Difference in Differences

Causal Inference Discussion Section

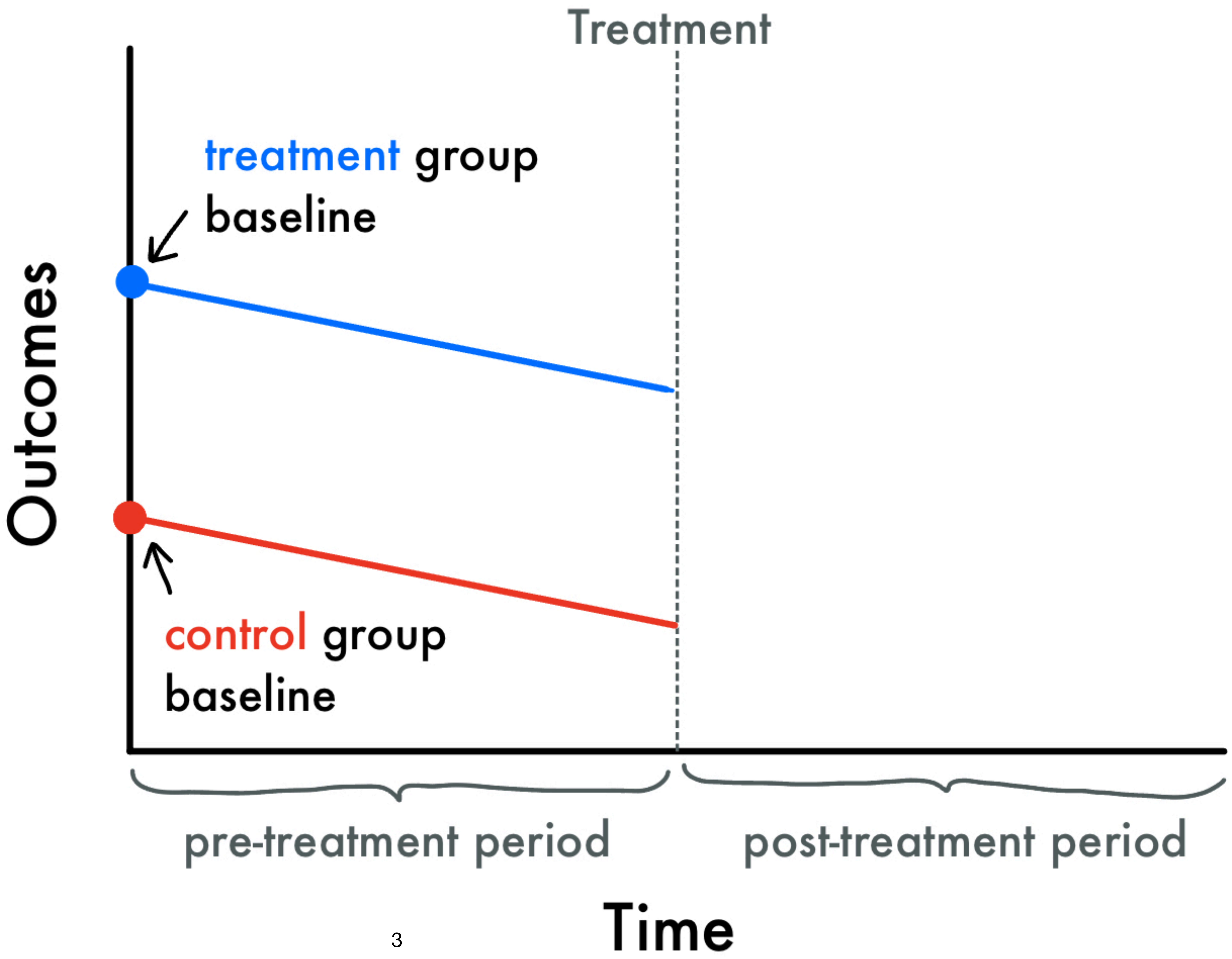
Wednesday, November 5th

Reminders & Announcements

- PS5 due Tuesday, Nov 11
- Project Groups and Part II instructions will be posted soon

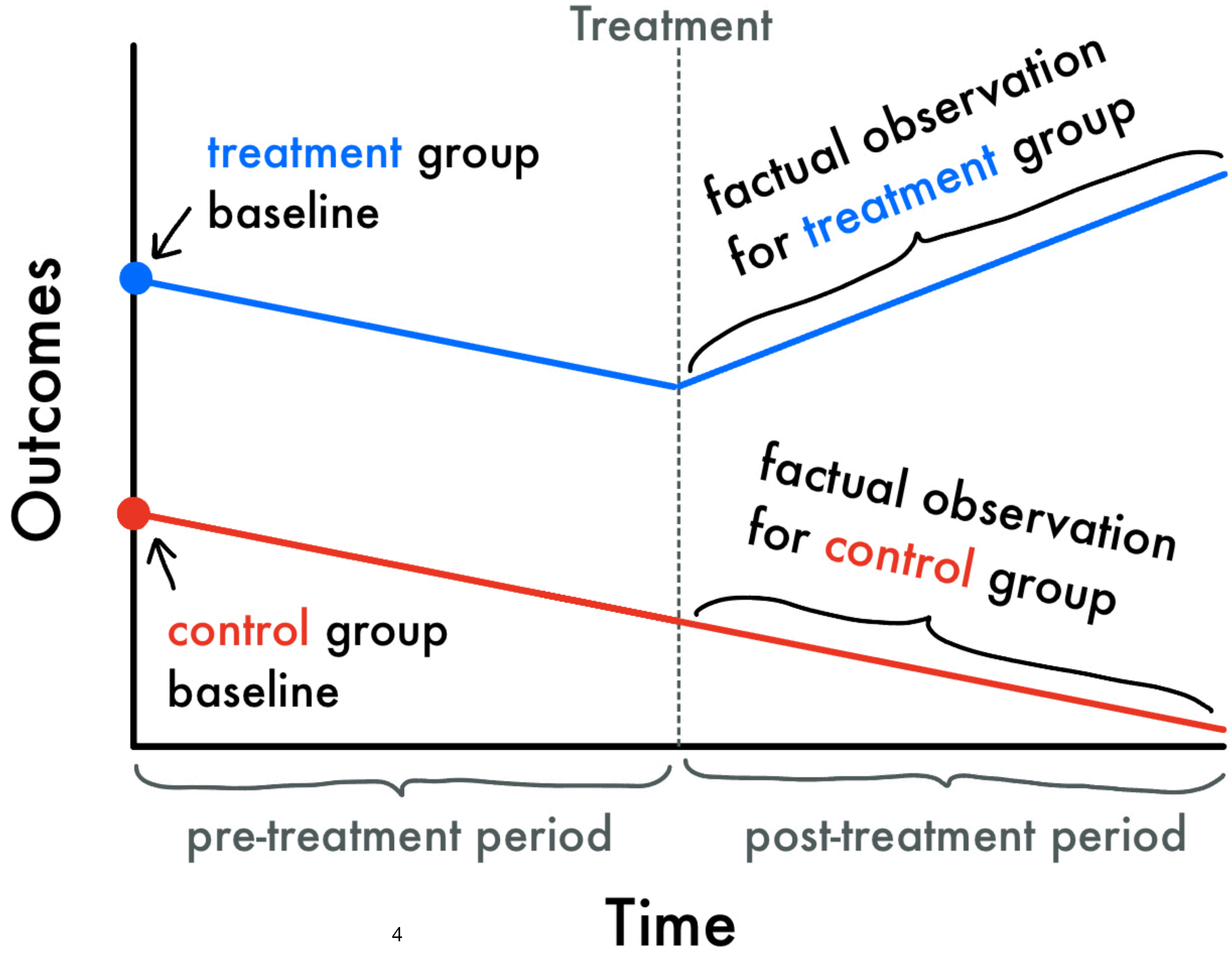
Difference in Differences Review

Visual illustration



Difference in Differences Review

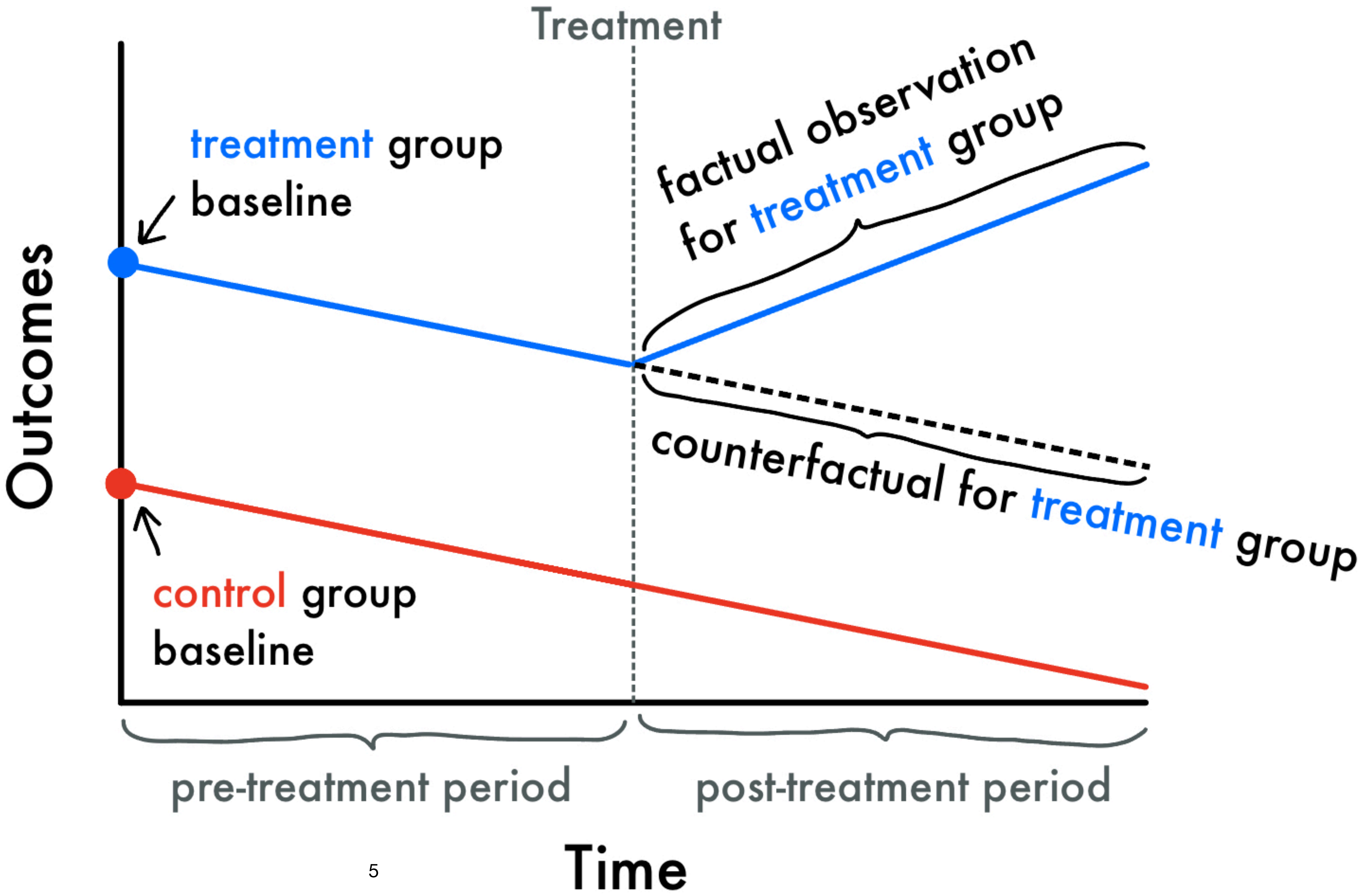
Visual illustration



Difference in Differences Review

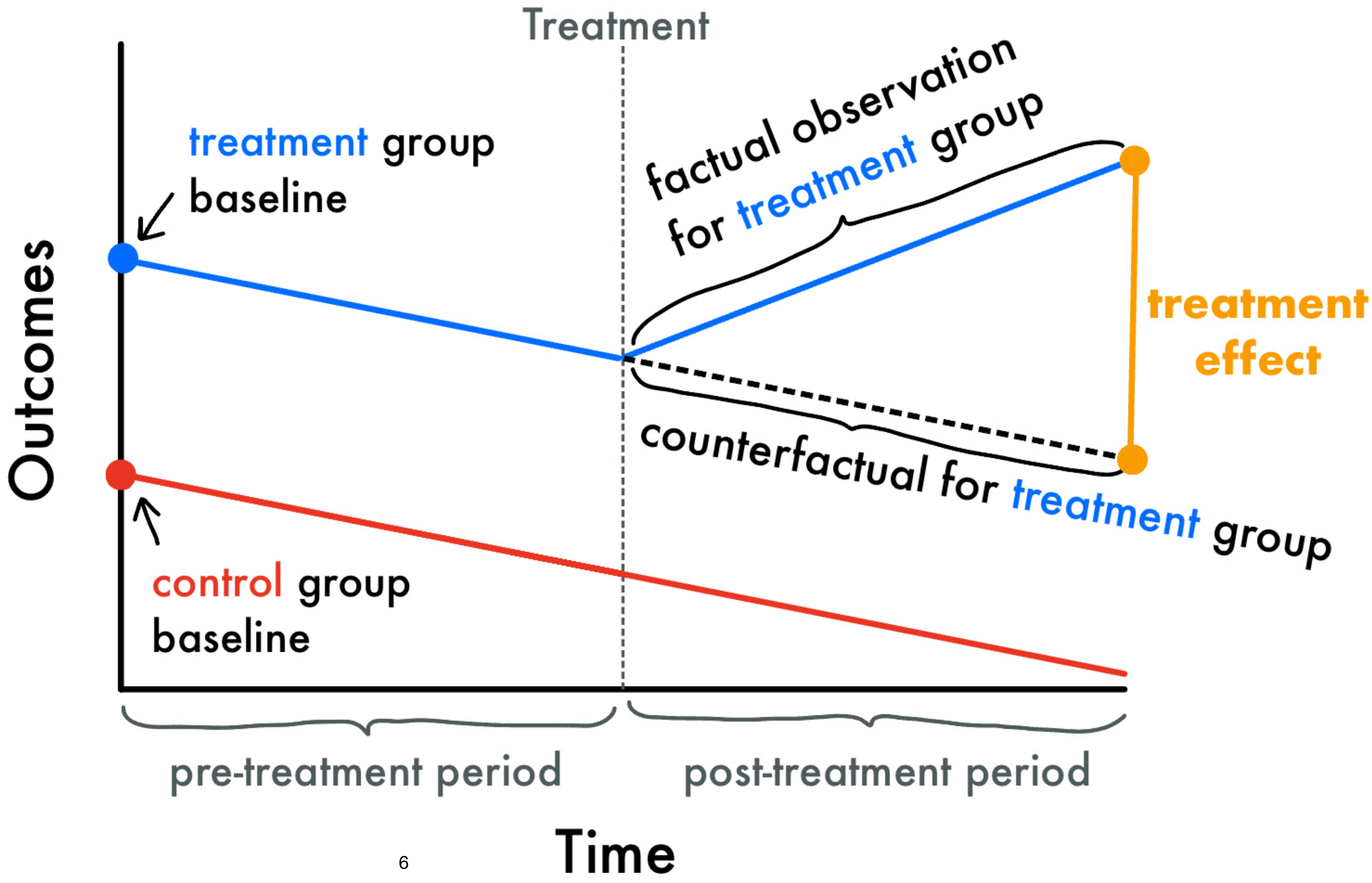
Visual illustration

Parallel trends assumption!!!



Difference in Differences Review

Visual illustration



Difference in Differences Review

Intuitive Idea

Group	Time Period	Outcome	Difference 1	Difference 2
Treatment	Pre-treatment			
	Post-treatment			
Control	Pre-treatment			
	Post-treatment			

Difference in Differences Review

Intuitive Idea

Different baseline outcomes

Group	Time Period	Outcome	Difference 1	Difference 2
Treatment	Pre-treatment	$Y = B_1$		
	Post-treatment			
Control	Pre-treatment	$Y = B_0$		
	Post-treatment			

Difference in Differences Review

Intuitive Idea

Time effect T

Treatment effect D

Group	Time Period	Outcome	Difference 1	Difference 2
Treatment	Pre-treatment	$Y = B_1$		
	Post-treatment	$Y = B_1 + T + D$		
Control	Pre-treatment	$Y = B_0$		
	Post-treatment	$Y = B_0 + T$		

Difference in Differences Review

Intuitive Idea

Parallel trends assumption!!!

(the same T)

Group	Time Period	Outcome	Difference 1	Difference 2
Treatment	Pre-treatment	$Y = B_1$	$T + D$	
	Post-treatment	$Y = B_1 + T + D$		
Control	Pre-treatment	$Y = B_0$	T	
	Post-treatment	$Y = B_0 + T$		

Difference in Differences Review

Intuitive Idea

<table>
<tr><th>Group</th><th>Time Period</th><th>Outcome</th><th>Difference 1</th><th>Difference 2</th></tr>
<tr><td rowspan="2">Treatment</td><td>Pre-treatment</td><td>$Y = B_1$</td><td rowspan="2">$T + D$</td><td rowspan="4">D</td></tr>
<tr><td>Post-treatment</td><td>$Y = B_1 + T + D$</td></tr>
<tr><td rowspan="2">Control</td><td>Pre-treatment</td><td>$Y = B_0$</td><td rowspan="2">T</td></tr>
<tr><td>Post-treatment</td><td>$Y = B_0 + T$</td></tr>
</table>

Difference in Differences Review

Using Regression

Consider the following linear model for outcomes:

$$Y_{i,t} = \alpha + \gamma\text{Treated} + \lambda\text{Time} + \delta(\text{Treated} \times \text{Time}) + \varepsilon_{i,t}$$

- Treated is a binary variable (1 if in treatment group, 0 if in control group)
- Time is a binary variable indicating if this is the post-treatment period (1) or the pre-treatment period (0)
- Treated $\times$ Time is an interaction term

Difference in Differences Review

Using Regression

Consider the following linear model for outcomes:

$$Y_{i,t} = \alpha + \gamma \text{Treated} + \lambda \text{Time} + \delta(\text{Treated} \times \text{Time}) + \varepsilon_{i,t}$$

- Control pre-treatment: α
- Control post-treatment: $\alpha + \lambda$
- Treated pre-treatment: $\alpha + \gamma$
- Treated post-treatment: $\alpha + \gamma + \lambda + \delta$

Difference in Differences Review

Using Regression

Consider the following linear model for outcomes:

$$Y_{i,t} = \alpha + \gamma \text{Treated} + \lambda \text{Time} + \delta(\text{Treated} \times \text{Time}) + \varepsilon_{i,t}$$

- Control pre-treatment: α
- Control post-treatment: $\alpha + \lambda$
- Treated pre-treatment: $\alpha + \gamma$
- Treated post-treatment: $\alpha + \gamma + \lambda + \delta$

$\boxed{\lambda}$

$\boxed{\lambda + \delta}$

1st difference(s)

Difference in Differences Review

Using Regression

Consider the following linear model for outcomes:

$$Y_{i,t} = \alpha + \gamma \text{Treated} + \lambda \text{Time} + \boxed{\delta(\text{Treated} \times \text{Time})} + \varepsilon_{i,t}$$

- Control pre-treatment: α
- Control post-treatment: $\alpha + \lambda$
- Treated pre-treatment: $\alpha + \gamma$
- Treated post-treatment: $\alpha + \gamma + \lambda + \delta$

$\boxed{\lambda}$

$\boxed{\lambda + \delta}$

1st difference(s)

$\boxed{\delta}$

2nd difference

Application

A Study of Decentralization on Public Services in Vietnam

- Looking at the effects of decentralizing government (`treatment`) on public services such as educational programs (`pro4`)
- Other variables in the data:
    - `year`: the year the data record is from (we'll focus on two periods, 2008 and 2010, since treatment was introduced in 2009)
    - `post_treat`: a binary variable indicated if the data record is from the pre-treatment period (0) or the post-treatment period (1)

Your Turn in RMarkdown

A Study of Decentralization on Public Services in Vietnam

- Implement a linear regression model to estimate the treatment effect using a simple difference in differences (DID) design
    - Filter your data so that you only keep the years 2008 and 2010
    - Build a linear regression model

$$Y_{i,t} = \alpha + \gamma \text{Treated} + \lambda \text{Time} + \delta(\text{Treated} \times \text{Time}) + \varepsilon_{i,t}$$

    - Interpret the results to get the treatment effect estimate