Analysis and interpretation of multidimensional regression discontinuity designs

Rocío Titiunik

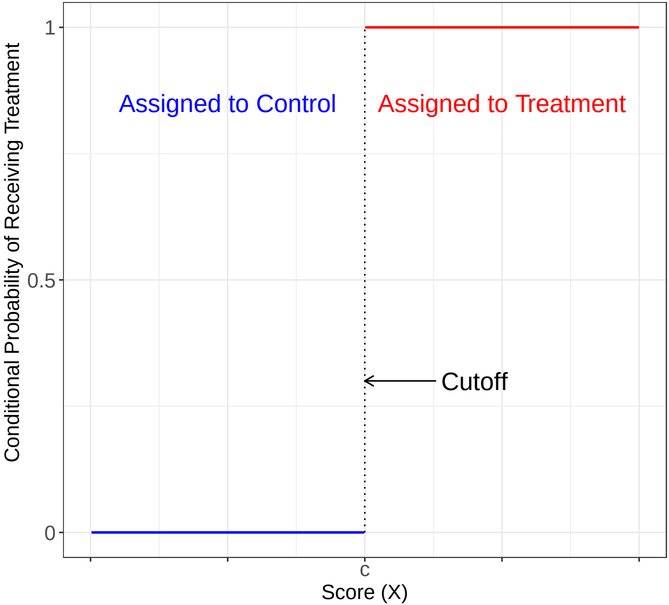
Professor
Department of Politics
Princeton University

Joint with Matias Cattaneo, Luke Keele and Gonzalo Vazquez-Bare

Prepared for 2022 Stata Economics Virtual Symposium

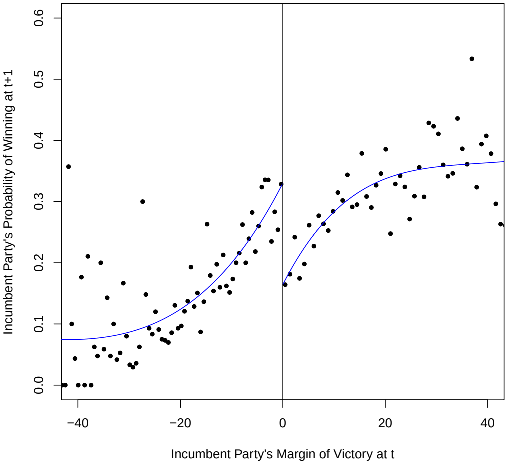
November 3, 2022

RD Design: score, cutoff, treatment



RD Example: Incumbency Effects in Brazil

- Effect of party winning election t on victory at election $t + 1$
- Mayor elections in Brazil, 1996-2012 (first past the post)
- Score: party's margin of victory at election t
- Cutoff: zero
- Outcome: victory at election $t + 1$



Standard RD Framework: Basics

- n units, $i = 1, 2, \ldots, n$
- Score is X_i, treatment is $D_i = \mathbb{1}(X_i \geq x_0)$, with cutoff x_0
- Potential outcomes

  Y_{1i}: outcome under treatment

  Y_{0i}: outcome under control

  $\tau_i = Y_{1i} - Y_{0i}$: individual "treatment effect"

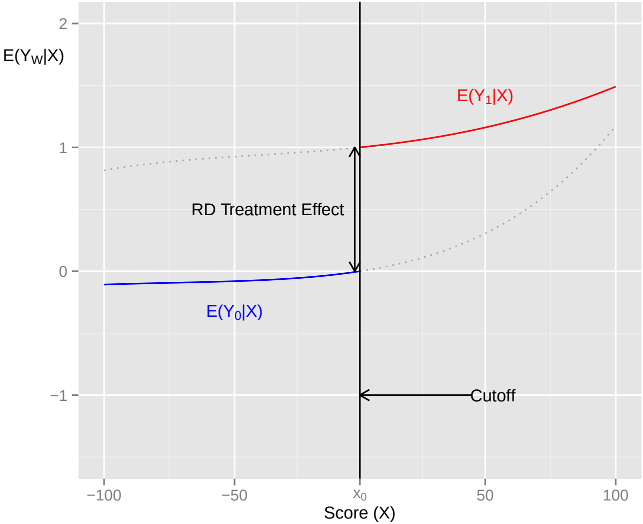
- Observed outcome (Fundamental problem of causal inference)

$$Y_i = \begin{cases} Y_{0i} & \text{if } X_i < x_0, \\ Y_{1i} & \text{if } X_i \geq x_0. \end{cases}$$

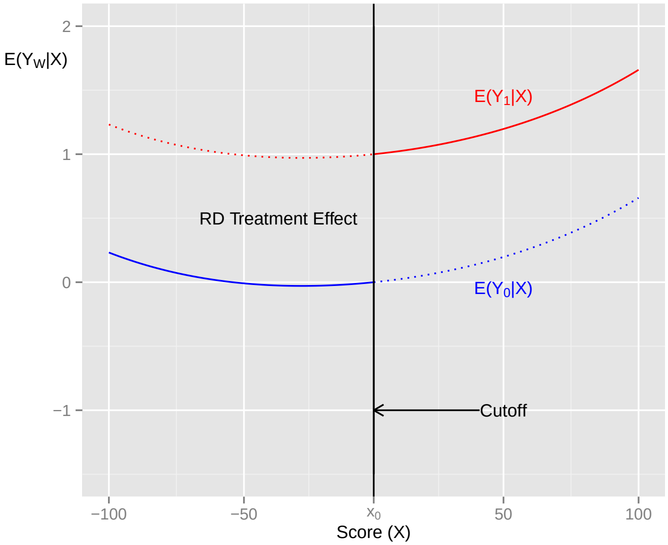
- Under smoothness,

$$\mathbb{E}[\tau_i \mid X_i = x_0] = \lim_{x \downarrow x_0} \mathbb{E}[Y_i \mid X_i = x] - \lim_{x \uparrow x_0} \mathbb{E}[Y_i \mid X_i = x]$$

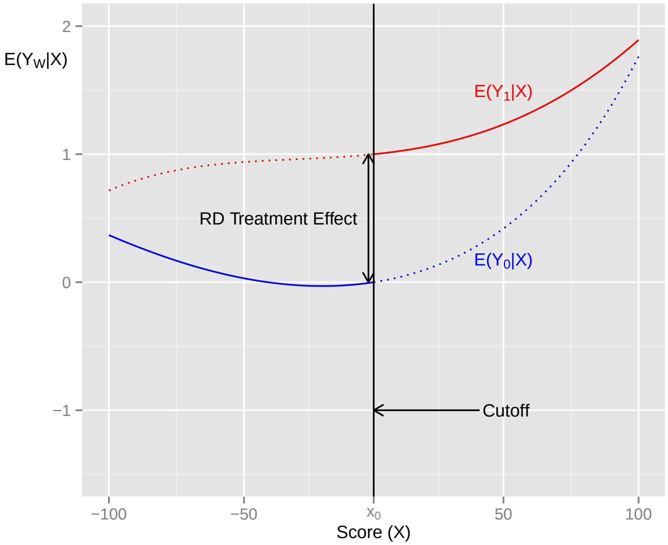
$$\tau = \underbrace{\mathbb{E}[Y_{1i} - Y_{0i} | X_i = x_0]}_{\text{Unobservable}} = \underbrace{\lim_{x \downarrow x_0} \mathbb{E}[Y_i | X_i = x]}_{\color{red}{\text{Observable}}} - \underbrace{\lim_{x \uparrow x_0} \mathbb{E}[Y_i | X_i = x]}_{\color{blue}{\text{Observable}}}$$



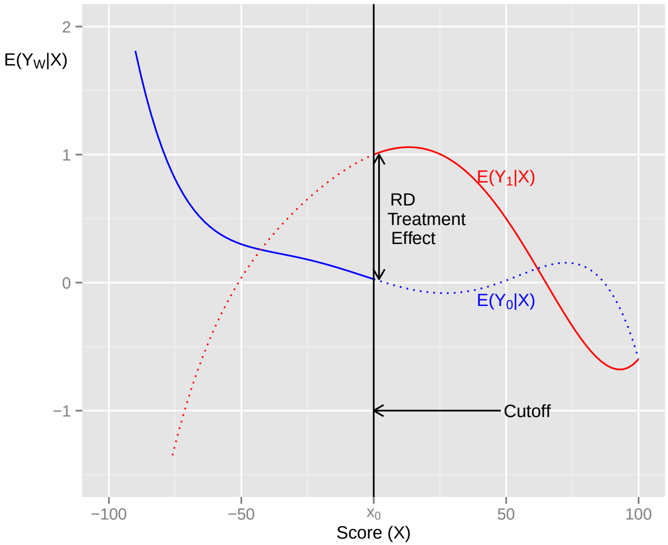
The RD Parameter: No Heterogeneity



The RD Parameter: Mild Heterogeneity



The RD Parameter: Wild Heterogeneity



RD Design with Multiple Dimensions

Multi-dimensional RD designs: treatment is assigned on the basis of more than one score and/or more than one cutoff.

Multi-Cutoff RD design:

- Treatment assigned on the basis of single score, but different groups of units face different cutoff values.
- Example: Mexican conditional cash transfer program Progresa based eligibility on poverty index; in rural areas, seven different cutoffs per geographic region.

Multi-Score RD design:

- Treatment assigned on the basis of two or more scores, often all scores simultaneously must exceed their respective cutoffs.
- Example: in education, scholarships given to students who score above a given cutoff in both a mathematics and a language exam.

Multi-Dimensional RD Design: Ser Pilo Paga (SPP)

Londoño-Vélez, Rodríguez, and Sánchez (2020) study Ser Pilo Paga (SPP), a governmental subsidy for post-secondary education in Colombia.

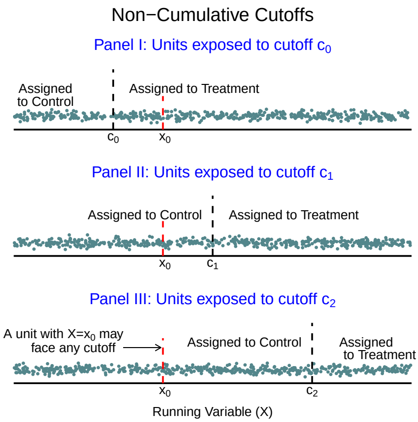
- Treatment: funding of full tuition of a 4-year or 5-year undergraduate program in any government-certified higher education institution (HEI)
- Assignment: eligibility depends on both merit and economic need:
  - Students must obtain a high grade in Colombia's national standardized high school exit exam, SABER 11 (top 9 percent of scores), and
  - they must also come from economically disadvantaged families, measured by a survey-based wealth index, SISBEN (below a region-specific threshold).
- Sample: students who took the SABER 11 test in the fall of 2014 (first cohort of beneficiaries of SPP).
- Ignore non-compliance, focus on intention-to-treat effects.

Multi-Cutoff RD: SPP

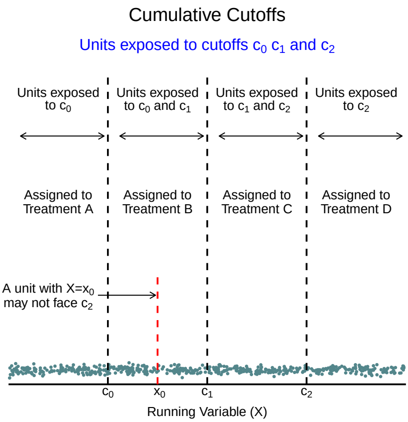
Subpopulation	Cutoff	Sample Size	Min X_i	Max X_i
Area 1 (14 metropolitan areas)	57.21	11,238	.98	83.15
Area 2 (other urban areas)	56.32	10,053	1.78	91.91
Area 3 (rural areas)	40.75	1,841	2.89	84.23

Note: Sample size is number of students in each area facing a unique cutoff. X_i is the SISBEN wealth score. Sample includes only students with SABER 11 above the cutoff and non-missing SISBEN.

Cumulative versus Non-cumulative Cutoffs



(a) Non-cumulative Cutoffs



(b) Cumulative Cutoffs

Multicutoff RD Setup

- Unit's score is X_i
- Cutoff is discrete random variable C_i

  $\mathbb{P}[C_i = c] = p_c \in [0,1]$ for $c \in \{c_1, c_2, ..., c_J\}$

  $f_{X|C}(x|c)$ is the conditional density of $X_i|C_i = c$

- Treatment is $D_i = \mathbb{1}(X_i \geq C_i)$
- Outcomes

  Potential $Y_i(1,c)$, $Y_i(0,c)$ for $c \in \mathcal{C}$

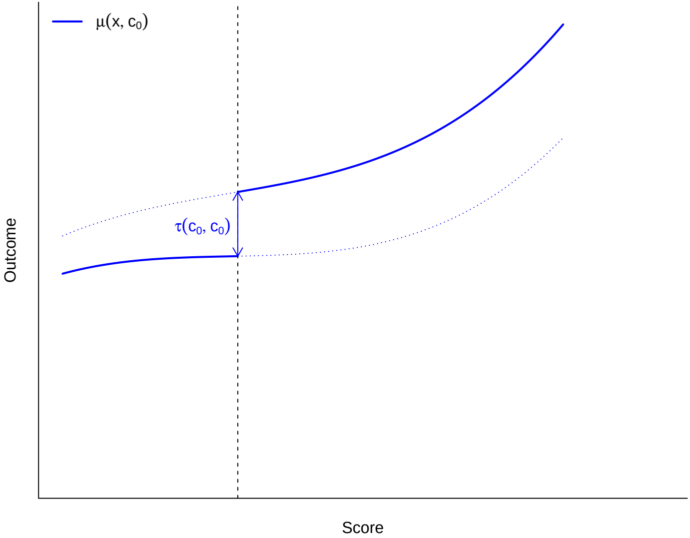
  Observed $Y_i = Y_i(1, C_i)D_i + Y_i(0, C_i)(1 - D_i)$

- Cutoffs may affect potential outcomes directly

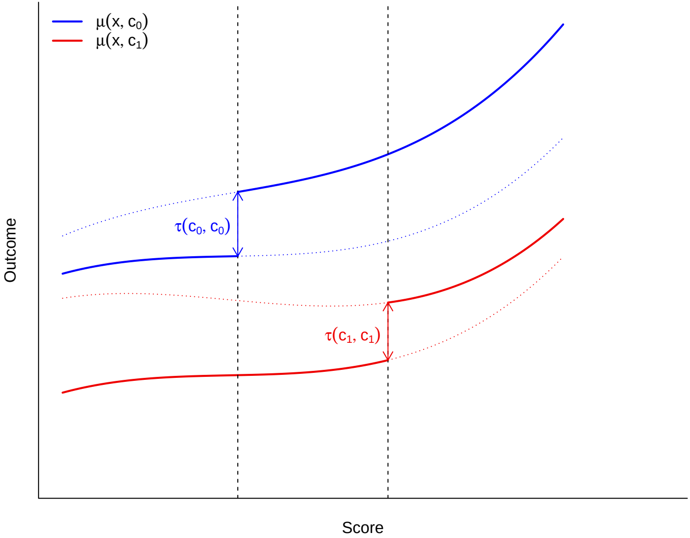
Multi-Cutoff RD Parameters

- Cutoff-specific effects
- Normalizing and pooling effect
- Far-from-cutoff effects

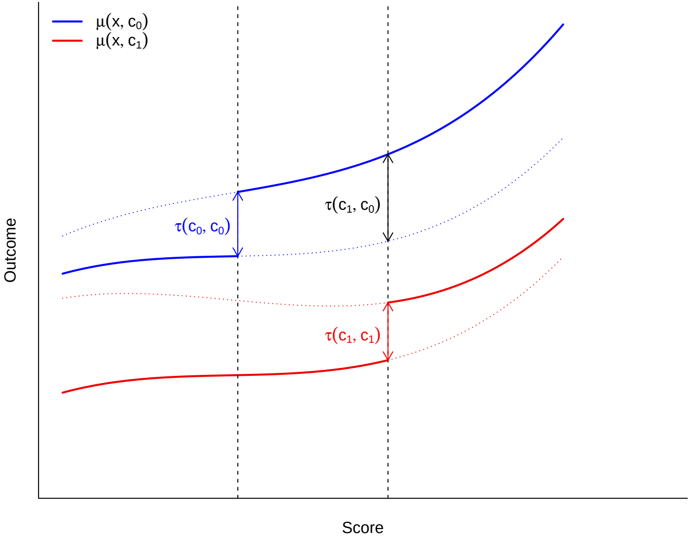
Exploiting multiple cutoffs: parameters of interest



Exploiting multiple cutoffs: parameters of interest



Exploiting multiple cutoffs: parameters of interest



Multi-Cutoff RD Parameters

- Cutoff-specific effects
- Normalizing and pooling effect

} Characterize heterogeneity

- Far-from-cutoff effects

} Extrapolation

Helpful to assess the external validity of RD parameters

Multi-Cutoff RD: Cutoff-specific Effects

- When cutoffs are non-cumulative, the cutoff-specific effects are defined in the same way as the single effect in the standard one-dimensional RD design,

$$\tau_{\texttt{SRD}}(c) \equiv \mathbb{E}[Y_i(1,c) - Y_i(0,c)|X_i = c]$$

  for $c \in \mathcal{C}$.

- Interpretation analogous to standard single-cutoff RD design.
- Because each $\tau_{\texttt{SRD}}(c)$ focuses on subpopulation exposed to c, a cutoff-specific analysis allows researchers to explore heterogeneity of treatment effect across the subpopulations exposed to different cutoffs.

Multi-Cutoff RD: Normalizing and Pooling Effect

- Normalize score $\tilde{X}_i := X_i - C_i$, single cutoff $\tilde{X}_i = 0$
- Treat as single-cutoff RD, $D_i = \mathbb{1}(\tilde{X}_i \geq 0)$
- Example: score is party's margin of victory, cutoff is zero

- The RD pooled estimand is

$$\tau_{pool} = \lim_{x \downarrow 0} \mathbb{E}[Y_i \mid \tilde{X}_i = x] - \lim_{x \uparrow 0} \mathbb{E}[Y_i \mid \tilde{X}_i = x]$$

- What parameter is this approach identifying?

Pooled Estimand: Identification

If the CEFs and $f_{X|C}(x|c)$ are continuous at the cutoffs,

$$\tau_{pool} = \sum_{c \in \mathcal{C}} \underbrace{\mathbb{E}[Y_{1i}(c) - Y_{0i}(c) \mid X_i = c, C_i = c]}_{\text{Cutoff-specific effect}} \cdot \underbrace{\frac{f_{X|C}(c|c)\mathbb{P}[C_i = c]}{\sum_{c \in \mathcal{C}} f_{X|C}(c|c)\mathbb{P}[C_i = c]}}_{\text{Weight}}$$

- Two components
  - Average treatment effect when score and cutoff equal take same value c
  - Weight determines how much each effect contributes to τ_{pool}

Multi-Cutoff RD Analysis: SPP

R Snippet 5.1

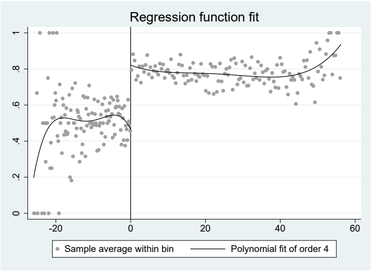
```
> out <- rdmc(data$spadies_any, data$sisben_score, data$cutoff)

Cutoff-specific RD estimation with robust bias-corrected inference
================================================================================
Cutoff     Coef.   P-value          95% CI          hl        hr       Nh   Weight
================================================================================
-57.210     0.346   0.000     0.269     0.452     5.083     5.083    2495    0.384
-56.320     0.203   0.000     0.112     0.282     10.605    10.605   3471    0.534
-40.750     0.209   0.112     -0.042    0.408     8.790     8.790     531    0.082
--------------------------------------------------------------------------------
Weighted    0.259   0.000     0.198     0.319         .         .    6497        .
Pooled      0.269   0.000     0.221     0.328     9.041     9.041    7785        .
================================================================================
```

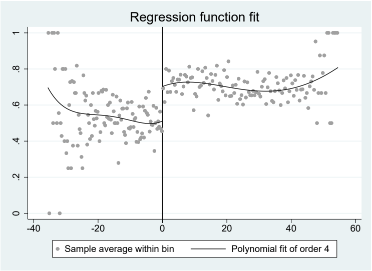
The normalizing-and-pooling parameter weights cutoff-specific effects using $\omega(c) = \mathbb{P}[C_i = c|\tilde{X}_i = 0]$, estimated for bandwidth $h > 0$ as

$$\hat{w}(c) = \hat{\mathbb{P}}(C_i = c|\tilde{X}_i = 0) = \frac{\sum_i \mathbb{1}(C_i = c, -h < \tilde{X}_i < h)}{\sum_i \mathbb{1}(-h < \tilde{X}_i < h)}.$$

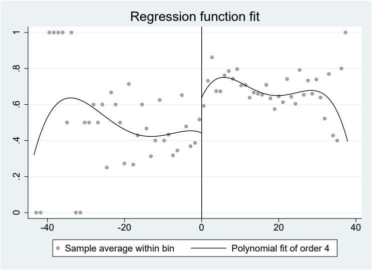
Multi-Cutoff RD Effects: SPP



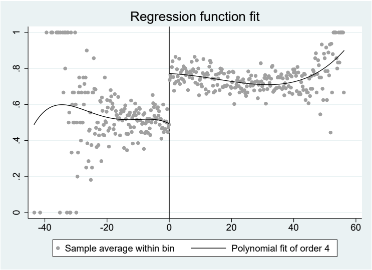
(a) Cutoff 57.210



(b) Cutoff 56.320



(c) Cutoff 40.750



(d) Normalizing-and-pooling

Multi-Cutoff RD Design for Extrapolation

Multi-Cutoff RD Design for Extrapolation

Difference between TE across subgroups

- Consider two cutoffs $c_0 < c_1$.
- For a given value of X_i, difference in ATEs has two components:
  - Direct effect: impact of moving a person from one cutoff to the other one.
  - Indirect effect: switching cutoffs shifts distribution of individual characteristics.
- SPP example:
  - Treatment is subsidy, score is SISBEN wealth, cutoff differs across regions, lower wealth cutoff in rural (40) than urban (57) regions.
  - Direct effect: subsidy received in rural areas where SISBEN wealth cutoff is 40 may have larger effect if poorer households face more severe credit constraints
  - Indirect effect: rural areas may have higher proportion of high school graduates who go to farming instead of college

Difference between TE across subgroups

- Formally:

$$\begin{aligned}
\tau(c_1, c_0) - \tau(c_1, c_1) &= \mathbb{E}[\tau_i | X_i = c_1, C_i = c_0] - \mathbb{E}[\tau_i | X_i = c_1, C_i = c_1] \\
&= \int \underbrace{[\tau(c_1, c_0, u) - \tau(c_1, c_1, u)]}_{\text{direct effect}} f_{U|X,C}(u|c_1, c_0) d\mu \\
&\quad + \int \tau(c_1, c_1, u) \underbrace{[f_{U|X,C}(u|c_1, c_0) - f_{U|X,C}(u|c_1, c_1)]}_{\text{indirect effect}} d\mu
\end{aligned}$$

Exploiting multiple cutoffs

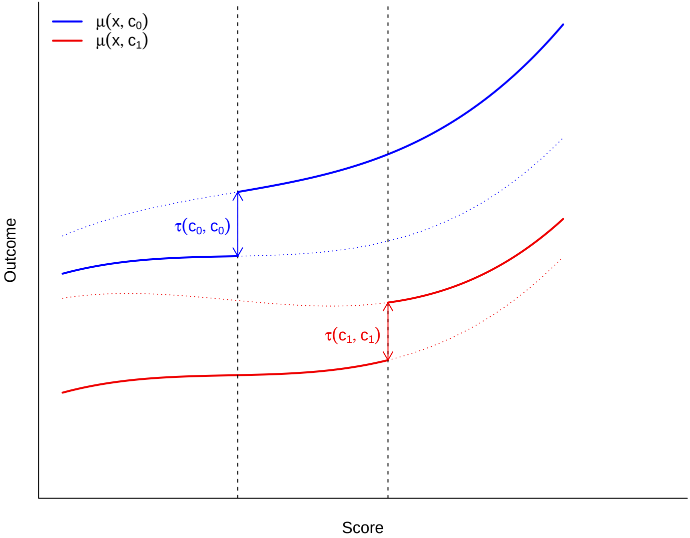
- What are the parameters of interest in this context?
- Potential CEFs:

$$\mu_d(x, c) := \mathbb{E}[Y_{di}|X_i = x, C_i = c], \qquad d \in \{0, 1\}$$

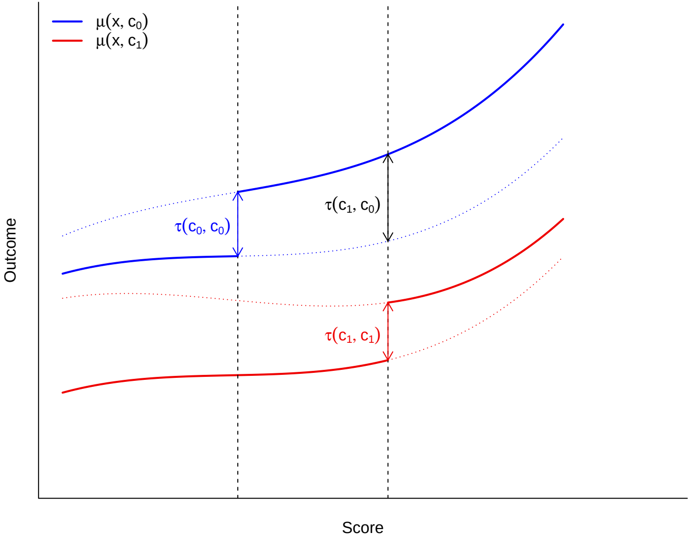
- (Conditional) ATE:

$$\tau(x, c) := \mathbb{E}[\tau_i \mid X_i = x, C_i = c] = \mu_1(x, c) - \mu_0(x, c)$$

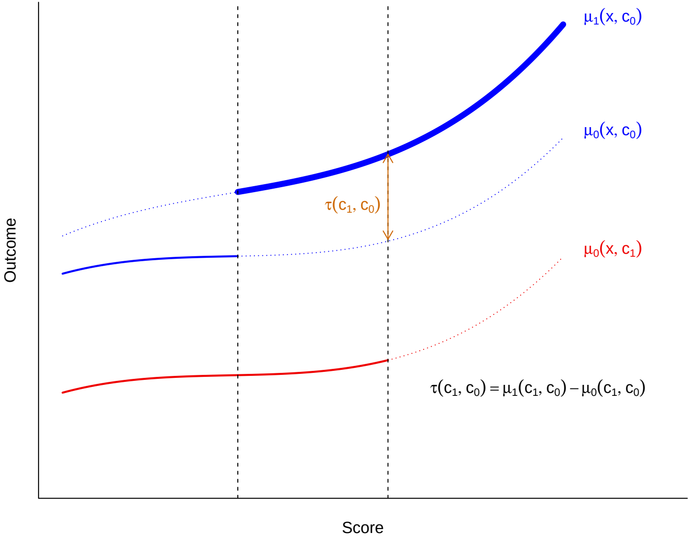
Multi-Cutoff RD Extrapolation: Two cutoffs



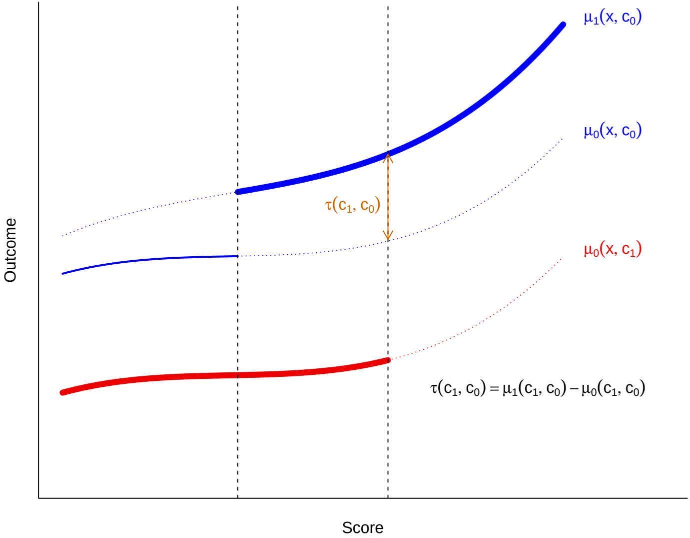
Multi-Cutoff RD Extrapolation: Two cutoffs



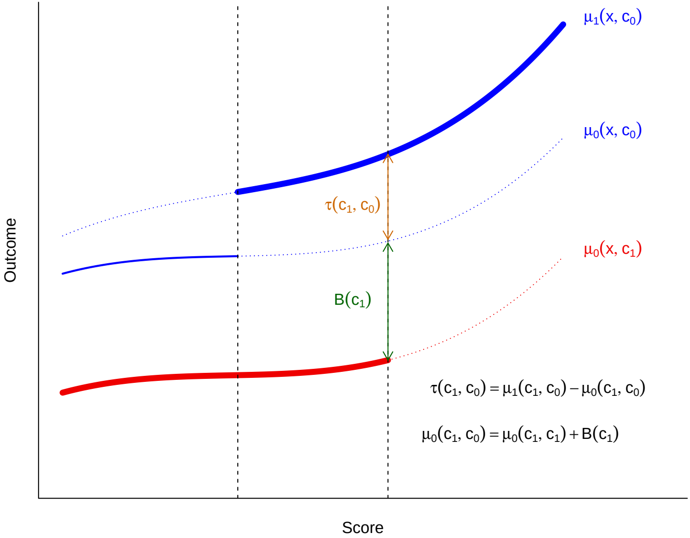
Multi-Cutoff RD Extrapolation: Two cutoffs



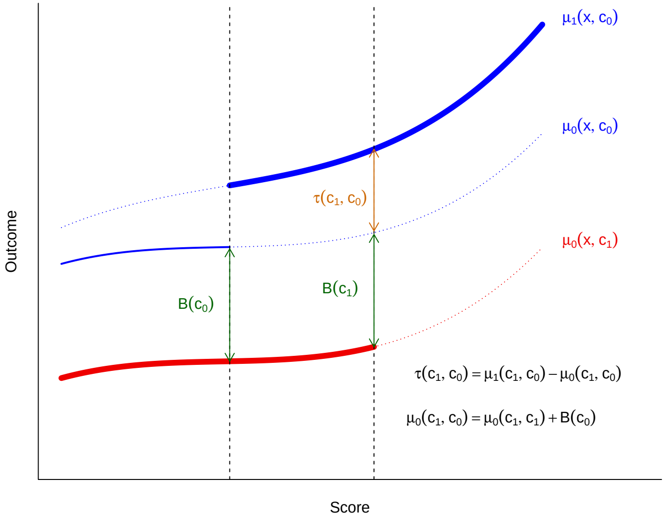
Multi-Cutoff RD Extrapolation: Two cutoffs



Multi-Cutoff RD Extrapolation: Two cutoffs



Multi-Cutoff RD Extrapolation: Two cutoffs



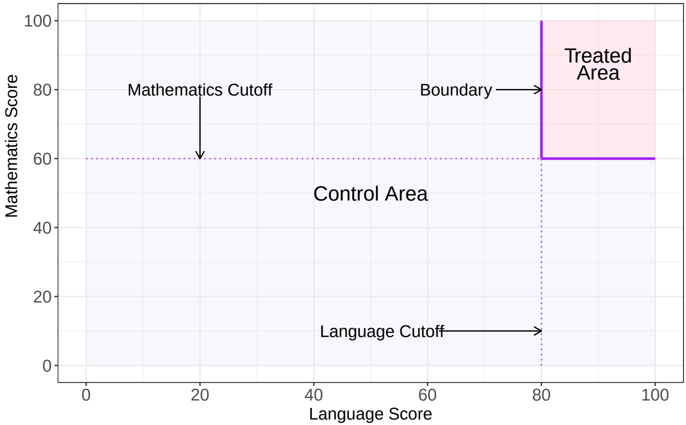
Go to conclusion. | Go to empirical example.

RD Design with Multiple Scores

- Each unit's score is a vector denoted by $\mathbf{X}_i = (X_{1i}, X_{2i})$.
- Treatment assignment is $T_i = T(\mathbf{X}_i)$.
- Common assignment rule is to require both scores above a cutoff, leading to $T(\mathbf{X}_i) = \mathbb{1}(X_{1i} > b_1) \cdot \mathbb{1}(X_{2i} > b_2)$ where b_1 and b_2 denote the cutoff points along each of the two dimensions.
- Assume potential outcome functions are $Y_i(1)$ and $Y_i(0)$ (e.g., no spill-overs in a geographic setting).

RD Design with Multiple Scores

Figure: Example of RD Design With Multiple Scores: Treated and Control Areas



RD Design with Multiple Scores

Parameters of interest:

- Point-specific effects.
- Normalizing and pooling effect.

RD Design with Multiple Scores: Point-specific effects

- Generalization of standard Sharp RD parameter,

  $$\tau_{\texttt{SRD}}(\mathbf{b}) = \mathbb{E}[Y_i(1) - Y_i(0) | \mathbf{X}_i = \mathbf{b}], \qquad \mathbf{b} \in \mathscr{B},$$

  where

  $\mathscr{A}_t = \{(x_1, x_2) : T_i(\mathbf{X}_i) = 1\}$ treated area

  $\mathscr{A}_c = \{(x_1, x_2) : T_i(\mathbf{X}_i) = 0\}$ control area

  $\mathscr{B} = \{(x_1, x_2) : (x_1, x_2) \in (\texttt{bd}(\mathscr{A}_t) \cap \texttt{bd}(\mathscr{A}_c))\}$, with $\texttt{bd}(B) \equiv \texttt{cl}(B) \setminus \texttt{int}(B)$

- In the example,
  $\mathscr{B} = \{(x_1, x_2) : (x_1 \geq 80 \text{ and } x_2 = 60) \text{ or } (x_1 = 80 \text{ and } x_2 \geq 60)\}$.

RD Design with Multiple Scores: Point-specific effects

Identification of Multi-Score RD effect analogous to single score case,

$$\tau_{\texttt{SRD}}(\mathbf{b}) = \lim_{\mathbf{x}\to\mathbf{b};\mathbf{x}\in\mathcal{A}_t} \mathbb{E}[Y_i|\mathbf{X}_i = \mathbf{x}] - \lim_{\mathbf{x}\to\mathbf{b};\mathbf{x}\in\mathcal{A}_c} \mathbb{E}[Y_i|\mathbf{X}_i = \mathbf{x}], \qquad \mathbf{b} \in \mathcal{B},$$

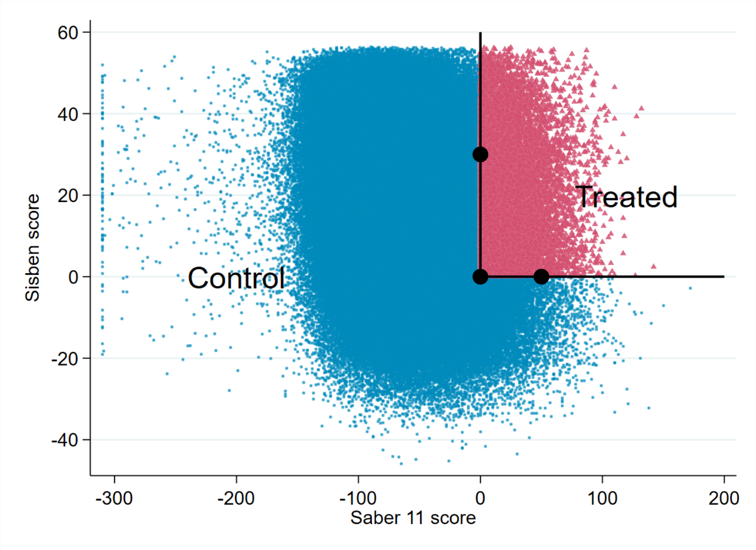
- Treatment effect at every point **b** along the boundary identifiable by observed bivariate regression functions for treated and control groups.
- Multi-Score RD designs generate a family or curve of treatment effects $\tau_{\texttt{SRD}}(\mathbf{b})$, one for each boundary point $\mathbf{b} \in \mathcal{B}$.
- For example, $\tau_{\texttt{SRD}}(80, 70)$ and $\tau_{\texttt{SRD}}(90, 60)$.

RD Design with Multiple Scores: Normalizing and pooling effect

Define running variable the shortest distance to boundary, then pooling all observations in one-dimensional RD analysis.

- Choose a distance metric, $d_i(\cdot)$.
- Using $d_i(\cdot)$, calculate for each i the shortest distance between i's score $\mathbf{X}_i$ and the boundary, denoted $d_{i\mathcal{B}}$.
- Define $\tilde{d}_{i\mathcal{B}} = d_{i\mathcal{B}}(\mathbf{b})T(X_{1i}, X_{2i}) - d_{i\mathcal{B}}(\mathbf{b})(1 - T(X_{1i}, X_{2i}))$ for all i.
- Implement one-dimensional RD analysis pooling all observations, using $\tilde{d}_{i\mathcal{B}}$ as running variable and zero as cutoff.

RD Design with Multiple Scores
SPP assignment



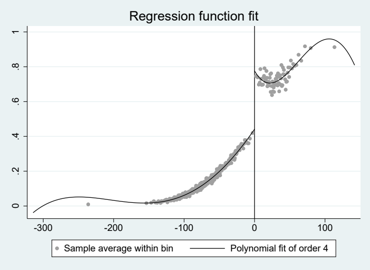
RD Design with Multiple Scores
SPP effects

R Snippet 5.5

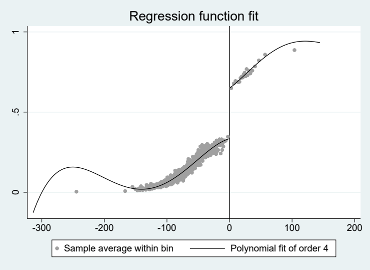
```
> cvec <- c(0, 30, 0)
> cvec2 <- c(0, 0, 50)
> out <- rdms(Y = data$spadies_any, X = data$running_sisben, X2 = data$running_saber11,
+ zvar = data$tr, C = cvec, C2 = cvec2)

================================================================================
Cutoff            Coef.   P-value          95% CI          hl       hr        Nh
================================================================================
(0.00,0.00)       0.323   0.000      0.293     0.379     30.701   30.701    41771
(30.00,0.00)       0.315   0.000      0.286     0.356     42.582   42.582    71579
(0.00,50.00)       0.229   0.000      0.144     0.351     27.762   27.762    5057
================================================================================
```

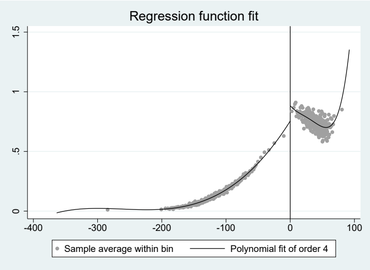
Multi-Score RD Effects: SPP



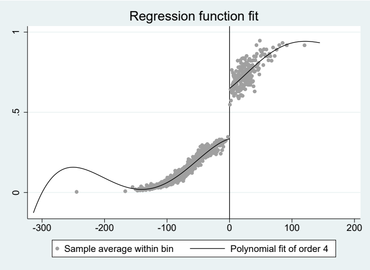
(a) Point (0,0)



(b) Point (0,30)



(c) Point (50,0)



(d) Normalizing-and-pooling

RD Design with Multiple Scores: Ongoing

- Estimation: $\widehat{\tau}_{\texttt{SRD}}(\mathbf{b})$ can be constructed two ways
  - Two-dimensional local polynomial of Y on each coordinate separately: $X_1 - b_1$, $X_2 - b_2$, $(X_1 - b_1)^2$, $(X_2 - b_2)^2$,...
  - One-dimensional local polynomial of Y on $d_i(\mathbf{b})$.

    Study rates of convergence of each case

- Inference:
  - Use strong approximations to make inferences about treatment effect curve $\tau_{\texttt{SRD}}(\mathbf{b})$

Concluding Remarks

- RD designs are observational studies: we are not in control of treatment assignment
- Must take threats to internal validity seriously
- But also threats to external validity: the identifiable RD parameter not decided by us
- Multiple dimensional RD designs allow us to explore heterogeneity and (under additional assumptions) study far-from-cutoff effects
- Both help with bolstering external validity of RD findings

Thanks!

My webpage

https://scholar.princeton.edu/titiunik

RD software at

https://rdpackages.github.io/

RD Software Packages

https://rdpackages.github.io/

- **rdrobust**: estimation, inference and graphical presentation using local polynomials, partitioning, and spacings estimators.
  - `rdrobust`, `rdbwselect`, `rdplot`.
- **rddensity**: discontinuity in density test at cutoff (a.k.a. manipulation testing) using novel local polynomial density estimator.
  - `rddensity`, `rdbwdensity`.
- **rdmulti**: RD plots, estimation, inference, and extrapolation with multiple cutoffs and scores.
  - `rdmc`, `rdmcplot`, `rdms`.
- **rdpower**: power calculations and survey/sample design.
  - `rdpower`, `rdsampsi`.
- **rdlocrand**: covariate balance, binomial tests, randomization inference methods (window selection & inference).
  - `rdrandinf`, `rdwinselect`, `rdsensitivity`, `rdrbounds`.

For Further Details

- Multi-Cutoff RD designs
  - Cattaneo, Keele, Titiunik, Vazquez-Bare, 2016, JOP.
  - Cattaneo, Keele, Titiunik, Vazquez-Bare, 2021, JASA.
  - Cattaneo, Idrobo, Titiunik, 2023, CUP Elements.
- RD Reviews:
  - Cattaneo, Idrobo, Titiunik, 2020, CUP Elements.
  - Cattaneo and Titiunik, 2022, Annual Review of Economics.

Thanks!

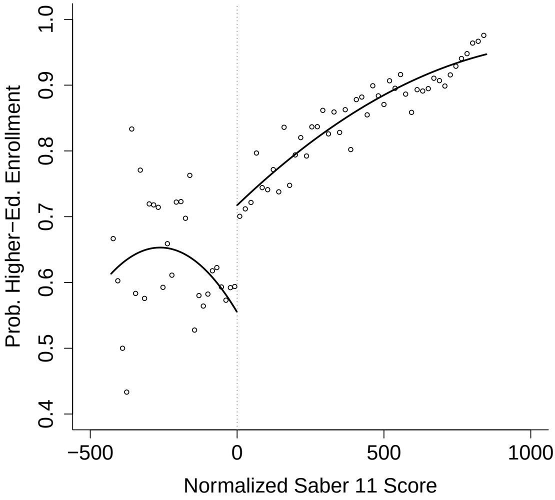
Effect of Access to Credit on Higher Education

- ACCES program in Colombia, which provides long-term credit to underprivileged populations to cover tuition of various post-secondary education programs such as technical or university degrees
- Eligibility for ACCES credit depends on scores in the Saber 11 exam
  - A mandatory exam for all students who wish to enter post-secondary education
  - Each semester of every year, the 1,000-quantiles of the Saber 11 score are calculated among all students who took the exam that semester. Students receive a score between 1 and 1,000 according to their position in the distribution (we call them Saber 11 position scores).
  - For example, a student whose Saber 11 score is between the top 0.1% and 0.2% of the distribution in that year and semester, receives a position score of 2.

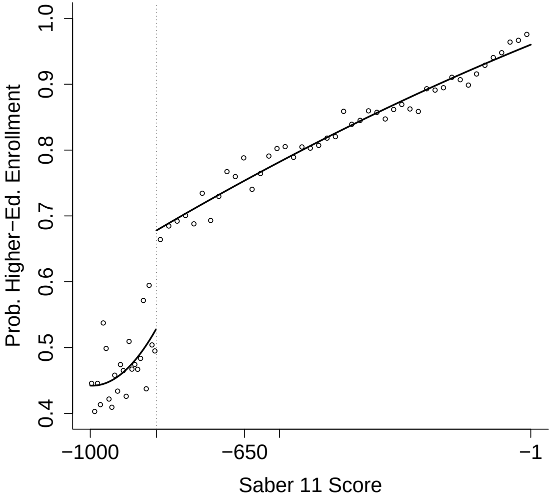
Effect of Access to Credit on Higher Education

- Eligibility for ACCES credit depends on scores in the Saber 11 exam, creating a RD design
  - Running variable: Saber 11 position scores
  - Treatment: Eligibility to receive ACESS credit
  - Outcome: Enrolling in a higher education program
  - Cutoff: 850 in 2002-2008, varies by department starting in 2009

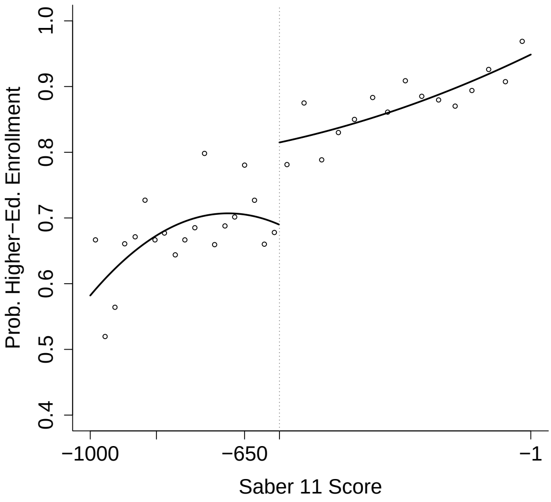
Effect of Access to Credit on Higher Education: Normalized and pooled effect



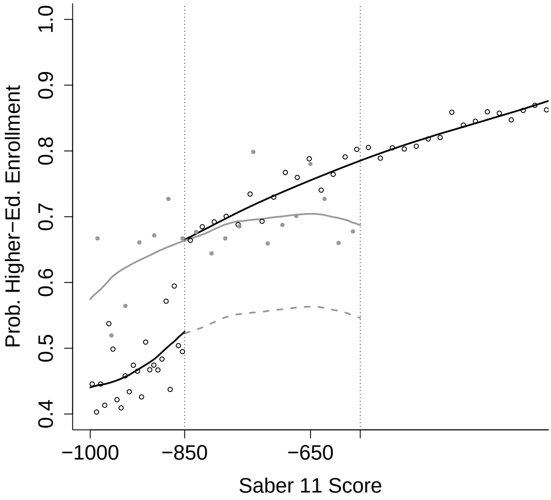
Effect of Access to Credit on Higher Education: Effect at cutoff -850



Effect of Access to Credit on Higher Education: Effect at cutoff -571



Effect of Access to Credit on Higher Education: Extrapolated Effect at cutoff -650



RD and Extrapolation Effects of ACCES Loan Eligibility on Higher Education Enrollment

	Estimate	Bw	Eff. N	Robust BC Inference p-value	Robust BC Inference 95% CI
RD effects					
$C = -850$	0.137	72.9	145	0.007	[0.036 , 0.231]
$C = -571$	0.170	135.4	208	0.101	[-0.038 , 0.429]
Pooled	0.125	145.5	514	0.028	[0.012 , 0.22]
Naive difference					
$\mu_\ell(-650)$	0.755	303.4	504		
$\mu_h(-650)$	0.706	137.4	208		
Difference	0.049			0.172	[-0.019 , 0.105]
Bias					
$\mu_\ell(-850)$	0.525	54.9	54		
$\mu_h(-850)$	0.666	149.5	237		
Difference	-0.141			0.004	[-0.273 , -0.053]
Extrapolation					
$\tau_\ell(-650)$	0.190			0.001	[0.079 , 0.334]

Note: estimates obtained using local linear regression with MSE-optimal bandwidth and robust bias-corrected p-values and confidence intervals.

Multi-Cutoff RD Extrapolation: Formalization

Definition (Cutoff Selection Bias)

For $c, c' \in \mathcal{C}$, let $B(x, c, c') = \mu_{0,c}(x) - \mu_{0,c'}(x)$. There is bias from exposure to different cutoffs if $B(x, c, c') \neq 0$ for some $c, c' \in \mathcal{C}$, $c \neq c'$ and for some $x \in \mathcal{X}$.

Multi-Cutoff RD Extrapolation: Formalization

Assumptions

- Standard continuity assumptions on the relevant regression functions

$$
\begin{aligned}
\mu_{0,c}(c) &= \lim_{\varepsilon \uparrow 0} \mu_c(c+\varepsilon) \quad \text{for } c \in \mathcal{C} = \{\ell, \hbar\} \\
\mu_{1,c}(c) &= \lim_{\varepsilon \downarrow 0} \mu_c(c+\varepsilon) \quad \text{for } c \in \mathcal{C} = \{\ell, \hbar\} \\
\mu_{0,\hbar}(x) &= \mu_{\hbar}(x) \quad \text{for all } x \in (\ell, \hbar) \\
\mu_{1,\ell}(x) &= \mu_{\ell}(x) \quad \text{for all } x \in (\ell, \hbar).
\end{aligned}
$$

- Main extrapolation assumption

Assumption (Constant Bias)

$B(\ell) = B(x)$ *for all* $x \in (\ell, \hbar)$.

Multi-Cutoff RD Extrapolation: Formalization

- The bias at the low cutoff ℓ can be written as

$$B(\ell) = \lim_{\varepsilon \uparrow 0} \mu_{\ell}(\ell + \varepsilon) - \mu_{\hbar}(\ell).$$

- Under constant bias assumption, we have

$$\mu_{0,\ell}(\bar{x}) = \mu_{\hbar}(\bar{x}) + B(\ell),$$

average control response for ℓ subpopulation equal to average observed response for $\hbar$ subpopulation, plus difference in average control responses between both subpopulations at low cutoff ℓ. This leads to our main identification result.

Multi-Cutoff RD Extrapolation: Formalization

Theorem (Extrapolation)

Under constant bias assumption and standard continuity assumptions in sharp RD designs, $\tau_\ell(\bar{x})$ is identifiable by

$$\tau_\ell(\bar{x}) = \mu_\ell(\bar{x}) - [\mu_h(\bar{x}) + B(\ell)],$$

for any point $\bar{x} \in (\ell, h)$.

Extensions

- Generalization of constant-bias assumption:

$$B(c_1) \approx B(c_0) + \sum_{s=1}^{p} \frac{1}{s!} B^{(s)}(c_0) \cdot [c_1 - c_0]^s$$

$\rightarrow$ account for differences in slopes, curvature, etc.

- Implementation with more than two cutoffs: "fixed effects" model.

$$\mu_0(x, c_j) = g(x) + \theta_j$$

- Combining both approaches:

$$\mu_0(x, c_j) = g(x) + p_k(x)' \theta_j$$