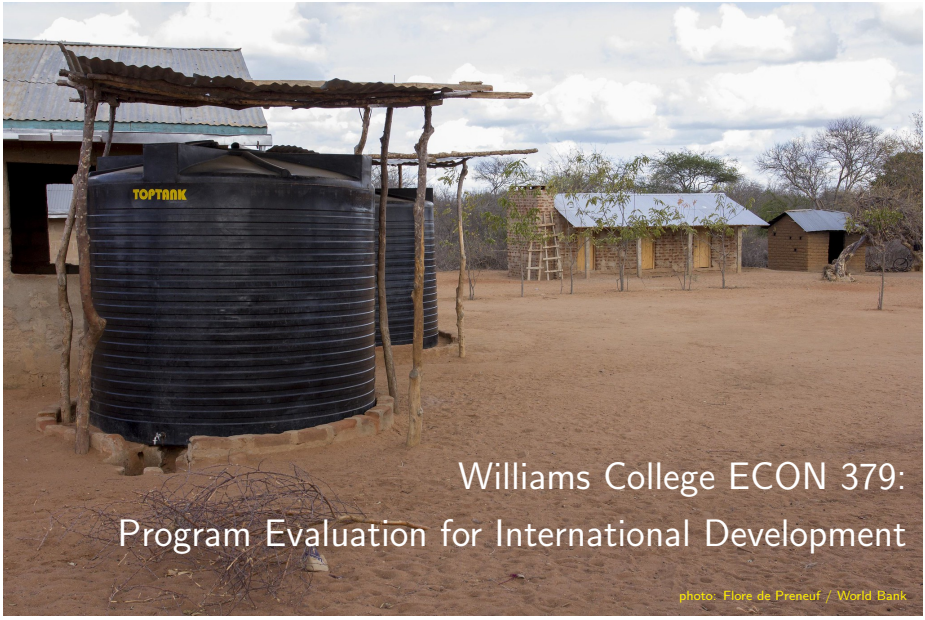
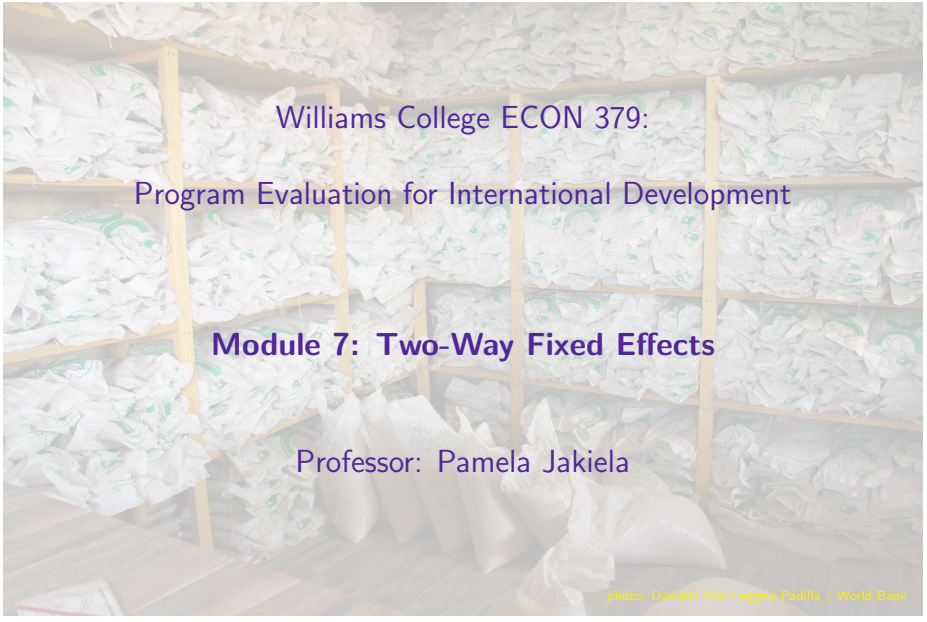


Williams College ECON 379:
Program Evaluation for International Development

photo: Flore de Preneuf / World Bank



Williams College ECON 379:

Program Evaluation for International Development

Module 7: Two-Way Fixed Effects

Professor: Pamela Jakiela

photo: Daniella Van Leggelo-Padilla / World Bank

Diff-in-Diff with Staggered Timing

Variation in Treatment Timing

Example: counties introduced food stamps at different times

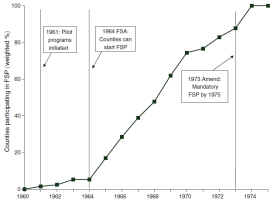


FIGURE 1. WEIGHTED PERCENT OF COUNTIES WITH FOOD STAMP PROGRAM, 1960–1975

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

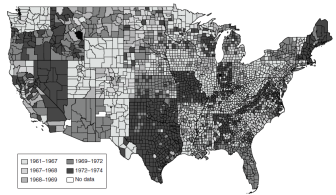


FIGURE 2. FOOD STAMP PROGRAM START DATE, BY COUNTY, 1961–1974

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

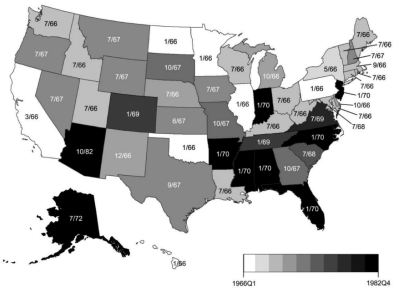


Figure 2.
Medicaid Adoption by Quarter
Notes: Adoption dates come from the Department of Health Education and Welfare (1970) & Social Security Administration (2013). The map is shaded relative to the quarter of adoption and states are labeled with the month and year of adoption.

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

Variation in Treatment Timing

Example: counties opening community health centers

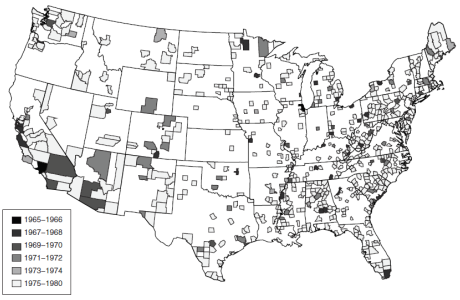


FIGURE 3. ESTABLISHMENT OF COMMUNITY HEALTH CENTERS BY COUNTY OF SERVICE DELIVERY, 1965–1980

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)

Variation in Treatment Timing

Example: African countries democratized at different times

FIGURE A.2. Geographical distribution of democratized countries since 1990. Black-colored countries are democratized since 1990: grey-colored countries the other countries in the sample for infant mortality analysis. Democratized countries include the Comoros, tiny islands to the northwest of Madagascar, which may not be visible as black-colored.

source: Kudamatsu (JEEA, 2012)

Two-Way Fixed Effects Estimates of β^{DD}

What exactly is β^{DD}?

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

unit fixed effects

time fixed effects

treatment dummy

that turns "on" at different times

What exactly is β^{DD} in TWFE?

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	
ID 1	3	3	5	5	5	treatment
ID 2	3	3	5	5	5	
ID 3	1	1	4	4	4	comparison
ID 4	1	1	4	4	4	

$$\Rightarrow \beta^{DD} = 5 - 3 - (4 - 1)$$

What exactly is β^{DD} in TWFE?

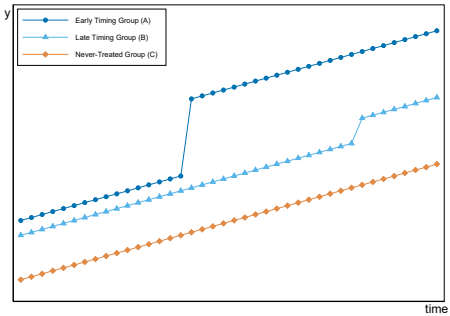
	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	
ID 1	3	3	5	5	5	treatment
ID 2	3	3	7	7	7	
ID 3	1	1	4	4	4	comparison
ID 4	1	1	4	4	4	

$\Rightarrow \beta^{DD} = 6 - 3 - (4 - 1)$ $\leftarrow$ average treatment effect

Multiple Treatment and Comparison Groups

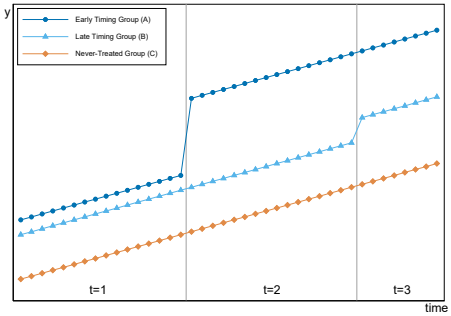
	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$
ID 1	0	0	1	1	1	1	1	1
ID 2	0	0	1	1	1	1	1	1
ID 3	0	0	0	0	0	1	1	1
ID 4	0	0	0	0	0	1	1	1
ID 5	0	0	0	0	0	0	0	0
ID 6	0	0	0	0	0	0	0	0

Decomposition into Timing Groups



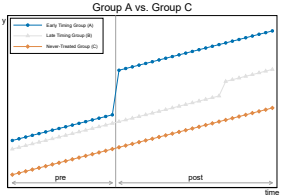
Panel with variation in treatment timing can be decomposed into distinct **timing groups** reflecting observed onset of treatment

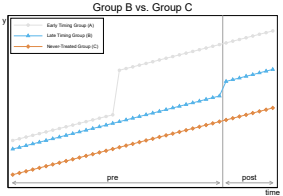
Decomposition into Timing Groups

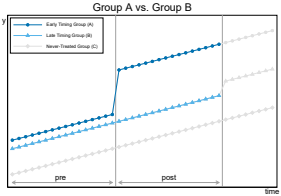


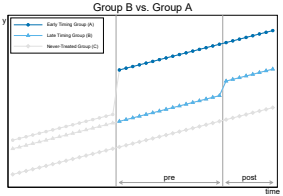
Example: with three timing groups (one of which is never treated), can construct three timing windows (pre, middle, post or $t = 1, 2, 3$)

Decomposition into Standard 2×2 DDs









Two-Way Fixed Effects β^{DD} as a Weighted Sum

The two-way fixed effects estimator β^{DD} is a weighted sum of all possible 2×2 diff-in-diff estimators across timing groups

- Some use an already-treated group as comparison
  - Creates problems if treatment effect grows/changes over time
  - If treatment effect gets increases over time within treated units, using already-treated units biases estimates of treatment effect
  - Bias intuitively similar to continuous treatment case
    - When treated units in comparison group, we are relying on assumptions about functional form (of treatment effect)

Easiest way to see this $\rightarrow$ unpack TWFE diff-in-diff estimator (as a weighted sum of Y values – but what are the weights?)

Two-Way Fixed Effects as Univariate Regression

Two-way fixed effects is equivalent to univariate regression:

$$\tilde{Y}_{it} = \alpha + \tilde{D}_{it} + \epsilon_{it}$$

where $\tilde{Y}_{it} = Y_{it} - \bar{Y}_t - \left(\bar{Y}_i - \bar{\bar{Y}} \right)$ and $\tilde{D}_{it}$ defined analogously

"grand mean"

(just the mean across i and t)

Two-Way Fixed Effects as Univariate Regression

Two-way fixed effects is equivalent to univariate regression:

$$\tilde{Y}_{it} = \alpha + \tilde{D}_{it} + \epsilon_{it}$$

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

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

$\Rightarrow$ Treatment "dummy" now continuous

$$\Rightarrow \beta^{OLS} = \sum_i \tilde{Y}_{it} \left(\tilde{D}_{it} - \bar{\tilde{D}}_{it} \right) \left(\frac{1}{\sum_i \left(\tilde{D}_{it} - \bar{\tilde{D}}_{it} \right)^2} \right)$$

Two-Way Fixed Effects as Univariate Regression

Two-way fixed effects is equivalent to univariate regression:

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

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

$\Rightarrow$ Treatment "dummy" now continuous

$\Rightarrow \beta^{OLS} = \sum_i \tilde{Y}_{it} \left(\tilde{D}_{it} - \bar{\tilde{D}}_{it} \right) \left(\frac{1}{\sum_i \left(\tilde{D}_{it} - \bar{\tilde{D}}_{it} \right)^2} \right)$

$\Rightarrow$ Observations where $\tilde{D}_{it} - \bar{\tilde{D}}_{it} < 0$ in comparison group

Diff-in-Diff with Staggered Treatment: Example

	$t = 1$	$t = 2$	$t = 3$	$t = 4$
ID 1	0	1	1	1
ID 2	0	0	0	1

Two timing groups, D_{it} is either 0 or 1

$\Rightarrow$ If we regressed Y_{it} on D_{it}, β_{OLS} is weighted sum of Y_{it} values

Diff-in-Diff with Staggered Treatment: Example

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	0	1	1	1
ID 2	0	0	0	1
$\bar{D}_t$	0	0.5	0.5	1

mean treatment in period t

Diff-in-Diff with Staggered Treatment: Example

	$t=1$	$t=2$	$t=3$	$t=4$	$\bar{D}_i$	$\bar{\bar{D}}$	$\bar{D}_i - \bar{\bar{D}}$
ID 1	0	1	1	1	0.75	0.5	0.25
ID 2	0	0	0	1	0.25	0.5	-0.25
$\bar{D}_t$	0	0.5	0.5	1			

mean treatment in period t

Diff-in-Diff with Staggered Treatment: Example

	D_{it} (treatment dummy)						
	$t=1$	$t=2$	$t=3$	$t=4$	$\bar{D}_i$	$\bar{\bar{D}}$	$\bar{D}_i - \bar{\bar{D}}$
ID 1	0	1	1	1	0.75	0.5	0.25
ID 2	0	0	0	1	0.25	0.5	-0.25
$\bar{D}_t$	0	0.5	0.5	1			

mean treatment in period t

Diff-in-Diff with Staggered Treatment: Example

$D_{it} - \bar{D}_t$

	$t=1$	$t=2$	$t=3$	$t=4$	$\bar{D}_i$	$\bar{\bar{D}}$	$\bar{D}_i - \bar{\bar{D}}$
ID 1	0	0.5	0.5	0	0.75	0.5	0.25
ID 2	0	−0.5	−0.5	0	0.25	0.5	−0.25
$\bar{D}_t$	0	0.5	0.5	1			

mean treatment in period t

Diff-in-Diff with Staggered Treatment: Example

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

	$t=1$	$t=2$	$t=3$	$t=4$	$\bar{D}_i$	$\bar{\bar{D}}$	$\bar{D}_i - \bar{\bar{D}}$
ID 1	-0.25	0.25	0.25	-0.25	0.75	0.5	0.25
ID 2	0.25	-0.25	-0.25	0.25	0.25	0.5	-0.25
$\bar{D}_t$	0	0.5	0.5	1			

mean treatment in period t

Diff-in-Diff with Staggered Treatment: Example

$$\tilde{D}_{it} = D_{it} - \bar{D}_t - \left(\bar{D}_i - \bar{\bar{D}}\right) \longrightarrow \text{need to divide by } \sum_i \left(\tilde{D}_{it} - \bar{\tilde{D}}_{it}\right)^2 \text{ to get OLS weights}$$

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	-0.25	0.25	0.25	-0.25
ID 2	0.25	-0.25	-0.25	0.25

Diff-in-Diff with Staggered Treatment: Example

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

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	-0.25	0.25	0.25	-0.25
ID 2	0.25	-0.25	-0.25	0.25

Treated cells are not all getting positive weight

Diff-in-Diff with Staggered Treatment: Example

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

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	-0.25	0.25	0.25	-0.25
ID 2	0.25	-0.25	-0.25	0.25

Treated cells are not all getting positive weight

$\Rightarrow$ Later treated periods in early adopter units negatively weighted

Diff-in-Diff with Staggered Treatment: Example

	Y_{it}			
	$t = 1$	$t = 2$	$t = 3$	$t = 4$
ID 1	0	10	10	10
ID 2	0	0	0	10

Let $Y_{it} = \gamma_i + \lambda_t + \delta_{it}$

Treated cells

Positive weights (treatment group)

Diff-in-Diff with Staggered Treatment: Example

	Y_{it}			
	$t = 1$	$t = 2$	$t = 3$	$t = 4$
ID 1	0	10	10	10
ID 2	0	0	0	10

Treated cells

Positive weights (treatment group)

Let $Y_{it} = \delta_{it}$

$\Rightarrow$ Y_{it} is treatment effect

Diff-in-Diff with Staggered Treatment: Example

$Y_{it} \longrightarrow \beta_{OLS} = 10$

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	0	10	10	10
ID 2	0	0	0	10

homogeneous impacts:
$E[\beta_{OLS}] = \text{ATE}$

Treated cells

Positive weights (treatment group)

Diff-in-Diff with Staggered Treatment: Example

$Y_{it} \longrightarrow \beta_{OLS} = 6$

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	0	2	2	2
ID 2	0	0	0	10

heterogeneous impacts:
$E[\beta_{OLS}]$ = weighted ATE

Treated cells

Positive weights (treatment group)

Diff-in-Diff with Staggered Treatment: Example

$Y_{it} \longrightarrow \beta_{OLS} = 6$

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	0	10	2	2
ID 2	0	0	0	2

heterogeneous impacts:
$E[\beta_{OLS}]$ = weighted ATE

(with mysterious weights)

Treated cells

Positive weights (treatment group)

Diff-in-Diff with Staggered Treatment: Example

$Y_{it} \longrightarrow \beta_{OLS} = -2$

	$t=1$	$t=2$	$t=3$	$t=4$
ID 1	0	2	2	10
ID 2	0	0	0	2

delayed impacts:

$E[\beta_{OLS}] =$ 

Treated cells

Positive weights (treatment group)

Practical Implications

Using diff-in-diff to evaluate staggered programs is complicated

- Avoid variation in treatment timing?
- A (substantial) never-treated group, many pre-periods helps
- Calculate your weights

Look at your data to assess common trends **after** treatment

- Graph the 2×2 diff-in-diffs
- Remove time FEs and plot the jumps at treatment

New tools for assessing the validity of diff-in-diff in these settings:

- Stata commands: `bacondecomp` and `fuzzydid`

The End!