In-person session 8

October 11, 2021

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

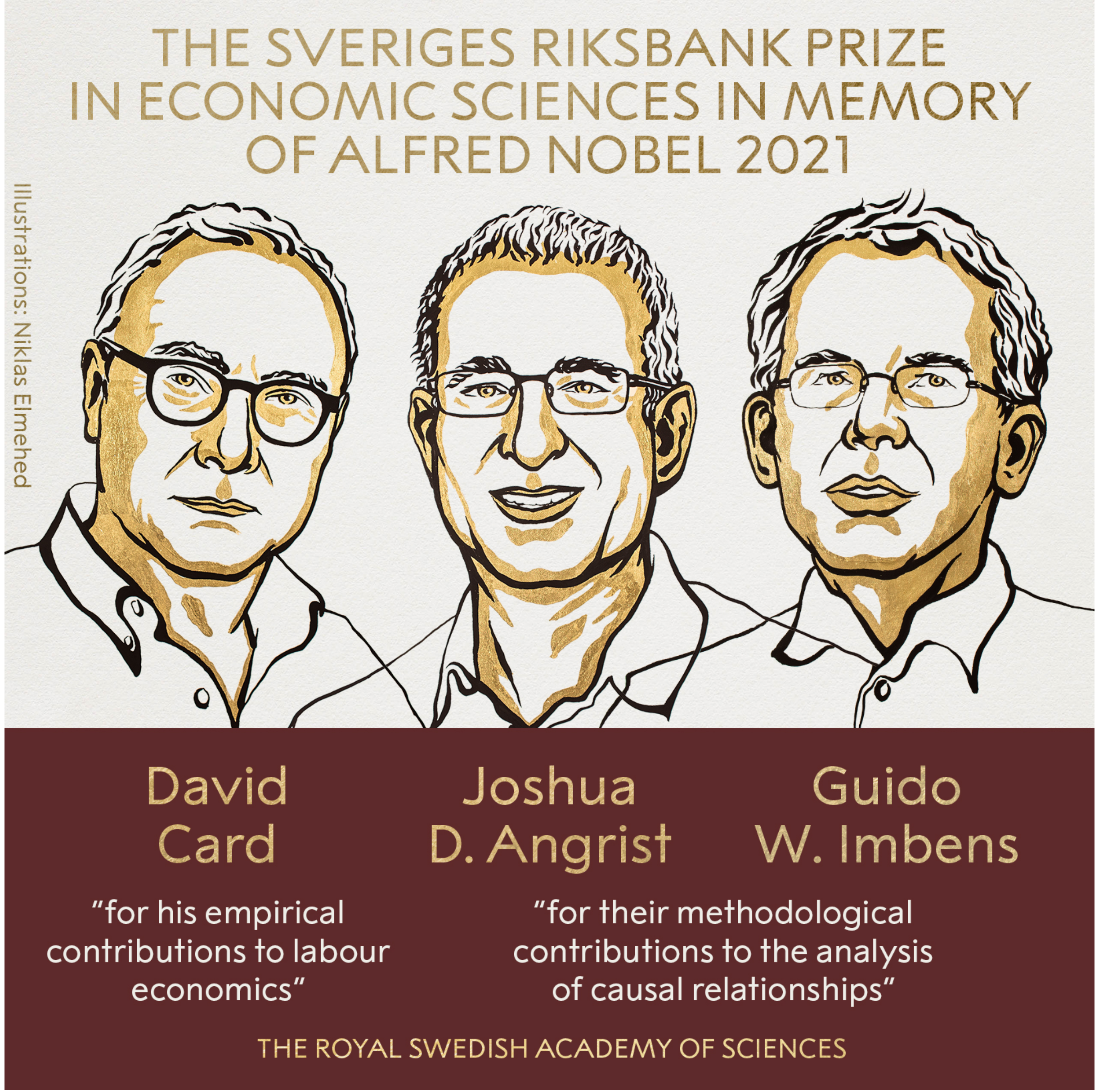
Plan for today

Econ Nobel!

Sensitivity analysis

Diff-in-diff FAQs

Econ Nobel

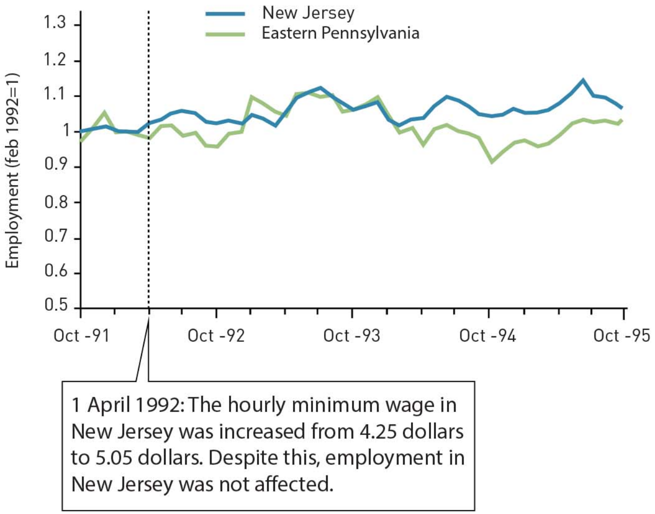

THE SVERIGES RIKSBANK PRIZE
IN ECONOMIC SCIENCES IN MEMORY
OF ALFRED NOBEL 2021
Illustrations: Niklas Elmehed
David
Card
Joshua
D. Angrist
Guido
W. Imbens
"for his empirical
contributions to labour
economics"
"for their methodological
contributions to the analysis
of causal relationships"
THE ROYAL SWEDISH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

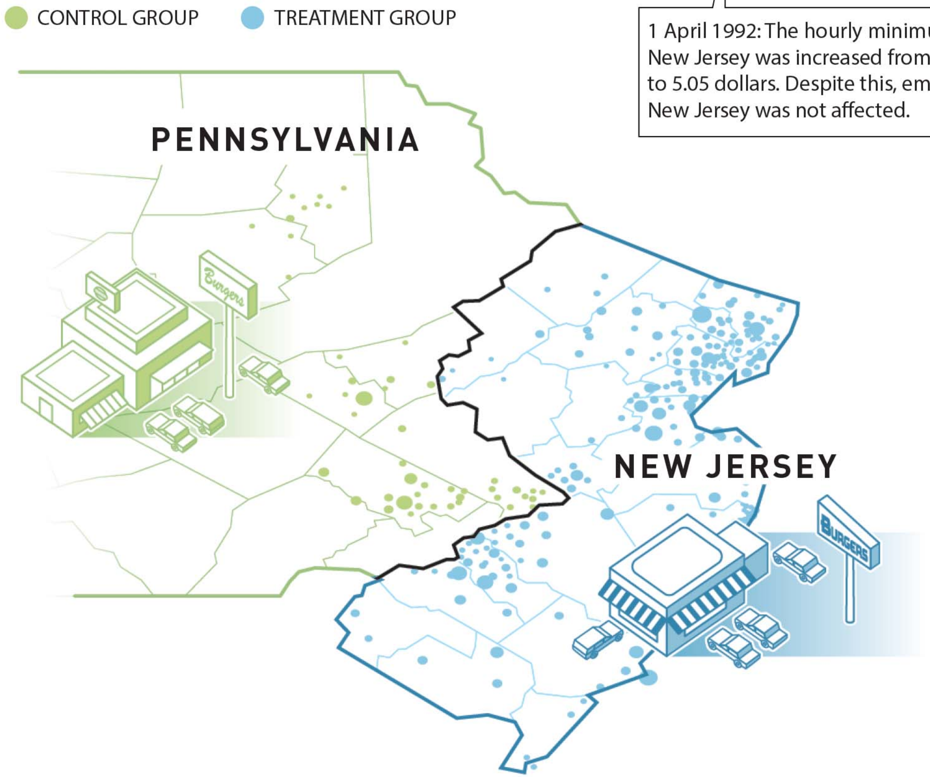

THE WHITE HOUSE
WASHINGTON

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.





Matt Blackwell
@matt_blackwell

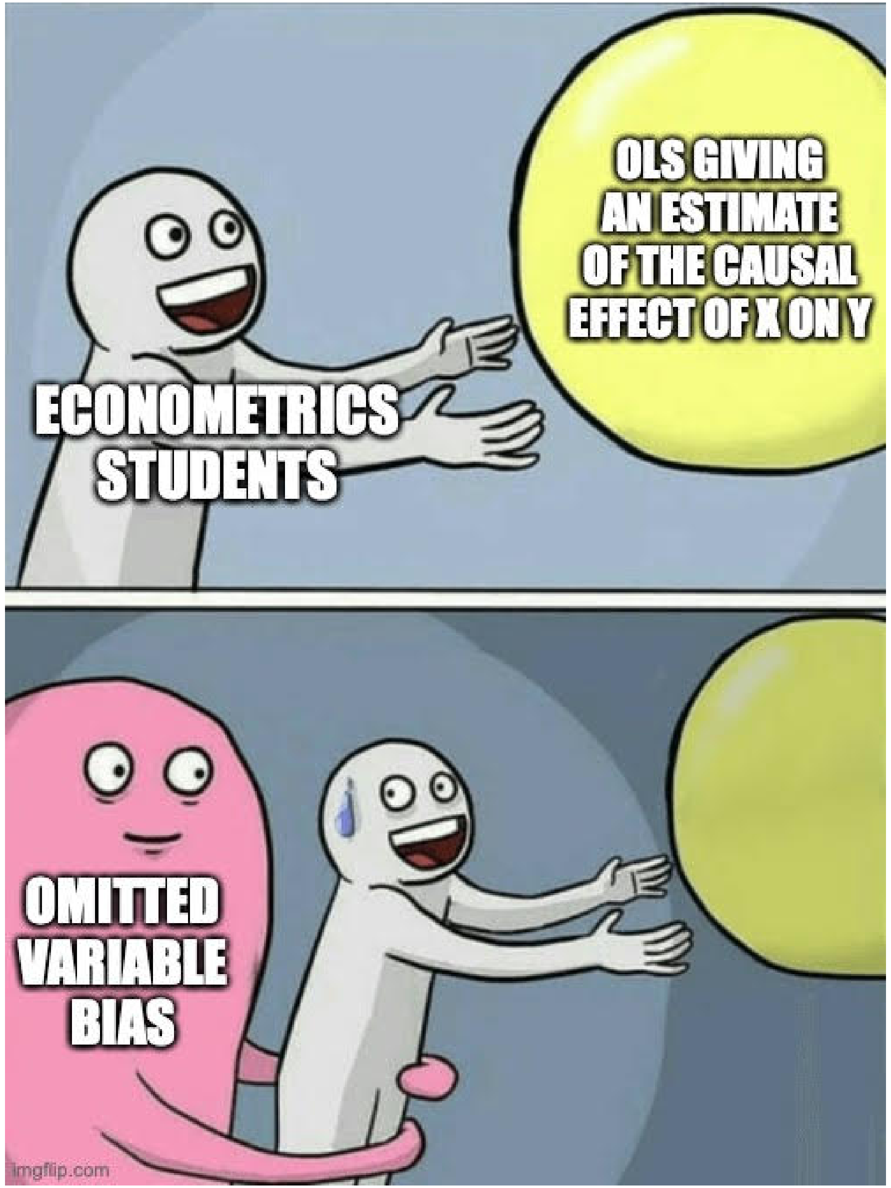
NPR reporter just said Card, Angrist, and Imbens won the Nobel for their analysis of “casual” relationships

7:04 AM · Oct 11, 2021 · Twitter for iPhone

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

Diff-in-diff FAQs

Design-based vs. model-based inference

Special situations vs. controlling for stuff

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

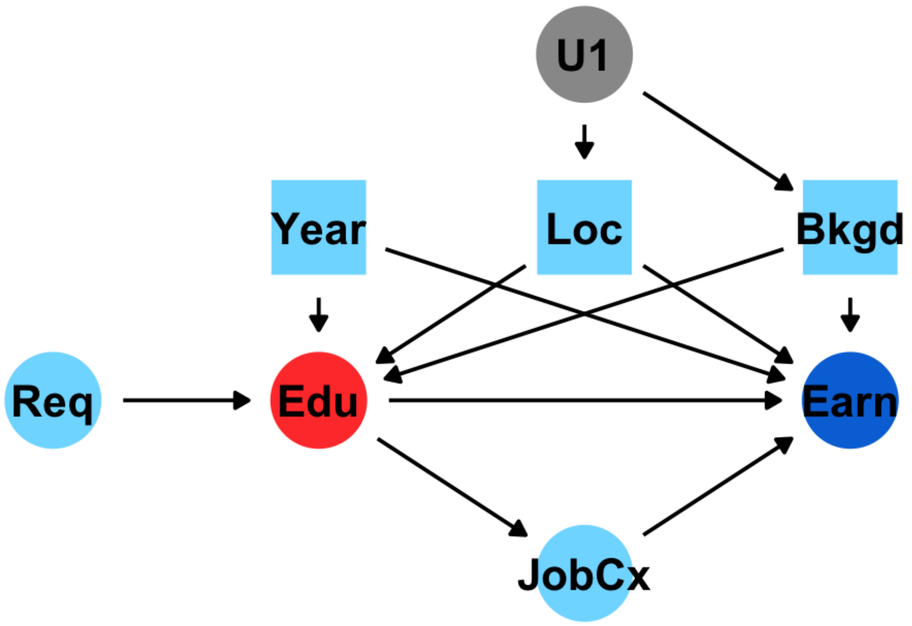
Difference-in-differences

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!

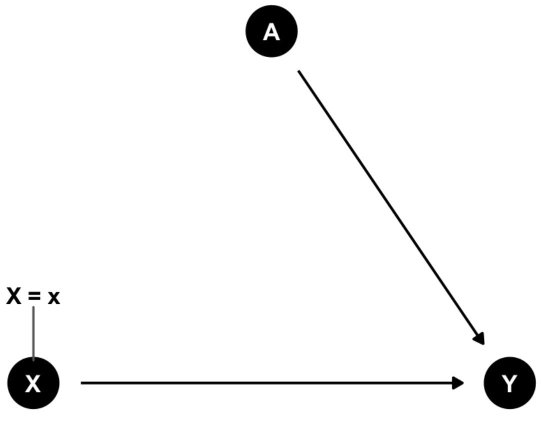
This is why lots of people don't like DAG-based adjustment

Design-based identification

Use a special situation to isolate arrow

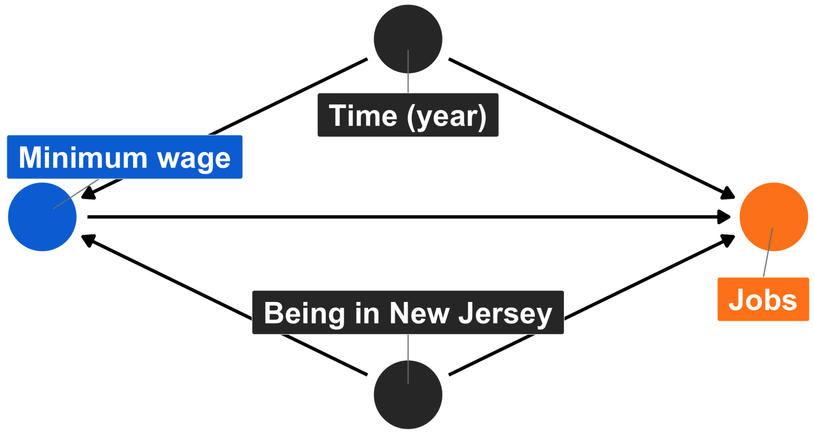
RCTs

Use randomization
to remove confounding

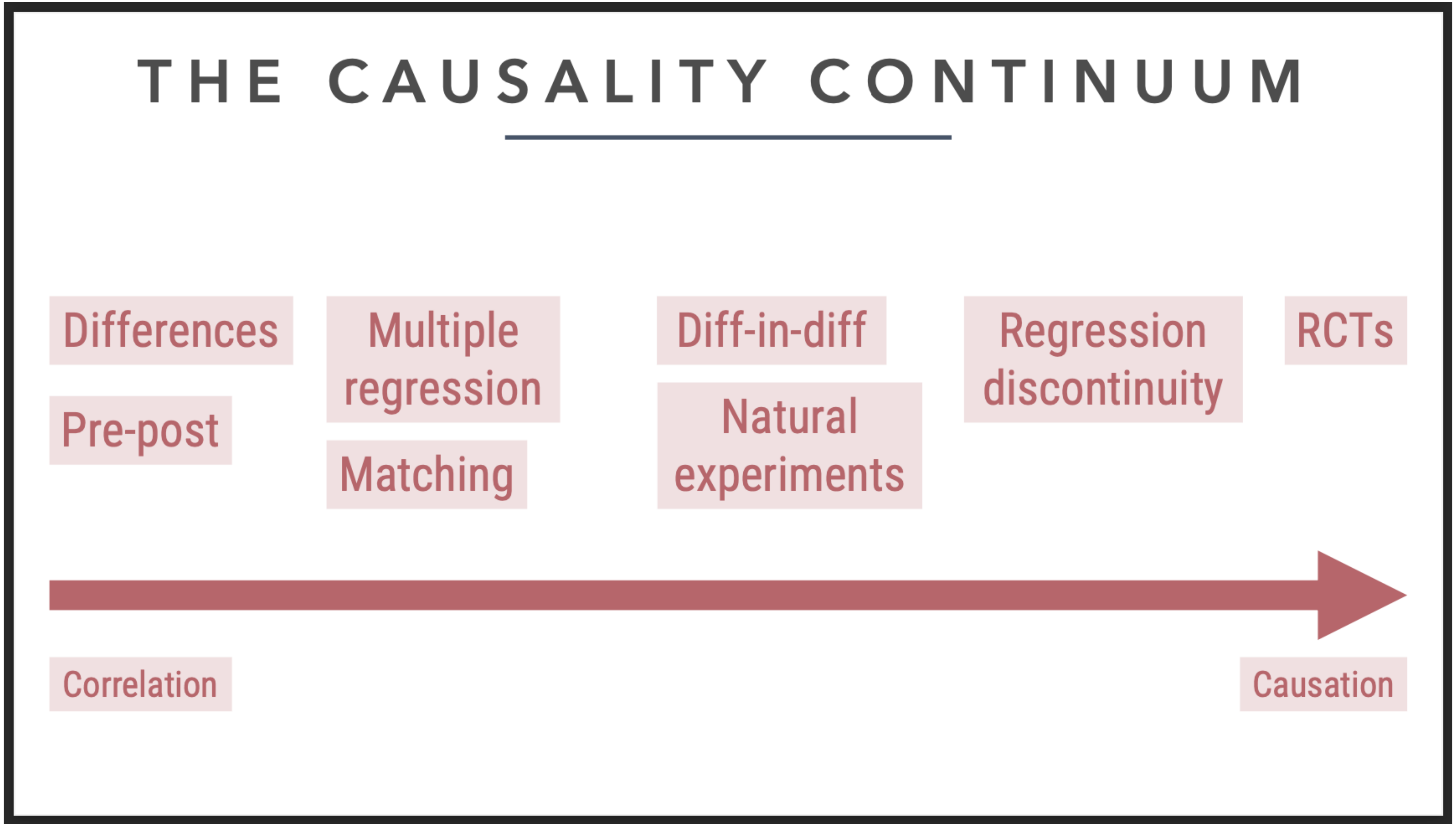


Difference-in-differences

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



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

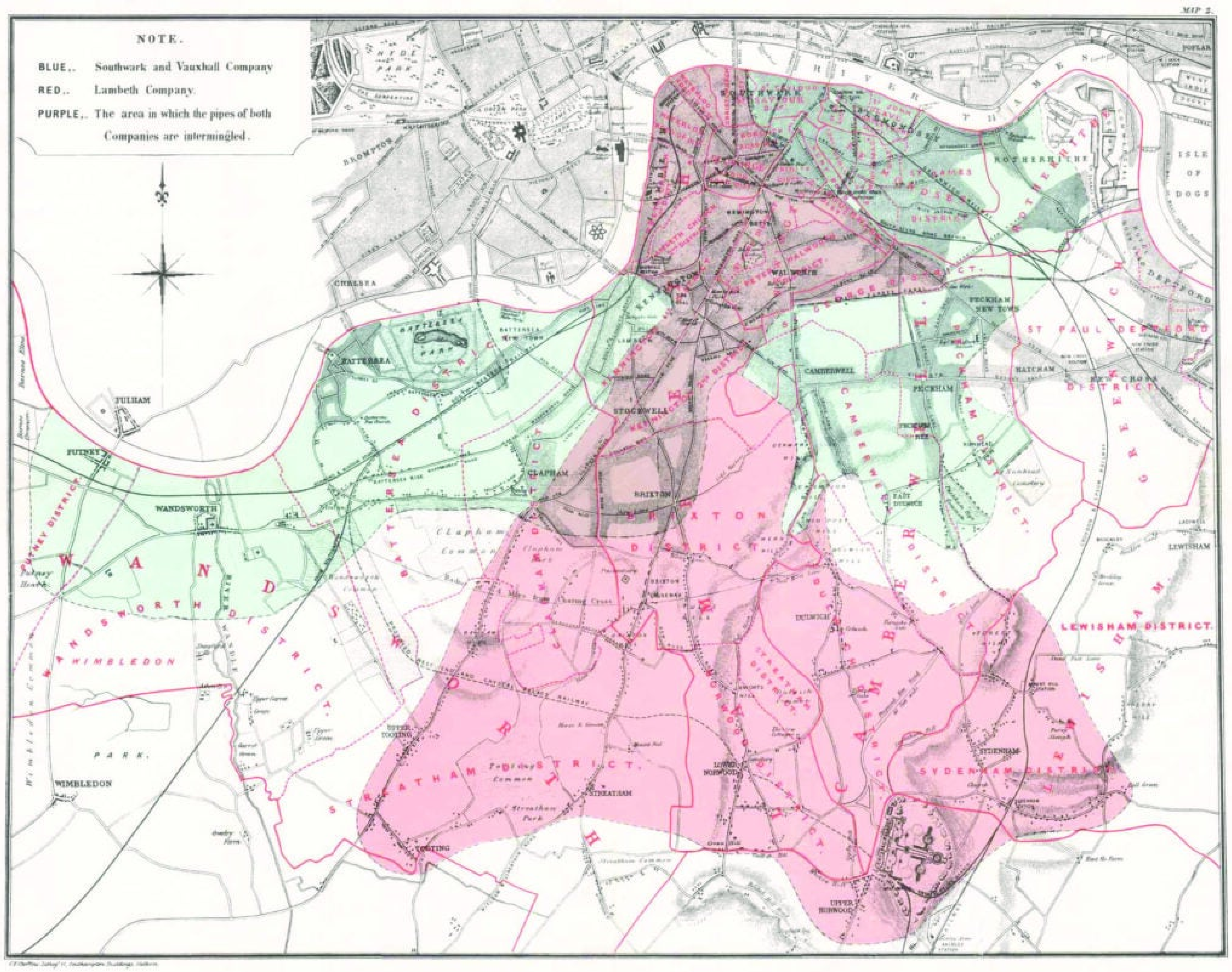

THE CAUSALITY CONTINUUM
Differences
Pre-post
Multiple regression
Matching
Diff-in-diff
Natural experiments
Regression discontinuity
RCTs
Correlation
Causation

There's no hierarchy!

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?

Can causal effects be negative or are they always positive?


MAP 2.
NOTE.
BLUE,. Southwark and Vauxhall Company
RED.. Lambeth Company.
PURPLE,. The area in which the pipes of both Companies are intermingled.
RIVER THAMES
ISLE OF DOGS
CHELSEA
BATTERSEA
FULHAM
PUTNEY
WANDSWORTH
CLAPHAM
STOCKWELL
BRIXTON
KENNINGTON
WALWORTH
CAMBERWELL
PECKHAM
ROTHERHITHE
HATCHAM
LEWISHAM
LEWISHAM DISTRICT.
ST PAUL DEPTFORD
DULWICH
STREATHAM
TOOTING
UPPER TOOTING
WIMBLEDON
SYDENHAM
LOWER NORWOOD
UPPER NORWOOD

1849

Cholera deaths per 100,000

Southwark & Vauxhall: 1,349

Lambeth: 847

1854

Cholera deaths per 100,000

Southwark & Vauxhall: 1,466

Lambeth: 193

Multiple adjustment sets

Where do we get all this data?

lolz

Data resources

See this

REPORT

FILE NOT FOUND

A generation that grew up with Google is forcing professors to rethink their lesson plans

By Monica Chin | @mcsquared96 | Sep 22, 2021, 8:00am EDT

Illustrations by Micha Huigen

Project structures

One approach

Another approach

Yet another approach

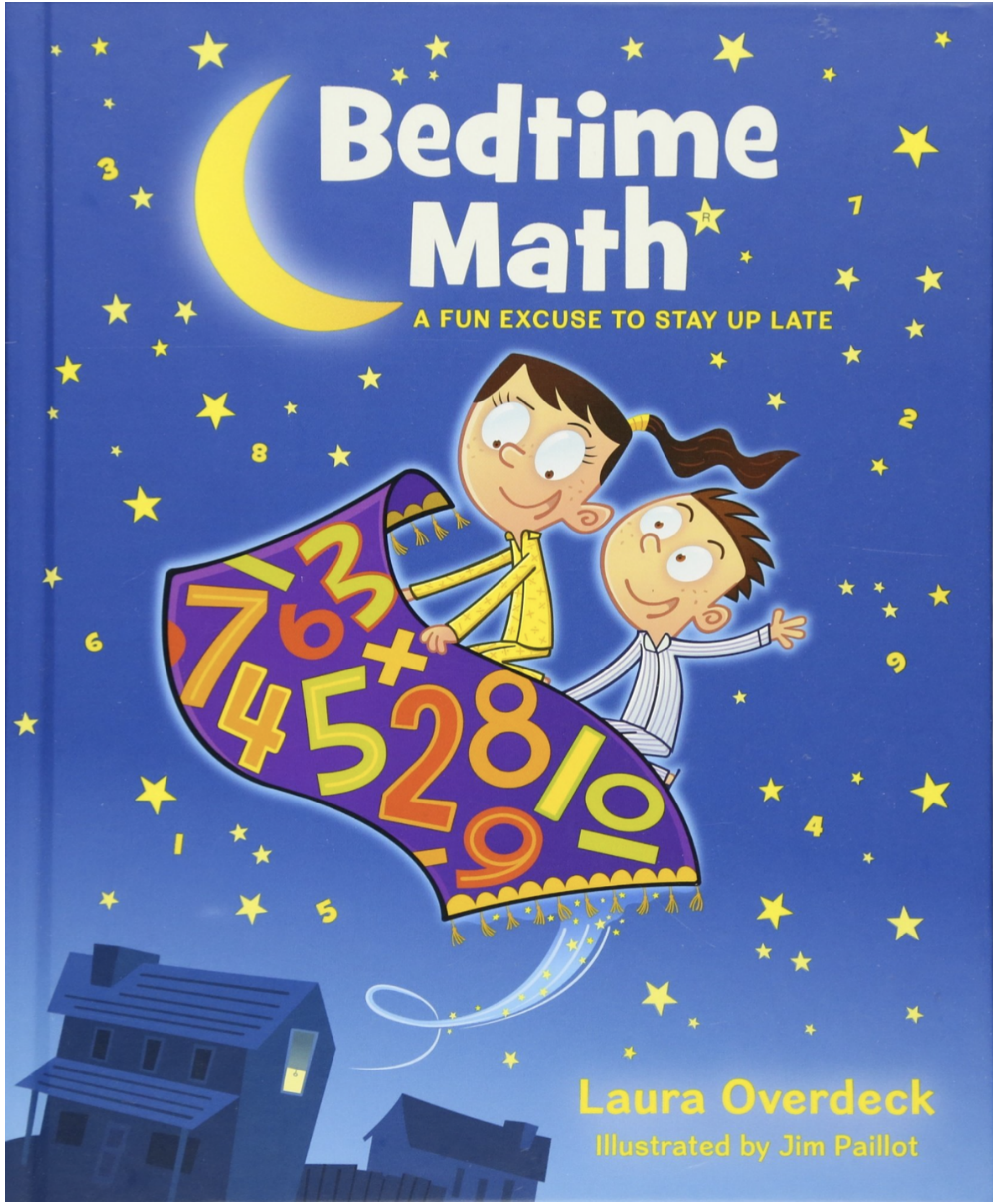
Another another approach

File types

Image types slides

CSV vs. Excel

`.docx` VS. `.txt` VS. `.md` VS `.Rmd`


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

If the control group changes in the same way, and the causal effect was zero, would we say that the treatment didn't work?

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 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

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

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

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

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.

Figure 5.4
An MLDA effect in states with parallel trends

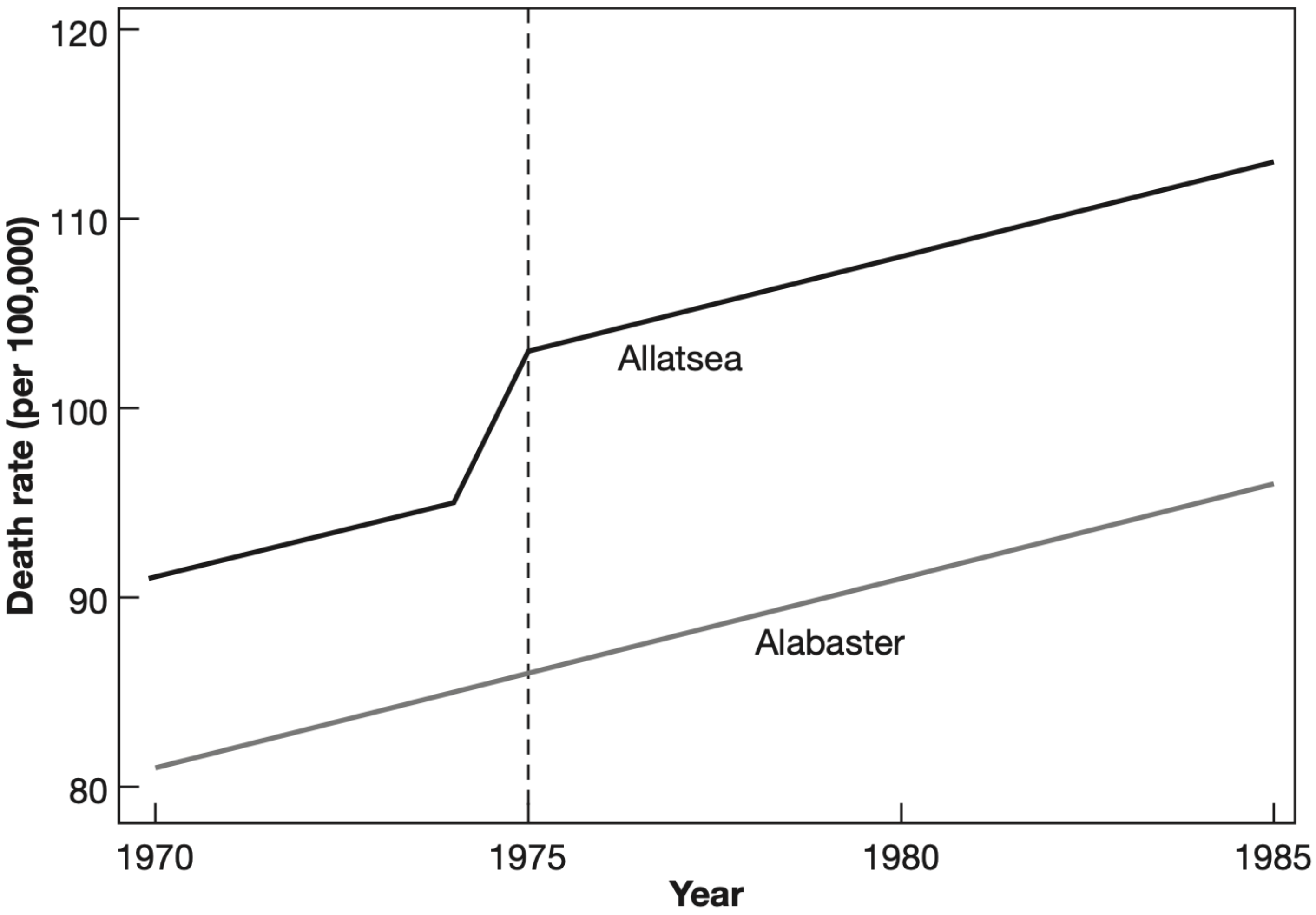


Figure 5.5
A spurious MLDA effect in states where trends are not parallel

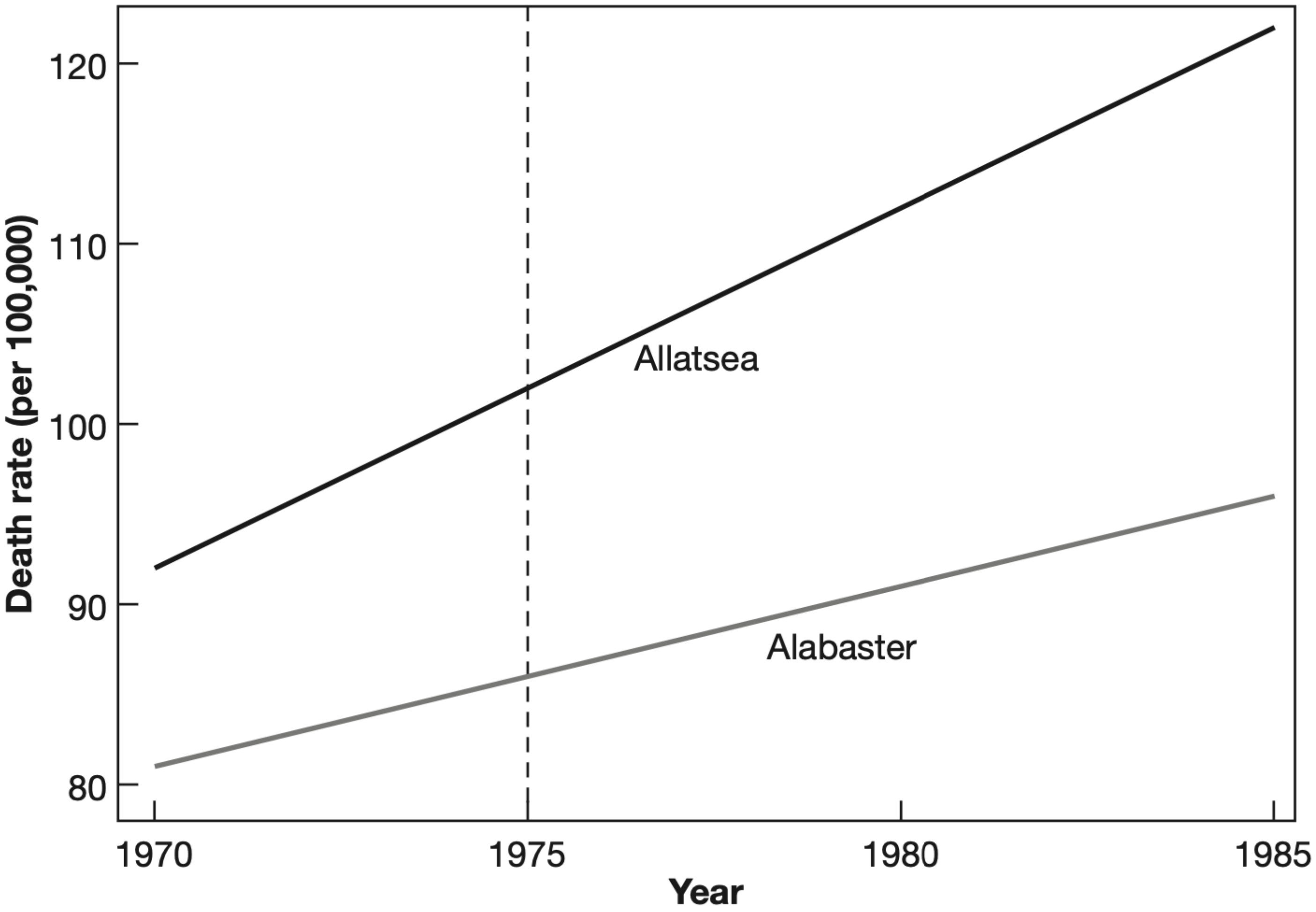
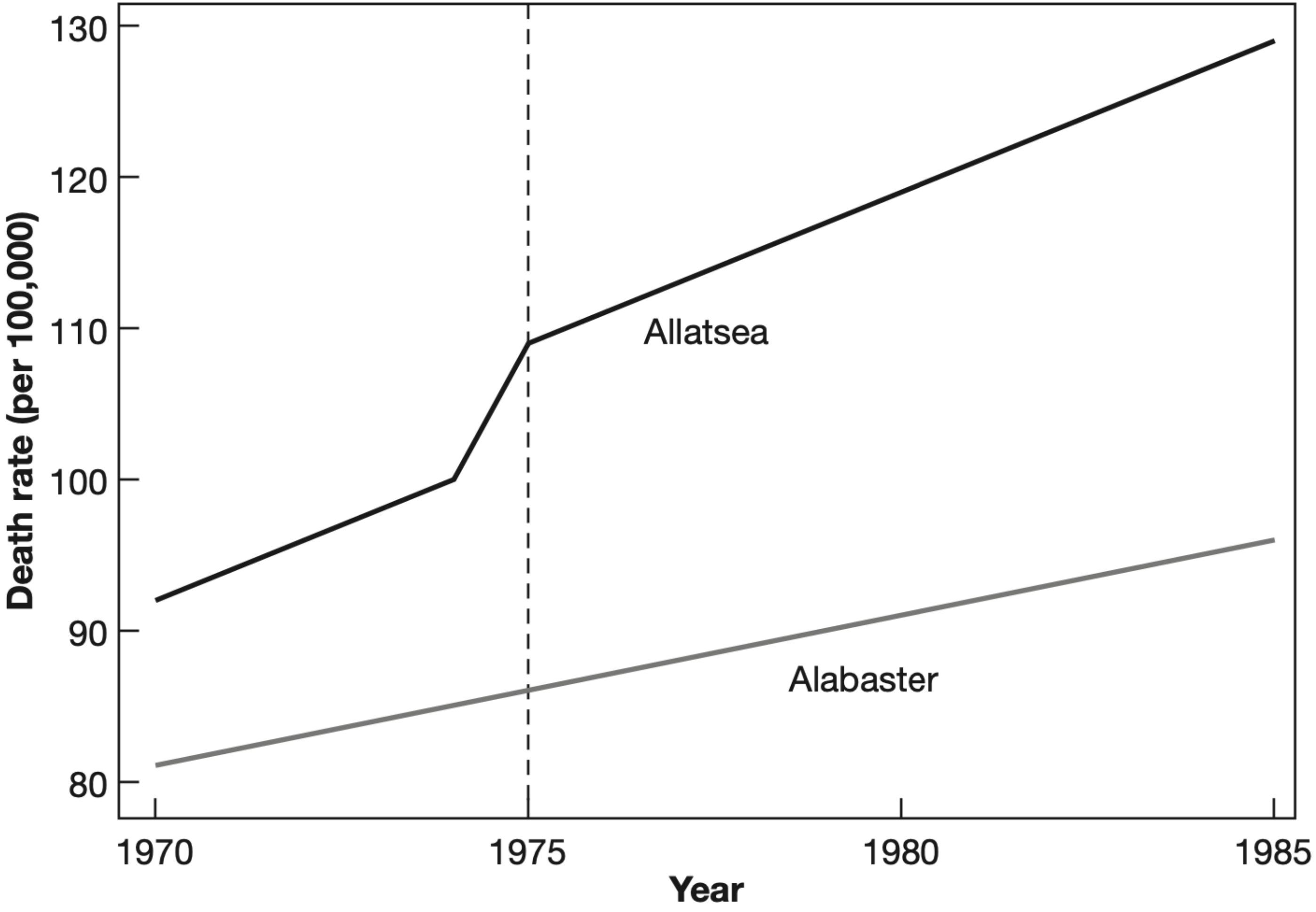


Figure 5.6
A real MLDA effect, visible even though trends are not parallel



What happened to confounding??

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

Is it reasonable to conduct sensitivity analysis when working with diff in diff?

How do we play with time to check for parallel trends?

What about this staggered treatment stuff?

See this