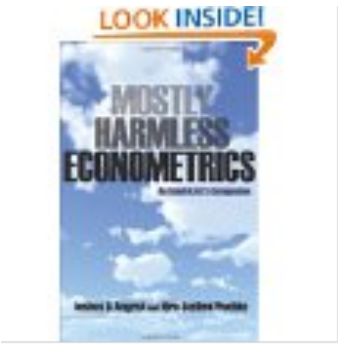
Empirical Methods in Development Economics



Other interesting references

Symposium in The Journal of Economic Perspectives, Volume 24, Number 2, Spring 2010. Starting with:

Angrist and Pischke:

"The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics"

See also:

Deaton (2009)

"Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development".

Banarjee and Duflo (2009)

"The experimental approach to development economics".

Causality with Potential Outcomes

Treatment

D_i: Indicator of treatment intake for *unit i*

$$D_i = \begin{cases} 1 & \text{if unit } i \text{ received the treatment} \\ 0 & \text{otherwise.} \end{cases}$$

Outcome

Y_i: Observed outcome variable of interest for unit i

Potential Outcomes

Y_{0i} and Y_{1i}: Potential outcomes for unit i

$$\begin{array}{ll} Y_{1i}: & \text{Potential outcome for unit } i \text{ with treatment} \\ Y_{0i}: & \text{Potential outcome for unit } i \text{ without treatment} \end{array}$$

Causality with Potential Outcomes

Treatment Effect

The treatment effect or causal effect of the treatment on the outcome for unit i is the difference between its two potential outcomes:

$$Y_{1i} - Y_{0i}$$

Observed Outcomes

Observed outcomes are realized as

$$Y_i = Y_{1i}D_i + Y_{0i}(1 - D_i) \quad \text{or} \quad Y_i = \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases}$$

Fundamental Problem of Causal Inference

Cannot observe both potential outcomes (Y_{1i}, Y_{0i})

Identification Problem for Causal Inference

Problem

Causal inference is difficult because it involves missing data. How can we find $Y_{1i} - Y_{0i}$?

- A large amount of homogeneity would solve this problem:
  - (Y_{1i}, Y_{0i}) constant across individuals
  - (Y_{1i}, Y_{0i}) constant across time
- Unfortunately, often there is a large degree of heterogeneity in the individual responses to participation in public programs

Stable Unit Treatment Value Assumption (SUTVA)

Assumption

Observed outcomes are realized as

$$Y_i = Y_{1i}D_i + Y_{0i}(1 - D_i)$$

- Implies that potential outcomes for unit i are unaffected by the treatment of unit j
- Rules out interference across units
- Examples:
  - Effect of fertilizer on plot yield
  - Effect of flu vaccine on hospitalization
- This assumption may be problematic, so we should choose the units of analysis to minimize interference across units.

Quantities of Interest (Estimands)

ATE

Average treatment effect is:

$$\alpha_{ATE} = E[Y_1 - Y_0]$$

ATET

Average treatment effect on the treated is:

$$\alpha_{ATET} = E[Y_1 - Y_0 | D = 1]$$

Average Treatment Effect (ATE)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0
$E[Y_1]$	1.5				
$E[Y_0]$		0.5			
$E[Y_1 - Y_0]$					1

$$\alpha_{ATE} = E[Y_1 - Y_0] = 3 \cdot (1/4) + 0 \cdot (1/4) + 1 \cdot (1/4) + 0 \cdot (1/4) = 1$$

Average Treatment Effect on the Treated (ATET)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0
$E[Y_1 \mid D=1]$	2				
$E[Y_0 \mid D=1]$		0.5			
$E[Y_1 - Y_0 \mid D=1]$					1.5

$\alpha_{ATET} = E[Y_1 - Y_0 \mid D=1] = 3 \cdot (1/2) + 0 \cdot (1/2) = 1.5$

Selection Bias

Problem

Comparisons of earnings for the treated and the untreated do not usually give the right answer:

$$\begin{aligned} E[Y \mid D=1] - E[Y \mid D=0] &= E[Y_1 \mid D=1] - E[Y_0 \mid D=0] \\ &= \underbrace{E[Y_1 - Y_0 \mid D=1]}_{ATET} + \underbrace{\{E[Y_0 \mid D=1] - E[Y_0 \mid D=0]\}}_{BIAS} \end{aligned}$$

- Bias term is not likely to be zero for most public policy applications
- Selection into treatment is often associated with the potential outcomes

Assignment Mechanism

Assignment Mechanism

Assignment mechanism is the procedure that determines which units are selected for treatment intake. Examples include:

- random assignment
- selection on observables
- selection on unobservables

Most models of causal inference attain identification of treatment effects by restricting the assignment mechanism in some way.

Key Ideas

- Causality is defined by potential outcomes, not by realized (observed) outcomes
- Observed association is neither necessary nor sufficient for causation
- Estimation of causal effects of a treatment (usually) starts with studying the assignment mechanism

An Example

Treatment:
A fertilizer program where fertilizers are given for free to some farmers.

Effect
= (Yield for the farmers who got fertilizer) – (Yield at *the same point in time* for *the same farmers* in absence of the program).

Problem:
We never observe the outcome of the same individual with and without program at the same point in time.

We cannot simply compare before and after

- Other things may happen over time so that we cannot separate the effect of the treatment and the effect of those other things.
- Even if *you* know "nothing else happened" it is hard to convince others.
- The burden of proof is on you.

We cannot simply compare with those who did not get fertilizers

- Some may *choose* not to participate.
- Those *not offered* the program may differ.
- Remember, the burden of proof is on you.

A solution

- Find a good proxy for what would have happened to the outcome in the absence of program
- Compare the farmer with someone who is exactly like her but who was not exposed to the intervention
- In other words, we must find a valid **Counterfactual**

→ **only reason** for different outcomes between treatment and counterfactual is the **intervention**

Techniques to be used in the course

- Randomized Experiments
- Instrumental variables
- Panel data and difference in differences.
- Regression discontinuity design.

Andreas's advice about empirics: be critical!