FLS 6441 - Methods III: Explanation and Causation

Week 8 - Difference-in-Differences

Jonathan Phillips

May 2019

Classification of Research Designs

		Independence of Treatment Assignment	**Researcher Controls Treatment Assignment?**
Controlled Experiments	Field Experiments	✓	✓
	Survey and Lab Experiments	✓	✓
Natural Experiments	Natural Experiments	✓	
	Instrumental Variables	✓	
	Discontinuities	✓	
Observational Studies	Difference-in-Differences		
	Controlling for Confounding		
	Matching		
	Comparative Cases and Process Tracing		

Section 1

Difference-in-Differences

Difference-in-Differences

- What if we have *NO* variation in treatment that is independent of potential outcomes?

Difference-in-Differences

- What if we have *NO* variation in treatment that is independent of potential outcomes?
- Then we have an *Observational* study

Difference-in-Differences

- Two types of observational studies:

Difference-in-Differences

- Two types of observational studies:

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**

Difference-in-Differences

- Two types of observational studies:

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - BUT Omitted variable bias

Difference-in-Differences

- Two types of observational studies:

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - BUT Omitted variable bias
2. **Time-series:** Compare outcomes of units **before** and **after** treatment

Difference-in-Differences

- Two types of observational studies:

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - BUT Omitted variable bias
2. **Time-series:** Compare outcomes of units **before** and **after** treatment
   - BUT Outcomes might change over time for reasons other than treatment ('Overall Trend Bias')

Difference-in-Differences

- But each approach also has advantages

Difference-in-Differences

- But each approach also has advantages

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**

Difference-in-Differences

- But each approach also has advantages

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - Allows us to compare units at the same point in time, removing 'Overall Trend Bias'

Difference-in-Differences

- But each approach also has advantages

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - Allows us to compare units at the same point in time, removing 'Overall Trend Bias'
2. **Time-series:** Compare outcomes of units **before** and **after** treatment

Difference-in-Differences

- But each approach also has advantages

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - Allows us to compare units at the same point in time, removing 'Overall Trend Bias'
2. **Time-series:** Compare outcomes of units **before** and **after** treatment
   - Allows us to keep the fixed characteristics of the same unit, removing Omitted Variable Bias

Difference-in-Differences

- But each approach also has advantages

1. **Cross-sectional:** Compare outcomes across different units, **treated** and **control**
   - Allows us to compare units at the same point in time, removing 'Overall Trend Bias'
2. **Time-series:** Compare outcomes of units **before** and **after** treatment
   - Allows us to keep the fixed characteristics of the same unit, removing Omitted Variable Bias
   - Even *unobserved* fixed characteristics

Difference-in-Differences

- What if we combine both approaches?

Difference-in-Differences

- What if we combine both approaches?
- Comparing **across units** and **across time**

Difference-in-Differences

- What if we combine both approaches?
- Comparing **across units** and **across time**
- Comparing **changes** instead of **levels**

Difference-in-Differences

- What if we combine both approaches?
- Comparing **across units** and **across time**
- Comparing **changes** instead of **levels**
- Removing the risks from both overall trends and omitted variables

Difference-in-Differences

	Balances time-invariant 'fixed' unit characteristics		**Balances time-varying characteristics**	
	Observed	Unobserved	Overall trends	Unit-specific trends
Field Experiments	✓	✓	✓	✓
Survey and Lab Experiments	✓	✓	✓	✓
Natural Experiments	✓	✓	✓	✓
Instrumental Variables	✓	✓	✓	✓
Regression Discontinuity	✓	✓	✓	✓
Cross-sectional comparisons	X	X	✓	X
Before-After comparisons	✓	✓	X	X
Difference-in-Differences	✓	✓	✓	X

Difference-in-Differences

- Example: How has the Brexit vote affected the UK's growth rate?

Difference-in-Differences

- Example: How has the Brexit vote affected the UK's growth rate?
  - Comparing with European growth rates is biased - UK growth is influenced by oil, different labour laws etc.

Difference-in-Differences

- Example: How has the Brexit vote affected the UK's growth rate?
  - Comparing with European growth rates is biased - UK growth is influenced by oil, different labour laws etc.
  - Comparing before and after the Brexit vote is biased - the world economy improved around the same time as Brexit (coincidentally)

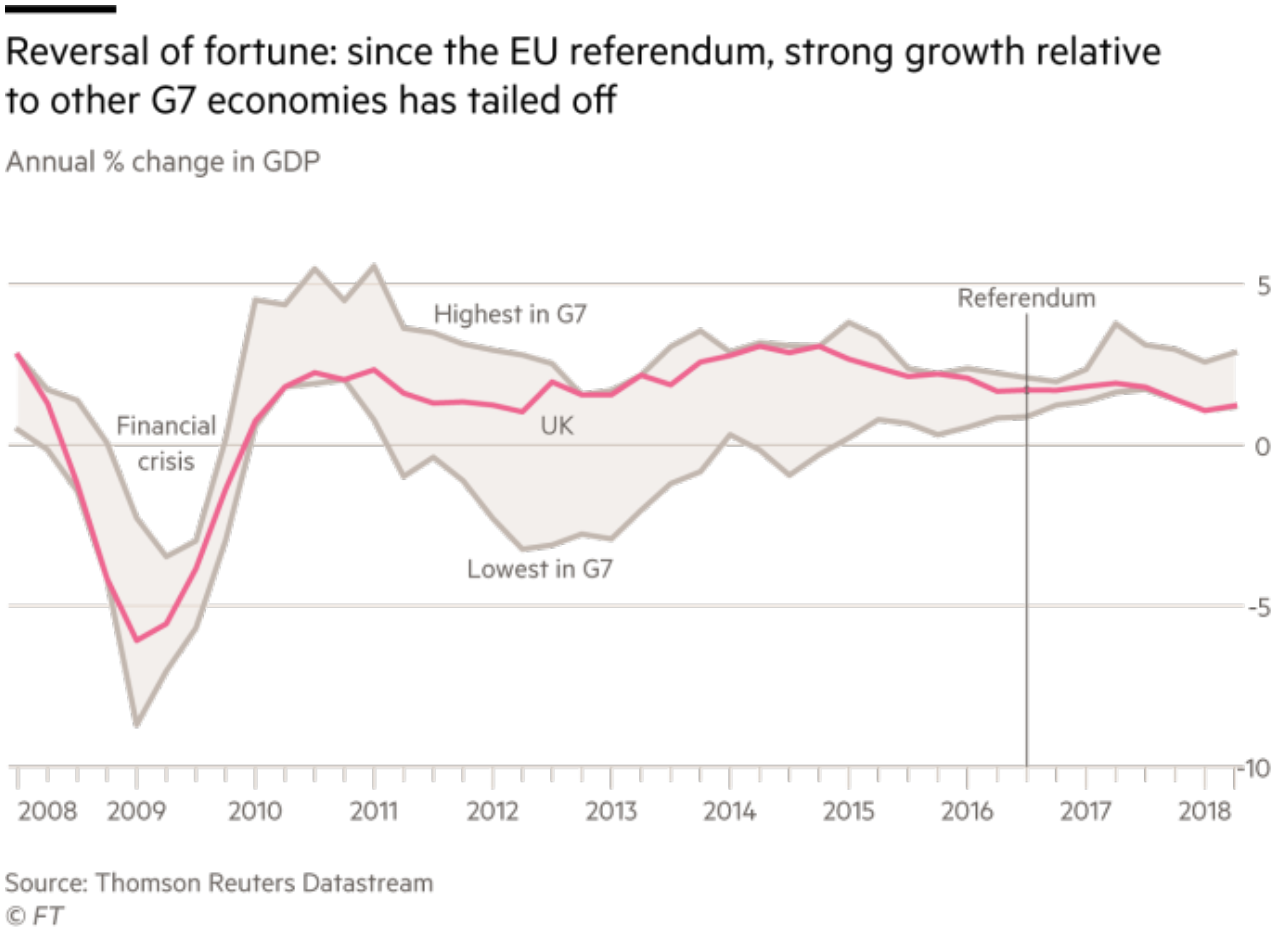
Difference-in-Differences

- Example: How has the Brexit vote affected the UK's growth rate?
  - Comparing with European growth rates is biased - UK growth is influenced by oil, different labour laws etc.
  - Comparing before and after the Brexit vote is biased - the world economy improved around the same time as Brexit (coincidentally)
  - But compare how European growth **changed** (+0.3%) and UK growth **changed** (-0.4%)

Difference-in-Differences

- Example: How has the Brexit vote affected the UK's growth rate?
  - Comparing with European growth rates is biased - UK growth is influenced by oil, different labour laws etc.
  - Comparing before and after the Brexit vote is biased - the world economy improved around the same time as Brexit (coincidentally)
  - But compare how European growth **changed** (+0.3%) and UK growth **changed** (-0.4%)
  - The net effect of Brexit is -0.7%

Difference-in-Differences



Difference-in-Differences

- But can we really say this was the effect of Brexit?

Difference-in-Differences

- But can we really say this was the effect of Brexit?

1. Maybe the UK was on a different **unit-specific trend** to the EU before Brexit?

Difference-in-Differences

- But can we really say this was the effect of Brexit?

1. Maybe the UK was on a different **unit-specific trend** to the EU before Brexit?
   - Diff-in-Diff does NOT control for **time-varying confounders**

Difference-in-Differences

- But can we really say this was the effect of Brexit?

1. Maybe the UK was on a different **unit-specific trend** to the EU before Brexit?
   - Diff-in-Diff does NOT control for **time-varying confounders**
   - We have to check for **Parallel pre-treatment trends**

Difference-in-Differences

- But can we really say this was the effect of Brexit?

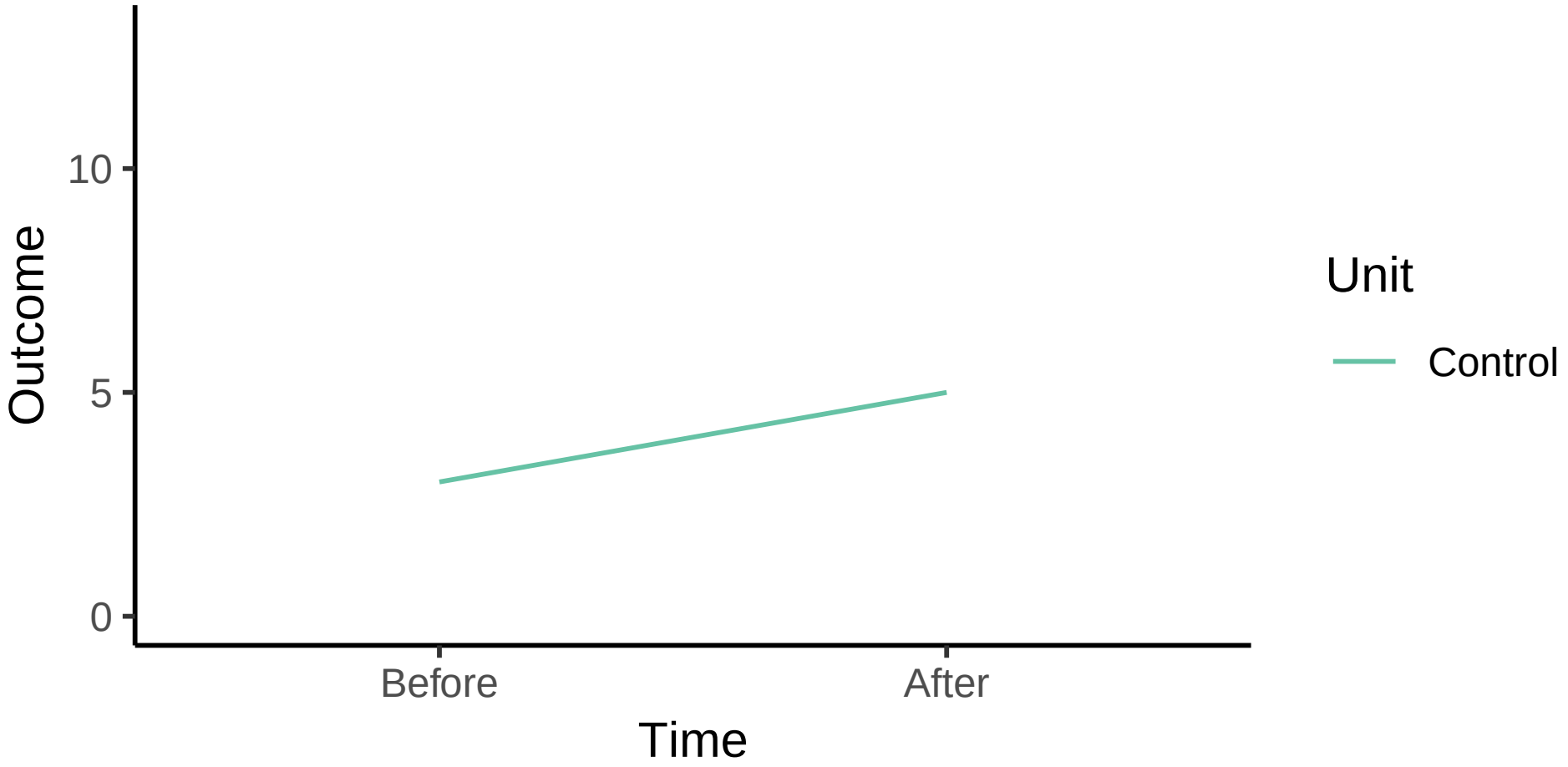
1. Maybe the UK was on a different **unit-specific trend** to the EU before Brexit?
   - Diff-in-Diff does NOT control for **time-varying confounders**
   - We have to check for **Parallel pre-treatment trends**
2. Maybe the UK passed other policies at the same time as Brexit?

Difference-in-Differences

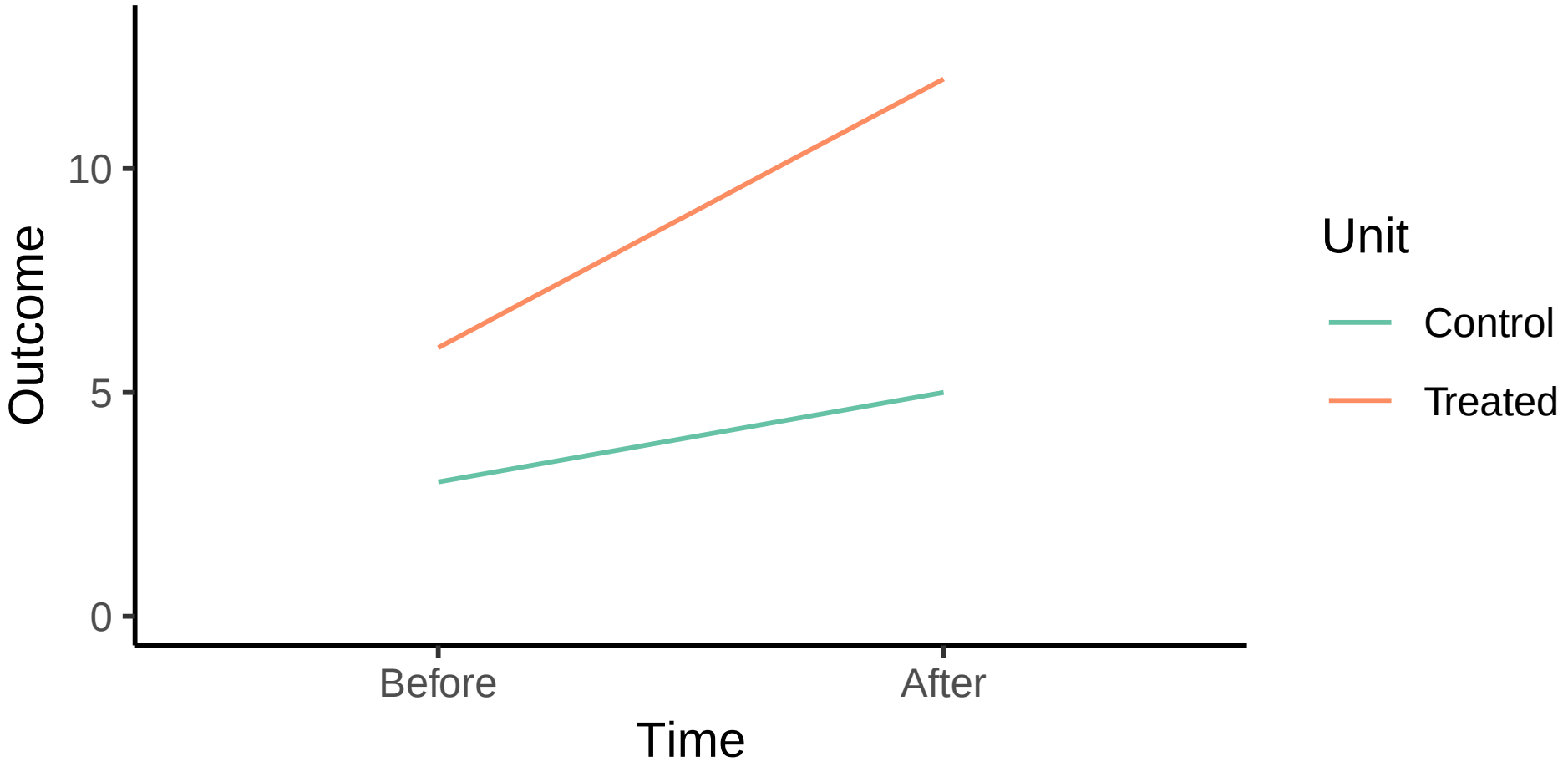
- But can we really say this was the effect of Brexit?

1. Maybe the UK was on a different **unit-specific trend** to the EU before Brexit?
   - Diff-in-Diff does NOT control for **time-varying confounders**
   - We have to check for **Parallel pre-treatment trends**
2. Maybe the UK passed other policies at the same time as Brexit?
   - We have to check there are no **compound treatments**

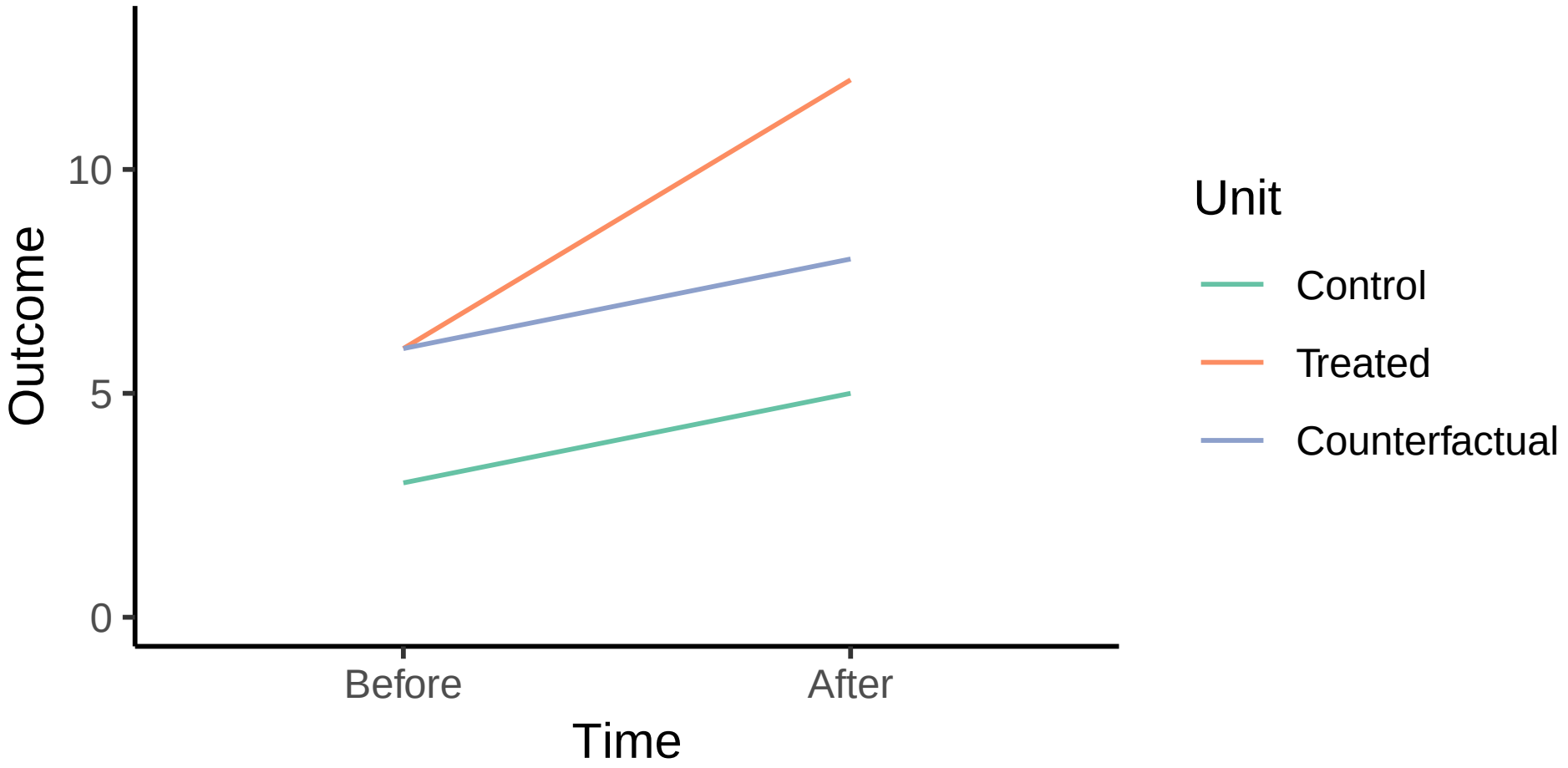
Difference-in-Differences



Difference-in-Differences



Difference-in-Differences



Estimating Difference-in-Differences

- Regression for the cross-unit effect of treatment

$$Y_{it} = \alpha + \gamma D_i$$

Estimating Difference-in-Differences

- Regression for the cross-unit effect of treatment

$$Y_{it} = \alpha + \gamma D_i$$

- Regression for the before-after treatment comparison

$$Y_{it} = \alpha + \delta T_i$$

Estimating Difference-in-Differences

- Regression for the cross-unit effect of treatment

$$Y_{it} = \alpha + \gamma D_i$$

- Regression for the before-after treatment comparison

$$Y_{it} = \alpha + \delta T_i$$

- The difference-in-differences estimate is just the *interaction* of time and treatment status

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

Estimating Difference-in-Differences

- Regression for the cross-unit effect of treatment

$$Y_{it} = \alpha + \gamma D_i$$

- Regression for the before-after treatment comparison

$$Y_{it} = \alpha + \delta T_i$$

- The difference-in-differences estimate is just the *interaction* of time and treatment status

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

- β is our **Average Treatment Effect** estimate

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$D = 0, T = 0 : E(Y) =$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) =$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) =$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$D = 0, T = 0 : E(Y) = \alpha$

$D = 0, T = 1 : E(Y) = \alpha + \delta$
$D = 1, T = 0 : E(Y) = \alpha + \gamma$
$D = 1, T = 1 : E(Y) =$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

$$\Delta(Y|D = 1) = E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0) =$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

$$\Delta(Y|D = 1) = E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0) = \delta + \beta$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

$$\Delta(Y|D = 1) = E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0) = \delta + \beta$$
$$\Delta(Y|D = 0) = E(Y|D = 0, T = 1) - E(Y|D = 0, T = 0) =$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

$$\Delta(Y|D = 1) = E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0) = \delta + \beta$$
$$\Delta(Y|D = 0) = E(Y|D = 0, T = 1) - E(Y|D = 0, T = 0) = \delta$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

$$\Delta(Y|D = 1) = E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0) = \delta + \beta$$
$$\Delta(Y|D = 0) = E(Y|D = 0, T = 1) - E(Y|D = 0, T = 0) = \delta$$

$$\Delta(Y|D = 1) - \Delta(Y|D = 0) =$$

Estimating Difference-in-Differences

$$Y_{it} = \alpha + \gamma D_i + \delta T_t + \beta D_i * T_t$$

$$D = 0, T = 0 : E(Y) = \alpha$$

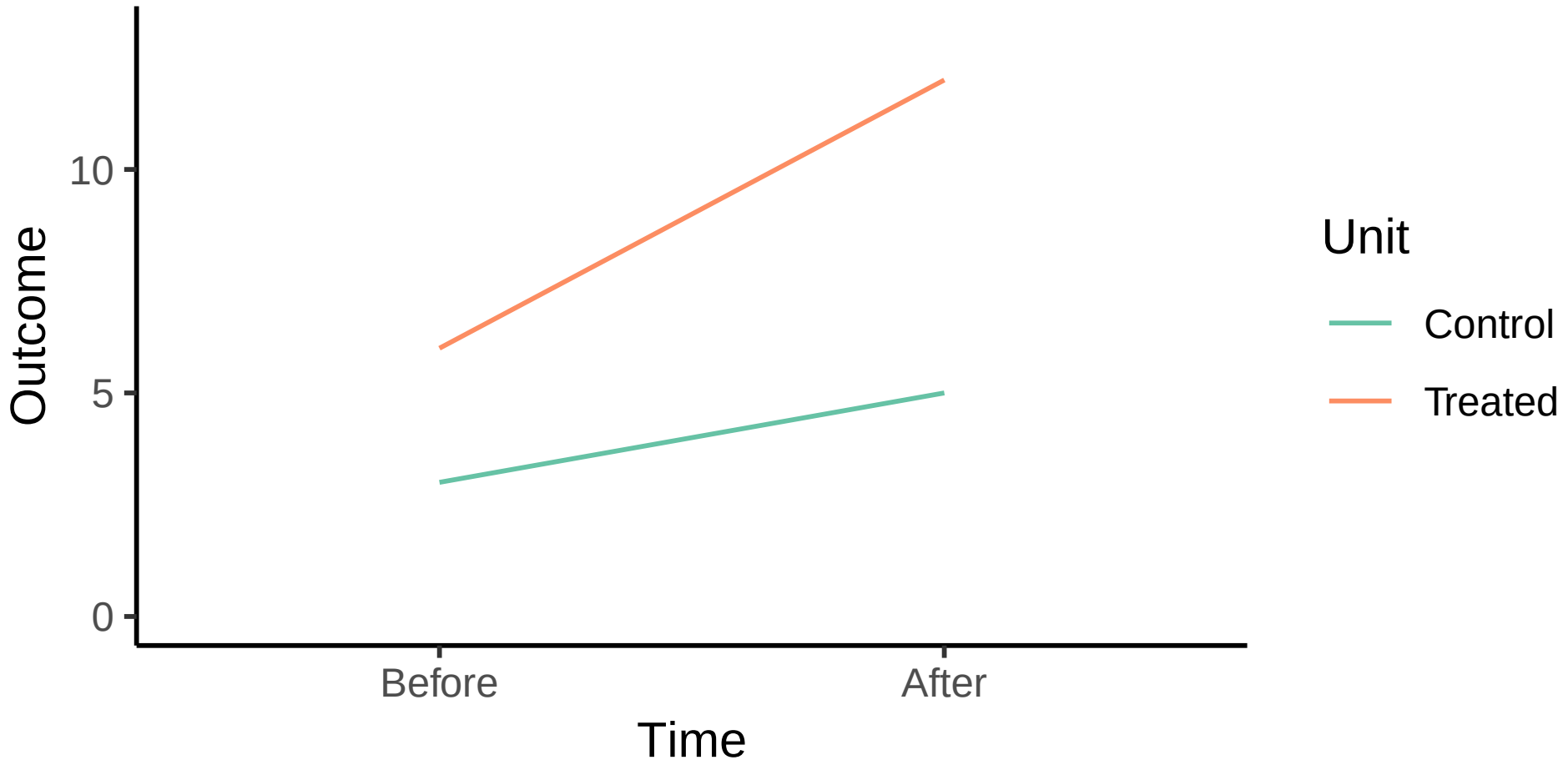
$$D = 0, T = 1 : E(Y) = \alpha + \delta$$
$$D = 1, T = 0 : E(Y) = \alpha + \gamma$$
$$D = 1, T = 1 : E(Y) = \alpha + \delta + \gamma + \beta$$

$$\Delta(Y|D = 1) = E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0) = \delta + \beta$$
$$\Delta(Y|D = 0) = E(Y|D = 0, T = 1) - E(Y|D = 0, T = 0) = \delta$$

$$\Delta(Y|D = 1) - \Delta(Y|D = 0) = \beta$$

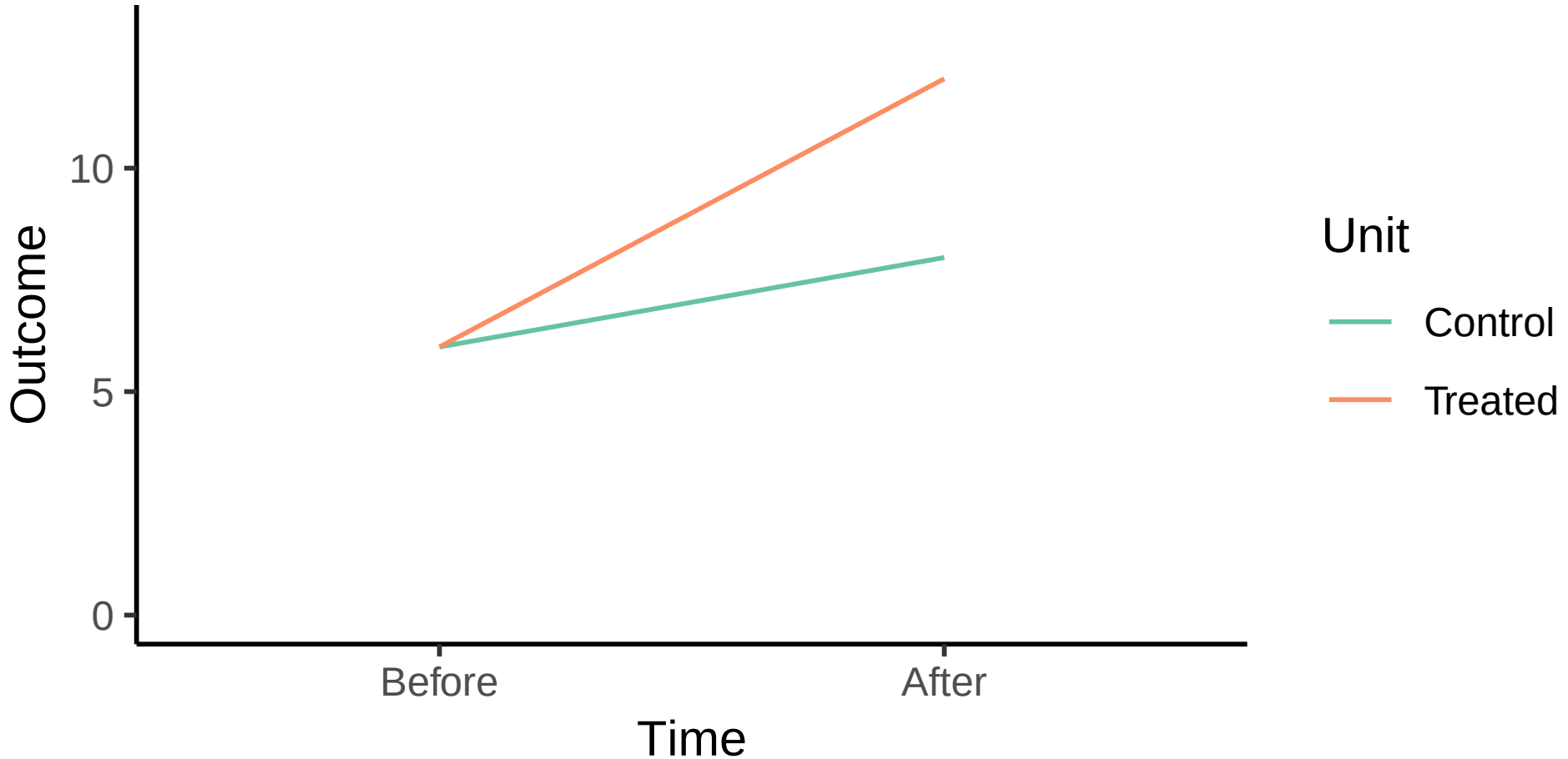
Difference-in-Differences

Raw Data:



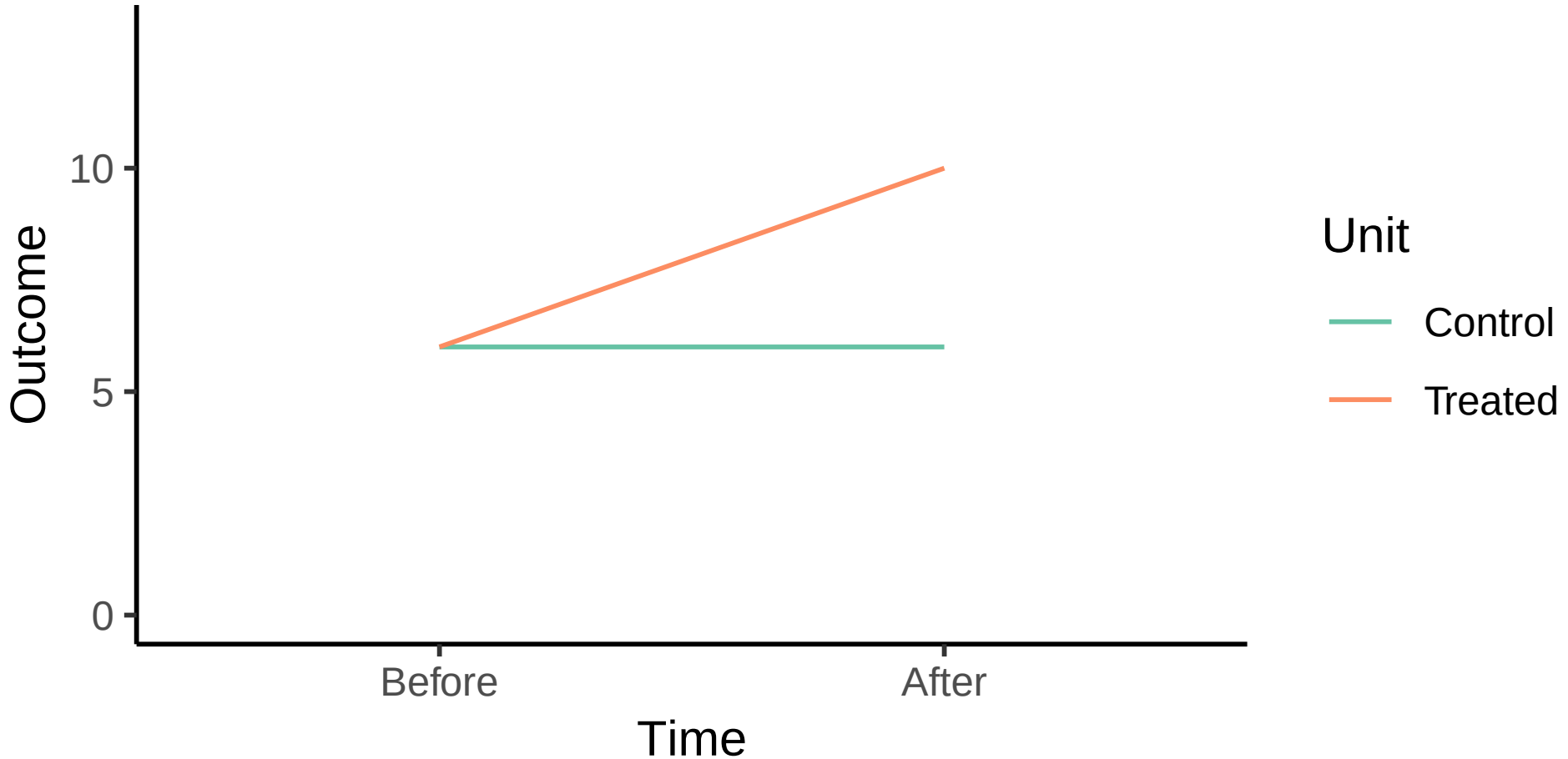
Difference-in-Differences

Add a variable (fixed effect) for treated/control:



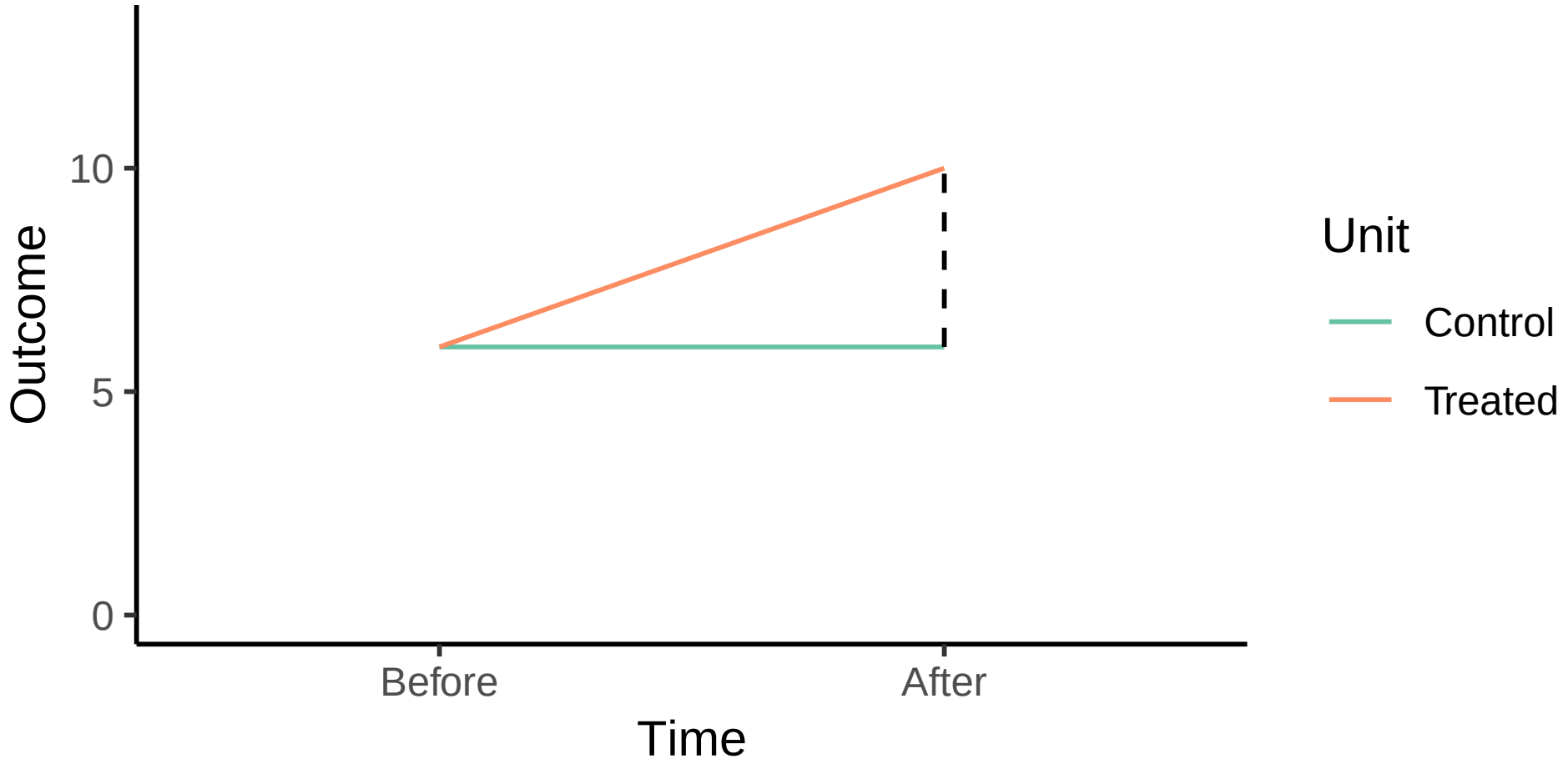
Difference-in-Differences

Add a variable (fixed effect) for time:



Difference-in-Differences

Add a variable (fixed effect) for time:



Estimating Difference-in-Differences

- With time-series data, we have temporal autocorrelation

Estimating Difference-in-Differences

- With time-series data, we have temporal autocorrelation
- Crucial to cluster standard errors by each cross-sectional unit (eg. each country)

Difference-in-Differences

- How do we know if there are **time-varying confounders**?

Difference-in-Differences

- How do we know if there are **time-varying confounders**?
- Selection into treatment is usually not just due to 'fixed' variables (eg. gender) but due to 'time-varying' variables (eg. income, employment etc.)

Difference-in-Differences

- How do we know if there are **time-varying confounders**?
- Selection into treatment is usually not just due to 'fixed' variables (eg. gender) but due to 'time-varying' variables (eg. income, employment etc.)
- Eg. training program participants' income has usually fallen a lot in the past few months

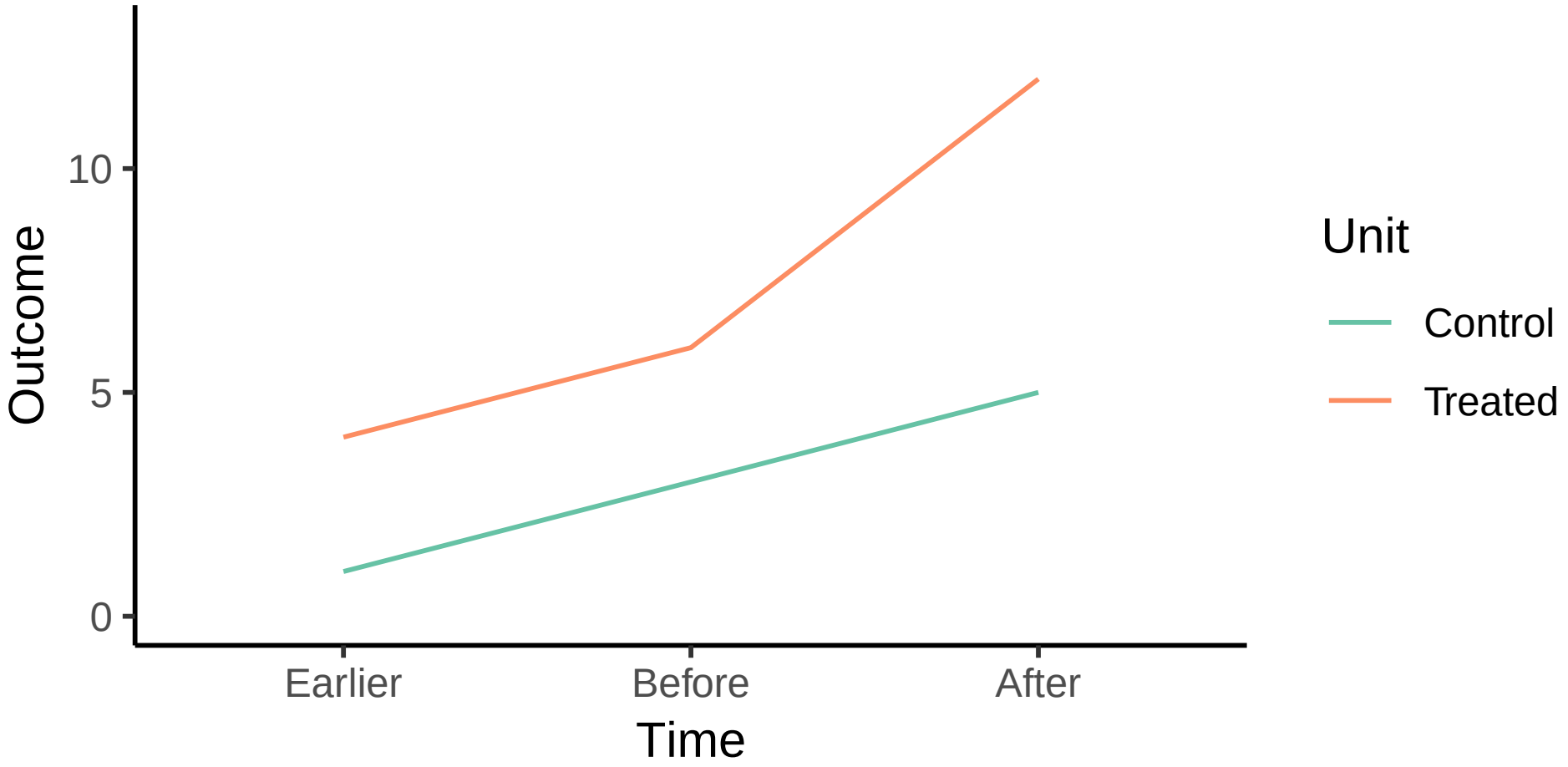
Difference-in-Differences

- How do we know if there are **time-varying confounders**?
- Selection into treatment is usually not just due to 'fixed' variables (eg. gender) but due to 'time-varying' variables (eg. income, employment etc.)
- Eg. training program participants' income has usually fallen a lot in the past few months
- We want the outcome for the treated group to have the same trend as the control group

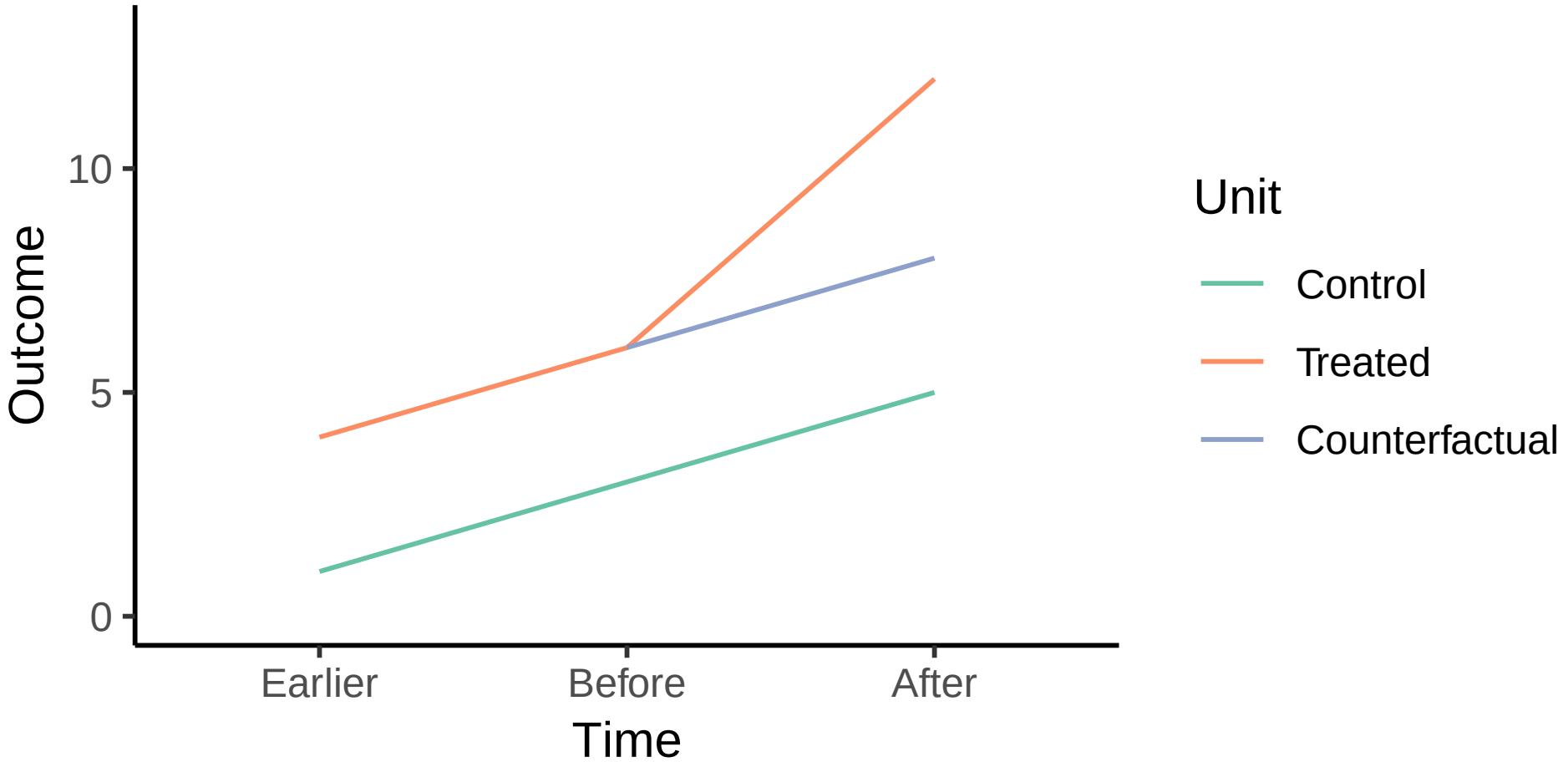
Difference-in-Differences

- How do we know if there are **time-varying confounders**?
- Selection into treatment is usually not just due to 'fixed' variables (eg. gender) but due to 'time-varying' variables (eg. income, employment etc.)
- Eg. training program participants' income has usually fallen a lot in the past few months
- We want the outcome for the treated group to have the same trend as the control group
- One test of this is to check if **pre-treatment trends are parallel**

Difference-in-Differences



Difference-in-Differences



Difference-in-Differences Assumptions

1. **Parallel pre-treatment trends** between treated and control units

Difference-in-Differences Assumptions

1. **Parallel pre-treatment trends** between treated and control units
2. **No compound treatment**

Difference-in-Differences Assumptions

1. **Parallel pre-treatment trends** between treated and control units
2. **No compound treatment**
3. **No spillovers** (SUTVA)

Difference-in-Differences Assumptions

1. **Parallel pre-treatment trends** between treated and control units
2. **No compound treatment**
3. **No spillovers** (SUTVA)
4. **Group membership is stable** (no migration from control to treatment)

Section 2

The Effect of Illegal Activities on Violence

Chimeli and Soares 2017

- How does an activity being illegal affect violence?

Chimeli and Soares 2017

- How does an activity being illegal affect violence?
- How did Brazil's ban on mahogany affect homicides?

Chimeli and Soares 2017

- How does an activity being illegal affect violence?
- How did Brazil's ban on mahogany affect homicides?
- What are the challenges to explanation?

Chimeli and Soares 2017

- How does an activity being illegal affect violence?
- How did Brazil's ban on mahogany affect homicides?
- What are the challenges to explanation?
  - Omitted variables, eg. State capacity

Chimeli and Soares 2017

- How does an activity being illegal affect violence?
- How did Brazil's ban on mahogany affect homicides?
- What are the challenges to explanation?
  - Omitted variables, eg. State capacity
  - Overall trends, eg. national decrease in homicides

Chimeli and Soares 2017

- How does an activity being illegal affect violence?
- How did Brazil's ban on mahogany affect homicides?
- What are the challenges to explanation?
  - Omitted variables, eg. State capacity
  - Overall trends, eg. national decrease in homicides
- Comparing the *change* in violence in mahogany-growing areas to the change in violence in non-mahogany areas

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:**

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:**

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany
- **Before:**

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany
- **Before:** Pre-1999

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany
- **Before:** Pre-1999
- **After:**

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany
- **Before:** Pre-1999
- **After:** Post-1999

Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany
- **Before:** Pre-1999
- **After:** Post-1999
- **Outcome:**

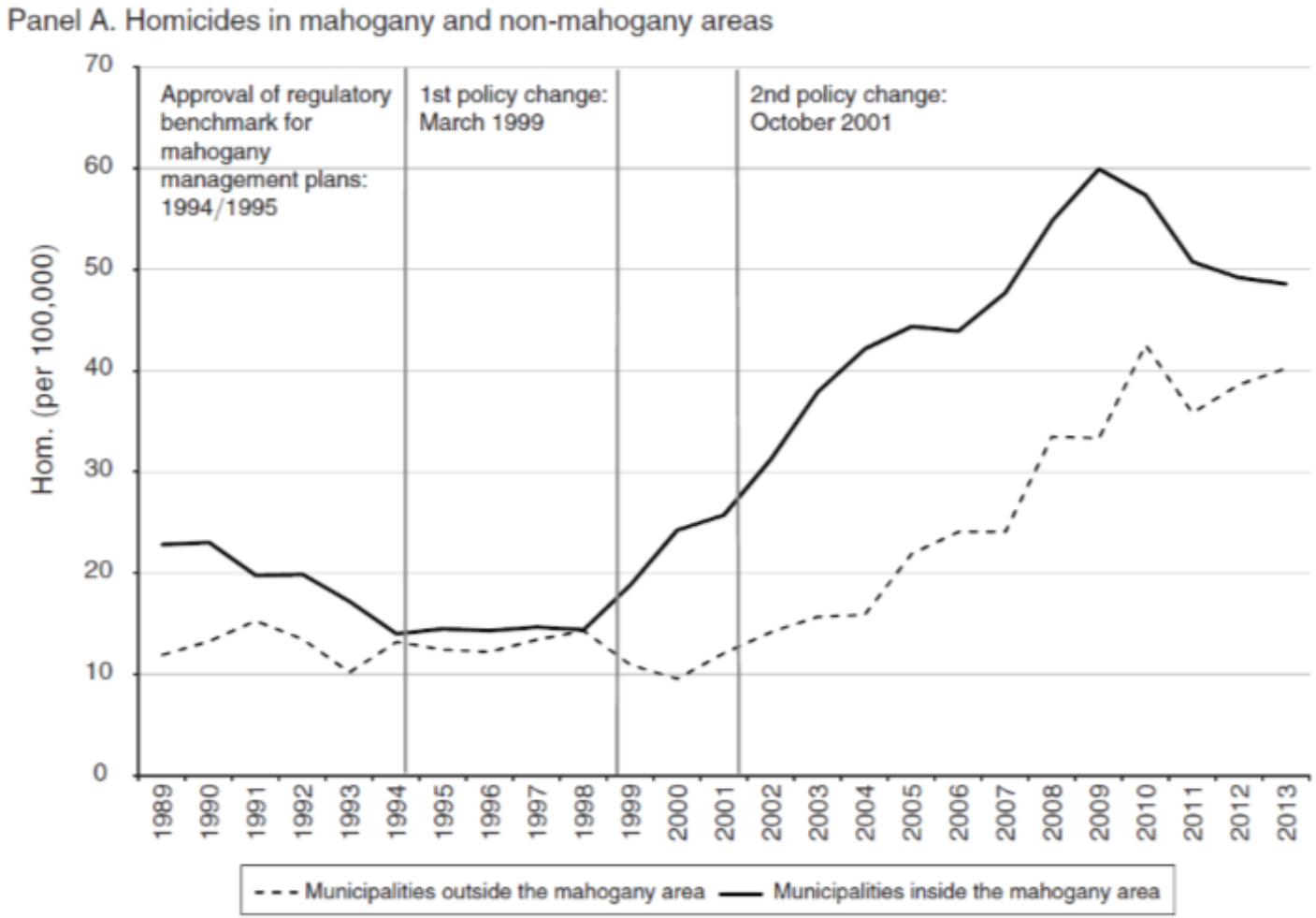
Chimeli and Soares 2017

- In the 'After' period we need treated **and** control units
- But the ban on mahogany applied to **all** of Brazil.
- So what are treatment and control here?
- **Treatment:** Municipalities with mahogany
- **Control:** Municipalities **without** mahogany
- **Before:** Pre-1999
- **After:** Post-1999
- **Outcome:** Homicides per 100,000 people

Chimeli and Soares 2017

- Multiple treatment timings:
  - Partial Ban on Mahogany exports
  - Full Ban on Mahogany exports
  - 'Reverse' treatment: Better policing of mahogany regulations

Difference-in-Differences



- Methodology:

- Methodology:

$$Homicides_{it} = \beta_1 Post-1999_t + \beta_2 Mahogany_i$$

$$+ \beta_3 (Post-1999_t * Mahogany_i) + \epsilon_i$$

- Methodology:

$$Homicides_{it} = \beta_1 Post-1999_t + \beta_2 Mahogany_i$$

$$+ \beta_3(Post-1999_t * Mahogany_i) + \epsilon_i$$

Cluster standard errors by municipality

- Methodology:

$$Homicides_{it} = \beta_1 Post-1999_t + \beta_2 Mahogany_i$$

$+ \beta_3(Post-1999_t * Mahogany_i) + \epsilon_i$
Cluster standard errors by municipality

Apply more complex state-specific trends for covariates to minimize risk of non-parallel trends

- Methodology:

$$Homicides_{it} = \beta_1 Post - 1999_t + \beta_2 Mahogany_i$$

$$+ \beta_3 (Post - 1999_t * Mahogany_i) + \epsilon_i$$

Cluster standard errors by municipality

Apply more complex state-specific trends for covariates to minimize risk of non-parallel trends

Supporting evidence: The 'extra' homicides were the type we'd expect from illegal activity

Chimeli and Soares 2017

- Interpretation

Chimeli and Soares 2017

- Interpretation
  - Illegal activity prevents ’peaceful’ contract enforcement
  - Competition between loggers
  - Contract enforcement with buyers
  - Intimidation of communities to not report logging