ECON 626: Applied Microeconomics

Lecture 3:

Difference-in-Differences

Professors: Pamela Jakiela and Owen Ozier

Intuition and Assumptions

False Counterfactuals

Before vs. After Comparisons:

- **Compares:** same individuals/communities before and after program
- **Drawback:** does not control for time trends

Participant vs. Non-Participant Comparisons:

- **Compares:** participants to those not in the program
- **Drawback:** selection — why didn't non-participants participate?

Two Wrongs Sometimes Make a Right

Difference-in-differences (or "diff-in-diff" or "DD") estimation combines the (flawed) pre vs. post and participant vs. non-participant approaches

- This can sometimes overcome the twin problems of [1] selection bias (on fixed traits) and [2] time trends in the outcome of interest
- The basic idea is to observe the (self-selected) treatment group and a (self-selected) comparison group before and after the program

Two Wrongs Sometimes Make a Right

Difference-in-differences (or "diff-in-diff" or "DD") estimation combines the (flawed) pre vs. post and participant vs. non-participant approaches

- This can sometimes overcome the twin problems of [1] selection bias (on fixed traits) and [2] time trends in the outcome of interest
- The basic idea is to observe the (self-selected) treatment group and a (self-selected) comparison group before and after the program

The diff-in-diff estimator is:

$$DD = \bar{Y}^{treatment}_{post} - \bar{Y}^{treatment}_{pre} - \left(\bar{Y}^{comparison}_{post} - \bar{Y}^{comparison}_{pre}\right)$$

DD Estimation: Early Examples

1849: London's worst cholera epidemic claims 14,137 lives

- Two companies supplied water to much of London: the Lambeth Waterworks Co. and the Southwark and Vauxhall Water Co.
  - Both got their water from the Thames

DD Estimation: Early Examples

1849: London's worst cholera epidemic claims 14,137 lives

- Two companies supplied water to much of London: the Lambeth Waterworks Co. and the Southwark and Vauxhall Water Co.
  - Both got their water from the Thames
- John Snow believed cholera was spread by contaminated water

DD Estimation: Early Examples

1849: London's worst cholera epidemic claims 14,137 lives

- Two companies supplied water to much of London: the Lambeth Waterworks Co. and the Southwark and Vauxhall Water Co.
  - Both got their water from the Thames
- John Snow believed cholera was spread by contaminated water

1852: Lambeth Waterworks moved their intake upriver

- Everyone knew that the Thames was dirty below central London

DD Estimation: Early Examples

1849: London's worst cholera epidemic claims 14,137 lives

- Two companies supplied water to much of London: the Lambeth Waterworks Co. and the Southwark and Vauxhall Water Co.
    - Both got their water from the Thames
- John Snow believed cholera was spread by contaminated water

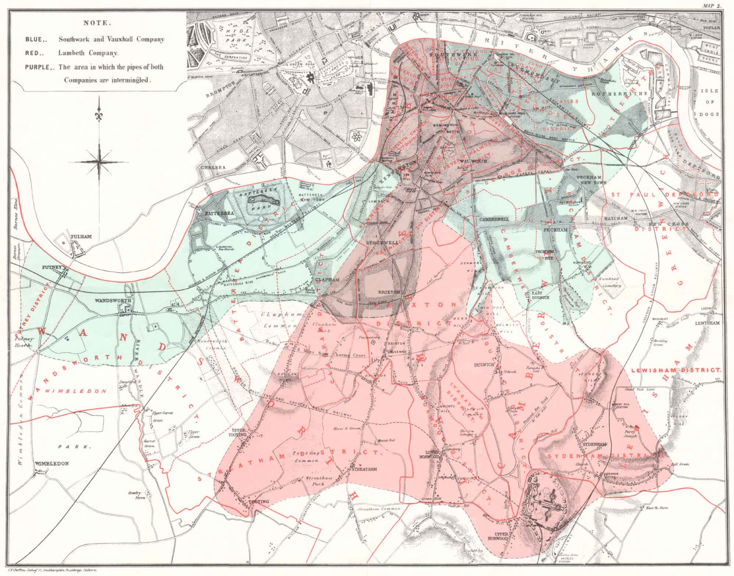
1852: Lambeth Waterworks moved their intake upriver

- Everyone knew that the Thames was dirty below central London

1853: London has another cholera outbreak

- Are Lambeth Waterworks customers less likely to get sick?

DD Estimation: Early Examples



Source: John Snow Archive and Research Companion

DD Estimation: Early Examples

John Snow's Grand Experiment:

- Mortality data showed that very few cholera deaths were reported in areas of London that were **only** supplied by the Lambeth Waterworks
- Snow hired John Whiting to visit the homes of the deceased to determine which company (if any) supplied their drinking water
- Using Whiting's data, Snow calculated the death rate
    - Southwark and Vauxhall: 71 cholera deaths/10,000 homes
    - Lambeth: 5 cholera deaths/10,000 homes

DD Estimation: Early Examples

John Snow's Grand Experiment:

- Mortality data showed that very few cholera deaths were reported in areas of London that were **only** supplied by the Lambeth Waterworks
- Snow hired John Whiting to visit the homes of the deceased to determine which company (if any) supplied their drinking water
- Using Whiting's data, Snow calculated the death rate
  - Southwark and Vauxhall: 71 cholera deaths/10,000 homes
  - Lambeth: 5 cholera deaths/10,000 homes
- **Southwark and Vauxhall responsible for 286 of 334 deaths**
  - Southwark and Vauxhall moved their intake upriver in 1855

DD Estimation: Early Examples

In the 1840s, observers of Vienna's maternity hospital noted that death rates from postpartum infections were higher in one wing than the other

- Division 1 patients were attended by doctors and trainee doctors
- Division 2 patients were attended by midwives and trainee midwives

DD Estimation: Early Examples

In the 1840s, observers of Vienna's maternity hospital noted that death rates from postpartum infections were higher in one wing than the other

- Division 1 patients were attended by doctors and trainee doctors
- Division 2 patients were attended by midwives and trainee midwives

Ignaz Semmelweis noted that the difference emerged in 1841, when the hospital moved to an "anatomical" training program involving cadavers

- Doctors received new training; midwives never handled cadavers
- Did the transference of "cadaveric particles" explain the death rate?

DD Estimation: Early Examples

In the 1840s, observers of Vienna's maternity hospital noted that death rates from postpartum infections were higher in one wing than the other

- Division 1 patients were attended by doctors and trainee doctors
- Division 2 patients were attended by midwives and trainee midwives

Ignaz Semmelweis noted that the difference emerged in 1841, when the hospital moved to an "anatomical" training program involving cadavers

- Doctors received new training; midwives never handled cadavers
- Did the transference of "cadaveric particles" explain the death rate?

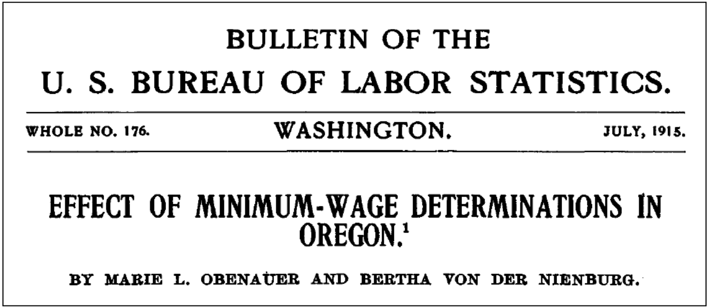
Semmelweis proposed an intervention: hand-washing with chlorine

- Policy implemented in May of 1847

DD Estimation: Early Examples

Year	Physicians' Division			Midwives' Division		
		Deaths			Deaths	
	Births	No.	%	Births	No.	%
1841	3036	237	7.7	2442	86	3.5
1842	3287	518	15.8	2659	202	7.5
1843	3060	274	8.9	2739	169	6.2
1844	3157	260	8.2	2956	68	2.3
1845	3492	241	6.8	3241	66	2.03
1846	4010	459	11.4	3754	105	2.7
1847				3306	32	0.9
January–May	2134	120	5.6			
	Intervention introduced in May					
June–December	1841	56	3.04			
1848	3556	45	1.27	3219	43	1.33

DD Estimation: Early Examples



BULLETIN OF THE
U. S. BUREAU OF LABOR STATISTICS.

WHOLE NO. 176. WASHINGTON. JULY, 1915.

EFFECT OF MINIMUM-WAGE DETERMINATIONS IN OREGON.[1]

BY MARIE L. OBENAUER AND BERTHA VON DER NIENBURG.

Source: Obenauer and Nienburg (1915)

DD Estimation: Early Examples

In 1913, Oregon increased the minimum wage for experienced women to \$9.25 per week, with a maximum of 50 hours of work per week

- Minimum wage for inexperienced women (and girls) also increased, but was new minimum (\$6/week) not seen as a binding constraint
- Obenauer and Nienburg obtain HR records of 40 firms
- Compare employment of experienced women before after minimum wage to law to employment of girls, inexperienced women, men

DD Estimation: Early Examples

TABLE 1.—ESTABLISHMENTS COVERED IN THE INVESTIGATION AND WOMEN AND MEN EMPLOYED DURING PERIOD STUDIED IN 1914.

[This table does not include extra male or female help whose identity from week to week could not be traced, such female help being equivalent to 3 women working full time; nor does it include 20 saleswomen whose regular employment began with the opening of a new department on the last day of the period covered in the investigation.]

Type of store.	Number of establishments covered.	Number of persons employed during period studied in 1914.	
		Women and girls.	Men.
PORTLAND.			
Department, dry-goods, and 5 and 10 cent stores	6	1,345	802
Specialty stores	11	181	49
Neighborhood stores	16	20	17
Total	33	1,546	868
SALEM.			
Dry-goods, specialty, and 5 and 10 cent stores	7	96	34
Grand total	40	1,642	902

[1] See note [1], p. 57.

[2] One firm, Olds, Wortman & King, a Portland department store, refused the Federal agents access to their records. They offered to furnish a summary statement, but the Bureau did not regard this as comparable with material obtained direct from other firms' books.

Source: Obenauer and Nienburg (1915)

DD Estimation: Early Examples

TABLE 1
EMPLOYMENT CHANGES IN PORTLAND RETAIL STORES, 1913–1914

	Men	Girls (Age 16–18)	Ratio (Girls/Men)	Women (Age > 18)	Ratio (Women/Men)	Women Age Unknown
Before (Mar/Ap 1913)	940	138	.1468	1543	1.641	152
After Mar/Ap 1914)	868	160	.1843	1327	1.529	59
Change	–72	22	.0375	–216	–0.113	–93
% Change	–7.7%	15.9%	23.6%	–14%	–6.3%	–61.2%

Source: Obenauer and Nienburg (1915, Table 3, pages 14–15).

Source: Kennan (1995)

Difference-in-Differences Estimation

	Treatment	**Comparison**
Pre-Program	$\bar{Y}_{pre}^{treatment}$	$\bar{Y}_{pre}^{comparison}$
Post-Program	$\bar{Y}_{post}^{treatment}$	$\bar{Y}_{post}^{comparison}$

Intuitively, diff-in-diff estimation is just a comparison of 4 cell-level means

- Only one cell is treated: **Treatment**$\times$**Post-Program**

Difference-in-Differences Estimation

The assumption underlying diff-in-diff estimation is that, in the absence of the program, individual i's outcome at time t is given by:

$$E[Y_i|D_i = 0, t = \tau] = \gamma_i + \lambda_\tau$$

Difference-in-Differences Estimation

The assumption underlying diff-in-diff estimation is that, in the absence of the program, individual i's outcome at time t is given by:

$$E[Y_i|D_i = 0, t = \tau] = \gamma_i + \lambda_\tau$$

There are two implicit identifying assumptions here:

- Selection bias relates to fixed characteristics of individuals (γ_i)
  - The magnitude of the selection bias term isn't changing over time
- Time trend (λ_t) same for treatment and control groups

Difference-in-Differences Estimation

The assumption underlying diff-in-diff estimation is that, in the absence of the program, individual i's outcome at time t is given by:

$$E[Y_i|D_i = 0, t = \tau] = \gamma_i + \lambda_\tau$$

There are two implicit identifying assumptions here:

- Selection bias relates to fixed characteristics of individuals (γ_i)
  - The magnitude of the selection bias term isn't changing over time
- Time trend (λ_t) same for treatment and control groups

Both necessary conditions for identification in diff-in-diff estimation

- Referred to as the **common trends** assumption

Difference-in-Differences Estimation

In the absence of the program, i's outcome at time τ is:

$$E[Y_{0i}|D_i = 0, t = \tau] = \gamma_i + \lambda_\tau$$

Difference-in-Differences Estimation

In the absence of the program, i's outcome at time τ is:

$$E[Y_{0i}|D_i = 0, t = \tau] = \gamma_i + \lambda_\tau$$

Outcomes in the comparison group:

$$E[\bar{Y}_{pre}^{comparison}] = E[Y_{0i}|D_i = 0, t = 1] = E[\gamma_i|D_i = 0] + \lambda_1$$

$$E[\bar{Y}_{post}^{comparison}] = E[Y_{0i}|D_i = 0, t = 2] = E[\gamma_i|D_i = 0] + \lambda_2$$

Difference-in-Differences Estimation

In the absence of the program, i's outcome at time τ is:

$$E[Y_{0i}|D_i = 0, t = \tau] = \gamma_i + \lambda_\tau$$

Outcomes in the comparison group:

$$E[\bar{Y}_{pre}^{comparison}] = E[Y_{0i}|D_i = 0, t = 1] = E[\gamma_i|D_i = 0] + \lambda_1$$

$$E[\bar{Y}_{post}^{comparison}] = E[Y_{0i}|D_i = 0, t = 2] = E[\gamma_i|D_i = 0] + \lambda_2$$

The comparison group allows us to estimate the **time trend**:

$$E[\bar{Y}_{post}^{comparison}] - E[\bar{Y}_{pre}^{comparison}] = E[\gamma_i|D_i = 0] + \lambda_2 - (E[\gamma_i|D_i = 0] + \lambda_1)$$

$$= \lambda_2 - \lambda_1$$

Difference-in-Differences Estimation

Let δ denote the true impact of the program:

$$\delta = E[Y_{1i}|D_i = 1, t = \tau] - E[Y_{0i}|D_i = 1, t = \tau]$$

which does not depend on the time period or i's characteristics

Difference-in-Differences Estimation

Let δ denote the true impact of the program:

$$\delta = E[Y_{1i}|D_i = 1, t = \tau] - E[Y_{0i}|D_i = 1, t = \tau]$$

which does not depend on the time period or i's characteristics

Outcomes in the treatment group:

$$E[\bar{Y}_{pre}^{treatment}] = E[Y_{0i}|D_i = 1, t = 1] = E[\gamma_i|D_i = 1] + \lambda_1$$

$$E[\bar{Y}_{post}^{treatment}] = E[Y_{1i}|D_i = 1, t = 2] = E[\gamma_i|D_i = 1] + \delta + \lambda_2$$

Difference-in-Differences Estimation

Let δ denote the true impact of the program:

$$\delta = E[Y_{1i}|D_i = 1, t = \tau] - E[Y_{0i}|D_i = 1, t = \tau]$$

which does not depend on the time period or i's characteristics

Outcomes in the treatment group:

$$E[\bar{Y}_{pre}^{treatment}] = E[Y_{0i}|D_i = 1, t = 1] = E[\gamma_i|D_i = 1] + \lambda_1$$

$$E[\bar{Y}_{post}^{treatment}] = E[Y_{1i}|D_i = 1, t = 2] = E[\gamma_i|D_i = 1] + \delta + \lambda_2$$

Differences in outcomes pre-treatment vs. post treatment cannot be attributed to the program; treatment effect is conflated with time trend

Difference-in-Differences Estimation

If we were to calculate a pre-vs-post estimator, we'd have:

$$E[\bar{Y}_{post}^{treatment}] - E[\bar{Y}_{pre}^{treatment}] = E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 1] + \lambda_1)$$

$$= \delta + \underbrace{\lambda_2 - \lambda_1}_{\text{time trend}}$$

Difference-in-Differences Estimation

If we were to calculate a pre-vs-post estimator, we'd have:

$$E[\bar{Y}^{treatment}_{post}] - E[\bar{Y}^{treatment}_{pre}] = E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 1] + \lambda_1)$$

$$= \delta + \underbrace{\lambda_2 - \lambda_1}_{\text{time trend}}$$

If we calculated a treatment vs. comparison estimator, we'd have:

$$E[\bar{Y}^{treatment}_{post}] - E[\bar{Y}^{comparison}_{post}] = E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 0] + \lambda_2)$$

$$= \delta + \underbrace{E[\gamma_i|D_i = 1] - E[\gamma_i|D_i = 0]}_{\text{selection bias}}$$

Difference-in-Differences Estimation

Substituting in the terms from our model:

$$
\begin{aligned}
DD &= \bar{Y}_{post}^{treatment} - \bar{Y}_{pre}^{treatment} - \left(\bar{Y}_{post}^{comparison} - \bar{Y}_{pre}^{comparison}\right) \\
&= E[Y_{1i}|D_i = 1, t = 2] - E[Y_{0i}|D_i = 1, t = 1] \\
&\qquad - \Big(E[Y_{0i}|D_i = 0, t = 2] - E[Y_{0i}|D_i = 0, t = 1]\Big) \\
&= E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 1] + \lambda_1) \\
&\qquad - \Big[E[\gamma_i|D_i = 0] + \lambda_2 - \Big(E[\gamma_i|D_i = 0] + \lambda_1\Big)\Big] \\
&= \delta
\end{aligned}
$$

Difference-in-Differences Estimation

Substituting in the terms from our model:

$$
\begin{aligned}
DD &= \bar{Y}^{treatment}_{post} - \bar{Y}^{treatment}_{pre} - \left(\bar{Y}^{comparison}_{post} - \bar{Y}^{comparison}_{pre}\right) \\
&= E[Y_{1i}|D_i = 1, t = 2] - E[Y_{0i}|D_i = 1, t = 1] \\
&\quad - \Big(E[Y_{0i}|D_i = 0, t = 2] - E[Y_{0i}|D_i = 0, t = 1]\Big) \\
&= E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 1] + \lambda_1) \\
&\quad - \Big[E[\gamma_i|D_i = 0] + \lambda_2 - \Big(E[\gamma_i|D_i = 0] + \lambda_1\Big)\Big] \\
&= \delta
\end{aligned}
$$

DD estimation recovers the true impact of the program on participants (as long as the common trends assumption isn't violated)

Difference-in-Differences Estimation

DD does not rely on assumption of homogeneous treatment effects

- When treatment effects are homogeneous, DD estimation yields **average treatment effect on the treated** (ATT)
- Averages across treated units and over time
  - When impacts change over time (within treated units), DD estimate of treatment effect may depend on choice of evaluation window

Example: A Natural Experiment in Education

In a famous paper in the *American Economic Review*, Esther Duflo examines the impacts of a large school construction program in Indonesia

Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment

By ESTHER DUFLO*

Between 1973 and 1978, the Indonesian government engaged in one of the largest school construction programs on record. Combining differences across regions in the number of schools constructed with differences across cohorts induced by the timing of the program suggests that each primary school constructed per 1,000 children led to an average increase of 0.12 to 0.19 years of education, as well as a 1.5 to 2.7 percent increase in wages. This implies estimates of economic returns to education ranging from 6.8 to 10.6 percent. (JEL I2, J31, O15, O22)

Example: A Natural Experiment in Education

The Sekolar Dasar INPRES program (1973–1979):

- Oil crisis creates large windfall for Indonesia
- Suharto uses oil money to fund school construction
- Close to 62,000 schools built by national gov't
  - Approximately 1 school built per 500 school-age children
- More schools built in areas which started with fewer schools
- Schools intended to promote equality, national identity

The Return to Education in Indonesia

Do children who were born into areas with more newly built INPRES primary schools get more education? Do they earn more as adults?

The Return to Education in Indonesia

Do children who were born into areas with more newly built INPRES primary schools get more education? Do they earn more as adults?

Strategy: **difference-in-differences estimation**

- Data on children born before and after program (pre vs. post)
  - Children aged 12 and up in 1974 did not benefit from program
  - Children aged 6 and under were young enough to be treated
- Data on children born in communities where many schools were built (treatment), those where few schools were built (comparison)
  - Partition sample based on residuals from a regression of the number of schools built (per district) on the number of school-aged children
- Difference-in-differences estimate of program impact compares pre vs. post differences in treatment vs. comparison communities

The Return to Education in Indonesia

The simplest difference-in-differences estimator is:

$$DD = \bar{Y}_{post}^{treatment} - \bar{Y}_{pre}^{treatment} - \left(\bar{Y}_{post}^{comparison} - \bar{Y}_{pre}^{comparison}\right)$$

The Return to Education in Indonesia

The simplest difference-in-differences estimator is:

$$DD = \bar{Y}^{treatment}_{post} - \bar{Y}^{treatment}_{pre} - \left(\bar{Y}^{comparison}_{post} - \bar{Y}^{comparison}_{pre}\right)$$

Dependent Variable: Years of Schooling

	Many Schools Built	**Few Schools Built**	**Difference**
Over 11 in 1974	8.02	9.40	-1.38
Under 7 in 1974	8.49	9.76	-1.27
Difference	0.47	0.36	0.12

Difference-in-differences estimation compares the change in years of schooling (i.e. the pre vs. post estimate) in treatment, control areas

- Program areas increased faster than comparison areas
- Difference is not statistically significant

The Return to Education in Indonesia

The simplest difference-in-differences estimator is:

$$DD = \bar{Y}_{post}^{treatment} - \bar{Y}_{pre}^{treatment} - \left(\bar{Y}_{post}^{comparison} - \bar{Y}_{pre}^{comparison}\right)$$

The Return to Education in Indonesia

The simplest difference-in-differences estimator is:

$$DD = \bar{Y}_{post}^{treatment} - \bar{Y}_{pre}^{treatment} - \left(\bar{Y}_{post}^{comparison} - \bar{Y}_{pre}^{comparison}\right)$$

Dependent Variable: Log (Wages)

	Many Schools Built	**Few Schools Built**	**Difference**
Over 11 in 1974	6.87	7.02	-0.15
Under 7 in 1974	6.61	6.73	-0.12
Difference	-0.26	-0.29	0.026

Difference-in-differences estimation compares the change in the log of adult wages (i.e. the pre vs. post estimate) in treatment, control areas

- Program had a modest impact on adult wages
- Difference is not statistically significant

DD in a Regression Framework

DD in a Regression Framework

To implement diff-in-diff in a regression framework, we estimate:

$$Y_{i,t} = \alpha + \beta D_i + \zeta Post_t + \delta \left(D_i * Post_t \right) + \varepsilon_{i,t}$$

where:

- $Post_i$ is an indicator equal to 1 if $t = 2$
- δ is the coefficient of interest (the treatment effect)
- $\alpha = E[\gamma_i | D_i = 0] + \lambda_1$ — pre-program mean in comparison group
- $\beta = E[\gamma_i | D_i = 1] - E[\gamma_i | D_i = 0]$ — selection bias
- $\zeta = \lambda_2 - \lambda_1$ — time trend

DD in a Regression Framework

Pooled OLS specification is equivalent to first differences:

$$Y_{i,2} - Y_{i,1} = \eta + \gamma D_i + \epsilon_{it}$$

where:

- $Y_{i,2} - Y_{i,1}$ is the change (pre vs. post) in the outcome of interest
- γ is the coefficient of interest (the treatment effect)
- η is the time trend

DD in a Regression Framework

We can also implement diff-in-diff in a panel data framework when more than two periods of data are available; this can increase statistical power*

$$Y_{i,t} = \alpha + \eta_i + \nu_t + \gamma D_{i,t} + \varepsilon_{i,t}$$

DD in a Regression Framework

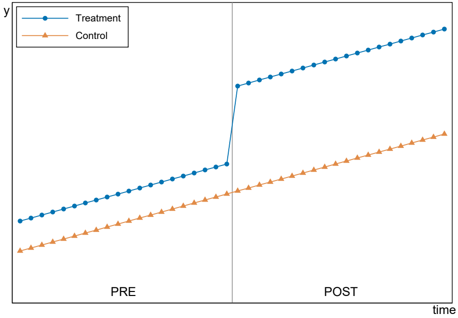
We can also implement diff-in-diff in a panel data framework when more than two periods of data are available; this can increase statistical power*

$$Y_{i,t} = \alpha + \eta_i + \nu_t + \gamma D_{i,t} + \varepsilon_{i,t}$$

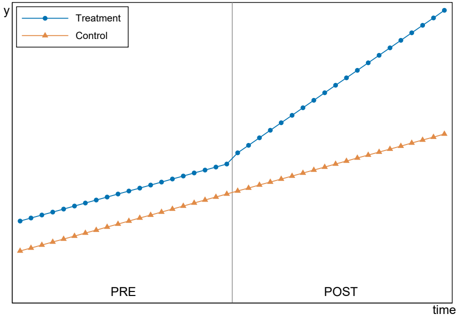
with some caveats:

- Variation in treatment timing?
- Allows for a credible defense of the common trends assumption
  - Unless the common trends assumption is violated
- Serial correlation in treatment and outcome variable is a problem

DD in a Regression Framework



DD in a Regression Framework



DD in a Regression Framework

Event study framework includes dummies for each post-treatment period:

$$Y_{i,t} = \alpha + \eta_i + \nu_t + \gamma_1 D1_{i,t} + \gamma_2 D2_{i,t} + \gamma_3 D3_{i,t} + \ldots + \varepsilon_{i,t}$$

DD in a Regression Framework

Event study framework includes dummies for each post-treatment period:

$$Y_{i,t} = \alpha + \eta_i + \nu_t + \gamma_1 D1_{i,t} + \gamma_2 D2_{i,t} + \gamma_3 D3_{i,t} + \ldots + \varepsilon_{i,t}$$

When treatment intensity is a continuous variable:

$$Y_{i,t} = \alpha + \beta Intensity_i + \zeta Post_t + \delta \left(Intensity_i * Post_t \right) + \varepsilon_{i,t}$$

Example: A Natural Experiment in Education

Main empirical specification in Duflo (2001):

$$S_{ijk} = \alpha + \eta_j + \beta_k + \gamma \left(Intensity_j * Young_i \right) + \mathbf{C}_j \delta + \varepsilon_{ijk}$$

where:

- S_{ijk} = education of individual i born in region j in year k
- η_j = region of birth fixed effect
- β_k = year of birth fixed effect
- $Young_i$ = dummy for being 6 or younger in 1974 (treatment group)
- $Intensity_j$ = INPRES schools per thousand school-aged children
- $\mathbf{C}_j$ = a vector of region-specific controls (that change over time)

Example: A Natural Experiment in Education

Dependent Variable: Years of Education

	Obs.	OLS (1)	OLS (2)	OLS (3)
Panel A: Entire Sample				
$Intensity_j * Young_i$	78,470	0.124	0.150	0.188
		(0.025)	(0.026)	(0.029)
Panel B: Sample of Wage Earners				
$Intensity_j * Young_i$	31,061	0.196	0.199	0.259
		(0.042)	(0.043)	(0.050)
Controls Included:				
YOB∗enrollment rate in 1971		No	Yes	Yes
YOB∗other INPRES programs		No	No	Yes

Sample includes individuals aged 2 to 6 or 12 to 17 in 1974. All Specifications include region of birth dummies, year of birth dummies, and interactions between the year of birth dummis and the number of children in the region of birth (in 1971). Standard errors are in parentheses.

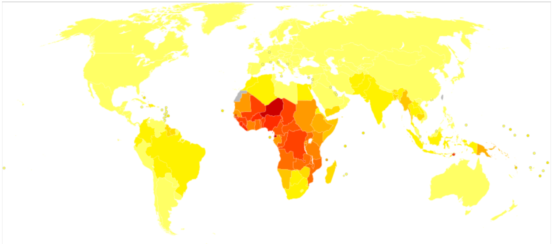
Example: A Natural Experiment in Education

Dependent Variable: Log Hourly Wages (as Adults)

	Obs.	OLS (1)	OLS (2)	OLS (3)
Panel A: Sample of Wage Earners				
$Intensity_j * Young_i$	31,061	0.0147	0.0172	0.027
		(0.007)	(0.007)	(0.008)
Controls Included:				
YOB∗enrollment rate in 1971		No	Yes	Yes
YOB∗other INPRES programs		No	No	Yes

Sample includes individuals aged 2 to 6 or 12 to 17 in 1974. All Specifications include region of birth dummies, year of birth dummies, and interactions between the year of birth dummis and the number of children in the region of birth (in 1971). Standard errors are in parentheses.

Malaria Eradication as a Natural Experiment



Malaria kills about 800,000 people per year

- Most are African children
- Repeated bouts of malaria may also reduce overall child health
- Countries with malaria are substantially poorer than other countries, but it is not clear whether malaria is the cause or the effect

Malaria Eradication as a Natural Experiment

Organized efforts to eradicate malaria are a natural experiment

- First the US (1920s) and then many Latin American countries (1950s) launched major (and successful) eradication campaigns
- Compare trends in adult income by birth cohort in regions which did, did not see major reductions in malaria because of campaigns

Malaria Eradication as a Natural Experiment

Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure†

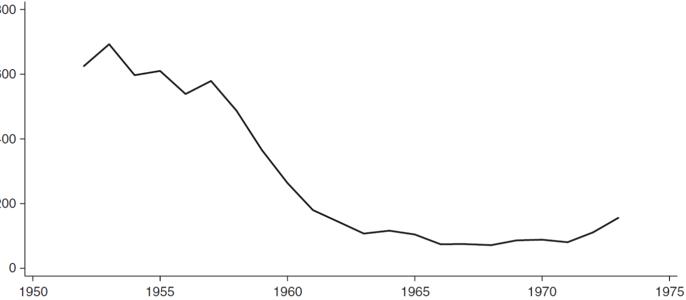
By HOYT BLEAKLEY*

This study uses the malaria-eradication campaigns in the United States (circa 1920) and in Brazil, Colombia, and Mexico (circa 1955) to measure how much childhood exposure to malaria depresses labor productivity. The campaigns began because of advances in health technology, which mitigates concerns about reverse causality. Malarious areas saw large drops in the disease thereafter. Relative to non-malarious areas, cohorts born after eradication had higher income as adults than the preceding generation. These cross-cohort changes coincided with childhood exposure to the campaigns rather than to pre-existing trends. Estimates suggest a substantial, though not predominant, role for malaria in explaining cross-region differences in income. (*JEL* I12, I18, J13, O15)

Malaria Eradication as a Natural Experiment

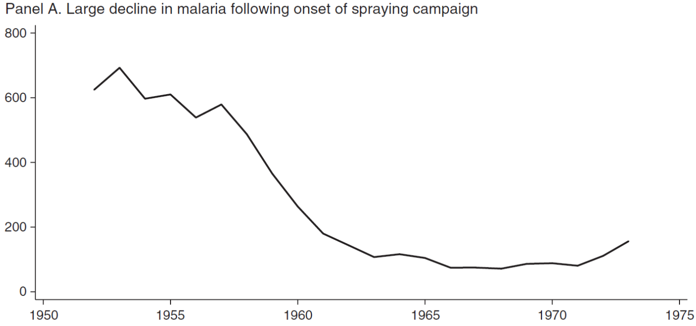
Colombia's malaria eradication campaign began in in the late 1950s...

Panel A. Large decline in malaria following onset of spraying campaign



Malaria Eradication as a Natural Experiment

Colombia's malaria eradication campaign began in in the late 1950s...



... and led to a huge decline in malaria morbidity

Malaria Eradication as a Natural Experiment

Areas with highest pre-program prevalence saw largest declines in malaria

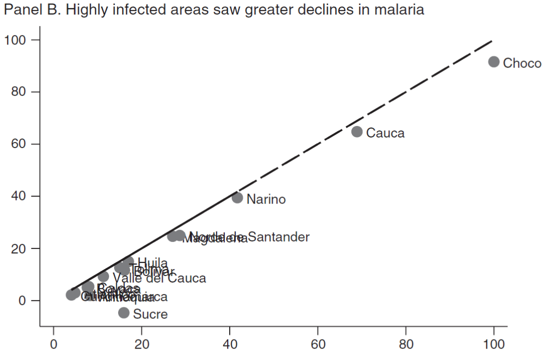


Figure 1. Malaria Incidence Before and After the Eradication Campaign, Colombia

Estimation Strategy

In this framework, treatment is a continuous variable

- Areas with higher pre-intervention malaria prevalence were, in essence "treated" more intensely by the eradication program
- Malaria-free areas should not benefit from eradication
- They can be used (implicitly) to measure the time trend

Estimation Strategy

In this framework, treatment is a continuous variable

- Areas with higher pre-intervention malaria prevalence were, in essence "treated" more intensely by the eradication program
- Malaria-free areas should not benefit from eradication
- They can be used (implicitly) to measure the time trend

Exposure (during childhood) also depends on one's year of birth

- Colombians born after 1957 were fully exposed to program
  - Did not suffer from chronic malaria in their early childhood
  - Did not miss school because of malaria
- Colombians born before 1940 were adults by the time the eradication campaign began, serve as the comparison group

Estimation Strategy

Regression specification:

$$Y_{j,post} - Y_{j,pre} = \alpha + \beta M_{j,pre} + \delta X_{j,pre} + \varepsilon_j$$

where

- $Y_{j,t}$ is an outcome of interest (eg literacy)
- $M_{j,pre}$ is pre-eradication malaria prevalence
- $X_{j,pre}$ is a vector of region-level controls
- ε_j is the noise term

The Impact of Childhood Exposure to Malaria

Regression specification:

$$Y_{j,post} - Y_{j,pre} = \alpha + \beta M_{j,pre} + \delta X_{j,pre} + \varepsilon_j$$

	Malaria ecology (Poveda)			Malaria ecology (Mellinger)		
Dependent variables: Differences across cohorts in…	Literacy	Years of schooling	Income index	Literacy	Years of schooling	Income index
Panel A. Alternative controls						
Additional controls:						
None (basic specification)	0.035***	0.168*	0.065***	0.071***	0.064	0.048***
	(0.013)	(0.088)	(0.011)	(0.016)	(0.108)	(0.014)
Conflict	0.032***	0.175*	0.063***	0.068***	0.068	0.046***
	(0.012)	(0.090)	(0.011)	(0.016)	(0.110)	(0.014)
Economic activity	0.008	0.194**	0.057***	0.043***	0.156	0.039***
	(0.010)	(0.089)	(0.012)	(0.013)	(0.110)	(0.014)
Other diseases	0.024*	0.180**	0.065***	0.058***	0.057	0.042***
	(0.013)	(0.089)	(0.012)	(0.016)	(0.114)	(0.015)
Full controls	0.006	0.165*	0.064***	0.046***	0.076	0.034**
	(0.011)	(0.095)	(0.013)	(0.015)	(0.117)	(0.015)

Defending the Common Trends Assumption

The Common Trends Assumption

Diff-in-diff does not identify the treatment effect if treatment and comparison groups were on different trajectories prior to the program

- This is the **common trends assumption**

The Common Trends Assumption

Diff-in-diff does not identify the treatment effect if treatment and comparison groups were on different trajectories prior to the program

- This is the **common trends assumption**

Remember the assumptions underlying diff-in-diff estimation:

- Selection bias relates to fixed characteristics of individuals (γ_i)
- Time trend (λ_t) same for treatment and control groups

The Common Trends Assumption

Diff-in-diff does not identify the treatment effect if treatment and comparison groups were on different trajectories prior to the program

- This is the **common trends assumption**

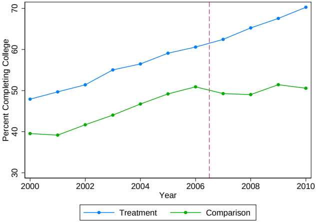
Remember the assumptions underlying diff-in-diff estimation:

- Selection bias relates to fixed characteristics of individuals (γ_i)
- Time trend (λ_t) same for treatment and control groups

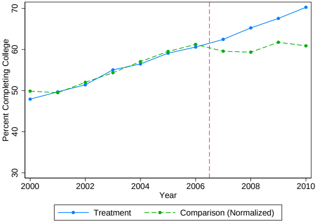
These assumptions guarantee that the common trends assumption is satisfied, but they cannot be tested directly — we have to trust!

- As with any identification strategy, it is important to think carefully about whether it checks out both intuitively and econometrically

The Common Trends Assumption

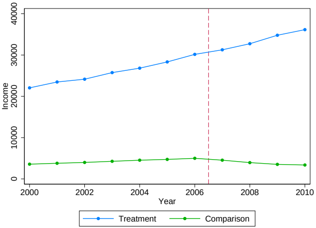


$\Rightarrow$

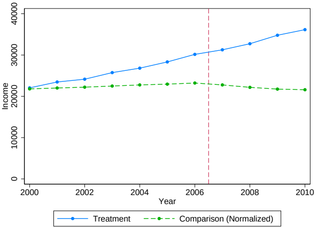


Sometimes, the common trends assumption is clearly OK

The Common Trends Assumption



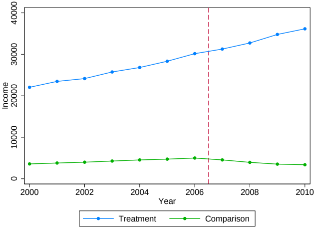
$\Rightarrow$



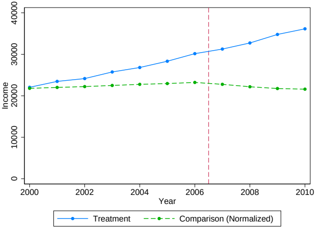
Other times, the common trends assumption is fairly clearly violated

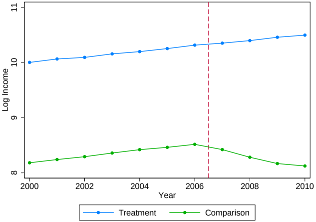
The Common Trends Assumption

Or is it? DD is robust to transformations of the outcome variable

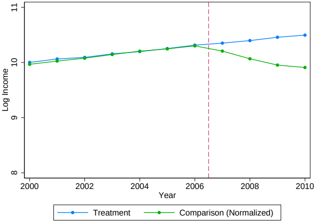


$\Rightarrow$





$\Rightarrow$



Defending the Common Trends Assumption

Three approaches:

1. A compelling graph
2. A falsification test or, analogously, a direct test in panel data
3. Controlling for time trends directly
    - Drawback: identification comes from functional form assumption

Defending the Common Trends Assumption

Three approaches:

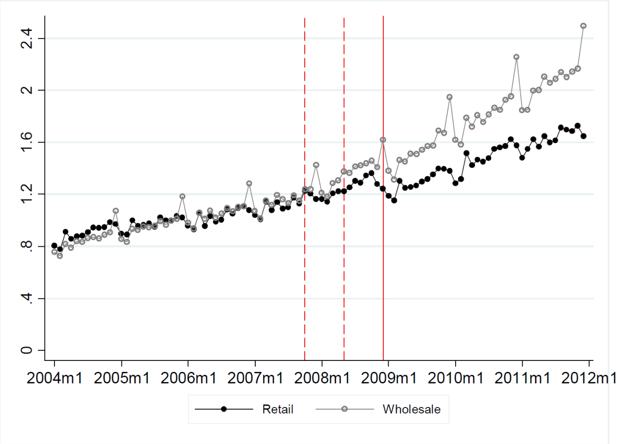
1. A compelling graph
2. A falsification test or, analogously, a direct test in panel data
3. Controlling for time trends directly
   - Drawback: identification comes from functional form assumption

None of these approaches are possible with two periods of data

Approach #1: DD Porn

Figure 4: Compliance Effect – Retail vs. Wholesale

a. Raw data: reported revenue changes



Source: Naritomi (2015)

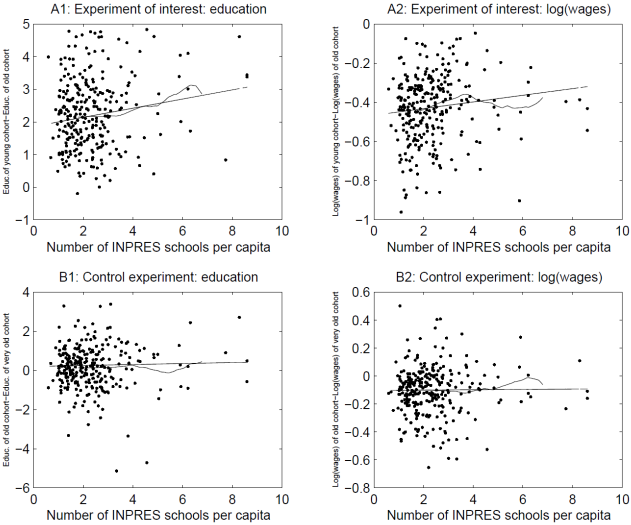
Approach #2: A Falsification Test

Dependent Variable: Years of Education

	Obs.	OLS (1)	OLS (2)	OLS (3)
Panel A: Entire Sample				
$Intensity_j * Younger_i$	78,488	0.009	0.018	0.008
		(0.026)	(0.027)	(0.030)
Panel B: Sample of Wage Earners				
$Intensity_j * Younger_i$	30,255	0.012	0.024	0.079
		(0.048)	(0.048)	(0.056)
Controls Included:				
YOB∗enrollment rate in 1971		No	Yes	Yes
YOB∗other INPRES programs		No	No	Yes

Sample includes individuals aged 12 to 24 in 1974. All Specifications include region of birth dummies, year of birth dummies, and interactions between the year of birth dummis and the number of children in the region of birth (in 1971). Standard errors are in parentheses.

Approach #2: A Falsification Test



Diff-in-Diff in a Panel Data Framework

Variation in Treatment Timing

Example: counties introduced food stamps at different times

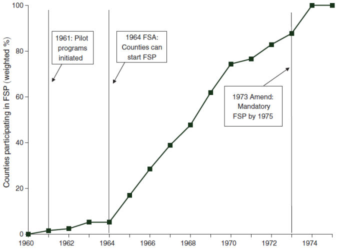


Figure 1. Weighted Percent of Counties with Food Stamp Program, 1960–1975

Source: Authors' tabulations of food stamp administrative data (US Department of Agriculture, various years). Counties are weighted by their 1960 population.

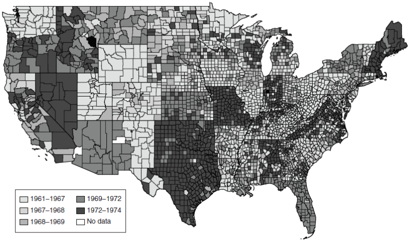


Figure 2. Food Stamp Program Start Date, by County, 1961–1974

Notes: Authors' tabulations of food stamp administrative data (US Department of Agriculture, various years). The shading corresponds to the county FSP start date, where darker shading indicates later county implementation.

Source: Almond, Hoynes, and Schanzenbach (AER, 2016)

Variation in Treatment Timing

Example: states adopted Medicaid at different times

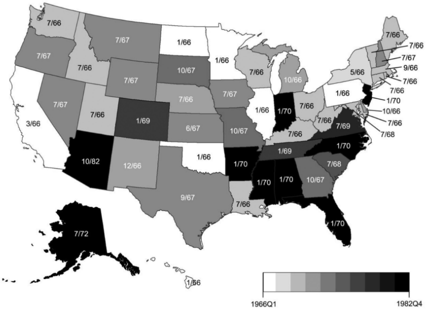


Figure 2.
Medicaid Adoption by Quarter
Notes: Adoption dates come from the Department of Health Education and Welfare (1970) & Social Security Administration (2013). The map is shaded relative to the quarter of adoption and states are labeled with the month and year of adoption.

Source: Boudreaux, Golberstein, and McAlpine (Journal of Health Economics, 2016)

Variation in Treatment Timing

Example: counties opened community health centers at different times

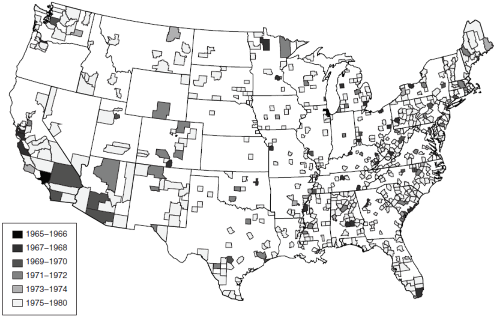


FIGURE 3. ESTABLISHMENT OF COMMUNITY HEALTH CENTERS BY COUNTY OF SERVICE DELIVERY, 1965–1980

Note: Dates are the first year that a CHC was established in the county.
Source: Information on CHCs drawn from NACAP and PHS reports.

Source: Bailey and Goodman-Bacon (AER, 2015)

Fixed Effects Estimates of β^{DD}

$$Y_{it} = \alpha_i + \gamma_t + \beta^{DD} D_{it} + \varepsilon_{ti}$$

Fixed Effects Estimates of β^{DD}

$$Y_{it} = \underbrace{\alpha_i}_{\text{unit fixed effects}} + \underbrace{\gamma_t}_{\text{time fixed effects}} + \beta^{DD} \underbrace{D_{it}}_{\text{treatment dummy}} + \varepsilon_{ti}$$

Fixed Effects Estimates of β^{DD}

What exactly is β^{DD}?

$$Y_{it} = \alpha_i + \gamma_t + \beta^{DD} D_{it} + \varepsilon_{ti}$$

unit fixed effects

time fixed effects

treatment dummy

Fixed Effects Estimates of β^{DD}

Frisch-Waugh (1933):

Two-way fixed effects regression is equivalent to univariate regression:

$$\tilde{Y}_{it} = \tilde{D}_{it} + \zeta_{ti}$$

where

$$\tilde{Y}_{it} = Y_{it} - \bar{Y}_i - \left(\bar{Y}_t - \bar{\bar{Y}}\right)$$

and

$$\tilde{D}_{it} = D_{it} - \bar{D}_i - \left(\bar{D}_t - \bar{\bar{D}}\right)$$

Fixed Effects Estimates of β^{DD}

Frisch-Waugh (1933):

Two-way fixed effects regression is equivalent to univariate regression:

$$\tilde{Y}_{it} = \tilde{D}_{it} + \zeta_{ti}$$

where

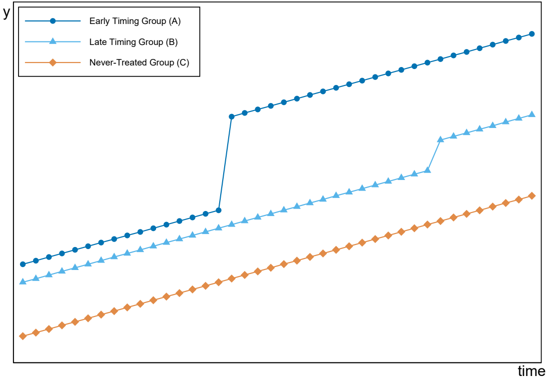
$$\tilde{Y}_{it} = Y_{it} - \bar{Y}_i - \left(\bar{Y}_t - \bar{\bar{Y}}\right)$$

and

$$\tilde{D}_{it} = D_{it} - \bar{D}_i - \left(\bar{D}_t - \bar{\bar{D}}\right)$$

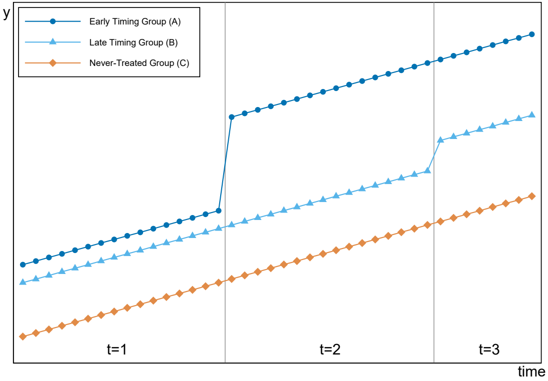
Which is cool, but doesn't really tell us what the estimand is

Decomposition into Timing Groups



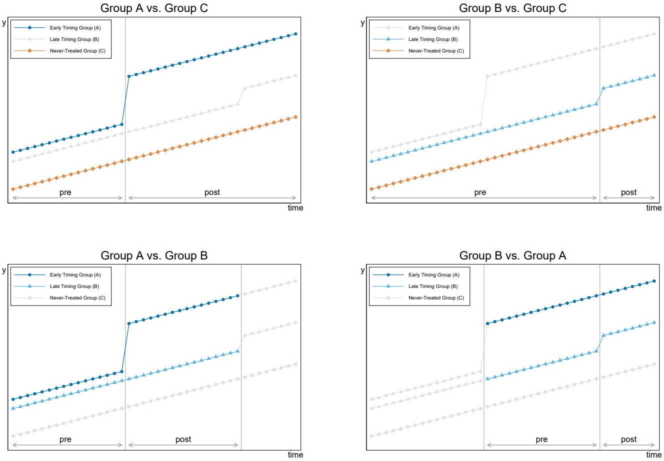
Goodman-Bacon (2019): panel with variation in treatment timing can be decomposed into **timing groups** reflecting observed onset of treatment

Decomposition into Timing Groups

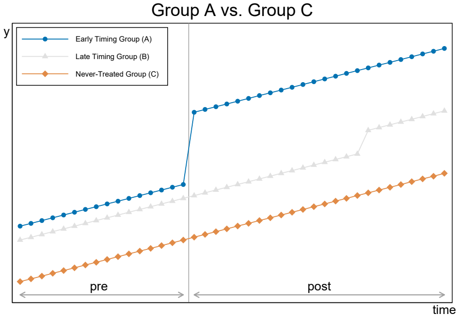


Example: with three timing groups (one of which is never treated), we can construct three timing windows (pre, middle, post or $t = 1, 2, 3$)

Decomposition into Standard 2×2 DDs



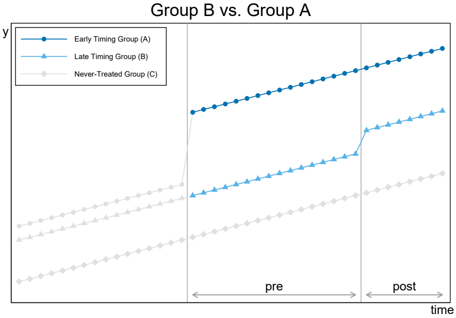
Decomposition into Standard 2×2 DDs



We know the DD estimate of the treatment effect for each timing group:

$$\hat{\beta}_{AC}^{DD} = \left(\bar{Y}_A^{POST} - \bar{Y}_C^{POST}\right) - \left(\bar{Y}_A^{PRE} - \bar{Y}_C^{PRE}\right)$$
$$= \left(\bar{Y}_A^{t=2,3} - \bar{Y}_C^{t=2,3}\right) - \left(\bar{Y}_A^{t=1} - \bar{Y}y_C^{t=1}\right)$$

Decomposition into Standard 2×2 DDs



We know the DD estimate of the treatment effect for each timing group:

$$\hat{\beta}^{DD}_{BA} = \left(\bar{Y}^{POST}_{B} - \bar{Y}^{POST}_{A}\right) - \left(\bar{Y}^{PRE}_{B} - \bar{Y}^{PRE}_{A}\right)$$
$$= \left(\bar{Y}^{t=3}_{B} - \bar{Y}^{t=3}_{A}\right) - \left(\bar{Y}^{t=2}_{B} - \bar{Y}{y}^{t=2}_{A}\right)$$

DD Decomposition Theorem (aka D^3 Theorem)

Theorem

Consider a data set comprising K timing groups ordered by the time at which they first receive treatment and a maximum of one never-treated group, U. The OLS estimate from a two-way fixed effects regression is:

$$\hat{\beta}^{DD} = \sum_{k \neq U} s_{kU} \hat{\beta}^{DD}_{kU} + \sum_{k \neq U} \sum_{j > k} \left[s_{kj} \hat{\beta}^{DD}_{kj} + s_{jk} \hat{\beta}^{DD}_{jk} \right]$$

In other words, the DD estimate from a two-way fixed effects regression is a weighted average of the (well-understood) 2×2 DD estimates

DD Decomposition Theorem (aka D^3 Theorem)

Weights depend on sample size, variance of treatment w/in each DD:

$$s_{kU} = \left[\frac{(n_k + n_U)^2}{\hat{V}^{\tilde{D}}}\right] \underbrace{n_{kU}\,(1 - n_{kU})\,\bar{D}_k(1 - \bar{D}_k)}_{\hat{Var}^{\tilde{D}}_{kU}}$$

$$s_{kj} = \left[\frac{\left((n_k + n_j)\,(1 - \bar{D}_j)\right)^2}{\hat{V}^{\tilde{D}}}\right] \underbrace{n_{kj}(1 - n_{kj})\left(\frac{\bar{D}_k - \bar{D}_j}{1 - \bar{D}_j}\right)\left(\frac{1 - \bar{D}_k}{1 - \bar{D}_j}\right)}_{\hat{Var}^{\tilde{D}}_{kj}}$$

$$s_{jk} = \left[\frac{\left((n_k + n_j)\,\bar{D}_k\right)^2}{\hat{V}^{\tilde{D}}}\right] \underbrace{n_{kj}(1 - n_{kj})\frac{\bar{D}_j}{\bar{D}_k}\left(\frac{\bar{D}_k - \bar{D}_j}{\bar{D}_k}\right)}_{\hat{Var}^{\tilde{D}}_{jk}}$$

where n_k is. . . , n_{kj} is . . . , and $\bar{D}_k$ is . . .

DD Decomposition Theorem (aka D^3 Theorem)

Weights depend on sample size, variance of treatment w/in each DD:

$$s_{kU} = \left[\frac{(n_k + n_U)^2}{\hat{V}^{\tilde{D}}}\right] \underbrace{n_{kU}\left(1 - n_{kU}\right)\bar{D}_k(1 - \bar{D}_k)}_{\hat{Var}^{\tilde{D}}_{kU}}$$

$$s_{kj} = \left[\frac{\left(\left(n_k + n_j\right)\left(1 - \bar{D}_j\right)\right)^2}{\hat{V}^{\tilde{D}}}\right] \underbrace{n_{kj}(1 - n_{kj})\left(\frac{\bar{D}_k - \bar{D}_j}{1 - \bar{D}_j}\right)\left(\frac{1 - \bar{D}_k}{1 - \bar{D}_j}\right)}_{\hat{Var}^{\tilde{D}}_{kj}}$$

$$s_{jk} = \left[\frac{\left(\left(n_k + n_j\right)\bar{D}_k\right)^2}{\hat{V}^{\tilde{D}}}\right] \underbrace{n_{kj}(1 - n_{kj})\frac{\bar{D}_j}{\bar{D}_k}\left(\frac{\bar{D}_k - \bar{D}_j}{\bar{D}_k}\right)}_{\hat{Var}^{\tilde{D}}_{jk}}$$

where n_k is. . . , n_{kj} is . . . , and $\bar{D}_k$ is . . .

Implications of the D^3 Theorem

1. When treatment effects are homogeneous, $\hat{\beta}^{DD}$ is the ATE
2. When treatment effects are heterogeneous across units (not time), $\hat{\beta}^{DD}$ is a variance-weighted treatment effect that is not the ATE
   - $\Rightarrow$ Weights on timing groups are sums of s_{kU}, s_{kj} terms
3. When treatment effects change over time, $\hat{\beta}^{DD}$ is biased
   - $\Rightarrow$ Changes in treatment effect bias DD coefficient
   - $\Rightarrow$ Event study, stacked DD more appropriate

Implications of the D^3 Theorem

DD in a potential outcomes framework assuming common trends:

$$Y_{it} = \begin{cases} Y_{0,it} \text{ if } D_{it} = 0 \\ Y_{0,it} + \delta_{it} \text{ if } D_{it} = 1 \end{cases}$$

Implications of the D^3 Theorem

DD in a potential outcomes framework assuming common trends:

$$Y_{it} = \begin{cases} Y_{0,it} \text{ if } D_{it} = 0 \\ Y_{0,it} + \delta_{it} \text{ if } D_{it} = 1 \end{cases}$$

$\hat{\beta}^{DD}_{kU}$ and $\hat{\beta}^{DD}_{kj}$ (where $k < j$) are familiar, but $\hat{\beta}^{DD}_{jk}$ is different:

$$\hat{\beta}^{DD}_{jk} = \bar{Y}^{POST}_{0,j} + \bar{\delta}^{POST}_{j} - \left(\bar{Y}^{POST}_{0,k} + \bar{\delta}^{POST}_{k}\right) - \left[\bar{Y}^{PRE}_{0,j} - \left(\bar{Y}^{PRE}_{0,k} + \bar{\delta}^{PRE}_{k}\right)\right]$$

$$= \bar{\delta}^{POST}_{j} + \underbrace{\left[\left(\bar{Y}^{POST}_{0,j} - \bar{Y}^{POST}_{0,k}\right) - \left(\bar{Y}^{PRE}_{0,j} - \bar{Y}^{PRE}_{0,k}\right)\right]}_{\text{common trends}} + \underbrace{\left(\bar{\delta}^{PRE}_{k} - \bar{\delta}^{POST}_{k}\right)}_{\Delta\delta_k}$$

Takeaways

1. Stack the 2×2 DDs to asses common trends (visually)

   $\Rightarrow$ Trends should look similar before and after treatment

   $\Rightarrow$ Treatment effect should be a level shift, no a trend break

   $\Rightarrow$ How much weight is placed on problematic timing groups?

2. Plot the relationship between the 2×2 DD estimates, weights

   $\Rightarrow$ No heterogeneity? No problems!

   $\Rightarrow$ Heterogeneity across units is an object of interest