Inference on Union Bounds

Xinyue Bei
<xinyue.bei@austin.utexas.edu>
UT Austin

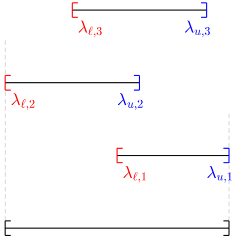
UCSD Econometrics Seminar
Sep 23, 2025

Motivation

In many empirical applications, the target object is in a **union bound**

$$\theta \in \left[\min_{b \in \mathcal{B}} \lambda_{\ell,b},\ \max_{b \in \mathcal{B}} \lambda_{u,b}\right]$$

- θ is the target object
- $(\lambda_\ell, \lambda_u) \in \mathbb{R}^{2|\mathcal{B}|}$ is unknown but estimable
- $\mathcal{B}$ is the set of indices: known and finite

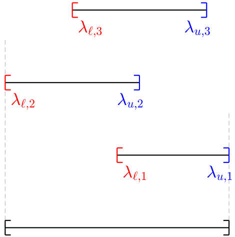


Motivation

In many empirical applications, the target object is in a **union bound**

$$\theta \in \left[\min_{b\in\mathcal{B}} \lambda_{\ell,b},\ \max_{b\in\mathcal{B}} \lambda_{u,b}\right]$$

- θ is the target object
- $(\lambda_\ell, \lambda_u) \in \mathbb{R}^{2|\mathcal{B}|}$ is unknown but estimable
- $\mathcal{B}$ is the set of indices: known and finite



The goal of this paper: Construct a confidence interval for θ

Example: Difference in Differences

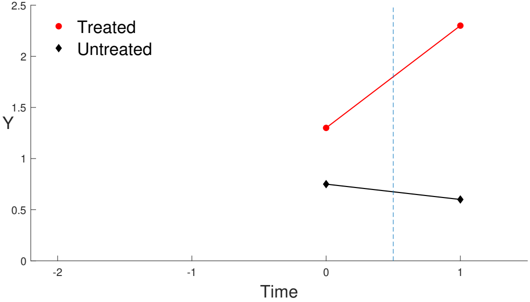
Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?

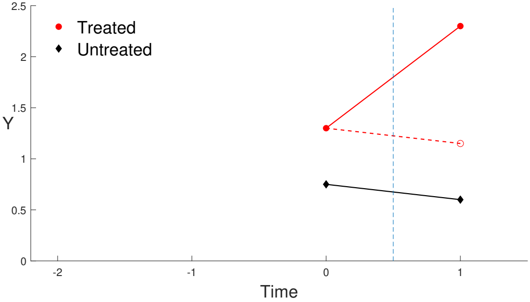


Units with $D = 1$ receive a treatment at $t = 1$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



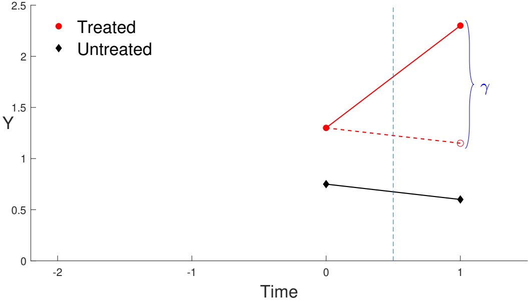
$Y_t(d)$: the potential outcome at t with treatment d
With parallel trends

$$0 = \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 1\right] - \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 0\right]$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



$Y_t(d)$: the potential outcome at t with treatment d
With parallel trends

$$0 = \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 1\right] - \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 0\right]$$

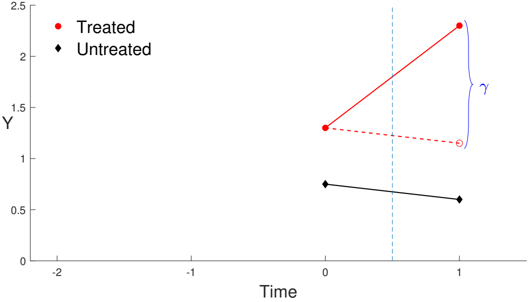
The DiD estimand γ is

$$\gamma = \underbrace{ATT}_{\theta}$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



$Y_t(d)$: the potential outcome at t with treatment d
Without parallel trends

$$\Delta_t = \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 1\right] - \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 0\right]$$

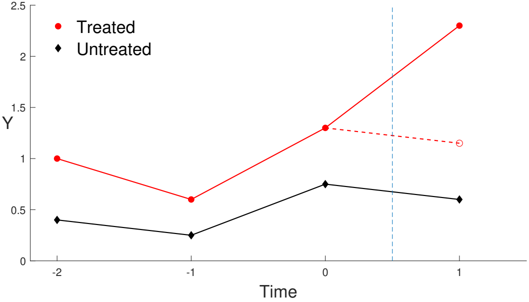
The DiD estimand γ is

$$\gamma = \underbrace{ATT}_{\theta} + \Delta_1$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



$Y_t(d)$: the potential outcome at t with treatment d
Without parallel trends

$$\Delta_t = \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 1\right] - \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 0\right]$$

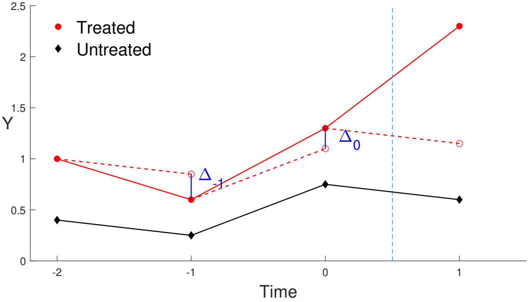
The DiD estimand γ is

$$\gamma = \underbrace{ATT}_{\theta} + \Delta_1$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



$Y_t(d)$: the potential outcome at t with treatment d
Without parallel trends

$$\Delta_t = \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 1\right] - \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid D = 0\right]$$

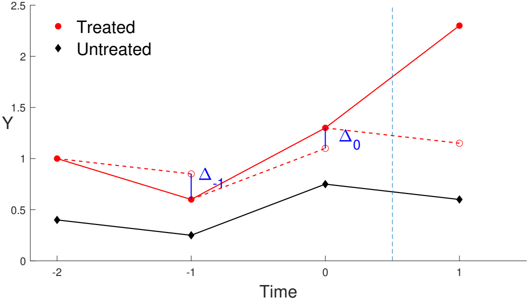
The DiD estimand γ is

$$\gamma = \underbrace{ATT}_{\theta} + \Delta_1$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



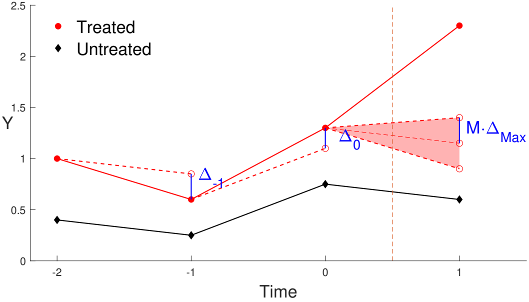
Relax the parallel trends assumption by

$$|\Delta_1| \leqslant M \cdot \max_{t=-T+1,\ldots,0} |\Delta_t|$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



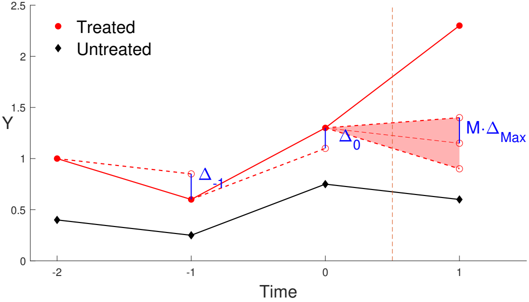
Relax the parallel trends assumption by

$$|\Delta_1| \leqslant M \cdot \max_{t=-T+1,\ldots,0} |\Delta_t|$$

Example: Difference in Differences

Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud)

What if the parallel trends assumption does not hold?



$$\theta = ATT \in \left[\min_{b \in \mathcal{B}} \lambda_{\ell,b},\ \max_{b \in \mathcal{B}} \lambda_{u,b} \right]$$

where $\mathcal{B} = \{-(T-1), ..., T-1, T\}$,

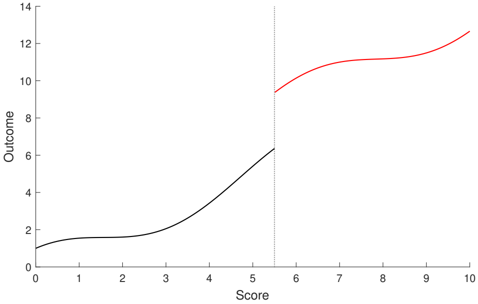
$$\lambda_{\ell,b} = \lambda_{u,b} = \begin{cases} \gamma - M\Delta_b & \text{if } b = -(T-1), ..., 0, \\ \gamma + M\Delta_{b-T} & \text{if } b = 1, ..., T \end{cases}$$

Example: Regression Discontinuity Design

Kolesár and Rothe (2018, AER)

Example: Regression Discontinuity Design

Kolesár and Rothe (2018, AER)



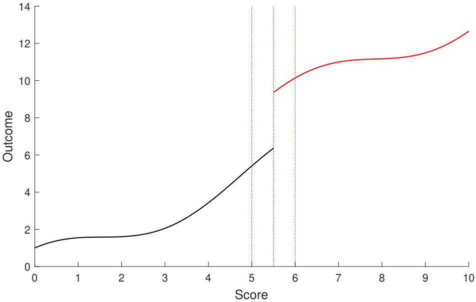
Treatment $D = 1\{X \geq k\}$ with running variable X

Let $\mu(X) = E[Y \mid X]$. The ATE at the threshold is

$$\theta = E\Big[Y(1) - Y(0) \mid X = 0\Big] = \lim_{x \downarrow k} \mu(x) - \lim_{x \uparrow k} \mu(x)$$

Example: Regression Discontinuity Design

Kolesár and Rothe (2018, AER)

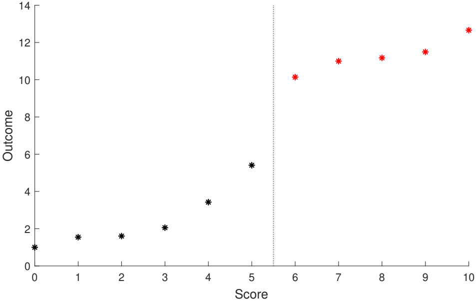


Estimate θ: local OLS of Y on $m(X)$ with $X \in [k-h, k+h]$ where

$$m(x) = (1\{x \geq k\}, 1\{x \geq k\}(x-k), \ldots, 1\{x \geq k\}(x-k)^p,\\ 1, x-k, \ldots, (x-k)^p)$$

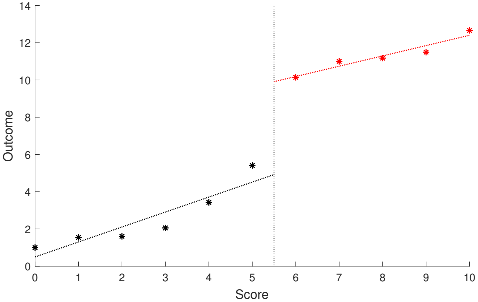
Example: Regression Discontinuity Design

Kolesár and Rothe (2018, AER)



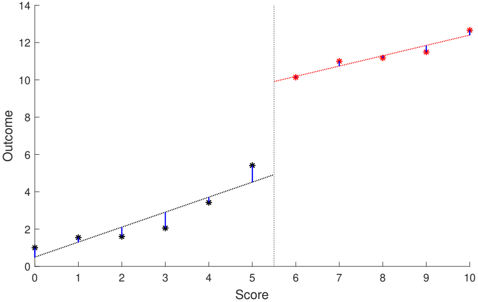
Example: Regression Discontinuity Design

Kolesár and Rothe (2018, AER)



Example: Regression Discontinuity Design

Kolesár and Rothe (2018, AER)



Assumption: bounds on specification errors at the threshold

$$|\lim_{x\uparrow k}\Delta(x)| \leq \max_{x'<k}|\Delta(x')|, \ |\lim_{x\downarrow k}\Delta(x)| \leq \max_{x'>k}|\Delta(x')|$$

Examples

- Difference-in-Differences: Manski and Pepper (2018, REStat), Rambachan and Roth (2023, REStud), Hasegawa, Small, Webster (2019, Epidemiology), Ye, Keele, Hasegawa, and Small (2023, JASA), Ban and Kédagni (2023, WP)
- Regression Discontinuity Design: Kolesár and Rothe (2018, AER)
- Misspecification Analysis: Masten and Poirier (2021, ECTA), Apfel and Windmeijer (2022, WP), Stoye (2022, WP)
- Bunching and Income Elasticity: Blomquist, Newey, Kumar, and Liang (2021, JPE)
- Sign Congruence: Brinch, Mogstad, Wiswall (2017, JPE), Kowalski (2022, RES), Kim (2024, WP), Molinari, Miller, Stoye (2024, WP)
- Mediation Effect: van Garderen and van Giersbergen (2024, REStat)
- Instrumental Variables: Machado, Shaikh, and Vytlacil (2019, JoE)

Main Contributions

I propose a novel CI based on modified conditional inference

- **Valid**: CI covers θ with prob. $\geq 1 - \alpha$ under mild regularity conditions
- **Short**: higher power than existing methods under a large set of DGPs

Main Contributions

I propose a novel CI based on modified conditional inference

- **Valid**: CI covers θ with prob. $\geq 1 - \alpha$ under mild regularity conditions
- **Short**: higher power than existing methods under a large set of DGPs

Main Contributions

I propose a novel CI based on modified conditional inference

- **Valid**: CI covers θ with prob. $\geq 1 - \alpha$ under mild regularity conditions
- **Short**: higher power than existing methods under a large set of DGPs

Compare with Existing Procedures

My method improves upon existing valid procedures

	My	Simple	Hybrid	Adj. Boot.
Adjust for union	✓	×	×	✓
$\sqrt{n}$ conv. rate to id set	✓	✓	✓	×

- Simple CI: Kolesár and Rothe (2018, AER), among others
- Hybrid CI: Rambachan and Roth (2023, RES)
- Adjusted Bootstrap: Ye, Keele, Hasegawa and Small (2023, JASA)

Contributions to Other Related Literature

Intersection Bounds & Moment Inequalities
Chernozhukov, Hong, and Tamer (2007), Romano and Shaikh (2008), Rosen (2008), D. Andrews and Guggenberger (2009), D. Andrews and Soares (2010), Chernozhukov, Lee, and Rosen (2013), D. Andrews and Shi (2013), Bugni, Canay and Shi (2015), among others

- This paper complements the intersection bounds ▸ Intersection

Contributions to Other Related Literature

Intersection Bounds & Moment Inequalities
Chernozhukov, Hong, and Tamer (2007), Romano and Shaikh (2008), Rosen (2008), D. Andrews and Guggenberger (2009), D. Andrews and Soares (2010), Chernozhukov, Lee, and Rosen (2013), D. Andrews and Shi (2013), Bugni, Canay and Shi (2015), among others

- This paper complements the intersection bounds ▸ Intersection

Directionally Differentiable Functions
Hirano and Porter (2012), Fang and Santos (2019), Fang (2018), Ponomarev (2022), among others

- This paper sheds light on inference of direct. diff. func. w/o convex null

Contributions to Other Related Literature

Intersection Bounds & Moment Inequalities

Chernozhukov, Hong, and Tamer (2007), Romano and Shaikh (2008), Rosen (2008), D. Andrews and Guggenberger (2009), D. Andrews and Soares (2010), Chernozhukov, Lee, and Rosen (2013), D. Andrews and Shi (2013), Bugni, Canay and Shi (2015), among others

- This paper complements the intersection bounds Intersection

Directionally Differentiable Functions

Hirano and Porter (2012), Fang and Santos (2019), Fang (2018), Ponomarev (2022), among others

- This paper sheds light on inference of direct. diff. func. w/o convex null

Conditional Inference

I. Andrews and Mikusheva (2016), I. Andrews, Roth and Pakes (2016), I. Andrews, Kitagawa, McCloskey (2021, 2023), Rambachan and Roth (2023), among others

- This paper widens the use of the conditional inference technique

Outline

Outline

Setting

The target object

$$\theta \in \left[\min_{b \in \mathcal{B}} \lambda_{\ell,b}, \ \max_{b \in \mathcal{B}} \lambda_{u,b} \right]$$

where λ_ℓ and λ_u are $|\mathcal{B}|$-dimensional vectors

Assume that $\widehat{\lambda}_\ell$, $\widehat{\lambda}_u$ are asymptotically normal

$$\sqrt{n} \begin{pmatrix} \widehat{\lambda}_\ell - \lambda_\ell \\ \widehat{\lambda}_u - \lambda_u \end{pmatrix} \xrightarrow{d} N(0, \Sigma)$$

Goal: construct a *uniformly valid* and *short* CI for θ

$$\liminf_n \inf_{P \in \mathcal{P}} \inf_{\theta \in [\lambda_{\ell,\min}, \lambda_{u,\max}]} P(\theta \in CI) \geq 1 - \alpha$$

A Simple Example

Consider the simplest possible case where

$$\theta \in [\min\{\lambda_1, \lambda_2\},\ \max\{\lambda_1, \lambda_2\}]$$

with

$$\begin{pmatrix} \widehat{\lambda}_1 \\ \widehat{\lambda}_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

A Simple Example

Construct the CI by inverting tests of the hypothesis

$$H_0 : \min\{\lambda_1, \lambda_2\} \leq \theta \leq \max\{\lambda_1, \lambda_2\}$$

A Simple Example

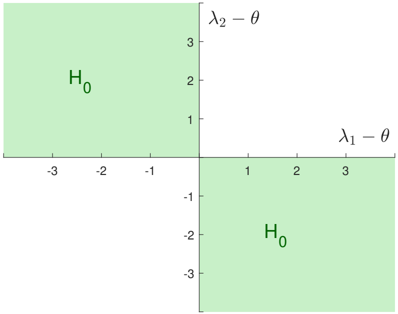
Construct the CI by inverting tests of the hypothesis

$$H_0 : \min\{\lambda_1 - \theta, \lambda_2 - \theta\} \leq 0 \leq \max\{\lambda_1 - \theta, \lambda_2 - \theta\}$$

A Simple Example

Construct the CI by inverting tests of the hypothesis

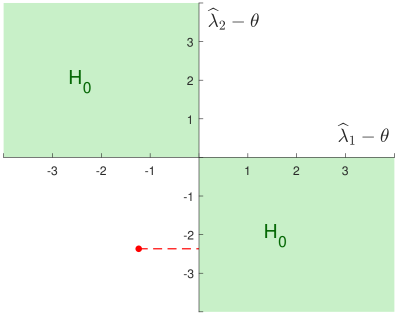
$$H_0 : \min\{\lambda_1 - \theta, \lambda_2 - \theta\} \leq 0 \leq \max\{\lambda_1 - \theta, \lambda_2 - \theta\}$$



A Simple Example

Construct the CI by inverting tests of the hypothesis

$$H_0 : \min\{\lambda_1 - \theta, \lambda_2 - \theta\} \leq 0 \leq \max\{\lambda_1 - \theta, \lambda_2 - \theta\}$$



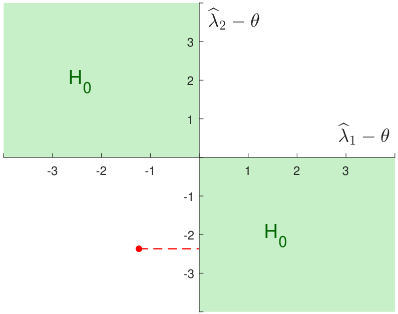
Test stat is

$$\widehat{T}(\theta) = \max\Big\{\min\{\widehat{\lambda}_1 - \theta, \widehat{\lambda}_2 - \theta\}, \min\{\theta - \widehat{\lambda}_1, \theta - \widehat{\lambda}_2\}\Big\}$$

A Simple Example

Construct the CI by inverting tests of the hypothesis

$$H_0 : \min\{\lambda_1 - \theta, \lambda_2 - \theta\} \leq 0 \leq \max\{\lambda_1 - \theta, \lambda_2 - \theta\}$$



Test stat is

$$\widehat{T}(\theta) = \max\Big\{\min\{\widehat{\lambda}_1 - \theta, \widehat{\lambda}_2 - \theta\}, \min\{\theta - \widehat{\lambda}_1, \theta - \widehat{\lambda}_2\}\Big\}$$

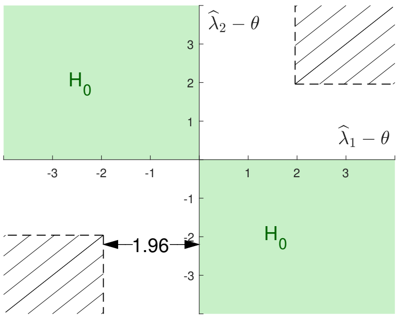
The CI is

$$\left[\min_{\widehat{T}(\theta)\leq c(\theta)} \theta, \max_{\widehat{T}(\theta)\leq c(\theta)} \theta\right]$$

A Simple Example

Construct the CI by inverting tests of the hypothesis

$$H_0 : \min\{\lambda_1 - \theta, \lambda_2 - \theta\} \leq 0 \leq \max\{\lambda_1 - \theta, \lambda_2 - \theta\}$$



Test stat is

$$\widehat{T}(\theta) = \max\Big\{\min\{\widehat{\lambda}_1 - \theta, \widehat{\lambda}_2 - \theta\}, \min\{\theta - \widehat{\lambda}_1, \theta - \widehat{\lambda}_2\}\Big\}$$

The CI is

$$\left[\min_{\widehat{T}(\theta) \leq c(\theta)} \theta, \max_{\widehat{T}(\theta) \leq c(\theta)} \theta\right]$$

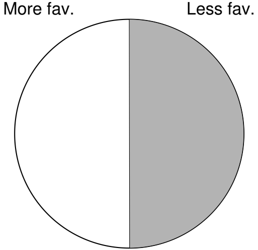
Current practice: $c^{\text{sim}} = \Phi^{-1}(1 - \frac{\alpha}{2})$, 1.96 for $\alpha = 0.05$

A Simple Example - Modified Conditional CV

To improve upon c^{sim}, I propose a modified conditional cv

A Simple Example - Modified Conditional CV

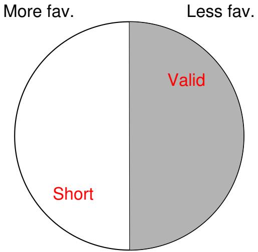
To improve upon c^{sim}, I propose a modified conditional cv



1. Define less and more favorable DGPs

A Simple Example - Modified Conditional CV

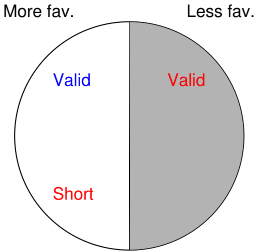
To improve upon c^{sim}, I propose a modified conditional cv



1. Define less and more favorable DGPs
2. Construct a conditional cv

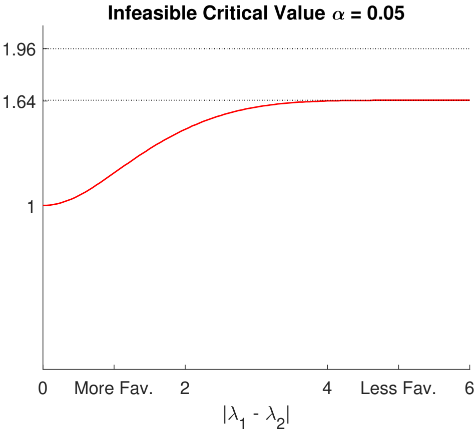
A Simple Example - Modified Conditional CV

To improve upon c^{sim}, I propose a modified conditional cv

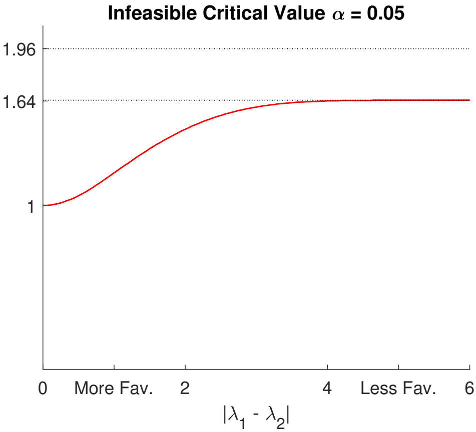


1. Define less and more favorable DGPs
2. Construct a conditional cv
3. Modify the conditional cv

A Simple Example - Step 1 More & Less fav. DGPs

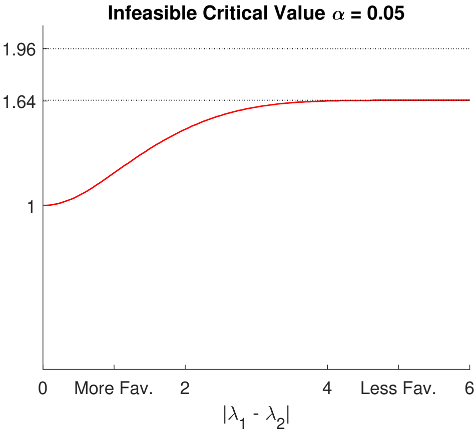


A Simple Example - Step 1 More & Less fav. DGPs



If $\lambda_1 = \lambda_2$, the infeasible CI is $\left[\min\{\widehat{\lambda}_1, \widehat{\lambda}_2\} - 1,\ \max\{\widehat{\lambda}_1, \widehat{\lambda}_2\} + 1\right]$

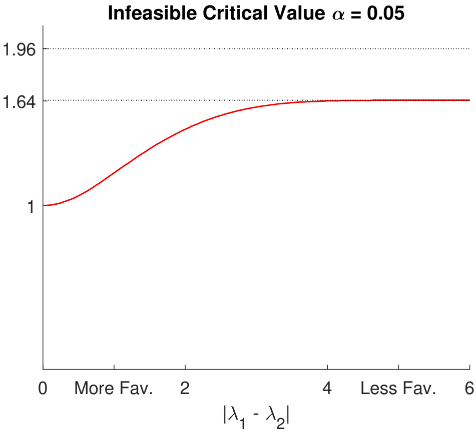
A Simple Example - Step 1 More & Less fav. DGPs



If $\lambda_1 = \lambda_2$, the infeasible CI is $\left[\min\{\widehat{\lambda}_1, \widehat{\lambda}_2\} - 1,\ \max\{\widehat{\lambda}_1, \widehat{\lambda}_2\} + 1\right]$

- we can use $\left[\widehat{\lambda}_1 - 1.96,\ \widehat{\lambda}_1 + 1.96\right]$ *or* $\left[\widehat{\lambda}_2 - 1.96,\ \widehat{\lambda}_2 + 1.96\right]$

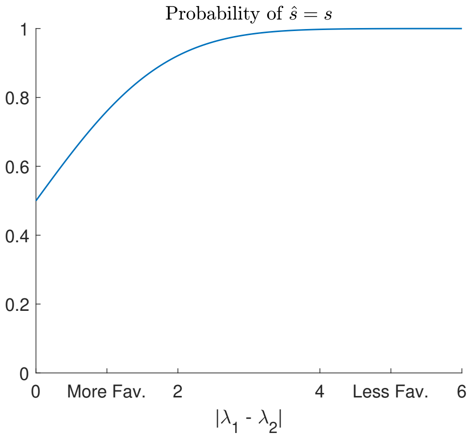
A Simple Example - Step 1 More & Less fav. DGPs



If $\lambda_1 = \lambda_2$, the infeasible CI is $\left[\min\{\widehat{\lambda}_1, \widehat{\lambda}_2\} - 1,\ \max\{\widehat{\lambda}_1, \widehat{\lambda}_2\} + 1\right]$

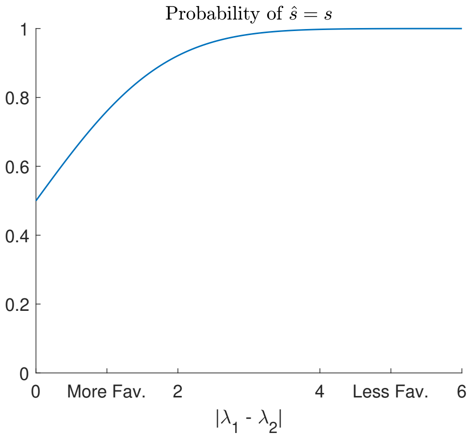
If $\lambda_1 \ll \lambda_2$, the infeasible CI is $\left[\min\{\widehat{\lambda}_1, \widehat{\lambda}_2\} - 1.64,\ \max\{\widehat{\lambda}_1, \widehat{\lambda}_2\} + 1.64\right]$

A Simple Example - Step 2 Conditional CV



where $s = \arg\min_{b=1,2} \lambda_b, \quad \widehat{s} = \arg\min_{b=1,2} \widehat{\lambda}_b$

A Simple Example - Step 2 Conditional CV



where $s = \arg\min_{b=1,2} \lambda_b$, $\quad \widehat{s} = \arg\min_{b=1,2} \widehat{\lambda}_b$

Construct conditional cv based on the distribution of $\widehat{T}(\theta)\Big|\,\widehat{s} = s$

A Simple Example - Step 2 Conditional CV

Construct conditional cv based on the distribution of $\widehat{T}(\theta)\Big|\,\widehat{s}=s$

A Simple Example - Step 2 Conditional CV

Construct conditional cv based on the distribution of $\widehat{T}(\theta)\Big|\widehat{s}=s$

Recall

$$\widehat{T}(\theta)=\max\left\{\min\left\{\theta-\widehat{\lambda}_1,\theta-\widehat{\lambda}_2\right\},\min\left\{\widehat{\lambda}_1-\theta,\widehat{\lambda}_2-\theta\right\}\right\}$$

A Simple Example - Step 2 Conditional CV

Construct conditional cv based on the distribution of $\widehat{T}(\theta)\Big|\widehat{s}=s$

Recall

$$\widehat{T}(\theta)=\max\left\{\min\left\{\theta-\widehat{\lambda}_1,\theta-\widehat{\lambda}_2\right\},\min\left\{\widehat{\lambda}_1-\theta,\widehat{\lambda}_2-\theta\right\}\right\}$$

Specifically, I consider

$$\widehat{T}(\theta)\,\Big|\,\widehat{T}(\theta)=\theta-\widehat{\lambda}_1,\widehat{s}=s$$

$$\widehat{T}(\theta)\,\Big|\,\widehat{T}(\theta)=\theta-\widehat{\lambda}_2,\widehat{s}=s$$

$$\widehat{T}(\theta)\,\Big|\,\widehat{T}(\theta)=\widehat{\lambda}_1-\theta,\widehat{s}=s$$

$$\widehat{T}(\theta)\,\Big|\,\widehat{T}(\theta)=\widehat{\lambda}_2-\theta,\widehat{s}=s$$

A Simple Example - Step 2 Conditional CV

Construct conditional cv based on the distribution of $\widehat{T}(\theta)\Big|\widehat{s}=s$

Recall

$$\widehat{T}(\theta)=\max\left\{\min\left\{\theta-\widehat{\lambda}_1,\theta-\widehat{\lambda}_2\right\},\min\left\{\widehat{\lambda}_1-\theta,\widehat{\lambda}_2-\theta\right\}\right\}$$

Specifically, I consider

$$\widehat{T}(\theta)\Big|\widehat{T}(\theta)=\theta-\widehat{\lambda}_1,\widehat{s}=s$$

$$\widehat{T}(\theta)\Big|\widehat{T}(\theta)=\theta-\widehat{\lambda}_2,\widehat{s}=s$$

$$\widehat{T}(\theta)\Big|\widehat{T}(\theta)=\widehat{\lambda}_1-\theta,\widehat{s}=s$$

$$\widehat{T}(\theta)\Big|\widehat{T}(\theta)=\widehat{\lambda}_2-\theta,\widehat{s}=s$$

A Simple Example - Step 2 Conditional CV

I illustrate with Con. CV

$$\widehat{T}(\theta) \,\Big|\, \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \,\Big|\, \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \quad \overset{\text{FOSD}}{\preceq} \quad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta)\Big|\widehat{T}(\theta)=\widehat{\lambda}_1-\theta,\widehat{s}=s \qquad \overset{\text{FOSD}}{\preceq} \qquad \mathcal{TN}\left(0,\left[\theta-\widehat{\lambda}_2,\widehat{\lambda}_2-\theta\right]\right)$$

Let $\alpha^{\text{con}} \in (\frac{\alpha}{2},\alpha)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1-\alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0,\left[\theta-\widehat{\lambda}_2,\widehat{\lambda}_2-\theta\right]\right)$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \,\Big|\, \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \qquad \overset{\text{FOSD}}{\preceq} \qquad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in (\frac{\alpha}{2}, \alpha)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

✓ Valid size under less favorable DGPs:

$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) = P\left(\widehat{T}(\theta) > \widehat{c}^{con}, \widehat{s} = s\right) + P\left(\widehat{T}(\theta) > \widehat{c}^{con}, \widehat{s} \neq s\right)$$

A Simple Example - Step 2 Conditional CV

I illustrate with ‣ Con. CV

$$\widehat{T}(\theta) \Big| \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \quad \overset{\text{FOSD}}{\preceq} \quad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in (\frac{\alpha}{2}, \alpha)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

- ✓ Valid size under less favorable DGPs:

$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) \leq P\left(\widehat{T}(\theta) > \widehat{c}^{con}, \widehat{s} = s\right) + P\left(\widehat{s} \neq s\right)$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \,\Big|\, \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \qquad \overset{\text{FOSD}}{\preceq} \qquad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in (\frac{\alpha}{2}, \alpha)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

✓ Valid size under less favorable DGPs:

$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) \leq P\left(\widehat{T}(\theta) > \widehat{c}^{con} \,|\widehat{s} = s\right) P\left(\widehat{s} = s\right) + P\left(\widehat{s} \neq s\right)$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \,\Big|\, \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \qquad \overset{\text{FOSD}}{\preceq} \qquad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in \left(\frac{\alpha}{2}, \alpha\right)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

✓ Valid size under less favorable DGPs:

$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) \leq \underbrace{P\left(\widehat{T}(\theta) > \widehat{c}^{con} \,|\widehat{s} = s\right)}_{\leq \alpha^{\text{con}} < \alpha} \underbrace{P\left(\widehat{s} = s\right)}_{\approx 1} + \underbrace{P\left(\widehat{s} \neq s\right)}_{\approx 0}$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \Big| \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \qquad \overset{\text{FOSD}}{\preceq} \qquad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in \left(\frac{\alpha}{2}, \alpha\right)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

✓ Valid size under less favorable DGPs:

$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) \leq \underbrace{P\left(\widehat{T}(\theta) > \widehat{c}^{con} \,|\widehat{s} = s\right)}_{\leq \alpha^{\mathbf{con}} < \alpha} \underbrace{P\left(\widehat{s} = s\right)}_{\approx 1} + \underbrace{P\left(\widehat{s} \neq s\right)}_{\approx 0} \leq \alpha$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \Big| \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \qquad \overset{\text{FOSD}}{\preceq} \qquad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in (\frac{\alpha}{2}, \alpha)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

✓ Valid size under less favorable DGPs:

$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) \leq P\left(\widehat{T}(\theta) > \widehat{c}^{con} \,|\widehat{s} = s\right) P\left(\widehat{s} = s\right) + P\left(\widehat{s} \neq s\right) \leq \alpha$$

✓ Smaller than c^{sim}

$$\widehat{c}^{con} = \Phi^{-1}\left(\alpha^{con} + (1 - 2\alpha^{con})\Phi\left(\widehat{\lambda}_2 - \theta\right)\right)$$

A Simple Example - Step 2 Conditional CV

I illustrate with ▸ Con. CV

$$\widehat{T}(\theta) \,\Big|\, \widehat{T}(\theta) = \widehat{\lambda}_1 - \theta, \widehat{s} = s \quad \overset{\text{FOSD}}{\preceq} \quad \mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$$

Let $\alpha^{\text{con}} \in \left(\frac{\alpha}{2}, \alpha\right)$ be the conditional rejection rate, recommend $\alpha^{\text{con}} = \frac{4}{5}\alpha$

Define $\widehat{c}^{\text{con}}$ as the $1 - \alpha^{\text{con}}$ quantile of $\mathcal{TN}\left(0, \left[\theta - \widehat{\lambda}_2, \widehat{\lambda}_2 - \theta\right]\right)$

- ✓ Valid size under less favorable DGPs:

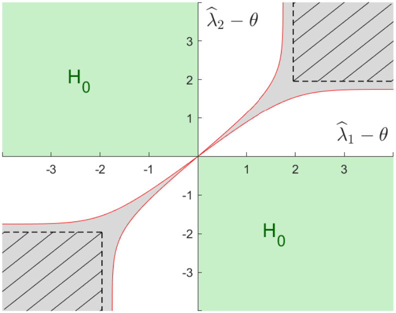
$$P\left(\widehat{T}(\theta) > \widehat{c}^{con}\right) \le P\left(\widehat{T}(\theta) > \widehat{c}^{con} \,|\widehat{s} = s\right) P\left(\widehat{s} = s\right) + P\left(\widehat{s} \neq s\right) \le \alpha$$

- ✓ Smaller than c^{sim}

$$\begin{aligned} \widehat{c}^{con} &= \Phi^{-1}\left(\alpha^{con} + (1 - 2\alpha^{con})\Phi\left(\widehat{\lambda}_2 - \theta\right)\right) \\ &\le \Phi^{-1}\left(\alpha^{con} + (1 - 2\alpha^{con})\right) \\ &= \Phi^{-1}\left(1 - \alpha^{con}\right) < \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = c^{sim} \end{aligned}$$

A Simple Example - Step 2 Conditional CV

The rejection region of $\hat{c}^{\text{con}}$ for $H_0 : \min\{\lambda_1, \lambda_2\} \leq \theta \leq \max\{\lambda_1, \lambda_2\}$



A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & \textit{if } \widehat{c}^{con}(\theta) \geq c^t \\ c^t & \textit{if } \widehat{c}^{con}(\theta) < c^t \end{cases}$$

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & \textit{if } \widehat{c}^{con}(\theta) \geq c^t \\ c^t & \textit{if } \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\mathrm{t}} = c$

$$P(\theta \notin CI; \lambda)$$

with $\lambda = (\lambda_1, \lambda_2)$

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & \text{if } \widehat{c}^{con}(\theta) \geq c^t \\ c^t & \text{if } \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\mathrm{t}} = c$

$$P(\theta \notin CI; \lambda) \leq \max\left\{P\left([\theta_\ell, \theta_{\text{mid}}] \not\subseteq CI; \lambda\right), P\left([\theta_{\text{mid}}, \theta_u] \not\subseteq CI; \lambda\right)\right\}$$

with $\lambda = (\lambda_1, \lambda_2)$, $\theta_\ell = \min\{\lambda_1, \lambda_2\}$, $\theta_u = \max\{\lambda_1, \lambda_2\}$, $\theta_{\text{mid}} = (\theta_\ell + \theta_u)/2$

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & if\ \widehat{c}^{con}(\theta) \geq c^t \\ c^t & if\ \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\mathrm{t}} = c$

$$\begin{aligned} P(\theta \notin CI; \lambda) \leq & \max \{ P\left([\theta_\ell, \theta_{\text{mid}}] \not\subseteq CI; \lambda \right), P\left([\theta_{\text{mid}}, \theta_u] \not\subseteq CI; \lambda \right) \} \\ \leq & \max \{ P\left(T(\theta_\ell) > c^{\mathrm{m}}(\theta_\ell, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda \right) \\ & P\left(T(\theta_u) > c^{\mathrm{m}}(\theta_u, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda \right) \} \end{aligned}$$

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & \textit{if } \widehat{c}^{con}(\theta) \geq c^t \\ c^t & \textit{if } \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\mathrm{t}} = c$

$$\begin{aligned} P(\theta \notin CI; \lambda) &\leq \max\left\{P\left([\theta_\ell, \theta_{\text{mid}}] \not\subseteq CI; \lambda\right), P\left([\theta_{\text{mid}}, \theta_u] \not\subseteq CI; \lambda\right)\right\} \\ &\leq \max\left\{P\left(T(\theta_\ell) > c^{\mathrm{m}}(\theta_\ell, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda\right)\right. \\ &\qquad\quad \left. P\left(T(\theta_u) > c^{\mathrm{m}}(\theta_u, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda\right)\right\} \\ &=: \bar{p}(c, \lambda) \end{aligned}$$

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & \textit{if } \widehat{c}^{con}(\theta) \geq c^t \\ c^t & \textit{if } \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\text{t}} = c$

$$\begin{aligned} P(\theta \notin CI; \lambda) \leq& \max\left\{P\left([\theta_\ell, \theta_{\text{mid}}] \not\subseteq CI; \lambda\right), P\left([\theta_{\text{mid}}, \theta_u] \not\subseteq CI; \lambda\right)\right\} \\ \leq& \max\left\{P\left(T(\theta_\ell) > c^{\text{m}}(\theta_\ell, c) \text{ or } T(\theta_{\text{mid}}) > c^{\text{m}}(\theta_{\text{mid}}, c); \lambda\right)\right. \\ & \left. P\left(T(\theta_u) > c^{\text{m}}(\theta_u, c) \text{ or } T(\theta_{\text{mid}}) > c^{\text{m}}(\theta_{\text{mid}}, c); \lambda\right)\right\} \\ =:& \bar{p}(c, \lambda) \end{aligned}$$

- $\bar{p}(c, \lambda)$ is easier to calculate

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & if\ \widehat{c}^{con}(\theta) \geq c^t \\ c^t & if\ \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\mathrm{t}} = c$

$$\begin{aligned} P(\theta \notin CI; \lambda) \leq& \max\left\{P\left([\theta_\ell, \theta_{\text{mid}}] \not\subseteq CI; \lambda\right), P\left([\theta_{\text{mid}}, \theta_u] \not\subseteq CI; \lambda\right)\right\} \\ \leq& \max\left\{P\left(T(\theta_\ell) > c^{\mathrm{m}}(\theta_\ell, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda\right)\right. \\ & \quad \left. P\left(T(\theta_u) > c^{\mathrm{m}}(\theta_u, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda\right)\right\} \\ =:& \bar{p}(c, \lambda) \end{aligned}$$

- $\bar{p}(c, \lambda)$ is easier to calculate
- $\bar{p}(c, \lambda)$ is not overly conservative

A Simple Example - Step 3 Modification

I introduce a novel modification

$$c^m(\theta, c^t) = \begin{cases} \widehat{c}^{con}(\theta) & \text{if } \widehat{c}^{con}(\theta) \geq c^t \\ c^t & \text{if } \widehat{c}^{con}(\theta) < c^t \end{cases}$$

To determine the value of c^{t}, fix $c^{\mathrm{t}} = c$

$$\begin{aligned} P(\theta \notin CI; \lambda) \leq & \max \left\{ P\left([\theta_\ell, \theta_{\text{mid}}] \not\subseteq CI; \lambda \right), P\left([\theta_{\text{mid}}, \theta_u] \not\subseteq CI; \lambda \right) \right\} \\ \leq & \max \left\{ P\left(T(\theta_\ell) > c^{\mathrm{m}}(\theta_\ell, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda \right) \right. \\ & \left. P\left(T(\theta_u) > c^{\mathrm{m}}(\theta_u, c) \text{ or } T(\theta_{\text{mid}}) > c^{\mathrm{m}}(\theta_{\text{mid}}, c); \lambda \right) \right\} \\ =: & \bar{p}(c, \lambda) \end{aligned}$$

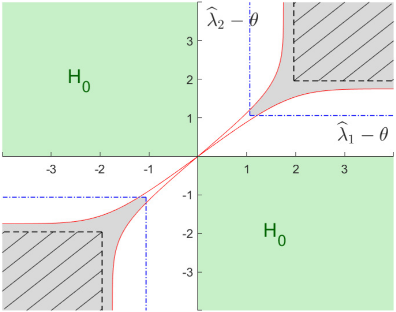
Thus it suffices to have

$$c^t = \inf \left\{ c \geq 0 : \sup_{\lambda \in \Lambda} \bar{p}(c, \lambda) \leq \alpha \right\}$$

This lower truncation guarantees uniform coverage

A Simple Example - Step 3 Modification

The rejection region of $\hat{c}^{\mathrm{m}}$ for $H_0 : \min\{\lambda_1, \lambda_2\} \le \theta \le \max\{\lambda_1, \lambda_2\}$



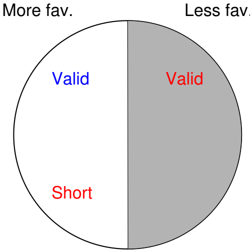
Larger power: the new CI has a larger rejection region

Valid: the lower truncation removes counter-intuitive rejection region

A Simple Example - Summary

The main idea of the modified conditional CI

1. Define less and more favorable DGPs
2. Construct a conditional cv
   - valid under less favorable DGPs
3. Modify the conditional cv
   - valid under more favorable DGPs



General Cases - Test Statistic

Construct CI by inverting

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta \leq \max_{b \in \mathcal{B}} \lambda_{u,b}$$

General Cases - Test Statistic

Construct CI by inverting

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta \leq \max_{b \in \mathcal{B}} \lambda_{u,b}$$

There are normal estimators $\widehat{\lambda}_\ell$, $\widehat{\lambda}_u$

$$\begin{pmatrix} \widehat{\lambda}_\ell - \lambda_\ell \\ \widehat{\lambda}_u - \lambda_u \end{pmatrix} \sim N(0, \Sigma)$$

General Cases - Test Statistic

Construct CI by inverting

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta \leq \max_{b \in \mathcal{B}} \lambda_{u,b}$$

There are normal estimators $\widehat{\lambda}_\ell$, $\widehat{\lambda}_u$

$$\begin{pmatrix} \widehat{\lambda}_\ell - \lambda_\ell \\ \widehat{\lambda}_u - \lambda_u \end{pmatrix} \sim N(0, \Sigma)$$

The test statistic

$$\widehat{T}(\theta) = \max \left\{ \min_{b \in \mathcal{B}} \frac{\widehat{\lambda}_{\ell,b} - \theta}{\sigma_{\ell,b}}, \quad \min_{b \in \mathcal{B}} \frac{\theta - \widehat{\lambda}_{u,b}}{\sigma_{u,b}} \right\}$$

where $\sigma_{\ell,b} = \sqrt{\text{var}(\widehat{\lambda}_{\ell,b})}$, $\sigma_{u,b} = \sqrt{\text{var}(\widehat{\lambda}_{u,b})}$

General Cases - Test Statistic

Construct CI by inverting

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta \leq \max_{b \in \mathcal{B}} \lambda_{u,b}$$

There are normal estimators $\widehat{\lambda}_\ell$, $\widehat{\lambda}_u$

$$\begin{pmatrix} \widehat{\lambda}_\ell - \lambda_\ell \\ \widehat{\lambda}_u - \lambda_u \end{pmatrix} \sim N(0, \Sigma)$$

The test statistic

$$\widehat{T}(\theta) = \max\left\{ \min_{b \in \mathcal{B}} \frac{\widehat{\lambda}_{\ell,b} - \theta}{\sigma_{\ell,b}}, \quad \min_{b \in \mathcal{B}} \frac{\theta - \widehat{\lambda}_{u,b}}{\sigma_{u,b}} \right\}$$

where $\sigma_{\ell,b} = \sqrt{\mathrm{var}(\widehat{\lambda}_{\ell,b})}$, $\sigma_{u,b} = \sqrt{\mathrm{var}(\widehat{\lambda}_{u,b})}$

The simple critical value $c^{\mathrm{sim}} = \Phi^{-1}(1 - \frac{\alpha}{2})$ gives a simple CI

$$CI^{\mathrm{sim}} = \left[\min_{b \in \mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1 - \frac{\alpha}{2}), \ \max_{b \in \mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1 - \frac{\alpha}{2}) \right]$$

General Cases - Test Statistic

Construct CI by inverting

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta \leq \max_{b \in \mathcal{B}} \lambda_{u,b}$$

There are normal estimators $\widehat{\lambda}_\ell$, $\widehat{\lambda}_u$

$$\begin{pmatrix} \widehat{\lambda}_\ell - \lambda_\ell \\ \widehat{\lambda}_u - \lambda_u \end{pmatrix} \sim N(0, \Sigma)$$

The test statistic

$$\widehat{T}(\theta) = \max \left\{ \min_{b \in \mathcal{B}} \frac{\widehat{\lambda}_{\ell,b} - \theta}{\sigma_{\ell,b}}, \quad \min_{b \in \mathcal{B}} \frac{\theta - \widehat{\lambda}_{u,b}}{\sigma_{u,b}} \right\}$$

where $\sigma_{\ell,b} = \sqrt{\text{var}(\widehat{\lambda}_{\ell,b})}$, $\sigma_{u,b} = \sqrt{\text{var}(\widehat{\lambda}_{u,b})}$

The simple critical value $c^{\text{sim}} = \Phi^{-1}(1 - \frac{\alpha}{2})$ gives a simple CI

$$CI^{\text{sim}} = \left[\min_{b \in \mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b} \Phi^{-1}(1 - \frac{\alpha}{2}), \ \max_{b \in \mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b} \Phi^{-1}(1 - \frac{\alpha}{2}) \right]$$

c^{sim} is valid but can be very conservative

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$P\left(\theta \notin CI^{\text{sim}}\right) = P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right)$$

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$
\begin{aligned}
P\left(\theta \notin CI^{\text{sim}}\right) =& P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta \notin \left[\hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2}),\ \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right)
\end{aligned}
$$

where

$$
b_\ell = \arg\min_{b\in\mathcal{B}} \lambda_{\ell,b}, \quad b_u = \arg\max_{b\in\mathcal{B}} \lambda_{u,b}
$$

- $\leq$ is conservative if $\lambda_{\ell,b_\ell} \approx \min_{b\neq b_\ell} \lambda_{\ell,b}$ or $\lambda_{u,b_u} \approx \min_{b\neq b_u} \lambda_{u,b}$

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$
\begin{aligned}
P\left(\theta \notin CI^{\text{sim}}\right) =& P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta \notin \left[\hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2}),\ \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta < \hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2})\right) + P\left(\theta > \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right)
\end{aligned}
$$

- $\leq$ is conservative if $\lambda_{\ell,b_\ell} \approx \min_{b\neq b_\ell} \lambda_{\ell,b}$ or $\lambda_{u,b_u} \approx \min_{b\neq b_u} \lambda_{u,b}$
- $\leq$ follows from $P(A\cup B) \leq P(A) + P(B)$

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$
\begin{aligned}
P\left(\theta \notin CI^{\text{sim}}\right) =& P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta \notin \left[\hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2}),\ \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta < \hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2})\right) + P\left(\theta > \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right) \\
=& P\left(\frac{\hat{\lambda}_{\ell,b_\ell} - \lambda_{\ell,b_\ell}}{\sigma_{\ell,b}} + \frac{\lambda_{\ell,b_\ell} - \theta}{\sigma_{\ell,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
&+ P\left(\frac{\lambda_{u,b_u} - \hat{\lambda}_{u,b_u}}{\sigma_{u,b}} + \frac{\theta - \lambda_{u,b_u}}{\sigma_{u,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right)
\end{aligned}
$$

- $\leq$ is conservative if $\lambda_{\ell,b_\ell} \approx \min_{b\neq b_\ell} \lambda_{\ell,b}$ or $\lambda_{u,b_u} \approx \min_{b\neq b_u} \lambda_{u,b}$
- $\leq$ follows from $P(A\cup B) \leq P(A) + P(B)$

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$
\begin{aligned}
P\left(\theta \notin CI^{\text{sim}}\right) =& P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta \notin \left[\hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2}),\ \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta < \hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2})\right) + P\left(\theta > \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right) \\
=& P\left(\frac{\hat{\lambda}_{\ell,b_\ell} - \lambda_{\ell,b_\ell}}{\sigma_{\ell,b}} + \frac{\lambda_{\ell,b_\ell} - \theta}{\sigma_{\ell,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
&+ P\left(\frac{\lambda_{u,b_u} - \hat{\lambda}_{u,b_u}}{\sigma_{u,b}} + \frac{\theta - \lambda_{u,b_u}}{\sigma_{u,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
\leq& \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha
\end{aligned}
$$

- $\leq$ is conservative if $\lambda_{\ell,b_\ell} \approx \min_{b\neq b_\ell} \lambda_{\ell,b}$ or $\lambda_{u,b_u} \approx \min_{b\neq b_u} \lambda_{u,b}$
- $\leq$ follows from $P(A\cup B) \leq P(A) + P(B)$
- $\leq$ is conservative if $\lambda_{u,b_u} - \lambda_{\ell,b_\ell} \gg 0$,
  see Imbens and Manski(2004), Stoye(2009)

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$
\begin{aligned}
P\left(\theta \notin CI^{\text{sim}}\right) =& P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta \notin \left[\hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2}),\ \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta < \hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2})\right) + P\left(\theta > \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right) \\
=& P\left(\frac{\hat{\lambda}_{\ell,b_\ell} - \lambda_{\ell,b_\ell}}{\sigma_{\ell,b}} + \frac{\lambda_{\ell,b_\ell} - \theta}{\sigma_{\ell,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
&+ P\left(\frac{\lambda_{u,b_u} - \hat{\lambda}_{u,b_u}}{\sigma_{u,b}} + \frac{\theta - \lambda_{u,b_u}}{\sigma_{u,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
\leq& \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha
\end{aligned}
$$

- $\leq$ is NOT conservative if $\lambda_{\ell,b_\ell} \ll \min_{b\neq b_\ell} \lambda_{\ell,b}$ and $\lambda_{u,b_u} \gg \min_{b\neq b_u} \lambda_{u,b}$
- $\leq$ follows from $P(A\cup B) \leq P(A) + P(B)$
- $\leq$ is NOT conservative if $\lambda_{u,b_u} = \lambda_{\ell,b_\ell}$
  see Imbens and Manski(2004), Stoye(2009)

General Cases - A Simple Critical Value

c^{sim} is valid but can be very conservative

$$
\begin{aligned}
P\left(\theta \notin CI^{\text{sim}}\right) =& P\left(\theta \notin \left[\min_{b\in\mathcal{B}} \hat{\lambda}_{\ell,b} - \sigma_{\ell,b}\Phi^{-1}(1-\frac{\alpha}{2}),\ \max_{b\in\mathcal{B}} \hat{\lambda}_{u,b} + \sigma_{u,b}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta \notin \left[\hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2}),\ \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right]\right) \\
\leq& P\left(\theta < \hat{\lambda}_{\ell,b_\ell} - \sigma_{\ell,b_\ell}\Phi^{-1}(1-\frac{\alpha}{2})\right) + P\left(\theta > \hat{\lambda}_{u,b_u} + \sigma_{u,b_u}\Phi^{-1}(1-\frac{\alpha}{2})\right) \\
=& P\left(\frac{\hat{\lambda}_{\ell,b_\ell} - \lambda_{\ell,b_\ell}}{\sigma_{\ell,b}} + \frac{\lambda_{\ell,b_\ell} - \theta}{\sigma_{\ell,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
& + P\left(\frac{\lambda_{u,b_u} - \hat{\lambda}_{u,b_u}}{\sigma_{u,b}} + \frac{\theta - \lambda_{u,b_u}}{\sigma_{u,b}} > \Phi^{-1}(1-\frac{\alpha}{2})\right) \\
\leq& \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha
\end{aligned}
$$

- c^{sim} is nearly optimal among *constant* critical values
- It is crucial to use a *data-dependent* critical value

General Cases - Conditional Critical Value

The test statistic

$$\widehat{T}(\theta) = \max\left\{\min_{b\in\mathcal{B}} \frac{\sqrt{n}(\widehat{\lambda}_{\ell,b}-\theta)}{\widehat{\sigma}_{\ell,b}}, \quad \min_{b\in\mathcal{B}} \frac{\sqrt{n}(\theta-\widehat{\lambda}_{u,b})}{\widehat{\sigma}_{u,b}}\right\}$$

where $\widehat{\sigma}_{\ell,b} = \sqrt{\text{var}(\widehat{\lambda}_{\ell,b})}$, $\widehat{\sigma}_{u,b} = \sqrt{\text{var}(\widehat{\lambda}_{u,b})}$

The conditional critical value

$$\widehat{c}^{con} = \begin{cases} \Phi^{-1}\left(\alpha^{con}\Phi\left(t_{\ell,1}(\theta,\widehat{b}_\ell)\right) + (1-\alpha^{con})\Phi\left(t_{\ell,2}(\theta,\widehat{b}_\ell)\right)\right) & \textit{if } \widehat{T}(\theta) = \frac{\widehat{\lambda}_{\ell,\widehat{b}_\ell}-\theta}{\widehat{\sigma}_{\ell,\widehat{b}_\ell}/\sqrt{n}} \\ \Phi^{-1}\left(\alpha^{con}\Phi\left(t_{u,1}(\theta,\widehat{b}_u)\right) + (1-\alpha^{con})\Phi\left(t_{u,2}(\theta,\widehat{b}_u)\right)\right) & \textit{if } \widehat{T}(\theta) = \frac{\theta-\widehat{\lambda}_{u,\widehat{b}_u}}{\widehat{\sigma}_{u,\widehat{b}_u}/\sqrt{n}} \end{cases}$$

where

$$\widehat{b}_\ell = \arg\min_{b\in\mathcal{B}} \frac{\widehat{\lambda}_{\ell,b}-\theta}{\widehat{\sigma}_{\ell,b}/\sqrt{n}}, \;\; \widehat{b}_u = \arg\min_{b\in\mathcal{B}} \frac{\theta-\widehat{\lambda}_{u,b}}{\widehat{\sigma}_{u,b}/\sqrt{n}}$$

$$t_{\ell,1}(\theta,b) = \min_{\tilde{b}\in\mathcal{B}} \left(1+\widehat{\rho}_{\ell u}(b,\tilde{b})\right)^{-1}\left(\frac{\theta-\widehat{\lambda}_{u,b}}{\widehat{\sigma}_{u,b}/\sqrt{n}} + \widehat{\rho}_{\ell u}(b,\tilde{b})\frac{\widehat{\lambda}_{\ell,b}-\theta}{\widehat{\sigma}_{\ell,b}/\sqrt{n}}\right)$$

$$t_{\ell,2}(\theta,b) = \min_{\tilde{b}\in\mathcal{B}} \left(1-\widehat{\rho}_{\ell}(b,\tilde{b})\right)^{-1}\left(\frac{\widehat{\lambda}_{\ell,b}-\theta}{\widehat{\sigma}_{\ell,b}/\sqrt{n}} - \widehat{\rho}_{\ell}(b,\tilde{b})\frac{\widehat{\lambda}_{\ell,b}-\theta}{\widehat{\sigma}_{\ell,b}/\sqrt{n}}\right)$$

General Cases - Conditional Critical Value

The test statistic

$$\widehat{T}(\theta) = \max\left\{ \min_{b\in\mathcal{B}} \frac{\sqrt{n}(\widehat{\lambda}_{\ell,b} - \theta)}{\widehat{\sigma}_{\ell,b}}, \quad \min_{b\in\mathcal{B}} \frac{\sqrt{n}(\theta - \widehat{\lambda}_{u,b})}{\widehat{\sigma}_{u,b}} \right\}$$

where $\widehat{\sigma}_{\ell,b} = \sqrt{\text{var}(\widehat{\lambda}_{\ell,b})}$, $\widehat{\sigma}_{u,b} = \sqrt{\text{var}(\widehat{\lambda}_{u,b})}$

The conditional critical value

$$\widehat{c}^{con} = \begin{cases} \Phi^{-1}\left(\alpha^{con}\Phi\left(t_{\ell,1}(\theta, \widehat{b}_\ell)\right) + (1-\alpha^{con})\Phi\left(t_{\ell,2}(\theta, \widehat{b}_\ell)\right)\right) & \textit{if } \widehat{T}(\theta) = \frac{\widehat{\lambda}_{\ell,\widehat{b}_\ell} - \theta}{\widehat{\sigma}_{\ell,\widehat{b}_\ell}/\sqrt{n}} \\ \Phi^{-1}\left(\alpha^{con}\Phi\left(t_{u,1}(\theta, \widehat{b}_u)\right) + (1-\alpha^{con})\Phi\left(t_{u,2}(\theta, \widehat{b}_u)\right)\right) & \textit{if } \widehat{T}(\theta) = \frac{\theta - \widehat{\lambda}_{u,\widehat{b}_u}}{\widehat{\sigma}_{u,\widehat{b}_u}/\sqrt{n}} \end{cases}$$

- If $\widehat{T}(\theta) = \frac{\widehat{\lambda}_{\ell,\widehat{b}_\ell} - \theta}{\widehat{\sigma}_{\ell,\widehat{b}_\ell}/\sqrt{n}}$, conditional on $\arg\min_{b\in\mathcal{B}} \frac{\widehat{\lambda}_{\ell,b} - \theta}{\widehat{\sigma}_{\ell,b}/\sqrt{n}} = \arg\min_{b\in\mathcal{B}} \frac{\lambda_{\ell,b} - \theta}{\sigma_{\ell,b}/\sqrt{n}}$

General Cases - Conditional Critical Value

The test statistic

$$\widehat{T}(\theta) = \max\left\{\min_{b\in\mathcal{B}} \frac{\sqrt{n}(\widehat{\lambda}_{\ell,b}-\theta)}{\widehat{\sigma}_{\ell,b}}, \quad \min_{b\in\mathcal{B}} \frac{\sqrt{n}(\theta-\widehat{\lambda}_{u,b})}{\widehat{\sigma}_{u,b}}\right\}$$

where $\widehat{\sigma}_{\ell,b} = \sqrt{\text{var}(\widehat{\lambda}_{\ell,b})}$, $\widehat{\sigma}_{u,b} = \sqrt{\text{var}(\widehat{\lambda}_{u,b})}$

The conditional critical value

$$\widehat{c}^{con} = \begin{cases} \Phi^{-1}\left(\alpha^{con}\Phi\left(t_{\ell,1}(\theta,\widehat{b}_\ell)\right) + (1-\alpha^{con})\Phi\left(t_{\ell,2}(\theta,\widehat{b}_\ell)\right)\right) & \text{if } \widehat{T}(\theta) = \frac{\widehat{\lambda}_{\ell,\widehat{b}_\ell}-\theta}{\widehat{\sigma}_{\ell,\widehat{b}_\ell}/\sqrt{n}} \\ \Phi^{-1}\left(\alpha^{con}\Phi\left(t_{u,1}(\theta,\widehat{b}_u)\right) + (1-\alpha^{con})\Phi\left(t_{u,2}(\theta,\widehat{b}_u)\right)\right) & \text{if } \widehat{T}(\theta) = \frac{\theta-\widehat{\lambda}_{u,\widehat{b}_u}}{\widehat{\sigma}_{u,\widehat{b}_u}/\sqrt{n}} \end{cases}$$

- If $\widehat{T}(\theta) = \frac{\widehat{\lambda}_{\ell,\widehat{b}_\ell}-\theta}{\widehat{\sigma}_{\ell,\widehat{b}_\ell}/\sqrt{n}}$, conditional on $\arg\min_{b\in\mathcal{B}} \frac{\widehat{\lambda}_{\ell,b}-\theta}{\widehat{\sigma}_{\ell,b}/\sqrt{n}} = \arg\min_{b\in\mathcal{B}} \frac{\lambda_{\ell,b}-\theta}{\sigma_{\ell,b}/\sqrt{n}}$
- If identified set large, $\max_{b\in\mathcal{B}}\lambda_{u,b} - \min_{b\in\mathcal{B}}\lambda_{\ell,b} \gg \frac{1}{\sqrt{n}}$,

$$\widehat{c}^{\text{con}} \leq \Phi^{-1}\left(1-\alpha^{\text{con}}\right) < c^{\text{sim}}$$

- In addition, if the bounds are not well separated, $\downarrow \widehat{c}^{\text{con}}$

General Cases - Modification

The modified conditional critical value

$$c^{\mathrm{m}}(\theta, c^{\mathrm{t}}) = \begin{cases} \widehat{c}^{\mathrm{con}}(\theta) & \text{if } \widehat{c}^{\mathrm{con}}(\theta) \geq c^{\mathrm{t}} \\ c^{\mathrm{t}} & \text{if } \widehat{c}^{\mathrm{con}}(\theta) < c^{\mathrm{t}} \end{cases}$$

The lower truncation c^{t} is

$$c^{\mathrm{t}} = \inf_{c} \left\{ c \geq 0 : \sup_{\lambda \in \Lambda} \bar{p}(c, \lambda) \leq \alpha \right\}$$

where $\theta_\ell = \min_{b \in \mathcal{B}} \lambda_{\ell,b}$, $\theta_u = \max_{b \in \mathcal{B}} \lambda_{u,b}$, $\theta_m = (\theta_\ell + \theta_u)/2$

$$\bar{p}(c, \lambda) = \max \left\{ P\left(\widehat{T}(\theta_\ell) > c^{\mathrm{m}}(\theta_\ell, c) \text{ or } \widehat{T}(\theta_m) > c^{\mathrm{m}}(\theta_m, c); (\lambda, \Sigma) \right), \right.$$
$$\left. P\left(\widehat{T}(\theta_m) > c^{\mathrm{m}}(\theta_m, c) \text{ or } \widehat{T}(\theta_u) > c^{\mathrm{m}}(\theta_u, c); (\lambda, \Sigma) \right) \right\}$$

General Cases - Modification

The modified conditional critical value

$$c^{\mathsf{m}}(\theta, c^{\mathsf{t}}) = \begin{cases} \widehat{c}^{\mathsf{con}}(\theta) & \text{if } \widehat{c}^{\mathsf{con}}(\theta) \geq c^{\mathsf{t}} \\ c^{\mathsf{t}} & \text{if } \widehat{c}^{\mathsf{con}}(\theta) < c^{\mathsf{t}} \end{cases}$$

The lower truncation c^{t} is

$$c^{\mathsf{t}} = \inf_{c} \left\{ c \geq 0 : \sup_{\lambda \in \Lambda_\eta} \bar{p}(c, \lambda) \leq \alpha - \eta \right\}$$

where $\theta_\ell = \min_{b \in \mathcal{B}} \lambda_{\ell,b}$, $\theta_u = \max_{b \in \mathcal{B}} \lambda_{u,b}$, $\theta_m = (\theta_\ell + \theta_u)/2$

$$\bar{p}(c, \lambda) = \max \Big\{ P\left(\widehat{T}(\theta_\ell) > c^{\mathsf{m}}(\theta_\ell, c) \text{ or } \widehat{T}(\theta_m) > c^{\mathsf{m}}(\theta_m, c); (\lambda, \Sigma) \right), \\ P\left(\widehat{T}(\theta_m) > c^{\mathsf{m}}(\theta_m, c) \text{ or } \widehat{T}(\theta_u) > c^{\mathsf{m}}(\theta_u, c); (\lambda, \Sigma) \right) \Big\}$$

Size & Power: Assumptions

Assumption (Known Singularity (KS))

There are known $|\mathcal{B}| \times J$ matrices A_ℓ, A_u such that for some $(\delta_P, \widehat{\delta}_n)$

$$\lambda_\ell = A_\ell \delta_P,\ \lambda_u = A_u \delta_P,\ \widehat{\lambda}_\ell = A_\ell \widehat{\delta}_n,\ \widehat{\lambda}_u = A_u \widehat{\delta}_n$$

Size & Power: Assumptions

Assumption (Known Singularity (KS))

There are known $|\mathcal{B}| \times J$ *matrices* A_ℓ, A_u *such that for some* $(\delta_P, \widehat{\delta}_n)$

$$\lambda_\ell = A_\ell \delta_P,\ \lambda_u = A_u \delta_P,\ \widehat{\lambda}_\ell = A_\ell \widehat{\delta}_n,\ \widehat{\lambda}_u = A_u \widehat{\delta}_n$$

Assumption (Asymptotic Normality (AN))

Let $\xi_P \sim \mathcal{N}(0, \Omega_P)$. *Assume*

$$\lim_{n\to\infty} \sup_{P\in\mathcal{P}} \sup_{f\in BL_1} \left| E_P\left[f\left(\sqrt{n}\left(\widehat{\delta}_n - \delta_P\right)\right)\right] - E\left[f\left(\xi_P\right)\right]\right| = 0$$

Assumption (Full Rank (FR))

For all $P \in \mathcal{P}$, $0 < \underline{e} \leq eig(\Omega_P) \leq \bar{e} < \infty$

Size & Power: Assumptions

Assumption (Known Singularity (KS))

There are known $|\mathcal{B}| \times J$ *matrices* A_ℓ, A_u *such that for some* $(\delta_P, \widehat{\delta}_n)$

$$\lambda_\ell = A_\ell \delta_P,\ \lambda_u = A_u \delta_P,\ \widehat{\lambda}_\ell = A_\ell \widehat{\delta}_n,\ \widehat{\lambda}_u = A_u \widehat{\delta}_n$$

Assumption (Asymptotic Normality (AN))

Let $\xi_P \sim \mathcal{N}(0, \Omega_P)$. *Assume*

$$\lim_{n\to\infty} \sup_{P\in\mathcal{P}} \sup_{f\in BL_1} \left| E_P \left[f\left(\sqrt{n} \left(\widehat{\delta}_n - \delta_P \right) \right) \right] - E\left[f(\xi_P) \right] \right| = 0$$

Assumption (Full Rank (FR))

For all $P \in \mathcal{P}$, $0 < \underline{e} \leq eig(\Omega_P) \leq \bar{e} < \infty$

Assumption (Consistent Covariance Estimator (CE))

For all $\varepsilon > 0$, $\lim_{n\to\infty} \sup_{P\in\mathcal{P}} P\left(\left\| \widehat{\Omega}_n - \Omega_P \right\| > \varepsilon \right) = 0$

Size & Power: Asymptotic Size Properties

Theorem (Uniform Coverage)

Suppose Assumptions KS, AN, FR, CE hold, for any $\alpha \in (0, 0.5)$,

$$\liminf_{n\to\infty} \inf_{P\in\mathcal{P}} \inf_{\theta\in\left[\lambda_{\ell,b_\ell},\lambda_{u,b_u}\right]} P\left(\theta \in CI^m\right) \geq 1-\alpha$$

The modified conditional CI has proper asymptotic coverage

Size & Power: Comparison with Simple CI

Theorem (Symmetric Or Large Bounds)

Suppose Assumptions KS, AN, FR, CE hold, $\alpha \in (0, \frac{1}{2})$

(Symmetric Bounds) *If* $corr(\widehat{\lambda}_{\ell,b_1}, \widehat{\lambda}_{\ell,b_2}) < \rho_1^*(\alpha, \alpha^c)$, $corr(\widehat{\lambda}_{\ell,b_\ell}, \widehat{\lambda}_{\ell,b_u}) < \rho_2^*(\alpha)$

$$\widehat{\lambda}_\ell = \widehat{\lambda}_u$$

Then

1. *My CI is Strictly Shorter:*
   There is $\alpha' > \alpha$ *such that*
   $$\liminf_n P\left(CI^m\left(\hat{\lambda}_n, \hat{\Sigma}_n/n; \alpha\right) \subseteq CI^{sim}\left(\hat{\lambda}_n, \hat{\Sigma}_n/n; \alpha'\right)\right) = 1$$
2. *My CI has Higher Power:*
   There is $\kappa \in (0, +\infty)$ *such that*
   $$\liminf_n P\left(\theta_n \notin CI^m\left(\hat{\lambda}, \hat{\Sigma}/n; \alpha\right)\right) - P\left(\theta_n \notin CI^{sim}\left(\hat{\lambda}, \hat{\Sigma}/n; \alpha\right)\right) > 0$$
   for some $\theta_n = \theta_\ell - \frac{\kappa}{\sqrt{n}}$.

▸ Uniform

Size & Power: Comparison with Simple CI

Theorem (Symmetric Or Large Bounds)

Suppose Assumptions KS, AN, FR, CE hold, $\alpha \in (0, \frac{1}{2})$

(Large Bounds)

$$\max_{b \in \mathcal{B}} \lambda_{u,b} - \min_{b \in \mathcal{B}} \lambda_{\ell,b} \geq \varepsilon > 0$$

Then

1. *My CI is Strictly Shorter:*
   There is $\alpha' > \alpha$ *such that*
   $$\liminf_n P\left(CI^m\left(\hat{\lambda}_n, \hat{\Sigma}_n/n; \alpha\right) \subseteq CI^{sim}\left(\hat{\lambda}_n, \hat{\Sigma}_n/n; \alpha'\right)\right) = 1$$
2. *My CI has Higher Power:*
   There is $\kappa \in (0, +\infty)$ *such that*
   $$\liminf_n P\left(\theta_n \notin CI^m\left(\hat{\lambda}, \hat{\Sigma}/n; \alpha\right)\right) - P\left(\theta_n \notin CI^{sim}\left(\hat{\lambda}, \hat{\Sigma}/n; \alpha\right)\right) > 0$$
   for some $\theta_n = \theta_\ell - \frac{\kappa}{\sqrt{n}}$.

Uniform

Connection to Moment Inequalities

Moment Inequalities

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Connection to Moment Inequalities

Moment Inequalities

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Connection to Moment Inequalities

Moment Inequalities

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated

Solution: $\sqrt{n}(\lambda_{\ell,b} - \theta) \leq 0 \Rightarrow$ get an upper bound $\sqrt{n}(\lambda_{\ell,b} - \theta)/\sqrt{\ln n}$

Connection to Moment Inequalities

Moment Inequalities

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated

Solution: $\sqrt{n}(\lambda_{\ell,b} - \theta) \leq 0 \Rightarrow$ get an upper bound $\sqrt{n}(\lambda_{\ell,b} - \theta)/\sqrt{\ln n}$

Union bounds

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \min_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \min_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Connection to Moment Inequalities

Moment Inequalities

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated

Solution: $\sqrt{n}(\lambda_{\ell,b} - \theta) \leq 0 \Rightarrow$ get an upper bound $\sqrt{n}(\lambda_{\ell,b} - \theta)/\sqrt{\ln n}$

Union bounds

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \min_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \min_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated & unknown sign

Connection to Moment Inequalities

We can write union bound problem as specification test in moment ineq

- Consider a one-sided union bound problem

$$\min\{\lambda_1, \lambda_2\} \leq 0$$

- This is equivalent to $\exists x \in [0, 1]$ such that

$$x\lambda_1 + (1 - x)\lambda_2 \leq 0$$

Connection to Moment Inequalities

We can write union bound problem as specification test in moment ineq

- Consider a one-sided union bound problem

$$\min \{\lambda_1, \lambda_2\} \leq 0$$

- This is equivalent to $\exists x \in [0, 1]$ such that

$$x\lambda_1 + (1 - x)\lambda_2 \leq 0$$

However, when both λ_1 and λ_2 are close to zero, the existing procedures in moment ineq do not perform well

Connection to Moment Inequalities

We can write union bound problem as specification test in moment ineq

- Consider a one-sided union bound problem

$$\min\{\lambda_1, \lambda_2\} \leq 0$$

- This is equivalent to $\exists x \in [0, 1]$ such that

$$x\lambda_1 + (1 - x)\lambda_2 \leq 0$$

However, when both λ_1 and λ_2 are close to zero, the existing procedures in moment ineq do not perform well

- Bugni, Canay, and Shi (2015, 2017): their minorant condition fails, no uniform validity

Connection to Moment Inequalities

We can write union bound problem as specification test in moment ineq

- Consider a one-sided union bound problem

$$\min\{\lambda_1, \lambda_2\} \le 0$$

- This is equivalent to $\exists x \in [0,1]$ such that

$$x\lambda_1 + (1-x)\lambda_2 \le 0$$

However, when both λ_1 and λ_2 are close to zero, the existing procedures in moment ineq do not perform well

- Bugni, Canay, and Shi (2015, 2017): their minorant condition fails, no uniform validity

- Kaido, Molinari, and Stoye (2019): a first-order approximation, the critical value reduces to the simple critical value

Connection to Moment Inequalities

We can write union bound problem as specification test in moment ineq

- Consider a one-sided union bound problem

$$\min\{\lambda_1, \lambda_2\} \leq 0$$

- This is equivalent to $\exists x \in [0, 1]$ such that

$$x\lambda_1 + (1 - x)\lambda_2 \leq 0$$

However, when both λ_1 and λ_2 are close to zero, the existing procedures in moment ineq do not perform well

- Bugni, Canay, and Shi (2015, 2017): their minorant condition fails, no uniform validity

- Kaido, Molinari, and Stoye (2019): a first-order approximation, the critical value reduces to the simple critical value

Connection to Moment Inequalities

We can write union bound problem as specification test in moment ineq

- Consider a one-sided union bound problem

$$\min\{\lambda_1, \lambda_2\} \leq 0$$

- This is equivalent to $\exists x \in [0, 1]$ such that

$$x\lambda_1 + (1 - x)\lambda_2 \leq 0$$

However, when both λ_1 and λ_2 are close to zero, the existing procedures in moment ineq do not perform well

- Bugni, Canay, and Shi (2015, 2017): their minorant condition fails, no uniform validity

- Kaido, Molinari, and Stoye (2019): a first-order approximation, the critical value reduces to the simple critical value

This suggests that my procedure can potentially be applied to improve existing specification tests when the constraint qualifications fail.

Outline

Setting

Relaxation of the parallel trend assumptions, where

$$\theta = ATT \in \left[\min_{b \in \mathcal{B}} \lambda_{\ell,b},\ \max_{b \in \mathcal{B}} \lambda_{u,b}\right]$$

where $\mathcal{B} = \{-(T-1), ..., T\}$,

$$\lambda_\ell = \lambda_u = \begin{cases} \gamma + \Delta_b & \text{if } b = -(T-1), ..., 0 \\ \gamma - \Delta_{-(b-1)} & \text{if } b = 1, ..., T \end{cases}$$

Setting

I conduct inference based on $\left(\widehat{\Delta}, \widehat{\gamma}, \Omega\right)$ where $(\widehat{\Delta}, \widehat{\gamma})$ is simulated from

$$\begin{pmatrix} \widehat{\Delta} \\ \widehat{\gamma} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \Delta \\ \gamma \end{pmatrix}, \Omega\right)$$

I use four Ω calibrated from

1. Benzarti and Carloni (2019): consumption tax cuts
2. Dustmann, Lindner, Schönberg, Umkehrer, Vom Berge (2022): minimum wage
3. Lovenheim and Willén (2019): teacher collective bargaining
4. Christensen, Keiser, Lade (2023): environmental crises

I normalized $\gamma = 0$ and use three Δ

1. Parallel trends assumption holds, i.e. $\Delta = 0_T$
2. Small pre-trends: Δ is calibrated
3. One large pre-trend: $\Delta = (10\omega_M, 0_{T-1})$

In sum, I use $4 \times 3 = 12$ empirically motivated DGPs ▸ Details

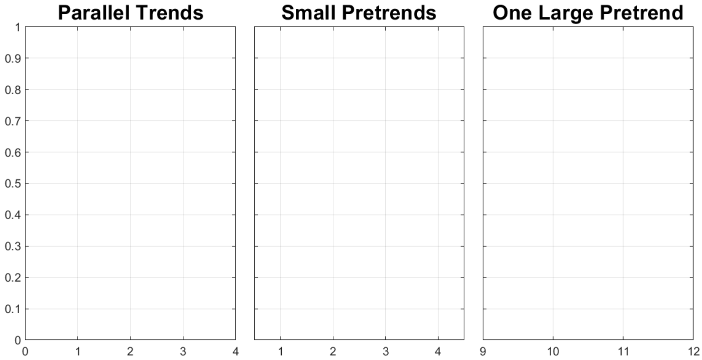
Alternative Methods

I compare my CI with

1. Simple CI in, e.g., Kolesár and Rothe (2018, AER)
2. Hybrid CI in Rambachan and Roth (2023, RES)
3. Adjusted bootstrap in Ye, Keele, Hasegawa and Small (2023, JASA)

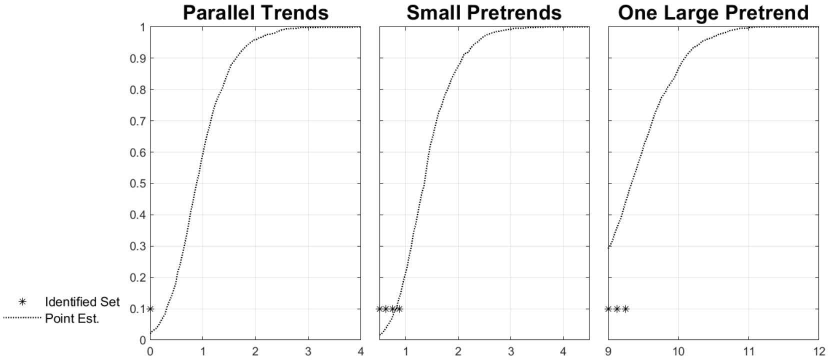
CDF of Upper Bounds of CIs

DGP: Ω from Lovenheim and Willén (2019)



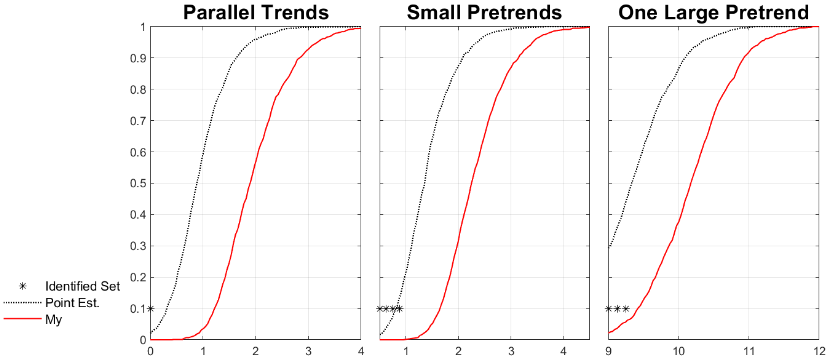
CDF of Upper Bounds of CIs

DGP: Ω from Lovenheim and Willén (2019)



CDF of Upper Bounds of CIs

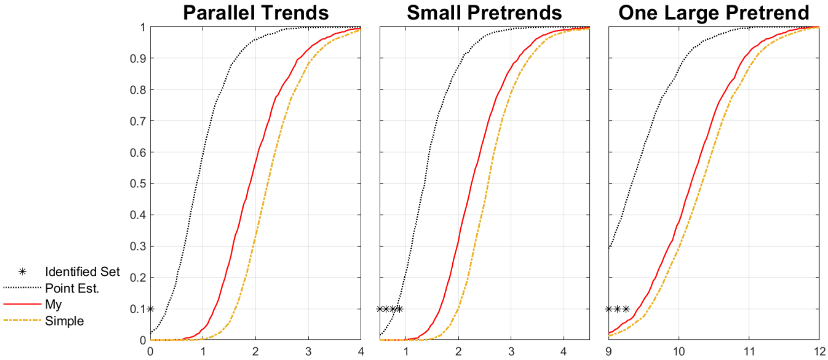
DGP: Ω from Lovenheim and Willén (2019)



Modified conditional CI has proper coverage

CDF of Upper Bounds of CIs

DGP: Ω from Lovenheim and Willén (2019)

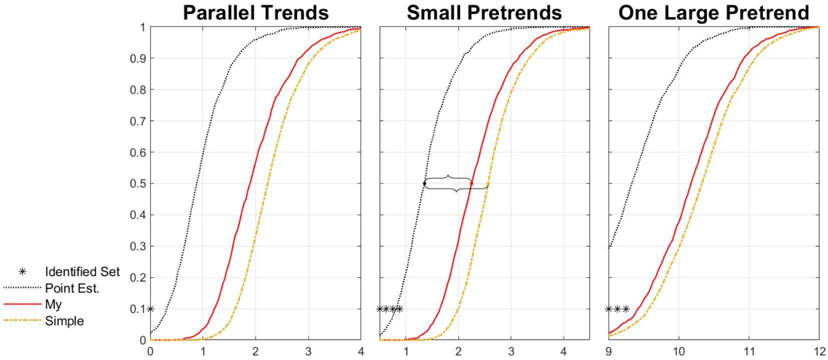


Modified conditional CI outperforms simple CI in all DGPs

- Reduces median simple CI (net of point est.) by **31%** under small violation

CDF of Upper Bounds of CIs

DGP: Ω from Lovenheim and Willén (2019)

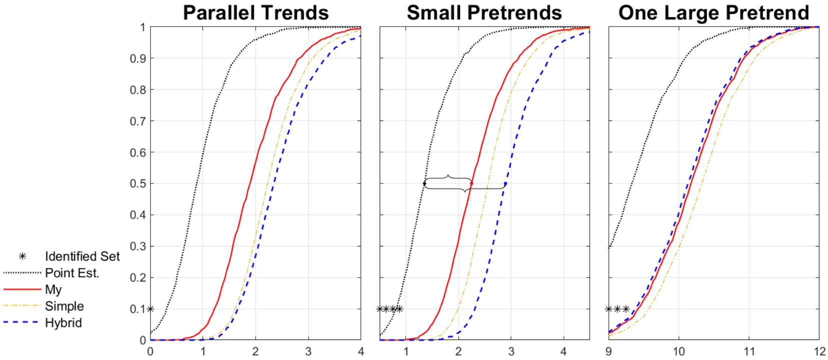


Modified conditional CI outperforms simple CI in all DGPs

- Reduces median simple CI (net of point est.) by **31%** under small violation

CDF of Upper Bounds of CIs

DGP: Ω from Lovenheim and Willén (2019)

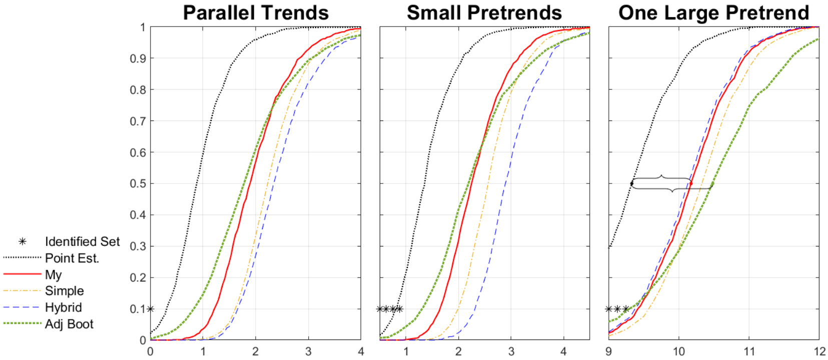


Modified conditional CI outperforms Hybrid CI under no or small violations

- Reduces median Hybrid CI (net of point est.) by **43%** under small violation
- Hybrid CI is efficient with large violation, but my CI is close

CDF of Upper Bounds of CIs

DGP: Ω from Lovenheim and Willén (2019)

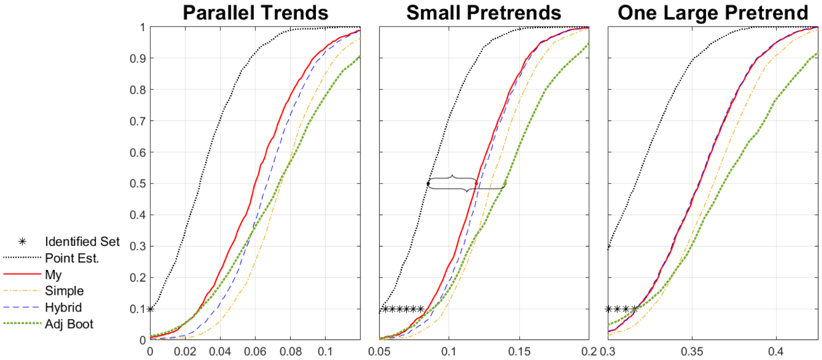


Modified conditional CI outperforms Adj Boot CI for relatively large alternatives

- Reduces median Adj Boot CI (net of point est.) by **27%** under large violation
- Power plot of Adj Boot will be flatter with larger *n*

CDF of Upper Bounds of CIs

DGP: Ω from Benzarti and Carloni (2019)



Modified conditional CI outperforms Adj Boot CI for relatively large alternatives

- Reduces median Adj Boot CI (net of point est.) by **38%** under small violation
- Power plot of Adj Boot CI will be flatter with larger n

Outline

The Effects of the Minimum Wage: Diff in Diff

RR23 in Dustmann, Lindner, Schönberg, Umkehrer, Vom Berge (2022, QJE)

What are the effects of the minimum wage?

- Addresses wage inequality
- Potential disemployment

The Effects of the Minimum Wage: Diff in Diff

RR23 in Dustmann, Lindner, Schönberg, Umkehrer, Vom Berge (2022, QJE)

What are the effects of the minimum wage?

- Addresses wage inequality $\Leftarrow$ Significant Wage Effect
- Potential disemployment $\Leftarrow$ Insignificant Employment Effect

The Effects of the Minimum Wage: Diff in Diff

To study the employment and wage effect, run

$$\log(\text{emp}_{rt}) = \sum_{\tau=2011,\tau\neq 2014}^{2016} \gamma_\tau^{\text{e}} \overline{GAP}_r 1\left[\tau = t\right] + \alpha_r^{\text{e}} + \xi_t^{\text{e}} + \varepsilon_{rt}^{\text{e}}$$

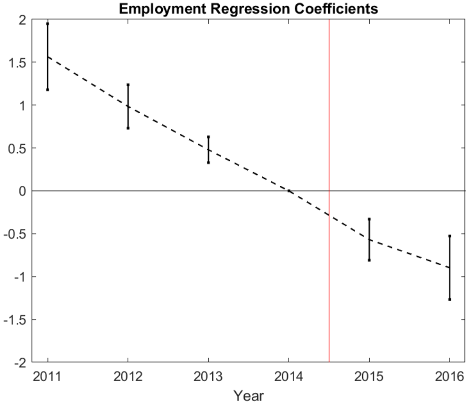
$$\log(\text{wage}_{rt}) = \sum_{\tau=2011,\tau\neq 2014}^{2016} \gamma_\tau^{\text{w}} \overline{GAP}_r 1\left[\tau = t\right] + \alpha_r^{\text{w}} + \xi_t^{\text{w}} + \varepsilon_{rt}^{\text{w}}$$

- $\log(\text{emp}_{rt})$ is the log employment in district r time t; $\log(\text{wage}_{rt})$ is log wage
- $\overline{GAP}_r$ is a measure of the exposure to the minimum wage
- Pre-policy year 2011-2014

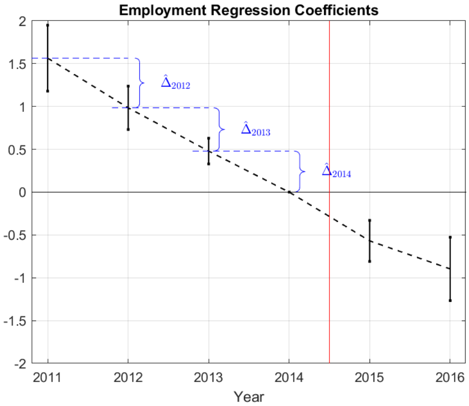
Question: whether the employment elasticity with respect to own wage is less than 1 in absolute value

$$\frac{\partial \log(\text{emp}_{rt})}{\partial \overline{GAP}_r} \geq -\frac{\partial \log(\text{wage}_{rt})}{\partial \overline{GAP}_r}$$

Employment Regression



Employment Regression



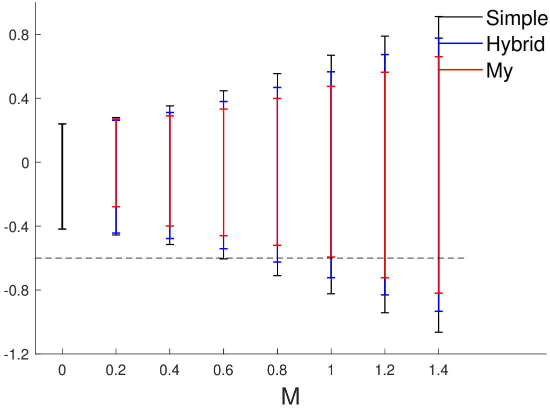
Employment Regression

Relax the parallel trends assumption by the second differences relative magnitudes

$$|\Delta_{2015} - \Delta_{2014}| \leqslant M \times \max_{t=2013,2014} |\Delta_t - \Delta_{t-1}|$$

The Effects of the Minimum Wage: Diff in Diff

The authors relax parallel trends as in RR23



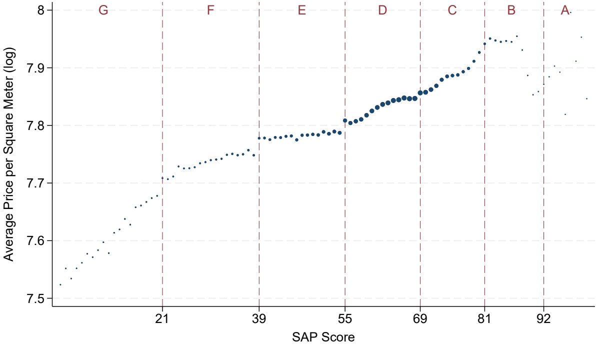
- Is employment effect ≥ -0.6 (wage effect) w/o parallel trends?
- The breakdown M: $M^{\text{My}} = 1$, $M^{\text{Hybrid}} = 0.75$, $M^{\text{sim}} = 0.6$

Energy Label Effects: Regression Discontinuity

Sejas-Portillo, Moro, and Stowasser (2025, AEJ) examine how energy labels influence property prices.

In the UK, residential properties for sale/rent report a SAP score, measured on a discrete scale from 1 to 100

- Rating bands from A to G are arbitrary and provide no additional information
- The simplified label may divert buyers' attention toward the rating bands



Energy Label Effects: Regression Discontinuity

The empirical strategy relies on a regression discontinuity design

$$P_i = \gamma_1 T_i + \gamma_2 T_i \times SAP_i + \gamma_2 T_i \times (SAP_i)^2 + \beta_0 + \beta_1 SAP_i + \beta_2 (SAP_i)^2 + \varepsilon_i$$

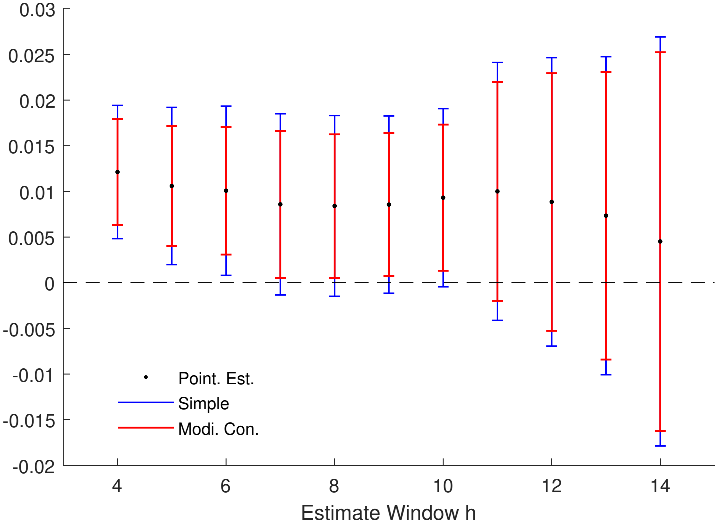
- P_i denotes the log of price per square meter
- T_i is an indicator for whether the SAP score has crossed a rating band cutoff
- γ_1 is a potentially biased estimand for the label effect

The running variable (the SAP score) is discrete. I follow Kolesár and Rothe (2018) and construct robust confidence intervals

Assumption: bounds on specification errors at the threshold

$$|\lim_{x \uparrow k} \Delta(x)| \leq \max_{x' < k} |\Delta(x')|, \ |\lim_{x \downarrow k} \Delta(x)| \leq \max_{x' > k} |\Delta(x')|$$

Energy Label Effects: Regression Discontinuity



For estimation window h, we have $4h(h+1)$ bounds

Computation takes approximately $2 \sim 20$ minutes per interval

Outline

Conclusion

This paper studies inference on **union bounds**

$$\theta \in \left[\min_{b \in \mathcal{B}} \lambda_{\ell,b}, \ \max_{b \in \mathcal{B}} \lambda_{u,b} \right]$$

I propose a CI based on *modified conditional* inference which

- Theory & Simulation: has shorter CI & larger local power under a large set of DGPs
- Empirical illustration: gives statistically significant results while the pre-existing alternatives do not
- Companion R package `UnionBounds` available online

Finite $\mathcal{B}$: Empirical Illustration

To study the employment and wage effect, run

$$\log(\mathrm{emp}_{rt}) = \sum_{\tau=2011,\tau\neq 2014}^{2016} \gamma_\tau^{\mathrm{e}} \overline{GAP}_r 1\left[\tau = t\right] + \alpha_r^{\mathrm{e}} + \xi_t^{\mathrm{e}} + \varepsilon_{rt}^{\mathrm{e}}$$

$$\log(\mathrm{wage}_{rt}) = \sum_{\tau=2011,\tau\neq 2014}^{2016} \gamma_\tau^{\mathrm{w}} \overline{GAP}_r 1\left[\tau = t\right] + \alpha_r^{\mathrm{w}} + \xi_t^{\mathrm{w}} + \varepsilon_{rt}^{\mathrm{w}}$$

- $\log(\mathrm{emp}_{rt})$ is the log employment in district r time t; $\log(\mathrm{wage}_{rt})$ is log wage
- $\overline{GAP}_r$ is a measure of the exposure to the minimum wage
- Pre-policy year 2011-2014

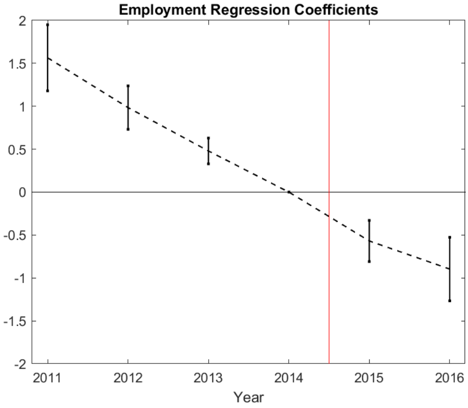
We are interested in the employment and wage effect at $t = 2015$ Back

Finite $\mathcal{B}$: Empirical Illustration

RR23 in Dustmann, Lindner, Schönberg, Umkehrer, Vom Berge (2022, QJE)

What are the effects of the minimum wage? Back

- Is employment effect ≥ -0.6 (wage effect) without parallel trends?



Finite $\mathcal{B}$: Empirical Illustration

RR23 in Dustmann, Lindner, Schönberg, Umkehrer, Vom Berge (2022, QJE)

What are the effects of the minimum wage? Back

- Is employment effect ≥ -0.6 (wage effect) without parallel trends?

Finite $\mathcal{B}$: Empirical Illustration

RR23 in Dustmann, Lindner, Schönberg, Umkehrer, Vom Berge (2022, QJE)

What are the effects of the minimum wage? Back

- Is employment effect ≥ -0.6 (wage effect) without parallel trends?

Relax the parallel trends assumption: $|\Delta_{2015} - \Delta_{2014}| \leqslant M \times \max_{t=2013,2014} |\Delta_t - \Delta_{t-1}|$

Simulation - Setting

I set $n = 5000$

Each sample $\{W_i\}_{i=1}^n$ and estimator is generated by

$$\begin{pmatrix} W_{\Delta,i} \\ W_{\gamma,i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \Delta \\ \gamma \end{pmatrix}, n\Omega\right)$$

The estimator is calculated by

$$\begin{pmatrix} \widehat{\Delta} \\ \widehat{\gamma} \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum W_{\Delta,i} \\ \frac{1}{n}\sum W_{\gamma,i} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \Delta \\ \gamma \end{pmatrix}, \Omega\right)$$

I conduct inference using pair $\left(\widehat{\Delta}, \widehat{\gamma}, \Omega\right)$

Back

Size & Power: Comparison with Simple CI

Theorem (Symmetric Or Large Bounds)

Suppose Assumptions KS, AN, FR, CE hold, $\alpha \in (0, \frac{1}{2})$

(Symmetric Bounds) *If* $corr(\widehat{\lambda}_{\ell,b_1}, \widehat{\lambda}_{\ell,b_2}) < \rho_1^*(\alpha, \alpha^c)$, $corr(\widehat{\lambda}_{\ell,b_\ell}, \widehat{\lambda}_{\ell,b_u}) < \rho_2^*(\alpha)$

$$\widehat{\lambda}_\ell = \widehat{\lambda}_u$$

1. *My CI is Strictly Shorter:*
   There is $\alpha' > \alpha$ *such that*
   $$\liminf_n \inf_{P \in \mathcal{P}_n} P\left(CI^m\left(\hat{\lambda}_n, \hat{\Sigma}_n/n; \alpha\right) \subseteq CI^{sim}\left(\hat{\lambda}_n, \hat{\Sigma}_n/n; \alpha'\right)\right) = 1$$
2. *My CI has Higher Power:*
   For all $P_n \in \mathcal{P}_n$, *there is a subsequence* P_{τ_n} *and* $\kappa \in (0, +\infty)$ *such that*
   $$\liminf_n P_{\tau_n}\left(\theta_{\tau_n} \notin CI^m\left(\hat{\lambda}, \hat{\Sigma}/\tau_n; \alpha\right)\right) - P_{\tau_n}\left(\theta_{\tau_n} \notin CI^{sim}\left(\hat{\lambda}, \hat{\Sigma}/\tau_n; \alpha\right)\right) > 0$$

for some $\theta_{\tau_n} = \theta_\ell - \frac{\kappa}{\sqrt{\tau_n}}$.

▸ Back

Size & Power: Comparison with Simple CI

Theorem (Symmetric Or Large Bounds)

Suppose Assumptions KS, AN, FR, CE hold, $\alpha \in (0, \frac{1}{2})$
(Large Bounds) *Let* $\kappa_n = o(\sqrt{n})$ *and* $\kappa_n \to \infty$*, and*

$$\mathcal{P}_n = \left\{ P \in \mathcal{P} : \lambda_{u,b_u} - \lambda_{\ell,b_\ell} \geq \frac{\kappa_n}{\sqrt{n}} \right\}$$

1. *My CI is Strictly Shorter:*
   There is $\alpha' > \alpha$ *such that*

$$\liminf_n \inf_{P \in \mathcal{P}_n} P \left(CI^m \left(\hat{\lambda}_n, \hat{\Sigma}_n / n; \alpha \right) \subseteq CI^{sim} \left(\hat{\lambda}_n, \hat{\Sigma}_n / n; \alpha' \right) \right) = 1$$

2. *My CI has Higher Power:*
   For all $P_n \in \mathcal{P}_n$*, there is a subsequence* P_{τ_n} *and* $\kappa \in (0, +\infty)$ *such that*

$$\liminf_n P_{\tau_n} \left(\theta_{\tau_n} \notin CI^m \left(\hat{\lambda}, \hat{\Sigma} / \tau_n; \alpha \right) \right) - P_{\tau_n} \left(\theta_{\tau_n} \notin CI^{sim} \left(\hat{\lambda}, \hat{\Sigma} / \tau_n; \alpha \right) \right) > 0$$

for some $\theta_{\tau_n} = \theta_\ell - \frac{\kappa}{\sqrt{\tau_n}}$.

Back

Intersection & Union

Intersection bounds

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Intersection & Union

Intersection bounds

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Intersection & Union

Intersection bounds

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated

Solution: $\sqrt{n}(\lambda_{\ell,b} - \theta) \leq 0 \Rightarrow$ get an upper bound $\sqrt{n}(\lambda_{\ell,b} - \theta)/\sqrt{\ln n}$

Intersection & Union

Intersection bounds

$$H_0 : \max_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated

Solution: $\sqrt{n}(\lambda_{\ell,b} - \theta) \leq 0 \Rightarrow$ get an upper bound $\sqrt{n}(\lambda_{\ell,b} - \theta)/\sqrt{\ln n}$

Union bounds

$$H_0 : \min_{b \in \mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \min_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \min_{b \in \mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Intersection & Union

Intersection bounds

$$H_0 : \max_{b\in\mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \max_{b\in\mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \max_{b\in\mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated

Solution: $\sqrt{n}(\lambda_{\ell,b} - \theta) \leq 0 \Rightarrow$ get an upper bound $\sqrt{n}(\lambda_{\ell,b} - \theta)/\sqrt{\ln n}$

Union bounds

$$H_0 : \min_{b\in\mathcal{B}} \lambda_{\ell,b} \leq \theta$$

An intuitive test statistic

$$\hat{T}(\theta) = \min_{b\in\mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \theta)}{\sigma_{\ell,b}} = \min_{b\in\mathcal{B}} \frac{\sqrt{n}(\hat{\lambda}_{\ell,b} - \lambda_{\ell,b})}{\sigma_{\ell,b}} + \frac{\sqrt{n}(\lambda_{\ell,b} - \theta)}{\sigma_{\ell,b}}$$

Key difficulty: $\sqrt{n}(\lambda_{\ell,b} - \theta)$ cannot be consistently estimated & unknown sign

Back

Step 2: Conditional CV

Lemma

$$\frac{\Phi\left(\hat{T}(\theta)\right) - \Phi\left(t_{\ell,1}(\theta, b_\ell)\right)}{\Phi\left(t_{\ell,2}(\theta, b_\ell)\right) - \Phi\left(t_{\ell,1}(\theta, b_\ell)\right)} \,\Big|\, \left\{ \hat{T}(\theta) = \mathcal{Z}_{\ell, b_\ell} \right\} \quad \overset{\text{FOSD}}{\preceq} \quad \text{Unif}(0, 1)$$

A Simple Example - Step 3 Conditional CV

I illustrate with Back

$$\hat{T}(\theta) \,\Big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s$$

A Simple Example - Step 3 Conditional CV

I illustrate with ▸ Back

$$
\begin{aligned}
&\hat{T}(\theta) \,\middle|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim &\hat{\lambda}_1 - \theta \,\middle|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s
\end{aligned}
$$

A Simple Example - Step 3 Conditional CV

I illustrate with ‣ Back

$$
\begin{aligned}
&\hat{T}(\theta) \,\Big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim &\hat{\lambda}_1 - \theta \,\Big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim &\hat{\lambda}_1 - \theta \,\Big|\, \theta - \hat{\lambda}_2 \leq \hat{\lambda}_1 - \theta \leq \hat{\lambda}_2 - \theta, \hat{s} = s
\end{aligned}
$$

A Simple Example - Step 3 Conditional CV

I illustrate with ▸ Back

$$
\begin{aligned}
& \hat{T}(\theta) \,\big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim & \hat{\lambda}_1 - \theta \,\big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim & \hat{\lambda}_1 - \theta \,\big|\theta - \hat{\lambda}_2 \le \hat{\lambda}_1 - \theta \le \hat{\lambda}_2 - \theta, \hat{s} = s \\
\sim & \hat{\lambda}_1 - \theta \,\big|\theta - \hat{\lambda}_2 \le \hat{\lambda}_1 - \theta \le \hat{\lambda}_2 - \theta, \text{ and } s = 1
\end{aligned}
$$

A Simple Example - Step 3 Conditional CV

I illustrate with ▸ Back

$$
\begin{aligned}
&\hat{T}(\theta)\,\big|\,\hat{T}(\theta)=\hat{\lambda}_1-\theta,\hat{s}=s\\
\sim&\hat{\lambda}_1-\theta\,\big|\,\hat{T}(\theta)=\hat{\lambda}_1-\theta,\hat{s}=s\\
\sim&\hat{\lambda}_1-\theta\,\big|\theta-\hat{\lambda}_2\leq\hat{\lambda}_1-\theta\leq\hat{\lambda}_2-\theta,\hat{s}=s\\
\sim&\hat{\lambda}_1-\theta\,\big|\theta-\hat{\lambda}_2\leq\hat{\lambda}_1-\theta\leq\hat{\lambda}_2-\theta,\text{ and } s=1\\
\sim&\mathcal{TN}\left(\lambda_1-\theta,\left[\theta-\hat{\lambda}_2,\hat{\lambda}_2-\theta\right]\right)
\end{aligned}
$$

A Simple Example - Step 3 Conditional CV

I illustrate with ▸ Back

$$
\begin{aligned}
&\hat{T}(\theta) \,\big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim &\hat{\lambda}_1 - \theta \,\big|\, \hat{T}(\theta) = \hat{\lambda}_1 - \theta, \hat{s} = s \\
\sim &\hat{\lambda}_1 - \theta \,\big|\theta - \hat{\lambda}_2 \leq \hat{\lambda}_1 - \theta \leq \hat{\lambda}_2 - \theta, \hat{s} = s \\
\sim &\hat{\lambda}_1 - \theta \,\big|\theta - \hat{\lambda}_2 \leq \hat{\lambda}_1 - \theta \leq \hat{\lambda}_2 - \theta, \text{ and } s = 1 \\
\sim &\mathcal{TN}\left(\lambda_1 - \theta, \left[\theta - \hat{\lambda}_2, \hat{\lambda}_2 - \theta\right]\right) \\
\preceq &\mathcal{TN}\left(0, \left[\theta - \hat{\lambda}_2, \hat{\lambda}_2 - \theta\right]\right)
\end{aligned}
$$

A Simple Example - Step 3 Conditional CV

Lemma

$$\frac{\Phi\left(\hat{T}(\theta)\right)-\Phi\left(t_{\ell,1}(\theta,b_\ell)\right)}{\Phi\left(t_{\ell,2}(\theta,b_\ell)\right)-\Phi\left(t_{\ell,1}(\theta,b_\ell)\right)}\,|\{\hat{T}(\theta)=\mathcal{Z}_{\ell,b_\ell}\} \overset{\text{FOSD}}{\preceq} \text{Unif}(0,1)$$

$$t_{\ell,1}(\theta,b)=\begin{cases}\min_{\tilde{b}\in\mathcal{B}}\left(1+\rho_{\ell u}(b,\tilde{b})\right)^{-1}\left(\mathcal{Z}_{u,\tilde{b}}+\rho_{\ell u}(b,\tilde{b})\mathcal{Z}_{\ell,b}\right), & \textit{if } \min_{\tilde{b}\in\mathcal{B}}\rho_{\ell u}(b,\tilde{b})>-1\\ -\infty & \textit{otherwise}\end{cases}$$

$$t_{\ell,2}(\theta,b)=\begin{cases}\min_{\tilde{b}\in\mathcal{B}:\rho_\ell(b,\tilde{b})<1}\left(1-\rho_{\ell}(b,\tilde{b})\right)^{-1}\left(\mathcal{Z}_{\ell,\tilde{b}}-\rho_{\ell}(b,\tilde{b})\mathcal{Z}_{\ell,b}\right) & \textit{if } \min_{\tilde{b}\in\mathcal{B}}\rho_{\ell}(b,\tilde{b})<1\\ +\infty & \textit{otherwise}\end{cases}$$

$$\rho_\ell(b_1,b_2)=\frac{\Sigma_{\ell,b_1b_2}}{\sigma_{\ell,b_1}\sigma_{\ell,b_2}},\quad \rho_{\ell u}(b_1,b_2)=\frac{\Sigma_{\ell u,b_1b_2}}{\sigma_{\ell,b_1}\sigma_{u,b_2}},$$

Back

How to choose α^{c}

Larger α^{c}: smaller $\hat{c}^{\mathrm{con}}$ & larger $\hat{c}^{\mathrm{t}}$ $\Rightarrow$ larger power w/ less fav. DGPs

How to choose α^{c}

Larger α^{c}: smaller $\hat{c}^{\mathrm{con}}$ & larger $\hat{c}^{\mathrm{t}}$ $\Rightarrow$ larger power w/ less fav. DGPs

I suggest $\alpha^{\mathrm{c}} = \frac{4}{5}\alpha$

- $CI^{\mathrm{m}} \subsetneqq CI^{\mathrm{sim}}$ if (i) the bound is wide; (ii) the bound is symmetric
- performs well in the simulation and empirical applications

How to choose α^{c}

Larger α^{c}: smaller $\hat{c}^{\text{con}}$ & larger $\hat{c}^{\text{t}}$ $\Rightarrow$ larger power w/ less fav. DGPs

I suggest $\alpha^{\mathsf{c}} = \frac{4}{5}\alpha$

- $CI^{\text{m}} \subsetneq CI^{\text{sim}}$ if (i) the bound is wide; (ii) the bound is symmetric
- performs well in the simulation and empirical applications

Special Case: $\lambda_{\ell,b_\ell} \ll \lambda_{u,b_u}$, $\lambda_{\ell,b_\ell} \ll \min_{b\in\mathcal{B}\setminus b_\ell} \lambda_{\ell,b}$, $\lambda_{u,b_u} \gg \min_{b\in\mathcal{B}\setminus b_u} \lambda_{u,b}$

$$\hat{c}^{\text{con}} \approx \Phi^{-1}(1-\alpha^{\mathsf{c}})$$

- By Imbens and Manski (2004), we only need to use $c = \Phi^{-1}(1-\alpha)$
- If $\alpha^{\mathsf{c}} = \alpha$, $c^t \approx \Phi^{-1}(1-\frac{\alpha}{2})$

How to choose α^{c}

Larger α^{c}: smaller $\hat{c}^{\mathsf{con}}$ & larger $\hat{c}^{\mathsf{t}}$ $\Rightarrow$ larger power w/ less fav. DGPs

I suggest $\alpha^{\mathsf{c}} = \frac{4}{5}\alpha$

- $CI^{\mathsf{m}} \subsetneqq CI^{\mathsf{sim}}$ if (i) the bound is wide; (ii) the bound is symmetric
- performs well in the simulation and empirical applications

Special Case: $\lambda_{\ell,b_\ell} \ll \lambda_{u,b_u}$, $\lambda_{\ell,b_\ell} \ll \min_{b\in\mathcal{B}\setminus b_\ell} \lambda_{\ell,b}$, $\lambda_{u,b_u} \gg \min_{b\in\mathcal{B}\setminus b_u} \lambda_{u,b}$

$$\hat{c}^{\mathsf{con}} \approx \Phi^{-1}(1-\alpha^{\mathsf{c}})$$

- By Imbens and Manski (2004), we only need to use $c = \Phi^{-1}(1-\alpha)$
- If $\alpha^{\mathsf{c}} = \alpha$, $c^t \approx \Phi^{-1}(1-\frac{\alpha}{2})$

Solution 1: depends on λ, e.g. test $\lambda_{\ell,b_\ell} \ll \lambda_{u,b_u}$

- need additional tuning parameter

How to choose α^c

Larger α^c: smaller $\hat{c}^{con}$ & larger $\hat{c}^t$ $\Rightarrow$ larger power w/ less fav. DGPs

I suggest $\alpha^c = \frac{4}{5}\alpha$

- $CI^m \subsetneq CI^{sim}$ if (i) the bound is wide; (ii) the bound is symmetric
- performs well in the simulation and empirical applications

Special Case: $\lambda_{\ell,b_\ell} \ll \lambda_{u,b_u}$, $\lambda_{\ell,b_\ell} \ll \min_{b\in\mathcal{B}\setminus b_\ell} \lambda_{\ell,b}$, $\lambda_{u,b_u} \gg \min_{b\in\mathcal{B}\setminus b_u} \lambda_{u,b}$

$$\hat{c}^{con} \approx \Phi^{-1}(1-\alpha^c)$$

- By Imbens and Manski (2004), we only need to use $c = \Phi^{-1}(1-\alpha)$
- If $\alpha^c = \alpha$, $c^t \approx \Phi^{-1}(1-\frac{\alpha}{2})$

Solution 1: depends on λ, e.g. test $\lambda_{\ell,b_\ell} \ll \lambda_{u,b_u}$

- need additional tuning parameter

Solution 2: depends on Σ, e.g. calculate weighted average power by simulation

- can be time consuming

Back