13. More on RDDs

ISS5096 || ECI

Jaewon ("Jay-one") Yoo
National Tsing Hua University

1/ Fuzzy Regression Discontinuity Designs

The Setup

- Recall treatment: $D_i = 1$ or $D_i = 0$ and forcing variable: X_i.
- **Fuzzy RD:** discontinuity in the probability of treatment.

$$\lim_{x \uparrow c} \mathbb{P}[D_i = 1 \mid X_i = x] \neq \lim_{x \downarrow c} \mathbb{P}[D_i = 1 \mid X_i = x]$$

  - No longer deterministic function of forcing variable.
  - SRD is a special case of the FRD.

The Setup

- Recall treatment: $D_i = 1$ or $D_i = 0$ and forcing variable: X_i.
- **Fuzzy RD:** discontinuity in the probability of treatment.

$$\lim_{x \uparrow c} \mathbb{P}[D_i = 1 \mid X_i = x] \neq \lim_{x \downarrow c} \mathbb{P}[D_i = 1 \mid X_i = x]$$

    - No longer deterministic function of forcing variable.
    - SRD is a special case of the FRD.
- Common use case: threshold allows participation in program.
    - Some might not participate even if allowed (noncompliance).

The Setup

- Recall treatment: $D_i = 1$ or $D_i = 0$ and forcing variable: X_i.
- **Fuzzy RD:** discontinuity in the probability of treatment.

$$\lim_{x \uparrow c} \mathbb{P}[D_i = 1 \mid X_i = x] \neq \lim_{x \downarrow c} \mathbb{P}[D_i = 1 \mid X_i = x]$$

  - No longer deterministic function of forcing variable.
  - SRD is a special case of the FRD.
- Common use case: threshold allows participation in program.
  - Some might not participate even if allowed (noncompliance).
- Forcing variable is an **instrument**:
  - Affects Y_i, but only through D_i (at the threshold).

Fuzzy RD in a Graph (Imbens and Leminux, 2008)

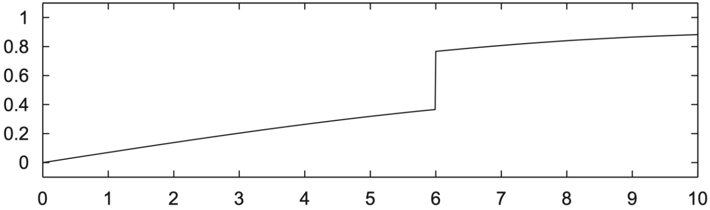


Fig. 3. Assignment probabilities (FRD).

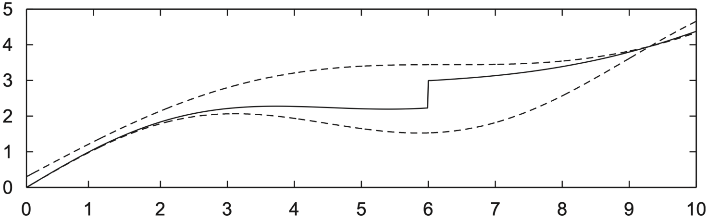


Fig. 4. Potential and observed outcome regression (FRD).

Source: *Figure 4* in Imbens, Guido W., and Thomas Lemieux. "Regression discontinuity designs: A guide to practice." Journal of econometrics 142, no. 2 (2008): 615-635.

Fuzzy RD in a Graph (Imbens and Leminux, 2008)

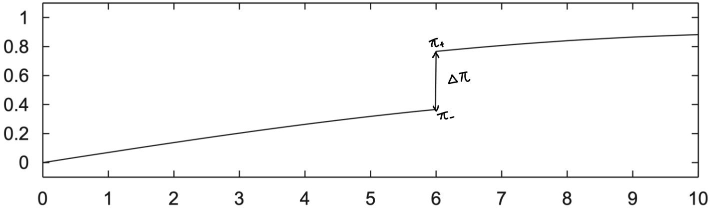


Fig. 3. Assignment probabilities (FRD).

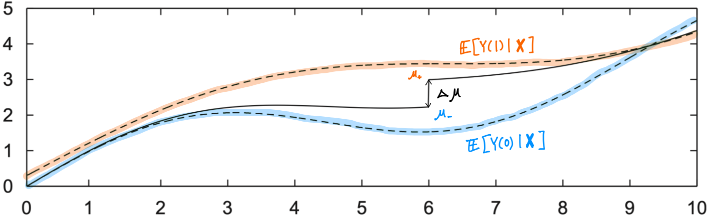


Fig. 4. Potential and observed outcome regression (FRD).

Source: *Figure 4* in Imbens, Guido W., and Thomas Lemieux. "Regression discontinuity designs: A guide to practice." Journal of econometrics 142, no. 2 (2008): 615-635.

Fuzzy RD assumptions

- $D_i(x)$ is potential value of treatment as cutoff changes around c.
  - $D_i(x) = 1$ if unit i would take treatment if cutoff were x.
  - $D_i(x) = 0$ if unit i would take control if cutoff were x.

Fuzzy RD assumptions

- $D_i(x)$ is potential value of treatment as cutoff changes around c.
    - $D_i(x) = 1$ if unit i would take treatment if cutoff were x.
    - $D_i(x) = 0$ if unit i would take control if cutoff were x.
- **Monotonicity assumption:** $D_i(x)$ is non-increasing in x.
    - Lowering the cutoff can only increase participation.

Fuzzy RD assumptions

- $D_i(x)$ is potential value of treatment as cutoff changes around c.
  - $D_i(x) = 1$ if unit i would take treatment if cutoff were x.
  - $D_i(x) = 0$ if unit i would take control if cutoff were x.
- **Monotonicity assumption:** $D_i(x)$ is non-increasing in x.
  - Lowering the cutoff can only increase participation.
- Compliers are those i such that for all $0 < e < \varepsilon$:

$$D_i(c - e) = 1 \quad \text{and} \quad D_i(c + e) = 0$$

  - Lowering or increasing the threshold would affect their treatment status.
  - Compliance status unobservable
    - $\rightsquigarrow$ Principal strata (Frangakis and Rubin, 2002. Biometrics).

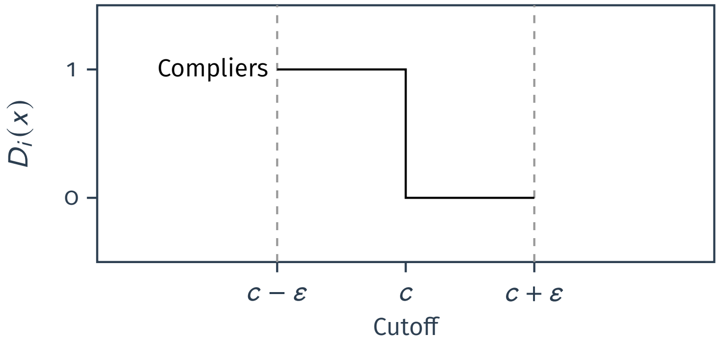
Fuzzy RD assumptions

- $D_i(x)$ is potential value of treatment as cutoff changes around c.
  - $D_i(x) = 1$ if unit i would take treatment if cutoff were x.
  - $D_i(x) = 0$ if unit i would take control if cutoff were x.
- **Monotonicity assumption:** $D_i(x)$ is non-increasing in x.
  - Lowering the cutoff can only increase participation.
- Compliers are those i such that for all $0 < e < \varepsilon$:

$$D_i(c - e) = 1 \quad \text{and} \quad D_i(c + e) = 0$$

  - Lowering or increasing the threshold would affect their treatment status.
  - Compliance status unobservable
    - $\rightsquigarrow$ Principal strata (Frangakis and Rubin, 2002. Biometrics).
- Example: college students that get above a certain GPA are encouraged to apply to grad school.
  - Compliers wouldn't apply if threshold were slightly higher.
  - Compliers would apply if the threshold were slightly lower.

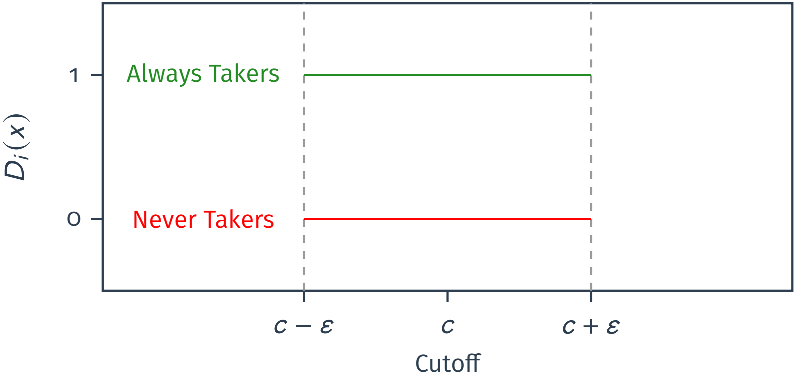
Compliance Graph



- Compliers would not take the treatment if they had $X_i = c$ and we increased the cutoff by some small amount.
- These are compliers at the threshold.

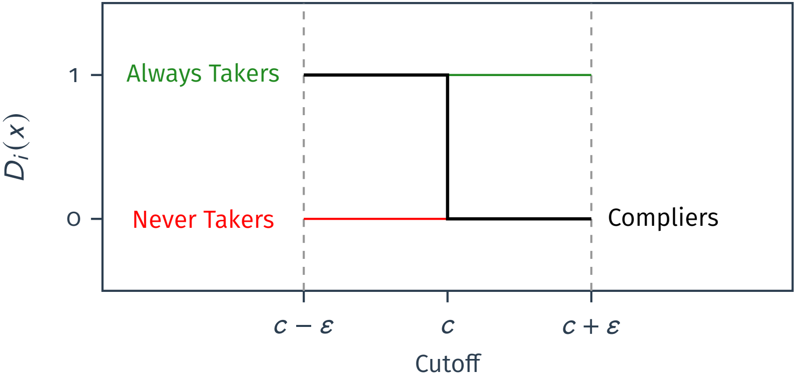
Compliance Groups

- Compliers: $D_i(c - e) = 1 \quad \text{and} \quad D_i(c + e) = 0$
- Always-takers: $D_i(c + e) = D_i(c - e) = 1$
- Never-takers: $D_i(c + e) = D_i(c - e) = 0$



Compliance Groups

- Compliers: $D_i(c - e) = 1 \quad$ and $\quad D_i(c + e) = 0$
- Always-takers: $D_i(c + e) = D_i(c - e) = 1$
- Never-takers: $D_i(c + e) = D_i(c - e) = 0$



LATE in the Fuzzy RD

- We can define an estimator that is in the spirit of IV:

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

$$= \frac{\text{effect of threshold on } Y_i}{\text{effect of threshold on } D_i}$$

LATE in the Fuzzy RD

- We can define an estimator that is in the spirit of IV:

$$\tau_{\text{FRD}} = \frac{\lim_{x \downarrow c} \mathbb{E}[Y_i \mid X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i \mid X_i = x]}{\lim_{x \downarrow c} \mathbb{E}[D_i \mid X_i = x] - \lim_{x \uparrow c} \mathbb{E}[D_i \mid X_i = x]}$$
$$= \frac{\text{effect of threshold on } Y_i}{\text{effect of threshold on } D_i}$$

- Under the FRD assumptions (continuity, consistency, and monotonicity), we can write that the estimator is equal to the effect at the threshold for compliers.

$$\tau_{\text{FRD}} = \mathbb{E}[\tau_i \mid i \text{ is a complier}, \; X_i = c]$$

  - Proof is very similar to the LATE proof.

LATE in the Fuzzy RD

- We can define an estimator that is in the spirit of IV:

$$\tau_{\text{FRD}} = \frac{\lim_{x \downarrow c} \mathbb{E}[Y_i \mid X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i \mid X_i = x]}{\lim_{x \downarrow c} \mathbb{E}[D_i \mid X_i = x] - \lim_{x \uparrow c} \mathbb{E}[D_i \mid X_i = x]}$$
$$= \frac{\text{effect of threshold on } Y_i}{\text{effect of threshold on } D_i}$$

- Under the FRD assumptions (continuity, consistency, and monotonicity), we can write that the estimator is equal to the effect at the threshold for compliers.

$$\tau_{\text{FRD}} = \mathbb{E}[\tau_i \mid i \text{ is a complier}, \;\; X_i = c]$$

  - Proof is very similar to the LATE proof.

- External validity? Doubly local $\rightsquigarrow$ careful about generalizing.

Estimation in FRD

- Remember that we had:

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

- Ratio of SRD estimands: use local linear regression for both.

$$\widehat{\tau}_{\text{FRD}} = \frac{\widehat{\tau}_{Y,\text{SRD}}}{\widehat{\tau}_{D,\text{SRD}}}$$

- Again, CCT provides robust bias correction, bandwidths.
  - Calonico, Cattaneo, and Titiunik (CCT; Econometrica 2014)

More practical FRD estimation

- The ratio estimator above is equivalent to a TSLS approach.
- Use the same specification as above with the following covariates:

$$V_i = \begin{pmatrix} 1 \\ \mathbf{1}\{X_i < c\}(X_i - c) \\ \mathbf{1}\{X_i \geq c\}(X_i - c) \end{pmatrix}$$

More practical FRD estimation

- The ratio estimator above is equivalent to a TSLS approach.
- Use the same specification as above with the following covariates:

$$V_i = \begin{pmatrix} 1 \\ \mathbf{1}\{X_i < c\}(X_i - c) \\ \mathbf{1}\{X_i \geq c\}(X_i - c) \end{pmatrix}$$

- First stage:

$$D_i = \delta_1' V_i + \rho \mathbf{1}\{X_i \geq c\} + \nu_i$$

- Second stage:

$$Y_i = \delta_2' V_i + \tau D_i + \eta_i$$

More practical FRD estimation

- The ratio estimator above is equivalent to a TSLS approach.
- Use the same specification as above with the following covariates:

$$V_i = \begin{pmatrix} 1 \\ \mathbf{1}\{X_i < c\}(X_i - c) \\ \mathbf{1}\{X_i \geq c\}(X_i - c) \end{pmatrix}$$

- First stage:

$$D_i = \delta_1' V_i + \rho \mathbf{1}\{X_i \geq c\} + \nu_i$$

- Second stage:

$$Y_i = \delta_2' V_i + \tau D_i + \eta_i$$

- Thus, being above the threshold is treated like an instrument, controlling for trends in X_i.

2/ Example R Codes

Free Tutoring Program and Academic Performance

- The Setting:
  - Students take an entrance exam at the beginning & end of a school year.
  - Those who scored below 70 are enrolled in a free tutoring program.

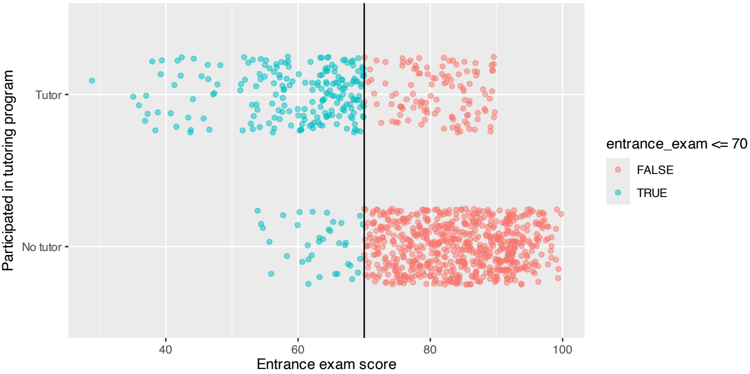
Name	Description
id	student id
entrance_exam	student's entrance exam score (out of 100)
exit_exam	student's exit exam score (out of 100)
tutoring	indicator showing if the student received tutoring

```
> pacman::p_load(tidyverse, broom, modelsummary, kableExtra, estimatr, rdrobust)
> tutoring <- read_csv("https://bit.ly/453ymbJ"); head(tutoring, 3)

# A tibble: 4 × 5
     id entrance_exam tutoring tutoring_text exit_exam
  <dbl>         <dbl> <lgl>    <chr>             <dbl>
1     1          92.4 FALSE    No tutor           78.1
2     2          72.8 FALSE    No tutor           58.2
3     3          53.7 TRUE     Tutor              62.0
```

Examine Compliance Around the Cutoff

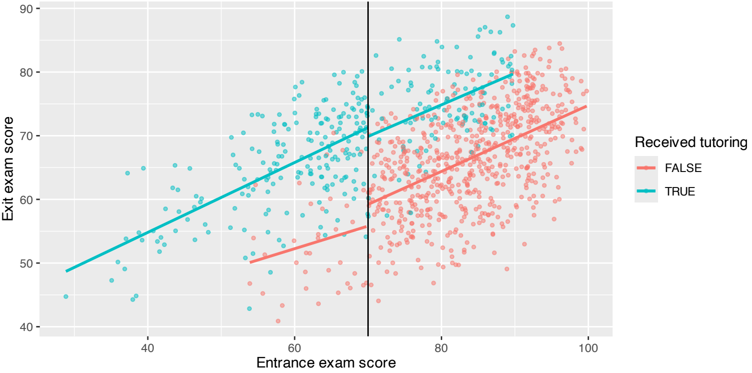
```
# Noncompliance around the cutoff
> tutoring |>
    group_by(tutoring, entrance_exam <= 70) |>
    summarize(count = n()) |>
    group_by(tutoring) |>
    mutate(prop = count / sum(count))

# A tibble: 4 × 4
# Groups:   tutoring [2]
  tutoring `entrance_exam <= 70` count   prop
  <lgl>    <lgl>                 <int>  <dbl>
1 FALSE    FALSE                   646 0.947
2 FALSE    TRUE                     36 0.0528
3 TRUE     FALSE                   116 0.365
4 TRUE     TRUE                    202 0.635

# Visualize Noncompliance
> ggplot(tutoring, aes(x = entrance_exam, y = tutoring_text, color = entrance_exam <= 70)) +
    geom_point(size = 1.5, alpha = 0.5,
               position = position_jitter(width = 0, height = 0.25, seed = 1234)) +
    geom_vline(xintercept = 70) +
    labs(x = "Entrance exam score", y = "Participated in tutoring program")
```

Visualizing Noncompliance around the Cutoff



Visualizing the Fuzzy Gap

```r
> ggplot(tutoring, aes(x = entrance_exam, y = tutoring_text, color = entrance_exam <= 70)) +
    geom_point(size = 1.5, alpha = 0.5,
               position = position_jitter(width = 0, height = 0.25, seed = 1234)) +
    geom_vline(xintercept = 70) +
    labs(x = "Entrance exam score", y = "Participated in tutoring program")
```



Measuring the Fuzzy Gap

- Center the forcing variable, `entrance_score` and generate V_i:

```
> tutoring_centered <- tutoring %>%
    mutate(entrance_centered = entrance_exam - 70,
           entrance_cnt_b70 = if_else(entrance_exam < 70, entrance_centered, 0),
           entrance_cnt_a70 = if_else(entrance_exam >= 70, entrance_centered, 0),
           below_cutoff = entrance_exam <= 70)
```

- If this was a Sharp RD:

```
# Bandwidth ±10
> sharp_parametric <- lm(exit_exam ~ entrance_centered + tutoring,
                         data = filter(tutoring_centered,
                                       entrance_centered >= -10 &
                                         entrance_centered <= 10))
> tidy(sharp_parametric)
# A tibble: 3 × 5
  term              estimate std.error statistic   p.value
  <chr>                <dbl>     <dbl>     <dbl>     <dbl>
1 (Intercept)         59.3      0.503     118.   9.75e-313
2 entrance_centered    0.511    0.0665      7.69 1.17e- 13
3 tutoringTRUE        11.5      0.744      15.4  1.77e- 42
```

- `tutoringTRUE` estimate measures the size of the jump at the cutoff.
    - There would still be confounding as we have noncompliance.

Measuring the Fuzzy Gap Correctly

- Use `estimatr::ivrobust()` for TSLS estimation:

```
> model_fuzzy <- iv_robust(
    exit_exam ~ entrance_cnt_b70 + entrance_cnt_a70 + tutoring |
      entrance_cnt_b70 + entrance_cnt_a70 + below_cutoff,
    data = filter(tutoring_centered, entrance_centered >= -10 & entrance_centered <= 10)
  )

> tidy(model_fuzzy, conf.int = F)
              term   estimate std.error statistic       p.value  df   outcome
1      (Intercept) 59.9731656 1.1165541 53.712726 9.984384e-185 399 exit_exam
2 entrance_cnt_b70  0.3773301 0.1855677  2.033383  4.267520e-02 399 exit_exam
3 entrance_cnt_a70  0.4715826 0.1344630  3.507156  5.044279e-04 399 exit_exam
4     tutoringTRUE  9.6265535 1.9424192  4.955961  1.065759e-06 399 exit_exam
```

Measuring the Fuzzy Gap Correctly

- Nonparametric estimation for FRD with `rdrobust()`:

```
> frd <- rdrobust::rdrobust(y = tutoring$exit_exam, x = tutoring$entrance_exam,
                    c = 70, fuzzy = tutoring$tutoring, kernel = "triangular", vce ="hc2")
> summary(frd)

Fuzzy RD estimates using local polynomial regression.

Number of Obs.                 1000
BW type                       mserd
Kernel                   Triangular
VCE method                      HC2

Number of Obs.                  238          762
Eff. Number of Obs.             170          347
Order est. (p)                    1            1
Order bias  (q)                   2            2
BW est. (h)                  12.961       12.961
BW bias (b)                  19.579       19.579
rho (h/b)                     0.662        0.662
Unique Obs.                     238          762

First-stage estimates.

=============================================================================
        Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
=============================================================================
  Conventional    -0.708     0.073    -9.672     0.000    [-0.851 , -0.565]
        Robust         -         -    -8.350     0.000    [-0.909 , -0.563]
=============================================================================

Treatment effect estimates.

=============================================================================
        Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
=============================================================================
  Conventional     9.685     1.957     4.948     0.000     [5.849 , 13.522]
        Robust         -         -     4.127     0.000     [5.075 , 14.257]
=============================================================================
```

Measuring the Fuzzy Gap Correctly

- Nonparametric estimation for FRD with `rdrobust()`:
- Report the “robust” estimates: $\widehat{\tau}_{\mathrm{BC}}$ with $\widehat{\sigma}^2_{\mathrm{robust}}$.

```
> cbind(frd$coef, frd$se)

                   Coeff Std. Err.
Conventional   9.685196  1.957433
Bias-Corrected 9.666085  1.957433
Robust         9.666085  2.342204
```

On to the Presentations & Discussions!

Contact Information:
jaewon.yoo@iss.nthu.edu.tw
https://j1yoo4.github.io/

Appendix

Kink RD

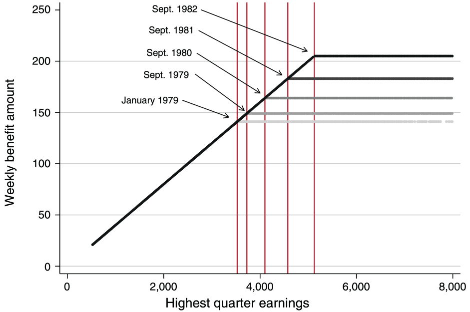


Figure 2. Louisiana: Schedule of UI Weekly Benefit Amount, Jan. 1979–Dec. 1983

Source: https://www.aeaweb.org/articles?id=10.1257/pol.20130248

Also see: https://blogs.worldbank.org/en/impactevaluations/tools-trade-regression-kink-design

- **Sharp Kink RD**: discontinuities in the first derivatives rather than levels.
    - Unemployment benefits as a function of prior earnings.
    - If there is a cap on benefits, there's a kink in the assignment.
    - Look for changes in the slope of $\mathbb{E}[Y_i \mid X_i = x]$ at threshold.
    - Estimation Similar, but better to use local quadratic regression.