

ECON 626: Applied Microeconomics

Lecture 3:

Difference-in-Difference Estimation

Professors: Pamela Jakiela and Owen Ozier

Department of Economics
University of Maryland, College Park

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-difference (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:

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

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
- John Snow believed cholera was spread by contaminated water

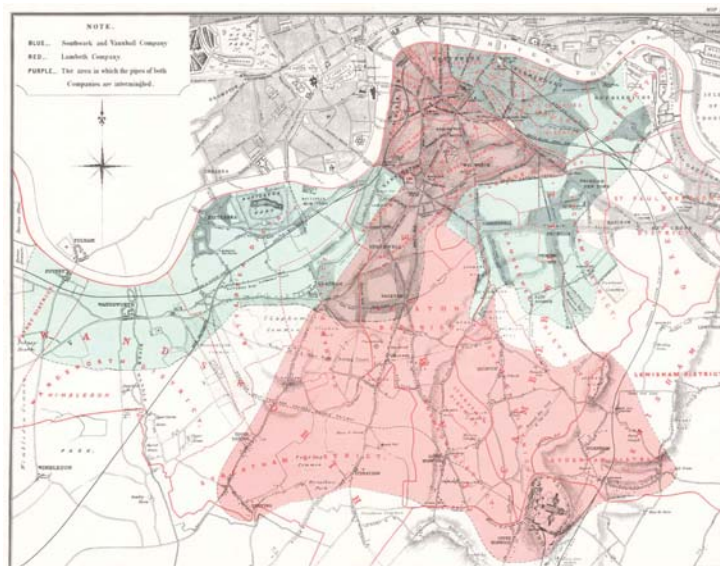
1852: Lambeth Waterworks moved their intake upriver

- 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



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

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		
	Births	Deaths		Births	Deaths	
		No.	%		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

Difference-in-Difference 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** × **Post-Program**

Difference-in-Difference 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

These two necessary conditions for identification in diff-in-diff estimation

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

Difference-in-Difference 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**:

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

Difference-in-Difference 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 are conflated with time trend

Difference-in-Difference Estimation

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

$$\begin{aligned} E[\bar{Y}_{post}^{treatment}] - E[\bar{Y}_{pre}^{treatment}] &= E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 1] + \lambda_1) \\ &= \delta + \underbrace{\lambda_2 - \lambda_1}_{\text{time trend}} \end{aligned}$$

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

$$\begin{aligned} E[\bar{Y}_{post}^{treatment}] - E[\bar{Y}_{post}^{comparison}] &= E[\gamma_i|D_i = 1] + \delta + \lambda_2 - (E[\gamma_i|D_i = 0] + \lambda_2) \\ &= \delta + \underbrace{E[\gamma_i|D_i = 1] - E[\gamma_i|D_i = 0]}_{\text{selection bias}} \end{aligned}$$

Difference-in-Difference Estimation

Substituting in the terms from our model:

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

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

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

Example: A Natural Experiment in Education

Research question: supply vs. demand constraints in education

- Supply side: governments need to provide more, better schools
 - ▶ Are there enough schools?
 - ▶ Are there enough well-trained teachers?
 - ▶ Are class sizes too large?
- Demand side: how large are the returns to education?
 - ▶ Are people dropping out at the optimal time given school quality, labor market conditions, and beliefs about the returns to education?

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-difference 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-difference estimate of program impact compares pre vs. post differences in treatment vs. comparison communities

The Return to Education in Indonesia

The simplest difference-in-difference estimator is:

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

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-difference 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-difference 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-difference 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

Diff-in-Diff in a Regression Framework

Diff-in-Diff in a Regression Framework

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

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

where:

- $Post_t$ 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

Diff-in-Diff 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

Diff-in-Diff in a Regression Framework

When treatment intensity is a continuous variable:

$$Y_{i,t} = \alpha + \beta \text{Intensity}_i + \zeta \text{Post}_t + \delta (\text{Intensity}_i * \text{Post}_t) + \varepsilon_{i,t}$$

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:

- 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

Example: A Natural Experiment in Education

Main empirical specification in Duflo (2001):

$$S_{ijk} = \alpha + \eta_j + \beta_k + \gamma (\text{Intensity}_j * \text{Young}_i) + \mathbf{C}_j \delta + \varepsilon_{ijk}$$

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)
<i>Panel A: Entire Sample</i>				
<i>Intensity_j * Young_i</i>	78,470	0.124 (0.025)	0.150 (0.026)	0.188 (0.029)
<i>Panel B: Sample of Wage Earners</i>				
<i>Intensity_j * Young_i</i>	31,061	0.196 (0.042)	0.199 (0.043)	0.259 (0.050)
<i>Controls Included:</i>				
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 dummies and the number of children in the region of birth (in 1971). Standard errors are in parentheses.

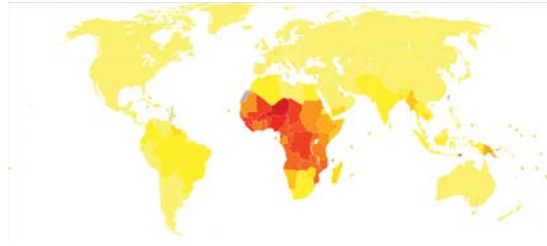
Example: A Natural Experiment in Education

Dependent Variable: Log Hourly Wages (as Adults)

	Obs.	OLS (1)	OLS (2)	OLS (3)
<i>Panel A: Sample of Wage Earners</i>				
<i>Intensity_j * Young_i</i>	31,061	0.0147 (0.007)	0.0172 (0.007)	0.027 (0.008)
<i>Controls Included:</i>				
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 dummies 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

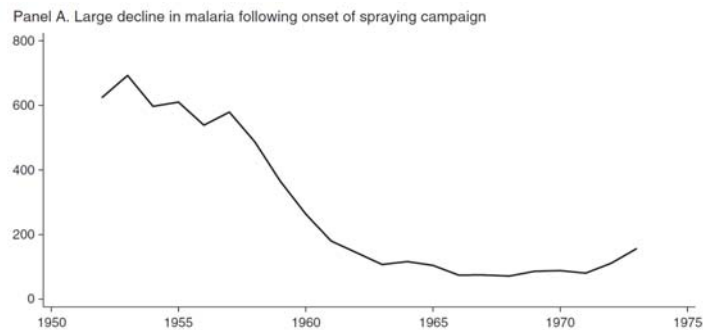
Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure[†]

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...



... 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

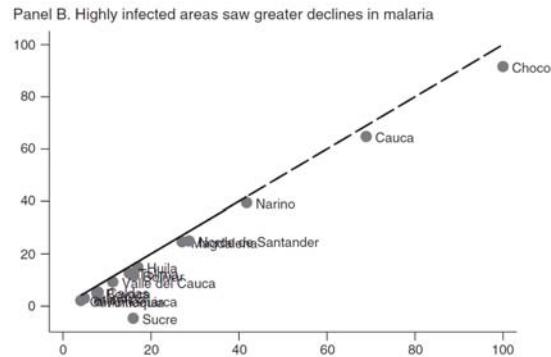


FIGURE 1. MALARIA INCIDENCE BEFORE AND AFTER THE ERADICATION CAMPAIGN, COLOMBIA

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

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_{ipt} is the noise term

The Impact of Childhood Exposure to Malaria

Regression specification:

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

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

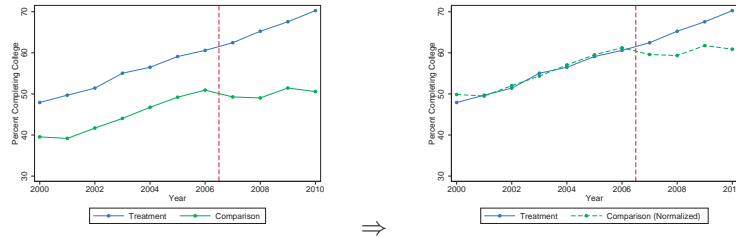
Remember the assumptions underlying diff-in-diff estimation:

- 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!

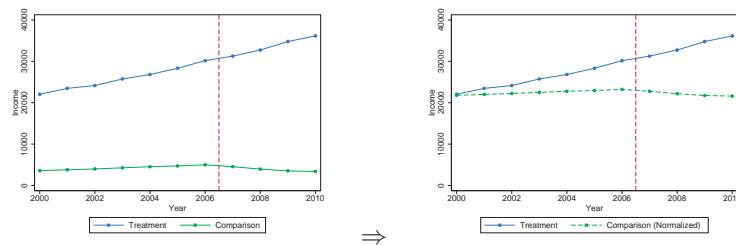
- As with any identification strategy, it is important to think carefully about whether it checks out both intuitively and econometrically

The Common Trends Assumption



Sometimes, the common trends assumption is clearly OK

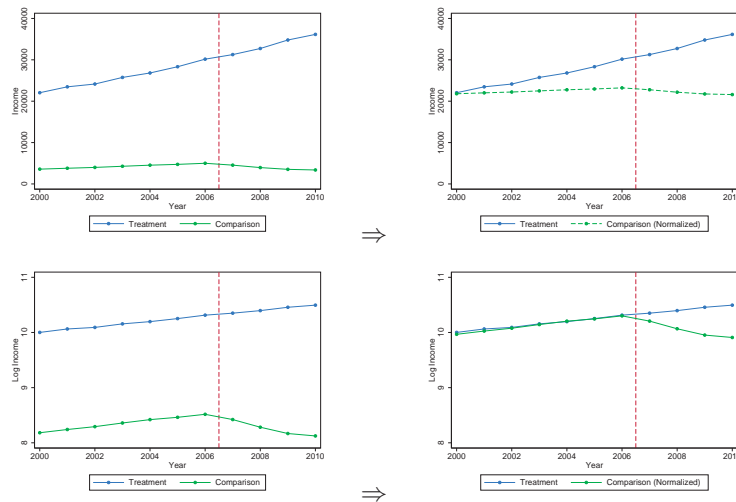
The Common Trends Assumption



Other times, the common trends assumption is fairly clearly violated

The Common Trends Assumption

Or is it? Diff-in-diff robust to transformations of the outcome variable



ECON 626: Applied Microeconomics

Lecture 3: Difference-in-Difference Estimation, Slide 41

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

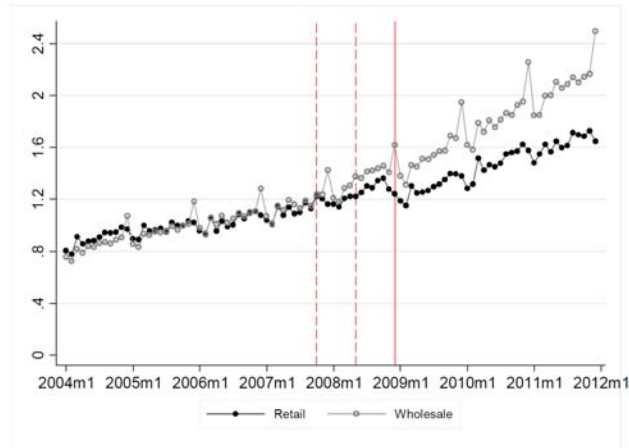
ECON 626: Applied Microeconomics

Lecture 3: Difference-in-Difference Estimation, Slide 42

Approach #1: DD Porn

Figure 4: Compliance Effect – Retail vs. Wholesale

a. Raw data: reported revenue changes



Source: Naritomi (2015)

ECON 626: Applied Microeconomics

Lecture 3: Difference-in-Difference Estimation, Slide 43

Approach #2: A Falsification Test

Dependent Variable: Years of Education

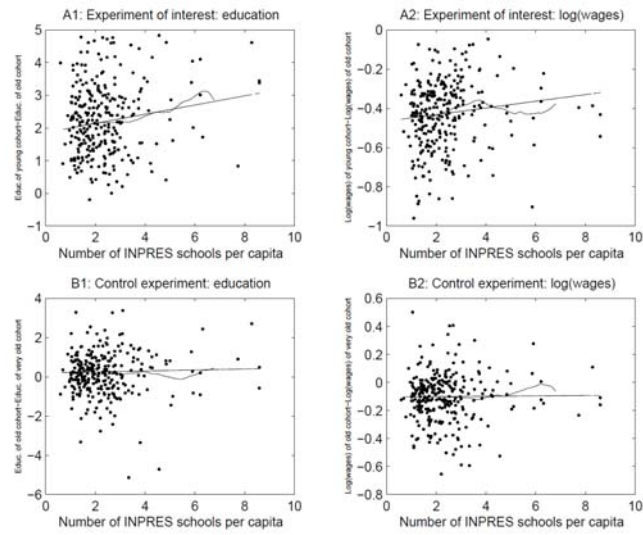
	Obs.	OLS (1)	OLS (2)	OLS (3)
<i>Panel A: Entire Sample</i>				
<i>Intensity_j * Younger_j</i>	78,488	0.009 (0.026)	0.018 (0.027)	0.008 (0.030)
<i>Panel B: Sample of Wage Earners</i>				
<i>Intensity_j * Younger_j</i>	30,255	0.012 (0.048)	0.024 (0.048)	0.079 (0.056)
<i>Controls Included:</i>				
YOB*enrollment 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 dummies and the number of children in the region of birth (in 1971). Standard errors are in parentheses.

ECON 626: Applied Microeconomics

Lecture 3: Difference-in-Difference Estimation, Slide 44

Approach #2: A Falsification Test

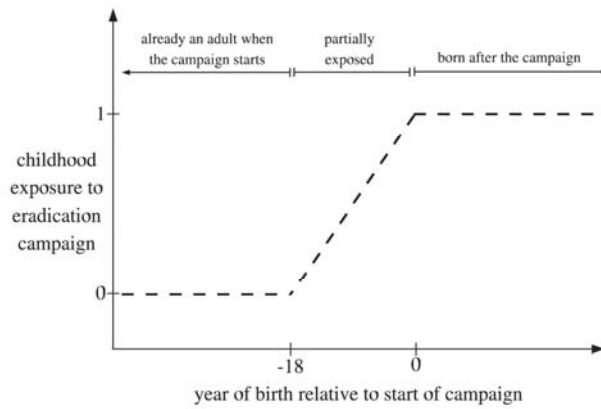


ECON 626: Applied Microeconomics

Lecture 3: Difference-in-Difference Estimation, Slide 45

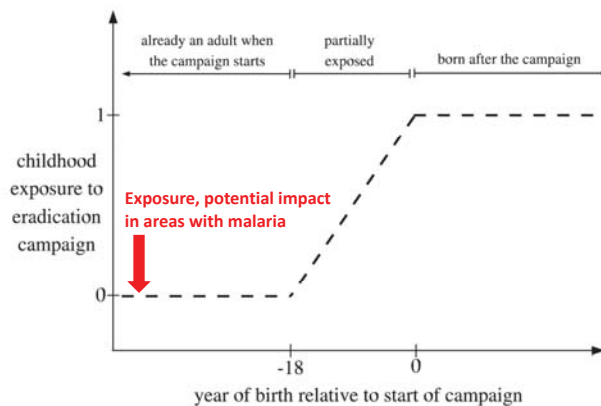
Diff-in-Diff in a Panel Data Framework

Exploiting (More) Variation by Birth Cohort



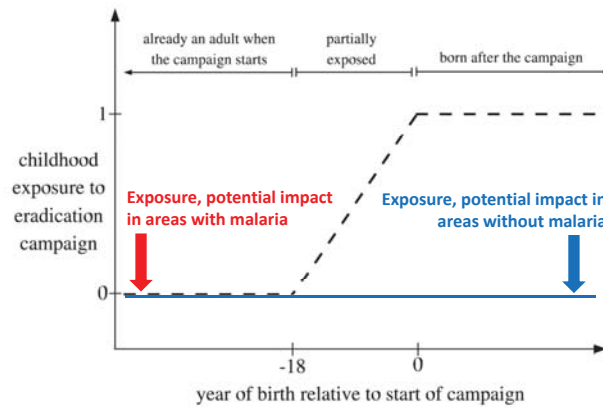
We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts

Exploiting (More) Variation by Birth Cohort



We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts

Exploiting (More) Variation by Birth Cohort



We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts

ECON 626: Applied Microeconomics Lecture 3: Difference-in-Difference Estimation, Slide 49

Panel Data Analysis

We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts

- Between 0 and 18 years of childhood post-eradication
- Interact exposure with pre-program malaria prevalence
- Treatment impacts should be larger for birth cohorts who spent more years of childhood malaria-free, areas with more initial malaria

Treat data set as a (YOB \times location) panel

- Control for region, YOB fixed effects

ECON 626: Applied Microeconomics Lecture 3: Difference-in-Difference Estimation, Slide 50

Panel Data Analysis

Regression specification:

$$Y_{jkt} = \alpha + \beta (M_j \times EXP_k) + \delta_k + \delta_j + \delta_t + \varepsilon_{jkt}$$

where

- M_j is pre-eradication malaria prevalence (by region)
- EXP_k is proportion of childhood post-eradication (by YOB)
- δ_k is a YOB fixed effect
- δ_j is a region of birth effect
- δ_t is a census year fixed effect
- ε_{jkt} is a conditionally mean-zero error term

Panel Data Analysis

Regression specification:

$$Y_{jkt} = \alpha + \beta (M_j \times EXP_k) + \delta_k + \delta_j + \delta_t + \varepsilon_{jkt}$$

Panel C. Colombia, log industrial income score

Baseline	0.052*** (0.010)	0.046*** (0.010)	0.046*** (0.010)	0.054*** (0.010)	0.048*** (0.010)	0.048*** (0.010)
Allow for birthplace × time effects	0.047*** (0.009)	0.046*** (0.010)	0.046*** (0.010)	0.050*** (0.010)	0.048*** (0.010)	0.049*** (0.010)
Add region × year × YOB effects	0.099*** (0.019)	0.076*** (0.019)	0.077*** (0.019)	0.081*** (0.020)	0.068*** (0.020)	0.068*** (0.020)

Panel Data Analysis

We can also look at the relationship between log (adult) wages and pre-eradication malaria rates separately by birth cohort

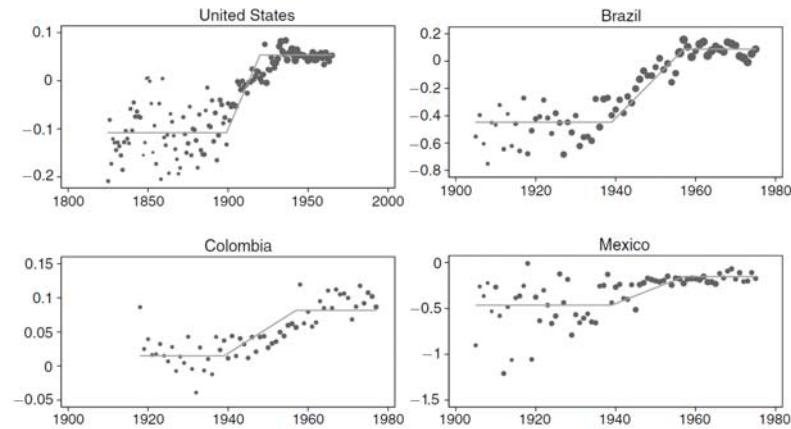


FIGURE 4. COHORT-SPECIFIC RELATIONSHIPS: INCOME AND PRE-CAMPAIGN MALARIA

Summary and Review

Diff-in-diff estimation can sometimes overcome the twin problems of [1] selection bias and [2] time trends in the outcomes of interest

- Is a credible impact evaluation strategy when the common trends assumption can be defended (i.e. you have a long panel)
 - ▶ Outcomes can be defined in levels or transformed levels
- Treatment need not be binary: often involves an interaction between pre-intervention conditions (e.g. malaria) and intervention timing