Topics in Applied Econometrics

Average treatment effects

Attila Gyetvai

Duke Economics
Summer 2020

How to tackle an empirical project

1. What causal effects are we interested in?
2. What ideal experiment would capture this effect?
3. What is our identification strategy?
4. What is our mode of statistical inference?

Average treatment effect

Average treatment effect (ATE)

$$ATE = \mathbb{E}(Y_{1i} - Y_{0i}) = \mathbb{E}(Y_{1i}) - \mathbb{E}(Y_{0i})$$

- Recall the fundamental problem: we can only observe the actual outcomes
- How can we measure the ATE from these observations?

Average treatment effect

- What is the average potential outcome in terms of observed outcomes?
- Using the law of total expectations:

$$\mathbb{E}(Y_{1i}) = \Pr(D_i = 1) \cdot \mathbb{E}(Y_{1i} \mid D_i = 1) + \Pr(D_i = 0) \cdot \mathbb{E}(Y_{1i} \mid D_i = 0)$$

- We can easily get the probabilities from the data
- Also, we can observe the potential outcome of treatment for the treated:

$$\mathbb{E}(Y_{1i} \mid D_i = 1) = \mathbb{E}(Y_i \mid D_i = 1)$$

- So what we need is the expected outcome of the treatment for the untreated
- Similarly, the expected outcome without treatment for the treated

Average treatment effect

- $\mathbb{E}(Y_{1i} \mid D_i = 0)$: expected outcome of the treatment for the untreated
- $\mathbb{E}(Y_{0i} \mid D_i = 1)$: expected outcome without treatment for the treated
- In a good experimental design, we can replace the counterfactuals with measurements
- *Example:* $\mathbb{E}(Y_{0i} \mid D_i = 1) = \mathbb{E}(Y_{0i} \mid D_i = 0) = \mathbb{E}(Y_i \mid D_i = 0)$
  - When can we reasonably assume that the potential outcome without treatment is the same for treated and untreated groups?
- However, ATE is hard to identify in most nonexperimental settings

Average treatment effect on the treated

- Maintaining two sets of assumptions is trickier than only one
- What if we can only measure $\mathbb{E}(Y_{0i} \mid D_i = 1)$ but not $\mathbb{E}(Y_{1i} \mid D_i = 0)$?

Average treatment effect on the treated (ATT)

$$ATT = \mathbb{E}(Y_{1i} - Y_{0i} \mid D_i = 1) = \mathbb{E}(Y_i \mid D_i = 1) - \mathbb{E}(Y_{0i} \mid D_i = 1)$$

- What is the effect of the treatment on treated units?
- ATT may be more important than ATE in program evaluation
  - *Example:* job training programs
  - We may not care about the effect of the program on nonparticipants

Average treatment effect on the untreated

- Analogously, what if we can only measure $\mathbb{E}(Y_{1i} \mid D_i = 0)$?

Average treatment effect on the untreated (ATU)

$$ATU = \mathbb{E}(Y_{1i} - Y_{0i} \mid D_i = 0) = \mathbb{E}(Y_{1i} \mid D_i = 0) - \mathbb{E}(Y_i \mid D_i = 0)$$

- What is the effect of the treatment on untreated units?
- ATU may be more important than ATE in nonexperimental settings
  - *Example:* college wage premium
  - How much more would people without a college degree earn if they had one?

Heterogeneity in treatment effects

- So far we have talked about ATE, the *average* of the treatment effects across units
- However, sometimes we would like to know the effect in certain subgroups
  - *Example:* pilot for no-strings-attached cash transfers
  - Effects are likely larger for poorer households $\implies$ eligibility thresholds
- We can simply measure ATE in a group with characteristics x

Conditional average treatment effect (ATE(x))

$$ATE(x) = \mathbb{E}(Y_{1i} - Y_{0i} \,|\, X_i = x)$$

- ATE(x) describes the whole distribution of treatment effects