Introduction and Experiments

Ethan Kaplan

1 Identification of Causation

- Godel's Theorem : Any logical system capable of whole number addition and multiplication will have question which can be asked and can not be answered true or false. Moreover, it is impossible to determine which questions can not be proven true or false.
  - Empirical Analogue: Empirical questions can be asked which may likely have no good answer.

- Ability to Understand the Past and Predict the Future
  - Making Progress on Identification to Important Questions: Acemoglu, Johnson and Robinson
  - "Clean Identification": Lee, Butler, and Moretti
  - Establishing Empirical Facts: Shleifer et al., Persson and Tabellini, Barro

- Hume's Problem
  - Need for the Ceterus Paribus Assumption for Internal Identification
  - Need for the Ceterus Paribus Assumption for External Validity

2 Empirical Approaches

- Identification of Causality:Experiments
  - Field
  - Laboratory
- Natural Experiments
  - Randomization
  - Conditional Randomization
- Identification of Parameters: Structural Estimation
- Stylized Fact Provision: Partial Sample Correlation

3 Course Outline

- Course Requirements
- Office Hours
- Syllabus

4 Rubin's Potential Outcome Model (1974)

- Treatment T is binary: $\{0, 1\}$.
- Outcome for Treatmeant is given by random variable Y_i^T and outcome for control: Y_i^C
- Impact of Treatment given by Potential Outcomes: $TY_i^T + (1 - T)\, Y_i^C(C)$
- Treatment Effect Without Randomization:

$$E\left[Y_i^T - Y_i^C\right] = E\left(Y_i^T|T\right) - E\left(Y_i^C|C\right) =$$

$$\left\{E\left(Y_i^T|T\right) - E\left(Y_i^C|T\right)\right\} + \left\{E\left(Y_i^C|T\right) - E\left(Y_i^C|C\right)\right\}$$

$\{$Treatment Effect (on the treated)$\}$ + $\{$Selection Bias$\}$

- Treatment Effect With Perfect Randomization (S is the Selection Criterion for the experiment):

$$E\left[Y_i^T - Y_i^C\right] = E\left(Y_i^T|S\right) - E\left(Y_i^C|S\right)$$

5 Experiments: An Introduction

- Experiments: Two Types
  - Laboratory: i.e. Iyengar, Going Negative, showing face pictures of politicians
  - Field: i.e. Bjorkman and Svensson: Randomly providing and making public report cards about health clinic performance

- Benefits
  - Ability to control for selection
  - Ability to design an experiment to ask exactly the question you wanted to answer
  - Ability to commit to a research design ahead of time and reduce degrees of freedom for manipulation

    * Subgroups (variables and/or strata) as Degrees of Potential Manipulation

    * Tradeoff Between Ex-Post Learning and Ex-Ante

- Costs

  - External Validity

    * Moral Constraints

    * Legal Constraints

      · Attrition

      · Substitution Bias (Heckman and Smith, 1995)

      · Randomization Bias and Selection in Experiment Participation (Heckman and Smith, 1995)

    * Effects of the Experiment Independent of the Treatment

        - Hawthorne Effects: Changes in behavior among the treatment group
        - John Henry Effects: Changes in behavior among the control group
    - Internal Validity
        * Attrition
        * Externalities: SUTVA (Stable Unit Treatment Value Assumption) "General Equilibrium": $E(Y_i^k|T)$ is independent of T_j.

$$\begin{aligned} E\left(Y_i^T|T\right) - E\left(Y_i^C|C\right) &= E\left(Y_i^T - Y_i^C|T\right) \\ &= E\left(Y_i^T - Y_i^C\right) \end{aligned}$$

        * Contamination
            - Treatment Doesn't Take Up
            - Control Takes Up
    - Small Sample Sizes:

    - Power
    - Identifying Heterogeneous Effects
    - Identifying Population Average Treatment Effects
- Monetary Costs of Implementation

6 Power Calculations

$$Y_i = \alpha + \beta T_i + \epsilon_i$$

OLS Estimator is:

$$\begin{aligned}
& \min_{\alpha,\beta} \sum_{i=1}^{I} \left(Y_i - \alpha - \beta T_i\right)^2 \\
\alpha \quad : \quad & -2 \sum_{i=1}^{I} \left(Y_i - \alpha - \beta T_i\right) = 0 \\
\beta \quad : \quad & -2 \sum_{i=1}^{I} \left(Y_i - \alpha - \beta T_i\right) T_i = 0
\end{aligned}$$

This implies the following estimators for α and β :

$$\begin{aligned}
\alpha &= \frac{\sum_{i=1}^{I} \left(Y_i - \beta T_i\right)}{I} = \bar{Y} - \beta \bar{T} \\
\hat{\beta} &= \frac{\sum_{i=1}^{I} \left(Y_i - \bar{Y}\right)\left(T_i - \bar{T}\right)}{\sum_{i=1}^{I} \left(T_i - \bar{T}\right)\left(T_i - \bar{T}\right)}
\end{aligned}$$

We can also compute the Standard Error by taking $V\left(\hat{\beta}\right)$ (and remembering that $P(T=1)=P$)

$$\frac{\sigma_{\epsilon}^{2}}{P\left(1-P\right)N}$$

- Size and Power
  - Size of a Test: Probability of a Type I Error (Probability of Failing to Reject a True Null) = 1 - Confidence Level.
  - Power of a Test: 1 - Probability of a Type II Error (Probability of Rejecting a False Null Hypothesis)

	Do Not Reject H_0	Reject H_0
		Type I Error
	Correct Decision	Size of Test
H_0 is True	$1-\alpha$:Confidence Level	α : Significance Level
	Type II Error	Correct Decision
H_0 is False	ω	$1-\omega$: Power of Test

- In order to reject a null hypothesis of no effect at an α level of confidence, we need:

$$\hat{\beta} > t_{\alpha} SE\left(\hat{\beta}\right)$$

- If we want power of $1-\omega$:

$$\hat{\beta} > \left(t_{1-\omega} + t_{\alpha}\right) SE\left(\hat{\beta}\right)$$

- $SE\left(\hat{\beta}\right) = \sqrt{V\left(\hat{\beta}\right)} = \sqrt{\frac{\sigma_{\epsilon}^2}{P(1-P)N}}$

- Therefore the Minimum Detectable Effect (MDE) where α is the size and $1-\omega$ is the power:

$$\hat{\beta} > \left(t_{1-\omega} + t_{\alpha}\right) \sqrt{\frac{\sigma_{\epsilon}^2}{P\left(1-P\right)N}}$$

- Now suppose we have grouped data with group effects: v_j. Then, we estimate:

$$Y_{ij} = \alpha + \beta T_{ij} + v_j + \epsilon_{ij}$$

- where there are J clusters of size n
- $v_j \tilde{} i.i.d.\ N\left(0, \tau^2\right)$ and $\epsilon_{ij} \tilde{} i.i.d.\ N\left(0, \sigma_\epsilon^2\right)$
- Then we get as our MDE:

$$\sqrt{\frac{1}{P\left(1-P\right)}}\sqrt{\frac{n\tau^2+\sigma_\epsilon^2}{nJ}}$$

- With individual randomization, we would get:

$$\sqrt{\frac{1}{P\left(1-P\right)}}\sqrt{\frac{\tau^2+\sigma_\epsilon^2}{nJ}}$$

- Ratio between the two $= D = \sqrt{1+\left(n-1\right)\rho}$ where $\rho = \frac{\tau^2}{\tau^2+\sigma_\epsilon^2}$

- With imperfect compliance:

$$\sqrt{\frac{1}{P\left(1-P\right)}}\sqrt{\frac{\sigma_\epsilon^2}{N}}\frac{1}{c-s}$$

- Generally (with spherical disturbances): Standard Errors Given By:

$$\sigma_{\epsilon}^{2}\left(X'X\right)^{-1}$$

- $V\left(\hat{J}\right) > V\left(Conditional\right) > V\left(Stratified\right)$: Imbens, King and Ridder.

- With stratification, this is a diagonal matrix in which case adding a dimension of stratification (constructed to be orthogonal to the other dimensions) will always reduce the standard errors.

7 Intention to Treat Estimates in Experiments

- Treatment: T, Assignment of Treatment: Z

- Average Treatment Effect (ATE):

$$E\left(Y_i^T - Y_i^C\right)$$

  - If you can actually randomize treatment

- Intention To Treat (ITT):

$$E\left(Y_i^T - Y_i^T|Z\right) = E\left(Y_i^T|Z=1\right) - E\left(Y_i^T|Z=0\right)$$

  - If you can randomize access to Treatment but not Treatment itself

- Local Average Treatment Effect (LATE):

$$E\left(Y_i^T - Y_i^T|T, X\right)$$

- Do we want the intention to treat estimate or the treatment effect?