end{aligned}$$

Motivation: Relaxing Rank Similarity

Rank similarity (Chernozhukov & Hansen 05):

$$F_{U_1|\eta} = F_{U_0|\eta}$$

Strong assumption with fragile empirical justification

- ▶ e.g., wages in US CPS for past decades (Massoumi & Wang 19)

Stronger than Condition 1.1 is **Condition 2**:

$$\begin{array}{ccc} & F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} & \\ \text{For arbitrary } A \text{ and } \tilde{A}, & \iff & \\ & F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} & \end{array}$$

We show rank similarity implies Condition 2 (within their model)

Motivation: Relaxing Rank Similarity

Rank similarity (Chernozhukov & Hansen 05):

$$F_{U_1|\eta} = F_{U_0|\eta}$$

Strong assumption with fragile empirical justification

- ▶ e.g., wages in US CPS for past decades (Massoumi & Wang 19)

Stronger than Condition 1.1 is **Condition 2**:

$$\begin{array}{ccc} & F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} & \\ \text{For arbitrary } A \text{ and } \tilde{A}, & \iff & \\ & F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} & \end{array}$$

We show rank similarity implies Condition 2 (within their model)

Therefore, we propose Condition 1.1 as substantial relaxation of rank similarity

Making Use of Condition 1

Condition 1.1:

For arbitrary A and \tilde{A} ,

$$\begin{aligned} F_{Y_1|\eta \in A} &\leq F_{Y_1|\eta \in \tilde{A}} \\ \implies \\ F_{Y_0|\eta \in \tilde{A}} &\leq F_{Y_0|\eta \in A} \end{aligned}$$

Key step: Find A and \tilde{A} such that

$$F_{Y_1|\eta \in A, D=1} \leq F_{Y_1|\eta \in \tilde{A}, D=1}$$

- ▶ more likely if Z can produce finer partition of $\text{Supp}(\eta)$
- ▶ therefore, multiple IVs can be helpful (makes sense given how much we give up by dropping rank similarity)

What We Partially Identify

Condition 1.1:

$$\begin{array}{l} \text{For arbitrary } A \text{ and } \tilde{A}, \\ F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} \\ \implies \\ F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} \end{array}$$

Condition 1.1 yields bounds for $QTE_\tau(1)$ and $ATE(1)$

What We Partially Identify

Condition 1.1:

$$\begin{array}{l} \text{For arbitrary } A \text{ and } \tilde{A}, \\ F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} \\ \implies \\ F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} \end{array}$$

Condition 1.1 yields bounds for $QTE_\tau(1)$ and $ATE(1)$

Condition 1.0 (with “ \Leftarrow ”) yields bounds for $QTE_\tau(0)$ and $ATE(0)$

What We Partially Identify

Condition 1.1:

$$\begin{array}{ccc} & F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} & \\ \text{For arbitrary } A \text{ and } \tilde{A}, & \implies & \\ & F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} & \end{array}$$

Condition 1.1 yields bounds for $QTE_\tau(1)$ and $ATE(1)$

Condition 1.0 (with “ \Leftarrow ”) yields bounds for $QTE_\tau(0)$ and $ATE(0)$

We provide models that...

- ▶ rationalize these conditions
- ▶ justify the *assumption-driven* parameters of interest for a policymaker concerning “conservative” populations

What We Partially Identify

Condition 1.1:

$$\begin{array}{l} \text{For arbitrary } A \text{ and } \tilde{A}, \\ F_{Y_1|\eta \in A} \leq F_{Y_1|\eta \in \tilde{A}} \\ \implies \\ F_{Y_0|\eta \in A} \leq F_{Y_0|\eta \in \tilde{A}} \end{array}$$

Condition 1.1 yields bounds for $QTE_\tau(1)$ and $ATE(1)$

Condition 1.0 (with “ \Leftarrow ”) yields bounds for $QTE_\tau(0)$ and $ATE(0)$

We provide models that...

- ▶ rationalize these conditions
- ▶ justify the *assumption-driven* parameters of interest for a policymaker concerning “conservative” populations

Condition 2 (with “ \Longleftrightarrow ”) yields bounds for QTE_τ and ATE

Previous Related Approaches

IV quantile models (with rank similarity):

- ▶ Chernozhukov & Hansen 05, Vuong & Xu 17

Triangular models (in the sense of explicit first stage):

- ▶ Control function approach with continuous D : Chesher 03, Lee 07, Imbens & Newey 09
- ▶ Threshold crossing first stage with binary D :
 - ▶ ATE (rank similarity): Shaikh & Vytlacil 11, Vytlacil & Yildiz 07
 - ▶ MTE: Mogstad et al. 08, Han & Yang 22

Local QTE with binary D :

- ▶ Abadie, Angrist & Imbens 02 (local parameter, strong homogeneity with multiple IVs (Mogstad et al. 21))

Generalized IV:

- ▶ Chesher & Rosen 17 (sharp bounds with IVs)

This Paper

This paper...

- ▶ relaxes rank similarity (without completely abandoning it) and
- ▶ constructs informative bounds on QTE and ATE for the treated or untreated
- ▶ for binary endogenous treatment
- ▶ using discrete IVs,
- ▶ bounds that are simple to calculate

Condition 1.1 differs from other assumptions on the relationship between Y_1 and Y_0

- ▶ e.g., stochastic increasing for distributional treatment effects (with experimental data) (Frandsen & Lefgren 21)

I. Key Conditions and Bounds on Treatment Effects

Maintained Assumptions

D : observed treatment indicator (endogenous)

Y_1 : counterfactual outcome of being treated

Y_0 : counterfactual outcome of not being treated

$$Y = DY_1 + (1 - D)Y_0$$

Z : vector of binary IVs or a multi-valued IV, taking values $\{z_1, \dots, z_L\}$

Suppress X for simplicity; easy to incorporate (in the paper)

Assume $D = h(Z, \eta)$ where $\eta \in \mathcal{T}$ has arbitrary dimensions

Define counterfactual treatment $D_z = h(z, \eta)$

Assumption Z

For $d \in \{0, 1\}$ and $z \in \{z_1, \dots, z_L\}$, (i) $Y_{d,z} = Y_d$;
(ii) $Z \perp (Y_d, D_z)$.

Preservation of Stochastic Dominance

Condition 1.1

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\implies \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

Preservation of Stochastic Dominance

Condition 1.1

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\implies \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

$\int w(t, x) F_{Y_d|\eta, X}(\cdot|t, x) dt$ is mixture of CDFs, and thus itself a CDF

Preservation of Stochastic Dominance

Condition 1.1 implies the following:

Condition 1.1*

For arbitrary nonnegative weight vectors (w_1, \dots, w_L) and $(\tilde{w}_1, \dots, \tilde{w}_L)$ s.t. $\sum_{\ell=1}^L w_\ell = \sum_{\ell=1}^L \tilde{w}_\ell = 1$,

$$\begin{aligned} \sum_{\ell=1}^L w_\ell P[Y_1 \leq \cdot | D = 1, Z = z_\ell] &\leq \sum_{\ell=1}^L \tilde{w}_\ell P[Y_1 \leq \cdot | D = 1, Z = z_\ell] \\ &\implies \\ \sum_{\ell=1}^L w_\ell P[Y_0 \leq \cdot | D = 1, Z = z_\ell] &\leq \sum_{\ell=1}^L \tilde{w}_\ell P[Y_0 \leq \cdot | D = 1, Z = z_\ell]. \end{aligned}$$

Preservation of Stochastic Dominance

Under Assumption Z, Condition 1.1' is equivalent to:

Condition 1.1*

For arbitrary positive weight vectors (w_1, \dots, w_L) and $(\tilde{w}_1, \dots, \tilde{w}_L)$
s.t. $\sum_{\ell=1}^L w_{\ell} = \sum_{\ell=1}^L \tilde{w}_{\ell} = 1$,

$$\sum_{\ell=1}^L w_{\ell} P[Y_1 \leq \cdot | D_{z_{\ell}} = 1] \leq \sum_{\ell=1}^L \tilde{w}_{\ell} P[Y_1 \leq \cdot | D_{z_{\ell}} = 1]$$

\implies

$$\sum_{\ell=1}^L w_{\ell} P[Y_0 \leq \cdot | D_{z_{\ell}} = 1] \leq \sum_{\ell=1}^L \tilde{w}_{\ell} P[Y_0 \leq \cdot | D_{z_{\ell}} = 1].$$

Preservation of Stochastic Dominance

Under Assumption Z, Condition 1.1' is equivalent to:

Condition 1.1*

For arbitrary positive weight vectors (w_1, \dots, w_L) and $(\tilde{w}_1, \dots, \tilde{w}_L)$
s.t. $\sum_{\ell=1}^L w_{\ell} = \sum_{\ell=1}^L \tilde{w}_{\ell} = 1$,

$$\begin{aligned} \sum_{\ell=1}^L w_{\ell} P[Y_1 \leq \cdot | D_{z_{\ell}} = 1] &\leq \sum_{\ell=1}^L \tilde{w}_{\ell} P[Y_1 \leq \cdot | D_{z_{\ell}} = 1] \\ &\implies \\ \sum_{\ell=1}^L w_{\ell} P[Y_0 \leq \cdot | D_{z_{\ell}} = 1] &\leq \sum_{\ell=1}^L \tilde{w}_{\ell} P[Y_0 \leq \cdot | D_{z_{\ell}} = 1]. \end{aligned}$$

$\{D_{z_{\ell}} = 1\}$ (for z_1, \dots, z_L) captures different compliance types

- e.g., when $L = 2$ with LATE monotonicity, then between always-takers and compliers

Preservation of Stochastic Dominance

Lemma 0

Under Assumption Z, Condition 1.1 implies Condition 1.1*.

Let $p(z) \equiv P[D = 1|Z = z]$, then $p(z) = P[\eta \in H(z)]$ with $H(z) \equiv \{\eta : 1 = h(z, \eta)\}$

Then, for example,

$$\begin{aligned} & \sum_{\ell} w_{\ell} P[Y_1 \leq \cdot | D = 1, Z = z_{\ell}] \\ &= \sum_{\ell} w_{\ell} P[Y_1 \leq \cdot | \eta \in H(z_{\ell})] \\ &= \int \frac{\sum_{\ell} w_{\ell} 1[t \in H(z_{\ell})]}{p(z_{\ell})} P[Y_1 \leq \cdot | \eta = t] dt \end{aligned}$$

Take $w(t) = \frac{\sum_{\ell} w_{\ell} 1[t \in H(z_{\ell})]}{p(z_{\ell})}$, which satisfies $\int w(t) dt = 1$

□

Bounds on $F_{Y_0|D=1}(\cdot)$

Recall $p(z_\ell) \equiv P[D = 1|Z = z_\ell]$

Let $\Gamma_p = \{(\gamma_1, \dots, \gamma_L) \in \mathbb{R}^L : \sum_{\ell=1}^L \gamma_\ell = 0 \text{ and } \sum_{\ell=1}^L p(z_\ell)\gamma_\ell = 1\}$

Theorem 1

Suppose Assumption Z and Condition 1.1* hold. For

$\gamma = (\gamma_1, \dots, \gamma_L) \in \Gamma_p$, suppose

$$P[Y \leq \cdot | D = 1] \leq \sum_{\ell=1}^L \gamma_\ell P[Y \leq \cdot, D = 1 | Z = z_\ell] \quad (1)$$

Then $F_{Y_0|D=1}(\cdot)$ is upper bounded by

$$P[Y_0 \leq \cdot | D = 1] \leq - \sum_{\ell=1}^L \gamma_\ell P[Y \leq \cdot, D = 0 | Z = z_\ell]$$

Analogous theorem for lower bound

Bounds on $F_{Y_0|D=1}(\cdot)$: Proof

WLOG, let $\gamma_\ell \leq 0$ for $\ell \leq \ell^*$ and $\gamma_\ell > 0$ for $\ell > \ell^*$

Let $q(z_\ell) \equiv P[Z = z_\ell | D = 1]$

Then (1) can be rewritten as

$$\begin{aligned} \sum_{\ell=1}^{\ell^*} \frac{q(z_\ell)}{a} P[Y_1 \leq y | D = 1, Z = z_\ell] \\ + \sum_{\ell=\ell^*+1}^L \frac{q(z_\ell)}{a} P[Y_1 \leq y | D = 1, Z = z_\ell] \\ \leq \sum_{\ell=1}^L \frac{\gamma_\ell p(z_\ell)}{a} P[Y_1 \leq y | D = 1, Z = z_\ell] \end{aligned}$$

where $a \equiv 1 - \sum_{\ell=1}^{\ell^*} p(z_\ell) \gamma_\ell$

Bounds on $F_{Y_0|D=1}(\cdot)$: Proof

WLOG, let $\gamma_\ell \leq 0$ for $\ell \leq \ell^*$ and $\gamma_\ell > 0$ for $\ell > \ell^*$

Let $q(z_\ell) \equiv P[Z = z_\ell | D = 1]$

Then (1) can be rewritten as

$$\begin{aligned} \sum_{\ell=1}^{\ell^*} \frac{q(z_\ell) - \gamma_\ell p(z_\ell)}{a} P[Y_1 \leq y | D = 1, Z = z_\ell] \\ + \sum_{\ell=\ell^*+1}^L \frac{q(z_\ell)}{a} P[Y_1 \leq y | D = 1, Z = z_\ell] \\ \leq \sum_{\ell=\ell^*+1}^L \frac{\gamma_\ell p(z_\ell)}{a} P[Y_1 \leq y | D = 1, Z = z_\ell] \end{aligned}$$

where $a \equiv 1 - \sum_{\ell=1}^{\ell^*} \gamma_\ell p(z_\ell)$ and the positive weights sum to 1

Bounds on $F_{Y_0|D=1}(\cdot)$: Proof

Therefore, by Condition 1.1*,

$$\begin{aligned} & \sum_{\ell=1}^{\ell^*} \frac{q(z_\ell) - \gamma_\ell p(z_\ell)}{a} P[Y_0 \leq y | D = 1, Z = z_\ell] \\ & \quad + \sum_{\ell=\ell^*+1}^L \frac{q(z_\ell)}{a} P[Y_0 \leq y | D = 1, Z = z_\ell] \\ & \leq \sum_{\ell=\ell^*+1}^L \frac{\gamma_\ell p(z_\ell)}{a} P[Y_0 \leq y | D = 1, Z = z_\ell] \end{aligned}$$

Bounds on $F_{Y_0|D=1}(\cdot)$: Proof

Equivalently, we have

$$\begin{aligned} & P[Y_0 \leq y | D = 1] \\ \leq & \sum_{\ell=1}^L \gamma_{\ell} P[Y_0 \leq y, D = 1 | Z = z_{\ell}] \\ = & \sum_{\ell=1}^L \gamma_{\ell} [P[Y_0 \leq y | Z = z_{\ell}] - P[Y_0 \leq y, D = 0 | Z = z_{\ell}]] \\ = & P[Y_0 \leq y] \sum_{\ell=1}^L \gamma_{\ell} - \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq y, D = 0 | Z = z_{\ell}] \\ = & - \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq y, D = 0 | Z = z_{\ell}] \end{aligned}$$

□

Bounds on $F_{Y_0|D=1}(\cdot)$

Finally, want to collect all γ that satisfy (1):

Corollary 1

Suppose Assumption Z and Condition 1.1* hold. Then,

$$F_{Y_0|D=1}^{UB}(y) = \min_{\gamma \in \Gamma_p: (1) \text{ holds}} - \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq y, D = 0 | Z = z_{\ell}]$$

Symmetric condition and bound can be derived for $F_{Y_0|D=1}^{LB}(\cdot)$

Then,

$$F_{Y_0|D=1}^{LB}(\cdot) \leq F_{Y_0|D=1}(\cdot) \leq F_{Y_0|D=1}^{UB}(\cdot)$$

Bounds on $QTE_\tau(1)$

Note that

$$QTE_\tau(1) = Q_{Y_1|D=1}(\tau) - Q_{Y_0|D=1}(\tau) = Q_{Y|D=1}(\tau) - Q_{Y_0|D=1}(\tau)$$

Worst case bounds for quantile (Manski 94, Blundell et al. 07):

$$Q_{Y_0|D=1}^{LB}(\tau) \leq Q_{Y_0|D=1}(\tau) \leq Q_{Y_0|D=1}^{UB}(\tau)$$

with $Q_{Y_0|D=1}^{LB}(\tau) = F_{Y_0|D=1}^{UB}(\tau)^{-1}$ and $Q_{Y_0|D=1}^{UB}(\tau) = F_{Y_0|D=1}^{LB}(\tau)^{-1}$

More on Theorem 1

Need to find $\gamma \in \Gamma_p$ that satisfies

$$P[Y \leq \cdot | D = 1] \leq \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq \cdot, D = 1 | Z = z_{\ell}] \quad (1)$$

► equiv. to finding (w, \tilde{w}) that satisfy “if” part of Condition 1.1*:

$$\sum_{\ell=1}^L w_{\ell} P[Y_1 \leq \cdot | D = 1, Z = z_{\ell}] \leq \sum_{\ell=1}^L \tilde{w}_{\ell} P[Y_1 \leq \cdot | D = 1, Z = z_{\ell}]$$

More on Theorem 1

Need to find $\gamma \in \Gamma_p$ that satisfies

$$P[Y \leq \cdot | D = 1] \leq \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq \cdot, D = 1 | Z = z_{\ell}] \quad (1)$$

► equiv. to finding (w, \tilde{w}) that satisfy “if” part of Condition 1.1*:

$$\sum_{\ell=1}^L w_{\ell} P[Y_1 \leq \cdot | D = 1, Z = z_{\ell}] \leq \sum_{\ell=1}^L \tilde{w}_{\ell} P[Y_1 \leq \cdot | D = 1, Z = z_{\ell}]$$

When $L = 2$ then $\gamma = \left(\frac{1}{p(z_1) - p(z_2)}, -\frac{1}{p(z_1) - p(z_2)} \right)$, and under LATE monotonicity, (1) is equiv. to

$$Y_1 | \{\text{always-takers}\} \prec_{FOSD} Y_1 | \{\text{compliers}\}$$

Why Multi-Valued IVs

$$\{\gamma \in \Gamma_p : \gamma \text{ satisfies (1)}\}$$

The size of this set determines the width of our bound

- ▶ a larger set \Rightarrow narrower bounds
- ▶ more values Z takes \Rightarrow ...
 - ▶ greater degree of freedom in Γ_p and
 - ▶ (1) is more likely to hold

Why Multi-Valued IVs

$$\{\gamma \in \Gamma_p : \gamma \text{ satisfies (1)}\}$$

The size of this set determines the width of our bound

- ▶ a larger set \Rightarrow narrower bounds
- ▶ more values Z takes \Rightarrow ...
 - ▶ greater degree of freedom in Γ_p and
 - ▶ (1) is more likely to hold

Multi-valued IV or multiple IVs are common in practice

- ▶ e.g., “continuous” IV, double eligibility in RCT

Why Multi-Valued IVs

$$\{\gamma \in \Gamma_p : \gamma \text{ satisfies (1)}\}$$

The size of this set determines the width of our bound

- ▶ a larger set \Rightarrow narrower bounds
- ▶ more values Z takes \Rightarrow ...
 - ▶ greater degree of freedom in Γ_p and
 - ▶ (1) is more likely to hold

Multi-valued IV or multiple IVs are common in practice

- ▶ e.g., “continuous” IV, double eligibility in RCT

Then (1) establishes FOSD btw the mixtures of F_{Y_1} conditional on various always-takers and compliers groups

Bounds on $QTE_\tau(0)$ and QTE_τ

Condition 1.0

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\iff \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

With Condition 1.0 (“ \iff ”), we can derive bounds on $QTE_\tau(0)$

Bounds on $QTE_\tau(0)$ and QTE_τ

Condition 1.0

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\Longleftarrow \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

With Condition 1.0 (“ \Longleftarrow ”), we can derive bounds on $QTE_\tau(0)$

With Condition 2 (“ \Longleftrightarrow ”), we can derive bounds on QTE_τ

Bounds on $QTE_\tau(0)$ and QTE_τ

Condition 1.0

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\iff \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

With Condition 1.0 (“ \Leftarrow ”), we can derive bounds on $QTE_\tau(0)$

With Condition 2 (“ \iff ”), we can derive bounds on QTE_τ

We also provide a condition for bounds on $ATE(d)$ and ATE (in the paper)

II. Structural Models and Policymaker's Problems

Sufficient Conditions: A Structural Model

To further interpret Conditions 1.1 (and 1.0), we propose a model that rationalizes it

Suppress “conditional on Z ” throughout

Model 1

$$Y_d = q(d, U_d) \quad \text{for } d \in \{0, 1\}$$

- (i) $q(d, \cdot)$ is continuous and monotone increasing
 - (ii) conditional on η , $U_d \stackrel{d}{=} U + \xi_d$ where $\xi_d \perp (\eta, U)$
 - (iii) ξ_0 is (weakly) more noisy than ξ_1 , i.e., $\xi_0 \stackrel{d}{=} \xi_1 + V$ for some V independent of ξ_1
- i.e., $U_0 \stackrel{d}{=} U_1 + V$

Sufficient Conditions: A Structural Model

To further interpret Conditions 1.1 (and 1.0), we propose a model that rationalizes it

Suppress “conditional on Z ” throughout

Model 1

$$Y_d = q(d, U_d) \quad \text{for } d \in \{0, 1\}$$

- (i) $q(d, \cdot)$ is continuous and monotone increasing
 - (ii) conditional on η , $U_d \stackrel{d}{=} U + \xi_d$ where $\xi_d \perp (\eta, U)$
 - (iii) ξ_0 is (weakly) more noisy than ξ_1 , i.e., $\xi_0 \stackrel{d}{=} \xi_1 + V$ for some V independent of ξ_1
- i.e., $U_0 \stackrel{d}{=} U_1 + V$

This model nests that in Chernozhukov and Hansen (2005)

- by taking $\xi_d = 0$ for all d , $U_0 \stackrel{d}{=} U_1 \stackrel{d}{=} U$ conditional on η

Sufficient Conditions: A Structural Model

Condition 1.1

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\implies \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

Lemma 1

Model 1 implies Condition 1.1

Model 1 with ξ_0 being *less* noisy than ξ_1 implies Condition 1.0

Example 1: Auction

Y : bid (which subsequently forms revenue)

D : participating in auction with different format

- ▶ $D = 1$ if online vs. $= 0$ if offline

$U_d \stackrel{d}{=} U + \xi_d$: valuation of the item

- ▶ U : common valuation (correlated with D)
- ▶ ξ_d : format specific random shocks, $\xi_d \perp (\eta, U)$
 - ▶ bidders may have limited info on certain features of auction that affect valuation
 - ▶ e.g., they know the distribution of ξ_d but not its realization

Example 1: Auction

Y : bid (which subsequently forms revenue)

D : participating in auction with different format

- ▶ $D = 1$ if online vs. $= 0$ if offline

$U_d \stackrel{d}{=} U + \xi_d$: valuation of the item

- ▶ U : common valuation (correlated with D)
- ▶ ξ_d : format specific random shocks, $\xi_d \perp (\eta, U)$
 - ▶ bidders may have limited info on certain features of auction that affect valuation
 - ▶ e.g., they know the distribution of ξ_d but not its realization

What justifies $\text{var}(\xi_0) > \text{var}(\xi_1)$?

- ▶ in offline auction, bidders may be more emotionally affected by others, which makes their bids more variable

Example 2: Insurance

Y : health outcome

D : getting insurance

▶ $D = 1$ if insured, $= 0$ if not

$U_d \stackrel{d}{=} U + \xi_d$: health conditions

▶ U : health conditions known to participant (and thus correlated with D)

▶ ξ_d : health conditions not fully known a priori

$\text{var}(\xi_0) > \text{var}(\xi_1)$: insurance by definition may ensure a certain level of health conditions

Example 3.1: Vaccination

Y : health outcome

D : getting vaccination (of established vaccine)

▶ $D = 1$ if vaccinated, $= 0$ if not

$U_d \stackrel{d}{=} U + \xi_d$: underlying health conditions

- ▶ U : health conditions known to participant (and thus correlated with D)
- ▶ ξ_d : vaccination status specific health conditions, which is not fully known a priori

$\text{var}(\xi_0) > \text{var}(\xi_1)$: when not vaccinated, there is risk of dying, while vaccination ensures certain level of immunity

Example 3.1: Vaccination

Y : health outcome

D : getting vaccination (of established vaccine)

▶ $D = 1$ if vaccinated, $= 0$ if not

$U_d \stackrel{d}{=} U + \xi_d$: underlying health conditions

- ▶ U : health conditions known to participant (and thus correlated with D)
- ▶ ξ_d : vaccination status specific health conditions, which is not fully known a priori

$\text{var}(\xi_0) > \text{var}(\xi_1)$: when not vaccinated, there is risk of dying, while vaccination ensures certain level of immunity

This scenario justifies Condition 1.1 \Rightarrow bounds on $QTE_\tau(1)$

Example 3.0: Frontier Medical Trial

A contrasting example would be risky medical trial

Y : health outcome

D : participating in medical trial

▶ $D = 1$ if participate, $= 0$ if not

$U_d \stackrel{d}{=} U + \xi_d$: underlying health conditions

▶ U : health conditions known to participant (and thus correlated with D)

▶ ξ_d : health conditions not fully known a priori

$\text{var}(\xi_0) < \text{var}(\xi_1)$: with newly developed medicine, there is high risk of unknown side effects

Example 3.0: Frontier Medical Trial

A contrasting example would be risky medical trial

Y : health outcome

D : participating in medical trial

▶ $D = 1$ if participate, $= 0$ if not

$U_d \stackrel{d}{=} U + \xi_d$: underlying health conditions

▶ U : health conditions known to participant (and thus correlated with D)

▶ ξ_d : health conditions not fully known a priori

$\text{var}(\xi_0) < \text{var}(\xi_1)$: with newly developed medicine, there is high risk of unknown side effects

This scenario justifies Condition 1.0 \Rightarrow bounds on $QTE_\tau(0)$

Policymaker's Problem

Assumption: The policymaker concerns risk averse individuals, which are the majority

Under this assumption, the policymaker wants to understand treatment effects for the target individuals in order to provide “insurance”

- ▶ literally insurance or policy that serves as insurance (e.g., vaccination, subsidies)

Our procedure provides a statistical tool for such a policymaker

Policymaker's Problem

Under Model 1, our procedure has the ability to bound treatment effect for individuals with $D = d$ such that $\text{var}(\xi_d) < \text{var}(\xi_{1-d})$

This is a unique feature of our setting:

- ▶ the plausibility of assumptions dictates the parameter of interest
- ▶ i.e., “assumption-driven” treatment parameters

Sufficient Condition: Model 1 with Rank Similarity

Condition 2

For arbitrary weight functions $w : \mathcal{T} \rightarrow \mathbb{R}_+$ and $\tilde{w} : \mathcal{T} \rightarrow \mathbb{R}_+$ s.t.
 $\int w(t)dt = \int \tilde{w}(t)dt = 1$,

$$\begin{aligned} \int w(x) F_{Y_1|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_1|\eta}(\cdot|t) dt \\ &\iff \\ \int w(t) F_{Y_0|\eta}(\cdot|t) dt &\leq \int \tilde{w}(t) F_{Y_0|\eta}(\cdot|t) dt. \end{aligned}$$

Lemma 2

Model 1 with $F_{U_0|\eta} = F_{U_1|\eta}$ (rank similarity) implies Condition 2.

Necessary and Sufficient Condition: Rank Linearity

Lemma 2

Model 1 with $F_{U_0|\eta} = F_{U_1|\eta}$ (rank similarity) implies Condition 2.

Converse is *not* true!

Counter-example:

Rank Linearity

$Y_d = q(d, U_d)$ (monotonic $q(d, \cdot)$) with

$$F_{Y_0|\eta}(\cdot|t) = \lambda(\cdot)F_{Y_1|\eta}(\psi(\cdot)|t)$$

where $\psi(\cdot)$ is one-to-one and onto mapping and $\lambda(\cdot)$ is consistent with $F_{Y_d|\eta}$ being a proper CDF

Rank linearity implies Condition 2 but is *weaker* than rank similarity

Necessary and Sufficient Condition: Rank Linearity

Rank Linearity

$Y_d = q(d, U_d)$ (monotonic $q(d, \cdot)$) with

$$F_{Y_0|\eta}(\cdot|t) = \lambda(\cdot)F_{Y_1|\eta}(\psi(\cdot)|t)$$

where $\psi(\cdot)$ is one-to-one and onto mapping and $\lambda(\cdot)$ is consistent with $F_{Y_d|\eta}$ being a proper CDF

Rank linearity implies Condition 2 but is *weaker* than rank similarity:

$$F_{U_0|\eta}(q^{-1}(0, y)|t) = \lambda(y)F_{U_1|\eta}(q^{-1}(1, \psi(y))|t)$$

and by choosing $\lambda(y) = 1$ and $\psi(y) = q(1, q^{-1}(0, y))$, we have

$$F_{U_0|\eta}(\cdot|t) = F_{U_1|\eta}(\cdot|t)$$

Necessary and Sufficient Condition: Rank Linearity

Rank Linearity

$Y_d = q(d, U_d)$ (monotonic $q(d, \cdot)$) with

$$F_{Y_0|\eta}(\cdot|t) = \lambda(\cdot)F_{Y_1|\eta}(\psi(\cdot)|t)$$

where $\psi(\cdot)$ is one-to-one and onto mapping and $\lambda(\cdot)$ is consistent with $F_{Y_d|\eta}$ being a proper CDF

Rank linearity implies Condition 2 but is *weaker* than rank similarity:

$$F_{U_0|\eta}(q^{-1}(0, y)|t) = \lambda(y)F_{U_1|\eta}(q^{-1}(1, \psi(y))|t)$$

and by choosing $\lambda(y) = 1$ and $\psi(y) = q(1, q^{-1}(0, y))$, we have

$$F_{U_0|\eta}(\cdot|t) = F_{U_1|\eta}(\cdot|t)$$

In fact, rank linearity may be equivalent to Condition 2

Necessary and Sufficient Condition: Rank Linearity

Suppose $Y_d \in \{y_{d,1}, \dots, y_{d,k_d}\}$ and $\eta \in \{t_1, \dots, t_{k_\eta}\}$

Lemma 3

For any \tilde{F}_d on $\{y_{d,1}, \dots, y_{d,k_d}\}$, suppose there always exists a nonnegative sequence $\{c_{d,1}, \dots, c_{d,k_\eta}\}$ s.t.

$$\tilde{F}_d(\cdot) = \sum_{j=1}^{k_\eta} c_{d,j} F_{Y_d|\eta}(\cdot|t_j). \quad (2)$$

Then, Condition 2 holds if and only if (i) $k_0 = k_1$ and (ii) for some one-to-one and onto mapping $\psi(\cdot)$ and $\lambda(\cdot) > 0$,

$$F_{Y_0|\eta}(\cdot|t_j) = \lambda(\cdot) F_{Y_1|\eta}(\psi(\cdot)|t_j), \quad \text{for } j = 1, \dots, k_\eta.$$

Necessary and Sufficient Condition: Rank Linearity

Suppose $Y_d \in \{y_{d,1}, \dots, y_{d,k_d}\}$ and $\eta \in \{t_1, \dots, t_{k_\eta}\}$

Lemma 3

For any \tilde{F}_d on $\{y_{d,1}, \dots, y_{d,k_d}\}$, suppose there always exists a nonnegative sequence $\{c_{d,1}, \dots, c_{d,k_\eta}\}$ s.t.

$$\tilde{F}_d(\cdot) = \sum_{j=1}^{k_\eta} c_{d,j} F_{Y_d|\eta}(\cdot|t_j). \quad (2)$$

Then, Condition 2 holds if and only if (i) $k_0 = k_1$ and (ii) for some one-to-one and onto mapping $\psi(\cdot)$ and $\lambda(\cdot) > 0$,

$$F_{Y_0|\eta}(\cdot|t_j) = \lambda(\cdot) F_{Y_1|\eta}(\psi(\cdot)|t_j), \quad \text{for } j = 1, \dots, k_\eta.$$

We conjecture an analogous result with continuous Y_d and η (in progress)

Necessary and Sufficient Condition: Proof

We only prove necessity

Recall Condition 2:

$$\begin{aligned}\sum_{j=1}^{k_\eta} w_j F_{Y_1|\eta}(\cdot|t_j) &\leq \sum_{j=1}^{k_\eta} \tilde{w}_j F_{Y_1|\eta}(\cdot|t_j) \\ &\iff \\ \sum_{j=1}^{k_\eta} w_j F_{Y_0|\eta}(\cdot|t_j) &\leq \sum_{j=1}^{k_\eta} \tilde{w}_j F_{Y_0|\eta}(\cdot|t_j)\end{aligned}$$

Necessary and Sufficient Condition: Proof

We only prove necessity

Recall Condition 2:

$$\sum_{j=1}^{k_\eta} \delta_j F_{\mathbf{Y}_1|\eta}(\cdot|t_j) \leq 0$$

$$\Longleftrightarrow$$

$$\sum_{j=1}^{k_\eta} \delta_j F_{\mathbf{Y}_0|\eta}(\cdot|t_j) \leq 0$$

Necessary and Sufficient Condition: Proof

We only prove necessity

Recall Condition 2:

$$\sum_{j=1}^{k_\eta} \delta_j F_{Y_1|\eta}(\cdot|t_j) \leq 0$$

$$\Longleftrightarrow$$

$$\sum_{j=1}^{k_\eta} \delta_j F_{Y_0|\eta}(\cdot|t_j) \leq 0$$

Define cone:

$$\Delta_d \equiv \left\{ \delta : \sum_{j=1}^{k_\eta} \delta_j F_{Y_d|\eta}(\cdot|t_j) \leq 0, \sum_{j=1}^{k_\eta} \delta_j = 0 \right\}$$

Then, by Condition 2, $\Delta_1 = \Delta_0$

Necessary and Sufficient Condition: Proof

Define polar cone:

$$\Delta_d^* \equiv \{F_d \in \mathbb{R}^{k_\eta} | F_d' \delta \leq 0, \forall \delta \in \Delta_d\}$$

Note Δ_d^* is a convex cone whose extreme ray is generated by

$$\left\{ \left(F_{Y_d|\eta}(y|t_1), \dots, F_{Y_d|\eta}(y|t_{k_\eta}) \right)' : y = y_{d,1}, \dots, y_{d,k_d} \right\}$$

which are linearly indep. by (2)

Necessary and Sufficient Condition: Proof

Define polar cone:

$$\Delta_d^* \equiv \{F_d \in \mathbb{R}^{k_\eta} \mid F_d' \delta \leq 0, \forall \delta \in \Delta_d\}$$

Note Δ_d^* is a convex cone whose extreme ray is generated by

$$\left\{ \left(F_{Y_d|\eta}(y|t_1), \dots, F_{Y_d|\eta}(y|t_{k_\eta}) \right)' : y = y_{d,1}, \dots, y_{d,k_d} \right\}$$

which are linearly indep. by (2)

Also, $\Delta_1 = \Delta_0$ implies $\Delta_1^* = \Delta_0^*$, and therefore $k_1 = k_0$

Necessary and Sufficient Condition: Proof

Define polar cone:

$$\Delta_d^* \equiv \{F_d \in \mathbb{R}^{k_\eta} | F_d' \delta \leq 0, \forall \delta \in \Delta_d\}$$

Note Δ_d^* is a convex cone whose extreme ray is generated by

$$\left\{ \left(F_{Y_d|\eta}(y|t_1), \dots, F_{Y_d|\eta}(y|t_{k_\eta}) \right)' : y = y_{d,1}, \dots, y_{d,k_d} \right\}$$

which are linearly indep. by (2)

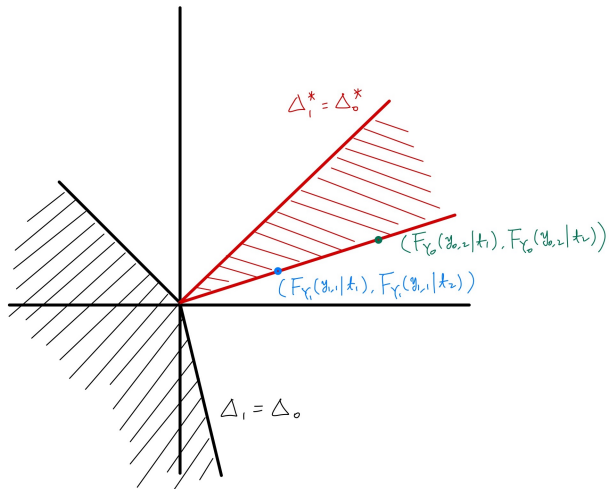
Also, $\Delta_1 = \Delta_0$ implies $\Delta_1^* = \Delta_0^*$, and therefore $k_1 = k_0$

Moreover, any extreme ray in Δ_0^* must be identical to some extreme ray in Δ_1^* : For any y

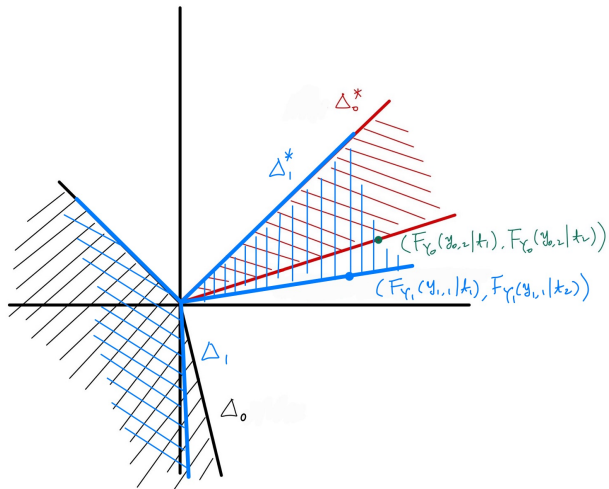
$$\left(F_{Y_0|\eta}(y|t_1), \dots, F_{Y_0|\eta}(y|t_{k_\eta}) \right) = \lambda \times \left(F_{Y_1|\eta}(y'|t_1), \dots, F_{Y_1|\eta}(y'|t_{k_\eta}) \right)$$

for some $\lambda > 0$ and y'

Necessary and Sufficient Condition: Proof



From Model 1 to Condition 1.1



Summary

Model 1 with rank similarity $\rightarrow QTE_\tau$



Model 1 with rank linearity



Condition 2 \rightarrow Bounds on QTE_τ



Condition 1.1



Condition 1.0



Condition 1.1*



Condition 1.0*



Bounds on $QTE_\tau(1)$



Bounds on $QTE_\tau(0)$

Summary

Model 1 with rank similarity $\rightarrow QTE_{\tau}$



Model 1



Condition 1.1



Condition 1.0



Condition 1.1*



Condition 1.0*



Bounds on $QTE_{\tau}(1)$



Bounds on $QTE_{\tau}(0)$

III. Systematic Calculation of Bounds

Computation of Bounds

Recall, the upper bound on $F_{Y_0|D=1}(\cdot)$ is

$$F_{Y_0|D=1}^{UB}(\cdot) = \min_{\gamma \in \Gamma_p: (1) \text{ holds}} - \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq \cdot, D = 0 | Z = z_{\ell}]$$

where

$$P[Y \leq \cdot | D = 1] \leq \sum_{\ell=1}^L \gamma_{\ell} P[Y \leq \cdot, D = 1 | Z = z_{\ell}] \quad (1)$$

and $\Gamma_p = \{(\gamma_1, \dots, \gamma_L) \in \mathbb{R}^L : \sum_{\ell=1}^L \gamma_{\ell} = 0 \text{ and } \sum_{\ell=1}^L p(z_{\ell}) \gamma_{\ell} = 1\}$

Computation of Bounds: Semi-Infinite Program

Simplifying notation:

$$p_{y,d|z} \equiv \{P[Y \leq y, D = d | Z = z_\ell]\}_{\ell=1}^L$$

$$p_{y|1} \equiv P[Y \leq y | D = 1]$$

Consider the following semi-infinite program:

$$\begin{aligned} F_{Y_0|D=1}^{UB}(\bar{y}) &= \min_{\gamma \in \Gamma_p} -p'_{\bar{y},0|z} \gamma \\ \text{s.t. } & p_{y|1} - p'_{y,1|z} \gamma \leq 0, \quad \forall y \in \mathcal{Y} \end{aligned} \quad (1)$$

- ▶ feasible as long as \exists such γ
- ▶ i.e., (1) is testable from data
- ▶ if Y is discrete, then we already have linear program (but not in general)

Computation of Bounds: Linear Program I

The semi-infinite program:

$$\begin{aligned} F_{Y_0|D=1}^{UB}(\bar{y}) &= \min_{\gamma \in \Gamma_p} -p'_{\bar{y},0|z} \gamma \\ \text{s.t. } & p_{y|1} - p'_{y,1|z} \gamma \leq 0, \quad \forall y \in \mathcal{Y} \end{aligned}$$

In practice, with i.i.d. $\{Y_i, D_i, Z_i\}_{i=1}^n$, we solve linear program:

$$\begin{aligned} \hat{F}_{Y_0|D=1}^{UB}(\bar{y}) &= \min_{\gamma \in \Gamma_{\hat{p}}} -\hat{p}'_{\bar{y},0|z} \gamma \\ \text{s.t. } & \hat{p}_{Y_i|1} - \hat{p}'_{Y_i,1|z} \gamma \leq 0, \quad \forall i = 1, \dots, n \end{aligned}$$

The optimization can sometimes be unstable

Computation of Bounds: Linear Program II

The semi-infinite program:

$$\begin{aligned} F_{Y_0|D=1}^{UB}(\bar{y}) &= \min_{\gamma \in \Gamma_p} -p'_{\bar{y},0|z} \gamma \\ \text{s.t. } & p_{y|1} - p'_{y,1|z} \gamma \leq 0, \quad \forall y \in \mathcal{Y} \end{aligned}$$

Dual program:

$$\begin{aligned} F_{Y_0|D=1}^{UB,\dagger}(\bar{y}) &= \sup_{\Lambda \succeq 0, \lambda \in \mathbb{R}^2} \int_{\mathcal{Y}} p_{y|1} d\Lambda(y) - [0 \quad 1] \lambda \\ \text{s.t. } & [1 \quad p] \lambda - \int_{\mathcal{Y}} p_{y,1|z} d\Lambda(y) - p_{\bar{y},0|z} = 0 \end{aligned}$$

► Λ is a nonnegative measure

Strong duality may hold (in progress)

Computation of Bounds: Linear Program II

Approximate $\lambda(y) \equiv d\Lambda(y)/dy$ using Bernstein polynomial:

$$\lambda(y) \approx \sum_{j=1}^J \theta_j b_j(y)$$

Then, results in linear program:

$$\begin{aligned} F_{Y_0|D=1}^{UB,\dagger\dagger}(\bar{y}) = & \max_{\theta \in \mathbb{R}_+^J, \lambda \in \mathbb{R}^2} \theta' b^1 - [0 \quad 1] \lambda \\ \text{s.t.} \quad & [1 \quad p] \lambda - B_1' \theta - \hat{p}_{\bar{y},0|z} = 0 \end{aligned}$$

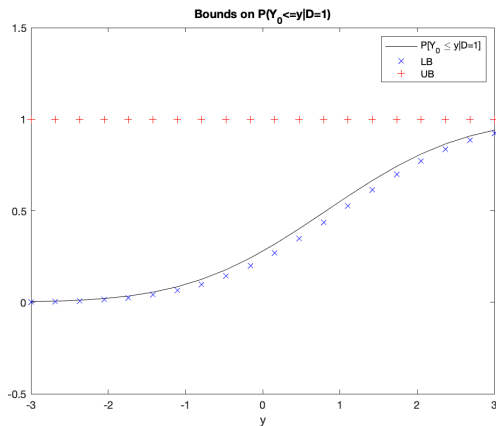
- ▶ $\theta \equiv (\theta_1, \dots, \theta_J)'$
- ▶ $b^d \equiv (b_1^d, \dots, b_J^d)'$ with $b_j^d \equiv \int_{\mathcal{Y}} b_j(y) \hat{p}_{y|d} dy$
- ▶ $\mathbf{b}_{d,j} \equiv (b_{d,j,1}, \dots, b_{d,j,L})'$ with $b_{d,j,\ell} \equiv \int_{\mathcal{Y}} b_j(y) \hat{p}_{y,d|z_\ell} dy$
- ▶ $B_d \equiv [\mathbf{b}_{d,1} \quad \dots \quad \mathbf{b}_{d,J}]$

IV. Numerical Studies

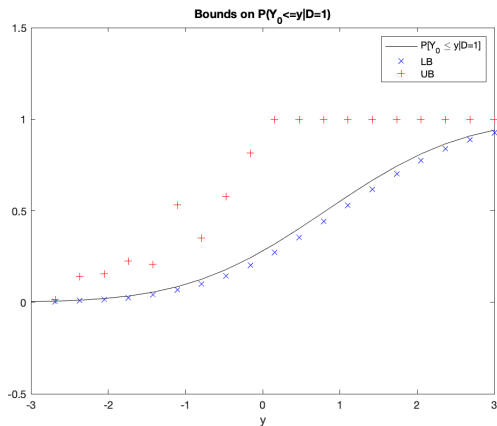
Numerical Illustration: Design

- ▶ $Y_d = q(d, U_d) = 1 - d + (d + 1)U_d$
 - ▶ $Y_1 = 2U_1$ and $Y_0 = 1 + U_0$
- ▶ $(U, \eta) \sim BVN((0, 0)', \Sigma)$
- ▶ $V \sim N(0, \sigma_V^2)$ and $\xi_1 \sim N(0, \sigma_V^2)$
- ▶ $\xi_0 = \xi_1 + V$
- ▶ $U_d = U + \xi_d$
- ▶ $Z \sim \text{Bin}(L - 1, p)/(L - 1) \in [0, 1]$
 - ▶ L is the number of values Z takes
 - ▶ Z is normalized
- ▶ $D = 1\{\pi_0 + \pi_1 Z \geq \eta\}$
- ▶ $Y = DY_1 + (1 - D)Y_0$

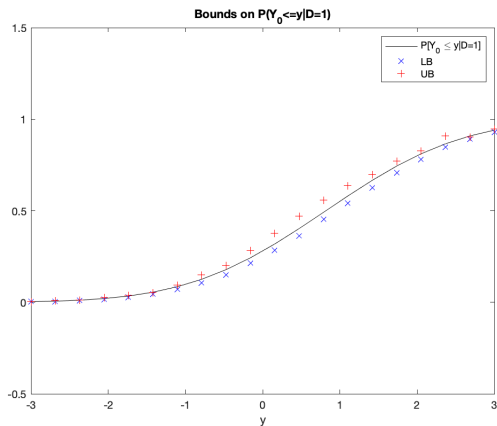
Bounds on $P[Y_0 \leq y | D = 1]$ when $L = 2$



Bounds on $P[Y_0 \leq y | D = 1]$ when $L = 5$



Bounds on $P[Y_0 \leq y | D = 1]$ when $L = 6$



V. Conclusions

Conclusions

The paper...

- ▶ proposes a way to weaken rank similarity and
- ▶ shows how to construct informative bounds on QTE and ATE
- ▶ for treated or untreated (e.g., risk-averse) populations
- ▶ using multi-valued IVs

Calculation of the bounds are simple

Inference is an open question

Thank You! 😊