FNCE 926
Empirical Methods in CF

Lecture 8 – Regression Discontinuity

Professor Todd Gormley

Announcements

- Rough draft of research proposal due next week…
  - Just 1-3 page (single-spaced) sketch of your proposal is fine…
    - Should clearly state your question
    - Should give me idea of where you're going with the identification strategy
    - *See grading template on Canvas*
  - Upload it to Canvas by noon next week
  - I will read and then send brief feedback

Background readings for today

- Roberts and Whited
  - *Section 5*
- Angrist and Pischke
  - *Chapter 6*

Outline for Today

- Quick review of last lecture on NE
- Discuss regression discontinuity
  - What is it? How is it useful?
  - How do we implement it?
  - What are underlying assumptions?
- Student presentations of "NE #2" papers

Quick Review *[Part 1]*

- Will adding controls affect diff-in-diff estimates if treatment assignment was random?
  - **Answer =** Not unless you've added 'bad controls', which are controls also affected by treatment. When you've done this, you're no longer estimating the causal effect of treatment
  - Controls (that are exogenous) will just improve precision, but shouldn't affect estimates

Quick Review *[Part 2]*

- What are some standard falsification tests you might want to run with diff-in-diff?
  - **Answers:**
    - Compare ex-ante characteristics of treated & untreated
    - Check timing of treatment effect
    - Run regression using dep. variables that shouldn't be affected by treatment (if it is what we think it is)
    - Check whether reversal of treatment has opposite effect
    - Triple-difference estimation

Quick Review *[Part 3]*

- If you find ex-ante differences in treated and treated, is internal validity gone?
  - **Answer =** Not necessarily but it could suggest non-random assignment of treatment that is problematic… E.g. observations with characteristic 'z' are more likely to be treated and observations with this characteristic **are also likely** to be <u>trending</u> differently for other reasons

Quick Review *[Part 4]*

- Does the absence of a pre-trend in diff-in-diff ensure that differential trends assumption holds and causal inferences can be made?
  - **Answer =** Sadly, no. We can never prove causality with 100% confidence. It could be that trend was going to change <u>**after**</u> treatment for reasons *unrelated* to treatment

Quick Review *[Part 5]*

- How are multiple events that affect multiple groups helpful?
  - **Answer =** Can check that treatment effect is similar across events; helps reduce concerns about violation of parallel trends since there would need to be violation for each event

Quick Review *[Part 6]*

- How are triple differences helpful and reducing concerns about violation of parallel trends assumption?
  - **Answer =** Before, an "identification policeman" would just need a story about why treated might be trending differently after event for other reasons… Now, he/she would need story about why that different trend would be particularly true for subset of firms that are more sensitive to treatment

Regression Discontinuity – *Outline*

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
- Estimating regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

Basic idea of RDD

- The basic idea of regression discontinuity (RDD) is the following:
  - Observations (e.g. firm, individual, etc.) are 'treated' based on known cutoff rule
    - E.g. for some observable variable, x, an observation is treated if $x \geq x'$
    - This cutoff is what creates the <u>discontinuity</u>
  - Researcher is interested in how this treatment affects outcome variable of interest, y

Examples of RDD settings

- If you think about it, these type of cutoff rules are commonplace in finance
  - A borrower FICO score > 620 makes securitization of the loan more likely
    - Keys, et al (QJE 2010)
  - Accounting variable x exceeding some threshold causes loan covenant violation
    - Roberts and Sufi (JF 2009)

RDD is like difference-in-difference…

- Has similar flavor to diff-in-diff natural experiment setting in that you can illustrate identification with a figure
  - Plot outcome y against independent variable that determines treatment assignment, x
  - Should observe sharp, discontinuous change in y at the cutoff value of x'

But, RDD is different…

- RDD has some key differences…
  - Assignment to treatment is **NOT** random; assignment is based on value of x
  - When treatment only depends on x (what I'll later call "sharp RDD", there is no overlap in treatment & controls; i.e. we never observe the same x for a treatment and a control

RDD randomization assumption

- Assignment to treatment and control isn't random, but whether *individual* observation is treated is assumed to be random
  - I.e. researcher assumes that observations (e.g. firm, person, etc.) can't perfectly manipulate their x value
  - Therefore, whether an observation's x falls immediately above or below key cutoff x' is random!

Regression Discontinuity – *Outline*

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
  - Notation & 'sharp' vs. fuzzy assumption
  - Assumption about local continuity
- Estimating regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

RDD terminology

- *x* is called the "forcing variable"
  - Can be a single variable or multiple variables; but for simplicity, we'll work with a single variable
- *x'* is called the "threshold"
- *y(0)* is outcome absent treatment
- *y(1)* is outcome with treatment

Two types of RDD

- Sharp RDD
  - Assignment to treatment only depends on x; i.e. if $x \geq x'$ you are treated with probability 1
- Fuzzy RDD
  - Having $x \geq x'$ only increases *probability* of treatment; i.e. other factors (besides x) will influence whether you are actually treated or not

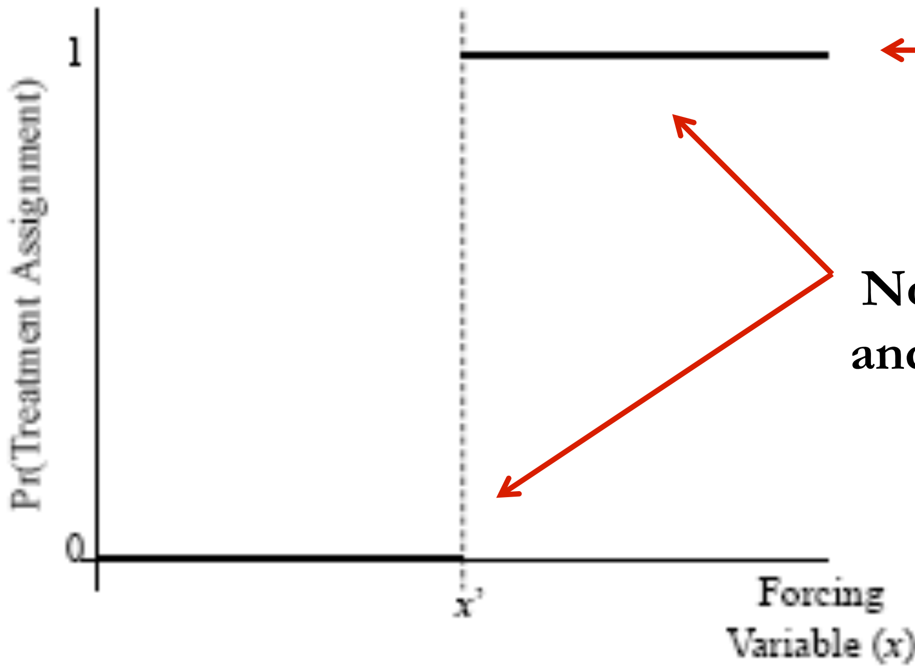
<u>**Sharp**</u> RDD assumption #1

- Assignment to treatment occurs through known and <u>deterministic</u> decision rule:

$$d = d(x) = \begin{cases} 1 & \text{if } x \geq x' \\ 0 & \text{otherwise} \end{cases}$$

  - Weak inequality and direction of treatment is unimportant *[i.e. could easily have x < x']*
  - <u>But</u>, it is important that there exists *x*'s around the threshold value

Sharp RDD assumption #1 – *Visually*



Probability of treatment moves from 0 to 1 around threshold value x'

No untreated for $x > x'$ and no treated for $x < x'$

<u>Only</u> x determines treatment

Figure is from Roberts and Whited (2010)

Sharp RDD – *Examples*

- Ex. #1 – PSAT score > *x'* means student receives national merit scholarship
  - Receiving scholarship was determined solely based on PSAT scores in the past
  - Thistlewaithe and Campbell (1960) used this to study effect of scholarship on career plans

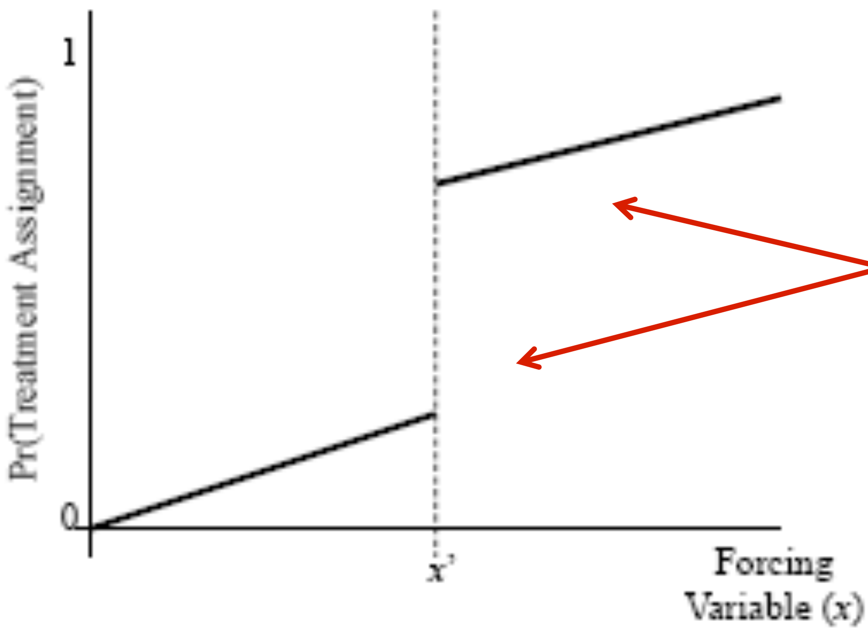
<u>**Fuzzy**</u> RDD assumption #1

- Assignment to treatment is <u>stochastic</u> in that only the *probability* of treatment has known discontinuity at *x'*

$$0 < \lim_{x \downarrow x'} \Pr(d = 1 \mid x) - \lim_{x \uparrow x'} \Pr(d = 1 \mid x) < 1$$

  - Can also go other way, i.e. probability of treatment drops at *x'*; all that is needed is jump in the <u>probability</u> of treatment at *x'*

Fuzzy RDD assumption #1 – *Visually*



Treatment probability increases at x'

Some untreated for $x > x'$ and some treated for $x < x'$

Treatment is <u>not</u> purely driven by x

Figure is from Roberts and Whited (2010)

Fuzzy RDD – *Example*

- Ex. #1 – FICO score > 620 increases likelihood of loan being securitized
  - But, extent of loan documentation, lender, etc., will matter as well…

Sharp *versus* Fuzzy RDD

- This subtle distinction affects exactly how you estimate the causal effect of treatment
  - With Sharp RDD, we will basically compare average *y* immediate above and below *x'*
  - With fuzzy RDD, the average change in *y* around threshold *understates* causal effect **[Why?]**
    - **Answer =** Comparison assumes all observations were treated, but this isn't true; if all observations had been treated, observed change in *y* would be even larger; we will need rescale based on change in probability

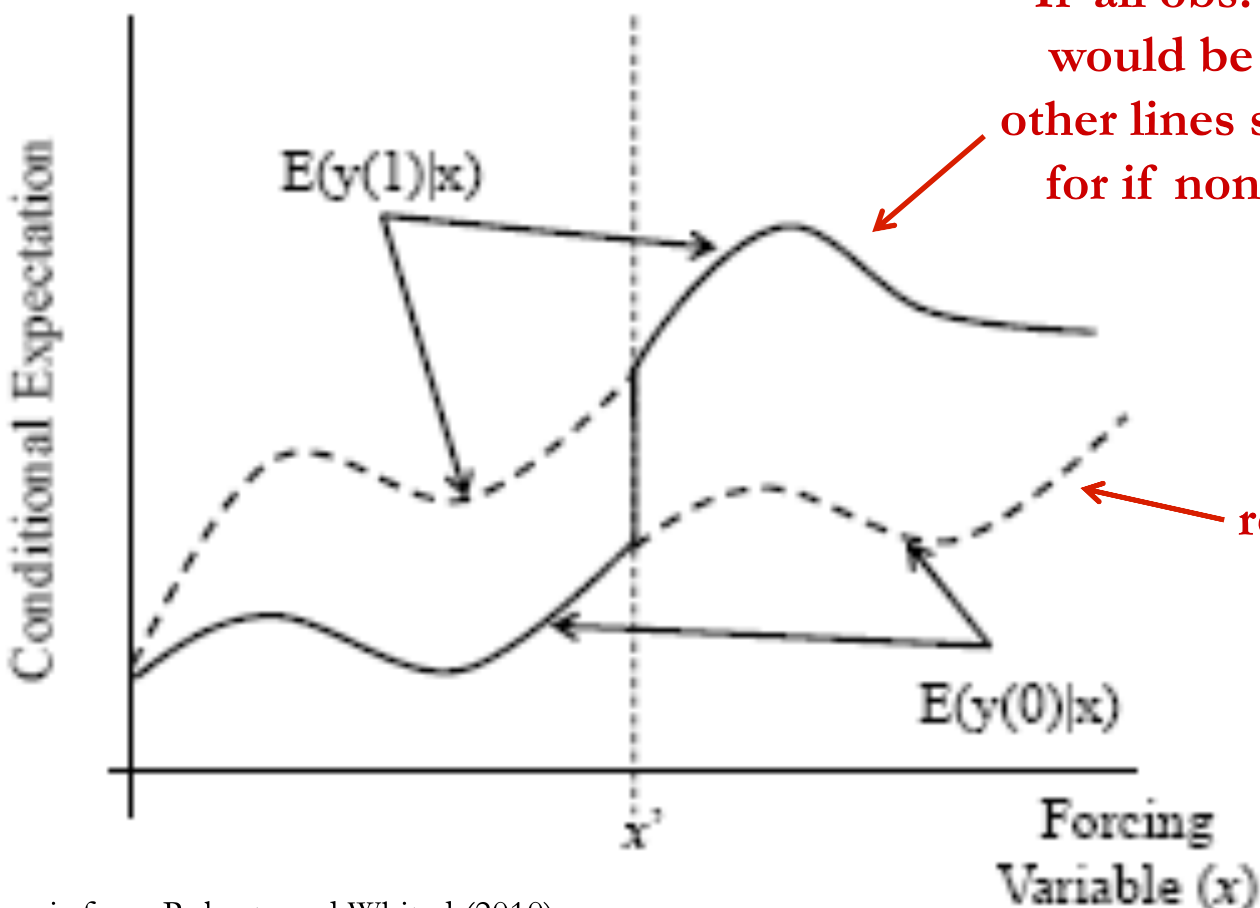
Regression Discontinuity – *Outline*

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
  - Notation & 'sharp' vs. fuzzy assumption
  - Assumption about local continuity
- Estimating regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

RDD assumption #2

- But, <u>both</u> RDDs share the following assumption about **local continuity**
- <u>Potential</u> outcomes, *y(0)* and *y(1)*, conditional on forcing variable, *x*, are continuous at threshold *x'*
  - **In words:** *y* would be a smooth function around threshold <u>absent</u> treatment; i.e. don't expect any jump in *y* at threshold *x'* <u>absent</u> treatment

RDD assumption #2 – *Visually*



If all obs. had been treated, *y* would be smooth around *x'*; other lines says equivalent thing for if none had been treated

Dashed lines represent unobserved counterfactuals

Figure is from Roberts and Whited (2010)

Regression Discontinuity – Outline

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
- Estimating regression discontinuity
  - Sharp regression discontinuity
  - Fuzzy regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

How **not** to do Sharp RDD…

- Given this setting, will the below estimation reveal causal effect of treatment, *d*, on *y*?

$$y_i = \beta_0 + \beta_1 d_i + u_i$$

  - **Answer** = Unlikely! *d* is correlated with *x*, and if *x* affects *y*, then there will be omitted variable!
    - E.g. Borrowers FICO score, used in Keys, et al (2010) affects likelihood of default… therefore, above regression can **NOT** be used to determine effect of securitization on default risk

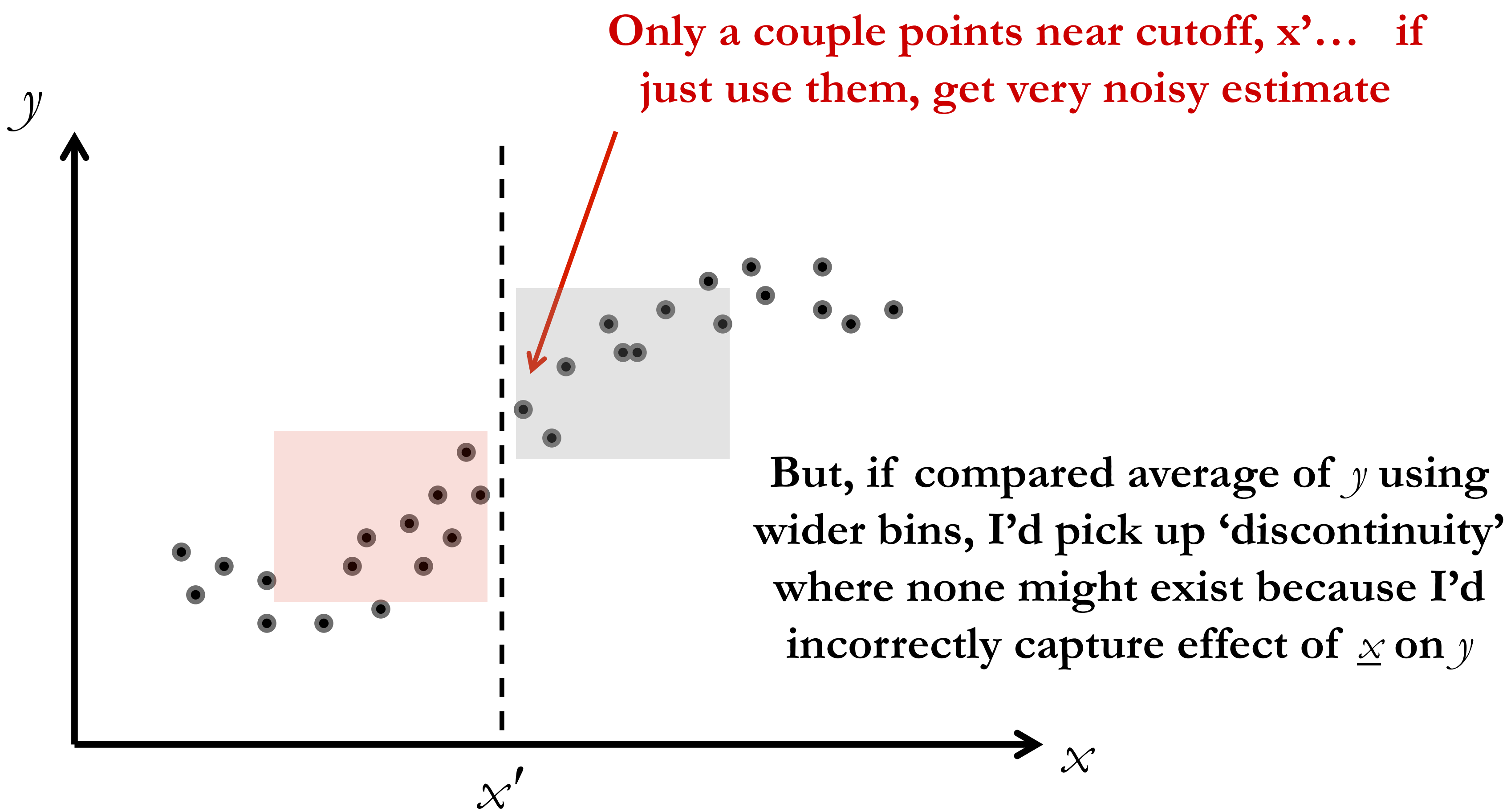
How not to do Sharp RDD… *[Part 2]*

- How can we modify previous regression to account for this omitted variable?
  - **Answer:** Control for *x*!
  - So, we could estimate: $y_i = \beta_0 + \beta_1 d_i + \beta_2 x_i + u_i$
  - **But, why might this still be problematic?**
    - **Answer:** (1) Assumes effect of *x* is linear, and (2) doesn't really make use of random assignment, which is really occurring near the threshold

Bias *versus* Noise

- Ideally, we'd like to compare average x right below and right above x'; what is tradeoff?
  - **Answer:** We won't have many observations and estimate will be very <u>noisy</u>. A wider range of x on each side reduces this noise, but increases risk of <u>bias</u> that observations further from threshold might vary for other reasons (including because of the direct effect of x on y)

Bias *versus* Noise – Visual



Estimating Sharp RDD

- There are generally two ways to do RDD that weigh that try to balance this tradeoff between bias and noise
  - Use all data, but control for effect of x on y in a very general and rigorous way
  - Use less rigorous controls for effect of x, but only use data in small window around threshold

Estimating Sharp RDD, *<u>Using all data</u>*

- First approach uses all the data available and estimates two separate regressions

$$y_i = \beta^b + f(x_i - x') + u_i^b$$

Estimate using only data <u>below</u> x'

$$y_i = \beta^a + g(x_i - x') + u_i^a$$

Estimate using only data <u>above</u> x'

  - Just let $f(\ )$ and $g(\ )$ be any continuous function of $x_i - x'$, where $f(0)=g(0)=0$
  - **Treatment effect =** $\boldsymbol{\beta}^a - \boldsymbol{\beta}^b$

Interpreting the Estimates…

$$y_i = \beta^b + f(x_i - x') + u_i^b$$

$$y_i = \beta^a + g(x_i - x') + u_i^a$$

- Why are *f( )* and *g( )* included?

  Answer = They are there to control for underlying effect of *x* on *y*

- What do β^b and β^a estimate?
  - **Answer =** β^b is $E[y|x=x']$ from below, and β^b is $E[y|x=x']$ from above

Easier way to do this estimation

- Can do all in one step; just use **all** the data at once and estimate:

$$y_i = \alpha + \beta d_i + f(x_i - x') + d_i \times g(x_i - x') + u_i$$

Recall:
d_i = indicator for $x \geq x'$

Estimate for β will equal $\beta^a - \beta^b$

Controls for relationship between x and y both above and below x'

What would we be assuming if we drop $d \times g(\ )$?

Tangent about dropping *g( )*

- **Answer:** If you drop $d_i \times g(x_i - x')$, you assume functional form between x and y is same above and below x'
  - Can be strong assumption, which is probably why it shouldn't be only specification used
  - But, Angrist and Pischke argue it usually doesn't make a big difference in practice

What should we use for *f( )* and *g( )*?

- In practice, a high-order polynomial function is used for both *f( )* and *g( )*
  - E.g. You might use a cubic polynomial

$$y_i = \alpha + \beta d_i + \sum_{s=1}^{3} \gamma_s^b (x_i - x')^s + \sum_{t=1}^{3} \gamma_t^a d_i (x_i - x')^t + u_i$$

  - **How might you determine the correct order of polynomial to use in practice?**

Sharp RDD – **Robustness Check**

- Ultimately, correct order of polynomial is unknown; so, **best to show robustness**
  - Should try to illustrate that findings are robust to different polynomial orders
  - Can do graphical analysis to provide a visual inspection that polynomial order is correct
    [I will cover graphical analysis in a second]

Estimating Sharp RDD, *Using Window*

- Do same RDD estimate as before, but…
  - Restrict analysis to smaller window around x'
  - Use lower polynomial order controls
- E.g. estimate below model in window $x' - \Delta \leq x \leq x' + \Delta$ for some $\Delta > 0$

$$y_i = \alpha + \beta d_i + \gamma^b (x_i - x') + \gamma^a d_i (x_i - x') + u_i$$

Controls are now just linear in this example

Practical issues with this approach

- What is appropriate window width and appropriate order of polynomial?
  - **Answer =** There is no right answer! But, it probably isn't as necessary to have as complicated of polynomial in smaller window
  - **But, best to just show robustness to choice of window width, Δ, and polynomial order**

Tradeoff between two approaches

- Approach with smaller window can be subject to greater noise, but advantage is…
  - Doesn't assume constant effect of treatment for all values of x in the sample; in essence you are estimating local avg. treatment effect
  - Less subject to risk of bias because correctly controlling for relationship between x and y is less important in the smaller window

Regression Discontinuity – Outline

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
- Estimating regression discontinuity
  - Sharp regression discontinuity
  - Graphical analysis
  - Fuzzy regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

Graphical Analysis of RDD

- Can construct a graph to visually inspect whether a discontinuity exists and whether chosen polynomial order seems to fit the data well
  - Always good idea to do this graph with RDD; provides sanity check and visual illustration of variation driving estimate

How to do RDD graphical analysis *[P1]*

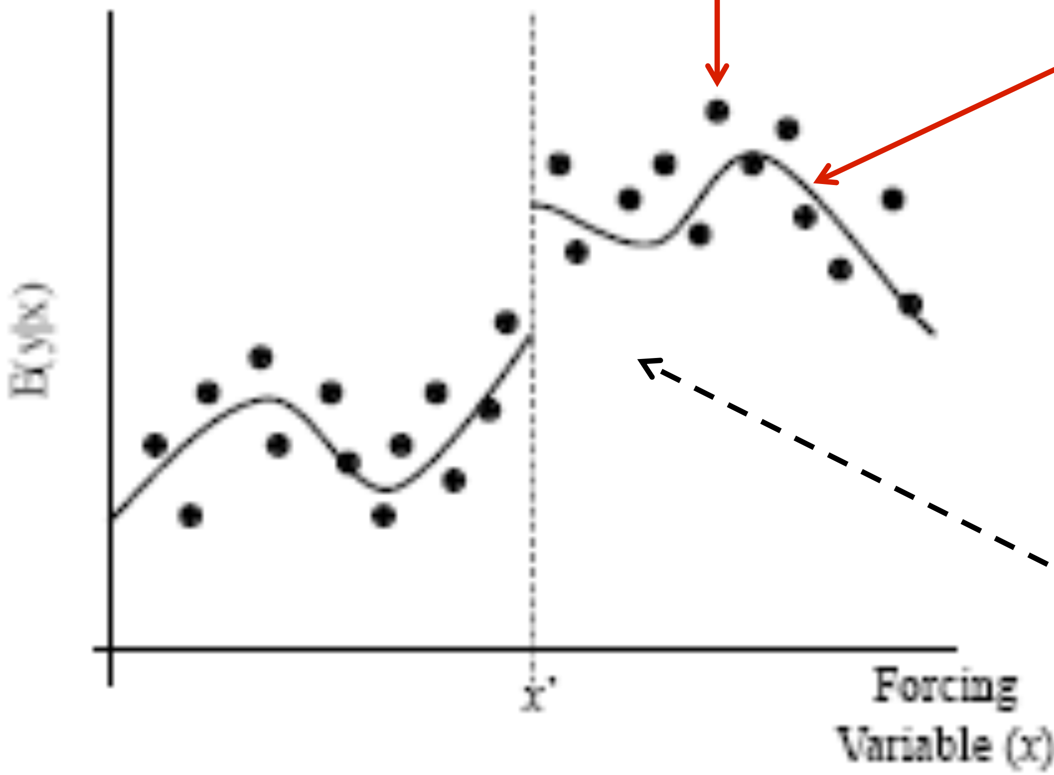
- First, divide x into bins, making sure no bin contains x' as an interior point
  - E.g., if x ranges between 0 and 10 and treatment occurs for $x \geq x' = 5$, you could construct 10 bins, $[0,1)$, $[1,2)$,…$[9,10]$
  - Or, if $x' = 4.5$, could use something like $[0,0.5)$, $[0.5,1.5)$, $[1.5, 2.5)$, etc.

How to do RDD graphical analysis *[P2]*

- Second, calculate average y in each bin, and plot this above midpoint for each bin
  - Plotted averages represent a non-parametric estimate of $E[y|x]$
- Third, estimate your RDD and plot predicted values of y from the estimation

Example of supportive graph

Each dot is average *y* for corresponding bin



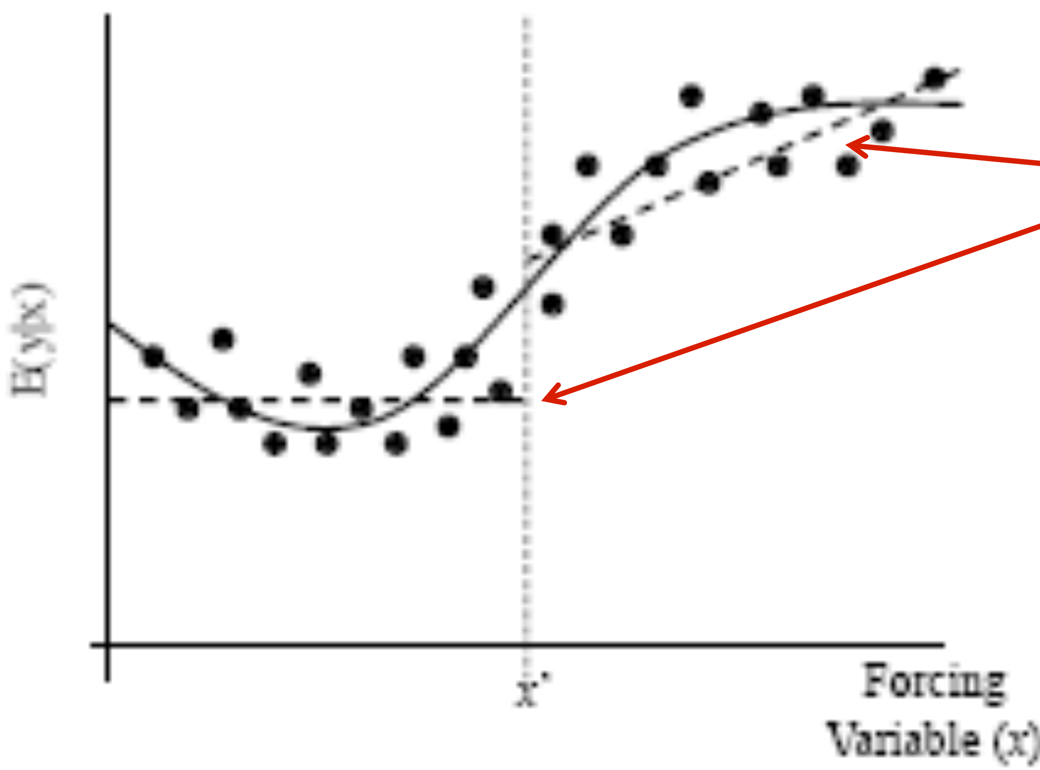
Solid line is predicted values of *y* from RDD regression

Fifth-order polynomial was needed to fit the non-parametric plot

Discontinuity is apparent in both estimation and non-parametric plot

Figure is from Roberts and Whited (2010)

Example of **non**-supportive graph



Dash lines would have been predicted values from linear RDD [i.e. polynomial of order 1]

But, looking at non-parametric graph would make clear that a cubic version (which is plotted as solid line) would show no effect!

Figure is from Roberts and Whited (2010)

RDD Graphs – *Miscellaneous Issues*

- Non-parametric plot shouldn't suggest jump in *y* at other points besides *x'* *[Why?]*
  - **Answer =** Calls into question internal validity of RDD; possible that jump at *x'* is driven by something else that is unrelated to treatment

Bin Width in RDD graphs

- What is optimal # of bins (i.e. bin width)? What is the tradeoff with smaller bins?
  - **Answer** = Choice of bin width is subjective because of tradeoff between precision and bias
    - By including more data points in each average, wider bins give us more precise estimate of $E[y|x]$ in that region of x
    - But, wider bins might be biased if $E[y|x]$ is not constant (i.e. has non-zero slope) within each of the wide bins

Test of overly wide graph bins

1. Construct indicator for each bin
2. Regress y on these indicators and their interaction with forcing variable, x
3. Do joint F-test of interaction terms
   - If fails, that suggests there is a slope in some of the bins… i.e. bins are too wide
   - See Lee and Lemieux (*JEL* 2010) for more details and another test

Regression Discontinuity – Outline

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
- Estimating regression discontinuity
  - Sharp regression discontinuity
  - Graphical analysis
  - Fuzzy regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

Intuition for Fuzzy RDD

- As noted earlier, comparison of average y immediately above and below threshold (as done in Sharp RDD) won't work
  - Again, not all observations above threshold are treated and not all below are untreated; $x > x'$ just increases probability of treatment…
- So, what can we do?
  - **Answer** = use $x \geq x'$ as IV for treatment!!!

Fuzzy RDD Notation

- Need to <u>relabel</u> a few variables
  - d_i = 1 if treated by event of interest; 0 otherwise
  - And, define new threshold indicator, T_i

$$T = T(x) = \begin{cases} 1 & \text{if } x \geq x' \\ 0 & \text{otherwise} \end{cases}$$

    - E.g. d_i = 1 if loan is securitized, T_i = 1 if FICO score is greater than 620, which increases probability loan is securitized

Estimating Fuzzy RDD *[Part 1]*

- Estimate the below 2SLS model

$$y_i = \alpha + \beta d_i + f(x_i - x') + u_i$$

  - Where you use T_i as IV for d_i
  - **What are necessary assumptions of IV?**
    - **Answer** = T_i affects probability of $d_i = 1$ *[relevance condition]* but is unrelated to y conditional on d_i and controls $f(\,)$ *[exclusion condition]*
    - These will be satisfied under earlier assumptions!

Estimating Fuzzy RDD *[Part 2]*

- Again, *f( )* is typically a polynomial function
- Unlike sharp RDD, it isn't as easy to allow functional form to vary above & below
  - So, if worried about different functional forms, what can you do to mitigate this concern?
  - **Answer =** Use a tighter window around event; this is less sensitive to functional form, *f(x)*

Fuzzy RDD – *Practical Issues*

- Exactly same practical issues arise
  - Correct polynomial order is unknown
  - Can also use small bandwidth (rather than all the data) with lower order polynomial order
- In general, show robustness to different specifications and show graphs!

Fuzzy RDD Graphs

- Do same graph of y on x as with sharp RDD
  - Again, should see discontinuity in y at x'
  - Should get sense that polynomial fit is good
- In fuzzy RDD, should also plot similar graph for treatment dummy, d, on x *[Why?]*
  - **Answer** = Helps make sure there is discontinuity of <u>treatment probability</u> at the threshold value

Regression Discontinuity – *Outline*

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
- Estimating regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

Robustness Tests for Internal Validity

- Already discussed a few…
  - Show graphical analysis *[picture is helpful!]*
  - Make sure finding robust to chosen polynomial
  - Make sure finding robust to chosen bandwidth
- Here are some others worth checking…

Additional check #1 – No manipulation

- Researcher should ask the following…
  - Is there any reason to believe threshold x' was chosen because of some *pre-existing* discontinuity in y or <u>lack of comparability</u> above and below x'?
    - If so… a clear violation of local continuity assumption
  - Is there any way or reason why subjects might manipulate their x around threshold?
    [Why ask this?]

Why manipulation can be problematic…

- **Answer** = Again, subjects' ability to manipulate x can cause violation of local continuity assumption
  - I.e. with manipulation, y might exhibit jump around x' absent treatment because of manipulation
    - E.g. in Keys, et al. (QJE 2010) default rate of loans at FICO = 620 might jump regardless if weak borrowers manipulate their FICO to get the lower interest rates that one gets immediately with FICO above 620

And, why it isn't always a problem

- Why isn't subjects' ability to manipulate x always a problem?
  - **Answer =** If they can't perfectly manipulate it, then there will still be randomness in treatment
    - I.e. in small enough bandwidth around x', there will still be randomness because idiosyncratic shocks will push some above and some below threshold even if they are trying to manipulate the x

An informal test for manipulation

- Look for bunching of observations immediately above or below threshold
  - Any bunching would suggest manipulation
  - But, why is this not a perfect test?
    - **Answer =** It assumes manipulation is monotonic; i.e. all subjects either try to get above or below x'. This need not be true in all scenarios

Additional check #2 – Balance tests

- RDD assumes observations near but on opposite sides of cutoff are comparable… **so, check this!**
  - I.e. using graphical analysis or RDD, make sure other observable factors that might affect y don't exhibit jump at threshold x'
  - Why doesn't this test prove validity of RDD?
    - **Answer:** There could be discontinuity in unobservables! Again, there is no way to prove causality

Using covariates instead…

- You could also just add these other variables that might affect *y* as controls
  - If RDD is internally valid, will these additional controls effect estimate, and if so, how?
  - **Answer:** Similar to NE, they should <u>only</u> affect precision of estimate. If they affect the estimated treatment effect, you've got bigger problems; **Why?**
    - You might have 'bad controls'
    - Or, observations around threshold aren't comparable ☹

Additional check #3 – Falsification Tests

- If threshold x' only existed in certain years or for certain types of observations…
  - E.g. law that created discontinuity was passed in a given year, but didn't exist before that, or maybe the law didn't apply to some firms
- Then, what is a good falsification test?
  - **Answer =** Make sure no effect in years where there was no discontinuity or for firms where there isn't supposed to be an effect!

Regression Discontinuity – Outline

- Basic idea of regression discontinuity
- Sharp *versus* fuzzy discontinuities
- Estimating regression discontinuity
- Checks on internal validity
- Heterogeneous effects & external validity

Heterogeneous effects (HE)

- If think treatment might differentially affect observations based on their x, then need a few additional assumptions for RDD to identify the local average treatment effect

1. Effect of treatment is locally continuous at x'

2. Likelihood of treatment is always weakly greater above threshold value x'

Note: Latter two only apply to Fuzzy RDD

3. Effect of treatment and whether observation is treated is independent of x near x'

HE assumption #1

- Assumption that treatment effect is locally continuous at x' is typically not problem
  - It basically just says that there isn't any jump in treatment's effect at x'; i.e. just again assuming observations on either side of x' are comparable
    - **Note:** This might violated if x' was chosen because effect of treatment was thought to be higher for $x>x'$

[E.g. law and/or regulation that creates discontinuity created threshold at that point because effect was known to be biggest there]

HE assumption #2

- Monotonic effect on likelihood of treatment usually not a problem either
  - Just says that having $x > x'$ doesn't make some observations less likely to be treated and others more likely to be treated
  - This is typically the case, but make sure that it makes sense in your setting as well

HE assumption #3

- Basically is saying 'no manipulation'
  - In practice, it means that observations where treatment effect is going to be larger aren't manipulating x to be above the threshold or that likelihood of treatment for individual observation depends on some variable that is correlated with magnitude of treatment effect

HE affects interpretation of estimate

- Key with heterogeneity is that you're only estimating a local average treatment effect
  - Assuming above assumptions hold, estimate only reveals effect of treatment around threshold, and for Fuzzy RDD, it only reveals effect on observations that change treatment status because of discontinuity
  - This limits external validity… **How?**

External validity and RDD *[Part 1]*

- **Answer #1:** Identification relies on observations close to the cutoff threshold
  - Effect of treatment might be different for observations further away from this threshold
  - I.e. don't make broad statements about how the effect would hold for observations further from the threshold value of x

External validity and RDD *[Part 2]*

- **Answer #2:** In fuzzy RDD, treatment is estimated using only "compliers"
  - I.e. we only pick up effect of those where discontinuity is what pushes them into treatment
    - E.g. Suppose you study effect of PhD on wages using GRE score > *x'* with a fuzzy RDD. If discontinuity only matters for students with mediocre GPA, then you only estimate effect of PhD for those students
  - Same as with IV… be careful to not extrapolate too much from the findings

Summary of Today *[Part 1]*

- RDD is yet another way to identify causal effect of some treatment on outcome *y*
  - Makes use of treatment assignment that isn't random, but where process follows some known and arbitrary cutoff rule
  - Very common scenario in practice, and estimator likely to be of increasing use

Summary of Today *[Part 2]*

- Two types of RDD: "sharp" and "fuzzy"
  - Sharp RDD is when treatment is deterministic and <u>only</u> depends on x
  - Fuzzy RDD is when treatment is stochastic and *probability* of treatment has discontinuity at x'
- Formal estimators are similar but different; 'fuzzy' RDD is really just an IV

Summary of Today *[Part 3]*

- Many checks for internal validity; e.g.
  - Graphical analysis with non-parametric plots
  - Check whether observations around cutoff appear to be comparable
- If treatment effect is heterogeneous, estimators interpretation is LATE

In First Half of Next Class

- Miscellaneous Issues
  - Common data problems
  - Industry-adjusting
  - High-dimensional FE
- Related readings… see syllabus

Assign papers for next week…

- Malenko and Shen (working paper 2015)
  - Role of proxy advisory firms
- Keys, et al. (QJE 2010)
  - Securitization and screening of loans
- Almeida, et al. (JFE Forthcoming)
  - Impact of share repurchases

Break Time

- Let's take our 10 minute break
- We'll do presentations when we get back