Estimation and Inference in Boundary Discontinuity Designs

Matias D. Cattaneo
Princeton University

Rocio Titiunik
Princeton University

Ruiqi (Rae) Yu
Princeton University

March 2025

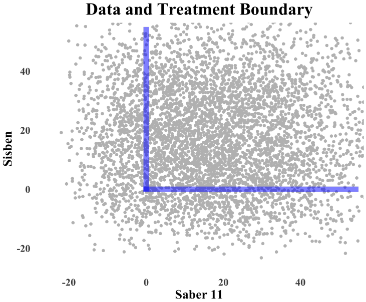
Outline

Introduction

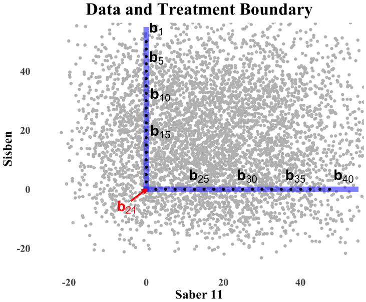
Boundary Discontinuity Designs are used in causal inference and policy evaluation.

- Multi-dimensional Regression Discontinuity (RD) designs.
  - Multi-score RD designs / Geographic RD designs.
- Two main approach for analysis in practice:
  - Local regression based on <u>univariate</u> distance to boundary.
  - Local regression based on <u>bivariate</u> location relative to boundary.
- Today: foundational, thorough study of Boundary Discontinuity Designs.
  - *Methodology*: guidance on valid and <u>invalid</u> current practices, and more.
  - *Theory*: novel strong approximation approach for uniform inference, and more.
  - *Practice*: new `R` software (`rd2d` package).

`https://rdpackages.github.io/`



- *Ser Pilo Paga* (SPP) Colombian policy program; students $i = 1, 2, \ldots, n$.
- $\mathbf{X}_i = (\texttt{SABER11}_i, \texttt{SISBEN}_i)^\top$; $\texttt{SABER11}_i$ = exam score and $\texttt{SISBEN}_i$ = wealth index.
- $\mathscr{B} = \{\texttt{SABER11} \geq 0 \text{ and } \texttt{SISBEN} = 0\} \cup \{\texttt{SABER11} = 0 \text{ and } \texttt{SISBEN} \geq 0\}$.
- $(Y_i(0), Y_i(1), \mathbf{X}_i)$, $i = 1, 2, \ldots, n$, random sample.
- $Y_i = \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_0) \cdot Y_i(0) + \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) \cdot Y_i(1)$; $\mathscr{A}_t$ group t's assignment area.



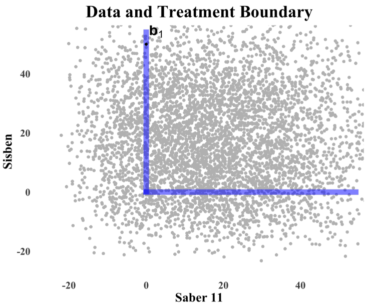
- Causal treatment effect along the assignment boundary:

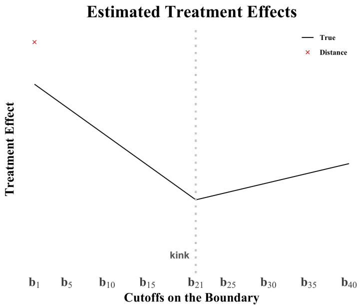
$$\tau(\mathbf{x}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{x}], \qquad \mathbf{x} \in \mathscr{B}.$$

- Estimation and Inference Approaches:
  - Local regression based on univariate distance to boundary:

$$D_i(\mathbf{x}) = \mathscr{d}(\mathbf{X}_i, \mathbf{x})(\mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) - \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_0)), \qquad \mathbf{x} \in \mathscr{B}.$$

  - Local regression based on bivariate location relative to boundary.

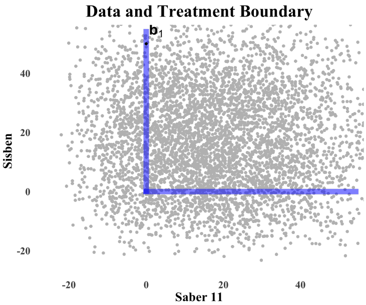


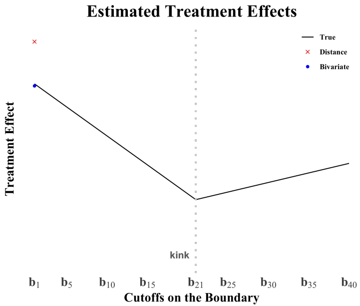


- Distance-based Estimator: $\widehat{\tau}_{\text{dis}}(\mathbf{b}_1) = \mathbf{e}_1^\top \widehat{\boldsymbol{\gamma}}_1(\mathbf{b}_1) - \mathbf{e}_1^\top \widehat{\boldsymbol{\gamma}}_0(\mathbf{b}_1)$, where

$$\widehat{\boldsymbol{\gamma}}_t(\mathbf{x}) = \arg\min_{\boldsymbol{\gamma}} \sum_{i=1}^{n} \Big[\big(Y_i - \mathbf{r}_p(D_i(\mathbf{x}))^\top \boldsymbol{\gamma} \big)^2 k_h(D_i(\mathbf{x})) \mathbb{1}(D_i(\mathbf{x}) \in \mathscr{I}_t) \Big].$$

  - $\mathbf{r}_p(u) = (1, u, u^2, \cdots, u^p)^\top$.
  - $k_h(u) = k(u/h)/h$, for univariate kernel $k(\cdot)$ and bandwidth h.
  - $\mathscr{I}_0 = (-\infty, 0)$ and $\mathscr{I}_1 = [0, \infty)$.
  - $D_i(\mathbf{x}) = \mathscr{d}(\mathbf{X}_i, \mathbf{x})(\mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) - \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_0))$.
  - $\mathbf{x} \in \mathscr{B}$ and $t \in \{0, 1\}$.

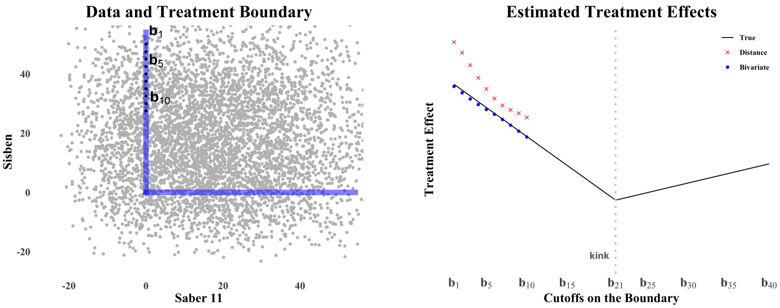




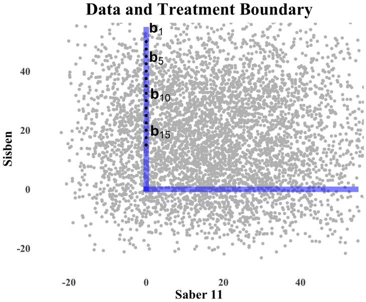
- Location-based Estimator: $\widehat{\tau}(\mathbf{b}_1) = \mathbf{e}_1^\top \widehat{\boldsymbol{\beta}}_1(\mathbf{b}_1) - \mathbf{e}_1^\top \widehat{\boldsymbol{\beta}}_0(\mathbf{b}_1)$ for $\mathbf{x} \in \mathscr{B}$,

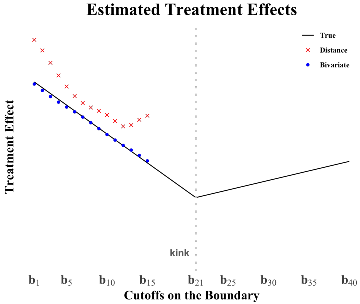
$$\widehat{\boldsymbol{\beta}}_t(\mathbf{x}) = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left(Y_i - \mathbf{R}_p(\mathbf{X}_i - \mathbf{x})^\top \boldsymbol{\beta}\right)^2 K_h\left(\frac{\mathbf{X}_i - \mathbf{x}}{h}\right) \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_t).$$

  - $\mathbf{R}_p(\mathbf{u}) = (1, u_1, u_2, u_1^2, u_2^2, u_1 u_2, \cdots, u_1^p, u_2^p)^\top$.
  - $K_h(\mathbf{u}) = K_h(u_1/h, u_2/h)/h^2$, for bivariate kernel $K(\cdot)$ and bandwidth h.
  - $\mathscr{A}_0$ = treatment region and $\mathscr{A}_1$ control region.
  - $\mathbf{X}_i$ bivariate score.
  - $\mathbf{x} \in \mathscr{B}$ and $t \in \{0, 1\}$.



- Distance-based Estimator: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$.
- Location-based Estimator: $\widehat{\tau}(\mathbf{x})$.
- Evaluation points along $\mathscr{B}$: $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{10}\}$.

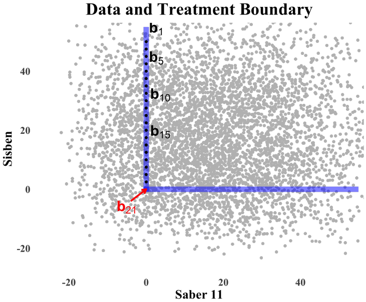


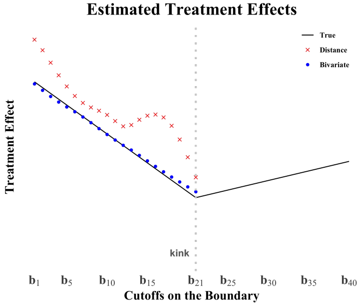


- Distance-based Estimator: $\widehat{\tau}_{\mathrm{dis}}(\mathbf{x})$.

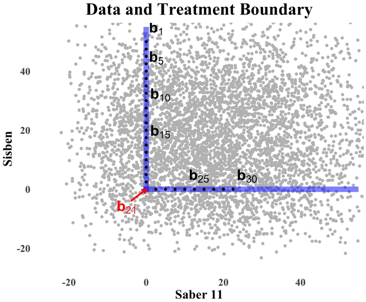
- Location-based Estimator: $\widehat{\tau}(\mathbf{x})$.

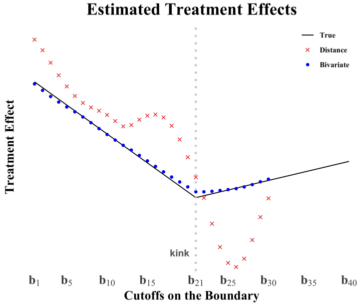
- Evaluation points along $\mathscr{B}$: $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{15}\}$.



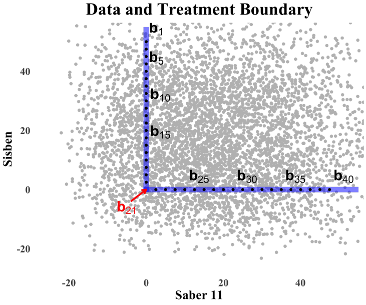


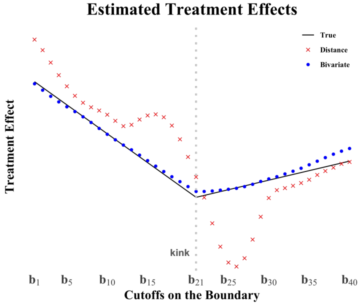
- Distance-based Estimator: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$.
- Location-based Estimator: $\widehat{\tau}(\mathbf{x})$.
- Evaluation points along $\mathscr{B}$: $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{21}\}$.



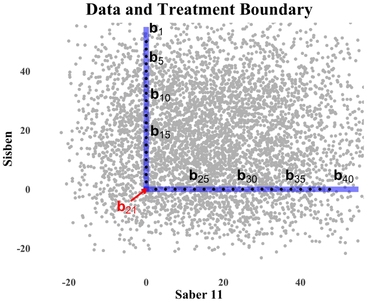


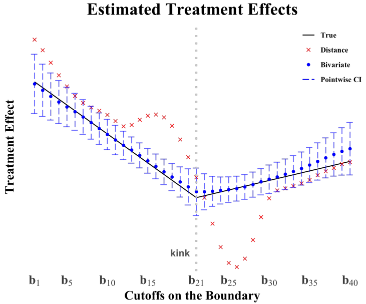
- Distance-based Estimator: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$.
- Location-based Estimator: $\widehat{\tau}(\mathbf{x})$.
- Evaluation points along $\mathscr{B}$: $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{30}\}$.





- Distance-based Estimator: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$.
- Location-based Estimator: $\widehat{\tau}(\mathbf{x})$.
- Evaluation points along $\mathscr{B}$: $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{40}\}$.



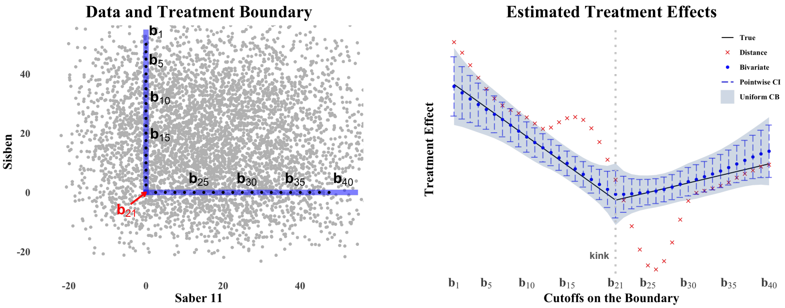


- Estimators: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$ and $\widehat{\tau}(\mathbf{x})$, for each $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{40}\}$.

- Uncertainty Quantification: Confidence Intervals. For each $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{40}\}$,

$$\widehat{I}(\mathbf{x};\alpha) = \left[\, \widehat{\tau}(\mathbf{x}) - \mathfrak{q}_\alpha \sqrt{\widehat{\Omega}_{\mathbf{x}}}\ ,\ \widehat{\tau}(\mathbf{x}) + \mathfrak{q}_\alpha \sqrt{\widehat{\Omega}_{\mathbf{x}}}\, \right].$$

  - $\mathfrak{q}_\alpha = \Phi^{-1}(1-\alpha/2)$, where $\Phi(x)$ be the standard Gaussian CDF.
  - $\mathfrak{q}_{0.95} \approx 1.96$.



- Estimators: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$ and $\widehat{\tau}(\mathbf{x})$, uniformly in $\mathbf{x} \in \mathscr{B}$.

- Uncertainty Quantification: Confidence Bands. Uniformly in $\mathbf{x} \in \mathscr{B}$,

$$\widehat{I}(\mathbf{x};\alpha) = \left[\, \widehat{\tau}(\mathbf{x}) - \mathfrak{q}_\alpha \sqrt{\widehat{\Omega}_{\mathbf{x}}}\ ,\ \widehat{\tau}(\mathbf{x}) + \mathfrak{q}_\alpha \sqrt{\widehat{\Omega}_{\mathbf{x}}} \,\right].$$

  - $\mathfrak{q}_\alpha = \inf\{c > 0 : \mathbb{P}[\sup_{\mathbf{x}\in\mathscr{B}} |\widehat{Z}_n(\mathbf{x})| \geq c|\text{data}] \leq \alpha\}$.
  - $(\widehat{Z}_n : \mathbf{x} \in \mathscr{B})$ is a Gaussian process conditional on data, with $\mathbb{E}[\widehat{Z}_n(\mathbf{x}_1)|\text{data}] = 0$ and an estimated covariance function $\mathbb{E}[\widehat{Z}_n(\mathbf{x}_1)\widehat{Z}_n(\mathbf{x}_2)|\text{data}]$ for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathscr{B}$.

Outline

Overview

- Analysis based on <u>univariate</u> distance to boundary: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$.
  1. Give sufficient conditions for identification.
  2. Show existence of "large" misspecification bias near a kink of $\mathscr{B}$.
  3. Show "small" misspecification bias when $\mathscr{B}$ is smooth.
  4. Establish pointwise and uniform convergence rates and distribution theory.
  5. Discuss connects and differences with standard univariate RD designs.

- Analysis based on <u>bivariate</u> location relative to boundary: $\widehat{\tau}(\mathbf{x})$.
  1. Identification and misspecification bias are standard.
  2. Mean square error expansions and bandwidth selection.
  3. Establish pointwise and uniform convergence rates and distribution theory.
  4. New methods for analysis of Boundary Discontinuity Designs.

- New strong approximation result for empirical processes.
  1. Finite polynomial moments.
  2. Incorporates possibly lower dimensional manifold structure $\mathscr{B} \subseteq \text{Supp}(\mathbf{X}_i)$.

Outline

Distance-Based Methods: Identification

- **Parameter**. $\tau(\mathbf{x}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{x}]$ for all $\mathbf{x} \in \mathscr{B}$.

- **Estimator**. $\widehat{\tau}_{\mathrm{dis}}(\mathbf{x}) = \mathbf{e}_1^\top \widehat{\gamma}_1(\mathbf{x}) - \mathbf{e}_1^\top \widehat{\gamma}_0(\mathbf{x})$ for $\mathbf{x} \in \mathscr{B}$, where

$$\widehat{\boldsymbol{\gamma}}_t(\mathbf{x}) = \arg\min_{\boldsymbol{\gamma}} \sum_{i=1}^{n} \Big[\big(Y_i - \mathbf{r}_p(D_i(\mathbf{x}))^\top \boldsymbol{\gamma}\big)^2 k_h(D_i(\mathbf{x})) \mathbb{1}(D_i(\mathbf{x}) \in \mathscr{I}_t) \Big].$$

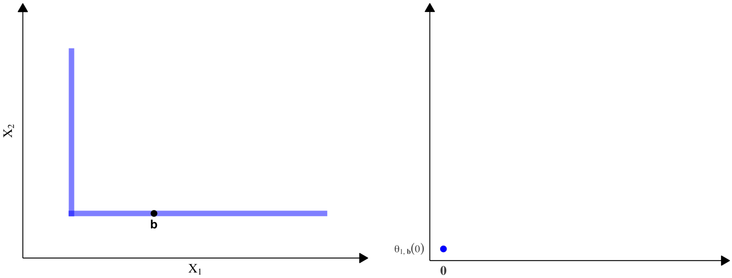
- **Assumption**. Let $t \in \{0, 1\}$.
  - $\mathscr{d} : \mathbb{R}^2 \mapsto [0, \infty)$ satisfies $\|\mathbf{x}_1 - \mathbf{x}_2\| \lesssim \mathscr{d}(\mathbf{x}_1, \mathbf{x}_2) \lesssim \|\mathbf{x}_1 - \mathbf{x}_2\|$ for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathscr{X}$.
  - $k : \mathbb{R} \to [0, \infty)$ is compact supported and Lipschitz continuous, or $k(u) = \mathbb{1}(u \in [-1, 1])$.
  - $\liminf_{h \downarrow 0} \inf_{\mathbf{x} \in \mathscr{B}} \int_{\mathscr{A}_t} k_h(\mathscr{d}(\mathbf{u}, \mathbf{x})) d\mathbf{u} \gtrsim 1$.

- **Identification**. For all $\mathbf{x} \in \mathscr{B}$,

$$\tau(\mathbf{x}) = \lim_{r \downarrow 0} \theta_{1,\mathbf{x}}(r) - \lim_{r \uparrow 0} \theta_{0,\mathbf{x}}(r)$$

with

$$\theta_{t,\mathbf{x}}(r) = \mathbb{E}[Y_i | D_i(\mathbf{x}) = r, D_i(\mathbf{x}) \in \mathscr{I}_t].$$



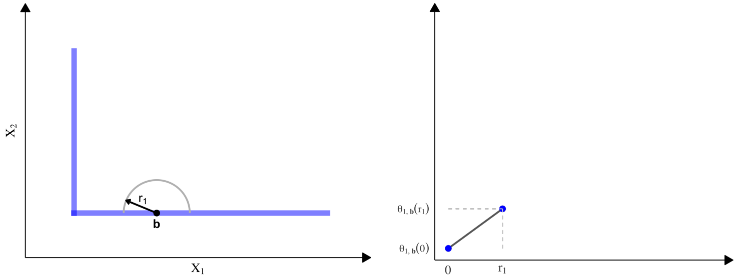
- **Best L_2 Approximation**. The distance-based estimator $\widehat{\tau}_{\text{dis}}(\mathbf{b})$ is sample analogue of

$$\tau^*_{\text{dis}}(\mathbf{b}) = \mathbf{e}_1^\top \boldsymbol{\gamma}_1^*(\mathbf{b}) - \mathbf{e}_1^\top \boldsymbol{\gamma}_0^*(\mathbf{b}),$$

where

$$\boldsymbol{\gamma}_t^*(\mathbf{x}) = \arg\min_{\boldsymbol{\gamma}} \mathbb{E}\Big[\big(Y_i - \mathbf{r}_p(D_i(\mathbf{x}))^\top \boldsymbol{\gamma}\big)^2 k_h(D_i(\mathbf{x}))\mathbb{1}(D_i(\mathbf{x}) \in \mathscr{I}_t)\Big]$$

for $t \in \{0, 1\}$.



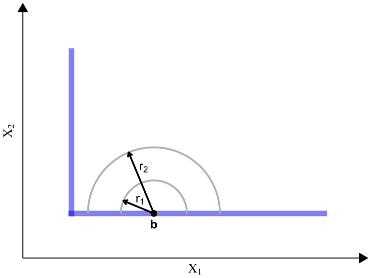
- $\theta^*_{t,\mathbf{b}}(0) = \mathbf{e}_1^\top \boldsymbol{\gamma}^*_t(\mathbf{b})$ is the best L_2-approx of $\theta_{t,\mathbf{b}}(r) = \mathbb{E}[Y_i | D_i(\mathbf{b}) = r, D_i(\mathbf{x}) \in \mathscr{I}_t]$.

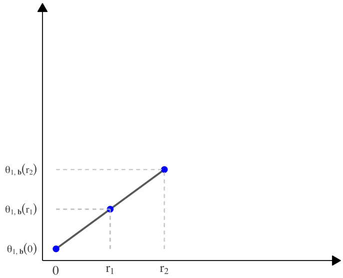
- **Bias**. Using the identification result,

$$\mathfrak{B}_n(\mathbf{b}) = \Big[\theta^*_{1,\mathbf{b}}(0) - \theta_{1,\mathbf{b}}(0)\Big] - \Big[\theta^*_{0,\mathbf{b}}(0) - \theta_{0,\mathbf{b}}(0)\Big] = \theta^*_{1,\mathbf{b}}(0) - \theta^*_{0,\mathbf{b}}(0) - \tau(\mathbf{b})$$

  is the best-L_2 misspecification bias of the estimator $\widehat{\tau}_{\mathrm{dis}}(\mathbf{b})$.

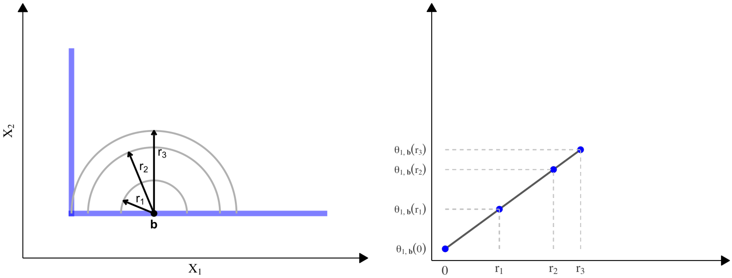
- **Smoothness**. If $r \mapsto \theta_{t,\mathbf{b}}(r)$ is locally to zero $(p+1)$th smooth, then $\mathfrak{B}_n(\mathbf{b}) \lesssim h^{p+1}$.





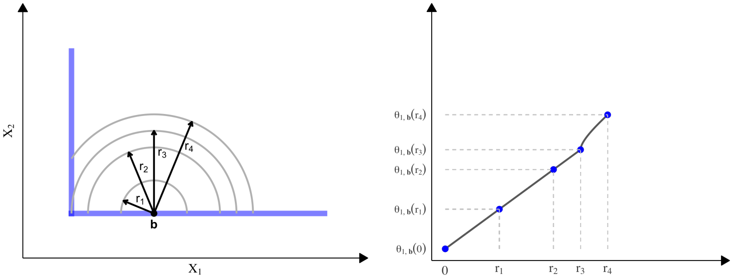
- **Bias**. $\mathfrak{B}_n(\mathbf{b}) = \Big[\theta^*_{1,\mathbf{b}}(0) - \theta_{1,\mathbf{b}}(0)\Big] - \Big[\theta^*_{0,\mathbf{b}}(0) - \theta_{0,\mathbf{b}}(0)\Big] = \theta^*_{1,\mathbf{b}}(0) - \theta^*_{0,\mathbf{b}}(0) - \tau(\mathbf{b})$.

- **Smoothness**. $r \mapsto \theta_{t,\mathbf{b}}(r)$ is locally to zero $(p+1)$th smooth, thus $\mathfrak{B}_n(\mathbf{b}) \lesssim h^{p+1}$.



- **Bias**. $\mathfrak{B}_n(\mathbf{b}) = \left[\theta^*_{1,\mathbf{b}}(0) - \theta_{1,\mathbf{b}}(0)\right] - \left[\theta^*_{0,\mathbf{b}}(0) - \theta_{0,\mathbf{b}}(0)\right] = \theta^*_{1,\mathbf{b}}(0) - \theta^*_{0,\mathbf{b}}(0) - \tau(\mathbf{b})$.

- **Smoothness**. $r \mapsto \theta_{t,\mathbf{b}}(r)$ is locally to zero $(p+1)$th smooth, thus $\mathfrak{B}_n(\mathbf{b}) \lesssim h^{p+1}$.

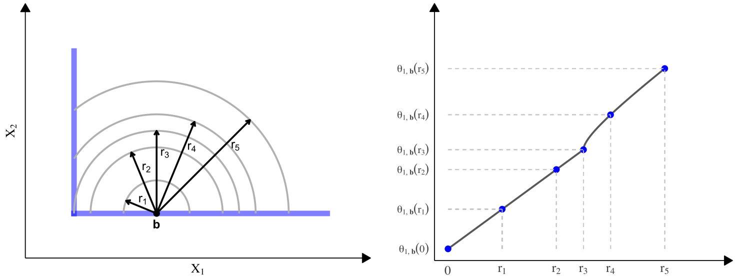


- **Bias**. $\mathfrak{B}_n(\mathbf{b}) = \left[\theta^*_{1,\mathbf{b}}(0) - \theta_{1,\mathbf{b}}(0)\right] - \left[\theta^*_{0,\mathbf{b}}(0) - \theta_{0,\mathbf{b}}(0)\right] = \theta^*_{1,\mathbf{b}}(0) - \theta^*_{0,\mathbf{b}}(0) - \tau(\mathbf{b})$.

- **Smoothness**. $r \mapsto \theta_{t,\mathbf{b}}(r)$ is locally to zero Lipschitz, thus $\mathfrak{B}_n(\mathbf{b}) \lesssim h$.

- **Derivatives**. $r \mapsto \theta_{t,\mathbf{b}}(r)$ is not differentiable for all $r \geq r_3$, and

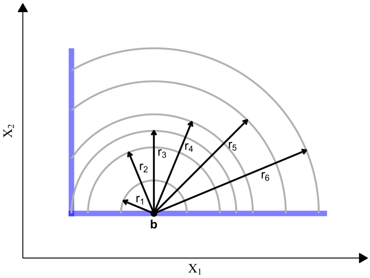
$$\lim_{r \uparrow r_3} \frac{d}{dr}\theta_{t,\mathbf{b}}(r) \neq \lim_{r \downarrow r_3} \frac{d}{dr}\theta_{t,\mathbf{b}}(r)$$

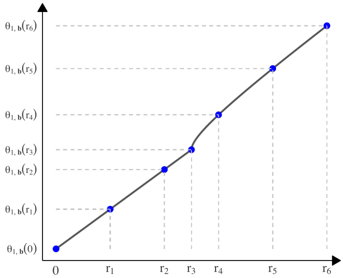


- **Bias**. $\mathfrak{B}_n(\mathbf{b}) = \Big[\theta^*_{1,\mathbf{b}}(0) - \theta_{1,\mathbf{b}}(0)\Big] - \Big[\theta^*_{0,\mathbf{b}}(0) - \theta_{0,\mathbf{b}}(0)\Big] = \theta^*_{1,\mathbf{b}}(0) - \theta^*_{0,\mathbf{b}}(0) - \tau(\mathbf{b})$.

- **Smoothness**. $r \mapsto \theta_{t,\mathbf{b}}(r)$ is locally to zero Lipschitz, thus $\mathfrak{B}_n(\mathbf{b}) \lesssim h$.

- **Pointwise Analysis**. Need to choose bandwidth $h \leq r_3 = d(\mathbf{b}, \text{kink})$.
  - Bandwidth must vary with $\mathbf{b} \in \mathscr{B}$, depending on "smoothness" of boundary!
  - The closer to a kink point on $\mathscr{B}$, the smaller the bandwidth h must be.

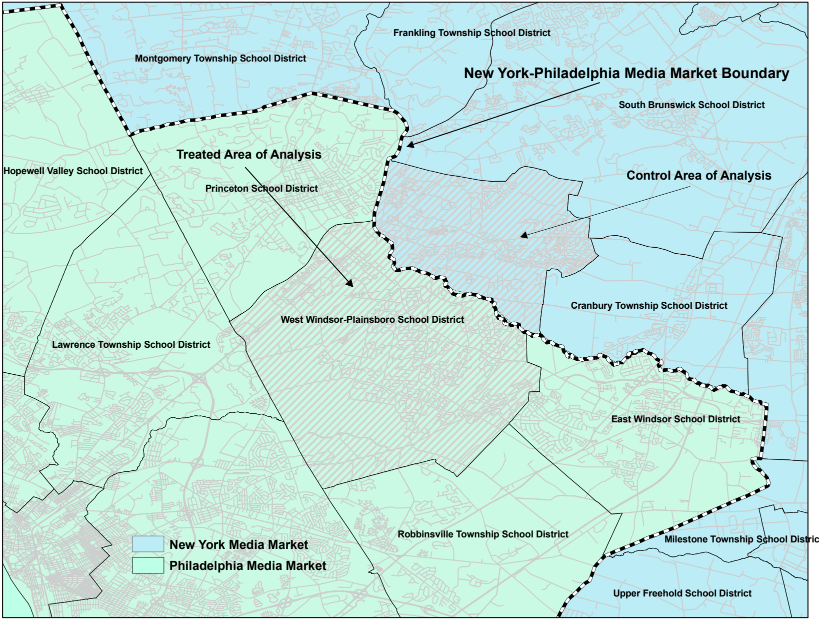




- **Uniform Analysis**. Under minimal regularity conditions, and for any $p \geq 1$,

$$1 \lesssim \liminf_{n\to\infty} \sup_{\mathbb{P}\in\mathscr{P}} \sup_{\mathbf{x}\in\mathscr{B}} \frac{\mathfrak{B}_n(\mathbf{x})}{h} \leq \limsup_{n\to\infty} \sup_{\mathbb{P}\in\mathscr{P}} \sup_{\mathbf{x}\in\mathscr{B}} \frac{\mathfrak{B}_n(\mathbf{x})}{h} \lesssim 1.$$

  - Bias cannot be better than order h (Lipschitz continuity) if $\mathscr{B}$ is non-smooth!
  - If $\mathscr{B}$ is smooth, then $\sup_{\mathbf{x}\in\mathscr{B}} \mathfrak{B}_n(\mathbf{x}) \lesssim h^{p+1}$.



Frankling Township School District
Montgomery Township School District
New York-Philadelphia Media Market Boundary
South Brunswick School District
Treated Area of Analysis
Hopewell Valley School District
Princeton School District
Control Area of Analysis
Cranbury Township School District
West Windsor-Plainsboro School District
Lawrence Township School District
East Windsor School District
Robbinsville Township School District
Milestone Township School Distric
New York Media Market
Philadelphia Media Market
Upper Freehold School District

Other Results for Distance-Based Methods

- **Regularity Condition**. $\sup_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[|Y_i(t)|^{2+v}|\mathbf{X}_i = \mathbf{x}] < \infty$ for some $v \geq 2$.

- **Convergence Rates**. Under minimal regularity conditions,

$$|\widehat{\tau}_{\mathrm{dis}}(\mathbf{x}) - \tau(\mathbf{x})| \lesssim_{\mathbb{P}} \frac{1}{\sqrt{nh^2}} + \frac{1}{n^{\frac{1+v}{2+v}} h^2} + |\mathfrak{B}_n(\mathbf{x})|, \qquad \mathbf{x} \in \mathscr{B},$$

  and

$$\sup_{\mathbf{x}\in\mathscr{B}} |\widehat{\tau}_{\mathrm{dis}}(\mathbf{x}) - \tau(\mathbf{x})| \lesssim_{\mathbb{P}} \sqrt{\frac{\log n}{nh^2}} + \frac{\log n}{n^{\frac{1+v}{2+v}} h^2} + \sup_{\mathbf{x}\in\mathscr{B}} |\mathfrak{B}_n(\mathbf{x})|.$$

- **Pointwise Inference**. Ignoring the potential bias problem when $\mathscr{B}$ is non-smooth, paper establishes distribution theory with valid standard errors for each $\mathbf{x} \in \mathscr{B}$. This result is fairly standard, up to handling $\mathscr{B}$.

- **Uniform Inference**. Ignoring the potential bias problem when $\mathscr{B}$ is non-smooth, paper establishes feasible uniform distribution theory via simulation. This result requires new technical tools, and requires careful handling of $\mathscr{B}$. More details later.

- **Practice**. Valid and invalid practices based on standard univariate RD designs methods.

Outline

Location-Based Methods: Setup

- **Parameter**. $\tau(\mathbf{x}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{x}]$ for all $\mathbf{x} \in \mathscr{B}$.

- **Estimator**. $\widehat{\tau}(\mathbf{b}_1) = \mathbf{e}_1^\top \widehat{\boldsymbol{\beta}}_1(\mathbf{b}_1) - \mathbf{e}_1^\top \widehat{\boldsymbol{\beta}}_0(\mathbf{b}_1)$ for $\mathbf{x} \in \mathscr{B}$,

$$\widehat{\boldsymbol{\beta}}_t(\mathbf{x}) = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left(Y_i - \mathbf{R}_p(\mathbf{X}_i - \mathbf{x})^\top \boldsymbol{\beta}\right)^2 K_h\left(\frac{\mathbf{X}_i - \mathbf{x}}{h}\right) \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_t).$$

- **Assumption**. Let $t \in \{0, 1\}$.
  - $K : \mathbb{R} \to [0, \infty)$ compact supported & Lipschitz continuous, or $K(\mathbf{u}) = \mathbb{1}(\mathbf{u} \in [-1, 1]^2)$.
  - $\liminf_{h\downarrow 0} \inf_{\mathbf{x}\in\mathscr{B}} \int_{\mathscr{A}_t} K_h(\mathbf{u} - \mathbf{x}) d\mathbf{u} \gtrsim 1$.

- **Identification**. For all $\mathbf{b} \in \mathscr{B}$,

$$\tau(\mathbf{b}) = \lim_{\mathbf{x}\to\mathbf{b}, \mathbf{x}\in\mathscr{A}_1} \mathbb{E}[Y_i|\mathbf{X}_i = \mathbf{x}] - \lim_{\mathbf{x}\to\mathbf{b}, \mathbf{x}\in\mathscr{A}_0} \mathbb{E}[Y_i|\mathbf{X}_i = \mathbf{x}].$$

  This is standard from the literature.

Point Estimation Results for Location-Based Methods

- **Regularity Condition**. $\sup_{\mathbf{x}\in\mathscr{X}} \mathbb{E}[|Y_i(t)|^{2+v}|\mathbf{X}_i = \mathbf{x}] < \infty$ for some $v \geq 2$.

- **Convergence Rates**. Under minimal regularity conditions,

$$|\widehat{\tau}(\mathbf{x}) - \tau(\mathbf{x})| \lesssim_{\mathbb{P}} |\widehat{\tau}(\mathbf{x}) - \tau(\mathbf{x})| \lesssim_{\mathbb{P}} \frac{1}{\sqrt{nh^2}} + \frac{1}{n^{\frac{1+v}{2+v}} h^2} + h^{p+1}, \qquad \mathbf{x} \in \mathscr{B},$$

  and

$$\sup_{\mathbf{x}\in\mathscr{B}} |\widehat{\tau}(\mathbf{x}) - \tau(\mathbf{x})| \lesssim_{\mathbb{P}} \sqrt{\frac{\log n}{nh^2}} + \frac{\log n}{n^{\frac{1+v}{2+v}} h^2} + h^{p+1}.$$

- **MSE Expansions**. Under minimal regularity conditions,

$$\mathbb{E}[(\widehat{\tau}(\mathbf{x}) - \tau(\mathbf{x}))^2|\mathbf{X}] = h^{2(p+1)}\mathbf{B}_{\mathbf{x}}^2 + \frac{1}{nh^2}\mathbf{V}_{\mathbf{x}} \qquad \mathbf{x} \in \mathscr{B},$$

  and

$$\int_{\mathscr{B}} \mathbb{E}[(\widehat{\tau}(\mathbf{x}) - \tau(\mathbf{x}))^2|\mathbf{X}]w(\mathbf{x})d\mathbf{x} = h^{2(p+1)} \int_{\mathscr{B}} \mathbf{B}_{\mathbf{x}}^2 dw(\mathbf{x}) + \frac{1}{nh^2} \int_{\mathscr{B}} \mathbf{V}_{\mathbf{x}} w(\mathbf{x})d\mathbf{x}$$

  - Standard bandwidth selection methods developed in the paper.

Inference Results for Location-Based Methods

- **Feasible t-test**. Using standard least squares algebra, $\widehat{\mathsf{T}}(\mathbf{x}) = \frac{\widehat{\tau}(\mathbf{x})-\tau(\mathbf{x})}{\sqrt{\widehat{\Omega}_{\mathbf{x},\mathbf{x}}}}$.

- **Uncertainty Quantification**. Confidence intervals and confidence bands,

$$\widehat{I}(\mathbf{x};\alpha) = \Big[\ \widehat{\tau}(\mathbf{x}) - \mathfrak{q}_\alpha\sqrt{\widehat{\Omega}_{\mathbf{x}}}\ ,\ \widehat{\tau}(\mathbf{x}) + \mathfrak{q}_\alpha\sqrt{\widehat{\Omega}_{\mathbf{x}}}\ \Big], \qquad \mathbf{x} \in \mathscr{B},$$

- **Pointwise Inference**. By standard CLT result, for each $\mathbf{x} \in \mathscr{B}$, set $\mathfrak{q}_\alpha = \Phi^{-1}(1-\alpha/2)$.

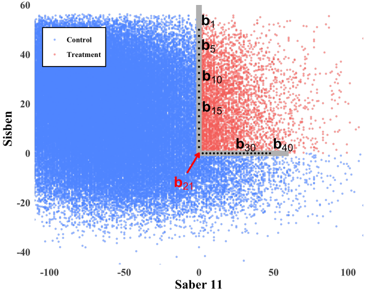
- **Uniform Inference**. Note that

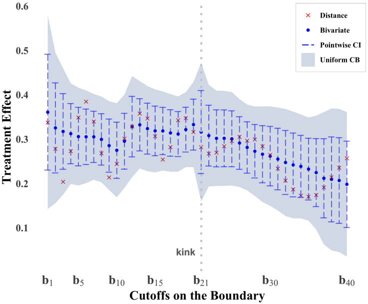
$$\mathbb{P}\Big[\tau(\mathbf{x}) \in \widehat{I}(\mathbf{x};\alpha)\ ,\ \text{for all } \mathbf{x} \in \mathscr{B}\Big] = \mathbb{P}\Big[\sup_{\mathbf{x}\in\mathscr{B}}\big|\widehat{\mathsf{T}}(\mathbf{x})\big| \leq \mathfrak{q}_\alpha\Big].$$

  1. Establish strong approximation for $(\widehat{\mathsf{T}}(\mathbf{x}) : \mathbf{x} \in \mathscr{B})$ by $(\widehat{Z}_n : \mathbf{x} \in \mathscr{B})$, a Gaussian process conditional on data.
  2. Deduce the distribution of $\sup_{\mathbf{x}\in\mathscr{B}}\big|\widehat{\mathsf{T}}(\mathbf{x})\big|$.
  3. Using simulations, set $\mathfrak{q}_\alpha = \inf\{c > 0 : \mathbb{P}[\sup_{\mathbf{x}\in\mathscr{B}}|\widehat{Z}_n(\mathbf{x})| \geq c|\text{data}] \leq \alpha\}$.

- **Implementation and Bias**. (I)MSE-optimal bandwidth selection for point estimation, robust bias correction for inference.

Outline





- *Ser Pilo Paga* (SPP) Colombian policy program; students $i = 1, 2, \dots, n$.
- $\mathbf{X}_i = (\texttt{SABER11}_i, \texttt{SISBEN}_i)^\top$; $\texttt{SABER11}_i$ = exam score and $\texttt{SABER11}_i$ = wealth index.
- $\mathscr{B} = \{\texttt{SABER11} \geq 0 \text{ and } \texttt{SISBEN} = 0\} \cup \{\texttt{SABER11} = 0 \text{ and } \texttt{SISBEN} \geq 0\}$.
- $(Y_i(0), Y_i(1), \mathbf{X}_i)$, $i = 1, 2, \dots, n$, random sample.
- $Y_i = \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_0) \cdot Y_i(0) + \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) \cdot Y_i(1)$; $\mathscr{A}_t$ group t's assignment area.

Outline

Conclusion

- Multi-dimensional RD designs are widely used across disciplines.
- Methodological and formal results lagging behind its popularity in practice.
- We offer a through treatment of Boundary Discontinuity Designs.
  - Distance-based methods may exhibit large bias when $\mathscr{B}$ is non-smooth.
  - Location-based methods do not suffer of this drawback.
  - We develop pointwise and uniform estimation and inference methods.
- rd2d package for `R`.

`https://rdpackages.github.io/`