



THE WORLD BANK



Technical Track Session V Regression Discontinuity (RD)

Christel Vermeersch
Beijing, China, 2009

Reminder: main objective of an evaluation.....

- 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

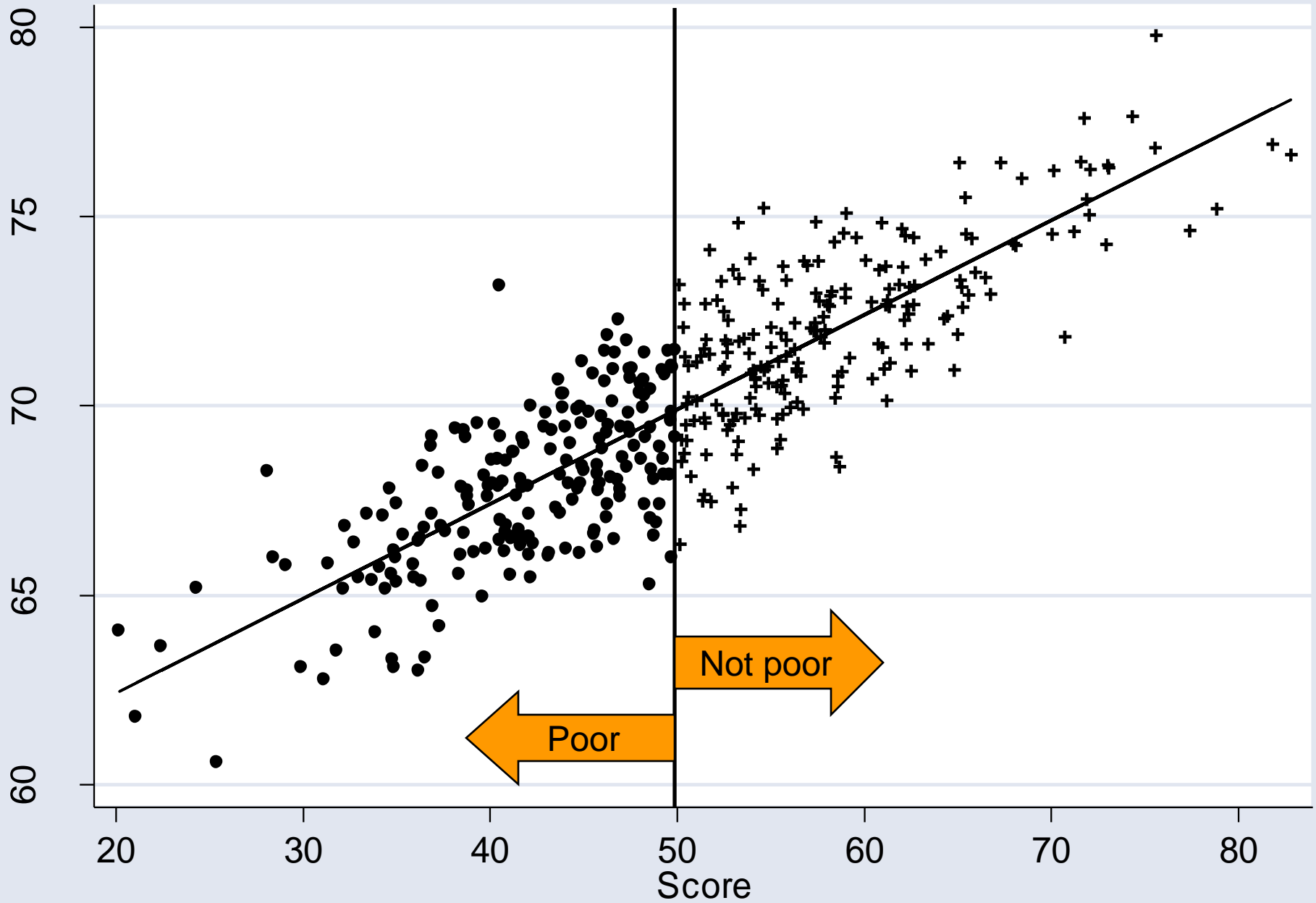
- When to use this method?
 - The beneficiaries/non-beneficiaries can be ordered along a quantifiable dimension.
 - This dimension can be used to compute a well-defined index or parameter.
 - The index/parameter has a cut-off point for eligibility.
 - The index value is what drives the assignment of 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.

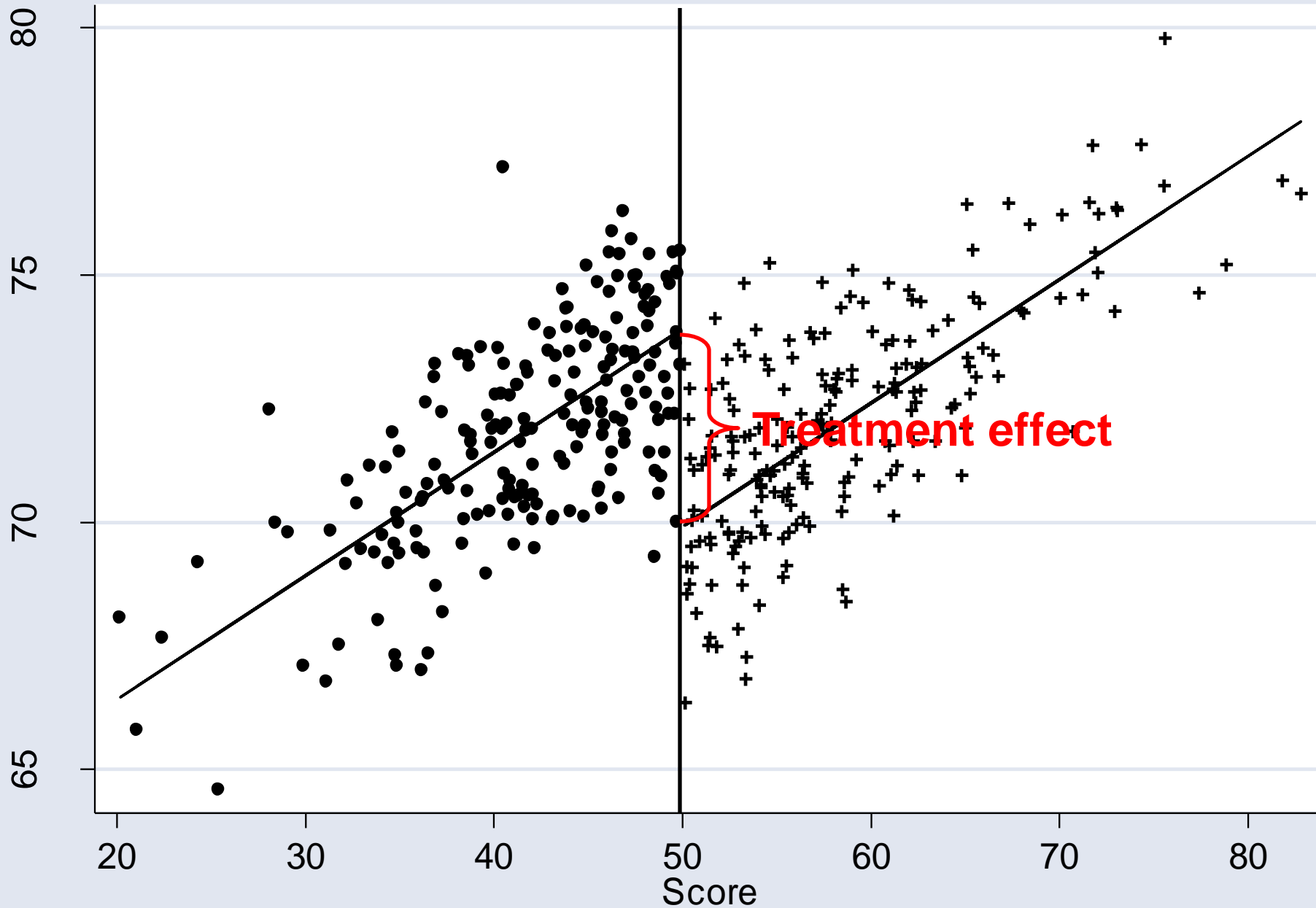
Example: effect of cash transfer on consumption

- **Goal:** Target transfer to poorest households
- **Method:**
 - Construct poverty index from 1 to 100 with pre-intervention characteristics
 - Households with a score ≤ 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

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

$D_i = 1$ if household i receives transfer

$D_i = 0$ if household i does not receive the transfer

$\delta(\text{score}_i)$ is a 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

$$y_i = \beta_0 + \beta_1 D_i + \delta(score_i) + \varepsilon_i$$

Where:

$D_i = 1$ if household receives transfer

$D_i = 0$ if household doesn't receive the transfer

BUT:

Treatment depends on - whether $score_i > < 50$

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: $D_i = \gamma_0 + \gamma_1 \underbrace{I(\text{score}_i > 50)}_{\text{dummy variable}} + \eta_i$

Second stage: $y_i = \beta_0 + \beta_1 \hat{D}_i + \underbrace{\delta(\text{score}_i)}_{\text{continuous function}} + \varepsilon_i$

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

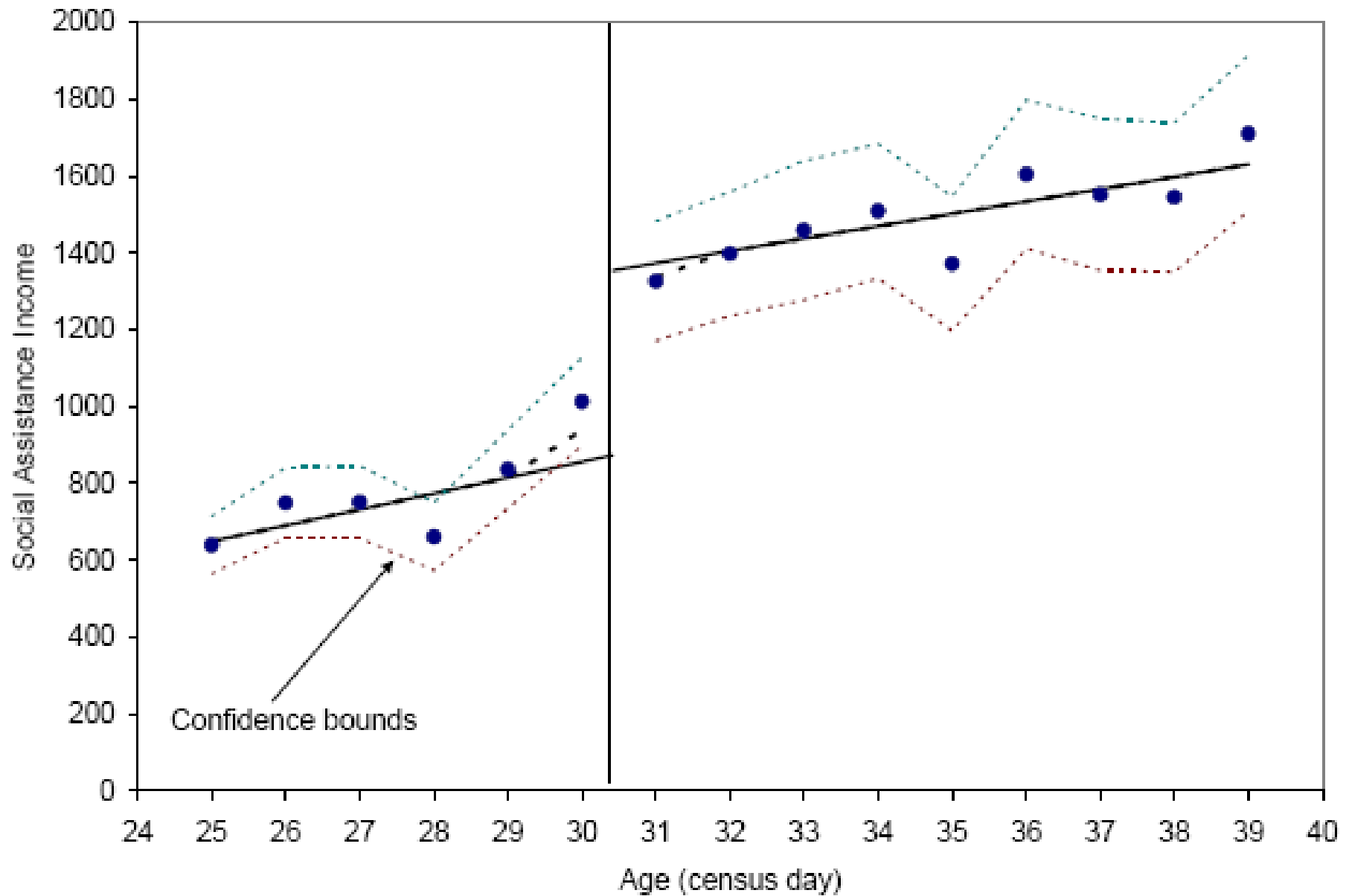
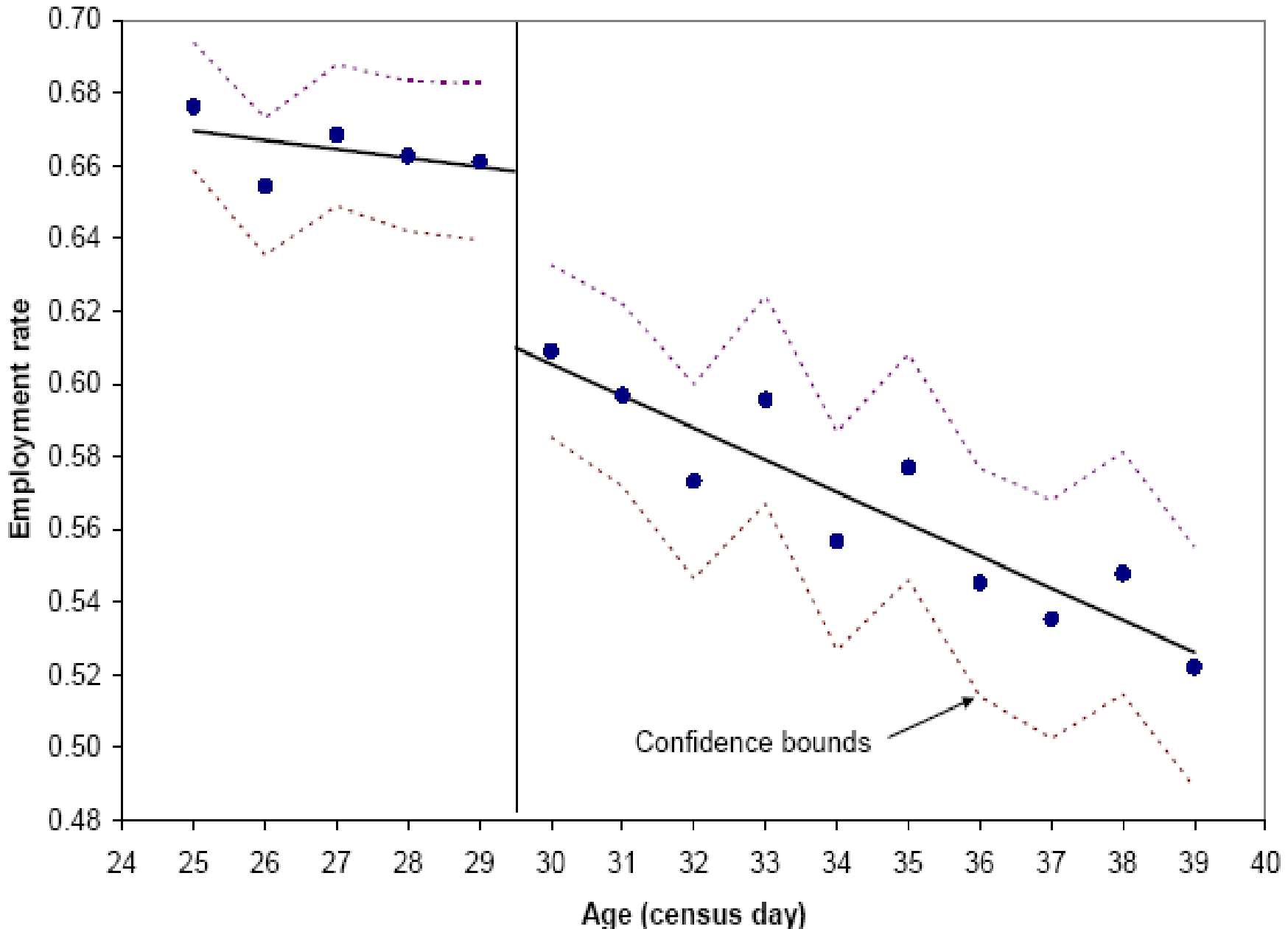


Figure 3: Employment Rate in Census Week, Quebec 1986



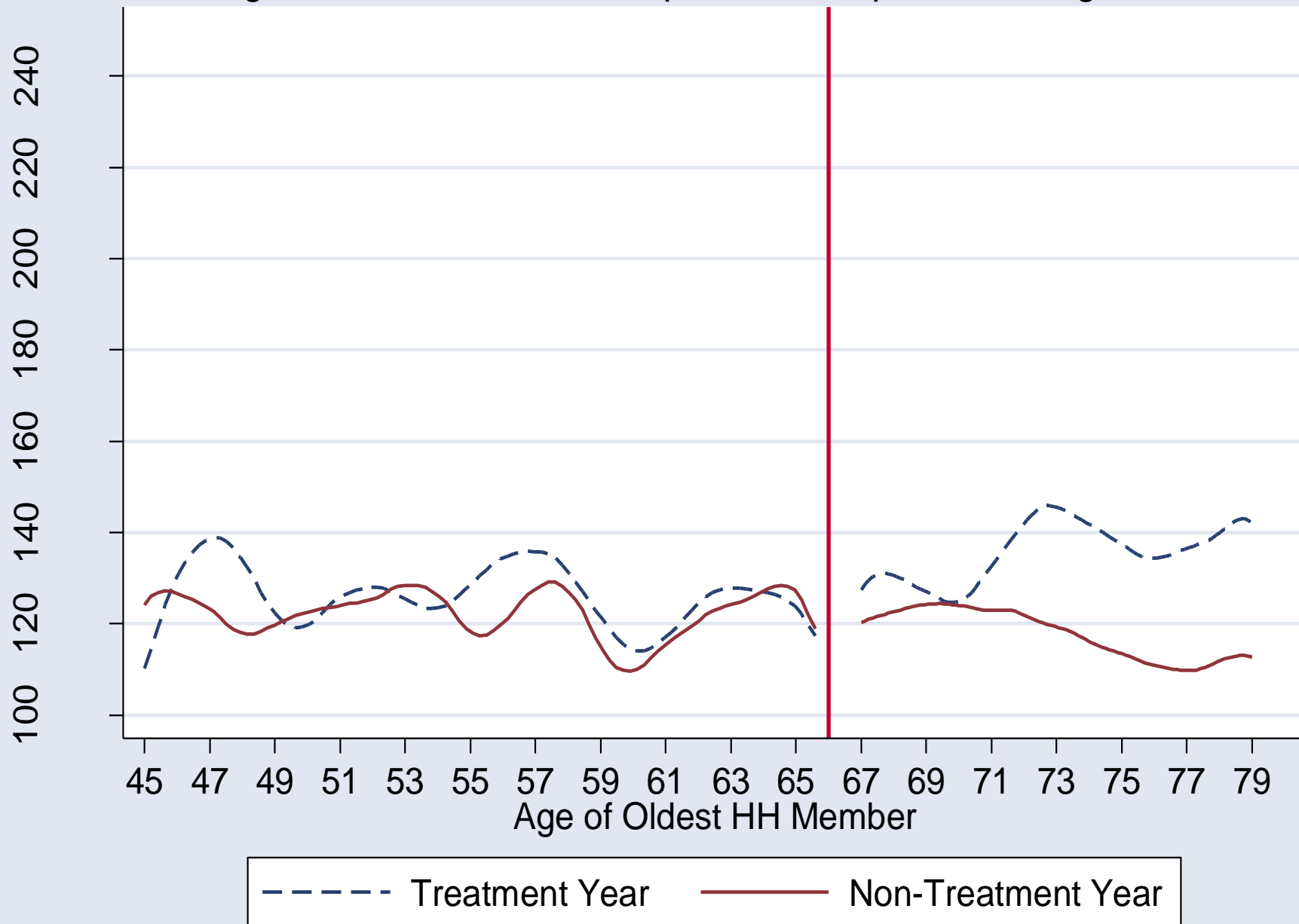
Example 2: Martinez: BONOSOL

- 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

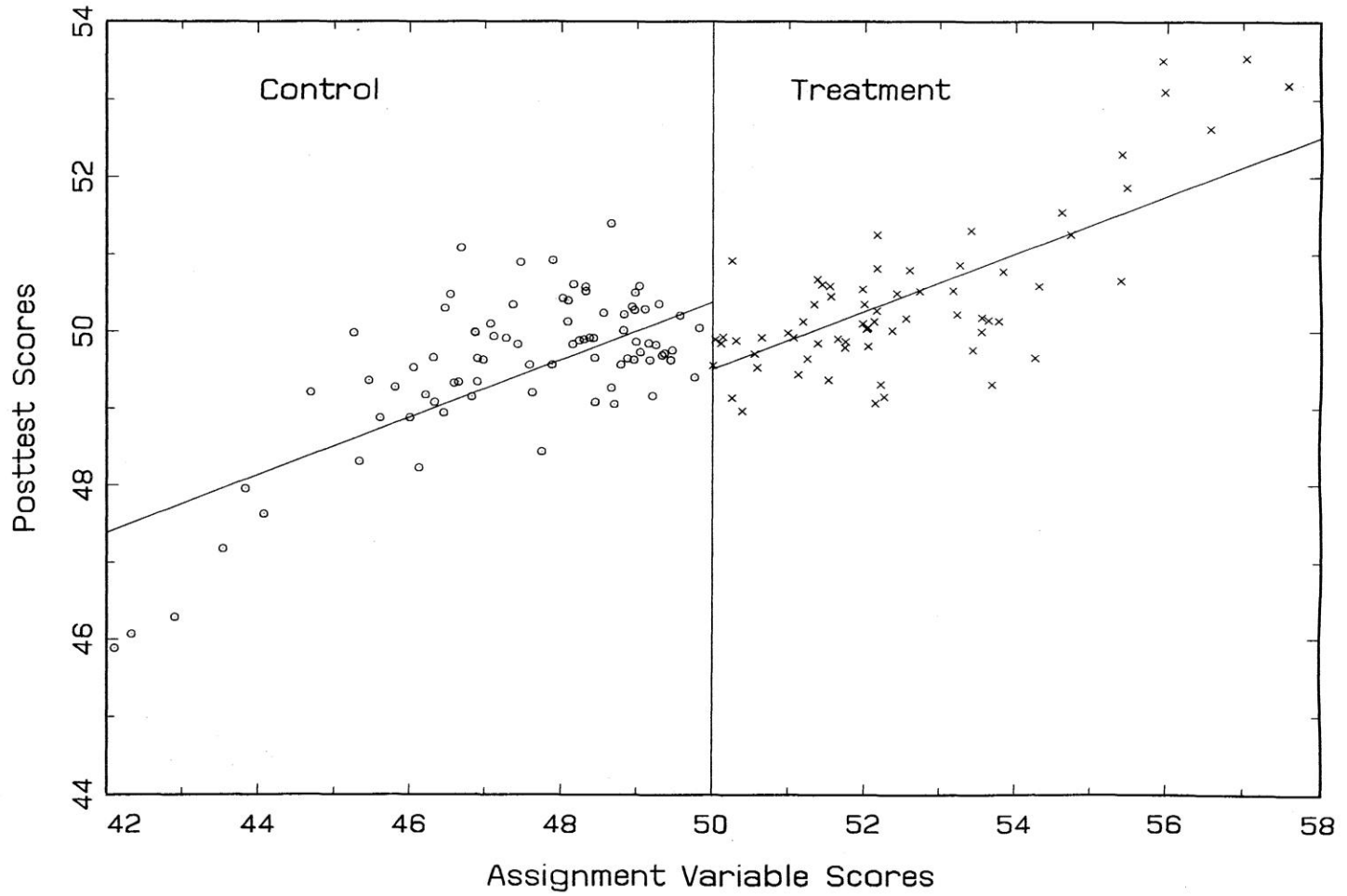
Figure 1.2b: Rural Consumption Per Capita - Fan regression



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:
 - Nonlinear Relationships
 - Interactions

False Regression Discontinuity Effect Due to Nonlinearity



7.5 7-9

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

Example 3: Free schooling program, Colombia

- **Goal:** estimate impact (causal!) of school fee reduction on school enrollment
- **Method:** Regression Discontinuity
- **Paper:** “The Effects of User Fee Reductions on Enrollment: Evidence from a quasi-experiment” (Barrera, Linden y Urquiola)

Context and *Free schooling* Program

- Each year the government issues a resolution that stipulates
 - which items schools may charge for
 - the maximum fee they can set for each of those items
- These expenses are between 7 and 29 monthly dollars, (between 6 and 25 percent of the minimum wage)

- The *Gratuidad* program reduces some of these fees.
- The program is targeted using the *Sisben* index.
- *Sisben* identifies the most vulnerable households in Colombia.
- The extent to which students benefit from these reductions is a function of their *Sisben* level.

What is *Sisben*?

- ❑ *Sisben* is an instrument used to focalize social assistance.
- ❑ First implemented in 1994
- ❑ Based on a survey about households'
 - infrastructure,
 - demographics and
 - human capital
- ❑ Each household receives an score between 0 and 100
- ❑ Using the score, each households is assigned to one of six “levels”, with 1= the poorest , and 6= richest.
 - Scores below a cutoff score of 11 → Level 1
 - Scores between 11 and 22 → Level 2.
 - Scores between 22 and 43 → Level 3

Free schooling Program Benefits

□ **Basic education (grades 1-9)**

- *Sisben 1* children: 100 percent reduction of complementary service fees
- *Sisben 2* and above: no reduction.

□ **High school (grades 10-11),**

- *Sisben 1* children: elimination of both academic and complementary services fees
- *Sisben 2*: approximately a 50 percent reduction
- *Sisben 3* and above: no reduction

Regression discontinuity analysis

- Where is the discontinuity in the regression?
 - Whether or not students benefit from the program is a discrete function of their score.
- Characteristics of the household (observable and unobservable) are continuously related to the score at the cutoff points
- They are similar for students just above and below the cutoff scores.
- Discrete differences in attendance rates between treated and untreated students close to the cutoff can be attributed to the fee reductions.
 - Students with scores of 21.5 might provide an adequate control group for students with scores of 22.5

Estimation

- The basic equation for the estimation, *close to the discontinuity*, is the following:

$$y_i = \alpha + \beta G_i + f(S_i) + \varepsilon_i$$

where y is the enrollment variable, G is a dummy that capture the level of *Sisben*, and S is the score of *Sisben*.

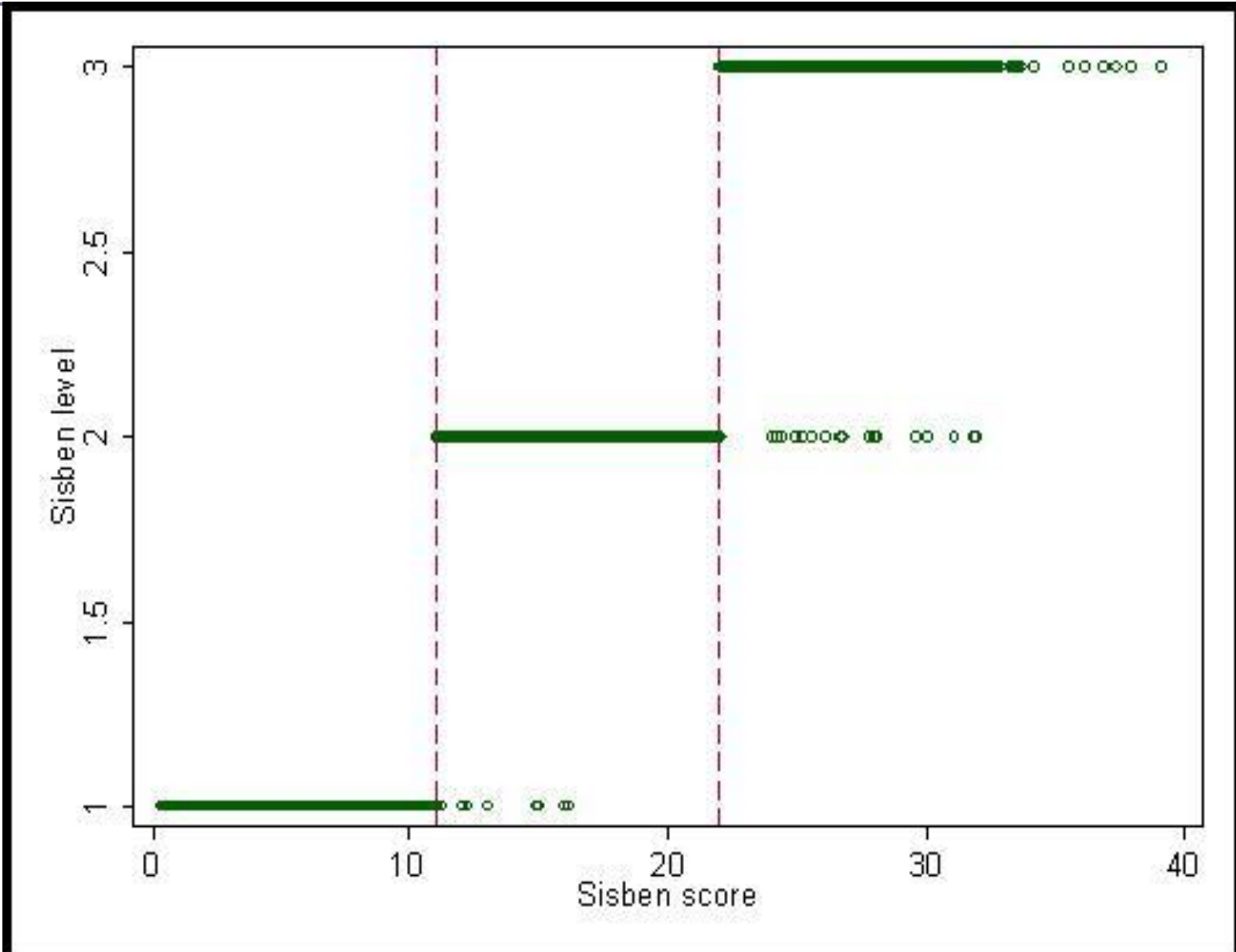
- β will consistently estimate the effect of the program.
- It can be estimated within arbitrarily narrow bands close to the cutoff point,

Validation of the RD strategy

- **First:** what are the properties of the assignment variable?
Is there a real discontinuity in assignment around the cutoff points of the score?
 - Is students' raw *Sisben* score (0-100) a good predictor of their level of benefits?
 - What is the magnitude of exclusion and inclusion errors?

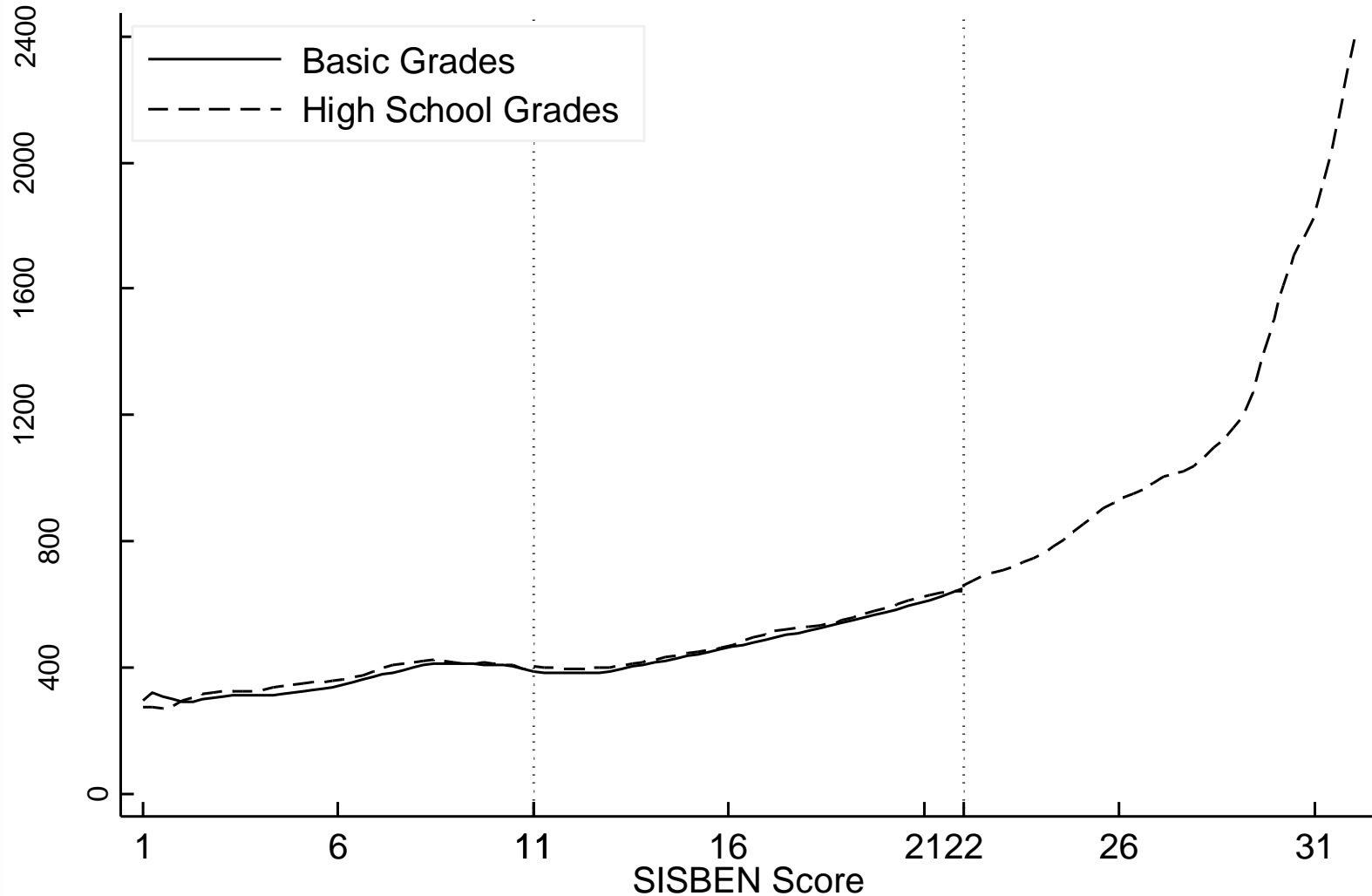
- **Second:** Are the characteristics of individuals smoothly around the cutoff points of the *Sisben* score?
 - E.g., are the beneficiaries and non-beneficiaries similar around the cutoff points?

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



Second Step validation example: Income:

Is it smooth around the cutoff points?

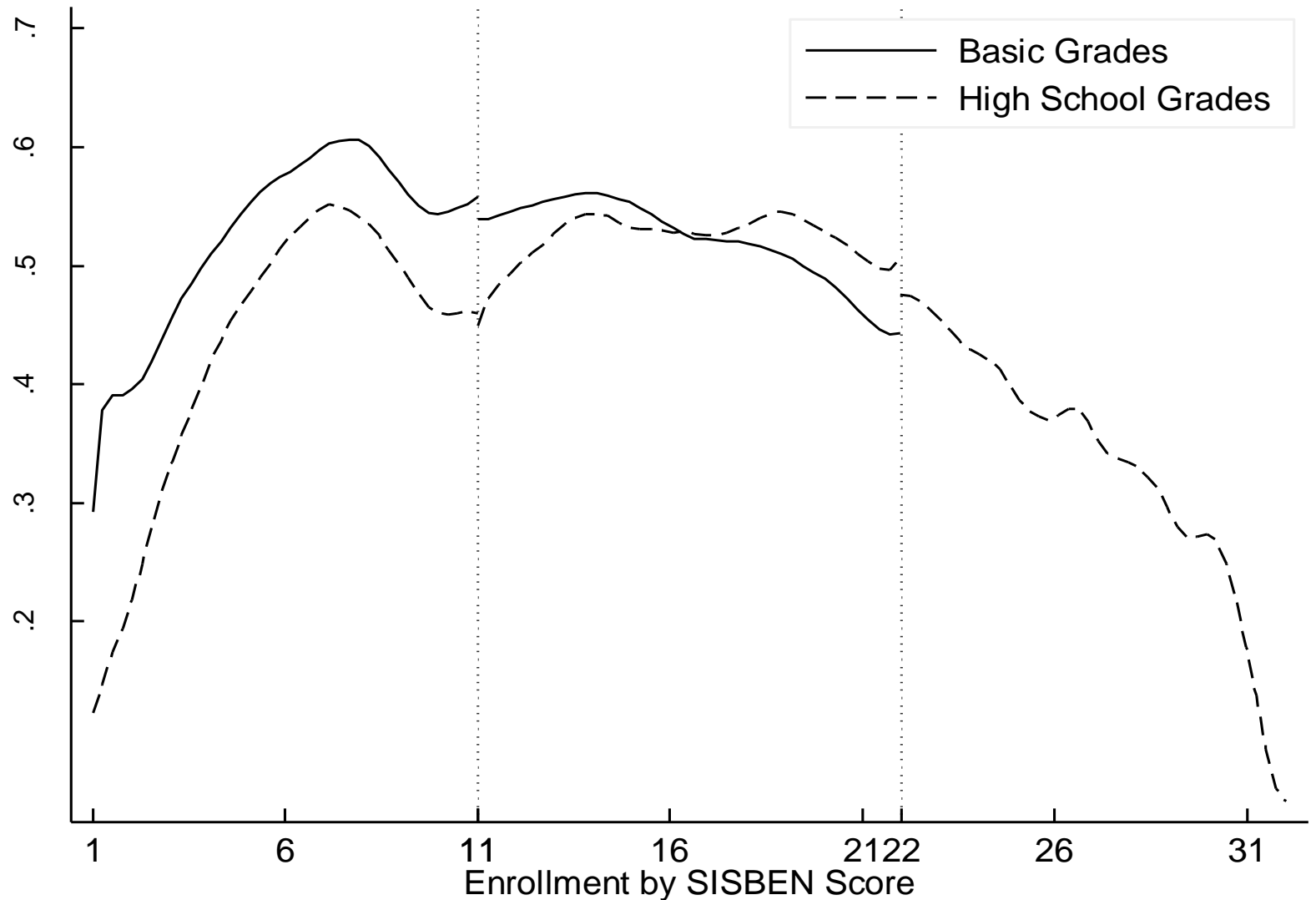


Second Step validation example: Years of education of household head : Is it smooth around the cutoff points?



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, Linden y Urquiola (2006), “The Effects of User Fee Reductions on Enrollment: Evidence from a quasi-experiment”