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

- 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 rate had declined to 22.6 among high school students.
- Note: 4.8 percentage point decline or a 17.5% reduction in teen smoking

- 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.

Difference in Difference Models

Bill Evans
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 – rates nationwide fell by a similar amount
Random assignment clinical trial

- New drug 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

<table>
<thead>
<tr>
<th></th>
<th>Before Treatment</th>
<th>3 months later</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 (Treatment)</td>
<td>$Y_{t1}$</td>
<td>$Y_{t2}$</td>
<td>$\Delta Y_t = Y_{t2} - Y_{t1}$</td>
</tr>
<tr>
<td>Group 2 (Control)</td>
<td>$Y_{c1}$</td>
<td>$Y_{c2}$</td>
<td>$\Delta Y_c = Y_{c2} - Y_{c1}$</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td>$\Delta\Delta Y$</td>
<td>$\Delta Y_t - \Delta Y_c$</td>
</tr>
</tbody>
</table>

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

Simple problem set up

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

• 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

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

- 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?

Difference in Difference

<table>
<thead>
<tr>
<th>Group 1 (Treat)</th>
<th>Before Change</th>
<th>After Change</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_{t1}$</td>
<td>$Y_{t2}$</td>
<td>$\Delta Y_t$</td>
<td>$= Y_{t2} - Y_{t1}$</td>
</tr>
<tr>
<td>Group 2 (Control)</td>
<td>$Y_{c1}$</td>
<td>$Y_{c2}$</td>
<td>$\Delta Y_c$</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td></td>
<td>$\Delta \Delta Y$</td>
</tr>
</tbody>
</table>

- 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?
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)

Difference in Difference

<table>
<thead>
<tr>
<th></th>
<th>Before MV</th>
<th>After MV</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>0.761</td>
<td>0.845</td>
<td>0.084</td>
</tr>
<tr>
<td>Group 2</td>
<td>0.834</td>
<td>0.867</td>
<td>0.033</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td></td>
<td>0.050</td>
</tr>
</tbody>
</table>

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)
• 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.

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

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

<table>
<thead>
<tr>
<th>Before</th>
<th>Change</th>
<th>After</th>
<th>Change</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>(Treat)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>(Control)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\Delta\Delta Y = \beta_3
\]

Meyer et al.

• Workers’ compensation
  – State run insurance program
  – Compensate workers for medical expenses and lost work due to on the job accident
• Premiums
  – Paid by firms
  – Function of previous claims and wages paid
• Benefits -- % of income w/ cap

• Typical benefits schedule
  – $\text{Min}(pY,C)$
  – $p=$percent replacement
  – $Y =$ earnings
  – $C =$ cap
  – e.g., 65% of earnings up to $400/week
• Concern: Moral hazard. Benefits will discourage return to work
• Empirical question: duration/benefits gradient
• Previous estimates
  \[ Y_i = \beta_0 + \beta_1 X_i + \beta_2 R_i + \epsilon_i \]
  - Y (duration)
  - R (replacement rate)
  - X (represents some other controls)
• Expect \( \beta_2 > 0 \) (Higher replacement, longer duration)

• Problem: Does realization of \( \epsilon_i \) convey any information about R?
  - Workers with longer duration tend to be higher income workers
  - They also have lower replacement rates (earnings put them above the weekly cap)
• \( \text{Cov}(R, \epsilon) < 0 \)
• Is \( \beta_2 \) over or under estimated?

Solution
• Quasi experiment in KY and MI
• Increased the earnings cap
  - Increased benefit for high-wage workers
    • (Treatment)
  - Did nothing to those already below original cap
    (comparision)
• Compare change in duration of spell before and after change for these two groups

\[\text{FIGURE 1. TEMPORARY TOTAL BENEFIT SCHEDULE BEFORE AND AFTER AN INCREASE IN THE MAXIMUM WEEKLY BENEFIT}\]
Data from Meyer et al.

- Data set kentucky.dta
- Key variables
  - durat (duration)
  - highearn (a high earning worker (treatment))
  - afchnge (after the law change)

* generate log duration
  . gen ldurat=ln(durat)
* sort the data by highearn and afchnge
  . sort highearn afchnge
* gets means of ldurat for
  . 2x2 table
  . by highearn afchnge: sum ldurat

<table>
<thead>
<tr>
<th>highearn = 0, afchnge = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>ldurat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>highearn = 0, afchnge = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>ldurat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>highearn = 1, afchnge = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>ldurat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>highearn = 1, afchnge = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>ldurat</td>
</tr>
</tbody>
</table>

Difference in Difference

<table>
<thead>
<tr>
<th>Mean average ln(duration)</th>
<th>Before change</th>
<th>After change</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>High earn (Treat)</td>
<td>1.356</td>
<td>1.599</td>
<td>0.243</td>
</tr>
<tr>
<td>Low earn (Control)</td>
<td>1.123</td>
<td>1.137</td>
<td>0.014</td>
</tr>
<tr>
<td>Difference</td>
<td>0.233</td>
<td>0.462</td>
<td>0.229</td>
</tr>
</tbody>
</table>
Model

- $Y_{it} = \text{duration of spell on WC}$
- $A_{it} = \text{period after benefits hike}$
- $H_{it} = \text{treated or high earnings group (Income>E_3)}$

- $Y_{it} = \beta_0 + H_{it}\beta_1 + A_{it}\beta_2 + A_{it}H_{it}\beta_3 + \varepsilon_{it}$
- Diff-in-diff estimate is $\beta_3$

More general model (allow for covariates)

- $Y_{it} = \beta_0 + H_{it}\beta_1 + A_{it}\beta_2 + A_{it}H_{it}\beta_3 + X_{1it}\alpha_1 + X_{2it}\alpha_2 + X_{3it}\alpha_3 + \ldots + X_{kit}\alpha_k + \varepsilon_{it}$

Results controlling for other covariates

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Kentucky (i)</th>
<th>Michigan (ii)</th>
<th>Kentucky (iii)</th>
<th>Michigan (iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>After-increase</td>
<td>0.016</td>
<td>-0.004</td>
<td>0.082</td>
<td>0.003</td>
</tr>
<tr>
<td>indicator variable</td>
<td>(0.045)</td>
<td>(0.038)</td>
<td>(0.084)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>High-earnings-group</td>
<td>-1.522</td>
<td>-0.594</td>
<td>5.577</td>
<td>3.607</td>
</tr>
<tr>
<td>indicator variable</td>
<td>(1.099)</td>
<td>(0.930)</td>
<td>(4.811)</td>
<td>(4.162)</td>
</tr>
<tr>
<td>After-increase x</td>
<td>0.215</td>
<td>0.162</td>
<td>0.157</td>
<td>0.203</td>
</tr>
<tr>
<td>high-earnings-group</td>
<td>(0.069)</td>
<td>(0.059)</td>
<td>(0.153)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>indicator variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Replicate results in Table 6, column (i)

```
reg ldurat highearn afchnge treat male married lage lprewage high_x_lpre _I*
Source |     SS     df        MS                      Number of obs = 5347
-------------+--------------------------------------------- F( 17, 5329) = 16.09
Model | 424.366738    17  24.9627493           Prob > F      = 0.0000
Residual | 8268.37878  5329  1.55158168           R-squared     =  0.0488
-------------+--------------------------------------------- Adj R-squared =  0.0458
Total | 8692.74552  5346  1.62602797           Root MSE      =  1.2456

------------------------------------------------------------------------------
ldurat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  highearn |  -1.522196   1.099035    -1.39   0.166    -3.676755    .6323633
  afchnge |   .0155081   .0447622     0.35   0.729     -.072244    .1032603
    treat |   .2146878   .0693106     3.10   0.002     .0788106     .350565
     male |  -.0722981    .046195    -1.57   0.118    -.1628593    .0182631
   married |   .0509362   .0409132     1.24   0.213    -.0292704    .1311427
     lage |   .2522586   .0522575     4.83   0.000     .1498124    .3547047
    lprewage |   .2582666   .1038422     2.49   0.013     .0546934    .4618397
high_x_lpre |   .2318765    .187021     1.24   0.215    -.1347612    .5985142
  _Iindustry_2 |   .2488065   .0593033     4.20   0.000     .1325478      1.0853
  _Iindustry_3 |   .1725146   .0410894     4.15   0.000     .0906449    .2540444
  _Injury_2 |    .780188   .155637     5.01   0.000     .4750758      1.0853

Questions to ask?

• What parameter is identified by the quasi-experiment? Is this an economically meaningful parameter?
• What assumptions must be true in order for the model to provide an unbiased estimate of $\beta_3$?
• Do the authors provide any evidence supporting these assumptions?

Minimum wage laws

• Minimum wage laws imposed by state, local and Federal governments
• “covered” sector includes most jobs
• States/locals can raise but not lower Federal minimum wage
### Federal Minimum Wages

- 01/01/1981 $3.35
- 04/01/1990 $3.80
- 04/01/1991 $4.25
- 10/01/1996 $4.75
- 09/01/1997 $5.15
- 07/24/2007 $5.85
- 07/24/2008 $6.55
- 07/24/2009 $7.25

### Some State Minimum Wage Laws

- WA $8.55
- OR $8.40
- CT/DC $8.25
- VT $8.04
- IL/MA/CA $8.00

---

### Textbook model of Minimum Wage

- Original conditions: $W_c$, $I_d$
- Minimum wage imposed, $W_m > W_c$
- Labor supply: higher wage encourages more work – labor supply increases to $I_s$
- Labor demand: higher wage is a shift along the demand curve to $I_d$
- New unemployment rate: $I_s - I_d$
- Job loss from minimum wage: $I_c - I_d$
Research Question?

- What happens to labor demand when minimum wage laws increased?
- Economic significance: test of theory of demand
- Policy significance: key question faced by lawmakers every time there is a proposed change in the minimum wage law.

NJ Minimum Wage Hike

- Federal MW stuck at $3.35 for most of the 90s
- Because of inflation, real value of MW fell considerably
- Nov 1989 law raised MW in 2 steps
  - To $3.80 on 4/1/90
  - To $4.25 on 4/1/91
- NJ law
  - Passed in early 1990
  - Went into effect April 1, 1992
  - Raised minimum wage from $4.25 - $5.05/hr, 18% increase

Questions

- In 1992, NJ slipped into a recession
- In March of 1992, State legislature voted to phase it in over two years,
  - Governor vetoed
  - Vote margin not large enough to override veto
- Law went into effect as planned
Why fast food industry?

• Leading employer of low wage workers
  – 25% employees in restaurant industry

• Fast food employers comply with Min wage laws
  – Their costs increase as a result of law

• Easy to generate a sample frame – use a phone book

• No tips paid to workers – so wage costs are known

Research methodology

• Examine employment before and after law goes into effect in NJ fast food restaurants

• Compare this change to changes in employment for employers not impacted by law
  – Fast food restaurants in PA
  – “Control group”

• Telephone interview of fast food restaurants before law goes into effect

• Ask store manager for basic information
  – Employees (full and part time)
  – Wages
  – Price of a basic meal

• Re-survey the same stores in November

Table 1: Sample Frame

<table>
<thead>
<tr>
<th></th>
<th>NJ Stores</th>
<th>PA Stores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contacted</td>
<td>Interview</td>
</tr>
<tr>
<td>Wave 1</td>
<td>364</td>
<td>331</td>
</tr>
<tr>
<td>Wave 2</td>
<td>331</td>
<td>321</td>
</tr>
</tbody>
</table>
Notes about sample

- Restaurants from 4 chains -- BK, KFC, Roy’s, Wendy’s – no McDonald’s
- Key outcome, Full time equivalents
  - FTE
  - FTE = Full time + .5* halftime

Table 2—Means of Key Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>NJ</th>
<th>PA</th>
<th>t*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Distribution of Store Types (percentages):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Burger King</td>
<td>41.1</td>
<td>44.3</td>
<td>−0.5</td>
</tr>
<tr>
<td>b. KFC</td>
<td>20.5</td>
<td>15.2</td>
<td>1.2</td>
</tr>
<tr>
<td>c. Roy Rogers</td>
<td>24.8</td>
<td>21.5</td>
<td>0.6</td>
</tr>
<tr>
<td>d. Wendy’s</td>
<td>13.6</td>
<td>19.0</td>
<td>−1.1</td>
</tr>
<tr>
<td>e. Company-owned</td>
<td>34.1</td>
<td>35.4</td>
<td>−0.2</td>
</tr>
</tbody>
</table>

2. Means in Wave 1:

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. FTE employment</td>
<td>20.4</td>
<td>23.3</td>
<td>−2.0</td>
</tr>
<tr>
<td>b. Percentage full-time employees</td>
<td>32.8</td>
<td>35.6</td>
<td>−0.7</td>
</tr>
<tr>
<td>c. Starting wage</td>
<td>4.61</td>
<td>4.63</td>
<td>−0.4</td>
</tr>
<tr>
<td>d. Wage =$4.25 (percentage)</td>
<td>30.5</td>
<td>32.5</td>
<td>−0.4</td>
</tr>
<tr>
<td>e. Price of full meal</td>
<td>3.35</td>
<td>3.04</td>
<td>0.07</td>
</tr>
<tr>
<td>f. Hours open (weekday)</td>
<td>14.4</td>
<td>14.5</td>
<td>−0.1</td>
</tr>
<tr>
<td>g. Recruiting bonus</td>
<td>23.6</td>
<td>29.1</td>
<td>−0.0</td>
</tr>
</tbody>
</table>

(0.51) (1.25) (0.7) (0.02) (0.04) (2.3) (5.3) (0.04) (1.2) (2.3) (5.3)
### Table 2 – Means at Wave 1

<table>
<thead>
<tr>
<th>Outcome</th>
<th>NJ</th>
<th>PA</th>
<th>t-stat on difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>%BK</td>
<td>41.1</td>
<td>44.3</td>
<td>-0.5</td>
</tr>
<tr>
<td>% Roys</td>
<td>24.8</td>
<td>21.5</td>
<td>0.6</td>
</tr>
<tr>
<td>FTE</td>
<td>20.4</td>
<td>23.3</td>
<td>-2.0</td>
</tr>
<tr>
<td>% full time</td>
<td>32.8</td>
<td>35.0</td>
<td>-0.7</td>
</tr>
<tr>
<td>Starting wage</td>
<td>4.61</td>
<td>4.63</td>
<td>-0.4</td>
</tr>
<tr>
<td>Hours open</td>
<td>14.4</td>
<td>14.5</td>
<td>-0.3</td>
</tr>
</tbody>
</table>

### Table 3 – row 4

#### Change in full time equivalent employment

<table>
<thead>
<tr>
<th></th>
<th>PA((\bar{x}_1))</th>
<th>NJ((\bar{x}_2))</th>
<th>Diff((\bar{x}_2 - \bar{x}_1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean and (standard error of mean)</td>
<td>-2.28</td>
<td>0.47</td>
<td>2.75</td>
</tr>
<tr>
<td></td>
<td>(1.25)</td>
<td>(0.48)</td>
<td>(1.34)</td>
</tr>
</tbody>
</table>
Why did employment increase

• Maybe PA is a poor control – notice that employment in NJ increased, but in PA it fell. Most of the effect is generated by an increase in the employment in PA
  – What would we like to know to help prove PA is a good control?
• Fast food is a monopsony?
  – Nah – fast food restaurants are all different

Alternative control groups

• Maybe PA is a bad control – are there other control groups available?
• High wage stores in NJ
  – Stores currently paying above the new MW
  – Will not be impacted by the new law – it is not binding
Table 3 – row 4

<table>
<thead>
<tr>
<th></th>
<th>High Wage stores in NJ (1)</th>
<th>Low Wage stores in NJ (2)</th>
<th>Diff $\Delta = (2)-(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-2.16</td>
<td>1.21</td>
<td>3.36</td>
</tr>
<tr>
<td>(standard error of mean)</td>
<td>(1.01)</td>
<td>(0.82)</td>
<td>(1.30)</td>
</tr>
</tbody>
</table>

- Reduced turnover?
  - High turnover of jobs in fast food – 400% in a year
  - Most due to quits
  - Higher wage reduces quits, decreases number of “open” jobs