Week 1: RCM and RCTs

JJ Chen

January 16, 2015

About Me and Sessions

- JJ Chen (2nd year Econ), jchen215@uic.edu
- Session Hour: Friday 10:00-10:50 a.m. and 3:00-3:50 p.m., room TBA
- Sessions are optional; I hope it will be valuable
- For every session, I was planning to
  1. review and reinforce class material
  2. go over problem sets if necessary
  3. perhaps provide a complementary perspective on the lecture material
  4. discuss papers assigned

Introduction

- Rubin's Casual Model: potential outcome framework, origins from Neyman's work in agricultural statistics
  - Not the only framework to think about causality, other competing framework: Pearl's Causal Model; decision theoretic...
- Generalize experimental setups to observational studies
- Useful to consider the effect of a cause (policy, law, program...)
  - this doesn't mean we can't study the cause of an effect
  - sometimes you get many hypothesis to test when thinking causes of an effect and might potentially lead to a successful project
  - e.g., Why health care spending in all developed countries, especially US, increase so dramatically?

Notations

- In Econ 534, we followed notational convention in linear algebra
  - scalar: β; vectors: $\boldsymbol{y}$, $\boldsymbol{x}$, $\boldsymbol{\beta}$; matrices: $\boldsymbol{X}$
- In Econ 535, Let's adopt notations in MHE, which follows convention in probability and statistics
  - random variable: Y_i, D_i; a vector of random variables: X_i; parameters: ρ, κ
  - realizations: y_i, d_i, $\boldsymbol{x}_i$
- RCM makes explicit notations for potential outcomes: Y_{1i}, Y_{0i}
  - Y_{1i} is the potential outcome **if** the i would have gotten treatment $\mathrm{D}_i = 1$ (confusion?)
  - another useful notation: $\mathrm{Y}_{1i} \equiv \mathrm{Y}_i \mid do(\mathrm{D}_i = 1)$, or $\mathrm{Y}_{1i} \equiv \mathrm{Y}_i \mid fix(\mathrm{D}_i = 1)$

Causal Effect and the Fundamental Difficulty

Table 1: The Fundamental Problem of Causal Inference

Groups (more groups?)	Y_{1i}	Y_{0i}
Treatment ($D_i = 1$)	Observable as Y_i	Counterfactural
Control ($D_i = 0$)	Counterfactural	Observable as Y_i

- Observed outcome: $Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$
- Individual Treatment Effect is defined as a linear difference
  - $Y_{1i} - Y_{0i} \mid D_i = 1$, $Y_{1i} - Y_{0i} \mid D_i = 0$
- Hopeless at the individual level, go to Average Treatment Effect
  - $ATE = E(Y_{1i} - Y_{0i}) = E(Y_{1i}) - E(Y_{0i})$

Prior Assumption of RCM

- RCM implicitly assume no general equilibrium effect and/or no social interaction
  - or Stable Unit Treatment Value Assumption as referred by Rubin (1986):

SUTVA is simply the a priori assumption that the value of Y for unit u when exposed to treatment twill be the same no matter what mechanism is used to assign treatment t to unit u and no matter what treatments the other units receive, and this holds for all $u = 1, \cdots, N$ and all $t = 1, \cdots, T$.

Linking Structural Eqn to RCM

- Suppose the structural equation of interest:
  - $\mathrm{Y}_i = f(\mathrm{D}_i, \eta_i) = \alpha + \rho \mathrm{D}_i + \eta_i$, where D_i is binary
- What are the assumptions we are imposing?
- Consider the potential outcome of i
  - $\mathrm{Y}_{0i} = \alpha + \eta_i$
  - $\mathrm{Y}_{1i} = \alpha + \rho + \eta_i$
  - individual causal effect: ρ (homogeneous, can be hetero ρ_i)
- Recall the observed outcomes: $\mathrm{Y}_i = \mathrm{D}_i \mathrm{Y}_{1i} + (1 - \mathrm{D}_i)\mathrm{Y}_{0i}$
  - $\implies \mathrm{Y}_i = \mathrm{Y}_{0i} + (\mathrm{Y}_{1i} - \mathrm{Y}_{0i})\mathrm{D}_i = \alpha + \rho_i \mathrm{D}_i + \epsilon_i$

Modeling the Assignment/Selection of D_i

- Tradition in statistics: Chance (Fisher 1925, Neyman 1923)
  - crops (or plots of land) are passive
  - D_i is randomly assigned; no need for modeling
  - automatically, $Y_{1i}, Y_{0i} \perp\!\!\!\perp D_i$
- When generalizing to observational studies, causal inference relies on a key assumption: CIA (other names: ignorability, unconfoundedness, exogeneity, selection-on-observables):
  - $Y_{1i}, Y_{0i} \perp\!\!\!\perp D_i | X_i$
  - X_i is a set of pretreatment variables; within strata given by X_i, variations in D_i is random
  - suitable for the single structure equation: $Y_i = \alpha + \rho D_i + X_i'\gamma + \eta_i$

Modeling the Assignment/Selection of D_i

- Tradition in econometrics: Choice
  - people are actively maximizing utilities s.t. constraints
  - reveal preference: even conditional on X_i, the fact that $d_i = 1$ and $d_j = 0$ are being chosen suggests i and j have different potential outcome (CIA is difficult to justify)
- Examples:
  1. the simplest Roy model: $\mathbb{1}[\mathrm{Y}_{1i} \geq \mathrm{Y}_{0i}]$
     - hunter or fisher? immigrate or not? college or not?
  2. discrete choice:
     $\Pr(\mathrm{D}_i = 1) = \Pr(\mathrm{U}_{1i} \geq \mathrm{U}_{0i}) = \Pr(\mathrm{V}_{1i} + \varepsilon_{1i} \geq \mathrm{V}_{0i} + \varepsilon_{0i})$
  3. selection into D_i is based on a latent variable: $\mathrm{D}_i = \mathbb{1}[\mathrm{D}_i^* \geq 0]$.
     - the latent variable D_i^* is based on an index function:
     - $\mathrm{D}_i^* = h(\mathrm{Z}_i, \xi_i) = \mathrm{Z}_i'\gamma + \xi_i$; Z_i observable, ξ_i unobservable

Interactions between Chance and Choice

- Is “Choice” irrelevant when you have “chance”?
  - Hawthorne effect; attrition; endognous response;
  - Example: consider the impact of universal preschool
    - kids’ cognitive outcome: $\text{Y} = f(\text{PreSchl}, \text{Parents}'\text{Input})$
    - parents’ choice: $\text{Parents}'\text{Input} = g(\text{Wealth}, \text{PreSchl})$
    - what will RCTs estimate? is it sufficient for answering policy question?
- Is “Chance” irrelevant when you consider “choice”?
  - hunting for exogenous shocks
  - Example: consider the impact of unskilled immigrants on native earnings
    - likely selection problem?
    - 1980 Mariel boatlift

More on observable and unobservable

- Selection-on-observables:
  - density: $f(y_{ji} \mid d_i) \neq f(y_{ji})$ but $f(y_{ji} \mid d_i, x_i) = f(y_{ji} \mid \boldsymbol{x}_i)$, $j = 1, 2$
  - mean independence: $\mathrm{E}(\mathrm{Y}_{ji} \mid \mathrm{D}_i) \neq \mathrm{E}(\mathrm{Y}_{ji})$ but $\mathrm{E}(\mathrm{Y}_{ji} \mid \mathrm{D}_i, \mathrm{X}_i) = \mathrm{E}(\mathrm{Y}_{ji} \mid \mathrm{X}_i)$, $j = 1, 2$
  - compactly: CIA: $\mathrm{Y}_{1i}, \mathrm{Y}_{0i} \perp\!\!\!\perp \mathrm{D}_i \mid \mathrm{X}_i$
- Selection-on-unobservables:
  - density: $f(y_{ji} \mid d_i, \boldsymbol{x}_i) \neq f(y_{ji} \mid \boldsymbol{x}_i)$ but $f(y_{ji} \mid d_i, \boldsymbol{x}_i, \xi_i) = f(y_{ji} \mid \boldsymbol{x}_i, \xi_i)$, $j = 1, 2$
  - mean independence: $\mathrm{E}(\mathrm{Y}_{ji} \mid \mathrm{D}_i, \mathrm{X}_i) \neq \mathrm{E}(\mathrm{Y}_i \mid \mathrm{X}_i)$ but $\mathrm{E}(\mathrm{Y}_i \mid \mathrm{D}_i, \mathrm{X}_i, \xi_i) = \mathrm{E}(\mathrm{Y}_i \mid \mathrm{X}_i, \xi_i)$, $j = 1, 2$

Two Examples

- Let's consider two interesting and important RCTs:
  - The Role of Simplification and Information in College Decisions: Results from the H&R Block FAFSA Experiment (Bettinger et al. (2012))
  - Health Insurance and the Demand for MedicalCare: Evidence from a Randomized Experiment (Manning et al. (1987))
- We will use these examples to review questions in the Week 1 study guide

What is the Selection Problem?

- We know that financial aid increases college enrollment
- But the take-up rate of federal financial aid is lower than one might expect. Why?
- This paper assesses whether complexity is an important barrier.
  - the FAFSA application is 8 pages long and has over 100 questions

The Experiment

- Program began in January 2008; 156 H & R Block tax preparation offices in Ohio and Charlotte, NC.
- The process:
  - a person uses HRB for tax preparation services
  - identify families with family income less than $45,000 (AGI from tax return)
  - family member between age 15 and 30 without a bachelor's degree
    - high school seniors and recent grads; adults
  - introduced program and received consent

Random Assignment into Three Groups

- FAFSA treatment
  - HRB helped families prepare the FAFSA using information from the tax return and with additional questions
  - computed eligible financial grants and loans & tuition at local public colleges
  - form was either mailed in by HRB or family took it home to mail it themselves
- Information-only treatment
  - calculate aid from the tax return
  - provided tuition information
  - encouraged families to summit on their own
- Control group
  - received brochure with basic information on the benefits of college and availability of financial aid

Result Discussion

- Table 1: Consent, Exit, and Processing Rates
  - What's the purpose of this table?
- Table 2: Discriptive Statistics
  - What's the implication of Table 2?
- Table 3: Treatment effect on FAFSA filing
  - How would you interpret the estimate 0.157 in the first column?
  - What does the estimated coefficient on information-only suggest?
- Table 4: Treatment effect on college
  - Combining with table 3, what can we say about independent participants?

Backgroup

- Health care costs have grown extremely fast over the last half-century
- Why is the cost of medical services rising so fast?
  - population aging
  - rising incomes
  - defensive medicine
  - technological change
  - expanding insurance generosity
    - moral hazard
    - but how responsive is spending on health to the out of pocket costs?

What is the Selection Problem?

- Suppose you compared health spending between people with no insurance, partial insurance, and full insurance
  - Who would spend the most?
  - Could you interpret this as a causal effect?

The RAND HIE

- "Between November 1974 and February 1977, the HIE enrolled families in six sites: Dayton, Ohio; Seattle, Washington; Fitchburg, Massachusetts; Franklin County, Massachusetts; Charleston, South Carolina; and Georgetown County, South Carolina"
- "The sites were selected to represent the four census regions; to represent the range of city sizes (a proxy for the complexity of the medical delivery system); to cover a range of waiting times to appointment and physician per capita ratios (to test for the sensitivity of demand elasticities to nonprice rationing); and to include both urban and rural sites in the North and the South."

The RAND HIE

- Families were randomly assigned to one of 14 fee-for-service plans or to a pre-paid group practice
  - Variation in the coinsurance rate (the percentage of costs that a person pays out of pocket): Zero, 25%, 50%, and 95%
  - Variation in the upper-limit of out-of-pocket expenses: 5%, 10%, or 15% of family income, up to $1000
- Enrolled for either three or five years

Discussion

- Table 1: Sample by plan and site
- Table 2: Sample means for use of medical services
- One of the really nice things in this papers is that the conclusion contains an lengthy discussion about how the results speak to the existing literature and policy debates
- What do these results imply about the role of insurance in explaining the sustained rise in health care costs?
  - "Because the free plan demand was only around 1.5 times that of the 95 percent plan, it appears that the change in insurance can explain only a small part, perhaps a tenth, of the factor of 7 change in health expenditure in the post-World War II period."

Introduction

- When we read most econometrics textbooks, we mainly learn methods of estimation and inference
  - to learn something about the population from a sample
  - OLS, GLS, FE, Std. Err., F Test, ...
  - admittedly, most of the textbooks authors do talk about causality, marginal effect; ceteris paribus
- When we read empirical economics papers and go to seminar, people talk about identifications, structural model, reduced form...
  - are there any formal definitions of these terms?
  - historically, terms like structural equation and reduced form came from simultaneous equations model, But now?

Usage of Structure

- Historically, economists rely on structural equations models (SEM) to talk about causality
  - an attempt to link Marshall's *ceteris paribus* to data
- "Structure" is referred to some mechanisms that generate the population distribution
  - we wish to "identify" the underlying structure at the population level
  - this is different from "estimating" a population parameter from a sample (more on this later)

Causality is in Our Mind

- But how do we know there exists some underlying structure?
  - We "model" the structure – People (especially economists) like to rationalize what they see in data and every day life
- Example from Pearl:
  - "Did you eat from that tree?"
  - "The woman whom you gave to be with me, she handed me the fruit from the tree; and I ate."
  - Pearl: "God did not ask for explanation, only for the fact... [C]ausal explanation is a **man-made concept**."

My Role as an TA

- Personally, I think many emprical work are exercises about "persuasion"; thus I see my role as an TA to share with you tools that I think can help us persuade our peers:
  - more "vocabularies" or "dialects" that I heard
  - equations that have clear interpretations
  - ID strategies that children can understand
  - tables and graphs that tells stories or fails to tell stories
  - robustness checks that actually do check
- More suggestions?

Sturture Objects

- We accept that we are starting with a model and try to identify it in econometrics ("Everyone has a model in their mind.")
- So what exactly is an structural object?
- Something that is *invariant* no matter what happens
  - examples: preference (economic primitives), law of demand
  - structural objects can be described as structural parameters or functions

Controverials

- How do we identify these structural objects?
  - Two main ways in econometrics: Model-based methods and Design-based methods (See Card's lecture)
- "Structural modeling": estimate all primitives and then simulate "structural relationship"/counterfactural
- "Reduced form": estimate structural relationship with exogenous variations
- These are two extremes, but you can also derive a small set of statistics sufficient to characterize the structural relationship of interest (Chetty (2008), Weyl (2014))

"Persuasive Activities"

- Best empirical workers are able to mix methods to identify structural relationship
- For more discussions of these "persuasive activities" see J. Angrist and Pischke (2010), Keane (2010), Sims (2010), Nevo and Whinston (2010), Deaton (2010), Imbens (2010), Wolpin (2013)
- To be more concrete, let's consider a few structural equations

Example 1: Single Structural Equation

- Example: $\mathrm{Y}_i = f(\mathrm{X}_i, \mathrm{U}_i) = \beta \mathrm{X}_i + \mathrm{U}_i$
  - assume linearity, homogeneous effect, and additive separability
  - Usually U is assumed to be followed some distribution
  - Note that this is not a regression; even the equal sign might not be the mathematical sign (much like an assignment operator)
- In MHE, one single structural equation keep popping out:
  - $\mathrm{Y}_i = \alpha + \rho \mathrm{S}_i + \mathrm{X}_i' \boldsymbol{\gamma} + \epsilon_i$
  - it would be our focus for at least half of the semester

Example 2: Linear Simultaneous Equations Model

- In general matrix notation: $\boldsymbol{Ay} = \boldsymbol{Bz} + \boldsymbol{u}$
  - $\boldsymbol{y}$: endogenous vector (eqm. outcome); $\boldsymbol{z}$: exogenous vector
- Example: supply and demand, or interaction between two agents
  - my diet habit: $p = \alpha_1 q + \beta_1 z_1 + u_1$,
  - my wife's diet habit: $q = \alpha_2 p + \beta_2 z_1 + u_2$
  - simplify to matrix notation?
- Reduced form: $\boldsymbol{y} = \boldsymbol{A}^{-1}\boldsymbol{Bz} + \boldsymbol{A}^{-1}\boldsymbol{u}$
  - $p = \frac{\beta_1 + \alpha_1 \beta_2}{1 - \alpha_1 \alpha_2} z_1 + v_1$
  - $q = \frac{\beta_1 \alpha_2 + \beta_2}{1 - \alpha_1 \alpha_2} z_1 + v_2$
- Point: running a regression based on the reduced form is not sufficient to identify the parameter of interest (α_1 or α_2)

Example 3: Triangular/Recursive System

- Consider the following structural system:
  - $Y_i = f(X_i, U_i)$
  - $X_i = g(Z_i, V_i)$
  - usually assume $Z \perp\!\!\!\perp (U, V)$
- The linear version of it is the traditional IV set up
  - $Y_i = \alpha + \rho X_i + U_i$
  - $X_i = \phi + \delta Z_i + V_i$
  - J. D. Angrist, Imbens, and Rubin (1996) relax many of the assumptions by introducing Rubin Causal Model

References I

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434). Taylor & Francis: 444–55.

Angrist, Joshua, and Jörn-Steffen Pischke. 2010. *The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con Out of Econometrics*. National Bureau of Economic Research.

Bettinger, Eric P, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa

References II

Experiment*." *The Quarterly Journal of Economics* 127 (3). Oxford University Press: 1205–42.

Chetty, Raj. 2008. *Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods*. National Bureau of Economic Research.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning About Development." *Journal of Economic Literature*. JSTOR, 424–55.

Imbens, Guido W. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of Economic Literature*. JSTOR, 399–423.

References III

Keane, Michael P. 2010. "A Structural Perspective on the Experimentalist School." *The Journal of Economic Perspectives*. JSTOR, 47–58.

Manning, Willard G, Joseph P Newhouse, Naihua Duan, Emmett B Keeler, and Arleen Leibowitz. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *The American Economic Review*. JSTOR, 251–77.

Nevo, Aviv, and Michael D Whinston. 2010. "Taking the Dogma Out of Econometrics: Structural Modeling and Credible Inference." *The Journal of Economic Perspectives*. JSTOR, 69–81.

References IV

Rubin, Donald B. 1986. "Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers." *Journal of the American Statistical Association* 81 (396). American Statistical Association: 961–62.

Sims, Christopher A. 2010. "But Economics Is Not an Experimental Science." *The Journal of Economic Perspectives*. JSTOR, 59–68.

Weyl, E Glen. 2014. "Price Theory: A Brief Summary." *Journal of Economic Literature, Forthcoming*.

Wolpin, Kenneth I. 2013. "The Limits of Inference Without Theory." *MIT Press Books* 1. The MIT Press.