FIN 620
Emp. Methods in Finance

Lecture 6 – Natural Experiment *[P1]*

Professor Todd Gormley

Announcements

- If interested in optional exercises, material in Exercise #3 covers today's stuff but also a bit of IV material
- **Rough draft of proposal due after next class (at noon)**
  - Should upload to Canvas
  - I'll try to give feedback by next week

Background readings

- Roberts and Whited
  - *Sections 2.2, 4*
- Angrist and Pischke
  - *Section 5.2*

Outline for Today

- Quick review of IV regressions
- Discuss natural experiments
  - How do they help?
  - What assumptions are needed?
  - What are their weaknesses?
- Student presentations of "IV" papers

Quick Review *[Part 1]*

- Two necessary conditions for an IV
  - **Relevance condition** – IV explains problematic regressor after conditioning on other *x*'s
  - **Exclusion restriction** – IV does not explain *y* after conditioning on other *x*'s
- We can only test relevance condition

Quick Review *[Part 2]*

- Angrist (1990) used randomness of Vietnam draft to study effect of military service on Veterans' earnings
  - Person's draft number (which was random) predicted likelihood of serving in Vietnam
  - He found, using draft # as IV, that serving in military reduced future earnings

Question: What might be a concern about the <u>external</u> validity of his findings, and why?

Quick Review *[Part 3]*

- **Answer =** IV only identifies effect of serving on those that served <u>because</u> of being drafted
  - E.g., his finding doesn't necessarily tell us what the effect of serving is for people that would serve *regardless* of whether they are drafted or not
  - Must keep this **local average treatment effect** (LATE) in mind when interpreting IV

Quick Review *[Part 4]*

- **Question:** Are more instruments necessarily a good thing? If not, why not?
  - **Answer** = …

Quick Review *[Part 5]*

- **Question:** How can overidentification tests be used to prove the IVs are valid?
  - **Answer** = Trick question! They cannot be used in such a way. They rely on the assumption that at least one IV is good. You must provide a convincing economic argument as to why your IVs make sense!

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
- Two types of simple differences
- Difference-in-differences

Recall… CMI assumption is key

- A violation of conditional mean independence (CMI), such that $E(u|x) \neq E(u)$ precludes our ability to make causal inferences

$$y = \beta_0 + \beta_1 x + u$$

  - $\text{Cov}(x,u) \neq 0$ implies CMI is violated

CMI violation implies non-randomness

- Another way to think about violation is that it indicates that our *x* is non-random
  - **I.e., the distribution of *x* (or the distribution of *x* after controlling for other observable covariates) isn't random**
    - E.g., firms with high *x* might have higher *y* (beyond just the effect of *x* on *y*) because high *x* is more likely for firms with some omitted variable contained in *u*…

Randomized experiments are great…

- In many of the "hard" sciences, the researcher can simply design experiment to achieve the necessary randomness
  - Ex. #1 – To determine effect of new drug, you randomly give it to certain patients
  - Ex. #2 – To determine effect of certain gene, you modify it in a random sample of mice

But we simply can't do them ☹

- We can't do this in corporate finance!
  - E.g., we can't randomly assign a firm's leverage to determine its effect on investment
  - And we can't randomly assign CEOs' # of options to determine their effect on risk-taking
- Therefore, we need to rely on what we call "Natural experiments"

Defining a Natural Experiment

- Natural experiment is basically when some event causes a random assignment of (or change in) a variable of interest, x
  - Ex. #1 – Some weather event increases leverage for a random subset of firms
  - Ex. #2 – Some change in regulation reduces usage of options at a random subset of firms

Nat. Experiments Provide Randomness

- We can use such "natural" experiments to ensure that randomness (i.e., CMI) holds and make causal inferences!
  - E.g., we use the randomness introduced into x by the natural experiment to uncover the causal effect of x on y

NEs can be used in many ways

- Technically, natural experiments can be used in many ways
  - Use them to construct IV
    - E.g., gender of first child being a boy used in Bennedsen, et al. (2007) is an example NE
  - Use them to construct regression discontinuity
    - E.g., cutoff for securitizing loans at credit score of 620 used in Keys, et al. (2010) is a NE

And the Difference-in-Differences…

- But admittedly, when most people refer to natural experiment, they are talking about a difference-in-differences (DiD) estimator
  - Basically, compares outcome y for a "treated" group to outcome y for "untreated" group where treatment is randomly assigned by the natural experiment
  - **This is how I'll use NE in this class**

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
  - Notation and definitions
  - Selection bias and why randomization matters
  - Regression for treatment effects
- Two types of simple differences
- Difference-in-differences

Treatment Effects

- Before getting into natural experiments in context of difference-in-differences, it is first helpful to describe "treatment effects"

Notation and Framework

- Let d equal a treatment indicator from the experiment we will study
  - $d = 0 \rightarrow$ untreated by experiment *(i.e., control group)*
  - $d = 1 \rightarrow$ treated by experiment *(i.e., treated group)*
- Let y be the potential outcome of interest
  - $y = y(0)$ for untreated group
  - $y = y(1)$ for treated group
  - Easy to show that $y = y(0) + d[y(1) - y(0)]$

Example treatments in corp. fin…

- Ex. #1 – Treatment might be that your firm's state passed anti-takeover law
  - $d = 1$ for firms incorporated in those states
  - y could be several things, e.g., ROA
- Ex. #2 – Treatment is that your firm discovers workers exposed to carcinogen
  - $d = 1$ if have exposed workers
  - y could be several things, like M&A

Average Treatment Effect (ATE)

- Can now define some useful things
  - **Average Treatment Effect** (ATE) is given by

$$\mathbf{E}[y(1) - y(0)]$$

    - What does this mean in words?
    - **Answer:** The expected change in y from being treated by the experiment; **this is the causal effect we are typically interested in uncovering!**

But ATE is unobservable

$$\mathbf{E}[y(1) - y(0)]$$

- Why can't we directly observe ATE?
  - **Answer** = We only observe one outcome…
    - If treated, we observe *y(1)*; if untreated, we observe *y(0)*. We never observe both.
    - E.g., we cannot observe the counterfactual of what your *y* would have been absent treatment

Defining ATT

- **Average Treatment Effect if Treated (ATT)** is given by $\mathbf{E}[y(1) - y(0) \mid d = 1]$
  - This is the effect of treatment on those that are treated; i.e., change in y we'd expect to find in treated random sample from a population of observations that are treated
  - **What don't we observe here?**
  - **Answer** $= y(0) \mid d = 1$

Defining ATU

- **Average Treatment Effect if Untreated (ATU)** is given by $\mathbf{E}[y(1) - y(0) \mid d = 0]$
  - This is what the effect of treatment would have been on those that are not treated by the experiment
  - We don't observe $y(1) \mid d = 0$

Uncovering ATE *[Part 1]*

- How do we estimate ATE, $E[y(1) - y(0)]$?
  - **Answer =** We instead rely on $E[y(1) \mid d = 1] - E[(y(0) \mid d = 0]$ as our way to *infer* the ATE

In words, what are we doing & what are we assuming?

Uncovering ATE *[Part 2]*

- In words, we compare average y of treated to average y of untreated observations
  - If we interpret this as the ATE, we are <u>assuming</u> that absent the treatment, the treated group would, on average, have had same outcome y as the untreated group
  - We can show this formally by simply working out $E[y(1) \mid d = 1] - E[y(0) \mid d = 0]$…

Uncovering ATE *[Part 3]*

$$\{E[y(1)\mid d=1]-E[y(0)\mid d=1]\}+\{E[y(0)\mid d=1]-E[y(0)\mid d=0]\}$$

First bracket is <u>ATT</u>

Just added and subtracted the same term

Second bracket is what we call the "selection bias"

- Simple comparison doesn't give us the ATE! In fact, the comparison is rather meaningless!
- What is the **"selection bias"** in words?

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
  - Notation and definitions
  - Selection bias and why randomization matters
  - Regression for treatment effects
- Two types of simple differences
- Difference-in-differences

Selection bias defined

- Selection bias: $E[y(0) | d = 1] - E[y(0) | d = 0]$
  - **Definition =** What the difference in average y would have been for treated and untreated observations <u>absent</u> any treatment
  - *We do not observe this counterfactual!*
- Now let's see why randomness is key!

Introducing random treatment

- A random treatment, *d*, implies that *d* is independent of potential outcomes; i.e.,

$$E[y(0) \mid d=1] = E[y(0) \mid d=0] = E[y(0)] \longleftarrow$$

$$and$$

$$E[y(1) \mid d=1] = E[y(1) \mid d=0] = E[y(1)]$$

In words, the expected value of *y* is the same for treated and untreated absent treatment

  - With this, easy to see that selection bias = 0
  - **And** remaining ATT is equal to ATE!

Random treatment makes life easy

- I.e., with random assignment of treatment, our simple comparison gives us the ATE!
  - This is why we like randomness!
  - But, absent randomness, we must worry that any observed difference is driven by selection bias

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
  - Notation and definitions
  - Selection bias and why randomization matters
  - Regression for treatment effects
- Two types of simple differences
- Difference-in-differences

ATE in Regression Format *[Part 1]*

- Can re-express everything in regression format

$$y = \beta_0 + \beta_1 d + u$$

This regression will only give consistent estimate of β_1 if cov(d, u) = 0; i.e., treatment, d, is random, and hence, uncorrelated with $y(0)$!

$$\text{where} \quad \begin{aligned} \beta_0 &= E[y(0)] \\ \beta_1 &= y(1) - y(0) \\ u &= y(0) - E[y(0)] \end{aligned}$$

  - If you plug-in, it will get you back to what the true model, $y = y(0) + d[y(1) - y(0)]$

ATE in Regression Format *[Part 2]*

- We are interested in $\mathrm{E}[y \mid d=1] - \mathrm{E}[y \mid d=0]$
  - But can easily show that this expression is equal to

$$\beta_1 + E[y(0) \mid d=1] - E[y(0) \mid d=0]$$

Our estimate will equal true effect plus selection bias term

Note: Selection bias term occurs only if CMI isn't true!

Adding additional controls *[Part 1]*

- Regression format also allows us to easily put in additional controls, X
  - Intuitively, comparison of treated and untreated just becomes $E[y(1) \mid d=1, X] - E[y(0) \mid d=0, X]$
  - Same selection bias term will appear if treatment, d, isn't random after conditioning on X
  - Regression version just becomes

$$y = \beta_0 + \beta_1 d + \Gamma X + u$$

Why might there still be a selection bias?

Adding additional controls *[Part 2]*

- Selection bias can still be present if treatment is correlated with unobserved variables
  - As we saw earlier, it is what we can't observe (and control for) that can be a problem!

Question: If we had truly randomized experiment, are controls necessary?

Adding additional controls *[Part 3]*

- **Answer:** No, controls are not necessary in truly randomized experiment
  - But they can be helpful in making the estimates more precise by absorbing residual variation… we'll talk more about this later

Treatment effect – *Example*

- Suppose compare leverage of firms with and without a credit rating [or equivalently, regress leverage on indicator for rating]
  - Treatment is having a credit rating
  - Outcome of interest is leverage

Why might our estimate not equal ATE of rating?

Why might controls not help us much?

Treatment effect – *Example Answer*

- **Answer #1:** Having a rating isn't random
  - Firms with rating likely would have had higher leverage anyway because they are larger, more profitable, etc.; selection bias will be positive
  - Selection bias is basically an omitted var.!
- **Answer #2:** Even adding controls might not help if firms also differ in <u>unobservable</u> ways, like investment opportunities

Heterogeneous Effects

- Allowing the effect of treatment to vary across individuals doesn't affect much
  - Just introduces additional bias term
  - Will still get ATE if treatment is random… broadly speaking, randomness is key

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
- Two types of simple differences
  - Cross-sectional difference & assumptions
  - Time-series difference & assumptions
  - Miscellaneous issues & advice
- Difference-in-differences

We actually just did this one!

Cross-sectional Simple Difference

- Very intuitive idea
  - Compare <u>post</u>-treatment outcome, *y*, for treated group to the untreated group
  - I.e., just run following regression…

In regression format…

- Cross-section simple difference

$$y_{i,t} = \beta_0 + \beta_1 d_i + u_{i,t}$$

  - $d = 1$ if observation i is in treatment group and equals zero otherwise
  - Regression only contains <u>post</u>-treatment time periods

What is needed for β_1 to capture the true (i.e., causal) treatment effect?

Identification Assumption

- **Answer:** $E(u \mid d) = 0$; i.e., treatment, d, is uncorrelated with the error
  - In words… after accounting for effect of treatment, the expected level of y in post-treatment period isn't related to whether you're in the treated or untreated group
  - ***I.e.***, expected y of treated group <u>would have been</u> same as untreated group *absent* treatment

Another way to see the assumption…

$$E[y \mid d=1] - E[y \mid d=0]$$

This is causal interpretation of coefficient on ***d***

$$\left(\beta_0 + \beta_1 + E[u \mid d=1]\right) - \left(\beta_0 + E[u \mid d=0]\right)$$

$$\beta_1 + E[u \mid d=1] - E[u \mid d=0]$$

CMI assumption ensures these last two terms cancel such that our interpretation matches causal β_1

- Then, plugging in for $u = y(0) - \mathrm{E}[y(0)]$, which is what true error is (see earlier slides)…

$$\beta_1 + E[y(0) \mid d=1] - E[y(0) \mid d=0]$$

I.e., we must assume no selection bias

Multiple time periods & SEs

- If have multiple post-treatment periods, need to be careful with standard errors
  - Errors $u_{i,t}$ and $u_{i,t+1}$ likely correlated if dependent variable exhibits serial correlation
    - E.g., we observe each firm (treated and untreated) for five years after treatment (e.g., regulatory change), and our post-treatment observations are not independent

Multiple time periods & SEs – *Solution*

- Should do one of two things
  - Collapse data to one post-treatment per unit; e.g., for each firm, use average of the firm's post-treatment observations
  - Or cluster standard errors at firm level *[We will come back to clustering in later lecture]*

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
- Two types of simple differences
  - Cross-sectional difference & assumptions
  - Time-series difference & assumptions
  - Miscellaneous issues & advice
- Difference-in-differences

Time-series Simple Difference

- Very intuitive idea
  - Compare pre- and post-treatment outcomes, *y*, for just the treated group *[i.e., pre-treatment period acts as 'control' group]*
  - I.e., run following regression…

In Regression Format

- Time-series simple difference

$$y_{i,t} = \beta_0 + \beta_1 p_t + u_{i,t}$$

  - $p_t = 1$ if period t occurs after treatment and equals zero otherwise
  - Regression contains only observations that are treated by "experiment"

What is needed for β_1 to capture the true (i.e., causal) treatment effect?

Identification Assumption

- **Answer:** $\mathrm{E}(u|p) = 0$; i.e., post-treatment indicator, p, is uncorrelated with the error
  - I.e., after accounting for effect of treatment, p, the expected level of y in post-treatment period wouldn't have been any different than expected y in pre-treatment period

Showing the assumption math…

This would be causal interpretation of coefficient on *p*

$$E[y \mid p=1] - E[y \mid p=0]$$

$$\left(\beta_0 + \beta_1 + E[u \mid p=1]\right) - \left(\beta_0 + E[u \mid p=0]\right)$$

$$\beta_1 + E[u \mid p=1] - E[u \mid p=0]$$

$$\beta_1 + E[y(0) \mid p=1] - E[y(0) \mid p=0]$$

Same selection bias term… our estimated coefficient on *p* only matches true causal effect if this is zero

Again, be careful about SEs

- Again, if you have multiple pre- and post-treatment periods, you need to be careful with estimating your standard errors
  - Either cluster SEs at level of each unit
  - Or collapse data down to one pre- and one post-treatment observation for each cross-section

Using a First-Difference (FD) Approach

- Could also run regression using first-differences specification

$$y_{i,t} - y_{i,t-1} = \beta_1 \left(p_t - p_{t-1} \right) + \left(u_{i,t} - u_{i,t-1} \right)$$

  - If just one pre- and one post-treatment period (i.e., *t-1* and *t*), then will get identical results
  - But, if more than one pre- and post-treatment period, the results will differ…

FD *versus* Standard Approach *[Part 1]*

- Why might these two models give different estimates of β_1 when there are more than one pre- and post-treatment periods?

$$y_{i,t} = \beta_0 + \beta_1 p_t + u_{i,t}$$

versus

$$y_{i,t} - y_{i,t-1} = \beta_1 \left(p_t - p_{t-1} \right) + \left(u_{i,t} - u_{i,t-1} \right)$$

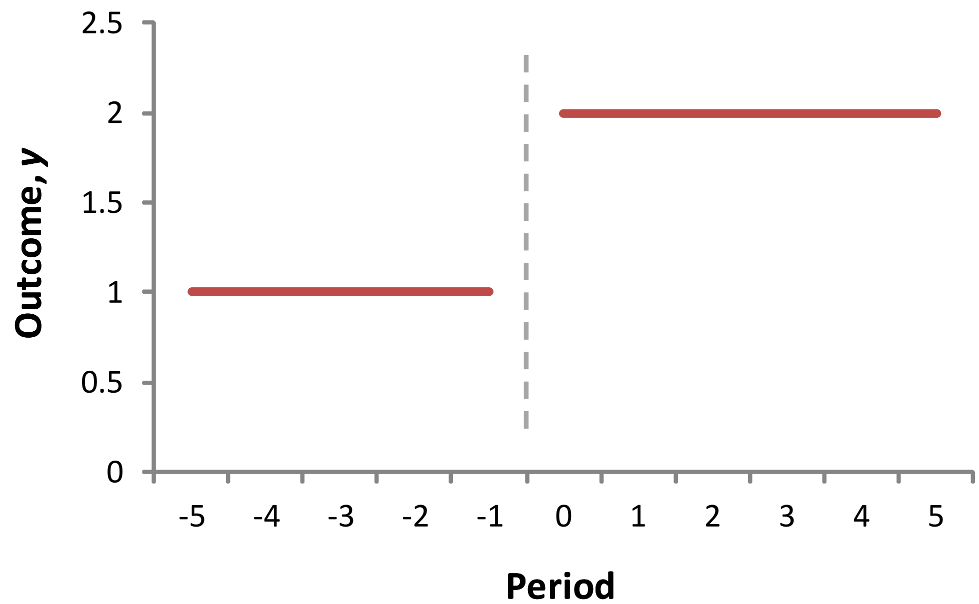
FD *versus* Standard Approach *[Part 2]*

How might this matter in practice?

- **Answer:**
  - In 1st regression, β_1 captures difference between avg. *y* pre-treatment *versus* avg. *y* post-treatment
  - In 2nd regression, β_1 captures difference in Δy immediately after treatment versus Δy in all other pre- and post-treatment periods
    - I.e., the Δp variable equals 1 only in immediate post-treatment period, and 0 for all other periods

FD *versus* Standard Approach *[Part 3]*

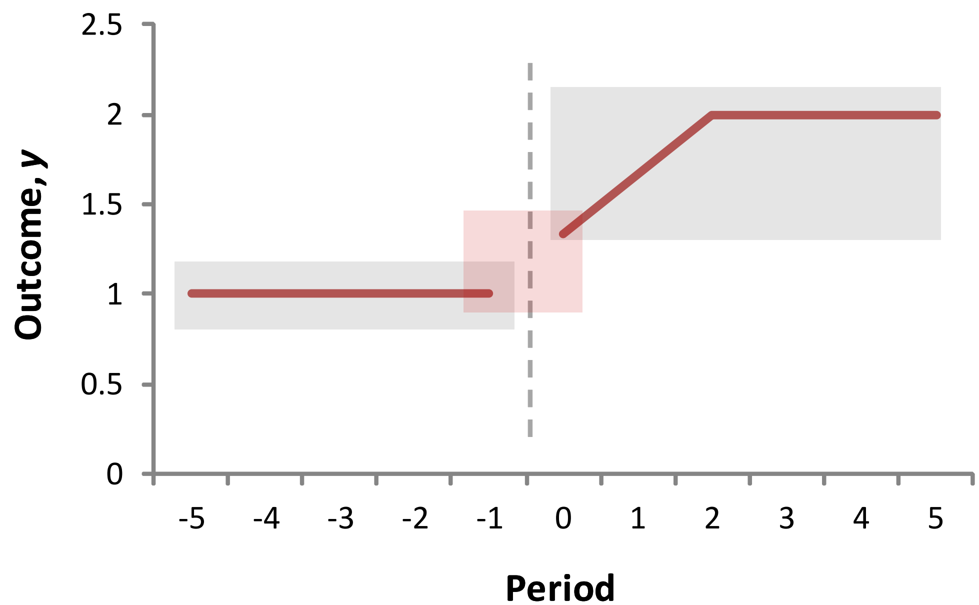
- Both approaches assume the effect of treatment is immediate and persistent, e.g.



In this scenario, both approaches give same estimate

FD *versus* Standard Approach *[Part 4]*

- But suppose the following is true…



In this scenario, FD approach gives much smaller estimate

1st approach compares avg. pre- versus post

FD compares Δy from t =0 to t =-1 against Δy elsewhere (which isn't always zero!)

Correct way to do difference

- Correct way to get a 'differencing' approach to match up with the more standard simple diff specification in multi-period setting is to instead use

$$\bar{y}_{i,post} - \bar{y}_{i,pre} = \beta_1 + \left(\bar{u}_{i,post} - \bar{u}_{i,pre}\right)$$

  - This is exactly the same as simple difference

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
- Two types of simple differences
  - Cross-sectional difference & assumptions
  - Time-series difference & assumptions
  - Miscellaneous issues & advice
- Difference-in-differences

Treatment effect isn't always immediate

- In prior example, the specification is wrong because the treatment effect only slowly shows up over time
  - Why might such a scenario be plausible?
  - **Answer =** Many reasons. E.g., firms might only slowly respond to change in regulation, or CEO might only slowly change policy in response to compensation shock

Accounting for a delay…

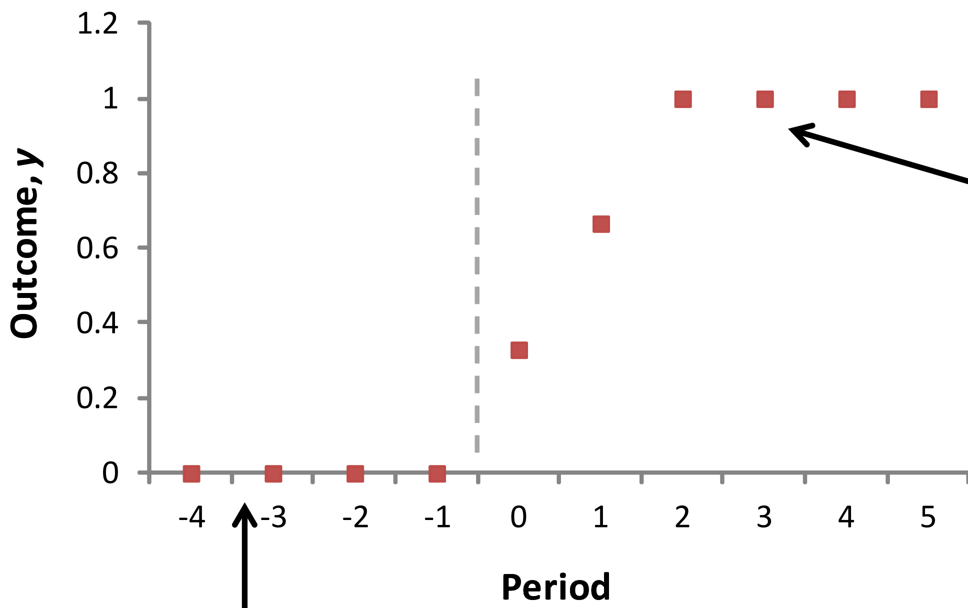
- Simple-difference misses this subtlety; it assumes effect was immediate
- For this reason, it is always helpful to run regression that allows effect to vary by period
  - **How can you do this?**
  - **Answer =** Insert indicators for each year relative to the treatment year *[see next slide]*

Non-parametric approach

- If have 5 pre- and 5 post-treatment obs.; could estimate : $y_{i,t} = \beta_0 + \sum_{t=-4}^{5} \beta_t p_t + u_{i,t}$
  - p_t is now an indicator that equals 1 if year = t and zero otherwise; e.g.
    - $t = 0$ is the period treatment occurs
    - $t = -1$ is period before treatment
  - β_t estimates change in y relative to <u>excluded</u> periods; you then plot these in graph

Non-parametric approach – *Graph*

- Plot estimates to trace out effect of treatment



These equal zero because *y* was same as *y* in excluded period (*t*-5)

Approach allows effect of treatment to vary by year!

Estimates capture change relative to excluded period (*t*-5)

Could easily plot confidence intervals as well

Simple Differences – *Advice*

- In general, simple differences are not that convincing in practice…
  - Cross-sectional difference requires us to assume the average *y* of treated and untreated would have been same absent treatment
  - Time-series difference requires us to assume the average *y* would have been same in post- and pre-treatment periods absent treatment
- **Is there a better way?**

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
- Two types of simple differences
- Difference-in-differences
  - Intuition & implementation
  - "Parallel trends" assumption

Difference-in-differences

- Yes, we can do better!
- We can do a difference-in-differences that combines the two simple differences
  - **Intuition** = compare <u>change</u> in *y* pre- versus post-treatment for treated group *[1st difference]* to <u>change</u> in *y* pre- versus post-treatment for untreated group *[2nd difference]*

Implementing diff-in-diff

- Difference-in-differences estimator

$$y_{i,t} = \beta_0 + \beta_1 p_t + \beta_2 d_i + \beta_3 \left(d_i \times p_t\right) + u_{i,t}$$

  - $p_t = 1$ if period t occurs after treatment and equals zero otherwise
  - $d_i = 1$ if unit is in treated group and equals zero otherwise

What do β$_1$, β$_2$, and β$_3$ capture?

Interpreting the estimates *[Part 1]*

- **Here is how to interpret everything…**
  - β_1 captures the average change in y from the pre- to post-treatment periods that is <u>common</u> to both treated and untreated groups
  - β_2 captures the average difference in level of y of the treated group that is <u>common</u> to both pre- and post-treatment periods

Interpreting the estimates *[Part 2]*

- β_3 captures the average <u>differential</u> change in *y* from the pre- to post-treatment period for the treatment group ***relative*** to the change in *y* for the untreated group

- β_3 is what we call the diff-in-diffs estimate

When does β_3 capture the causal effect of the treatment?

Natural Experiments – *Outline*

- Motivation and definition
- Understanding treatment effects
- Two types of simple differences
- Difference-in-differences
  - Intuition & implementation
  - "Parallel trends" assumption

"Parallel trends" assumption

- Identification assumption is what we call the **parallel trends assumption**
  - Absent treatment, the change in y for treated would not have been different than the change in y for the untreated observations
    - To see why this is the underlying identification assumption, it is helpful to re-express the diff-in-diffs…

Differences estimation

- Equivalent way to do difference-in-differences is to instead estimate the following:

$$\bar{y}_{i,post} - \bar{y}_{i,pre} = \beta_0 + \beta_1 d_i + \left(\bar{u}_{i,post} - \bar{u}_{i,pre}\right)$$

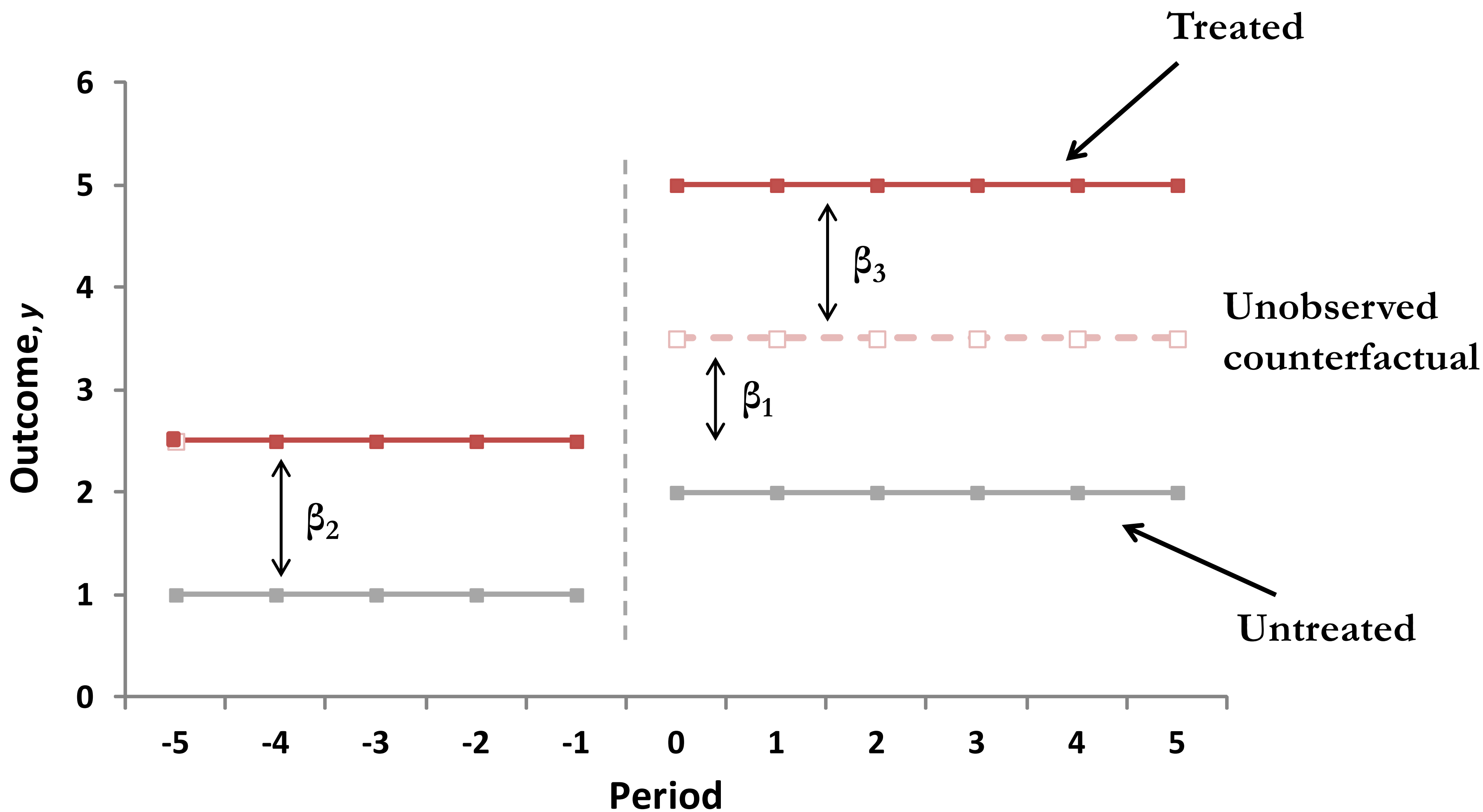
  - β_1 gives the difference-in-differences estimate
    - In practice, don't do this because an adjustment to standard errors is necessary to get right *t*-stat
    - And remember! This is not the same as taking first-differences; FD can give misleading results

Difference-in-differences – *Visually*

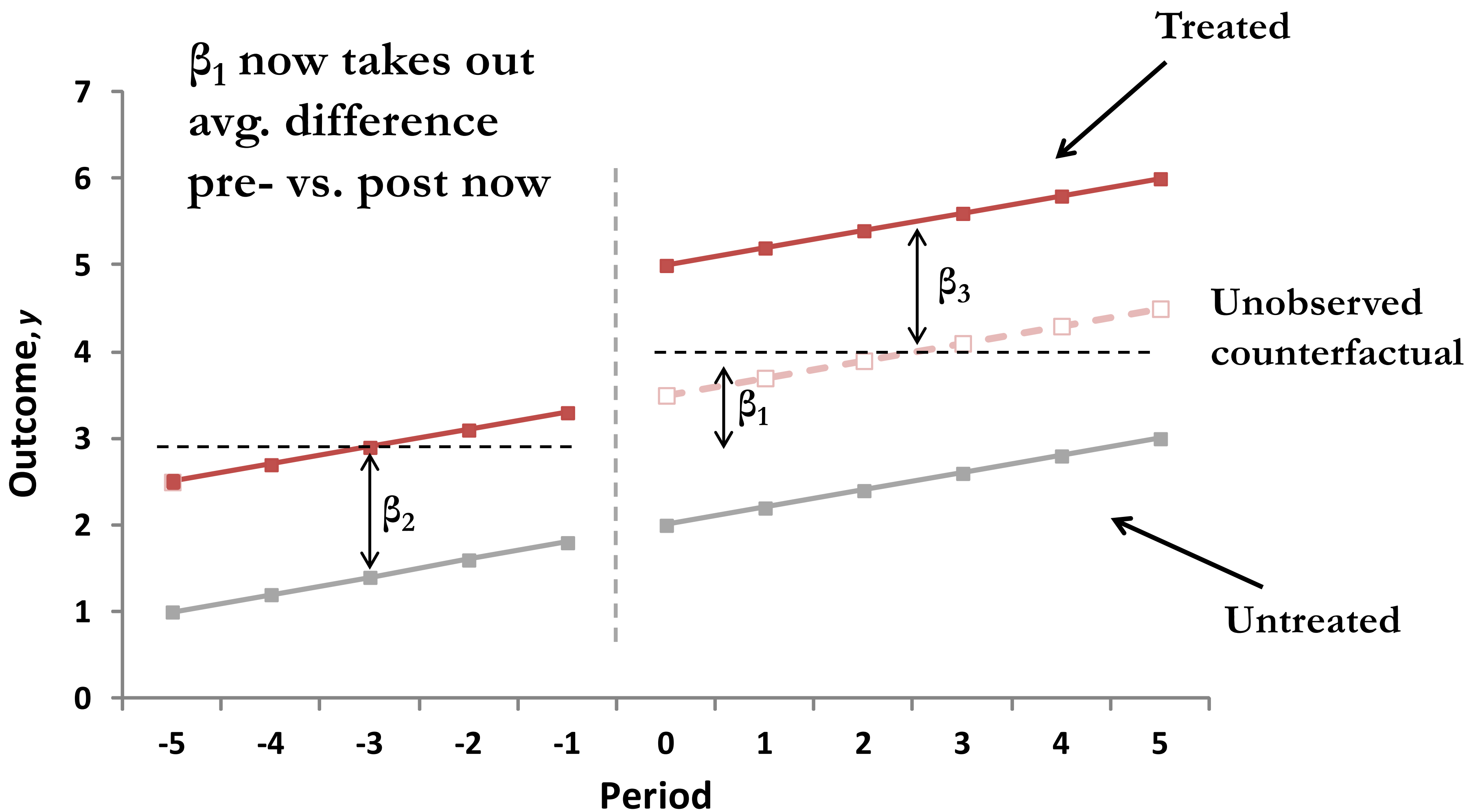
- Looking at what difference-in-differences estimate is doing in graphs will also help you see why the parallel trends assumption is key

$$y_{i,t} = \beta_0 + \beta_1 p_t + \beta_2 d_i + \beta_3 \left(d_i \times p_t\right) + u_{i,t}$$

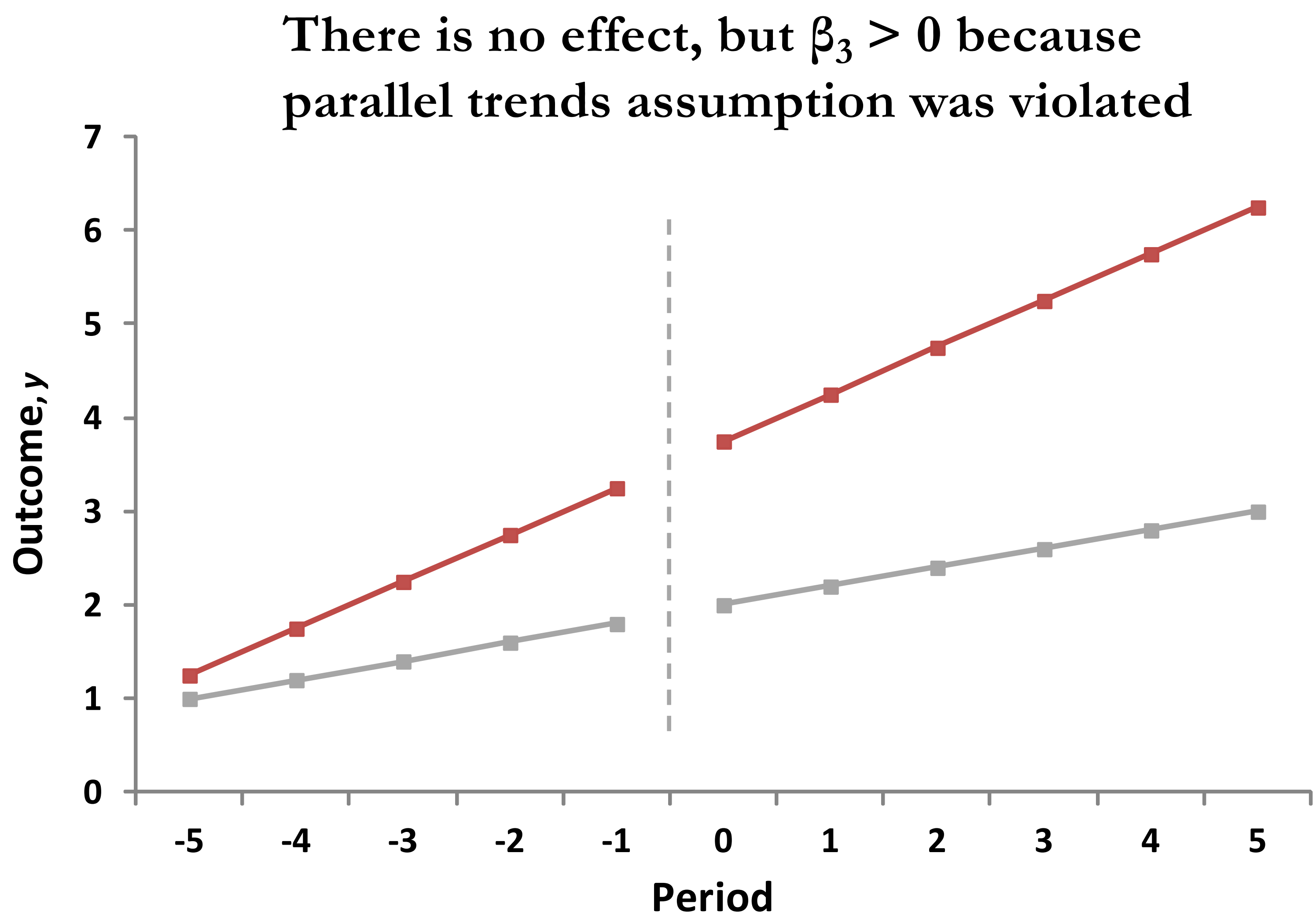
Diff-in-diffs – *Visual Example #1*



Diff-in-diff – *Visual Example #2*



Violation of parallel trends – *Visual*



Why we like diff-in-diff *[Part 1]*

- With simple difference, any of the below arguments would prevent causal inference
  - <u>Cross-sectional diff</u> – "Treatment and untreated avg. y could be different for reasons a, b, and c, that just happen to be correlated with whether you are treated or not"
  - <u>Time-series diff</u> – "Treatment group's avg. y could change post- treatment for reasons a, b, and c, that just happen to be correlated with the timing of treatment"

Why we like diff-in-diff *[Part 2]*

- But now the required argument to suggest the estimate isn't causal is…
  - "The change in *y* for treated observations after treatment would have been different than change in *y* for untreated observations for reasons a, b, and c, that just happen to be correlated with **both** whether you are treated and when the treatment occurs"

This is (usually) a harder story to tell

Example…

- Bertrand & Mullainathan (JPE 2003) uses state-by-state changes in regulations that made it harder for firms to do M&A
  - They compare wages at firms pre- versus post-regulation in treated versus untreated states
  - Are the below valid concerns about their difference-in-differences…

Are these concerns for internal validity?

- The regulations were passed during a time period of rapid growth of wages nationally…
  - **Answer = No.** Indicator for post-treatment accounts for common growth in wages
- States that implement regulation are more likely have unions, and hence, higher wages…
  - **Answer =** No. Indicator for treatment accounts for this average difference in wages

Example continued…

- However, ex-ante average differences is troublesome in some regard…
  - Suggests treatment wasn't random
  - And ex-ante differences can be problematic if we think that their effect may vary with time…
    - Time-varying omitted variables **are** problematic because they can cause violation of "parallel trends"
    - E.g., states with more unions were trending differently at that time because of changes in union power

Summary of Today *[Part 1]*

- Natural experiment provides random variation in x that allows causal inference
  - Can be used in IV, regression discontinuity, but most often associated with "treatment" effects
- Two types of simple differences
  - Post-treatment comparison of treated & untreated
  - Pre- and post-treatment comparison of treated

Summary of Today *[Part 2]*

- Simple differences require strong assumptions; typically, not plausible
- Difference-in-differences helps with this
  - Compares change in y pre- versus post-treatment for treated to change in y for untreated
  - Requires "parallel trends" assumption

In First Half of Next Class

- Natural experiments *[Part 2]*
  - How to handle multiple events
  - Triple differences
  - Common robustness tests that can be used to test whether internal validity is likely to hold
- Related readings… see syllabus

Assign papers for next week…

- Jayaratne and Strahan (QJE 1996)
  - Bank deregulation and economic growth
- Bertrand and Mullainathan (JPE 2003)
  - Governance and managerial preferences
- Hayes, Lemmon, and Qiu (JFE 2012)
  - Stock options and managerial incentives

Break Time

- Let's take our 10-minute break
- We'll do presentations when we get back