Evaluating Public Policy

Problems (easy) vs Solutions (hard)

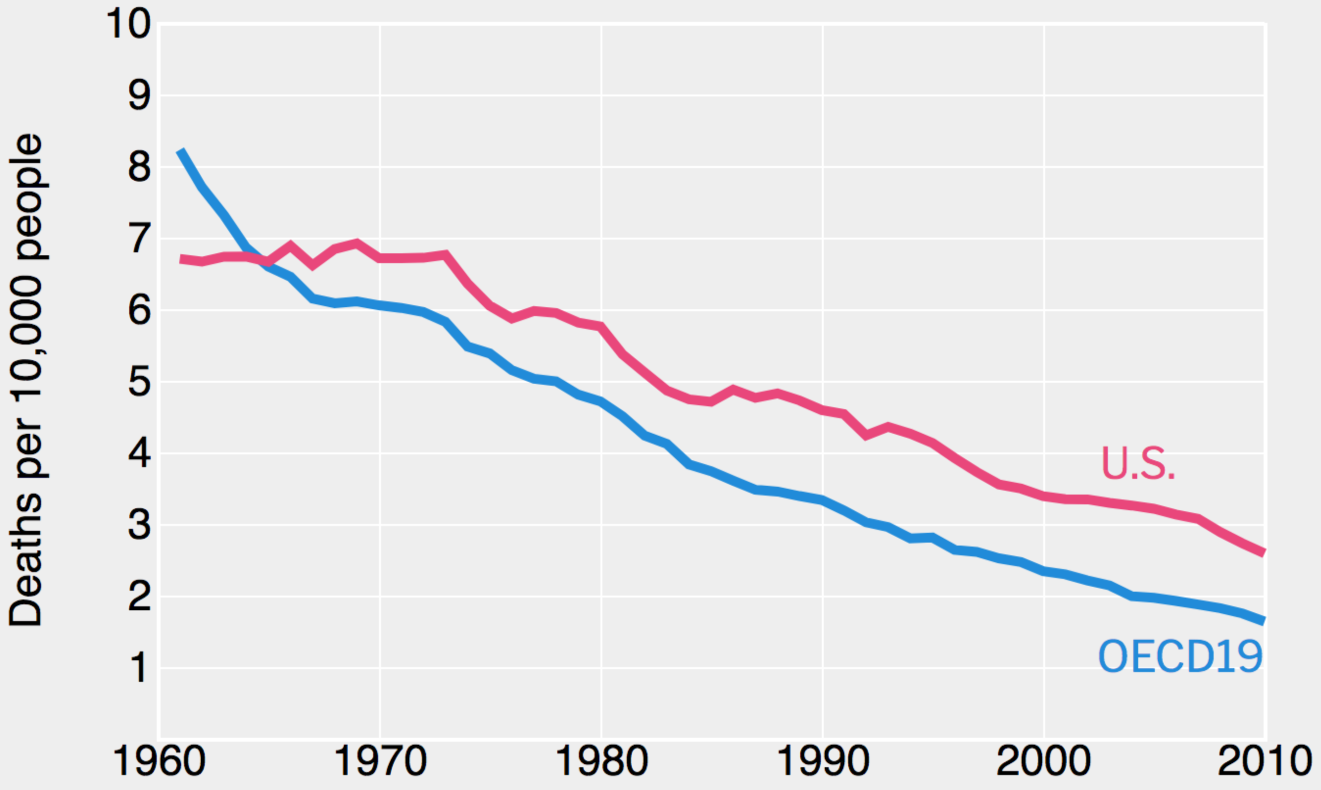
Understanding Political Numbers

April 1, 2019

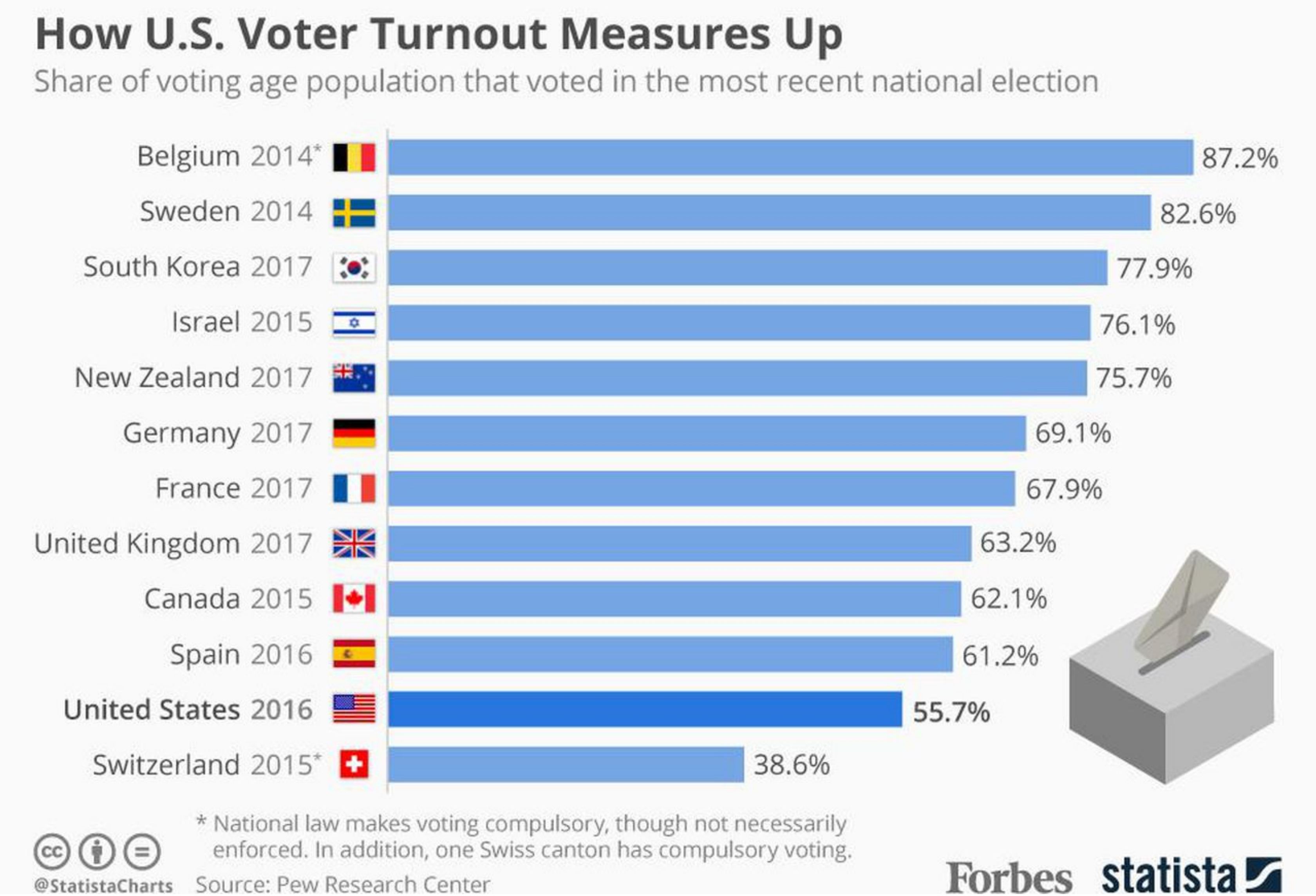
No shortage of bad things in the world

Child mortality in the U.S. vs. OECD19

Ages 1-19



Source: Authors' analysis of data from Human Mortality Database


How U.S. Voter Turnout Measures Up
Share of voting age population that voted in the most recent national election
Belgium 2014* 87.2%
Sweden 2014 82.6%
South Korea 2017 77.9%
Israel 2015 76.1%
New Zealand 2017 75.7%
Germany 2017 69.1%
France 2017 67.9%
United Kingdom 2017 63.2%
Canada 2015 62.1%
Spain 2016 61.2%
United States 2016 55.7%
Switzerland 2015* 38.6%
* National law makes voting compulsory, though not necessarily enforced. In addition, one Swiss canton has compulsory voting.
@StatistaCharts
Source: Pew Research Center
Forbes
statista

Policy "problems"

Are data in context?

What's causing the problem?

What policy could address that mechanism?

Policy "problems"

Are data in context?

What's causing the problem?

What policy could address that mechanism?

Policy "solutions"

Has the policy been implemented elsewhere?

Is there evidence that the policy works? How do you know?

Can the evidence be *generalized?*

What if the policy hasn't been enacted anywhere else?

Unintended consequences of policy

As economists say, "externalities"

Voter Identification Requirements

Problem? Election security

- Are there fraudulent ballots?
- Voters have shaky confidence that votes are counted as intended

Solution? Tighter ID requirements

- Bills, bank statements, birth certificates, SS cards not good enough
- Student photo ID? Not good enough (no signature)
- Wisconsin: must have photo and a signature

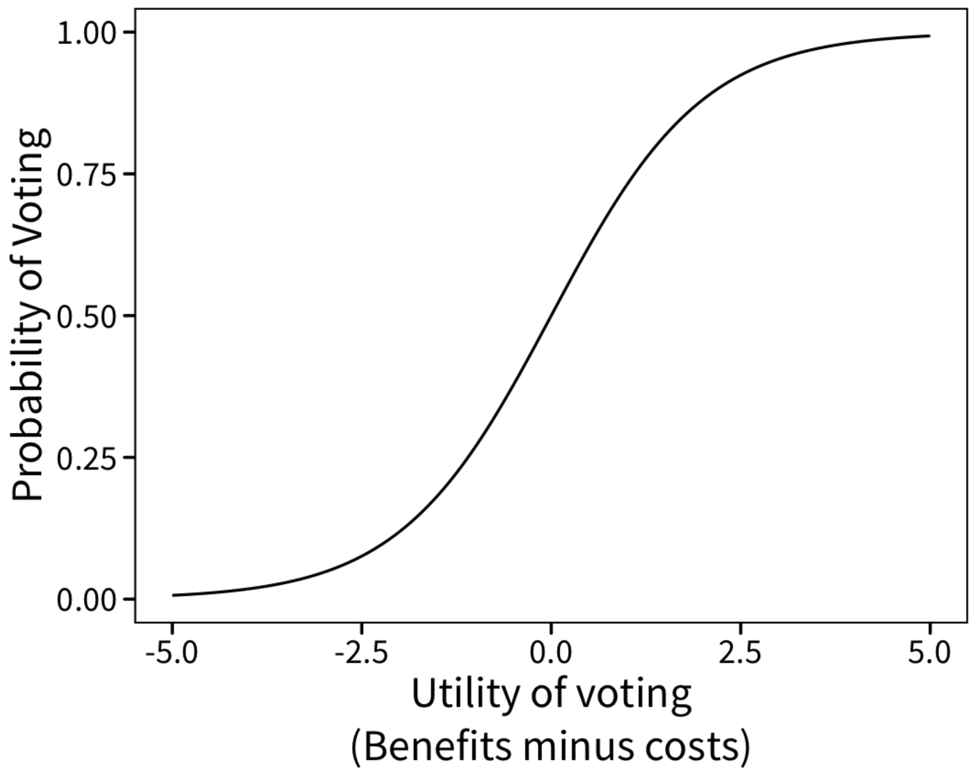
Problem? Election security

- Are there fraudulent ballots?
- Voters have shaky confidence that votes are counted as intended

Solution? Tighter ID requirements

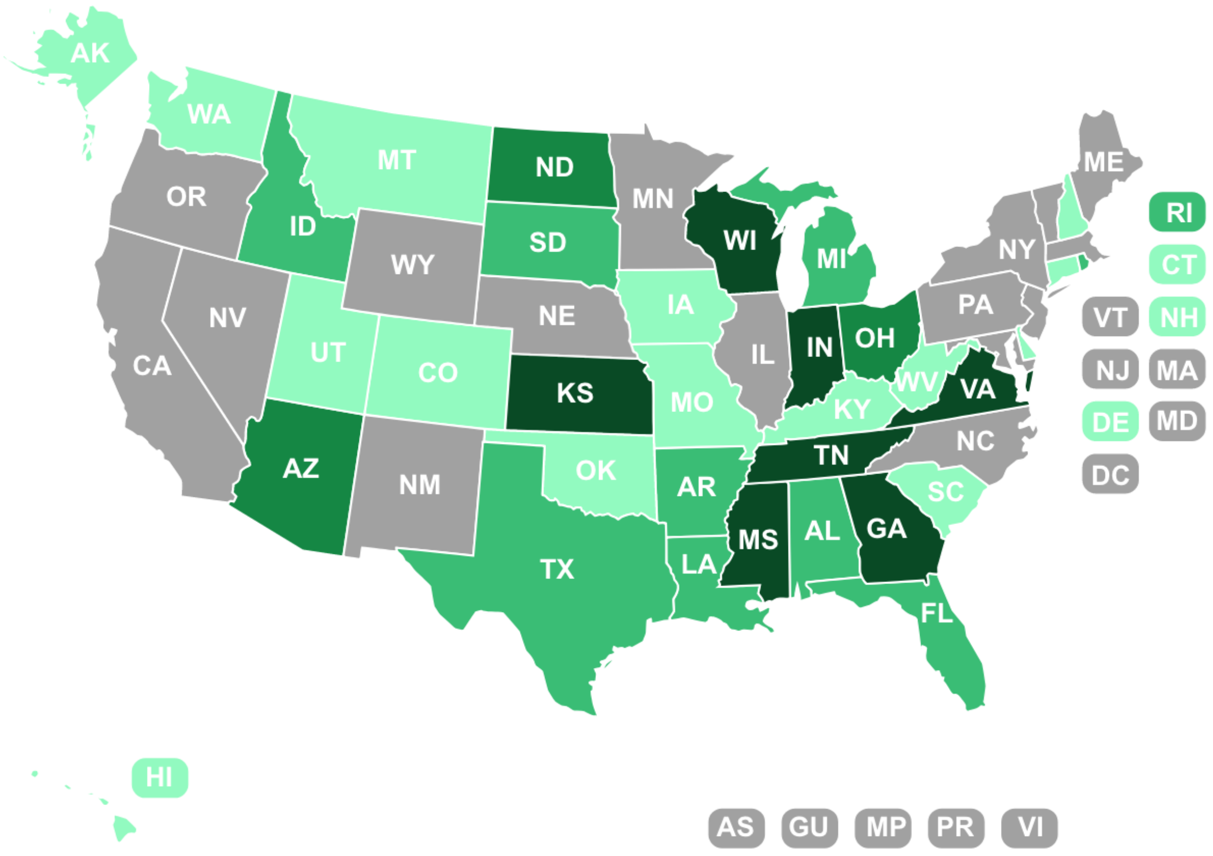
- Bills, bank statements, birth certificates, SS cards not good enough
- Student photo ID? Not good enough (no signature)
- Wisconsin: must have photo and a signature

Externalities?



Voter Identification Laws in Effect in 2019

Strict Photo ID	Strict Non-Photo ID	Photo ID requested	ID requested; photo not required	No document required to vote

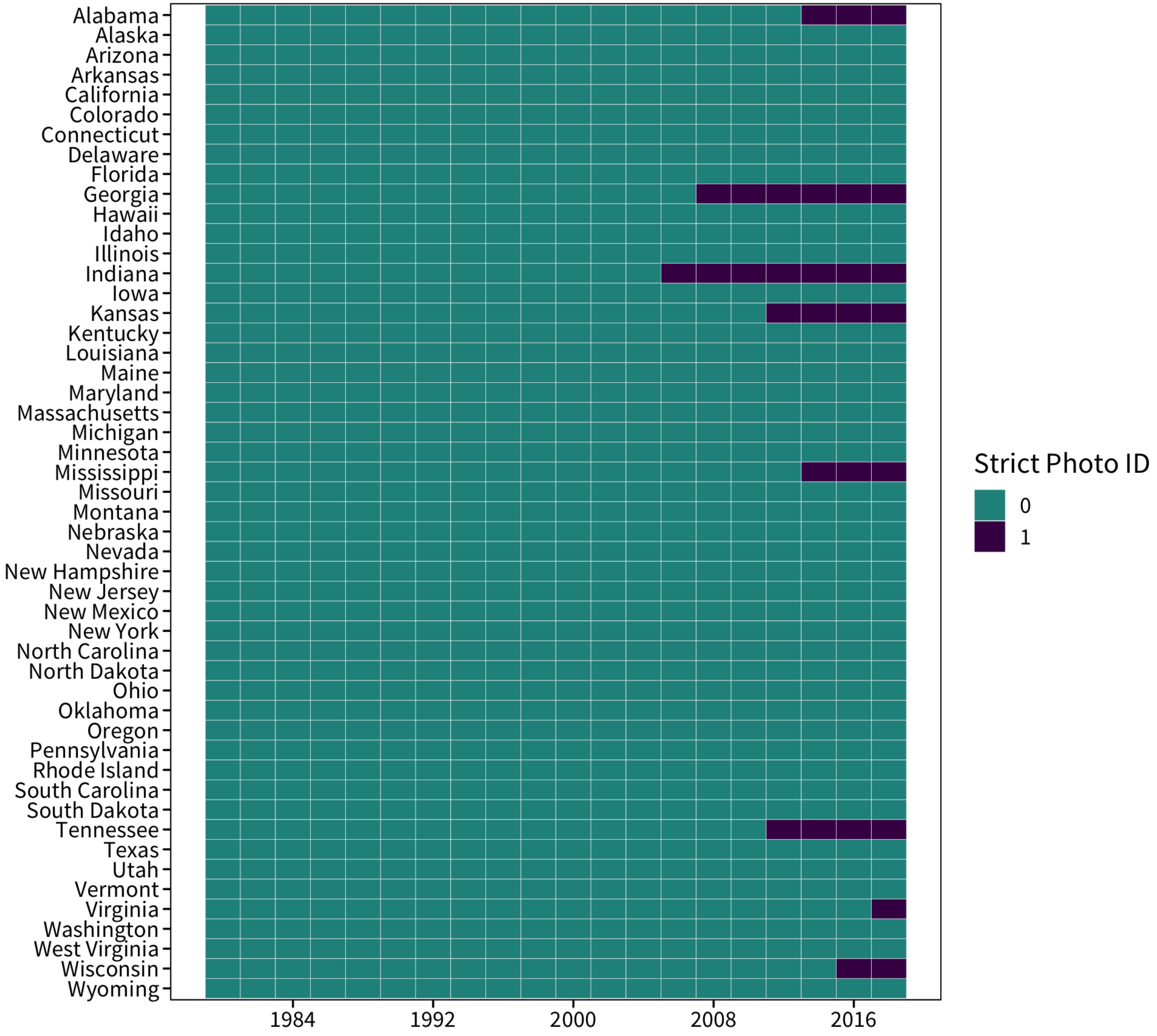


Do voter ID requirements reduce turnout?

Unit of analysis: states (state-years)

Independent variable: state has voter ID law (1) or not (0)

Dependent variable: turnout rate (% VEP)



Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming
1984
1992
2000
2008
2016
Strict Photo ID
0
1

Sticky Causality

Turnout in 2008

strict_id	turnout
0	63.3
1	60.8

Voter ID states (IN and GA) had 2.5 percent lower turnout in 2008

Sticky Causality

Turnout in 2008

strict_id	turnout
0	63.3
1	60.8

Voter ID states (IN and GA) had 2.5 percent lower turnout in 2008

Turnout in 2004, *before* any strict photo ID laws

strict_id	turnout
0	62.5
1	55.5

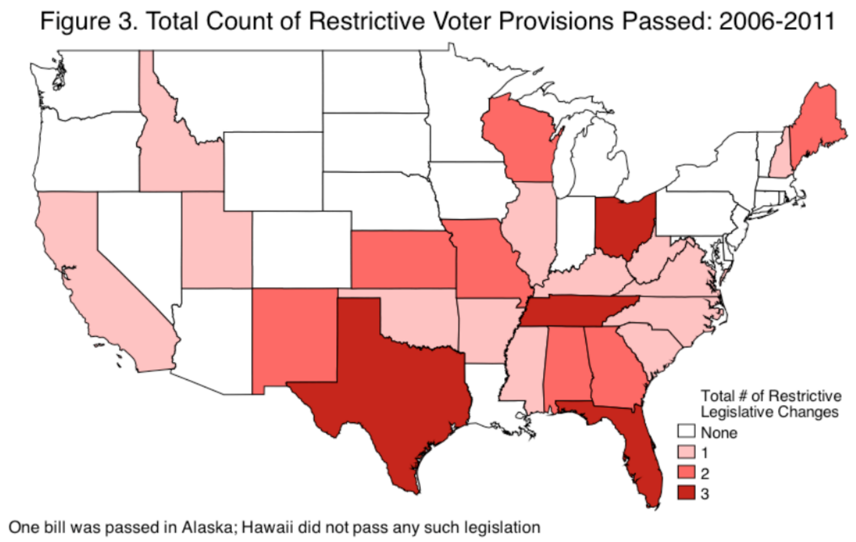
IL and GA had 7 percent lower turnout *before* strict photo ID

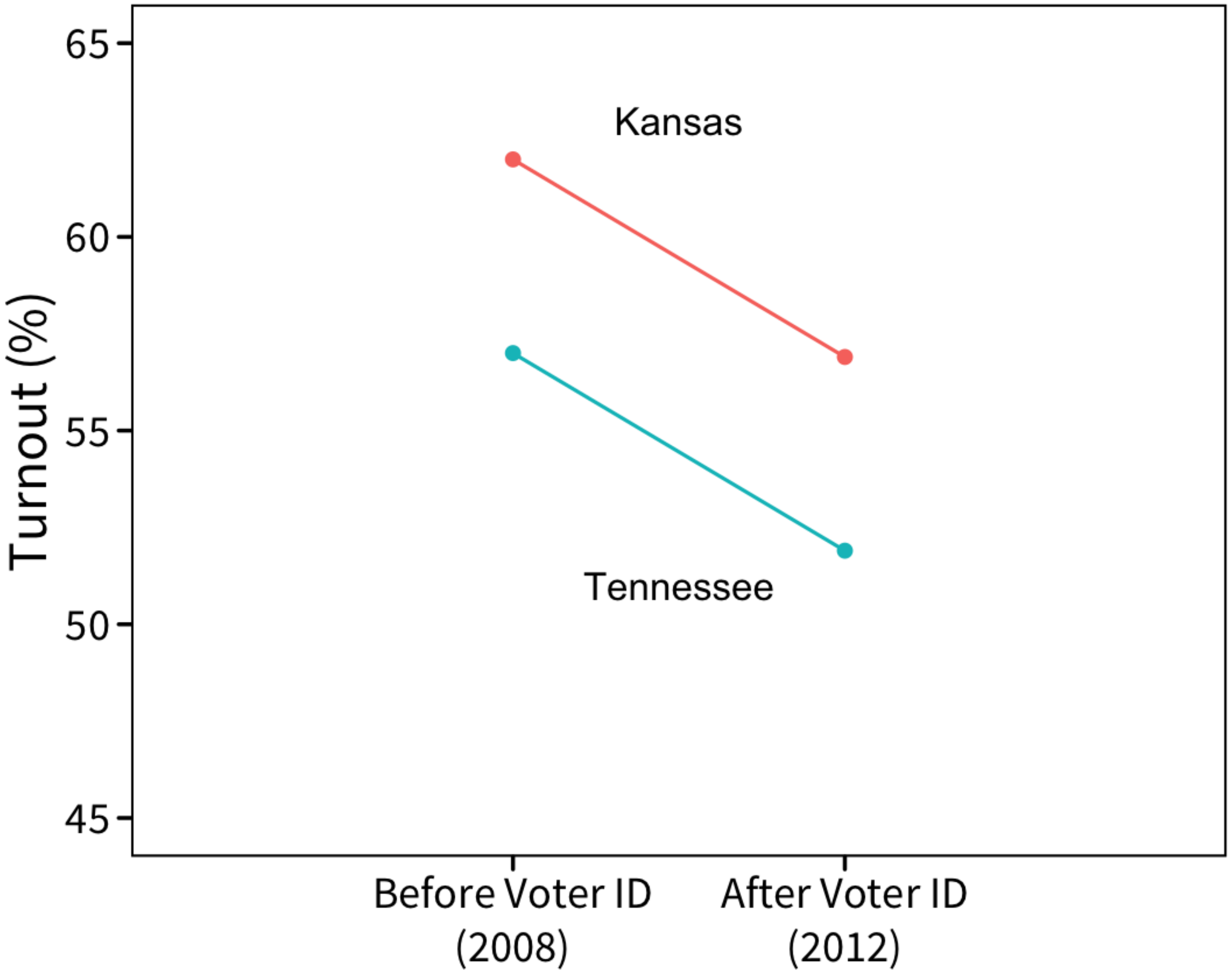
Confounding and selection bias

There's something about states (U) that make them adopt voter ID requirements (X)

Is that something related to turnout (Y) also?

We should compare states *against themselves*



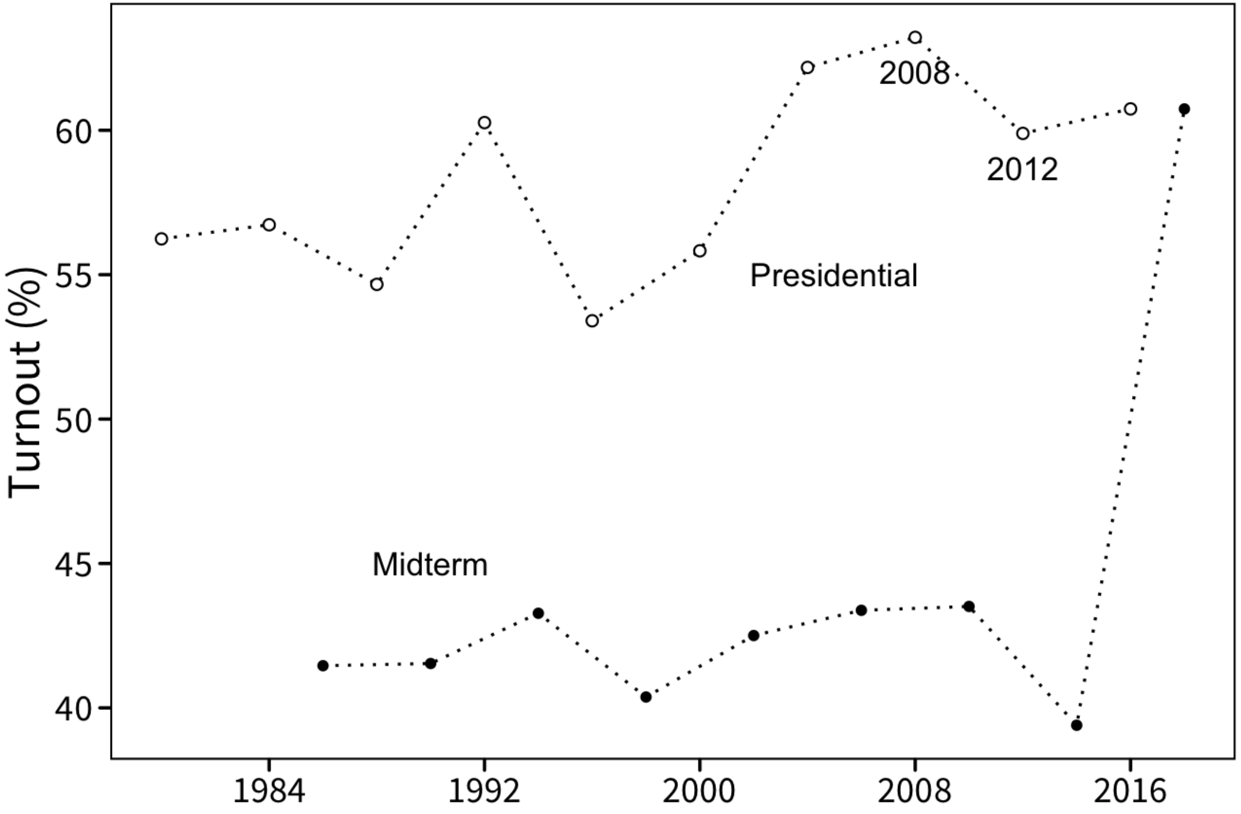

Kansas
Tennessee
Turnout (%)
65
60
55
50
45
Before Voter ID
(2008)
After Voter ID
(2012)

Turnout naturally fluctuates

Turnout goes down in voter ID states, so what?

Is that different from non-voter ID states?

We want to compare across time *and* across states



Difference-in-differences

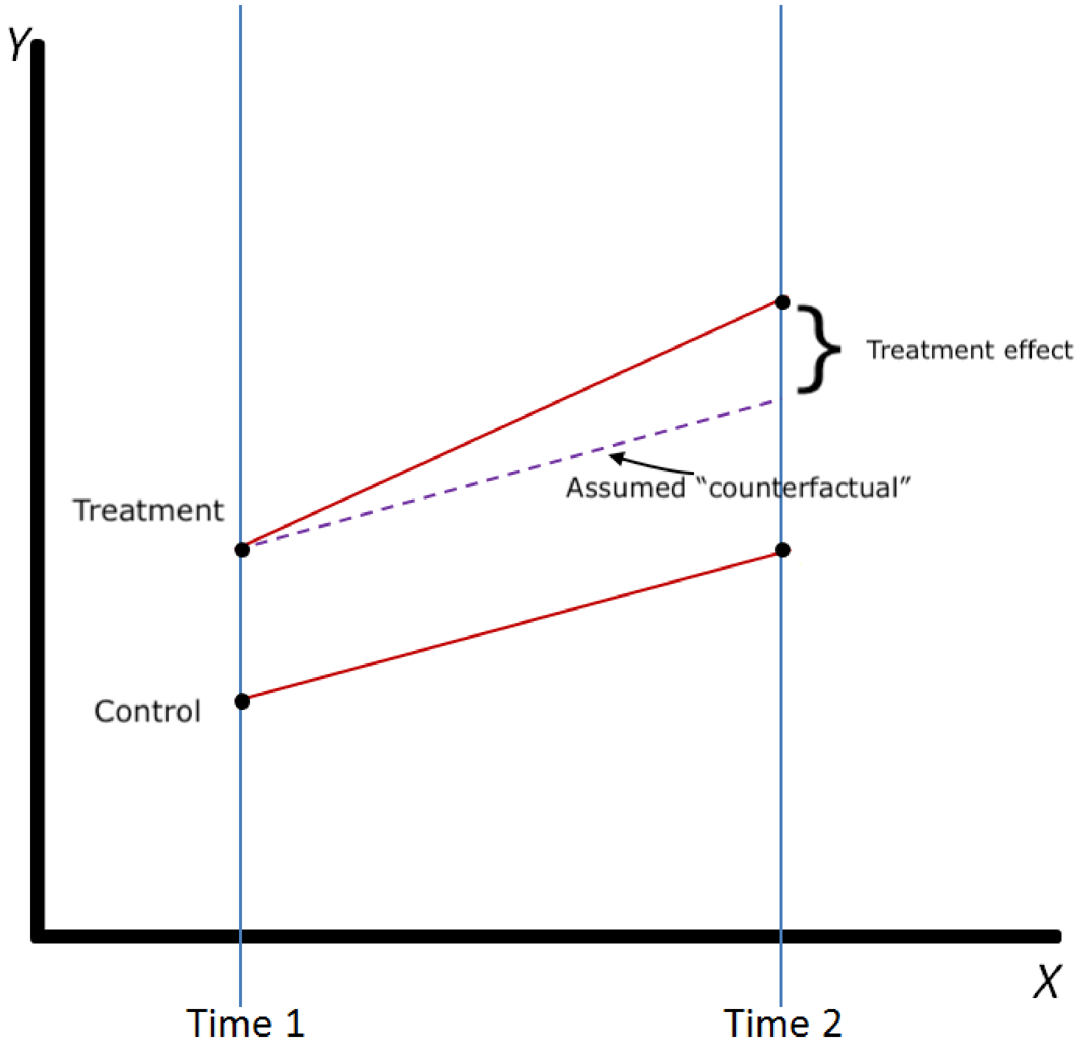
Background characteristics

Difference 1: difference over time

Difference 2: are treatment and control changes the same?

Counterfactuals

$$\hat{y} = a + b_1\,(\text{Treat}) + b_2\,(\text{After}) + b_3\,(\text{Treat} \times \text{After})$$



Understanding the Diff-in-diff model

$$\hat{y} = a + b_1\,(\text{Treat}) + b_2\,(\text{After}) + b_3\,(\text{Treat} \times \text{After})$$

Control group, before treatment (after = 0, treat = 0): a

Treatment group, before treatment (after = 0, treat = 1): $a + b_1$

Control group, after treatment (after = 1, treat = 0): $a + b_2$

Treatment group, after treatment (after = 1, treat = 1): $a + b_1 + b_2 + b_3$

Understanding the Diff-in-diff model

$$\hat{y} = a + b_1\,(\text{Treat}) + b_2\,(\text{After}) + b_3\,(\text{Treat} \times \text{After})$$

Control group, before treatment (after = 0, treat = 0): a

Treatment group, before treatment (after = 0, treat = 1): $a + b_1$

Control group, after treatment (after = 1, treat = 0): $a + b_2$

Treatment group, after treatment (after = 1, treat = 1): $a + b_1 + b_2 + b_3$

b_3 is the effect of treatment, controlling for over-time change ($b_2\text{After}$) and background characteristics of the treatment and control groups ($b_1\text{Treat}$)

With real data

$$\widehat{\text{turnout}} = a + d_{year} + d_{states} + b\,(\text{ID. Law})$$

With real data

$$\hat{\text{turnout}} = a + d_{year} + d_{states} + b\,(\text{ID. Law})$$

```
turnout_data
```

```
## # A tibble: 1,000 x 4
##     year state       turnout strict_id
##    <dbl> <chr>         <dbl>     <dbl>
##  1  1980 Alabama        49.2         0
##  2  1980 Alaska         58.7         0
##  3  1980 Arizona        46.2         0
##  4  1980 Arkansas       52           0
##  5  1980 California     55           0
##  6  1980 Colorado       57.2         0
##  7  1980 Connecticut    63.9         0
##  8  1980 Delaware       56           0
##  9  1980 Florida        52           0
## 10  1980 Georgia        42.1         0
## # … with 990 more rows
```

With real data

$$\widehat{\text{turnout}} = a + d_{year} + d_{states} + b\,(\text{ID. Law})$$

```
turnout_data
```

```
## # A tibble: 1,000 x 4
##     year state       turnout strict_id
##    <dbl> <chr>         <dbl>     <dbl>
##  1  1980 Alabama        49.2         0
##  2  1980 Alaska         58.7         0
##  3  1980 Arizona        46.2         0
##  4  1980 Arkansas       52           0
##  5  1980 California     55           0
##  6  1980 Colorado       57.2         0
##  7  1980 Connecticut    63.9         0
##  8  1980 Delaware       56           0
##  9  1980 Florida        52           0
## 10  1980 Georgia        42.1         0
## # … with 990 more rows
```

```
# within lm(), as.factor() makes a sequence of dummies
# one category is automatically omitted (alphanumerically)
diff_diff <-
  lm(turnout ~ strict_id + as.factor(state) + as.factor(year),
     data = filter(turnout_data))
```

```
summary(diff_diff)
```

Each state/year effect is comparison to omitted category (Alabama or 1980)

e.g. CA is (on avg) 2% higher turnout than AL

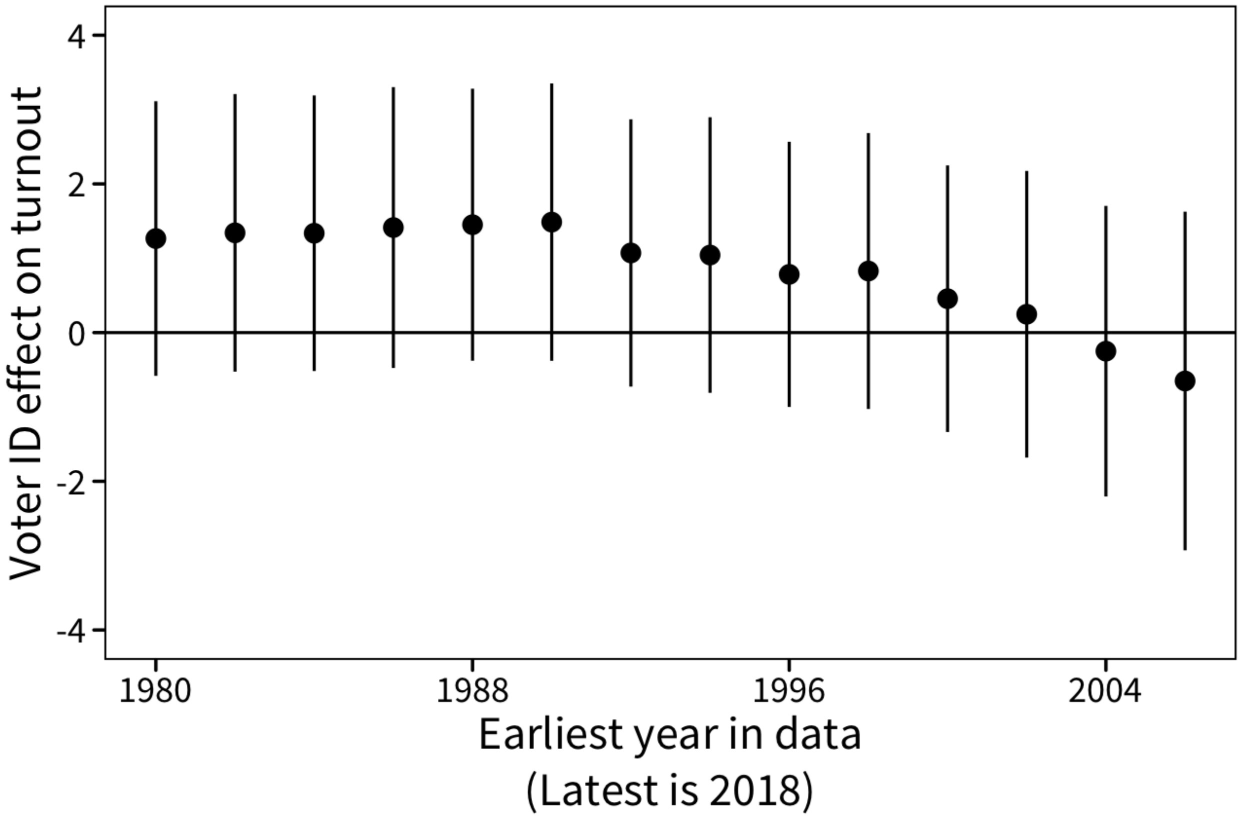
Turnout in 2008 was (on avg) 7 pts higher than 1980

Turnout under voter ID is (on avg) 1.2 pts higher (not sig.)

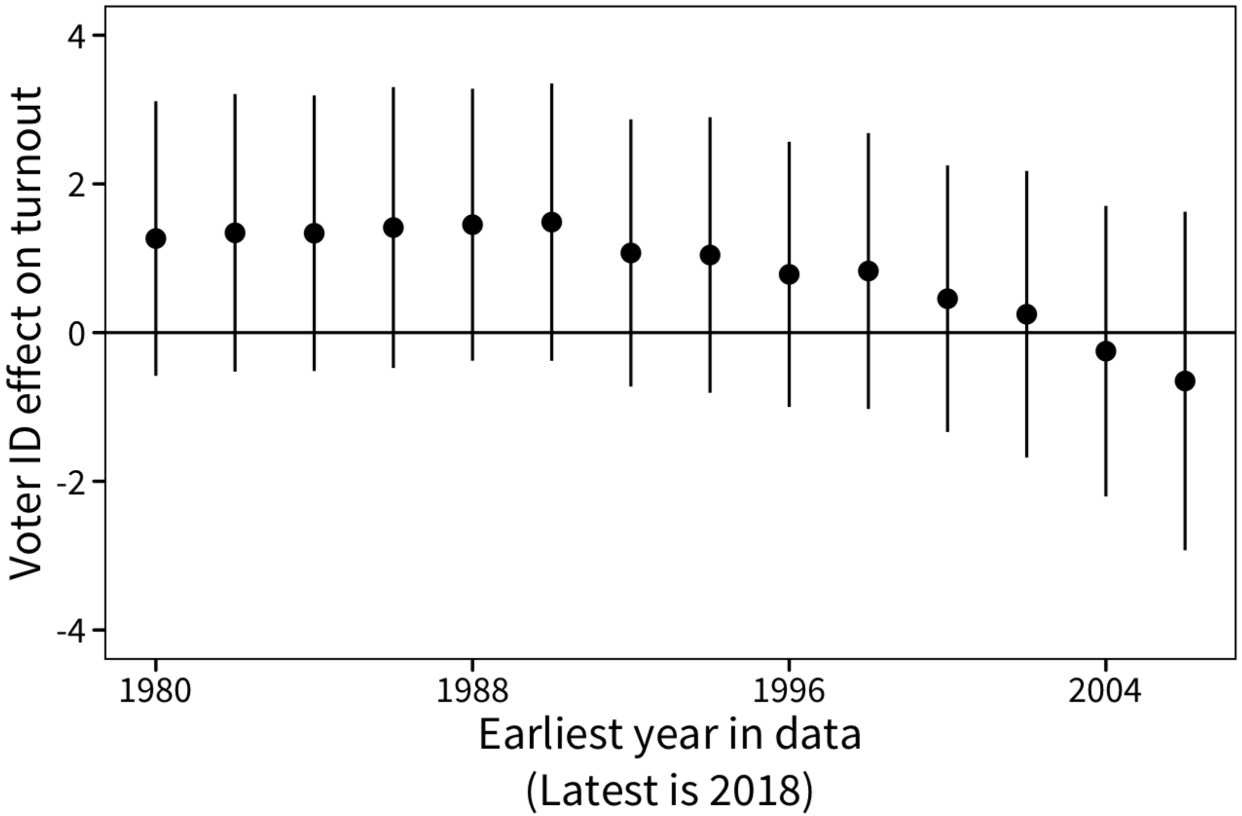
```
##
## Call:
## lm(formula = turnout ~ strict_id + as.factor(state) + as.factor(year)
##     data = filter(turnout_data))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.3247  -2.4630   0.1195   2.4472  15.0295
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     52.7747     1.1226  47.012  < 2e-16 *
## strict_id                        1.2652     0.9409   1.345 0.179076
## as.factor(state)Alaska          11.0648     1.3519   8.184 8.97e-16 *
## as.factor(state)Arizona         -2.8552     1.3519  -2.112 0.034959 *
## as.factor(state)Arkansas        -1.1102     1.3519  -0.821 0.411742
## as.factor(state)California       2.3348     1.3519   1.727 0.084504 .
## as.factor(state)Colorado         8.0348     1.3519   5.943 3.95e-09 *
## as.factor(state)Connecticut      7.6148     1.3519   5.632 2.35e-08 *
## as.factor(state)Delaware         2.7698     1.3519   2.049 0.040768 *
## as.factor(state)Florida          2.5298     1.3519   1.871 0.061629 .
## as.factor(state)Georgia         -4.9248     1.3519  -3.643 0.000285 *
## as.factor(state)Hawaii          -2.9402     1.3519  -2.175 0.029896 *
## as.factor(state)Idaho            7.5148     1.3519   5.558 3.56e-08 *
## as.factor(state)Illinois         3.6398     1.3519   2.692 0.007225 *
## as.factor(state)Indiana         -1.2780     1.3577  -0.941 0.346768
## as.factor(state)Iowa            10.0698     1.3519   7.448 2.16e-13 *
## as.factor(state)Kansas           3.8467     1.3454   2.859 0.004342 *
## as.factor(state)Kentucky        -2.0802     1.3519  -1.539 0.124222
```

Threats of time-varying confounders?

Threats of time-varying confounders?



Threats of time-varying confounders?



Counter-mobilization, Battleground states, Other legal changes

Exogenous vs endogenous policy change

Policy change comes from within the system (endogenous)

Policy change is imposed from external forces (exogenous)

Exogenous policy change are easier to study

Looking ahead

On Wednesday:

- experiments (and ethics)
- Discuss in-section presentation

Section:

- More multiple regression

Next week:

- Big data (by Michael)
- Elections, campaigns, polls...