Difference-in-Difference Models

Bill Evans
Health Economics

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 $200m to run anti-smoking campaign aimed at kids
- Florida Tobacco Pilot Program (FTPP)
- Precursor to the national 'truth' campaign

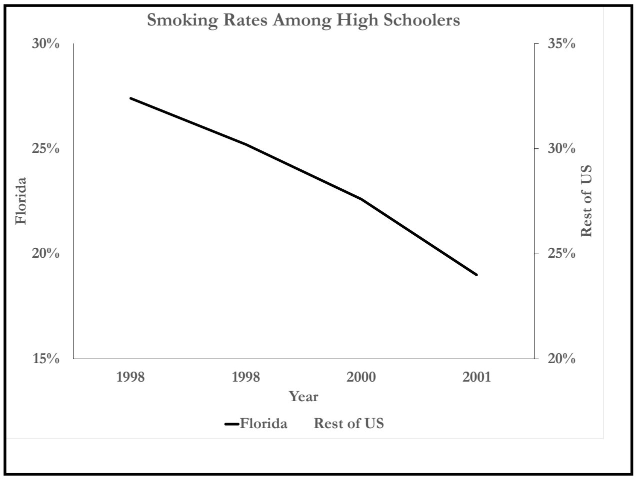
2

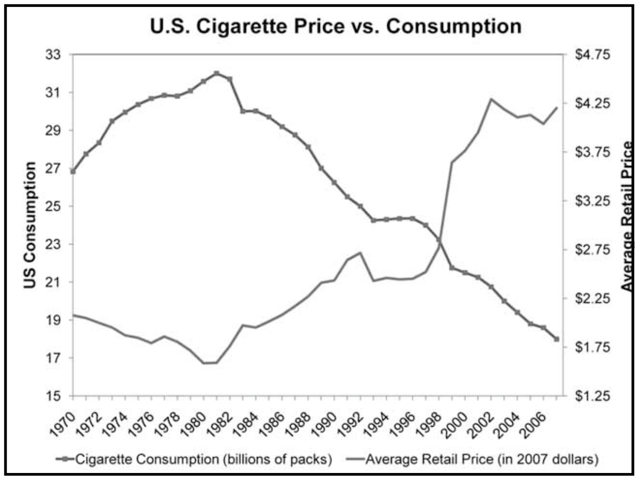
- Florida's edgy "Truth" advertising campaign continues to have a significant impact in reducing teen smoking, a team of researchers concluded from a new study that examines the impact of the state's anti-tobacco advertising.

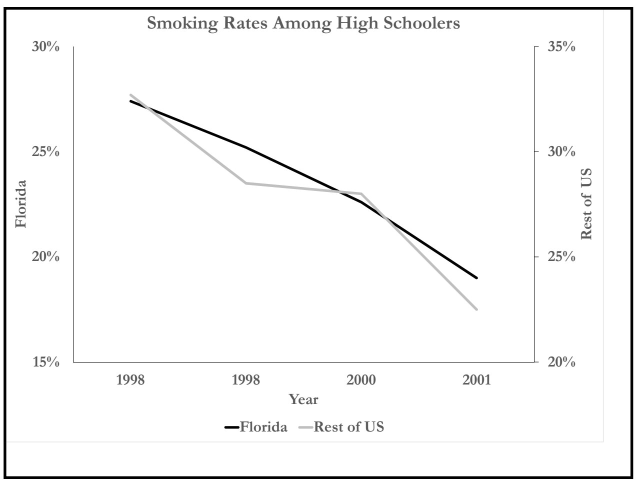
3

- 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%
- Rates in Florida fell by 4.8 percentage points – similar to what was happening in the nation as a whole

8

Difference in difference models

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

9

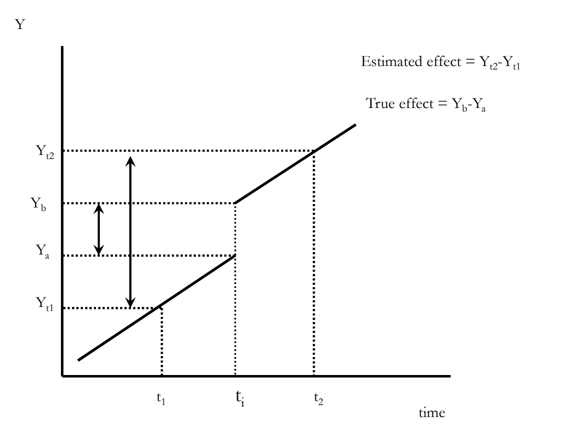
Simple problem set up

- 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

10

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

11



12

- 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?

13

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?

14

Difference in Difference

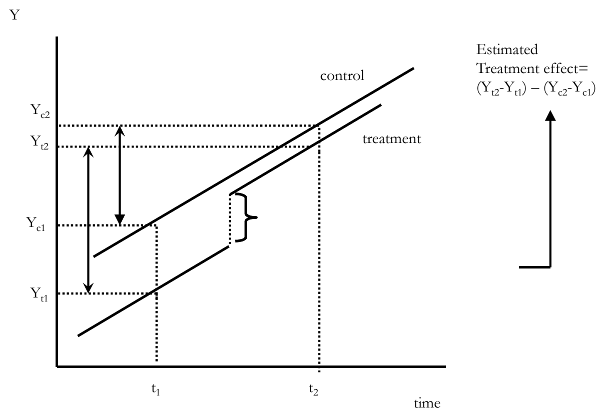
	Before Change	After Change	Difference
Group 1 (Treat)	Y_{t1}	Y_{t2}	$\Delta Y_t = Y_{t2}-Y_{t1}$
Group 2 (Control)	Y_{c1}	Y_{c2}	$\Delta Y_c = Y_{c2}-Y_{c1}$
Difference			$\Delta\Delta Y$ $\Delta Y_t – \Delta Y_c$

15

Motor Voter Example

- Federal law change in 1993 – allows people to register when they get theior driver's license
- Designed to increase voter registration
- Some states had motor voter before 1993
- Data on voting rates in 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)

16

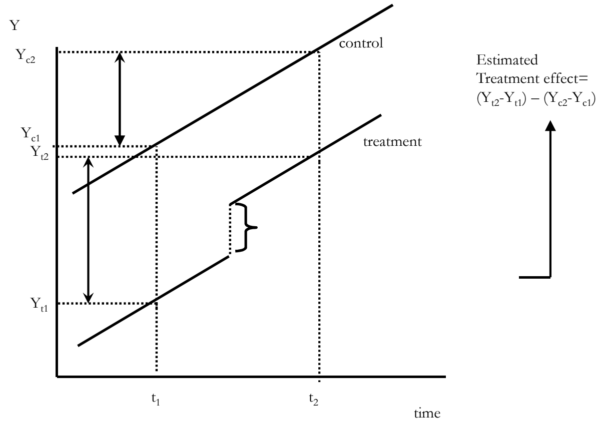


17

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)

18



19

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.)

20

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

21

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

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

22

Making the model more complicated

- So far, a very simple model
  - Two groups
  - Two periods
- However, the "treatment" may cover more than 1 group
- The treatment may happen at very different time periods across groups
- How to generalize this type of model for
  - Many treatments
  - Multiple groups being treated

23

Example: States as laboratories

- Tremendous variation across states in their laws
  - Variation across states in any given year
  - Variation over time within a state
- Examples
  - Minimum wages, welfare policy, Medicaid coverage, traffic safety laws, use of death penalty, drinking age, cigarette taxes

24

Panel Data at the State Level

- Data is in two dimensions
- y_{it} outcome for state i in year t
  - $i=1,2,\ldots n$
  - $t=1,2,\ldots T$
- Example: all states from 1990-2019
  - 30 years
  - 51 states
  - 1530 obs.

Empirical example: Motorcycle Helmet laws

- 1967, Feds require states to have helmet law to get all federal highway money
- By 1975, all states have qualifying law
- 1976, Congress responds to state pressure and eliminate penalties
  - 20 states weaken their law and only require coverage for teens
  - 8 states repeal law completely

26

- 1991 Federal law again provides incentives for laws covering everyone
  - A bunch of states pass universal laws
- Congress changes its mind and in 1995 eliminate penalties
  - Again many states drop the law
- Currently
  - 20 states have universal law
  - 27 have teen coverage only

27

- Helmets are estimated to reduce the likelihood of death in a motorcycle crash by 37%. (Center for Disease Control)
- http://www.cdc.gov/motorvehiclesafety/pdf/mc2012/MotorcycleSafetyBook.pdf

28

Problem

- Time series correlation
  - many laws came into effect during a boom/recession
  - Motorcycle fatalities are pro-cyclic
  - Need to control for the time series
- Motorcycle drivers HATE helmet laws
  - Laws are much less prevalent in states with lots of motorcyclists
  - Simple political economy

$define:$

$$i = 1, 2, \ldots n; \quad t = 1, 2, \ldots T$$

$$S_i = 1 \; if \;\; state \; i, = 0 \; otherwise$$

$$W_t = 1 \; if \;\; year \; t, = 0 \; otherwise$$

$$x_{it} = some \; \text{controls}$$

$$Law_{it} = 1 \; if \;\; state \; i \; has \; helmet \; law$$

$$in \; year \; t, = 0 \; otherwise$$

$$y_{it} = \beta_0 + LAW_{it}\beta_1 + x_i\beta_2 + \sum_{j=2}^{n} S_j\alpha_j + \sum_{k=2}^{T} W_k\lambda_k + \varepsilon_{it}$$

30

- Why k=2 to N and j=2 to T?
- What does α measure?
- What does λ measure?

31

- Question: impact of MC helmet laws on motorcycle fatalities
- Data: 48 states, 18 years (1988-2005), 864 observations
- Outcome ln(motor cycle death rate)
  - Death rates = deaths/100,000 population
- Treatment variable: =1 if state i has a motor cycle law in year t, =0 otherwise

32

```
Contains data from motorcycles.dta
  obs:           864
 vars:            12                          10 Nov 2012 09:27
 size:        49,248 (99.6% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
year            int    %9.0g                  year
mcfatals        double %9.0g                  total motor cycle fatalities
state           str2   %2s                    2 digit postal code, AL, CA,
                                                etc.
fips            byte   %8.0g                  2 digit numeric fips code
helmet_law      float  %9.0g                  =1 if motorcycle helmet law, =0
                                                otherwise
speed65         float  %9.0g                  =1 if speed limit is 65, 0
                                                otherwise
speed70p        float  %9.0g                  =1 if speed limit is 70 plus, 0
                                                otherwise
bac_10          float  %9.0g                  drunk driving defined as
                                                bac>=0.1, =0 otherwise
bac_08          float  %9.0g                  drunk driving defined as
                                                bac>=0.08, =0 otherwise
unemp           float  %9.0g                  state unemployment rate, 5 is 5%
population      float  %9.0g                  state population
mregs           double %10.0g                 motor cycle registrations
-------------------------------------------------------------------------------
Sorted by:
```

33

```
.* construct dummy variables for state and year
. xi i.state i.year
i.state           _Istate_1-48        (_Istate_1 for state==AL omitted)
i.year            _Iyear_1988-2005    (naturally coded; _Iyear_1988 omitted)
```

34

```
. * run the difference  in difference model
. reg mcdrl speed65 speed70p unemp bac_08 bac_10 _I* helmet_law

      Source |       SS       df       MS              Number of obs =     864
-------------+------------------------------           F( 70,   793) =   30.54
       Model |  139.812929    70  1.99732756           Prob > F      =  0.0000
    Residual |  51.8558902   793  .065392043           R-squared     =  0.7295
-------------+------------------------------           Adj R-squared =  0.7056
       Total |   191.66882   863  .222095967           Root MSE      =  .25572

------------------------------------------------------------------------------
       mcdrl |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     speed65 |  -.0577686   .0552537    -1.05   0.296    -.1662293    .0506922
    speed70p |  -.0855586   .0815308    -1.05   0.294    -.2456004    .0744831
       unemp |  -.0117339   .0118625    -0.99   0.323    -.0350195    .0115517
      bac_08 |   .1423512   .0725064     1.96   0.050     .0000241    .2846783
      bac_10 |   .1163134   .0628129     1.85   0.064    -.0069859    .2396127
   _Istate_2 |   -.038139   .0889074    -0.43   0.668    -.2126606    .1363826
              delete some results
  _Istate_47 |   .2392712   .0896769     2.67   0.008     .0632391    .4153033
  _Istate_48 |   .3987819     .09788     4.07   0.000     .2066474    .5909164
 _Iyear_1989 |  -.2367341    .052373    -4.52   0.000    -.3395401   -.1339281
              delete some results
 _Iyear_2005 |   .1032509   .0703676     1.47   0.143    -.0348778    .2413796
  helmet_law |  -.3728078   .0458932    -8.12   0.000    -.4628943   -.2827213
        cons |   .5393718   .1275965     4.23   0.000     .2889049    .7898387
------------------------------------------------------------------------------
```

35