Class 10: Difference-in-Differences (Basics)

POLS 6388: Causal Inference

Michael Kistner
Department of Political Science, University of Houston

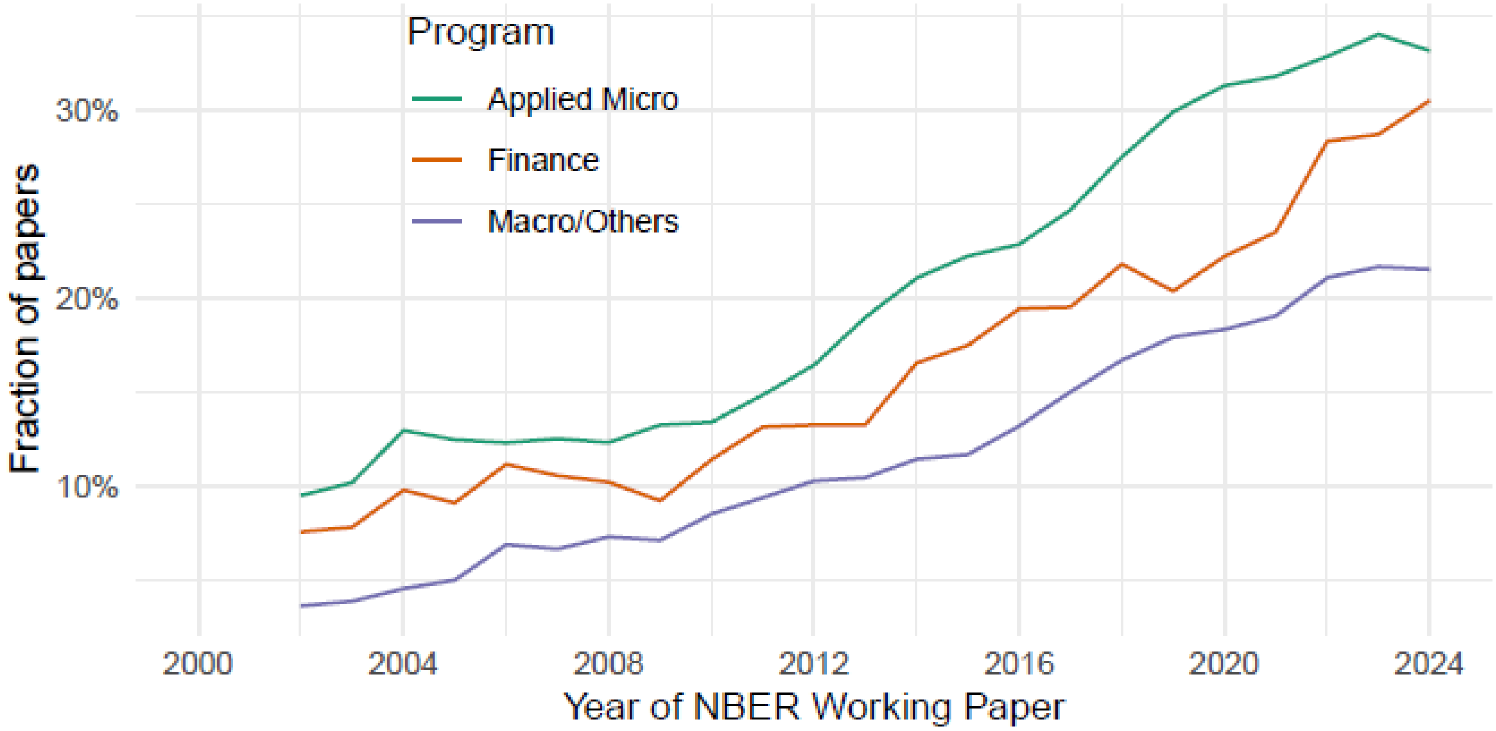
March 25, 2025

Goals of Today's Class

1. Introduce what a difference-in-differences (DiD) design is and what it estimates
2. Learn how to implement a DiD in the two-period case
3. Understand the parallel trends assumption and when it is plausible
4. Learn how to conduct an event study to examine pre-trends and dynamic DiD effects

Introducing the Difference-in-Differences Design

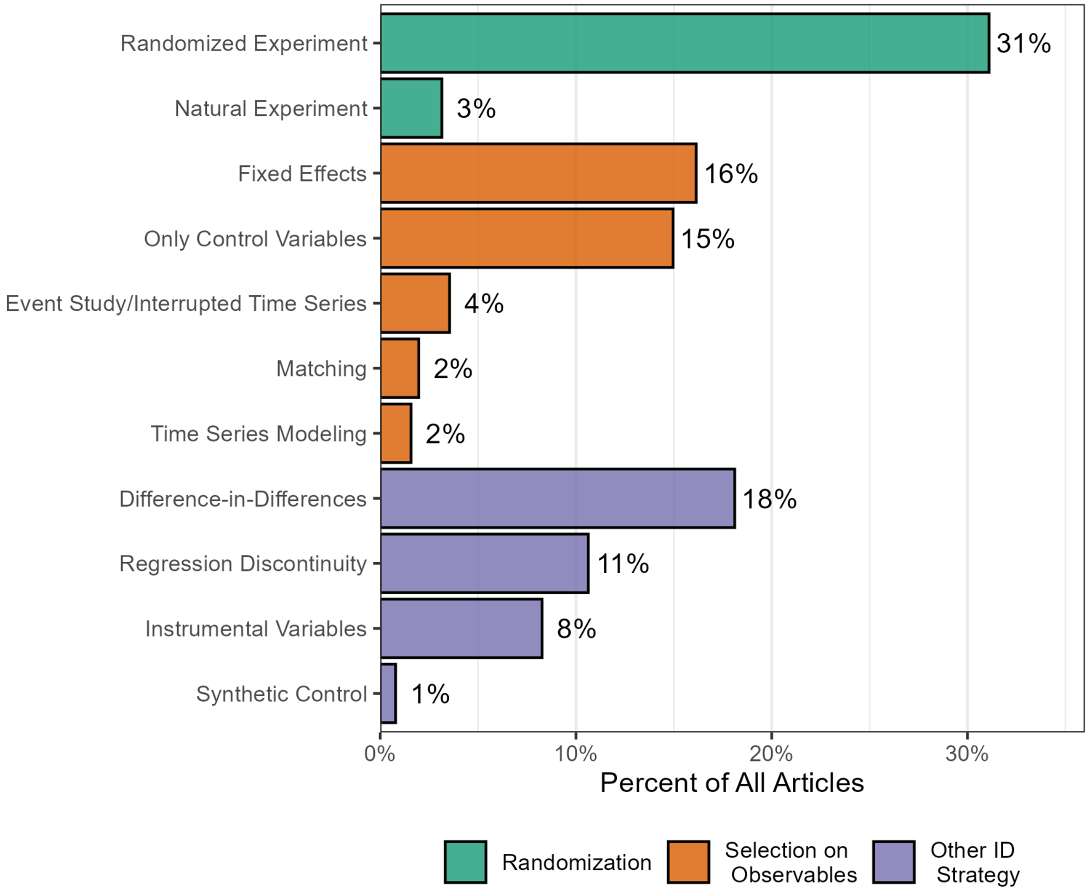
Difference-in-Differences in Economics



(a) Difference-in-differences

Percent of NBER Papers Using DiD

Difference-in-Differences in Political Science



Percent of Political Science Articles (2022 - 2024) Using DiD (And Other ID Strategies)

Combining Panel Data and the Interrupted Time Series

Recall the setup for the *interrupted time series* design…

- A treatment ($Post_t$) occurs at some point in time $t = t^*$, s.t.

$$Post_t = \begin{cases} 0 & t < t^* \text{(pre-periods)} \\ 1 & t \geq t^* \text{(post-periods)} \end{cases}$$

Recall also the concept of *panel data*

- Data where we have multiple units (i) observed for multiple time periods (t)

Extending the Interrupted Time Series

Let's extend the ITS design for the panel context…

- For some units (*the treated group*), $Post_{it} = 1$ iff $t \geq t^*$
- For other units (*the control group*), $Post_{it} = 0 \; \forall t$

Suppose that there are only two time periods: $t \in \{1, 2\}$ and $t^* = 2$

How could we estimate the TE? Two obvious comparisons:

- **The Within-Comparison** ($Y_{i2}^{Treated} - Y_{i1}^{Treated}$)
- **The Between-Comparison** ($Y_{i2}^{Treated} - Y_{i2}^{Control}$)

The Difference-in-Differences Design

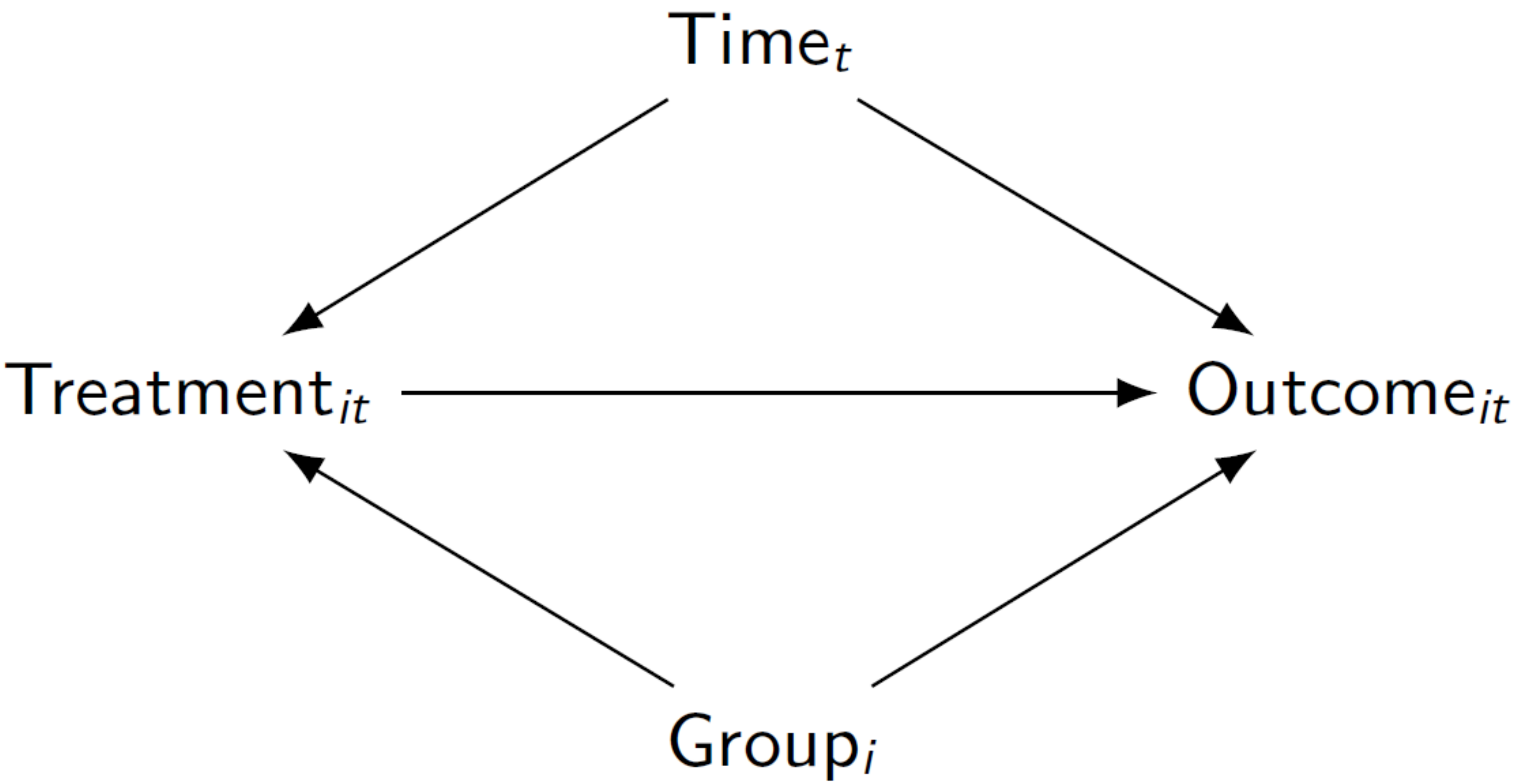
What are the advantages of each comparison?

- **Within-Comparison** – Differences between treated and control groups can't be a confounder
- **Between-Comparison** – Concurrent unrelated changes can't be a confounder

Each comparison is useful for different reasons. Can we use both?

Difference-in-Differences Design – a design that compares the *change* in outcome (Y_{it}) among treated units to the *change* in outcome among never treated units.

Why the DiD Design is So Powerful (1/2)



DAG Displaying the DiD Setting

Why the DiD Design is So Powerful (2/2)

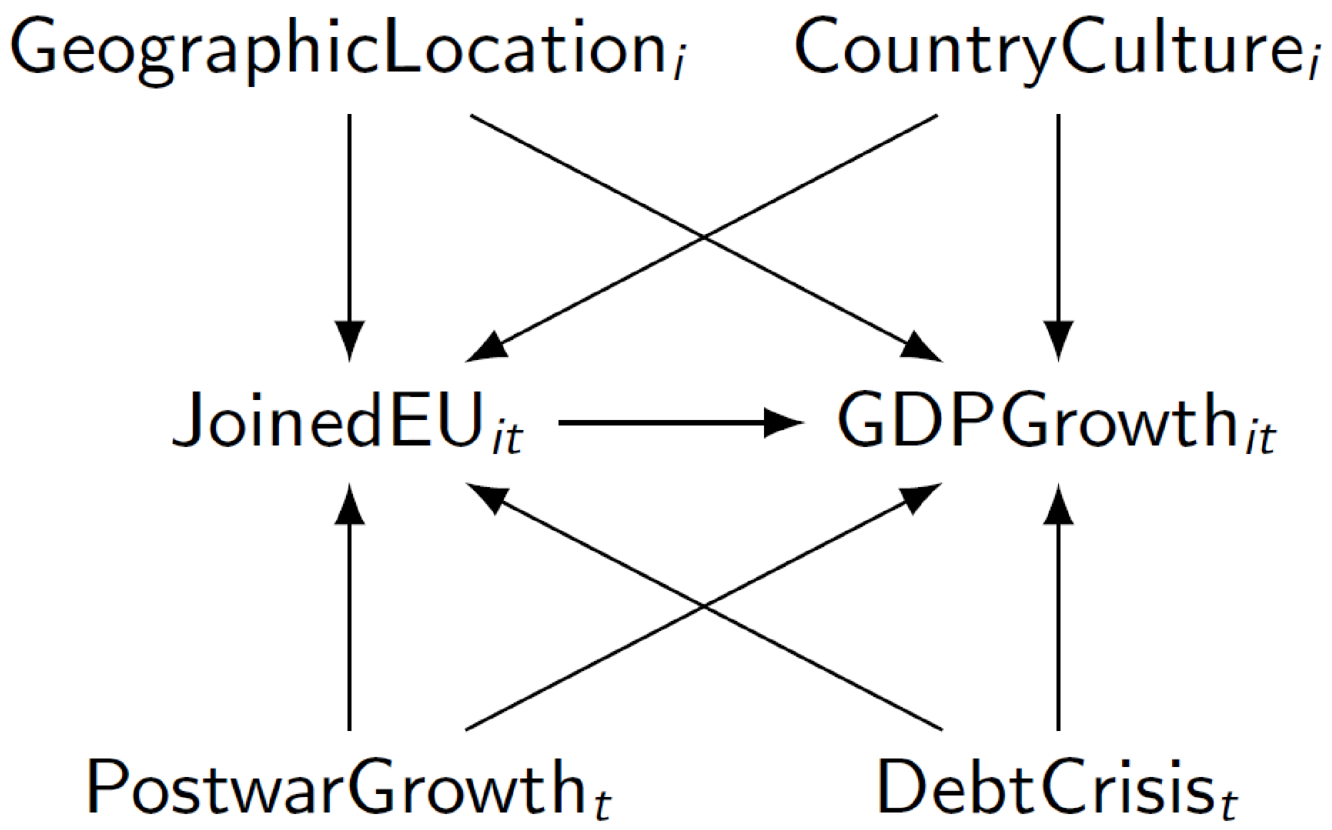
To control for both sets of factors simultaneously, we…

1. **Isolate** the within-variation (pre- vs. post-treatment periods) for treated and untreated
   - How much did each group change?
   - This is the second part of the name (the "-differences")
2. **Compare** the within-variation between treated and untreated by subtracting untreated within-variation from treated
   - How much more/less did the treated group change post-treatment than the control group?
   - This is the second part of the name (the "difference-in")

The Class of ’04

The Largest European Union Expansion of All Time

A DAG of Multidimensional Confounding



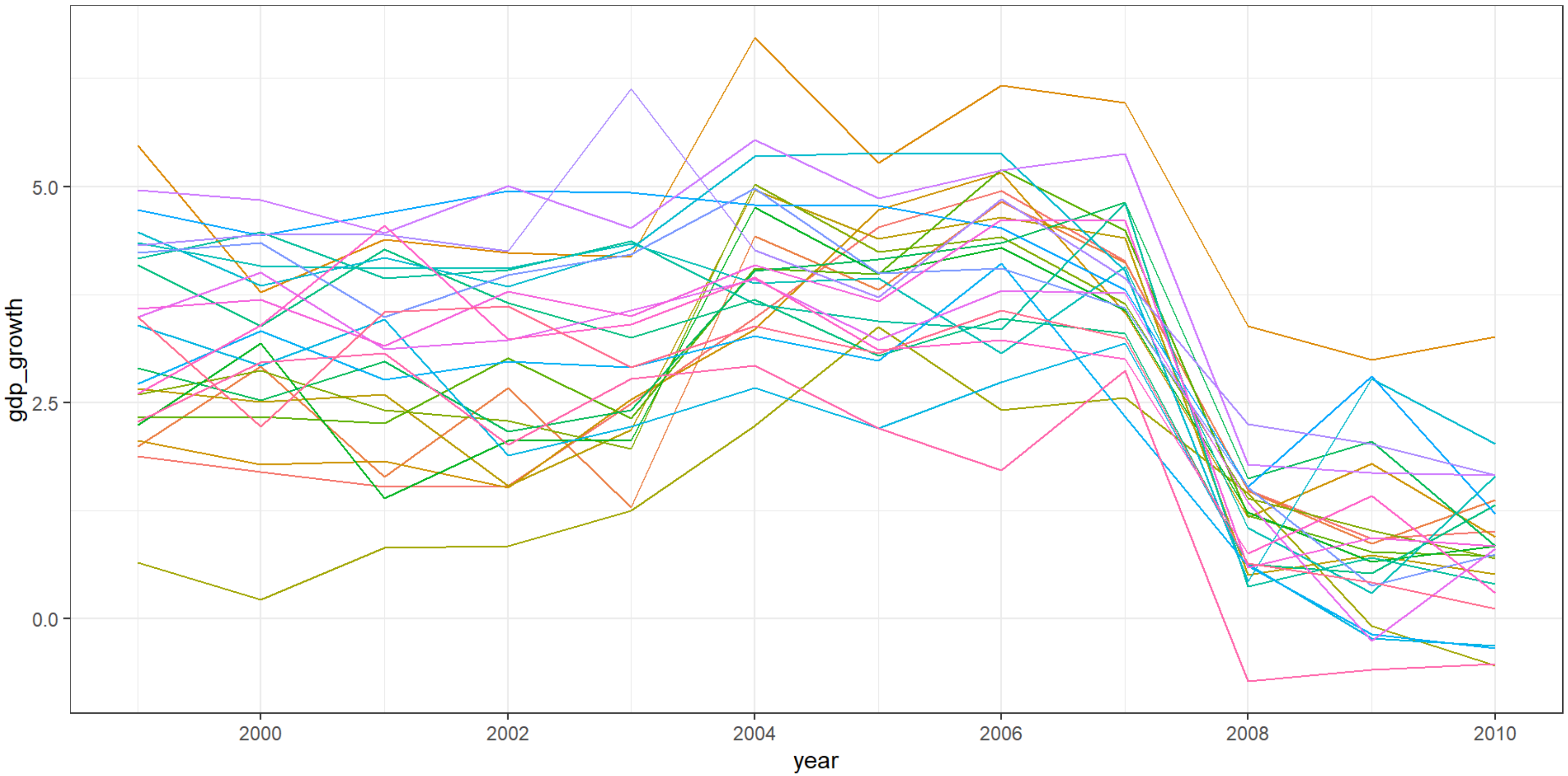
DAG of EU Admission with Country and Year Confounding

Generating Data Using the DAG

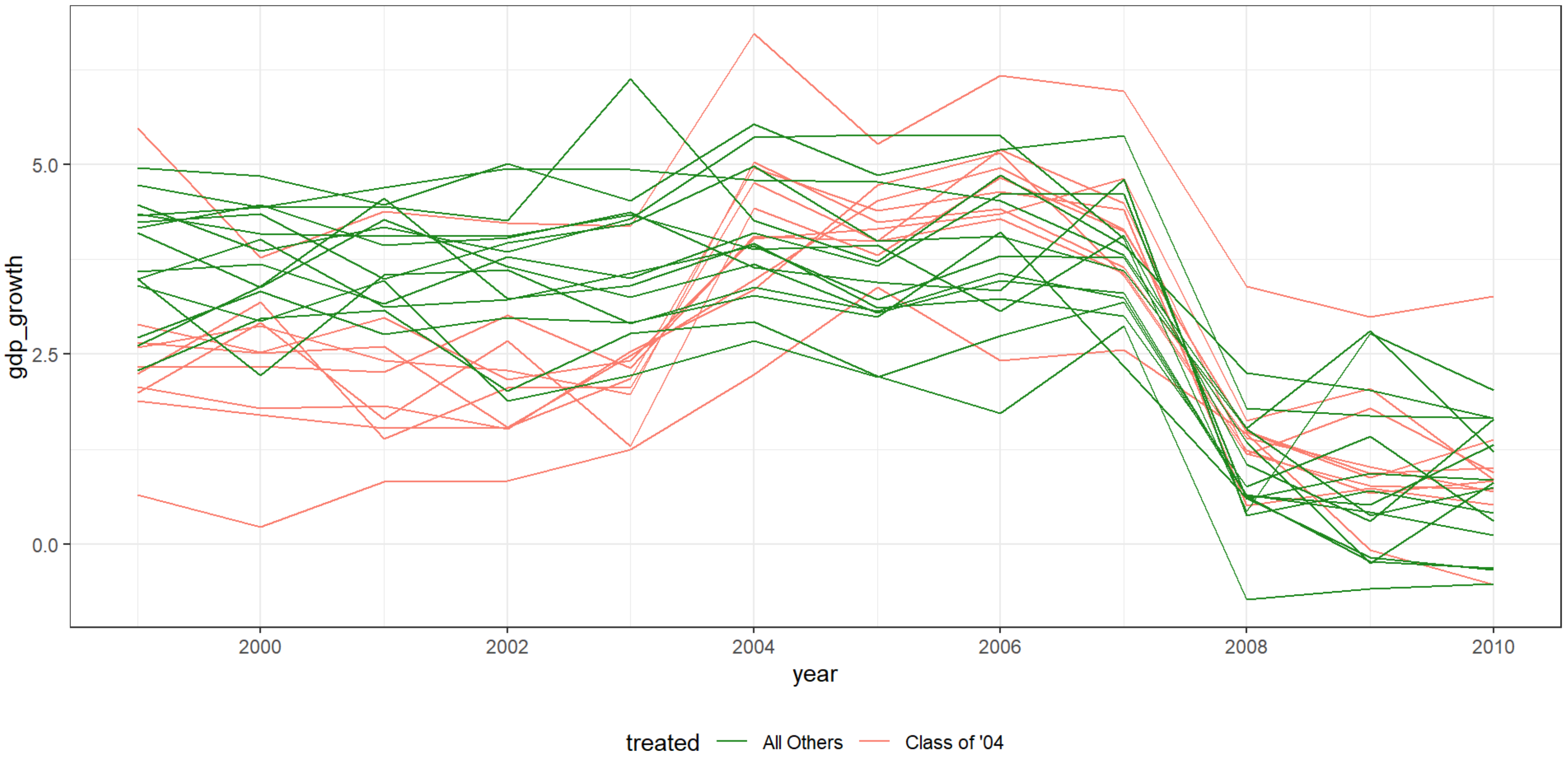
```r
class_04 <- c(1:10)

eu_df <- data.frame(
  country_id = as.factor(rep(1:25, 12)),
  culture = rep(rnorm(25, 0, 0.5), 12),
  eastern = rep(c(rbernoulli(10, 0.9),
                  rbernoulli(15, 0.2)), 12),
  year = sort(rep(1999:2010, 25)),
  debt_crisis = c(rep(0, 225), rep(1, 75))
) %>%
  rowwise() %>%
  mutate(joined_eu = ifelse(country_id %in% class_04 & year >= 2004, 1, 0),
         gdp_growth = culture + 2 * joined_eu -
           2 *eastern - 3 * debt_crisis + rnorm(1, 4, 0.5))
```

Implies the treatment effect $\tau = 2\%$ GDP growth per year

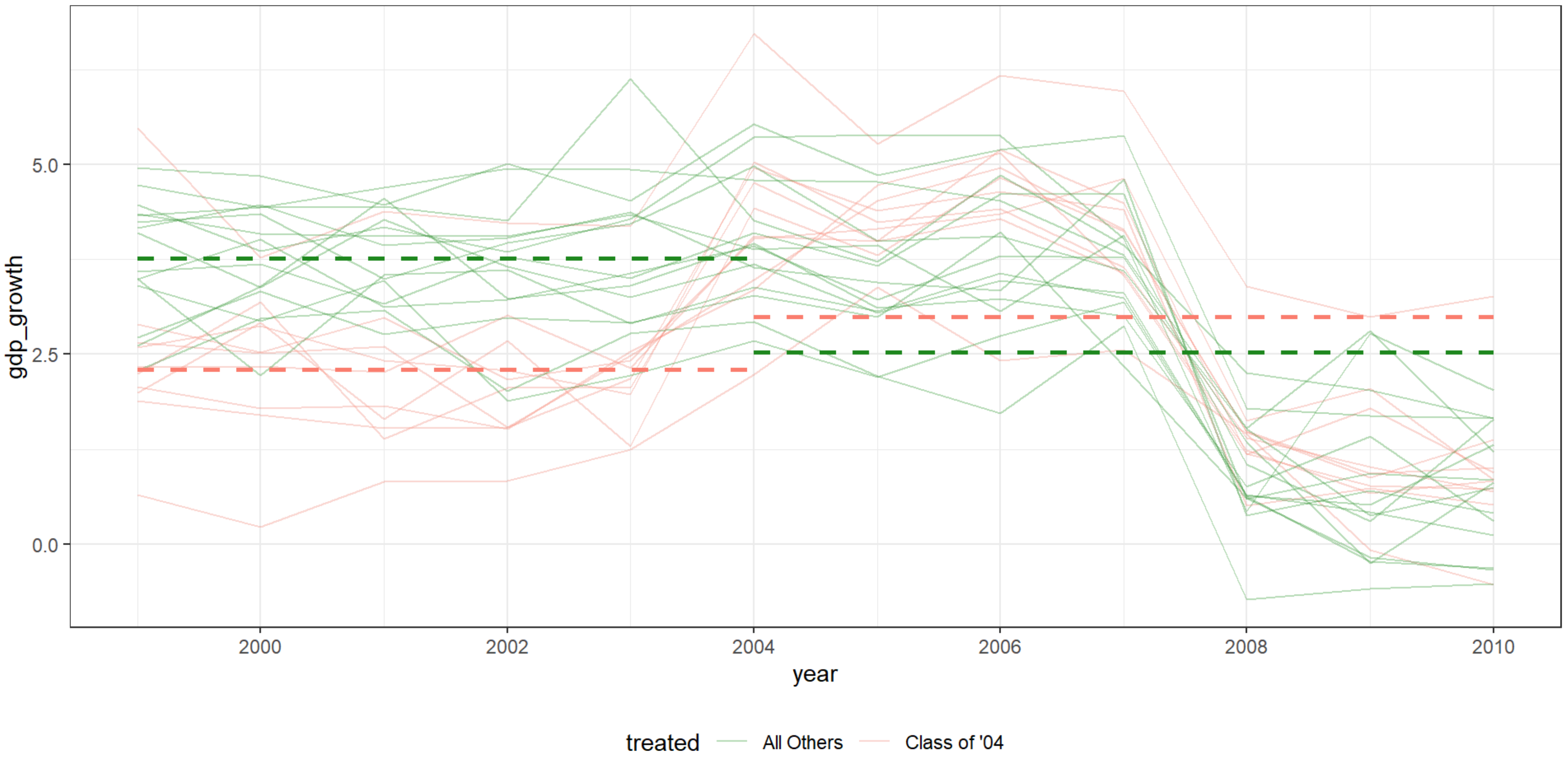
Visualizing the Data



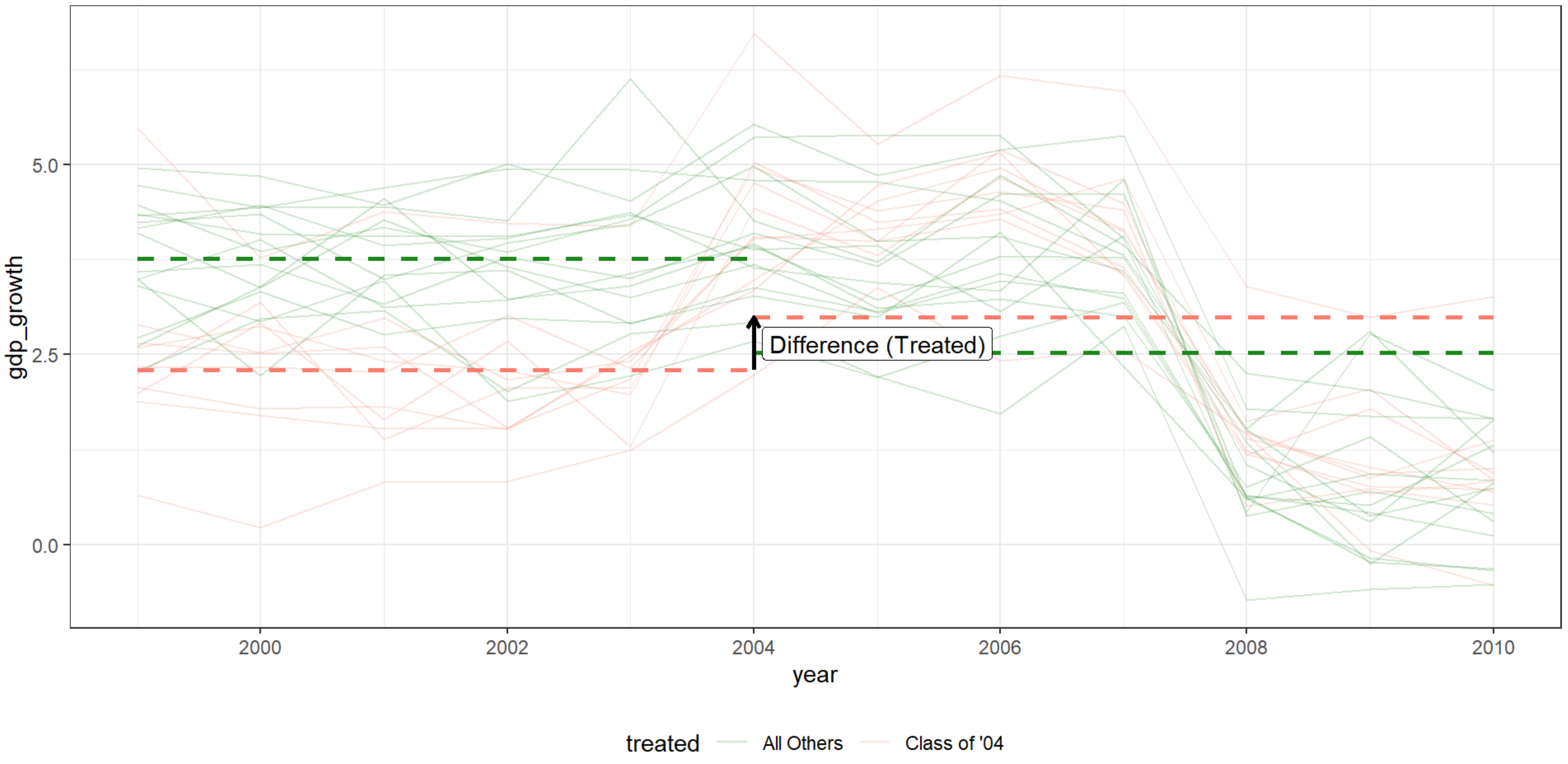
Visualizing Treated Versus Untreated



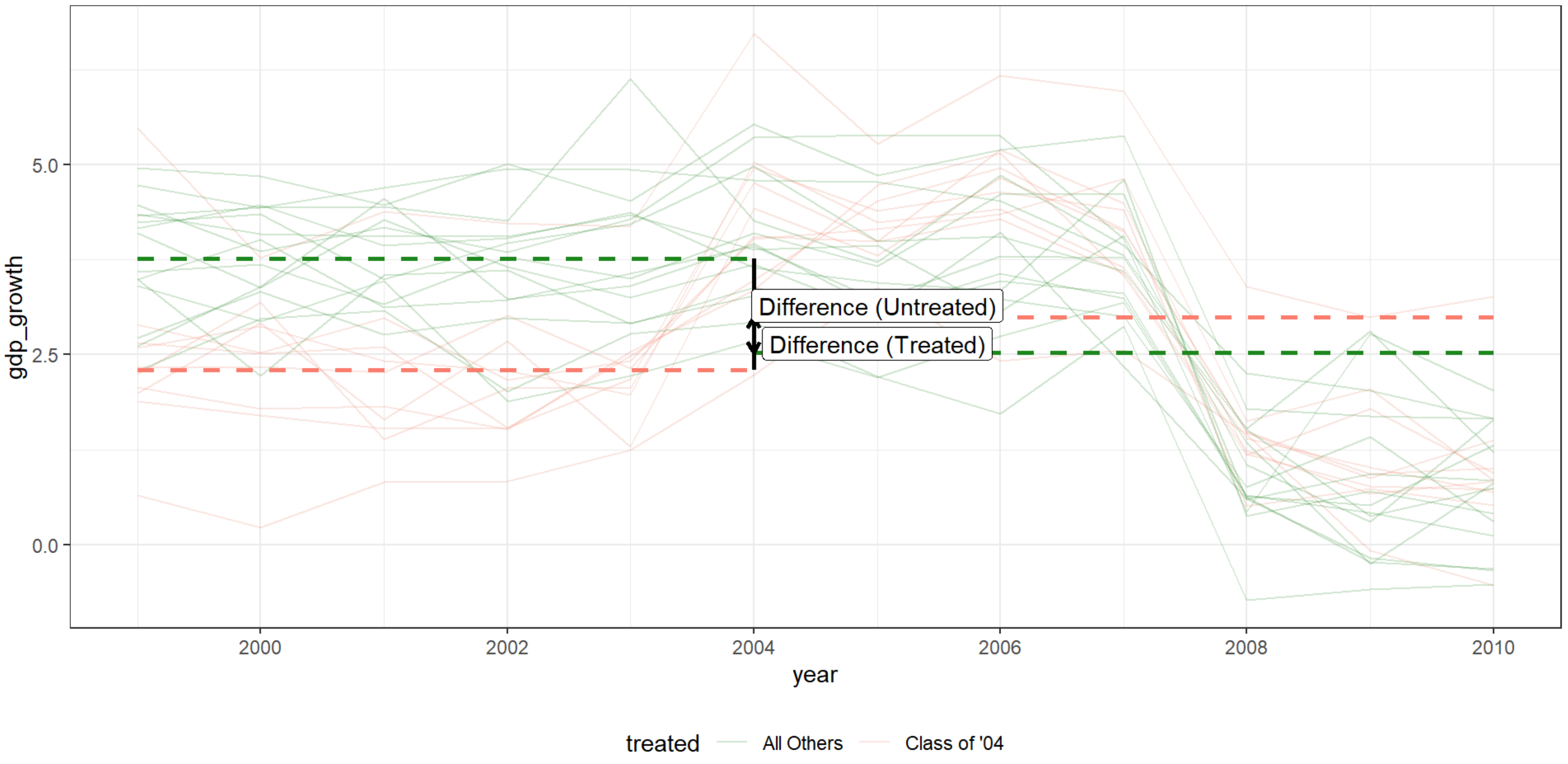
Pre-/Post- Means by Treatment Group



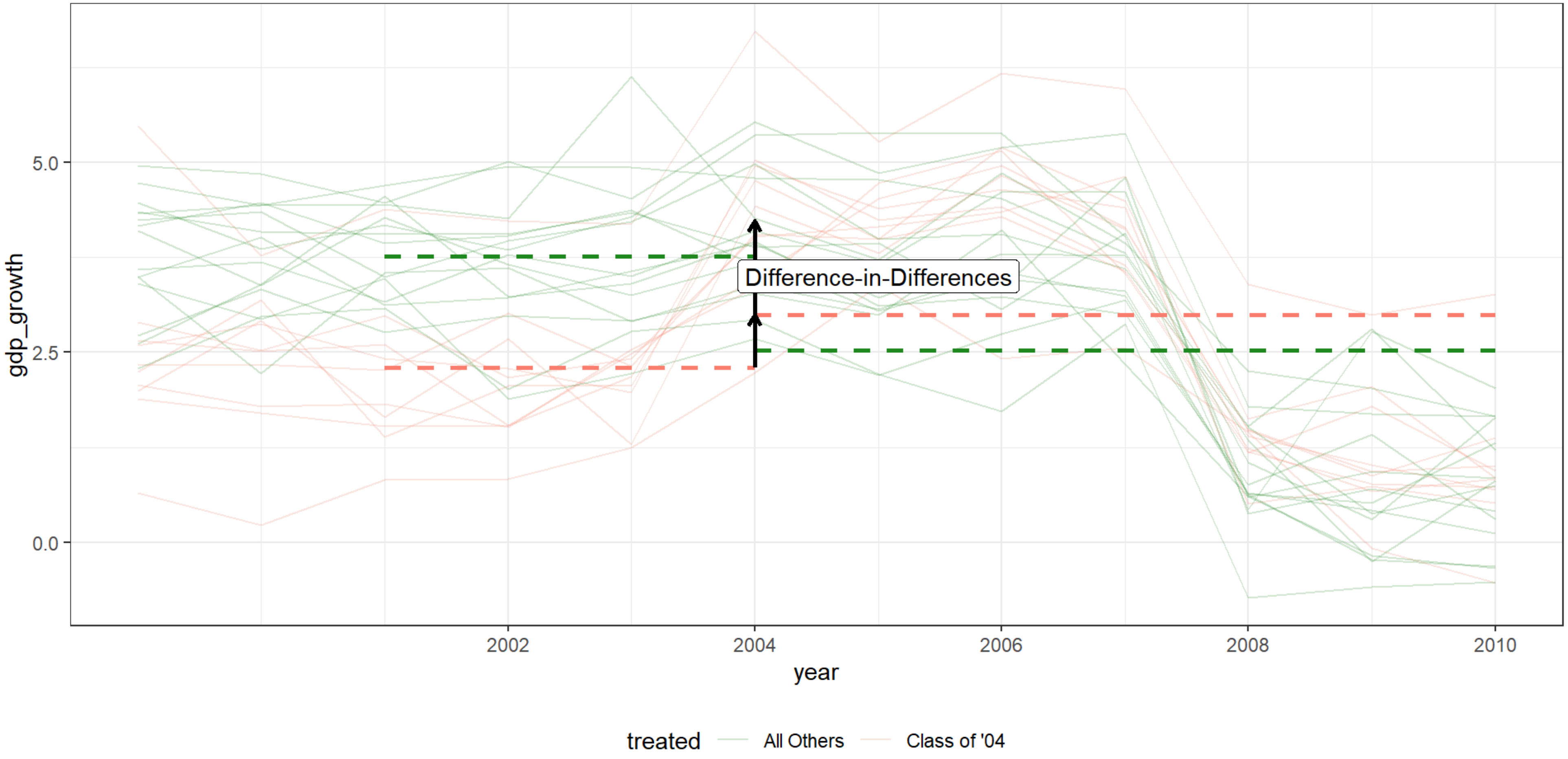
Calculating Difference #1



Calculating Difference #2



Calculating the DiDs



Implementing a Difference-in-Differences Analysis

The Two-By-Two Setup

Some of the earliest DiD models used only two periods (one before treatment, and one after) with two groups (the treated and the untreated)

- **Example:** Card and Krueger (1993) Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania

The two-by-two estimator compares:

$$\hat{\tau}_{2x2} = (\bar{Y}_{i1}^{Treated} - \bar{Y}_{i0}^{Treated}) - (\bar{Y}_{i1}^{Control} - \bar{Y}_{i0}^{Control})$$

The Two-By-Two Regression Estimator

Denote treatment versus control groups using $D_i = \{1, 0\}$, respectively

Denote pre- versus post-treatment using $Post_t = \{0, 1\}$, respectively

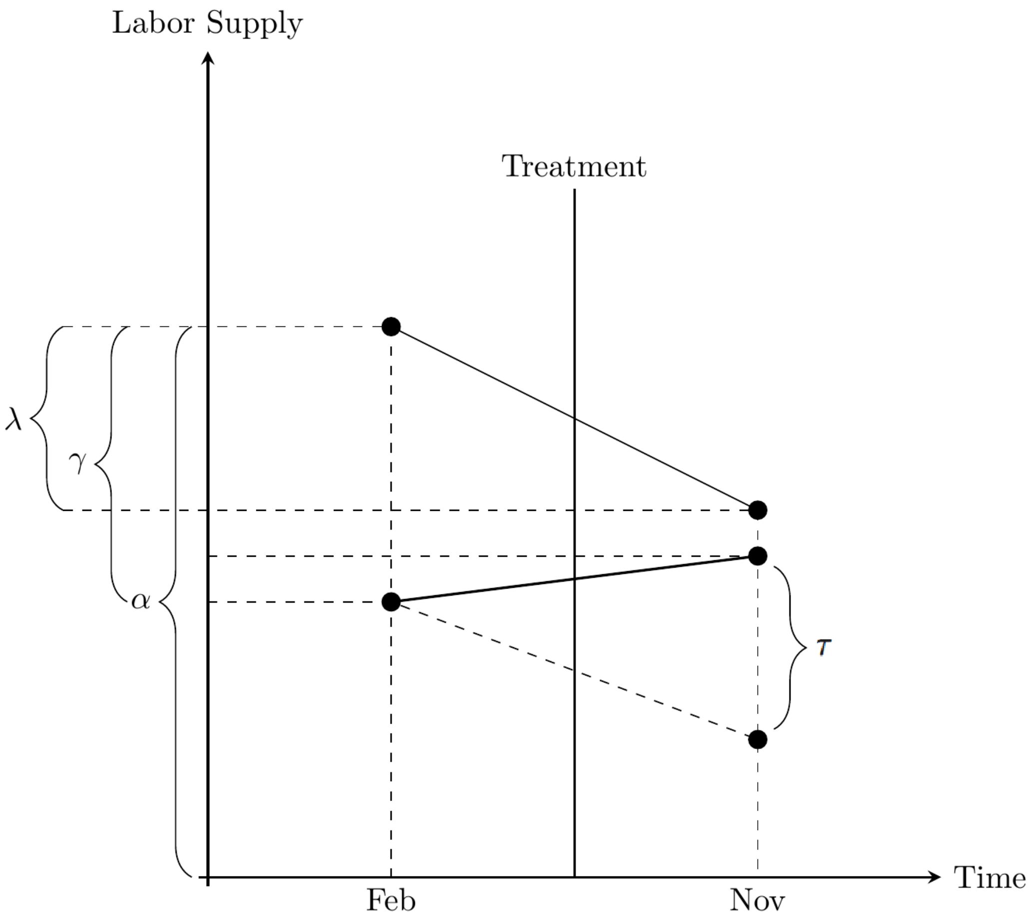
Then we can estimate the two-by-two setup as:

$$Y_{it} = \hat{\alpha} + \hat{\gamma} D_i + \hat{\lambda} Post_t + \hat{\tau}_{2x2} \left(D_i \times Post_t \right) + \epsilon_{it}$$

Here:

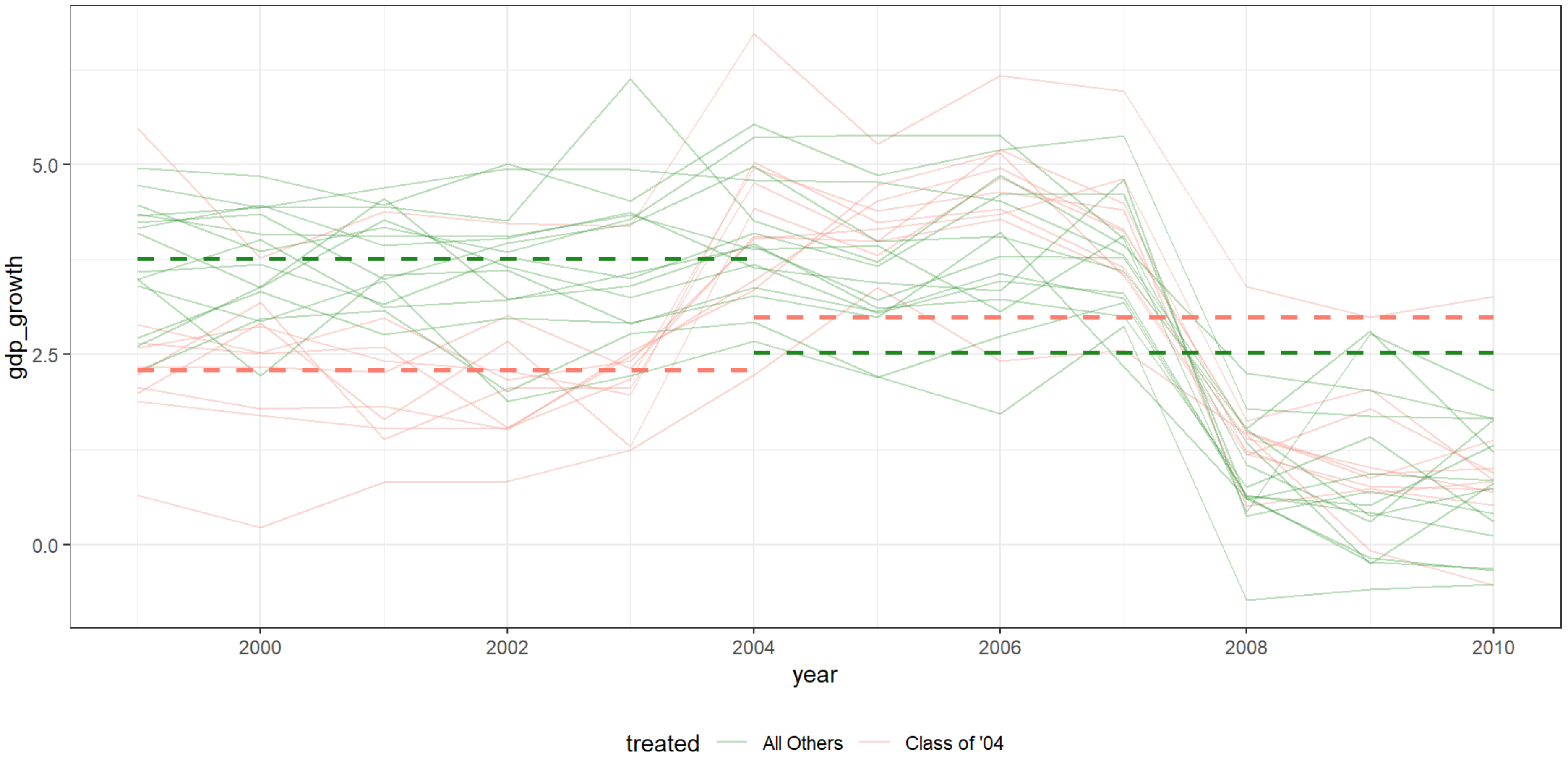
- $\hat{\gamma}$ estimates $\bar{Y}_{i0}^{Treated} - \bar{Y}_{i0}^{Control}$ (pre-existing differences)
- $\hat{\lambda}$ estimates $\bar{Y}_{i1}^{Control} - \bar{Y}_{i0}^{Control}$ (non-treatment time differences)
- $\hat{\tau}_{2x2}$ is the difference in slope for the treated (versus the control)

Mapping Coefficients To Estimands

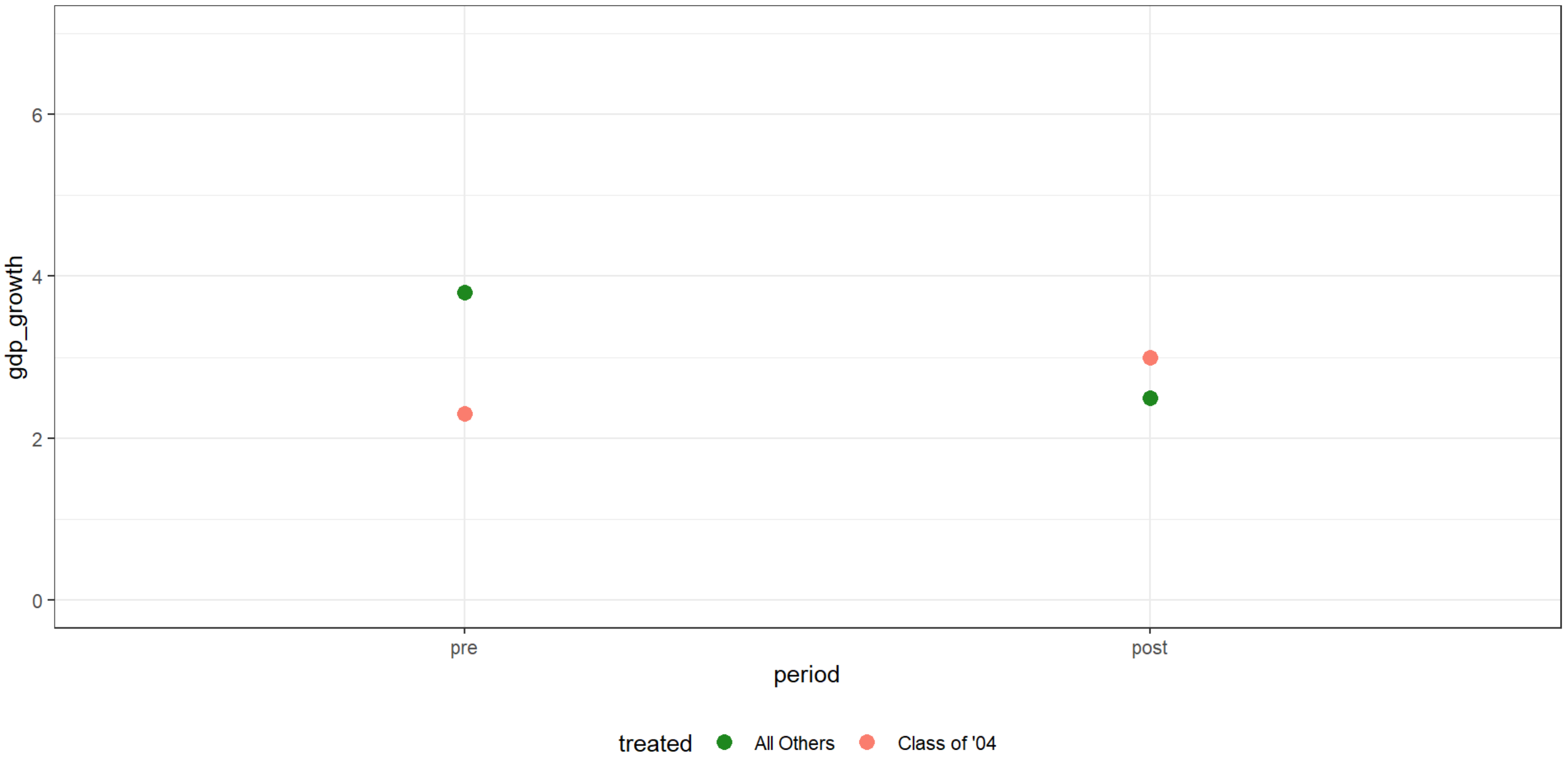


The 2X2 Estimator for the Minimum Wage Example

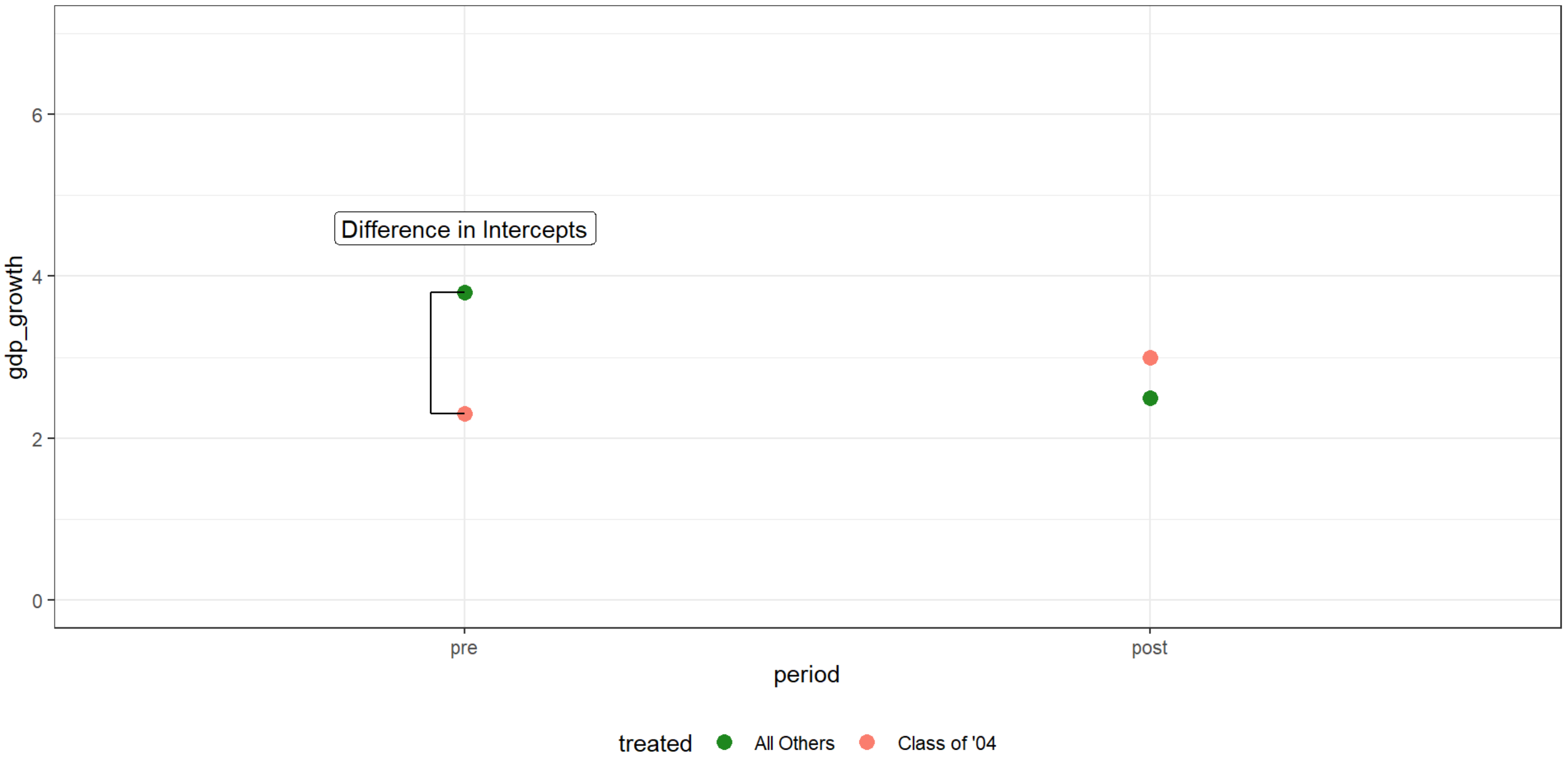
EU Membership as a 2x2 (1/5)



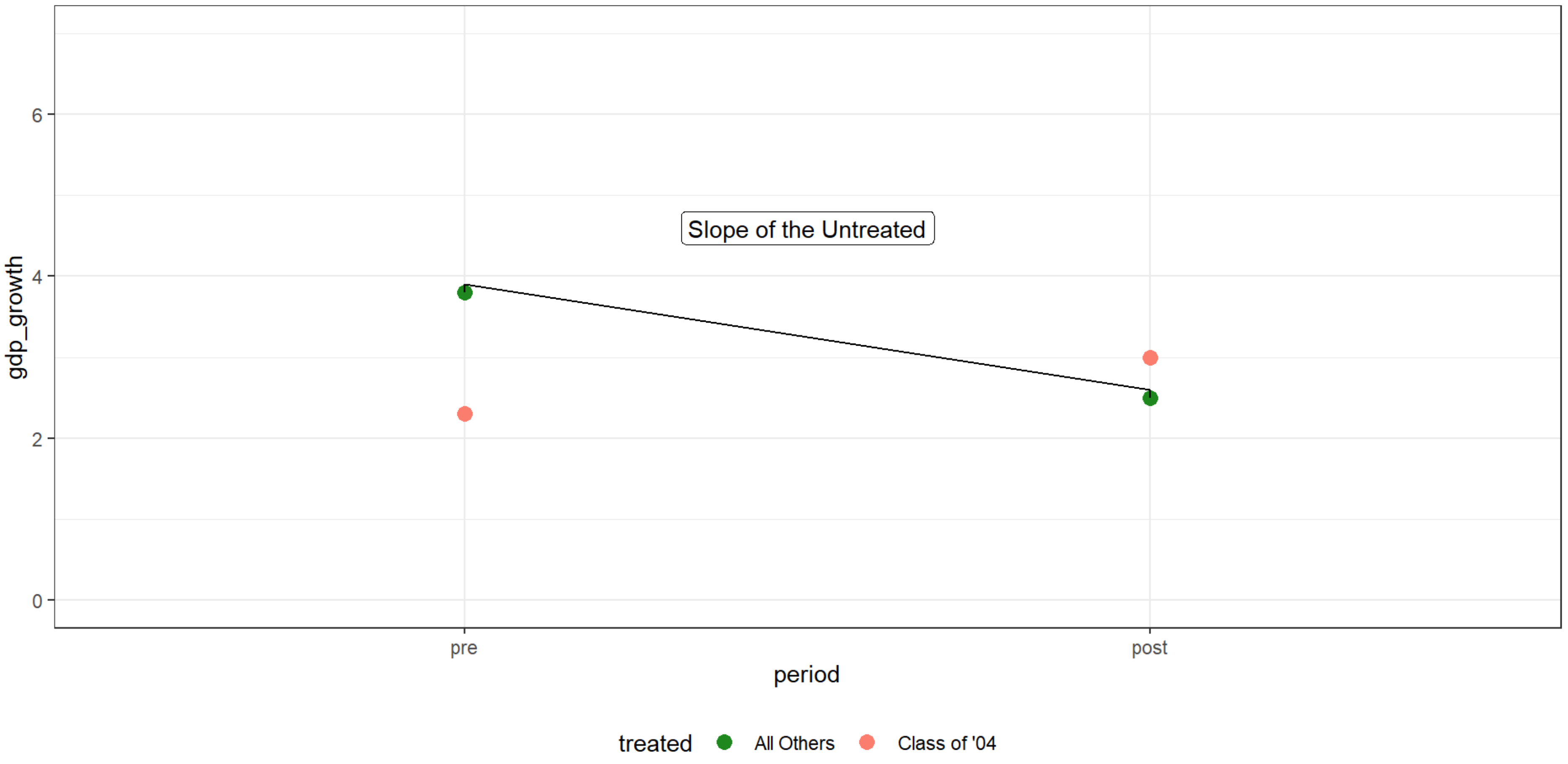
EU Membership as a 2x2 (2/5)



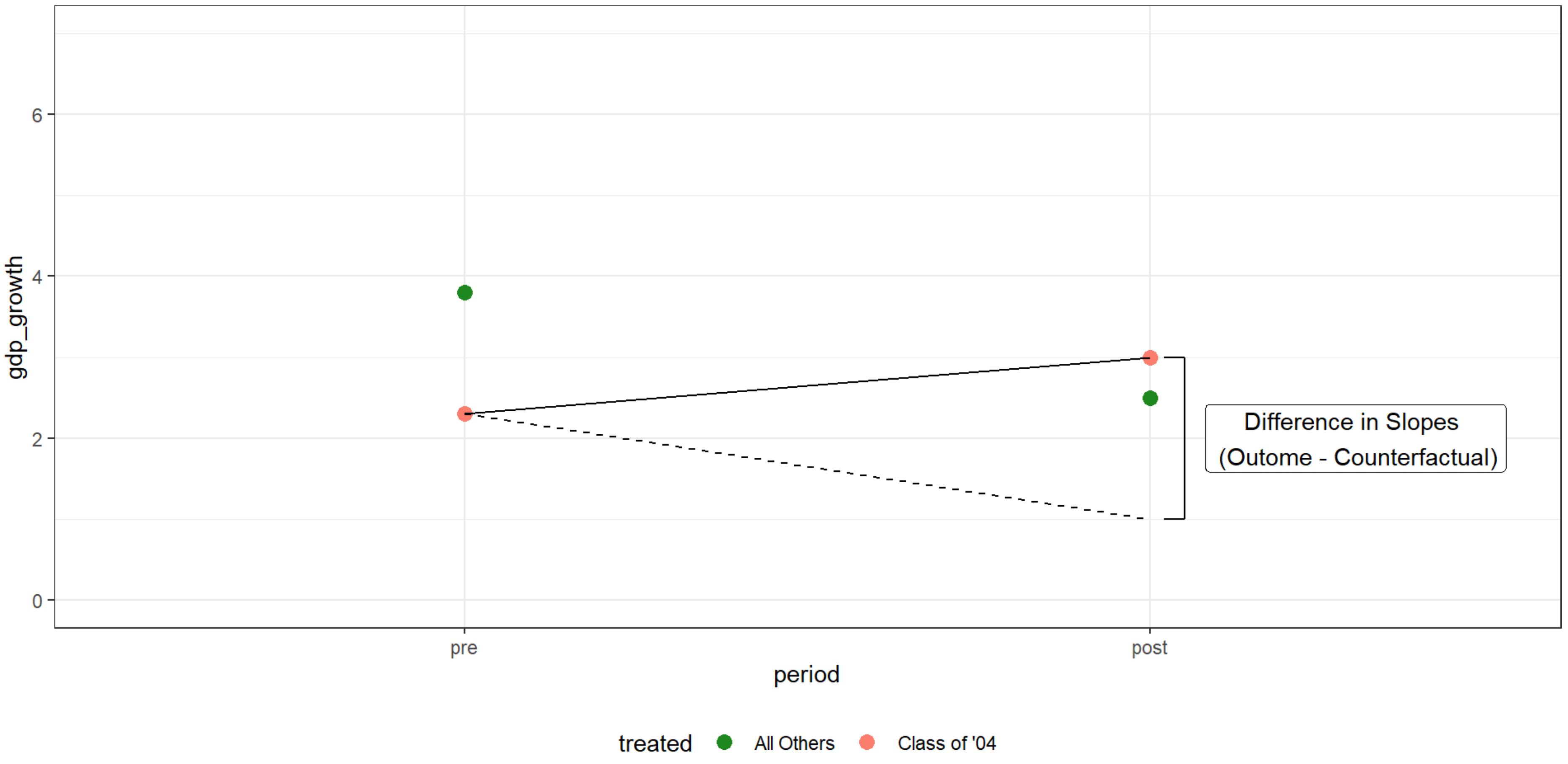
EU Membership as a 2x2 (3/5)



EU Membership as a 2x2 (4/5)



EU Membership as a 2x2 (5/5)



Estimating the 2x2 Design Using OLS

```r
prepost_df <- mutate(eu_df, period = ifelse(year < 2004, "1", "2")) %>%
  group_by(period, country_id) %>%
  summarize(gdp_growth = mean(gdp_growth),
            treated = first(treated),
            joined_eu = first(joined_eu))

reg_2x2 <- lm(gdp_growth ~ period + treated + period*treated,
              data = prepost_df)
```

	(1)
Class of '04	−1.46*
	(0.31)
Post-2004	−1.24*
	(0.28)
Class of '04 x Post-2004	1.93*
	(0.45)
Num.Obs.	50
R2	0.381

* $p < 0.05$

DiD and the 2WFE Estimator

A Common (But Risky) Alternative The *Two-Way Fixed Effects (2WFE) Estimator*

- Discussed this in last week's lecture (on fixed effects):

$$Y_{it} = \hat{\alpha}_i + \hat{\gamma}_t + \hat{\tau}_{2WFE} Post_{it} + \epsilon_{it}$$

Why Risky? In the two-period case, it's actually not.

- The 2x2 and the 2WFE estimators will return the exact same estimates

With more than two time periods, the methods return different answers

- The 2WFE estimator will use non-intuitive weights and make "bad" comparisons (we'll talk more about this next week)

Comparing the Two Estimators

```r
reg_2WFE<- lm(gdp_growth ~ joined_eu +
                as.factor(year) + as.factor(country_id), # the fixed effects
              data = eu_df)
```

	2x2	2WFE
Joined EU		1.93*
		(0.12)
Class of '04 x Post-2004	1.93*	
	(0.45)	
Num.Obs.	50	300
R2	0.381	0.908
Period FEs	N	Y
Country FEs	N	Y

* $p < 0.05$

Inference with Difference-in-Differences

With panel data and difference-in-differences, we have two concerns:

- Errors might be systematically larger or smaller for particular units (within *clusters*)
- Errors are likely to be similar in size and direction for consecutive time periods (*autocorrelation*)

One solution: two-way clustering (cluster on unit and year)

- Usually will lead to more conservative (i.e., larger standard errors)
- Requires a sufficient # of clusters (≥ 50; see Cameron and Miller 2014)
- With few clusters (< 50), use Wild Bootstrap (Cameron et. al. 2008)

Inference Using Different SE types

```r
model_1 <- feols(gdp_growth ~ joined_eu | country_id + year,
                 vcov = "iid", data = eu_df)

model_2 <- feols(gdp_growth ~ joined_eu | country_id + year,
                 cluster = "country_id", data = eu_df)

model_3 <- feols(gdp_growth ~ joined_eu | country_id + year,
                 vcov = "newey_west", panel.id = ~ country_id + year, data = eu_df)

model_4 <- feols(gdp_growth ~ joined_eu | country_id + year,
                 vcov = "twoway", data = eu_df)
```

	SEs: iid	SEs: clustered	SEs: autocorrelated	SEs: twoway
Joined EU	1.93*	1.93*	1.93*	1.93*
	(0.12)	(0.12)	(0.11)	(0.10)
Num.Obs.	300	300	300	300
R2	0.908	0.908	0.908	0.908

* $p < 0.05$

Estimands And Assumptions

What Counterfactual Does the DiD Recover? (1/3)

What is the theoretical estimand we're recovering with $\hat{\tau}_{2x2}$?

Let's begin by writing out the empirical estimand:

$$\hat{\tau}_{2x2} = (\bar{Y}_{i1}^{Treated} - \bar{Y}_{i0}^{Treated}) - (\bar{Y}_{i1}^{Control} - \bar{Y}_{i0}^{Control})$$

We can rewrite the right-hand side as a set of conditional expectations:

$$\big(\mathbf{E}[Y^T|Post] - \mathbf{E}[Y^T|Pre\big) - \big(\mathbf{E}[Y^C|Post] - \mathbf{E}[Y^C|Pre]\big)$$

Let's express this in terms of the potential outcomes (we observe):

$$\big(\mathbf{E}[Y^T(1)|Post] - \mathbf{E}[Y^T(0)|Pre\big) - \big(\mathbf{E}[Y^C(0)|Post] - \mathbf{E}[Y^C(0)|Pre]\big)$$

What Counterfactual Does the DiD Recover? (2/3)

Now we'll use a familiar trick: adding zero

$$
\begin{aligned}
&\big(\mathbf{E}[Y^T(1)|Post] - \mathbf{E}[Y^T(0)|Pre\big) - \big(\mathbf{E}[Y^C(0)|Post] - \mathbf{E}[Y^C(0)|Pre]\big) \\
&\quad + \mathbf{E}[Y^T(0)|Post] - \mathbf{E}[Y^T(0)|Post]
\end{aligned}
$$

Finally, we can re-arrange terms:

$$
\begin{aligned}
\hat{\tau}_{2x2} = &\underbrace{\mathbf{E}[Y^T(1)|Post] - \mathbf{E}[Y^T(0)|Post]+}_{\text{ATT}} \\
&\underbrace{\big(\mathbf{E}[Y^T(0)|Post] - \mathbf{E}[Y^T(0)|Pre]\big) - \big(\mathbf{E}[Y^C(0)|Post] - \mathbf{E}[Y^C(0)|Pre]\big)}_{\text{Non-Parallel Trend Bias}}
\end{aligned}
$$

What Counterfactual Does the DiD Recover? (3/3)

DiD estimates the TE by comparing *observed outcomes for the treated in the post-period* to *counterfactual outcomes for the treated in the post-period* (without treatment)

To construct this counterfactual, we must be willing to make the *parallel trends assumption*:

- Parallel Trends Assumption – Absent treatment, change in the outcome of treated units would have been the same as change for the untreated units
- Compare this to the CEA of the ITS. How is this better?

Example: Would [Poland/Estonia/Latvia…] have had the same economic growth as [Ukraine/Belarus/Moldova…] if they didn't join the EU?

Violation of the Parallel Trends Assumption

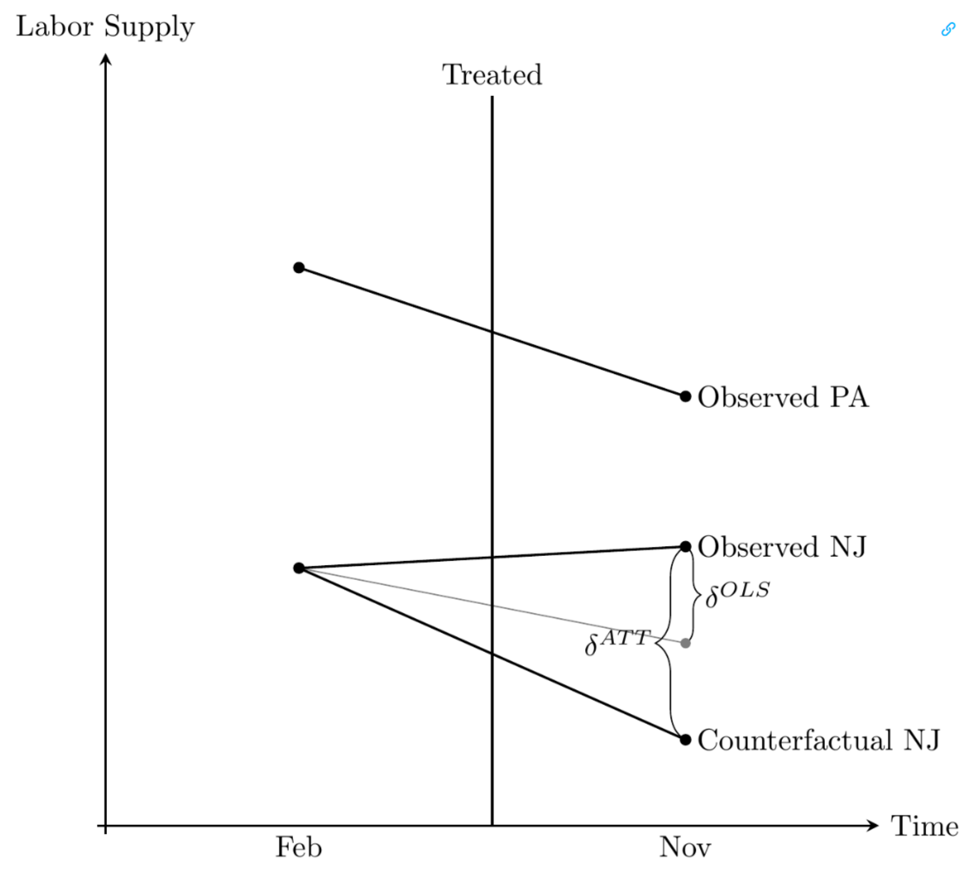


Figure 9.3: DD regression diagram without parallel trends

Non-Parallel Trends in the Minimum Wage Example

Comparing Pre-Trends to Assess Plausibility

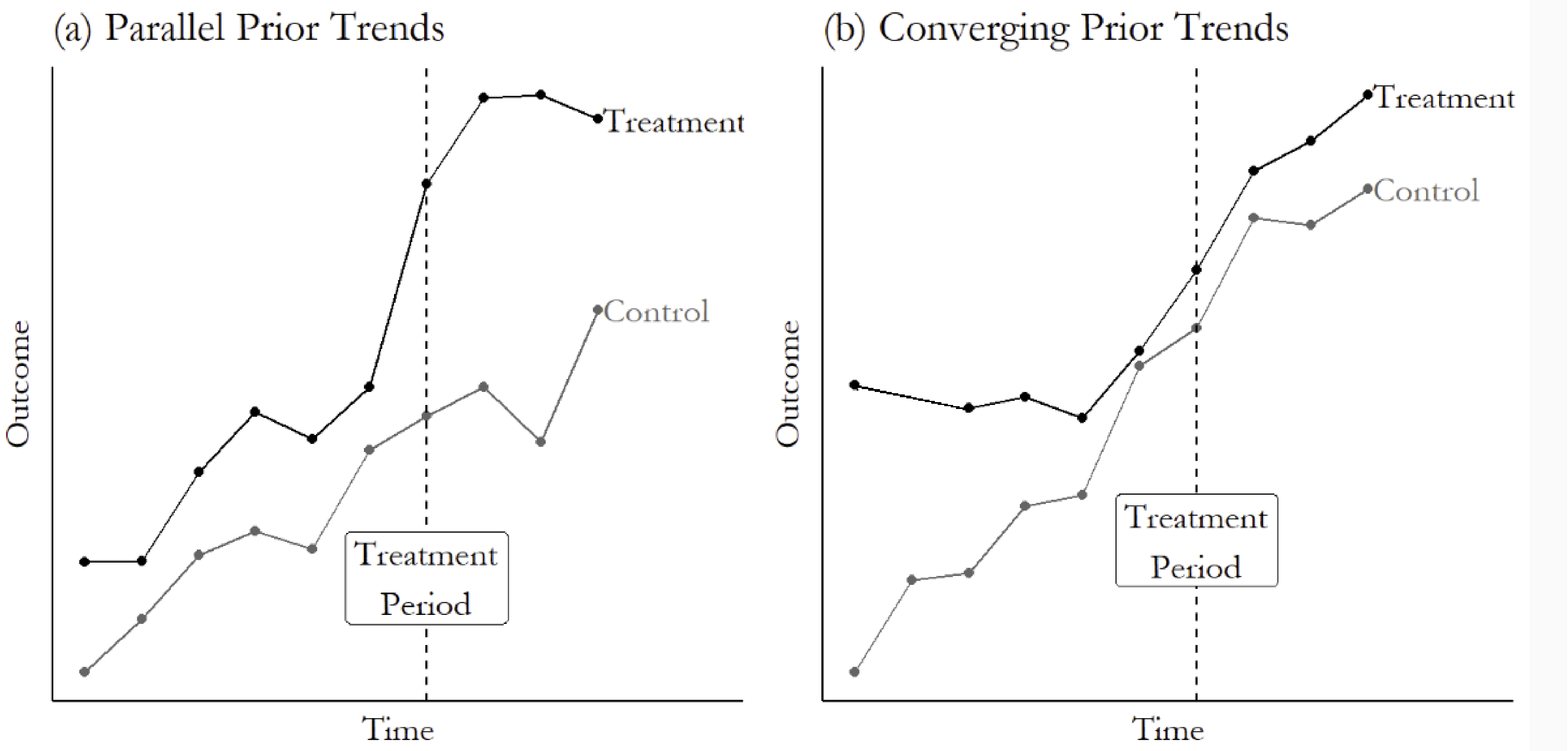
The parallel trends assumption is fundamentally untestable. We never observe the counterfactual trend.

A test that can bolster credibility: Are the trends parallel *prior to treatment*?

- *Logic:* If the trends are already different pre-treatment, why would they magically become similar post-treatment?

To conduct this test, can plot the mean of treated versus non-treated groups in each time period

Parallel vs. non-Parallel Pre-Trends



Evaluating Prior Trends in Our Data

Dynamic DiD Using the Event Study

Event Studies and Per-Period Effects

What if treatment effects might vary over time ($\tau_t \neq \tau_{t-1}$)?

- Effects may *increase over time* (e.g., learning, accumulation)
- Effects may *diminish over time* (e.g., forgetting, adaptation)
- Effects may also be constant

Oftentimes we want to understand these dynamics.

Event Study Design – A version of the DiD design that estimates per-period treatment effects before and after treatment

- "Effects" before treatment indicate a violation of the parallel trend assumption

The Event Study Regression Estimator

The Event Study Regression Estimator can be given as:

$$Y_{it} = \hat{\gamma}_i + \hat{\lambda}_t + \sum_{t=-T^{Pre}}^{T^{Post}} \hat{\tau}_t Post_{it} + \epsilon_{it}$$

where t is an "event time" integer s.t. treatment occurs at $t^* = 0$, and:

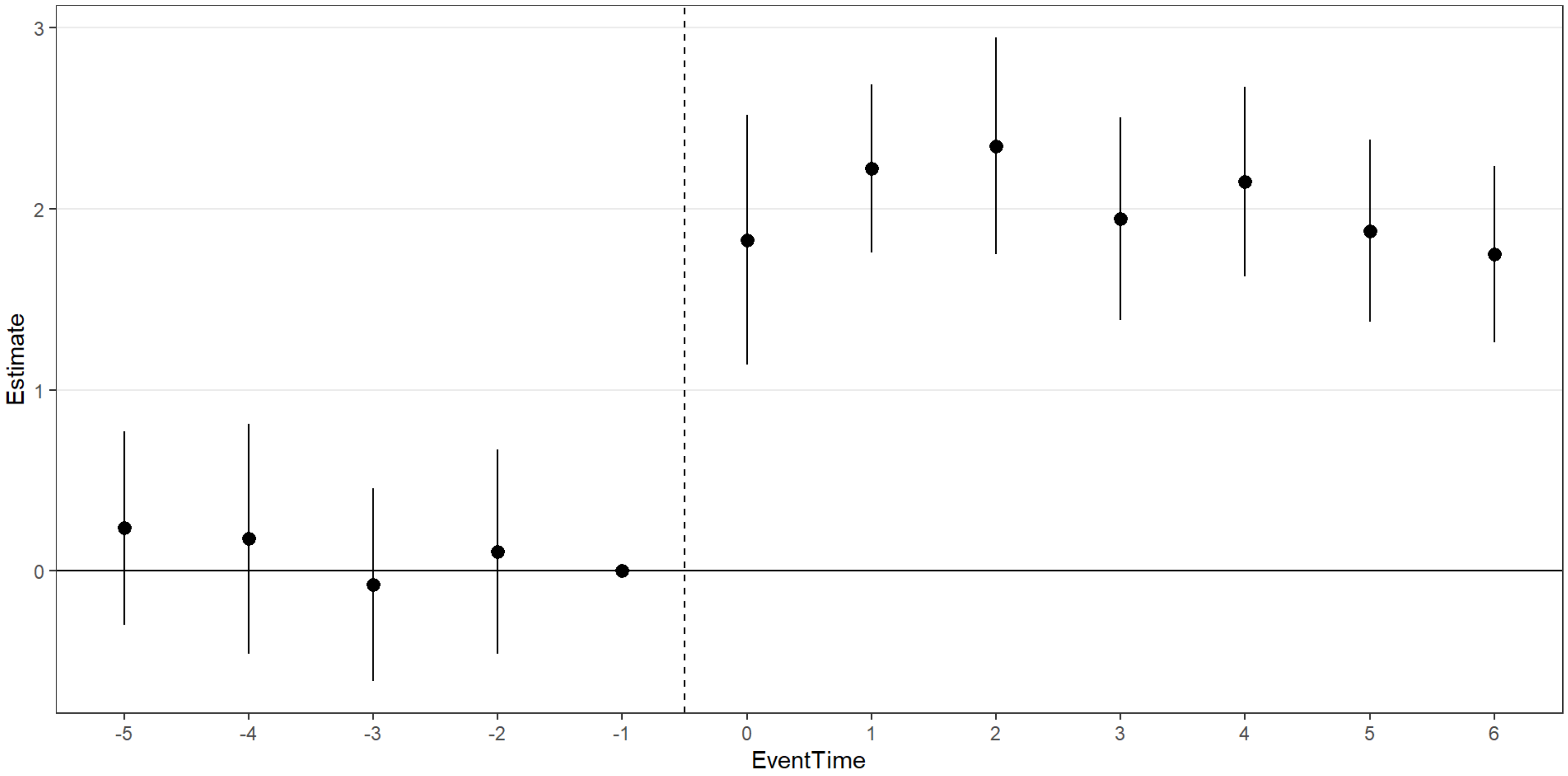
- $\hat{\gamma}_i$ adjusts for pre-existing differences between units
- $\hat{\lambda}_i$ adjusts for non-treatment time differences
- T^{Pre}, T^{Post} are the # of time periods before/after treatment to investigate
- $\hat{\tau}_t$ represent the per-period effects before and after treatment

An Event Study Analysis for the Class of ’04

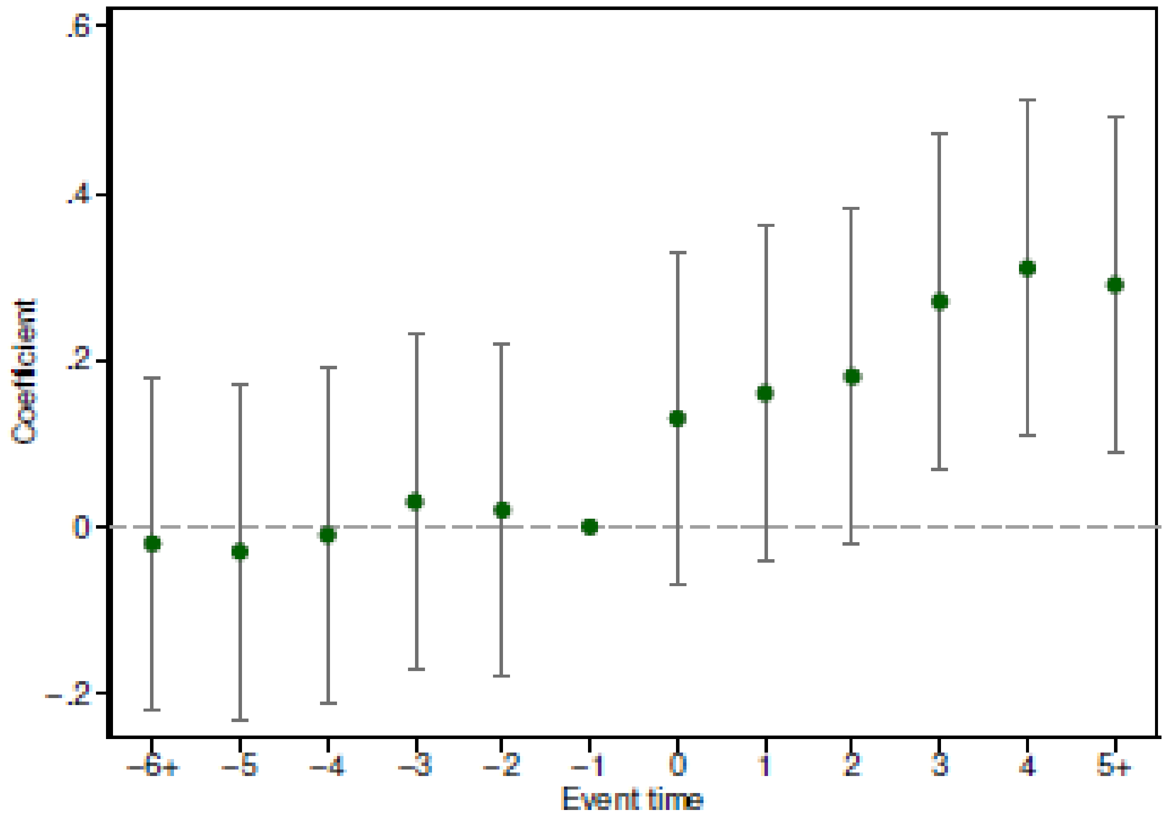
```r
eu_df <- mutate(eu_df,
                event_time = year - 2004,
                event_time = as.factor(event_time),
                event_time = relevel(event_time, "-1"))

reg_es <- feols(gdp_growth ~ treated * event_time |
                  country_id + event_time,
                data = eu_df)

summary(reg_es)
```

```
OLS estimation, Dep. Var.: gdp_growth
Observations: 300
Fixed-effects: country_id: 25,  event_time: 12
Standard-errors: Clustered (country_id)
                                  Estimate Std. Error   t value         Pr(>|t|)
treatedClass of '04:event_time-5  0.237093   0.272709  0.869400 0.3932405297795
treatedClass of '04:event_time-4  0.176915   0.323454  0.546954 0.5894601521499
treatedClass of '04:event_time-3 -0.077853   0.272010 -0.286214 0.7771701899975
treatedClass of '04:event_time-2  0.104705   0.287499  0.364192 0.7189034143898
treatedClass of '04:event_time0   1.828140   0.351637  5.198936 0.0000251321086
treatedClass of '04:event_time1   2.224783   0.236932  9.389950 0.0000000016562
treatedClass of '04:event_time2   2.347677   0.305209  7.692036 0.0000000627762
treatedClass of '04:event_time3   1.945820   0.285025  6.826841 0.0000004631523
treatedClass of '04:event_time4   2.150697   0.267253  8.047433 0.0000000284200
treatedClass of '04:event_time5   1.877966   0.257252  7.300098 0.0000001533779
treatedClass of '04:event_time6   1.748149   0.248023  7.048347 0.0000002751325
```

Plotting the Event Study Coefficients

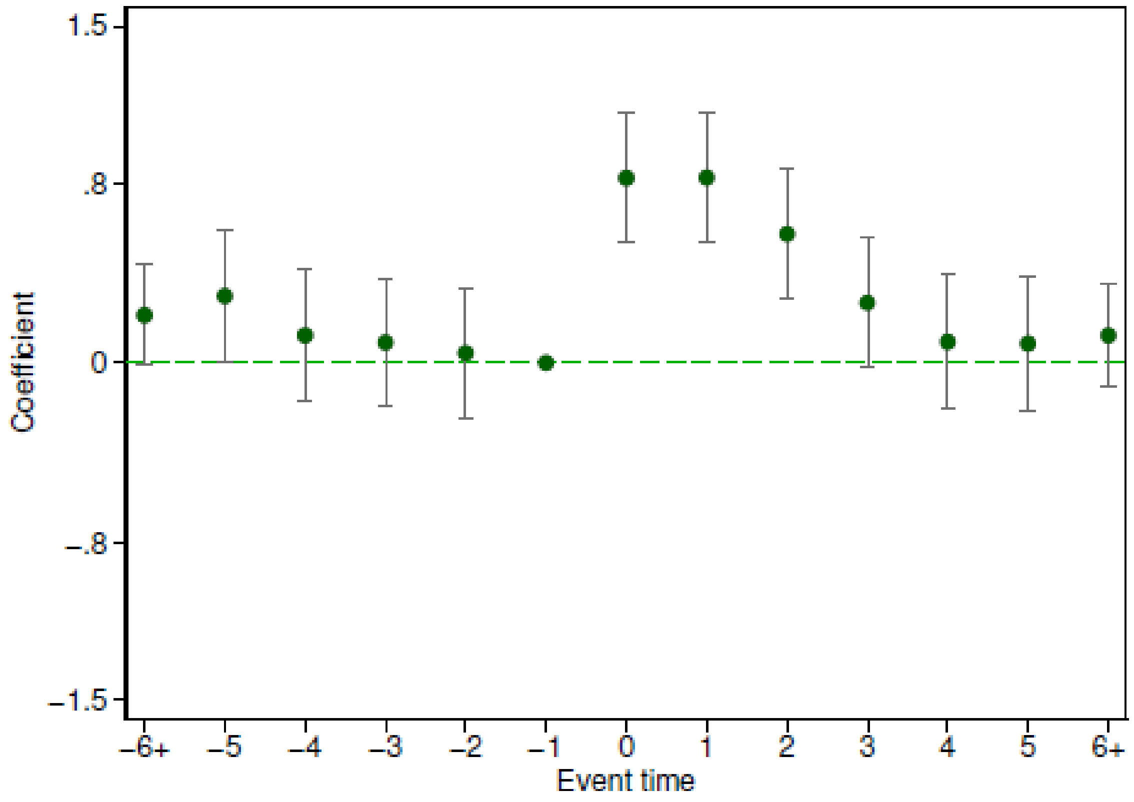


Another Example Event Study Plot (1/3)



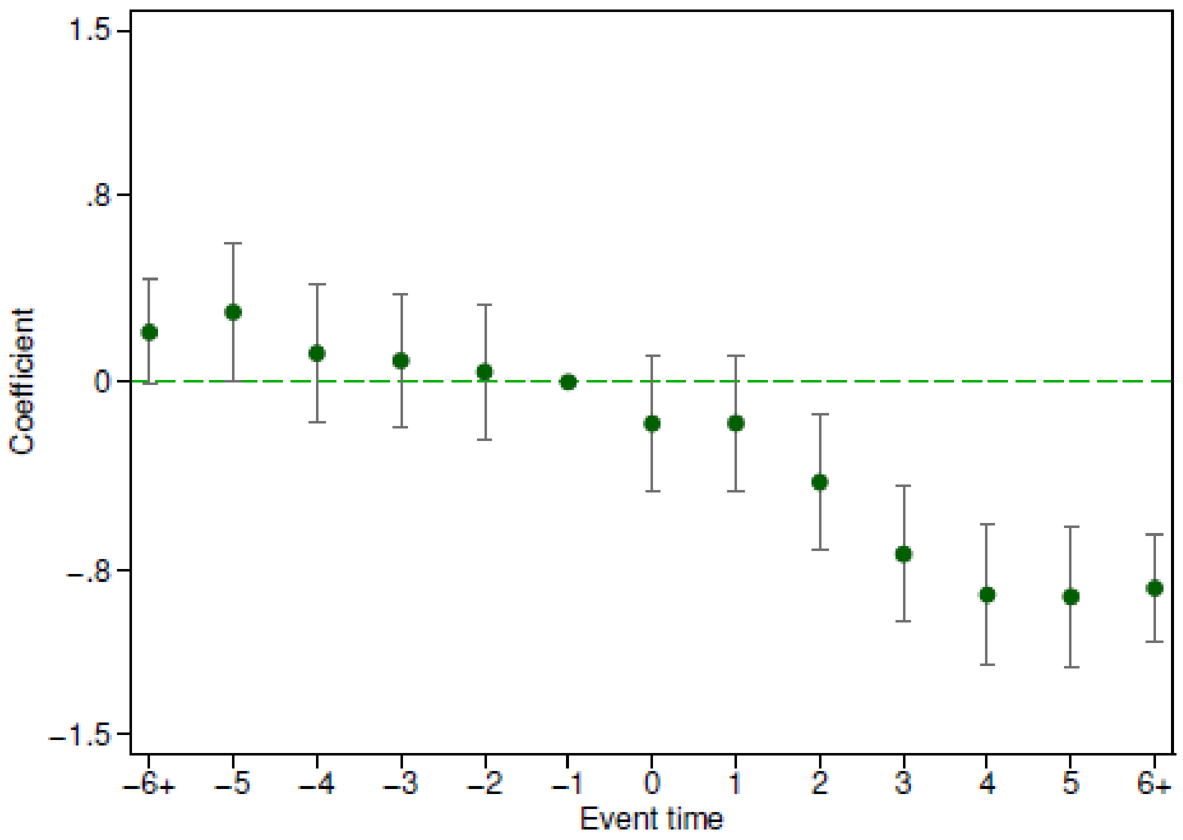
A Steadily Increasing Treatment Effect

Another Example Event Study Plot (2/3)



A Sudden Jump, Then Decaying Treatment Effect

Another Example Event Study Plot (3/3)



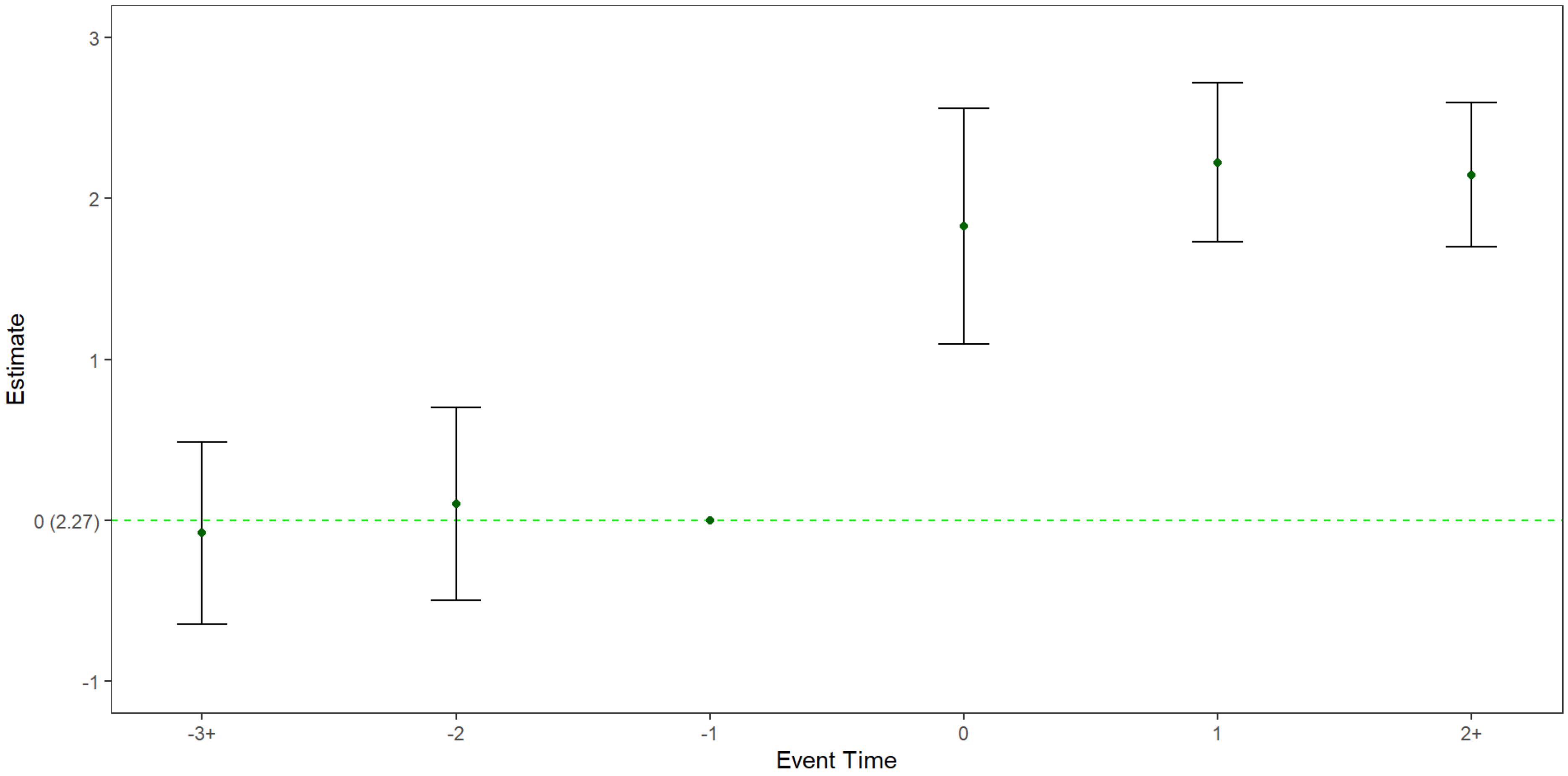
A Continuing Pre-Trend

Additional Event Study Plot Considerations

Some considerations when plotting an event study:

1. Because event time is relative, can use with a *staggered treatment*
    - Make sure time fixed effects (λ_t) are defined using actual time period, however
    - We'll have more to say about staggered treatment DiD's next week…
2. If T^{Pre} or T^{Post} are large, can combine time periods before/after a specified distance from treatment
3. For interpeting effect sizes, it's helpful to denote what the baseline (0) represents on the y-axis
    - A common choice: $\mathbf{E}[Y_{i,t-1}]$

Modifying the Class of ’04 ES Plot



Inference with the Event Study

Freyaldenhoven et. al. 2021 develop some tools for inference with the event study. These include:

1. Multiple hypothesis testing-adjusted confidence intervals
2. Hypothesis test of no pre-trend (and leveling-off of effects)
3. “Least Wiggly Path” sensitivity analysis

These authors created the `eventstudyr` package, which automatically estimates an event study, constructs the plot, and conducts inference.

- We’ll see how to use this in the data lab…

Adjusting for Multiple Hypothesis Testing

Researchers commonly plot (pointwise) 95% confidence intervals for the $\hat{\tau}_t$

- $H_0 : \tau_t = 0$

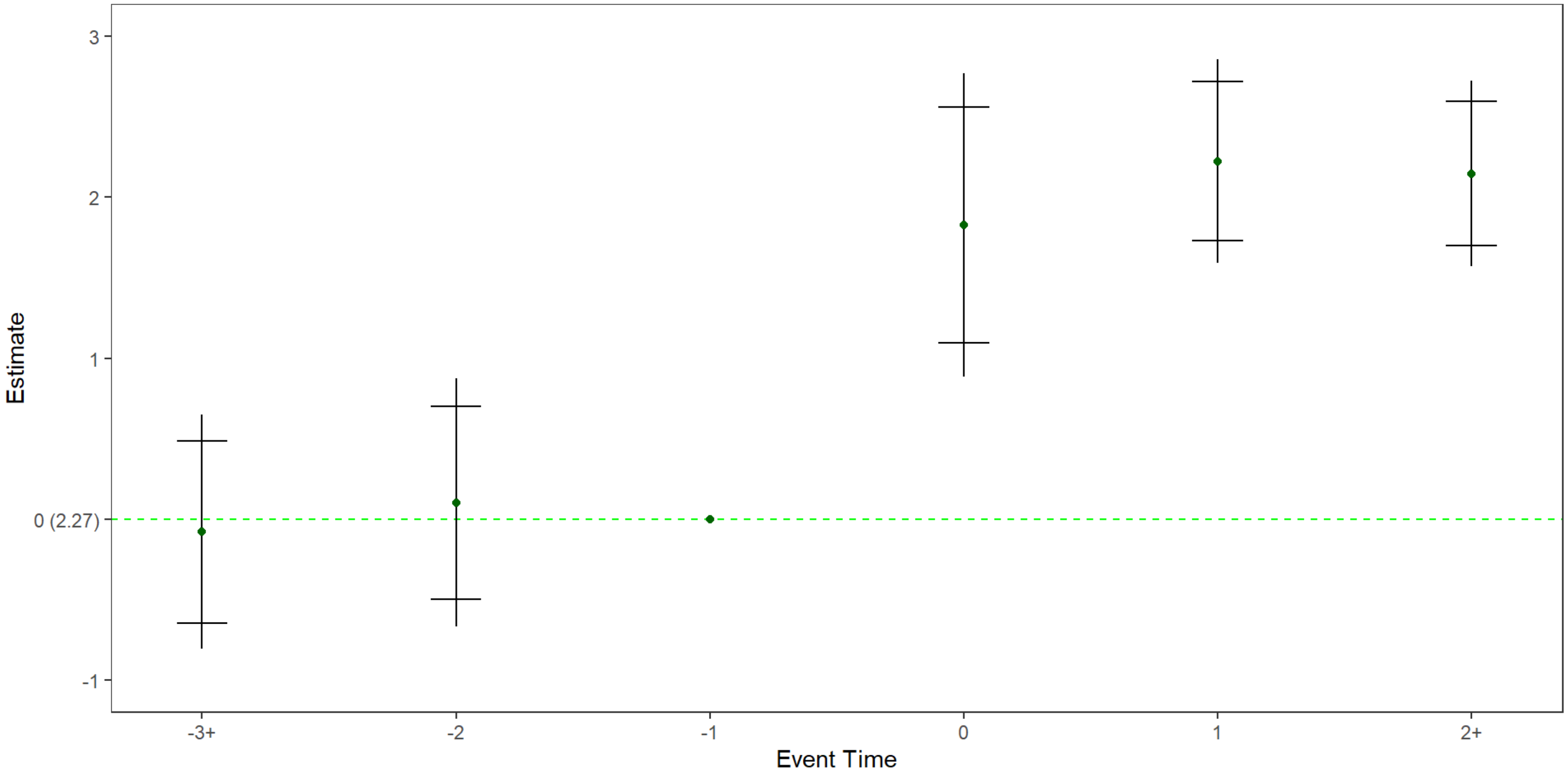
As # of pre-/post-periods increases, $\Pr(1+ \text{ Rejections})$ increases as well

- **Example:** With 10 time periods, $\Pr(1+ \text{ Rejections}|H_0) = 1 - 0.95^{10} > 0.4$

Instead, we can use *sup-t* confidence bands (for parameter set)

- $H_0 : \tau_t = 0 \quad \forall \quad t$
- Event-time paths that do not pass through all *sup-t* confidence intervals are inconsistent with the data

Plotting *sup-t* confidence bands

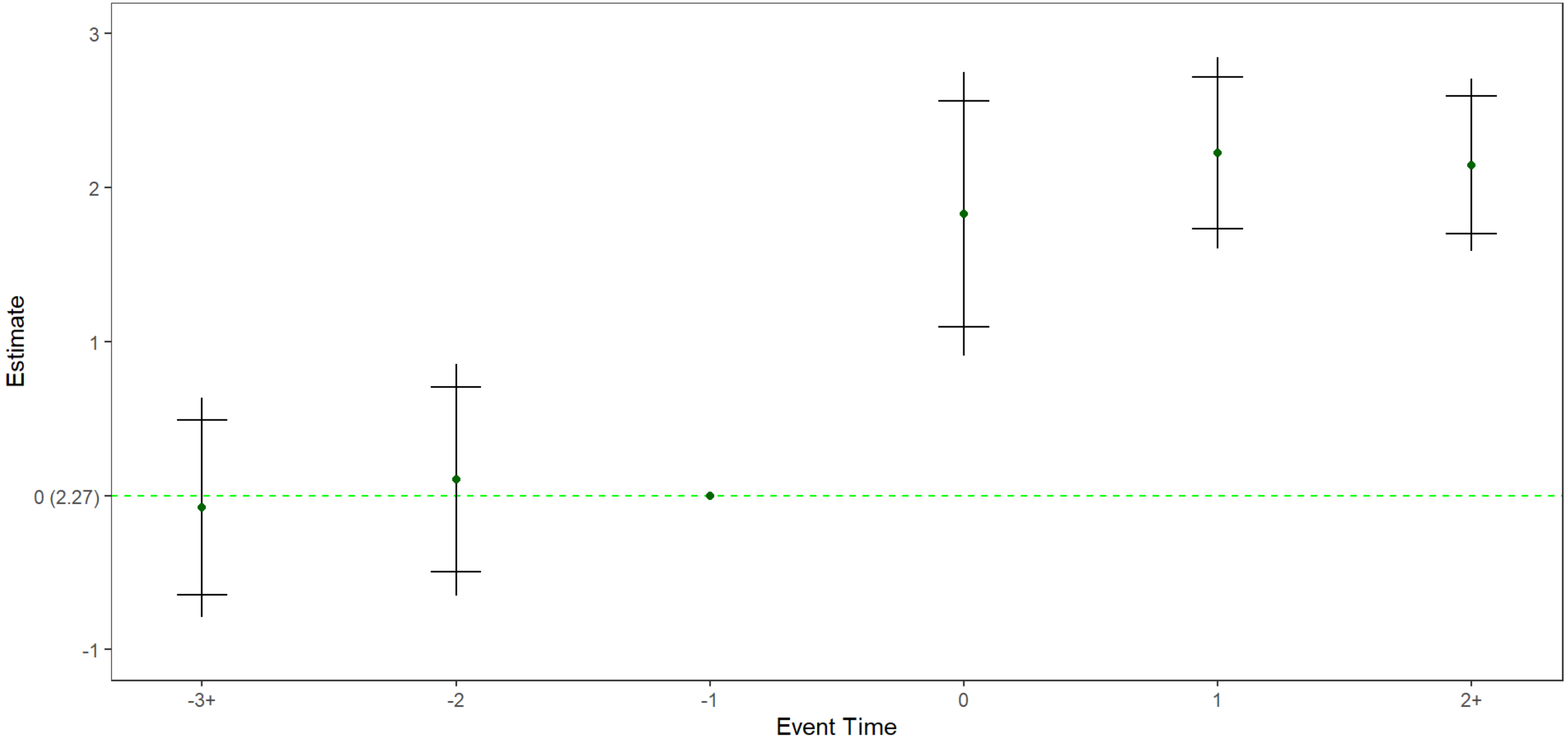


Conducting a Pre-Trend Hypothesis Test

Using the same principle, can conduct a hypothesis test for whether there are differing pre-trends

- **Null hypothesis:** outcome trend for treated units pre-treatment does not differ from outcome trend for untreated units ($H_0 : \tau_t = 0 \quad \forall \quad t < 0$)
- If the *p*-value of this test < 0.05, this is evidence pre-trends differ (DiD unlikely to give unbiased estimate of the ATT)

Displaying the Pre-Trend Hypothesis Test



Pretrends p-value = 0.83

“Least Wiggly Path” Sensitivity Analysis

Another way the DiD analysis can mislead: an omitted, time-varying confounder among the treated

- Must only (or differentially) affect treated units

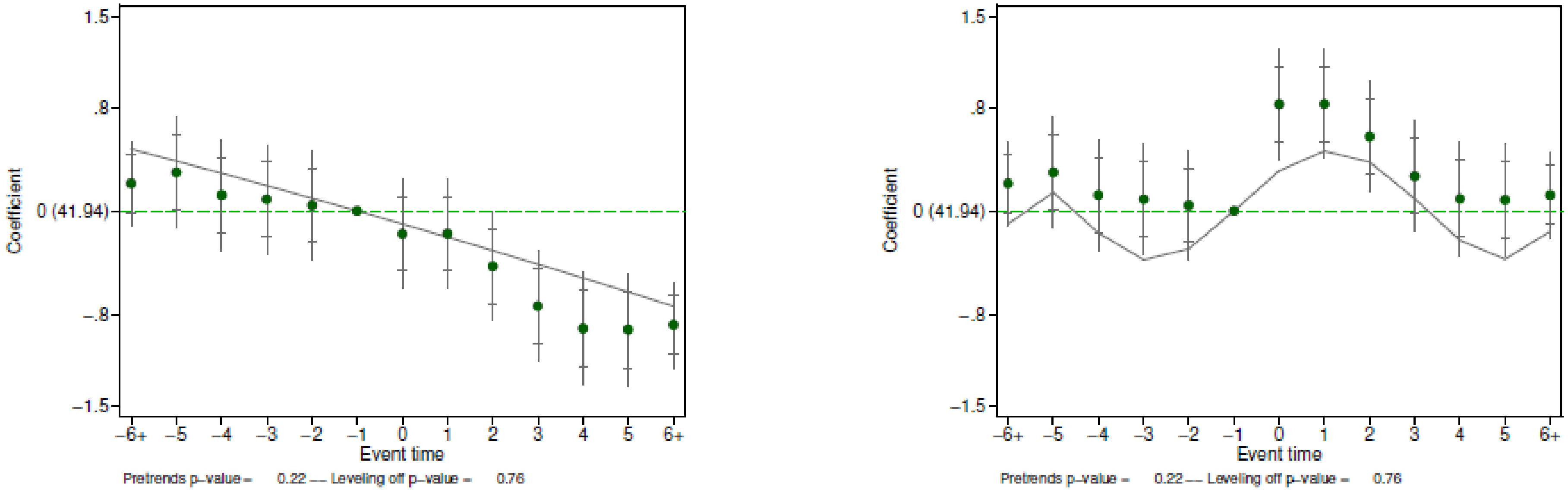
Event study paths provide insight into how unlikely confounding is

- *Intuition:* If a jump is observed exactly at (and only at) the time of treatment, a “just-so” confounding story is necessary

We can formalize this intuition by determining the smallest order p polynomial regression contained in the *sup-t* confidence band (the “least wiggly path”)

- How small does p have to be to be concerned? It depends…

How Wiggly Is The Path?



(a) "Smooth" event-time trend

(b) "Jump" at the time of the event

Figure 6: Least "wiggly" path of confound. Exemplary event-study plot for two possible datasets. Relative to Figure 5, a curve has been added that illustrates the least "wiggly" confound that is consistent with the event-time path of the outcome, in accordance with Suggestion 6.

How Wiggly Is *Our* Path?

```
Smoothest path note: The lowest order such that a polynomial is in confidence region is 4.
```