Regression Discontinuity Designs

Teppei Yamamoto

Keio University

Introduction to Causal Inference
Spring 2016

Regression Discontinuity Design (RDD)

- RDD: A fairly old idea (Thistlethwaite and Campbell, 1960)
- Recently experienced a renaissance because of the formal theoretical foundation given by the causal inference framework

Regression Discontinuity Design (RDD)

- RDD: A fairly old idea (Thistlethwaite and Campbell, 1960)
- Recently experienced a renaissance because of the formal theoretical foundation given by the causal inference framework
- Applicable when treatment is assigned according to a *rule* based on another variable (called the forcing or running variable)
- Often useful for analysis in a "rule-based" world (administrative programs, elections, etc.)

Regression Discontinuity Design (RDD)

- RDD: A fairly old idea (Thistlethwaite and Campbell, 1960)
- Recently experienced a renaissance because of the formal theoretical foundation given by the causal inference framework
- Applicable when treatment is assigned according to a *rule* based on another variable (called the forcing or running variable)
- Often useful for analysis in a “rule-based” world (administrative programs, elections, etc.)
- High internal validity: One of the few observational designs that reproduced an experimental benchmark (Cook and Wong 2008)
- Limited external validity: Effect is only identified for a small subpopulation

Sharp RDD: Basic Setup

- $D_i \in \{0, 1\}$: Treatment
- X_i: Forcing variable that perfectly determines the value of D_i with cutpoint c

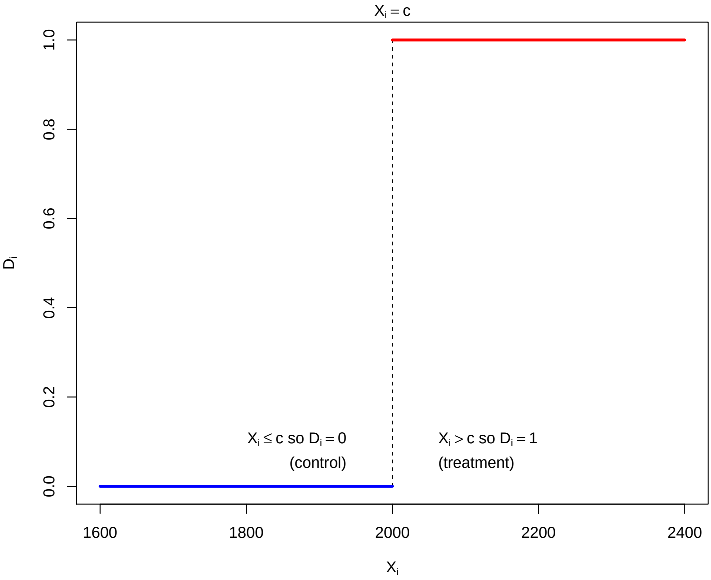
$$D_i = \mathbf{1}\{X_i > c\} \quad \text{or equivalently} \quad D_i = \begin{cases} 1 & \text{if } X_i > c \\ 0 & \text{if } X_i \leq c \end{cases}$$

- X_i may be correlated with $Y_i(0)$ and $Y_i(1)$, either directly or via other unobserved confounders
- Simply adjusting for X_i does not work because of lack of common support
- Basic intuition: Right at the cutpoint $X_i = c$, assignment to D_i may be as-if random

A Hypothetical Example: Effect of Scholarship

- Thistlethwaite and Campbell (1960) study the effects of college scholarships on later students' achievements
- Scholarships are given on the basis of whether or not a student's test score exceeds some threshold c
  - Treatment D_i is scholarship
  - Forcing variable X_i is SAT score with cutoff c
  - Outcome Y_i is subsequent earnings
  - $Y_i(0)$: potential earnings without the scholarship
  - $Y_i(1)$: potential earnings with the scholarship
- $Y_i(1)$ and $Y_i(0)$ are correlated with X_i: on average, students with higher SAT scores obtain higher earnings

Probability of Treatment in Sharp RDD



Identification of the Threshold Causal Effect

Key assumption: Continuity of average potential outcomes

$$\mathbb{E}[Y_i(d) \mid X_i = x] \text{ is continuous in } x \text{ around } X_i = c \text{ for } d = 0, 1$$

Identification of the Threshold Causal Effect

Key assumption: Continuity of average potential outcomes

$$\mathbb{E}[Y_i(d) \mid X_i = x] \text{ is continuous in } x \text{ around } X_i = c \text{ for } d = 0, 1$$

Causal estimand: Local ATE at the threshold

$$\tau_{SRD} \;=\; \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = c]$$

Identification of the Threshold Causal Effect

Key assumption: Continuity of average potential outcomes

$$\mathbb{E}[Y_i(d) \mid X_i = x] \text{ is continuous in } x \text{ around } X_i = c \text{ for } d = 0, 1$$

Causal estimand: Local ATE at the threshold

$$\tau_{SRD} \; = \; \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = c]$$

Identification result: If the continuity assumption holds, τ_{SRD} is nonparametrically identified as

$$\tau_{SRD} \; = \; \lim_{x \downarrow c} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i | X_i = x]$$

Identification of the Threshold Causal Effect

Key assumption: Continuity of average potential outcomes

$$\mathbb{E}[Y_i(d) \mid X_i = x] \text{ is continuous in } x \text{ around } X_i = c \text{ for } d = 0, 1$$

Causal estimand: Local ATE at the threshold

$$\tau_{SRD} \;=\; \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = c]$$

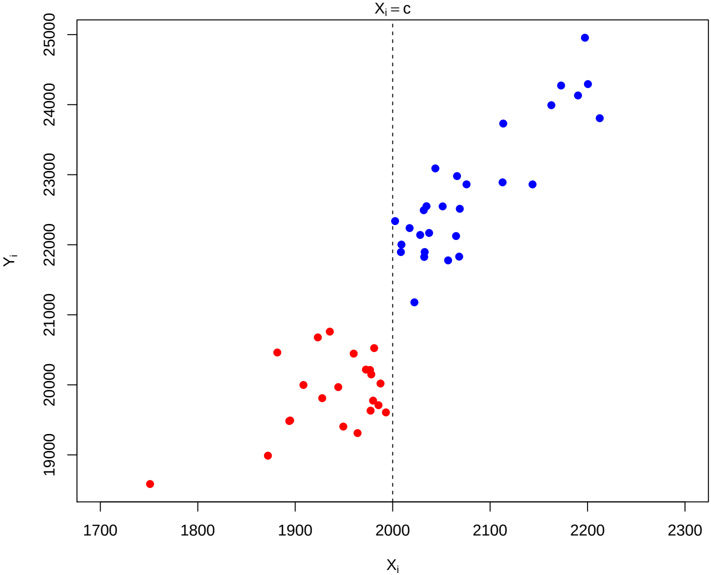
Identification result: If the continuity assumption holds, τ_{SRD} is nonparametrically identified as

$$\tau_{SRD} \;=\; \lim_{x \downarrow c} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i | X_i = x]$$

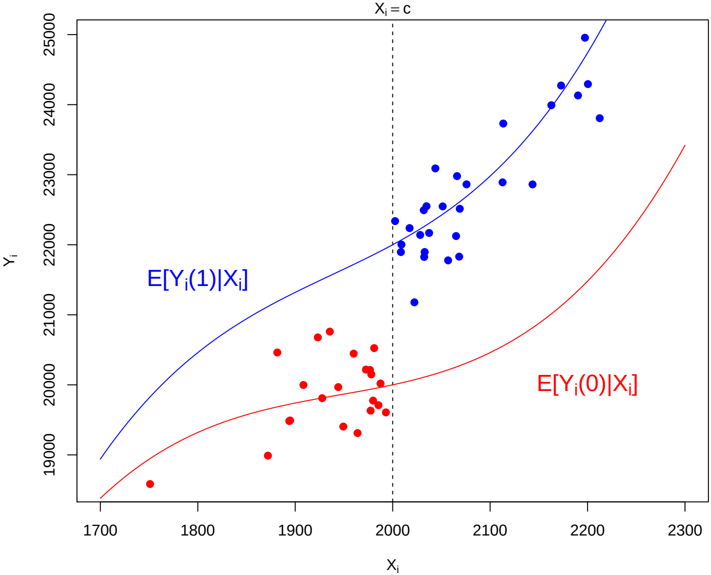
A “proof”:

- D_i is wholly determined by X_i, so contional ignorability is trivially satisfied given X_i: $Y_i(1), Y_i(0) \perp\!\!\!\perp D_i \mid X_i$
- However, *there is no common support*, so conditioning on X_i in a usual way won’t work.
- The continuity assumption allows us to do a tiny bit of extrapolation and compensate for the lack of common support at the threshold.

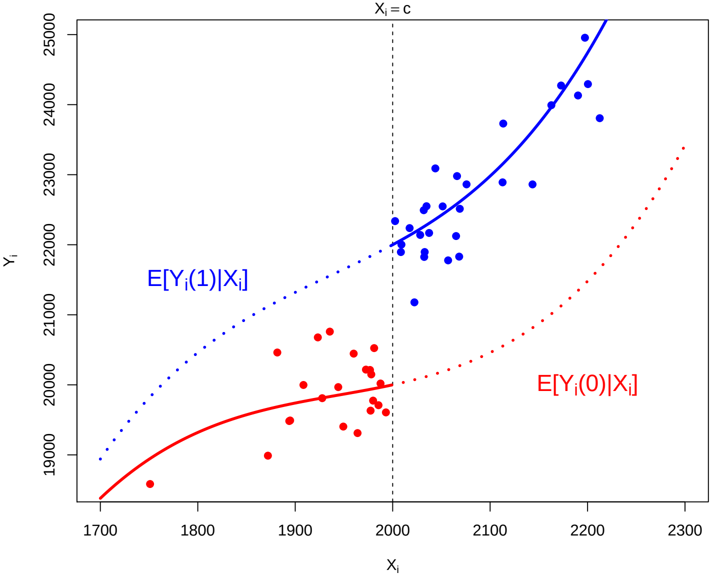
Graphical Illustration



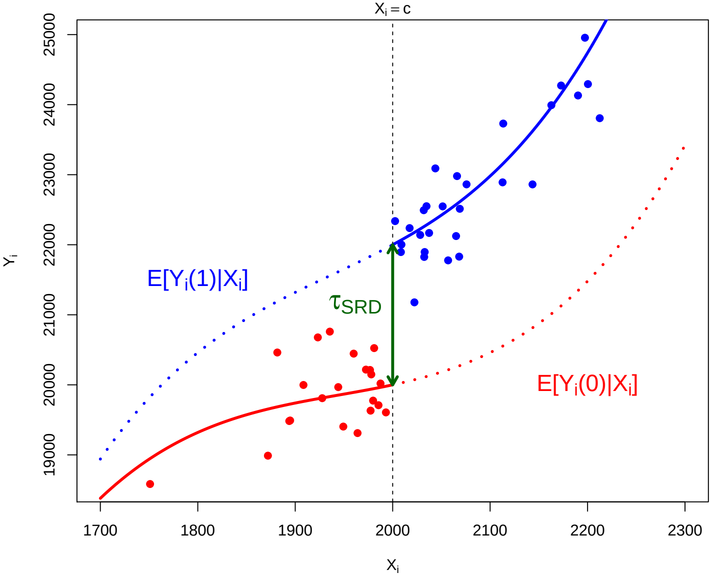
Graphical Illustration: Continuous $\mathbb{E}[Y_i(d) \mid X_i]$



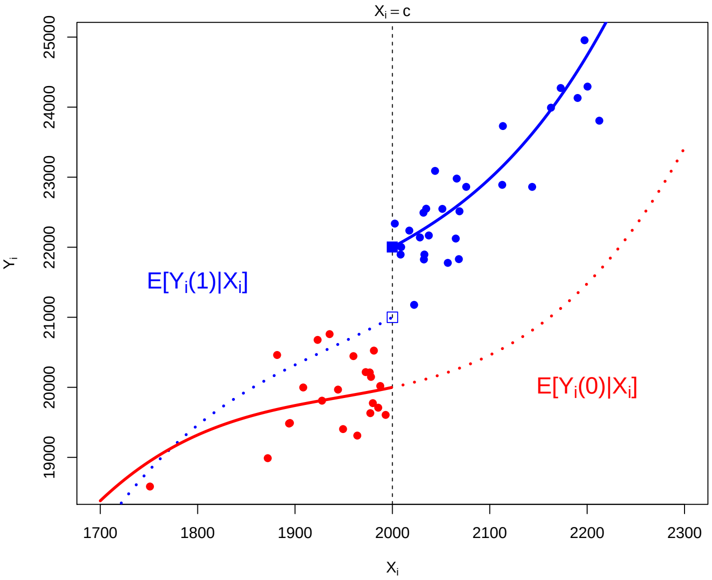
Graphical Illustration: Continuous $\mathbb{E}[Y_i(d) \mid X_i]$



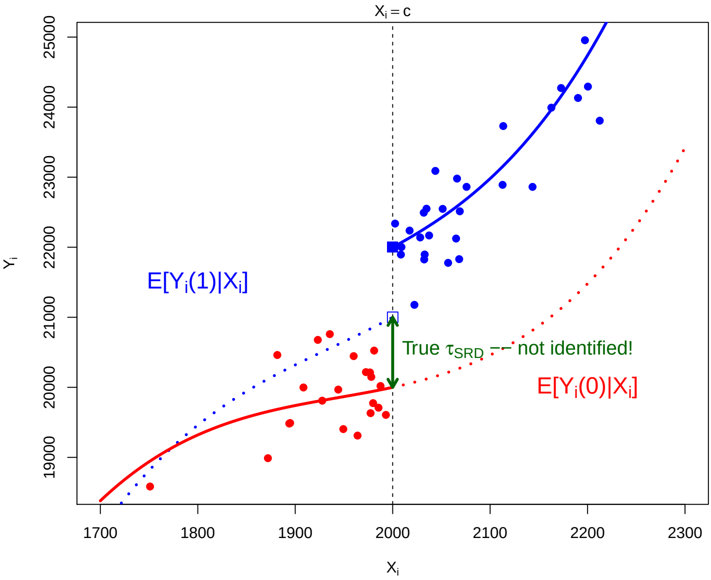
Graphical Illustration: Continuous $\mathbb{E}[Y_i(d) \mid X_i]$



Graphical Illustration: Discontinuous $\mathbb{E}[Y_i(d) \mid X_i]$



Graphical Illustration: Discontinuous $\mathbb{E}[Y_i(d) \mid X_i]$



Estimation of the LATE at the Threshold

1. Trim the sample to a reasonable window around the threshold c (discontinuity sample)
   - $c - h \leq X_i \leq c + h$, were $h > 0$ determines the width of the window
2. Recode forcing variable to deviations from threshold: $\tilde{X}_i = X_i - c$
   - $\tilde{X}_i = 0$ if $X_i = c$
   - $\tilde{X}_i > 0$ if $X_i > c$ and thus $D_i = 1$
   - $\tilde{X}_i < 0$ if $X_i < c$ and thus $D_i = 0$
3. Decide on a model for $\mathbb{E}[Y_i|\tilde{X}_i]$:
   - linear, common slope for $\mathbb{E}[Y_i \mid \tilde{X}_i < 0]$ and $\mathbb{E}[Y_i|\tilde{X}_i > 0]$
   - linear, different slopes
   - non-linear

   - each model corresponds to a particular set of assumptions about the potential outcomes
   - always start with visual inspection (e.g. scatter plot with lowess) to check which model is plausible

Estimation with a Linear Model with a Common Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0) | X_i = x]$ is linear in x
  2. Treatment effect, τ, does not depend on X_i

  i.e.,

Estimation with a Linear Model with a Common Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ is linear in x
  2. Treatment effect, τ, does not depend on X_i

  i.e.,

$$\mathbb{E}[Y_i(0)|X_i] = \alpha + \beta X_i \quad \text{and} \quad \mathbb{E}[Y_i(1) - Y_i(0)|X_i] = \tau$$

  which implies

Estimation with a Linear Model with a Common Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ is linear in x
  2. Treatment effect, τ, does not depend on X_i

  i.e.,

$$\mathbb{E}[Y_i(0)|X_i] = \alpha + \beta X_i \quad \text{and} \quad \mathbb{E}[Y_i(1) - Y_i(0)|X_i] = \tau$$

  which implies

$$\mathbb{E}[Y_i(1)|X_i] = \tau + \mathbb{E}[Y_i(0)|X_i] = \tau + \alpha + \beta X_i$$

Estimation with a Linear Model with a Common Slope

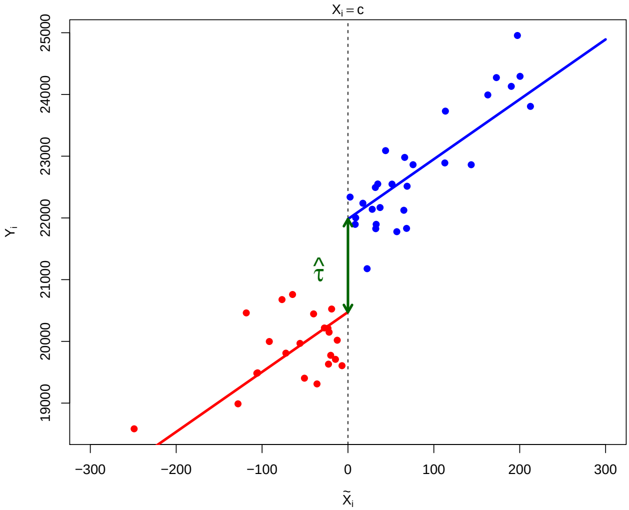
- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ is linear in x
  2. Treatment effect, τ, does not depend on X_i

  i.e.,
  $$\mathbb{E}[Y_i(0)|X_i] = \alpha + \beta X_i \quad \text{and} \quad \mathbb{E}[Y_i(1) - Y_i(0)|X_i] = \tau$$
  which implies
  $$\mathbb{E}[Y_i(1)|X_i] = \tau + \mathbb{E}[Y_i(0)|X_i] = \tau + \alpha + \beta X_i$$

- Therefore, the model for the observed outcome should be:
  $$\begin{aligned}\mathbb{E}[Y_i|X_i, D_i] &= D_i \cdot \mathbb{E}[Y_i(1)|X_i] + (1 - D_i) \cdot \mathbb{E}[Y_i(0)|X_i] \\ &= \alpha + \tau D_i + \beta X_i \\ &= \tilde{\alpha} + \tau D_i + \beta \tilde{X}_i \quad (\text{where } \tilde{\alpha} = \alpha + \beta c)\end{aligned}$$

- So we just regress the observed outcome (Y_i) on D_i and $\tilde{X}_i$

Estimation with a Linear Model with a Common Slope



Estimation with a Linear Model with a Different Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ and $\mathbb{E}[Y_i(1)|X_i = x]$ are both linear in x
  2. But we now allow treatment effect to vary with X_i

  i.e.,

Estimation with a Linear Model with a Different Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ and $\mathbb{E}[Y_i(1)|X_i = x]$ are both linear in x
  2. But we now allow treatment effect to vary with X_i

  i.e.,
  $$\mathbb{E}[Y_i(0)|X_i] = \alpha_0 + \beta_0 X_i \quad \text{and} \quad \mathbb{E}[Y_i(1)|X_i] = \alpha_1 + \beta_1 X_i$$
  such that
  $$\mathbb{E}[Y_i(1) - Y_i(0)|X_i] \; = \; (\alpha_1 - \alpha_0) + (\beta_1 - \beta_0) X_i$$

Estimation with a Linear Model with a Different Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ and $\mathbb{E}[Y_i(1)|X_i = x]$ are both linear in x
  2. But we now allow treatment effect to vary with X_i

  i.e.,

  $$\mathbb{E}[Y_i(0)|X_i] = \alpha_0 + \beta_0 X_i \quad \text{and} \quad \mathbb{E}[Y_i(1)|X_i] = \alpha_1 + \beta_1 X_i$$

  such that

  $$\mathbb{E}[Y_i(1) - Y_i(0)|X_i] \; = \; (\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)X_i$$

- The observed outcome model is therefore:

$$\begin{aligned}
\mathbb{E}[Y_i|X_i, D_i] &= D_i \cdot \mathbb{E}[Y_i(1)|X_i] + (1 - D_i) \cdot \mathbb{E}[Y_i(0)|X_i] \\
&= \alpha_0 + \beta_0 X_i + (\alpha_1 - \alpha_0)D_i + (\beta_1 - \beta_0)D_i X_i \\
&= (\alpha_0 + \beta_0 c) + \beta_0 \tilde{X}_i \\
&\qquad + \{(\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)c\} D_i + (\beta_1 - \beta_0)D_i \tilde{X}_i \\
&\equiv \tilde{\alpha} + \beta_0 \tilde{X}_i + \tau D_i + \tilde{\beta} D_i \tilde{X}_i
\end{aligned}$$

Estimation with a Linear Model with a Different Slope

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ and $\mathbb{E}[Y_i(1)|X_i = x]$ are both linear in x
  2. But we now allow treatment effect to vary with X_i

  i.e.,

$$\mathbb{E}[Y_i(0)|X_i] = \alpha_0 + \beta_0 X_i \quad \text{and} \quad \mathbb{E}[Y_i(1)|X_i] = \alpha_1 + \beta_1 X_i$$

  such that

$$\mathbb{E}[Y_i(1) - Y_i(0)|X_i] \ = \ (\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)X_i$$

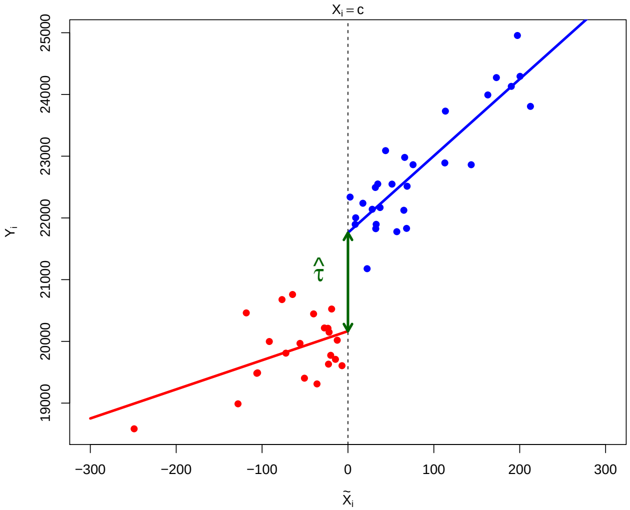
- The observed outcome model is therefore:

$$\begin{aligned}
\mathbb{E}[Y_i|X_i, D_i] &= D_i \cdot \mathbb{E}[Y_i(1)|X_i] + (1 - D_i) \cdot \mathbb{E}[Y_i(0)|X_i] \\
&= \alpha_0 + \beta_0 X_i + (\alpha_1 - \alpha_0)D_i + (\beta_1 - \beta_0)D_i X_i \\
&= (\alpha_0 + \beta_0 c) + \beta_0 \tilde{X}_i \\
&\qquad + \{(\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)c\}\, D_i + (\beta_1 - \beta_0)D_i\tilde{X}_i \\
&\equiv \tilde{\alpha} + \beta_0 \tilde{X}_i + \tau D_i + \tilde{\beta} D_i \tilde{X}_i
\end{aligned}$$

  Note that $\tau = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = c]$, LATE at the threshold

- So, regress Y_i on $\tilde{X}_i$, D_i and the interaction $D_i\tilde{X}_i$

Estimation with a Linear Model with a Different Slope



Estimation with a Nonlinear Model

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ and $\mathbb{E}[Y_i(1)|X_i = x]$ are now allowed to be non-linear in X_i, but must be correctly specified
  2. Treatment effect is allowed to vary across X_i
- Include quadratic, cubic, etc. terms in $\tilde{X}_i$ and their interactions with D_i in the equation

Estimation with a Nonlinear Model

- Assumptions:
  1. $\mathbb{E}[Y_i(0)|X_i = x]$ and $\mathbb{E}[Y_i(1)|X_i = x]$ are now allowed to be non-linear in X_i, but must be correctly specified
  2. Treatment effect is allowed to vary across X_i
- Include quadratic, cubic, etc. terms in $\tilde{X}_i$ and their interactions with D_i in the equation
- The specification with quadratic terms:

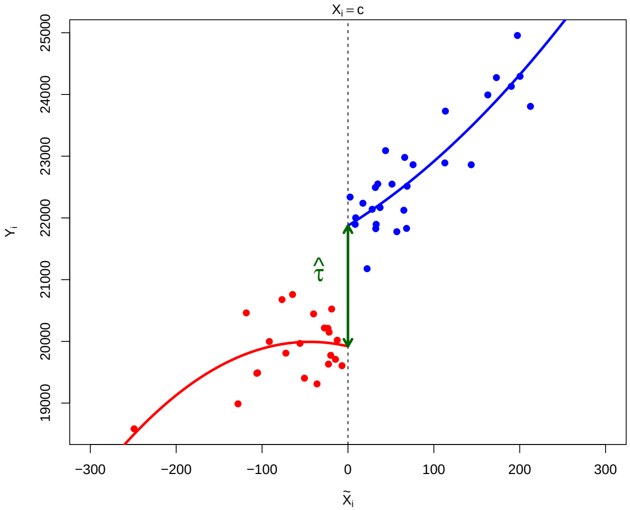
$$\mathbb{E}[Y_i|X_i, D_i] = \gamma_0 + \gamma_1 \tilde{X}_i + \gamma_2 \tilde{X}_i^2 + \tau D_i + \alpha_1 \tilde{X}_i D_i + \alpha_2 \tilde{X}_i^2 D_i$$

  The specification with cubic terms is

$$\begin{aligned}\mathbb{E}[Y_i|X_i, D_i] &= \gamma_0 + \gamma_1 \tilde{X}_i + \gamma_2 \tilde{X}_i^2 + \gamma_3 \tilde{X}_i^3 \\ &\quad + \alpha_0 D_i + \tau \tilde{X}_i D_i + \alpha_2 \tilde{X}_i^2 D_i + \alpha_3 \tilde{X}_i^3 D_i\end{aligned}$$

- In both cases, the coefficient on D_i corresponds to the LATE at the threshold: $\tau = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = c]$

Estimation with a Nonlinear Model



Model Selection and Choice of Bandwidth in RDD

- How should we pick the "right" model and bandwidth?

Model Selection and Choice of Bandwidth in RDD

- How should we pick the “right” model and bandwidth?
- No ex-ante correct answer usually, but several data-driven procedures are available

Model Selection and Choice of Bandwidth in RDD

- How should we pick the “right” model and bandwidth?
- No ex-ante correct answer usually, but several data-driven procedures are available

Model choice:

- A tradeoff between bias and variance
- Standard practice: Use different specs and show robustness
- Local linear regression with a kernel smoother is a popular choice

Model Selection and Choice of Bandwidth in RDD

- How should we pick the “right” model and bandwidth?
- No ex-ante correct answer usually, but several data-driven procedures are available

Model choice:

- A tradeoff between bias and variance
- Standard practice: Use different specs and show robustness
- Local linear regression with a kernel smoother is a popular choice

Bandwidth selection:

- Imbens-Kalyanaraman (IK) algorithm: Pick h that minimizes (a first-order approximation of) the MSE in $\hat{\tau}_{SRD}$
- Cross-validation: See Imbens and Lemieux (2008) for details

Example: Party Incumbency Advantage (Lee, 2006)

- What is the effect of incumbency status on vote shares?
- D_{itj}: Incumbency status of party j in district i at election t
- $V_{itj}(d)$: Potential vote share as incumbent ($d = 1$) or non-incumbent ($d = 0$)
- Observed vote share: $V_{itj} = V_{itj}(D_{itj})$

Example: Party Incumbency Advantage (Lee, 2006)

- What is the effect of incumbency status on vote shares?
- D_{itj}: Incumbency status of party j in district i at election t
- $V_{itj}(d)$: Potential vote share as incumbent ($d = 1$) or non-incumbent ($d = 0$)
- Observed vote share: $V_{itj} = V_{itj}(D_{itj})$

- Party incumbency effect for party j: $V_{itj}(1) - V_{itj}(0)$
- (Observed) margin of victory for party j: $Z_{itj} = V_{itj} - V_{itk}$ where k indicates the strongest opposition party

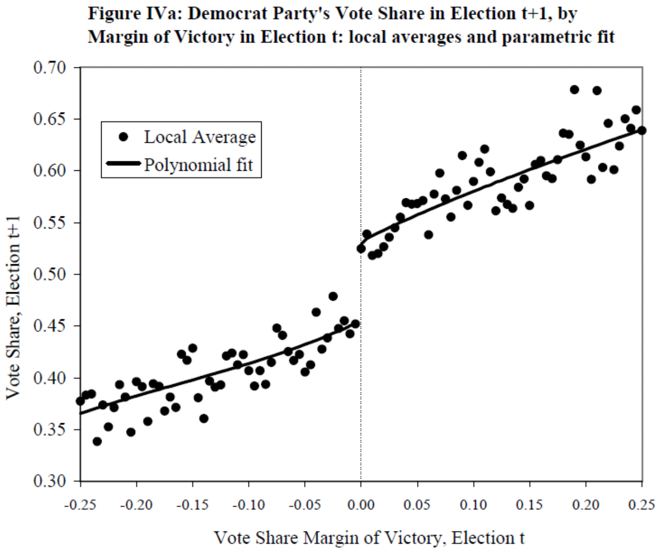
Example: Party Incumbency Advantage (Lee, 2006)

- What is the effect of incumbency status on vote shares?
- D_{itj}: Incumbency status of party j in district i at election t
- $V_{itj}(d)$: Potential vote share as incumbent ($d = 1$) or non-incumbent ($d = 0$)
- Observed vote share: $V_{itj} = V_{itj}(D_{itj})$
- Party incumbency effect for party j: $V_{itj}(1) - V_{itj}(0)$
- (Observed) margin of victory for party j: $Z_{itj} = V_{itj} - V_{itk}$ where k indicates the strongest opposition party
- Party incumbency status in election $t + 1$ is then assigned by Z_{itj}:

$$D_{ij(t+1)} = \mathbf{1}\{Z_{itj} > 0\} \quad \text{or equivalently} \quad D_{ij(t+1)} = \begin{cases} 1 & \text{if } Z_{itj} > 0 \\ 0 & \text{if } Z_{itj} < 0 \end{cases}$$

- With only two parties we can also use $Z_{itj} = V_{itj} - c$ with $c = 0.5$

Example: Party Incumbency Advantage



Other Recent Examples

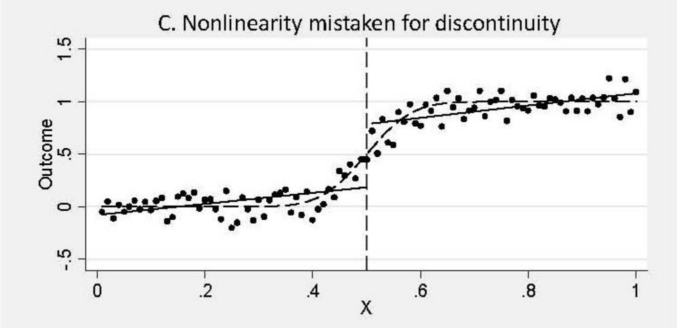
- Effect of class size on student achievement (class size is determined by a cutoff in class size)
- Effect of access to credit on development outcomes (loan offer is determined by credit score threshold)
- Effect of party democratic versus republican mayor
- Effect of wages increase for mayors on policy performance (wage jumps at population cutoffs)
- Effect of an additional night in the hospital, a newborn delivered at 12:05 a.m. will have an extra night of reimbursable care
- Effect of school district boundaries on home values
- Effect of colonial borders on development outcomes
- Effect of electronic voting on incumbent vote shares

Falsification Checks

Diagnoze the robustness of your results via falsification checks:

1. Sensitivity: Are results sensitive to alternative specifications?
2. Balance checks: Does any covariate Z_i jump at the threshold?
3. Check if jumps occur at placebo thresholds c^*?
4. Sorting: Do units sort around the threshold?

Sensitivity to Specification



- RDD requires specification of the functional form and bandwidth
- Misspecification of either can lead to a spurious jump
- Take care not to confuse a nonlinear relation with a discontinuity!
- More flexibility (e.g. polynomials) creates more bias but less efficiency
- Check sensitivity to size of bandwidth *h*

Balance Checks: Covariates as Placebo Outcomes

- Test for comparability of agents around the cutoff:
  - Visual tests: Plot $\mathbb{E}[Z_i|X_i, D_i]$ and look for jumps
  - Relation between covariates and treatment should be smooth around threshold
  - Use Z_i as a placebo outcome and see if there is inbalance:

$$\mathbb{E}[Z_i|X_i, D_i] = \beta_0 + \beta_1 \tilde{X}_i + \tau_z D_i + \beta_3 \tilde{X}_i D_i$$

  - $\tau_z = 0$ if Z_i is balanced at the threshold

Balance Checks: Covariates as Placebo Outcomes

- Test for comparability of agents around the cutoff:
  - Visual tests: Plot $\mathbb{E}[Z_i|X_i, D_i]$ and look for jumps
  - Relation between covariates and treatment should be smooth around threshold
  - Use Z_i as a placebo outcome and see if there is inbalance:

$$\mathbb{E}[Z_i|X_i, D_i] = \beta_0 + \beta_1 \tilde{X}_i + \tau_z D_i + \beta_3 \tilde{X}_i D_i$$

  - $\tau_z = 0$ if Z_i is balanced at the threshold
- Discontinuity in Z_i indicates evidence of discontinuous $\mathbb{E}[Y_i(d) \mid X_i = x]$, violating the key assumption
- Inbalance can be addressed by incorporating Z_i in the analysis:
  - Use Z_i as an additional covariate in the model
  - Alternatively, regress Y_i on Z_i and use the residuals in the model, instead of Y_i itself

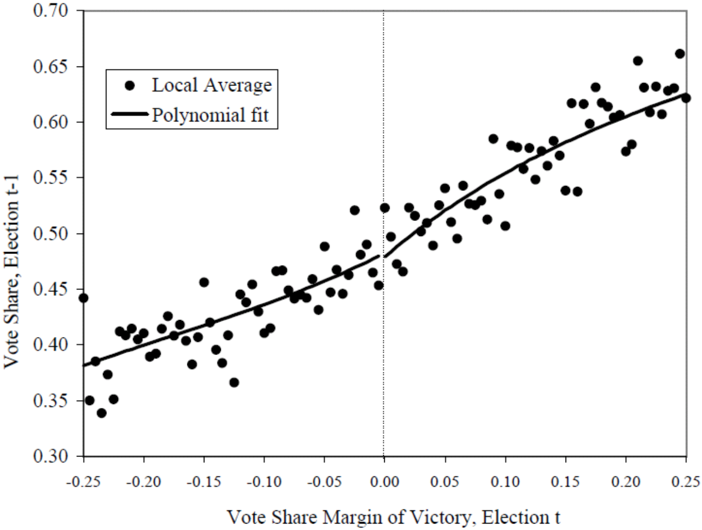
Balance Checks: Covariates as Placebo Outcomes

- Test for comparability of agents around the cutoff:
  - Visual tests: Plot $\mathbb{E}[Z_i|X_i, D_i]$ and look for jumps
  - Relation between covariates and treatment should be smooth around threshold
  - Use Z_i as a placebo outcome and see if there is inbalance:

$$\mathbb{E}[Z_i|X_i, D_i] = \beta_0 + \beta_1 \tilde{X}_i + \tau_z D_i + \beta_3 \tilde{X}_i D_i$$

  - $\tau_z = 0$ if Z_i is balanced at the threshold
- Discontinuity in Z_i indicates evidence of discontinuous $\mathbb{E}[Y_i(d) \mid X_i = x]$, violating the key assumption
- Inbalance can be addressed by incorporating Z_i in the analysis:
  - Use Z_i as an additional covariate in the model
  - Alternatively, regress Y_i on Z_i and use the residuals in the model, instead of Y_i itself
- Balance checks address only observables, not unobservables

Placebo Outcome: Lee (2006)



Placebo Thresholds

- Test whether the treatment effect is zero when it should be
- Concern: the relation may be fundamentally discontinuous and jump at the cutoff is contaminated by other factors

Placebo Thresholds

- Test whether the treatment effect is zero when it should be
- Concern: the relation may be fundamentally discontinuous and jump at the cutoff is contaminated by other factors

- c^*: a placebo thershold value for X_i
- $\tilde{X}_i^* = X_i - c^*$
- Then use $\tilde{X}_i^*$ instead of $\tilde{X}_i$:

$$\mathbb{E}[Y_i | X_i, D_i] = \beta_0 + \beta_1 \tilde{X}_i^* + \tau^* D_i + \beta_3 \tilde{X}_i^* D_i$$

  and check if τ^* is large and significant
- Use only observations on the same side of the actual threshold c (to avoid misspecification by imposing a zero jump at c)

Placebo Thresholds

- Test whether the treatment effect is zero when it should be
- Concern: the relation may be fundamentally discontinuous and jump at the cutoff is contaminated by other factors

- c^*: a placebo thershold value for X_i
- $\tilde{X}_i^* = X_i - c^*$
- Then use $\tilde{X}_i^*$ instead of $\tilde{X}_i$:

$$\mathbb{E}[Y_i|X_i, D_i] = \beta_0 + \beta_1 \tilde{X}_i^* + \tau^* D_i + \beta_3 \tilde{X}_i^* D_i$$

  and check if τ^* is large and significant
- Use only observations on the same side of the actual threshold c (to avoid misspecification by imposing a zero jump at c)

- Note that a large placebo jump does not directly imply a violation of the identification assumption
- But it requires an explanation

Sorting Around the Threshold

- Agents' behavior can invalidate the continuity assumption:
  - Agents may exercise control over their values of X_i to fall on the beneficial side of the threshold
  - Administrators may strategically choose what X_i to use or which threshold to use
  - Such sorting of agents invalidate the continuity assumption

Sorting Around the Threshold

- Agents' behavior can invalidate the continuity assumption:
  - Agents may exercise control over their values of X_i to fall on the beneficial side of the threshold
  - Administrators may strategically choose what X_i to use or which threshold to use
  - Such sorting of agents invalidate the continuity assumption
- When this occurs, distribution of X_i will discontinuously change at the threshold

Sorting Around the Threshold

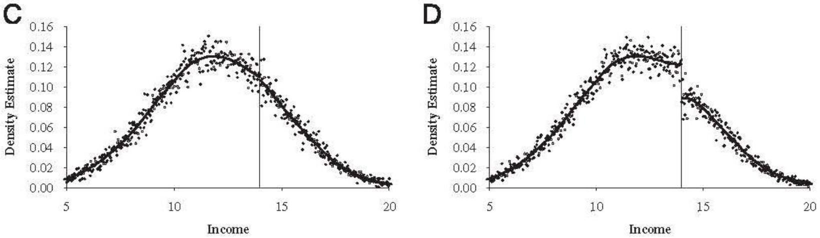
- Agents' behavior can invalidate the continuity assumption:
  - Agents may exercise control over their values of X_i to fall on the beneficial side of the threshold
  - Administrators may strategically choose what X_i to use or which threshold to use
  - Such sorting of agents invalidate the continuity assumption
- When this occurs, distribution of X_i will discontinuously change at the threshold
- Diagnostics:
  1. Visual inspection of Histograms (make sure no bin overlaps with the threshold!)
  2. Formal tests (e.g. McCrary 2008)

Sorting Around the Threshold

- Agents' behavior can invalidate the continuity assumption:
  - Agents may exercise control over their values of X_i to fall on the beneficial side of the threshold
  - Administrators may strategically choose what X_i to use or which threshold to use
  - Such sorting of agents invalidate the continuity assumption
- When this occurs, distribution of X_i will discontinuously change at the threshold
- Diagnostics:
  1. Visual inspection of Histograms (make sure no bin overlaps with the threshold!)
  2. Formal tests (e.g. McCrary 2008)
- A related problem: Other treatments assigned by the exact same X_i and c (e.g. geographic boundary)

Example: Job Training Program

- Beneficial job training program offered to agents with income $< c$
- Concern: People may withhold labor to lower their income just below the cutoff to gain access to the program



Generalization to Imperfect Thresholds

- Threshold may not perfectly determine treatment exposure, but it may still create a discontinuity in the *probability* of treatment

Generalization to Imperfect Thresholds

- Threshold may not perfectly determine treatment exposure, but it may still create a discontinuity in the *probability* of treatment
- Hypothetical example: SAT and scholarship
  - Scholarship may not be wholly determined by SAT scores X_i, but the rule specifies a threshold $X_i = c$ to become eligible
  - Students above c get a chance of receiving the scholarship, but may not actually get it
  - Students below c is excluded from consideration, perhaps except ones who are otherwise remarkable

Generalization to Imperfect Thresholds

- Threshold may not perfectly determine treatment exposure, but it may still create a discontinuity in the *probability* of treatment
- Hypothetical example: SAT and scholarship
  - Scholarship may not be wholly determined by SAT scores X_i, but the rule specifies a threshold $X_i = c$ to become eligible
  - Students above c get a chance of receiving the scholarship, but may not actually get it
  - Students below c is excluded from consideration, perhaps except ones who are otherwise remarkable
- That is, an encouragement to receive treatment is discontinuously determined by X_i at the threshold c
- Fuzzy RDD can therefore be thought of an instrumental variable version of an RDD

Fuzzy RDD: Setup and Identification

- $Z_i \in \{0, 1\}$: Encouragement
- X_i: Forcing variable that perfectly determines the value of Z_i with cutpoint c

$$Z_i = \mathbf{1}\{X_i > c\} \quad \text{or equivalently} \quad Z_i = \begin{cases} 1 & \text{if } X_i > c \\ 0 & \text{if } X_i \leq c \end{cases}$$

Fuzzy RDD: Setup and Identification

- $Z_i \in \{0, 1\}$: Encouragement
- X_i: Forcing variable that perfectly determines the value of Z_i with cutpoint c

$$Z_i = \mathbf{1}\{X_i > c\} \quad \text{or equivalently} \quad Z_i = \begin{cases} 1 & \text{if } X_i > c \\ 0 & \text{if } X_i \leq c \end{cases}$$

Identification assumptions:

- Both $\mathbb{E}[D_i(z) \mid X_i = x]$ (potential treatment) and $\mathbb{E}[Y_i(z) \mid X_i = x]$ (potential outcome) are continuous in x around $X_i = c$ for $z = 0, 1$
- IV assumptions: Monotonicity, exclusion restriction, relevance of Z_i

Fuzzy RDD: Setup and Identification

- $Z_i \in \{0, 1\}$: Encouragement
- X_i: Forcing variable that perfectly determines the value of Z_i with cutpoint c

$$Z_i = \mathbf{1}\{X_i > c\} \quad \text{or equivalently} \quad Z_i = \begin{cases} 1 & \text{if } X_i > c \\ 0 & \text{if } X_i \leq c \end{cases}$$

Identification assumptions:

- Both $\mathbb{E}[D_i(z) \mid X_i = x]$ (potential treatment) and $\mathbb{E}[Y_i(z) \mid X_i = x]$ (potential outcome) are continuous in x around $X_i = c$ for $z = 0, 1$
- IV assumptions: Monotonicity, exclusion restriction, relevance of Z_i

Causal estimand: Local ATE for compliers at the threshold

$$\tau_{FRD} \;=\; \mathbb{E}[Y_i(1) - Y_i(0) \mid \text{unit } i \text{ is a complier and } X_i = c]$$

Fuzzy RDD: Setup and Identification

- $Z_i \in \{0, 1\}$: Encouragement
- X_i: Forcing variable that perfectly determines the value of Z_i with cutpoint c

$$Z_i = \mathbf{1}\{X_i > c\} \quad \text{or equivalently} \quad Z_i = \begin{cases} 1 & \text{if } X_i > c \\ 0 & \text{if } X_i \leq c \end{cases}$$

Identification assumptions:

- Both $\mathbb{E}[D_i(z) \mid X_i = x]$ (potential treatment) and $\mathbb{E}[Y_i(z) \mid X_i = x]$ (potential outcome) are continuous in x around $X_i = c$ for $z = 0, 1$
- IV assumptions: Monotonicity, exclusion restriction, relevance of Z_i

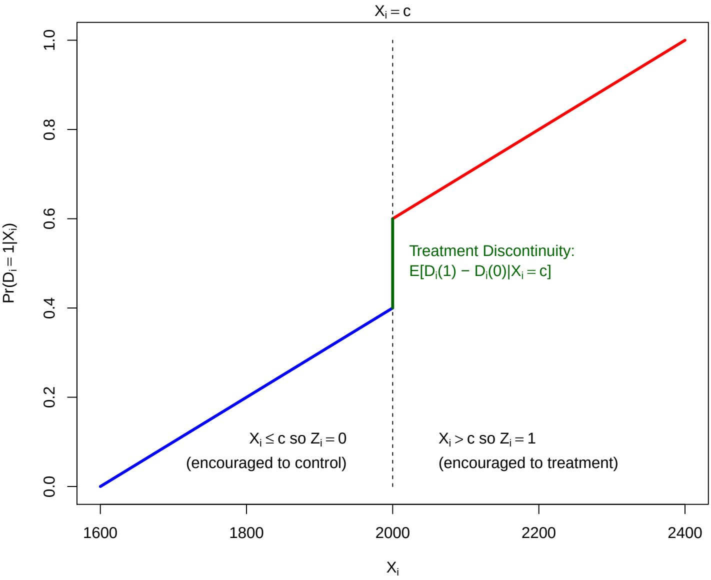
Causal estimand: Local ATE for compliers at the threshold

$$\tau_{FRD} = \mathbb{E}[Y_i(1) - Y_i(0) \mid \text{unit } i \text{ is a complier and } X_i = c]$$

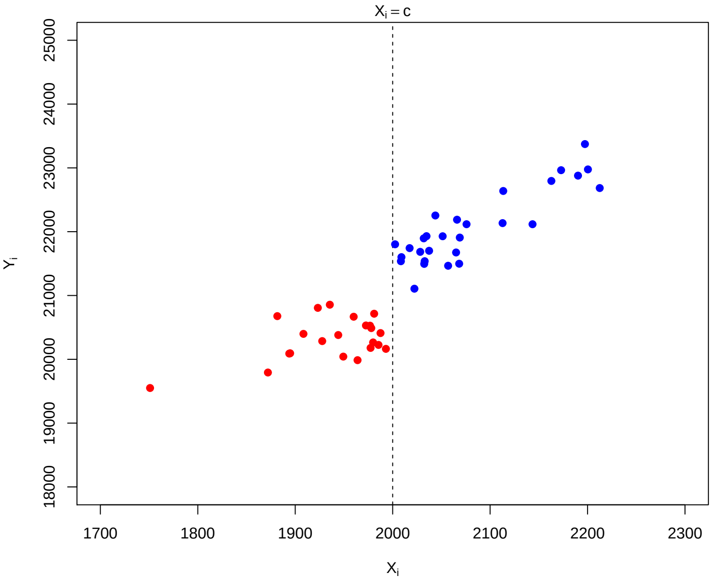
Identification result: Under the above assumptions, τ_{FRD} is identified as

$$\tau_{FRD} = \frac{\lim_{x \downarrow c} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i | X_i = x]}{\lim_{x \downarrow c} \mathbb{E}[D_i | X_i = x] - \lim_{x \uparrow c} \mathbb{E}[D_i | X_i = x]}$$

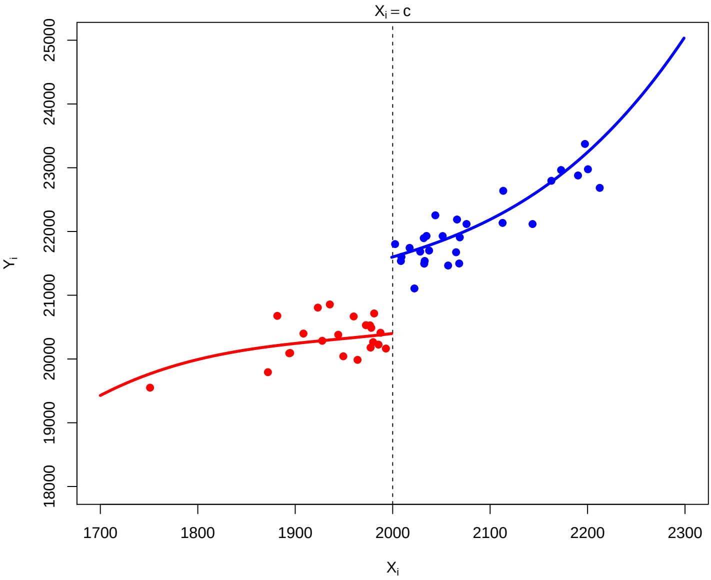
Probability of Treatment in Fuzzy RDD



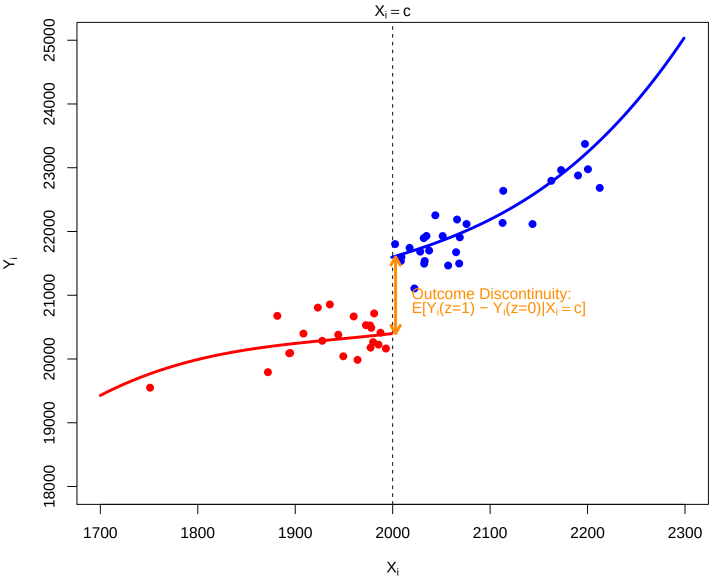
Fuzzy RDD: Discontinuity in $\mathbb{E}[Y|X]$



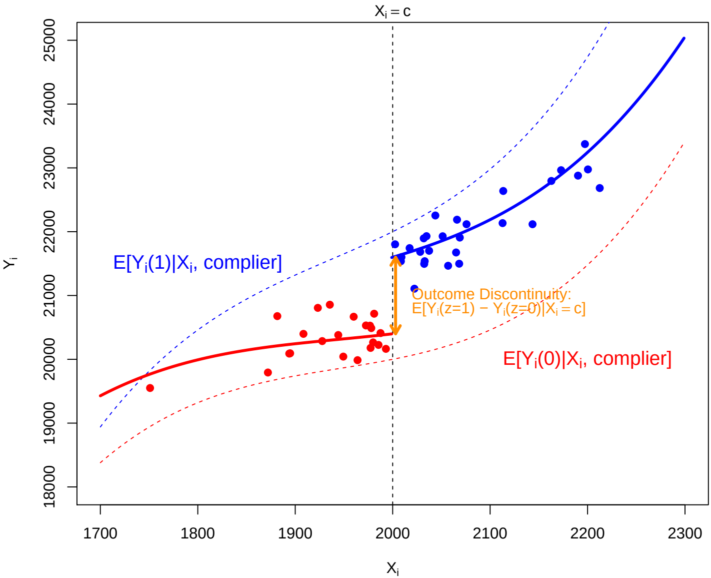
Fuzzy RDD: Discontinuity in $\mathbb{E}[Y|X]$



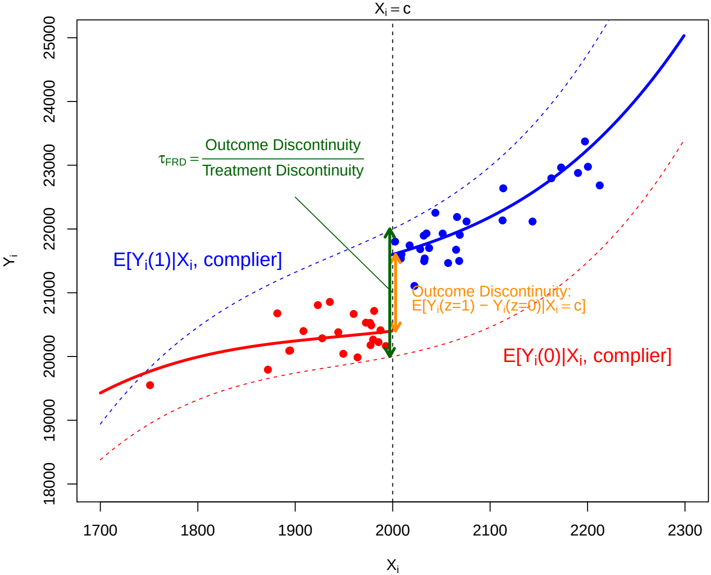
Fuzzy RDD: Discontinuity in $\mathbb{E}[Y|X]$



Fuzzy RDD: Discontinuity in $\mathbb{E}[Y|X]$



Fuzzy RDD: Discontinuity in $\mathbb{E}[Y|X]$



Fuzzy RDD: Estimation

1. Trim the sample to a reasonable window above and below the threshold c (discontinuity sample)
2. Code the encouragement indicator: $Z_i = \mathbf{1}\{X_i > c\}$
3. Recode the forcing variable to deviation from c: $\tilde{X}_i = X_i - c$
4. Estimate the outcome model using two-stage least squares:

$$Y_i = \beta_0 + \beta_1 \tilde{X}_i + \beta_2 Z_i \tilde{X}_i + \tau D_i + \varepsilon_i,$$

   where D_i is instrumented by Z_i
   - More flexible specifications can be used (e.g. polynomials of $\tilde{X}_i$)
5. Then $\hat{\tau}_{2SLS}$ consistently (but not unbiasedly) estimates τ_{FRD}

Fuzzy RDD: Estimation

1. Trim the sample to a reasonable window above and below the threshold c (discontinuity sample)
2. Code the encouragement indicator: $Z_i = \mathbf{1}\{X_i > c\}$
3. Recode the forcing variable to deviation from c: $\tilde{X}_i = X_i - c$
4. Estimate the outcome model using two-stage least squares:

$$Y_i = \beta_0 + \beta_1 \tilde{X}_i + \beta_2 Z_i \tilde{X}_i + \tau D_i + \varepsilon_i,$$

   where D_i is instrumented by Z_i
   - More flexible specifications can be used (e.g. polynomials of $\tilde{X}_i$)
5. Then $\hat{\tau}_{2SLS}$ consistently (but not unbiasedly) estimates τ_{FRD}

- In addition, it is also helpful to separately plot (and estimate) the outcome discontinuity and treatment discontinuity for interpretation
- Usual diagnostics can be applied to check plausibility of assumptions

Early Release Program (HDC)

- Prison system in many countries is faced with overcrowding and high recidivism rates after release.
- Early discharge of prisoners on electronic monitoring has become a popular policy
- Difficult to estimate impact of early release program on future criminal behaviour: best behaved inmates are usually the ones to be released early.
- Marie (2008) considers Home Detention Curfew (HDC) scheme in England and Wales:
  - Fuzzy RDD: Only offenders sentenced to more than three months (88 days) in prison are eligible for HDC, but obviously, not all those with longer sentences are offered HDC

Table 1: Descriptive Statistics for Prisoners Released by Length of Sentence and HDC and Non HDC Discharges

Panel A - Released Before 3 Months:			
Discharge Type	**Non HDC**	**HDC**	**Total**
Percentage Female	12.2	-	12.2
Mean Age	29.5	-	29.5
Percentage Incarcerated for Violence	17.6	-	17.6
Mean Number Previous Offences	8.8	-	8.8
Recidivism within 12 Months	52.4	-	52.4
Sample Size	42,987	0	42,987
Panel B - Released Between 3 and 6 Months:			
Discharge Type	**Non HDC**	**HDC**	**Total**
Percentage Female	8.8	8.8	8.8
Mean Age at Release	27.6	30.8	28.4
Percentage Incarcerated for Violence	20.3	18.3	19.8
Mean Number Previous Offences	10	6.5	9.1
Recidivism within 12 Months	60	30.2	52.6
Sample Size	52,091	17,222	69,313

Table 2: Descriptive Statistics for Prisoners Released by Length of Sentence and HDC and Non HDC Discharges and +/-7 Days Around Discontinuity Threshold

Panel A - Released +/- 7 Days of 3 Months (88 Days) Cut-off:			
Discharge Type	**Non HDC**	**HDC**	**Total**
Percentage Female	10.5	9.7	10.3
Mean Age at Release	28.9	30.7	29.3
Percentage Incarcerated for Violence	19.8	18.2	19.4
Mean Number Previous Offences	9.5	5.7	8.7
Recidivism within 12 Months	54.6	28.1	48.8
Sample Size	18,928	5,351	24,279
Panel B - Released +/- 7 Days of 3 Months (88 Days) Cu-off:			
Day of Release around Cut-off	**- 7 Days**	**+ 7 Days**	**Total**
Percentage Female	11	10.2	10.3
Mean Age at Release	28.8	29.4	29.3
Percentage Incarcerated for Violence	17.1	19.7	19.4
Mean Number Previous Offences	9.1	8.6	8.7
Recidivism within 12 Months	56.8	47.9	48.8
Percentage Released on HDC	0	24.4	22
Sample Size	2,333	21,946	24,279

Figure 1: Proportion Discharged on HDC by Sentence Length

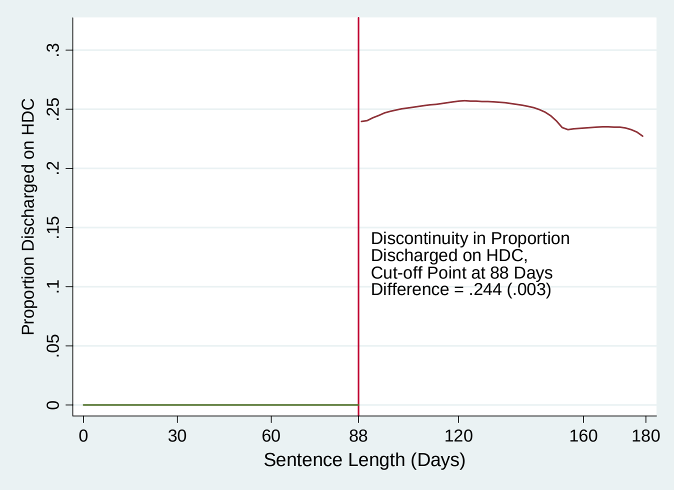


Figure 2: Mean Number of Previous Offence by Sentence Length

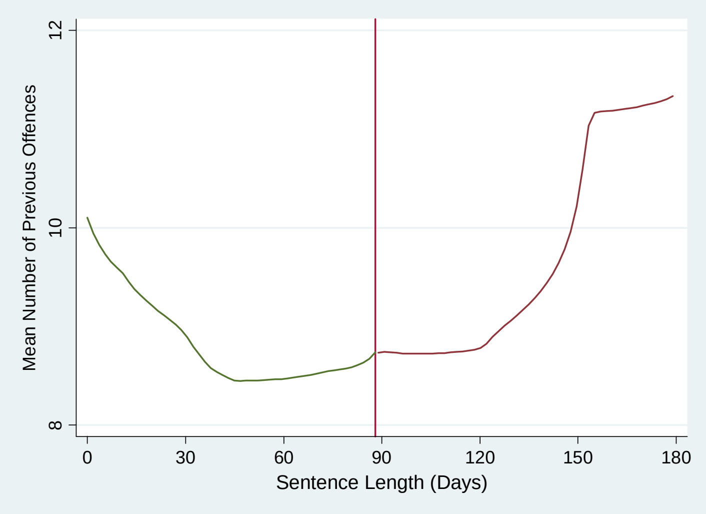


Figure 4: Recidivism within 1 Year by Sentence Length

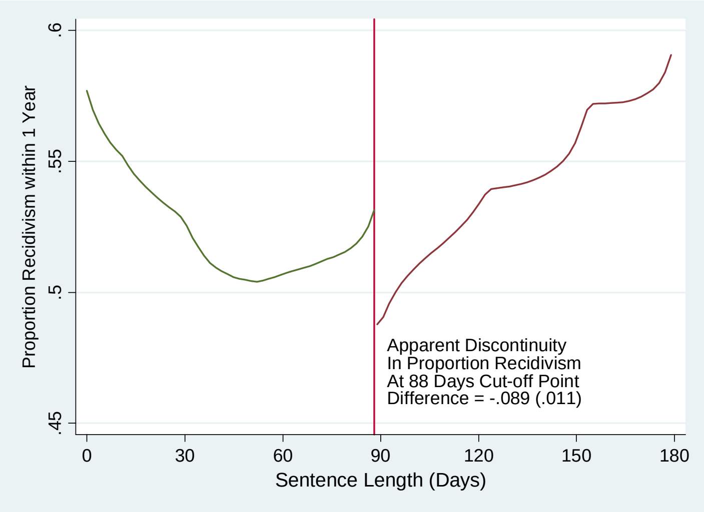


Table 4: RDD Estimates of HDC Impact on Recidivism – Around Threshold

	Dependent Variable = Recidivism Within 12 Months Estimation on Individuals Discharged +/- 7 Days of 88 Days Threshold		
	(1)	(2)	(3)
Estimated Discontinuity of HDC Participation at Threshold ($HDC^{+} - HDC^{-}$)	.243 (.009)	.223 (.009)	.243 (.003)
Estimated Difference in Recidivism Around Threshold ($Rec^{+} - Rec^{-}$)	-.089 (.011)	-.059 (.009)	-.044 (.014)
Estimated Effect of HDC on Recidivism Participation ($Rec^{+} - Rec^{-}$)/ ($HDC^{+} - HDC^{-}$)	-.366 (.044)	-.268 (.044)	-.181 (n.a.)
Controls	No	Yes	No
PSM	No	No	Yes
Sample Size	24,279	24,279	24,279

Concluding Remarks: Internal and External Validity

- At best, RDD estimates the average effect of the sub-population with X_i close to c
- Only with additional assumptions (e.g. constant treatment effects) can we estimate the overall ATE
- External validity is even more limited for fuzzy RDD
- Estimand refers to compliers at the threshold only