The Power Asymmetry in Fuzzy Regression Discontinuity Designs

Daniel Kaliski
Birkbeck U. of London

Michael P. Keane
John Hopkins

Timothy Neal
U. of New South Wales

Presentation by Marinho Bertanha (U. of Notre Dame)
2025 Annual Health Econometrics Workshop in Ann Arbor, MI
Nov 7^{th}, 2025

Motivation

- Keane and Neal (JoE 2023 and ARoE 2024) show asymmetric power curves for the conventional 2SLS t-test even when instruments are "considered strong".

Motivation

- Keane and Neal (JoE 2023 and ARoE 2024) show asymmetric power curves for the conventional 2SLS t-test even when instruments are "considered strong".

- When the true effect is negative but OLS says it is positive, the 2SLS t-test thas little power to detect the true effect.

Motivation

- Keane and Neal (JoE 2023 and ARoE 2024) show asymmetric power curves for the conventional 2SLS t-test even when instruments are "considered strong".
- When the true effect is negative but OLS says it is positive, the 2SLS t-test thas little power to detect the true effect.
- Fuzzy RD is quite similar to 2SLS.
- Does the conventional fuzzy RD t-test have the same issues?

What Does the Paper Do?

- Reviews differences and similarities between fuzzy RD and IV.

What Does the Paper Do?

- Reviews differences and similarities between fuzzy RD and IV.
- Runs several simulations using two popular testing procedures for fuzzy RD.
    1. R and STATA `rdrobust`, Cattaneo et al ECMA 2014.
    2. R and STATA `RDHonest`, Armstrong & Kolesar QE 2020.

What Does the Paper Do?

- Reviews differences and similarities between fuzzy RD and IV.
- Runs several simulations using two popular testing procedures for fuzzy RD.
  1. R and STATA `rdrobust`, Cattaneo et al ECMA 2014.
  2. R and STATA `RDHonest`, Armstrong & Kolesar QE 2020.
- A set of those simulations is grounded on a real-world example, Ambrus et al AER 2020.

What Does the Paper Find?

- Power asymmetry for the conventional fuzzy RD t-test even for conventionally "strong" first stages.

What Does the Paper Find?

- Power asymmetry for the conventional fuzzy RD t-test even for conventionally "strong" first stages.

- Simulations based on Ambrus et al AER 2020: when the true effect is negative and "OLS" overestimates it, the fuzzy RD t-test fails to detect the true effect.

What Does the Paper Find?

- Power asymmetry for the conventional fuzzy RD t-test even for conventionally "strong" first stages.
- Simulations based on Ambrus et al AER 2020: when the true effect is negative and "OLS" overestimates it, the fuzzy RD t-test fails to detect the true effect.
- **General pattern:** when bias of "OLS" is positive, there is a negative relationship between fuzzy RD estimate and standard errors. And vice-versa.

What Does the Paper Find?

- Power asymmetry for the conventional fuzzy RD t-test even for conventionally "strong" first stages.
- Simulations based on Ambrus et al AER 2020: when the true effect is negative and "OLS" overestimates it, the fuzzy RD t-test fails to detect the true effect.
- **General pattern:** when bias of "OLS" is positive, there is a negative relationship between fuzzy RD estimate and standard errors. And vice-versa.
- **Easy fix**: Anderson-Rubin test, that is, the conventional t-test based on the reduced-form for the outcome.

What Does the Paper Find?

- Power asymmetry for the conventional fuzzy RD t-test even for conventionally "strong" first stages.
- Simulations based on Ambrus et al AER 2020: when the true effect is negative and "OLS" overestimates it, the fuzzy RD t-test fails to detect the true effect.
- **General pattern:** when bias of "OLS" is positive, there is a negative relationship between fuzzy RD estimate and standard errors. And vice-versa.
- **Easy fix**: Anderson-Rubin test, that is, the conventional t-test based on the reduced-form for the outcome.
- **Interesting**: the way you run RD matters; lack of bias correction and high-order polynomials tend to aggravate the problem.

A Simple Example

- Have you ever estimated a parameter that is a ratio ?

$$\beta = \frac{a}{b}$$

A Simple Example

- Have you ever estimated a parameter that is a ratio ?

$$\beta = \frac{a}{b}$$

- Yes, all the time!

A Simple Example

- Have you ever estimated a parameter that is a ratio ?

$$\beta = \frac{a}{b}$$

- Yes, all the time!
    1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.

A Simple Example

- Have you ever estimated a parameter that is a ratio ?

$$\beta = \frac{a}{b}$$

- Yes, all the time!
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
  2. Relative effects of `educ` and `exper` in a linear regression of `wages` on `educ`, `exper`, etc, $\beta = \frac{\beta_{\texttt{educ}}}{\beta_{\texttt{exper}}}$.

A Simple Example

- Have you ever estimated a parameter that is a ratio ?

$$\beta = \frac{a}{b}$$

- Yes, all the time!
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
  2. Relative effects of `educ` and `exper` in a linear regression of `wages` on `educ`, `exper`, etc, $\beta = \frac{\beta_{\texttt{educ}}}{\beta_{\texttt{exper}}}$.
  3. IV, $\beta = \frac{COV[Z,Y]}{COV[Z,X]}$.

A Simple Example

- Have you ever estimated a parameter that is a ratio ?

$$\beta = \frac{a}{b}$$

- Yes, all the time!
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
  2. Relative effects of `educ` and `exper` in a linear regression of `wages` on `educ`, `exper`, etc, $\beta = \frac{\beta_{\texttt{educ}}}{\beta_{\texttt{exper}}}$.
  3. IV, $\beta = \frac{COV[Z,Y]}{COV[Z,X]}$.
  4. Fuzzy RD, $\beta = \frac{\mathbb{E}[Y|X=c^+]-\mathbb{E}[Y|X=c^-]}{\mathbb{E}[D|X=c^+]-\mathbb{E}[D|X=c^-]}$.

How Do We Learn About Ratios?

- We often perform a hypothesis test for

$$\mathbb{H}_0 : \beta = \frac{a}{b} = 0 \quad \text{vs.} \quad \mathbb{H}_1 : \beta = \frac{a}{b} \neq 0.$$

How Do We Learn About Ratios?

- We often perform a hypothesis test for

$$\mathbb{H}_0 : \beta = \frac{a}{b} = 0 \quad \text{vs.} \quad \mathbb{H}_1 : \beta = \frac{a}{b} \neq 0.$$

- This means we need an approximation for the distribution of

$$\widehat{\beta} = \frac{\widehat{a}}{\widehat{b}},$$

where $\widehat{a} \xrightarrow{p} a$ and $\widehat{b} \xrightarrow{p} b$.

How Do We Learn About Ratios?

- We often perform a hypothesis test for

$$\mathbb{H}_0 : \beta = \frac{a}{b} = 0 \quad \text{vs.} \quad \mathbb{H}_1 : \beta = \frac{a}{b} \neq 0.$$

- This means we need an approximation for the distribution of

$$\widehat{\beta} = \frac{\widehat{a}}{\widehat{b}},$$

where $\widehat{a} \xrightarrow{p} a$ and $\widehat{b} \xrightarrow{p} b$.

- Typically, we rely on the delta method/Slutsky to get:

$$\widehat{\beta} = \frac{\widehat{a}}{\widehat{b}} \approx \frac{\mathcal{N}(a, \sigma_a^2)}{b}.$$

How Do We Learn About Ratios?

- We often perform a hypothesis test for

$$\mathbb{H}_0 : \beta = \frac{a}{b} = 0 \quad \text{vs.} \quad \mathbb{H}_1 : \beta = \frac{a}{b} \neq 0.$$

- This means we need an approximation for the distribution of

$$\widehat{\beta} = \frac{\widehat{a}}{\widehat{b}},$$

where $\widehat{a} \xrightarrow{p} a$ and $\widehat{b} \xrightarrow{p} b$.

- Typically, we rely on the delta method/Slutsky to get:

$$\widehat{\beta} = \frac{\widehat{a}}{\widehat{b}} \approx \frac{\mathcal{N}(a, \sigma_a^2)}{b}.$$

- Is this always a good approximation?

What is The Distribution of $\widehat{a}/\widehat{b}$?

- For simplicity, assume that:

$$\begin{bmatrix} \widehat{a} \\ \widehat{b} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix} \; ; \; \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

What is The Distribution of $\widehat{a}/\widehat{b}$?

- For simplicity, assume that:

$$\begin{bmatrix} \widehat{a} \\ \widehat{b} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix} ; \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

- When $\mu_a = \mu_b = \rho = 0$, $\frac{\widehat{a}}{\widehat{b}} \sim$ Std. Cauchy Distribution
  1. No moments!
  2. Symmetric around 0.

What is The Distribution of $\widehat{a}/\widehat{b}$?

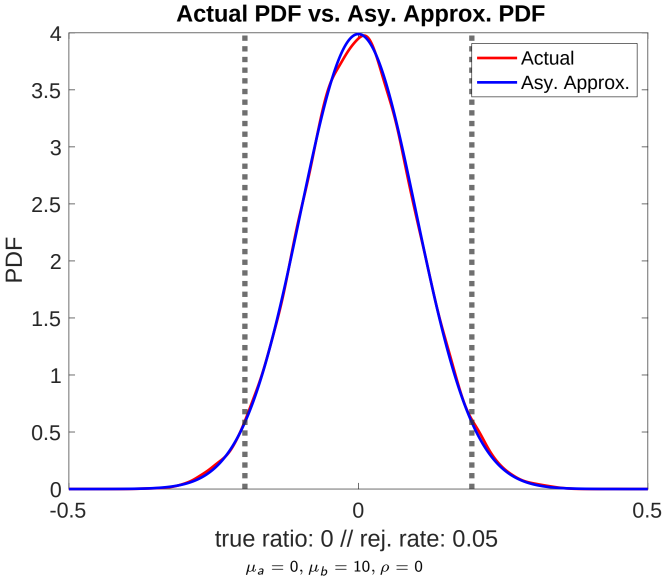
- For simplicity, assume that:

$$\begin{bmatrix} \widehat{a} \\ \widehat{b} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix} \; ; \; \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

- When $\mu_a = \mu_b = \rho = 0$, $\frac{\widehat{a}}{\widehat{b}} \sim$ Std. Cauchy Distribution
  1. No moments!
  2. Symmetric around 0.
- When $\mu_b \neq 0$ or $\rho \neq 0$, the distribution becomes asymmetric.

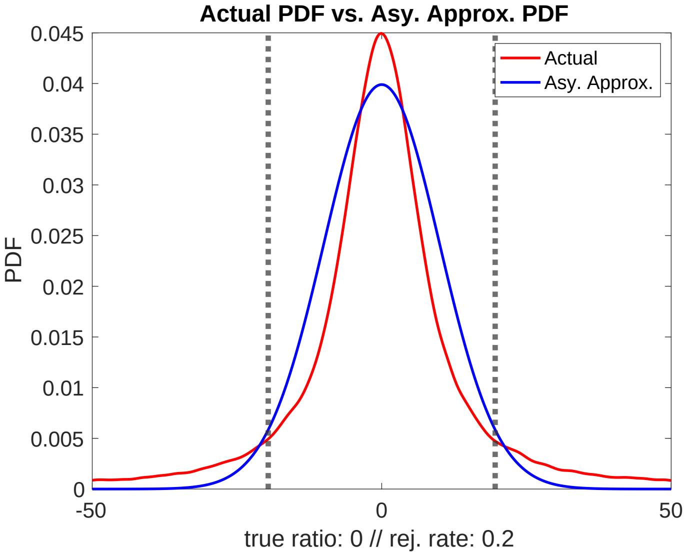
Approximating The Distribution of $\widehat{a}/\widehat{b}$

Ideal Case: μ_b is Large



Approximating The Distribution of $\widehat{a}/\widehat{b}$

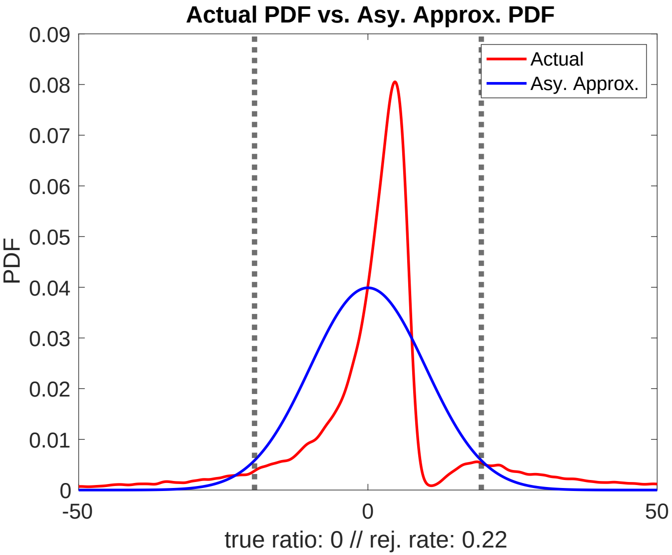
Problem: μ_b is Small



$\mu_a = 0, \mu_b = 0.1, \rho = 0$

Approximating The Distribution of $\widehat{a}/\widehat{b}$

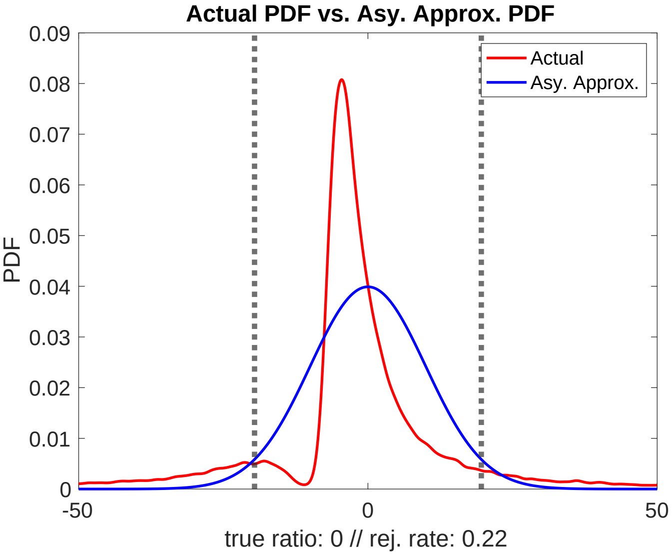
Asymmetry with $\rho > 0$



$\mu_a = 0$, $\mu_b = 0.1$, $\rho = 0.95$

Approximating The Distribution of $\widehat{a}/\widehat{b}$

Asymmetry with $\rho < 0$



$\mu_a = 0$, $\mu_b = 0.1$, $\rho = -0.95$

Takeaways So Far

- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.

Takeaways So Far

- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.
- You tend to reject more on one of the two sides depending on the sign of the correlation between $\widehat{a}$ and $\widehat{b}$.

Takeaways So Far

- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.

- You tend to reject more on one of the two sides depending on the sign of the correlation between $\widehat{a}$ and $\widehat{b}$.

- Do our examples suffer from this?

Takeaways So Far

- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.
- You tend to reject more on one of the two sides depending on the sign of the correlation between $\widehat{a}$ and $\widehat{b}$.
- Do our examples suffer from this?
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
     Okay, $VAR[X]$ is positive and typically "large"...

Takeaways So Far

- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.

- You tend to reject more on one of the two sides depending on the sign of the correlation between $\widehat{a}$ and $\widehat{b}$.

- Do our examples suffer from this?
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
     Okay, $VAR[X]$ is positive and typically "large"...
  2. Relative effects of `educ` and `exper` $\beta = \frac{\beta_{\texttt{educ}}}{\beta_{\texttt{exper}}}$.
     Issue when $\beta_{\texttt{exper}}$ is small.

Takeaways So Far

- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.

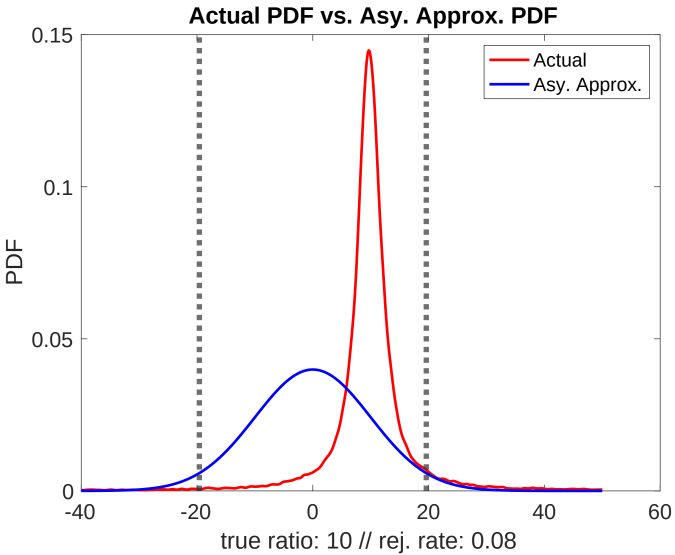
- You tend to reject more on one of the two sides depending on the sign of the correlation between $\widehat{a}$ and $\widehat{b}$.

- Do our examples suffer from this?
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
     Okay, $VAR[X]$ is positive and typically "large"...
  2. Relative effects of educ and exper $\beta = \frac{\beta_{\texttt{educ}}}{\beta_{\texttt{exper}}}$.
     Issue when $\beta_{\texttt{exper}}$ is small.
  3. IV, $\beta = \frac{COV[Z,Y]}{COV[Z,X]}$.
     Issue when instrument is weak, i.e., $COV[Z, X]$ is small.

Takeaways So Far

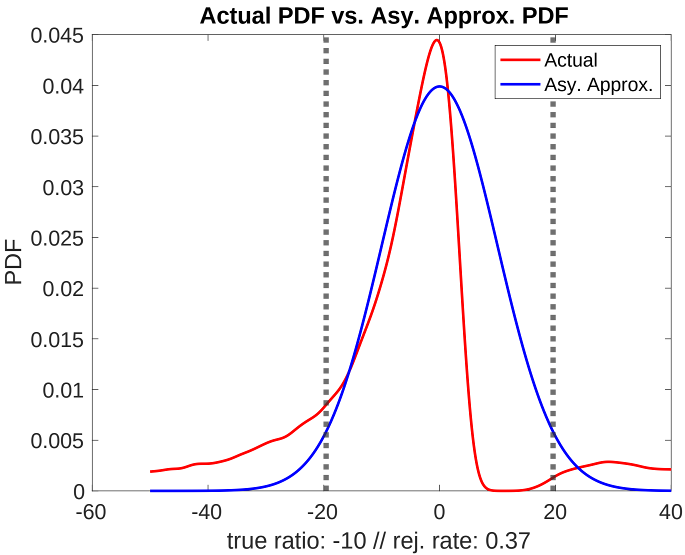
- Problem with the size of t-tests based on the conventional approximation of the distribution of $\widehat{a}/\widehat{b}$ when μ_b is small.
- You tend to reject more on one of the two sides depending on the sign of the correlation between $\widehat{a}$ and $\widehat{b}$.
- Do our examples suffer from this?
  1. OLS, $\beta = \frac{COV[X,Y]}{VAR[X]}$.
     Okay, $VAR[X]$ is positive and typically "large"...
  2. Relative effects of educ and exper $\beta = \frac{\beta_{\texttt{educ}}}{\beta_{\texttt{exper}}}$.
     Issue when $\beta_{\texttt{exper}}$ is small.
  3. IV, $\beta = \frac{COV[Z,Y]}{COV[Z,X]}$.
     Issue when instrument is weak, i.e., $COV[Z,X]$ is small.
  4. Fuzzy RD, $\beta = \frac{\mathbb{E}[Y|X=c^+]-\mathbb{E}[Y|X=c^-]}{\mathbb{E}[D|X=c^+]-\mathbb{E}[D|X=c^-]}$.
     Issue when compliance to threshold rule is small, i.e., $\mathbb{E}[D|X=c^+]-\mathbb{E}[D|X=c^-]$ is small.

What About Power?



$\mu_a = 1, \mu_b = 0.1, \rho = 0.95$

What About Power?



$\mu_a = -1$, $\mu_b = 0.1$, $\rho = 0.95$

The Power of Ratios

- The conventional t-test rejects null effects $\sim$ two times more often than small positive effects; **X**

The Power of Ratios

- The conventional t-test rejects null effects $\sim$ two times more often than small positive effects; ✗
- On the other hand, it rejects negative effects $\sim$ two times more often than null effects. ✓

The Power of Ratios

- The conventional t-test rejects null effects $\sim$ two times more often than small positive effects; X
- On the other hand, it rejects negative effects $\sim$ two times more often than null effects. ✓
- What is a solution to this problem? Given that $b \neq 0$,

$$\mathbb{H}_0 : \beta = \frac{a}{b} = 0 \Leftrightarrow \mathbb{H}_0 : a = 0.$$

The Power of Ratios

- The conventional t-test rejects null effects $\sim$ two times more often than small positive effects; **X**
- On the other hand, it rejects negative effects $\sim$ two times more often than null effects. ✓
- What is a solution to this problem? Given that $b \neq 0$,

$$\mathbb{H}_0 : \beta = \frac{a}{b} = 0 \Leftrightarrow \mathbb{H}_0 : a = 0.$$

- The Anderson-Rubin test is a t-test of $\mathbb{H}_0 : a = 0$.
  It is easier to approximate the distribution of $\widehat{a}$.

Figure 2 of Keane and Neal JoE 2023 shows the power curves.

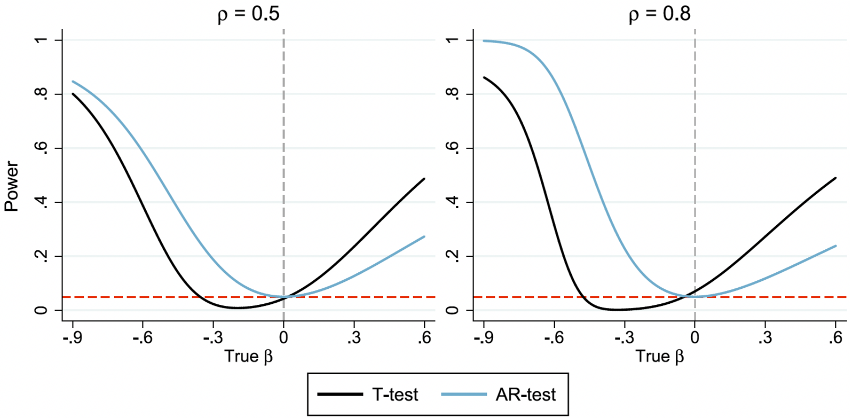


Fig. 2. Power of t-Test vs. AR-Test when $C = 10$ ($F_{5\%}$=23.1).

Fuzzy RD

- How do we estimate the ratio?

Fuzzy RD

- How do we estimate the ratio?
  1. Use local linear regression with triangular kernel to obtain:

     $$\widehat{\mathbb{E}}[Y|X=c^{+}],\ \widehat{\mathbb{E}}[Y|X=c^{-}],\ \widehat{\mathbb{E}}[D|X=c^{+}],\ \widehat{\mathbb{E}}[D|X=c^{-}],$$

     and

     $\widehat{\beta}=\frac{\widehat{\mathbb{E}}[Y|X=c^{+}]-\widehat{\mathbb{E}}[Y|X=c^{-}]}{\widehat{\mathbb{E}}[D|X=c^{+}]-\widehat{\mathbb{E}}[D|X=c^{-}]}.$

Fuzzy RD

- How do we estimate the ratio?
  1. Use local linear regression with triangular kernel to obtain:

     $\widehat{\mathbb{E}}[Y|X = c^+]$, $\widehat{\mathbb{E}}[Y|X = c^-]$, $\widehat{\mathbb{E}}[D|X = c^+]$, $\widehat{\mathbb{E}}[D|X = c^-]$,

     and

     $\widehat{\beta} = \frac{\widehat{\mathbb{E}}[Y|X=c^+]-\widehat{\mathbb{E}}[Y|X=c^-]}{\widehat{\mathbb{E}}[D|X=c^+]-\widehat{\mathbb{E}}[D|X=c^-]}$.
  2. Choose what "local" means, i.e., choose the bandwidth.

Fuzzy RD

- How do we estimate the ratio?
  1. Use local linear regression with triangular kernel to obtain:

     $$\widehat{\mathbb{E}}[Y|X=c^+],\ \widehat{\mathbb{E}}[Y|X=c^-],\ \widehat{\mathbb{E}}[D|X=c^+],\ \widehat{\mathbb{E}}[D|X=c^-],$$

     and

     $\widehat{\beta} = \frac{\widehat{\mathbb{E}}[Y|X=c^+]-\widehat{\mathbb{E}}[Y|X=c^-]}{\widehat{\mathbb{E}}[D|X=c^+]-\widehat{\mathbb{E}}[D|X=c^-]}$.
  2. Choose what "local" means, i.e., choose the bandwidth.
  3. Perform bias correction.

Fuzzy RD

- How do we estimate the ratio?
  1. Use local linear regression with triangular kernel to obtain:

     $\widehat{\mathbb{E}}[Y|X=c^+]$, $\widehat{\mathbb{E}}[Y|X=c^-]$, $\widehat{\mathbb{E}}[D|X=c^+]$, $\widehat{\mathbb{E}}[D|X=c^-]$,

     and

     $\widehat{\beta} = \frac{\widehat{\mathbb{E}}[Y|X=c^+]-\widehat{\mathbb{E}}[Y|X=c^-]}{\widehat{\mathbb{E}}[D|X=c^+]-\widehat{\mathbb{E}}[D|X=c^-]}$.
  2. Choose what "local" means, i.e., choose the bandwidth.
  3. Perform bias correction.
- There is some disagreement on points (2) and (3).

Fuzzy RD

- How do we estimate the ratio?
  1. Use local linear regression with triangular kernel to obtain:

     $$\widehat{\mathbb{E}}[Y|X=c^+], \widehat{\mathbb{E}}[Y|X=c^-], \widehat{\mathbb{E}}[D|X=c^+], \widehat{\mathbb{E}}[D|X=c^-],$$

     and

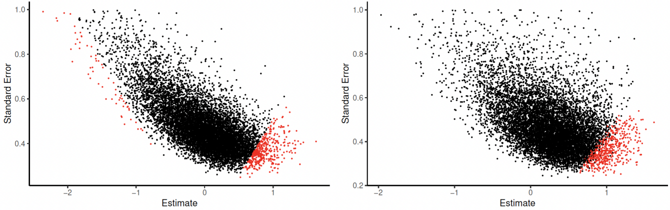
     $$\widehat{\beta} = \frac{\widehat{\mathbb{E}}[Y|X=c^+]-\widehat{\mathbb{E}}[Y|X=c^-]}{\widehat{\mathbb{E}}[D|X=c^+]-\widehat{\mathbb{E}}[D|X=c^-]}.$$

  2. Choose what "local" means, i.e., choose the bandwidth.
  3. Perform bias correction.
- There is some disagreement on points (2) and (3).
- This paper studies several methods but focuses on two:
  1. RBC - methods by Cattaneo et al.
  2. BSD - methods by Kolesar et al.

Simulation Results

t-test

Figure 9: **DGP 4: Effect of Admission to Education Program on Wages, $E[Y|R]$ Quadratic in R.**



Notes: Based on 10,000 replications with 2000 observations each. *Left:* Robust bias-corrected inference (RBC) via the rdrobust R package. *Right:* Bounded second derivative (BSD) inference via the RDHonest R package. Both use a uniform kernel. We plot $\hat{\beta}_1$ against $se(\hat{\beta}_1)$. Runs with a standard error > 1 excluded. Red dots indicate $H_0{:}\beta = 0$ rejected at 5% level.

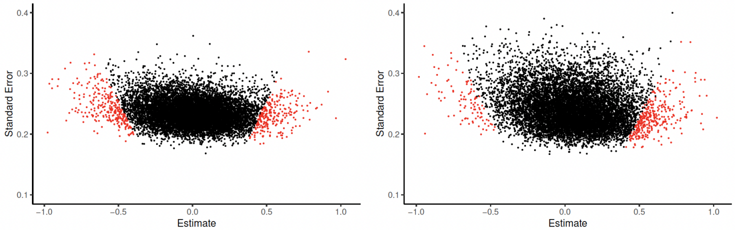
Table 1: **RBC and BSD Inference with Uniform Kernel on DGPs 1 to 4**

	RBC							BSD						
DGP	Reject	$\% > 0$	Median: $\hat{\beta}$	SE	F	N^e		Reject	$\% > 0$	Median: $\hat{\beta}$	SE	F	N^e	
1	5.1%	97.7%	-.003	.207	32.04	508		9.8%	100%	.094	.185	24.40	334	
2	4.5%	76.1%	-.005	.417	63.17	498		3.3%	99.7%	.139	.413	44.93	341	
3	4.5%	79.7%	.004	.442	49.59	451		3.2%	99.7%	.157	.438	36.92	303	
4	4.7%	84.0%	.012	.456	40.00	357		5.3%	100%	.284	.451	33.41	289	

Notes: Summary results for 10,000 artificial datasets of size $N = 2000$ each. The 4 rows report results for the 4 DGPs discussed in Sections 4.1 to 4.5. RBC and BSD indicate results from the rdrobust and RDHonest packages, respectively. Both use a uniform kernel. We report the rate of rejecting $H_0{:}\beta = 0$, the fraction of these rejections that occur when $\hat{\beta} > 0$, and the medians of the estimate, estimated standard error, first stage F, and effective observations N^e.

Simulation Results
AR test

Figure 12: **Reduced Form Inference: DGP 4, $E[Y|R]$ Quadratic in R**



Notes: Based on 10,000 replications with N=2000 each. *Left:* RBC inference via the rdrobust R package. *Right:* BSD inference via the RDHonest R package. Both use a uniform kernel. We plot $\hat{\xi}$ against $se(\hat{\xi})$. Red dots indicate $H_0{:}\beta = 0$ rejected at 5% level. For AR inference this corresponds to cases where $\hat{\xi}_1$ is significant at the 5% level.

Discussion

- Connection with literature on impossible inference, e.g., Bertanha and Moreira JoE 2020.

Discussion

- Connection with literature on impossible inference, e.g., Bertanha and Moreira JoE 2020.
  - Type A, trivial power, e.g., testing means, sharp RD;
  - Type B: unbounded confidence sets, e.g., ratios, IV;
  - Fuzzy RD has both!

Discussion

- Connection with literature on impossible inference, e.g., Bertanha and Moreira JoE 2020.
  - Type A, trivial power, e.g., testing means, sharp RD;
  - Type B: unbounded confidence sets, e.g., ratios, IV;
  - Fuzzy RD has both!
- Noack and Rothe ECMA 2024 propose AR test + BSD as a solution to both problems but simulations here show they do a little worse than AR test + RBC.

Discussion

- Connection with literature on impossible inference, e.g., Bertanha and Moreira JoE 2020.
  - Type A, trivial power, e.g., testing means, sharp RD;
  - Type B: unbounded confidence sets, e.g., ratios, IV;
  - Fuzzy RD has both!
- Noack and Rothe ECMA 2024 propose AR test + BSD as a solution to both problems but simulations here show they do a little worse than AR test + RBC.
- Why?
  - Plot simulated power curves;
  - Optimality of BSD?
  - Run simulations under worst-case DGP.