

Regression Discontinuity Design

Dr. Katrina Kosec
Senior Research Fellow
IFPRI, Washington, DC, USA

Regional Training Course on Applied Econometric Analysis

June 12, 2018

WIUT, Tashkent, Uzbekistan

Outline

- Introduction and Intuition
- Origins of RDD
- Sharp vs. Fuzzy RDD
- Notation and Estimation
- Local Rather than Global Inference
- Illustrations from recent literature:
 - Of the discontinuities
 - Of the impacts of treatment
- Spatial RDD
- Conclusions

Introduction

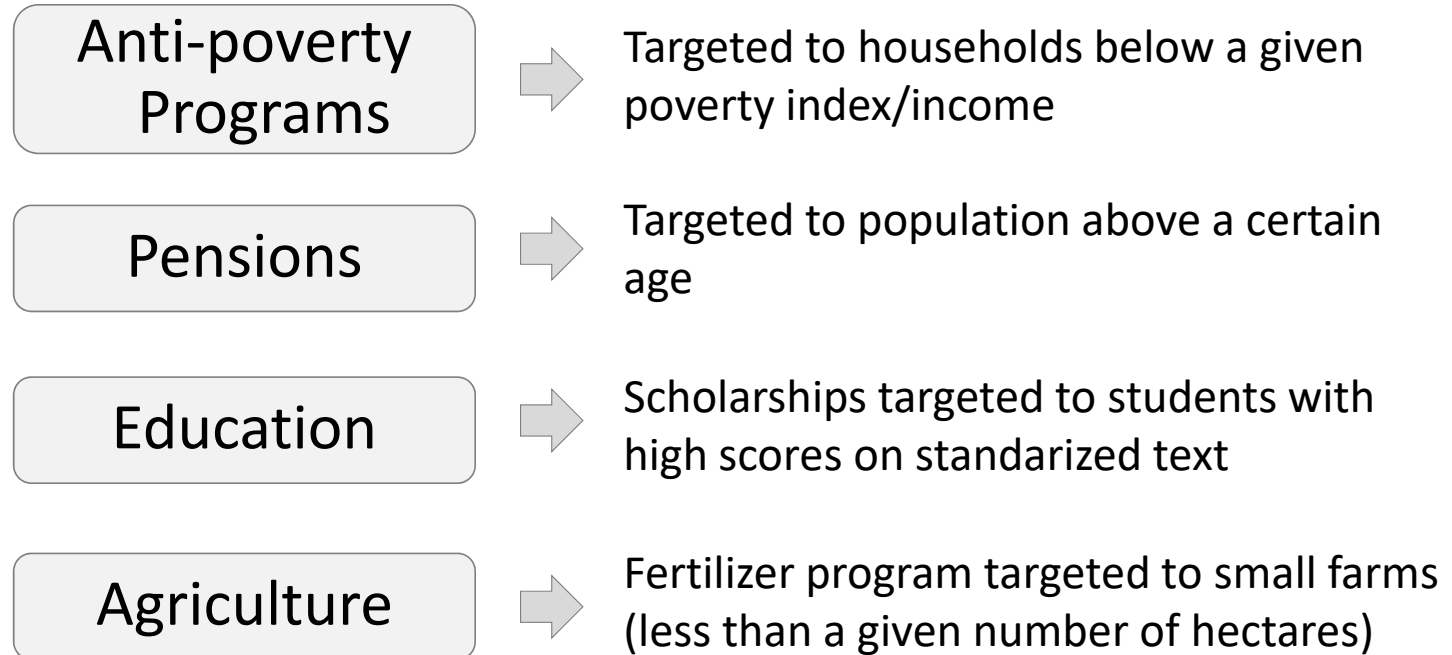
- A method of addressing endogeneity that has been around since 1960, but became especially popular relatively recently
- RDD is a quasi-experimental technique
 - We say ‘quasi’ because it is not actually an experiment; that is, assignment to treatment and control is *not* random...
 - ...But the abrupt change in treatment status at some “cutoff” point of a given variable (e.g., you get a scholarship if your test score exceeds 750) does mimic an experiment, at least around the cutoff (e.g., students getting a 749 and a 751 should be otherwise the same, except only one of them got a scholarship)
 - Has been described as being a “local” randomized control trial

Introduction

- Exploits knowledge of the rules determining treatment (specifically, a known “cutoff” point above or below which treatment is received, and otherwise not)
- Internal validity is high, provided that there are not other factors changing discontinuously at the cutoff
- There are several nice features of RDDs:
 - RDDs are quite common when you look for them
 - You can create very illustrative graphs – both of how treatment status ‘jumps’ at the cutoff and of how outcome variables ‘jump’ at the cutoff
- RDD tells us about the effects of treatment around the “cutoff”—which limits its external validity (a specific type of local average treatment effect, or LATE)

The Basics: Where Might We Use a Regression Discontinuity Design (RDD)?

Many social programs select beneficiaries using an **index** or **score**:



A Group Activity to Understand the Intuition

- Goal: Help everyone be 2 meters tall
- Intervention: A nutrition program provided in primary schools
- Easiest way to evaluate it: Randomly assign among the people most in need
- But random assignment may not be feasible!
 - Involves excluding individuals in great need of the program (in order to get a 'control' group that is similar to the treatment group)
 - Some recipients of the program will be richer and less in need of the program than non-recipients, due to random chance
 - What if the program has already initiated? (Too late to randomize!)

But What Can We Do if Random Assignment is Not Feasible?



One possibility: RDD

How does it work? Let's do a group activity

Summary of Group Activity

- Lined up in height order (such a variable is known as an “assignment” or “running” variable)
- Examined the amount of resources (in our case packages of crackers) I had; gave crackers out to individuals from shortest to tallest until I ran out
- This creates a ‘cutoff’ height:
 - If you’re shorter than it, you get the crackers every day for a year
 - If you’re taller than it, you do not get the crackers
- Individuals far from the cutoff are not comparable, but those close to the cutoff are comparable
- After a year, compare the heights of individuals near the cutoff height; this gives us the effect of the program (...for people near the cutoff)

RDD Example: Scholarships to Improve School Enrolment

Goal

Improve school attendance for poor students

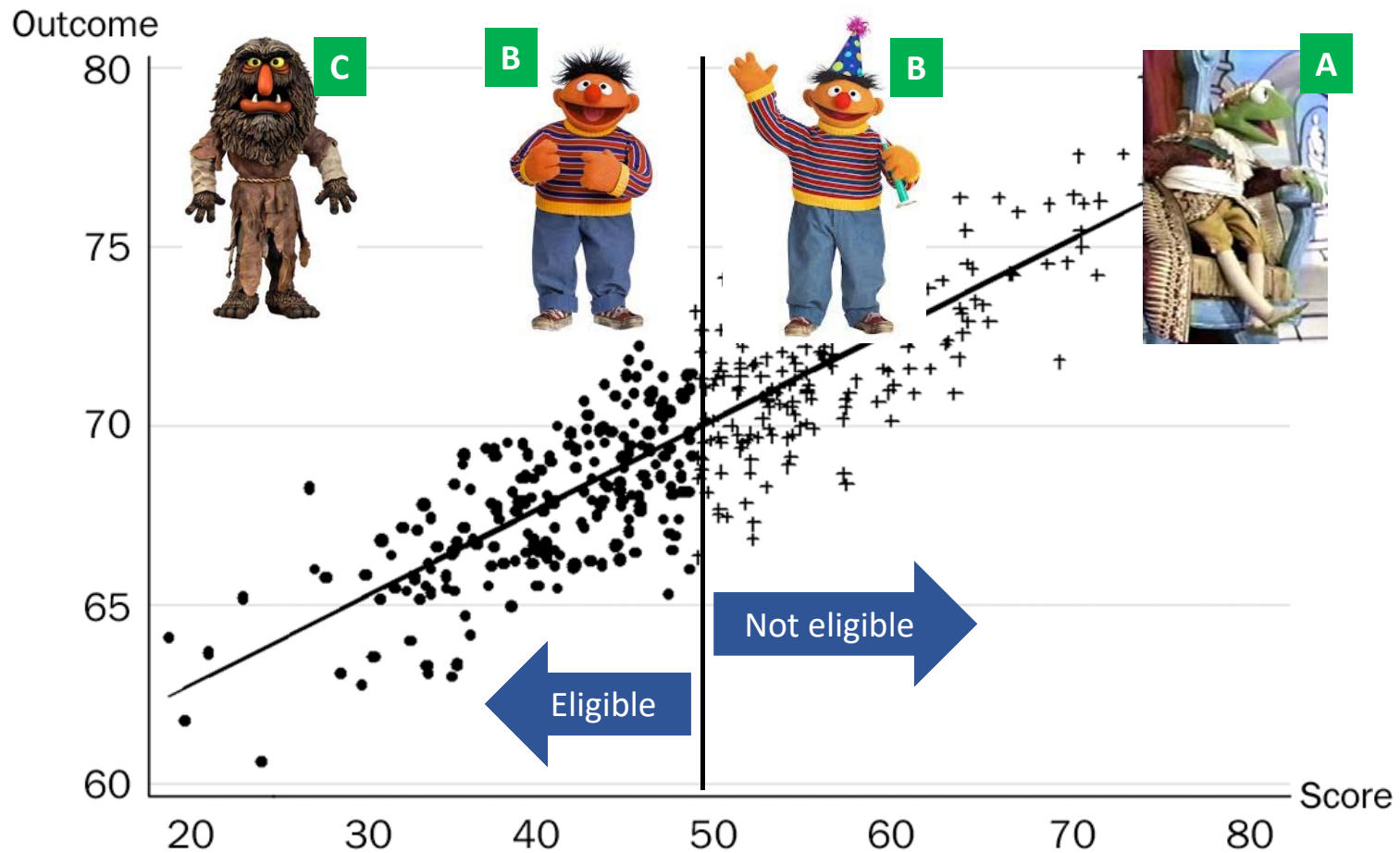
Method

- Households with a score (Pa) of assets ≤ 50 are **poor**
- Households with a score (Pa) of assets > 50 are not poor

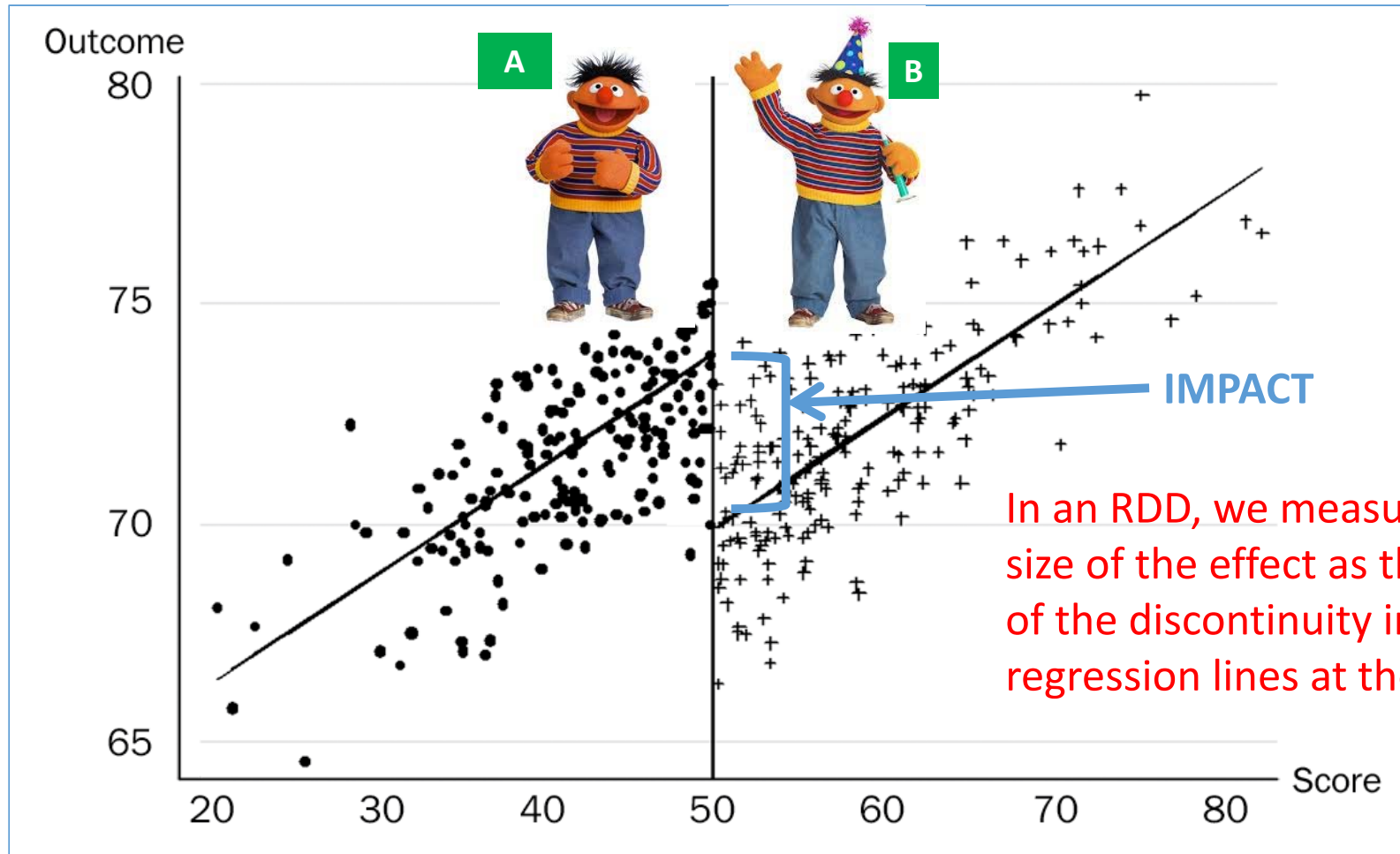
Intervention

Poor households receive scholarships to send children to school

RDD: Pre-Intervention (Baseline)



RDD: Post-Intervention



RDD: The Intuition

- We have a continuous eligibility index (i.e. “assignment” or “running” variable) with a defined cut-off
 - Households with a score \leq cutoff are **eligible**
 - Households with a score $>$ cutoff are **not eligible** (or **vice-versa**)
- Intuitive explanation of the method:
 - Units just above the cut-off are very similar (on *many* dimensions!) to units just below (comparable!)
 - Compare outcomes **Y** for units just above and below the cut-off



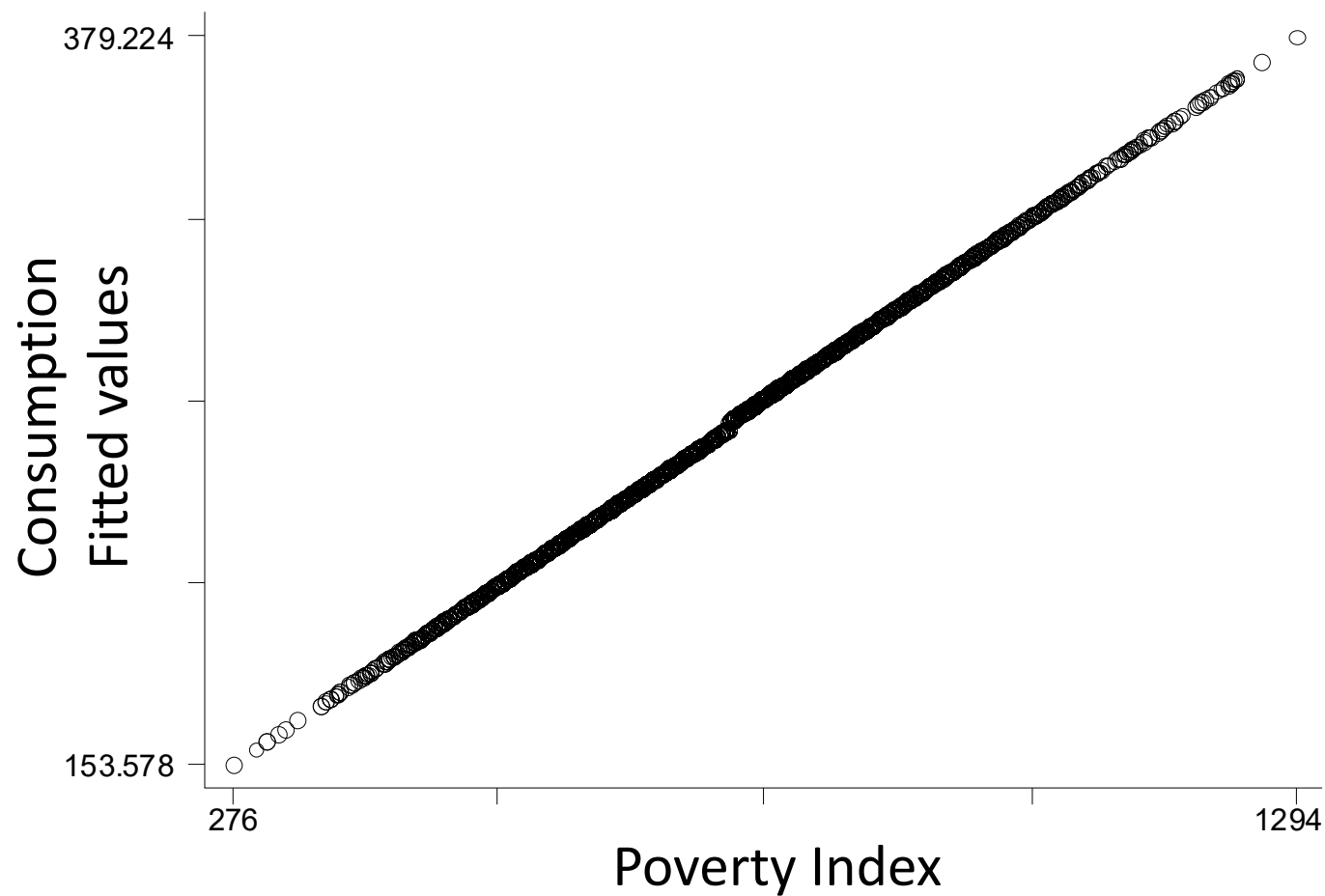
For a discontinuity design, you need:

- 1) A **continuous eligibility** index
- 2) A clearly-defined **eligibility cut-off**

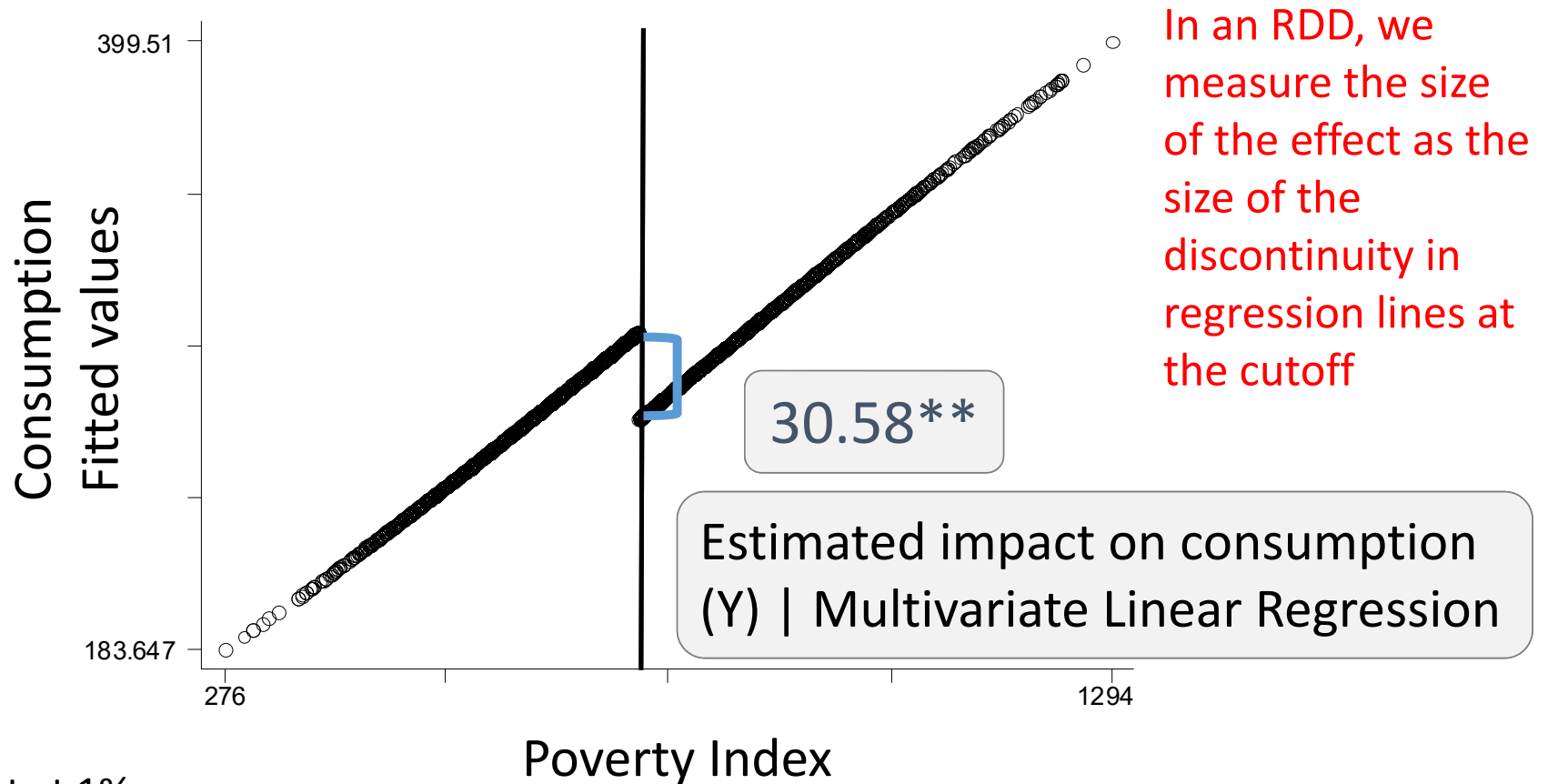
RDD Example: Mexico's Social Protection Program, Progresa

- **Continuous eligibility index:** Progresa is based on a national poverty index (score for each household)
- **Eligibility cutoff:** household is poor if score ≤ 750
 - Eligible=1 if score ≤ 750
 - Eligible=0 if score > 750

Progresa: Poverty Score vs. Consumption at Baseline



Progresa: Poverty Score vs. Consumption Post-Intervention



Thinking About the Assignment Variable

- Notice that consumption is smoothly increasing as the poverty index increases
- RDD does NOT make use of these continuous changes in consumption with the poverty index; what we are interested in is **the discontinuous jump in the outcome at the cutoff point**
- How do we model this empirically?
 - As we will show later, we control for the (continuous) assignment variable, but also include a dummy for being treated (i.e. above or below the cutoff)
 - We are interested in this treatment dummy

Keep in Mind



Discontinuity Design requires continuous eligibility criteria with a clear cut-off

Gives unbiased estimate of the treatment effect: *Observations just across the cutoff are good comparisons*

No need to exclude a group of eligible households/ individuals from treatment

Can sometimes use it for programs that are already ongoing

Keep in Mind



Regression Discontinuity Design

produces a *local* estimate:

- Effect of the program around the cut-off point/discontinuity
- This is not always generalizable

To have enough statistical power to make inferences:

- Need many observations around the cut-off point

Tells us the impact for Ernies but not for Sweetums or Kermits



??



Origins of RDD: Thistlethwaite and Campbell (1960)

- The first application of RDD
- Examine impact of merit awards, allocated on the basis of United States high school students' test scores, on future academic outcomes (career aspirations, enrollment in postgraduate programs, etc.)
 - $\text{Score} > c \rightarrow \text{receive award}$
 - $\text{Score} \leq c \rightarrow \text{do not receive award}$
- Naive approach: compare those who received the award to those who did not
- Why wrong?
 - These two groups are not 'otherwise similar'/ comparable; many observable and unobservable factors influence both test scores and future academic outcomes (HH income, parents' education, motivation, ability, etc)
- Thistlethwaite and Campbell thus compare those just above and below the cutoff

Origins of RDD: Thistlethwaite and Campbell (1960) (continued...)

- Assignment mechanism is known; treatment jumps to 1 if test score $\geq c$
- Individuals near cutoff point are comparable and there appears to be no reason, other than the merit award, for future academic outcomes to change *abruptly* at the cutoff
- This suggests attributing the *discontinuous* jump in outcome variables at c to the causal effect of merit awards
- If treated and untreated individuals are similar near the cutoff point, then data can be analyzed as if they came from a randomized experiment performed among individuals on either side of the cutoff (“high internal validity”)

Internal vs. External Validity in RDD

- Internal validity means our estimate has a low level of bias
- We can show that our study has high internal validity by checking if there are any individual characteristics that change discontinuously, rather than smoothly, at the cutoff
 - Effectively amounts to running your main model with individual- or household-level characteristics as the outcome variable
 - Any abrupt changes suggest that other things rather than just treatment are changing at the cutoff (so we cannot say we isolate the effects of treatment by examining people on either side of the cutoff)
- Despite high internal validity, external validity (i.e. the ability to generalize findings and make conclusions about a broader population or another context) is diminished by the fact that we're estimating an effect for individuals near the cutoff

Notation

- Each household has two *potential* outcomes
 - $Y_i(1)$ denotes the outcome of household i if in the treated group
 - $Y_i(0)$ denotes the outcome of household i if in the non-treated group
- The effect of treatment for household i is
$$Y_i(1) - Y_i(0)$$
- The fundamental problem of **causal inference** is that we cannot observe the pair $Y_i(1)$ and $Y_i(0)$ **simultaneously** (a household is either treated or not—we cannot see what a household looks like when treated and also see what *the same* household looks like when untreated!)
- The *observed* outcome is:
$$Y_i = t_i Y_i(1) + (1 - t_i) Y_i(0)$$
Where t_i is an indicator for household i receiving treatment

Regression Representation

- The *observed* outcome is:

$$Y_i = \alpha + \beta_i t_i + u_i$$

Where t_i is an indicator for household i receiving treatment

- This implies that:

1. $Y_i(1) = \alpha + \beta_i + u_i \rightarrow \beta_i = Y_i(1) - \alpha - u_i$

2. $Y_i(0) = \alpha + u_i \rightarrow \alpha = Y_i(0) - u_i$

- Substituting 2 into 1 gives: $\beta_i = Y_i(1) - Y_i(0)$
- Taking expectations in 2 gives: $E(\alpha) = E(Y_i(0))$

Average Treatment Effect (ATE)

- Since we cannot assign the same household to two conditions and compare them, what we are able to estimate is the average outcome for the treated *minus* the average outcome for the non-treated:

$$E(Y_i(1) | t_i=1) - E(Y_i(0) | t_i=0)$$

- In our framework:

Average outcome for treated: $E(Y_i(1) | t_i=1) = \alpha + E(\beta_i | t_i=1) + E(u_i | t_i=1)$

Average outcome for non-treated: $E(Y_i(0) | t_i=0) = \alpha + E(u_i | t_i=0)$

Difference: $E(\beta_i | t_i=1) + E(u_i | t_i=1) - E(u_i | t_i=0)$

- Note: $E(\beta_i) = \Pr(t_i=1) E(\beta_i | t_i=1) + \Pr(t_i=0) E(\beta_i | t_i=0)$
 $= (1 - \Pr(t_i=0)) E(\beta_i | t_i=1) + \Pr(t_i=0) E(\beta_i | t_i=0)$
 $= E(\beta_i | t_i=1) - \Pr(t_i=0) [E(\beta_i | t_i=1) - E(\beta_i | t_i=0)]$
- This implies that $E(\beta_i | t_i=1) = E(\beta_i) + \Pr(t_i=0) [E(\beta_i | t_i=1) - E(\beta_i | t_i=0)]$

Average Treatment Effect (ATE)

- Substituting back into the Difference gives:

$$\begin{aligned} E(Y_i(1) | t_i=1) - E(Y_i(0) | t_i=0) = & E(\beta_i) + \\ & \text{Pr}(t_i=0) [E(\beta_i | t_i=1) - E(\beta_i | t_i=0)] + \\ & E(u_i | t_i=1) - E(u_i | t_i=0) \end{aligned}$$

- What does this imply?
 - The difference in the average outcomes of treated and non-treated units may not equal the average treatment effect, $E(\beta_i)$ (if either of the terms **in blue** is non-zero, it will not!)
- What would lead the terms **in blue** to be non-zero?
 - Average outcome gains resulting from treatment were different for both groups of individuals
 - Average outcomes for recipients and non-recipients differed even in the absence of treatment

Average Treatment Effect (ATE)

- Randomized assignment would have guaranteed us that these two last terms are 0; our comparison of the outcomes of treated vs. non-treated units would produce the average treatment effect (ATE)
- But, since we do not have randomization, we cannot just compare outcomes for treated with untreated units and take this as the causal effect of treatment
- What to do? RDD estimation! Basically, this means we:
 - Put X (the assignment, or running variable), or some function of it (e.g., a quadratic or a cubic in X) on to the right-hand-side of the regression to “control” for confounding variation in X (i.e. smooth changes in Y as X varies)
 - Include a dummy (indicator) for being above the cutoff

Sharp and Fuzzy RDs

- Sharp RD
 - The discontinuity precisely determines treatment, i.e., probability of treatment jumps from 0 to 1 when the assignment (i.e. running) variable crosses the cutoff
 - Equivalent to random assignment in the neighborhood around the cutoff
 - Example: U.S. social security (retirement) payments depends directly and immediately on a person's age (few if any administrative errors, due to good records systems)
- Fuzzy RD
 - The discontinuity (cutoff) is highly correlated with receipt of treatment
 - Probability of treatment jumps by less than one when the assignment (i.e. running) variable crosses the cutoff
 - Example: Rules determine eligibility but there is a margin of administrative error
 - Use the assignment rule as an instrumental variable for program participation

Sharp RD

- Sharp RD is used when treatment status is a deterministic and discontinuous function of the assignment (i.e. running) variable, x_i
- Specifically:

$$D_i = \begin{cases} 1 & \text{if } x_i \geq x_c \\ 0 & \text{if } x_i < x_c \end{cases}$$

where x_c is a known a cutoff

Basic Specification of Sharp RD

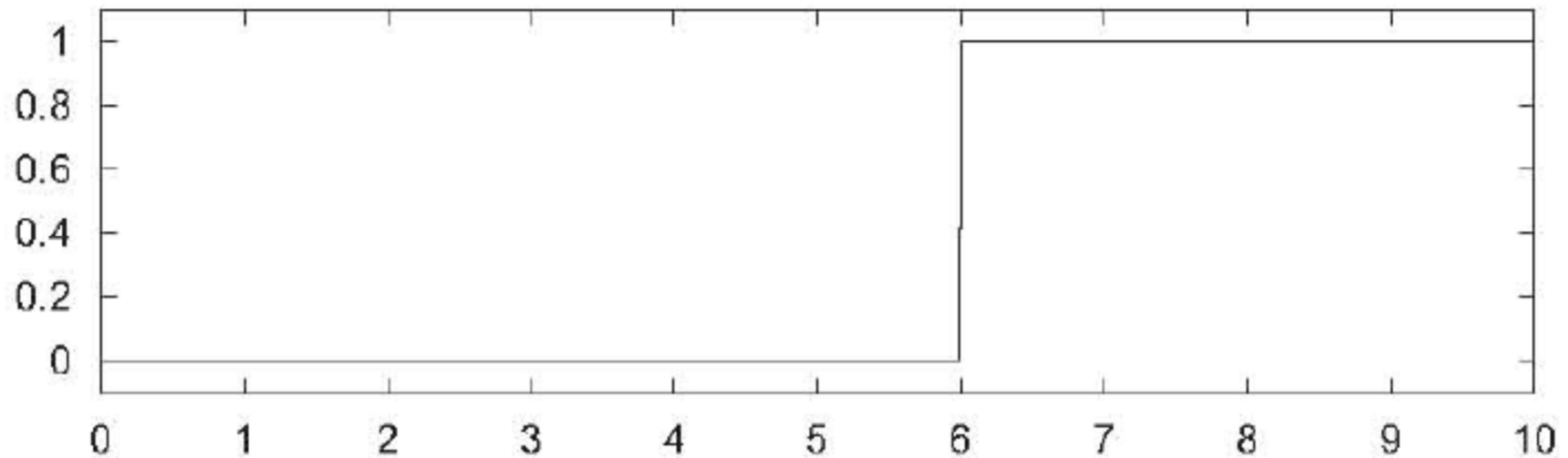
- Estimate the following:

$$Y_i = \beta_0 + \beta_1 D_i + f(X_i - X_c) + u_i$$

Where:

- Y is the outcome
- X is the assignment, or running variable
- D is the treatment dummy variable (takes on 0 or 1)
- X_c is the cutoff (subtract it so that $X_i - X_c = 0$ at the cutoff, and this term thus drops out when we are at the cutoff, positive past the cutoff, negative before)
- $f(\cdot)$ can be any smooth function of X (no functional form restrictions)
- $\widehat{\beta}_1$ is the treatment effect at the cutoff

Imbens and Lemieux (2008) Example of Sharp RD



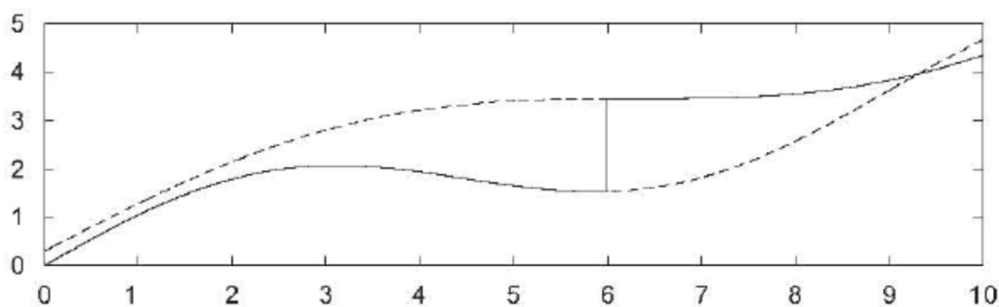
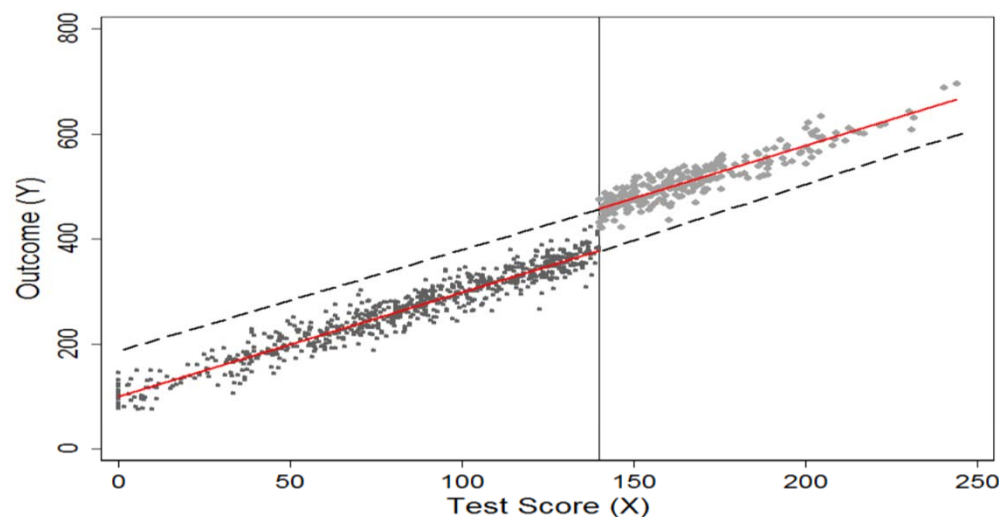
- Vertical axis = conditional probability of treatment
- Horizontal axis = assignment (i.e. running) variable
- Cutoff for treatment assignment is $x_c = 6$
- Probability of assignment jumps from 0 to 1 at cutoff

Sharp RD Example

- National Merit Scholarship (privately-funded college scholarship) awards in United States
 - The National Merit Scholarship Corporation (NMSC) uses scores from a test called the PSAT/NMSQT to screen over 1.5 million applicants, computing a national selection index score for each
 - NMSC decides on a 'qualifying' score (critical reading + math + writing skills scores) – usually set so that students at the 96th percentile or higher are recognized and get a scholarship
 - A small difference in test score means a discontinuous jump in scholarship access

Extrapolation in a Sharp RD Setting

- Estimation of treatment effects in sharp RD depends on extrapolation, since:
 - To the left of cutoff point, have only non-treated observations
 - To the right of cutoff point, have only treated observations (or vice versa)
- The dashed lines in these examples show us the type of extrapolation our estimator does (first is linear, second non-linear)
 - Depends on our choice of $f(\cdot)$ (e.g., control for X to get a linear extrapolation, or do a quadratic/ cubic/ higher order polynomial)
- To avoid excessive extrapolation, focus on observations around the cutoff



Extrapolation in a Sharp RD Setting (continued...)

- Key assumption of Sharp RD: Local Continuity
 - This implies that units close to the cutoff point X_c (i.e. which have similar X values) are **comparable**, meaning units just above and below cut-off have *similar potential outcomes* (they will obviously have very different *observed outcomes*)
- Comparing average outcomes just above and below the cutoff identifies the average effect of treatment *for subjects close to the cut-off*
 - Equivalently: You are computing the difference between two regression functions *at a point*
- This effect may be very different from treatment effects at values of X *far* from the cutoff

Importance of the Local Continuity Assumption

- Even if treatment receipt is determined solely by the cutoff, this is insufficient for identification
- Why?
 - There may be coincidental functional discontinuities in relationship between Y and X ; e.g., other programs that use the same assignment variable and cutoff
- So, we need the continuity assumption as well
 - And thus want to show that other variables that might affect the outcome are changing smoothly at the cutoff

Summary: When to Use Sharp RD Design

- 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
- Probability of assignment to treatment jumps from 0 to 1 at cut-off (i.e. is discontinuous at the cut-off)

Fuzzy RD

- Practically speaking, imagine incentives to participate changing discontinuously at some cutoff value X_c , but that the cutoff is not powerful enough to move all subjects from non-treated to treated status
- We can think of this as:
 - The presence of mis-assignment relative to the cutoff value in a sharp RD
 - Some values of X near the cutoff appear in both treated and non-treated groups
 - Assignment is based on variables observed by the administrator but not the researcher (e.g. not only a test score, but also letters of recommendation, decide who gets a scholarship)
 - A random experiment with:
 - No-shows: treatment group members who do not receive treatment; and/or
 - Spillovers/ contamination: control group members who nonetheless receive treatment

Identification for Fuzzy RDD

$$Y_i = \beta_0 + \beta_1 D_i + f(X_i - X_c) + u_i$$

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

Similar setup, but:

Treatment depends on whether $X_i \geq X_c$ or $X_i < X_c$ and other factors (observed or not)

Because of the other factors, we approach this how we do in an instrumental variables setup, with being above the cutoff (or below, depending on what constitutes treatment in your study) as the instrumental variable for receiving treatment.

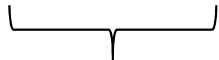
Identification for Fuzzy RDD (continued...)

$$Y_i = \theta_0 + \theta_1 \hat{D}_i + f(X_i - X_c) + \varepsilon_i$$

IV estimation

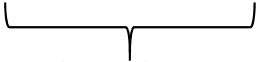
First stage:

$$D_i = \alpha_0 + \alpha_1 I(X_i \geq X_c) + \mu_i$$


Dummy (indicator) variable

Second stage:

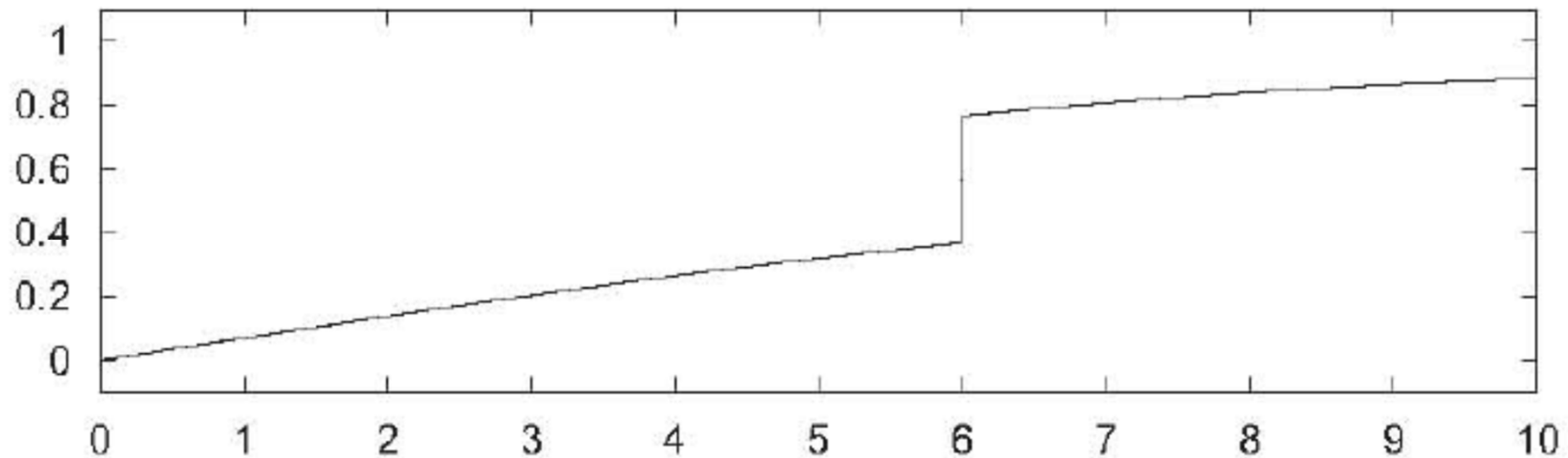
$$y_i = \theta_0 + \theta_1 \hat{D}_i + f(X_i - X_c) + \varepsilon_i$$


Continuous
function in X

Which Local Average Treatment Effect (LATE) Does Fuzzy RDD Estimate?

- Since the probability of treatment jumps by less than one at the threshold, the relationship between Y and D *cannot* be interpreted as the average treatment effect (ATE) at the cutoff
- A fuzzy RD *instead* measures the average treatment effect *at the cutoff* (just as for the case of sharp RD) AND *on the compliers* (i.e. those whose treatment status is influenced by which side they fall on in relation to the cutoff, X_c)

Imbens and Lemieux (2008) Example of Fuzzy RD



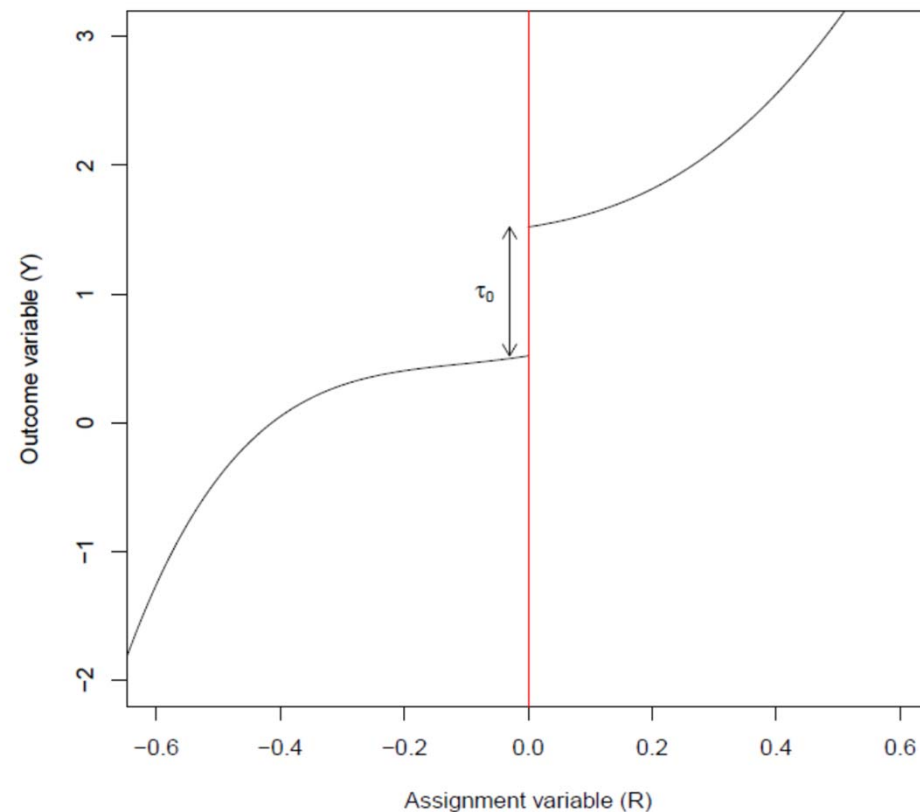
- Vertical axis = conditional probability of treatment
- Horizontal axis = Assignment (i.e. running) variable
- Cutoff for treatment assignment is $x_c = 6$
- Probability of assignment jumps from 0.3 to 0.7 at cut-off

Relationship Between the Assignment Variable and Outcomes

- We need to correctly model this relationship; as mentioned earlier, it may not be linear
- We can plot the relationship to visually inspect whether or not it is linear (e.g., should we control for only the assignment (running) variable, or also its square?)

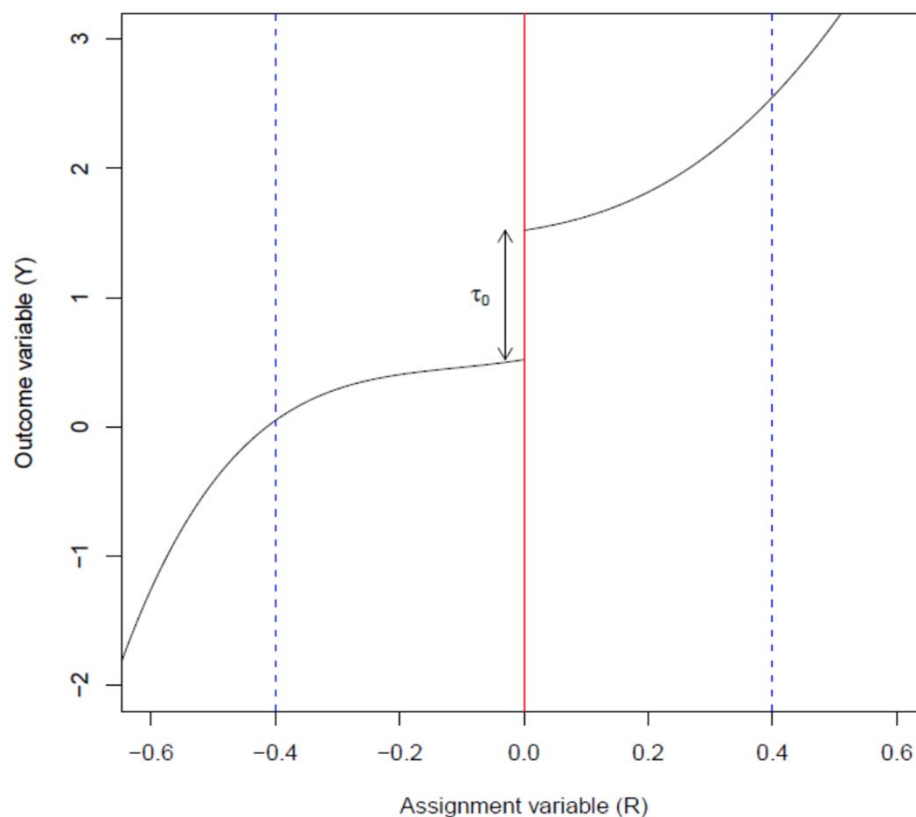
Choosing a Bandwidth: Calonico, Cattaneo, and Titiunik (*Econometrica*, 2014)

- Doing global inference uses all of the data shown



Choosing a Bandwidth: Calonico, Cattaneo, and Titiunik (*Econometrica*, 2014), continued...

- Local inference truncates the data by choosing some “optimal” bandwidth (enough data have the power to make inference, but not so far from the cutoff that units are not comparable) and looks only at units of analysis within that bandwidth
- Additionally, within that bandwidth, you can put more weight on units *near* the cutoff than on those further away

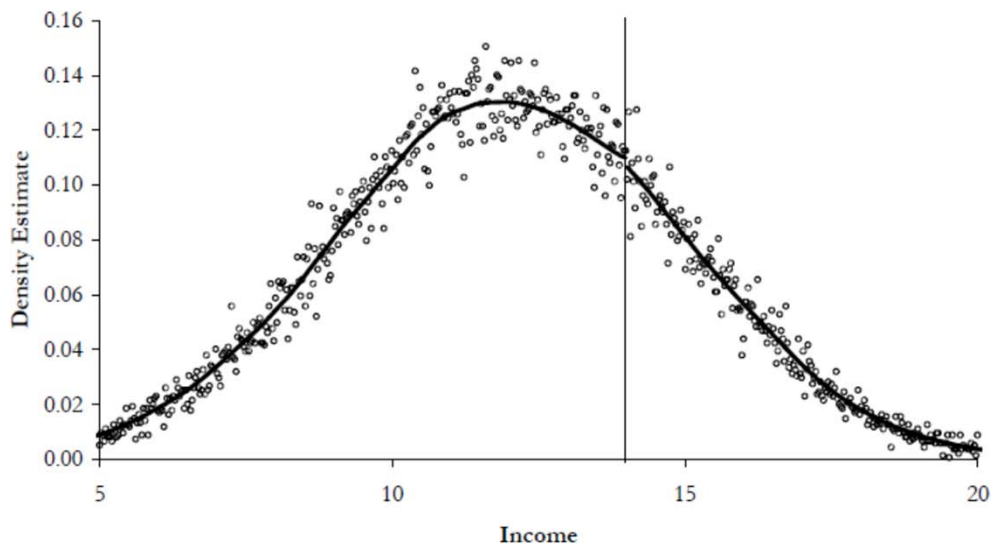


McCrary (2007): Manipulation Checks

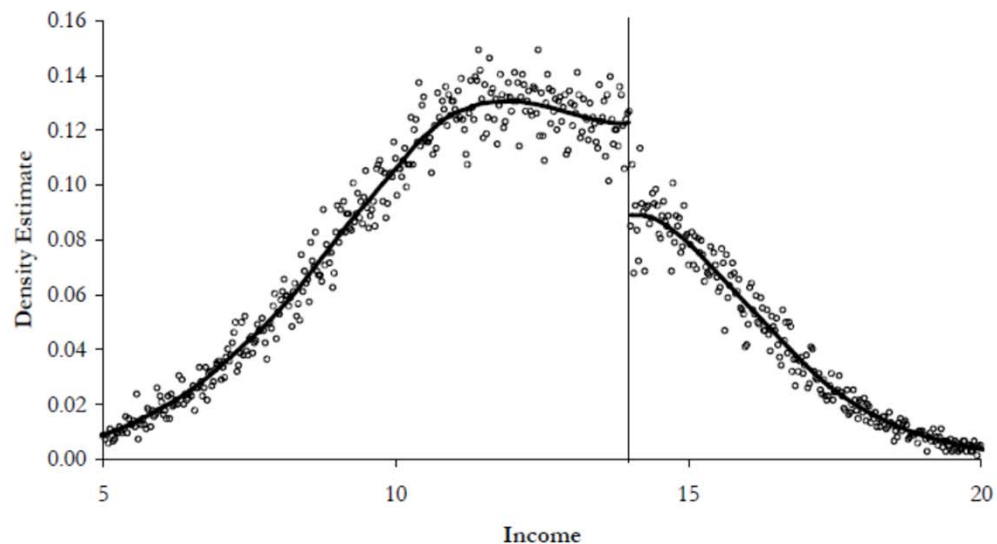
- One worry about RD: people can manipulate their treatment assignment (or someone can manipulate it for them)
- E.g., around the cutoff, perhaps you push your friends' wealth scores down very slightly so that they are *just* under the cutoff and get some cash transfer program
- Problematic as then treatment is not 'as if random' at the cutoff!
- What to do? Check for manipulation!
- What does manipulation look like? More people on the 'favorable' side of the cutoff (barely got the program) than the unfavorable side (just missed the program)

Examples of No Manipulation and Manipulation from McCrary (2007)

C. Density of Income
with No Pre-Announcement and No Manipulation



D. Density of Income
with Pre-Announcement and Manipulation



The distribution of income *should* be approx. normal, so the break at the right shows clear manipulation of incomes to make sure more people get the program.

Overview of RDD Packages in Stata

Before class tomorrow, please look at each of these:

1. <https://sites.google.com/site/rddpackages/>

- As noted at the link above, type the following in the Stata command line to install rdrobust:

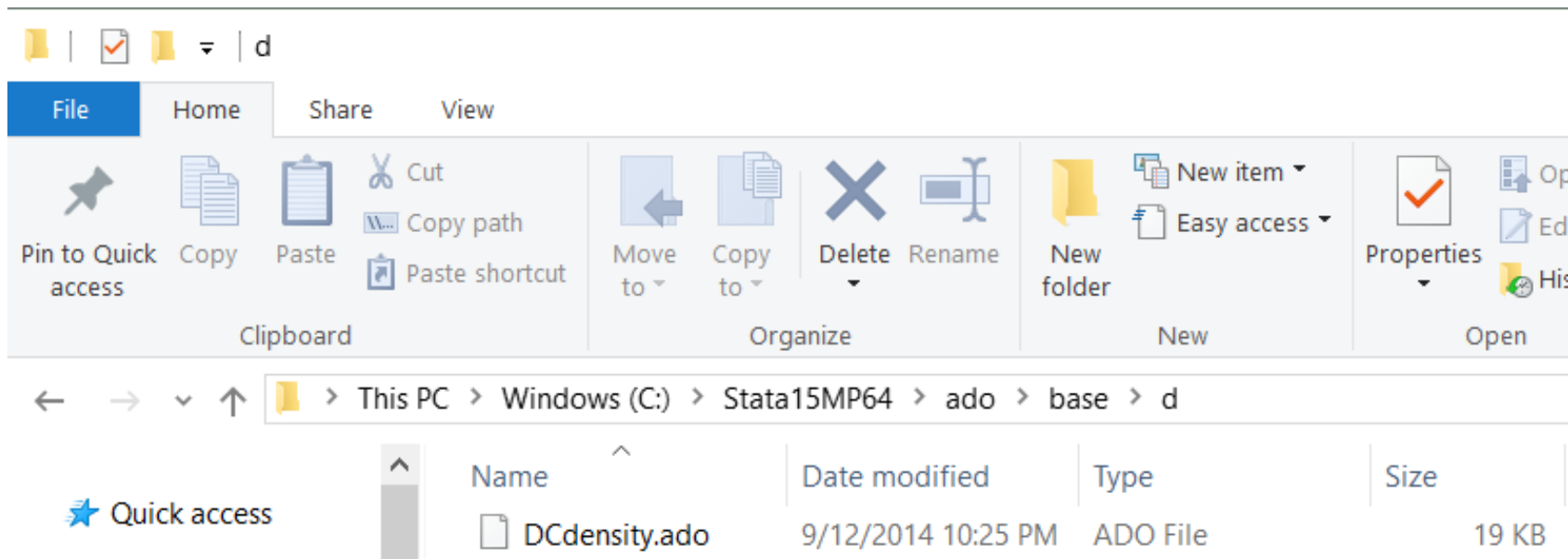
`net install rdrobust, from(https://sites.google.com/site/rddpackages/rdrobust/stata)` replace

- rdrobust: estimation, inference and graphical presentation
- rdrobust calls on the program rdbwselect and optimally selects a bandwidth
- rdplot: part of rdrobust, and plots data for RD analyses

Overview of RDD Packages in Stata, continued...

2. <https://eml.berkeley.edu/~jmccrary/DCdensity/>

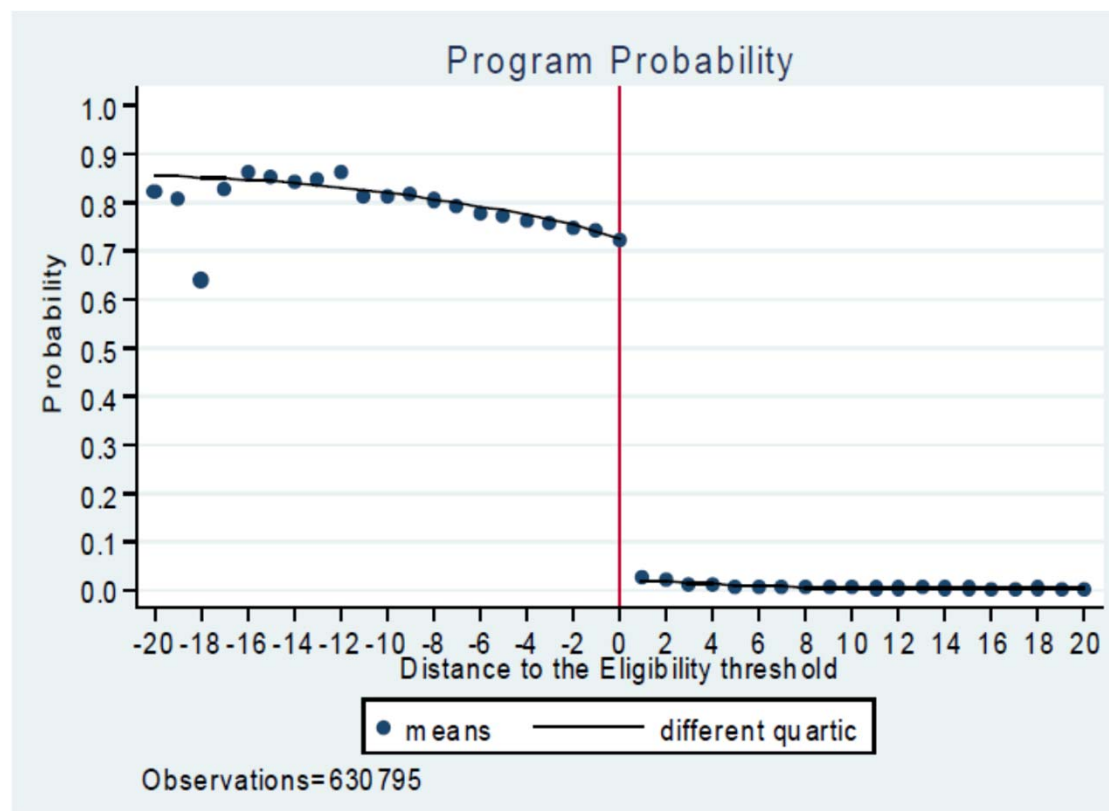
- Download the .ado file and save it with the Stata ado files on your computer. This is a typical location for the ado files (often filed by first letter in file name):



Illustrations of the Discontinuities Several Recent Papers Exploit

Baez and Camacho (2011)

- Estimate the impacts of households in Colombia receiving up to 9 years of conditional cash transfer program payments, provided to those with a sufficiently low poverty index score, on high school completion and test scores
- **Finding:** participant children are 4-8 percentage points more likely than non-participants to finish high school, but have similar test scores



Bauhoff, Hotchkiss, and Smith (2011)

- Estimate the impacts of having publicly-provided medical insurance in the Republic of Georgia, given freely to those with a sufficiently low proxy means test score, on healthcare and health-related outcomes
- **Finding:** The program did not affect utilization of health services, but decreased mean out-of-pocket expenditures for some groups; no systematic impacts on health behavior

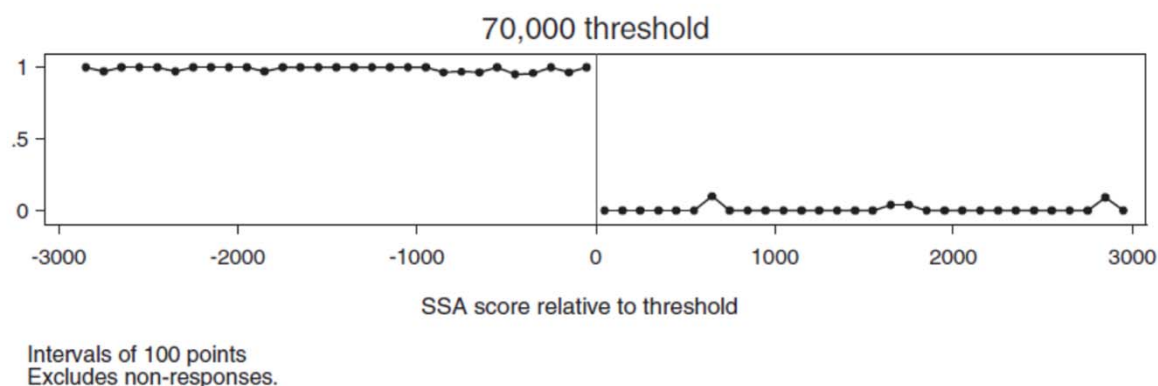


Figure 2. Share of households enrolled in MIP

Manacorda et al. (2011)

- Estimate the impacts of households in Uruguay receiving cash transfer payments, provided to those with a sufficiently low poverty-related eligibility score, on political support
- **Finding:** beneficiary households are 11 to 13 percentage points more likely to favor the current government relative to the previous government

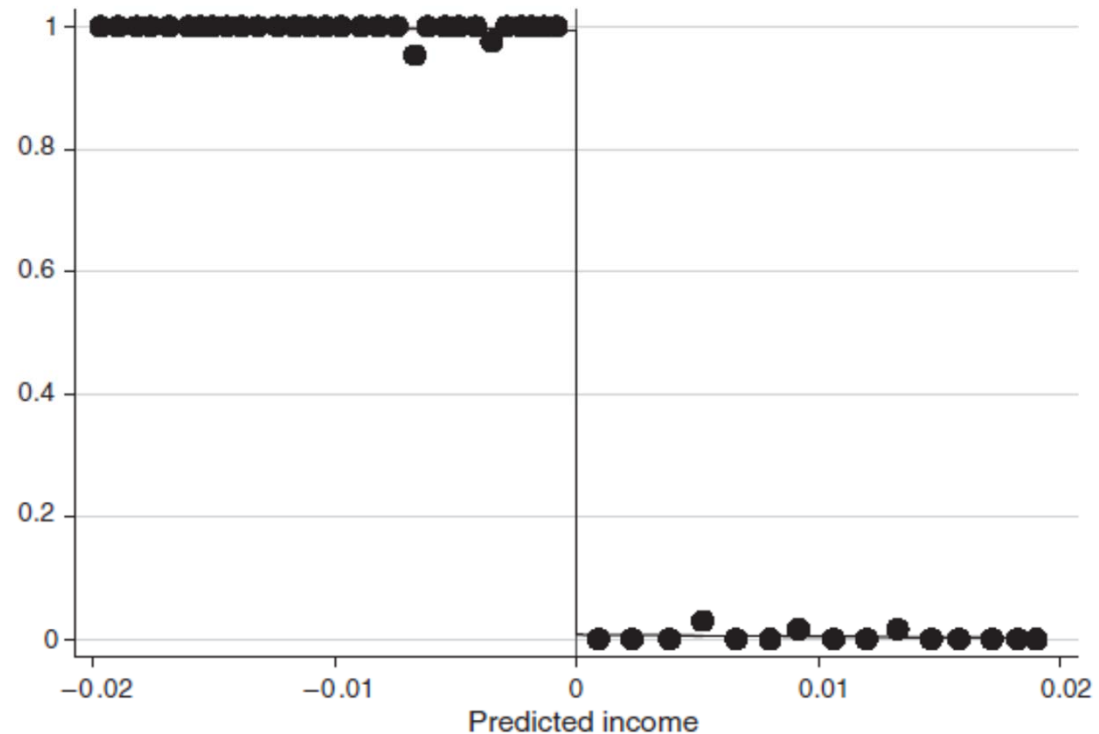


FIGURE 2. PANES PROGRAM ELIGIBILITY AND PARTICIPATION

Kazianga et al. (2013)

- Estimate the effects of constructing a “girl-friendly” primary school in Burkina Faso, provided to the top half of villages in each department based on a score related to the number of students that would benefit from a new school, on student enrollment and test scores
- **Finding:** After 2.5 years, the program increased enrollment by 19 percentage points and increased test scores by 0.41 SD

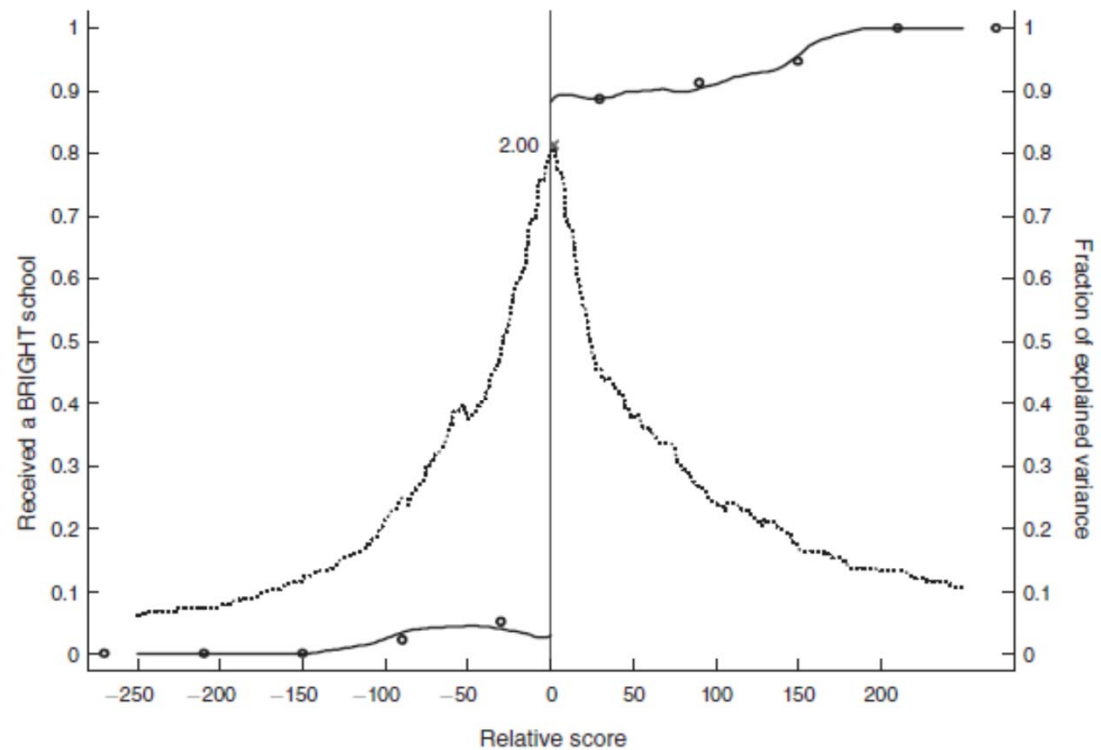
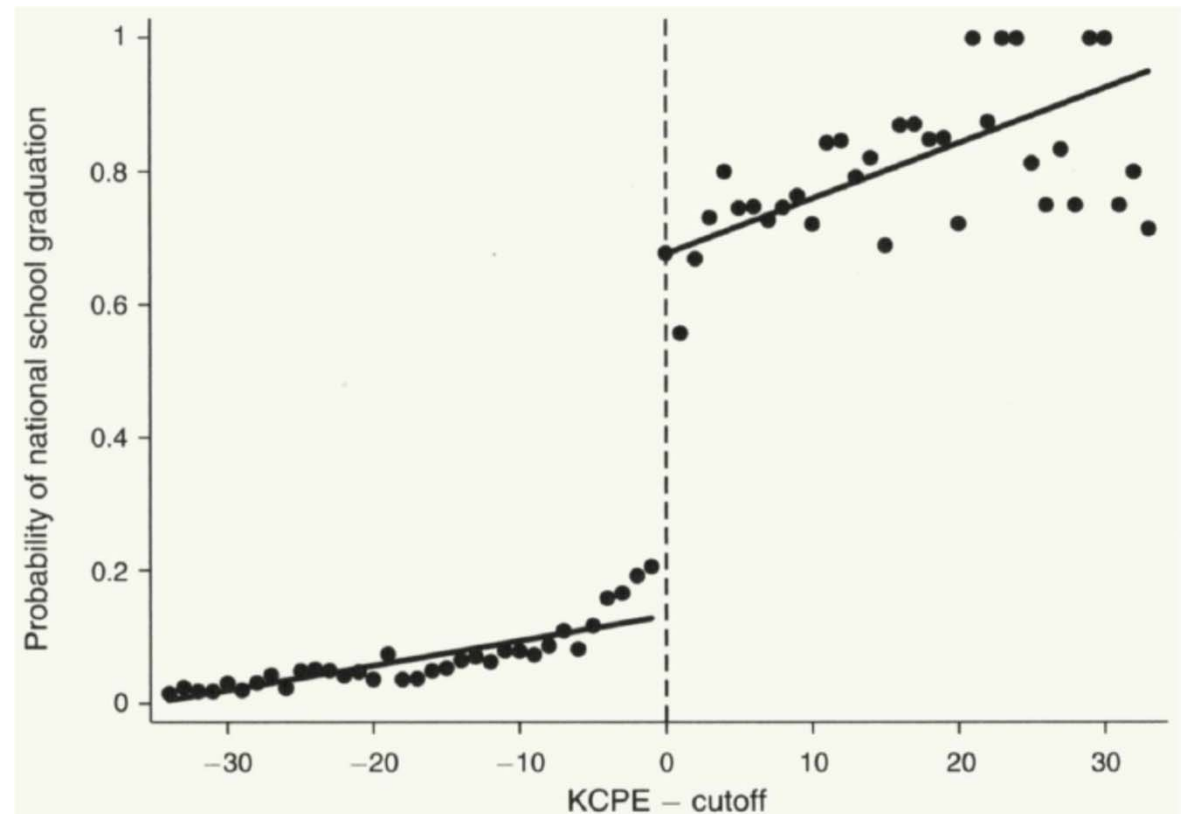


FIGURE 1. INCLUSION IN THE BRIGHT PROGRAM

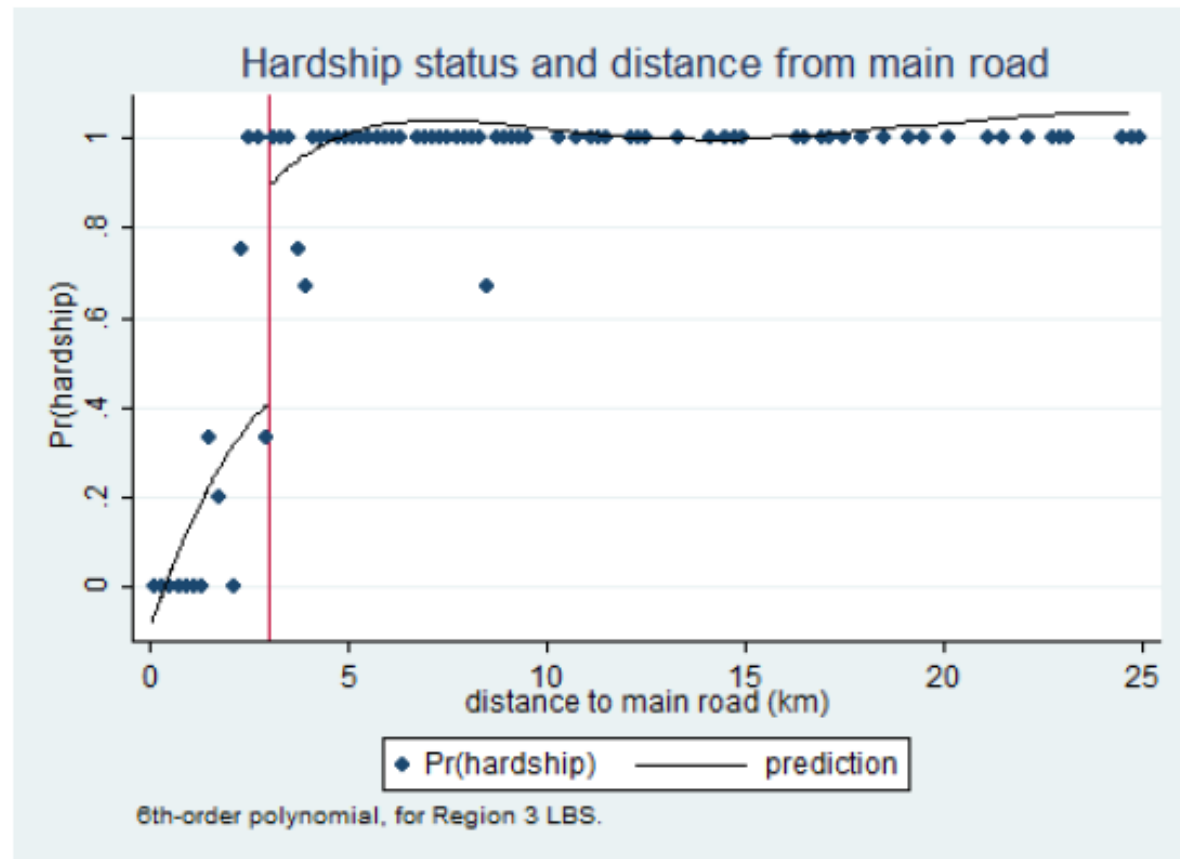
Lucas and Mbiti (2014)

- Estimate the effects of admission to an elite government secondary school in Kenya, for which admission is determined by a (location-specific) minimum test score, on student progression and test scores in secondary school
- **Finding:** Despite their reputations, little evidence of positive impacts on learning outcomes



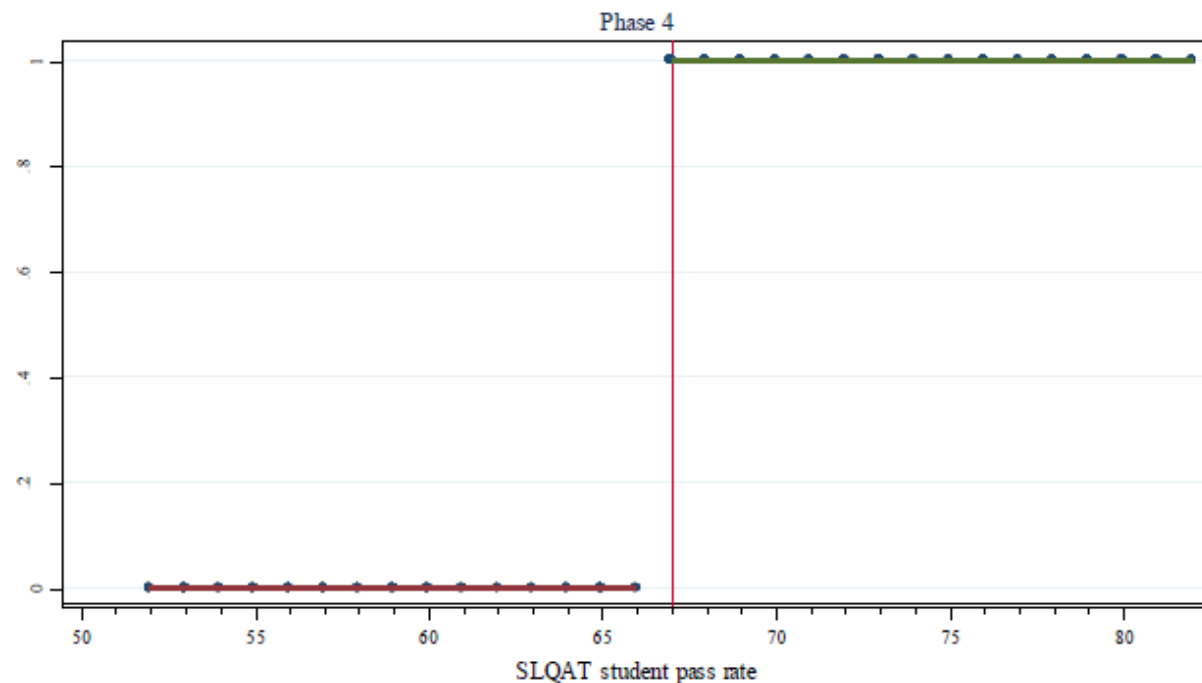
Pugatch & Schroeder (2014)

- Analyze the impact of “hardship allowance” (a 30-40% salary premium) given to teachers in the Gambia if they work above a certain distance from the main road; does it improve teacher quality?
- **Finding:** Increased the share of qualified (certified) teachers by 10 percentage points in remote areas



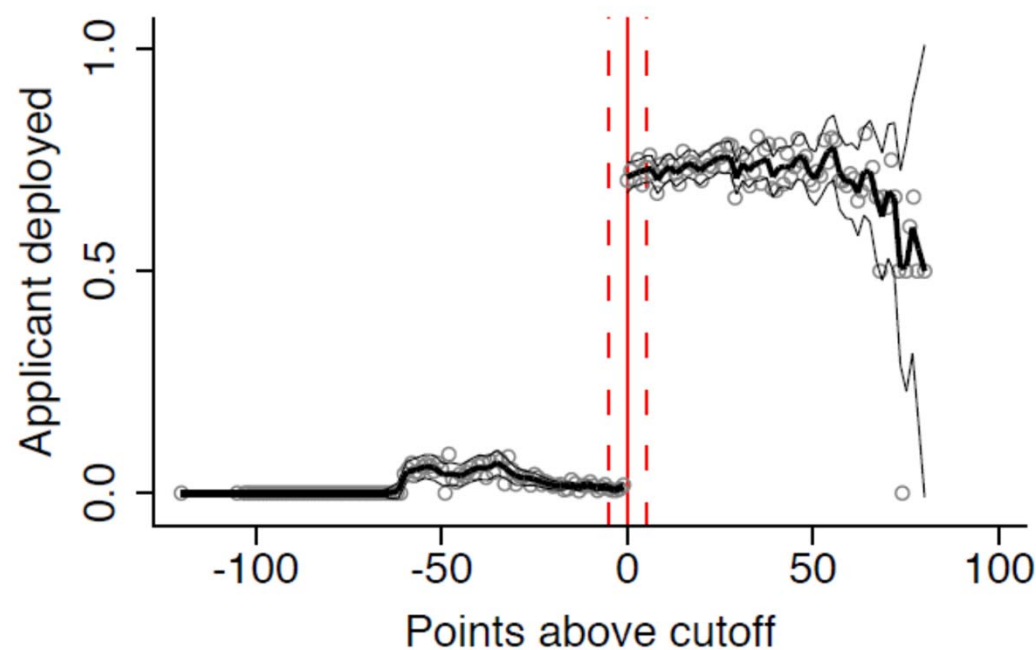
Barrera-Osorio and Raju (2015)

- Estimate the impacts of public per-student subsidies provided to low-cost private schools in Punjab, Pakistan that achieved a minimum pass rate on a specially-designed academic test; consider impacts on student enrolment and school inputs
- **Finding:** Large, positive impacts of the program on school enrolment, number of teachers, and other inputs



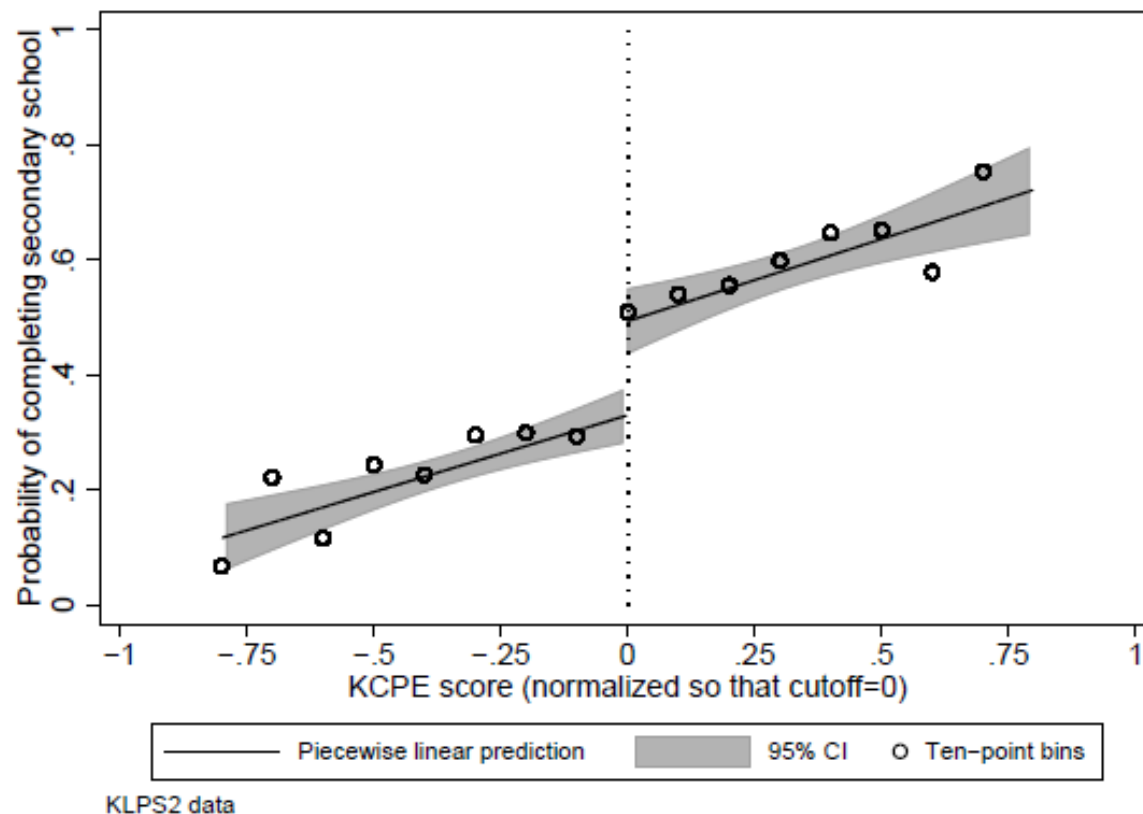
Clemens and Tiongson (2017)

- Estimate the impacts of temporary, partial-household migration to perform high-wage jobs in Korea, which individuals receive if they pass a Korean language test, on Filipino households' spending, finance, and labor supply decisions
- **Finding:** Migration changes households' spending and saving patterns—due to both remittances and shifts in decision-making power back home; it does not reduce labor supply by non-migrants



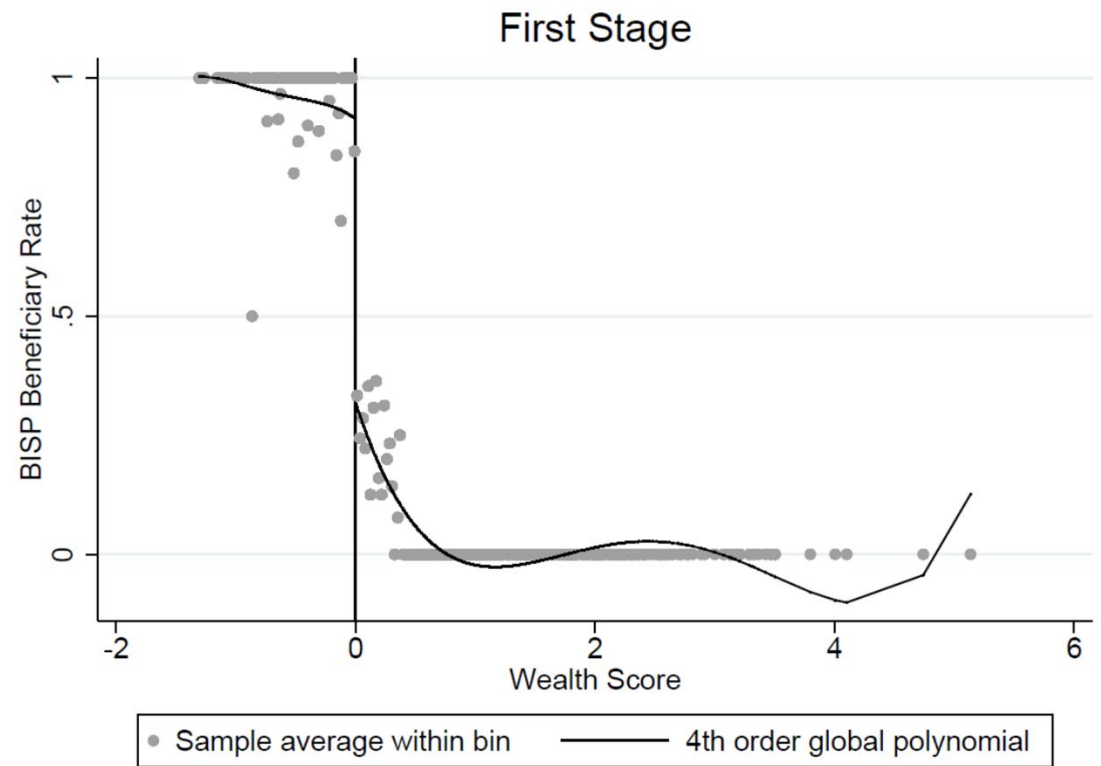
Ozier (2018)

- Estimate the impacts of secondary schooling in Kenya, which students are more likely to receive if they get an above-mean score on a standardized national exam, on several outcomes related to human capital, fertility, and occupational choice
- **Finding:** Reduced incidence of low-skill self-employment for men and lower incidence of teen pregnancy for women



Kosec and Mo (2018)

- Estimate the impacts of receipt of a cash transfer program in Pakistan, which households are more likely to receive if they have a sufficiently low wealth score, on support for government
- **Finding:** Receipt of cash transfers increased support for government, and these effects are significantly larger among individuals primed to feel relatively poor

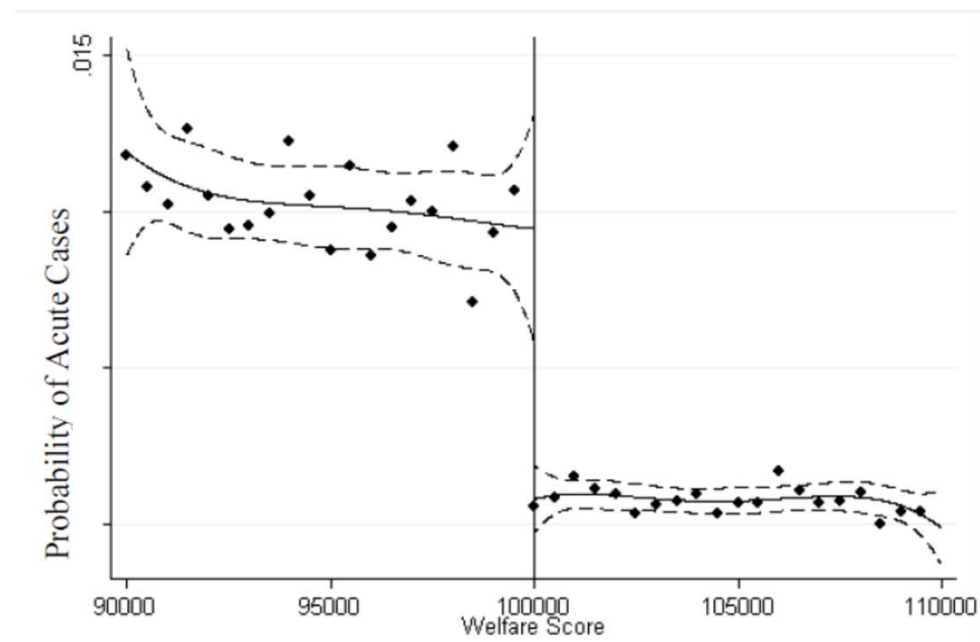


Illustrations of the Impacts Identified by Several Recent Papers

Hou and Chao (2008)

- Estimate the impacts of having publicly-provided medical insurance in the Republic of Georgia, given freely to those with a sufficiently low proxy means test score, on healthcare and health-related outcomes
- **Finding:** Program significantly increased utilization of acute surgeries/ inpatient services by the poor, as shown at right

Figure 3: The Effect of MAP on Utilization of Acute Surgeries/In-Patient Services



Manacorda et al. (2011)

- Estimate the impacts of households in Uruguay receiving cash transfer payments, provided to those with a sufficiently low poverty-related eligibility score, on political support
- **Finding:** beneficiary households are 11 to 13 percentage points more likely to favor the current government relative to the previous government

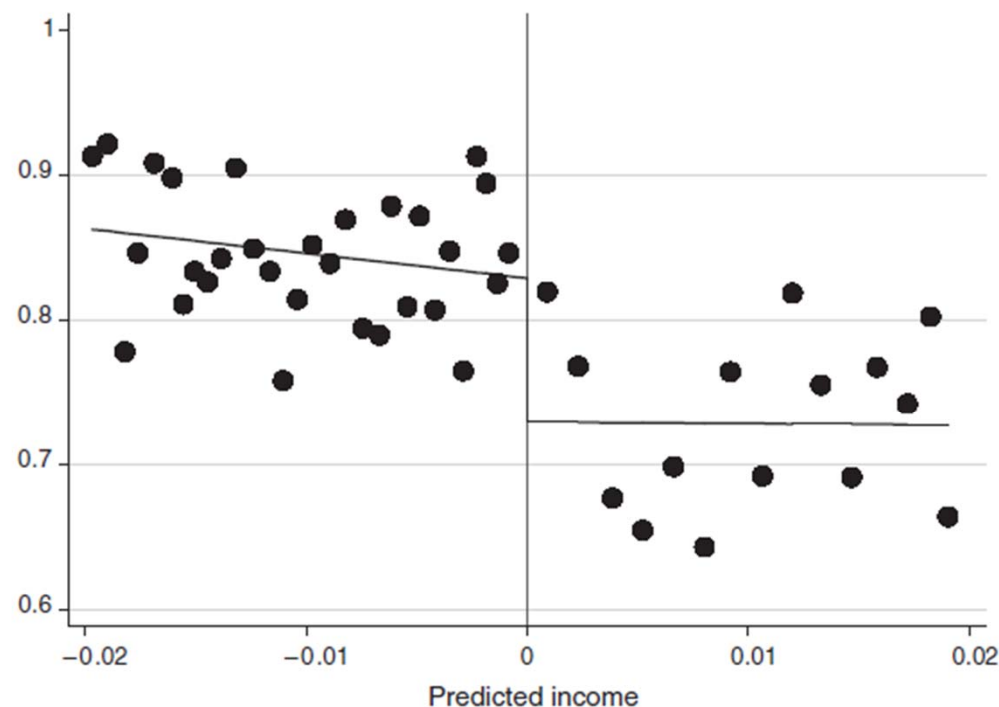
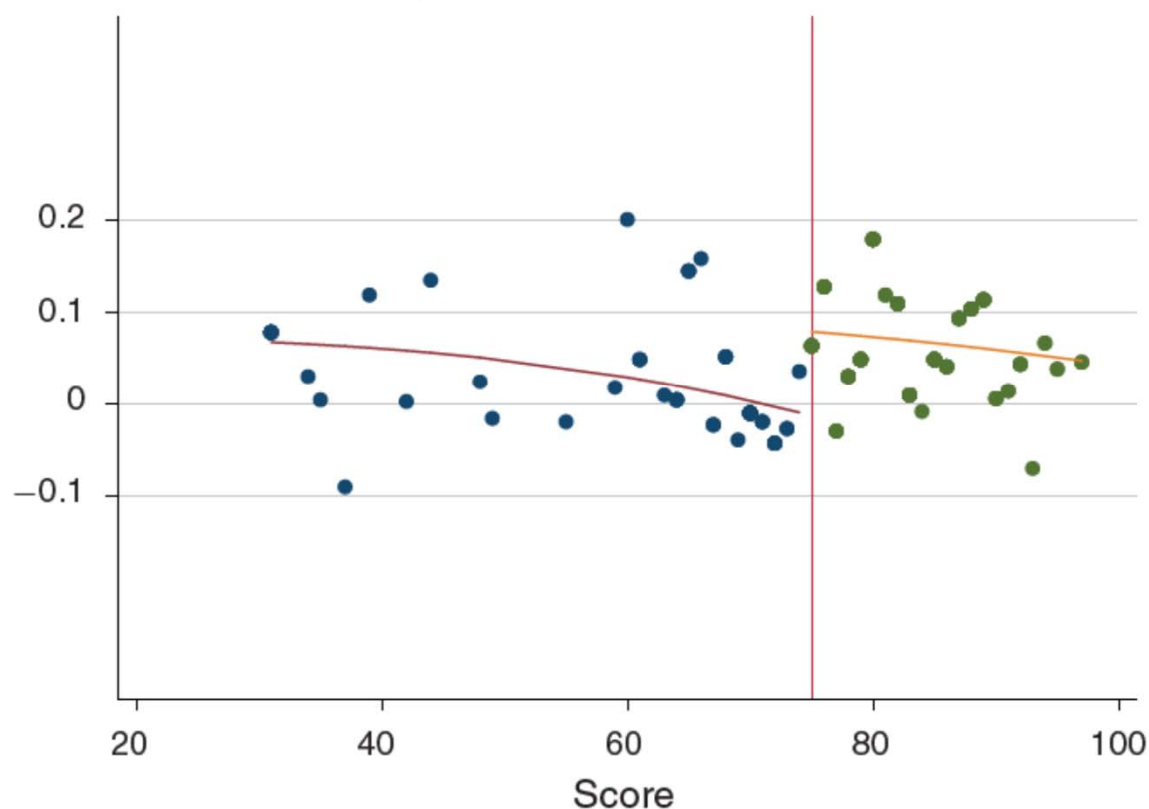


FIGURE 4. PANES PROGRAM ELIGIBILITY AND POLITICAL SUPPORT FOR THE GOVERNMENT, 2008 FOLLOW-UP SURVEY ROUND

Bronzini and Lachini (2014)

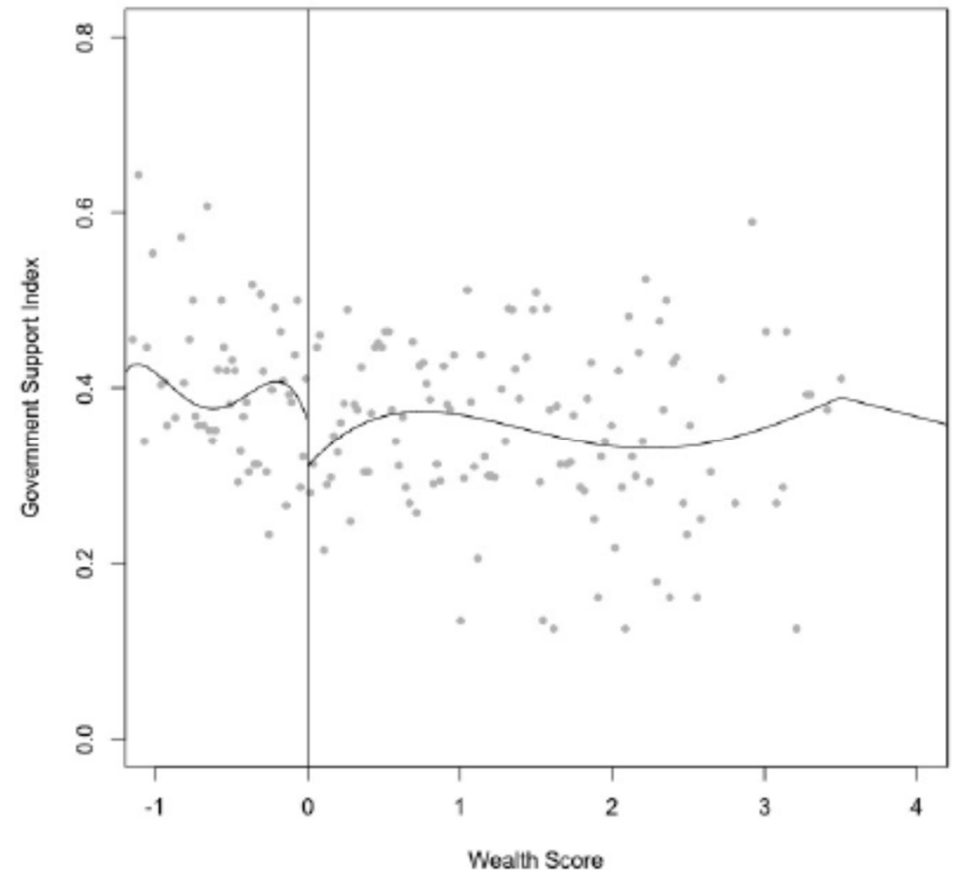
- Evaluate the effects of a unique R&D subsidy program implemented in northern Italy, for which firms were invited to submit proposals for new projects and only those which scored above a certain threshold received the subsidy, on subsequent investment spending by firms
- **Finding:** Null overall impacts, but significant increases in investment by small firms (pictured at right)

Panel A. Total investment/pre-program sales



Kosec and Mo (2018)

- Estimate the impacts of receipt of a cash transfer program in Pakistan, which households are more likely to receive if they have a sufficiently low wealth score, on support for government
- **Finding:** Receipt of cash transfers increased support for government, and these effects are significantly larger among individuals primed to feel relatively poor



Spatial RDD

- Use a very particular cutoff for access to treatment: an administrative or other boundary (border)
- Usually a sharp RD
 - E.g., on one side of the border, one policy is in place, while a different policy is in place on the other
- Key challenge: convincing readers that there are not other factors that change discontinuously at the boundary and which would themselves influence outcomes
 - E.g., many other policies may change at borders

Spatial RD Example: Holmes (1998)

- Examines the impact of being in a state with a right-to-work law (which restricts union activity/ is pro-business) on manufacturing activity
- Policy changes discretely at state borders; the authors use counties within 25 miles of policy change borders (shown in black) and do a spatial RD:

$$Y_i = \left\{ \sum_{j=0}^4 [\gamma_{1,j} A_i + \gamma_{2,j} \cdot (1 - A_i)] B_i^j \right\} + \delta D_l + \varepsilon_i$$

where:

- i = county index, l = state index
- Y = share of employment in manufacturing
- A = dummy for being on segment 2
- D = treatment: dummy for being a pro-business state
- B = continuous boundary marker (distance along policy change boundary from G for segment 1, from C for segment 2)

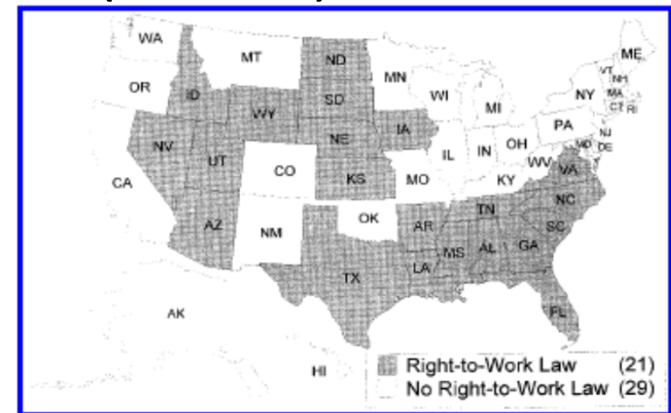


FIG. 1.—Geography of right-to-work laws

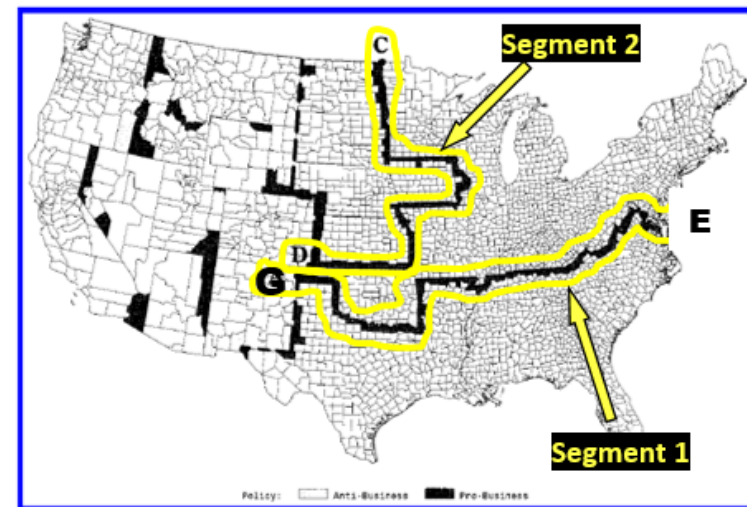
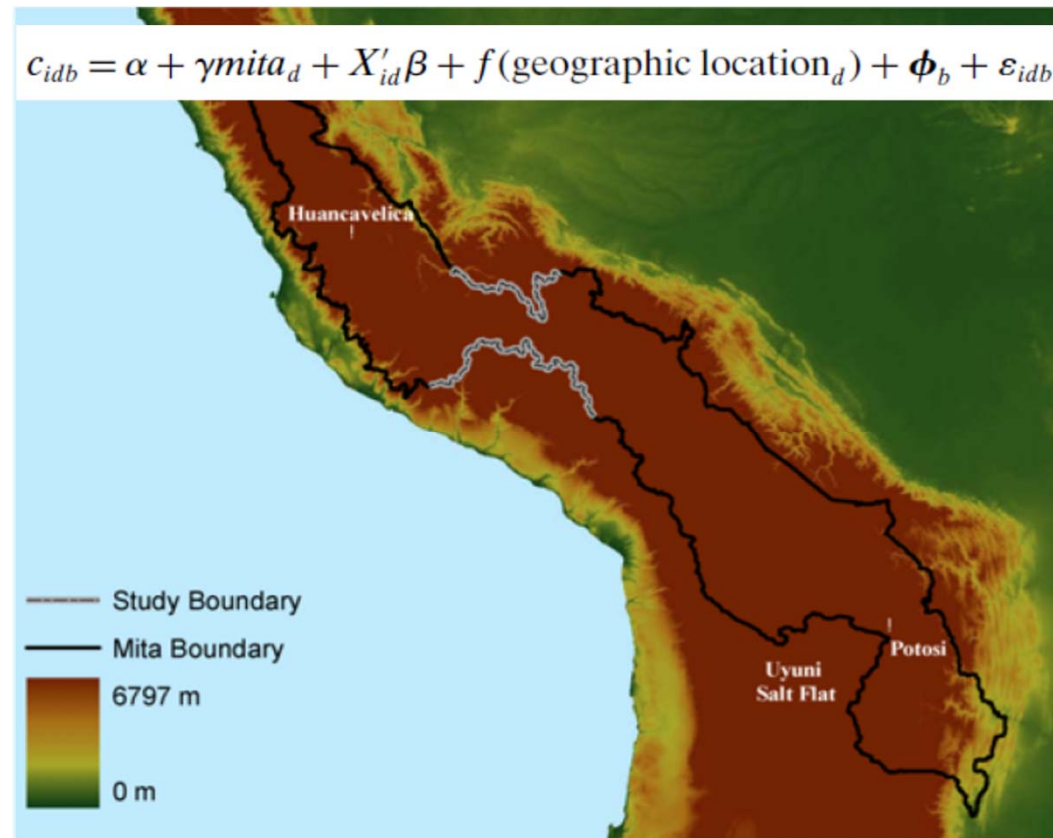


FIG. 3.—Counties within 25 miles of the policy change border

Spatial RD Example: Dell (2010)

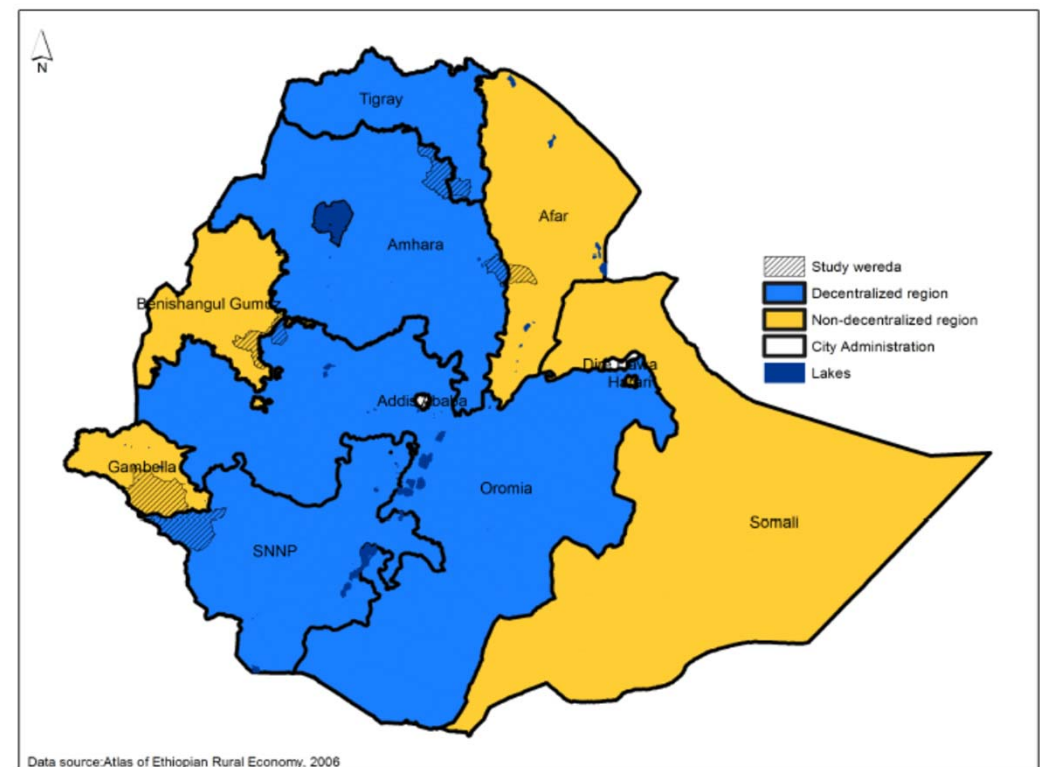
- Examines the long-run impacts of the mining *mita*, a forced labor system instituted by the Spanish in Peru and Bolivia during 1573 – 1812, on modern-day household consumption and child stunting
- Policy changed discretely at the boundary of the subjected region: on one side, all communities had to send 1/7 of adult males to work in silver and mercury mines, while on the other side, all communities were exempt. Author considers the two study boundaries shown.
- Permits spatial RD with the *mita* boundary forming a discontinuity in longitude–latitude space
- Regress outcome on mita dummy, a polynomial in latitude and longitude (primary specification: latitude, longitude, latitude squared, longitude squared, and latitude x longitude), and boundary dummies



Spatial RD Example: Kosec and Mogues (2018)

- In 2001, the central government in Ethiopia decentralized responsibilities for service delivery down to the district level in half of Ethiopia's regions (in blue) but not in the other half (in yellow)
- Permits spatial RD with the regional boundaries between blue and yellow regions forming a discontinuity. Authors consider three survey sites straddling 'policy change' borders
- Control for a running variable which is distance to the regional border and examine the impact of decentralization (a dummy) on the quality of rural service delivery (agricultural services and drinking water services, specifically)
- Use specifications mirroring Holmes (1998) and Dell (2010); both specifications show improvements in services due to decentralization

Figure 1: Map of regions of Ethiopia and their decentralization status as of 2001



Looking for Opportunities to Use RDD

- Look for treatment that is assigned by cutoff
- Cutoffs can be:
 1. Naturally occurring in the world; or
 2. Ones you propose to policymakers or others decisionmakers who are reluctant to use random assignment
- Be sure that outcomes of both treated *and* non-treated units are measured (e.g., you cannot do much if data were only collected for those who got a program)
- Sample as densely as you can around the cutoff