Boundary Discontinuity Designs: Theory and Practice

Matias D. Cattaneo, Princeton University
Rocio Titiunik, Princeton University
Ruiqi (Rae) Yu, Princeton University

August 2025

Talk based on:

- "Identification, Estimation, and Inference for Boundary Average Treatment Effects"
- "Estimation and Inference in Boundary Discontinuity Designs: Distance-Based Methods"
- "Estimation and Inference in Boundary Discontinuity Designs: Location-Based Methods"
- "rd2d: Causal Inference in Boundary Discontinuity Designs"

Outline

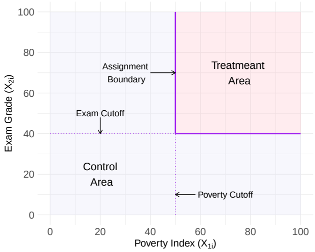
Introduction

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

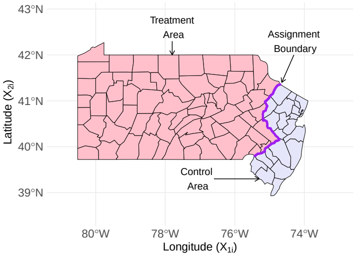
- Multi-dimensional Regression Discontinuity (RD) designs.
  - Multi-score RD designs / Geographic RD designs.
- Three approaches for analysis in practice.
  - Local regression using pooled data near the boundary.
  - Local regression using univariate distance to boundary point.
  - Local regression using bivariate location relative to boundary point.
- Today: foundational, thorough study of Boundary Discontinuity Designs.
  - *Methodology*: guidance on current practices, and more.
  - *Theory*: convergence over manifolds, minimax estimation, and strong approximation.
  - *Practice*: new `R` software (`rd2d` package).

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

Motivation: Basic Setup



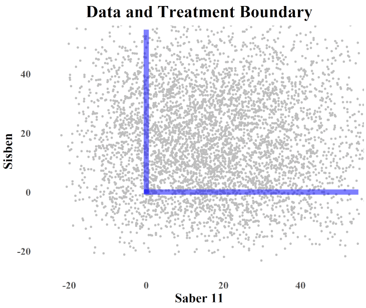
(a) Two-Score RD Design.



(b) Geographic RD Design.

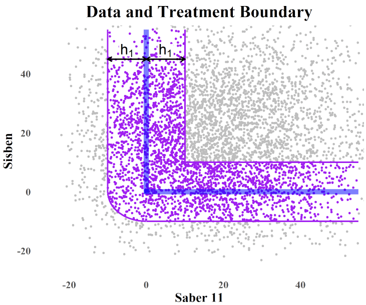
- $\mathscr{B}$ = assignment boundary; $\mathscr{A}_0$ = control assignment; $\mathscr{A}_1$ = treatment assignment.
- $\mathbf{X}_i$ = bivariate score; $\quad Y_i = (1 - T_i) \cdot Y_i(0) + T_i \cdot Y_i(1); \quad T_i = \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1)$.
- $D_i(\mathbf{b}) = (2 \cdot T_i - 1)\mathscr{d}(\mathbf{X}_i, \mathbf{b})$ <u>distance to point</u> $\mathbf{b} \in \mathscr{B}$; $\quad$ e.g., $\mathscr{d}(\mathbf{x}, \mathbf{b}) = \|\mathbf{x} - \mathbf{b}\|$.
- $D_i = (2 \cdot T_i - 1)d(\mathbf{X}_i, \mathscr{B})$ <u>distance to boundary</u> $\mathscr{B}$; $\quad d(\mathbf{x}, \mathscr{B}) = \inf_{\mathbf{b} \in \mathscr{B}} \mathscr{d}(\mathbf{x}, \mathbf{b})$.

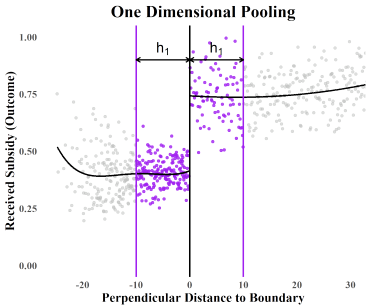
Motivation: *Ser Pilo Paga* (SPP) Colombian Policy Program



- High-school graduates $i = 1, 2, \ldots, n$ offered cash transfer to attend college ($T_i = 1$).
- $\mathbf{X}_i = (\texttt{SABER11}_i, \texttt{SISBEN}_i)^\top$; $\texttt{SABER11}_i$ = exam score; $\texttt{SISBEN}_i$ = wealth index.
- $\mathscr{B} = \{\texttt{SABER11} \geq 0 \text{ and } \texttt{SISBEN} = 0\} \cup \{\texttt{SABER11} = 0 \text{ and } \texttt{SISBEN} \geq 0\}$.
- $Y_i = 1$ if first year of college completed, $= 0$ otherwise.

Motivation: Boundary Average Treatment Effect





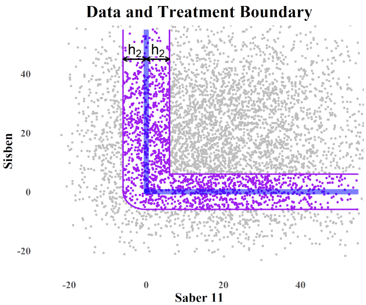
- Basic approach:

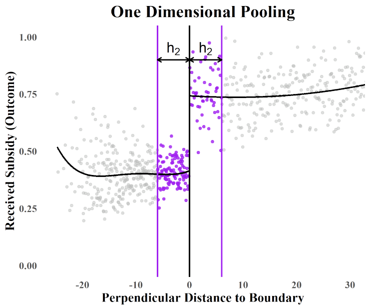
$$(\widehat{\zeta}, \widehat{\tau}) = \arg\min_{\zeta,\tau} \sum_{i=1}^{n} \big(Y_i - \zeta - T_i\tau\big)^2 \mathbb{1}(|D_i| \leq h),$$

$$T_i = \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) = \mathbb{1}(D_i \geq 0)$$

$$D_i = (2 \cdot T_i - 1) d(\mathbf{X}_i, \mathscr{B}), \qquad d(\mathbf{x}, \mathscr{B}) = \inf_{\mathbf{b} \in \mathscr{B}} \mathscr{d}(\mathbf{x}, \mathbf{b})$$

Motivation: Boundary Average Treatment Effect





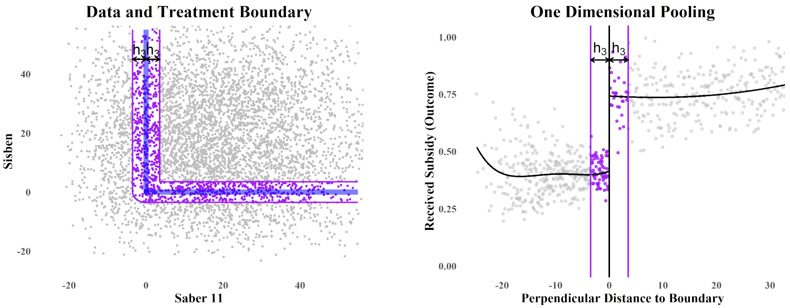
- Basic approach + boundary-segment FE + poly-expansion:

$$(\widehat{\boldsymbol{\zeta}}, \widehat{\tau}, \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \boldsymbol{\iota}(S_i)^\top \boldsymbol{\zeta} - T_i \tau - \mathbf{q}_p(D_i, T_i)^\top \boldsymbol{\beta}\right)^2 \mathbb{1}(|D_i| \leq h),$$

$$\boldsymbol{\iota}_L(S_i) = [\mathbb{1}(S_i = 1), \cdots, \mathbb{1}(S_i = L)]^\top, \qquad S_i = \underset{1\leq \ell \leq L}{\arg\min}\, d(\mathbf{X}_i, \mathscr{B}_\ell), \qquad \mathscr{B} = \sqcup_{1\leq \ell \leq L} \mathscr{B}_\ell$$

$$\mathbf{q}_p(u, t) = [(u, u^2, \cdots, u^p)^\top, t(u, u^2, \cdots, u^p)^\top]^\top$$

Motivation: Boundary Average Treatment Effect

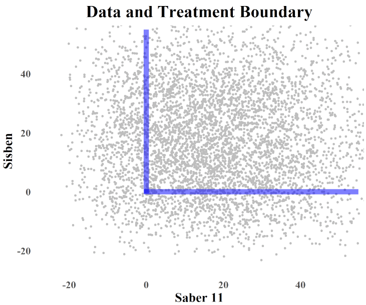


- Identification/Interpretation:

$$\widehat{\tau} \to_{\mathbb{P}} \tau = \frac{\int_{\mathscr{B}} \tau(\mathbf{b}) f(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} f(\mathbf{b}) d\mathbf{b}}, \qquad \tau(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0) | \mathbf{X}_i = \mathbf{b}].$$

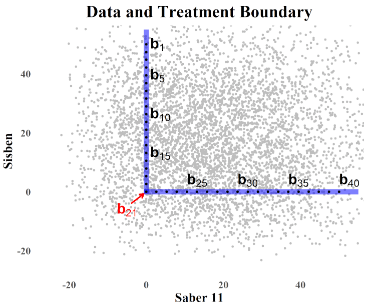
- How to define the above integral/convergence appropriately?
- What conditions on $\mathscr{B}$ and $\mathscr{d}(\cdot)$ are needed/sufficient?
- Estimation and inference results to help practice.

Motivation: *Ser Pilo Paga* (SPP) Colombian Policy Program



- High-school graduates $i = 1, 2, \dots, n$ offered cash transfer to attend college ($T_i = 1$).
- $\mathbf{X}_i = (\texttt{SABER11}_i, \texttt{SISBEN}_i)^\top$; $\texttt{SABER11}_i =$ exam score; $\texttt{SISBEN}_i =$ wealth index.
- $\mathscr{B} = \{\texttt{SABER11} \geq 0 \text{ and } \texttt{SISBEN} = 0\} \cup \{\texttt{SABER11} = 0 \text{ and } \texttt{SISBEN} \geq 0\}$.
- $Y_i = 1$ if first year of college completed, $= 0$ otherwise.

Motivation: Distance-Based vs. Location-Based

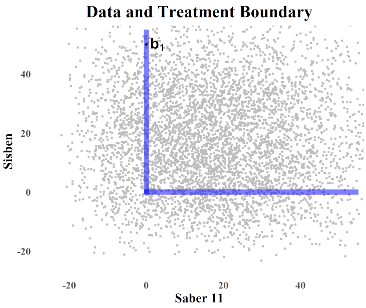


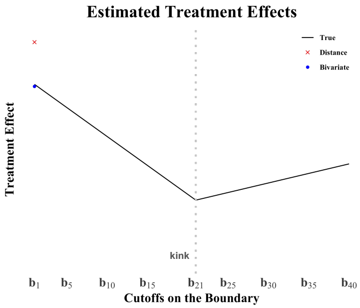
- Average treatment effect curve along the boundary:

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

- Estimation and Inference Approaches:
  - Local regression based on univariate distance to point on boundary.
  - Local regression based on bivariate location relative to point on boundary.

Motivation: Distance-Based vs. Location-Based





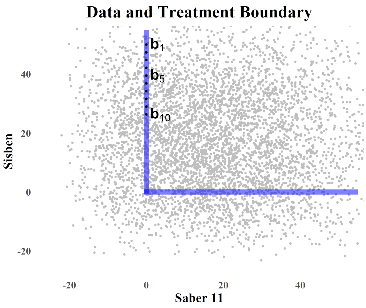
- Distance-based Estimator:

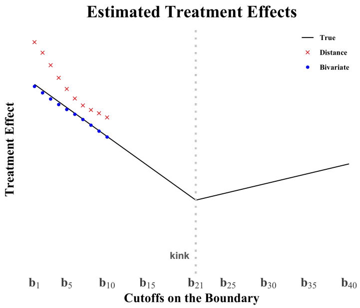
$$(\widehat{\zeta}, \widehat{\tau}_{\text{dis}}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\zeta,\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}_j), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b}_j)}{h}\right).$$

- Location-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\zeta,\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}_j, T_i)^\top \boldsymbol{\beta}\right)^2 K\left(\frac{\mathbf{X}_i - \mathbf{b}_j}{h}\right).$$

Motivation: Distance-Based vs. Location-Based





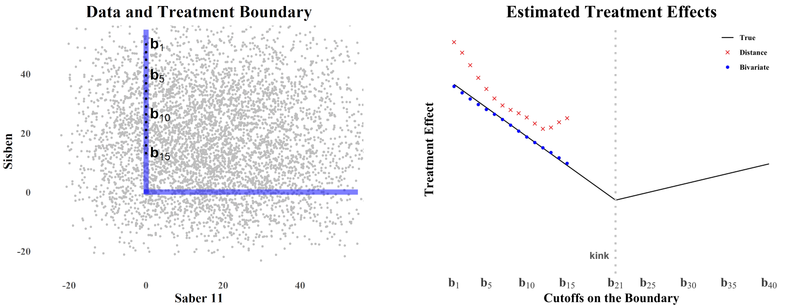
- Distance-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}_{\text{dis}}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}_j), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b}_j)}{h}\right).$$

- Location-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}_j, T_i)^\top \boldsymbol{\beta}\right)^2 K\left(\frac{\mathbf{X}_i - \mathbf{b}_j}{h}\right).$$

Motivation: Distance-Based vs. Location-Based



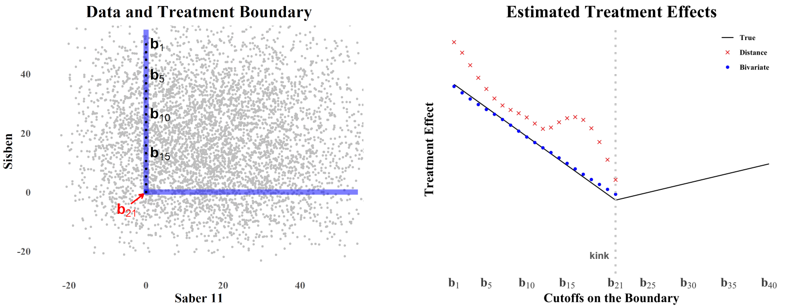
- Distance-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}_{\text{dis}}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}_j), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b}_j)}{h}\right).$$

- Location-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}_j, T_i)^\top \boldsymbol{\beta}\right)^2 K\left(\frac{\mathbf{X}_i - \mathbf{b}_j}{h}\right).$$

Motivation: Distance-Based vs. Location-Based



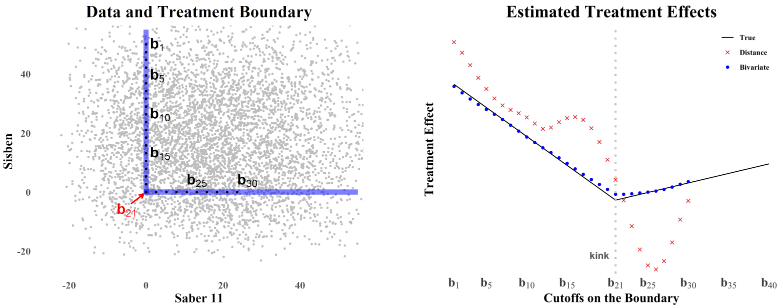
- Distance-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}_{\mathrm{dis}}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\zeta,\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}_j), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b}_j)}{h}\right).$$

- Location-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\zeta,\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}_j, T_i)^\top \boldsymbol{\beta}\right)^2 K\left(\frac{\mathbf{X}_i - \mathbf{b}_j}{h}\right).$$

Motivation: Distance-Based vs. Location-Based



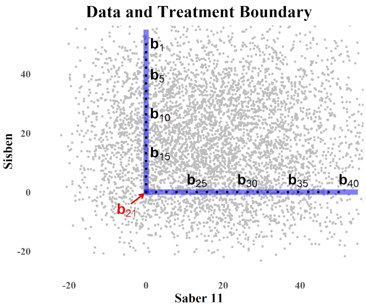
- Distance-based Estimator:

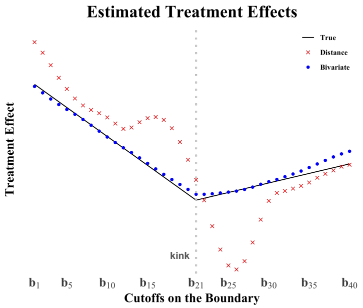
$$(\widehat{\zeta}, \widehat{\tau}_{\text{dis}}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}_j), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b}_j)}{h}\right).$$

- Location-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}_j, T_i)^\top \boldsymbol{\beta}\right)^2 K\left(\frac{\mathbf{X}_i - \mathbf{b}_j}{h}\right).$$

Motivation: Distance-Based vs. Location-Based





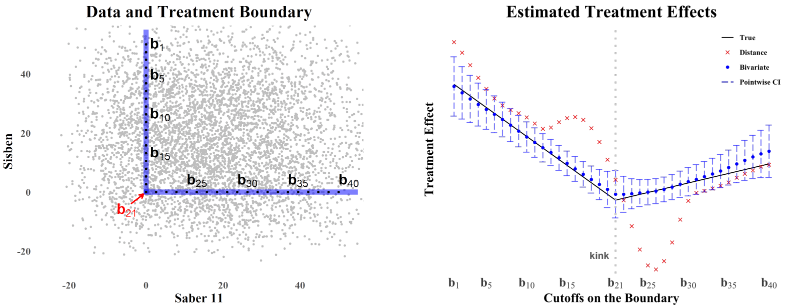
- Distance-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}_{\mathrm{dis}}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \operatorname*{arg\,min}_{\zeta,\tau,\boldsymbol{\beta}} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}_j), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b}_j)}{h}\right).$$

- Location-based Estimator:

$$(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}_j), \widehat{\boldsymbol{\beta}}) = \operatorname*{arg\,min}_{\zeta,\tau,\boldsymbol{\beta}} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}_j, T_i)^\top \boldsymbol{\beta}\right)^2 K\left(\frac{\mathbf{X}_i - \mathbf{b}_j}{h}\right).$$

Motivation: Distance-Based vs. Location-Based

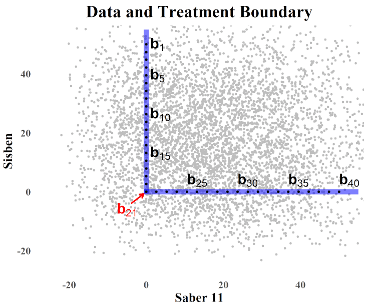


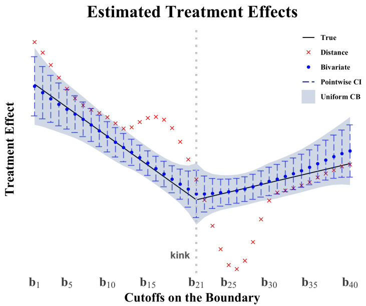
- Estimators: $\widehat{\tau}_{\text{dis}}(\mathbf{x})$ and $\widehat{\tau}(\mathbf{x})$, for each $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{40}\}$.
- Uncertainty Quantification: Confidence Intervals. For each $\mathbf{x} \in \{\mathbf{b}_1, \ldots, \mathbf{b}_{40}\}$,

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

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

Motivation: Distance-Based vs. Location-Based





- 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 $\mathbf{x} \in \mathscr{B}$,

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

  - $\mathfrak{q}_\alpha = \inf\{c > 0 : \mathbb{P}[\sup_{\mathbf{x}\in\mathscr{B}} |\widehat{Z}_n(\mathbf{x})| \geq c|\text{data}] \leq \alpha\}$.
  - $(\widehat{Z}_n : \mathbf{x} \in \mathscr{B})$ is a Gaussian process conditional on data.

Summary of Results

- Pooled Boundary Average Treatment Effects: $\widehat{\tau}$.
  1. Necessary and sufficient conditions for identification.
  2. Sufficient conditions for local polynomial debiasing.
  3. Estimation and inference.
  4. Similarities and differences with standard univariate RD designs.
- Univariate distance-based methods: $\widehat{\tau}_{\mathrm{dis}}(\mathbf{x})$.
  1. Sufficient conditions for identification.
  2. "Large" misspecification bias when $\mathscr{B}$ is non-smooth (e.g., near a kink).
  3. "Small" misspecification bias when $\mathscr{B}$ is smooth.
  4. Pointwise and uniform convergence rates and distribution theory.
  5. Discuss connects and differences with standard univariate RD designs.
- Bivariate location-based methods: $\widehat{\tau}(\mathbf{x})$.
  1. Identification, estimation, and inference (pointwise and uniform over $\mathscr{B}$) are standard.
  2. Additional (mild) regularity on $\mathscr{B}$ is needed.
  3. New methods for analysis of Boundary Discontinuity Designs.
- Tools: convergence over manifolds, minimax estimation, and strong approximation.

Outline

BATE: Technical Setup

- **Goal**: Properly define existence of and converge to

$$\tau_w = \frac{\int_{\mathscr{B}} \tau(\mathbf{b})w(\mathbf{b})d\mathbf{b}}{\int_{\mathscr{B}} w(\mathbf{b})d\mathbf{b}}, \qquad \tau(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{b}].$$

- **Challenge**: Assignment boundary $\mathscr{B}$ can be non-smooth.
  - Recall: $D_i = (2 \cdot T_i - 1)d(\mathbf{X}_i, \mathscr{B})$ distance to boundary $\mathscr{B}$; $\quad d(\mathbf{x}, \mathscr{B}) = \inf_{\mathbf{b} \in \mathscr{B}} \mathscr{d}(\mathbf{x}, \mathbf{b})$.

- **Approach**: Geometric measure theory.
  - $\mathscr{B}$ is a 1-dimensional rectifiable curve.
  - $\int_{\mathscr{B}} m(\mathbf{b})d\mathbf{b} = \int_{\mathscr{B}} m \; d\mathfrak{H}^1$, with $\mathfrak{H}^1$ the 1-dimensional Hausdorff measure.

- **Question**: When does the following limit of integrals is well-defined?

$$\lim_{\epsilon \downarrow 0} \int_{\mathscr{T}(\epsilon)} \frac{1}{\epsilon} g\Big(\frac{d(\mathbf{x}, \mathscr{B})}{\epsilon}\Big) m(\mathbf{x}) d\mathbf{x},$$

where

$$\mathscr{T}(\epsilon) = \big\{\mathbf{x} \in \mathscr{X} : d(\mathbf{x}, \mathscr{B}) \leq \epsilon\big\}, \qquad \epsilon \geq 0.$$

BATE: Workhorse Technical Tool

$$\mathscr{T}(\epsilon) = \{\mathbf{x} \in \mathscr{X} : d(\mathbf{x}, \mathscr{B}) \leq \epsilon\}, \qquad d(\mathbf{x}, \mathscr{B}) = \inf_{\mathbf{b} \in \mathscr{B}} \mathscr{d}(\mathbf{x}, \mathbf{b}).$$

- **Technical Lemma**.
  - $\mathscr{d} : \mathbb{R}^2 \mapsto [0, \infty)$ 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 $\mathbf{x}_1, \mathbf{x}_2 \in \mathscr{X}$.
  - $\mathrm{J}_2 d(\cdot, \mathscr{B}) \neq 0$ $\mathfrak{m}$-almost everywhere on $\mathscr{X}$, and $\int_{\mathscr{X}} |\frac{m(\mathbf{x})}{\mathrm{J}_2 d(\mathbf{x}, \mathscr{B})}| d\mathbf{x} < \infty$.
  - There exists a constant $\mathsf{c}_{\mathscr{B}}$ such that $\lim_{\epsilon \downarrow 0} M(\epsilon) = \mathsf{c}_{\mathscr{B}} \cdot M(0)$ is finite, where

$$M(r) = \int_{d(\mathbf{x}, \mathscr{B}) = r} \frac{m(\mathbf{x})}{\mathrm{J}_2 d(\mathbf{x}, \mathscr{B})} d\mathfrak{H}^1(\mathbf{x}), \qquad r \geq 0.$$

Then,

$$\lim_{\epsilon \downarrow 0} \int_{\mathscr{T}(\epsilon)} \frac{1}{\epsilon} g\Big(\frac{d(\mathbf{x}, \mathscr{B})}{\epsilon}\Big) m(\mathbf{x}) d\mathbf{x} = \mathsf{c}_{\mathscr{B}} \cdot \int_0^1 g(s) ds \cdot \int_{\mathscr{B}} \frac{m(\mathbf{x})}{\mathrm{J}_2 d(\mathbf{x}, \mathscr{B})} d\mathfrak{H}^1(\mathbf{x}).$$

BATE: Workhorse Technical Tool – Proof

- The function $\mathbf{x} \mapsto d(\mathbf{x}, \mathscr{B})$ is C_u-Lipschitz.
- The level sets $\mathscr{L}(\epsilon) = \{\mathbf{x} \in \mathscr{X} : d(\mathbf{x}, \mathscr{B}) = \epsilon\}$, $\epsilon \geq 0$, are 1-dimensional rectifiable sets.
- The function $\mathbf{x} \mapsto \epsilon^{-1} g(\frac{d(\mathbf{x},\mathscr{B})}{\epsilon}) m(\mathbf{x})\, \mathrm{J}_2 d(\mathbf{x}, \mathscr{B})^{-1}$ is $\mathfrak{m}$-summable over $\mathscr{X}$.
- Using the Coarea formula,

$$
\begin{aligned}
\lim_{\epsilon \downarrow 0} \int_{\mathscr{T}(\epsilon)} \frac{1}{\epsilon} g\Big(\frac{d(\mathbf{x}, \mathscr{B})}{\epsilon}\Big) m(\mathbf{x}) d\mathbf{x} &= \lim_{\epsilon \downarrow 0} \int_{\mathscr{T}(\epsilon)} \frac{1}{\epsilon} g\Big(\frac{d(\mathbf{x}, \mathscr{B})}{\epsilon}\Big) m(\mathbf{x}) \frac{1}{\mathrm{J}_2 d(\mathbf{x}, \mathscr{B})} \mathrm{J}_2 d(\mathbf{x}, \mathscr{B}) d\mathbf{x} \\
&= \lim_{\epsilon \downarrow 0} \int_0^{\epsilon} \int_{d(\mathbf{x},\mathscr{B})=r} \frac{1}{\epsilon} g\Big(\frac{r}{\epsilon}\Big) \frac{m(\mathbf{x})}{\mathrm{J}_2 d(\mathbf{x}, \mathscr{B})} d\mathfrak{H}^1(\mathbf{x}) dr \\
&= \lim_{\epsilon \downarrow 0} \int_0^{\epsilon} \frac{1}{\epsilon} g\Big(\frac{r}{\epsilon}\Big) M(r) dr \\
&= \lim_{\epsilon \downarrow 0} \int_0^{1} g(u) M(\epsilon u) du \\
&= \mathsf{c}_{\mathscr{B}} \cdot \int_0^1 g(s) ds \cdot M(0),
\end{aligned}
$$

BATE: Identification, Estimation, and Inference

- Basic approach + poly-expansions of D_i:

$$(\widehat{\zeta}, \widehat{\tau}, \widehat{\boldsymbol{\beta}}) = \underset{\zeta,\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i, T_i)^\top \boldsymbol{\beta}\right)^2 \mathbb{1}(|D_i| \leq h).$$

- **Identification**.

  (i) $r \mapsto \mathbb{E}[Y_i|D_i = r, T_i = t] = \mathbb{E}[Y_i(t)|d(\mathbf{X}_i, \mathscr{B}) = |r|]$ is continuous at 0, for $t = 0, 1$.

  (ii) $\mathrm{J}_2 d(\cdot, \mathscr{B}) \neq 0$ $\mathfrak{m}$-almost everywhere on $\mathscr{X}$, and $\int_{\mathscr{X}} |\frac{f(\mathbf{x})}{\mathrm{J}_2 d(\mathbf{x},\mathscr{B})}| d\mathbf{x} < \infty$.

  (iii) $\lim_{\epsilon \downarrow 0} F(\epsilon) = F(0)$ is finite, where $F(r) = \int_{d(\mathbf{x},\mathscr{B})=r} \frac{f(\mathbf{x})}{\mathrm{J}_2 d(\mathbf{x},\mathscr{B})} d\mathfrak{H}^1(\mathbf{x})$, $r \geq 0$.

  Then,

$$\widehat{\tau} \to_{\mathbb{P}} \tau = \frac{\int_{\mathscr{B}} \tau(\mathbf{b}) f(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} f(\mathbf{b}) d\mathbf{b}}, \qquad \tau(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{b}]$$

- **Convergence Rate and Feasible CLT**.

$$\left|\widehat{\tau} - \tau\right| \lesssim_{\mathbb{P}} \sqrt{\frac{1}{nh}} + |\mathfrak{B}_n|, \qquad \frac{\widehat{\tau} - \mathbb{E}[\widehat{\tau}|\mathbf{D}]}{\sqrt{\widehat{\mathbb{V}}[\widehat{\tau}|\mathbf{D}]}} \rightsquigarrow \mathsf{Normal}(0, 1).$$

- **Valid Inference**. Need to understand bias $\mathfrak{B}_n$.

BATE: Verifying Sufficient Conditions + Extensions

- **Identification: Sufficient conditions**.
  - $\mathscr{d}(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|$, and $\mathscr{B}$ is piecewise linear.
  - $\mu_0(\mathbf{x}) = \mathbb{E}[Y(0)|\mathbf{X}_i = \mathbf{x}]$, $\mu_1(\mathbf{x}) = \mathbb{E}[Y(1)|\mathbf{X}_i = \mathbf{x}]$, and $f(\mathbf{x})$ are continuous on $\mathscr{X}$.

  Then, identification conditions (i)–(iii) hold for small enough $\epsilon > 0$.

- **Estimation and Inference (Bias): Sufficient conditions**.
  - $\mathscr{d}(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|$, and $\mathscr{B}$ is piecewise linear.
  - μ_0, μ_1, and f are s-times continuously differentiable on $\mathscr{X}$.

  Then,

  $$r \mapsto \mathbb{E}[Y_i|D_i = r, T_i = t] = \mathbb{E}[Y_i(t)|\mathscr{d}(\mathbf{X}_i, \mathscr{B}) = |r|]$$

  is s-times continuously differentiable on $[0, \epsilon]$. Therefore,

  $$|\mathfrak{B}_n| \lesssim_{\mathbb{P}} h^{\min\{s,p+1\}}$$

- **Extensions**. Boundary-segment FE, fuzzy (IV) setting, etc.

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**:

$$(\widehat{\zeta}, \widehat{\tau}_{\mathrm{dis}}(\mathbf{b}), \widehat{\boldsymbol{\beta}}) = \underset{\boldsymbol{\zeta},\tau,\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - \zeta - T_i\tau - \mathbf{q}_p(D_i(\mathbf{b}), T_i)^\top \boldsymbol{\beta}\right)^2 k\left(\frac{D_i(\mathbf{b})}{h}\right).$$

  - $D_i(\mathbf{b}) = (2 \cdot T_i - 1)\mathscr{d}(\mathbf{x}, \mathbf{b})$ <u>distance to point</u> $\mathbf{b} \in \mathscr{B}$.

- **Assumptions**: 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 $k(u) = \mathbb{1}(|u| \leq 1)$.
  - $\liminf_{h\downarrow 0} \inf_{\mathbf{x}\in\mathscr{B}} \int_{\mathscr{A}_t} k(\mathscr{d}(\mathbf{u}, \mathbf{x})/h) d\mathbf{u} \gtrsim 1$.

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

$$\begin{aligned}\tau(\mathbf{b}) &= \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{b}] \\ &= \lim_{r\downarrow 0} \mathbb{E}[Y_i|D_i(\mathbf{b}) = r, T_i = t] - \lim_{r\uparrow 0} \mathbb{E}[Y_i|D_i(\mathbf{b}) = r, T_i = t]\end{aligned}$$

Distance-Based Methods: Estimation and Inference Results

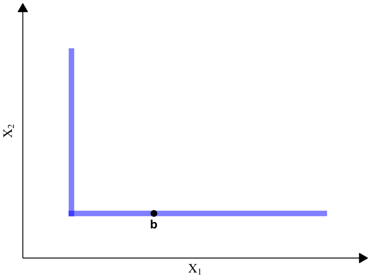
- **Convergence Rates**.

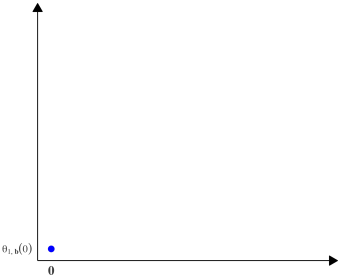
$$\left|\widehat{\tau}_{\mathrm{dis}}(\mathbf{x}) - \tau(\mathbf{x})\right| \lesssim_{\mathbb{P}} \sqrt{\frac{1}{nh^2}} + |\mathfrak{B}_n(\mathbf{x})|, \qquad \mathbf{x} \in \mathscr{B}.$$

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

- Putting aside the potential bias problem when $\mathscr{B}$ is non-smooth.
  - **Pointwise Inference**: Fairly standard, up to handling $\mathscr{B}$.
  - **Uniform Inference**: Requires new technical tools, and careful handling of $\mathscr{B}$.
    - <u>De Giorgi Perimeter</u>: captures notion of "wiggleness" of $\mathscr{B}$.
- **Practice**. Valid and invalid practices based on standard univariate RD methods.

Distance-Based Methods: Implied Smoothness and Bias





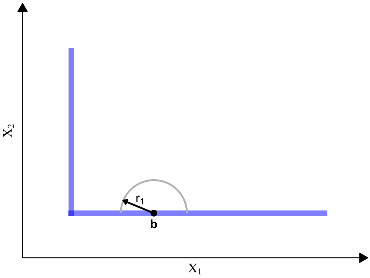
- **Treatment Group**. Bivariate vs. univariate (distance-induced) expectations:

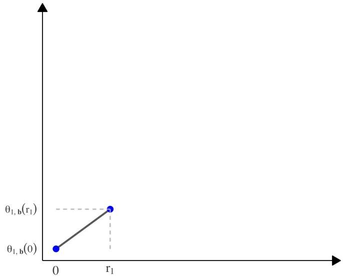
$$\mathbb{E}[Y(1)|\mathbf{X}_i = \mathbf{b}, \mathbf{X}_i \in \mathscr{A}_1] = \lim_{r \downarrow 0} \theta_{1,\mathbf{b}}(r)$$

where

$$\theta_{t,\mathbf{b}}(r) = \mathbb{E}[Y_i | D_i(\mathbf{b}) = r, D_i(\mathbf{b}) \geq 0] = \mathbb{E}[Y_i(1) | \mathscr{d}(\mathbf{X}_i, \mathbf{x}) = r].$$

Distance-Based Methods: Implied Smoothness and Bias





- **Distance range**: $[0, r_1)$. If $\mathbf{x} \mapsto \mathbb{E}[Y(1)|\mathbf{X}_i = \mathbf{b}, \mathbf{X}_i \in \mathscr{A}_1]$ is smooth, then

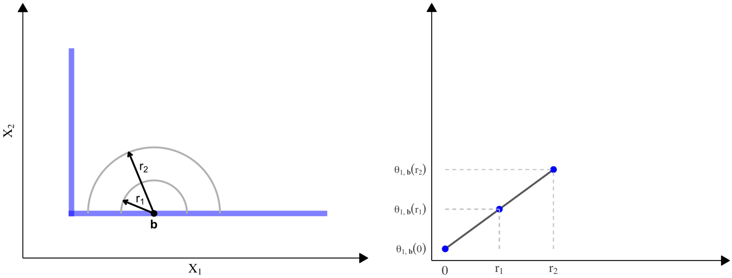
$$r \mapsto \theta_{t,\mathbf{b}}(r) = \mathbb{E}[Y_i|D_i(\mathbf{b}) = r, D_i(\mathbf{b}) \geq 0] = \mathbb{E}[Y_i(1)|\mathscr{d}(\mathbf{X}_i, \mathbf{x}) = r]$$

is also smooth.

- Thus, distance-based local polynomial estimator misspecification bias is

$$\left|\mathfrak{B}_n(\mathbf{x})\right| \lesssim_{\mathbb{P}} h^{p+1}.$$

Distance-Based Methods: Implied Smoothness and Bias



- **Distance range**: $[0, r_2)$. If $\mathbf{x} \mapsto \mathbb{E}[Y(1)|\mathbf{X}_i = \mathbf{b}, \mathbf{X}_i \in \mathscr{A}_1]$ is smooth, then

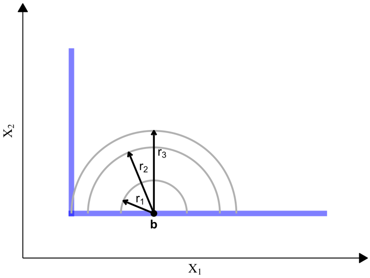
$$r \mapsto \theta_{t,\mathbf{b}}(r) = \mathbb{E}[Y_i|D_i(\mathbf{b}) = r, D_i(\mathbf{b}) \geq 0] = \mathbb{E}[Y_i(1)|\mathscr{d}(\mathbf{X}_i, \mathbf{x}) = r]$$

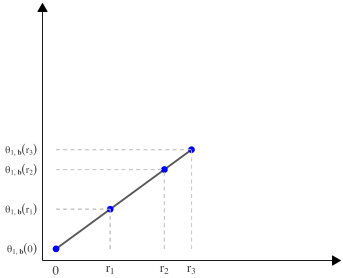
  is also smooth.

- Thus, distance-based local polynomial estimator misspecification bias is

$$\left|\mathfrak{B}_n(\mathbf{x})\right| \lesssim_{\mathbb{P}} h^{p+1}.$$

Distance-Based Methods: Implied Smoothness and Bias





- **Distance range**: $[0, r_3)$. If $\mathbf{x} \mapsto \mathbb{E}[Y(1)|\mathbf{X}_i = \mathbf{b}, \mathbf{X}_i \in \mathscr{A}_1]$ is smooth, then

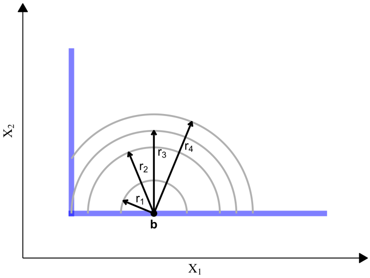
$$r \mapsto \theta_{t,\mathbf{b}}(r) = \mathbb{E}[Y_i|D_i(\mathbf{b}) = r, D_i(\mathbf{b}) \geq 0] = \mathbb{E}[Y_i(1)|\mathscr{d}(\mathbf{X}_i, \mathbf{x}) = r]$$

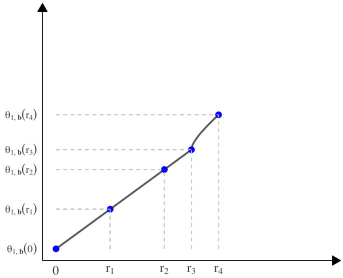
  is also smooth.

- Thus, distance-based local polynomial estimator misspecification bias is

$$\left|\mathfrak{B}_n(\mathbf{x})\right| \lesssim_{\mathbb{P}} h^{p+1}.$$

Distance-Based Methods: Implied Smoothness and Bias





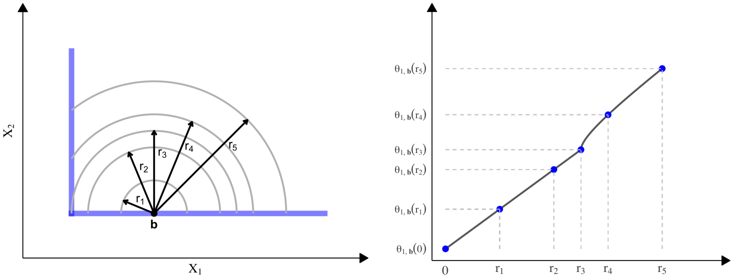
- **Distance range**: $[0, r_4)$. If $\mathbf{x} \mapsto \mathbb{E}[Y(1)|\mathbf{X}_i = \mathbf{b}, \mathbf{X}_i \in \mathscr{A}_1]$ is smooth, then

$$r \mapsto \theta_{t,\mathbf{b}}(r) = \mathbb{E}[Y_i|D_i(\mathbf{b}) = r, D_i(\mathbf{b}) \geq 0] = \mathbb{E}[Y_i(1)|\mathscr{d}(\mathbf{X}_i, \mathbf{x}) = r]$$

  is non-smooth.

- **Smoothness**. $r \mapsto \theta_{t,\mathbf{b}}(r)$ is locally to zero Lipschitz, regardless underlying smoothness.

Distance-Based Methods: Implied Smoothness and Bias



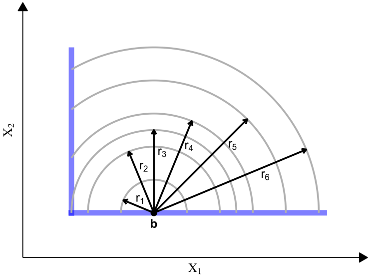
- **Distance range**: $[0, r_5)$. Distance-based local poly estimator misspecification bias is

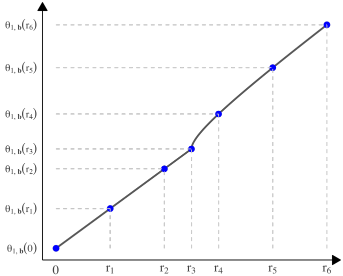
$$\left|\mathfrak{B}_n(\mathbf{x})\right| \lesssim_{\mathbb{P}} h$$

  regardless p used! Not of order h^{p+1} as expected given standard smoothness.

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

Distance-Based Methods: Implied Smoothness and Bias





- **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 $\sup_{\mathbf{x}\in\mathscr{B}} |\mathfrak{B}_n(\mathbf{x})| \lesssim h^{p+1}$.

Distance-Based Methods: Minimax Result

- Is the "large" bias with non-smooth $\mathscr{B}$ a general problem? Yes!

- **Impossibility Result**. Under standard regularity conditions:

$$\liminf_{n\to\infty} n^{1/4} \inf_{T_n \in \mathscr{T}} \sup_{\mathbb{P}\in\mathscr{P}_{\mathtt{NP}}} \mathbb{E}_{\mathbb{P}}\left[\sup_{\mathbf{x}\in\mathscr{B}} \left|T_n(\mathbf{U}_n(\mathbf{x})) - \mu(\mathbf{x})\right|\right] \gtrsim 1,$$

  where

  - $\mathscr{T}$ denotes the class of all distance-based estimators $T_n(\mathbf{U}_n(\mathbf{x}))$ with $\mathbf{U}_n(\mathbf{x}) = [(Y_i, D_i(\mathbf{x}) = \|\mathbf{X}_i - \mathbf{x}\|) : 1 \leq i \leq n]$ for each $\mathbf{x} \in \mathscr{X}$,
  - $\mathscr{B}$ is assumed to be rectifiable, and
  - $\mathscr{P}_{\mathtt{NP}}$ includes q-smooth $\mu(\mathbf{x})$ functions.

- **Stone (1982)**. Under the same conditions:

$$\liminf_{n\to\infty} \left(\frac{n}{\log n}\right)^{\frac{q}{2q+2}} \inf_{S_n \in \mathcal{S}} \sup_{\mathbb{P}\in\mathscr{P}_{\mathtt{NP}}} \mathbb{E}_{\mathbb{P}}\left[\sup_{\mathbf{x}\in\mathscr{B}} \left|S_n(\mathbf{x}; \mathbf{W}_n) - \mu(\mathbf{x})\right|\right] \gtrsim 1,$$

  where

  - $\mathcal{S}$ is the (unrestricted class) of all estimators based on $(\mathbf{W}_n = (Y_i, \mathbf{X}_i^\top)^\top : 1 \leq i \leq n)$.

Outline

1. Introduction

2. Boundary Average Treatment Effects

3. Distance-Based Methods

4. Location-Based Methods

5. Empirical Application

6. Conclusion

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**:

$$\big(\widehat{\zeta}, \widehat{\tau}(\mathbf{b}), \widehat{\boldsymbol{\beta}}\big) = \underset{\boldsymbol{\zeta}, \tau, \boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \big(Y_i - \zeta - T_i \tau - \mathbf{q}_p(\mathbf{X}_i - \mathbf{b}, T_i)^\top \boldsymbol{\beta}\big)^2 K\Big(\frac{\mathbf{X}_i - \mathbf{b}}{h}\Big).$$

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

Location-Based Methods: Point Estimation

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

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

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

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

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

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

  - Standard bandwidth selection methods developed in the paper.

Location-Based Methods: Inference

- **Uncertainty Quantification.**

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

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

- **Confidence Band.**

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

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. Deduce the distribution of $\sup_{\mathbf{x}\in\mathscr{B}} \big|\widehat{\mathsf{T}}(\mathbf{x})\big|$.
2. Using simulations, set $\mathfrak{q}_\alpha = \inf\{c > 0 : \mathbb{P}[\sup_{\mathbf{x}\in\mathscr{B}} |\widehat{Z}_n(\mathbf{x})| \ge c|\text{data}] \le \alpha\}$.
3. Key regularity condition: De Giorgi Perimeter of $\mathscr{B}$.

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

Boundary Treatment Effects

- **Recall: Pooling Approach.**

$$\widehat{\tau} \to_{\mathbb{P}} \tau = \frac{\int_{\mathscr{B}} \tau(\mathbf{b}) f(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} f(\mathbf{b}) d\mathbf{b}}, \qquad \tau(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0) | \mathbf{X}_i = \mathbf{b}].$$

- **Location-Based Methods.**
  - Boundary Average Treatment Effects:

$$\widehat{\tau}_{\mathrm{loc},\mathscr{B}} = \frac{\int_{\mathscr{B}} \widehat{\tau}(\mathbf{b}) w(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} w(\mathbf{b}) d\mathbf{b}} \approx \frac{\sum_{j=1}^{J} \widehat{\tau}(\mathbf{b}_j) w(\mathbf{b}_j)}{\sum_{j=1}^{J} w(\mathbf{b}_j)} \to_{\mathbb{P}} \tau = \frac{\int_{\mathscr{B}} \tau(\mathbf{b}) w(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} w(\mathbf{b}) d\mathbf{b}}.$$

    - IMSE-optimal bandwidth choice is more natural.
    - Choice of $w(\cdot)$ changes causal interpretation.
    - Convergence rate may change from $\frac{1}{nh^2}$ to $\frac{1}{nh}$; distribution theory follows.
    - Natural connection with pooled OLS analysis (for a specific choice of $w(\cdot)$).

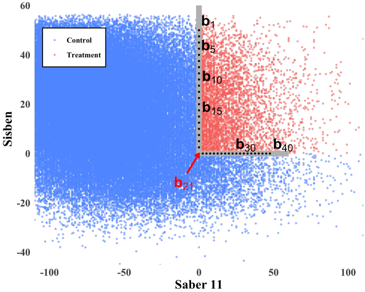
  - Boundary Largest Treatment Effect:

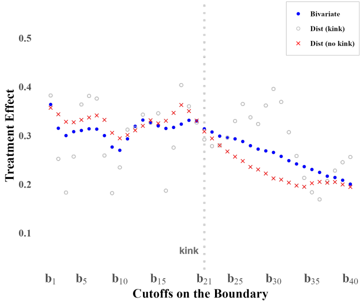
$$\widehat{\tau}_{\mathrm{loc},\max} = \sup_{\mathbf{b} \in \mathscr{B}} \widehat{\tau}(\mathbf{b}) \approx \max_{j=1,\ldots,J} \widehat{\tau}(\mathbf{b}_j) \to_{\mathbb{P}} \tau_{\max} = \sup_{\mathbf{b} \in \mathscr{B}} \tau(\mathbf{b}).$$

    - Inference based on strong approximations and related methods.

Outline

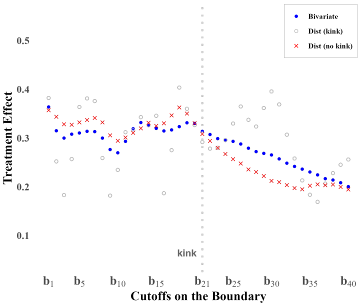
Ser Pilo Paga (SPP) Colombian Policy Program

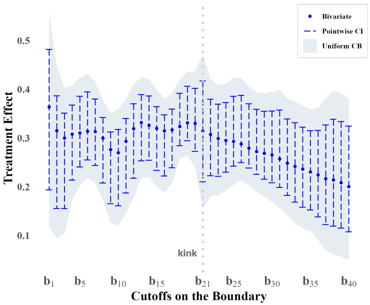


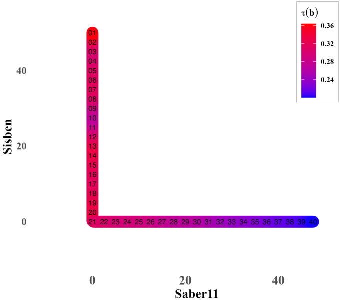


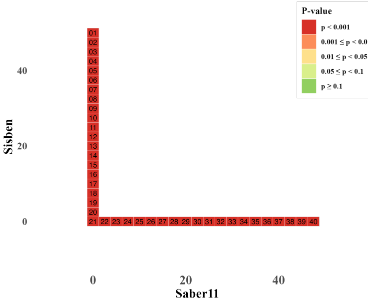
- High-school graduates $i = 1, 2, \ldots, n$ offered cash transfer to attend college ($T_i = 1$).
- $\mathbf{X}_i = (\texttt{SABER11}_i, \texttt{SISBEN}_i)^\top$; $\texttt{SABER11}_i$ = exam score; $\texttt{SISBEN}_i$ = wealth index.
- $\mathscr{B} = \{\texttt{SABER11} \geq 0 \text{ and } \texttt{SISBEN} = 0\} \cup \{\texttt{SABER11} = 0 \text{ and } \texttt{SISBEN} \geq 0\}$.
- $Y_i = 1$ if first year of college completed, $= 0$ otherwise.

Heterogeneous Treatment Effects Along the Boundary









Boundary Average Treatment Effects

$$\tau = \frac{\int_{\mathscr{B}} \tau(\mathbf{b}) f(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} f(\mathbf{b}) d\mathbf{b}}, \qquad \tau(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0) | \mathbf{X}_i = \mathbf{b}].$$

Method	h	Estimate	p-value	CI
$\widehat{\tau}$	5	0.381	0.000	(0.3669, 0.3957)
	10	0.406	0.000	(0.3952, 0.4171)
	15	0.419	0.000	(0.4095, 0.4287)
	20	0.433	0.000	(0.4239, 0.4416)
$\widehat{\tau}_{\text{loc},\mathscr{B}}$	5	0.278	0.000	(0.1748, 0.3370)
	10	0.294	0.000	(0.2425, 0.3282)
	15	0.295	0.000	(0.2601, 0.3248)
	20	0.302	0.000	(0.2636, 0.3180)

$$\text{Pooled Approach :} \quad (\widehat{\zeta}, \widehat{\tau}) = \underset{\zeta,\tau}{\arg\min} \sum_{i=1}^{n} \big(Y_i - \zeta - T_i \tau\big)^2 \mathbb{1}(|D_i| \leq h).$$

$$\text{Location-Based Approach :} \quad \widehat{\tau}_{\text{loc},\mathscr{B}} = \frac{\int_{\mathscr{B}} \widehat{\tau}(\mathbf{b}) w(\mathbf{b}) d\mathbf{b}}{\int_{\mathscr{B}} w(\mathbf{b}) d\mathbf{b}} \approx \frac{\sum_{j=1}^{J} \widehat{\tau}(\mathbf{b}_j) w(\mathbf{b}_j)}{\sum_{j=1}^{J} w(\mathbf{b}_j)}.$$

Outline

Conclusion

- Multi-dimensional RD designs are widely used across disciplines.
- Methodological and formal results lagging behind its popularity in practice.
- Through treatment of Boundary Discontinuity Designs:
  - Boundary average treatment effects using pooled data near $\mathscr{B}$.
  - Distance-based methods to each point on $\mathscr{B}$.
  - Location-based methods to each point on $\mathscr{B}$.
  - Pointwise and uniform estimation and inference methods.
  - Aggregation of heterogeneous treatment effects along boundary.
- rd2d package for `R`.

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