Lecture 3:
Difference in Difference 1 of 2

January 28, 2026

Course Administration

1. Summaries are graded
    - I invited you to a google folder
    - All returned work will go there
2. Thanks for proposals – will grade by next week
3. Problem set 2 posted
4. Will post answers to PS 1, including code
5. Any other issues?

Today

Problem Set 1 quiz

Today

Problem Set 1 quiz

Diff-in-diff overview

1. When to use diff-in-diff
2. Motivating example
3. Diff-in-diff version 1
4. With potential outcomes notation
5. Writing and interpreting a DD regression
6. Failure of parallel trends

Milligan and the Stork

1. Estimation problem
2. Data
3. Diff-in-diff in chart
4. Diff-in-diff in table
5. Diff-in-diff in regression

Next time: DD problems and including a “trend”

1. Motivating Diff-in-Diff

Motivating Diff-in-Diff

1. When should you use diff-in-diff?
2. Motivating example
3. Diff-in-diff v1
4. With potential outcomes notation
5. Writing and interpreting a diff-in-diff regression
6. Failure of parallel trends

Next time: validity tests and trends

0. Why Bother? Or, Why Not Regression with Covariates?

- OLS with covariates is unlikely to deliver a causal estimate of $\hat{\beta}$

0. Why Bother? Or, Why Not Regression with Covariates?

- OLS with covariates is unlikely to deliver a causal estimate of $\hat{\beta}$
- So we need a causal strategy
- Diff-in-diff is a causal strategy

1.1 When to use diff-in-diff?

When to Use a Difference in Difference Methodology?

- To evaluate the impact of a policy at an aggregate level
- Where you have some potential control group

When to Use a Difference in Difference Methodology?

- To evaluate the impact of a policy at an aggregate level
- Where you have some potential control group
- Groups are frequently but not necessarily geographic
- For example: national policy that affects some groups by not others
- Examples?

When to Use a Difference in Difference Methodology?

- To evaluate the impact of a policy at an aggregate level
- Where you have some potential control group
- Groups are frequently but not necessarily geographic
- For example: national policy that affects some groups by not others
- Examples? EITC evaluation that compares women with children versus those without

1.2 Motivating Example

Motivating Example: Card and Krueger, *AER*, 1991

Policy

- April 1992
  - NJ and PA have the same minimum wage of \$4.25/hour
- April 1992 onward
  - NJ raises state minimum wage to \$5.05/hour, no change in PA

Motivating Example: Card and Krueger, *AER*, 1991

Policy

- April 1992
  - NJ and PA have the same minimum wage of \$4.25/hour
- April 1992 onward
  - NJ raises state minimum wage to \$5.05/hour, no change in PA

Data

- C&K collect data on employment and wages at fast food places in NJ and E PA
- observe data from February to November 1992

1.3 Diff-in-Diff Version 1

With This Setup, How Do We Estimate?

We observe

- Employment in NJ before and after
  - NJ_B and NJ_A
- Employment in PA before and after
  - PA_B and PA_A

With This Setup, How Do We Estimate?

Why not $(NJ_A - NJ_B)$?

We observe

- Employment in NJ before and after
  - NJ_B and NJ_A
- Employment in PA before and after
  - PA_B and PA_A

With This Setup, How Do We Estimate?

Why not $(NJ_A - NJ_B)$?
Estimating

We observe

- Employment in NJ before and after
  - NJ_B and NJ_A
- Employment in PA before and after
  - PA_B and PA_A

With This Setup, How Do We Estimate?

We observe

- Employment in NJ before and after
  - NJ_B and NJ_A
- Employment in PA before and after
  - PA_B and PA_A

Why not $(NJ_A - NJ_B)$?
Estimating

$$(NJ_A - NJ_B) - (PA_A - PA_B)$$

With This Setup, How Do We Estimate?

We observe

- Employment in NJ before and after
  - NJ_B and NJ_A
- Employment in PA before and after
  - PA_B and PA_A

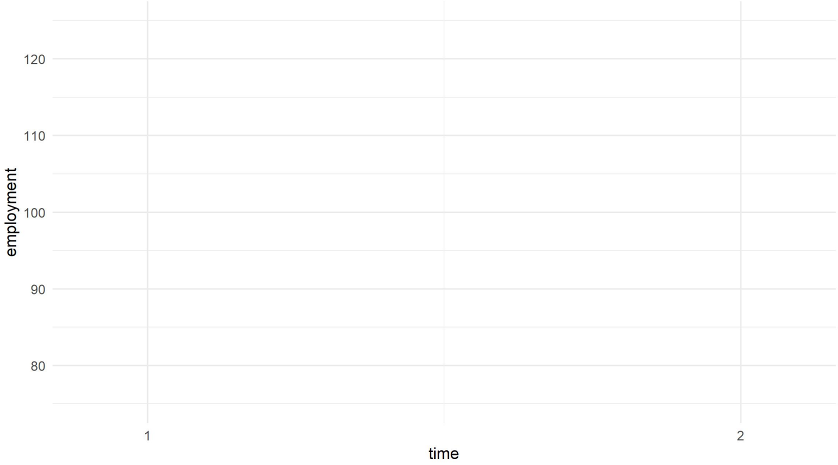
Why not $(NJ_A - NJ_B)$?
Estimating

$$(NJ_A - NJ_B) - (PA_A - PA_B)$$

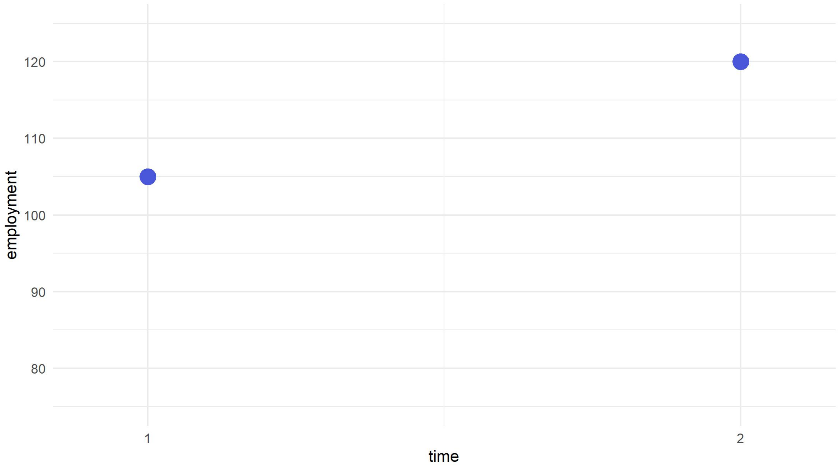
or

$$(NJ_A - PA_A) - (NJ_B - PA_B)$$

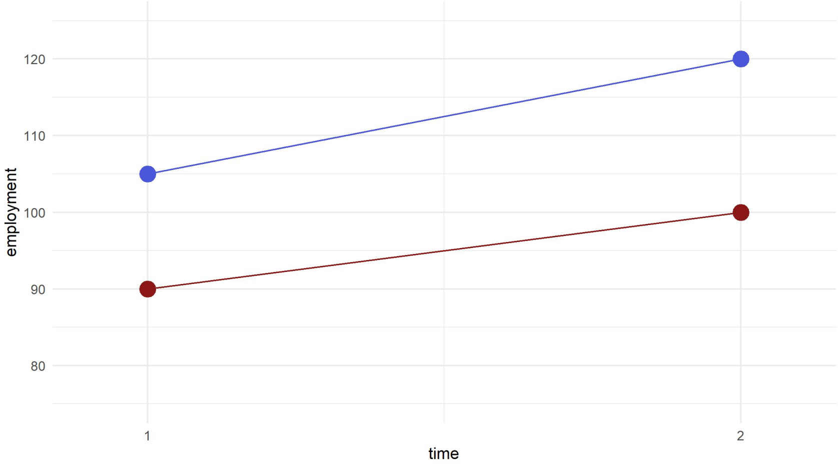
In Graph Form



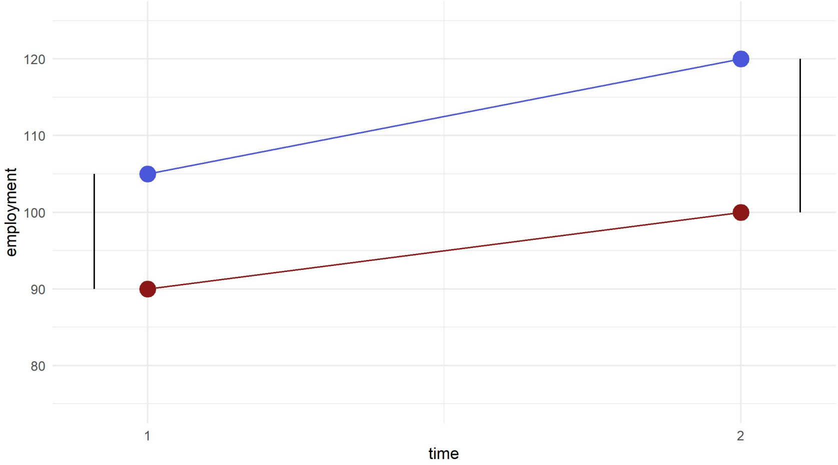
This is NJ Only – Why Not This Comparison?



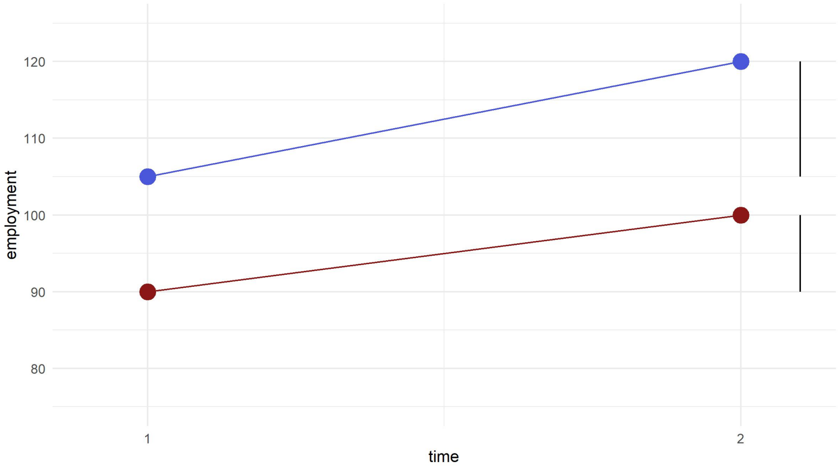
Here are Both: Where is Double Difference?



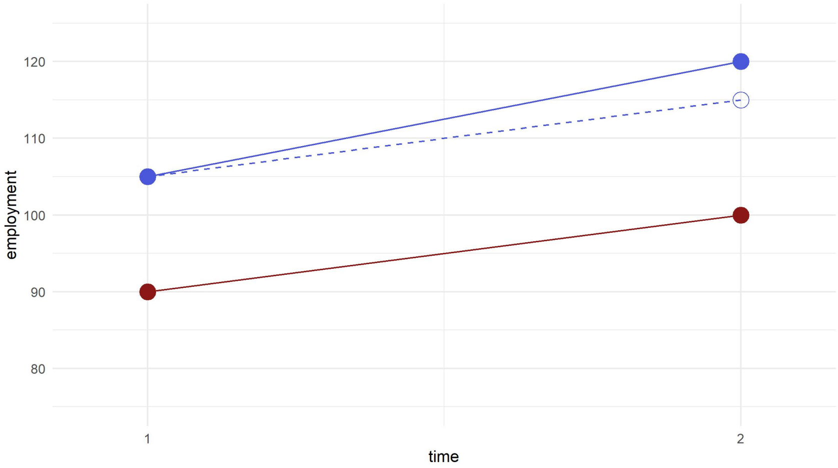
Double Difference v.1



Double Difference v.2



Or, the Implicit Comparison



1.4 Potential Outcomes Framework

Card and Krueger in a Potential Outcomes Framework

- $Y_{0ist} \equiv$ fast food employment at restaurant i, state s, period t with the low minimum wage
- $Y_{1ist} \equiv$ fast food employment at restaurant i, state s, period t with the high minimum wage
- Recall that we only observe one of these for any given t
- State $s \in \{\text{NJ}, \text{PA}\}$
- Time period $t \in \{\text{before}, \text{after}\}$

Key Assumptions

1. $E[Y_{0ist}|s,t] = \gamma_s + \lambda_t$

Key Assumptions

1. $E[Y_{0ist}|s,t] = \gamma_s + \lambda_t$
   - $\gamma_s \equiv$ state fixed effects
   - $\lambda_t \equiv$ time fixed effects

Key Assumptions

1. $E[Y_{0ist}|s,t] = \gamma_s + \lambda_t$
   - $\gamma_s \equiv$ state fixed effects
   - $\lambda_t \equiv$ time fixed effects
   - In words: the outcome, conditional on state and time, can be explained by something fixed about the state, and something fixed in a given time period for all states
   - Note that γ_s does not have to be the same for all states
   - Give an example where you think this isn't true

Key Assumptions

1. $E[Y_{0ist}|s, t] = \gamma_s + \lambda_t$
    - $\gamma_s \equiv$ state fixed effects
    - $\lambda_t \equiv$ time fixed effects
    - In words: the outcome, conditional on state and time, can be explained by something fixed about the state, and something fixed in a given time period for all states
    - Note that γ_s does not have to be the same for all states
    - Give an example where you think this isn't true
    - This is the "common" or "parallel trends" assumption

Key Assumptions

1. $E[Y_{0ist}|s,t] = \gamma_s + \lambda_t$
   - $\gamma_s \equiv$ state fixed effects
   - $\lambda_t \equiv$ time fixed effects
   - In words: the outcome, conditional on state and time, can be explained by something fixed about the state, and something fixed in a given time period for all states
   - Note that γ_s does not have to be the same for all states
   - Give an example where you think this isn't true
   - This is the "common" or "parallel trends" assumption
2. $E[Y_{1ist} - Y_{0ist}|s,t] = \delta$
   - Change between treated and untreated states is a level difference – it's additive, not multiplicative, or some other function

1.5 Difference in difference estimation

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

and note that $NJ_s * d_t$ is the treatment

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment

- Break down this equation
  - PA before?

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after?

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before?

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after?

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after? $\alpha + \gamma + \lambda + \delta$

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after? $\alpha + \gamma + \lambda + \delta$
  - Can you see diff-in-diff?

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after? $\alpha + \gamma + \lambda + \delta$
  - Can you see diff-in-diff?

$$(\text{NJ after} - \text{NJ before}) - (\text{PA after} - \text{PA before}) \quad =$$

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after? $\alpha + \gamma + \lambda + \delta$
  - Can you see diff-in-diff?

$$\begin{aligned}(\text{NJ after} - \text{NJ before}) - (\text{PA after} - \text{PA before}) &= \\ ((\alpha + \gamma + \lambda + \delta) - (\alpha + \gamma)) - ((\alpha + \lambda) - (\alpha))\end{aligned}$$

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after? $\alpha + \gamma + \lambda + \delta$
  - Can you see diff-in-diff?

$$\begin{aligned}
\text{(NJ after} - \text{NJ before)} - \text{(PA after} - \text{PA before)} &= \\
((\alpha + \gamma + \lambda + \delta) - (\alpha + \gamma)) - ((\alpha + \lambda) - (\alpha)) &= \delta
\end{aligned}$$

Regression Specification and Interpretation

- In the regression world, we write the regression equation as

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta NJ_s * d_t + \epsilon_{ist}$$

  and note that $NJ_s * d_t$ is the treatment
- Break down this equation
  - PA before? α
  - PA after? $\alpha + \lambda$
  - NJ before? $\alpha + \gamma$
  - NJ after? $\alpha + \gamma + \lambda + \delta$
  - Can you see diff-in-diff?

$$\begin{aligned}
(\text{NJ after} - \text{NJ before}) - (\text{PA after} - \text{PA before}) &= \\
((\alpha + \gamma + \lambda + \delta) - (\alpha + \gamma)) - ((\alpha + \lambda) - (\alpha)) &= \delta
\end{aligned}$$

- Note that you can estimate this with sample means! A very good place to start, for reasons we will talk about next week

1.6 Failure in Parallel Trends

Example: Clean Air Act Analysis

What is the impact of the Clean Air Act?

- Clean Air Act regulations pollution, including PM 2.5
- If a county has PM 2.5 that is "too high" it is a "non-attainment area"
- When a county is so designated, the EPA can limit funding or permitting
- Does this improve air quality?

Example: Clean Air Act Analysis

What is the impact of the Clean Air Act?

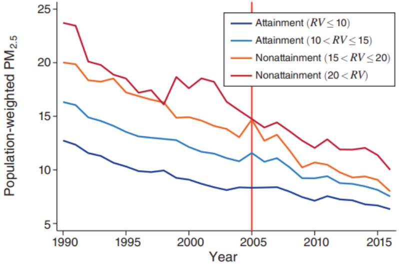
- Clean Air Act regulations pollution, including PM 2.5
- If a county has PM 2.5 that is "too high" it is a "non-attainment area"
- When a county is so designated, the EPA can limit funding or permitting
- Does this improve air quality?
- Compare attainment to non-attainment areas

Example: Clean Air Act Analysis

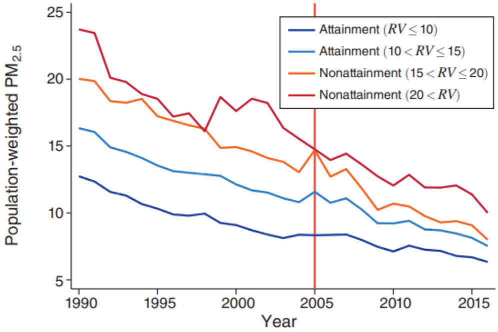
What is the impact of the Clean Air Act?

- Clean Air Act regulations pollution, including PM 2.5
- If a county has PM 2.5 that is "too high" it is a "non-attainment area"
- When a county is so designated, the EPA can limit funding or permitting
- Does this improve air quality?
- Compare attainment to non-attainment areas

Does it look so bad?

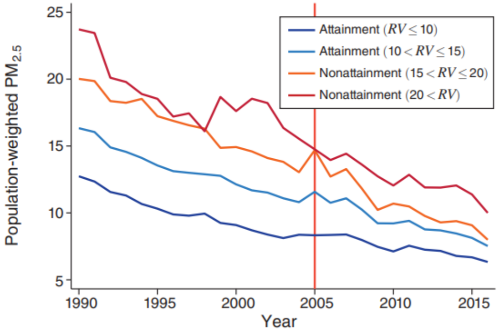


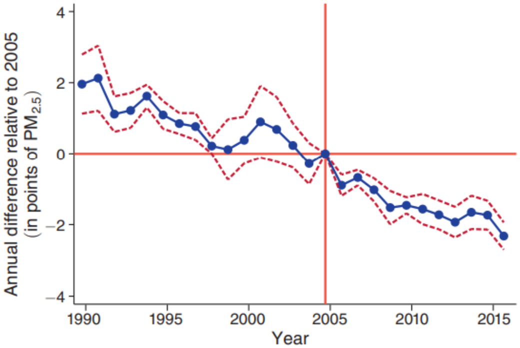
Comparing Attainment to Non-Attainment Areas



Translate into a regression to compare.

Comparing Attainment to Non-Attainment Areas

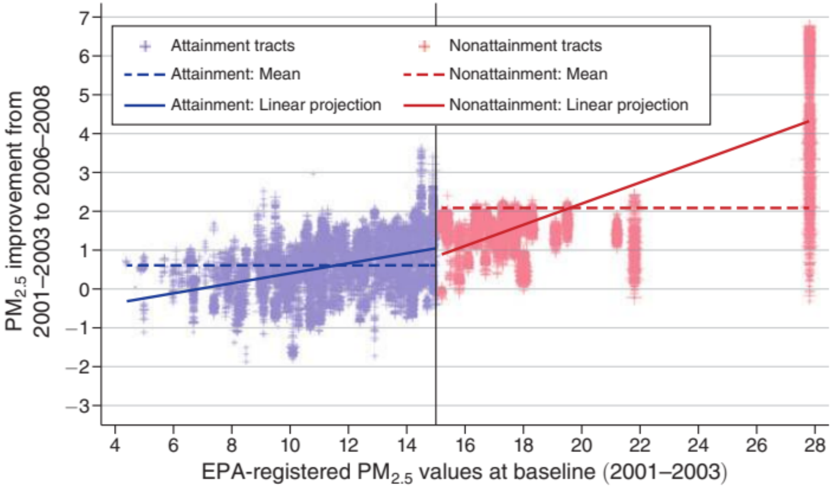




Translate into a regression to compare.

$$
\begin{aligned}
PM_{i,t} = & \quad \beta_0 + \beta_1 \mathrm{NA}_{i,t} + \beta_2 \mathrm{time}_t \\
+ & \quad \beta_3 \mathrm{NA}_{i,t} * \mathrm{time}_t + \epsilon_{i,t}
\end{aligned}
$$

Decreases are Larger Where PM 2.5 is Higher



Recap: Key Parts

Key Assumption

- In the absence of treatment, treatment and control observations would have evolved in parallel fashion
- AKA, "parallel trends"
- Fundamentally untestable
- Phrased differently: the only difference between treatment and control, apart from any level differences, is treatment

Recap: Key Parts

Key Assumption

- In the absence of treatment, treatment and control observations would have evolved in parallel fashion
- AKA, "parallel trends"
- Fundamentally untestable
- Phrased differently: the only difference between treatment and control, apart from any level differences, is treatment

Why a regression?

- a convenient way to get estimates and standard errors
- can do more policies (e.g. put in value of wage changes)
- can add controls, if parallel trend assumption is only valid conditionally, or if we want to reduce variance

2. Setting up the Milligan et al paper

Research Question and Estimation Problem

1. What is the research question?

Research Question and Estimation Problem

1. What is the research question?
2. What does this mean? "... empirical researchers have shown great interest in trying to uncover evidence of a relationship between prices and fertility. The endogeneity of key variables has frustrated this effort.

Research Question and Estimation Problem

1. What is the research question?
2. What does this mean? "... empirical researchers have shown great interest in trying to uncover evidence of a relationship between prices and fertility. The endogeneity of key variables has frustrated this effort. Women may have unobserved proclivities for different family sizes.

Research Question and Estimation Problem

1. What is the research question?
2. What does this mean? "... empirical researchers have shown great interest in trying to uncover evidence of a relationship between prices and fertility. The endogeneity of key variables has frustrated this effort. Women may have unobserved proclivities for different family sizes. If differences in these proclivities lead to different human capital accumulation and marital decisions, then the opportunity cost of time out of the labor market will be jointly determined with fertility."

Omitted Variable Bias

- What are the two components of an omitted variable problem/story? An omitted variable is

Omitted Variable Bias

- What are the two components of an omitted variable problem/story? An omitted variable is
  1. correlated with the treatment
  2. and with the outcome conditional on covariates (aka the error)

Omitted Variable Bias

- What are the two components of an omitted variable problem/story? An omitted variable is
  1. correlated with the treatment
  2. and with the outcome conditional on covariates (aka the error)
- Give an example of a potential omitted variable in this paper

What's the Argument for How the ANC Solves This Problem?

What's the Argument for How the ANC Solves This Problem?

- **Large** subsidy – at peak $>$ \$15k today
- Plausibly not related to fertility
- Everyone in Quebec treated similarly
- $\rightarrow$ compare Quebec and Rest of Canada
- Can also use a triple difference

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

- from birth records

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

- from birth records
  - $\rightarrow$ unit of observation is woman in a year

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

- from birth records
  - $\rightarrow$ unit of observation is woman in a year
- aggregates to fertility rates by cohort/province/parity

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

- from birth records
  - $\rightarrow$ unit of observation is woman in a year
- aggregates to fertility rates by cohort/province/parity
  - $\rightarrow$ unit of analysis is province/year

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

- from birth records
  - $\rightarrow$ unit of observation is woman in a year
- aggregates to fertility rates by cohort/province/parity
  - $\rightarrow$ unit of analysis is province/year

Canadian Census data

Data and Units of Observation and Analysis

What are the two data sources?

Vital statistics data

- from birth records
  - → unit of observation is woman in a year
- aggregates to fertility rates by cohort/province/parity
  - → unit of analysis is province/year

Canadian Census data

- from 1991 and 1996
- covering five prior years
- unit of observation and analysis is family

Basic Diff-in-diff

- We need
  - treated and untreated
  - before and after

Basic Diff-in-diff

- We need
  - treated and untreated
  - before and after
- What are these here?

Basic Diff-in-diff

- We need
  - treated and untreated
  - before and after
- What are these here?
  - before and after: before ANC and during ANC

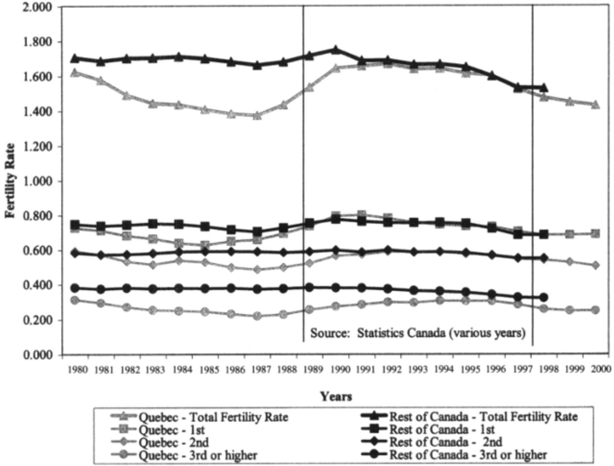
Basic Diff-in-diff

- We need
  - treated and untreated
  - before and after
- What are these here?
  - before and after: before ANC and during ANC
  - treated and untreated: Quebec and Rest of Canada

3. Estimation in Milligan

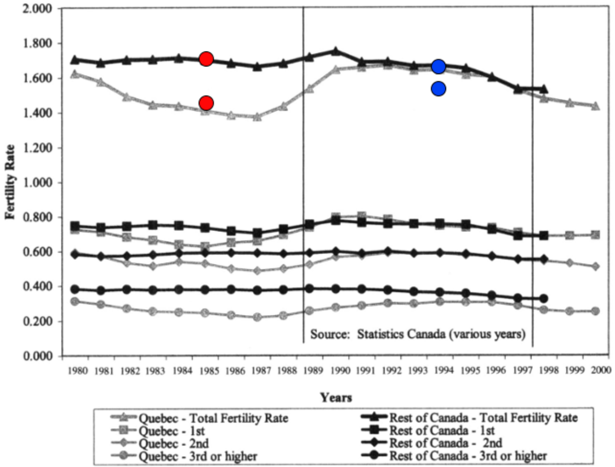
Basic Diff-in-diff in Figure 1

For the simplest diff-in-diff, what are the two comparisons?



Basic Diff-in-diff in Figure 1

For the simplest diff-in-diff, what are the two comparisons?



The Diff-in-diff in Table 5

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

- How do you calculate 0.418?

The Diff-in-diff in Table 5

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

- How do you calculate 0.418?
- And 0.441?

The Diff-in-diff in Table 5

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

- How do you calculate 0.418?
- And 0.441?
- And Col. 3, 0.033?

The Diff-in-diff in Table 5

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

- How do you calculate 0.418?
- And 0.441?
- And Col. 3, 0.033?

- How do we find 0.024?

The Diff-in-diff in Table 5

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

- How do you calculate 0.418?
- And 0.441?
- And Col. 3, 0.033?

- How do we find 0.024?
- And 5.5%?

The Diff-in-diff in Table 5

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

- How do you calculate 0.418?
- And 0.441?
- And Col. 3, 0.033?

- How do we find 0.024?
- And 5.5%? (0.024)/(0.418+0.009)

What Regression Equation Parallels the Diff-in-diff?

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

What Regression Equation Parallels the Diff-in-diff?

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

$$\text{fertility}_{i,j,t} = \beta_0 + \beta_1 \text{Quebec}_i * 1\{t = 1996\}_t + \beta_2 \text{Quebec}_i + \beta_3 1\{t = 1996\}_t + \beta_4 X_{i,j,t} + \epsilon_{i,j,t}$$

What Regression Equation Parallels the Diff-in-diff?

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)
A. All Parities					
Quebec	0.418	0.451	0.033		
	(0.003)	(0.004)	(0.005)		
n	20,285	16,453			
Rest of Canada	0.432	0.441	0.009	0.024	5.5%
	(0.002)	(0.002)	(0.003)	(0.006)	
n	54,115	46,032			

$$\text{fertility}_{i,j,t} = \beta_0 + \beta_1 \text{Quebec}_i * 1\{t = 1996\}_t + \beta_2 \text{Quebec}_i + \beta_3 1\{t = 1996\}_t + \beta_4 X_{i,j,t} + \epsilon_{i,j,t}$$

When estimated without covariates $X_{i,j,t}$, β_1 is **the same** as the estimate above.

What is the Underlying Assumption Here?

Can state the assumption many different ways

What is the Underlying Assumption Here?

Can state the assumption many different ways

- Only differences between Quebec and ROC are time-invariant
- Fertility in Quebec would have evolved like that in ROC absent the policy
- There are no pre-treatment trends in fertility in Quebec

What is the Underlying Assumption Here?

Can state the assumption many different ways

- Only differences between Quebec and ROC are time-invariant
- Fertility in Quebec would have evolved like that in ROC absent the policy
- There are no pre-treatment trends in fertility in Quebec

Can you test with Census data?

What is the Underlying Assumption Here?

Can state the assumption many different ways

- Only differences between Quebec and ROC are time-invariant
- Fertility in Quebec would have evolved like that in ROC absent the policy
- There are no pre-treatment trends in fertility in Quebec

Can you test with Census data? no – we only have one pre-treatment period

Table 6: Regression Version

Independent Variable	(a)	(b)
Pseudo R^2	0.0003	0.058
1996 dummy × Quebec	0.024* (0.005)	0.034* (0.006)
1996 dummy	0.009 (0.005)	0.013* (0.006)
Implied percentage increase in probability of having a child	5.6%	7.8%
Quebec	−0.014* (0.007)	−0.021* (0.007)
One older child	—	0.205* (0.016)
Two or more older children	—	−0.163* (0.011)
Female age 25–34	—	0.187* (0.009)
Female immigrant	—	0.032* (0.007)
Female Francophone	—	−0.047* (0.010)
Female Anglophone	—	−0.049* (0.012)
Female high school	—	−0.015* (0.006)
Female post–high school	—	−0.086* (0.004)
Female university degree	—	−0.192* (0.005)
Male age 25–34	—	—
Male age 35–44	—	—
Male age 45+	—	—
Male immigrant	—	—
Male Francophone	—	—
Male Anglophone	—	—
Male high school	—	—
Male post–high school	—	—
Male university degree	—	—
Married	—	—
Lives in urban area	—	—
Family income (C$10,000)	—	—
Provincial GDP growth	—	—
Provincial migration rate	—	—
Provincial education spending	—	—

- interpret coefficient for 1996 dummy
- interpret coefficient for 1996 dummy X Quebec

And a Triple Difference!

Region	Mean 1991 (1)	Mean 1996 (2)	Trend Difference in Means, (2) − (1) = (3)	Difference in Differences (4)	Percentage Increase (5)	Triple Difference (6)
A. All Parities						
Quebec	0.418	0.451	0.033			
	(0.003)	(0.004)	(0.005)			
n	20,285	16,453				
Rest of Canada	0.432	0.441	0.009	0.024	5.5%	
	(0.002)	(0.002)	(0.003)	(0.006)		
n	54,115	46,032				
B. *Zero older children*						
Quebec	0.393	0.418	0.025			
	(0.004)	(0.004)	(0.006)			
n	15,017	12,399				
Rest of Canada	0.398	0.407	0.009	0.016	4.0%	
	(0.002)	(0.003)	(0.003)	(0.007)		
n	38,754	33,338				
C. *One older child*						
Quebec	0.627	0.677	0.050			
	(0.009)	(0.009)	(0.013)			
n	3,207	2,475				
Rest of Canada	0.691	0.681	−0.010	0.060	9.7%	
	(0.005)	(0.006)	(0.008)	(0.015)		
n	8,262	7,088				
D. *Two or more older children*						
Quebec	0.278	0.353	0.075			
	(0.010)	(0.012)	(0.015)			
n	2,061	1,579				
Rest of Canada	0.321	0.344	0.023	0.052	17.2%	0.036
	(0.006)	(0.006)	(0.008)	(0.018)		(0.020)
n	7,099	5,606				

Next Lecture

- Read
  - Janssen and Zhang, selected pages only through Section 4
  - Abadie et al, through Section 3
- Summary due next week if you're on the list