STATISTICAL HORIZONS

Treatment Effects Analysis

Stephen Vaisey, Ph.D.

Upcoming Seminar:
May 17-18, 2018, Los Angeles, California

Treatment Effects Analysis

Stephen Vaisey
December 2016

1

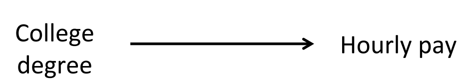
A Note on Software

- I use Stata for examples
- You will see concepts *implemented* in Stata but understanding them will allow you to implement them correctly in other software
- R and SAS have matching capabilities (but more dispersed than Stata)

2

PART I: THE POTENTIAL OUTCOMES FRAMEWORK

Asking causal questions



Does more education
cause higher pay?

Asking causal questions

Job training → Employment

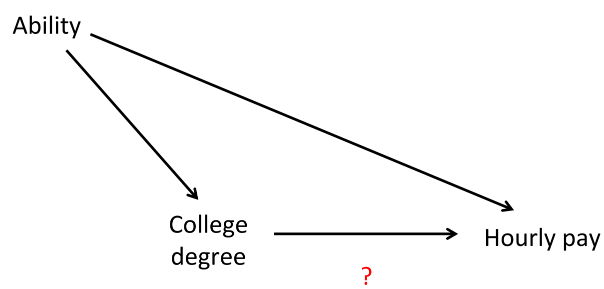
Does participating in a job training program *cause* a higher probability of employment?

Asking causal questions

Boycott → Share price

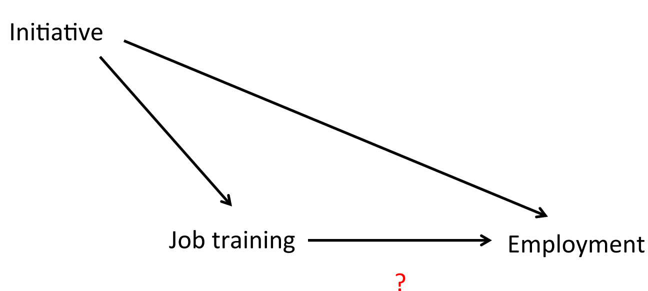
Does a boycott of a company affect its share price?

Threats to causal inference



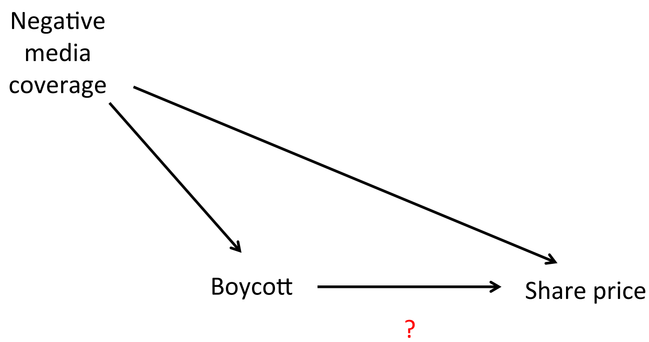
7

Threats to causal inference



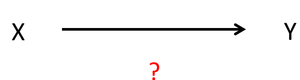
8

Threats to causal inference



Key Term: Identification

How do we *identify* a causal effect of a treatment on an outcome?



Experiments and the idea of an "independent variable"

What is an experiment?

How do experiments solve the problem we just talked about?

Example: Audit Study

- Correll et al. sent resumes and letters to 300+ employers
- Treatment categories
  – **on resume**: PTA officer vs. alumni association officer
  – **in letter**: mention relocating "with family" vs. mention relocating

Results

TABLE 6
PROPORTIONS OF APPLICANTS RECEIVING CALLBACKS BY GENDER AND PARENTAL STATUS

	Callbacks/Total Jobs	Proportion Called Back
Mothers	10/320	.0313
Childless women	21/320	$.0656^{++}$
Fathers	16/318	.0503
Childless men	9/318	.0283

NOTE.—Mothers and childless women applied to the same 320 jobs; fathers and childless men applied to the same 318 jobs. See text for variable descriptions.
$^{++}$ $Z < .05$, test for difference in proportions between parents and nonparents.

They find evidence of a (negative) causal effect of being a mother on the probability that a female applicant gets called back.

13

Experiments are great!

Assuming successful randomization (which is more likely the larger your sample), you **know** it's the treatment that is causing the effect

14

Experiments can't do everything

- Ethical concerns
- External validity concerns
- Harder on units of analysis other than people (e.g., firms, countries)
- Sometimes things we want to understand are hard or impossible to assign randomly (education, motherhood, divorce, boycotts)

15

Why are experiments able to identify causal effects?

16

Some notation

D	a binary treatment variable (0 or 1)
Y	the value of the outcome we observe
Y^0	the value the outcome would take if D=0
Y^1	the value the outcome would take if D=1

17

Some notation

D	a binary treatment variable (0 or 1)
Y	the value of the outcome we observe
Y^0	the value the outcome would take if D=0
Y^1	the value the outcome would take if D=1

Let's think about these a little more carefully

18

The world *before* the experiment

Subject	Y^0	Y^1	D	Y
Andrew	2	3		
Barbara	3	4		
Catherine	3	4		
David	2	3		

What do these numbers mean?

19

The world *after* the experiment

Subject	Y^0	Y^1	D	Y
Andrew		3	1	3
Barbara	3		0	3
Catherine		4	1	4
David	2		0	2

$$Y = DY^1 + (1-D)Y^0$$

20

That is…

$Y = Y^1$ for D=1

$Y = Y^0$ for D=0

We **can't know** Y^1 for those D=0

We **can't know** Y^0 for those D=1

Potential outcomes and counterfactuals

Y^0 and Y^1 are *potential outcomes*

In the real world, D is either 1 or 0 for each case
We see Y^1 **or** Y^0 for each case, never both

When D=0, Y^1 is *counterfactual*
When D=1, Y^0 is *counterfactual*

What do we want to know?

We really care about the difference between Y^0 and Y^1 (why?)

$$\text{Let } \delta_i = y_i^1 - y_i^0$$

$$E[\delta] = E[Y^1 - Y^0]$$

$$E[\delta] = E[Y^1] - E[Y^0]$$

This is the definition of a ***treatment effect***

23

Assume an experiment: What is *E[δ]* here?

Subject	Y^0	Y^1	*D*	*Y*
Andrew	2	3		
Barbara	3	4		
Catherine	3	4		
David	2	3		

Why can't we make these calculations in real life?

24

Assume an experiment: What is *E[δ]* here?

Subject	Y^0	Y^1	*D*	*Y*
Andrew		3	1	3
Barbara	3		0	3
Catherine		4	1	4
David	2		0	2

Why can we do this with experiments?

$$D \perp Y^0$$

$$D \perp Y^1$$

$$E[Y^0 \mid D = 0] = E[Y^0 \mid D = 1]$$

$$E[Y^1 \mid D = 0] = E[Y^1 \mid D = 1]$$

Or in other words…

In a properly executed experiment, there is no association between the potential outcome variables and treatment assignment.

$$E[Y \mid D=0] \; ; \; E[Y^0]$$

$$E[Y \mid D=1] \; ; \; E[Y^1]$$

$$so\ldots$$

$$E[\delta] = E[Y \mid D=1] - E[Y \mid D=0]$$

27

What is this treatment effect?

$E[\delta]$ is the expected value (mean) of the difference between each unit's value of Y^1 and Y^0. It is the ***average treatment effect***.

Even though the individual differences are unobservable (because either Y^0 or Y^1 will be counterfactual for each unit), we can estimate the mean difference via experiment.

28

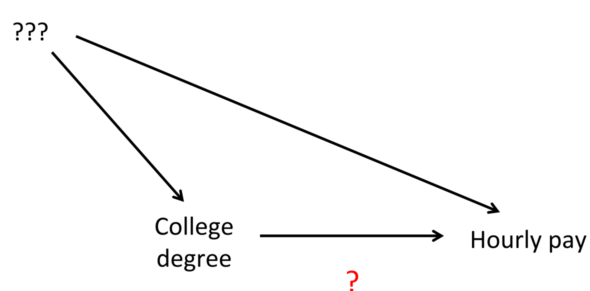
Life is not always so simple

Two sources of bias with observational data:

1. Treated and control cases would ***be different*** from each other even in the same treatment state (baseline bias)
2. Treated and control cases would ***respond differently*** to treatment (TE heterogeneity)

29

An observational example



30