Lecture 19: Difference-in-Difference Models

ECON 480 - Econometrics - Fall 2018

Ryan Safner

December 5, 2018

Difference-in-Difference Models

Difference-in-Difference Models

- Often, we want to examine the consequences of a change, such as a law or policy

- Often, we want to examine the consequences of a change, such as a law or policy
    - e.g. States that implemented law X saw a change in Y

- Often, we want to examine the consequences of a change, such as a law or policy
  - e.g. States that implemented law X saw a change in Y
  - **Treatment**: States that implement law X

- Often, we want to examine the consequences of a change, such as a law or policy
  - e.g. States that implemented law X saw a change in Y
  - **Treatment**: States that implement law X
  - **Control**: States that did not implement law X

- Often, we want to examine the consequences of a change, such as a law or policy
    - e.g. States that implemented law X saw a change in Y
    - **Treatment**: States that implement law X
    - **Control**: States that did not implement law X
    - If we have **panel data** with observations for all states before *and* after the change:

- Often, we want to examine the consequences of a change, such as a law or policy
    - e.g. States that implemented law X saw a change in Y
    - **Treatment**: States that implement law X
    - **Control**: States that did not implement law X
    - If we have **panel data** with observations for all states before *and* after the change:
- Simple logic: compare difference in outcomes of treatment group (before and after treatment) with those of non-treated group (before and after same treatment period)

Difference-in-Difference Models II

- The difference-in-difference model (aka ”diff-in-diff”” or ”DND”) identifies treatment effect by differencing the difference pre- and post-treatment between treatment and control groups

$$\widehat{Y_{it}} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i * \text{After}_t) + \epsilon_{it}$$

- The difference-in-difference model (aka "diff-in-diff"" or "DND") identifies treatment effect by differencing the difference pre- and post-treatment between treatment and control groups

$$\widehat{Y_{it}} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i * \text{After}_t) + \epsilon_{it}$$

  - $Treated_i = \begin{cases} 1 \text{ if } i \text{ is in treatment group} \\ 0 \text{ if } i \text{ is not in treatment group} \end{cases}$

- The difference-in-difference model (aka "diff-in-diff"" or "DND") identifies treatment effect by differencing the difference pre- and post-treatment between treatment and control groups

$$\widehat{Y_{it}} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i * \text{After}_t) + \epsilon_{it}$$

  - $Treated_i = \begin{cases} 1 \text{ if } i \text{ is in treatment group} \\ 0 \text{ if } i \text{ is not in treatment group} \end{cases}$
  - $After_t = \begin{cases} 1 \text{ if } t \text{ is after treatment period} \\ 0 \text{ if } t \text{ is before treatment period} \end{cases}$

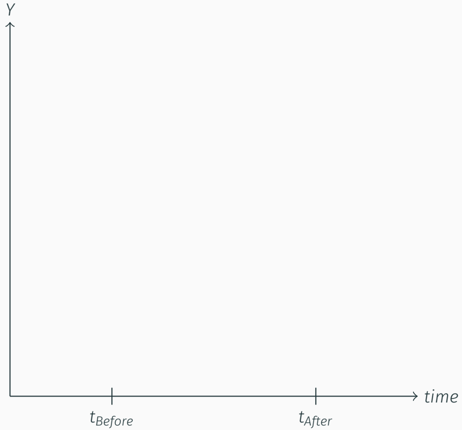
- The difference-in-difference model (aka "diff-in-diff"" or "DND") identifies treatment effect by differencing the difference pre- and post-treatment between treatment and control groups

$$\widehat{Y_{it}} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i * \text{After}_t) + \epsilon_{it}$$

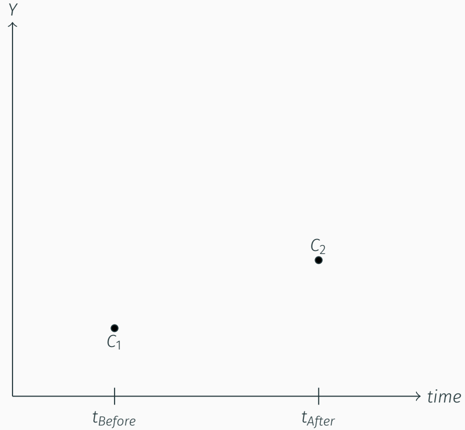
  - $Treated_i = \begin{cases} 1 \text{ if } i \text{ is in treatment group} \\ 0 \text{ if } i \text{ is not in treatment group} \end{cases}$
  - $After_t = \begin{cases} 1 \text{ if } t \text{ is after treatment period} \\ 0 \text{ if } t \text{ is before treatment period} \end{cases}$

	Control	Treatment	Group Diff. (ΔY_i)
Before	β_0	$\beta_0 + \beta_1$	β_1
After	$\beta_0 + \beta_2$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_1 + \beta_3$
Time Diff. (ΔY_t)	β_2	$\beta_2 + \beta_3$	β_3
			Diff-in-diff ($\Delta_i \Delta_t Y$)

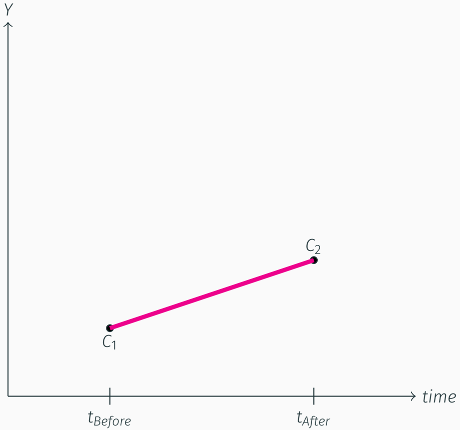
$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$

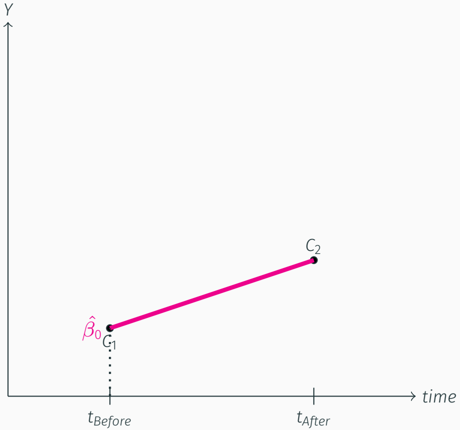


$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



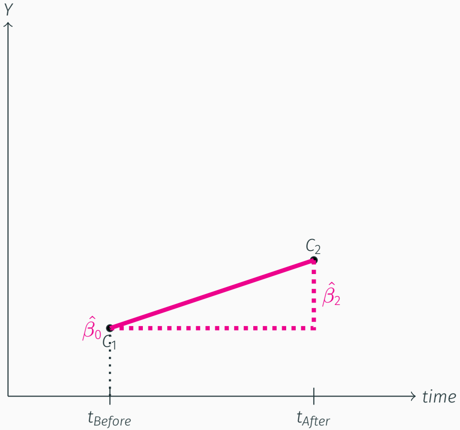
- Control (*Treated* = 0) group

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon_{it}}$$



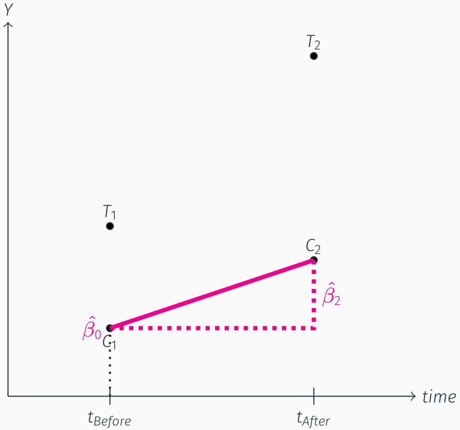
- Control ($Treated = 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



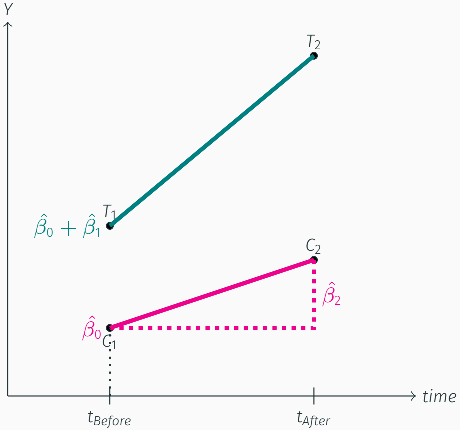
- Control ($Treated = 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment
- $\hat{\beta}_2$: time difference (for control group)

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



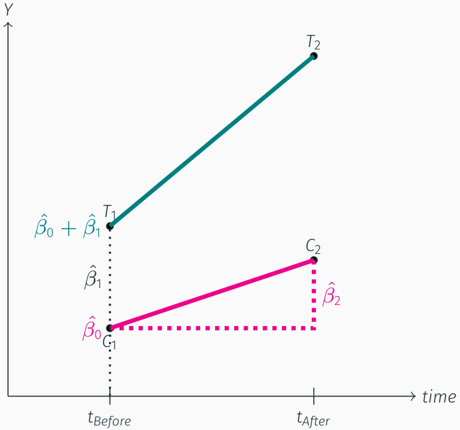
- Control ($Treated = 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment
- $\hat{\beta}_2$: time difference (for control group)

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



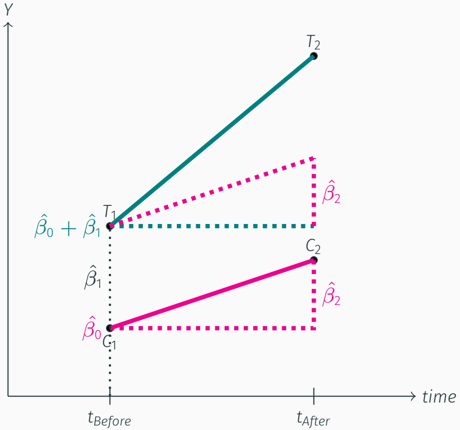
- Treatment ($Treated = 1$) group
- Control ($Treated = 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment
- $\hat{\beta}_2$: time difference (for control group)

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



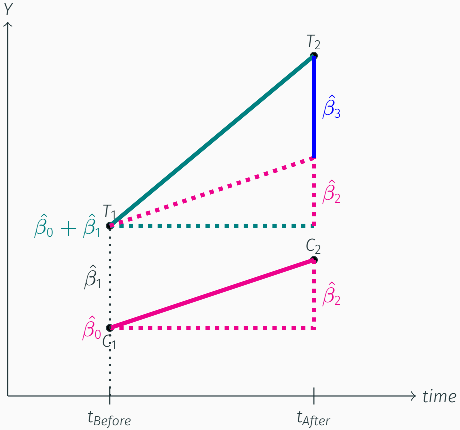
- Treatment (*Treated* $= 1$) group
- Control (*Treated* $= 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment
- $\hat{\beta}_1$: difference between treatment and control (before treatment)
- $\hat{\beta}_2$: time difference (for control group)

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3(\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



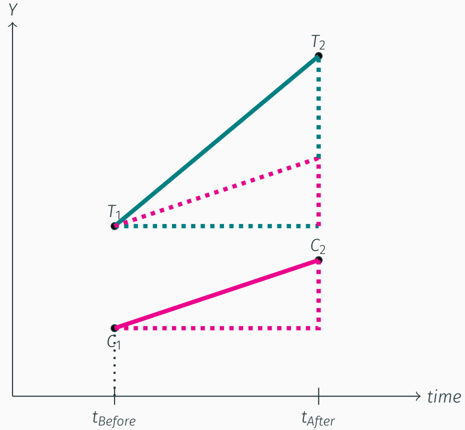
- Treatment ($Treated = 1$) group
- Control ($Treated = 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment
- $\hat{\beta}_1$: difference between treatment and control (before treatment)
- $\hat{\beta}_2$: time difference (for control group)

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$

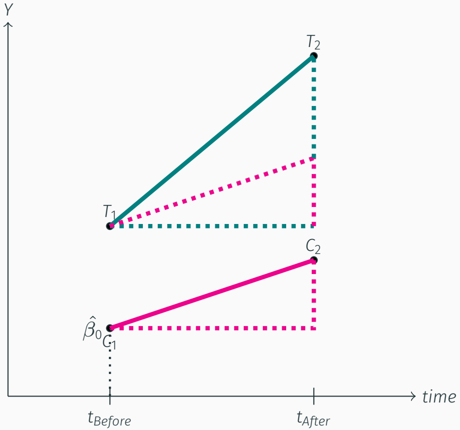


- Treatment ($Treated = 1$) group
- Control ($Treated = 0$) group
- $\hat{\beta}_0$: value of Y for control before treatment
- $\hat{\beta}_1$: difference between treatment and control (before treatment)
- $\hat{\beta}_2$: time difference (for control group)
- $\hat{\beta}_3$: difference-in-difference: effect of treatment

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$

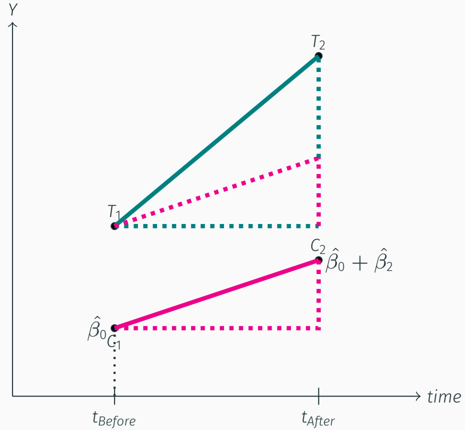


$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



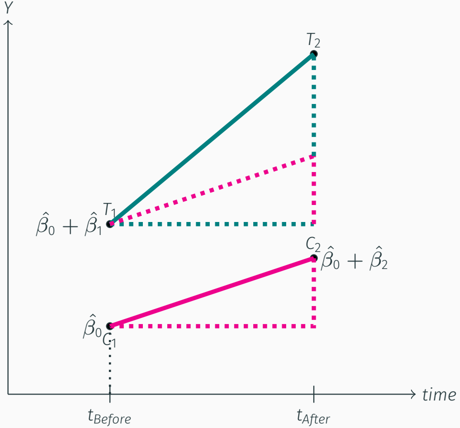
- Y for Control Group Before: $\hat{\beta}_0$

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



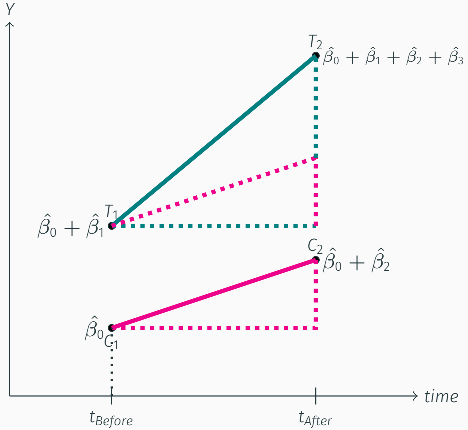
- Y for Control Group Before: $\hat{\beta}_0$
- Y for Control Group After: $\hat{\beta}_0 + \hat{\beta}_2$

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon_{it}}$$



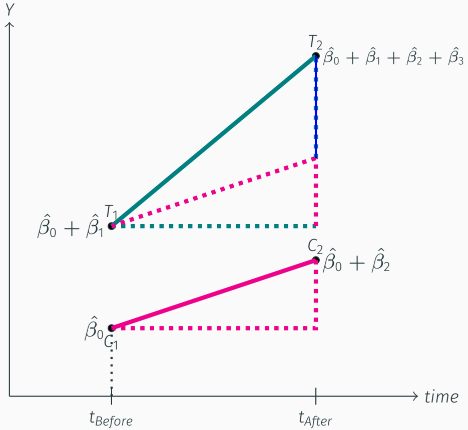
- Y for Control Group Before: $\hat{\beta}_0$
- Y for Control Group After: $\hat{\beta}_0 + \hat{\beta}_2$
- Y for Treatment Group Before: $\hat{\beta}_0 + \hat{\beta}_1$

$$\widehat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



- Y for Control Group Before: $\hat{\beta}_0$
- Y for Control Group After: $\hat{\beta}_0 + \hat{\beta}_2$
- Y for Treatment Group Before: $\hat{\beta}_0 + \hat{\beta}_1$
- Y for Treatment Group After: $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



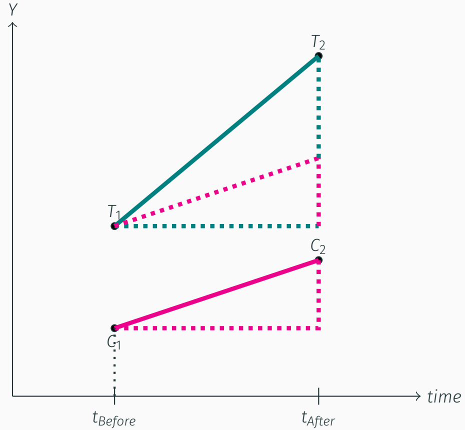
- Y for Control Group Before: $\hat{\beta}_0$
- Y for Control Group After: $\hat{\beta}_0 + \hat{\beta}_2$
- Y for Treatment Group Before: $\hat{\beta}_0 + \hat{\beta}_1$
- Y for Treatment Group After: $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$
- Treatment Effect: $\hat{\beta}_3$

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$

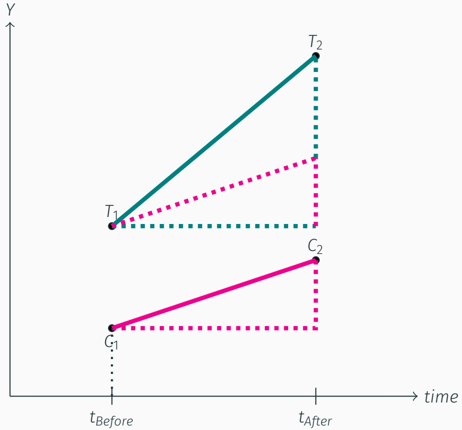
	Control	Treatment	Group Diff. (ΔY_i)
Before	β_0	$\beta_0 + \beta_1$	β_1
After	$\beta_0 + \beta_2$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_1 + \beta_3$
Time Diff. (ΔY_t)	β_2	$\beta_2 + \beta_3$	β_3
			Diff-in-diff ($\Delta_i \Delta_t Y$)

Key Assumption About Counterfactual

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



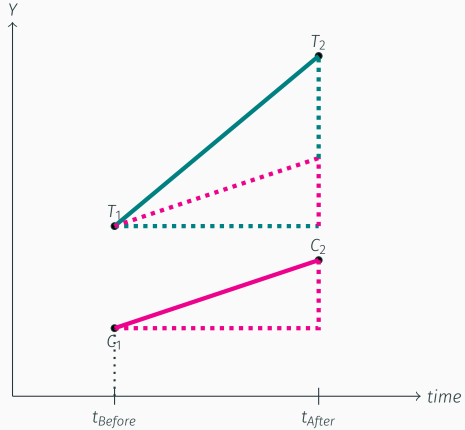
$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon_{it}}$$



- Key assumption in DND models is that the time trend is parallel

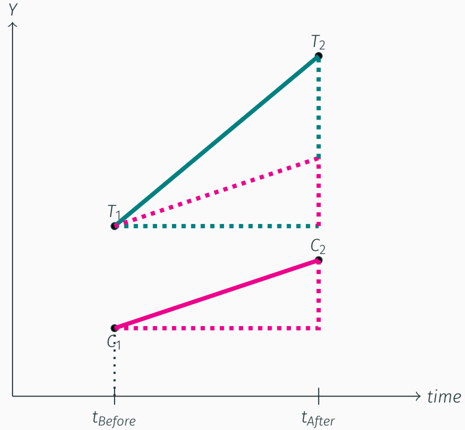
Key Assumption About Counterfactual

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



- Key assumption in DND models is that the time trend is parallel
- Treatment & control groups must be similar over time *except for treatment*

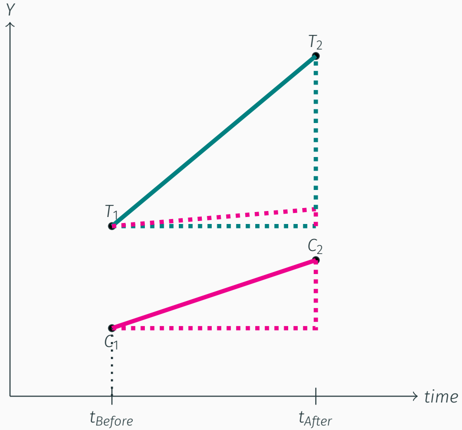
$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



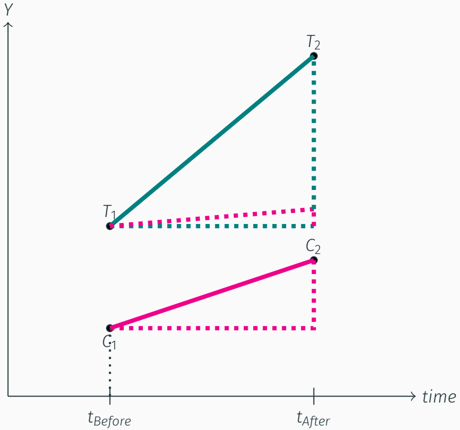
- Key assumption in DND models is that the time trend is parallel
- Treatment & control groups must be similar over time *except for treatment*
- Counterfactual if the treatment group were *not* treated, they would change the same as control group over time $\hat{\beta}_2$

Key Assumption about Counterfactual II

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$

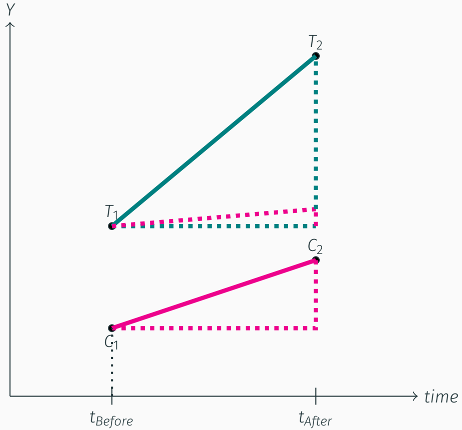


$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon}_{it}$$



- If time trend is different between treatment and control groups

$$\widehat{Y_{it}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Treated}_i + \hat{\beta}_2 \text{After}_t + \hat{\beta}_3 (\text{Treated}_i \times \text{After}_t) + \hat{\epsilon_{it}}$$



- If time trend is different between treatment and control groups
- Treatment effect may be over/under-estimated!

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- Dynarski, Susan (2000), "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance" micro-level data on 4,291 young individuals:

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- Dynarski, Susan (2000), "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance" micro-level data on 4,291 young individuals:
  - $\text{InCollege}_{it} = \begin{cases} 1 \text{ if } i \text{ is in college during year } t \\ 0 \text{ if } i \text{ is not in college during year } t \end{cases}$

HOOD COLLEGE

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- Dynarski, Susan (2000), "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance" micro-level data on 4,291 young individuals:
    - $\text{InCollege}_{it} = \begin{cases} 1 \text{ if } i \text{ is in college during year } t \\ 0 \text{ if } i \text{ is not in college during year } t \end{cases}$
    - $\text{Georgia}_{i} = \begin{cases} 1 \text{ if } i \text{ is a Georgia resident} \\ 0 \text{ if } i \text{ is not a Georgia resident} \end{cases}$

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- Dynarski, Susan (2000), "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance" micro-level data on 4,291 young individuals:
  - $\text{InCollege}_{it} = \begin{cases} 1 \text{ if } i \text{ is in college during year } t \\ 0 \text{ if } i \text{ is not in college during year } t \end{cases}$
  - $\text{Georgia}_{i} = \begin{cases} 1 \text{ if } i \text{ is a Georgia resident} \\ 0 \text{ if } i \text{ is not a Georgia resident} \end{cases}$
  - $\text{After}_{t} = \begin{cases} 1 \text{ if } t \text{ is after 1992} \\ 0 \text{ if } t \text{ is after 1992} \end{cases}$

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- Dynarski, Susan (2000), "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance" micro-level data on 4,291 young individuals:
  - $\text{InCollege}_{it} = \begin{cases} 1 \text{ if } i \text{ is in college during year } t \\ 0 \text{ if } i \text{ is not in college during year } t \end{cases}$
  - $\text{Georgia}_{i} = \begin{cases} 1 \text{ if } i \text{ is a Georgia resident} \\ 0 \text{ if } i \text{ is not a Georgia resident} \end{cases}$
  - $\text{After}_{t} = \begin{cases} 1 \text{ if } t \text{ is after 1992} \\ 0 \text{ if } t \text{ is after 1992} \end{cases}$
  - Note: With a dummy *dependent* variable (Y), coefficients estimate the probability $Y = 1$, i.e. the *probability* a person is enrolled in college

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

Example
In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- We can use a diff-in-diff model to measure the effect of HOPE scholarship on enrollments

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- We can use a diff-in-diff model to measure the effect of HOPE scholarship on enrollments
- Georgia and nearby states: if not for HOPE, changes should be the same over time

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- We can use a diff-in-diff model to measure the effect of HOPE scholarship on enrollments
- Georgia and nearby states: if not for HOPE, changes should be the same over time
- Treatment period: 1992

Diff-in-Diff Example II

Example
In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- We can use a diff-in-diff model to measure the effect of HOPE scholarship on enrollments
- Georgia and nearby states: if not for HOPE, changes should be the same over time
- Treatment period: 1992
- Treatment: Georgia

Example

In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- We can use a diff-in-diff model to measure the effect of HOPE scholarship on enrollments
- Georgia and nearby states: if not for HOPE, changes should be the same over time
- Treatment period: 1992
- Treatment: Georgia

$$\Delta_i \Delta_t Enrolled = (\text{GA}_{after} - \text{GA}_{before}) - (\text{neighbors}_{after} - \text{neighbors}_{before})$$

Example
In 1993 Georgia initiated a HOPE scholarship program to let state residents with at least a B average in high school attend public college in Georgia for free. Did it increase college enrollment?

- We can use a diff-in-diff model to measure the effect of HOPE scholarship on enrollments
- Georgia and nearby states: if not for HOPE, changes should be the same over time
- Treatment period: 1992
- Treatment: Georgia

$$\Delta_i \Delta_t Enrolled = (\mathrm{GA}_{after} - \mathrm{GA}_{before}) - (\mathrm{neighbors}_{after} - \mathrm{neighbors}_{before})$$

$$\widehat{Enrolled}_{it} = \beta_0 + \beta_1 Georgia_{it} + \beta_2 After_{it} + \beta_3 Georgia_{it} \times After_{it}$$

```r
DND<-lm(InCollege~Georgia+After+AfterGeorgia, data=HOPE)
summary(DND)
```

```
##
## Call:
## lm(formula = InCollege ~ Georgia + After + AfterGeorgia, data = HOPE)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4058 -0.4058 -0.4013  0.5942  0.6995
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.40578    0.01092  37.146  < 2e-16 ***
## Georgia      -0.10524    0.03778  -2.785  0.00537 **
## After        -0.00446    0.01585  -0.281  0.77848
## AfterGeorgia  0.08933    0.04889   1.827  0.06776 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4893 on 4287 degrees of freedom
## Multiple R-squared:  0.001872,   Adjusted R-squared:  0.001174
## F-statistic: 2.681 on 3 and 4287 DF,  p-value: 0.04528
```

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_i - 0.004 After_t + 0.08$$

$$\widehat{Enrolled}_{it} = 0.406 - 0.105Georgia_i - 0.004After_t + 0.089Georgia_i \times After_t$$

- β_0: A non-Georgian before 1992 was 40.6% likely to be a college student

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_i - 0.004 After_t + 0.089 Georgia_i \times After_t$$

- β_0: A non-Georgian before 1992 was 40.6% likely to be a college student
- β_1: Georgians are 10.5% less likely to be college students than neighboring states

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_i - 0.004 After_t + 0.089 Georgia_i \times After_t$$

- β_0: A non-Georgian before 1992 was 40.6% likely to be a college student
- β_1: Georgians are 10.5% less likely to be college students than neighboring states
- β_2: After 1992, young Americans are 0.4% less likely to be college students

$$\widehat{Enrolled}_{it} = 0.406 - 0.105Georgia_i - 0.004After_t + 0.089Georgia_i \times After_t$$

- β_0: A non-Georgian before 1992 was 40.6% likely to be a college student
- β_1: Georgians are 10.5% less likely to be college students than neighboring states
- β_2: After 1992, young Americans are 0.4% less likely to be college students
- β_3: After 1992, Georgians are 8.9% more likely to enroll in colleges than neighboring states

-Treatment effect: HOPE increased enrollment likelihood by 8.9%

Diff-in-Diff Example: Comparing Group Means

$$\widehat{Enrolled}_{it} = 0.406 - 0.105Georgia_i - 0.004After_t + 0.089Georgia_i \times After_t$$

- A group mean for a dummy Y is $E[Y = 1]$, i.e. the probability a student is enrolled:

Diff-in-Diff Example: Comparing Group Means

$$\widehat{Enrolled}_{it} = 0.406 - 0.105Georgia_i - 0.004After_t + 0.089Georgia_i \times After_t$$

- A group mean for a dummy Y is $E[Y = 1]$, i.e. the probability a student is enrolled:
  - **Non-Georgian enrollment probability pre-1992**: $\beta_0 = 0.406$

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_i - 0.004 After_t + 0.089 Georgia_i \times After_t$$

- A group mean for a dummy Y is $E[Y = 1]$, i.e. the probability a student is enrolled:
  - **Non-Georgian enrollment probability pre-1992**: $\beta_0 = 0.406$
    - **Georgian enrollment probability pre-1992**: $\beta_0 + \beta_1 = 0.406 - 0.105 = 0.301$

Diff-in-Diff Example: Comparing Group Means

$$\widehat{Enrolled}_{it} = 0.406 - 0.105Georgia_i - 0.004After_t + 0.089Georgia_i \times After_t$$

- A group mean for a dummy Y is $E[Y = 1]$, i.e. the probability a student is enrolled:
    - **Non-Georgian enrollment probability pre-1992**: $\beta_0 = 0.406$
        - **Georgian enrollment probability pre-1992**: $\beta_0 + \beta_1 = 0.406 - 0.105 = 0.301$
        - **Non-Georgian enrollment probability post-1992**: $\beta_0 + \beta_2 = 0.406 - 0.004 = 0.402$

HOOD COLLEGE

Diff-in-Diff Example: Comparing Group Means

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_i - 0.004 After_t + 0.089 Georgia_i \times After_t$$

- A group mean for a dummy Y is $E[Y=1]$, i.e. the probability a student is enrolled:
  - **Non-Georgian enrollment probability pre-1992**: $\beta_0 = 0.406$
    - **Georgian enrollment probability pre-1992**: $\beta_0 + \beta_1 = 0.406 - 0.105 = 0.301$
    - **Non-Georgian enrollment probability post-1992**: $\beta_0 + \beta_2 = 0.406 - 0.004 = 0.402$
    - **Georgian enrollment probability post-1992**: $\beta_0 + \beta_1 + \beta_2 + \beta_3 = 0.406 - 0.105 - 0.004 + 0.089 = 0.386$

Diff-in-Diff Example: Comparing Group Means in R

```r
# group mean for non-Georgian before 1992
HOPE %>%
  filter(Georgia==0 & After==0) %>%
  summarize(prob=mean(InCollege))
```

```
##        prob
## 1 0.4057827
```

```r
# group mean for Georgian before 1992
HOPE %>%
  filter(Georgia==1 & After==0) %>%
  summarize(prob=mean(InCollege))
```

```
##        prob
## 1 0.3005464
```

```r
# group mean for non-Georgian AFTER 1992
HOPE %>%
  filter(Georgia==0 & After==1) %>%
  summarize(prob=mean(InCollege))
```

```
##       prob
## 1 0.401323
```

```r
# group mean for Georgian AFTER 1992
HOPE %>%
  filter(Georgia==1 & After==1) %>%
  summarize(prob=mean(InCollege))
```

```
##        prob
## 1 0.3854167
```

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_{it} - 0.004 After_{it} + 0.089 Georgia_{it} \times After_{it}$$

	Neighbors	Georgia	Group Diff. (ΔY_i)
Before	0.406	0.301	-0.105
After	0.402	0.386	0.016
Time Diff. (ΔY_t)	-0.004	0.085	0.089
			Diff-in-diff ($\Delta\Delta Y$)

$$\widehat{Enrolled}_{it} = 0.406 - 0.105 Georgia_{it} - 0.004 After_{it} + 0.089 Georgia_{it} \times After_{it}$$

	Neighbors	Georgia	Group Diff. (ΔY_i)
Before	0.406	0.301	-0.105
After	0.402	0.386	0.016
Time Diff. (ΔY_t)	-0.004	0.085	0.089
			Diff-in-diff ($\Delta\Delta Y$)

$$\begin{aligned}
\Delta\Delta Enrolled &= (\text{GA}_{after} - \text{GA}_{before}) - (\text{neighbors}_{after} - \text{neighbors}_{before}) \\
&= (0.386 - 0.301) - (0.402 - 0.406) \\
&= (0.085) - (-0.004) \\
&= 0.089
\end{aligned}$$