In-person session 8

March 2, 2023

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

Models vs. designs

Interactions and regression

Simple diff-in-diff

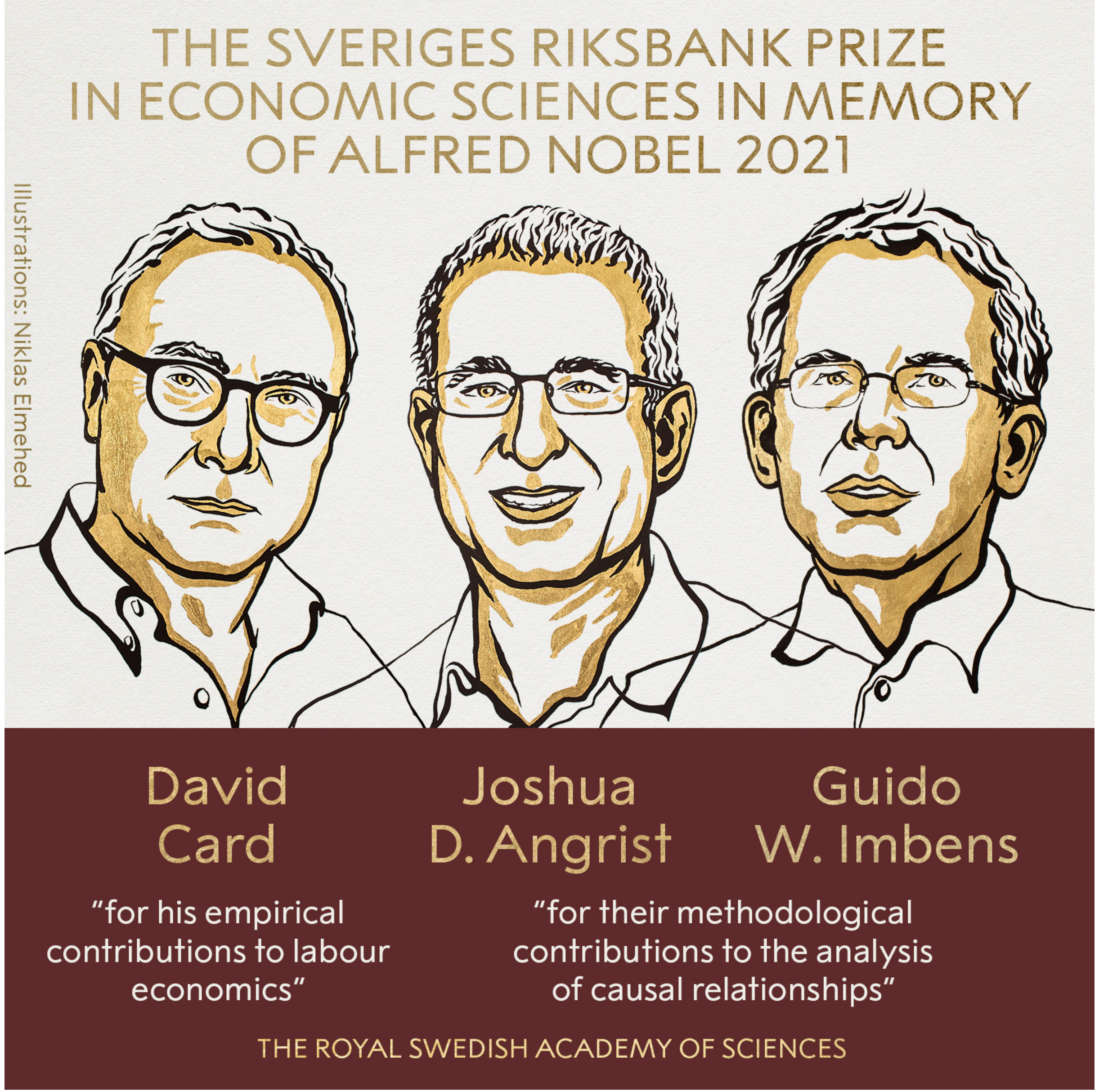
Two-way fixed effects

Two quick things

Chunk names

Correct SEs for IPW

Models vs. designs

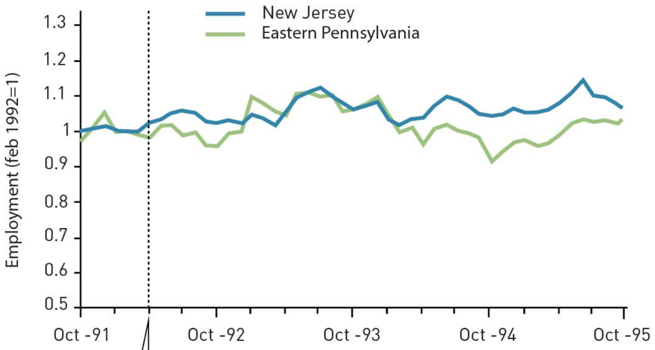

THE SVERIGES RIKSBANK PRIZE
IN ECONOMIC SCIENCES IN MEMORY
OF ALFRED NOBEL 2021
Illustrations: Niklas Elmehed
David Card
Joshua D. Angrist
Guido W. Imbens
"for his empirical contributions to labour economics"
"for their methodological contributions to the analysis of causal relationships"
THE ROYAL SWEDISH ACADEMY OF SCIENCES


How the Estimated $200 Billion Reduction in Consumer Spending
in 2013 Would Likely Be Spread Out
THE WHITE HOUSE
WASHINGTON

The effect of increasing the minimum wage

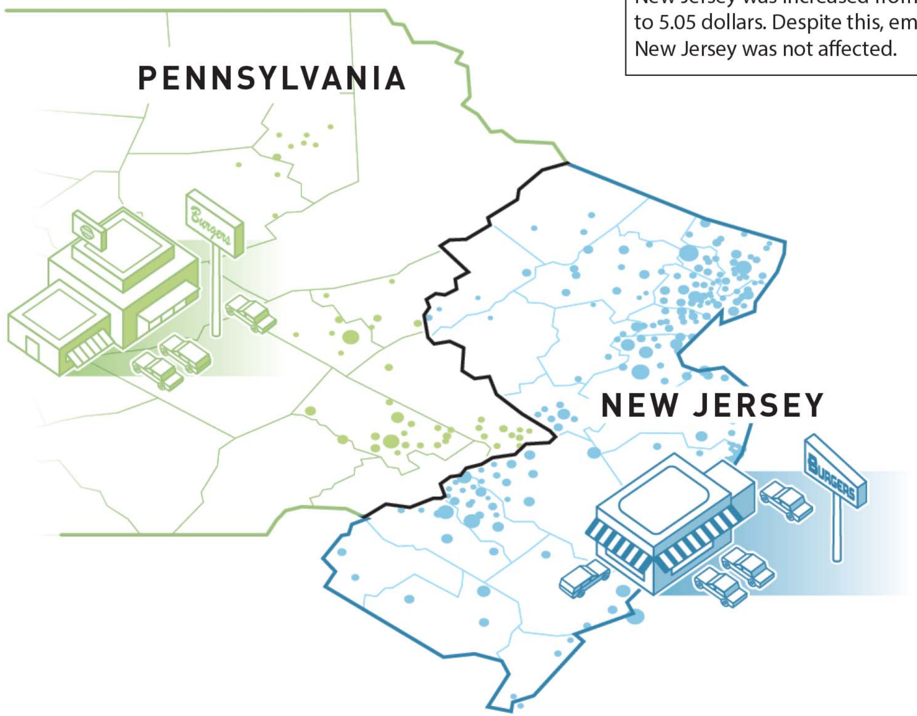
Card and Krueger used a natural experiment to study how increasing the minimum wage affects employment.

The researchers identified a treatment group (restaurants in New Jersey) and a control group (restaurants in eastern Pennsylvania) to measure the effect of increasing the minimum wage.



1 April 1992: The hourly minimum wage in New Jersey was increased from 4.25 dollars to 5.05 dollars. Despite this, employment in New Jersey was not affected.

CONTROL GROUP TREATMENT GROUP



Design-based vs. model-based inference

Special situations vs. controlling for stuff

How would you know when it is appropriate to use a quasi-experiment over an RCT?

Identification strategies

The goal of *all* these methods is to isolate (or **identify**) the arrow between treatment → outcome

Model-based identification

DAGs

Matching

Inverse probability weighting

Design-based identification

Randomized controlled trials

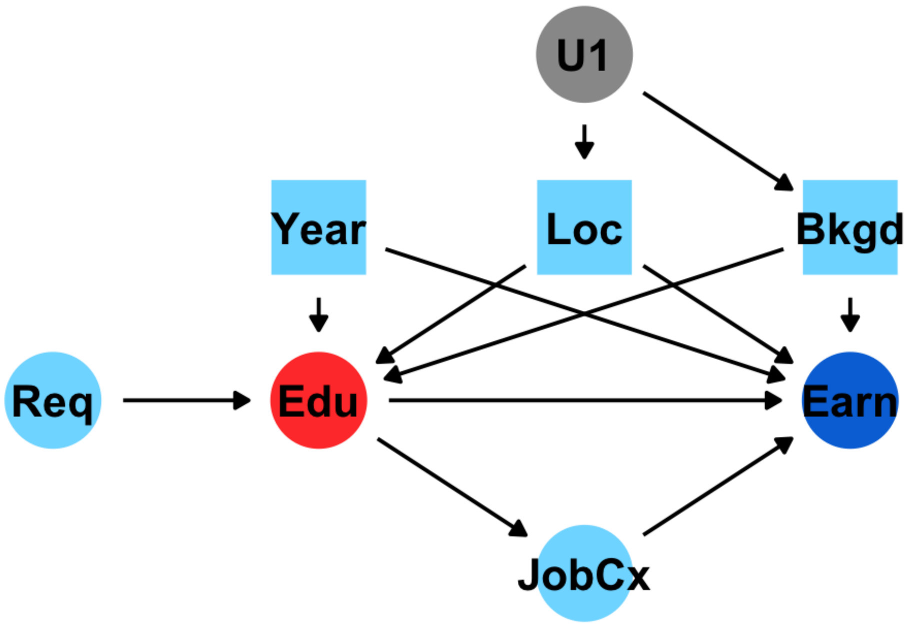
Difference-in-differences

Regression discontinuity

Instrumental variables

Model-based identification

Use a DAG and *do*-calculus to isolate arrow



Core assumption: selection on observables

Everything that needs to be adjusted is measurable; no unobserved confounding

Big assumption!

This is why lots of people don't like DAG-based adjustment

~~Prince~~ King Charles

- Male
- Born in 1948
- Raised in the UK
- Married twice
- Lives in a castle
- Wealthy & famous

Ozzy Osbourne

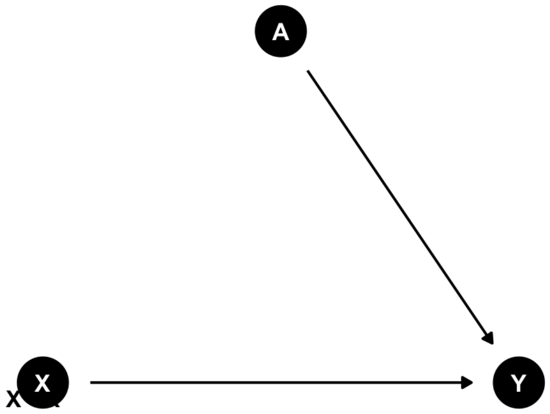
- Male
- Born in 1948
- Raised in the UK
- Married twice
- Lives in a castle
- Wealthy & famous

Design-based identification

Use a special situation to isolate arrow

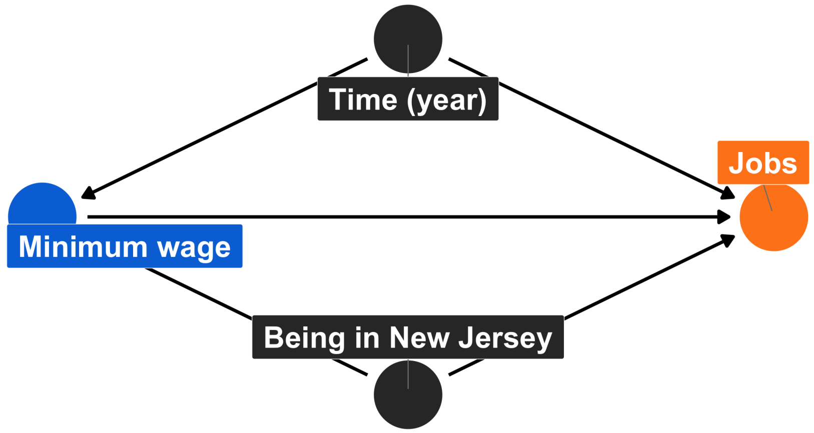
RCTs

Use randomization
to remove confounding

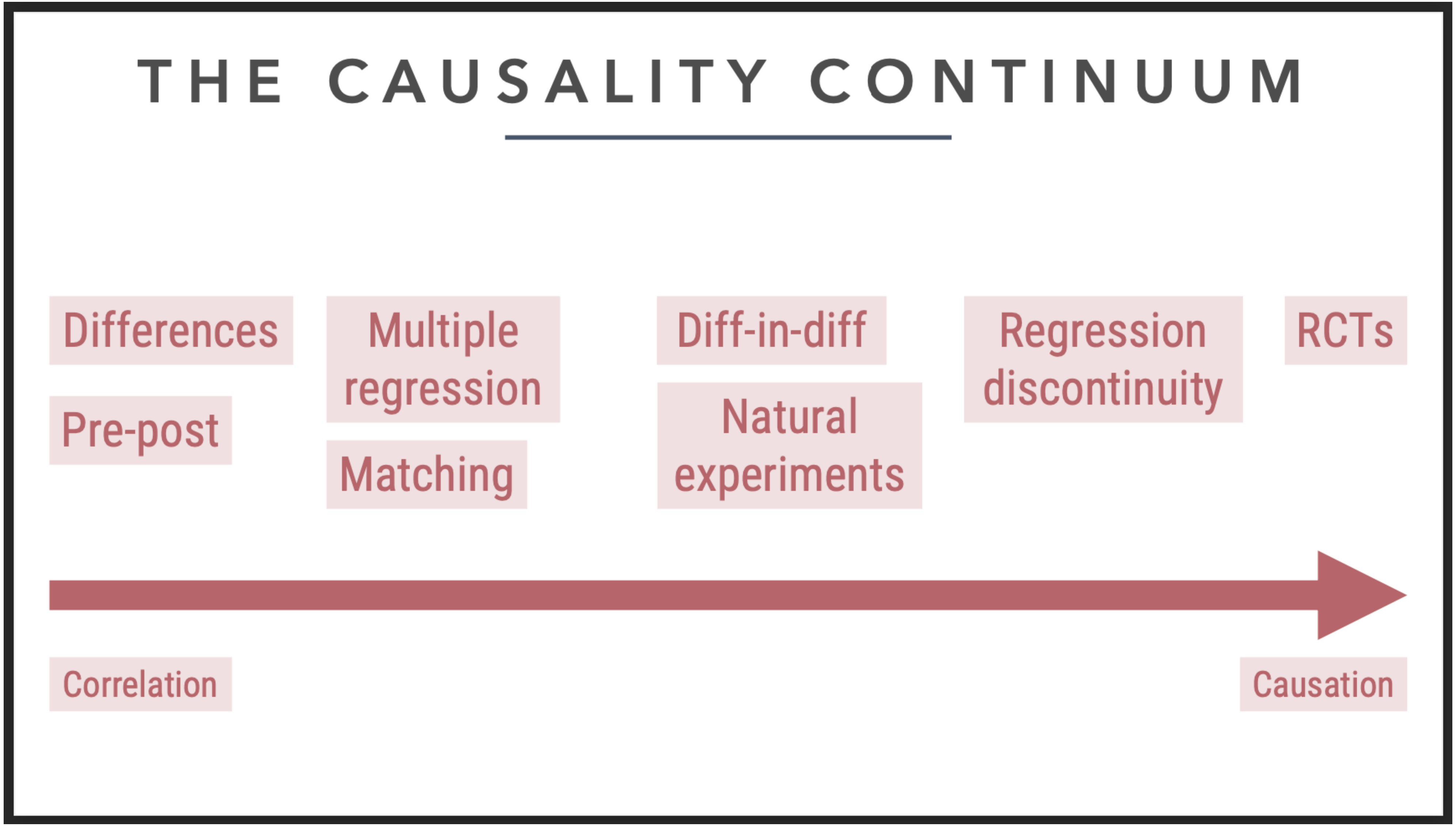


Difference-in-differences

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



Which is better or more credible? RCTs, quasi experiments, or DAG-based models?



THE CAUSALITY CONTINUUM
Differences
Pre-post
Multiple regression
Matching
Diff-in-diff
Natural experiments
Regression discontinuity
RCTs
Correlation
Causation

There's no hierarchy!

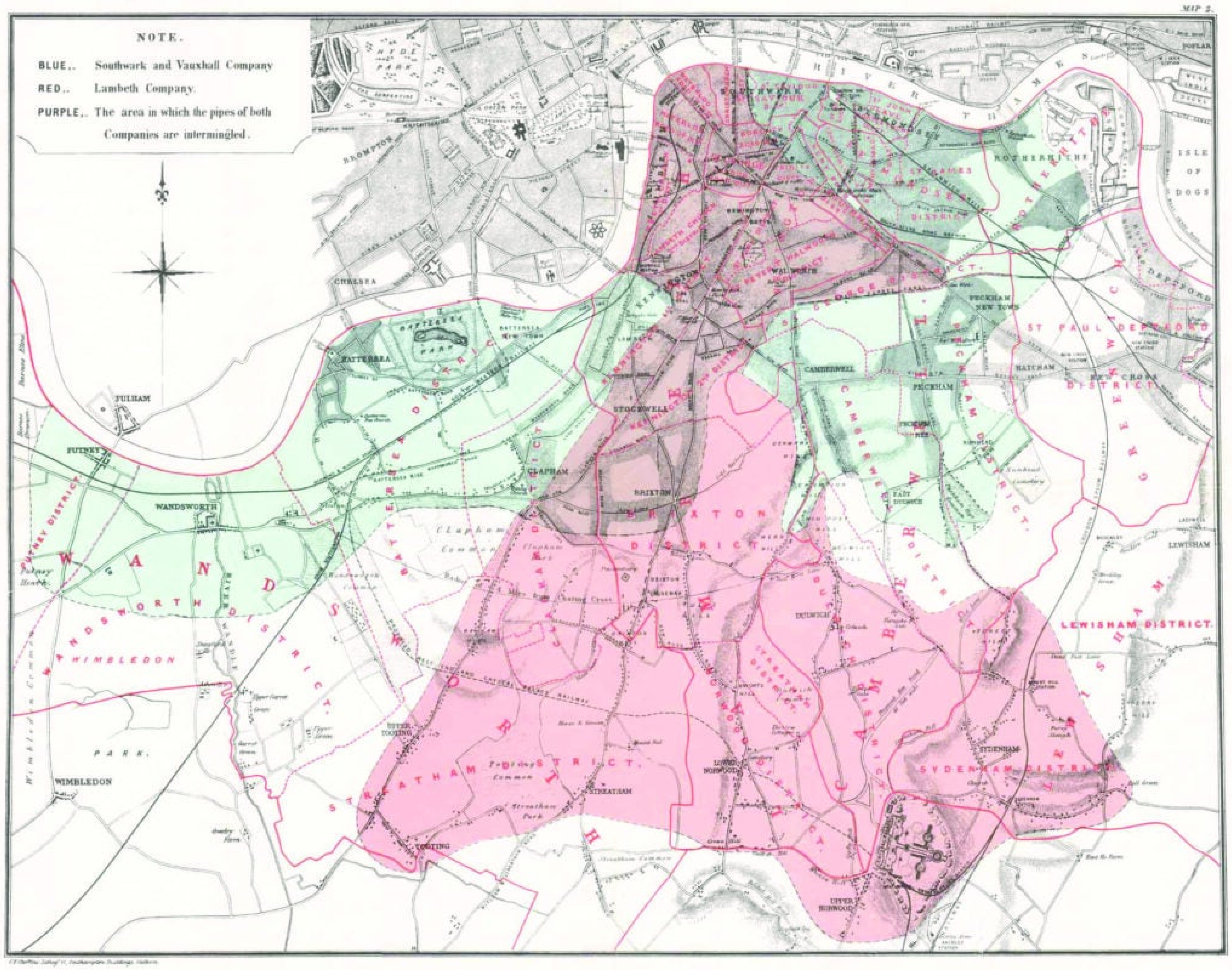
Interactions and regression

Can we talk more about interaction terms and how to interpret them?

Are interaction effects in regression always more accurate of a difference than running a "regular" regression without them?

Regression is just fancy averages!

Simple diff-in-diff


MAP 2.
NOTE.
BLUE,. Southwark and Vauxhall Company
RED.. Lambeth Company.
PURPLE,. The area in which the pipes of both Companies are intermingled.

1849

Cholera deaths per 100,000

Southwark & Vauxhall: 1,349

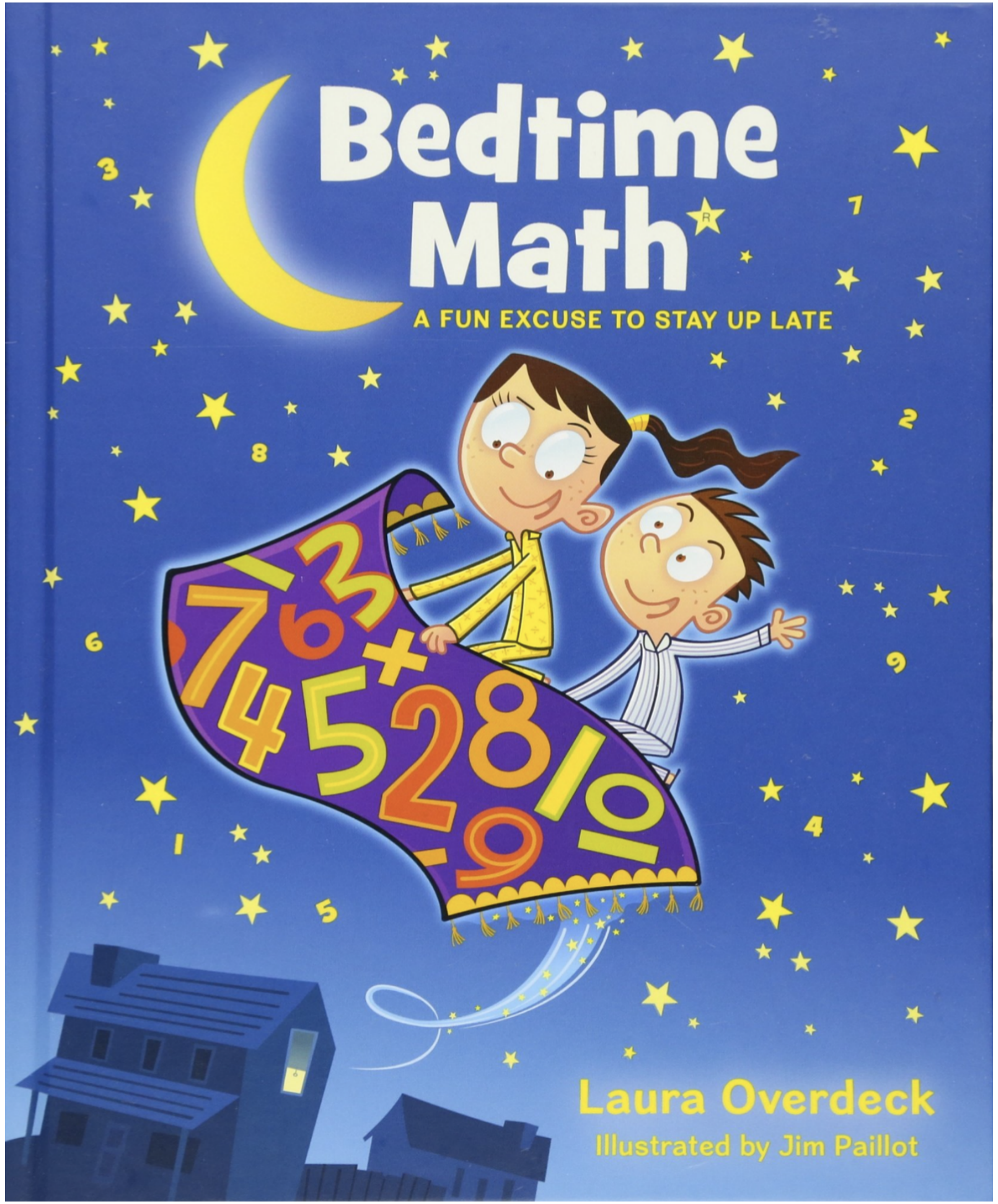
Lambeth: 847

1854

Cholera deaths per 100,000

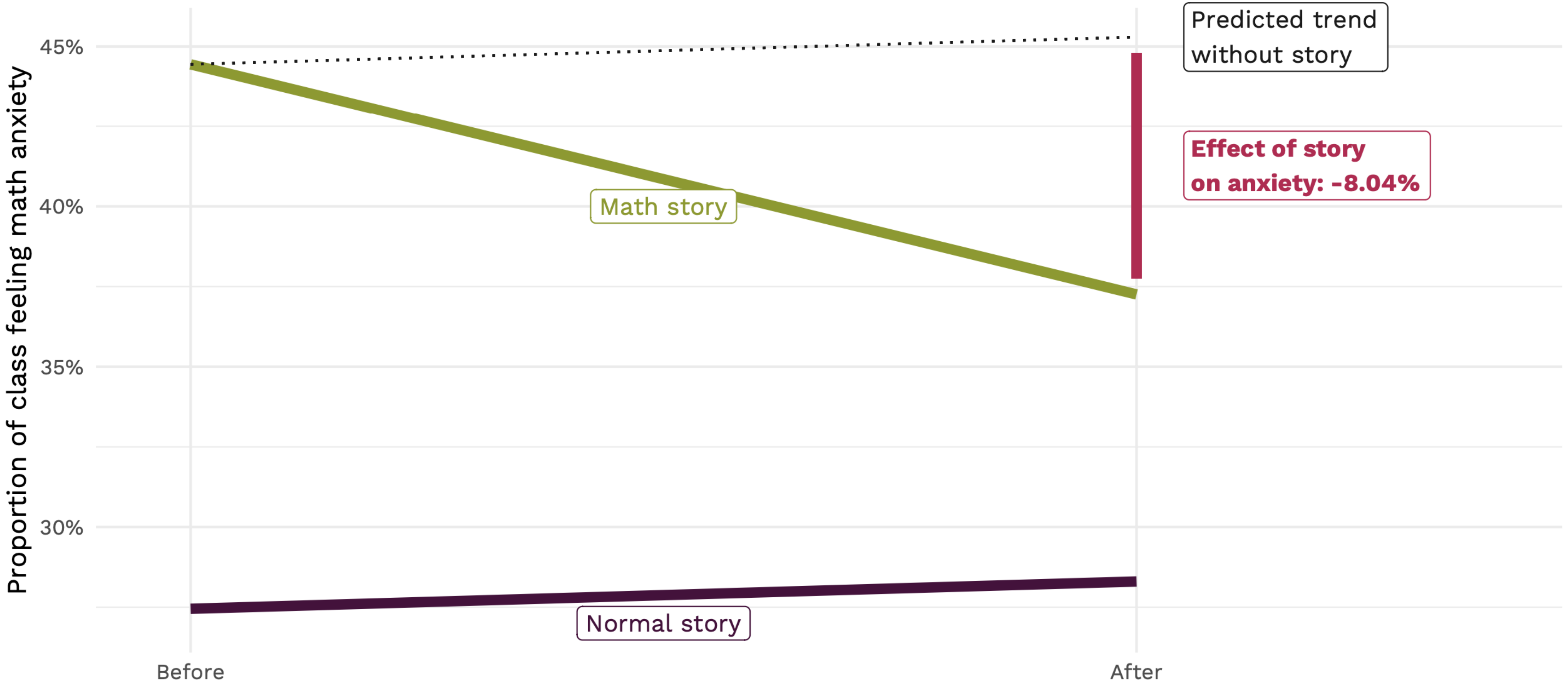
Southwark & Vauxhall: 1,466

Lambeth: 193


Bedtime Math
A FUN EXCUSE TO STAY UP LATE
Laura Overdeck
Illustrated by Jim Paillot

Reading a story about math reduces math anxiety

Experiment in four 4th grade classes



When doing your subtracting to get your differences in the matrix, is it better to do the vertical or horizontal subtractions?

Are there situations where one is preferable to the other?

Why are we learning two ways to do diff-in-diff? (2x2 matrix vs. `lm()`)

What happened to confounding??

Now we're only looking at just two "confounders"?

What group level is best for comparison? For example, if we are looking at policy change in NJ, is it best to compare with just one or two similar states? How similar do the populations need to be?

Wouldn't matching be better?

Do we have to think about balance when dealing with observational data in diff in diff?

Two-way fixed effects (TWFE)

Minimum legal drinking age

Figure 5.4
An MLDA effect in states with parallel trends

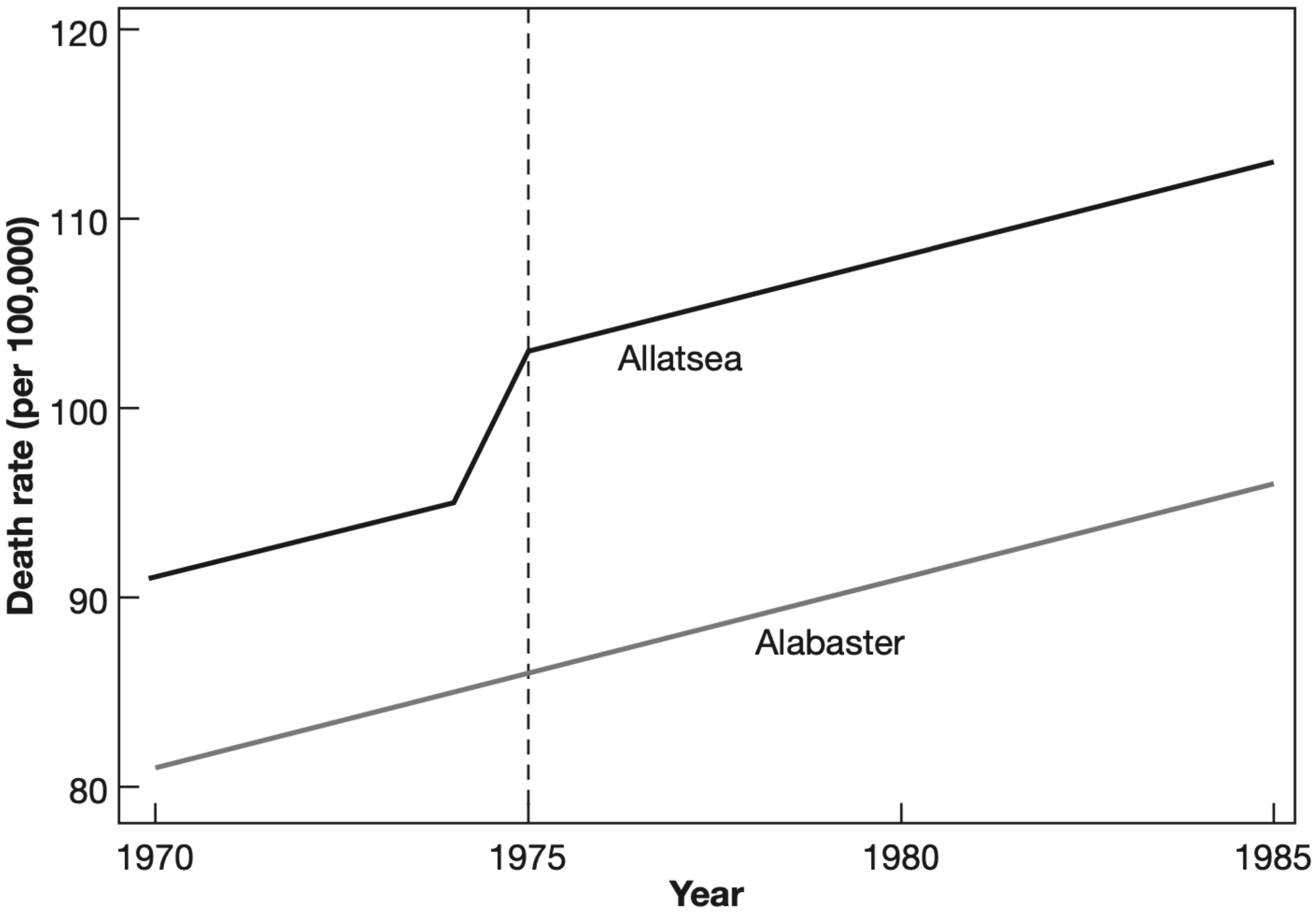


FIGURE 5.5

A spurious MLDA effect in states where trends are not parallel

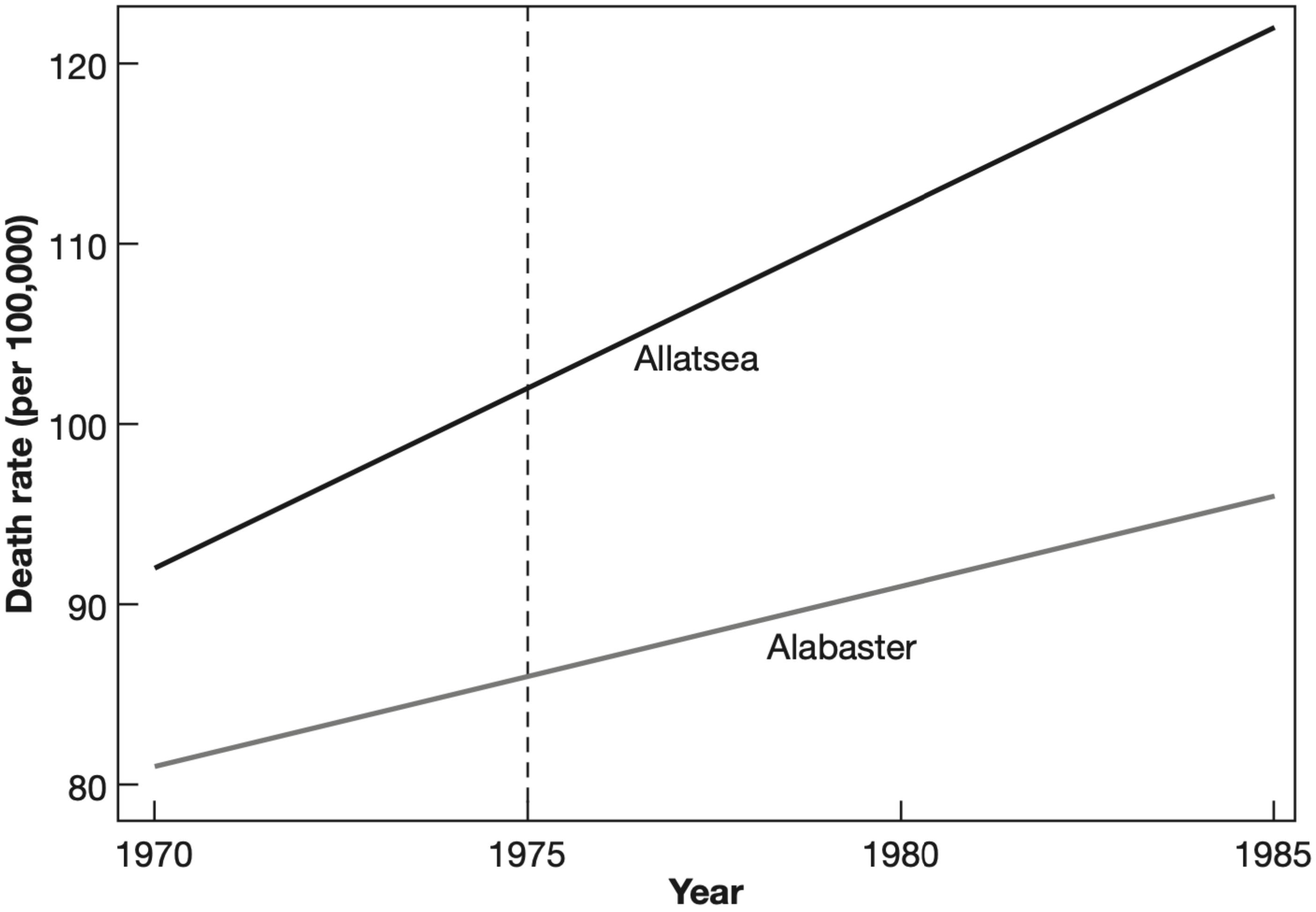
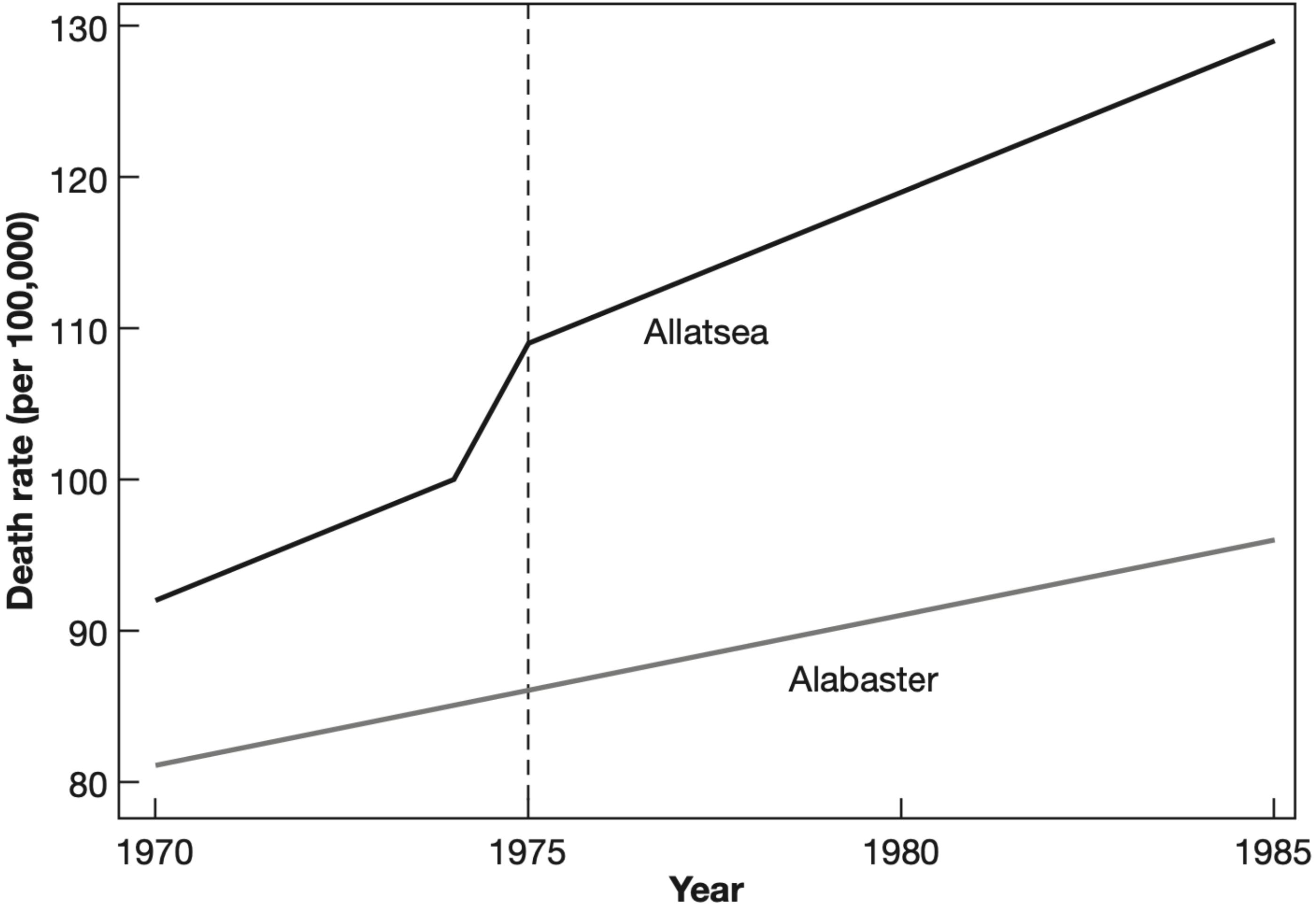


Figure 5.6
A real MLDA effect, visible even though trends are not parallel



MLDA reduction

Two states: Alabama vs. Arkansas

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

Organan donations

Two states: California vs. New Jersey

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

Two-way fixed effects (TWFE)

Two states: Alabama vs. Arkansas

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

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

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

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

vs.

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

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

vs.

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

vs.

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

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.

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

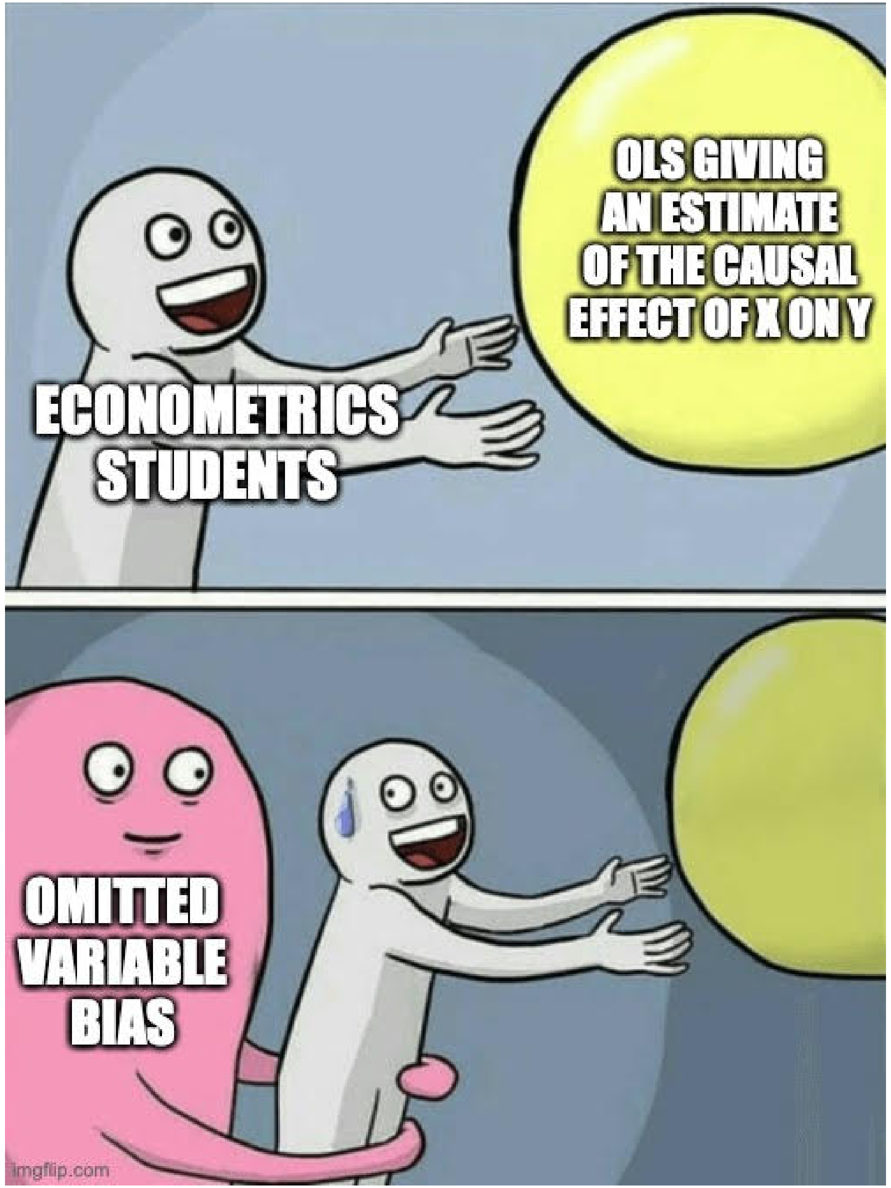
What about this staggered treatment stuff?

See this

Sensitivity analysis

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