Public Policy Evaluation
Lecture 1: Randomised Control Trial

Yao Thibaut Kpegli

Licence 3 in Economics, ENS Paris-Saclay
2023–2024

March 2024

Outline

Set up

Targeted population

- Job seekers

Treatment

- Enhanced support ($T_i = 1$) rather than standard support ($T_i = 0$)

Potential outcomes

- Employment status with enhanced support: Y_{1i}
- Employment status without enhanced support: Y_{0i}

Treatment Effect

$$\Delta_i = Y_{1i} - Y_{0i}$$

Random draw

Independence — law of large numbers

$$E[X|T=1]=E[X|T=0]=E[X] \quad \text{for all} \quad X$$

Examples:

- Observables:
  $E[\textbf{household}|T=1]=E[\textbf{household}|T=0]=E[\textbf{household}]$
  $E[\textbf{education}|T=1]=E[\textbf{education}|T=0]=E[\textbf{education}]$
  $E[\textbf{age}|T=1]=E[\textbf{age}|T=0]=E[\textbf{age}]$

  [...]
- Unobservables:

  $E[\textbf{motivation}|T=1]=E[\textbf{motivation}|T=0]=E[\textbf{motivation}]$

  $E[\textbf{network}|T=1]=E[\textbf{network}|T=0]=E[\textbf{network}]$
  - including potentials outcomes:
    $E[Y_0|T=1]=E[Y_0|T=0]=E[Y_0]$
    $E[Y_1|T=1]=E[Y_{1i}|T=0]=E[Y_1]$

NE = ATT = ATE

- Naive estimator

$$\begin{aligned} \text{NE} &= E[Y|T=1] - E[Y|T=0] \\ &= E[Y_1|T=1] - E[Y_0|T=0] \end{aligned}$$

- As $E[Y_0|T=1] = E[Y_0|T=0]$, then:

$$\begin{aligned} \text{NE} &= E[Y_1|T=1] - E[Y_0|T=1] \\ &= \text{ATT} \end{aligned}$$

- As $E[Y_0|T=1] = E[Y_0]$ and $E[Y_1|T=1] = E[Y_1]$, then:

$$\begin{aligned} \text{NE} &= E[Y_1] - E[Y_0] \\ &= \text{ATE} \end{aligned}$$

Outline

Regression equivalent

- Linear model

$$\boldsymbol{Y}_i = a + b\boldsymbol{T}_i + e_i$$

$$b_{OLS} = E[Y_i|T_i = 1] - E[Y_i|T_i = 0]$$

- OLS is unbiased:
  - out of random draw $E[\epsilon_i|T_i] = 0$
- Often, we add controls:

$$\boldsymbol{Y}_i = a + b\boldsymbol{T}_i + \boldsymbol{X}_i c + u_i$$

- But why to add controls ?
  - randomization may fail (in small samples)
  - with interactions, it allows to estimate heterogeneous treatment effects
  - additional controls decreases the variance of b_{OLS} if they have explanatory power over $\boldsymbol{Y}_i$ (see next slide)

Controls reduce variance

- Consider the two models:

$$
\begin{array}{ccccccccccc}
\boldsymbol{Y_i} & = & \text{a} & + & \text{b}\,\boldsymbol{T}_i & + & & & e_i & & (1) \\
\boldsymbol{Y_i} & = & \text{a} & + & \text{b}\,\boldsymbol{T}_i & + & \boldsymbol{X}_i c & + & u_i & & (2)
\end{array}
$$

- If treatment assignment is truly random, then conditioning on $_{X\,i}$ does not affect point estimate: b_{OLS}
- But controls reduce variance:
  - $V(b_{OLS}) = V(\epsilon_i)(\boldsymbol{TT'})^{-1}$
  - $V(\epsilon_i) = V(\boldsymbol{X}_i c + u_i) = V(\boldsymbol{X}_i c) + V(u_i) > V(u_i)$

RCT in pratice

1. Define the population of interest
2. Randomly divide the population of interest into treatment and control groups
3. Collect baseline data (before the treatment is implemented)
4. Check the randomization (balance check)
5. Process to the treatment implementation
6. Collect endline data
7. Run OLS regressions

Outline

Application

The Demand and Impact of Learning HIV Status - ecampus

Thank you for your attention !