Difference in Differences

Jacob M. Montgomery

Quantitative Political Methodology

Causality in Regression and Difference in Differences

Road map

Where we have been:

- What is regression?
- How to interpret coefficients?
- Interactions/Dummies

Road map

Where we have been:

- What is regression?
- How to interpret coefficients?
- Interactions/Dummies

Today:

- Using regression for causal inference
- Using difference-in-differences to make causal claims

Causal inference

- We will use T to represent a treatment variable.
- For a categorical treatment

$$T_i = \begin{array}{l} 1 \text{ if unit i receives the treatment} \\ 0 \text{ if unit i receives the control} \end{array}$$

Causal inference

- We will use T to represent a treatment variable.
- For a categorical treatment

$$T_i = \begin{matrix} \text{1 if unit i receives the treatment} \\ \text{0 if unit i receives the control} \end{matrix}$$

- We let y_i^1 represent the outcome of the ith unit if the treatment is given.

Causal inference

- We will use T to represent a treatment variable.
- For a categorical treatment

$$T_i = \begin{array}{l} 1 \text{ if unit i receives the treatment} \\ 0 \text{ if unit i receives the control} \end{array}$$

- We let y_i^1 represent the outcome of the ith unit if the treatment is given.
- We let y_i^0 represent the outcome of the ith unit if the control is given.

Causal inference

- We will use T to represent a treatment variable.
- For a categorical treatment

$$T_i = \begin{array}{l} \text{1 if unit i receives the treatment} \\ \text{0 if unit i receives the control} \end{array}$$

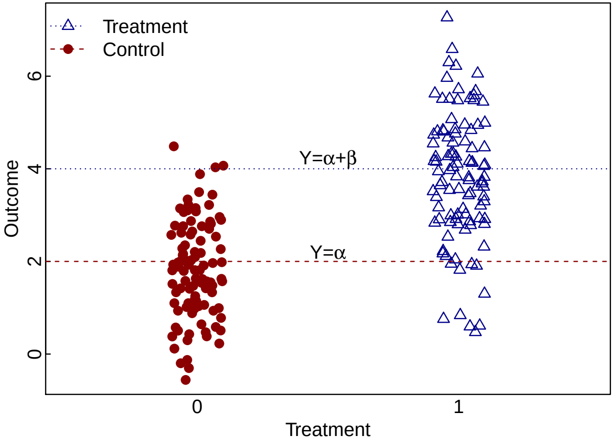
- We let y_i^1 represent the outcome of the ith unit if the treatment is given.
- We let y_i^0 represent the outcome of the ith unit if the control is given.
- One of these is observed, the other is the counterfactual – what would have been observed if the other treatment have been given?

Causal inference

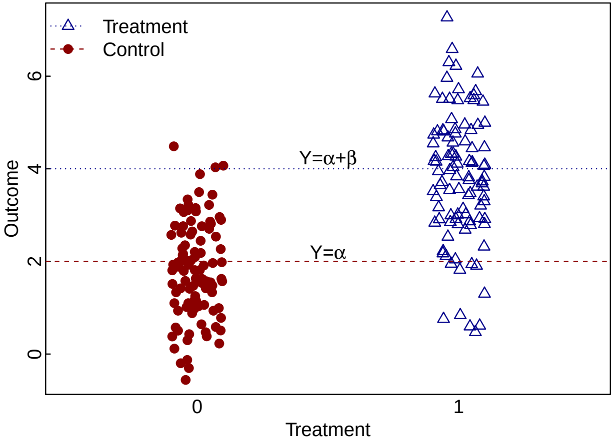
- We will use T to represent a treatment variable.
- For a categorical treatment

$$T_i = \begin{array}{l} 1 \text{ if unit i receives the treatment} \\ 0 \text{ if unit i receives the control} \end{array}$$

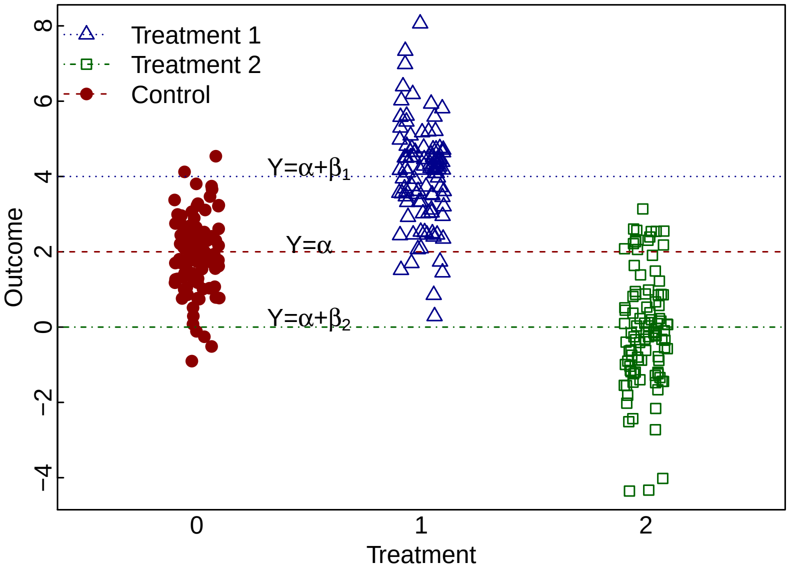
- We let y_i^1 represent the outcome of the ith unit if the treatment is given.
- We let y_i^0 represent the outcome of the ith unit if the control is given.
- One of these is observed, the other is the counterfactual – what would have been observed if the other treatment have been given?
- The causal effect of T_i on observation i will then be $y_i^1 - y_i^0$



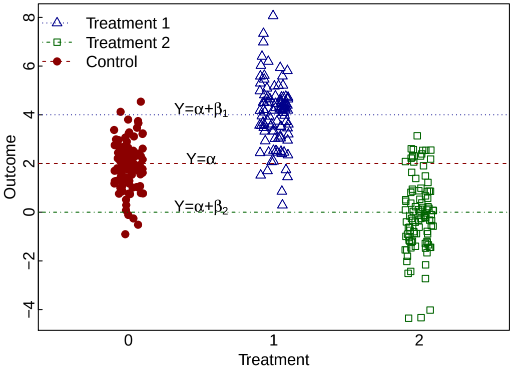
Treatment
Control
Y=α+β
Y=α
Outcome
Treatment
0
2
4
6
0
1



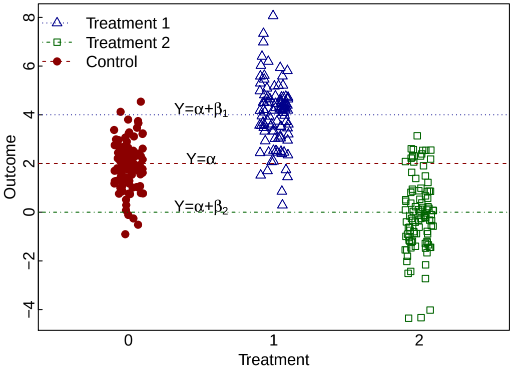
- $ATE = mean(y_i^1) - mean(y_i^0)$
- $ATE = (\alpha + \beta) - (\alpha)$
- $ATE = \beta$



Treatment 1
Treatment 2
Control
Y=α+β₁
Y=α
Y=α+β₂
Outcome
Treatment
8
6
4
2
0
−2
−4
0
1
2

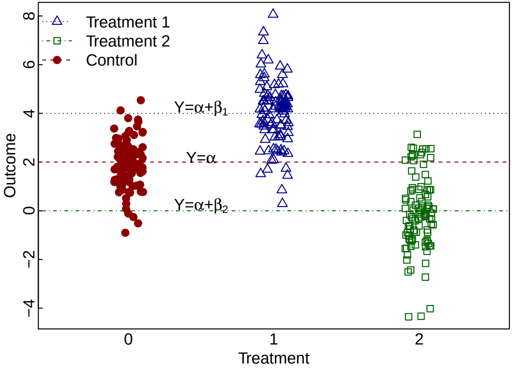


- $Y = \alpha + T_1\beta_1 + T_2\beta_2$

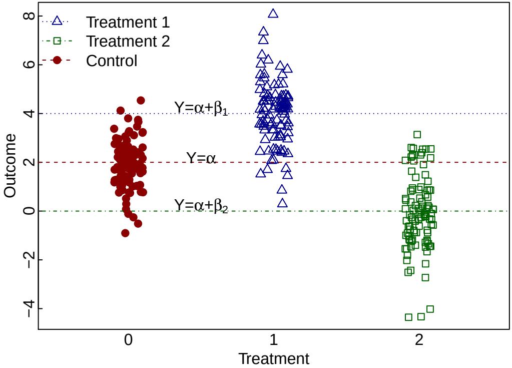


- $Y = \alpha + T_1\beta_1 + T_2\beta_2$

- Control: $Y = \alpha$
- Treatment 1: $Y = \alpha + \beta_1$
- Treatment 2: $Y = \alpha + \beta_2$

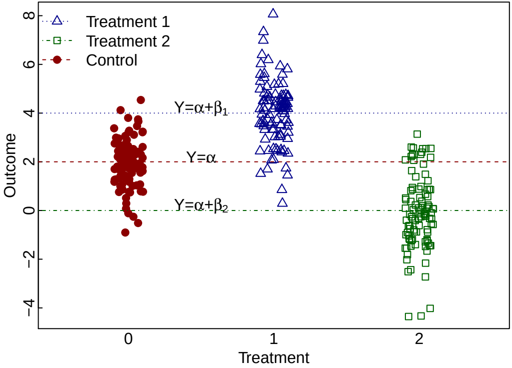


- $Y = \alpha + T_1\beta_1 + T_2\beta_2$
- $ATE_1 = mean(y_i^1) - mean(y_i^0)$
- $ATE_1 = (\alpha + \beta_1) - (\alpha)$
- $ATE_1 = \beta_1$



- $ATE_2 = mean(y_i^2) - mean(y_i^0)$
- $ATE_2 = (\alpha + \beta_2) - (\alpha)$
- $ATE_2 = \beta_2$

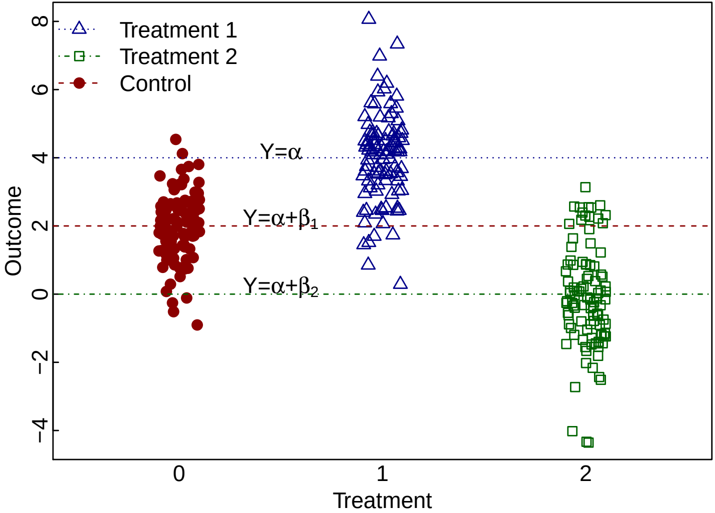
Is blue Different than green?



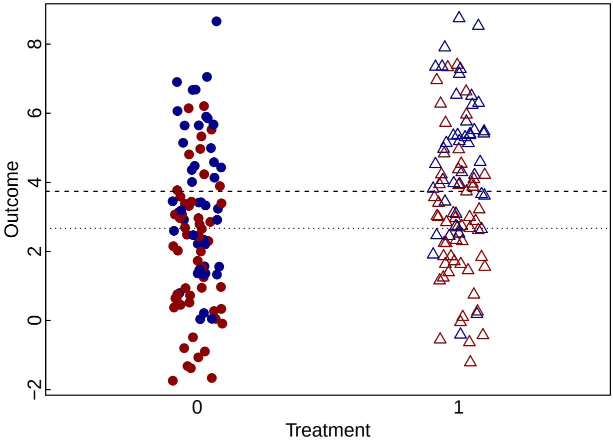
- $ATE_{2vs.1} = mean(y_i^2) - mean(y_i^1)$
- $ATE_{2vs.1} = (\alpha + \beta_2) - (\alpha + \beta 1)$
- $ATE_{2vs.1} = \beta_2 - \beta_1$

Is blue Different than green?

A much easier way to do this is to change what is the “control” and leave that out.

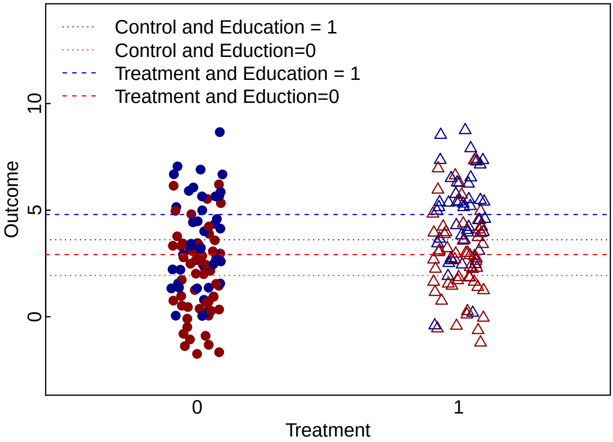


Control variables



$$Y = \alpha + T\beta_1 + X\beta_2$$

Control variables

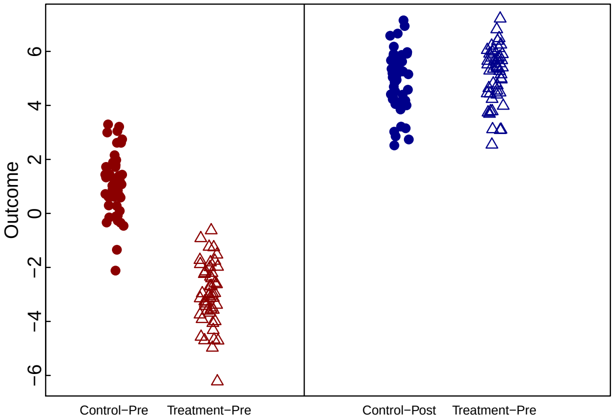


$$Y = \alpha + T\beta_1 + X\beta_2$$

Adding control variables improves precision

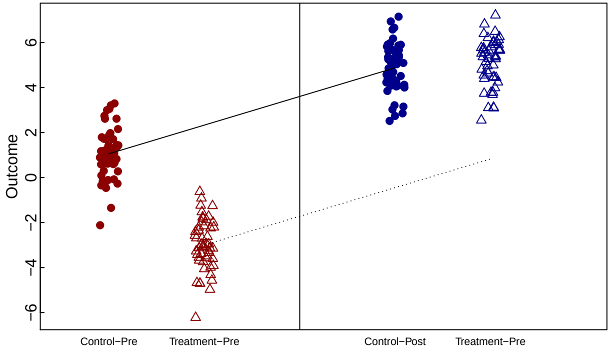
	No Control	Control
Intercept	2.67	1.89
	(0.22)	(0.24)
Treatment	1.43	1.42
	(0.31)	(0.29)
Covariate		1.78
		(0.29)
R^2	0.056	0.136

Difference in differences

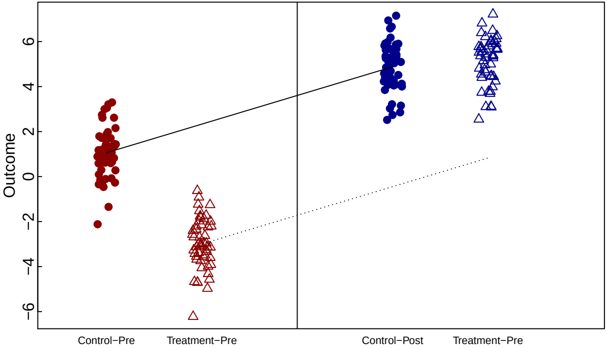


$$Y = \alpha + Treatment\beta_1 + Pre/Post\beta_2 + (T \times PP)\beta_3$$

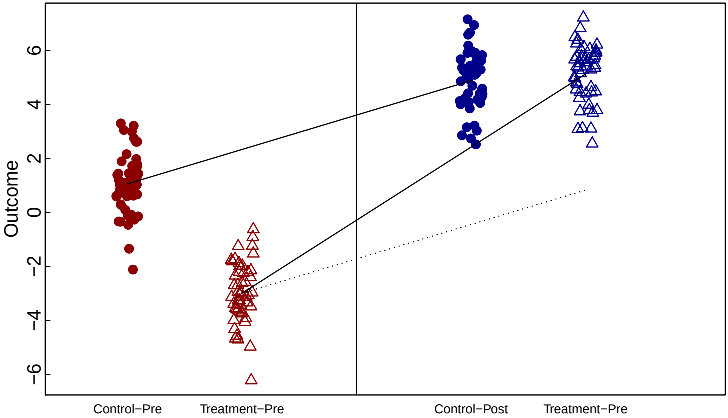
Difference in differences



Difference in differences



- Assume parallel paths
- The treatment group would have moved in parallel to the control group in the absence of the intervention.
- We can estimate the effect of the treatment on the treated



- We need to model both what happened and what would have happened
- $Y = \alpha + Treatment\beta_1 + Pre/Post\beta_2 + (T \times PP)\beta_3$

Diff-in-diff model

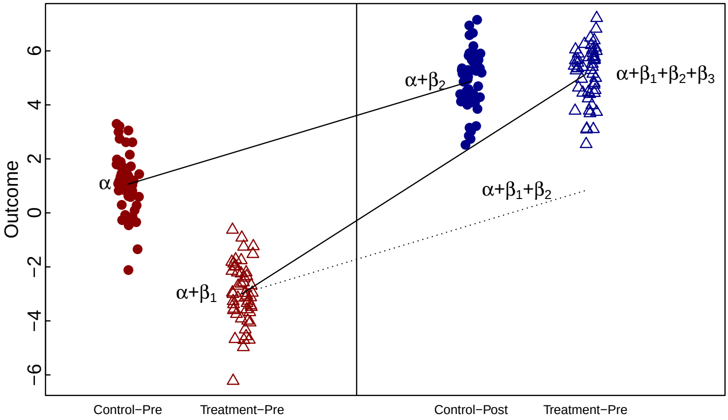
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.0591	0.1547	6.84	0.0000
Assignment	-4.0522	0.2188	-18.52	0.0000
PrePost	3.8202	0.2188	17.46	0.0000
Assignment:PrePost	4.2994	0.3095	13.89	0.0000

Diff-in-diff model

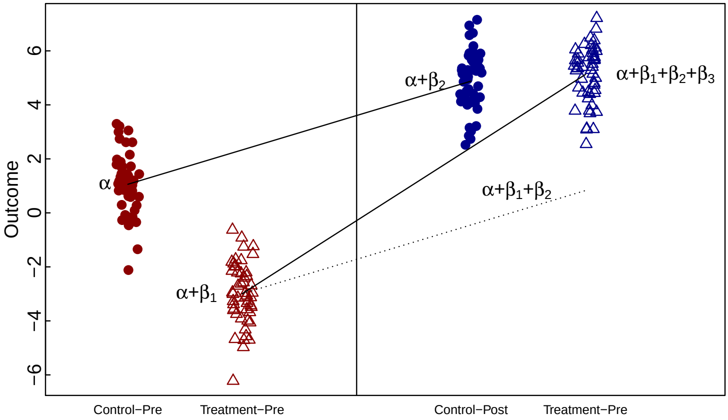
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.0591	0.1547	6.84	0.0000
Assignment	-4.0522	0.2188	-18.52	0.0000
PrePost	3.8202	0.2188	17.46	0.0000
Assignment:PrePost	4.2994	0.3095	13.89	0.0000

Let's write down the prediction equation on the board.

Difference in differences



Difference in differences



The difference between what we observed $(\alpha + \beta_1 + \beta_2 + \beta_3)$ and what we would have observed $(\alpha + \beta_1 + \beta_2)$ is

$$(\alpha + \beta_1 + \beta_2 + \beta_3) - (\alpha + \beta_1 + \beta_2) = \beta_3$$

Difference in difference: Discussion

- Despite being tricky to think about it, it is simple to estimate.

Difference in difference: Discussion

- Despite being tricky to think about it, it is simple to estimate.
- Does *not require* that the treatment and control group be similar, only that they would have had parallel paths.

Difference in difference: Discussion

- Despite being tricky to think about it, it is simple to estimate.
- Does *not require* that the treatment and control group be similar, only that they would have had parallel paths.
- You can (and should) also control for other factors, which will help improve your estimates.

Applied Example: Minimum wage laws

What is the effect of increasing the minimum wage on low-wage workers?

- Side 1?
- Side 2?

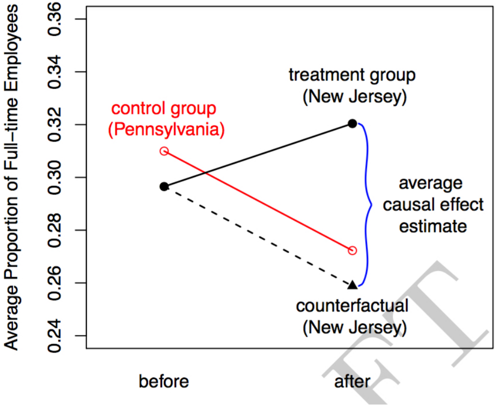
Study basics:

- In 1992, New Jersey increased the minimum wage from \$4.25/hour to \$5.05/hour.
- Researchers collected information about the number of full time employees (as a proportion of total number of employees) at various fast food chains before and after the reform.

Study basics:

- In 1992, New Jersey increased the minimum wage from \$4.25/hour to \$5.05/hour.
- Researchers collected information about the number of full time employees (as a proportion of total number of employees) at various fast food chains before and after the reform.

- Researchers *also* collected the same information from the same chains in Pennsylvania to serve as a counterfactual.
- Are New Jersey and Pennsylvania perfect counterfactuals for each other?



- Divide up into your groups.
- Interpret this plot. What was the effect of the wage increase in New Jersey? Be ready to answer.