Recitation 10: Local average treatment effects

Seung-hun Lee

Columbia University
Undergraduate Introduction to Econometrics Recitation

December 8th, 2022

Local average treatment effect

Setting up LATE: Selection on unobservables

- What if random assignment and the unconfoundedness assumptions fail?
- Solution: IV approach which uses exogenous variation that dictates the assignment into the treatment
- Z_i: Binary variable equal to 1 if individual i qualifies for a randomized eligibility status to be treated
- X_i: (observed) actual assignment into the treatment (participation).
- X_i is now a function of Z_i and this is also counterfactual; For a given individual i, you can only observe one of $X_i(1) = X_i(Z_i = 1)$ or $X_i(0) = X_i(Z_i = 1)$.
- Based on this idea, we can write

$$\begin{aligned} X_i &= X_i(1)Z_i + X_i(0)(1 - Z_i) \\ &= X_i(0) + Z_i(X_i(1) - X_i(0)) \end{aligned}$$

- Similarly: $1 - X_i = 1 - X_i(0) - Z_i(X_i(1) - X_i(0))$

Rewriting Y_i in a potential outcome framework

- Y_i depends on two parameters: X_i (whether i participated in the treatment) and Z_i (whether i was eligible)
- We can write the potential outcome framework as follows (But note that $Y_i(1) = Y_i(X_i = 1)$ and that $X_i(1) = X_i(Z_i = 1)$)

$$
\begin{aligned}
Y_i &= Y_i(1)X_i + Y_i(0)(1 - X_i) \\
&= Y_i(1)[X_i(0) + Z_i(X_i(1) - X_i(0))] + Y_i(0)[1 - X_i(0) - Z_i(X_i(1) - X_i(0))] \\
&= Y_i(1)X_i(0) + Z_i Y_i(1)(X_i(1) - X_i(0)) + Y_i(0) - Y_i(0)X_i(0) - Z_i Y_i(0)(X_i(1) - X_i(0)) \\
&= Y_i(0) + X_i(0)(Y_i(1) - Y_i(0)) + Z_i(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))
\end{aligned}
$$

Local average treatment effect

- Local average treatment effect (LATE) is defined as

$$LATE = E[Y_i(1) - Y_i(0) | X_i(1) - X_i(0) = 1]$$

- Treatment effect on 'compliers'
- We will need to set assumptions and categorize the individuals into 4 groups - those who always participate (always takers, AT), those who never participate (never takers, NT), those who participate only if they are eligible (compliers, CP), and those who participate only if they are ineligible (defiers)

LATE assumptions that are similar to IV

- LATE1: Z_i **is independent of** $(Y_i(1), Y_i(0), X_i(1), X_i(0))$
  $\rightarrow$ Broken down into
  1. $Z_i \perp (\underbrace{Y_i(1,1), Y_i(0,1)}_{Y_i(1)}, \underbrace{Y_i(1,0), Y_i(0,0)}_{Y_i(0)}, X_i(1), X_i(0))$, where we have $Y_i(z,x)$ now
  2. $Y_i(z,x) = Y_i(z',x)$ For all z, z', x

  $\rightarrow$ Randomization of Z_i guarantees the first subcondition.
  $\rightarrow$ The second subcondition (and LATE1 in general) is an exogeneity (or exclusion) condition in that Z_i only affects outcome through treatment status X_i (so Z_i has no direct impact on outcome variable)
- LATE2: **Relevance,** $\Pr(X_1 = 1|Z_i = 1) \neq \Pr(X_i = 1|Z_i = 0)$
  $\rightarrow$ Z_i dictates the likelihood of participating in the treatment (relevance)

LATE assumptions that are uniquie

- LATE3: **No defiers,** $X_i(1) \geq X_i(0)$
  $\rightarrow$ Rules out the case that i participates if ineligible but does not participate if eligible.
  $\rightarrow$ Under this assumption, we have the three types of participants
  1. Always takers: $X_i(1) = 1, X_i(0) = 1$
  2. Never takers: $X_i(1) = 0, X_i(0) = 0$
  3. Compliers: $X_i(1) = 1, X_i(0) = 0$

  $\rightarrow$ So $X_i(z)$ is only allowed to increase with z_i in this setup.
  4. Defiers: $X_i(1) = 0, X_i(0) = 1$

  $\rightarrow$ Here $X_i(z)$ decreases with Z_i. By imposing LATE3, we rule out this movement (so this is also called no two-way movement condition)

Categorizing observations

- So if we categorize the observations based on the values of $X_i(z)$ for each z, we can write

	$X_i(0) = 0$	$X_i(0) = 1$
$X_i(1) = 0$	Never taker	Defier
$X_i(1) = 1$	Complier	Always taker

- But $X_i(z)$ also suffers from missing data problem since we can only observe only one of $X_i(1)$ or $X_i(0)$ for each i.
- If we write the above tables for the observables X_i and Z_i

	$Z_i = 0$	$Z_i = 1$
$X_i = 0$	Never taker *or* Complier	Never taker *or* Defier
$X_i = 1$	Always taker *or* Defier	Always taker *or* Complier

Categorizing observations: Finding the shares

- By LATE3 assumption, we can rule out defiers and identify the share of always takers(AT), never takers(NT), and compliers(C) as follows

$$\begin{aligned}\pi_{AT} + \pi_{NT} + \pi_C &= 1 \\ \pi_{AT} + \pi_C &= E[X_i|Z_i = 1] \\ \pi_{NT} + \pi_C &= 1 - E[X_i|Z_i = 0]\end{aligned}$$

- So we get

$$1 + \pi_C = 1 + E[X_i|Z_i = 1] - E[X_i|Z_i = 0] \implies \pi_C = E[X_i|Z_i = 1] - E[X_i|Z_i = 0]$$

and by replacing π_C, we can back out π_{AT} and π_{NT}

$$\begin{aligned}\pi_{AT} &= E[X_i|Z_i = 0] \\ \pi_{NT} &= 1 - E[X_i|Z_i = 1]\end{aligned}$$

Identifying LATE: Rewrite our target equation

- To back out the LATE, we need to rework the definition as follows

$$
\begin{aligned}
E[Y_i|Z_i] &= E[Y_i(0)|Z_i] + E[X_i(0)(Y_i(1) - Y_i(0))|Z_i] + Z_i E[(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))|Z_i] \\
&= E[Y_i(0)] + E[X_i(0)(Y_i(1) - Y_i(0))] + Z_i E[(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))] \; (\because LATE1)
\end{aligned}
$$

- This means

$$
\begin{aligned}
E[Y_i|Z_i = 0] &= E[Y_i(0)] + E[X_i(0)(Y_i(1) - Y_i(0))] \\
E[Y_i|Z_i = 1] &= E[Y_i(0)] + E[X_i(0)(Y_i(1) - Y_i(0))] + E[(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))]
\end{aligned}
$$

- Thus, $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = E[(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))]$

Identifying LATE: Break down $E[(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))]$

- This is equal to

$$\begin{aligned} E[(Y_i(1) - Y_i(0))(X_i(1) - X_i(0))] &= 1 \times E[Y_i(1) - Y_i(0) | X_i(1) - X_i(0) = 1] \Pr(X_i(1) - X_i(0) = 1) \\ &+ 0 \times E[Y_i(1) - Y_i(0) | X_i(1) - X_i(0) = 0] \Pr(X_i(1) - X_i(0) = 0) \\ &= E[Y_i(1) - Y_i(0) | X_i(1) - X_i(0) = 1] \Pr(X_i(1) - X_i(0) = 1) \end{aligned}$$

- We now have treatment effect for complier groups only
- Use counterfactuals for X_i to get

$$\begin{aligned} & E[X_i | Z_i] = Z_i E[X_i(1) | Z_i] + (1 - Z_i) E[X_i(0) | Z_i] \\ \implies & E[X_i | Z_i = 1] = E[X_i(1) | Z_i = 1] = E[X_i(1)] \\ \implies & E[X_i | Z_i = 0] = E[X_i(0) | Z_i = 0] = E[X_i(0)] \end{aligned}$$

- We can get this probability using observables!

$$E[X_i | Z_i = 1] - E[X_i | Z_i = 0] = E[X_i(1)] - E[X_i(0)] = E[X_i(1) - X_i(0)] = \Pr(X_i(1) - X_i(0) = 1)$$

Identifying LATE: Combine all steps

- If we combine all the steps, we get

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = E[Y_i(1) - Y_i(0)|X_i(1) - X_i(0) = 1](E[X_i|Z_i = 1] - E[X_i|Z_i = 0])$$
$$\implies \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[X_i|Z_i = 1] - E[X_i|Z_i = 0]} = E[Y_i(1) - Y_i(0)|X_i(1) - X_i(0) = 1] = LATE$$

- Denominator is only defined if LATE2 assumption is satisfied. (Replace $E[X_i|\cdot]$ with $\Pr(X_i = 1|\cdot)$ to get this, which works since X_i is binary)

Estimating LATE

- We can find a sample analogue (separate sample for $Z = 1$ and $Z = 0$)
- We can also get this as a Wald estimate - ratio of two separate reduced form regression coefficients. Start with

$$Y_i = \beta_0 + \beta_1 Z_i + e_i$$
$$X_i = \gamma_0 + \gamma_1 Z_i + u_i$$

Then, assuming that Z_i is independent of both e_i and u_i, we get

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = \beta_1, \quad E[X_i|Z_i = 1] - E[X_i|Z_i = 0] = \gamma_1$$

Therefore, $LATE = \frac{\beta_1}{\gamma_1}$, whose estimates can be obtained by $\frac{\hat{\beta}_1}{\hat{\gamma}_1}$

Notes on applying LATE

- This can be applied for RD estimation - fuzzy or sharp.
- $Z_i = 1$ if a running variable crosses the threshold so that i is eligible
- X_i is a binary variable for participation.
- If RD is sharp, $\gamma_1 = 1$. If RD is fuzzy, $\gamma_1 < 1$.
- β_1 would be the treatment effect difference between those who passes the threshold vs. those who do not.
- The treatment effect in RD can be obtained by dividing between β_1 and γ_1.
- In general, LATE identifies a treatment effect for an unidentifiable segment of the population. IRL, difficult to identify compliers!
- Furthermore, the definition of compliers change with the definition of Z_i. So interpretation of LATE is tricky.