# Fuzzy RD designs

## Northwestern Causal Inference Workshop

Gonzalo Vazquez-Bare

Department of Economics, UC Santa Barbara

August 1, 2024

# Participation in the Head Start program

- In 1965, some counties received assistance to develop HS proposals
    - HS participation 50-100% higher in counties that received assistance
- We have been using HS assistance and participation interchangeably
- Some counties with assistance may not have participated
- Some counties without assistance may have participated
- We have *imperfect compliance*

# Other examples

- Antipoverty programs
  - Eligibility based on poverty index
  - Administrators may decide to treat some ineligible households
  - Some eligible households may fail to get treatment
- Scholarship for higher education
  - Assigned to students who score above some cutoff
  - Students above $c$ with other funding sources may refuse treatment
  - Students below $c$ may manage to get a second chance

# Overview

- Sharp RD: score perfectly determines treatment status
  - All units scoring above the cutoff receive the treatment
  - $D_i = \mathbb{1}(X_i \geq c)$
  - Probability of treatment jumps from 0 to 1 at $c$
- Fuzzy designs: imperfect compliance
  - Some units below $c$ may be treated or vice versa
  - Jump in probability at $c$ may be $< 1$ (but $> 0$)

# Setup

- Treatment assignment $Z_i = \mathbb{1}(X_i \geq c)$
  - Assigned to treatment if above the cutoff
- Treatment status $D_i = D(X_i)$
  - Endogenous: result of individual decisions
- In a sharp design, $D_i = Z_i$
- The ideas in FRD are borrowed from the IV literature
  - Imbens and Angrist (1994), Angrist, Imbens and Rubin (1996)

# Setup

Treatment status of unit $i$ "slightly" above / below $c$:

$$D_{1i} = \lim_{x \downarrow c} D_i(x), \quad D_{0i} = \lim_{x \uparrow c} D_i(x)$$

- Always-takers: $D_{1i} = D_{0i} = 1$
- Never-takers: $D_{1i} = D_{0i} = 0$
- Compliers: $D_{1i} = 1, D_{0i} = 0$
- Defiers: $D_{1i} = 0, D_{0i} = 1$

# Intention-to-treat (ITT) parameter

- ITT: effect of being assigned to treatment
- Sharp RD design on the treatment assignment variable

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

- Under some continuity assumptions,

$$\tau_{\mathsf{ITT}} = \mathbb{E}[\underbrace{(Y_i(1) - Y_i(0))}_{\tau_i}(\underbrace{D_{1i} - D_{0i}}_{\substack{= \ 1 \text{ for compliers} \\ =-1 \text{ for defiers} \\ = \ 0 \text{ for at, nt}}})|X_i = c]$$

# Intention-to-treat (ITT) parameter

$$D_{1i} - D_{0i} = \begin{cases} 1 & \text{for compliers} \\ -1 & \text{for defiers} \\ 0 & \text{for always-takers, never-takers} \end{cases}$$

Identification problem:

- Defiers receive the treatment below the cutoff but not above
- TE for compliers and defiers enter with opposite signs
- ITT could be $\leq 0$ even if $\tau_i > 0$ for all $i$

# The monotonicity assumption

- To restore identification, we will rule out the presence of defiers:

$$\mathbb{P}[\text{defier}|X_i = c] = 0$$

- This assumption is called *monotonicity*, since it implies that:

$$D_{1i} \geq D_{0i}, \quad \forall\, i$$

- Intuition: $X_i \geq c$ does not decrease the probability of treatment

# Intention-to-treat (ITT) parameter

- $D_{1i} - D_{0i} = 1$ for compliers, 0 for always-takers and never-takers
- Then

$$\tau_{\text{ITT}} = \underbrace{\mathbb{E}[Y_i(1) - Y_i(0)|X_i = c, D_{1i} > D_{0i}]}_{\text{ATE on compliers: LATE}} \times \underbrace{\mathbb{P}[D_{1i} > D_{0i}|X_i = c]}_{\text{prop of compliers}}$$

- ITT can be $\approx 0$ even if LATE is large
- But still a policy relevant parameter:
  - Effect of offering the treatment

# First stage

- First stage: effect of treatment assignment on treatment status:

$$\tau_{\mathsf{FS}} = \lim_{x \downarrow c} \mathbb{E}[D_i | X_i = x] - \lim_{x \uparrow c} \mathbb{E}[D_i | X_i = x]$$

- Under monotonicity,

$$\tau_{\mathsf{FS}} = \mathbb{P}[D_{1i} > D_{0i} | X_i = c] = \mathbb{P}[\text{complier} | X_i = c]$$

- First stage identifies the proportion of compliers at the cutoff

# Recovering the ATE on compliers

- From the previous reasoning,

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

  where

$$\tau_{\mathsf{FRD}} = \mathbb{E}[Y_i(1) - Y_i(0) | X_i = c, D_{1i} > D_{0i}]$$

- Fuzzy RD parameter is “doubly local”:
  - At the cutoff
  - On the subpopulation of compliers

# Estimation in fuzzy designs

- ITT and FS are sharp RD estimators
- FRD is a ratio of two sharp RDs
  - For local constant regression, equivalent to 2SLS
- Can adapt all previous tools to this case
  - Data driven bandwidth selection
  - Local polynomial estimation
  - Robust bias-corrected inference

# Empirical example

- Amarante, Manacorda, Miguel and Vigorito (AEJ, 2016)
  - Effect of cash transfer on health and employment
- *Plan de Atención Nacional a la Emergencia Social* (PANES)
  - Cash transfer to poor households in Uruguay
  - Assigned to households below an income threshold
  - Eligibility rules were not perfectly enforced