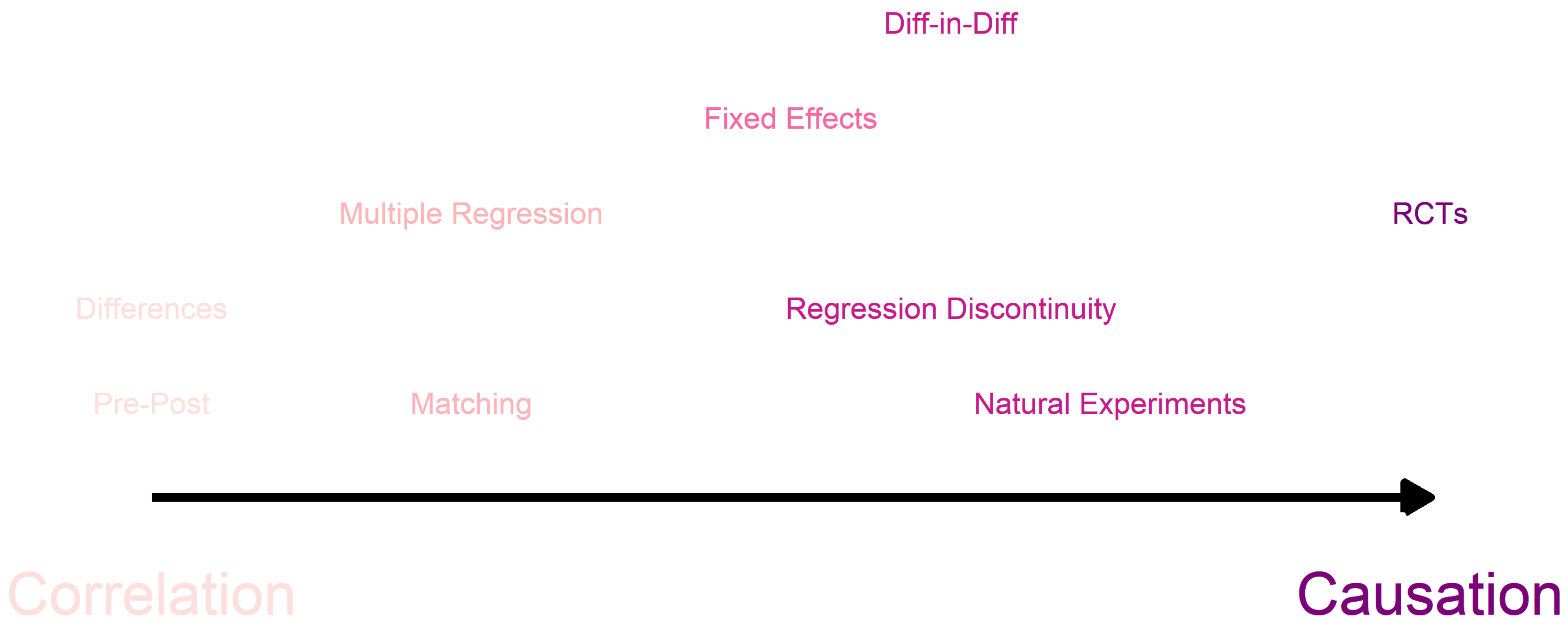
QM — Difference in Differences Estimators

Dr. Zahid Asghar
Professor of Economics
zasghar@qau.edu.pk
zahedasghar
zahidasghar.com

Contents

Clever Research Designs Identify Causality

Again, **this toolkit** of research designs to **identify causal effects** is the economist's **comparative advantage** that firms and governments want!



Difference-in-Differences Models

Natural Experiments

Difference-in-Differences Models I

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

Difference-in-Differences Models I

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

Example

- How do States that implement policy X see changes in Y
  - **Treatment**: States that implement X
  - **Control**: States that did not implement X

- If we have **panel data** with observations for all states **before** and **after** the change...
- Find the *difference* between treatment & control groups *in* their *differences* before and after the treatment period

Difference-in-Differences Models I

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

Example

- How do States that implement policy X see changes in Y
  - **Treatment**: States that implement X
  - **Control**: States that did not implement X

- If we have **panel data** with observations for all states **before** and **after** the change...
- Find the *difference* between treatment & control groups *in* their *differences* before and after the treatment period

Difference-in-Differences Models II

- The **difference-in-differences** (aka **"diff-in-diff"** or **"DND"**) estimator identifies treatment effect by differencing the difference pre- and post-treatment values of Y between treatment and control groups

$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$

- $\text{Treated}_i = \begin{cases} 1 \text{ if } i \text{ is in treatment group} \\ 0 \text{ if } i \text{ is not in treatment group} \end{cases}$ $\quad \text{After}_t = \begin{cases} 1 \text{ if } t \text{ is after treatment period} \\ 0 \text{ if } t \text{ is before treatment period} \end{cases}$

	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** $(\Delta_i \Delta_t)$

Example: Hot Dogs

- Is there a discount when you get cheese *and* chili?

```
# A tibble: 4 × 3
  price cheese chili
  <dbl>  <dbl> <dbl>
1  2         0     0
2  2.35      1     0
3  2.35      0     1
4  2.7       1     1
```

Example: Hot Dogs

- Is there a discount when you get cheese *and* chili?

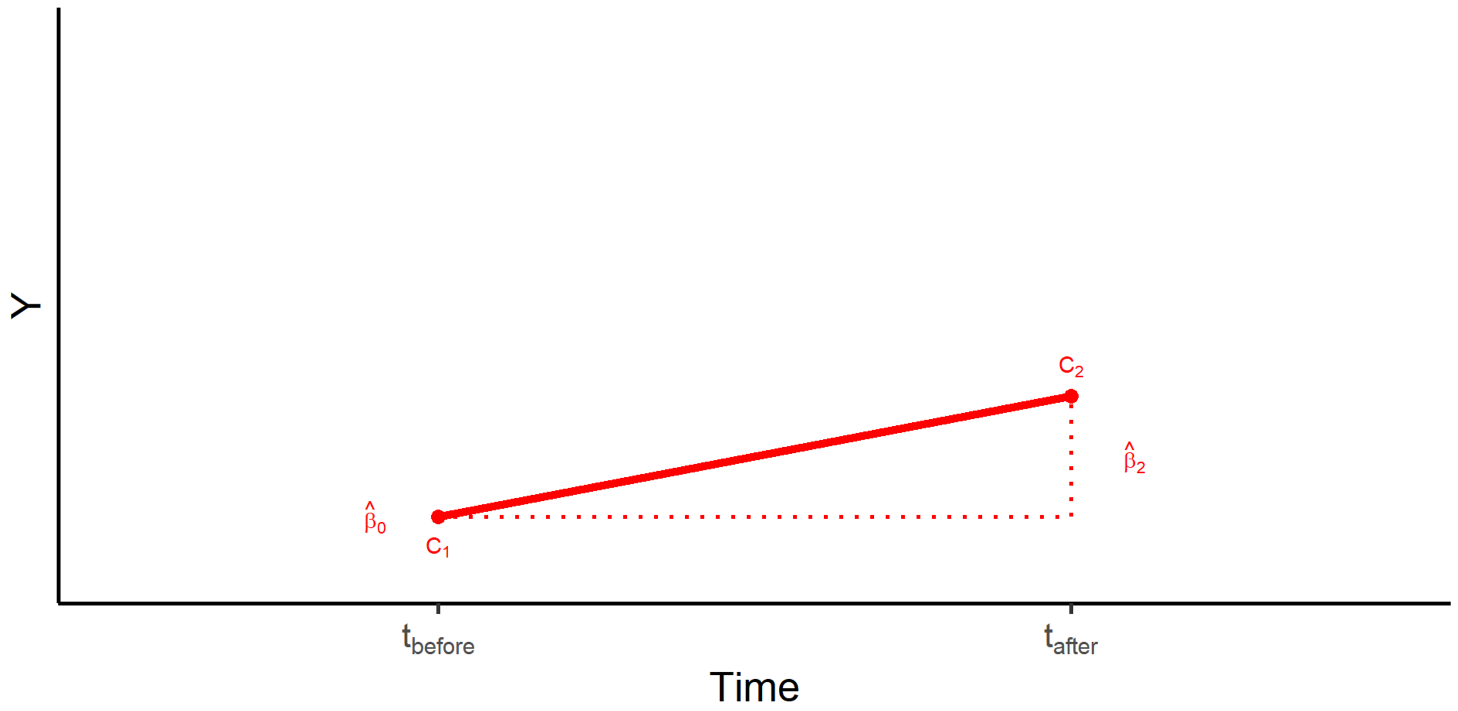
```r
lm(price ~ cheese + chili + cheese*chili,
   data = hotdogs) %>%
  tidy()
```

```
# A tibble: 4 × 2
  term          estimate
  <chr>            <dbl>
1 (Intercept)       2
2 cheese            0.35
3 chili             0.35
4 cheese:chili      0
```

- Diff-n-diff is just a model with an interaction term between two dummies!

Visualizing Diff-in-Diff

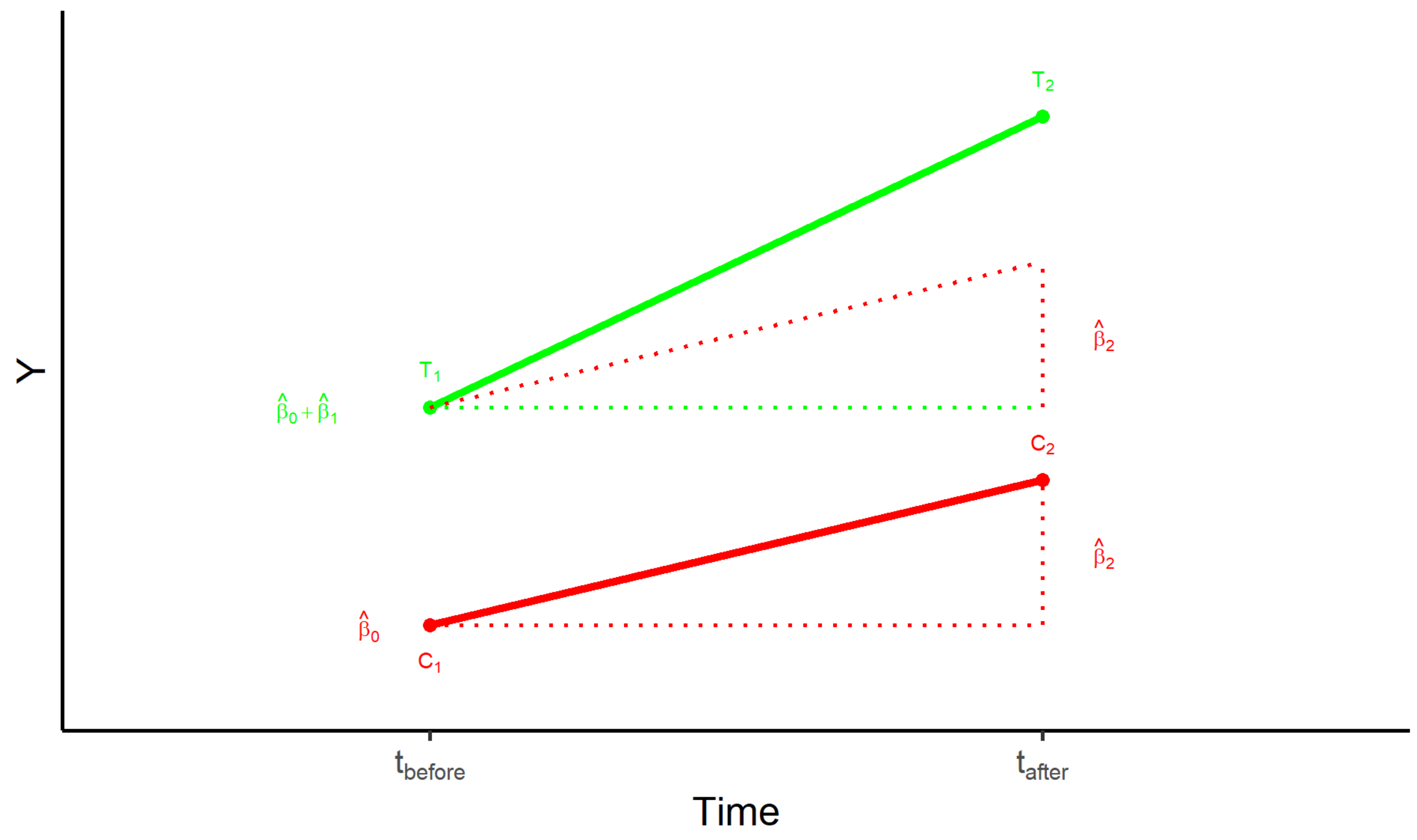
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i \times \text{After}_t) + u_{it}$$



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

Visualizing Diff-in-Diff

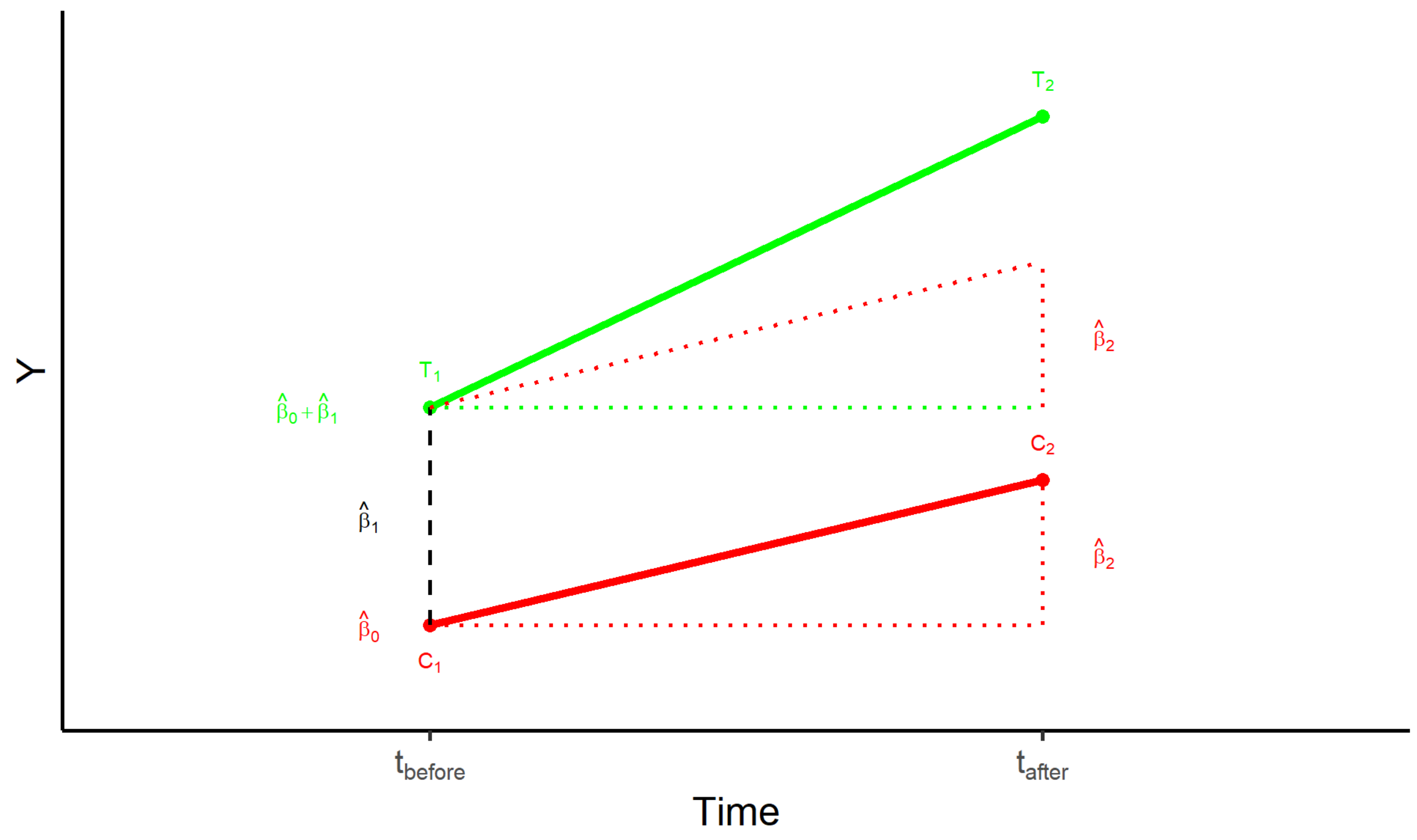
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i \times \text{After}_t) + u_{it}$$



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

Visualizing Diff-in-Diff

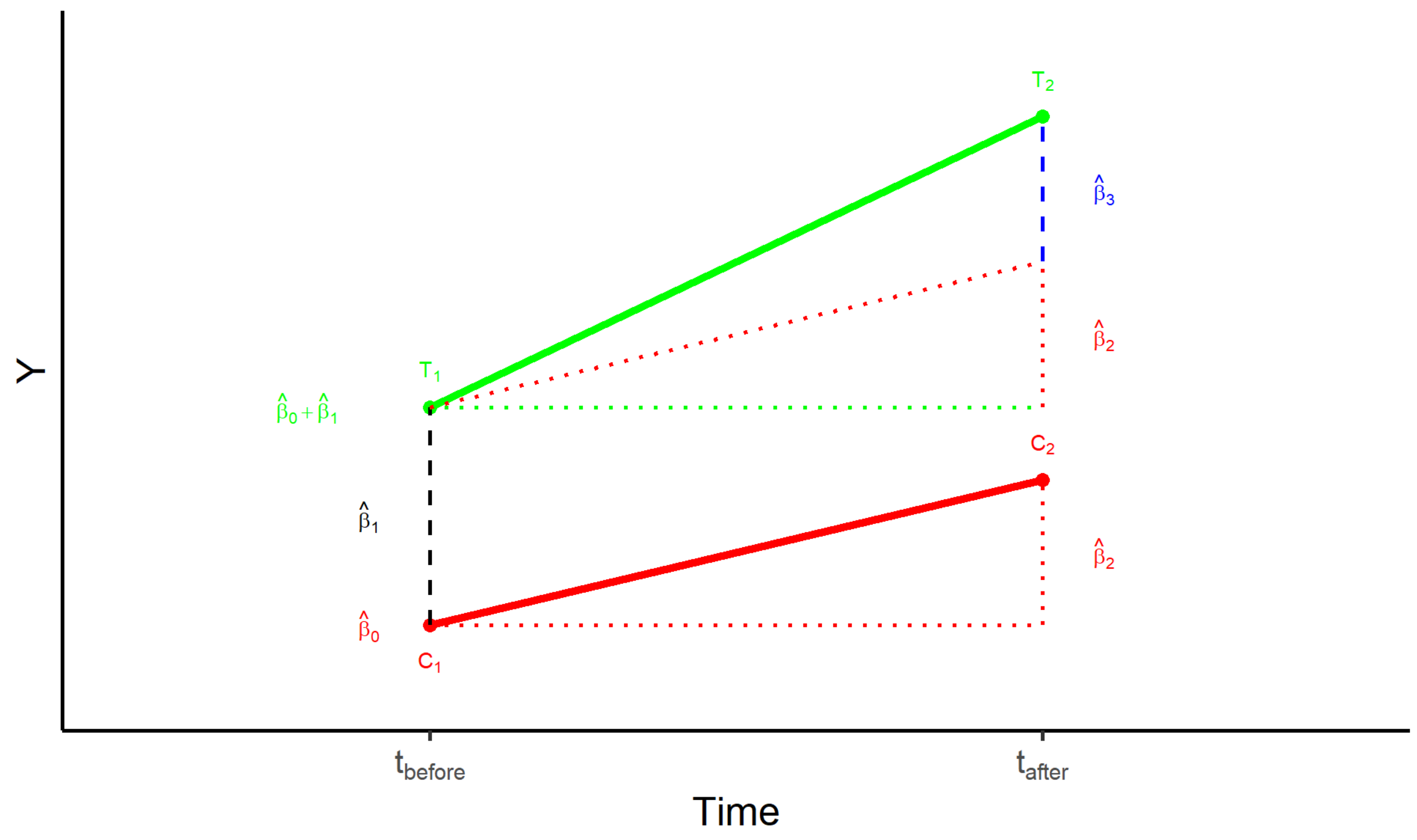
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i \times \text{After}_t) + u_{it}$$



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

Visualizing Diff-in-Diff

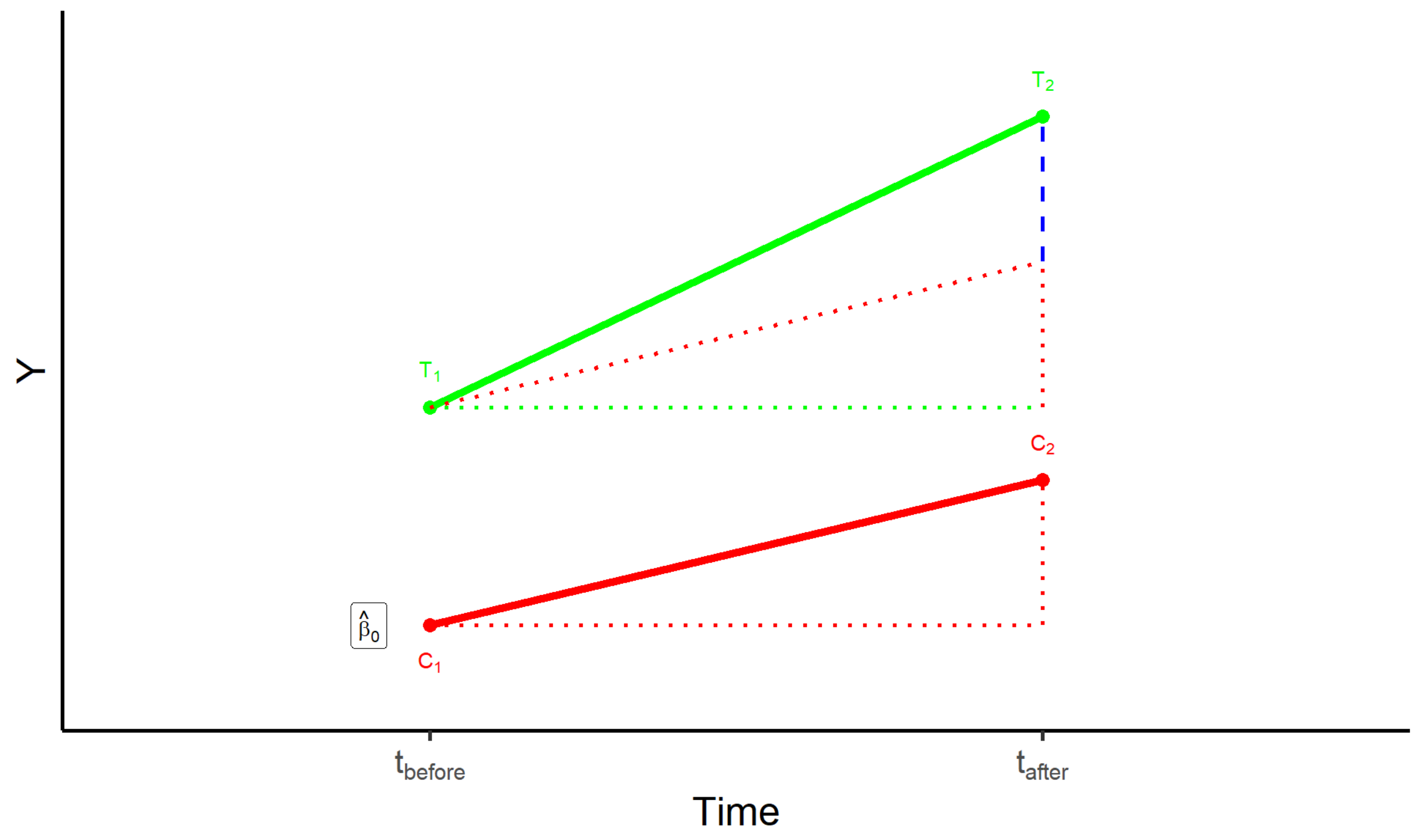
$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$



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

Visualizing Diff-in-Diff II

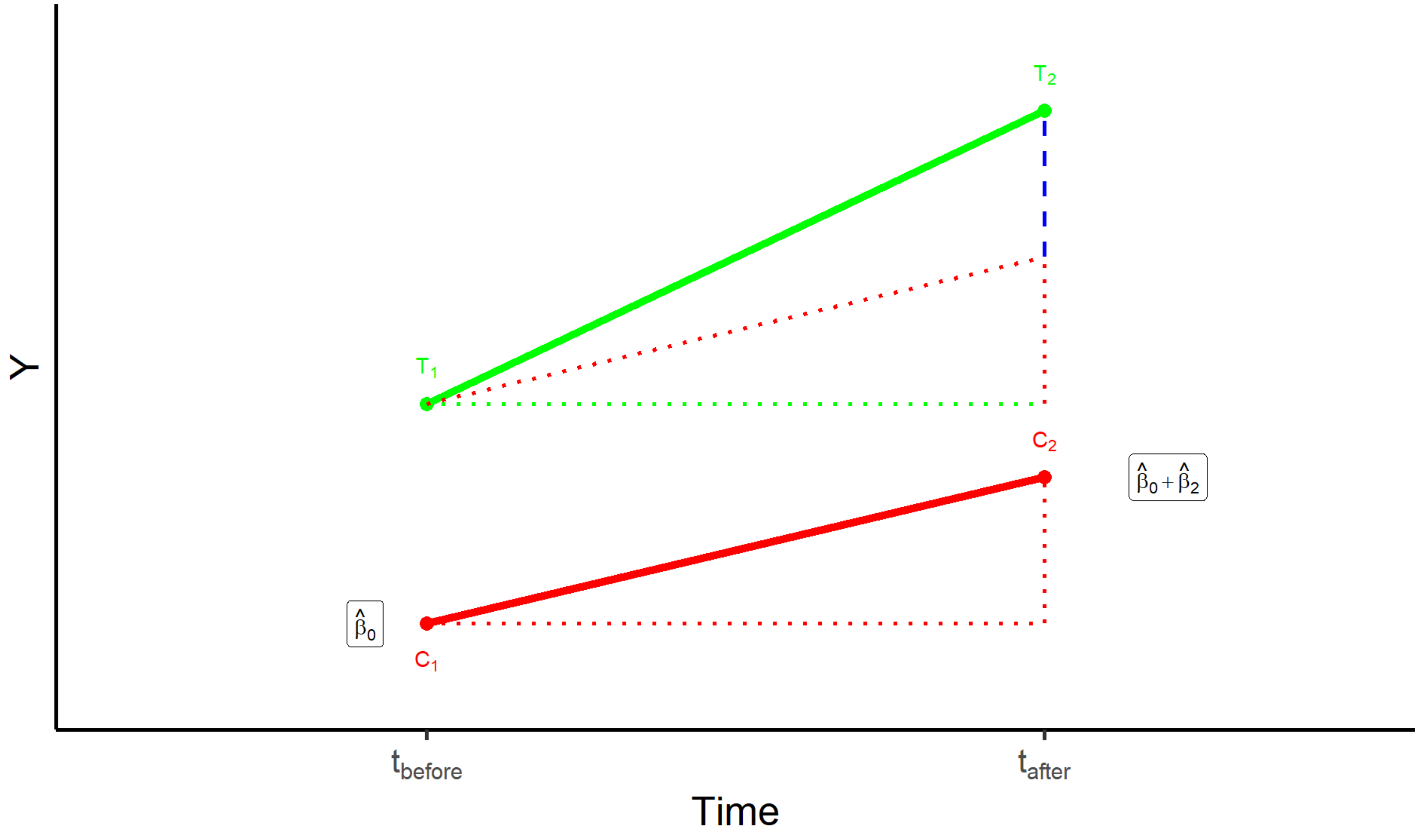
$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$

Visualizing Diff-in-Diff II

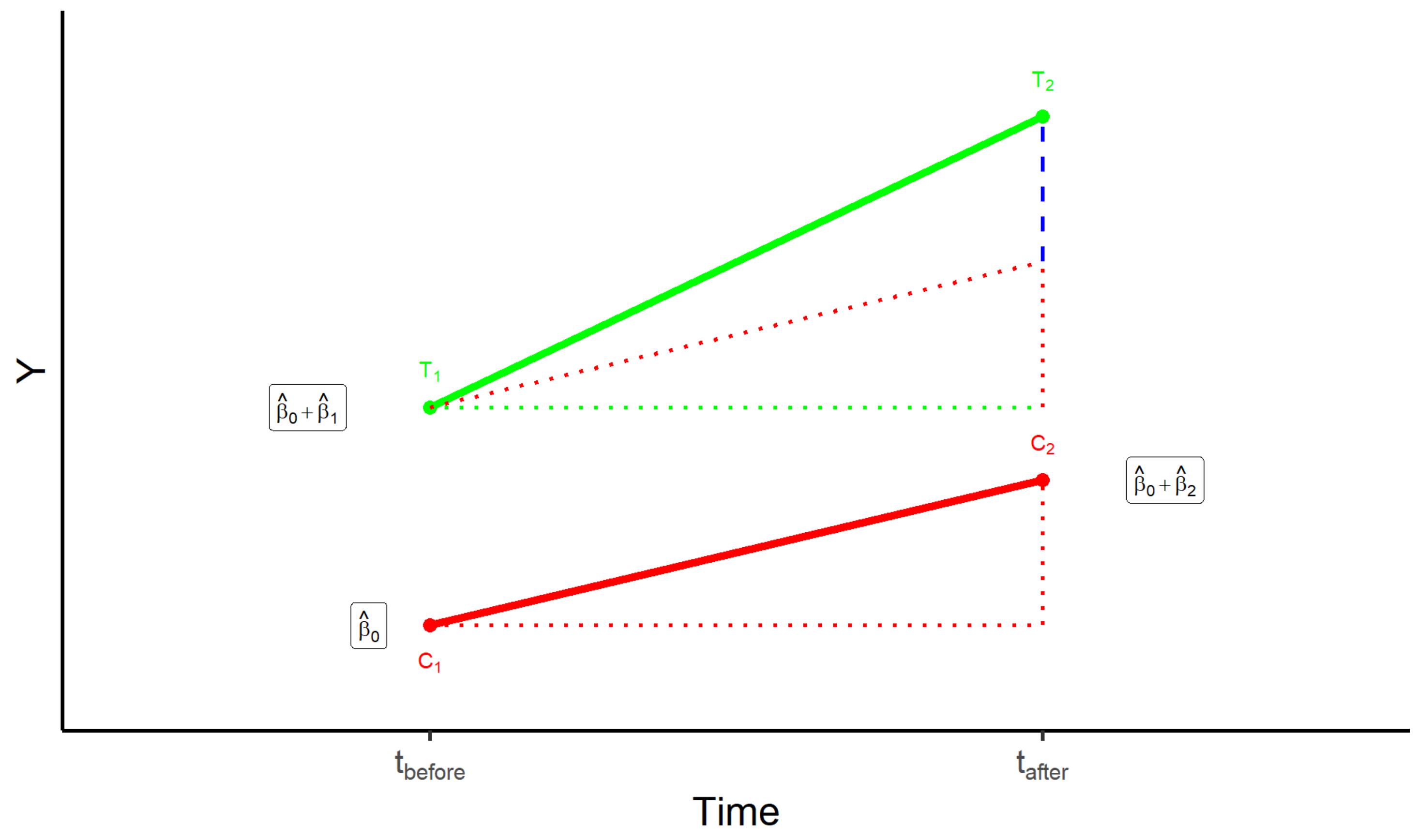
$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$
- $\bar{Y}_i$ for **Control** group **after**: $\hat{\beta}_0 + \hat{\beta}_2$

Visualizing Diff-in-Diff II

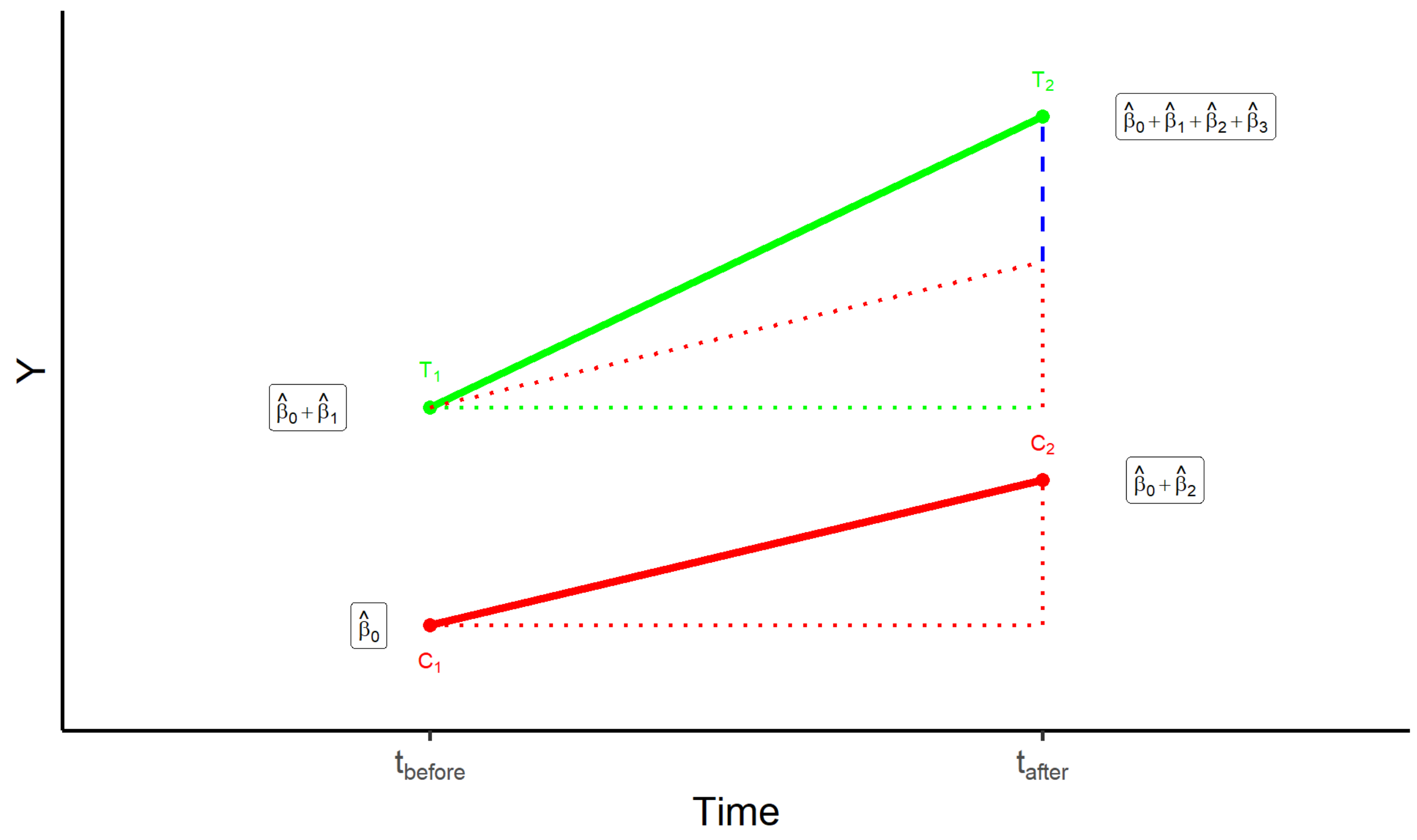
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i \times \text{After}_t) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$
- $\bar{Y}_i$ for **Control** group **after**: $\hat{\beta}_0 + \hat{\beta}_2$
- $\bar{Y}_i$ for **Treatment** group **before**: $\hat{\beta}_0 + \hat{\beta}_1$

Visualizing Diff-in-Diff II

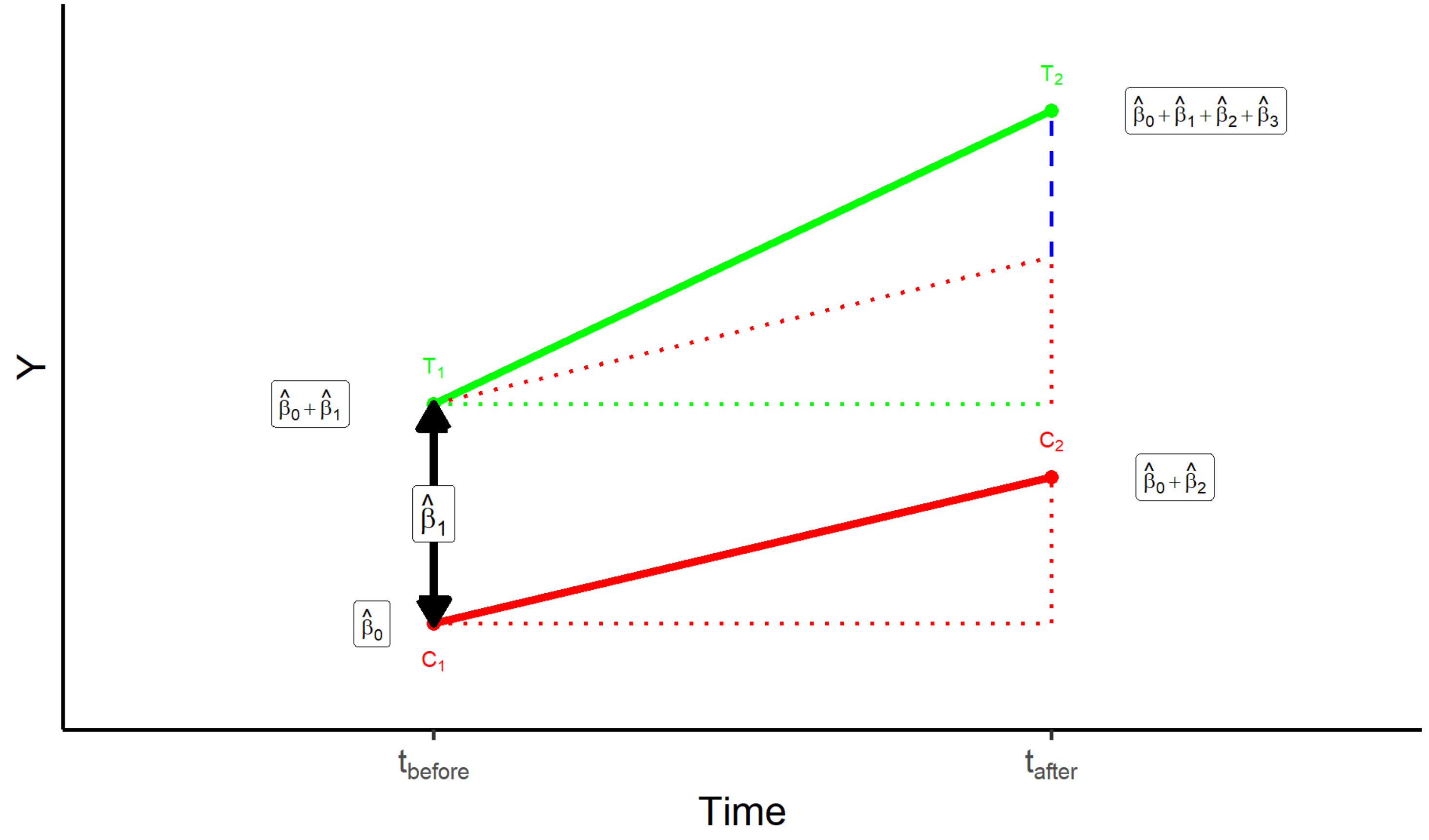
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treated}_i \times \text{After}_t) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$
- $\bar{Y}_i$ for **Control** group **after**: $\hat{\beta}_0 + \hat{\beta}_2$
- $\bar{Y}_i$ for **Treatment** group **before**: $\hat{\beta}_0 + \hat{\beta}_1$
- $\bar{Y}_i$ for **Treatment** group **after**: $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$

Visualizing Diff-in-Diff II

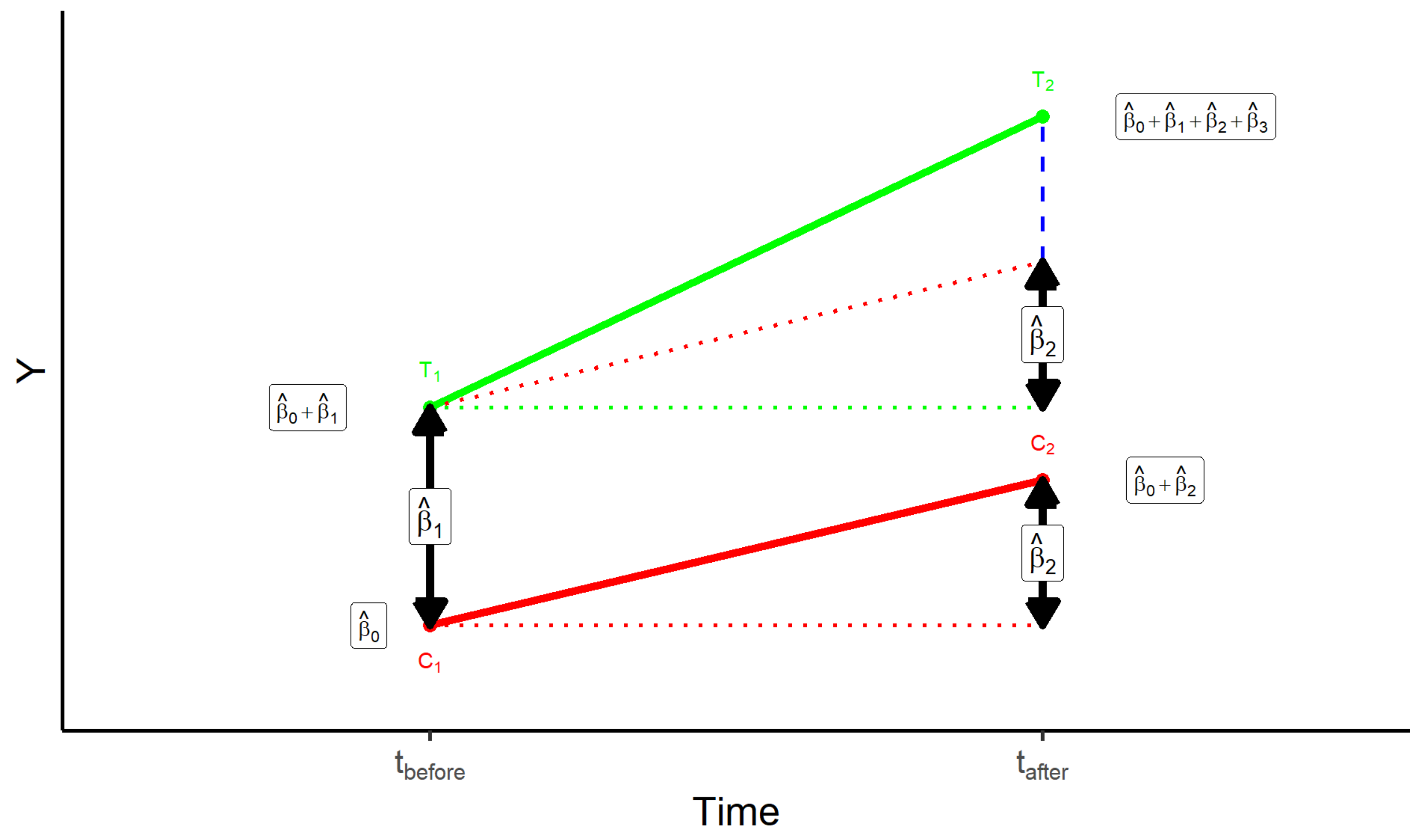
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{ Treated}_i + \beta_2 \text{ After}_t + \beta_3 \,(\text{Treated}_i \times \text{After}_t) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$
- $\bar{Y}_i$ for **Control** group **after**: $\hat{\beta}_0 + \hat{\beta}_2$
- $\bar{Y}_i$ for **Treatment** group **before**: $\hat{\beta}_0 + \hat{\beta}_1$
- $\bar{Y}_i$ for **Treatment** group **after**: $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$
- **Group Difference (before)**: $\hat{\beta}_1$

Visualizing Diff-in-Diff II

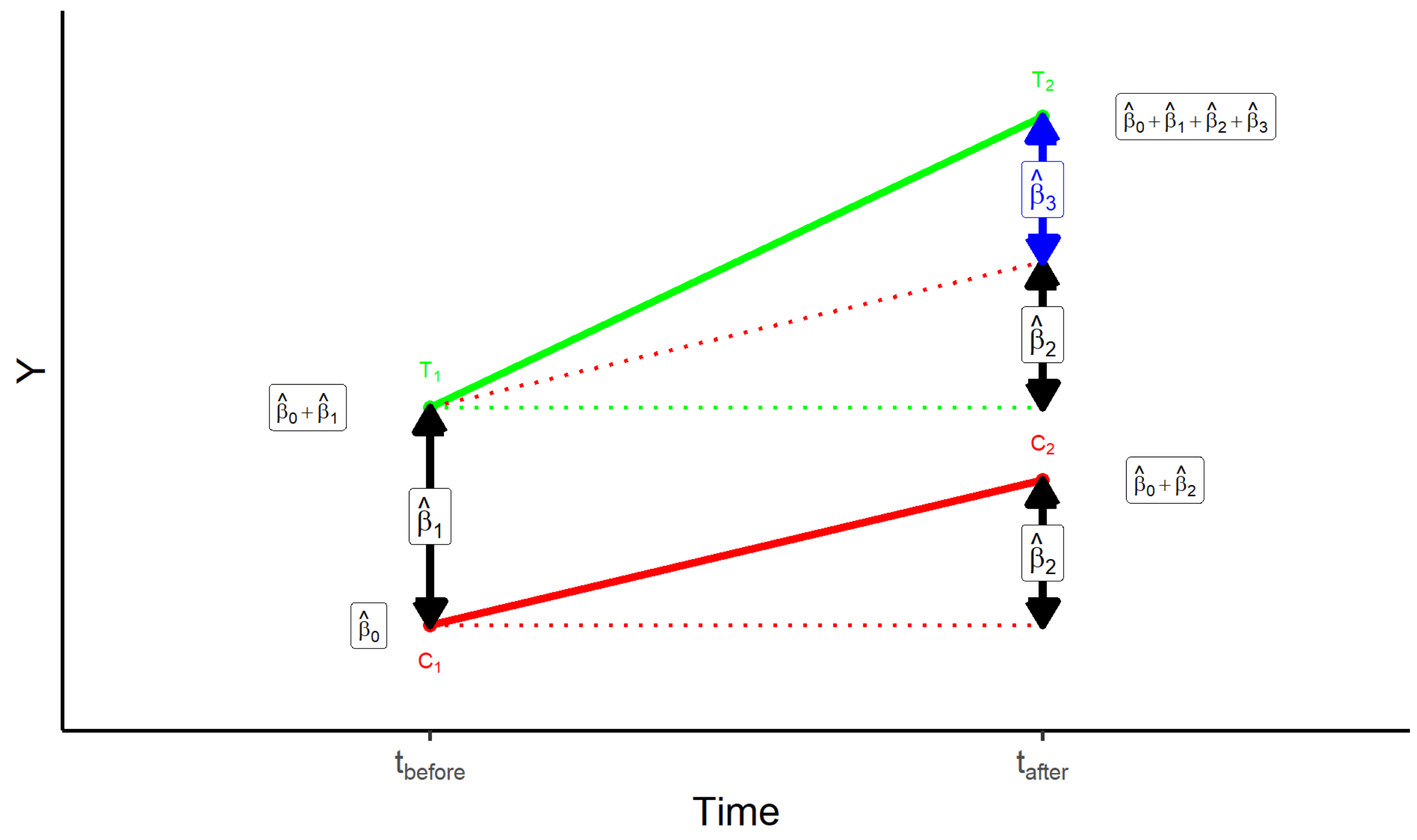
$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{ Treated}_i + \beta_2 \text{ After}_t + \beta_3 \left(\text{Treated}_i \times \text{After}_t\right) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$
- $\bar{Y}_i$ for **Control** group **after**: $\hat{\beta}_0 + \hat{\beta}_2$
- $\bar{Y}_i$ for **Treatment** group **before**: $\hat{\beta}_0 + \hat{\beta}_1$
- $\bar{Y}_i$ for **Treatment** group **after**: $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$
- **Group Difference (before)**: $\hat{\beta}_1$
- **Time Difference**: $\hat{\beta}_2$

Visualizing Diff-in-Diff II

$$\hat{Y}_{it} = \beta_0 + \beta_1 \text{ Treated}_i + \beta_2 \text{ After}_t + \beta_3 \left(\text{Treated}_i \times \text{After}_t\right) + u_{it}$$



- $\bar{Y}_i$ for **Control** group **before**: $\hat{\beta}_0$
- $\bar{Y}_i$ for **Control** group **after**: $\hat{\beta}_0 + \hat{\beta}_2$
- $\bar{Y}_i$ for **Treatment** group **before**: $\hat{\beta}_0 + \hat{\beta}_1$
- $\bar{Y}_i$ for **Treatment** group **after**: $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$
- **Group Difference (before)**: $\hat{\beta}_1$
- **Time Difference**: $\hat{\beta}_2$
- **Difference-in-differences**: $\hat{\beta}_3$ (treatment effect)

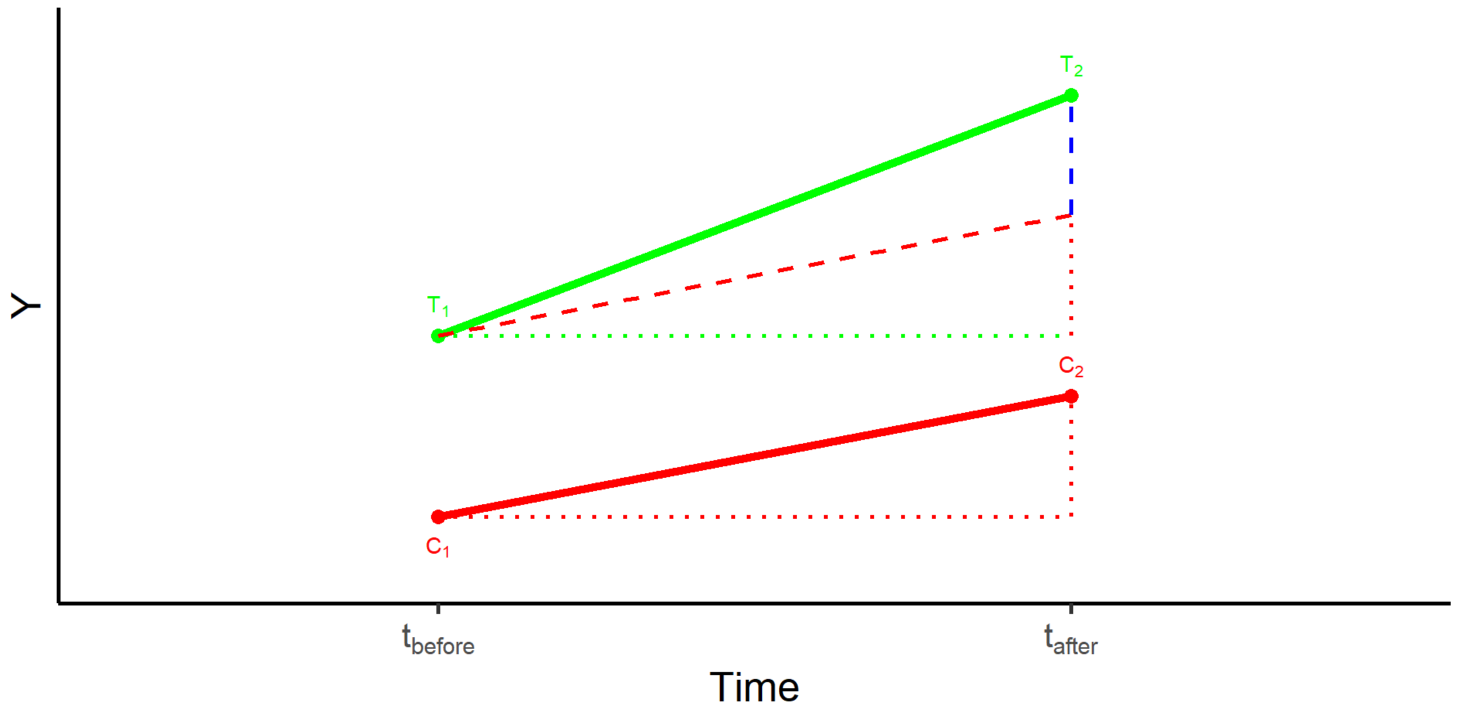
Comparing Group Means (Again)

$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$

	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	$\beta_2 + \beta_3$	**Diff-in-diff** $\Delta_i \Delta_t : \beta_3$

Key Assumption: Counterfactual

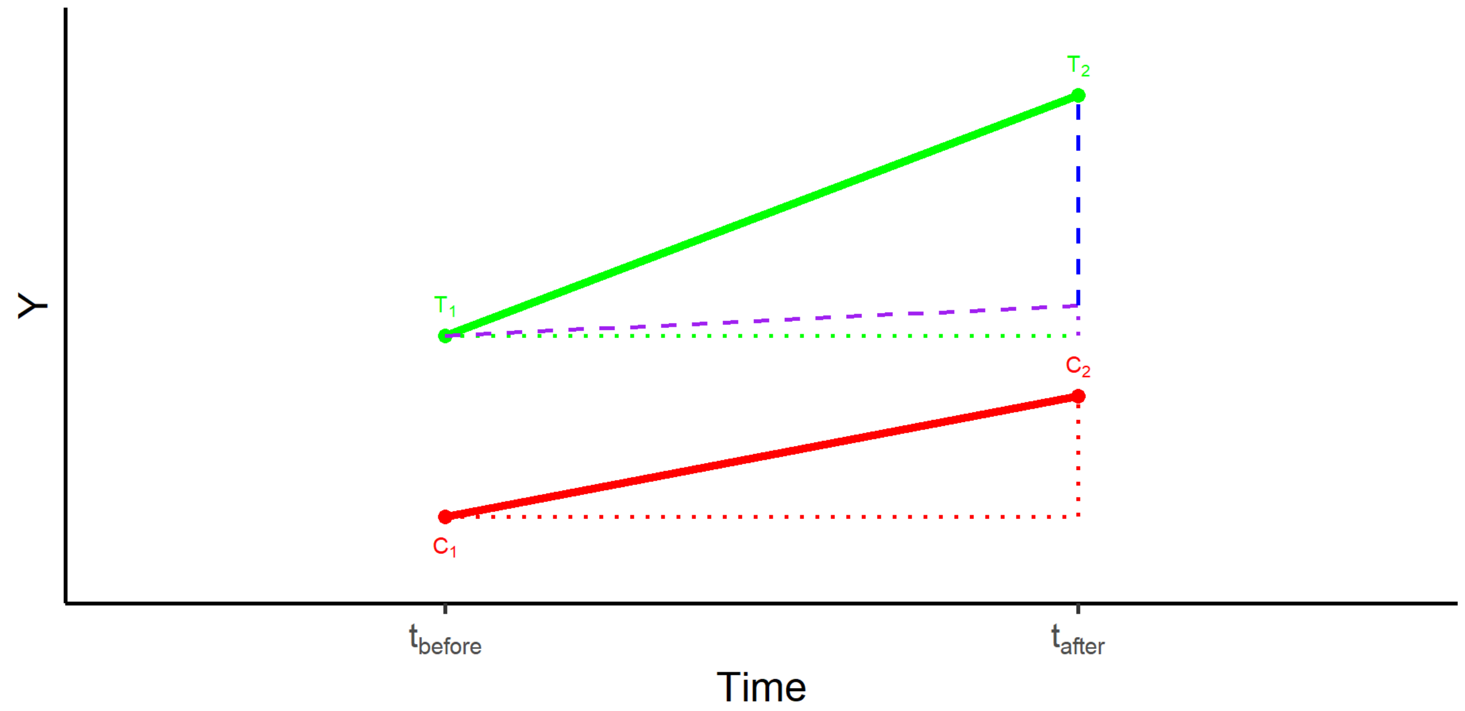
$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$



- Key assumption for DND: **time trends** (for treatment and control) are **parallel**
- Treatment and control groups assumed to be identical over time on average, **except for treatment**
- **Counterfactual**: if the treatment group had not recieved treatment, it would have changed identically over time as the control group ($\hat{\beta}_2$)

Key Assumption: Counterfactual

$$\hat{Y}_{it} = \beta_0 + \beta_1 \, \text{Treated}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{Treated}_i \times \text{After}_t) + u_{it}$$



- If the time-trends would have been *different*, a **biased** measure of the treatment effect ($\hat{\beta}_3$)!

Example I: HOPE in Georgia

Diff-in-Diff Example I

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?

- Micro-level data on 4,291 young individuals
- $\text{InCollege}_{it} = \begin{cases} 1 \text{ if } i \text{ is in college during year } t \\ 0 \text{ if } i \text{ is not in college during year } t \end{cases}$ [1]
- $\text{Georgia}_i = \begin{cases} 1 \text{ if } i \text{ is a Georgia resident} \\ 0 \text{ if } i \text{ is not a Georgia resident} \end{cases}$
- $\text{After}_t = \begin{cases} 1 \text{ if } t \text{ is after 1992} \\ 0 \text{ if } t \text{ is after 1992} \end{cases}$

1. Note: With a dummy *dependent* (Y) variable, coefficients estimate the probability $Y = 1$, i.e. the *probability* a person is enrolled in college

Diff-in-Diff Example II

- We can use a DND 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: after 1992
- Treatment: Georgia
- Difference-in-differences:

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

- Regression equation:

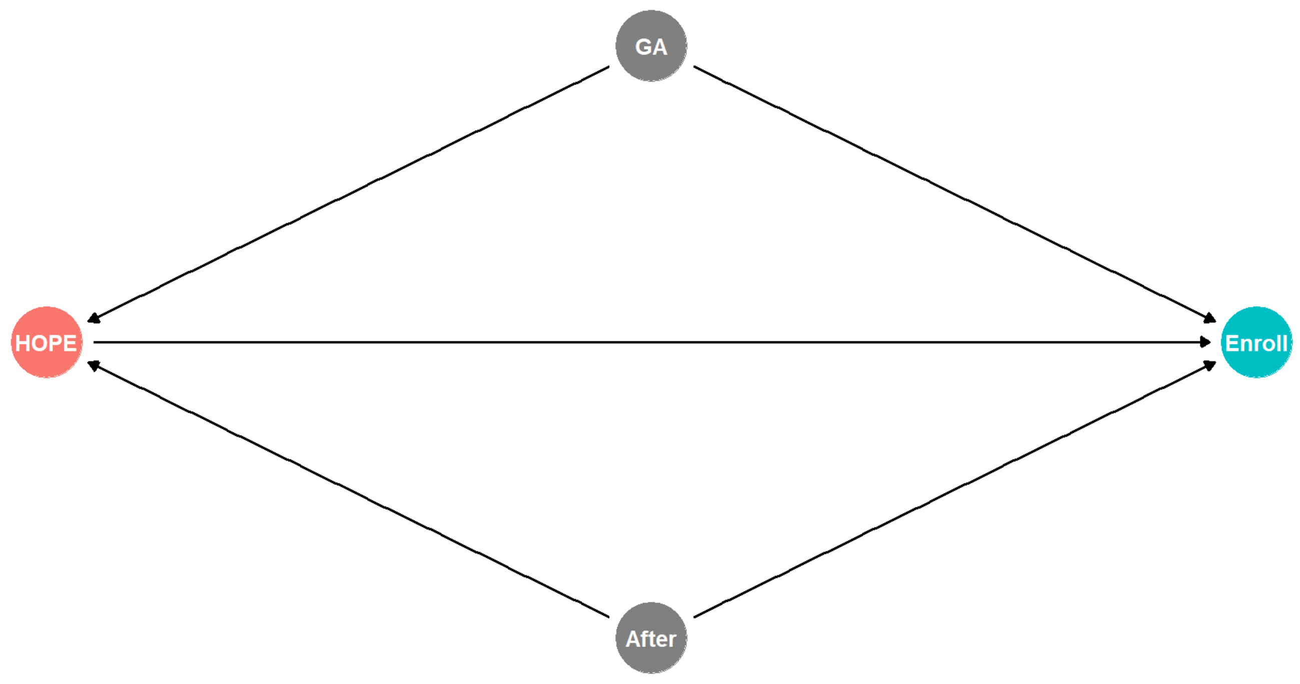
$$\widehat{\text{Enrolled}}_{it} = \beta_0 + \beta_1 \,\text{Georgia}_i + \beta_2 \,\text{After}_t + \beta_3 \,(\text{Georgia}_i \times \text{After}_t)$$

Example: Data

```
1  hope
```

```
# A tibble: 4,291 × 11
   StateC…¹   Age Year  Weight Age18 LowIn…² InCol…³ After Georgia After…⁴ Black
   <fct>    <dbl> <fct>  <dbl> <dbl>   <dbl>   <dbl> <dbl>   <dbl>   <dbl> <dbl>
 1 56          19 89      1396     0       1       1     0       0       0     0
 2 56          19 89      1080     0      NA       1     0       0       0     1
 3 56          18 89       924     1       1       1     0       0       0     0
 4 56          19 89       891     0       0       1     0       0       0     0
 5 56          19 89      1395     0      NA       0     0       0       0     0
 6 56          18 89      1106     1       1       1     0       0       0     1
 7 56          19 89       965     0      NA       0     0       0       0     0
 8 56          18 89       958     1      NA       0     0       0       0     0
 9 56          19 89      1006     0      NA       0     0       0       0     0
10 56          19 89      1183     0       1       1     0       0       0     0
# … with 4,281 more rows, and abbreviated variable names ¹StateCode,
```

Example: Data



The effect of HOPE is identified by differences between Georgia and the rest of the southeastern United States in the time pattern of college attendance rates. I use difference–in–differences estimation, comparing attendance rates before and after HOPE was introduced, within Georgia and in the rest of the region. This calculation can be made using ordinary least squares:

$$[7] \quad y_i = \alpha_1 + \beta_1(Georgia_i * After_i) + \delta_1 Georgia_i + \theta_1 After_i + v_{i1}$$

where the dependent variable is a binary measure of college attendance, $Georgia_i$ is a binary variable that is set to one if a youth is a Georgia resident and $After_i$ is a

Dynarski, Susan, 1999, "Hope for Whom? Financial Aid for the Middle Class and its Impact on College Attendance," *National Tax Journal* 53(3): 629-661

Example: Regression

```r
DND_reg <- lm(InCollege ~ Georgia + After + Georgia*After, data = hope)
DND_reg %>% tidy()
```

```
# A tibble: 4 × 5
  term          estimate std.error statistic   p.value
  <chr>            <dbl>     <dbl>     <dbl>     <dbl>
1 (Intercept)    0.406      0.0109    37.1   4.22e-262
2 Georgia       -0.105      0.0378    -2.79  5.37e-  3
3 After         -0.00446    0.0159    -0.281 7.78e-  1
4 Georgia:After  0.0893     0.0489     1.83  6.78e-  2
```

$$\widehat{\text{Enrolled}}_{it} = 0.406 - 0.105\,\text{Georgia}_i - 0.004\,\text{After}_t + 0.089\,(\text{Georgia}_i \times \text{After}_t)$$

Example: Interpretting the Regression

$$\widehat{\text{Enrolled}}_{it} = 0.406 - 0.105\,\text{Georgia}_i - 0.004\,\text{After}_t + 0.089\,(\text{Georgia}_i \times \text{After}_t)$$

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

Example: Comparing Group Means

$$\widehat{\text{Enrolled}}_{it} = 0.406 - 0.105\,\text{Georgia}_i - 0.004\,\text{After}_t + 0.089\,(\text{Georgia}_i \times \text{After}_t)$$

- A group mean for a dummy Y is $\mathbb{E}[Y=1]$, 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**:
  $\beta_0 + \beta_1 + \beta_2 + \beta_3 = 0.406 - 0.105 - 0.004 + 0.089 = 0.386$

Example: Comparing Group Means in R

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

```
# A tibble: 1 × 1
   prob
  <dbl>
1 0.406
```

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

```
# A tibble: 1 × 1
   prob
  <dbl>
1 0.401
```

Example: Comparing Group Means in R

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

```
# A tibble: 1 × 1
   prob
  <dbl>
1 0.301
```

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

```
# A tibble: 1 × 1
   prob
  <dbl>
1 0.385
```

Example: Diff-in-Diff Summary

$$\widehat{\text{Enrolled}}_{it} = 0.406 - 0.105\,\text{Georgia}_i - 0.004\,\text{After}_t + 0.089\,(\text{Georgia}_i \times \text{After}_t)$$

$\dots$

	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	**Diff-in-diff**: 0.089

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

Diff-in-Diff Summary & Data

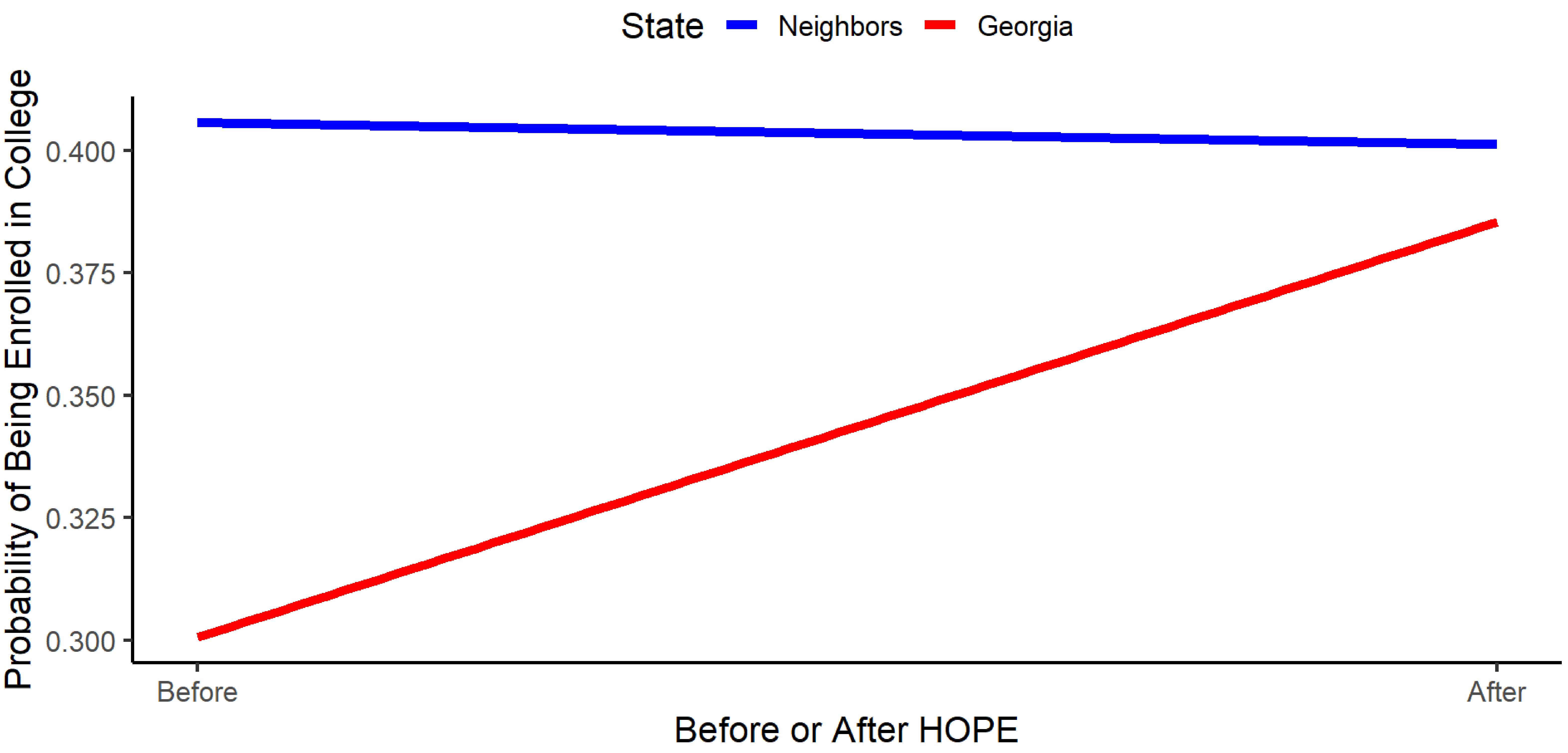
TABLE 2
DIFFERENCE–IN–DIFFERENCES
SHARE OF 18–19–YEAR–OLDS ATTENDING COLLEGE
OCTOBER CPS, 1989–97

	Before 1993	1993 and After	Difference
Georgia	0.300	0.378	0.078
Rest of Southeastern States	0.415	0.414	–0.001
Difference	0.115	0.036	0.079

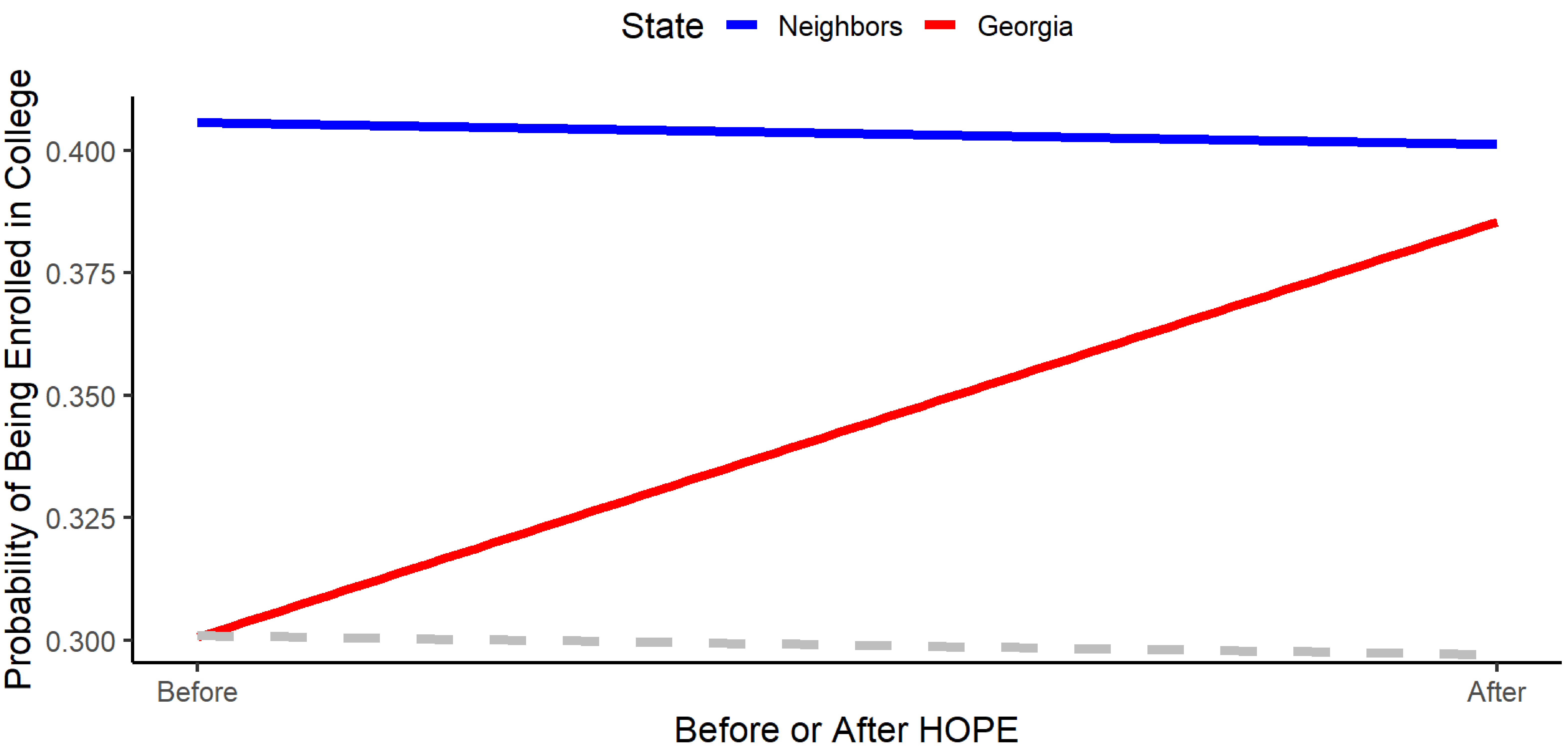
Note: Means are weighted by CPS sample weights. The Southeastern states are defined in the note to Table 1.

Dynarski, Susan, 1999, "Hope for Whom? Financial Aid for the Middle Class and its Impact on College Attendance," *National Tax Journal* 53(3): 629-661

Example: Diff-in-Diff Graph



Example: Diff-in-Diff Graph



Generalizing DND Models

Generalizing DND Models

- DND can be **generalized** with a **two-way fixed effects** model:

$$\hat{Y}_{it} = \beta_1 \left(\text{Treated}_i \times \text{After}_t\right) + \alpha_i + \theta_t + \nu_{it}$$

  - α_i: **group fixed effects** (treatments/control groups)
  - θ_t: **time fixed effects** (pre/post treatment)
  - β_1: diff-in-diff (interaction effect, β_3 from before)

- Flexibility: *many* periods (not just before/after), *many* different treatment(s)/groups, and treatment(s) can occur at different times to different units (so long as some do not get treated)

- Can also add control variables that vary within units and over time

$$\hat{Y}_{it} = \beta_1 \left(\text{Treated}_i \times \text{After}_t\right) + \beta_2 X_{it} + \cdots + \alpha_i + \theta_t + \nu_{it}$$

Our Example, Generalized I

$$\widehat{\text{Enrolled}}_{it} = \beta_1 \left(\text{Georgia}_i \times \text{After}_t\right) + \alpha_i + \theta_t +$$

- `StateCode` is a variable for the State $\implies$ create State fixed effect (α_i)
- `Year` is a variable for the year $\implies$ create year fixed effect (θ_t)

Our Example, Generalized II

Using LSDV method:

```r
DND_fe <- lm(InCollege ~ Georgia*After + factor(StateCode) + factor(Year),
             data = hope)
DND_fe %>% tidy()
```

```
# A tibble: 17 × 5
   term                 estimate std.error statistic  p.value
   <chr>                   <dbl>     <dbl>     <dbl>    <dbl>
 1 (Intercept)           0.418      0.0226    18.5    1.73e-73
 2 Georgia              -0.142      0.0394    -3.59   3.28e- 4
 3 After                 0.0753     0.0313     2.41   1.61e- 2
 4 factor(StateCode)57 -0.0142      0.0274    -0.518  6.05e- 1
 5 factor(StateCode)58 NA          NA         NA     NA
 6 factor(StateCode)59 -0.0624      0.0195    -3.19   1.42e- 3
 7 factor(StateCode)62 -0.133       0.0281    -4.73   2.35e- 6
 8 factor(StateCode)63 -0.00510     0.0263    -0.194  8.46e- 1
 9 factor(Year)90        0.0466     0.0283     1.64   1.00e- 1
10 factor(Year)91        0.0323     0.0286     1.13   2.59e- 1
11 factor(Year)92        0.0235     0.0298     0.789  4.30e- 1
12 factor(Year)93       -0.0452     0.0330    -1.37   1.71e- 1
13 factor(Year)94       -0.0608     0.0333    -1.83   6.79e- 2
14 factor(Year)95       -0.0786     0.0344    -2.29   2.22e- 2
15 factor(Year)96       -0.0967     0.0348    -2.78   5.50e- 3
16 factor(Year)97      NA          NA         NA     NA
17 Georgia:After         0.0914     0.0488     1.87   6.09e- 2
```

Our Example, Generalized II

Using `fixest`

```r
library(fixest)
DND_fe_2 <- feols(InCollege ~ Georgia*After | factor(StateCode) + factor(Year),
            data = hope)
DND_fe_2 %>% tidy()
```

```
# A tibble: 1 × 5
  term          estimate std.error statistic   p.value
  <chr>            <dbl>     <dbl>     <dbl>     <dbl>
1 Georgia:After   0.0914   0.00564      16.2 0.0000163
```

$$\widehat{\text{InCollege}}_{it} = 0.091\,(\text{Georgia}_i \times \text{After}_{it}) + \alpha_i + \theta_t$$

Our Example, Generalized, with Controls II

Using LSDV Method

```r
DND_fe_controls <- lm(InCollege ~ Georgia*After + factor(StateCode) + factor(Year) + Black + LowIncome,
          data = hope)
DND_fe_controls %>% tidy()
```

```
# A tibble: 19 × 5
   term                 estimate std.error statistic    p.value
   <chr>                   <dbl>     <dbl>     <dbl>      <dbl>
 1 (Intercept)           0.736      0.0299    24.6    1.16e-121
 2 Georgia              -0.109      0.0477    -2.29   2.23e-  2
 3 After                -0.00575    0.0374    -0.154  8.78e-  1
 4 factor(StateCode)57 -0.0434      0.0305    -1.42   1.54e-  1
 5 factor(StateCode)58 NA          NA         NA     NA
 6 factor(StateCode)59 -0.0532      0.0231    -2.31   2.12e-  2
 7 factor(StateCode)62 -0.116       0.0328    -3.54   4.12e-  4
 8 factor(StateCode)63  0.00739     0.0306     0.242  8.09e-  1
 9 factor(Year)90        0.0394     0.0333     1.18   2.37e-  1
10 factor(Year)91        0.0292     0.0335     0.873  3.83e-  1
11 factor(Year)92        0.0512     0.0348     1.47   1.41e-  1
12 factor(Year)93        0.0324     0.0388     0.835  4.04e-  1
13 factor(Year)94       -0.0193     0.0396    -0.486  6.27e-  1
14 factor(Year)95       -0.0492     0.0408    -1.21   2.28e-  1
15 factor(Year)96       -0.0713     0.0410    -1.74   8.20e-  2
16 factor(Year)97       NA          NA         NA     NA
17 Black                -0.0940     0.0206    -4.57   5.06e-  6
```

Our Example, Generalized, with Controls II

Using `fixest`

```r
DND_fe_controls_2 <- feols(InCollege ~ Georgia*After + Black + LowIncome | factor(StateCode) + factor(Year),
            data = hope)
DND_fe_controls_2 %>% tidy()
```

```
# A tibble: 3 × 5
  term          estimate std.error statistic  p.value
  <chr>            <dbl>     <dbl>     <dbl>    <dbl>
1 Black          -0.0940    0.0127     -7.38 0.000717
2 LowIncome      -0.302     0.0307     -9.84 0.000185
3 Georgia:After   0.0234    0.0128      1.83 0.127
```

$$\widehat{\text{InCollege}}_{it} = 0.023\,(\text{Georgia}_i \times \text{After}_{it}) - 0.094\,\text{Black}_{it} - 0.302\,\text{LowIncome}_{it}$$

Our Example, Generalized, with Controls III

	No FE	TWFE	TWFE
Constant	0.40578***		
	(0.01092)		
Georgia	−0.10524***		
	(0.03778)		
After	−0.00446		
	(0.01585)		
Georgia x After	0.08933*	0.09142***	0.02344
	(0.04889)	(0.00564)	(0.01282)
Black			−0.09399***
			(0.01273)
LowIncome			−0.30172***
			(0.03066)
n	4291	4291	2967
Adj. R^2	0.00		
SER	0.49	0.49	0.47

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The Findings

TABLE 3
COLLEGE ATTENDANCE OF 18–19–YEAR–OLDS
OCTOBER CPS, 1989–97
CONTROL GROUP: SOUTHEASTERN STATES

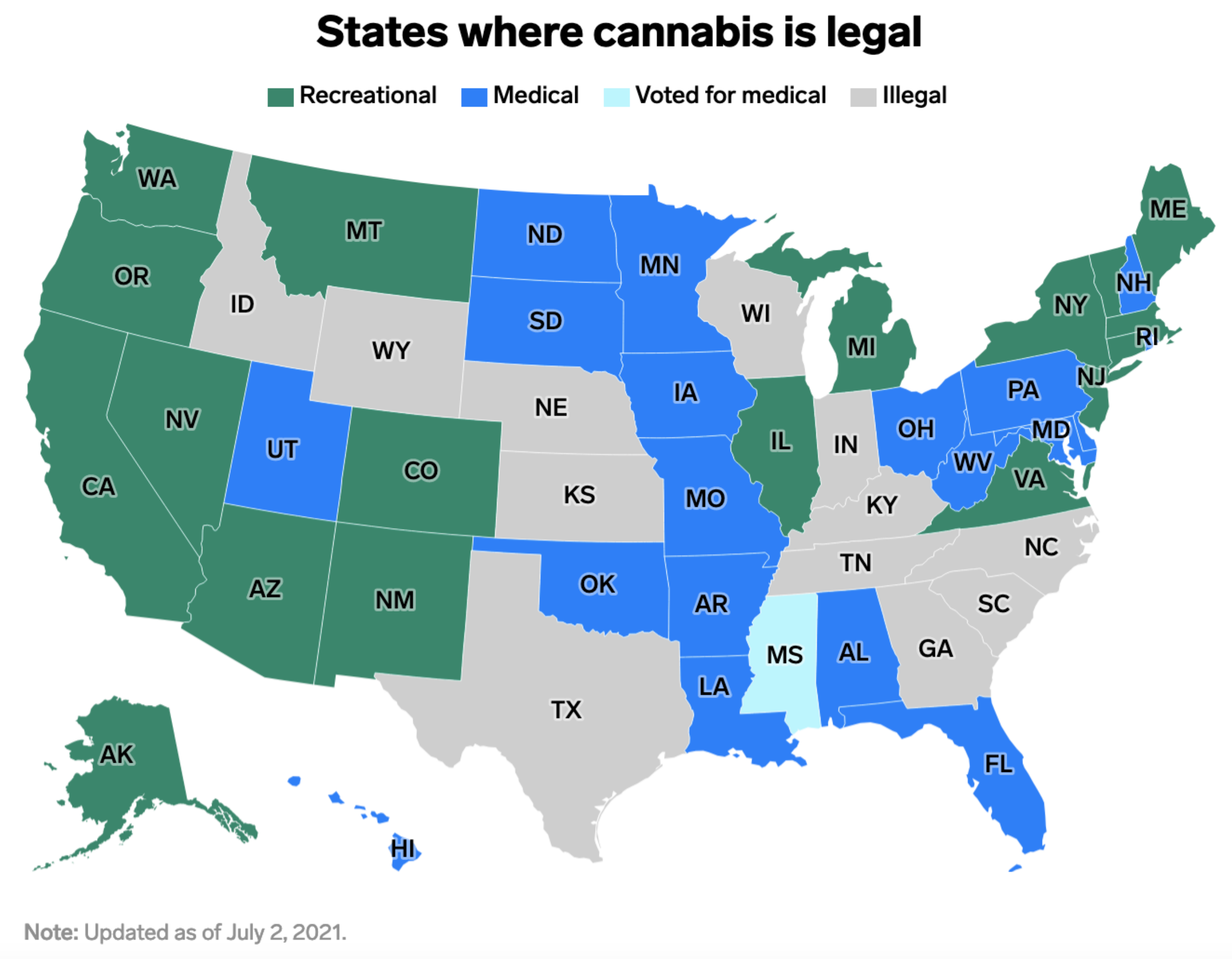
	(1) Difference–in–Differences	(2) Add Covariates	(3) Add Local Economic Conditions Controls
After*Georgia	0.079	0.075	0.070
	(0.029)	(0.030)	(0.030)
Georgia	–0.115	–0.100	–0.097
	(0.023)	(0.019)	(0.018)
After	–0.001		
	(0.018)		
Age 18		–0.042	–0.042
		(0.014)	(0.016)
Metro Resident		0.042	0.042
		(0.016)	(0.015)
Black		–0.134	–0.133
		(0.014)	(0.015)
State Unemployment Rate			0.005
			(0.007)
Year Dummies		Yes	Yes
R^2	0.003	0.023	0.023
N	6,811	6,811	6,811

Note: Regressions are weighted by CPS sample weights. Standard errors are adjusted for heteroskedasticity and correlation within state–year cells. The Southeastern states are defined in the note to Table 1.

Dynarski, Susan, 1999, “Hope for Whom? Financial Aid for the Middle Class and its Impact on College Attendance,” *National Tax Journal* 53(3): 629-661

Intuition behind DND

- Diff-in-diff models are the quintessential example of exploiting **natural experiments**
- A major change at a point in time (change in law, a natural disaster, political crisis) separates groups where one is affected and another is not—identifies the effect of the change (treatment)
- One of the cleanest and clearest causal **identification strategies**



Example II: “The” Card-Kreuger Minimum Wage Study

Example: “The” Card-Kreuger Minimum Wage Study I

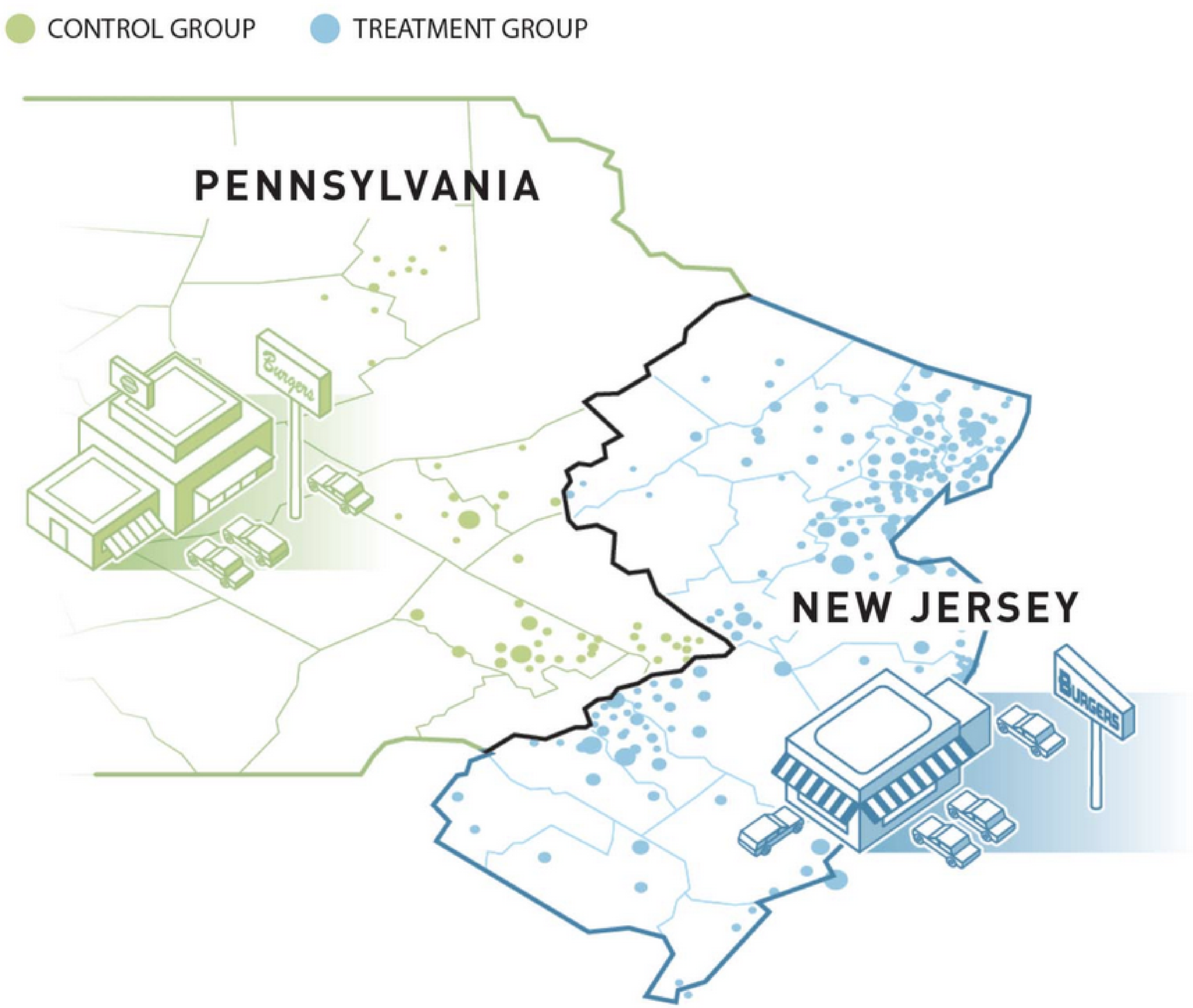
Example

The controversial minimum wage study, Card & Kreuger (1994) is a quintessential (and clever) diff-in-diff.]

Card, David, Krueger, Alan B, (1994), “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania,” *American Economic Review* 84 (4): 772–793

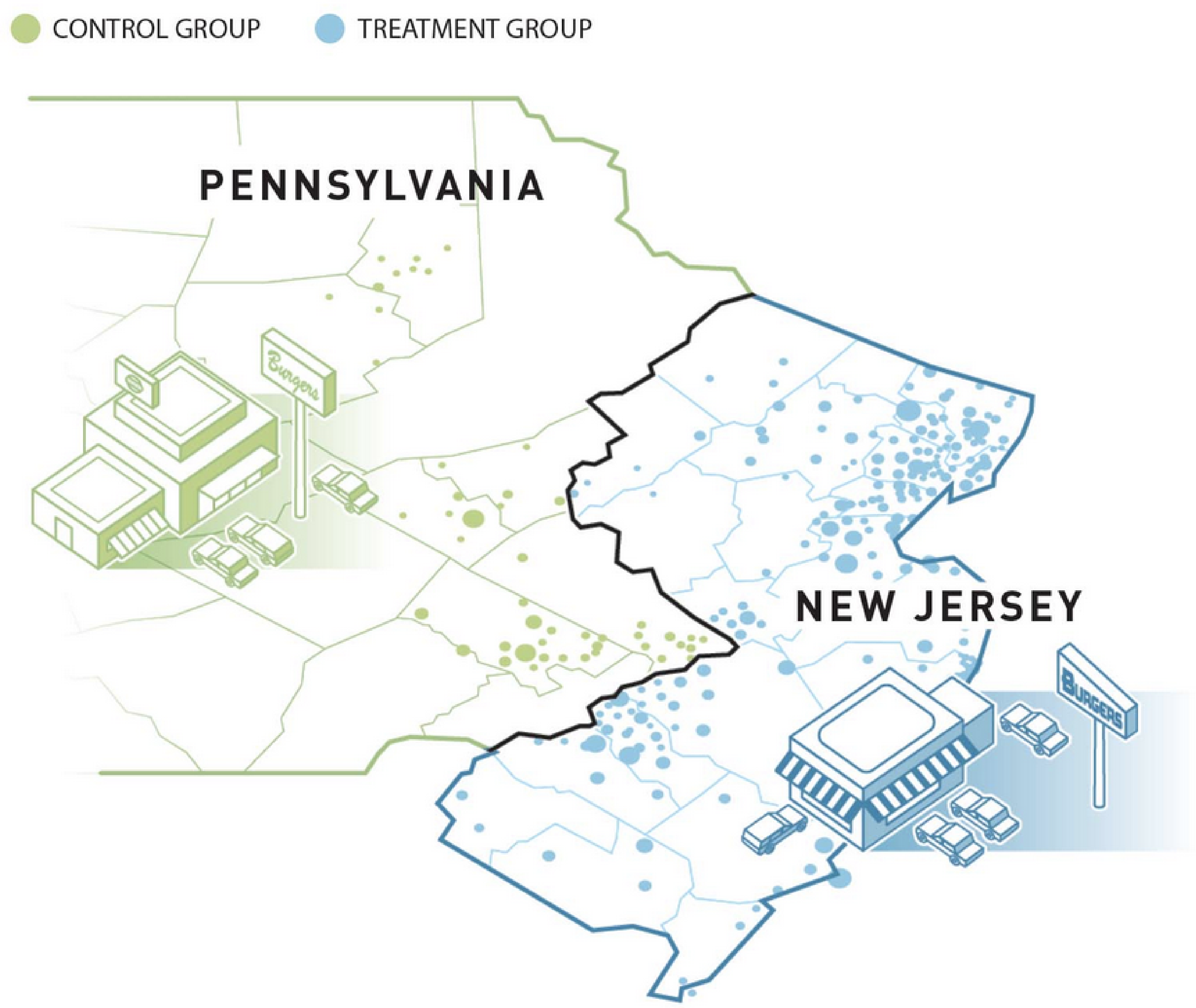
Card & Kreuger (1994): Background I

- Card & Kreuger (1994) compare employment in fast food restaurants on New Jersey and Pennsylvania sides of border between February and November 1992.
- Pennsylvania & New Jersey both had a minimum wage of $4.25 before February 1992
- In February 1992, New Jersey raised minimum wage from $4.25 to $5.05



Card & Kreuger (1994): Background II

- If we look only at New Jersey before & after change:
  - **Omitted variable bias**: macroeconomic variables (there's a recession!), weather, etc.
    - Including PA as a control will control for these time-varying effects if they are national trends
- Surveyed 400 fast food restaurants on each side of the border, before & after min wage increase
  - **Key assumption**: Pennsylvania and New Jersey follow parallel trends,
    - **Counterfactual**: if not for the minimum wage increase, NJ employment would have changed similar to PA employment



Card & Kreuger (1994): Comparisons

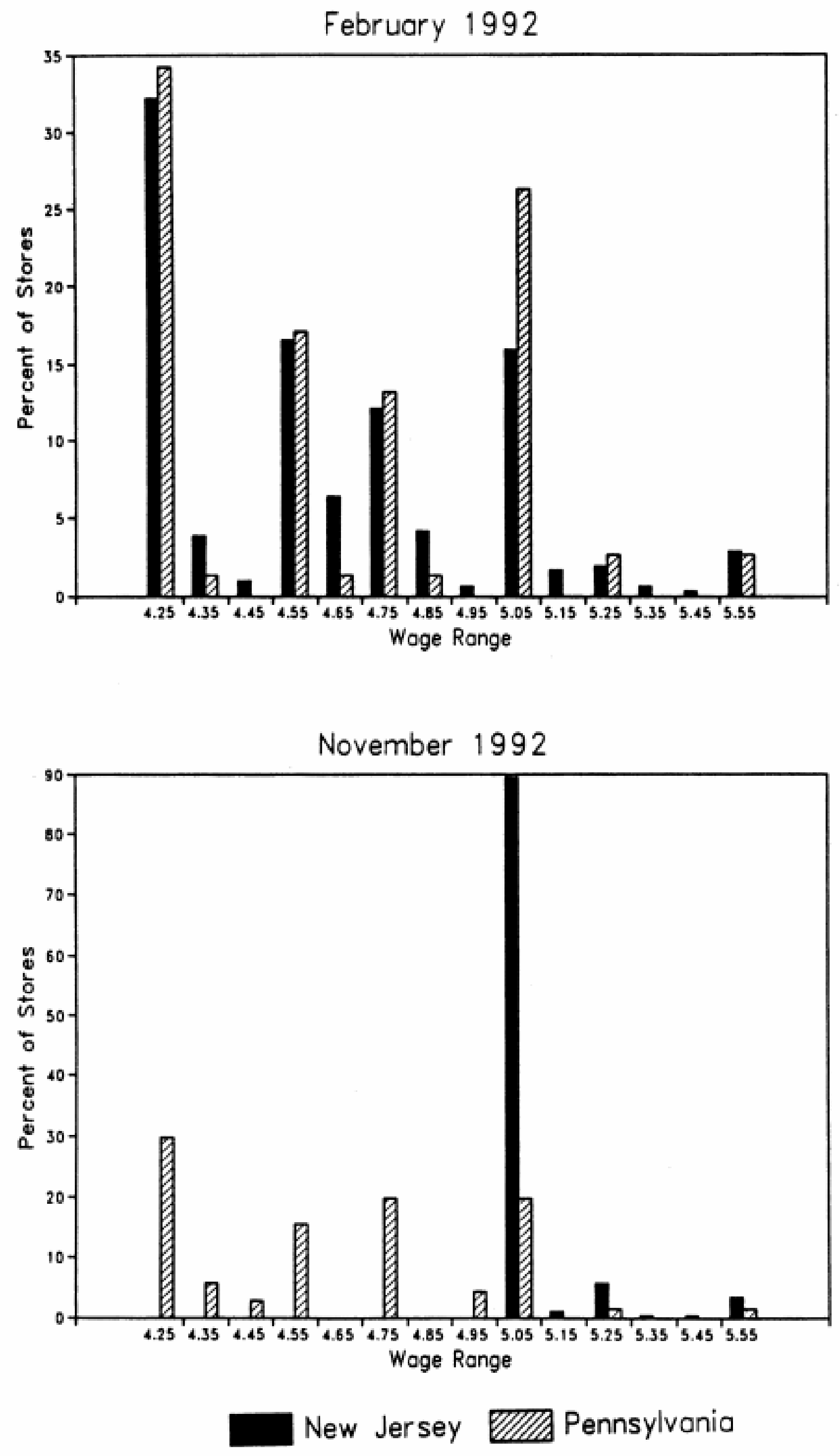


FIGURE 1. DISTRIBUTION OF STARTING WAGE RATES

Card & Kreuger (1994): Summary I

TABLE 1—SAMPLE DESIGN AND RESPONSE RATES

	All	Stores in: NJ	Stores in: PA
Wave 1, February 15 – March 4, 1992:			
Number of stores in sample frame:[a]	473	364	109
Number of refusals:	63	33	30
Number interviewed:	410	331	79
Response rate (percentage):	86.7	90.9	72.5
Wave 2, November 5 – December 31, 1992:			
Number of stores in sample frame:	410	331	79
Number closed:	6	5	1
Number under rennovation:	2	2	0
Number temporarily closed:[b]	2	2	0
Number of refusals:	1	1	0
Number interviewed:[c]	399	321	78

Card & Kreuger (1994): Summary II

TABLE 2—MEANS OF KEY VARIABLES

Variable	Stores in: NJ	PA
1. *Distribution of Store Types (percentages):*		
a. Burger King	41.1	44.3
b. KFC	20.5	15.2
c. Roy Rogers	24.8	21.5
d. Wendy's	13.6	19.0
e. Company-owned	34.1	35.4

Card & Kreuger (1994): Model

$$\widehat{\text{Employment}}_{it} = \beta_0 + \beta_1 \, \text{NJ}_i + \beta_2 \, \text{After}_t + \beta_3 \, (\text{NJ}_i \times After_t)$$

- PA Before: β_0
- PA After: $\beta_0 + \beta_2$
- NJ Before: $\beta_0 + \beta_1$
- NJ After: $\beta_0 + \beta_1 + \beta_2 + \beta_3$
- **Diff-in-diff**: $(\text{NJ}_{after} - \text{NJ}_{before}) - (\text{PA}_{after} - \text{PA}_{before})$

	PA	**NJ**	**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	$\beta_2 + \beta_3$	**Diff-in-diff** $\Delta_i \Delta_t$: β_3

Card & Kreuger (1994): Results

Variable	Stores by state PA (i)	NJ (ii)	Difference, NJ – PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	–2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	–0.14 (1.07)
3. Change in mean FTE employment	–2.16 (1.25)	0.59 (0.54)	2.76 (1.36)