

# Noise-Induced Randomization in Regression Discontinuity Designs

Dean Eckles, Nikolaos Ignatiadis, Stefan Wager, Han Wu

Presented by: Sai Zhang

November 18, 2022

# Outline

# Confidence Intervals

# Asymptotic Normality

$$\hat{\tau} = \frac{\sum_i \gamma_+(Z_i) Y_i}{\sum_i \gamma_+(Z_i)} - \frac{\sum_i \gamma_-(Z_i) Y_i}{\sum_i \gamma_-(Z_i)}$$
$$\hat{\tau}_\gamma \xrightarrow{p} \theta_\gamma = \frac{\mathbb{E} [\alpha_{(1)}(U) h(U, \gamma_+)]}{\mathbb{E} [h(U, \gamma_+)]} - \frac{\mathbb{E} [\alpha_{(0)}(U) h(U, \gamma_-)]}{\mathbb{E} [h(U, \gamma_-)]}$$

# Asymptotic Normality

## Theorem: Asymptotic Normality of $\hat{\tau}$

Suppose the sequence of weighting kernels  $\gamma_+^{(n)}$  and  $\gamma_-^{(n)}$  is deterministic, and  $\exists \beta \in (0, \frac{1}{2})$ ,  $C, C' > 0$  s.t.  $\forall n$  large enough:

$$\sup_z \left| \gamma_\diamond^{(n)}(z) \right| < C n^\beta \mathbb{E} \left[ \gamma_\diamond^{(n)}(Z_i) \right] \quad \sup_u \left| h(u, \gamma_\diamond^{(n)}) \right| < C' \mathbb{E} \left[ \gamma_\diamond^{(n)}(Z_i) \right], \quad \diamond = \{+, -\}$$

Then

$$\frac{\sqrt{n} (\hat{\tau}_\gamma - \theta_\gamma)}{\sqrt{V_\gamma}} \xrightarrow{d} \mathcal{N}(0, 1)$$

where

$$V_\gamma = \frac{\mathbb{E} \left[ \gamma_+^2(Z_i) (Y_i - \mu_{\gamma,+})^2 \right]}{\mathbb{E} [\gamma_+(Z_i)]^2} + \frac{\mathbb{E} \left[ \gamma_-^2(Z_i) (Y_i - \mu_{\gamma,-})^2 \right]}{\mathbb{E} [\gamma_-(Z_i)]^2}$$

# Asymptotic Normality

## Theorem: Asymptotic Normality of $\hat{\tau}$

Suppose the sequence of weighting kernels  $\gamma_+^{(n)}$  and  $\gamma_-^{(n)}$  is deterministic, and  $\exists \beta \in (0, \frac{1}{2})$ ,  $C, C' > 0$  s.t.  $\forall n$  large enough:  $\sup_z |\gamma_\diamond^{(n)}(z)| < C n^\beta \mathbb{E}[\gamma_\diamond^{(n)}(Z_i)]$ ,  $\sup_u |h(u, \gamma_\diamond^{(n)})| < C' \mathbb{E}[\gamma_\diamond^{(n)}(Z_i)]$  where  $\diamond = \{+, -\}$ . Then

$$\frac{\sqrt{n}(\hat{\tau}_\gamma - \theta_\gamma)}{\sqrt{V_\gamma}} \xrightarrow{d} \mathcal{N}(0, 1)$$

$$\text{where } V_\gamma = \frac{\mathbb{E}[\gamma_+^2(Z_i)(Y_i - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2} + \frac{\mathbb{E}[\gamma_-^2(Z_i)(Y_i - \mu_{\gamma,-})^2]}{\mathbb{E}[\gamma_-(Z_i)]^2}$$

Assumption:

- The response  $Y_i$  is bounded:  $Y_i \in [0, 1]$
- $\inf_z \text{Var}[Y_i | Z_i = z] > 0$

## Asymptotic Normality: A Sketch of the Proof

$$\frac{\sqrt{n} [(\hat{\mu}_{\gamma,+} - \hat{\mu}_{\gamma,-}) - (\mu_{\gamma,+} - \mu_{\gamma,-})]}{\sqrt{\frac{\mathbb{E}[\gamma_+^2(Z_i)(Y_i - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2} + \frac{\mathbb{E}[\gamma_-^2(Z_i)(Y_i - \mu_{\gamma,-})^2]}{\mathbb{E}[\gamma_-(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

## Asymptotic Normality: A Sketch of the Proof

$$\frac{\sqrt{n} [(\hat{\mu}_{\gamma,+} - \hat{\mu}_{\gamma,-}) - (\mu_{\gamma,+} - \mu_{\gamma,-})]}{\sqrt{\frac{\mathbb{E}[\gamma_+^2(Z_i)(Y_i - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2} + \frac{\mathbb{E}[\gamma_-^2(Z_i)(Y_i - \mu_{\gamma,-})^2]}{\mathbb{E}[\gamma_-(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

$$\frac{\sqrt{n} (\hat{\mu}_{\gamma,+} - \mu_{\gamma,+})}{\sqrt{\frac{\mathbb{E}[\gamma_+(Z_i)^2(Y_i(1) - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$



## Asymptotic Normality: A Sketch of the Proof

$$\frac{\sqrt{n} [(\hat{\mu}_{\gamma,+} - \hat{\mu}_{\gamma,-}) - (\mu_{\gamma,+} - \mu_{\gamma,-})]}{\sqrt{\frac{\mathbb{E}[\gamma_+^2(Z_i)(Y_i - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2} + \frac{\mathbb{E}[\gamma_-^2(Z_i)(Y_i - \mu_{\gamma,-})^2]}{\mathbb{E}[\gamma_-(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

$$\frac{\sqrt{n} (\hat{\mu}_{\gamma,+} - \mu_{\gamma,+})}{\sqrt{\frac{\mathbb{E}[\gamma_+(Z_i)^2(Y_i(1) - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

$$\frac{\sqrt{n} \left( \frac{\sum_i \gamma_+(Z_i) Y_i}{\sum_i \gamma_+(Z_i)} - \mu_{\gamma,+} \right)}{\sqrt{\frac{\mathbb{E}[\gamma_+(Z_i)^2(Y_i(1) - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

## Asymptotic Normality: A Sketch of the Proof

$$\frac{\sqrt{n} \left( \frac{\sum_i \gamma_+(Z_i) Y_i}{\sum_i \gamma_+(Z_i)} - \mu_{\gamma,+} \right)}{\sqrt{\frac{\mathbb{E}[\gamma_+(Z_i)^2 (Y_i(1) - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

## Asymptotic Normality: A Sketch of the Proof

$$\frac{\sqrt{n} \left( \frac{\sum_i \gamma_+(Z_i) Y_i}{\sum_i \gamma_+(Z_i)} - \mu_{\gamma,+} \right)}{\sqrt{\frac{\mathbb{E}[\gamma_+(Z_i)^2 (Y_i(1) - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

$$\frac{\frac{\sum_i \gamma_+(Z_i) (Y_i(1) - \mu_{\gamma,+})}{\sum_i \gamma_+(Z_i)}}{\sqrt{\frac{n \mathbb{E}[\gamma_+(Z_i)^2 (Y_i(1) - \mu_{\gamma,+})^2]}{n \mathbb{E}[\gamma_+(Z_i)]}}} = \underbrace{\frac{\sum_i \gamma_+(Z_i) (Y_i(1) - \mu_{\gamma,+})}{\sqrt{n \mathbb{E}[\gamma_+(Z_i)^2 (Y_i(1) - \mu_{\gamma,+})^2]}}}_{\xrightarrow{d} \mathcal{N}(0,1)} \cdot \underbrace{\frac{1}{\frac{\frac{1}{n} \sum_i \gamma_+(Z_i)}{\mathbb{E}[\gamma_+(Z_i)]}}}_{=1+o_p(1)} \xrightarrow{d} \mathcal{N}(0, 1)$$

# Asymptotic Normality: A Sketch of the Proof

$$\frac{\sum_i \gamma_+(Z_i) (Y_i(1) - \mu_{\gamma,+})}{\sqrt{n \mathbb{E} [\gamma_+(Z_i)^2 (Y_i(1) - \mu_{\gamma,+})^2]}} \xrightarrow{d} \mathcal{N}(0, 1)$$

- $\underline{\sigma}^2 = \inf_z \text{Var} [Y_i \mid Z_i = z] > 0$ :  $\text{Var} [\gamma_+(Z_i) (Y_i(1) - \mu_{\gamma,+})] \geq \underline{\sigma}^2 \mathbb{E} [\gamma_+(Z_i)^2]$
- $\sup_u \left| h(u, \gamma_\diamond^{(n)}) \right| < C' \mathbb{E} [\gamma_\diamond^{(n)}(Z_i)]$  and  $Y_i \in [0, 1]$ :  $|\mu_{\gamma,+}| = \left| \frac{\mathbb{E}[\alpha_{(1)}(U)h(U, \gamma_+)]}{\mathbb{E}[\gamma_+(Z_i)]} \right| \leq C'$
- $\sup_z \left| \gamma_\diamond^{(n)}(z) \right| < Cn^\beta \mathbb{E} [\gamma_\diamond^{(n)}(Z_i)]$  and Lyapunov CLT condition:

$$\frac{n \mathbb{E} [|\gamma_+(Z_i) (Y_i(1) - \mu_{\gamma,+})|^{2+q}]}{(n \text{Var} [\gamma_+(Z_i) (Y_i(1) - \mu_{\gamma,+})])^{\frac{2+q}{2}}} \leq \left( \frac{C' + 1}{\underline{\sigma}} \right)^{2+q} \left( Cn^{\beta - \frac{1}{2}} \right)^q \xrightarrow{0 \rightarrow \infty} 0$$

## Asymptotic Normality: A Sketch of the Proof

$$\frac{\frac{1}{n} \sum_i \gamma_+(Z_i)}{\mathbb{E}[\gamma_+(Z_i)]} \xrightarrow{p} 1$$

■  $\left[ \sup_z \left| \gamma_{\diamond}^{(n)}(z) \right| < C n^{\beta} \mathbb{E} \left[ \gamma_{\diamond}^{(n)}(Z_i) \right] \right]$  and Chebyshev's inequality:

$$\begin{aligned} \mathbf{P} \{ |\mathbb{E}_n[\gamma_+(Z_i)] - \mathbb{E}[\gamma_+(Z_i)]| \geq \epsilon \mathbb{E}[\gamma_+(Z_i)] \} &\leq \frac{\text{Var}[\gamma_+(Z_i)]}{n \epsilon^2 \mathbb{E}[\gamma_+(Z_i)]^2} \\ &\leq \left( \frac{C}{\epsilon} \cdot n^{\beta - \frac{1}{2}} \right)^2 \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

## Asymptotic Normality: A Sketch of the Proof

$$\begin{aligned}
 \frac{\frac{\sum_i \gamma_+(Z_i)(Y_i(1) - \mu_{\gamma,+})}{\sum_i \gamma_+(Z_i)}}{\sqrt{\frac{n\mathbb{E}[\gamma_+(Z_i)^2(Y_i(1) - \mu_{\gamma,+})^2]}{n\mathbb{E}[\gamma_+(Z_i)]}}} &= \underbrace{\frac{\sum_i \gamma_+(Z_i)(Y_i(1) - \mu_{\gamma,+})}{\sqrt{n\mathbb{E}[\gamma_+(Z_i)^2(Y_i(1) - \mu_{\gamma,+})^2]}}}_{\xrightarrow{d} \mathcal{N}(0,1)} \cdot \underbrace{\frac{1}{\frac{\frac{1}{n} \sum_i \gamma_+(Z_i)}{\mathbb{E}[\gamma_+(Z_i)]}}}_{=1+o_p(1)} \xrightarrow{d} \mathcal{N}(0,1) \\
 \Rightarrow \frac{\sqrt{n} \left( \frac{\sum_i \gamma_+(Z_i)Y_i}{\sum_i \gamma_+(Z_i)} - \mu_{\gamma,+} \right)}{\sqrt{\frac{\mathbb{E}[\gamma_+(Z_i)^2(Y_i(1) - \mu_{\gamma,+})^2]}{\mathbb{E}[\gamma_+(Z_i)]^2}}} &\xrightarrow{d} \mathcal{N}(0,1) \Rightarrow \frac{\sqrt{n}(\hat{\tau}_\gamma - \theta_\gamma)}{\sqrt{V_\gamma}} \xrightarrow{d} \mathcal{N}(0,1)
 \end{aligned}$$

## Plug-in Estimator for $V_\gamma$

$$V_\gamma = \frac{\mathbb{E} \left[ \gamma_+^2 (Z_i) (Y_i - \mu_{\gamma,+})^2 \right]}{\mathbb{E} [\gamma_+ (Z_i)]^2} + \frac{\mathbb{E} \left[ \gamma_-^2 (Z_i) (Y_i - \mu_{\gamma,-})^2 \right]}{\mathbb{E} [\gamma_- (Z_i)]^2}$$

## Plug-in Estimator for $V_\gamma$

$$V_\gamma = \frac{\mathbb{E} \left[ \gamma_+^2 (Z_i) (Y_i - \mu_{\gamma,+})^2 \right]}{\mathbb{E} [\gamma_+ (Z_i)]^2} + \frac{\mathbb{E} \left[ \gamma_-^2 (Z_i) (Y_i - \mu_{\gamma,-})^2 \right]}{\mathbb{E} [\gamma_- (Z_i)]^2}$$

### Proposition: Plug-in Estimator $\hat{V}_\gamma$

Under the same assumptions, we have  $\frac{\hat{V}_\gamma}{V_\gamma} = 1 + o_p(1)$  where

$$\hat{V}_\gamma = \frac{\frac{1}{n} \sum_{i=1}^n \gamma_+^2 (Z_i) (Y_i - \hat{\mu}_{\gamma,+})^2}{\left[ \frac{1}{n} \sum_{i=1}^n \gamma_+ (Z_i) \right]^2} + \frac{\frac{1}{n} \sum_{i=1}^n \gamma_-^2 (Z_i) (Y_i - \hat{\mu}_{\gamma,-})^2}{\left[ \frac{1}{n} \sum_{i=1}^n \gamma_- (Z_i) \right]^2}$$



## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

$$\begin{aligned} a\text{Bias} = \theta_\gamma - \tau_w &= \underbrace{\int \left( \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} - \frac{h(u, \gamma_-)}{\mathbb{E}_G[h(U, \gamma_-)]} \right) \alpha_{(0)}(u) dG(u)}_{\text{Confounding bias}} \\ &+ \underbrace{\int \left( \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} - \frac{w(u)}{\mathbb{E}_G[w(U)]} \right) \tau(u) dG(u)}_{\text{CATE heterogeneity bias}} \end{aligned}$$

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

$$\begin{aligned} a\text{Bias} = \theta_\gamma - \tau_w = & \underbrace{\int \left( \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} - \frac{h(u, \gamma_-)}{\mathbb{E}_G[h(U, \gamma_-)]} \right) \alpha_{(0)}(u) dG(u)}_{\text{Confounding bias}} \\ & + \underbrace{\int \left( \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} - \frac{w(u)}{\mathbb{E}_G[w(U)]} \right) \tau(u) dG(u)}_{\text{CATE heterogeneity bias}} \end{aligned}$$

The distribution of  $G(\cdot)$  is **unknown**

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

Bound the **worst-case** bias:

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

Bound the **worst-case** bias:

- Back out the class of latent variable distribution from  $\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(Z_i \leq t)$ , the empirical distribution of  $Z_i$  (**massart1990tight**):

$$\mathcal{G}_n = \left\{ G(\cdot) : \sup_{t \in \mathbb{R}} |F_G(t) - \hat{F}_n(t)| \leq \sqrt{\frac{\log(2/\alpha_n)}{2n}} \right\}, \quad \alpha_n = \min \{0.05, n^{-1/4}\}$$

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

Bound the **worst-case** bias:

- Back out the class of latent variable distribution from  $\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(Z_i \leq t)$ , the empirical distribution of  $Z_i$  (**massart1990tight**):

$$\mathcal{G}_n = \left\{ G(\cdot) : \sup_{t \in \mathbb{R}} |F_G(t) - \hat{F}_n(t)| \leq \sqrt{\frac{\log(2/\alpha_n)}{2n}} \right\}, \quad \alpha_n = \min \{0.05, n^{-1/4}\}$$

- Take treatment effect heterogeneity into consideration:

$$\mathcal{T}_M = \{\tau(\cdot) \mid \tau(u) = \bar{\tau} + \Delta(u), \bar{\tau} \in \mathbb{R}, |\Delta(u)| \leq M\}, \quad M \in [0, 1]$$

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

Bound the **worst-case** bias:

- Back out the class of latent variable distribution from  $\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(Z_i \leq t)$ , the empirical distribution of  $Z_i$  (**massart1990tight**):

$$\mathcal{G}_n = \left\{ G(\cdot) : \sup_{t \in \mathbb{R}} |F_G(t) - \hat{F}_n(t)| \leq \sqrt{\frac{\log(2/\alpha_n)}{2n}} \right\}, \quad \alpha_n = \min \{0.05, n^{-1/4}\}$$

- Take treatment effect heterogeneity into consideration:

$$\mathcal{T}_M = \{\tau(\cdot) \mid \tau(u) = \bar{\tau} + \Delta(u), \bar{\tau} \in \mathbb{R}, |\Delta(u)| \leq M\}, \quad M \in [0, 1]$$

- $\mathcal{T}_0$  ( $M = 0$ ): constant CATE
- $\mathcal{T}_1$  ( $M = 1$ ): no assumptions
- $\mathcal{T}_{1/2}$  ( $M = 0$ ): a conservative choice for a monotonicity restriction

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

### Proposition: Upper Bound for the Potential Bias $\hat{B}_{\gamma,M}$

Under asymptotic normality and all necessary assumptions, for  $\tau(\cdot) \in \mathcal{T}_M$ , the upper bound of bias is

$$\hat{B}_{\gamma,M} = \sup \left\{ |\text{Bias}[\gamma_\pm, \tau_w; \alpha_0(\cdot), \tau(\cdot), G]| : G \in \mathcal{G}_n, \alpha_{(0)}(\cdot) \in [0, 1], \tau(\cdot) \in \mathcal{T}_M \right\}$$

then  $\mathbf{P} \left( |b_\gamma| \leq \hat{B}_{\gamma,M} \right) \xrightarrow{n \rightarrow \infty} 1$

## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

### Proposition: Upper Bound for the Potential Bias $\hat{B}_{\gamma,M}$

Under asymptotic normality and all necessary assumptions, for  $\tau(\cdot) \in \mathcal{T}_M$ , the upper bound of bias is

$$\hat{B}_{\gamma,M} = \sup \left\{ |\text{Bias}[\gamma_\pm, \tau_w; \alpha_0(\cdot), \tau(\cdot), G]| : G \in \mathcal{G}_n, \alpha_{(0)}(\cdot) \in [0, 1], \tau(\cdot) \in \mathcal{T}_M \right\}$$

then  $\mathbf{P} \left( |b_\gamma| \leq \hat{B}_{\gamma,M} \right) \xrightarrow{n \rightarrow \infty} 1$

$$\{G \in \mathcal{G}_n\} \subset \left\{ |b_\gamma| \leq \hat{B}_{\gamma,M} \right\} \Rightarrow \mathbf{P} \left( |b_\gamma| \leq \hat{B}_{\gamma,M} \right) \geq \mathbf{P} (G \in \mathcal{G}_n)$$



## Upper Bound for the Potential Bias $|b_\gamma| = |\theta_\gamma - \tau_w|$

### Proposition: Upper Bound for the Potential Bias $\hat{B}_{\gamma,M}$

Under asymptotic normality and all necessary assumptions, for  $\tau(\cdot) \in \mathcal{T}_M$ , the upper bound of bias is

$$\hat{B}_{\gamma,M} = \sup \left\{ |\text{Bias}[\gamma_\pm, \tau_w; \alpha_0(\cdot), \tau(\cdot), G]| : G \in \mathcal{G}_n, \alpha_{(0)}(\cdot) \in [0, 1], \tau(\cdot) \in \mathcal{T}_M \right\}$$

then  $\mathbf{P} \left( |b_\gamma| \leq \hat{B}_{\gamma,M} \right) \xrightarrow{n \rightarrow \infty} 1$

$$\{G \in \mathcal{G}_n\} \subset \{|b_\gamma| \leq \hat{B}_{\gamma,M}\} \Rightarrow \mathbf{P} \left( |b_\gamma| \leq \hat{B}_{\gamma,M} \right) \geq \mathbf{P} (G \in \mathcal{G}_n)$$

and for  $G \in \mathcal{G}_n$ , Dvoretzky-Kiefer-Wolfowitz (DKW) inequality gives:

$$\mathbf{P} (G \in \mathcal{G}_n) \geq \mathbf{P} \left[ \sup_{t \in \mathbb{R}} |F_G(t) - \hat{F}_n(t)| \leq \sqrt{\frac{\log(2/\alpha_n)}{2n}} \right] \geq 1 - \alpha_n \xrightarrow{n \rightarrow \infty} 1$$

## Recap: What We Have Derived

## Recap: What We Have Derived

- Asymptotic limit:

$$\hat{\tau}_{\gamma} \xrightarrow{p} \theta_{\gamma} = \frac{\mathbb{E} [\alpha_{(1)} (U) h (U, \gamma_+)]}{\mathbb{E} [h (U, \gamma_+)]} - \frac{\mathbb{E} [\alpha_{(0)} (U) h (U, \gamma_-)]}{\mathbb{E} [h (U, \gamma_-)]}$$

## Recap: What We Have Derived

- Asymptotic limit:

$$\hat{\tau}_{\gamma} \xrightarrow{p} \theta_{\gamma} = \frac{\mathbb{E} [\alpha_{(1)} (U) h (U, \gamma_{+})]}{\mathbb{E} [h (U, \gamma_{+})]} - \frac{\mathbb{E} [\alpha_{(0)} (U) h (U, \gamma_{-})]}{\mathbb{E} [h (U, \gamma_{-})]}$$

- Asymptotic normality:

$$\frac{\sqrt{n} (\hat{\tau}_{\gamma} - \theta_{\gamma})}{\sqrt{\hat{V}_{\gamma}}} \xrightarrow{d} \mathcal{N} (0, 1)$$

## Recap: What We Have Derived

- Asymptotic limit:

$$\hat{\tau}_{\gamma} \xrightarrow{p} \theta_{\gamma} = \frac{\mathbb{E} [\alpha_{(1)} (U) h (U, \gamma_+)]}{\mathbb{E} [h (U, \gamma_+)]} - \frac{\mathbb{E} [\alpha_{(0)} (U) h (U, \gamma_-)]}{\mathbb{E} [h (U, \gamma_-)]}$$

- Asymptotic normality:

$$\frac{\sqrt{n} (\hat{\tau}_{\gamma} - \theta_{\gamma})}{\sqrt{\hat{V}_{\gamma}}} \xrightarrow{d} \mathcal{N} (0, 1)$$

- Upper bound of the asymptotic bias:

$$\hat{B}_{\gamma, M} = \sup \left\{ |\text{Bias} [\gamma_{\pm}, \tau_w; \alpha_0 (\cdot), \tau (\cdot), G]| : G \in \mathcal{G}_n, \alpha_{(0)} (\cdot) \in [0, 1], \tau (\cdot) \in \mathcal{T}_M \right\}$$

# Bias-aware Confidence Intervals

## Corollary: Valid Confidence Intervals

Under asymptotic normality and all necessary assumptions, for  $\tau(\cdot) \in \mathcal{T}_M$ , consider the CIs

$$\hat{\tau}_\gamma \pm l_\alpha, \quad l_\alpha = \min \left\{ l : \mathbf{P} \left[ \left| b + n^{-\frac{1}{2}} \hat{V}_\gamma^{\frac{1}{2}} \tilde{Z} \right| \leq l \right] \geq 1 - \alpha, \forall |b| \leq \hat{B}_{\gamma, M} \right\}$$

where

- $\tilde{Z}$  is a standard Gaussian random variable
- $\alpha \in (0, 1)$  is the significant level
- $\hat{V}_\gamma$  is an estimate of the sampling variance  $V_\gamma$

then

$$\liminf_{n \rightarrow \infty} \mathbf{P} [\tau_w \in \hat{\tau}_\gamma \pm l_\alpha] \geq 1 - \alpha$$

# Bias-aware Confidence Intervals

CLT (with bias considered) is

$$\frac{\sqrt{n}(\hat{\tau}_\gamma - \theta_\gamma)}{\hat{V}_\gamma^{1/2}} = \frac{\sqrt{n}(\hat{\tau}_\gamma - \tau_w - b_\gamma)}{\hat{V}_\gamma^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1)$$

where  $b_\gamma = \theta_\gamma - \tau_w$ , then let  $\tilde{Z} \sim \mathcal{N}(0, 1)$  we have

$$\begin{aligned} \mathbf{P}[\tau_w \in \hat{\tau}_\gamma \pm l_\alpha] &= \mathbf{P}[-l_\alpha - b_\gamma \leq \hat{\tau}_\gamma - \tau_w - b_\gamma \leq l_\alpha - b_\gamma] \\ &= \mathbf{P}\left[-\sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha + b_\gamma) \leq \sqrt{n}\hat{V}_\gamma^{-1/2}(\hat{\tau}_\gamma - \tau_w - b_\gamma) \leq \sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha - b_\gamma)\right] \\ &= \mathbb{E}\left(\mathbf{P}\left[-\sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha + b_\gamma) \leq \tilde{Z} \leq \sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha - b_\gamma)\right] \mid \hat{V}_\gamma, \hat{B}_{\gamma, M}, \hat{\tau}_\gamma\right) + o(1) \\ &= \mathbb{E}\left[\mathbf{P}\left(-l_\alpha \leq n^{-1/2}\hat{V}_\gamma^{1/2}\tilde{Z} + b_\gamma \leq l_\alpha\right) \mid \hat{V}_\gamma, \hat{B}_{\gamma, M}, \hat{\tau}_\gamma\right] + o(1) \\ &= 1 - \alpha + o(1) \end{aligned}$$

## Bias-aware Confidence Intervals

CLT (with bias considered) is

$$\frac{\sqrt{n}(\hat{\tau}_\gamma - \theta_\gamma)}{\hat{V}_\gamma^{1/2}} = \frac{\sqrt{n}(\hat{\tau}_\gamma - \tau_w - b_\gamma)}{\hat{V}_\gamma^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1)$$

where  $b_\gamma = \theta_\gamma - \tau_w$ , then let  $\tilde{Z} \sim \mathcal{N}(0, 1)$  we have

$$\begin{aligned} \mathbf{P}[\tau_w \in \hat{\tau}_\gamma \pm l_\alpha] &= \mathbf{P}[-l_\alpha - b_\gamma \leq \hat{\tau}_\gamma - \tau_w - b_\gamma \leq l_\alpha - b_\gamma] \\ &= \mathbf{P}\left[-\sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha + b_\gamma) \leq \sqrt{n}\hat{V}_\gamma^{-1/2}(\hat{\tau}_\gamma - \tau_w - b_\gamma) \leq \sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha - b_\gamma)\right] \\ &= \mathbb{E}\left(\mathbf{P}\left[-\sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha + b_\gamma) \leq \tilde{Z} \leq \sqrt{n}\hat{V}_\gamma^{-1/2}(l_\alpha - b_\gamma)\right] \mid \hat{V}_\gamma, \hat{B}_{\gamma, M}, \hat{\tau}_\gamma\right) + o(1) \\ &= \mathbb{E}\left[\mathbf{P}\left(-l_\alpha \leq n^{-1/2}\hat{V}_\gamma^{1/2}\tilde{Z} + b_\gamma \leq l_\alpha\right) \mid \hat{V}_\gamma, \hat{B}_{\gamma, M}, \hat{\tau}_\gamma\right] + o(1) \\ &= 1 - \alpha + o(1) \end{aligned}$$



## Robustness to CATE Heterogeneity Misspecification

$$\hat{\tau}_{\gamma} \pm l_{\alpha}, l_{\alpha} = \min \left\{ l : \mathbf{P} \left[ \left| b + n^{-\frac{1}{2}} \hat{V}_{\gamma}^{\frac{1}{2}} \tilde{Z} \right| \leq l \right] \geq 1 - \alpha, \forall |b| \leq \hat{B}_{\gamma, M} \right\}$$

where

$$\mathcal{T}_M = \{ \tau(\cdot) \mid \tau(u) = \bar{\tau} + \Delta(u), \bar{\tau} \in \mathbb{R}, |\Delta(u)| \leq M \}, \quad M \in [0, 1]$$

Consider an extreme misspecification of CATE heterogeneity:  $M = 0$ , are the CIs robust?

## Robustness to CATE Heterogeneity Misspecification

$$\hat{\tau}_{\gamma} \pm l_{\alpha}, l_{\alpha} = \min \left\{ l : \mathbf{P} \left[ \left| b + n^{-\frac{1}{2}} \hat{V}_{\gamma}^{\frac{1}{2}} \tilde{Z} \right| \leq l \right] \geq 1 - \alpha, \forall |b| \leq \hat{B}_{\gamma, M} \right\}$$

where

$$\mathcal{T}_M = \{ \tau(\cdot) \mid \tau(u) = \bar{\tau} + \Delta(u), \bar{\tau} \in \mathbb{R}, |\Delta(u)| \leq M \}, \quad M \in [0, 1]$$

Consider an extreme misspecification of CATE heterogeneity:  $M = 0$ , are the CIs robust?

### Corollary: Robustness to CATE Heterogeneity Misspecification

The CIs under the misspecification of  $M = 0$  is still valid, but only for the convenience-weighted treatment effect:

$$\tau_{h,+} := \int \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} \tau(u) dG(u)$$

# Robustness to CATE Heterogeneity Misspecification

## Corollary: Robustness to CATE Heterogeneity Misspecification

The CIs under the misspecification of  $M = 0$  is still valid, but only for:

$$\tau_{h,+} := \int \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} \tau(u) dG(u)$$

$$\begin{aligned} a\text{Bias} = \theta_\gamma - \tau_w &= \int \left( \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} - \frac{h(u, \gamma_-)}{\mathbb{E}_G[h(U, \gamma_-)]} \right) \alpha_{(0)}(u) dG(u) \\ &\quad + \underbrace{\int \left( \frac{h(u, \gamma_+)}{\mathbb{E}_G[h(U, \gamma_+)]} - \frac{w(u)}{\mathbb{E}_G[w(U)]} \right) \tau(u) dG(u)}_{\text{CATE heterogeneity bias}} \end{aligned}$$

# References I

Thank you!