Diff-in-diff II

March 11, 2020

PMAP 8521: Program Evaluation for Public Service
Andrew Young School of Policy Studies
Spring 2020

Fill out your reading report on iCollege!

Plan for today

Quick talk about COVID-19

DiD review

DiD full example

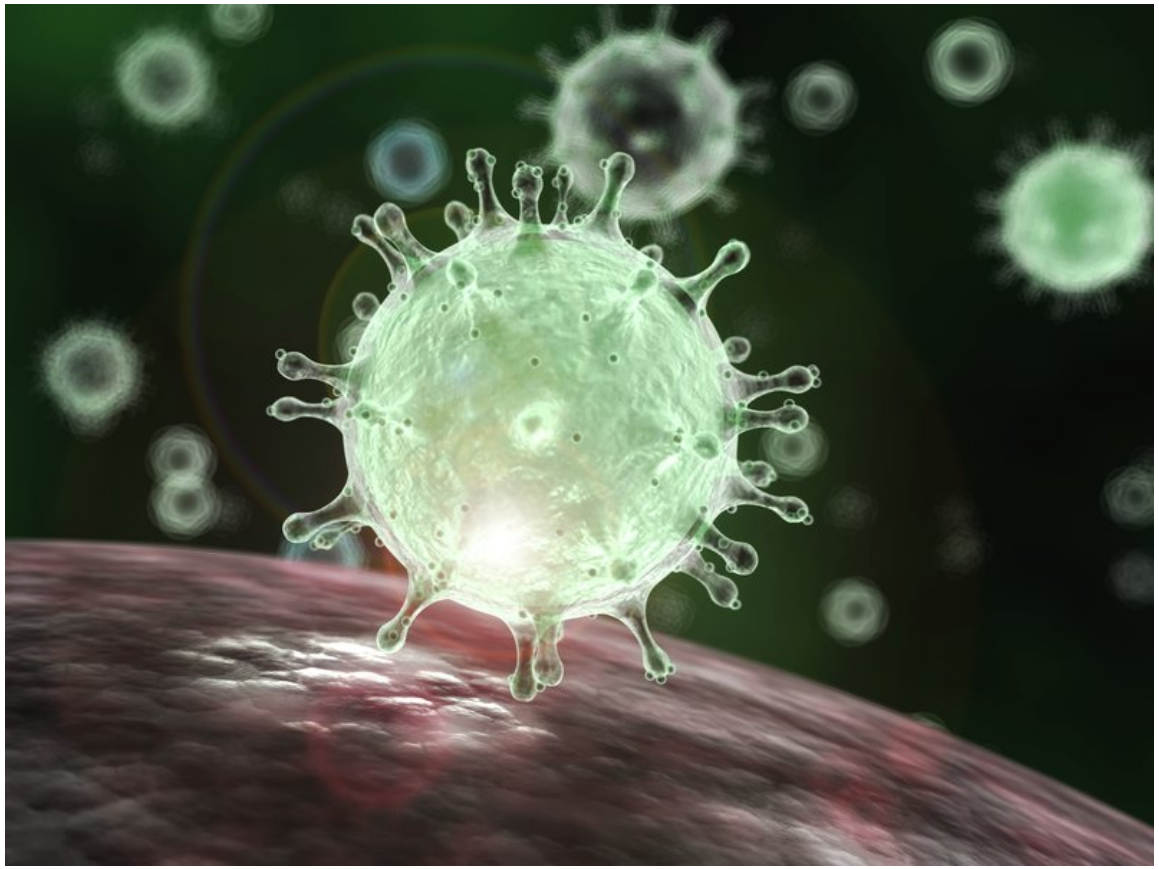
Quick talk about COVID-19

What is all this?

New virus in the coronavirus family

Officially named SARS-COV-2

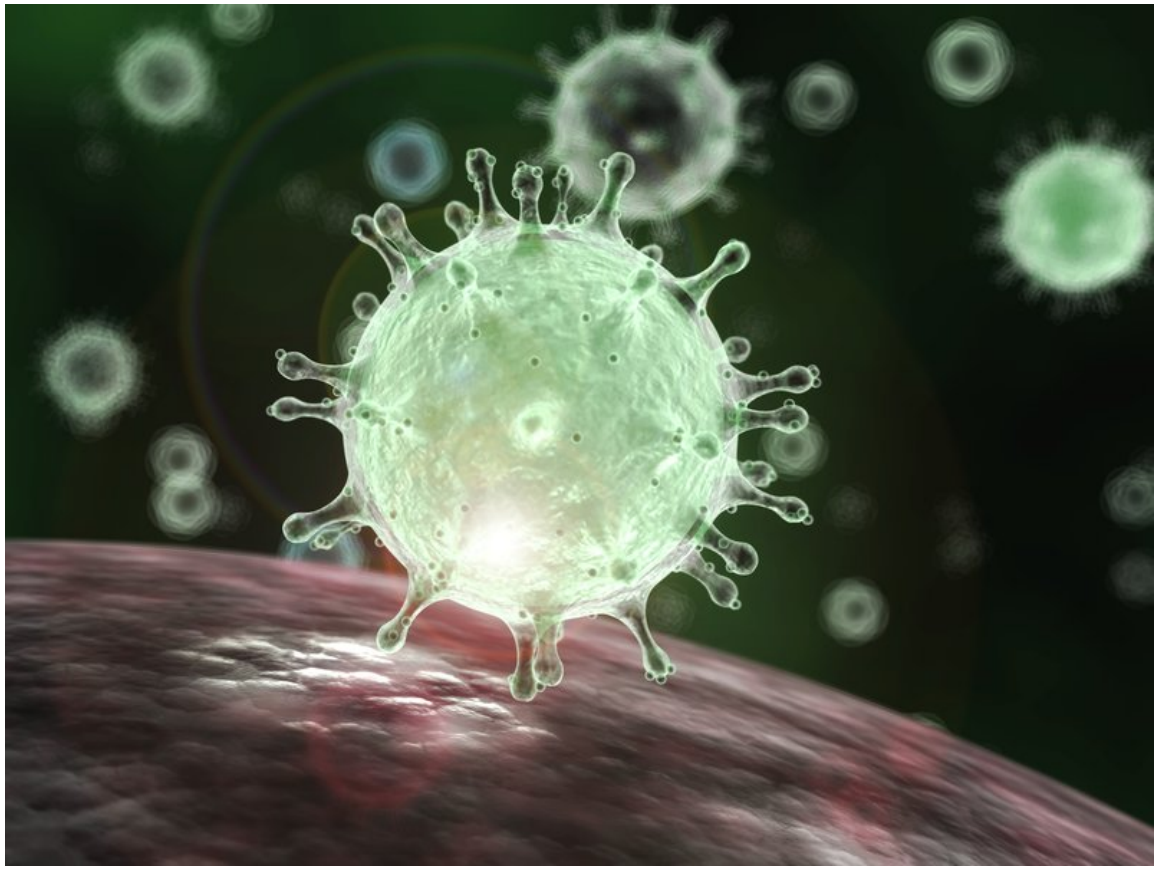
Causes respiratory disease named COVID-19



What is all this?

Originated in Wuhan, Hubei Province, China

Don’t call it “Chinese Coronavirus” or “Kung Flu” or other xenophobic names!



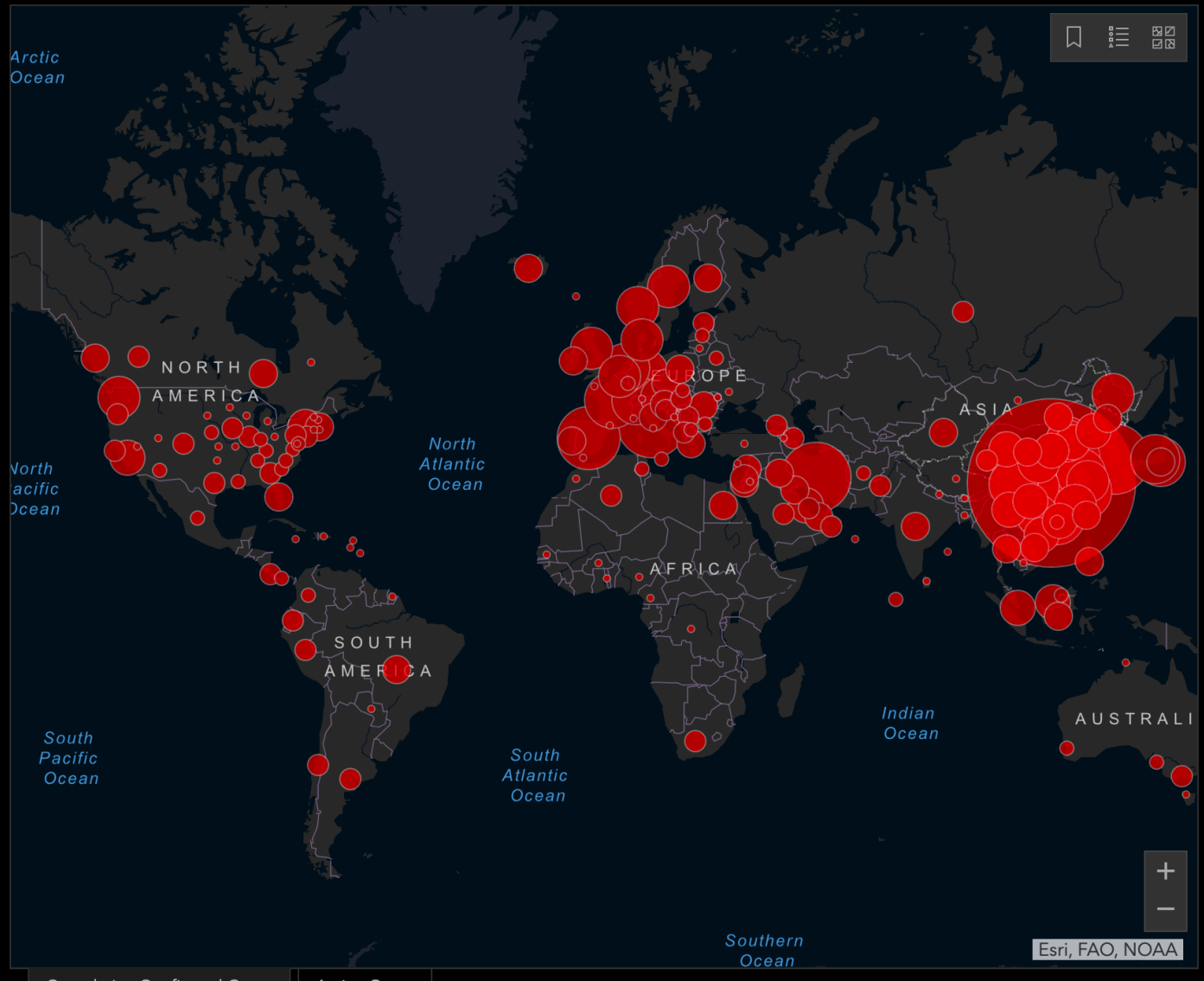
Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)

Total Confirmed
121,564

Confirmed Cases by Country/Region
- 80,969 China
- 10,149 Italy
- 9,000 Iran (Islamic Republic of)
- 7,755 Republic of Korea
- 2,174 Spain
- 1,784 France
- 1,629 Germany
- 1,050 US
- 696 Others
- 613 Switzerland
- 581 Japan
- 503 Netherlands
- 477 Sweden
- 440 Norway
- 382 UK
- 340 Denmark
- 314 Belgium

Country/Region | St/Prov

Last Updated at (M/D/YYYY)
3/11/2020, 10:33:04 AM



Cumulative Confirmed Cases | Active Cases

118
countries/regions

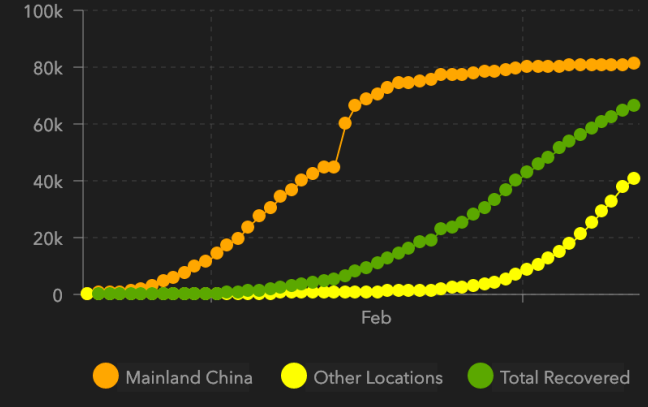
Lancet Inf Dis Article: Here. Mobile Version: Here. Visualization: JHU CSSE. Automation Support: Esri Living Atlas team and JHU APL.
Data sources: WHO, CDC, ECDC, NHC and DXY and local media reports. Read more in this blog. Contact US.
Downloadable database: GitHub: Here. Feature layer: Here.
Confirmed cases include **presumptive positive** cases.

Total Deaths
4,373

- 3,046 deaths **Hubei** China
- 631 deaths Italy
- 354 deaths Iran (Islamic Republic of)
- 54 deaths Republic of Korea
- 49 deaths Spain
- 33 deaths France
- 23 deaths **Washington** US
- 22 deaths **Henan** China
- 13 deaths **Heilongjiang** China

Total Recovered
66,239

- 49,134 recovered **Hubei** China
- 2,959 recovered Iran (Islamic Republic of)
- 1,282 recovered **Guangdong** China
- 1,249 recovered **Henan** China
- 1,195 recovered **Zhejiang** China
- 995 recovered **Hunan** China
- 984 recovered **Anhui** China
- 932 recovered **Jiangxi** China
- 726 recovered **Shandong** China



Actual | Logarithmic | Daily Cases

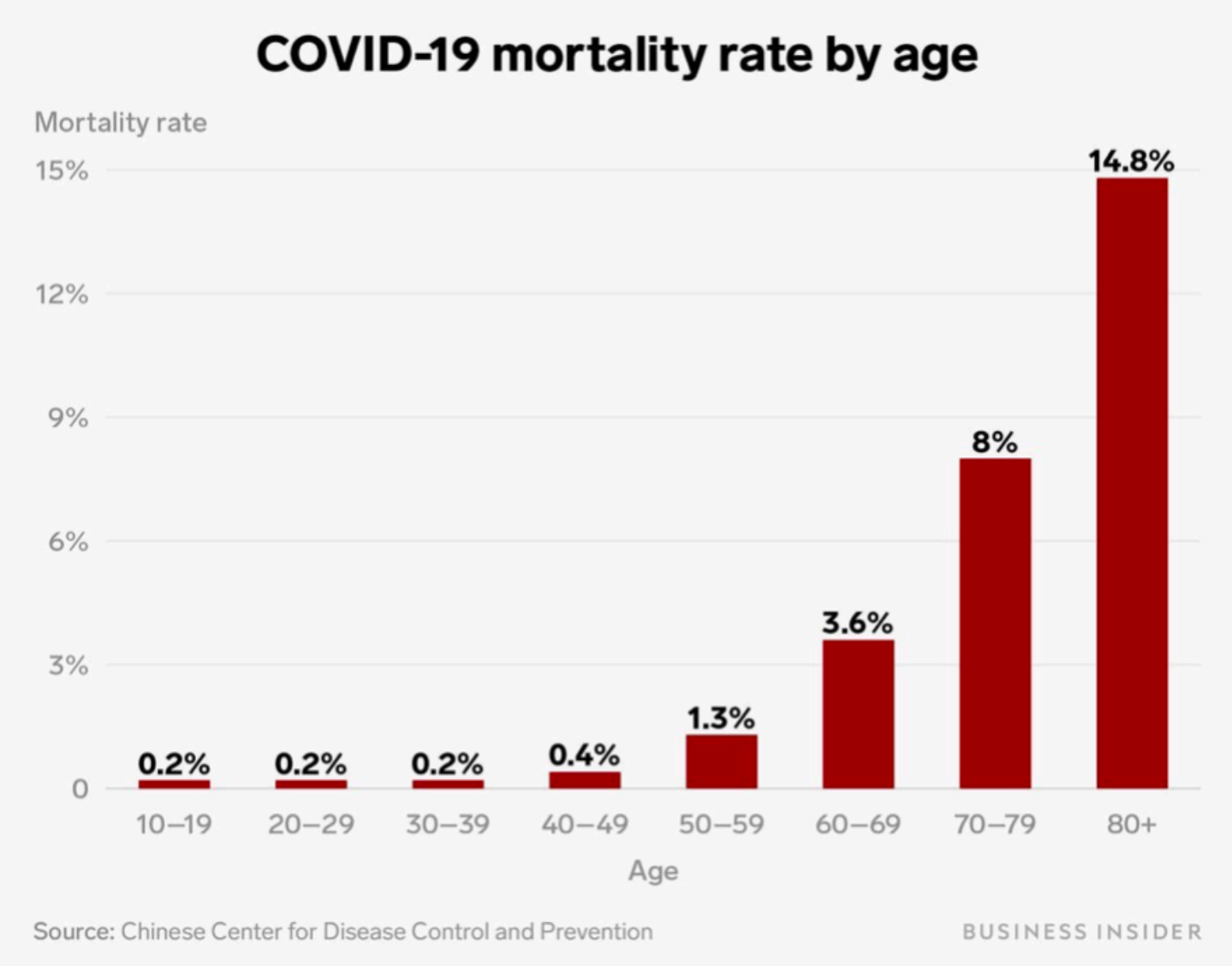
Symptoms

Fever and dry cough initially; pneumonia-like respiratory failure later for vulnerable people

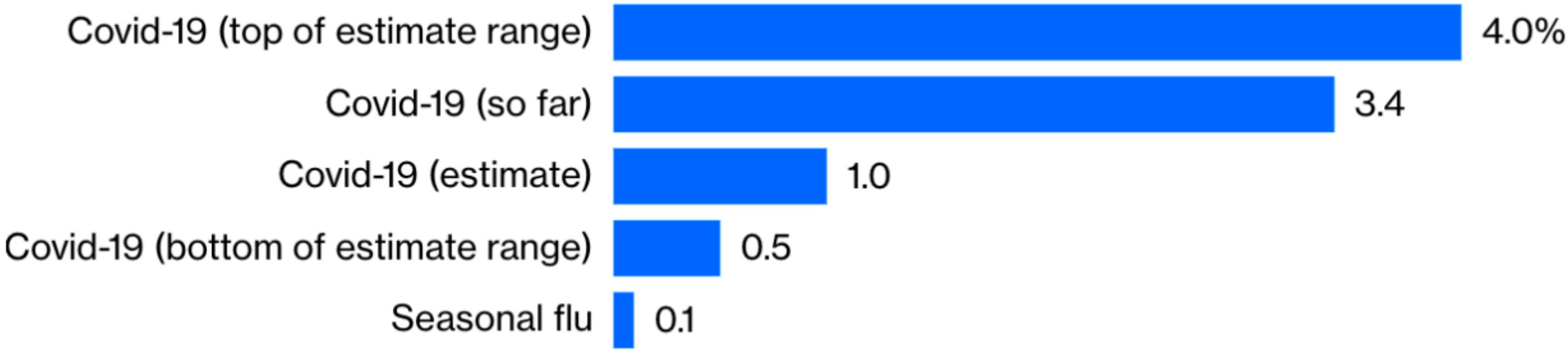
Up to two weeks can pass between exposure and symptoms

Asymptomatic transmission likely possible

Lethality

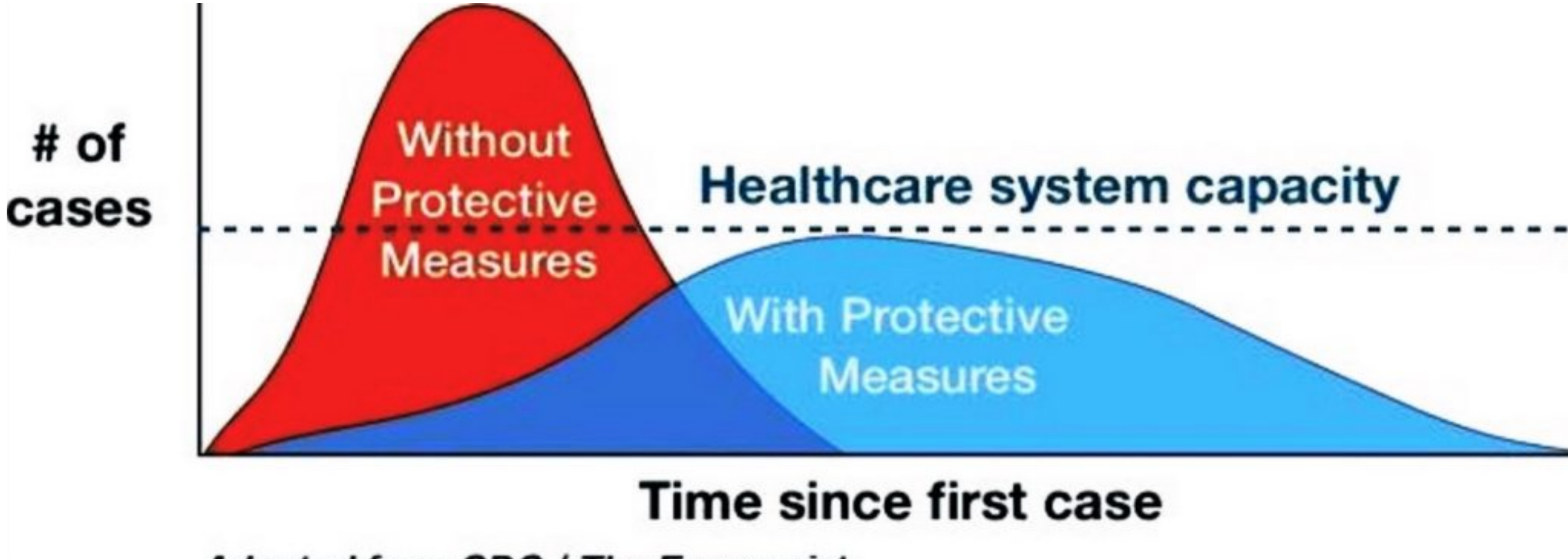

COVID-19 mortality rate by age
Mortality rate
15%
12%
9%
6%
3%
0
0.2%
0.2%
0.2%
0.4%
1.3%
3.6%
8%
14.8%
10–19
20–29
30–39
40–49
50–59
60–69
70–79
80+
Age
Source: Chinese Center for Disease Control and Prevention
BUSINESS INSIDER

Lethality


Covid-19 (top of estimate range)
4.0%
Covid-19 (so far)
3.4
Covid-19 (estimate)
1.0
Covid-19 (bottom of estimate range)
0.5
Seasonal flu
0.1

Why is everything shutting down?

Flattening the curve



Adapted from CDC / The Economist

If you're young and healthy, all these cancellations and precautions are not about you!

Social distancing, staying home, washing your hands, etc. protects the vulnerable

Huge collective action problem!

What you can do

Wash hands for 20 seconds, disinfect phone, don't touch your face

Stay home if you're sick

Practice social distancing

Limit non-essential travel

Don't buy masks

Stock up on essentials but don't hoard

flattenthecurve.com

What does this mean for our class?

I HAVE NO IDEA YET

GSU hasn't made any official decisions

I'm committed to helping you all succeed and keep learning!

I'll continue to stream class via WebEx

2-week late work window is eliminated

Two wrongs make a right


I
federalism
(for the natural experiments)

Raising the minimum wage

What happens if you raise the minimum wage?

Economic theory says there should be fewer jobs

New Jersey in 1992

$4.25 → $5.05

Before vs. after

Average fast food jobs in NJ

Before: 20.44

After: 21.03

Δ: 0.59

Is this the causal effect?

Treatment vs. control

Average fast food jobs in states

PA_{after}: 21.17

NJ_{after}: 21.03

Δ: −0.14

Is this the causal effect?

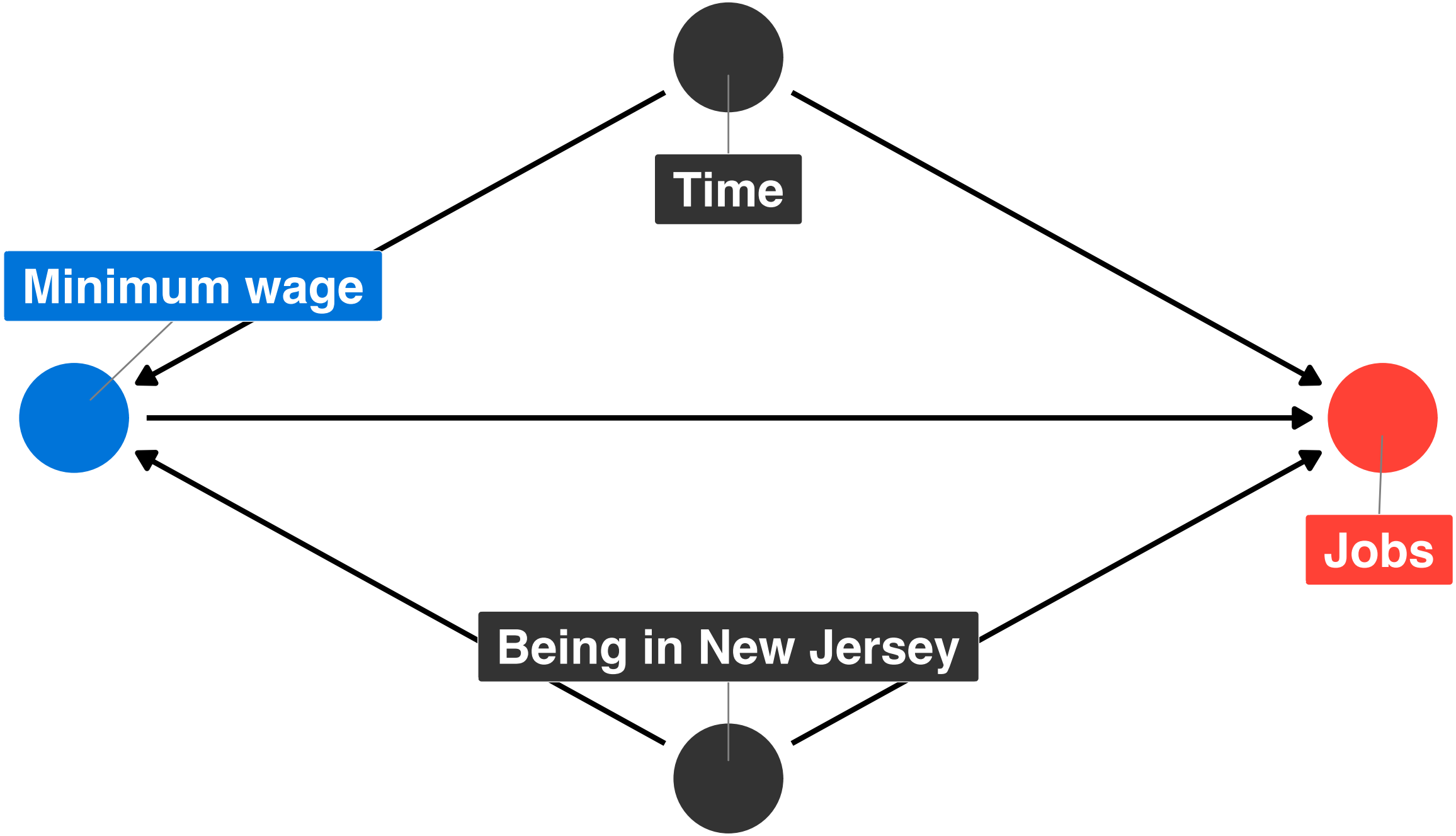
Problems

Comparing only before/after

Impossible to know if growth happened because of treatment or just naturally

Comparing only treatment/control

Impossible to know if any changes happened because of natural growth


Time
Minimum wage
Jobs
Being in New Jersey

	Pre mean	Post mean
Treatment	A (not yet treated)	B (treated)
Control	C (never treated)	D (never treated)

	Pre mean	Post mean	Δ (post–pre)
Treatment	A (not yet treated)	B (treated)	B–A
Control	C (never treated)	D (never treated)	D–C

Growth!

	Pre mean	Post mean
Treatment	A (not yet treated)	B (treated)
Control	C (never treated)	D (never treated)
Δ (trtmt–ctrl)	A–C	B–D

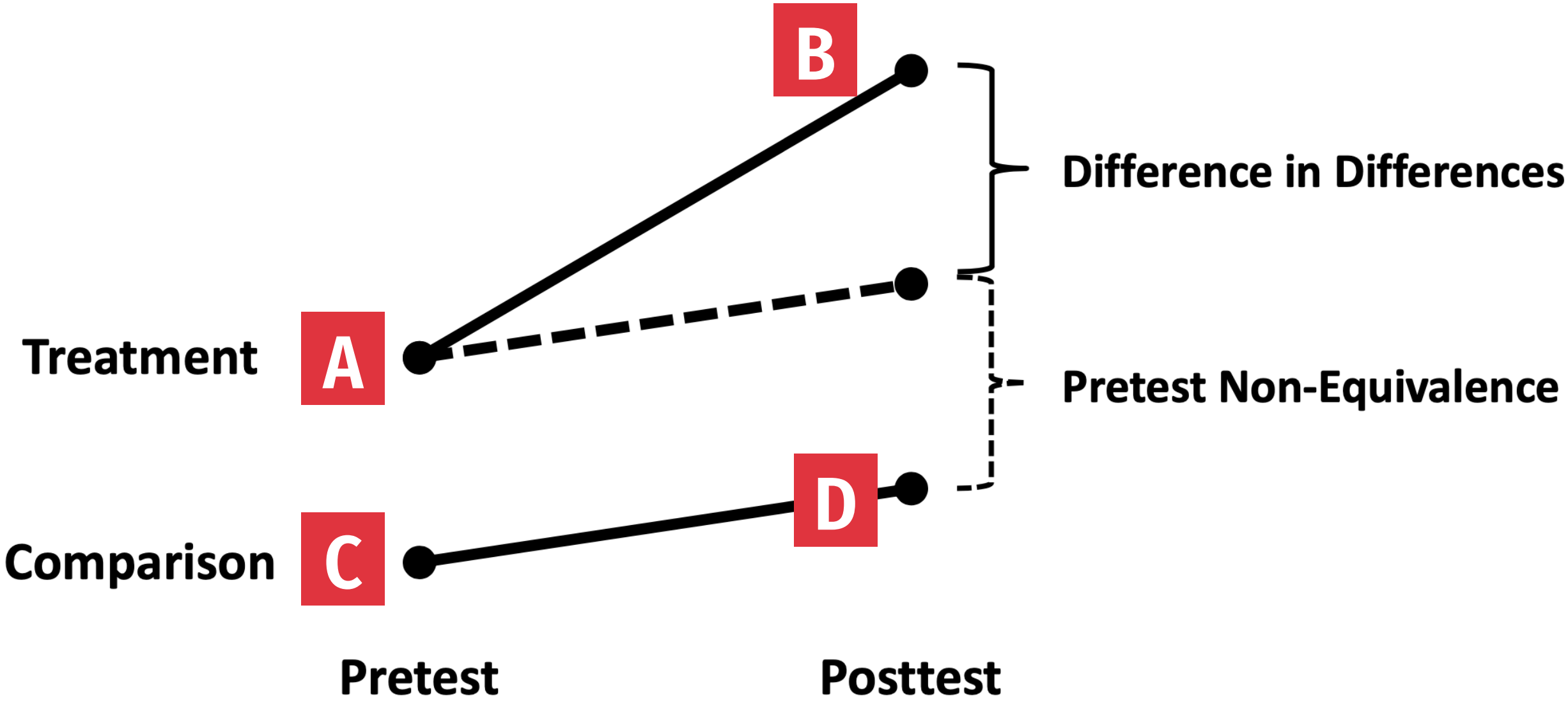
Within-group effects

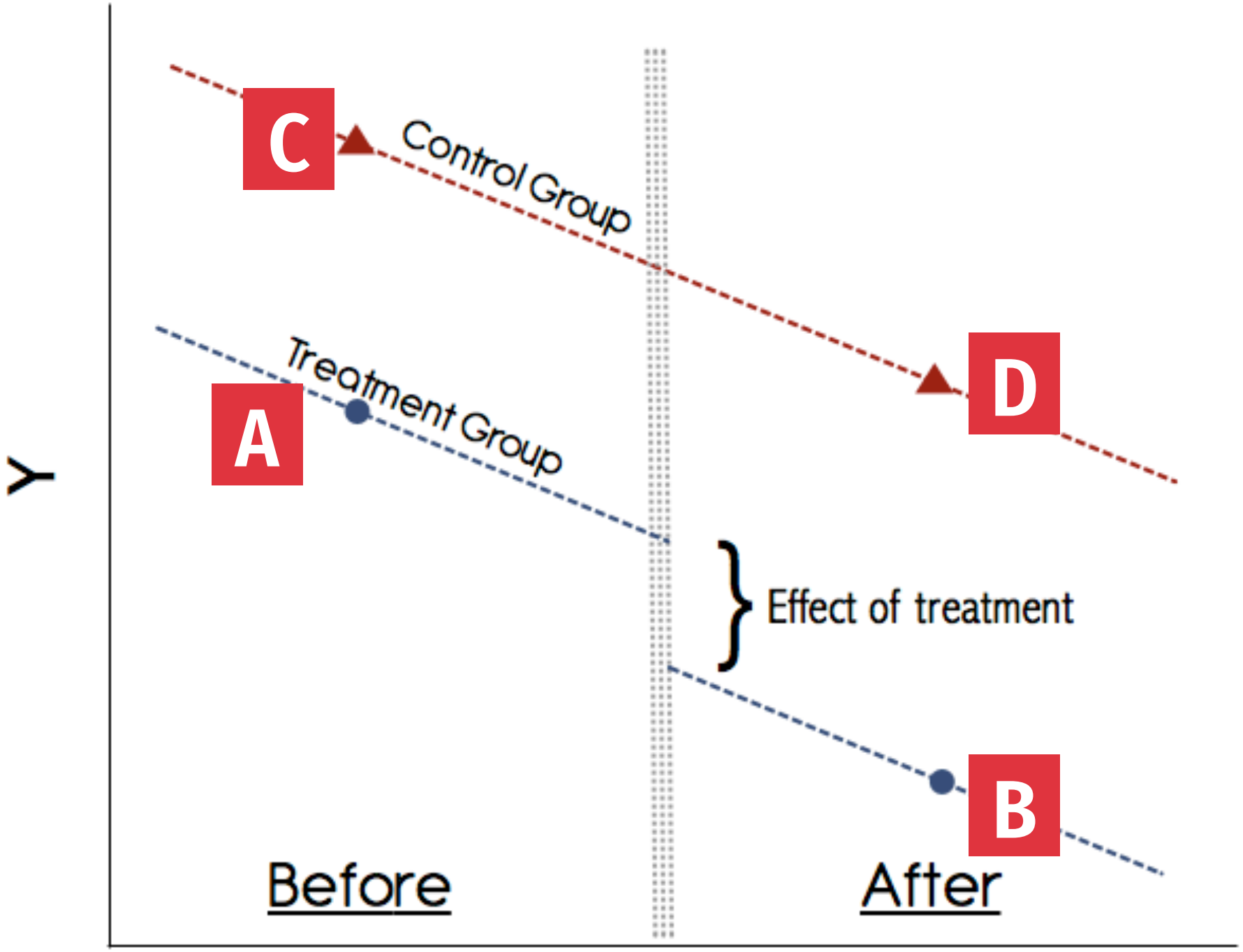
	Pre mean	Post mean	Δ (post–pre)
Treatment	A (not yet treated)	B (treated)	B–A
Control	C (never treated)	D (never treated)	D–C
Δ (trtmt–ctrl)	A–C	B–D	(B–A) – (D–C)

Growth of treatment – growth of control (DiD!)

$$\begin{aligned}\text{DD} =& (\bar{x}_{\text{treatment, post}} - \bar{x}_{\text{treatment, pre}}) \\ & - (\bar{x}_{\text{control, post}} - \bar{x}_{\text{control, pre}})\end{aligned}$$

	Pre mean	Post mean	Δ (post−pre)
NJ	A **20.44**	B **21.03**	B−A **0.59**
PA	C **23.33**	D **21.17**	D−C **−2.16**
Δ (trtmt−ctrl)	A−C **−2.89**	B−D **−0.14**	**(0.59) − (−2.16) =** **2.76**

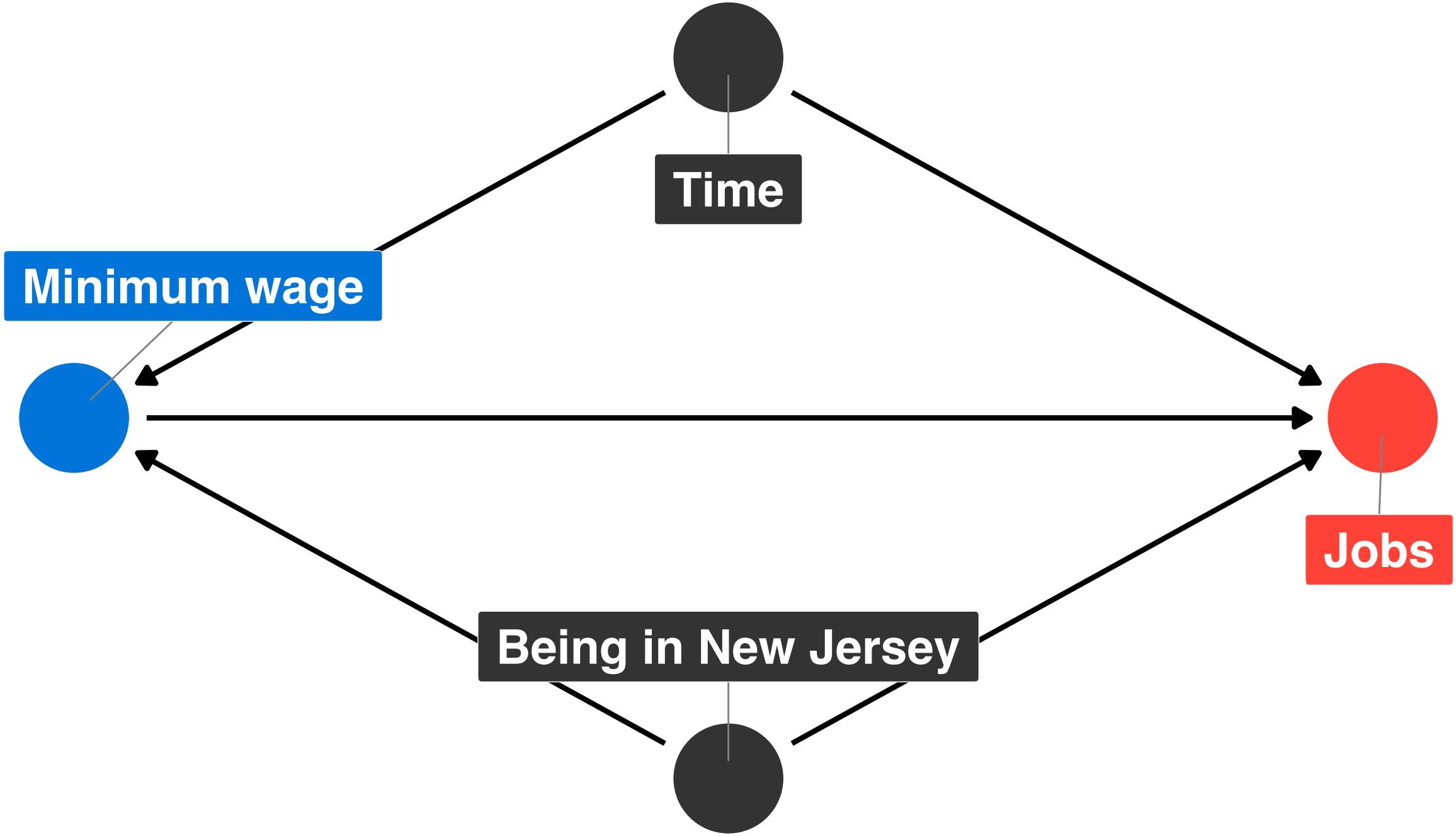

B
Difference in Differences
Treatment
A
Pretest Non-Equivalence
D
Comparison
C
Pretest
Posttest


C
Control Group
D
Treatment Group
A
Y
Effect of treatment
B
Before
After

Finding all the group means is tedious though!

What if there are other backdoors to worry about?

Regression to the rescue!


Time
Minimum wage
Jobs
Being in New Jersey

$$Y_{it} = \alpha + \beta\ \text{Group}_i + \gamma\ \text{Time}_t + \delta\ (\text{Group}_i \times \text{Time}_t) + \epsilon_{it}$$

```
model <- lm(outcome ~ group + time + group * time)
```

Group = 1/TRUE if treatment

Time = 1/TRUE if after

$$Y_{it} = \alpha + \beta\ \text{Group}_i + \gamma\ \text{Time}_t + \delta\ (\text{Group}_i \times \text{Time}_t) + \epsilon_{it}$$

```
model <- lm(outcome ~ group + time + group * time)
```

α = Mean of control, pre-treatment

β = Increase in outcome across groups

γ = Increase in outcome across time

δ = Difference in differences!

$$Y_{it} = \alpha + \beta\ \text{Group}_i + \gamma\ \text{Time}_t + \delta\ (\text{Group}_i \times \text{Time}_t) + \epsilon_{it}$$

	Pre mean	Post mean	Δ (post−pre)
Control	α	α + γ	γ
Treatment	α + β	α + β + γ + δ	γ + δ
Δ (trtmt−ctrl)	β	β + δ	δ

R time!