In-person session 9

October 17, 2022

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

General questions

Final project

Simple diff-in-diff

Two-way fixed effects

Clusters and sensitivity

General questions

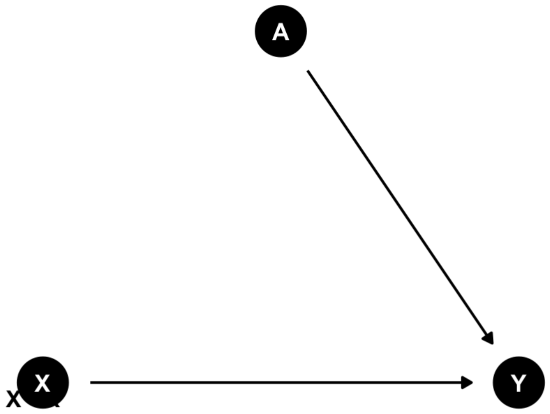
Should we control for variables to close as many backdoors as possible in our diff-in-diff model?

Design-based identification

Use a special situation to isolate arrow

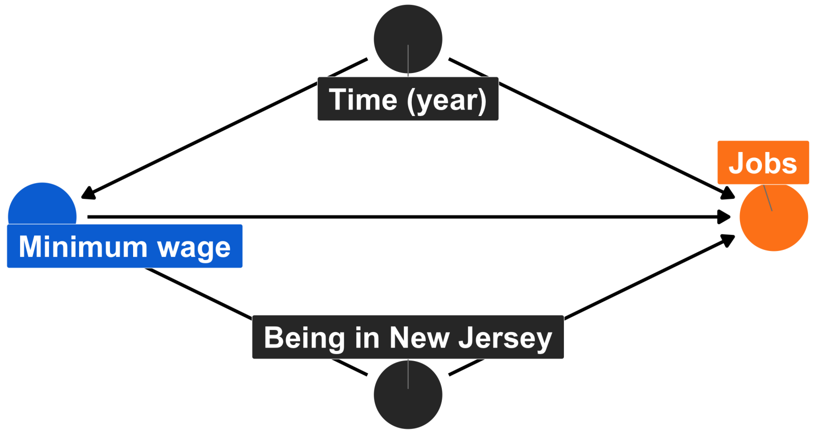
RCTs

Use randomization
to remove confounding



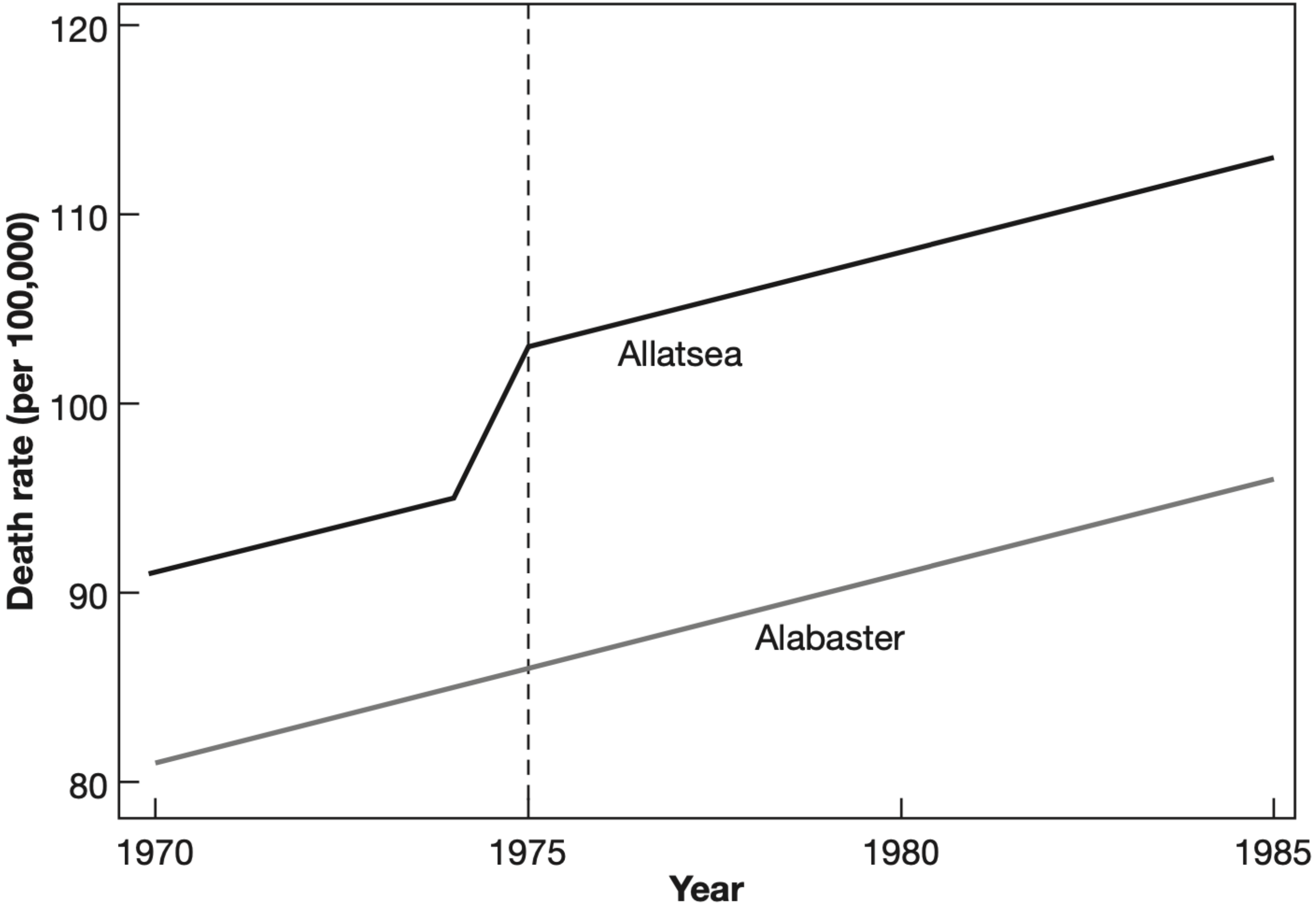
Difference-in-differences

Use before/after & treatment/control
differences to remove confounding



How does moving time back let us check for parallel trends?

Figure 5.4
An MLDA effect in states with parallel trends



Can you conduct diff-in-diff with a binary outcome?

I keep reading about estimates, estimands, and estimators.

What are these and are they the same thing?

Estima(and|or|ate)s

Estimand

Theoretical thing you want to know (β)

Estimator

Process for guessing the thing (e.g., diff-in-diff with interaction term)

Estimate

The guess (β-hat)

estimand

estimate

Ingredients

- 150g unsalted butter, plus extra for greasing
- 150g plain chocolate, broken into pieces
- 150g plain flour
- ½ tsp baking powder
- ½ tsp bicarbonate of soda
- 200g light muscovado sugar
- 2 large eggs

Method

1. Heat the oven to 160C/140C fan/gas 3. Grease and base line a 1 litre heatproof glass pudding basin and a 450g loaf tin with baking parchment.

2. Put the butter and chocolate into a saucepan and melt over a low heat, stirring. When the chocolate has all melted remove from the heat.

estimator

Figure 2.5: The Identification-Estimation Flowchart – a flowchart that illustrates the process of moving from a target causal estimand to a corresponding estimate, through identification and estimation.

1. Theoretical, unobservable estimand (τ):

$$\tau = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{\substack{\text{Mean over} \\ \text{all countries} \\ \text{eligible for aid}}} \Bigg[\underbrace{Y_{it}(x'_{i,t-1:T})}_{\substack{\text{Potential foreign aid} \\ \text{with alternative} \\ \text{NGO legal history}}} - \underbrace{Y_{it}(x_{i,t-1:T})}_{\substack{\text{Potential foreign aid} \\ \text{with actual} \\ \text{NGO legal history}}} \Bigg] \tag{6}$$

$$\tau = \underbrace{\mathbf{E}}_{\substack{\text{Expectation} \\ \text{for all countries} \\ \text{eligible for aid}}} \Bigg[\underbrace{Y_{it} \mid \text{do}(x'_{i,t-1:T})}_{\substack{\text{Causal effect} \\ \text{with alternative} \\ \text{NGO legal history}}} - \underbrace{Y_{it} \mid \text{do}(x_{i,t-1:T})}_{\substack{\text{Causal effect of} \\ \text{actual NGO legal} \\ \text{history on foreign aid}}} \Bigg] \tag{7}$$

2. Empirical estimand (θ):

$$\theta = \underbrace{\mathbf{E}_Z}_{\substack{\text{Expectation} \\ \text{conditional} \\ \text{on observed} \\ \text{confounders } Z}} \Bigg[\underbrace{\mathbf{E}[Y_{it} \mid X_{i,t-1:T} = X_{i,t-1:T} + 1, Z]}_{\substack{\text{Observed mean aid given} \\ \text{total NGO laws in } t-1 \\ \textbf{plus one hypothetical extra law}}} - \underbrace{\mathbf{E}[Y_{it} \mid X_{i,t-1:T,Z}]}_{\substack{\text{Observed mean aid given} \\ \text{total NGO laws in } t-1}} \Bigg] \tag{8}$$

3. Estimate of estimand ($\hat{\theta}$):

$$\hat{\theta} = \hat{g}(x_{t-1:T} + 1; \hat{\beta}) - \hat{g}(x_{t-1:T}; \hat{\beta}),$$

where

$$\hat{\beta} = (\hat{\beta}_0 + \hat{\beta}_1 x_{i,t-1}) \times \mathbf{IPTW}_{i,t-1:T}, \tag{9}$$

which simplifies to

$$\hat{\theta} = \underbrace{\hat{\beta}_1}_{\substack{\text{average} \\ \text{causal} \\ \text{effect}}}$$

Final project

Tell us more about the final project!

Simple diff-in-diff

Minimum legal drinking age

MLDA reduction

Two states: Alabama vs. Arkansas

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Alabama} + \beta_2 \text{ After 1975} + \\ \beta_3 \text{ (Alabama} \times \text{After 1975)}$$

Organ donations

Two states: California vs. New Jersey

$$\begin{aligned}\text{Donation rate} = &\ \beta_0 + \beta_1 \text{ California} + \beta_2 \text{ After Q22011} + \\ &\ \beta_3 \left(\text{California} \times \text{After Q22011}\right)\end{aligned}$$

Two-way fixed effects (TWFE)

Two states: Alabama vs. Arkansas

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Alabama} + \beta_2 \text{ After 1975} + \beta_3 \text{ (Alabama} \times \text{After 1975)}$$

All states: `Treatment == 1`
if legal for 18-20-year-olds to drink

$$\text{Mortality} \; = \; \beta_0 + \beta_1 \; \text{Treatment} + \beta_2 \; \text{State} + \beta_3 \; \text{Year}$$

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Alabama} + \beta_2 \text{ After 1975} + \beta_3 \text{ (Alabama} \times \text{After 1975)}$$

vs.

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Treatment} + \beta_2 \text{ State} + \beta_3 \text{ Year}$$

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Alabama} + \beta_2 \text{ After 1975} + \beta_3 \text{ (Alabama} \times \text{After 1975)}$$

vs.

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Treatment} + \beta_2 \text{ State} + \beta_3 \text{ Year}$$

vs.

$$\text{Mortality} = \beta_0 + \beta_1 \text{ Treatment} + \beta_2 \text{ State} + \beta_3 \text{ Year} + \beta_4 \text{ (State} \times \text{Year)}$$

Table 5.2

Regression DD estimates of MLDA effects on death rates

Dependent variable	(1)	(2)	(3)	(4)
All deaths	10.80	8.47	12.41	9.65
	(4.59)	(5.10)	(4.60)	(4.64)
Motor vehicle accidents	7.59	6.64	7.50	6.46
	(2.50)	(2.66)	(2.27)	(2.24)
Suicide	.59	.47	1.49	1.26
	(.59)	(.79)	(.88)	(.89)
All internal causes	1.33	.08	1.89	1.28
	(1.59)	(1.93)	(1.78)	(1.45)
State trends	No	Yes	No	Yes
Weights	No	No	Yes	Yes

Notes: This table reports regression DD estimates of minimum legal drinking age (MLDA) effects on the death rates (per 100,000) of 18–20-year-olds. The table shows coefficients on the proportion of legal drinkers by state and year from models controlling for state and year effects. The models used to construct the estimates in columns (2) and (4) include state-specific linear time trends. Columns (3) and (4) show weighted least squares estimates, weighting by state population. The sample size is 714. Standard errors are reported in parentheses.

$$\text{Donation rate} = \beta_0 + \beta_1 \text{ California} + \beta_2 \text{ After Q22011} + \beta_3 \text{ (California} \times \text{After Q22011)}$$

vs.

$$\text{Donation rate} = \beta_0 + \beta_1 \text{ Treatment} + \beta_2 \text{ State} + \beta_3 \text{ Quarter}$$

What about this staggered treatment stuff?

See this

What are random effects?

See this

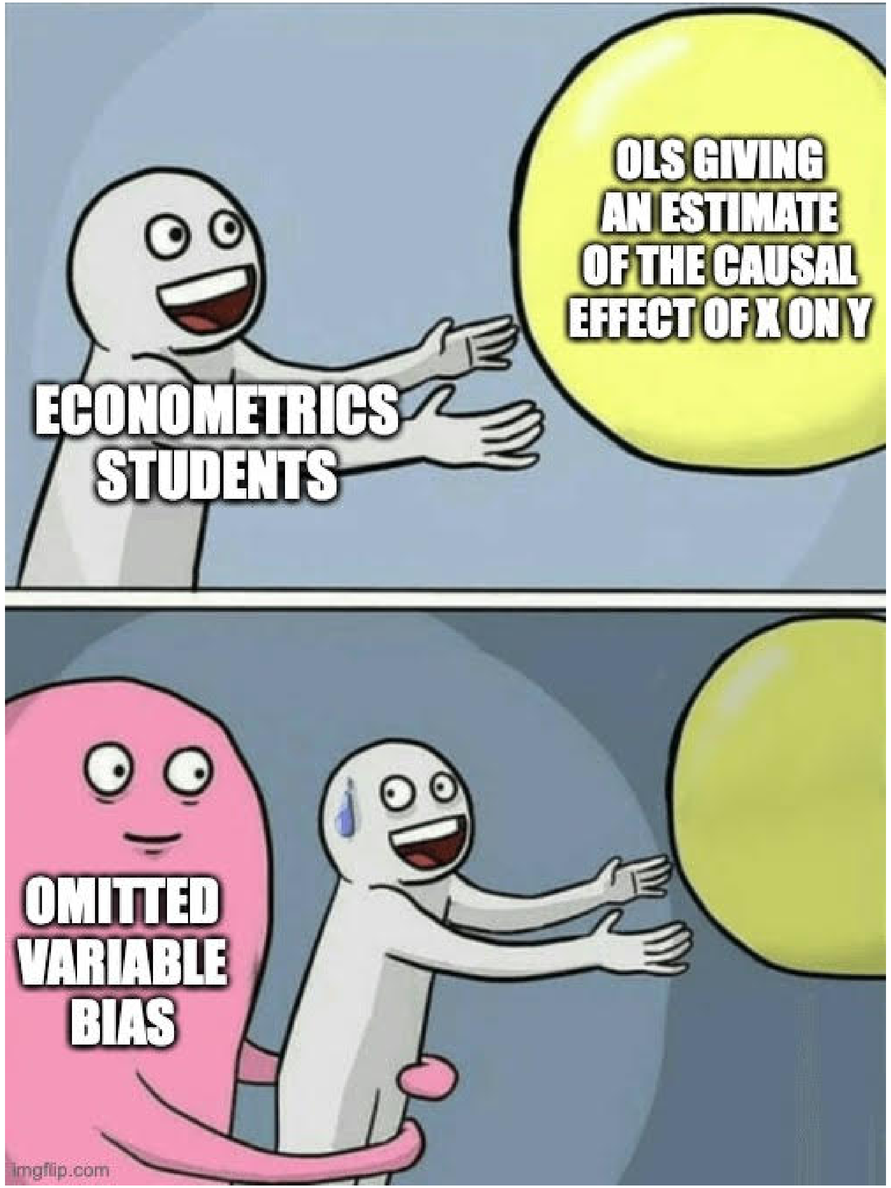
Clusters and sensitivity

What are robust standard errors?

What are clustered standard errors?

How do we know when we've got the right confounders in our DAG?

How do we solve the fact that we have so many unknowns in our DAG?


OLS GIVING AN ESTIMATE OF THE CAUSAL EFFECT OF X ON Y
ECONOMETRICS STUDENTS
OMITTED VARIABLE BIAS
imgflip.com