



# REGRESSION DISCONTINUITY (RD)

## Technical Track Session V

**Emanuela Galasso**  
The World Bank

These slides were developed by Christel Vermeersch and modified by Emanuela Galasso for the purpose of this workshop

## Main Objective of an Evaluation (Reminder)

- Estimate the effect of an intervention  $D$  on a results indicator  $Y$ .
- For example:
  - What is the effect of an increase in the minimum wage on employment?
  - What is the effect of a school meals program on learning achievement?
  - What is the effect of a job training program on employment and on wages?



## Indexes are common in targeting of social programs

- Anti-poverty programs → targeted to households below a given poverty index.
- Pension programs → targeted to population above a certain age.
- Scholarships → targeted to students with high scores on standardized test.
- CDD programs → awarded to NGOs that achieve highest scores.



## Regression discontinuity

- When to use this method?
  - The beneficiaries/non-beneficiaries can be ordered along a **quantifiable dimension**.
  - This dimension can be used to compute a well-defined **index or parameter**.
  - The index/parameter has a **cut-off point** for eligibility.
  - The **index value** is what drives the assignment of a potential beneficiary to the treatment (or to non-treatment.)
- Intuitive explanation of the method:
  - The potential beneficiaries (units) just above the cut-off point are very similar to the potential beneficiaries just below the cut-off.
  - We compare outcomes for units just above and below the cutoff.



## Example: Effect of cash transfer on consumption

**Goal** Target transfer to poorest households

## Method

- Construct poverty index from 1 to 100 with pre-intervention characteristics
- Households with a score  $\leq 50$  are poor
- Households with a score  $>50$  are non-poor

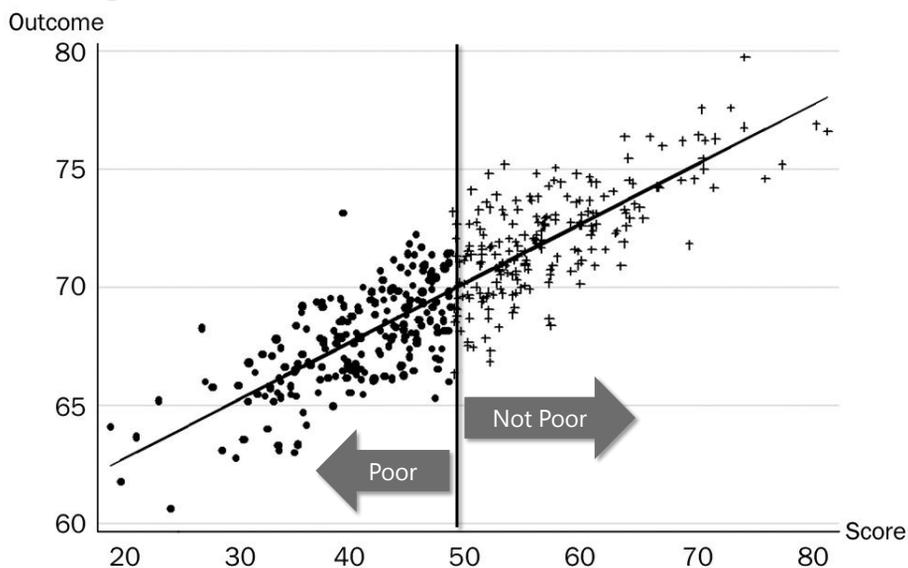
**Implementation** Cash transfer to poor households

## Evaluation

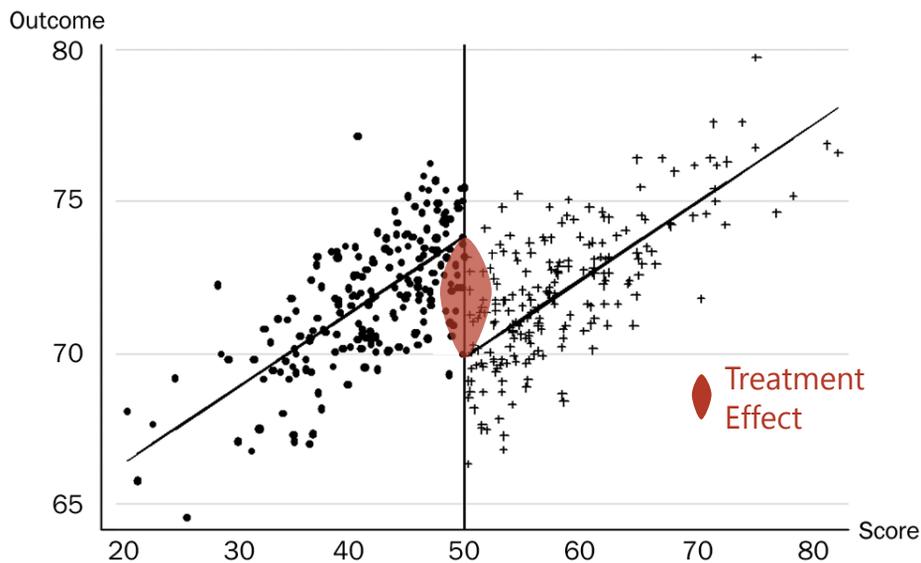
Measure outcomes (i.e. consumption, school attendance rates) before and after transfer, comparing households just above and below the cut-off point.



## Regression Discontinuity Design-Baseline



## Regression Discontinuity Design-Post Intervention



## Sharp and Fuzzy Discontinuity

### Sharp discontinuity

- The discontinuity precisely determines treatment
- Equivalent to random assignment in a neighborhood
- E.g. Social security payment depend directly and immediately on a person's age

### Fuzzy discontinuity

- Discontinuity is highly correlated with treatment .
- E.g. Rules determine eligibility but there is a margin of administrative error.
- Use the assignment as an IV for program participation.

## Identification for sharp discontinuity

$$y_i = \beta_0 + \beta_1 D_i + \delta(\text{score}_i) + \varepsilon_i$$

$$D_i = \begin{cases} 1 & \text{If household } i \text{ receives transfer} \\ 0 & \text{If household } i \text{ does not receive transfer} \end{cases}$$

$\delta(\text{score}_i)$  = Function that is continuous around the cut-off point

- Assignment rule under sharp discontinuity:

$$D_i = 1 \leftrightarrow \text{score}_i \leq 50$$

$$D_i = 0 \leftrightarrow \text{score}_i > 50$$



## Identification for fuzzy discontinuity

$$y_i = \beta_0 + \beta_1 D_i + \delta(\text{score}_i) + \varepsilon_i$$

$$D_i = \begin{cases} 1 & \text{If household receives transfer} \\ 0 & \text{If household does not receive transfer} \end{cases}$$

- But  
Treatment depends on whether  $\text{score}_i >$  or  $<$  50
- And  
Endogenous factors



## Identification for fuzzy discontinuity

$$y_i = \beta_0 + \beta_1 D_i + \delta(\text{score}_i) + \varepsilon_i$$

### IV estimation

● First stage: 
$$D_i = \gamma_0 + \gamma_1 I(\text{score}_i > 50) + \eta_i$$

Dumy variable

● Second stage: 
$$y_i = \beta_0 + \beta_1 D_i + \delta(\text{score}_i) + \varepsilon_i$$

Continuous function



## Main identification assumption

- Cutoff needs to be exogenous. Agents cannot choose or manipulate the score to become eligible to the program
  - Ex. Colombia CCT: eligibility based on a proxy means test (Camacho, Conover 2009)
  - algorithm for the score was made available to the municipal administrators sometime after 1997
  - → appearance of a discontinuity after 1998
  - Amount of manipulation (+) correlated with the extent of political competition in local elections and (-) with presence of monitoring institutions



## Examples

- Effect of transfers on labor supply  
(*Lemieux and Milligan, 2005*)
- Effect of old age pensions on consumption:  
BONOSOL in Bolivia  
(*Martinez, 2005*)
- The Effects of User Fee Reductions on  
School Enrollment  
(*Barrera, Linden and Urquiola, 2006*)

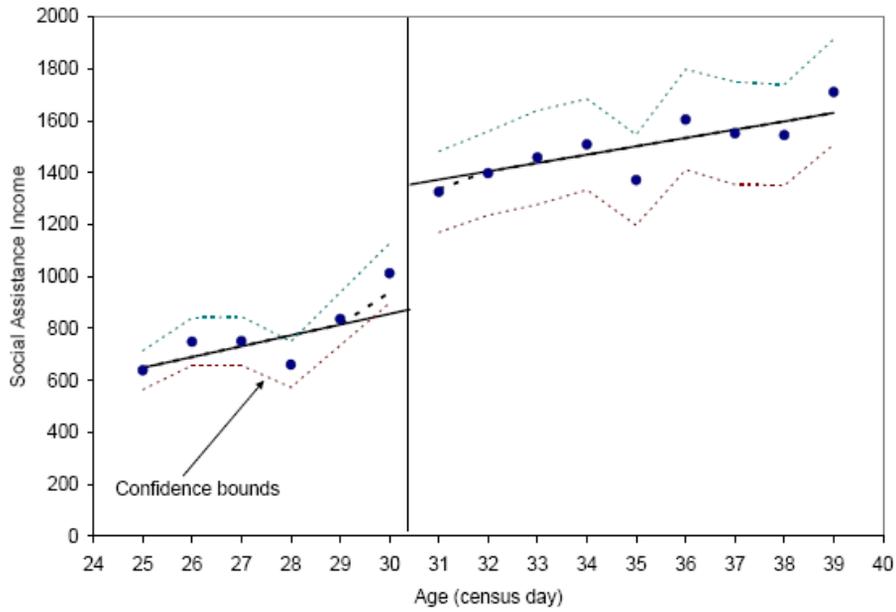


### **Example 1: Lemieux & Milligan: Incentive Effects of Social Assistance**

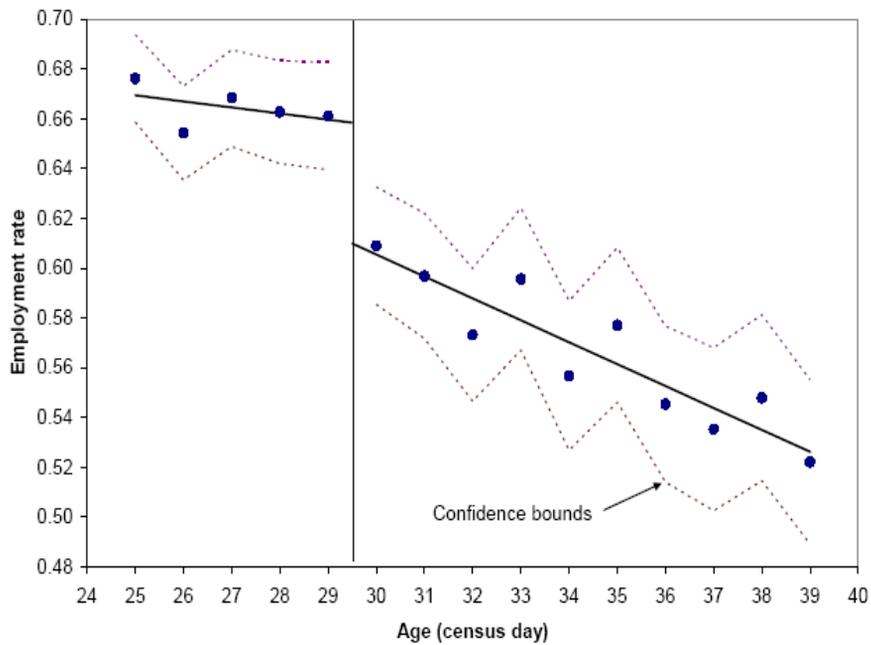
- Social assistance to the unemployed:
  - Low social assistance payments to individuals under 30
  - Higher payments for individuals 30 and over
- What is the effect of increased social assistance on employment?



**Figure 6: Social Assistance Income, Quebec 1986**

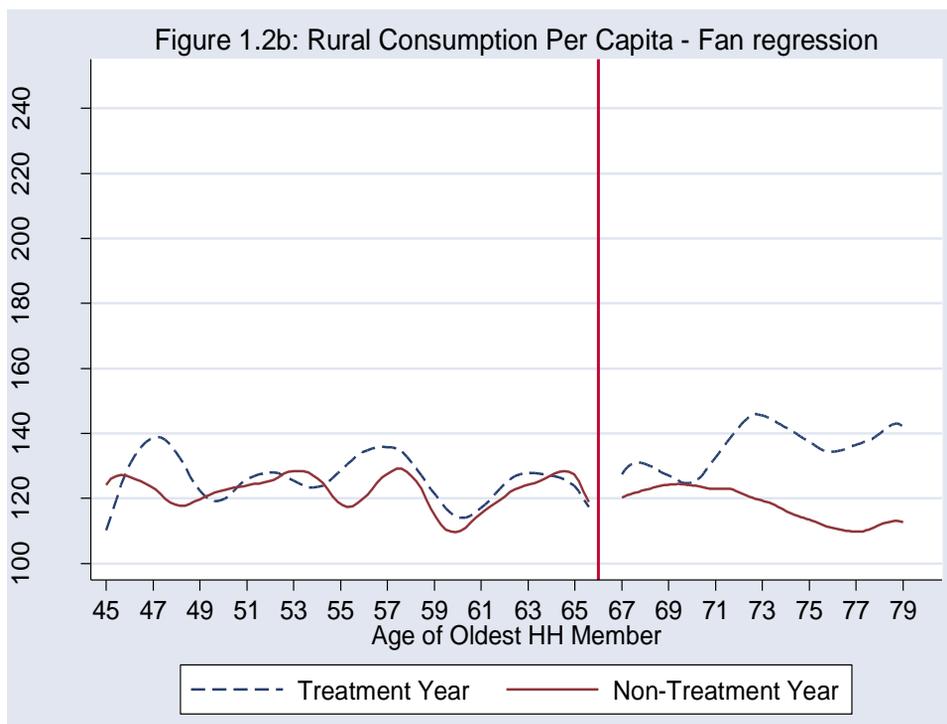


**Figure 3: Employment Rate in Census Week, Quebec 1986**



## Example 2: Martinez: BONOSOL

- Old age pension to all Bolivians:
  - Pension transfer to large group of poor households
  - Pensions paid as of 2001
  - Known eligibility criteria: 65+ years
- Have pre-(1999) and post-(2002) data on consumption
- **Goal:** Estimate effect of BONOSOL on consumption

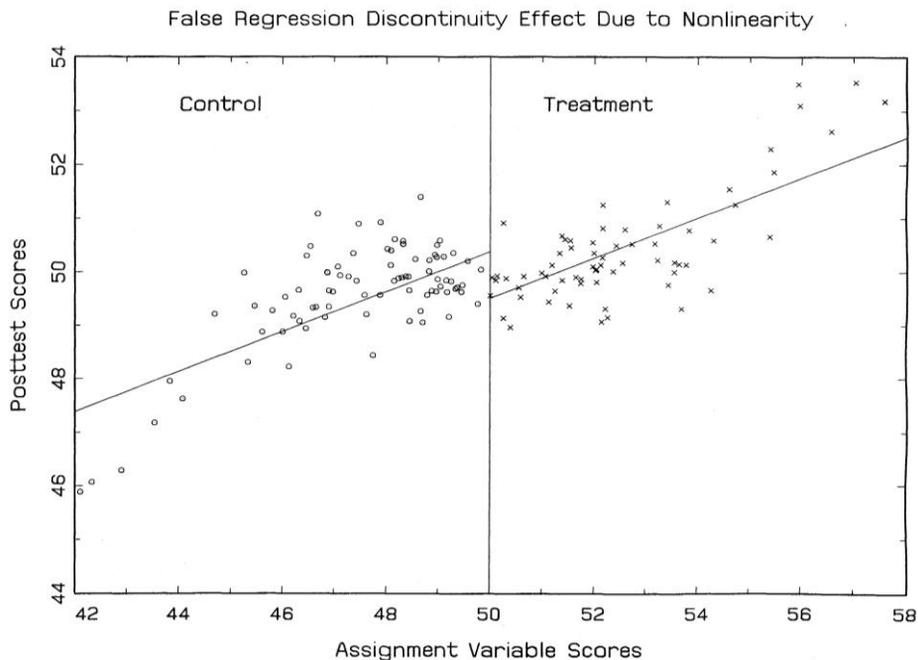


## Potential Disadvantages of RD

- Local average treatment effects
  - We estimate the effect of the program around the cut-off point
  - This is not always generalizable
- Power:
 

The effect is estimated at the discontinuity, so we generally have fewer observations than in a randomized experiment with the same sample size.
- Specification can be sensitive to functional form:
 

Make sure the relationship between the assignment variable and the outcome variable is correctly modeled, including: (1) Nonlinear Relationships and (2) Interactions.

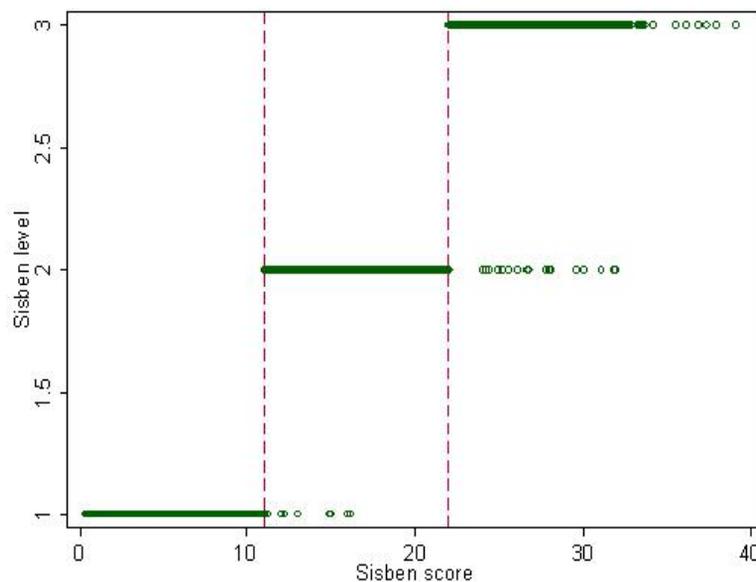


# Advantages of RD for Evaluation

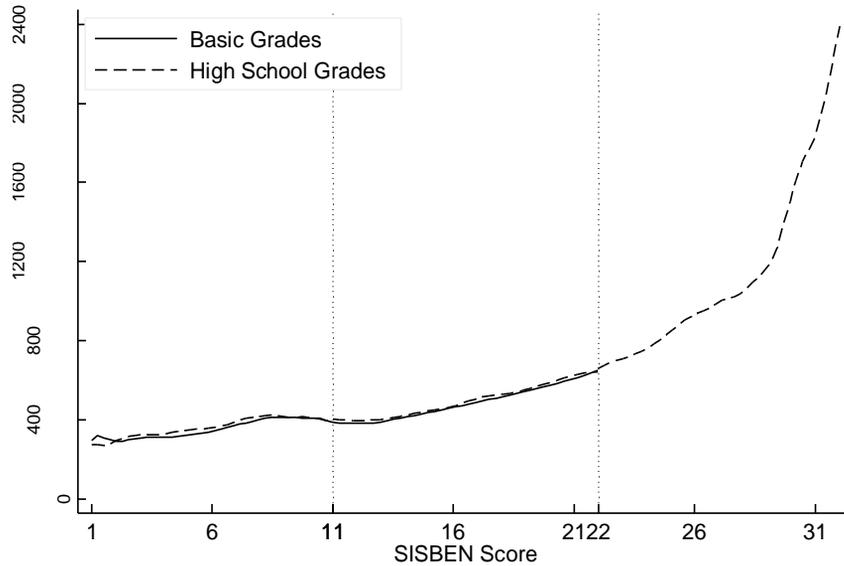
- RD yields an unbiased estimate of treatment effect at the discontinuity
- Can take advantage of a known rule for assigning the benefit
  - This is common in the design of social interventions
  - No need to “exclude” a group of eligible households/individuals from treatment



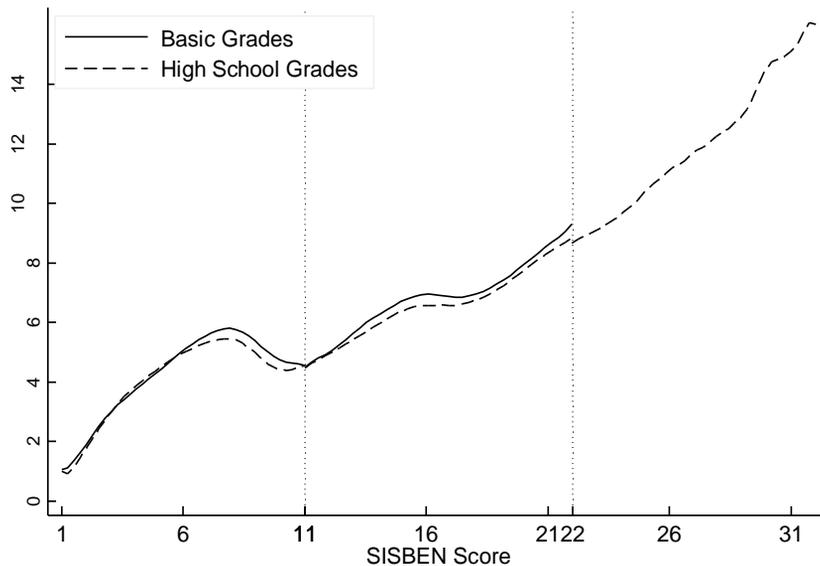
**First step validation: *Sisben* score versus benefit level: is the discontinuity sharp around the cutoff points?**



**Second step validation: Income: *Is it smooth around the cutoff points?***

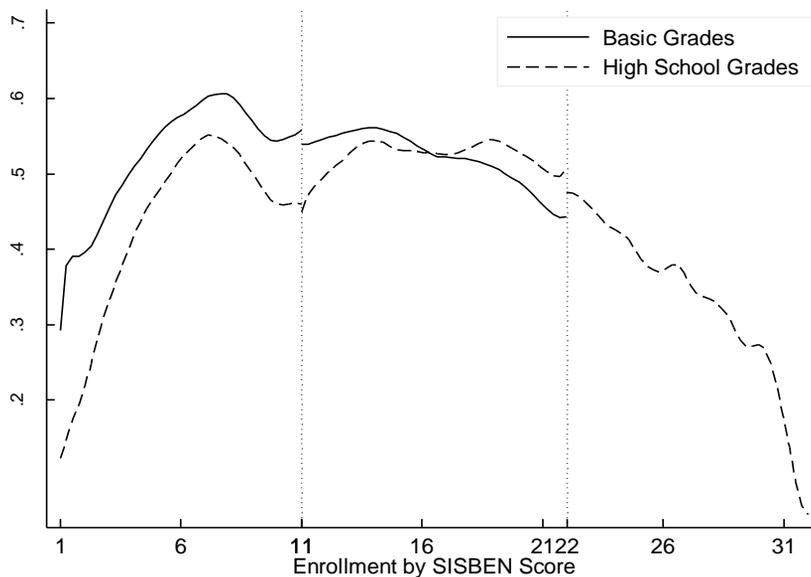


**Second step validation: Years of education of household head : *Is it smooth around the cutoff points?***



## RD Results: *Sisben* vs. school enrollment

### Graphic results



## References

- Angrist, J. and V. Lavy "Using Maimonodes Rule to Estimate the Effect of Class Size on Scholastic Achievement" Quarterly Journal of Economics, 114, 533-575
- Lemieux, T. and K. Milligan "Inentive Effects of Social Assistance: A Regression Discontinuity Approach". NBER working paper 10541.
- Hahn, J., P. Todd, W. Van der Klaauw. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design". Econometrica, Vol 69, 201-209.
- Barrera-Osorio, F., L.L. Linden and M. Urquiola, "The Effects of User Fee Reductions on Enrollment: Evidence from a quasi-experiment" (2007), mimeo, The World Bank.
- Conover, A. and Camacho "Political Manipulation of Social Program Eligibility", mimeo, Universidad del los Andes, Colombia





# Thank You



# Q & A