

# Notes on Empirical Methods in Business

## Lecture 0: Introduction

Hang XU

hxuch@connect.ust.hk

June 19, 2024

# Introduction

- During the first year of my PhD, I have taken a course on Empirical Methods in Business: Modeling and Estimation.
- This course is designed to provide students with a comprehensive understanding of the most commonly used empirical methods in business research.
- The main topics of the course can be seen in section 3. I mainly used handwritten notes when I was taking the lectures, which is hard to formalize.
- As a PhD researcher specializing in empirical studies, it is necessary to have a clear understanding of common empirical methods.
- Therefore, I review the course content and summarize it in a more formal way to help others who are interested in empirical methods in business research.
- Worth to mention that, all the faults in the notes are mine, and I will try my best to make it accurate and clear. If you find any mistakes or have any suggestions, please feel free to contact me.

# Research Classification

Traditional classifications in empirical research:

- Controlled data: Lab, AFE, FFE
  - Field experiment: AFE (artefactual field experiment), FFE (framed field experiment)
  - Lab experiment
- Naturally occurring / observational data
  - Natural experiment: NE, NFE (natural field experiment<sup>1</sup>)
  - Market data: IV, PSM, STR (Structural modeling)

---

<sup>1</sup>Field experiment that happens naturally, and people do not realize the experiment they are in.

# Causal Treatment Effects

Identify the causal treatment effects has been the main focus of empirical research in business.

- The golden rule for identification: **Randomization of treatment status.**
  - $y_i = \alpha + \gamma T_i + \epsilon_i$ , where  $T_i$  is the treatment status.
  - Randomization makes  $E(\epsilon_i | T_i = 0) = E(\epsilon_i | T_i = 1)$
  - Thus,  $\gamma$  can identify the causal effect of treatment.
- No endogenous issues:
  - People cannot quit or switch the groups;
  - No spillover effect:
    - ▷ Across sides: two-sided platform, sellers and buyers switch – no reverse causality;
    - ▷ Across groups in one side: individuals in each group do not aware they are treated or controlled. i.e., no information spillovers.

# Identifying Causal Effects with Market Data

Market data cannot be randomized, so we need to use other methods to identify the causal effect of treatment:

- Statistical methods: Approximating the experiments: e.g., DiD
- Econometric methods:
  - Control methods
  - Instrument variables
  - Structural models

# Research Considerations

Many researchers focus more on fancy methods, ignoring the data and assumptions, making the story less reliable. Questions need to think before digging into the research:

- What is the data? Can it help identify the causal effects?
- What are the identification assumptions? Are they reasonable?

# Key Components in Empirical Research

4 key components in empirical research:

- **Research Questions**

- Why are your research questions important?
- What is the use for business/consumers/regulators?
- What is your contribution to the literature?

- **Data**

- Can your data help address your research questions?

- **Model**

- What is  $Y$ ? What are your  $X$ 's?
- What is the relationship between  $Y$  and  $X$ 's?
- What is the data generating process (DGP)?
- How does your model address your research questions?

- **Estimation**

- OLS / NLS? MLE? Method of moments? Other more advanced methods?
- What is the identification of model parameters?

# Main Methods Covered

- Issues in Regression
  - Multicollinearity
  - Heteroskedasticity
  - Causality
- Causal Inference and Treatment Effect Models
  - Instrument Variables
  - Panel Data with Fixed Effects
  - Treatment Effects
    - ▷ Matching
    - ▷ Propensity Score Matching
    - ▷ Inverse Probability Weighting
    - ▷ Difference-in-Differences
    - ▷ Synthetic Control
    - ▷ Synthetic Difference-in-Differences
    - ▷ Regression Discontinuity
- Choice Model
  - Binary Choice Model
  - Multinomial Choice - Ordered
  - Multinomial Choice - Non-Ordered