Difference in Differences

Causal Inference Discussion Section

Wednesday, November 1st

Reminders & Announcements

- PS5 due tomorrow by 5pm
- Peer reviews will be due next Thursday
- Final project write-up due Tuesday, November 21 @ 5pm
  - Reach our to your group if no one else has!!!
  - Let us know if you have questions/concerns

Group Activity

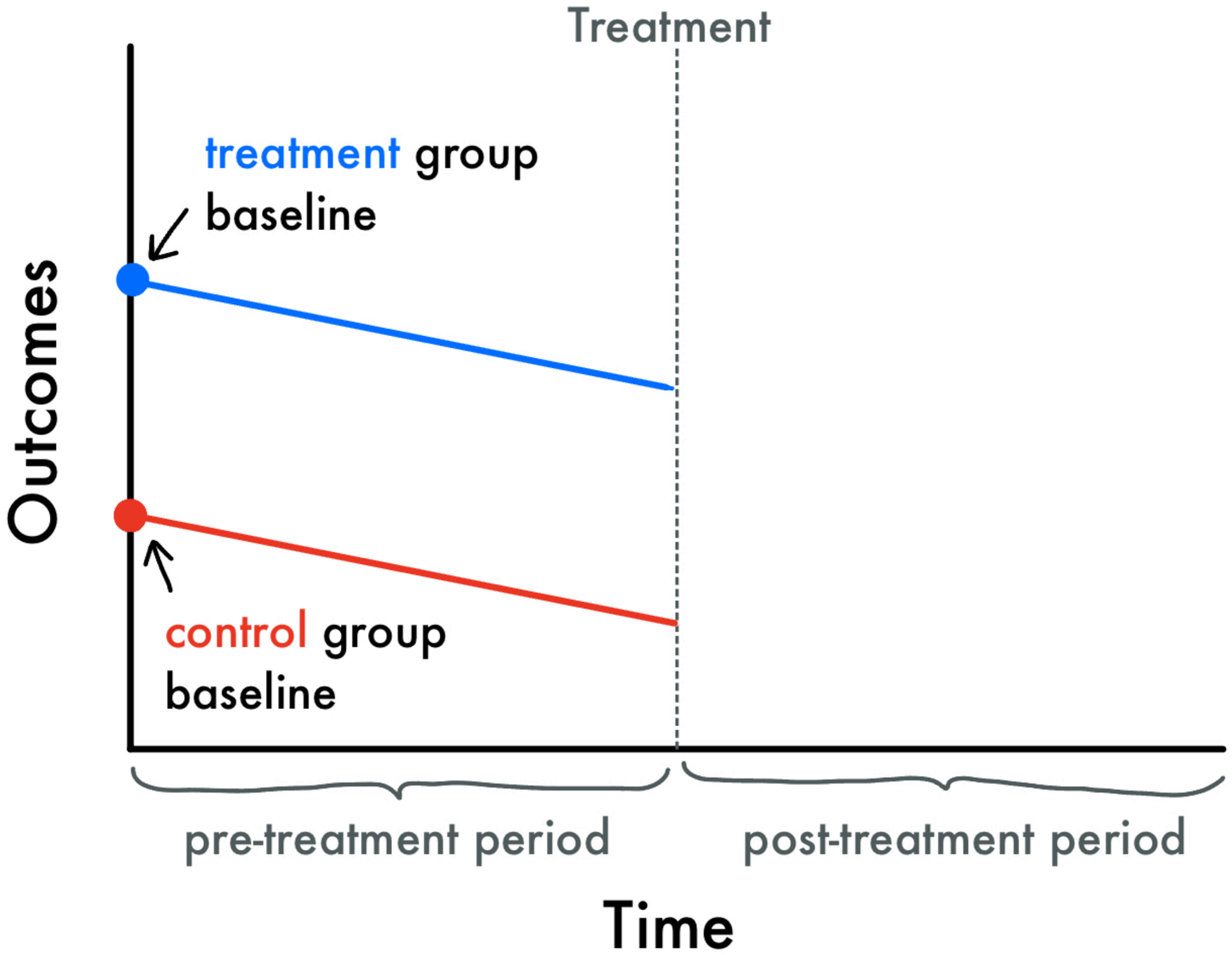
Difference in Differences (DID) - Parallel Trends

In groups, discuss the parallel trends assumption for DID designs.

- What do we mean by parallel trends?
- Is this an assumption on the data in the pre-treatment or post-treatment periods?
- Is this an assumption about the treatment group, the control group, or both?

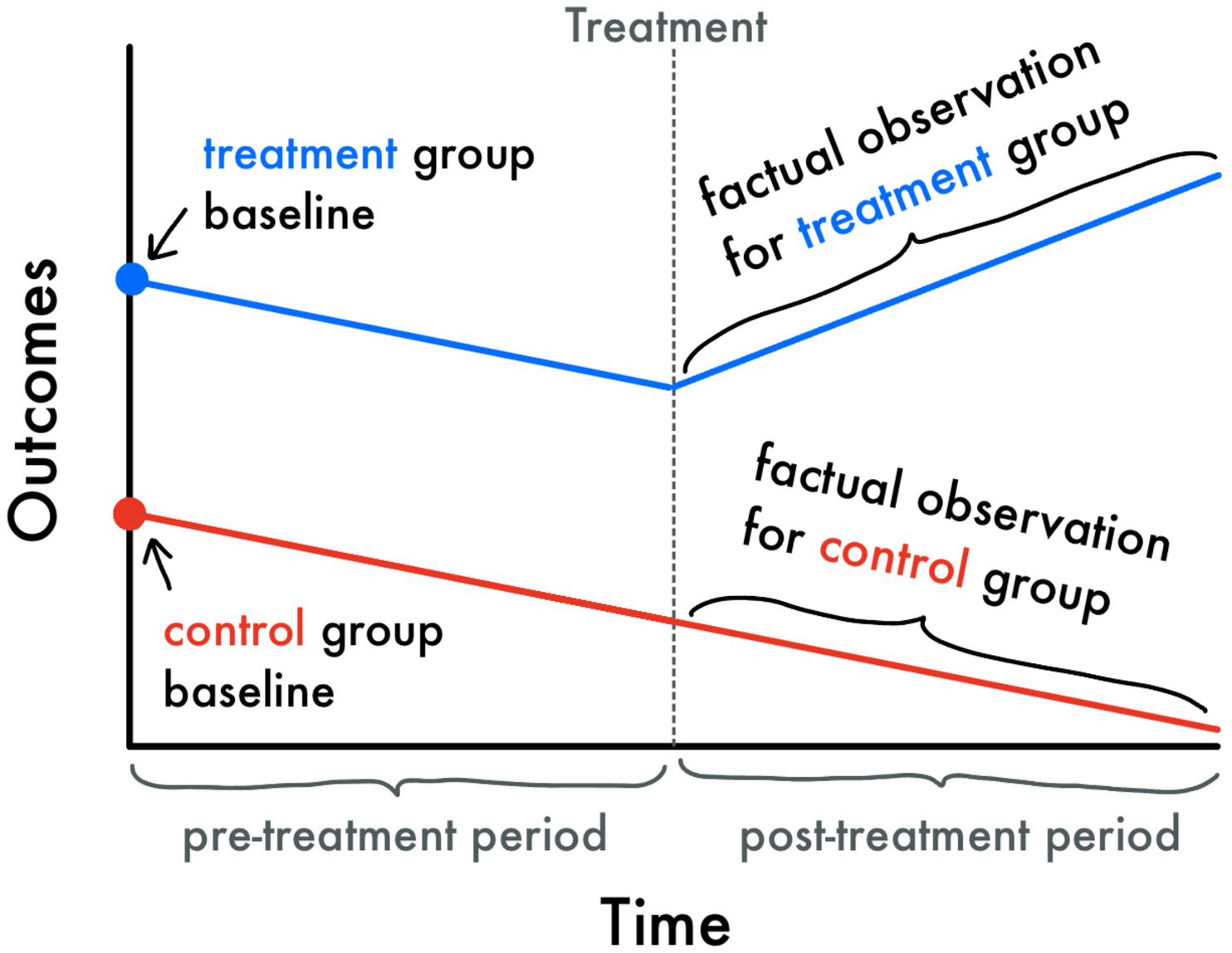
Difference in Differences Review

Visual illustration



Difference in Differences Review

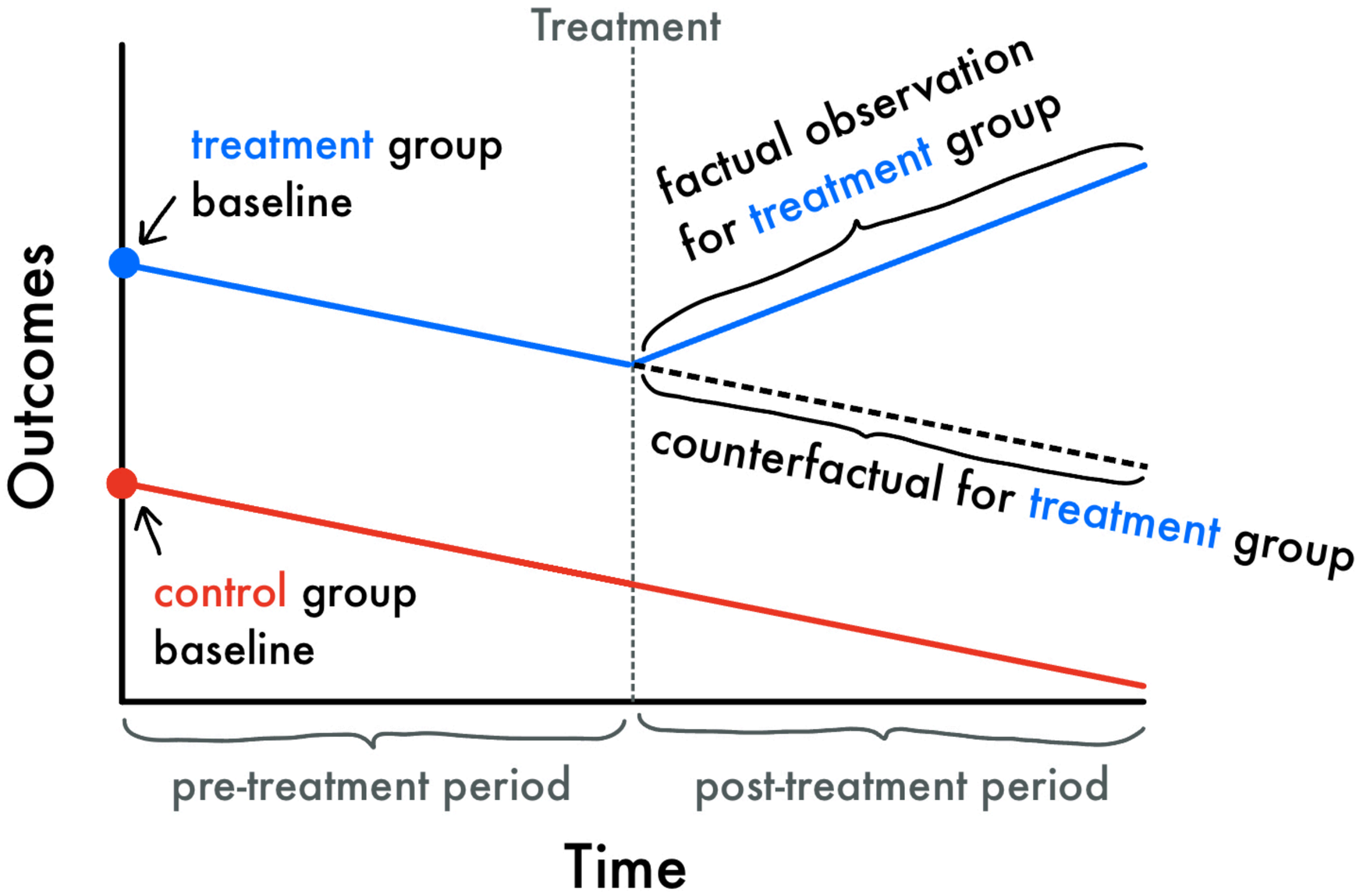
Visual illustration



Difference in Differences Review

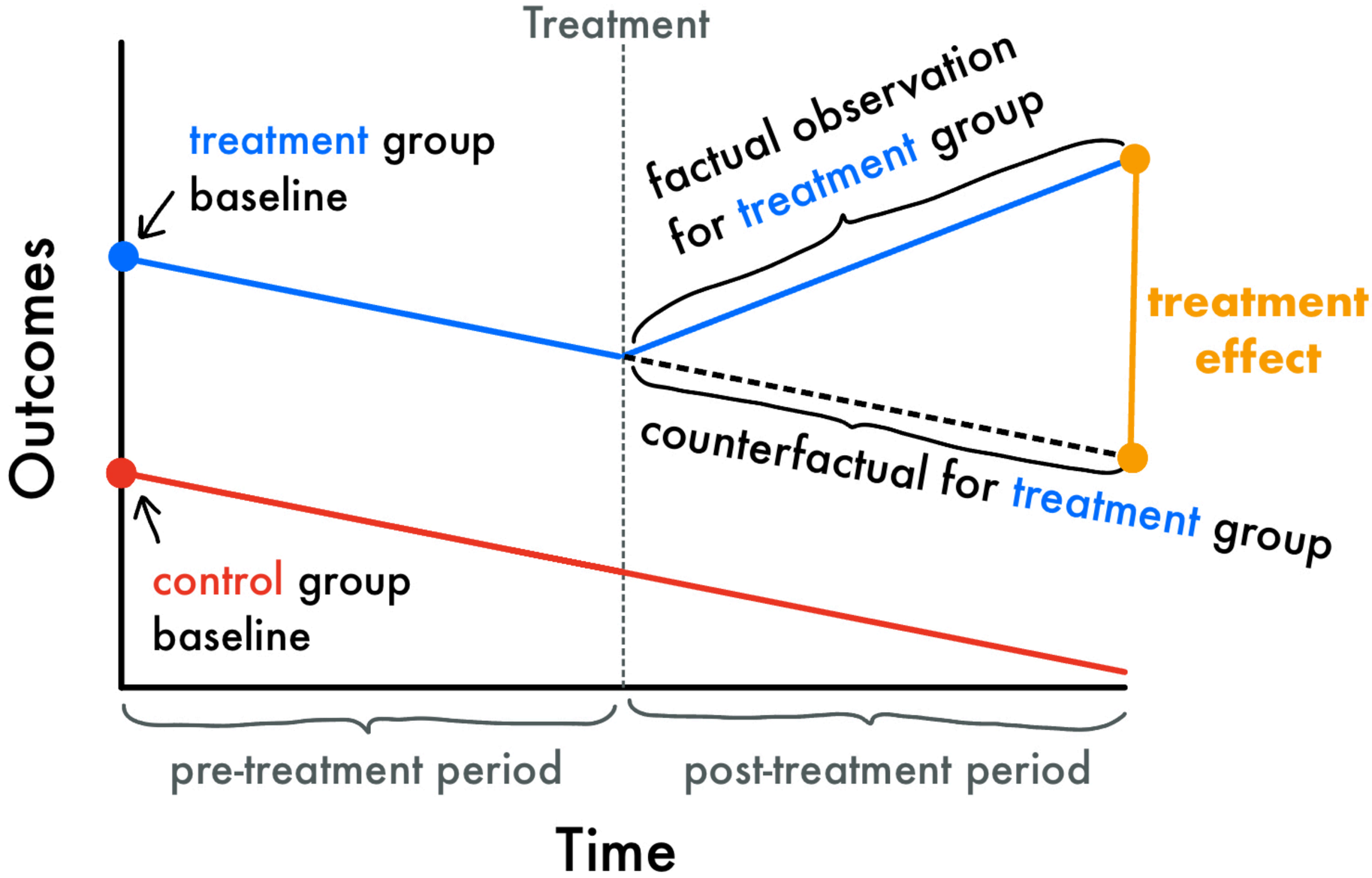
Visual illustration

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

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

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

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

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: $\alpha + \gamma + \lambda + \delta$

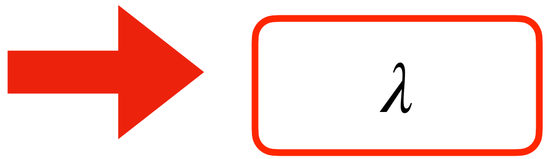
Difference in Differences Review

Using Regression

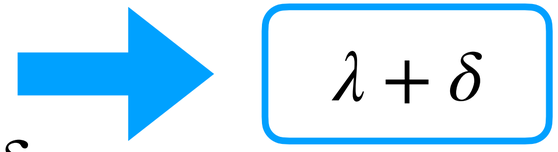
Consider the following linear model for outcomes:

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

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



λ



$\lambda + \delta$

1st difference(s)

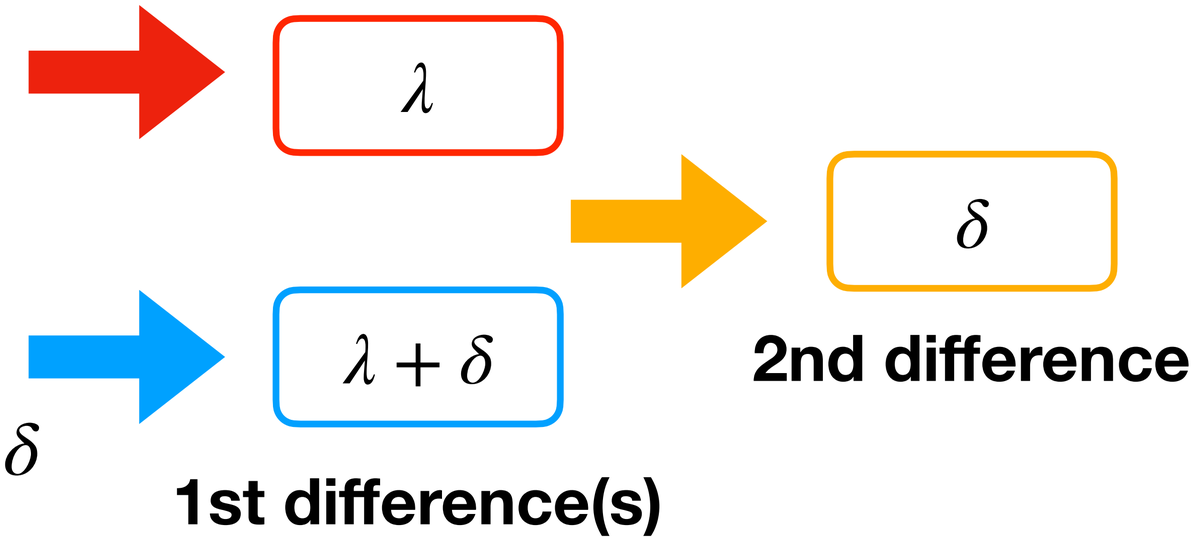
Difference in Differences Review

Using Regression

Consider the following linear model for outcomes:

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

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



Application

A Study of Recentralization on Public Services in Vietnam

- Introduced in lecture yesterday
- Looking at the effects of decentralizing government (`treatment`) on public services such as educational programs (`pro4`), tap water (`tapwater`), and agricultural centers (`agrext`) in Vietnam
- 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 Recentralization 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