Fuzzy Regression Discontinuity Design without Monotonicity

Yi Cui

UNC Chapel Hill
Department of Economics
November 8, 2024

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Table of Contents

Introduction

AER publication (2005-2024)

Study	Mono	Indep	IV	FRD
Allcott et al. (2020)		✓	✓	
Andrews (2016)	✓	✓		✓
Autor et al. (2019)	✓	✓	✓	
Dinkelman (2011)		✓	✓	
...	...	...	...	...
Simcoe (2012)			✓	

Table: Summary statistics of AER using IV/FRD (2005-2024)[1]

1. Note: Only 19 of these 39 articles (50%) that explicitly identify a local average treatment effect using either an IV or FRD approach include the word "monotonicity"!

Definition

Framework (FRD and LATE)

Let $(\Omega, \mathcal{F}, P)$ be the probability space
$D(\cdot,\cdot) : \mathcal{R} \times \Omega \to \{0,1\}$ is the observed binary treatment assignment,
$Y(\cdot,\cdot) : \mathcal{R} \times \Omega \to \mathcal{Y}$ is the observed outcome of interest,
$R(\cdot) : \Omega \to \mathcal{R}$ is a continuous running variable with a known cut-off r_0.
For the outcome Y, we have $Y := Y_1(r_0)D + Y_0(r_0)(1 - D)$. ($D$ is binary.)

Motivation

Key Assumption: Local Monotonicity (Arai et al., 2022 and Hsu et al., 2024)

There exists a small $\epsilon > 0$ such that $T_\epsilon \in \{a, c, n\}$ almost surely.

- Also called no-defier assumption
- Under local continuity and local monotonicity,
  - $\theta_{FRD} = \frac{\mathbb{E}[Y|R=r_0^+]-\mathbb{E}[Y|R=r_0^-]}{\mathbb{E}[D|R=r_0^+]-\mathbb{E}[D|R=r_0^-]}$ identifies LATE for compliers (Hahn et al, 2001), where R is the running variable and r_0 is the cutoff point.
- Question
  - Existence of defiers?
  - What if local monotonicity fails to hold?

Existence of defiers (McCrary and Royer, 2011)

Key variables (McCrary and Royer, 2011)

Treatment : D=1 if student's year of education is above certain threshold, like 9 years

R : An individual's day of birth relative to the school entry date for the state in which the individual begins school, the cutoff points: two reforms in 1947 and 1972

- Existence of defiers (Fiorini and Stevens, 2021)
  - If a woman's parents would choose to delay her entrance into school if she were born before the school entry date, but would choose to petition the school district to allow her to begin school early if she were born after the school entry date

Existence of defiers (Kirkeboen et al, 2016)

Key variables (Kirkeboen et al, 2016)

Treatment : D=1 if student graduates with a degree in science rather than a degree in humanities

R : student's performance score for college admission, the cut-off is an admission threshold to a competitive major (say, science) rather than less competitive majors (say, humanities)

- Existence of defiers (Arai et al., 2022)
    - if some students, who tend to be attracted by nonmajored subjects and/or change their minds about their career choices, always switch from their assigned major to the other based on revisions of their beliefs or preferences

Literature Review

- LATE framework literature and relaxations of assumptions
  - Imbens and Angrist (1994), Angrist et al. (1996) and Vytlacil (2002)
  - Test the validity of IV: Huber and Mellace (2015), Kitagawa (2015), Mourifié and Wan (2017)
  - Independence: Kédagni (2023)
  - Exclusion restriction: Cui et al. (2024)
  - Monotonicity (related): Huber et al. (2017), de Chaisemartin (2017), Noack (2021), Fiorini and Stevens (2021), Cui et al. (2024)
- RD framework literature (SRD and FRD)
  - Thistlethwaite and Campbell (1960): SRD
  - RD Application: Angrist and Lavy (1999), van der Klaauw (2002), Lee et al. (2004), DiNardo and Lee (2011), Choi and Lee (2023)
  - FRD framework: Arai et al. (2022), Hsu et al. (2024)
    - Local continuity and local monotonicity

Contributions

1. Extend the LATE framework without monotonicity to FRD
2. Construct the sharp bounds on LATE for compliers and defiers
3. Apply the estimation on our bounds
4. Show through simulations how informative the bounds can be

Model and Identification

Definition (Four strata of people add "i")

$$T_\epsilon = \begin{cases} a, & \text{if } D(r) = 1, \text{ for } r \in B_\epsilon, \\ n, & \text{if } D(r) = 0, \text{ for } r \in B_\epsilon, \\ c, & \text{if } D(r) = 1\{r \geq r_0\}, \text{ for } r \in B_\epsilon, \\ df, & \text{if } D(r) = 1\{r < r_0\}, \text{ for } r \in B_\epsilon, \\ i, & \text{otherwise,} \end{cases} \tag{1}$$

Existence Example

where $B_\epsilon = \{r \in \mathcal{R} : |r - r_0| \leq \epsilon\}$ of the cutoffs.

Model Setting

Assumption 1. Local continuity in distributions

For $d \in \{0, 1\}, t \in \{a, c, n, df\}$, and all measurable subset $A \subseteq \mathcal{Y}$,

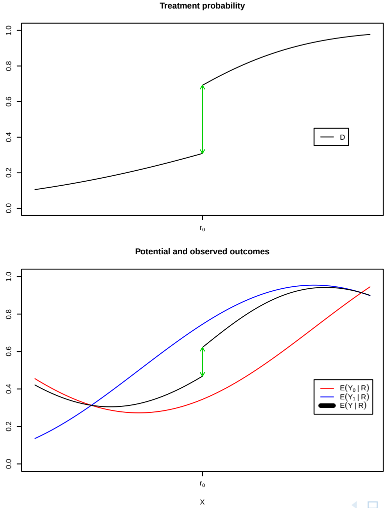
$$\lim_{\varepsilon \to 0} \mathbb{P}\left(Y_d \in A, T_\varepsilon = t \mid R = r_0 - \varepsilon\right) = \lim_{\varepsilon \to 0} \mathbb{P}\left(Y_d \in A, T_\varepsilon = t \mid R = r_0 + \varepsilon\right). \tag{2}$$

Assumption 2: Local no indefinite type

For $t = i$,

$$\lim_{\varepsilon \to 0} P\left(T_\varepsilon = t \mid Z = c - \varepsilon\right) = \lim_{\varepsilon \to 0} P\left(T_\varepsilon = t \mid Z = c + \varepsilon\right) = 0 \tag{3}$$

Model Setting (Fuzzy RD)



Model Setting

For any Borel set $A \in \mathcal{Y}$,

$$
\begin{aligned}
\mathbb{P}(Y \in A, D = 1 | R = r_0^+) &= \mathbb{P}(T = c | R = r_0^+)\mathbb{P}(Y_1(r_0) \in A | T = c, R = r_0^+) \\
&\quad + \mathbb{P}(T = a | R = r_0^+)\mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0^+) \\
&= \mathbb{P}(T = c | R = r_0)\mathbb{P}(Y_1(r_0) \in A | T = c, R = r_0) \\
(Assumption\ 1) &\quad + \mathbb{P}(T = a | R = r_0)\mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0), \\
\mathbb{P}(Y \in A, D = 1 | R = r_0^-) &= \mathbb{P}(T = df | R = r_0^-)\mathbb{P}(Y_1(r_0) \in A | T = df, R = r_0^-) \\
&\quad + \mathbb{P}(T = a | R = r_0^-)\mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0^-) \\
&= \mathbb{P}(T = df | R = r_0)\mathbb{P}(Y_1(r_0) \in A | T = df, R = r_0) \\
(Assumption\ 1) &\quad + \mathbb{P}(T = a | R = r_0)\mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0),
\end{aligned} \tag{4}
$$

Similar results hold for $D = 0$. ▸ Visualization

LATE

Proposition 1: Sharp bounds for $p_{df|r_0}$: $\Theta_I(p_{df|r_0})$ ▸ Proof ▸ DGP Example

$$\max\left\{\max\left\{\sup_A\{\mathbb{P}(Y\in A, D=1|R=r_0^-)-\mathbb{P}(Y\in A, D=1|R=r_0^+)\},\right.\right.$$
$$\left.\left.\sup_A\{\mathbb{P}(Y\in A, D=0|R=r_0^+)-\mathbb{P}(Y\in A, D=0|R=r_0^-)\}\right\},0\right\} \tag{5}$$
$$\leq \mathbb{P}(T=df|R=r_0)\equiv p_{df|r_0}\leq \min\{\mathbb{E}[D|R=r_0^-],\mathbb{E}[1-D|R=r_0^+]\}.$$

Proposition 2: Sharp bounds for $p_{a|r_0}$ and $p_{n|r_0}$ ▸ Proof

$$p_{a|r_0}=\mathbb{E}[D\mid R=r_0^-]-p_{df|r_0},\; p_{n|r_0}=\mathbb{E}[1-D\mid R=r_0^+]-p_{df|r_0}$$

LATE (Cont')

Notation: The identified sets for $F_{dtr} \equiv F_{Y_d(r)|T=t,R=r}$, $d \in \{0,1\}$, $t \in \{c, df\}$ are given in Proposition 3 Proof.

Proposition 3: Pointwise sharp bounds for the distributions F_{dtr_0}

For a given p_{df} interior point of $\Theta_I(p_{df})$, pointwise sharp bounds for the distributions F_{dtr_0} are given below:

$$\max\left\{F_{1ar_0}^{LB_+}(y), F_{1ar_0}^{LB_-}(y)\right\} \leq F_{1ar_0}(y) \leq \min\left\{F_{1ar_0}^{UB_+}(y), F_{1ar_0}^{UB_-}(y)\right\},$$

$$\max\left\{F_{0nr_0}^{LB_+}(y), F_{0nr_0}^{LB_-}(y)\right\} \leq F_{0nr_0}(y) \leq \min\left\{F_{0nr_0}^{UB_+}(y), F_{0nr_0}^{UB_-}(y)\right\},$$

$$F_{1cr_0}(y) = \frac{\mathbb{P}(Y \leq y, D=1|R=r_0^+) - p_a F_{1ar_0}(y)}{p_c}, \quad F_{0cr_0}(y) = \frac{\mathbb{P}(Y \leq y, D=0|R=r_0^-) - p_n F_{0nr_0}(y)}{p_c},$$

$$F_{1dfr_0}(y) = \frac{\mathbb{P}(Y \leq y, D=1|R=r_0^-) - p_a F_{1ar_0}(y)}{p_{df}}, \quad F_{0dfr_0}(y) = \frac{\mathbb{P}(Y \leq y, D=0|R=r_0^+) - p_n F_{0nr_0}(y)}{p_{df}}$$

LATE (Cont')

Let $\mu_{dtr} \equiv \mathbb{E}[Y_d | T = t, R = r]$, μ_F is the expected value of a given cdf F.

Lemma 1 (Lee's bound), $\Delta \in \{+, -\}$

$$\mu_{F_{1ar_0}^{LB_\Delta}} = \mathbb{E}\left[Y | D = 1, R = r_0^\Delta, Y > F^{-1}_{Y|D=1,R=r_0^\Delta}\left(1 - \frac{p_a}{\mathbb{E}[D|R = r_0^\Delta]}\right)\right],$$

$$\mu_{F_{1ar_0}^{UB_\Delta}} = \mathbb{E}\left[Y | D = 1, R = r_0^\Delta, Y < F^{-1}_{Y|D=1,R=r_0^\Delta}\left(\frac{p_a}{\mathbb{E}[D|R = r_0^\Delta]}\right)\right],$$

$$\mu_{F_{0nr_0}^{LB_\Delta}} = \mathbb{E}\left[Y | D = 0, R = r_0^\Delta, Y > F^{-1}_{Y|D=0,R=r_0^\Delta}\left(1 - \frac{p_n}{\mathbb{E}[1 - D|R = r_0^\Delta]}\right)\right],$$

$$\mu_{F_{0nr_0}^{UB_\Delta}} = \mathbb{E}\left[Y | D = 0, R = r_0^\Delta, Y < F^{-1}_{Y|D=0,R=r_0^\Delta}\left(\frac{p_n}{\mathbb{E}[1 - D|R = r_0^\Delta]}\right)\right].$$

Visualization Explanation (Lee, 2009)

$$\mathbb{E}[Y|D=1,R=r_0^+] = \frac{p_{a|r_0}}{\mathbb{E}(D|R=r_0^+)}\mu_{1ar_0} + \left(1-\frac{p_{a|r_0}}{\mathbb{E}(D|R=r_0^+)}\right)\mu_{1cr_0}$$

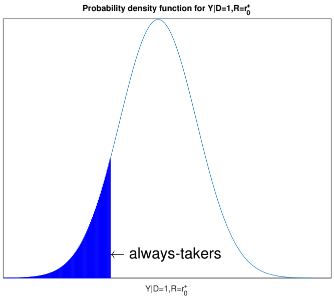


Figure: Worst-case scenario 1

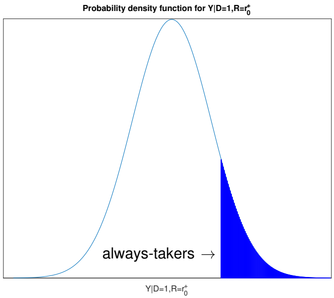


Figure: Worst-case scenario 2

LATE (Cont')

Theorem 1: Sharp bounds for μ_{1cr_0} and μ_{1dfr_0} Proof

Suppose p_{df} is an interior point of $\Theta_I(p_{df})$. Under Assumptions 1 to 2, sharp bounds for μ_{1cr_0} and μ_{1dfr_0} are as follows:

$$\mu_{1cr_0}(p_a) = \frac{\mathbb{E}[YD|R = r_0^+] - p_a\mu_{1ar_0}(p_a)}{\mathbb{E}[D|R = r_0^+] - p_a},$$

$$\mu_{1dfr_0}(p_a) = \frac{\mathbb{E}[YD|R = r_0^-] - p_a\mu_{1ar_0}(p_a)}{\mathbb{E}[D|R = r_0^-] - p_a},$$

where $p_a = \mathbb{E}[D|R = r_0^-] - p_{df}$, $p_{df} \in \mathring{\Theta}_I(p_{df})$ and $\mu_{1ar_0}(p_a) \in \left[\mu_{F^{UB}_{1ar_0}}(p_a), \mu_{F^{LB}_{1ar_0}}(p_a)\right]$.

$\mu_{0cr_0}(p_n)$ and $\mu_{0dfr_0}(p_n)$ could be solved similarly.

Estimation

Estimation

First, we focus on the lower bound for $\mu_{1cr_0}(p_a)$:

$$\mu_{1cr_0}(p_a) = \frac{\mathbb{E}[YD|R = r_0^+] - p_a \mu_{1ar_0}(p_a)}{\mathbb{E}[D|R = r_0^+] - p_a}$$

Following Hahn et al. (2001), we use local linear kernel regression to estimate these limits. Let $K(\cdot)$ and h_n denote the kernel function and bandwidth, respectively. For estimation of $\mathbb{E}[YD|R = r_0^+]$, the local linear regression is

$$\begin{aligned}(\hat{a}_r, \hat{b}_r) = \arg\min_{a_r, b_r} \sum_{i=1}^{n} (Y_i D_i - a_r - (R_i - r_0) b_r)^2 \\ \times K\left(\frac{R_i - r_0}{h}\right) 1\{R_i \geq r_0\}\end{aligned} \tag{6}$$

Estimation: Bandwidth Selection

For estimation of the limits, we are interested in the bandwidth h that minimizes

$$L^{\text{right}}(r,h) = \mathbb{E}\left[\left(\lim_{r\downarrow r_0} Y(r)D(r) - \hat{a}_r(r,h)\right)^2\right]$$

Now define the cross-validation criterion for the right limit as

$$CV_{YD}^{\text{right}}(h) = \frac{1}{n}\sum_{i=1}^{n}\Big(Y_iD_i1\{r_0 < r \leq r_0 + h\} - \hat{a}_r\,(r,h)\Big)^2$$

with the corresponding cross-validation choice for the bandwidth

$$h_{CV}^{opt} = \arg\min_h CV_{YD}^{\text{right}}(h).$$

Also, we define

$$CV_{YD}(h) = \big(CV_{YD}^{\text{right}}(h), CV_{YD}^{\text{left}}(h)\big).$$

Estimation: Bandwidth Selection (Cont')

Our parameters of interest are as follows:

$$\begin{aligned}&\mathbb{E}[YD|R = r_0^\Delta],\\&\mathbb{E}[D|R = r_0^\Delta],\\&\mathbb{E}[Y(1-D)|R = r_0^\Delta],\\&\mathbb{E}[1-D|R = r_0^\Delta].\end{aligned}$$

where $\Delta = +/-$, meaning the right or left limit.
Thus, we have the valid bandwidth h as

$$\begin{aligned}h_{CV}^{\text{valid}} = \min\Big(&\arg\min_h CV_{YD}(h), \arg\min_h CV_D(h),\\&\arg\min_h CV_{Y(1-D)}(h), \arg\min_h CV_{1-D}(h)\Big).\end{aligned}$$

Simulation Results

DGP 1

Consider the following data-generating process

$$
\begin{cases}
R & \sim & N(0,1) \text{ truncated at -2 and 2} \\
Y \mid (D=1, R=r) & \sim & N(1,1) \\
Y \mid (D=0, R=r) & \sim & N(0,1) \\
P(D=1 \mid R=r) & = & 1\{-2 \le r < 0\}\frac{(r+2)^2}{16} + 1\{0 \le r \le 2\}\left(1 - \frac{(r-2)^2}{16}\right)
\end{cases} \tag{7}
$$

where $r_0 = 0$.

Check for discontinuity:

$$
\begin{cases}
\lim_{r \downarrow r_0} P(D=1 \mid R=r) & \neq & \lim_{r \uparrow r_0} P(D=1 \mid R=r) \\
\lim_{r \downarrow r_0} P(D=0 \mid R=r) & \neq & \lim_{r \uparrow r_0} P(D=0 \mid R=r)
\end{cases} \tag{8}
$$

DGP 1: Bounds for LATE with IV

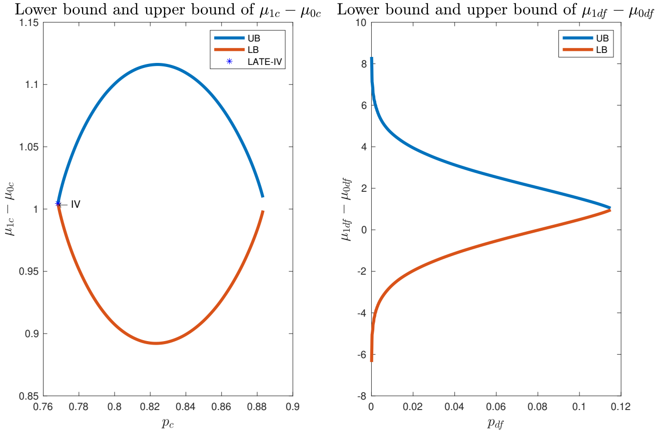


Figure: LATE bounds for compliers and defiers

DGP 2

Consider the following data-generating process

$$\begin{cases} R & \sim & N(0,1) \text{ truncated at -2 and 2} \\ Y \mid (D=1, R=r) & \sim & N(0.01, 1) \\ Y \mid (D=0, R=r) & \sim & N(0,1) \\ P(D=1 \mid R=r) & = & 1\{-2 \leq r < 0\} \times 0.45 + 1\{0 \leq r \leq 2\} \times 0.5 \end{cases} \tag{9}$$

where $r_0 = 0$.

Check for discontinuity:

$$\begin{cases} \lim_{r \downarrow r_0} P(D=1 \mid R=r) & \neq & \lim_{r \uparrow r_0} P(D=1 \mid R=r) \\ \lim_{r \downarrow r_0} P(D=0 \mid R=r) & \neq & \lim_{r \uparrow r_0} P(D=0 \mid R=r) \end{cases} \tag{10}$$

DGP 2: Bounds for LATE with IV

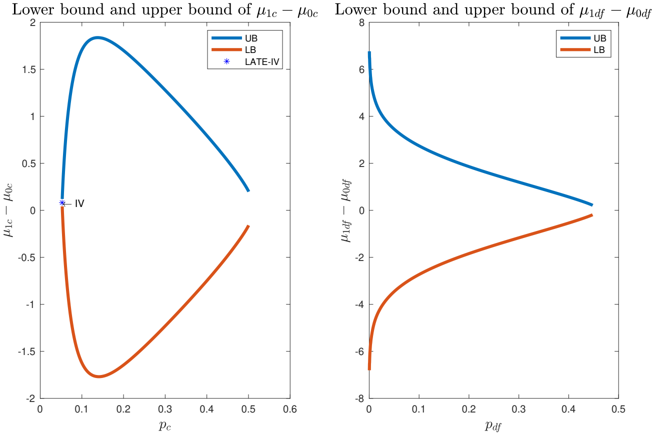


Figure: LATE bounds for compliers and defiers

Conclusion

Conclusion

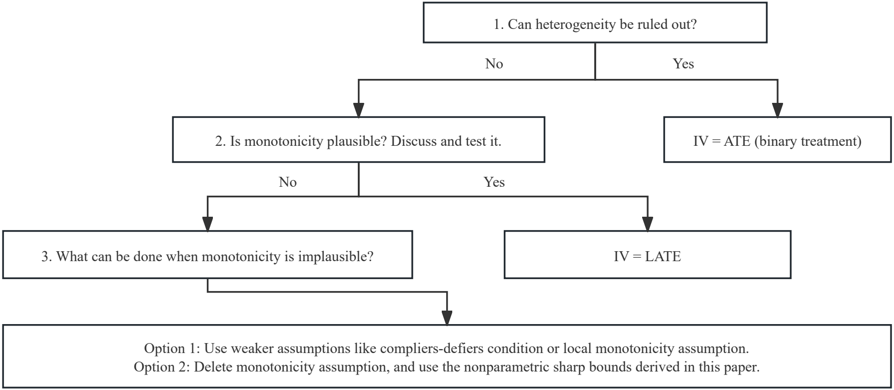


Figure: Recommendations for empirical researchers

Conclusion

Main contributions and strengths:

1. Extend the LATE framework without monotonicity to FRD
2. Derive the sharp bounds for LATE for compliers and defiers

Future work:

1. Finish the empirical examples
2. Finish the inference part (Chernozhukov et al. (2011) precision-corrected estimator)

Thanks!

Appendix

Definition (cont')

For the existence of type "indefinite", we can consider the single-threshold crossing specification of potential treatment like Hsu et al. (2024),

$$D(r) = 1\{r + 1\{r \geq r_0\} + X < 0\}, r \in [r_0 - \epsilon, r_0 + \epsilon]$$

where $X \sim N(0, 1)$. For simplicity, we fix $\epsilon > 0$ and $r_0 = 0$, the support of X can be divided into the groups from figure below. back

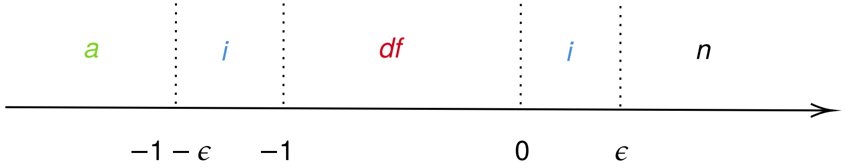


Figure: The distribution of types under the support of X

Visualization of Types

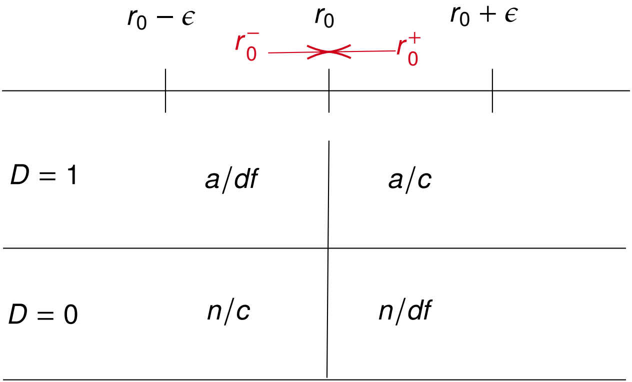


Figure: Visualization of types

back

Proof of Proposition 1 (Lower bound)

Recall Equation 4:

$$\mathbb{P}(Y \in A, D = 1|R = r_0^-) - \mathbb{P}(Y \in A, D = 1|R = r_0^+) =$$
$$p_{df|r_0}\mathbb{P}(Y_1(r_0) \in A|T = df, R = r_0) - p_{c|r_0}\mathbb{P}(Y_1(r_0) \in A|T = c, R = r_0).$$

Note that

$$\mathbb{P}(Y \in A, D = 1|R = r_0^-) - \mathbb{P}(Y \in A, D = 1|R = r_0^+)$$
$$= p_{df|r_0}\mathbb{P}(Y_1(r_0) \in A \mid T = df, R = r_0) - \underbrace{p_{c|r_0}\mathbb{P}(Y_1(r_0) \in A \mid T = c, R = r_0)}_{\geq 0}$$
$$\leq p_{df|r_0}\mathbb{P}(Y_1(r_0) \in A \mid T = df, R = r_0)$$
$$\leq p_{df|r_0}$$

back

Proof of Proposition 1 (Upper bound)

Recall Equation 4:

$$\mathbb{P}(Y \in A, D = 0 | R = r_0^+) =$$
$$p_{df|r_0} \mathbb{P}(Y_0(r_0) \in A | T = df, R = r_0) + p_{n|r_0} \mathbb{P}(Y_0(r_0) \in A | T = n, R = r_0).$$
$$\mathbb{P}(Y \in A, D = 1 | R = r_0^-) =$$
$$p_{df|r_0} \mathbb{P}(Y_1(r_0) \in A | T = df, R = r_0) + p_{a|r_0} \mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0).$$

If we let $A = \mathcal{Y}$:

$$\begin{aligned} \mathbb{P}(D = 0 \mid R = r_0^+) &= p_{df|r_0} + p_{n|r_0} \\ \mathbb{P}(D = 1 \mid R = r_0^-) &= p_{df|r_0} + p_{a|r_0} \end{aligned} \tag{11}$$

$$p_{df|r_0} \leq \min\{\mathbb{P}(D = 0 \mid R = r_0^+), \mathbb{P}(D = 1 \mid R = r_0^-)\}$$

back

Violation of Testable Implication

Consider the following data-generating process:

- Let $R \sim N(0,1)$ truncated at -2 and 2. For a value greater than 0.1, we take the reverse sign of itself.
- The propensity score:
  $P(D=1 \mid R=r) = 1\{-2 \leq r < 0\}\frac{(r+2)^2}{16} + 1\{0 \leq r \leq 2\}\left(1 - \frac{(r-2)^2}{16}\right)$.
- $Y \mid (D=1, R=r) \sim N(1,1)$ for all r and $Y \mid (D=0, R=r) \sim N(0,1)$ for all r. back

We have $\min\{\mathbb{E}[D|R=r_0^-], \mathbb{E}[1-D|R=r_0^+]\} = 0.0000024$.

$$\max\left\{\max\left\{\sup_A\{\mathbb{P}(Y \in A, D=1|R=r_0^-) - \mathbb{P}(Y \in A, D=1|R=r_0^+)\},\right.\right.$$

$$\left.\left.\sup_A\{\mathbb{P}(Y \in A, D=0|R=r_0^+) - \mathbb{P}(Y \in A, D=0|R=r_0^-)\}\right\}, 0\right\} = 0.0000037.$$

Proof of Proposition 2

$$\begin{aligned}
\mathbb{P}(D=1|R=r_0^-) &= \mathbb{P}(D=1,T=df|R=r_0^-)+\mathbb{P}(D=1,T=a|R=r_0^-) \\
&\quad \text{(Law of total probability)} \\
&= \mathbb{P}(T=df|R=r_0^-)\mathbb{P}(D=1|T=df,R=r_0^-) \\
&\quad +\mathbb{P}(T=a|R=r_0^-)\mathbb{P}(D=1|T=a,R=r_0^-) \\
&= \mathbb{P}(T=df|R=r_0^-)+\mathbb{P}(T=a|R=r_0^-) \\
&= \mathbb{P}(T=df|R=r_0)+\underbrace{\mathbb{P}(T=a|R=r_0)}_{p_{a|r_0}} \quad \text{(Assumption 1)}
\end{aligned}$$

back

Proof of Proposition 3

To proceed, we are going to derive sharp bounds for $\mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0)$ using Equations (4). For simplicity, suppose p_{df} is an interior point of $\Theta_I(p_{df})$. Equations (4) implies

$$\mathbb{P}(Y_1(r_0) \in A | T = a, R = r_0) = \frac{\mathbb{P}(Y \in A, D = 1 | R = r_0^+) - p_c \mathbb{P}(Y_1(r_0) \in A | T = c}{p_a}$$

Since $\mathbb{P}(Y_1(r_0) \in A | T = a/c, R = r_0) \in [0, 1]$, we have

$$\begin{aligned} \max\left\{\frac{\mathbb{P}(Y \in A, D = 1 \mid R = r_0^+) - p_c}{p_a}, 0\right\} &\leq \mathbb{P}(Y_1(r_0) \in A \mid T = a, R = r_0) \\ &\leq \min\left\{\frac{\mathbb{P}(Y \in A, D = 1 \mid R = r_0^+)}{p_a}, 1\right\} \end{aligned} \tag{13}$$

back

Proof of Theorem 1

For the mixture representation:

$$\mathbb{E}[Y|D=1,R=r_0^+] = \frac{p_{a|r_0}}{\mathbb{E}(D|R=r_0^+)}\mu_{1ar_0} + \frac{\mathbb{E}(D|R=r_0^+) - p_{a|r_0}}{E(D|R=r_0^+)}\mu_{1cr_0},$$

$$\mu_{1cr_0} = \frac{\mathbb{E}[YD|R=r_0^+] - p_{a|r_0}\mu_{1ar_0}}{\mathbb{E}[D|R=r_0^+] - p_{a|r_0}}$$

back