

Difference in Difference Models

Bill Evans

1

Florida

- 8/25/1997, State of Florida settles out of court in their suits against tobacco manufacturers
- Awarded \$13 billion over 25 years
- Use \$200 to run anti-smoking campaign aimed at kids
- Florida Tobacco Pilot Program (FTPP)
- Precursor to the national 'truth' campaign

2

- **Florida's edgy "Truth" advertising campaign continues to have a significant impact in reducing teen smoking, a team of researchers concluded from a new study that examines the impact of the state's anti-tobacco advertising.**

3

- **in 1998, when surveillance began for tobacco use among Florida youth, 27.4 percent of high school students were current cigarette smokers. by 2000, this rates had declined to 22.6 among high school students.**
- 4.8 percentage point decline or a 17.5% reduction in teen smoking

4

Nationwide

- Teen smoking rates fell from 36.5 to 31.4%
- A 5.1 percentage point decline or roughly 14%

5

Random assignment clinical trial

- New drugs that lowers cholesterol
- Recruit N people with high cholesterol
 - ½ in treatment (receive active ingredient)
 - ½ in control (placebo)
- Measure cholesterol levels
 - before the start of treatment
 - Then again after a specified time

6

Difference in Difference

| | Before Treatment | 3 months later | Difference |
|----------------------------|------------------|----------------|--|
| Group 1 (Treatment) | Y_{t1} | Y_{t2} | $\Delta Y_t = Y_{t2} - Y_{t1}$ |
| Group 2 (Control) | Y_{c1} | Y_{c2} | $\Delta Y_c = Y_{c2} - Y_{c1}$ |
| Difference | | | $\Delta\Delta Y = \Delta Y_t - \Delta Y_c$ |

7

Difference in difference models

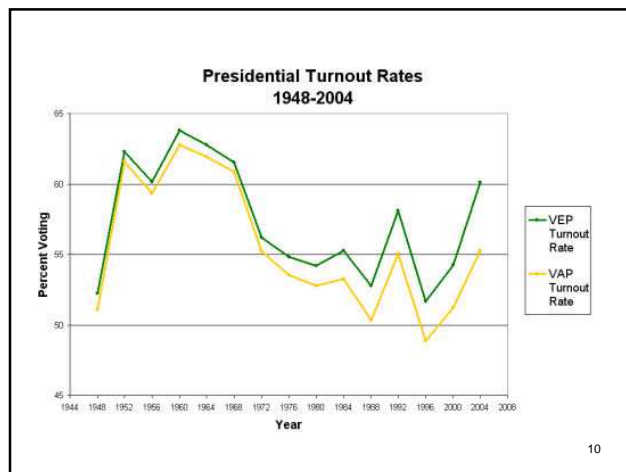
- Maybe the most popular “identification strategy” in applied statistical work in econometrics
- Attempts to mimic random assignment with treatment and “comparison” sample

8

Simple problem set up

- One group is 'treated' with intervention
- Have pre & post treatment data for group receiving intervention
- Can examine time-series changes but,
- Unsure how much of the change is due to secular changes

9



10

Example

- 1993 Federal government passes Motor Voter
 - Register to vote when you get drivers license
 - Designed to decrease the cost of voting
- Some states had state Motor Voter Law prior to 1993
- Suppose you compare outcomes in states before/after 1993 (1992 vs 1996 elections)

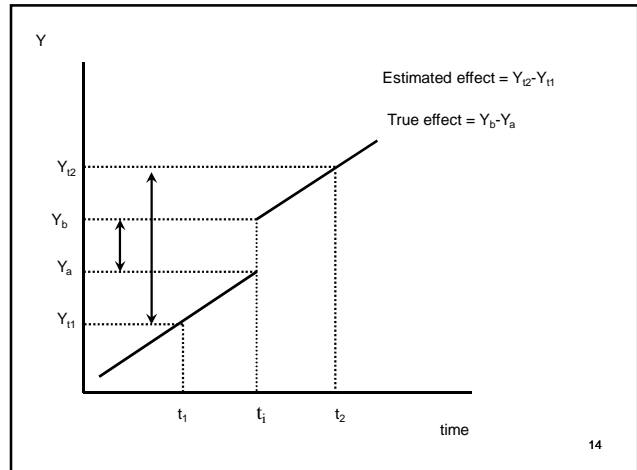
11

- States with new law saw an increase in voter registration of 8.4% points
 - 1992: 76.1%
 - 1996: 84.5%
- Question: how much of the increase was the law and how much was it secular trends?
- Cannot say without controlling for factors that impact these trends

12

- If the outcome of interest is trending over time, before/after comparisons will provide a biased estimate of the law
- Look at this graphically

13



14

- Intervention occurs at time period t_1
- True effect of law
– $Y_b - Y_a$
- Only have data at t_1 and t_2
– If using time series, estimate of the effectiveness of the law is $Y_{t1} - Y_{t2}$
- Solution?

15

Difference in difference models

- Pool cross-sectional and time series data
- Use time series of “untreated” group to establish “trends”
- What would have occurred in the treatment states in the absence of the intervention?

16

Difference in Difference

| | Before Change | After Change | Difference |
|-------------------|---------------|--------------|--|
| Group 1 (Treat) | Y_{t1} | Y_{t2} | ΔY_t $= Y_{t2} - Y_{t1}$ |
| Group 2 (Control) | Y_{c1} | Y_{c2} | ΔY_c $= Y_{c2} - Y_{c1}$ |
| Difference | | | $\Delta \Delta Y$ $\Delta Y_t - \Delta Y_c$ |

17

Motor Voter Example

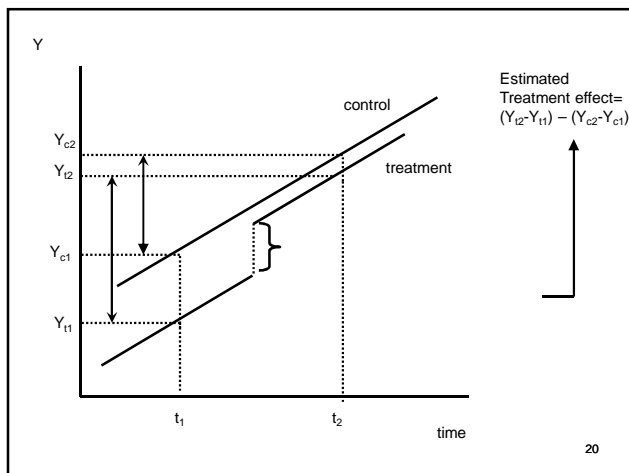
- Data in two years
 - 1992 Presidential (before MV)
 - 1996 Presidential (after)
- Two groups of states
 - Treated group (states that got MV through federal law in 1993)
 - Control group (states that had MV laws already)

18

Difference in Difference

| | Before MV | After MV | Difference |
|-------------------|-----------|----------|------------|
| Group 1 (Treat) | 0.761 | 0.845 | 0.084 |
| Group 2 (Control) | 0.834 | 0.867 | 0.033 |
| Difference | | | 0.050 |

19

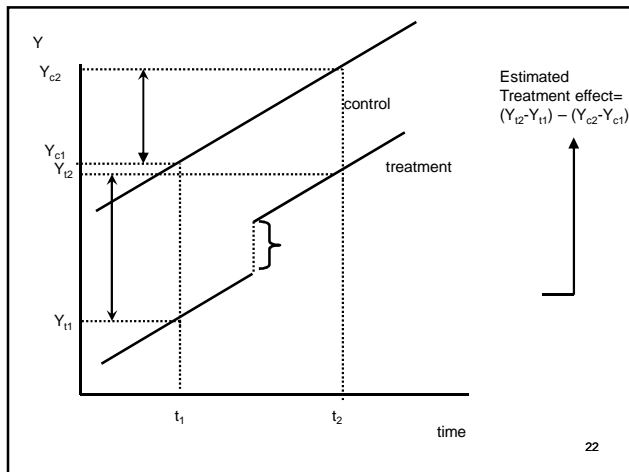


20

Key Assumption

- Control group identifies the time path of outcomes that would have happened in the absence of the treatment
- In this example, Y falls by $Y_{c2}-Y_{c1}$ even without the intervention
- Note that underlying 'levels' of outcomes are not important (return to this in the regression equation)

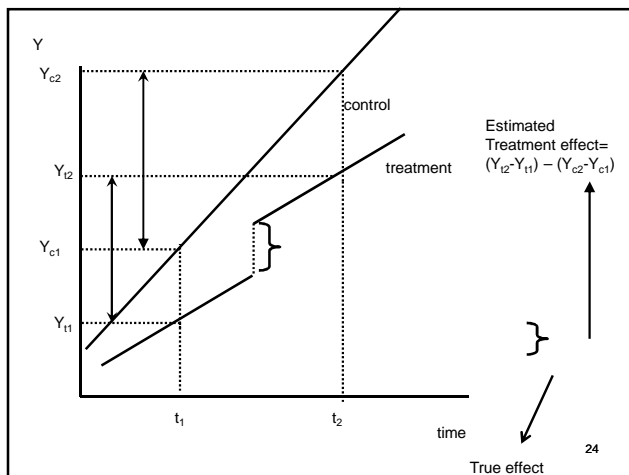
21



22

- In contrast, what is key is that the time trends in the absence of the intervention are the same in both groups
- If the intervention occurs in an area with a different trend, will under/over state the treatment effect
- In this example, suppose intervention occurs in area with faster falling Y

23



24

Basic Econometric Model

- Data varies by
 - state (i)
 - time (t)
 - Outcome is Y_{it}
- Only two periods
- Intervention will occur in a group of observations (e.g. states, firms, etc.)

25

- Three key variables
 - $T_{it}=1$ if obs i belongs in the state that will eventually be treated
 - $A_{it}=1$ in the periods when treatment occurs
 - $T_{it}A_{it}$ -- interaction term, treatment states after the intervention
- $Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 A_{it} + \beta_3 T_{it}A_{it} + \epsilon_{it}$

26

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 A_{it} + \beta_3 T_{it}A_{it} + \epsilon_{it}$$

| | Before Change | After Change | Difference |
|----------------------|------------------|-----------------|------------|
| Group 1 (Treat) | | | |
| Group 2 (Control) | | | |
| Difference | | | |

27