EEP/IAS 118 - Introductory Applied Econometrics, Section 11

Pierre Biscaye and Jed Silver

November 2021

Agenda

1. ATE, ITT, & TOT
2. Regression discontinuity

Impact Evaluation: Average Treatment Effect

- Randomization helps solve MLR4 $\Rightarrow$ gain confidence in *causal* interpretation of estimates
  - Recall that randomization ensures balance between treatment and control *in expectation*.
  - In any given sample, we might observe some spurious correlations between treatment and other characteristics.
  - Test for balance, potentially control for variables with differences by treatment.
- Estimate impact of treatment T on outcome Y by regressing $Y_i = \alpha + \beta_1 T_i + u_i$
- If take-up of treatment is perfectly aligned with assignment, the coefficient β_1 is referred to as the **Average Treatment Effect (ATE)**.
  - Exactly what it sounds like: the difference in the outcome between the treatment and control groups, on average

Treatment Take-up Compliance

- Sometimes we don't have perfect **compliance** among the treatment and control groups
    - Some observations assigned to the treatment group do not take the treatment and some assigned to control do take it
    - Think of a microcredit program. You can randomlya assign loan *offers*, but you can't force people to take them or prevent people from getting loans elsewhere.
- The choice to "not comply" with the research design is almost certainly correlated with other unobservable charactersitics.
- Comparing those in the treatment group who took the treatment, to those in the control group who did not means we are no longer comparing the randomly assigned groups $\Rightarrow$ we won't get a causal estimate.

	Assigned to Treatment	Assigned to Control
Treated	T Complier	Always-Taker
Not Treated	Never-Taker	C Complier

Impact Evaluation: Intention To Treat

- What to do if we do not have perfect compliance with a treatment?
  - For example, we can randomly assign whether a person receives a college scholarship, but we cannot force them to enroll and accept the scholarship.
  - Maybe some entrepreneurial students in the control group find scholarships from other sources
- We cannot compare those that actually used a scholarship to those that didn't because these groups may be different in a way that is correlated with treatment (OVB!)
- We compare outcomes between the two groups to which people were originally assigned. This is called the **Intention-to-Treat** (ITT) estimator

ITT vs TOT

- The ITT is the causal effect of being *assigned* to treatment.
  - If we have properly randomized, we can always calculate this.
  - If the treatment only affects an outcome through one channel (e.g. if a scholarship only affects wages through education), then the ITT is the effect of education on wages, diluted by incomplete take-up of the education-increasing treatment.
  - It is the sum of a 0 effect among non-compliers (who don't increase education) and an effect among compliers, weighted by the share of compliers.
- What if we want an undiluted estimate?
  - We can calculate the **Treatment on the Treated** or TOT, by dividing the ITT by the share of compliers: those that took the treatment because it was assigned to them.
  - Formally, the TOT is the effect of additional education on wages for "compliers" – those who would not have gotten additional education without the program.

TOT

- How do we know who is a complier?
  - We don't observe this for individuals, but the *share* of compliers is the difference in means of takeup between treatment and control
  - For example, if 24% of students in the control group obtain a scholarship anyway and 54% of treatment students accept their offered scholarship, then the share of compliers is 30%.
  - We consider 24% of treatment students would have been "Always-Takers" since treatment assignment was random. Thus the treatment really only increases education for the other 30% of treatment students that would not have gotten a scholarship otherwise.
  - The TOT is then the ITT divided by .3.
- Interpret the TOT as the effect of the treatment on an outcome for those who were affected by being assigned to treatment.

TOT vs ATE

- When is the TOT the same as the ATE?
  - The treatment only affects outcomes through direct participation (e.g., a scholarship offer only increases wages by increasing schooling) AND
  - The treatment effects for compliers are the same as people would experience on average: $E[y_T|Complier] - E[y_C|Complier] = E[y_T] - E[y_C]$
- Are these likely to hold?
  - If the characteristics of compliers are very different from those of the full population, the TOT may look very different from the ATE.
  - Anytime we don't have perfect compliance with treatment we need to be concerned about heterogeneity in treatment effects.
  - Could the treatment have other effects besides through the targeted mechanism? Spillovers or unintended consequences?

Treatment Impacts: Example

- Think of this in the context of a medical randomized control trial.
- We assign 50 people to control, and 50 people to treatment. The treatment group is given a pill that is supposed to help with energy levels.
- Suppose treatment compliance is perfect. We measure energy levels on a scale from 0-100, and we find that the mean among control individuals is 52 and the mean among treatment individuals is 74.
- What is the ATE?

Treatment Impacts: Example

- Think of this in the context of a medical randomized control trial.
- We assign 50 people to control, and 50 people to treatment. The treatment group is given a pill that is supposed to help with energy levels.
- Suppose treatment compliance is perfect. We measure energy levels on a scale from 0-100, and we find that the mean among control individuals is 52 and the mean among treatment individuals is 74.
- What is the ATE?
- $ATE = 74 - 52 = 22$

Treatment Impacts: Example

- Now suppose that among the 50 people in the treatment group, 40 people comply and take the pill but 10 fail to comply and refuse to take the pill.
- Among the 50 people in the control group, 10 people find a way to get the pill and 40 people comply and don't take pill.
- What is the share of always-takers? Compliers?

Treatment Impacts: Example

- Now suppose that among the 50 people in the treatment group, 40 people comply and take the pill but 10 fail to comply and refuse to take the pill.
- Among the 50 people in the control group, 10 people find a way to get the pill and 40 people comply and don't take pill.
- What is the share of always-takers? Compliers?
  - 20% of control individuals take the pill $\Rightarrow$ 20% of population are always-takers
  - 80% of treatment individuals take the pill - 20% of always-takers $\Rightarrow$ 60% are compliers (for treatment and control)
- We measure energy levels on a scale from 0-100, and we find that the mean among control compliers is 46 and the mean among treatment compliers is 76. What is the ITT?

Treatment Impacts: Example

- Now suppose that among the 50 people in the treatment group, 40 people comply and take the pill but 10 fail to comply and refuse to take the pill.
- Among the 50 people in the control group, 10 people find a way to get the pill and 40 people comply and don't take pill.
- What is the share of always-takers? Compliers?
  - 20% of control individuals take the pill $\Rightarrow$ 20% of population are always-takers
  - 80% of treatment individuals take the pill - 20% of always-takers $\Rightarrow$ 60% are compliers (for treatment and control)
- We measure energy levels on a scale from 0-100, and we find that the mean among control compliers is 46 and the mean among treatment compliers is 76. What is the ITT?
- $ITT = E[Complier] * (E[y_T|Complier] - E[y_C|Complier]) = .6 * (76 - 46) = 18$

Treatment Impacts: Example

- We have
  $ITT = E[Complier] * (E[y_T|Complier] - E[y_C|Complier]) = .6 * (76 - 46) = 18$
- What is the TOT?
- What does the TOT tell us?
- Why might we expect that $TOT \neq ATE$?

Treatment Impacts: Example

- We have $ITT = E[Complier] * (E[y_T|Complier] - E[y_C|Complier]) = .6 * (76 - 46) = 18$
- What is the TOT?
- $TOT = (E[y_T|Complier] - E[y_C|Complier]) = ITT/E[Complier] = 30$
- What does the TOT tell us?
- This is the average impact of the treatment among compliers.
- Why might we expect that $TOT \neq ATE$?
- If the treatment individuals who took the energy pill were those who thought they could benefit most from a boost in their energy levels, the TOT might be greater than the ATE. If being in the treatment group convinced treated inviduals to exercise and find other ways to increase their energy levels besides taking the pills, the TOT would also be larger than the ATE and we would be concerned about omitted variable bias.

Alternate Methods of Impact Evaluation: Intro

Sometimes an RCT can't be conducted because:

- Impossible (i.e., central bank policy)
- Intervention already happened
- Unethical

Without a formal RCT, we have to assume the treatment was **not** randomly assigned:

- Receiving treatment was dependent on (unobserved) characteristics of the subjects
- If we can't control for these variables we will have **OVB**

What can we try in these cases?

- Selection on observables (what we've talked about so far for OVB)
- Regression Discontinuity (RD) - today
- Difference-in-differences (DD) - later
- Other approaches

Regression Discontinuity: Intro

In an RD design, we take advantage of policy quirks where treatment was assigned based on some threshold value of a "running variable." Examples include:

- Age
- Test scores
- Poverty line
- Number of students in a classroom

Basic idea of an RD is to compare the outcome variable for observations just below and just above the threshold.

- The expectation is that people just below and above the threshold are identical in all observable and non-observable characteristics, except for program participation.

Regression Discontinuity: Estimation

$$y_i = \beta_0 + \beta_1 T_i + \beta_2 (RunningVar_i - threshold)$$
$$+\beta_3 T_i \times (RunningVar_i - threshold) + u_i$$

- $RunningVar_i$ is the running variable
- $threshold$ is the threshold value for being treated or not treated
- T_i is a dummy variable indicating whether the observation has a value of the running variable such that it received treatment

Question: What coefficient tells us the effect of the treatment?

Regression Discontinuity: Estimation

$$y_i = \beta_0 + \beta_1 T_i + \beta_2 (RunningVar_i - threshold)$$
$$+\beta_3 T_i \times (RunningVar_i - threshold) + u_i$$

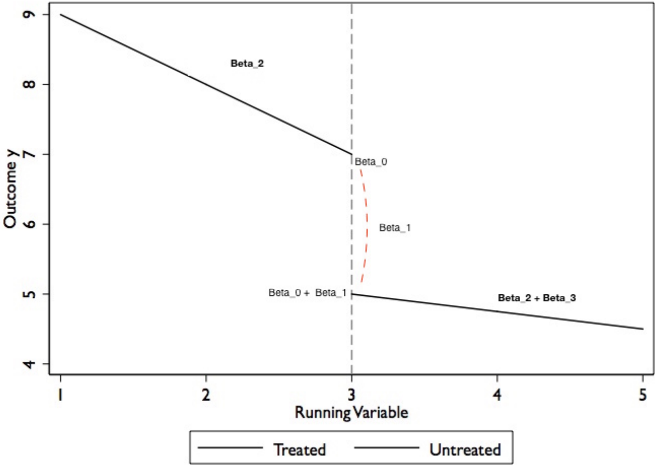
- $RunningVar_i$ is the running variable
- $threshold$ is the threshold value for being treated or not treated
- T_i is a dummy variable indicating whether the observation has a value of the running variable such that it received treatment

$\hat{\beta}_1$ **captures the effect of of the treatment**

We call it a **Local Average Treatment Effect (LATE)** around the threshold

Regression Discontinuity: Estimation

$$y_i = \beta_0 + \beta_1 T_i + \beta_2 (RunningVar_i - threshold)$$
$$+\beta_3 T_i \times (RunningVar_i - threshold) + u_i$$



Regression Discontinuity: Assumptions

Key Assumption:

- Relationship between outcome and running variable would be continuous around the threshold if it were not for the treatment
- In other words we need to assume that the treatment is the only reason why there would be a discontinuity

This assumption might be violated if:

- Participants in the program can manipulate the value of their running variable (e.g. mis-report income to receive subsidy).
- Characteristics of observations just above and just below threshold likely to be different (in ways that might be correlated with the outcome) if there is manipulation of the running variable.

Regression Discontinuity: Test Assumption

- Check for bunching of the running variable around the threshold - a sign of possible manipulation
- Test there are no discontinuities around the running variable threshold for relevant variables **other** than the treatment and the outcome variables
- Look at the averages of observable characteristics of household just above and below the threshold and make sure they're similar (kind of like in RCT)

$$x_i = \beta_0 + \beta_1 T_i + \beta_2 (RunningVar_i - threshold)$$
$$+ \beta_3 T_i \times (RunningVar_i - threshold) + u_i$$

Here we want to find a coefficient of zero for our estimated $\hat{\beta}_1$, to encourage us that observations just above and below the threshold look similar on observables.

Regression Discontinuity: Example

If a youth who is less than 18 years old commits an offense, the case is sent to the more lenient juvenile courts. However, if a youth commits an offense after his/her 18th birthday, the case is sent to the much harsher adult criminal court. You have a cross-sectional data set of youths of ages 16-20 in Florida in 2005. This data set includes the birthday, gender, family income, and whether or not the youth had been arrested for committing an offense in the past year.

1. Write a regression to estimate the causal effect of harsher punishments on the probability of committing a crime. Which coefficient will give you the estimated causal effect?
2. What key assumption do you need to make for your regression to estimate the causal effect of harsher punishments on the probability of committing a crime?
3. What test can we conduct in support of our assumption?

Regression Discontinuity: Example

1. Write a regression to estimate the causal effect of harsher punishments on the probability of committing a crime (assume likelihood of being arrested for crime doesn't vary with age). Which coefficient will give you the estimated causal effect?

Regression Discontinuity: Example

1. Write a regression to estimate the causal effect of harsher punishments on the probability of committing a crime (assume likelihood of being arrested for crime doesn't vary with age). Which coefficient will give you the estimated causal effect?

$$arrest_i = \beta_0 + \beta_1 Over18_i + \beta_2(age_i - 6574) + \beta_3 Over18_i \times (age_i - 6574) + u_i$$

Where $arrest_i$ is a dummy variable equal to 1 if youth i was arrested for a criminal offense, age_i is age in days, and $Over18_i$ is a dummy variable equal to 1 if youth i is 18 or older. Note that 6574 is the number of days in 18 years. We could equivalently have set this up using age in months or weeks as the running variable rather than age in days. Age in years would likely be too broad of a range. The coefficient $\hat{\beta}_1$ will give us the LATE estimate of the effect of harsher punishments on the probability of committing a criminal offense.

Regression Discontinuity: Example

2. What key assumption do you need to make for your regression to estimate the causal effect of harsher punishments on the probability of committing a crime?

Regression Discontinuity: Example

2. What key assumption do you need to make for your regression to estimate the causal effect of harsher punishments on the probability of committing a crime?

We have to assume that without the "treatment" of harsher punishments after age 18, the probability of committing an offense is a continuous function of age. That is, we have to assume that there are no other discontinuities in observable or unobservable characteristics around age 18, making the group just under the threshold on average similar to the group just above the threshold—except for the treatment. This makes those about to turn 18 a suitable counterfactual for those who just turned 18.

Regression Discontinuity: Example

3. What test can we conduct in support of our assumption?

Regression Discontinuity: Example

3. What test can we conduct in support of our assumption?

A test of the validity of the approach is that there are no discontinuities around the threshold for relevant variables other than the treatment and the outcome variable. With the data we have, we could run the same regression as in part 1) for family income.

Sharp vs Fuzzy RD

- Compliance with treatment assignments under a threshold eligibility rule may not be perfect.
  - In the above example, compliance would be perfect if once a youth turns 18 they are sent to the adult instead of the juvenile court without exceptions.
- We call situations of perfect compliance “Sharp” RD. We can use our RD approach to estimate the LATE.
- Imperfect compliance with a threshold-based treatment is sometimes called a “Fuzzy” RD. We can estimate a local ITT and TOT.
  - In Professor Magruder’s study of irrigation canals on hillsides in Rwanda, about 5% of plots above the canal used irrigation, while only about 23% of plots under the canal used it.
- We can check whether we do in fact have a discontinuity around the threshold for treatment by running the RD regression with treatment take-up (in this example, use of irrigation) as the outcome variable.

What do We Need to Worry About With RD?

1. Is the effect really there? Should be demonstrable with a figure: look for a discontinuity around the threshold.
2. Does the cutoff matter? Run the RD specification with treatment takeup as the outcome variable.
3. Can the running variable be manipulated? Look for "bunching" around the threshold.
4. Does anything else change sharply at the cutoff?
5. Who are we identifying the effect for? We know that we identify a local ATE (LATE). This might not be the same as the effect globally. For example, estimates of the impact of the harshness of punishment on crime among youths around age 18 might not apply for older adults.