GOV 51 Section

Week 1: Introduction, Notation, DiD

Pranav Moudgalya

Harvard College

Introduction (me)

- Third-year undergraduate in the Government Dept. (joint with ESPP)

Introduction (me)

- Third-year undergraduate in the Government Dept. (joint with ESPP)
- Political economy of energy, public opinion, methods

Introduction (me)

- Third-year undergraduate in the Government Dept. (joint with ESPP)
- Political economy of energy, public opinion, methods
- Orange County, CA

Introduction (me)

- Third-year undergraduate in the Government Dept. (joint with ESPP)
- Political economy of energy, public opinion, methods
- Orange County, CA
- Taught GOV 50 as a Senior CA/TF twice

Introduction (you)

- Harvard introduction (year, program, hometown, why you are taking GOV 51)

Introduction (you)

- Harvard introduction (year, program, hometown, why you are taking GOV 51)
- GOV 51 attracts diverse students (across Harvard schools, across concentrations, we want everyone to learn something new)

Introduction (you)

- Harvard introduction (year, program, hometown, why you are taking GOV 51)
- GOV 51 attracts diverse students (across Harvard schools, across concentrations, we want everyone to learn something new)
- Fulfills Statistics elective credit, Government methods requirement (for undergrads), part of "Data Science" track

Introduction (course)

- GOV 51 is about developing the statistical frameworks to more rigorously engage in *empirical* political science (& other social science) research (this means reading and most importantly, doing)

Introduction (course)

- GOV 51 is about developing the statistical frameworks to more rigorously engage in *empirical* political science (& other social science) research (this means reading and most importantly, doing)
- Is designed to be a mature (but certainly doable) jump from GOV 50

Introduction (course)

- GOV 51 is about developing the statistical frameworks to more rigorously engage in *empirical* political science (& other social science) research (this means reading and most importantly, doing)
- Is designed to be a mature (but certainly doable) jump from GOV 50
- Can be useful to check the course "prefresher" (tinyurl.com/GOV51prefresher)

Course Elements & Structure

- Course homebase: Naijia's website (NOT Canvas)

Course Elements & Structure

- Course homebase: Naijia's website (NOT Canvas)
- Gradescope (for assignments, exam, and project upload)

Course Elements & Structure

- Course homebase: Naijia's website (NOT Canvas)
- Gradescope (for assignments, exam, and project upload)
- Grading is as follows: participation (20%), problem sets (30%, 10% each), midterm (20%), final project (30%)

Course Elements & Structure

- Course homebase: Naijia's website (NOT Canvas)
- Gradescope (for assignments, exam, and project upload)
- Grading is as follows: participation (20%), problem sets (30%, 10% each), midterm (20%), final project (30%)
- Expectation is that you come to class and section and are engaged

A Quick Note on AI

- Tools like Claude, ChatGPT, DeepSeek are here to stay — and they are powerful; for every part of the course, you are allowed to use AI technology (though we reserve the right to amend this course policy at any time)

A Quick Note on AI

- Tools like Claude, ChatGPT, DeepSeek are here to stay — and they are powerful; for every part of the course, you are allowed to use AI technology (though we reserve the right to amend this course policy at any time)
  - **The good**: answer questions efficiently, get conceptual explanations, trouble shoot code, etc.

A Quick Note on AI

- Tools like Claude, ChatGPT, DeepSeek are here to stay — and they are powerful; for every part of the course, you are allowed to use AI technology (though we reserve the right to amend this course policy at any time)
  - **The good**: answer questions efficiently, get conceptual explanations, trouble shoot code, etc.
  - **The bad**: hallucinations (as of now, most LLMs cannot "reason" — they fail miserably at many simple tasks) $\rightsquigarrow$ use at your own risk

A Quick Note on AI

- Tools like Claude, ChatGPT, DeepSeek are here to stay — and they are powerful; for every part of the course, you are allowed to use AI technology (though we reserve the right to amend this course policy at any time)
  - **The good**: answer questions efficiently, get conceptual explanations, trouble shoot code, etc.
  - **The bad**: hallucinations (as of now, most LLMs cannot "reason" — they fail miserably at many simple tasks) $\rightsquigarrow$ use at your own risk
  - **The ugly**: overdependence on AI reduces how much you *learn* and how *creatively* you think

A Quick Note on AI

- Tools like Claude, ChatGPT, DeepSeek are here to stay — and they are powerful; for every part of the course, you are allowed to use AI technology (though we reserve the right to amend this course policy at any time)
  - **The good**: answer questions efficiently, get conceptual explanations, trouble shoot code, etc.
  - **The bad**: hallucinations (as of now, most LLMs cannot "reason" — they fail miserably at many simple tasks) $\rightsquigarrow$ use at your own risk
  - **The ugly**: overdependence on AI reduces how much you *learn* and how *creatively* you think
  - *Use it, don't abuse it*

Section

- My full section syllabus can be found on the GOV 51 course site

Section

- My full section syllabus can be found on the GOV 51 course site
- Section materials to be distributed after section

Section

- My full section syllabus can be found on the GOV 51 course site
- Section materials to be distributed after section
  - Materials sourced from James Jolin, Jeremiah Cha, Prof. Liu, and others; but all errors are my own!

Section

- My full section syllabus can be found on the GOV 51 course site
- Section materials to be distributed after section
  - Materials sourced from James Jolin, Jeremiah Cha, Prof. Liu, and others; but all errors are my own!
- My office hours:

Section

- My full section syllabus can be found on the GOV 51 course site
- Section materials to be distributed after section
  - Materials sourced from James Jolin, Jeremiah Cha, Prof. Liu, and others; but all errors are my own!
- My office hours:
  - Mondays, 7-9pm in Leverett Dining Hall
  - Thursdays, 7:45-9pm in Leverett Dining Hall (right after section)
  - Walk in!

Section

- My full section syllabus can be found on the GOV 51 course site
- Section materials to be distributed after section
  - Materials sourced from James Jolin, Jeremiah Cha, Prof. Liu, and others; but all errors are my own!
- My office hours:
  - Mondays, 7-9pm in Leverett Dining Hall
  - Thursdays, 7:45-9pm in Leverett Dining Hall (right after section)
  - Walk in!
- Other office hours

Section

- My full section syllabus can be found on the GOV 51 course site
- Section materials to be distributed after section
  - Materials sourced from James Jolin, Jeremiah Cha, Prof. Liu, and others; but all errors are my own!
- My office hours:
  - Mondays, 7-9pm in Leverett Dining Hall
  - Thursdays, 7:45-9pm in Leverett Dining Hall (right after section)
  - Walk in!
- Other office hours
  - Sima (head TF) & Ben (CA) by appointment, see course website

Causal Inference

- What do we mean when we say "causality" / "causal inference"; why is this not an issue in the natural sciences?

Causal Inference

- What do we mean when we say "causality" / "causal inference"; why is this not an issue in the natural sciences?
- Causal inference is the study of causal questions

Causal Inference

- What do we mean when we say “causality” / “causal inference”; why is this not an issue in the natural sciences?
- Causal inference is the study of causal questions
  - Does an electoral challenge increase legislative productivity?

Causal Inference

- What do we mean when we say "causality" / "causal inference"; why is this not an issue in the natural sciences?
- Causal inference is the study of causal questions
  - Does an electoral challenge increase legislative productivity?
  - Do protests lead to lasting institutional change?

Causal Inference

- What do we mean when we say "causality" / "causal inference"; why is this not an issue in the natural sciences?
- Causal inference is the study of causal questions
  - Does an electoral challenge increase legislative productivity?
  - Do protests lead to lasting institutional change?
- The fundamental problem of causal inference is that at most one of the potential outcomes can be realized and thus observed. Why?

Causal Inference

- What do we mean when we say "causality" / "causal inference"; why is this not an issue in the natural sciences?
- Causal inference is the study of causal questions
  - Does an electoral challenge increase legislative productivity?
  - Do protests lead to lasting institutional change?
- The fundamental problem of causal inference is that at most one of the potential outcomes can be realized and thus observed. Why?
- Causal effect: this is hypothetical; it is the comparison of potential outcomes, for the same unit, at the same moment in time post-treatment

Notation

- Promise: notation makes our lives easier, not harder

Notation

- Promise: notation makes our lives easier, not harder
- Treatment indicator, T_i, is a binary variable indicating the treatment status (treatment, control) for the ith unit.

Notation

- Promise: notation makes our lives easier, not harder
- Treatment indicator, T_i, is a binary variable indicating the treatment status (treatment, control) for the ith unit.
- The outcome indicator, Y_i is the value of the potential outcome for the ith unit.

Notation

- Promise: notation makes our lives easier, not harder
- Treatment indicator, T_i, is a binary variable indicating the treatment status (treatment, control) for the ith unit.
- The outcome indicator, Y_i is the value of the potential outcome for the ith unit.
- Causal effect quantity of interest: $Y_i(1) - Y_i(0)$, the difference in the potential outcomes under treatment for the *same* unit i.

Notation

- Promise: notation makes our lives easier, not harder
- Treatment indicator, T_i, is a binary variable indicating the treatment status (treatment, control) for the ith unit.
- The outcome indicator, Y_i is the value of the potential outcome for the ith unit.
- Causal effect quantity of interest: $Y_i(1) - Y_i(0)$, the difference in the potential outcomes under treatment for the *same* unit i.
- **BUT** $Y_i(1) - Y_i(0)$ can **never** be observed $\rightsquigarrow$ We must infer the missing counterfactuals (E.g. randomize treatment!)

Notation

- Promise: notation makes our lives easier, not harder
- Treatment indicator, T_i, is a binary variable indicating the treatment status (treatment, control) for the ith unit.
- The outcome indicator, Y_i is the value of the potential outcome for the ith unit.
- Causal effect quantity of interest: $Y_i(1) - Y_i(0)$, the difference in the potential outcomes under treatment for the *same* unit i.
- **BUT** $Y_i(1) - Y_i(0)$ can **never** be observed $\rightsquigarrow$ We must infer the missing counterfactuals (E.g. randomize treatment!)
- $\mathbb{E}[X]$ is the expectation, average, of random variable X (it is a *weighted* average)

Dataset Viz.

Voters	Age	Gender	Canvassed	Turnout		Causal Effect
i	X_1	X_2	T_i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
1	19	M	1	0	???	???
2	56	F	0	???	1	???
3	89	F	0	???	0	???

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of you**u** and **me**"

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of yo**u** and **me**"
- But we have to. . .

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of yo**u** and **me**"
- But we have to...
- **Causal ordering:** treatment T causes Y and not the other way round

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of yo**u** and **me**"
- But we have to...
- **Causal ordering:** treatment T causes Y and not the other way round
- **Consistency**: $Y_i = Y_i(T_i = t)$ whenever $T_i = t$. That is, *there is no hidden variation in treatment*. Example?

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of yo**u** and **me**"
- But we have to. . .
- **Causal ordering:** treatment T causes Y and not the other way round
- **Consistency**: $Y_i = Y_i(T_i = t)$ whenever $T_i = t$. That is, *there is no hidden variation in treatment*. Example?
  - This is violated if, for example, I defined treatment as $t =$ "receive tutoring once a week" but some kiddos are receiving tutoring twice per week and $Y_i = Y_i(\text{tutor once per week}) \neq Y_i(\text{tutor twice per week})$

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of yo**u** and **me**"
- But we have to...
- **Causal ordering:** treatment T causes Y and not the other way round
- **Consistency**: $Y_i = Y_i(T_i = t)$ whenever $T_i = t$. That is, *there is no hidden variation in treatment*. Example?
  - This is violated if, for example, I defined treatment as $t =$ "receive tutoring once a week" but some kiddos are receiving tutoring twice per week and $Y_i = Y_i(\text{tutor once per week}) \neq Y_i(\text{tutor twice per week})$
- **No interference**: The potential outcome of a unit i is independent of any other unit's $i = i'$ treatment

Let's Make Some Assumptions

- Famous saying: "When you **assume**, you make an **ass** out of yo**u** and **me**"
- But we have to...
- **Causal ordering:** treatment T causes Y and not the other way round
- **Consistency**: $Y_i = Y_i(T_i = t)$ whenever $T_i = t$. That is, *there is no hidden variation in treatment*. Example?
  - This is violated if, for example, I defined treatment as $t =$ "receive tutoring once a week" but some kiddos are receiving tutoring twice per week and $Y_i = Y_i(\text{tutor once per week}) \neq Y_i(\text{tutor twice per week})$
- **No interference**: The potential outcome of a unit i is independent of any other unit's $i = i'$ treatment
- Last two assumptions comprise **SUTVA**

More Assumptions

- Assumptions on assignment mechanism

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**: $P(T_i = 1 | X_i, X_{i'}) = P(T_i = 1 | X_i)$

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**:
    $P(T_i = 1|X_i, X_{i'}) = P(T_i = 1|X_i)$
    - That is, the probability of a unit receiving a treatment depends only on that unit's characteristics, not on the characteristics or treatment status of other units.

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**: $P(T_i = 1|X_i, X_{i'}) = P(T_i = 1|X_i)$
    - That is, the probability of a unit receiving a treatment depends only on that unit's characteristics, not on the characteristics or treatment status of other units.
  - **Probabilistic assignment**: $0 < P(T_i = 1|X_i) < 1$

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**:
    $P(T_i = 1|X_i, X_{i'}) = P(T_i = 1|X_i)$
    - That is, the probability of a unit receiving a treatment depends only on that unit's characteristics, not on the characteristics or treatment status of other units.
  - **Probabilistic assignment**: $0 < P(T_i = 1|X_i) < 1$
    - That is, there is always some chance that any unit receives the treatment, but it is neither guaranteed nor impossible.

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**: $P(T_i = 1|X_i, X_{i'}) = P(T_i = 1|X_i)$
    - That is, the probability of a unit receiving a treatment depends only on that unit's characteristics, not on the characteristics or treatment status of other units.
  - **Probabilistic assignment**: $0 < P(T_i = 1|X_i) < 1$
    - That is, there is always some chance that any unit receives the treatment, but it is neither guaranteed nor impossible.
  - **Unconfounded assignment**: $P(T_i = 1|X_i) = P(T_i = 1|X_i, Y_i(0), Y_i(1))$

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**: $P(T_i = 1|X_i, X_{i'}) = P(T_i = 1|X_i)$
    - That is, the probability of a unit receiving a treatment depends only on that unit's characteristics, not on the characteristics or treatment status of other units.
  - **Probabilistic assignment**: $0 < P(T_i = 1|X_i) < 1$
    - That is, there is always some chance that any unit receives the treatment, but it is neither guaranteed nor impossible.
  - **Unconfounded assignment**: $P(T_i = 1|X_i) = P(T_i = 1|X_i, Y_i(0), Y_i(1))$
    - That is, treatment assignment is independent of the potential outcomes once we account for observable characteristics.

More Assumptions

- Assumptions on assignment mechanism
  - **Individualistic assignment**: $P(T_i = 1|X_i, X_{i'}) = P(T_i = 1|X_i)$
    - That is, the probability of a unit receiving a treatment depends only on that unit's characteristics, not on the characteristics or treatment status of other units.
  - **Probabilistic assignment**: $0 < P(T_i = 1|X_i) < 1$
    - That is, there is always some chance that any unit receives the treatment, but it is neither guaranteed nor impossible.
  - **Unconfounded assignment**: $P(T_i = 1|X_i) = P(T_i = 1|X_i, Y_i(0), Y_i(1))$
    - That is, treatment assignment is independent of the potential outcomes once we account for observable characteristics.
- Don't worry about the notation (for now); focus on the ideas!

Difference-in-Difference

- As in lecture, difference-in-difference allows us to "infer what would have happened to the treatment group without treatment"
- What assumptions are necessary for identification? Well, visually...

DiD Assumptions

- **Parallel trends**: "treated and control groups will share the same difference if the treated were *not* to be treated at $t = 1$"

DiD Assumptions

- **Parallel trends**: "treated and control groups will share the same difference if the treated were *not* to be treated at $t = 1$"
- Parallel trends assumption is fundamentally untestable—but researchers can assess its plausibility using statistical methods (or more naively, a graphical check)
- Hang on. . . why can't you just do a simple subtraction of means?

DiD Assumptions

- **Parallel trends**: "treated and control groups will share the same difference if the treated were *not* to be treated at $t = 1$"
- Parallel trends assumption is fundamentally untestable—but researchers can assess its plausibility using statistical methods (or more naively, a graphical check)
- Hang on... why can't you just do a simple subtraction of means?
  - You can do a difference-in-difference analysis without parallel trends (e.g. just simple algebra subtracting means) – BUT your estimand will not be **identified**, so you cannot make causal interpretations

Esta-what?

- Before we jump into the *mathy* parts of DiD, let's discuss some terms

Esta-what?

- Before we jump into the *mathy* parts of DiD, let's discuss some terms
- **Estimand**: The true value we want to find (often unobservable)

Esta-what?

- Before we jump into the *mathy* parts of DiD, let's discuss some terms
- **Estimand**: The true value we want to find (often unobservable)
- **Estimator**: The method or formula we use to calculate an estimate (must be observable), often denoted with a hat.

Esta-what?

- Before we jump into the *mathy* parts of DiD, let's discuss some terms
- **Estimand**: The true value we want to find (often unobservable)
- **Estimator**: The method or formula we use to calculate an estimate (must be observable), often denoted with a hat.
- **Estimate**: The actual number we get when we apply the estimator to data.

Estimand, Estimator, Estimate (in action)

- Think of trying to measure the average height of all students in a school.

Estimand, Estimator, Estimate (in action)

- Think of trying to measure the average height of all students in a school.
- **Estimand**: The actual (but unknown) average height of all students.

Estimand, Estimator, Estimate (in action)

- Think of trying to measure the average height of all students in a school.
- **Estimand**: The actual (but unknown) average height of all students.
- **Estimator**: The formula we use, like taking the mean of a random sample.

Estimand, Estimator, Estimate (in action)

- Think of trying to measure the average height of all students in a school.
- **Estimand**: The actual (but unknown) average height of all students.
- **Estimator**: The formula we use, like taking the mean of a random sample.
- **Estimate**: The actual average height we compute from our sample.

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
    - $Y_{i1}(0)$:
    - $Y_{i0}(0)$:
    - $Y_{i1}(1)$:
    - $Y_{i0}(1)$:
    - $Y_{i1}(0)|T_i = 0$:
    - $Y_{i1}(0)|T_i = 1$:

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
  - $Y_{i1}(0)$: Potential outcome in period after treatment under control

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
  - $Y_{i1}(0)$: Potential outcome in period after treatment under control
  - $Y_{i0}(0)$: Potential outcome in period before treatment under control

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
  - $Y_{i1}(0)$: Potential outcome in period after treatment under control
  - $Y_{i0}(0)$: Potential outcome in period before treatment under control
  - $Y_{i1}(1)$: Potential outcome in period after treatment under treatment

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
  - $Y_{i1}(0)$: Potential outcome in period after treatment under control
  - $Y_{i0}(0)$: Potential outcome in period before treatment under control
  - $Y_{i1}(1)$: Potential outcome in period after treatment under treatment
  - $Y_{i0}(1)$: Potential outcome in period before treatment under treatment

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
  - $Y_{i1}(0)$: Potential outcome in period after treatment under control
  - $Y_{i0}(0)$: Potential outcome in period before treatment under control
  - $Y_{i1}(1)$: Potential outcome in period after treatment under treatment
  - $Y_{i0}(1)$: Potential outcome in period before treatment under treatment
  - $Y_{i1}(0)|T_i = 0$: Potential outcome in period after treatment under control **given** the units are **not treated**. This is observed!

Back to DiD

- With DiD, we're going to use some slightly new notation. Namely, instead of merely $Y_i(1)$, we're going to use $Y_{it}(1)$, where everything is the same as usual, but t is an indicator of **time period**, pre-treatment $(t = 0)$ or post-treatment $(t = 1)$.
- **Quiz! What are the following?**
  - $Y_{i1}(0)$: Potential outcome in period after treatment under control
  - $Y_{i0}(0)$: Potential outcome in period before treatment under control
  - $Y_{i1}(1)$: Potential outcome in period after treatment under treatment
  - $Y_{i0}(1)$: Potential outcome in period before treatment under treatment
  - $Y_{i1}(0)|T_i = 0$: Potential outcome in period after treatment under control **given** the units are **not treated**. This is observed!
  - $Y_{i1}(0)|T_i = 1$: Potential outcome in period after treatment under control **given** the units are **treated**. **Never Observed**

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0) | T_i = 1]$$

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0) | T_i = 1]$$

- **What's the problem?**

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0) | T_i = 1]$$

- **What's the problem?**
  - We can never observe $Y_{i1}(0)|T_i = 1$ (after treatment is given, the treatment group HAS to be treated!).

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0) | T_i = 1]$$

- **What's the problem?**
  - We can never observe $Y_{i1}(0) | T_i = 1$ (after treatment is given, the treatment group HAS to be treated!).
- **Parallel trend to the rescue!**

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0) | T_i = 1]$$

- **What's the problem?**
  - We can never observe $Y_{i1}(0)|T_i = 1$ (after treatment is given, the treatment group HAS to be treated!).
- **Parallel trend to the rescue!**
  - Let's use the trend in the control group as a stand-in for $Y_{i1}(0)$ for the control group!

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0)|T_i = 1]$$

- **What's the problem?**
  - We can never observe $Y_{i1}(0)|T_i = 1$ (after treatment is given, the treatment group HAS to be treated!).
- **Parallel trend to the rescue!**
  - Let's use the trend in the control group as a stand-in for $Y_{i1}(0)$ for the control group!
  - In other words: $\mathbb{E}[Y_{i1}(0) - Y_{i0}(1)|T_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|T_i = 0]$

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0) | T_i = 1]$$

- **What's the problem?**
  - We can never observe $Y_{i1}(0)|T_i = 1$ (after treatment is given, the treatment group HAS to be treated!).
- **Parallel trend to the rescue!**
  - Let's use the trend in the control group as a stand-in for $Y_{i1}(0)$ for the control group!
  - In other words: $\mathbb{E}[Y_{i1}(0) - Y_{i0}(1) | T_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0) | T_i = 0]$
  - Meaning: treated and control group will share the same difference, if the treated were not to be treated at $t = 1$.

DiD Estimand

- **What's the DID estimand (what we want to estimate)?**
  - $\Rightarrow$ It's not an ATE; it's an ATT (average treatment effect on the treated units)!
  - Our estimand is:

$$\mathbb{E}[Y_{i1}(1) - Y_{i1}(0)|T_i = 1]$$

- **What's the problem?**
  - We can never observe $Y_{i1}(0)|T_i = 1$ (after treatment is given, the treatment group HAS to be treated!).
- **Parallel trend to the rescue!**
  - Let's use the trend in the control group as a stand-in for $Y_{i1}(0)$ for the control group!
  - In other words: $\mathbb{E}[Y_{i1}(0) - Y_{i0}(1)|T_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|T_i = 0]$
  - Meaning: treated and control group will share the same difference, if the treated were not to be treated at $t = 1$.
- Estimator: $\widehat{DiD} = \mathbb{E}[Y_{i1}(1) - Y_{i0}(1)|T_i = 1] - \mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|T_i = 0]$

The Proof

- Highly encourage to review over the weekend

The Proof

- Highly encourage to review over the weekend
- Notes will be sent out going over it formally

The Proof

- Highly encourage to review over the weekend
- Notes will be sent out going over it formally
- Employs basic algebraic manipulation and the "linearity of expectation"

Next Week

- More coding!

Next Week

- More coding!
- Instrumental variables

Next Week

- More coding!
- Instrumental variables
- Advice: think about what intrigues you about the world (that you could reasonably answer using the techniques of this class - final project is a wonderful opportunity)

Next Week

- More coding!
- Instrumental variables
- Advice: think about what intrigues you about the world (that you could reasonably answer using the techniques of this class - final project is a wonderful opportunity)