

Difference in difference

Soc 114

Learning goals for today

At the end of class, you will be able to:

1. Estimate causal effects with unmeasured confounding using difference in difference
2. Understand the required data structure: one or more units become treated and one or more units do not
3. Explain the parallel trends assumption
4. Evaluate credibility of the parallel trends assumption

Card, D., & Krueger, A. B. (1994).

Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania.

The American Economic Review, 84(4), 772-793.

Economic theory

When the minimum wage rises, how might employment change?

Economic theory

When the minimum wage rises, how might employment change?

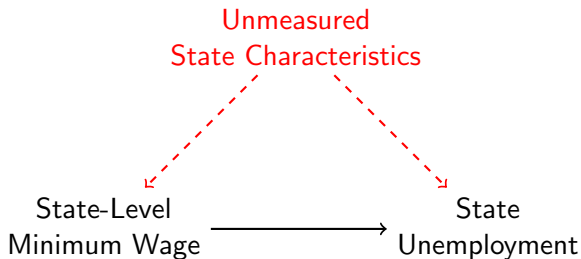
- ▶ employees cost more

Economic theory

When the minimum wage rises, how might employment change?

- ▶ employees cost more
- ▶ employers might get by with fewer employees

Problem: Unmeasured confounding



The setting

The setting

- ▶ Federal minimum wage
 - ▶ \$3.80 on April 1, 1990
 - ▶ \$4.25 on April 1, 1991

The setting

- ▶ Federal minimum wage
 - ▶ \$3.80 on April 1, 1990
 - ▶ \$4.25 on April 1, 1991
- ▶ New Jersey minimum wage
 - ▶ \$5.05 on April 1, 1992

NJ introduces a high minimum wage.

How would you study the effect on employment?

Source: Wikimedia

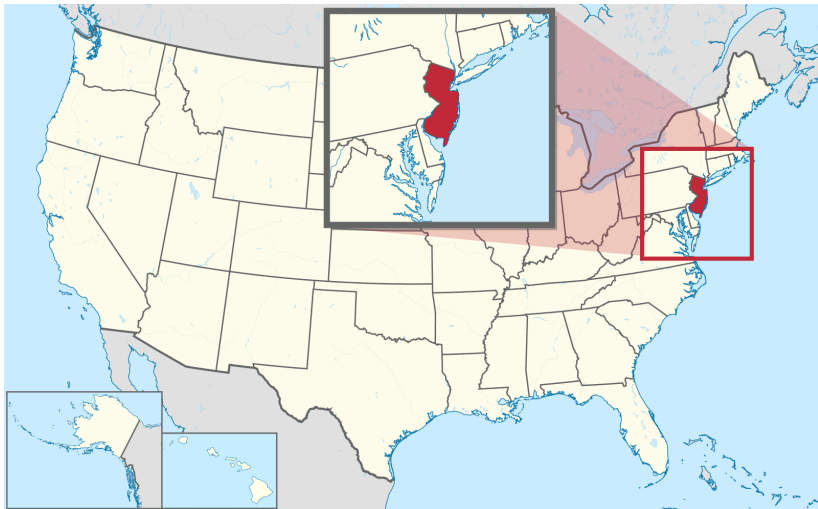




Photo by James Loesch - <https://www.flickr.com/photos/jal33/49113053632/>
CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=87207834>



New Jersey
Minimum Wage
Rose

Photo by James Loesch - <https://www.flickr.com/photos/jal33/49113053632/>
CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=87207834>



New Jersey
Minimum Wage
Rose

Pennsylvania
No change

Photo by James Loesch - <https://www.flickr.com/photos/jal33/49113053632/>
CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=87207834>





171 stores



80 stores



99 stores



60 stores



171 stores



80 stores



99 stores



60 stores

Phone interview: Feb-Mar 1992 before minimum wage rose
Nov-Dec 1992 after minimum wage rose



171 stores



80 stores



99 stores



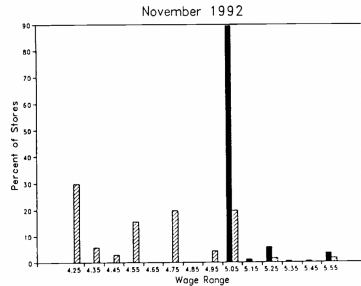
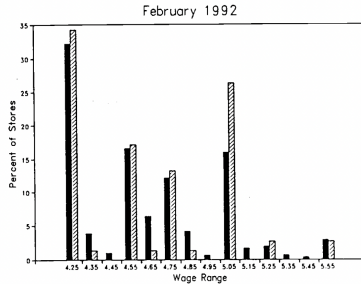
60 stores

Phone interview: Feb-Mar 1992 before minimum wage rose
Nov-Dec 1992 after minimum wage rose

Recorded: How many full-time equivalent employees?

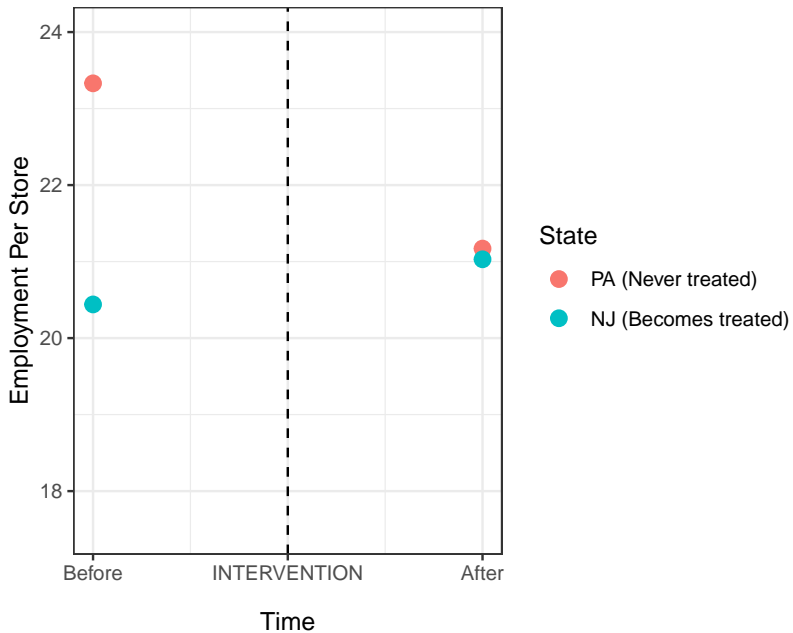
Did starting wages rise in NJ?

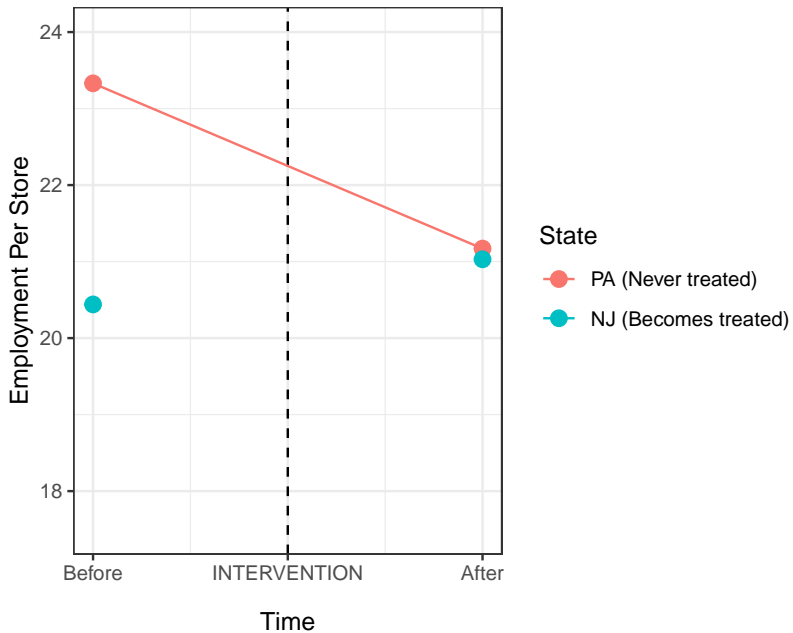
Figure 1
Distribution of Starting Wage Rates

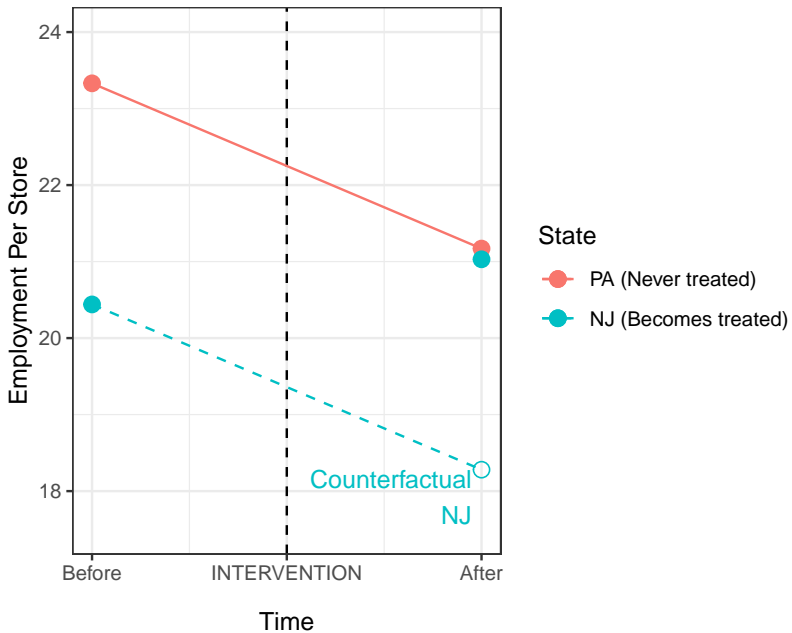


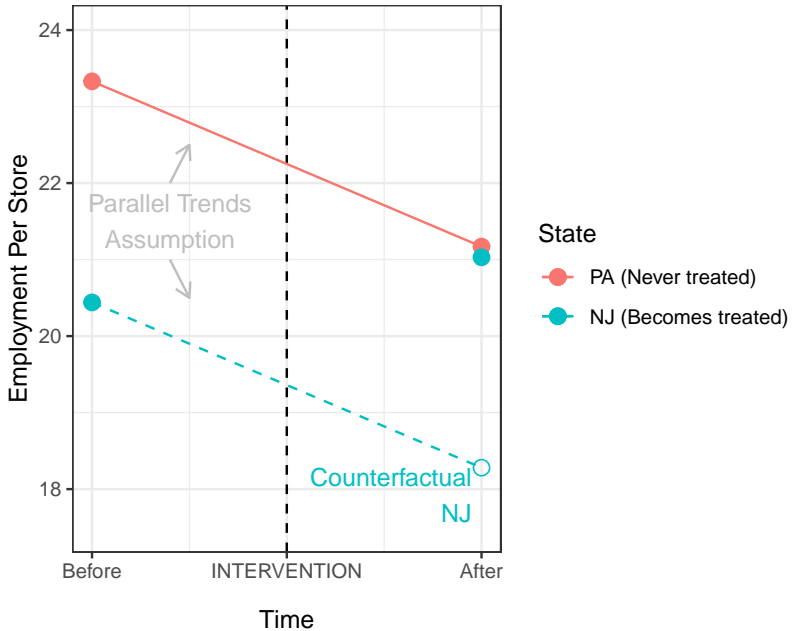
■ New Jersey ▨ Pennsylvanic

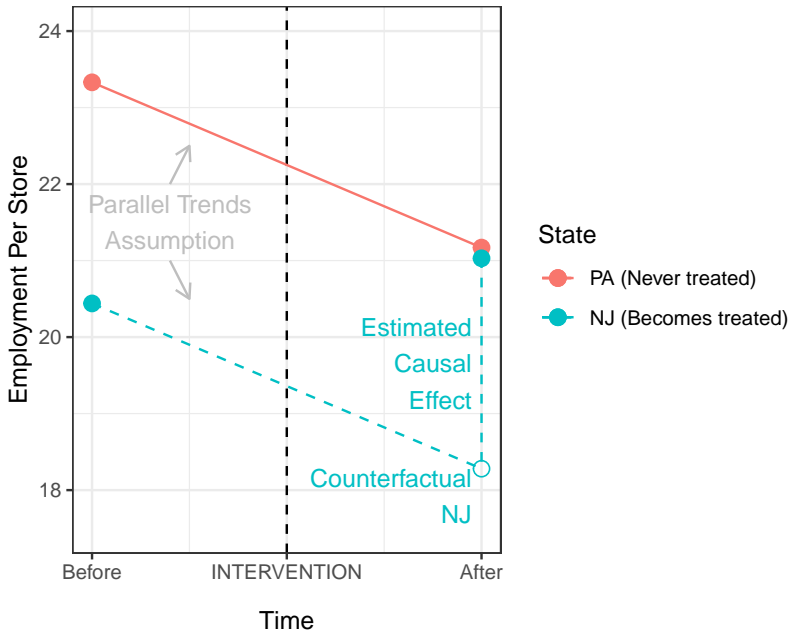
How did employment change?











Variable	Stores by state		
	PA (i)	NJ (ii)	Difference, NJ - PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	-2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	-0.14 (1.07)
3. Change in mean FTE employment	-2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

“Contrary to the central prediction of the textbook model of the minimum wage,...we find no evidence that the rise in New Jersey’s minimum wage reduced employment at fast-food restaurants in the state.”

Card & Krueger 1994, p. 792

- ▶ simple study
- ▶ well-executed
- ▶ upended conventional wisdom

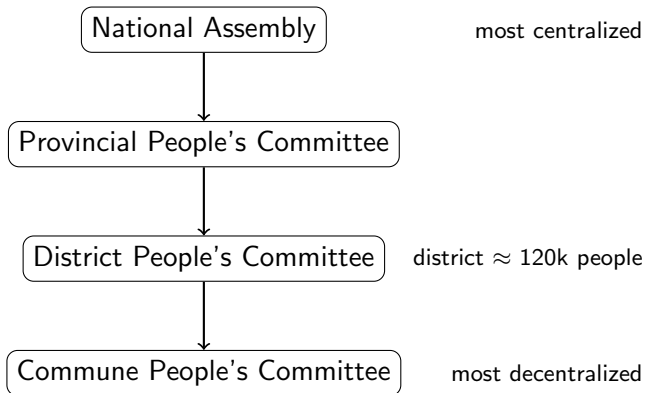
- ▶ simple study
- ▶ well-executed
- ▶ upended conventional wisdom

Decades of papers followed suit. Here is one.

Malesky, E. J., Nguyen, C. V., & Tran, A. (2014).
The impact of recentralization on public services:
A difference-in-differences analysis of the abolition of elected
councils in Vietnam.
American Political Science Review, 108(1), 144-168.

Does government work better when it is centralized or decentralized?

Vietnam setting: A study of recentralization



Vietnam setting: A study of recentralization

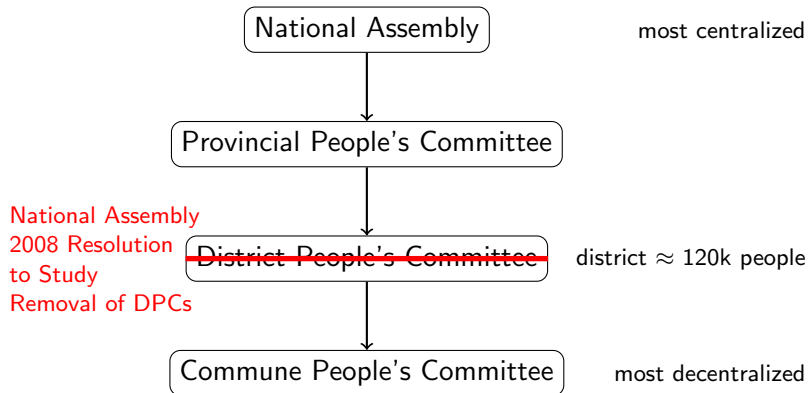
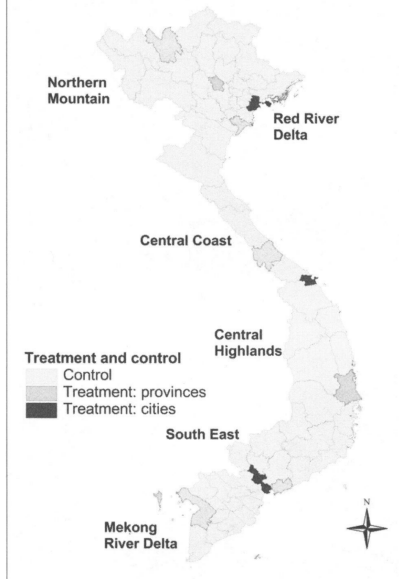


FIGURE 2. Map of Treatment Provinces and National-Level Cities



Vietnam setting: A study of recentralization

Input from social scientists

Vietnam setting: A study of recentralization

Input from social scientists

1. Enough treated units to study

Vietnam setting: A study of recentralization

Input from social scientists

1. Enough treated units to study
2. Sampling stratified by region

Vietnam setting: A study of recentralization

Input from social scientists

1. Enough treated units to study
2. Sampling stratified by region
3. Sampling stratified by
 - ▶ city versus rural
 - ▶ lowland versus highland
 - ▶ midland versus inter-nationally bordered land

Vietnam setting: A study of recentralization

Input from social scientists

1. Enough treated units to study
2. Sampling stratified by region
3. Sampling stratified by
 - ▶ city versus rural
 - ▶ lowland versus highland
 - ▶ midland versus inter-nationally bordered land
4. Sampling stratified by socioeconomic and public administration performance

Vietnam Household Living Standards Survey
Reports by each local commune by commune leaders

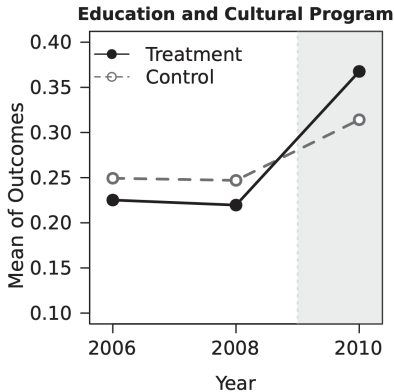
- ▶ 2006 and 2008: Before DPC abolition
- ▶ 2010: After DPC abolition

One outcome we will examine:

Is there the following project in the commune?

- ▶ Investment on culture and education

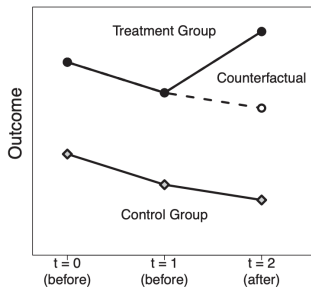
1. What is the parallel trends assumption here?
2. Is the parallel trends assumption credible?



Screenshot of re-analysis by [Egami & Yamauchi 2023](#), Fig 3

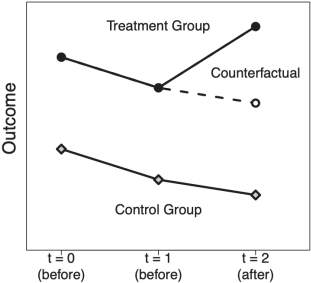
Assessing parallel trends

- ▶ The parallel trends assumption is that the treated group in the treatment period would have had the same trend as the control group
- ▶ It is most credible (but not guaranteed) when the two groups previously had similar trends

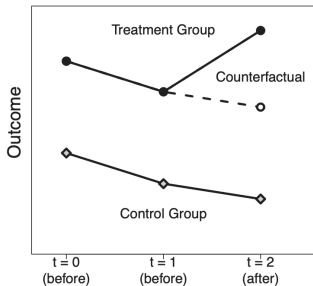


Egami, N., & Yamauchi, S. (2023).
Using multiple pretreatment periods to improve
difference-in-differences and staggered adoption designs.
Political Analysis, 31(2), 195-212.

Difference in difference



Difference in difference

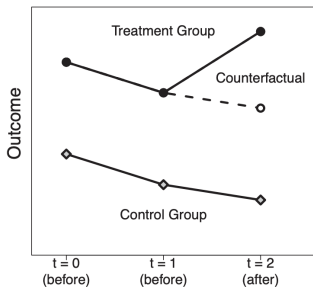


Notation

$Y_{(\text{unit})(\text{time})}^{\text{treatment value}}$

Example: Y_{i1}^0
is unit i at time 1
under treatment 0

Difference in difference



Notation

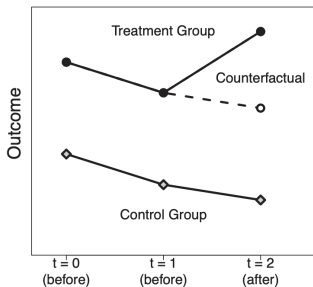
$Y_{(\text{unit})(\text{time})}^{\text{treatment value}}$

Example: Y_{i1}^0
is unit i at time 1
under treatment 0

Parallel Trends Assumption
(untestable)

$$E(Y_{\text{Treated},2}^0 - Y_{\text{Treated},1}^0)$$
$$=$$
$$E(Y_{\text{Control},2}^0 - Y_{\text{Control},1}^0)$$

Difference in difference



Notation

$Y_{(\text{unit})(\text{time})}^{\text{treatment value}}$

Example: Y_{i1}^0
is unit i at time 1
under treatment 0

Parallel Trends Assumption
(untestable)

$$E(Y_{\text{Treated},2}^0 - Y_{\text{Treated},1}^0) \\ = \\ E(Y_{\text{Control},2}^0 - Y_{\text{Control},1}^0)$$

Extended Parallel Trends
(testable)

$$E(Y_{\text{Treated},1}^0 - Y_{\text{Treated},0}^0) \\ = \\ E(Y_{\text{Control},1}^0 - Y_{\text{Control},0}^0)$$

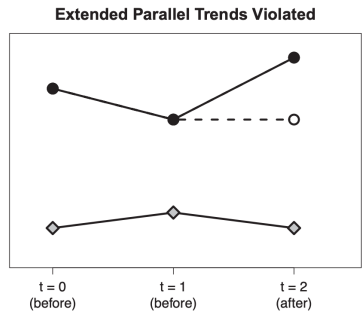
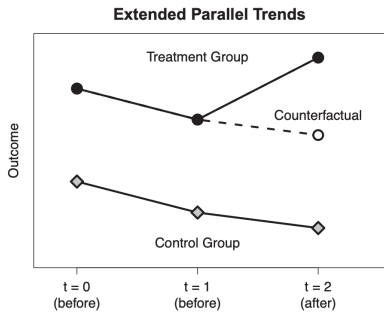


FIGURE 2. Map of Treatment Provinces and National-Level Cities

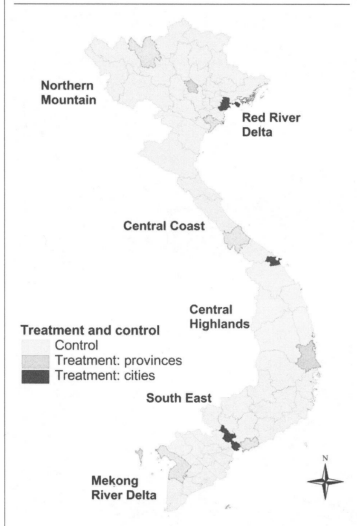
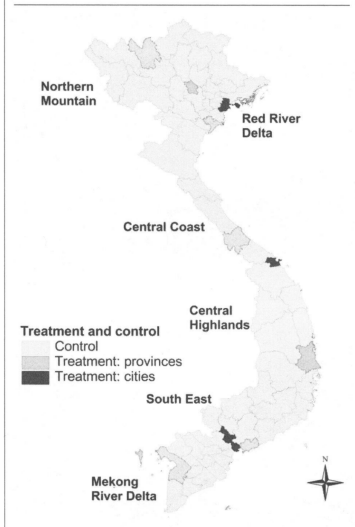


FIGURE 2. Map of Treatment Provinces and National-Level Cities



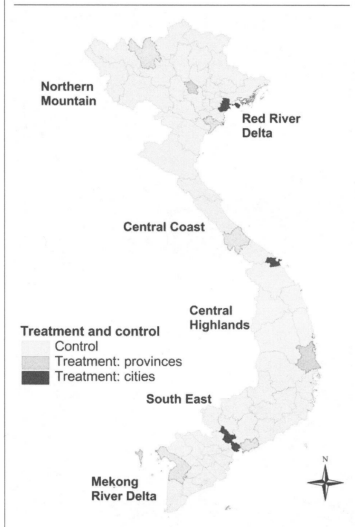
Outcome 1

Education and cultural programs

Is there the following project in the commune?

Investment on culture and education

FIGURE 2. Map of Treatment Provinces and National-Level Cities



Outcome 2

Tap water

Is there the following project in the commune?

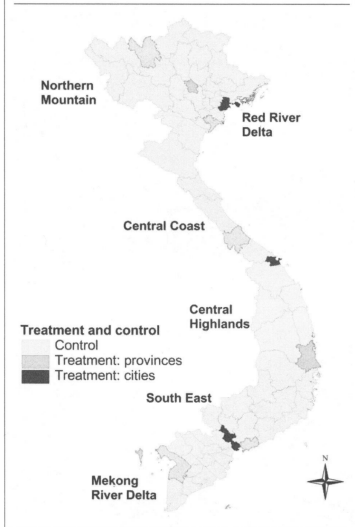
Coded 1

Indoor private piped water
Outdoor private piped water
Public piped water

Coded 0

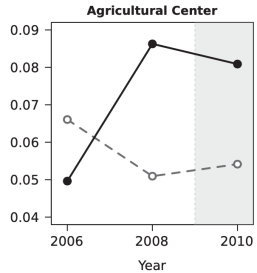
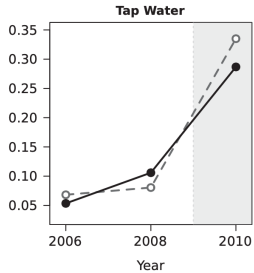
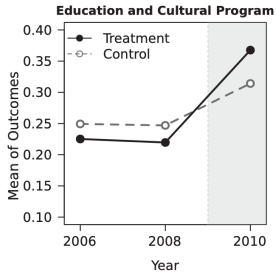
Well water
Well with protection walls
Well without protection walls
Stream water with protection
Stream water without protection
Rainwater
Bottled water
Water brought by pedicab
Tank water
river lake pond

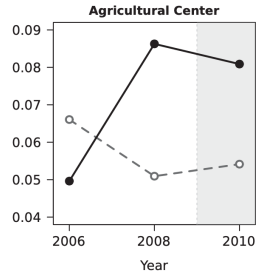
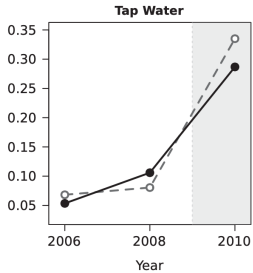
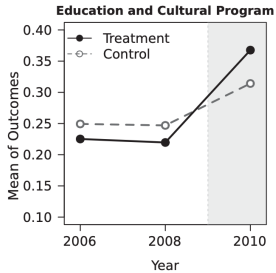
FIGURE 2. Map of Treatment Provinces and National-Level Cities



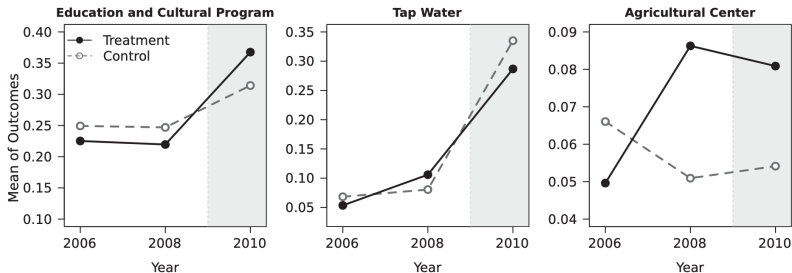
Outcome 3 Agricultural center

Is there any agriculture extension center in this commune?





In each case, do you believe parallel trends?



In each case, do you believe parallel trends?

Table 2. Assessing underlying assumptions using the pretreatment outcomes.

	Estimate	Std. error	p-value	95% Std. equivalence CI
Education and cultural program	-0.007	0.096	0.940	[-0.166, 0.166]
Tap water	0.166	0.083	0.045	[-0.302, 0.302]
Agricultural center	0.198	0.082	0.015	[-0.332, 0.332]

Benefit 1: Assessing assumptions

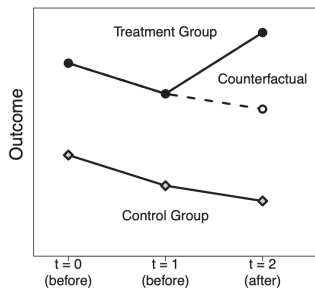
Pre-treatment periods enable us to
assess underlying assumptions

Parallel trends is untestable, but being parallel
in the pre-treatment period builds confidence

Benefit 2: Improving efficiency

Pre-treatment periods also enable us to
improve estimation accuracy
when parallel trends holds

Benefit 2: Improving efficiency



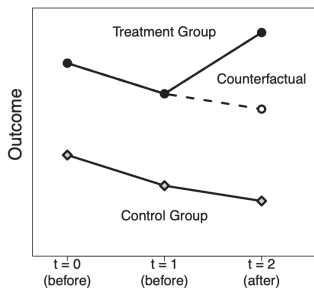
Estimator 1

Estimator 2

Notation

$y_{(\text{unit})(\text{time})}$ treatment value

Benefit 2: Improving efficiency



Estimator 1

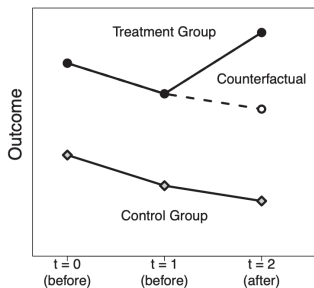
$$\underbrace{(\bar{Y}_{T2}^1 - \bar{Y}_{T1}^0)}_{\text{Treatment Group Time 2 - Time 1}} - \underbrace{(\bar{Y}_{C2}^0 - \bar{Y}_{C1}^0)}_{\text{Control Group Time 2 - Time 1}}$$

Estimator 2

Notation

$\bar{Y}_{(\text{unit})(\text{time})}$ treatment value

Benefit 2: Improving efficiency



Estimator 1

$$\underbrace{(\bar{Y}_{T2}^1 - \bar{Y}_{T1}^0)}_{\text{Treatment Group Time 2 - Time 1}} - \underbrace{(\bar{Y}_{C2}^0 - \bar{Y}_{C1}^0)}_{\text{Control Group Time 2 - Time 1}}$$

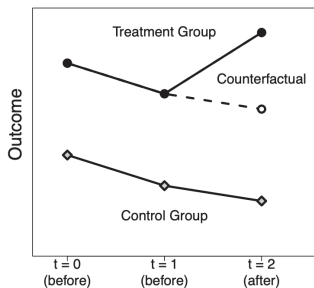
Estimator 2

$$\underbrace{(\bar{Y}_{T2}^1 - \bar{Y}_{T0}^0)}_{\text{Treatment Group Time 2 - Time 0}} - \underbrace{(\bar{Y}_{C2}^0 - \bar{Y}_{C0}^0)}_{\text{Control Group Time 2 - Time 0}}$$

Notation

$\bar{Y}_{(\text{unit})}^{\text{treatment value}}(\text{time})$

Benefit 2: Improving efficiency



Estimator 1

$$\underbrace{(\bar{Y}_{T2}^1 - \bar{Y}_{T1}^0)}_{\text{Treatment Group Time 2 - Time 1}} - \underbrace{(\bar{Y}_{C2}^0 - \bar{Y}_{C1}^0)}_{\text{Control Group Time 2 - Time 1}}$$

Estimator 2

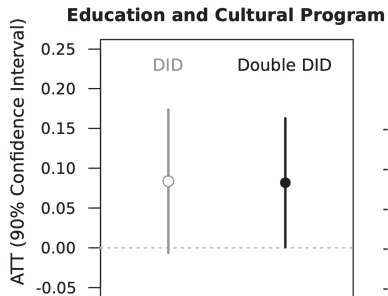
$$\underbrace{(\bar{Y}_{T2}^1 - \bar{Y}_{T0}^0)}_{\text{Treatment Group Time 2 - Time 0}} - \underbrace{(\bar{Y}_{C2}^0 - \bar{Y}_{C0}^0)}_{\text{Control Group Time 2 - Time 0}}$$

Notation

$\bar{Y}_{(\text{unit})}^{\text{treatment value}}(\text{time})$

Pooled estimator:
Average the two!

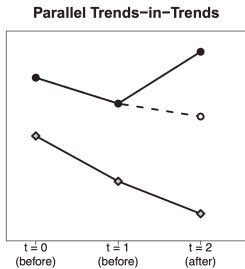
Benefit 2: Improving efficiency



Benefit 3: A more flexible assumption

Pre-treatment periods make it possible to
allow for a more flexible parallel trends assumption

Benefit 3: A more flexible assumption

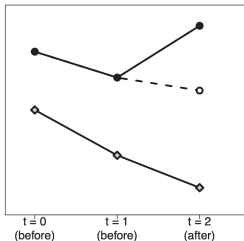


Trend of Treatment Group
(-2, -1)

Trend of Control Group
(-3.5, -2.5)

Benefit 3: A more flexible assumption

Parallel Trends-in-Trends



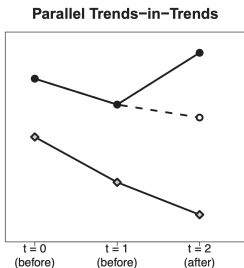
Trend of Treatment Group
(-2, -1)

Trend of Control Group
(-3.5, -2.5)

ASSUMPTION 3 (Parallel Trends-in-Trends)

$$\begin{aligned} & \underbrace{\{E[Y_{i2}(0) | G_i = 1] - E[Y_{i1}(0) | G_i = 1]\}}_{\text{Trend of the treatment group from } t=1 \text{ to } t=2} - \underbrace{\{E[Y_{i1}(0) | G_i = 1] - E[Y_{i0}(0) | G_i = 1]\}}_{\text{Trend of the treatment group from } t=0 \text{ to } t=1} \\ &= \underbrace{\{E[Y_{i2}(0) | G_i = 0] - E[Y_{i1}(0) | G_i = 0]\}}_{\text{Trend of the control group from } t=1 \text{ to } t=2} - \underbrace{\{E[Y_{i1}(0) | G_i = 0] - E[Y_{i0}(0) | G_i = 0]\}}_{\text{Trend of the control group from } t=0 \text{ to } t=1}. \end{aligned}$$

Benefit 3: A more flexible assumption



Trend of Treatment Group
(-2, -1)

Trend of Control Group
(-3.5, -2.5)

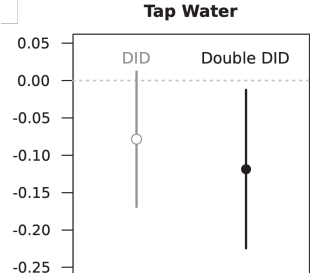
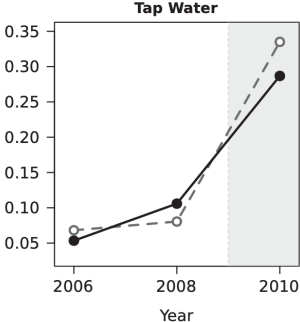
ASSUMPTION 3 (Parallel Trends-in-Trends)

$$\begin{aligned}
 & \underbrace{\{E[Y_{i2}(0) | G_i = 1] - E[Y_{i1}(0) | G_i = 1]\}}_{\text{Trend of the treatment group from } t=1 \text{ to } t=2} - \underbrace{\{E[Y_{i1}(0) | G_i = 1] - E[Y_{i0}(0) | G_i = 1]\}}_{\text{Trend of the treatment group from } t=0 \text{ to } t=1} \\
 &= \underbrace{\{E[Y_{i2}(0) | G_i = 0] - E[Y_{i1}(0) | G_i = 0]\}}_{\text{Trend of the control group from } t=1 \text{ to } t=2} - \underbrace{\{E[Y_{i1}(0) | G_i = 0] - E[Y_{i0}(0) | G_i = 0]\}}_{\text{Trend of the control group from } t=0 \text{ to } t=1}.
 \end{aligned}$$

Sequential DID Estimator

$$\begin{aligned}
 \hat{\tau}_{s-DID} = & \left\{ \left(\frac{\sum_{i: G_i=1} Y_{i2}}{n_{12}} - \frac{\sum_{i: G_i=1} Y_{i1}}{n_{11}} \right) - \left(\frac{\sum_{i: G_i=0} Y_{i2}}{n_{02}} - \frac{\sum_{i: G_i=0} Y_{i1}}{n_{01}} \right) \right\} \\
 & - \left\{ \left(\frac{\sum_{i: G_i=1} Y_{i1}}{n_{11}} - \frac{\sum_{i: G_i=1} Y_{i0}}{n_{10}} \right) - \left(\frac{\sum_{i: G_i=0} Y_{i1}}{n_{01}} - \frac{\sum_{i: G_i=0} Y_{i0}}{n_{00}} \right) \right\},
 \end{aligned}$$

Benefit 3: A more flexible assumption

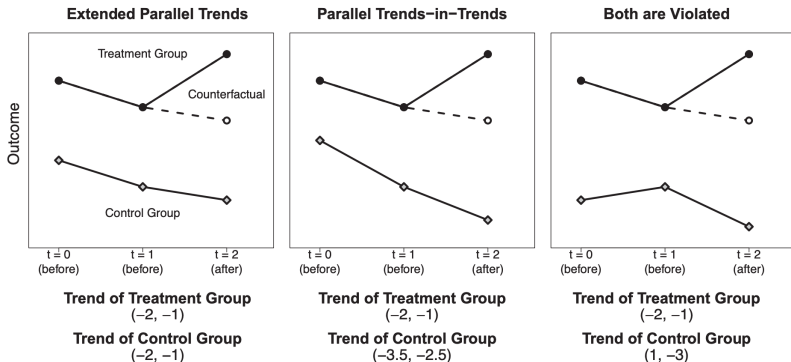


Benefits of multiple pre-treatment periods

1. assess underlying assumptions
2. improve estimation accuracy
3. allow for a more flexible parallel trends assumption

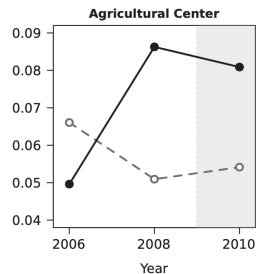
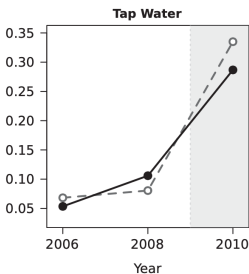
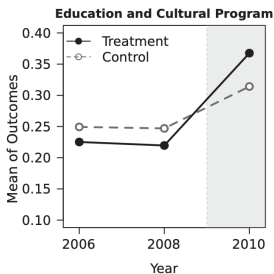
Benefits of multiple pre-treatment periods

1. assess underlying assumptions
2. improve estimation accuracy
3. allow for a more flexible parallel trends assumption



Benefits of multiple pre-treatment periods

1. assess underlying assumptions
2. improve estimation accuracy
3. allow for a more flexible parallel trends assumption



Learning goals for today

At the end of class, you will be able to:

1. Estimate causal effects with unmeasured confounding using difference in difference
2. Understand the required data structure: one or more units become treated and one or more units do not
3. Explain the parallel trends assumption
4. Evaluate credibility of the parallel trends assumption