Potential Outcomes Framework
Selection Bias

Md Moshi Ul Alam

March 24, 2025
Clark University

Topics that you have learnt in Econ 351

- Bivariate and multivariate linear regression models
- OLS estimator and assumptions from which it derives its Unbiasedness property
- Variance of the OLS estimator and assumptions on the variances of error terms: homoskedasticity and heteroskedasticity
- Gauss-Markov Theorem: OLS is BLUE under certain assumptions
- Interpretations of estimates of regressions: partial effects, ceteris paribus, partialling out, being careful of the dimensions of the vectors.
- Regressions with dummy and categorical variables, polynomials, interactions
- From sample to population: basics of OLS Asymptotics

Causal Inference

Overview

- Course on **causal inference**
- **Design based methods** in applied micro-econometric research
- **Math meets Economic Intuition:** Understanding the assumptions of different econometric frameworks and contexts where they can be credible applied to answer a causal question
- Very dynamic and evolving field!

Treatment Effects

Ask causal questions in terms of **treatment effects** within a **potential outcome framework**

Examples of various treatments

Treatment Effects

Ask causal questions in terms of **treatment effects** within a **potential outcome framework**

Examples of various treatments

- Going to preschool
- Working form home
- Providing inheritance rights for females in developing countries
- Being in a job training program
- Increasing the minimum wage
- Introducing gender quotas in boardrooms of firms

Treatment Effects

Ask causal questions in terms of **treatment effects** within a **potential outcome framework**

Examples of various treatments

- Going to preschool
- Working form home
- Providing inheritance rights for females in developing countries
- Being in a job training program
- Increasing the minimum wage
- Introducing gender quotas in boardrooms of firms

Object of interest: *What is the casual impact of the treatment (D) on an outcome variable of interest (Y)?*

Potential outcomes framework

Given a treatment D_i, some units in the population get treated $D_i = 1$ and some do not $D_i = 0$.

For each unit i in the population, there are two potential outcomes:

- $Y_i(1)$: i's potential outcome in the treated state $D_i = 1$
- $Y_i(0)$: i's potential outcome in the untreated state $D_i = 0$

Standard notation:

- i's observed outcome = Y_i
- i's potential outcomes = $(Y_i(0), Y_i(1))$

Parameters of interest

The common effect world:
- Treatment effects that are the same for all units
- In notation: $Y_i(1) - Y_i(0) = \delta_i = \delta$ for all i

Parameters of interest

The common effect world:
- Treatment effects that are the same for all units
- In notation: $Y_i(1) - Y_i(0) = \delta_i = \delta$ for all i

The heterogeneous effects world:
- The three most common parameters of interest are:

  - ATE $= E\left[Y_i(1) - Y_i(0)\right] =$ avg. treatment effect in the population
  - ATT $= E\left[Y_i(1) - Y_i(0) \mid D_i = 1\right] =$ avg. treatment effect on the treated
  - ATNT $= E\left[Y_i(1) - Y_i(0) \mid D_i = 0\right] =$ avg. treat. effect on the untreated

The problem of Causal Inference

- Each unit i either gets treated $D_i = 1$ or untreated $D_i = 0$
- So for each unit i we either observe $Y_i(0)$ or $Y_i(1)$
- **We do not observe the counterfactual**!

[Show with ATT & ATNT]

The problem of Causal Inference

- Each unit i either gets treated $D_i = 1$ or untreated $D_i = 0$
- So for each unit i we either observe $Y_i(0)$ or $Y_i(1)$
  - **We do not observe the counterfactual**!

[Show with ATT & ATNT]

This course will entirely focus on how we can
credibly estimate the unobseved counterfactual : IDENTIFICATION STRATEGIES

Statistical Inference vs Causal Inference

Statistical inference is about learning about the properties of a population using an estimator when we only observe a sample
- Gets easier as sample size $\uparrow$

Statistical Inference vs Causal Inference

Statistical inference is about learning about the properties of a population using an estimator when we only observe a sample
- Gets easier as sample size $\uparrow$

Causal inference is about learning about $Y_i(1) - Y_i(0)$ from a sample of observations of only $(., Y_i(1))$ or $(Y_i(0), .)$

- A larger sample does not solve the fundamental missing-counterfactual problem
- We will study a number of identification strategies that rely on different assumptions to estimate this missing counterfactual.
- Assumptions may be true or not true in any given context.
- Often settings where something like an experiment or a natural experiment occurred

Why have an entire course on this?

Why have an entire course on this?

Well there could be multiple courses on this and we will still not be done.

The credibility revolution has changed the face of economic research in the last two decades

- Units who get treated could be unobservably very different from those who do not get treated. [Give examples]
- May end comparing apples to oranges which will bias the estimates
- Need to have an "apples-to-apples" comparison!

Let's work with an example

Do hospitals make people healthier?

Group	Sample Size	Mean health status	Std. Error
Hospital	7774	2.79	0.014
No Hospital	90049	2.07	0.003

Figure: Source: NHIS 2005 and MHE. Measure of health: 1 to excellent and a 5 to poor

Do hospitals make people healthier?

Group	Sample Size	Mean health status	Std. Error
Hospital	7774	2.79	0.014
No Hospital	90049	2.07	0.003

Figure: Source: NHIS 2005 and MHE. Measure of health: 1 to excellent and a 5 to poor

$$E(Y_i|D_i = 1) - E(Y_i|D_i = 0) = 0.72 \;\; with \;\; t - stat = 58.9 \tag{1}$$

Selection bias

Identification strategies

- Randomized Control Trials
- Regressions
- Instrumental Variables
- Fixed effects
- Difference-in-differences
- Regression Discontinuity Design

Focus on contexts where SUTVA holds

- The "Stable Unit Treatment Value Assumption"
  - SUTVA: Treated and untreated outcomes for each unit do not depend on which units get treated or how many units get treated
- SUTVA rules out all equilibrium and spillover effects
- Examples:
  - Displacement effects in job search assistance programs
  - Increased policing in one location displaces crime in other locations