In-person session 11

March 28, 2024

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

General questions

IV questions

IV fun times

Synthetic data fun times

General questions

The final project instructions mention that we need to preregister our questions.

Do we really need to do that?

Tell us more about exam 2

Minor schedule change!

IV questions

Can you review endogeneity and exogeneity?

Slide from lecture

Can you review the three IV conditions?

Slide from lecture

Are there certain disciplines that tend to use instrumental variables more than others, like how we learned that DAGs are big in public health and diff-in-diff and RDD are all the craze in econ?

Why are you making us do 2SLS manually when `iv_robust()` exists?

Given the strict criteria for instrumental variables, they seem pretty impractical and uncommon (especially compared to diff-in-diff and RDD).

Why do you include instrumental variables as a part of this course?

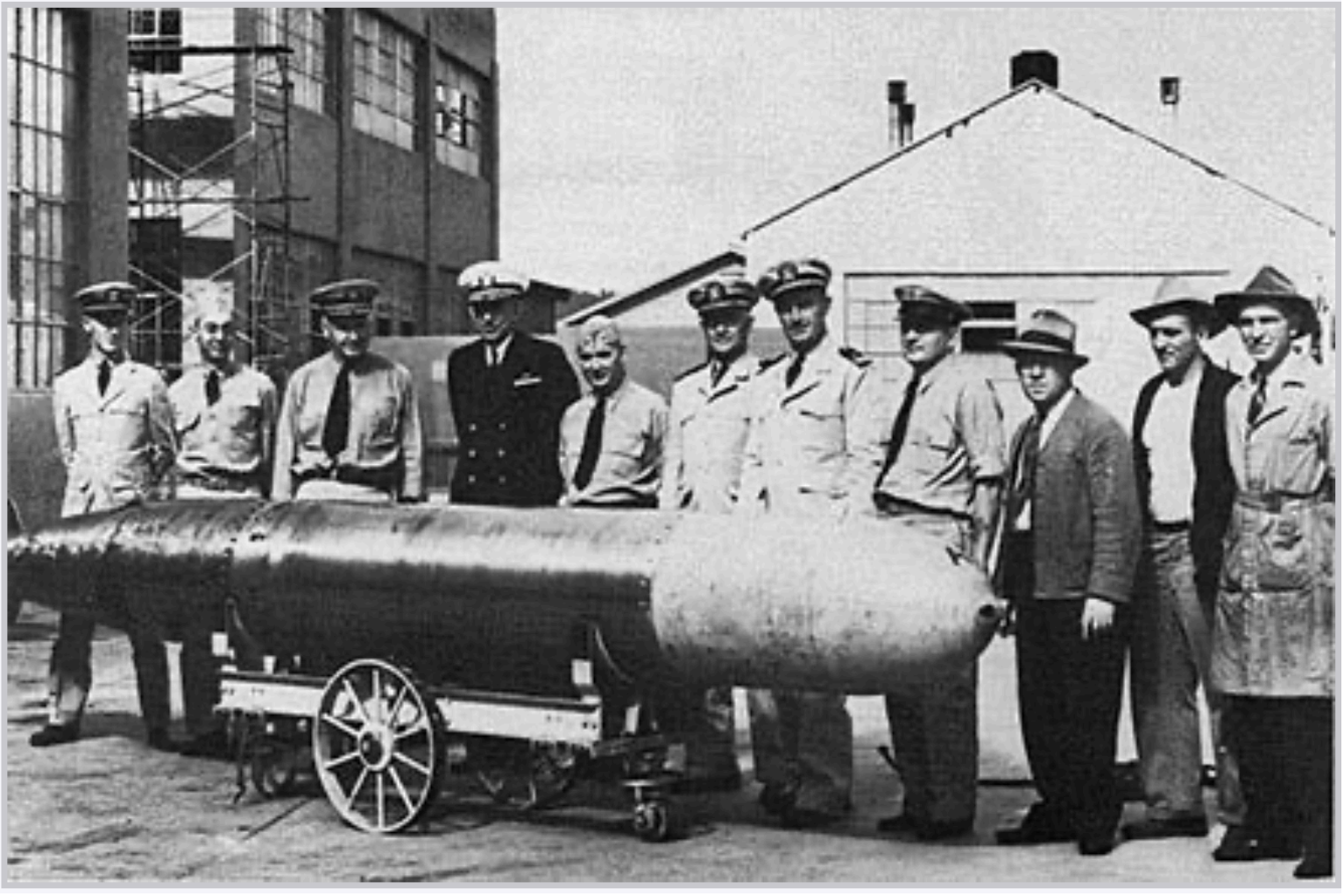
What's really the point of doing IV if finding instruments is so difficult and easy to mess up?

Why even bother?

Is there like a “bank” of good IVs?

Do you have a method that helps you think of instruments, or a popular process that people usually use to come up with ideas?

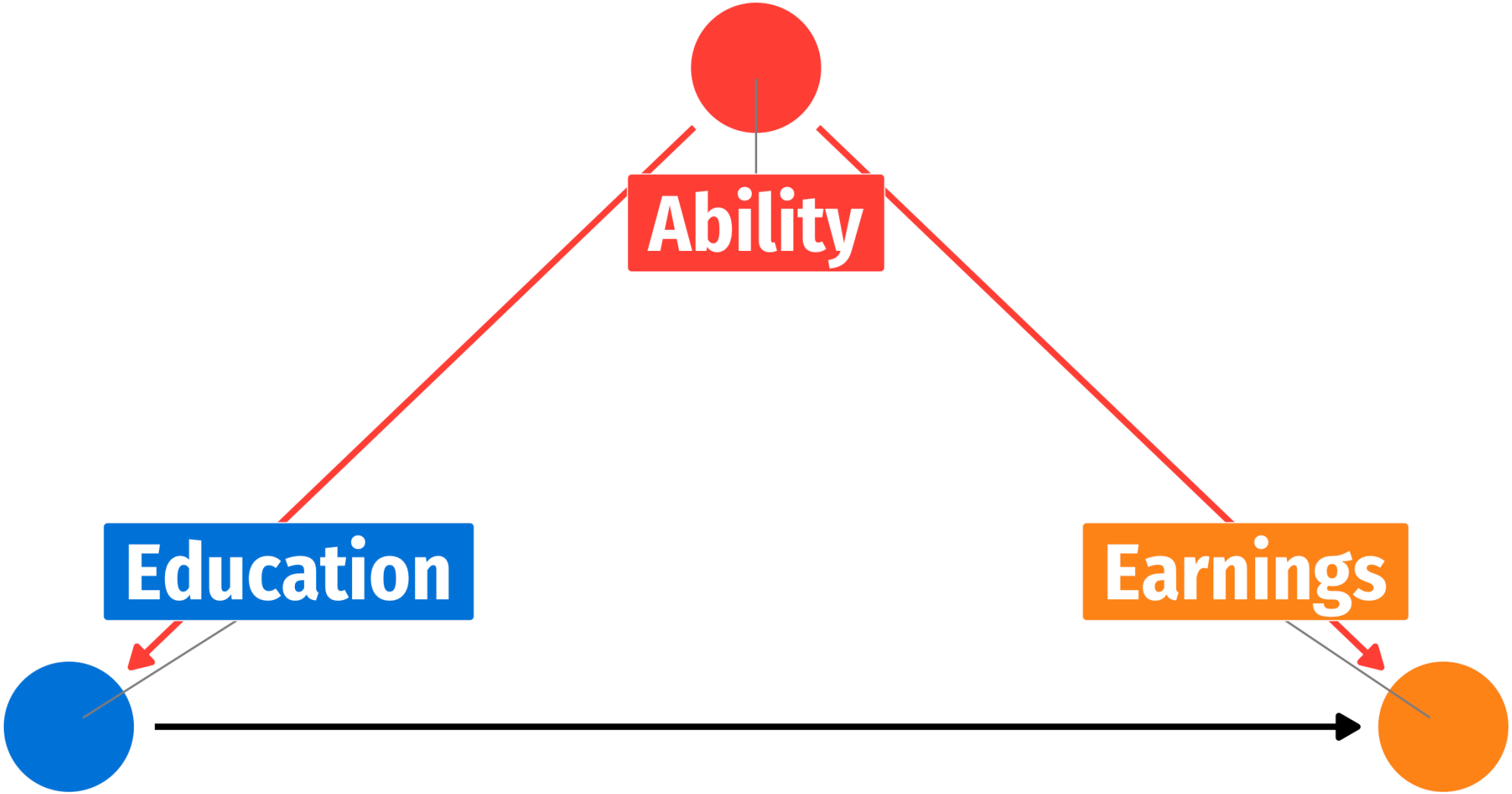
What does it mean to tell a good story about excludability and exogeneity?



Why would we use instrumental variables when we can simply use DAGs to control for things instead?

What is the advantage of using an IV versus trying to find a variable that is correlated with the excluded variable of interest? For example, trying to find a variable highly correlated with "ability"? This seems also difficult but less difficult than finding a true IV.

Could we assume that things like “ability” are latent variables and then try to model them directly?


Ability
Education
Earnings

In the lecture you mentioned that the instrumental variable should be weird (or make people say huh?). However, in *The Effect*, the author states that the instrument should be relevant. This is a bit mixed messaging. Which approach should we use?

Formal definitions of relevancy, excludability, and exogeneity

Why are things like weather, distance, or terrain bad instruments? How do they violate the exclusion restriction?

Lecture slides

Fuzzy RDD requires an instrument, but instruments seem impossible to find, so can we ever really do fuzzy RDD?

IV fun times

Exam 2

Synthetic data fun times!

Basic process

1: Draw a DAG

2: Create standalone exogenous columns

3: Connect endogenous columns

4: Polish columns

Iterate. Iterate so so much.