In-person session 13

April 14, 2022

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

Exam and final project stuff

Control variables and confounding

Ethics + pre-analysis plans

Sharing your stuff

Simulating data

Exam and final project stuff

Control variables and confounding

Do we really not need to interpret every coefficient?!

Why do we teach stats with the idea of throwing in a bunch of control variables if that's really not the best way to evaluate impact?

What is the point of using control variables if they aren't going to be interpreted?

How do we know which coefficients we need to ultimately worry about and interpret?

Why did we control for things in the RCT section of problem set 8?

How can you NOT have unobserved confounders unless you're doing a laboratory study where you can control everything?

Design-based inference seems easier since there's no unobserved confounding?

Why don't we just do that all the time?

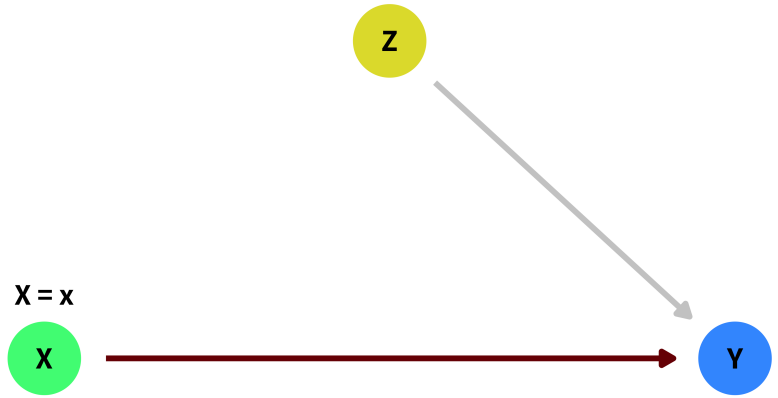
You said that choosing between model-based inference versus design-based inference is based on the situation that we are in.

Which one of the approaches do you think will be used the most in the real world?

Do we not need to use a DAG if we use one of the design-based methods?

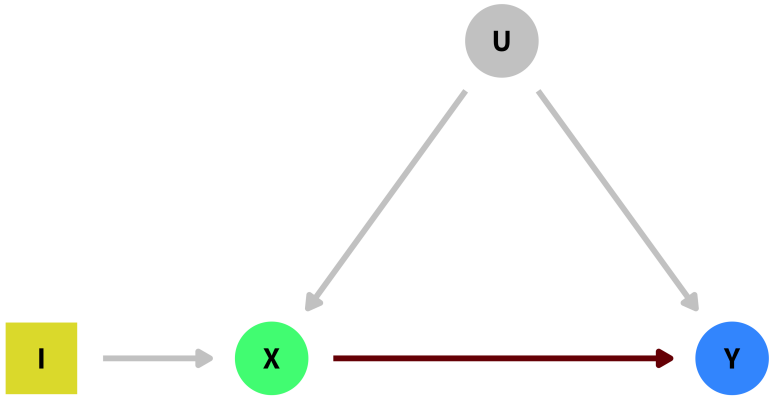
Randomized trial

Randomization deletes all arrows into X



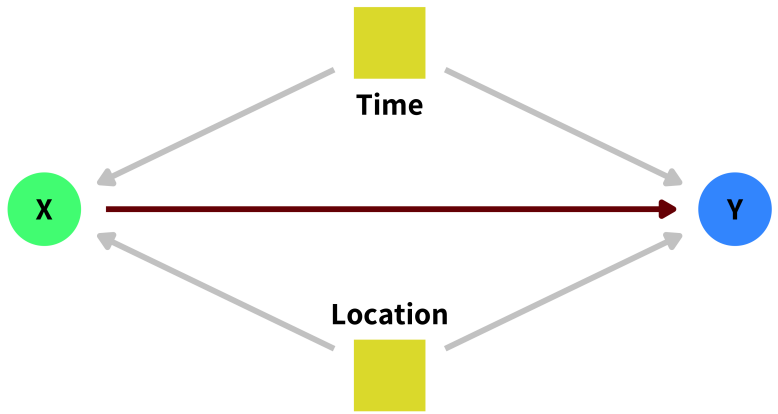
Instrumental variable

Find effect of instrument (I) on X, then find effect of (X | I) on Y



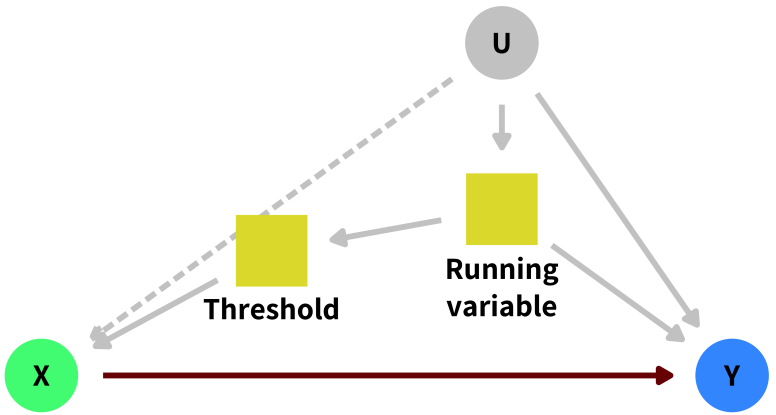
Difference-in-differences

Adjust for both time (e.g., year) and location (e.g., country, state)



Regression discontinuity

Adjust for both the running variable and the threshold



Ethics + pre-analysis plans

Where can one publish their assumptions before running the numbers in research?

OSF.io AsPredicted.org

Do people really provide all their code and analysis in advance? What if you realize you made a mistake or want to explore something later?

What if we have additional variables to investigate that we missed out during our simulation of fake data and pre-registration/pre-analysis stage? Can we add them once they have been registered?

Does pre-registering include interactions or quadratics you want to test?

Standard operating procedures

—

Departures from preregistered plan

—

Example

Should you share the synthetic data as part of a preanalysis plan?

Does pre-registration kill creativity and "ah-ha" moments?

Example of confirmatory vs. exploratory preregistration

Sharing your stuff

What can we put on our resumes now?

Can I really just post R stuff on a website or on Twitter?

It seems like that's what the experts do, not me!

(Public work)

Websites for sharing R stuff

Super easiest (but least flexible): RPubs

Easiest (but less flexible): R Markdown websites

Example; example

Medium-est: Distill

Example

Hardest (but most flexible): Blogdown

Example; example

Simulating data