

Introduction to Empirical Economics

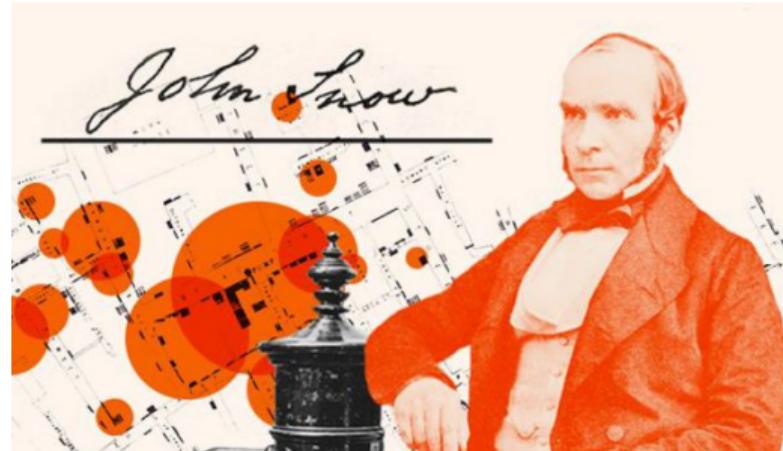
Difference-in-Differences & Regression Discontinuity

Léo Zabrocki

leo.zabrocki@psemail.eu

<https://lzabrocki.github.io/>

École Normale Supérieure



Program

Today's course on RDD and DiD

One course on the replication crisis

Then courses on specific topics

Program

Today's course on RDD and DiD

One course on the replication crisis

Then courses on specific topics

Program

Today's course on RDD and DiD

One course on the replication crisis

Then courses on specific topics

Program

Today's course on RDD and DiD

One course on the replication crisis

Then courses on specific topics

Materials of the course

[https://github.com/lzabrocki/
empirical_economics](https://github.com/lzabrocki/empirical_economics)

Today's slides are heavily based on
Causal Mixtape, Thomas S. Coleman's
slides and *Mastering Metrics*

Who can summarize the previous class?

Difference-in-Differences Design

Cholera Epidemic in 19th-century London

1831-32: 6,526 deaths

1849: 14,137 deaths

1853-54: 10,738 deaths



Cholera Epidemic in 19th-century London

1831-32: 6,526 deaths

1849: 14,137 deaths

1853-54: 10,738 deaths



Cholera Epidemic in 19th-century London

1831-32: 6,526 deaths

1849: 14,137 deaths

1853-54: 10,738 deaths



Cholera Epidemic in 19th-century London

1831-32: 6,526 deaths

1849: 14,137 deaths

1853-54: 10,738 deaths



Cholera Epidemic in 19th-century London

1831-32: 6,526 deaths

1849: 14,137 deaths

1853-54: 10,738 deaths



Mysterious Origin

Bad air

Bad breeding

Bad ground



A COURT FOR KING CHOLERA.

Mysterious Origin

Bad air

Bad breeding

Bad ground



A COURT FOR KING CHOLERA.

Mysterious Origin

Bad air

Bad breeding

Bad ground



A COURT FOR KING CHOLERA.

Mysterious Origin

Bad air

Bad breeding

Bad ground



A COURT FOR KING CHOLERA.

Mysterious Origin

Bad air

Bad breeding

Bad ground



What is Cholera?

Bacterium

Severe diarrhea

Source: contaminated water



Vibrio cholerae

Victorian London



The silent highwayman

Hero of the day: John Snow

Physician (1813-1858)

Fecal transmission through water

Very clever demonstration

No mention of cholera in his obituary



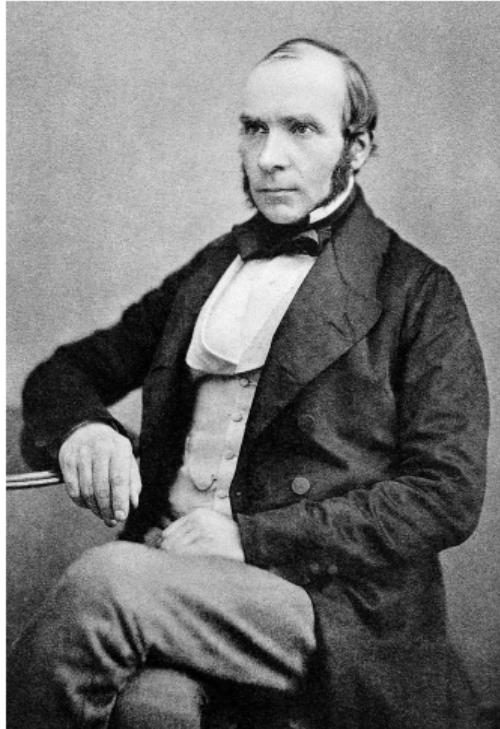
Hero of the day: John Snow

Physician (1813-1858)

Fecal transmission through water

Very clever demonstration

No mention of cholera in his obituary



Medical Opinion of the Time

Miasma

Snow at first accepted the theory

Tried to block these airborne poisons

Even used burlap bags!



Medical Opinion of the Time

People kept getting sick and dying

He changed his mind

...and began look for a new explanation



Puzzling Anecdotes

A sailor on a ship would not get sicked if he remained docked

Two apartment buildings next to one another

Several evidences question quality of the water

A Quasi-Natural Experiment

In the 1800s, several water companies served different areas of London

Took their water from the Thames

Treatment: in 1852, the Lambeth water company had moved its intake pipes upstream

Control: Southwark and Vauxhall Waterworks Company

A Quasi-Natural Experiment

Snow spent an entire book documenting similarities:

In many cases a single house has a supply different from that on either side. Each company supplies both rich and poor, both large houses and small; there is no difference in the condition or occupation of the persons receiving the water of the different companies

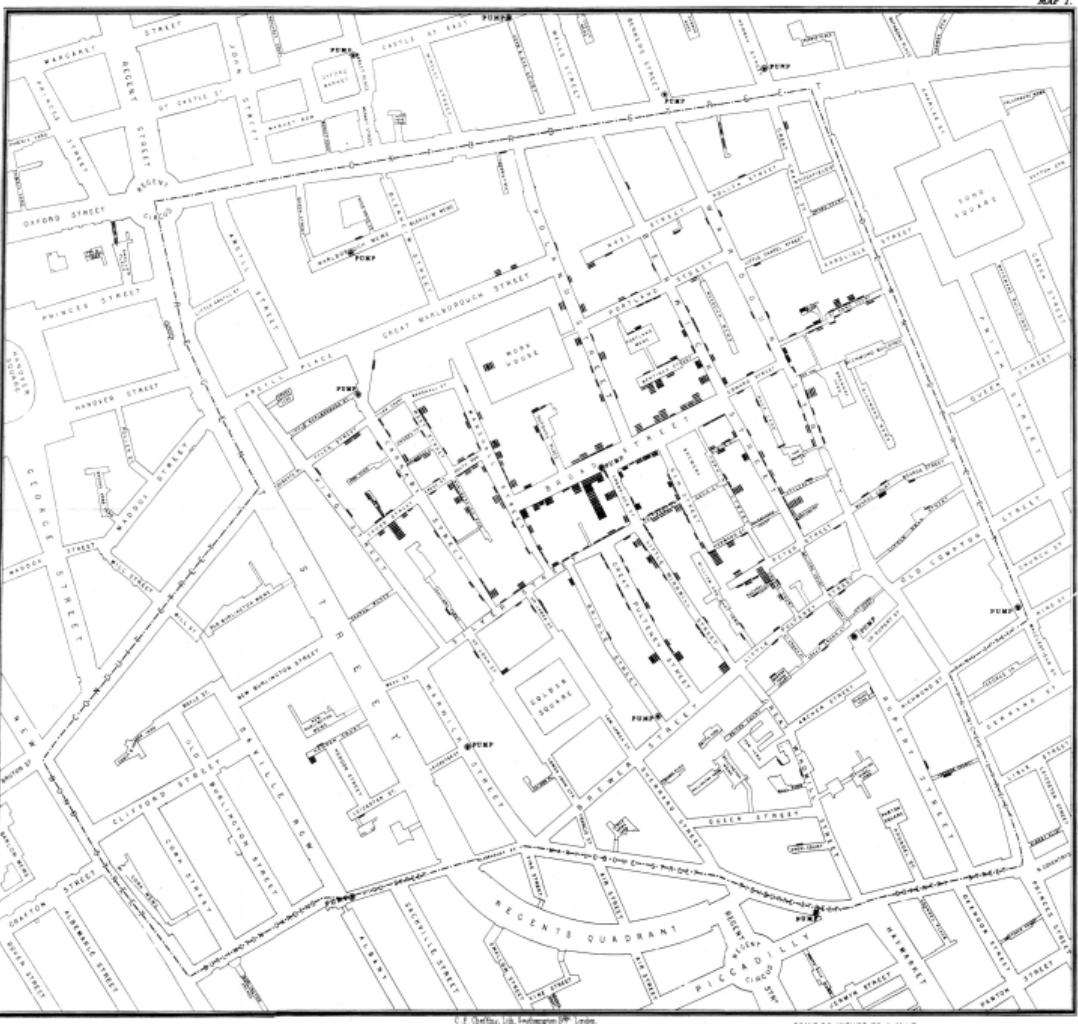
A Quasi-Natural Experiment

Incredible effort to collect data:

Door to door asking household heads the name of their utility company

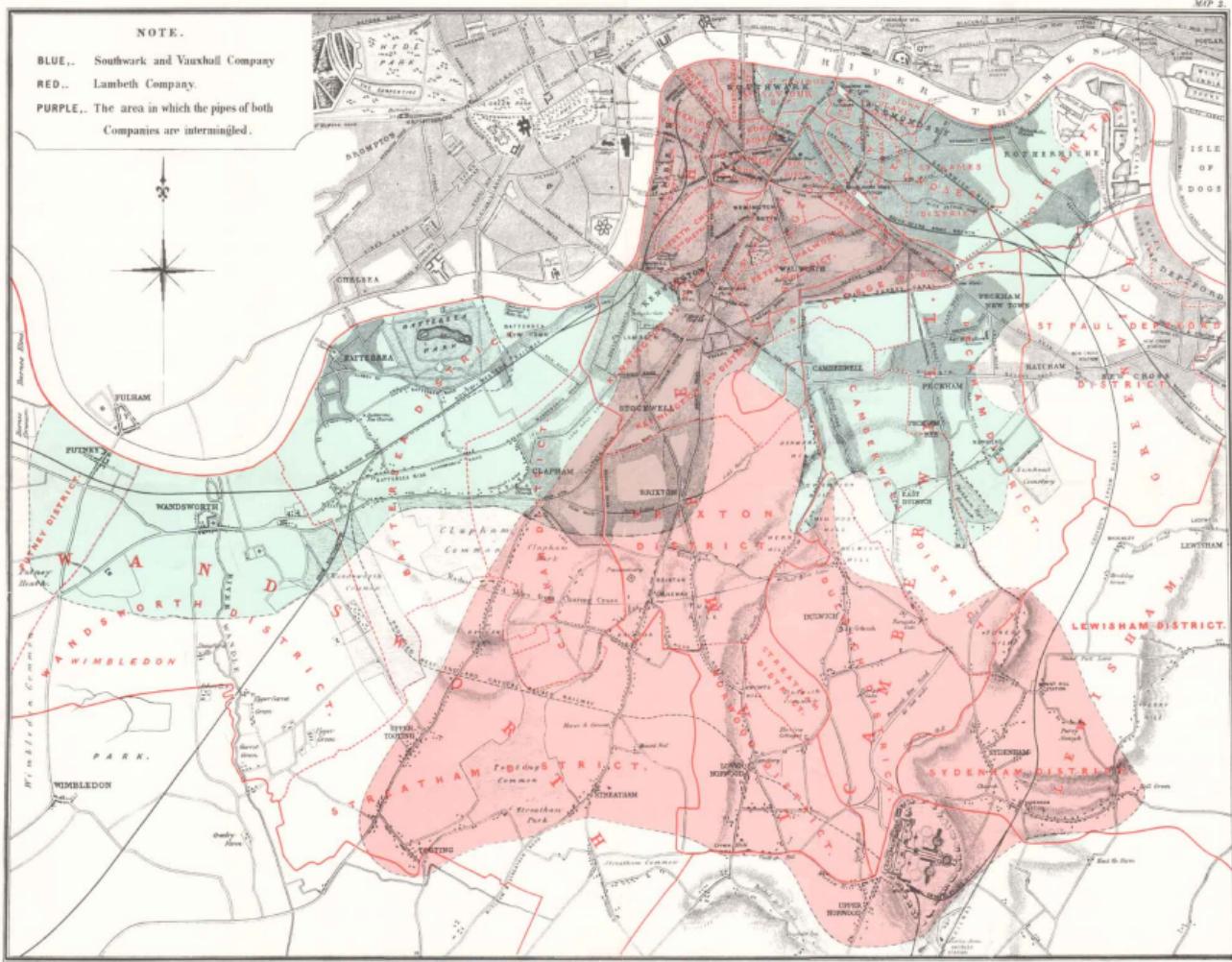
If people did not remember, he used a saline test!

Matched those data with the city's data on the cholera death rates at the household level



NOTE.

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



Data

Company name	1849	1854
Southwark and Vauxhall	135	147
Lambeth	85	19

Cholera Deaths per 10,000

Set-Up

D : clean water

Y : number of cholera deaths

Can we identify the causal effect of D if we just compare the post-treatment 1854 Lambeth cholera death values to that of the 1854 Southwark and Vauxhall values?

Warning

A simple difference in outcome only collapsed to the ATE if the treatment had been randomized!

Introducing Fixed Effects

Company	Outcome	1854
Lambeth	$Y=L+D$ $Y=L+D$	147
Southwark and Vauxhall	$Y=SV$	19

L and SV : time-invariant unique differences between the two companies as it relates to cholera deaths in 1854

Outcomes difference: $D + (L - SV)$

$L - SV$: selection bias!

Before-and-After Difference

Company	Time	Outcome
Lambeth	Before	$Y=L$
	After	$Y=L+(T+D)$

This procedure successfully eliminates the Lambeth fixed effect

But does not eliminate T , the natural changes in the cholera deaths over time

A Difference-in-Differences

Companies	Time	Outcome	First \neq	Second \neq
Lambeth	Before	$Y=L$		
	After	$Y=L+T+D$	$T+D$	D
Southwark and Vauxhall	Before	$Y=SV$		
	After	$Y=SV+T$	T	

The first difference eliminates the unit-specific fixed effects

The difference the differences is then an unbiased estimate of D

But what is the key assumption for this method to work?

A Key Assumption

We are assuming that there is no time-variant company specific unobservables

Nothing unobserved in Lambeth households that is changing between these two periods that *also* determines cholera deaths.

Parallel trends assumption

What is the Causal Effect of Clean Water?

Company name	1849	1854
Southwark and Vauxhall	135	147
Lambeth	85	19

Cholera Deaths per 10,000

A Powerful Design

Using repeated observations on a treatment and control unit (usually several units)

We can eliminate the unobserved heterogeneity

Widely used in economics, social sciences and public health



More Theory

The 2×2 DD design:

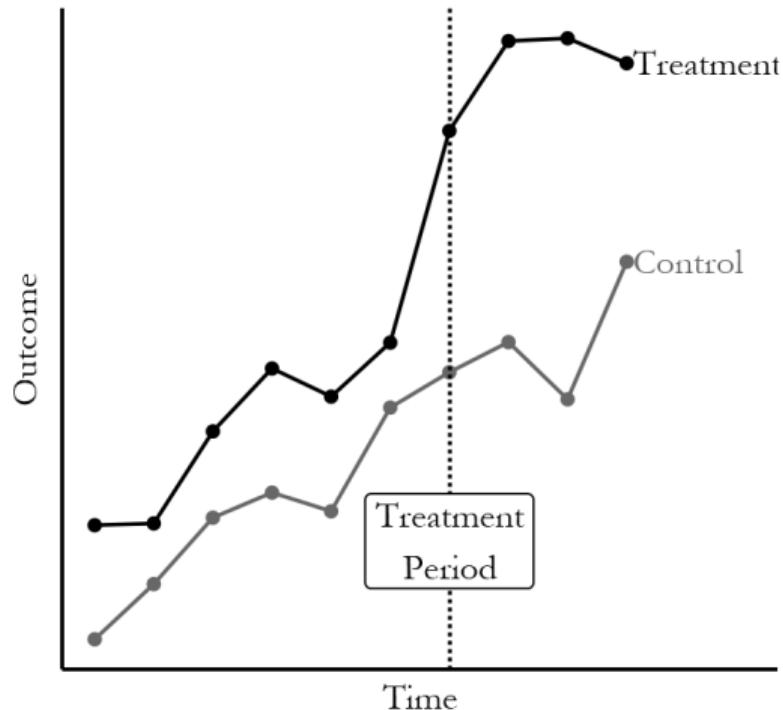
$$\text{ATT} = (E[Y_k|\text{Post}] - E[Y_k|\text{Pre}]) - (E[Y_U|\text{Post}] - E[Y_U|\text{Pre}])$$

With the potential outcomes notation:

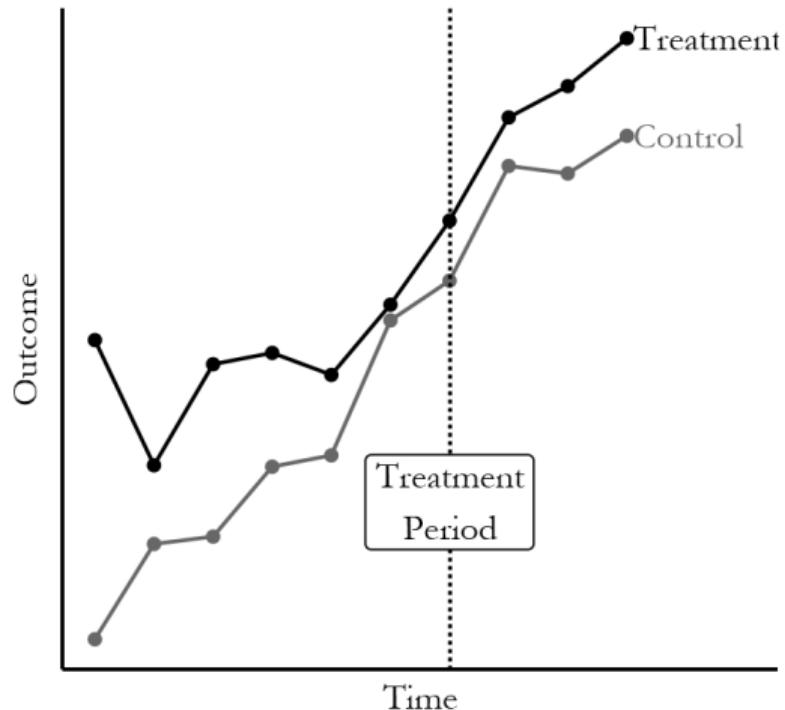
$$\begin{aligned}\text{ATT} = & (E[Y_k^1|\text{Post}] - E[Y_k^0|\text{Post}]) + \\ & (E[Y_k^0|\text{Post}] - E[Y_k^0|\text{Pre}]) - (E[Y_U^0|\text{Post}] - E[Y_U^0|\text{Pre}])\end{aligned}$$

More Theory

(a) Parallel Prior Trends



(b) Converging Prior Trends



A Mississippi Experiment



A bank in Tchula (1939)

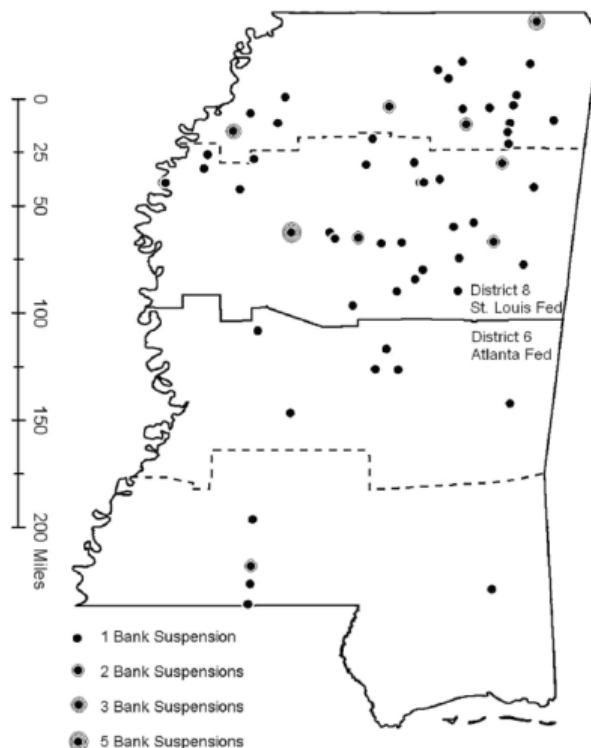
Caldwell crashed down in November 1930.

Mississippi was supervised by two regional Federal Reserve Bank.

The Atlanta Fed ran the Sixth District and lent money to banks.

The St. Louis Fed ran the Eighth District and restricted credits.

One Mississippi, Two Mississippi

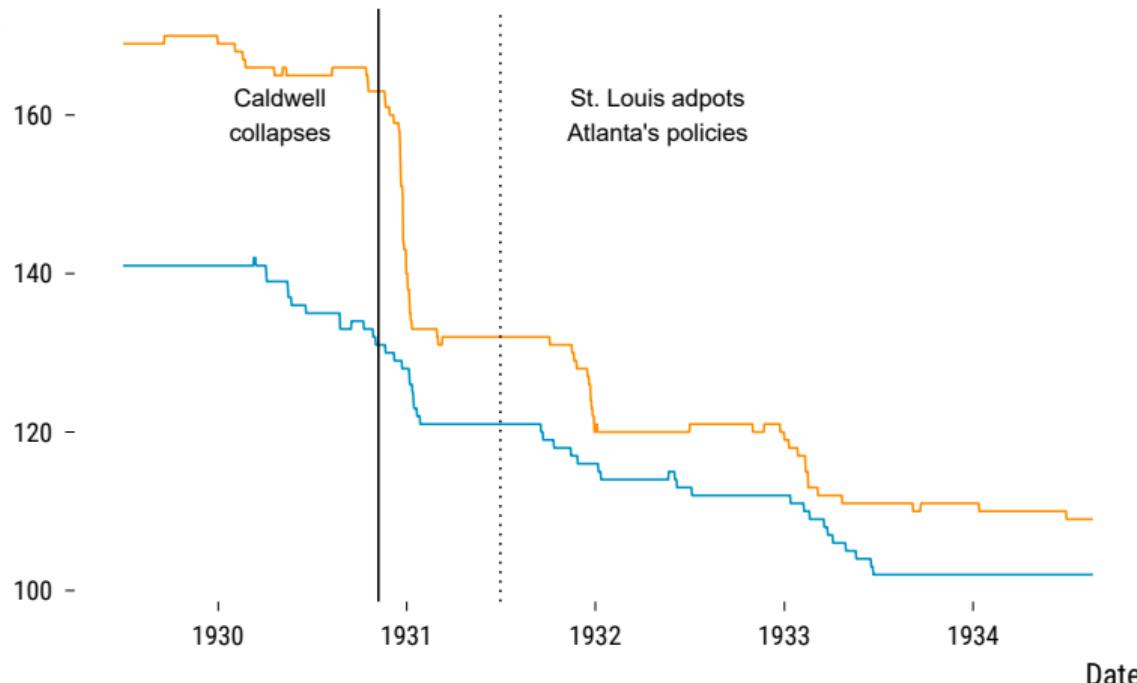


Intuition

Number of Banks in Business and in Operations in the 6th and 8th Federal Reserve Districts in Mississippi, July 1929 to June 1933

District 6 Atlanta 8 St. Louis

Number of Banks



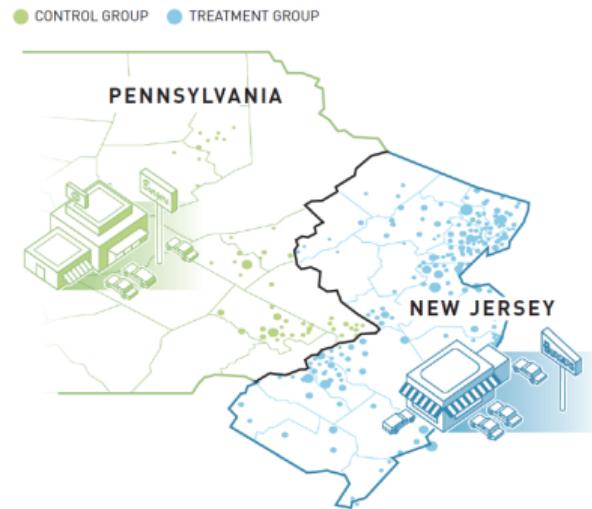
CASE STUDY

Minimum Wage Employment

Card and Krueger (1994)

Negative effects of the minimum wage?

James Buchanan: "*camp following whores*"



Minimum Wage Employment

In competitive labor markets, an increase in the minimum wage would decrease employment

But, in a monopsony market, minimum wages can increase employment

The effect of minimum wage is an empirical question

... hard to answer it with a RCT

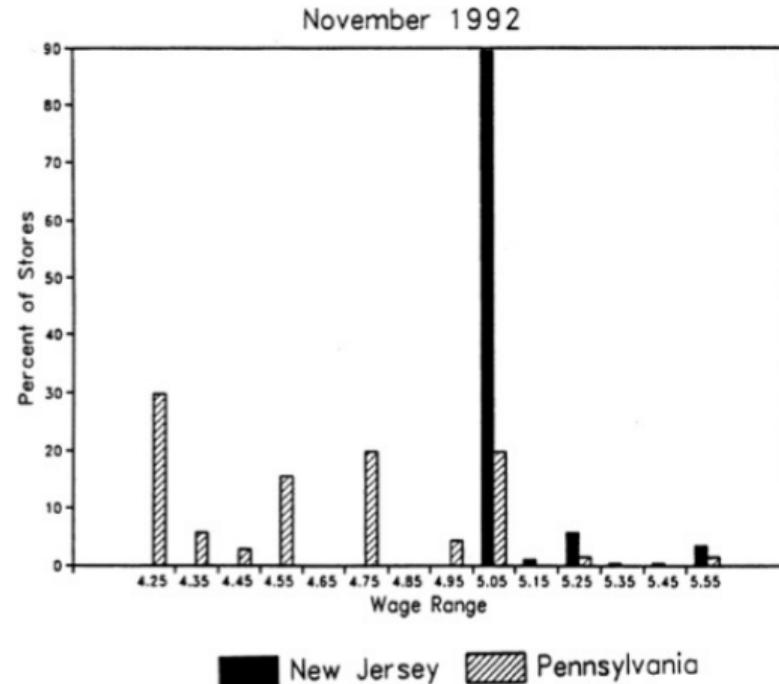
A Quasi-Natural Experiment

In November 1992, New Jersey increased the state minimum wage from \$4.25 to \$5.05

Pennsylvania's minimum wage was staying at \$4.25

Card and Krueger fielded a survey of about four hundred fast-food restaurants in both states—once in February 1992 (before) and again in November (after).

Distribution of wages for NJ and PA in November 1992



Difference-in-Differences

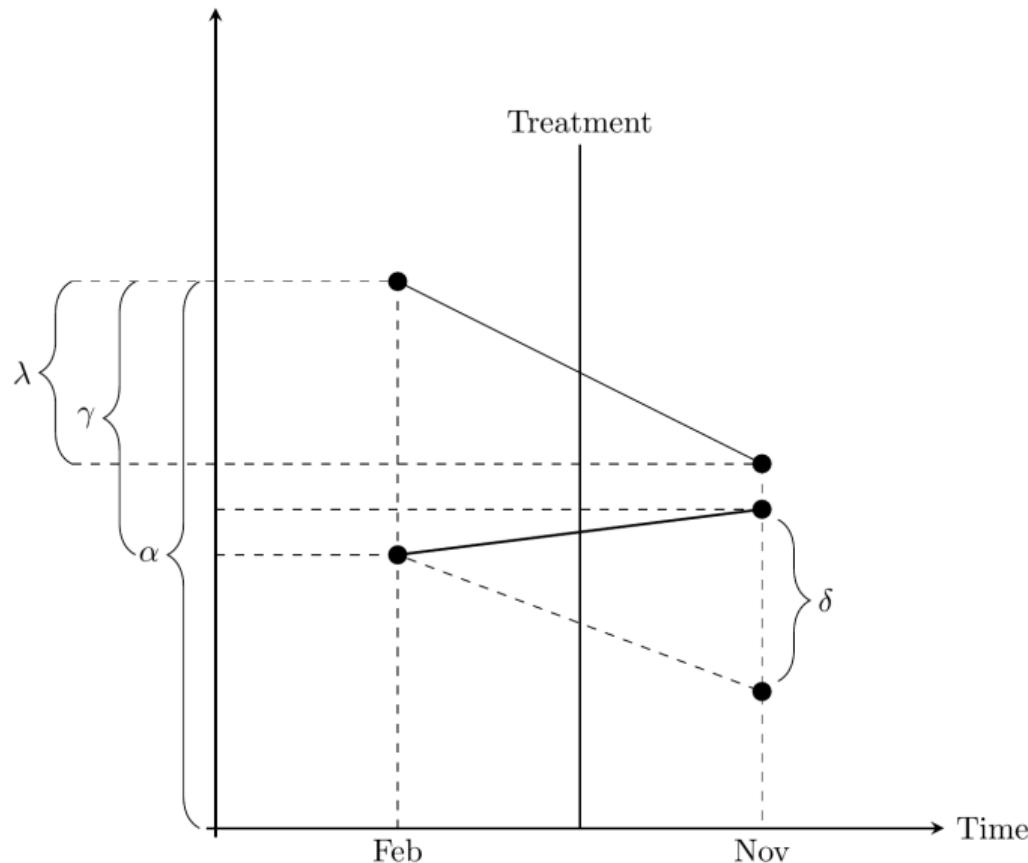
Dependent Variable	Stores by State		
	PA	NJ	NJ – PA
FTW before	23.3	20.44	-2.89
	(1.35)	(0.51)	(1.44)
FTE after	21.147	21.03	-0.14
	(0.94)	(0.52)	(1.07)
Change in mean FTE	-2.16	0.59	2.76
	(1.25)	(0.54)	(1.36)

Difference-in-Differences with Regression

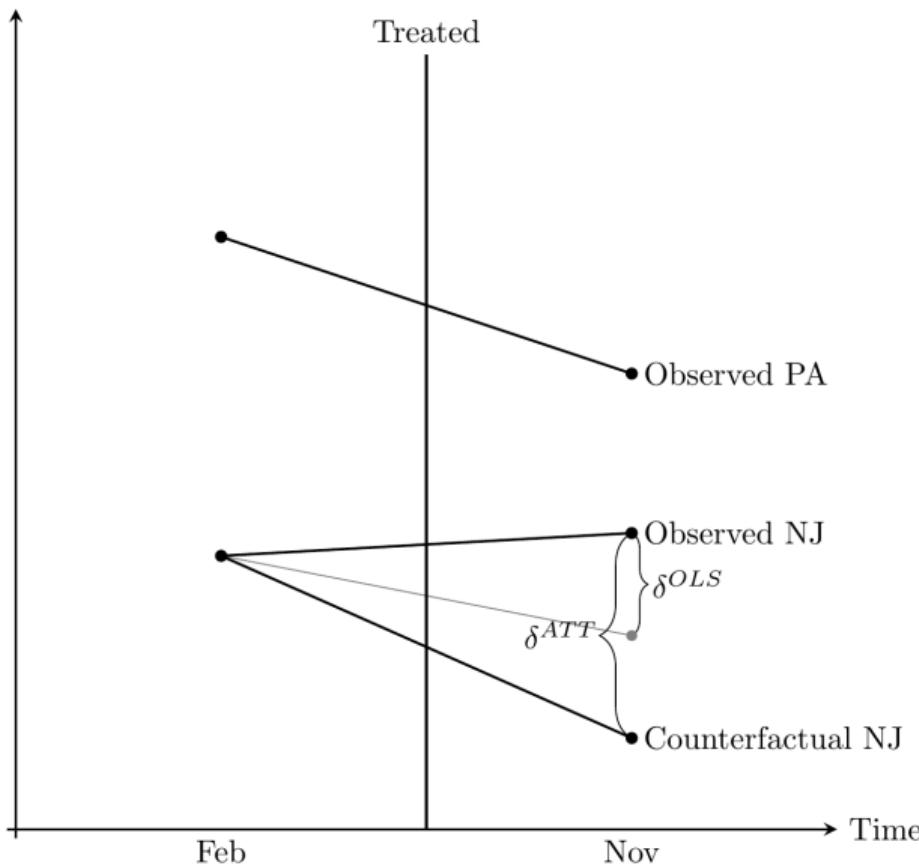
$$Y_{its} = \alpha + \gamma NJ_s + \lambda D_t + \delta (NJ \times D)_{st} + \epsilon_{its}$$

1. PA Pre: α
2. PA Post: $\alpha + \lambda$
3. NJ Pre: $\alpha + \gamma$
4. NJ Post: $\alpha + \gamma + \lambda + \delta$

Labor Supply



Labor Supply



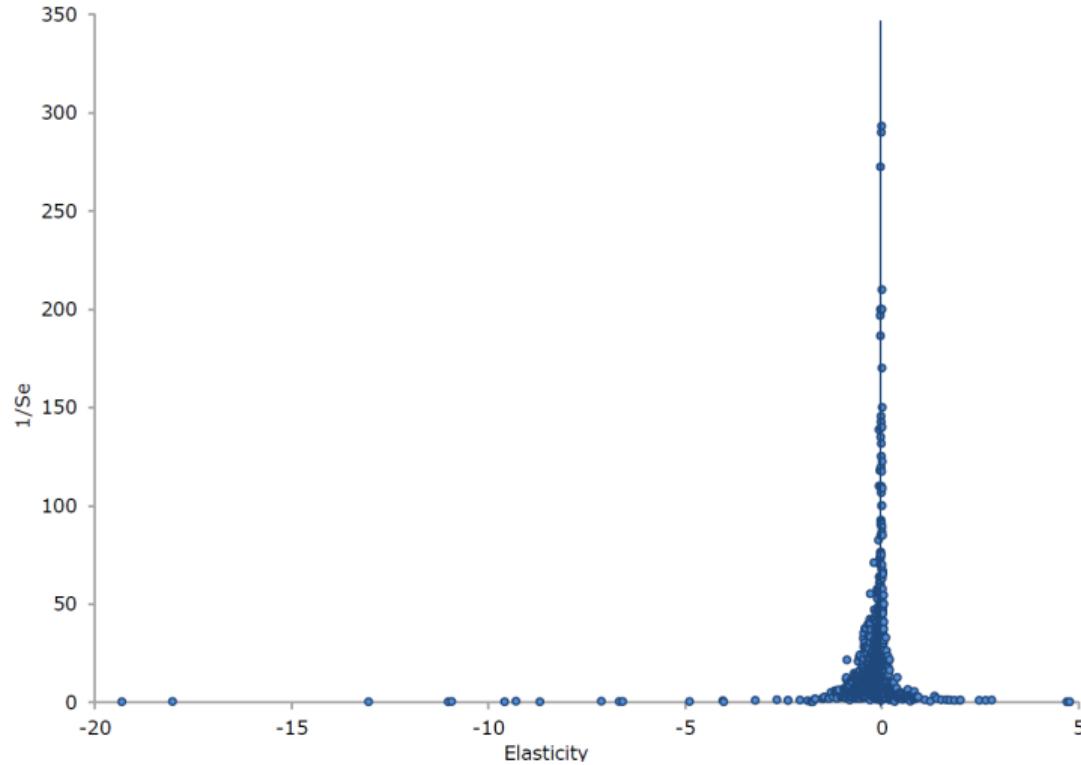
Why Does the Minimum Wage Have No Discernible Effect on Employment?

John Schmitt (2013):

The weight of that evidence points to little or no employment response to modest increases in the minimum wage.

Channels: reductions in labor turnover; improvements in organizational efficiency; reductions in wages of higher earners ("wage compression"); and small price increases

Meta-Analysis



Doucouliagos and Stanley (2009)

Regression Discontinuity Design

The Importance of Rules in our Lives

State of California limits elementary school class size to 32 students

Potential armed forces recruits with test scores in the lower deciles are ineligible for American military service

Number of people admitted to the ENS...

We can exploit these rules for causal inference!

The Importance of Rules in our Lives

State of California limits elementary school class size to 32 students

Potential armed forces recruits with test scores in the lower deciles are ineligible for American military service

Number of people admitted to the ENS...

We can exploit these rules for causal inference!

21st Birthday in the US

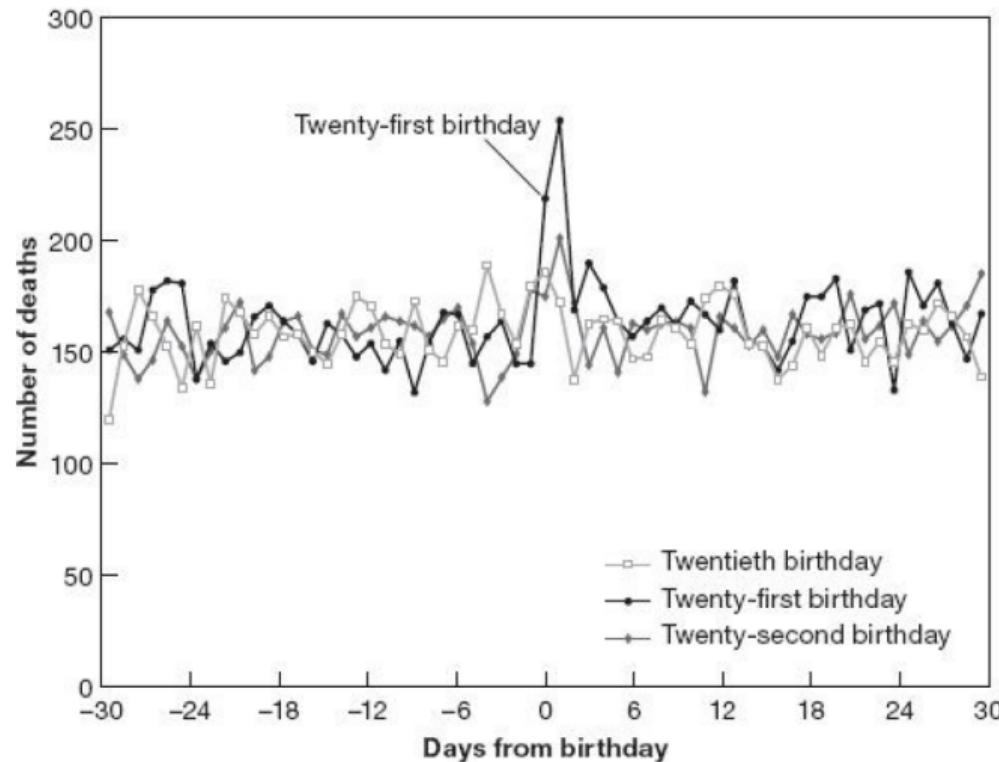
Can drink *legally*

Amethyst Initiative

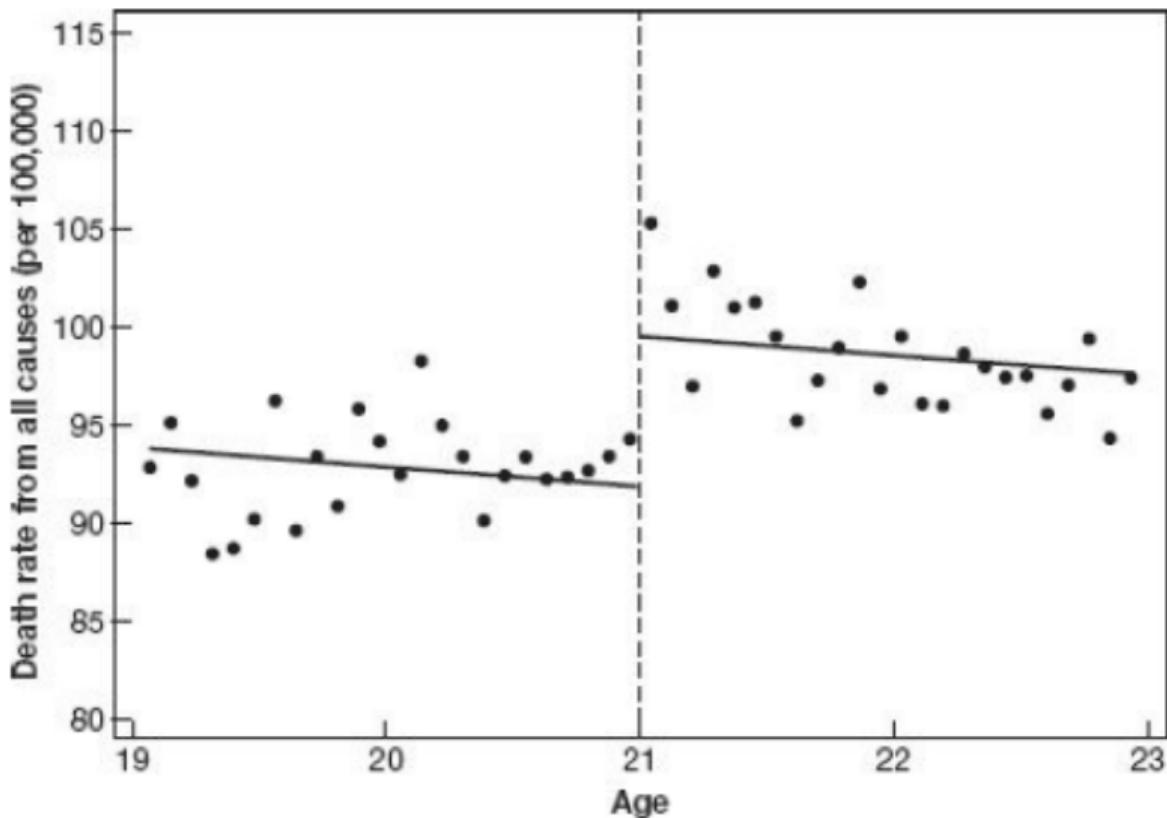
MLDA to 18



A Natural Experiment for MLDA



Sharp RD



Sharp RD

Treatment: D_a where $D_a = 1$ if legal drinking age and 0 otherwise

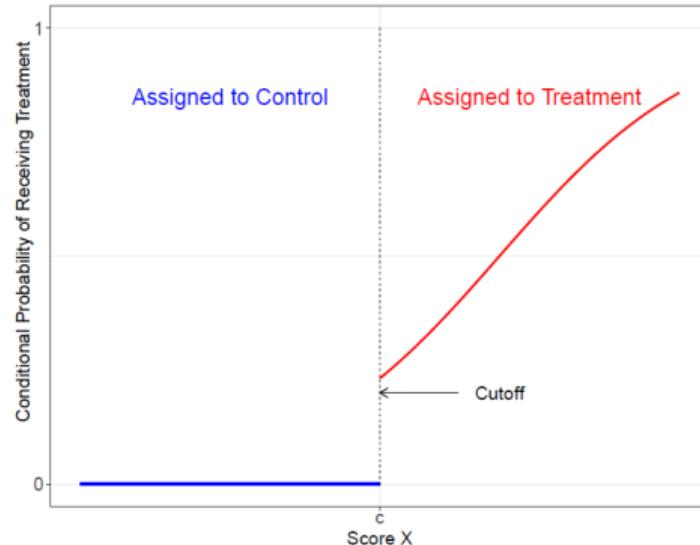
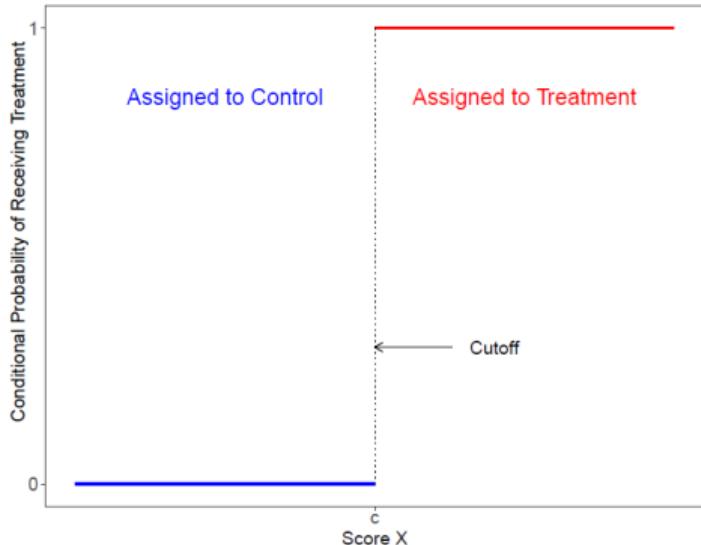
Treatment status is a deterministic function of age

Treatment status is a discontinuous function of age

Here: *sharp* design

When the probability of intensity of treatment jumps at a cutoff, *fuzzy* design (IV).

Why Does It Work?



Important points

Mortality changes with age for reasons unrelated to the MLDA

Death rates from disease-related causes like cancer are low but increasing

Deaths from external causes decrease

To separate this trend variation from any possible MLDA effects, an RD analysis controls for smooth variation in death rates generated by age.

Simple RD Analysis

$$\bar{M}_a = \alpha + \rho D_a + \gamma a + e_a$$

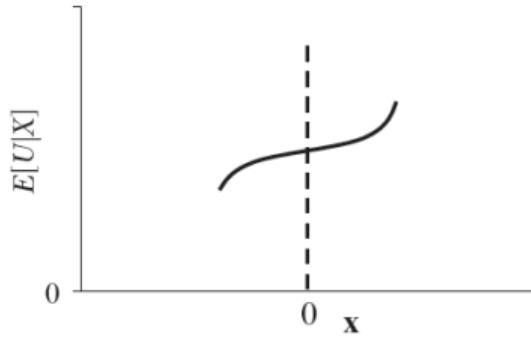
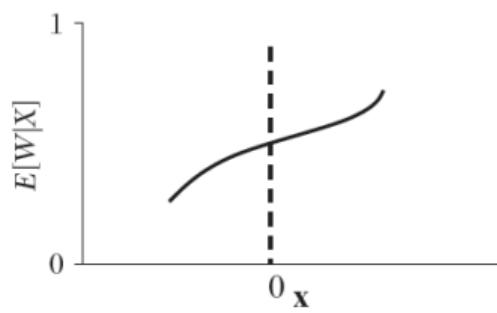
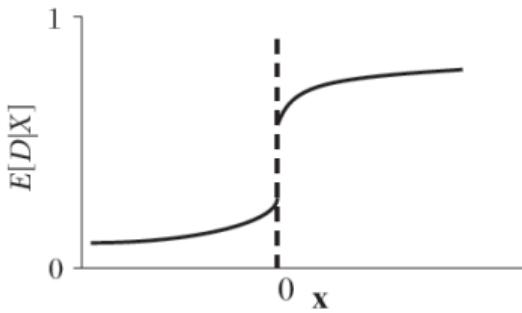
\bar{M}_a death in a month a

Fitted values produce the lines of the previous graph

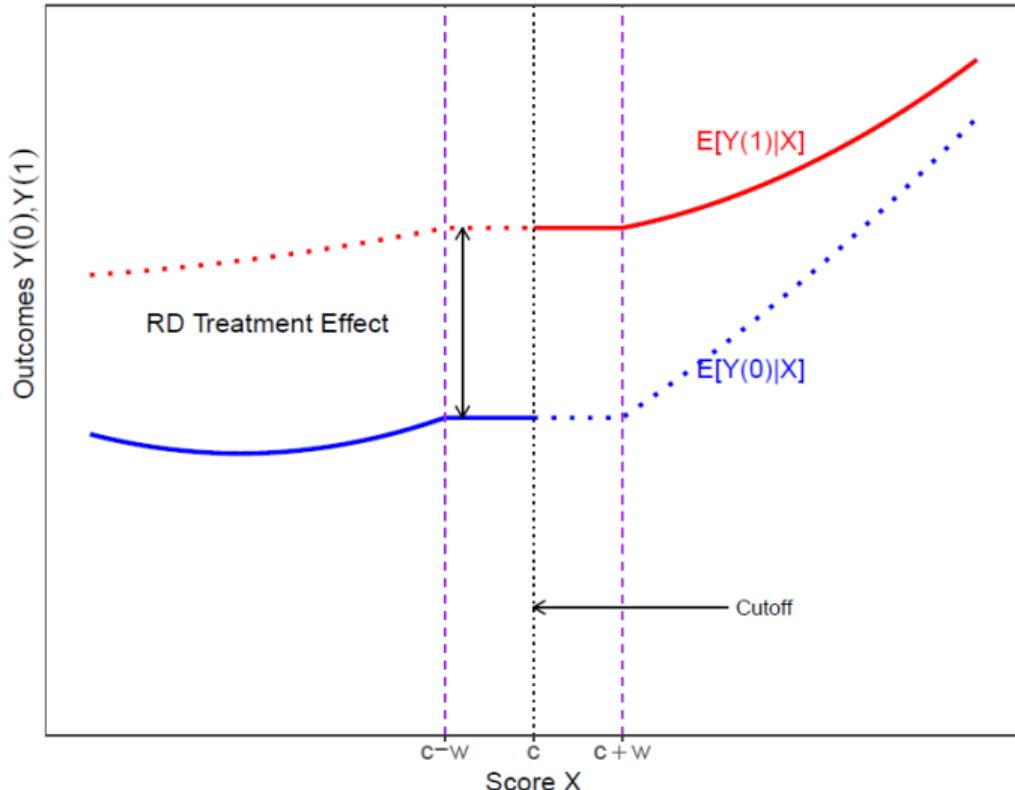
γ reflects the declining death rates as young people mature

ρ is equal to 7.7 is the estimate for the treatment of interest

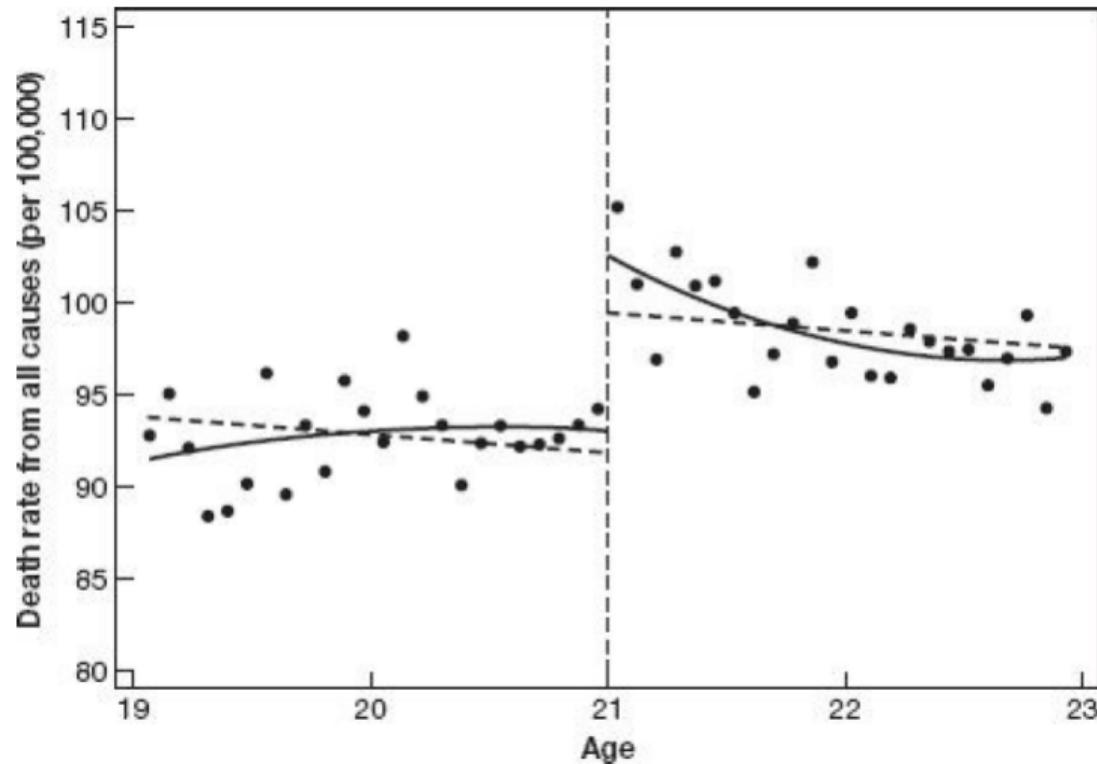
Why Does It Work?



Extrapolation Issue



Why Does It Work?



Final Result

