Week 9: Regression in the Social Sciences

Brandon Stewart[1]

Princeton

November 14 and 19, 2018

[1] These slides are heavily influenced by Matt Blackwell, Justin Grimmer, Jens Hainmueller, Erin Hartman and Kosuke Imai.

Where We've Been and Where We're Going...

- Last Week
  - diagnostics
- This 'Week'
  - Wednesday:
    - making an argument in social sciences
    - causal inference
  - Monday:
    - more causal inference
    - visualization
- Next Week
  - selection on observables
- Long Run
  - probability $\rightarrow$ inference $\rightarrow$ regression $\rightarrow$ causal inference

Questions?

Why Are We Doing All of This Again?

Why Are We Doing All of This Again?

- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.

Why Are We Doing All of This Again?

- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.
- Quantitative methods are an increasingly big part of that so whether you are reading or actively doing quantitative analysis it is going to be there.

Why Are We Doing All of This Again?

- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.
- Quantitative methods are an increasingly big part of that so whether you are reading or actively doing quantitative analysis it is going to be there.
- So why all the math?

Why Are We Doing All of This Again?

- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.
- Quantitative methods are an increasingly big part of that so whether you are reading or actively doing quantitative analysis it is going to be there.
- So why all the math? We are taking a future-oriented approach. We want to prepare you for the next big thing

Why Are We Doing All of This Again?

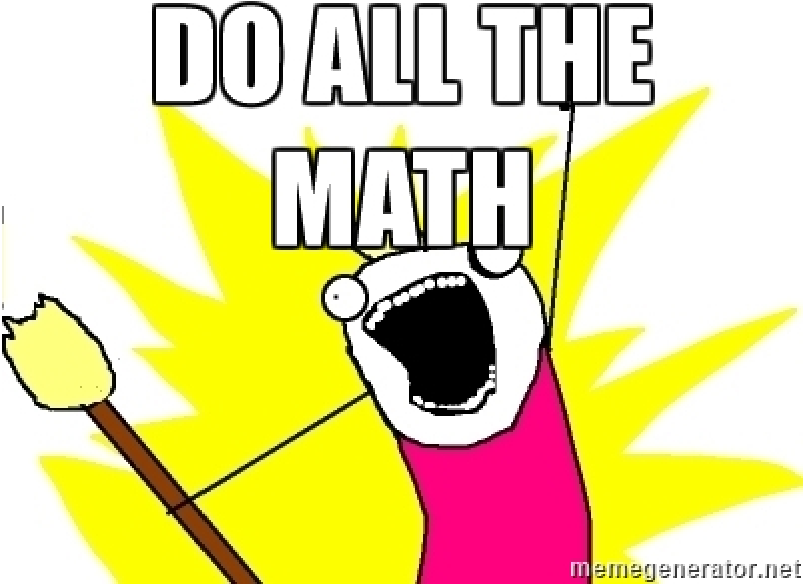
- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.
- Quantitative methods are an increasingly big part of that so whether you are reading or actively doing quantitative analysis it is going to be there.
- So why all the math? We are taking a future-oriented approach. We want to prepare you for the next big thing
- Methods that became popular in the social sciences since I took the equivalent of this class: machine learning, text-as-data, Bayesian nonparametrics, design-based inference, DAG-based causal inference, deep learning

Why Are We Doing All of This Again?

- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.
- Quantitative methods are an increasingly big part of that so whether you are reading or actively doing quantitative analysis it is going to be there.
- So why all the math? We are taking a future-oriented approach. We want to prepare you for the next big thing
- Methods that became popular in the social sciences since I took the equivalent of this class: machine learning, text-as-data, Bayesian nonparametrics, design-based inference, DAG-based causal inference, deep learning
- A technical foundation prepares you to learn new methods for the rest of your career. Trust me now is the time to invest.

Why Are We Doing All of This Again?

- We are all here because we are trying to do some social science, that is, we are in the business of knowledge production.
- Quantitative methods are an increasingly big part of that so whether you are reading or actively doing quantitative analysis it is going to be there.
- So why all the math? We are taking a future-oriented approach. We want to prepare you for the next big thing
- Methods that became popular in the social sciences since I took the equivalent of this class: machine learning, text-as-data, Bayesian nonparametrics, design-based inference, DAG-based causal inference, deep learning
- A technical foundation prepares you to learn new methods for the rest of your career. Trust me now is the time to invest.
- Knowing how methods work also makes you a better reader of work.


DO ALL THE
MATH
memegenerator.net

Quantitative Social Science

Quantitative Social Science

- Three components of quantitative social science:

Quantitative Social Science

- Three components of quantitative social science:
  1. Argument
  2. Research Design
  3. Presentation

Quantitative Social Science

- Three components of quantitative social science:
  1. Argument
  2. Research Design
  3. Presentation
- These two classes we will focus on:

Quantitative Social Science

- Three components of quantitative social science:
  1. Argument
  2. Research Design
  3. Presentation
- These two classes we will focus on:
  - identification and causal inference (argument, design)
  - visualization and quantities of interest (argument, presentation)

Quantitative Social Science

- Three components of quantitative social science:
  1. Argument
  2. Research Design
  3. Presentation
- These two classes we will focus on:
  - identification and causal inference (argument, design)
  - visualization and quantities of interest (argument, presentation)
- My core argument: to have a hope of success we need to be clear about the estimand. The implicit estimand is often (but not always) causal.

Quantitative Social Science

- Three components of quantitative social science:
  1. Argument
  2. Research Design
  3. Presentation
- These two classes we will focus on:
  - identification and causal inference (argument, design)
  - visualization and quantities of interest (argument, presentation)
- My core argument: to have a hope of success we need to be clear about the estimand. The implicit estimand is often (but not always) causal.

We will mostly talk about statistical methods here (it is a statistics class!) but the best work is a combination of substantive and statistical theory.

What is Causal Inference?

What is Causal Inference?

- A causal inference is a statement about counterfactuals — it is a statement about the difference between what did and didn't happen

What is Causal Inference?

- A causal inference is a statement about counterfactuals — it is a statement about the difference between what did and didn't happen
- The core puzzle of causal inference is how you get the information about what didn't happen

What is Causal Inference?

- A causal inference is a statement about counterfactuals — it is a statement about the difference between what did and didn't happen
- The core puzzle of causal inference is how you get the information about what didn't happen
- The difference between prediction and causal inference is the intervention on the system under study

What is Causal Inference?

- A causal inference is a statement about counterfactuals — it is a statement about the difference between what did and didn't happen
- The core puzzle of causal inference is how you get the information about what didn't happen
- The difference between prediction and causal inference is the intervention on the system under study
- Like it or not, social science theories are almost always expressed as causal claims: e.g. "an increase in X causes an increase in Y" (or something more opaque meaning the same thing)

What is Causal Inference?

- A causal inference is a statement about counterfactuals — it is a statement about the difference between what did and didn't happen
- The core puzzle of causal inference is how you get the information about what didn't happen
- The difference between prediction and causal inference is the intervention on the system under study
- Like it or not, social science theories are almost always expressed as causal claims: e.g. "an increase in X causes an increase in Y" (or something more opaque meaning the same thing)
- The study of causal inference helps us understand the assumptions we need to make this kind of claim.

Identification

Identification

- A quantity of interest is identified when (given stated assumptions) access to infinite data would result in the estimate taking on only a single value

Identification

- A quantity of interest is identified when (given stated assumptions) access to infinite data would result in the estimate taking on only a single value
- For example, having all dummy variables in a linear model is not statistically identified because they cannot be distinguished from the intercept.

Identification

- A quantity of interest is identified when (given stated assumptions) access to infinite data would result in the estimate taking on only a single value
- For example, having all dummy variables in a linear model is not statistically identified because they cannot be distinguished from the intercept.
- Causal identification is what we can learn about a causal effect from available data.

Identification

- A quantity of interest is identified when (given stated assumptions) access to infinite data would result in the estimate taking on only a single value
- For example, having all dummy variables in a linear model is not statistically identified because they cannot be distinguished from the intercept.
- Causal identification is what we can learn about a causal effect from available data.
- If an effect is not identified, no estimation method will recover it.

Identification

- A quantity of interest is identified when (given stated assumptions) access to infinite data would result in the estimate taking on only a single value
- For example, having all dummy variables in a linear model is not statistically identified because they cannot be distinguished from the intercept.
- Causal identification is what we can learn about a causal effect from available data.
- If an effect is not identified, no estimation method will recover it.
- This means the relevant question is "what's your identification strategy?" or what are the set of assumptions that let you claim you will be able to estimate a causal effect from the observed data?

Identification

- A quantity of interest is identified when (given stated assumptions) access to infinite data would result in the estimate taking on only a single value
- For example, having all dummy variables in a linear model is not statistically identified because they cannot be distinguished from the intercept.
- Causal identification is what we can learn about a causal effect from available data.
- If an effect is not identified, no estimation method will recover it.
- This means the relevant question is "what's your identification strategy?" or what are the set of assumptions that let you claim you will be able to estimate a causal effect from the observed data?
- As we will see this is not a conversation about estimation: in other words, if someone answers "regression" they have made a category error

Identification vs. Estimation

- Identification: How much can you learn about the estimand if you have an infinite amount of data?
- Estimation: How much can you learn about the estimand from a finite sample?

Identification vs. Estimation

- Identification: How much can you learn about the estimand if you have an infinite amount of data?
- Estimation: How much can you learn about the estimand from a finite sample?
- Identification precedes estimation

Identification vs. Estimation

- Identification: How much can you learn about the estimand if you have an infinite amount of data?
- Estimation: How much can you learn about the estimand from a finite sample?
- Identification precedes estimation

The role of assumptions:

- Often identification requires (hopefully minimal) assumptions

Identification vs. Estimation

- Identification: How much can you learn about the estimand if you have an infinite amount of data?
- Estimation: How much can you learn about the estimand from a finite sample?
- Identification precedes estimation

The role of assumptions:

- Often identification requires (hopefully minimal) assumptions
- Even when identification is possible, estimation may impose additional assumptions (i.e. that the linear approximation to the CEF is good enough)

Identification vs. Estimation

- Identification: How much can you learn about the estimand if you have an infinite amount of data?
- Estimation: How much can you learn about the estimand from a finite sample?
- Identification precedes estimation

The role of assumptions:

- Often identification requires (hopefully minimal) assumptions
- Even when identification is possible, estimation may impose additional assumptions (i.e. that the linear approximation to the CEF is good enough)
- Law of Decreasing Credibility (Manski): The credibility of inference decreases with the strength of the assumptions maintained

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals
- Introduction of randomization as the "reasoned basis" for inference by Fisher (1925)

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals
- Introduction of randomization as the "reasoned basis" for inference by Fisher (1925)
  - *p*-values and permutation inference

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals
- Introduction of randomization as the "reasoned basis" for inference by Fisher (1925)
  - p-values and permutation inference
- Causal effects defined at the unit level, allowing for effects to be defined without a known assignment mechanism by Rubin (1974)

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals
- Introduction of randomization as the "reasoned basis" for inference by Fisher (1925)
  - *p*-values and permutation inference
- Causal effects defined at the unit level, allowing for effects to be defined without a known assignment mechanism by Rubin (1974)
- Potential outcomes expanded to observational studies by Rubin (1974)

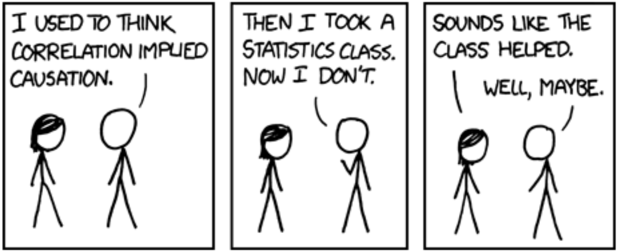
Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals
- Introduction of randomization as the "reasoned basis" for inference by Fisher (1925)
  - *p*-values and permutation inference
- Causal effects defined at the unit level, allowing for effects to be defined without a known assignment mechanism by Rubin (1974)
- Potential outcomes expanded to observational studies by Rubin (1974)
- Formalization of the assignment mechanism in potential outcomes by Rubin (1975, 1978)

Brief History of Potential Outcomes and Causal Inference

- Introduction of potential outcomes in randomized experiments by Neyman (1923)
  - Super-population inference and confidence intervals
- Introduction of randomization as the "reasoned basis" for inference by Fisher (1925)
  - *p*-values and permutation inference
- Causal effects defined at the unit level, allowing for effects to be defined without a known assignment mechanism by Rubin (1974)
- Potential outcomes expanded to observational studies by Rubin (1974)
- Formalization of the assignment mechanism in potential outcomes by Rubin (1975, 1978)
- Pearl (1995) develops graphical models for causal inference

Causation


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Potential Outcomes

Definitions:

Potential Outcomes

Definitions:

T_i: Dichotomous Treatment assignment for unit i (multi-valued treatments–just more potential outcomes for each unit)

$$T_i = \begin{cases} 1 & \text{Unit is assigned to treatment} \\ 0 & \text{Unit is not assigned to treatment} \end{cases}$$

Potential Outcomes

Definitions:

T_i: Dichotomous Treatment assignment for unit i (multi-valued treatments–just more potential outcomes for each unit)

$$T_i = \begin{cases} 1 & \text{Unit is assigned to treatment} \\ 0 & \text{Unit is not assigned to treatment} \end{cases}$$

Y_i: Outcome for unit i

Potential Outcomes

Definitions:

T_i: Dichotomous Treatment assignment for unit i (multi-valued treatments–just more potential outcomes for each unit)

$$T_i = \begin{cases} 1 & \text{Unit is assigned to treatment} \\ 0 & \text{Unit is not assigned to treatment} \end{cases}$$

Y_i: Outcome for unit i

Potential outcomes for unit i:

$$Y_i(T_i) = \begin{cases} Y_i(1) & \text{Potential outcome for unit } i \text{ with treatment} \\ Y_i(0) & \text{Potential outcome for unit } i \text{ without treatment} \end{cases}$$

Potential Outcomes

Definitions:

T_i: Dichotomous Treatment assignment for unit i (multi-valued treatments–just more potential outcomes for each unit)

$$T_i = \begin{cases} 1 & \text{Unit is assigned to treatment} \\ 0 & \text{Unit is not assigned to treatment} \end{cases}$$

Y_i: Outcome for unit i

Potential outcomes for unit i:

$$Y_i(T_i) = \begin{cases} Y_i(1) & \text{Potential outcome for unit } i \text{ with treatment} \\ Y_i(0) & \text{Potential outcome for unit } i \text{ without treatment} \end{cases}$$

Pre-treatment covariates X_i

Potential Outcomes

Definitions:

T_i: Dichotomous Treatment assignment for unit i (multi-valued treatments–just more potential outcomes for each unit)

$$T_i = \begin{cases} 1 & \text{Unit is assigned to treatment} \\ 0 & \text{Unit is not assigned to treatment} \end{cases}$$

Y_i: Outcome for unit i

Potential outcomes for unit i:

$$Y_i(T_i) = \begin{cases} Y_i(1) & \text{Potential outcome for unit } i \text{ with treatment} \\ Y_i(0) & \text{Potential outcome for unit } i \text{ without treatment} \end{cases}$$

<u>Pre</u>-treatment covariates X_i

τ_i: The treatment effect

$$\tau_i = Y_i(1) - Y_i(0)$$

Potential Outcomes – Aspirin Example

Definitions:

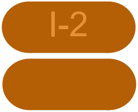
Potential Outcomes – Aspirin Example

Definitions:

T_i: Unit assigned to:

$$T_i = \begin{cases} 1 & \text{Receive Aspirin} \\ 0 & \text{Receive Placebo} \end{cases}$$

$(T_i = 1)$



$(T_i = 0)$

Potential Outcomes – Aspirin Example

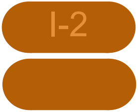
Definitions:

T_i: Unit assigned to:

$$T_i = \begin{cases} 1 & \text{Receive Aspirin} \\ 0 & \text{Receive Placebo} \end{cases}$$

Y_i: Outcome for unit i – Patient has headache, or not

$(T_i = 1)$



$(T_i = 0)$

$(Y_i = 1)$

$(Y_i = 0)$

Potential Outcomes – Aspirin Example

Definitions:

T_i: Unit assigned to:

$$T_i = \begin{cases} 1 & \text{Receive Aspirin} \\ 0 & \text{Receive Placebo} \end{cases}$$

$(T_i = 1)$ $(T_i = 0)$

$(Y_i = 1)$ $(Y_i = 0)$

Y_i: Outcome for unit i – Patient has headache, or not

Potential outcomes for unit i:

$$Y_i(T_i) = \begin{cases} Y_i(1) & \text{Headache (or not) for unit } i \text{ with Aspirin} \\ Y_i(0) & \text{Headache (or not) for unit } i \text{ with placebo} \end{cases}$$

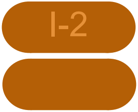
Potential Outcomes – Aspirin Example

Definitions:

T_i: Unit assigned to:

$$T_i = \begin{cases} 1 & \text{Receive Aspirin} \\ 0 & \text{Receive Placebo} \end{cases}$$

$(T_i = 1)$



$(T_i = 0)$

Y_i: Outcome for unit i – Patient has headache, or not

$(Y_i = 1)$

$(Y_i = 0)$

Potential outcomes for unit i:

$$Y_i(T_i) = \begin{cases} Y_i(1) & \text{Headache (or not) for unit } i \text{ with Aspirin} \\ Y_i(0) & \text{Headache (or not) for unit } i \text{ with placebo} \end{cases}$$

Pre-treatment covariates X_i

Illustrated potential outcomes here and later courtesy of Erin Hartman

What is random in the potential outcomes framework?

Note that potential outcomes are thought of as fixed, and that they, and the difference between them, can vary by arbitrary amounts for each unit i. There is some true distribution of potential outcomes across the population.

What is random in the potential outcomes framework?

Note that potential outcomes are thought of as fixed, and that they, and the difference between them, can vary by arbitrary amounts for each unit i. There is some true distribution of potential outcomes across the population.

Treatment assignment is the source of randomness

Causal Inference is a Missing Data Problem

Definition: Observed Outcome

$$Y_i = T_i * Y_i(1) + (1 - T_i) * Y_i(0)$$

Causal Inference is a Missing Data Problem

Definition: Observed Outcome

$$Y_i = T_i * Y_i(1) + (1 - T_i) * Y_i(0)$$

Inherently, since we cannot observe both treatment and control for unit i, thus we only observe Y_i, causal inference suffers from a **missing data problem**.

Causal Inference is a Missing Data Problem

Definition: Observed Outcome

$$Y_i = T_i * Y_i(1) + (1 - T_i) * Y_i(0)$$

Inherently, since we cannot observe both treatment and control for unit i, thus we only observe Y_i, causal inference suffers from a **missing data problem**.

No methodology allows us to simultaneously observe both potential outcomes, $Y_i(1)$ and $Y_i(0)$, making τ_i unobservable–and unidentifiable without additional assumptions (Fundamental Problem of Causal Inference Holland (1986))

Causal Inference is a Missing Data Problem

Example: Asprin's Impact on Headaches

Patient i		Pill T_i	Headache Status Y_i(PLACEBO)	Y_i(I-2)	Y_i	Age X_{1i}	Academic X_{2i}
1						25	Y
2						55	N
3						62	Y
4						80	N
5						32	Y
6						45	B
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n						71	N

Causal Inference is a Missing Data Problem

Example: Asprin's Impact on Headaches

Patient i		Pill T_i	Headache Status Y_i(PLACEBO)	Y_i(I-2)	Y_i	Age X_{1i}	Academic X_{2i}
1						25	Y
2						55	N
3						62	Y
4						80	N
5						32	Y
6						45	B
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n						71	N

Step 1: (Randomly) sample units

Causal Inference is a Missing Data Problem

Example: Asprin's Impact on Headaches

Patient i		Pill T_i	Headache Status Y_i(PLACEBO)	Y_i(I-2)	Y_i	Age X_{1i}	Academic X_{2i}
1		I-2				25	Y
2						55	N
3						62	Y
4		PLACEBO				80	N
5		I-2				32	Y
6						45	B
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n		PLACEBO				71	N

Step 2: Randomly assign treatment

Causal Inference is a Missing Data Problem

Example: Asprin's Impact on Headaches

Patient		Pill	Headache Status			Age	Academic
i		T_i	$Y_i($PLACEBO$)$	$Y_i($I-2$)$	Y_i	X_{1i}	X_{2i}
1		I-2				25	Y
2						55	N
3						62	Y
4		PLACEBO				80	N
5		I-2				32	Y
6						45	B
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n		PLACEBO				71	N

Step 3: Measure revealed potential outcome

Causal Inference is a Missing Data Problem

Example: Asprin's Impact on Headaches

Patient		Pill	Headache Status			Age	Academic
i		T_i	$Y_i(0)$	$Y_i(1)$	Y_i	X_{1i}	X_{2i}
1		1	0	0	0	25	Y
2			0	1		55	N
3			1	1		62	Y
4		0	1	1	1	80	N
5		1	0	1	1	32	Y
6			1	0		45	N
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n		0	0	0	0	71	N

Built in Assumptions

The notation implies three assumptions:

Built in Assumptions

The notation implies three assumptions:

- No simultaneity

Built in Assumptions

The notation implies three assumptions:

- No simultaneity
- No interference
  - We are implicitly stating that the potential outcomes for that unit are unaffected by the treatment status of other units
  - If this is not true, the number of potential outcomes for unit i grows
  - Ex: in an experiment with 3 units, if the potential outcomes for unit i depend on the treatment assignment of units j and k, the potential outcomes for unit i are defined by Y_{ijk}:

$$\begin{array}{ll} Y_{100} & Y_{000} \\ Y_{110} & Y_{010} \\ Y_{101} & Y_{001} \\ Y_{111} & Y_{011} \end{array}$$

Built in Assumptions

The notation implies three assumptions:

- No simultaneity
- No interference
  - We are implicitly stating that the potential outcomes for that unit are unaffected by the treatment status of other units
  - If this is not true, the number of potential outcomes for unit i grows
  - Ex: in an experiment with 3 units, if the potential outcomes for unit i depend on the treatment assignment of units j and k, the potential outcomes for unit i are defined by Y_{ijk}:

$$\begin{array}{cc} Y_{100} & Y_{000} \\ Y_{110} & Y_{010} \\ Y_{101} & Y_{001} \\ Y_{111} & Y_{011} \end{array}$$

- Same version of the treatment

How do we proceed?

Combined, the previous assumptions give us

How do we proceed?

Combined, the previous assumptions give us

- **S**table **U**nit **T**reatment **V**alue **A**ssumption (SUTVA)

How do we proceed?

Combined, the previous assumptions give us

- **S**table **U**nit **T**reatment **V**alue **A**ssumption (SUTVA)
- Potential violations:
  - feedback effects
  - spill-over effects, carry-over effects
  - different treatment administration

How do we proceed?

Combined, the previous assumptions give us

- **S**table **U**nit **T**reatment **V**alue **A**ssumption (SUTVA)
- Potential violations:
  - feedback effects
  - spill-over effects, carry-over effects
  - different treatment administration

We also need to assume Positivity $0 < p(T_i) < 1 \forall i$

Identification by assumption

Homogeneity

Identification by assumption

Homogeneity

- If the potential outcomes are constant across all i (i.e. $Y_i(1)$ and $Y_i(0)$ are the same for all individuals), then cross-sectional comparisons of treatment and control groups will recover τ_i

Identification by assumption

Homogeneity

- If the potential outcomes are constant across all i (i.e. $Y_i(1)$ and $Y_i(0)$ are the same for all individuals), then cross-sectional comparisons of treatment and control groups will recover τ_i
- If the potential outcomes are constant across time (i.e. $Y_i(1)$ and $Y_i(0)$ are time-invariant for all individuals), then pre-/ post-treatment comparisons will recover τ_i

Identification by assumption

Homogeneity

- If the potential outcomes are constant across all i (i.e. $Y_i(1)$ and $Y_i(0)$ are the same for all individuals), then cross-sectional comparisons of treatment and control groups will recover τ_i
- If the potential outcomes are constant across time (i.e. $Y_i(1)$ and $Y_i(0)$ are time-invariant for all individuals), then pre-/ post-treatment comparisons will recover τ_i

This is highly implausible in most social science

How do we proceed?

Identification by randomization:

How do we proceed?

Identification by randomization:

- If treatment is randomized, then treatment is unrelated to any and all underlying characteristics, observed and unobserved

How do we proceed?

Identification by randomization:

- If treatment is randomized, then treatment is unrelated to any and all underlying characteristics, observed and unobserved
- Randomization therefore means treatment assignment is independent of the potential outcomes $Y_i(1)$ and $Y_i(0)$, i.e.

$$\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i$$

How do we proceed?

Identification by randomization:

- If treatment is randomized, then treatment is unrelated to any and all underlying characteristics, observed and unobserved
- Randomization therefore means treatment assignment is independent of the potential outcomes $Y_i(1)$ and $Y_i(0)$, i.e.

$$\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i$$

- This is sometimes called unconfoundedness or ignorability

How do we proceed?

Identification by randomization:

- If treatment is randomized, then treatment is unrelated to any and all underlying characteristics, observed and unobserved
- Randomization therefore means treatment assignment is independent of the potential outcomes $Y_i(1)$ and $Y_i(0)$, i.e.

$$\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i$$

- This is sometimes called unconfoundedness or ignorability
- Another way of thinking of it: The distributions of the potential outcomes $(Y_i(1), Y_i(0))$ are the same for the treatment and control group.

How do we proceed?

Identification by randomization:

- If treatment is randomized, then treatment is unrelated to any and all underlying characteristics, observed and unobserved
- Randomization therefore means treatment assignment is independent of the potential outcomes $Y_i(1)$ and $Y_i(0)$, i.e.

$$\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i$$

- This is sometimes called unconfoundedness or ignorability
- Another way of thinking of it: The distributions of the potential outcomes $(Y_i(1),\ Y_i(0))$ are the same for the treatment and control group.
- Yet another way of thinking of it: The treatment and control group are exchangeable, or balanced (on observables and unobservables) on average

How do we proceed?

Identification by conditional independence:

- If treatment is not randomized, then treatment may be related underlying characteristics, observed and unobserved, which are related to the potential outcomes

How do we proceed?

Identification by conditional independence:

- If treatment is not randomized, then treatment may be related underlying characteristics, observed and unobserved, which are related to the potential outcomes
- Therefore, we need to assume that treatment assignment is independent of the potential outcomes $Y_i(1)$ and $Y_i(0)$, conditional on some pre-treatment characteristics X, i.e.

$$\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i \mid X_i$$

How do we proceed?

Identification by conditional independence:

- If treatment is not randomized, then treatment may be related underlying characteristics, observed and unobserved, which are related to the potential outcomes
- Therefore, we need to assume that treatment assignment is independent of the potential outcomes $Y_i(1)$ and $Y_i(0)$, conditional on some pre-treatment characteristics X, i.e.

$$\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i \mid X_i$$

- Conditioning set should yield $Y_i(0), Y_i(1)$ and T_i conditionally independent

Some Estimands of Interest

- **Sample average treatment effect (SATE)**
  $\frac{1}{n}\sum_{i=1}^{n}(Y_i(1) - Y_i(0))$

Some Estimands of Interest

- **Sample average treatment effect (SATE)**
  $\frac{1}{n}\sum_{i=1}^{n}(Y_i(1) - Y_i(0))$
- **Population average treatment effect (PATE)**
  $\frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$

Some Estimands of Interest

- **Sample average treatment effect (SATE)**
  $\frac{1}{n}\sum_{i=1}^{n}(Y_i(1) - Y_i(0))$
- **Population average treatment effect (PATE)**
  $\frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$
- **Population average treatment effect for the treated (PATT)**
  $\mathbb{E}(Y_i(1) - Y_i(0) \mid T_i = 1)$

Some Estimands of Interest

- **Sample average treatment effect (SATE)**
  $\frac{1}{n}\sum_{i=1}^{n}(Y_i(1) - Y_i(0))$
- **Population average treatment effect (PATE)**
  $\frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$
- **Population average treatment effect for the treated (PATT)**
  $\mathbb{E}(Y_i(1) - Y_i(0) \mid T_i = 1)$
- **Population conditional average treatment effect (CATE)**
  $\mathbb{E}(Y_i(1) - Y_i(0) \mid X_i = x)$

Some Estimands of Interest

- **Sample average treatment effect (SATE)**
  $\frac{1}{n}\sum_{i=1}^{n}(Y_i(1) - Y_i(0))$
- **Population average treatment effect (PATE)**
  $\frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$
- **Population average treatment effect for the treated (PATT)**
  $\mathbb{E}(Y_i(1) - Y_i(0) \mid T_i = 1)$
- **Population conditional average treatment effect (CATE)**
  $\mathbb{E}(Y_i(1) - Y_i(0) \mid X_i = x)$
- Treatment effect heterogeneity: Zero ATE doesn't mean zero effect for everyone

Three Big Assumptions

1. SUTVA
2. Positivity
3. (Conditional) Ignorability

The Selection Problem

- Why is this difficult? selection bias

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

$$
\begin{aligned}
& E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
& \quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0]
\end{aligned}
$$

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

$$
\begin{aligned}
&E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]
\end{aligned}
$$

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

$$\begin{aligned}
&E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = \underbrace{E[Y_i(1) - Y_i(0)|D_i = 1]}_{\text{Average Treatment Effect on Treated}} + \underbrace{E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]}_{\text{selection bias}}
\end{aligned}$$

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

$$
\begin{aligned}
&E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = \underbrace{E[Y_i(1) - Y_i(0)|D_i = 1]}_{\text{Average Treatment Effect on Treated}} + \underbrace{E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]}_{\text{selection bias}}
\end{aligned}
$$

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

$$
\begin{aligned}
& E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
& \quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
& \quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
& \quad = \underbrace{E[Y_i(1) - Y_i(0)|D_i = 1]}_{\text{Average Treatment Effect on Treated}} + \underbrace{E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]}_{\text{selection bias}}
\end{aligned}
$$

- Naive estimator = Average Treatment Effect on Treated + Selection Bias

The Selection Problem

- Why is this difficult? selection bias
- The core idea is that the people who get treatment might look different from those who get control and thus they are not good counterfactuals for each other.
- Let's look at what we get from a naive difference in means with a binary treatment:

$$
\begin{aligned}
&E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0] \\
&\quad = \underbrace{E[Y_i(1) - Y_i(0)|D_i = 1]}_{\text{Average Treatment Effect on Treated}} + \underbrace{E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]}_{\text{selection bias}}
\end{aligned}
$$

- Naive estimator = Average Treatment Effect on Treated + Selection Bias
- Selection bias: how different the treated and control groups are in terms of their potential outcome under control.

Selection Makes Us Care About Assignment Mechanisms

Selection Makes Us Care About Assignment Mechanisms

Assignment Mechanism

"The process that determines which units receive which treatments, hence which potential outcomes are realized and thus can be observed, and, conversely, which potential outcomes are missing."
(Imbens and Rubin, 2015, p. 31)

Selection Makes Us Care About Assignment Mechanisms

Assignment Mechanism

"The process that determines which units receive which treatments, hence which potential outcomes are realized and thus can be observed, and, conversely, which potential outcomes are missing."
(Imbens and Rubin, 2015, p. 31)

Key Assumptions:

- Individualistic assignment: Limits the dependence of a particular unit's assignment probability on the values of the covariates and potential outcomes for other units

Selection Makes Us Care About Assignment Mechanisms

Assignment Mechanism

"The process that determines which units receive which treatments, hence which potential outcomes are realized and thus can be observed, and, conversely, which potential outcomes are missing."
(Imbens and Rubin, 2015, p. 31)

Key Assumptions:

- Individualistic assignment: Limits the dependence of a particular unit's assignment probability on the values of the covariates and potential outcomes for other units
- Probabilistic assignment: Requires the assignment mechanism to imply a non-zero probability for each treatment value, for every unit

Selection Makes Us Care About Assignment Mechanisms

Assignment Mechanism

"The process that determines which units receive which treatments, hence which potential outcomes are realized and thus can be observed, and, conversely, which potential outcomes are missing."
(Imbens and Rubin, 2015, p. 31)

Key Assumptions:

- Individualistic assignment: Limits the dependence of a particular unit's assignment probability on the values of the covariates and potential outcomes for other units
- Probabilistic assignment: Requires the assignment mechanism to imply a non-zero probability for each treatment value, for every unit
- Unconfounded assignment: Disallows dependence of the assignment mechanism on the potential outcomes

What Gets to Be a Cause?

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion
- Effect of race on income

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion
- Effect of race on income

Consider causal effect of gender on promotion:

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion
- Effect of race on income

Consider causal effect of gender on promotion:

- Do we mean gender reassignment surgery?

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion
- Effect of race on income

Consider causal effect of gender on promotion:

- Do we mean gender reassignment surgery?
- Do we mean randomly assigning at birth? (a lot of other stuff different)

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion
- Effect of race on income

Consider causal effect of gender on promotion:

- Do we mean gender reassignment surgery?
- Do we mean randomly assigning at birth? (a lot of other stuff different)
- one idea: manipulate perceptions–women evaluated differently on paper

What Gets to Be a Cause?

We can imagine a world where individual i is assigned to treatment and control conditions

What is the Hypothetical Experiment?

Problem: Immutable (or difficult to change) characteristics

- Effect of gender on promotion
- Effect of race on income

Consider causal effect of gender on promotion:

- Do we mean gender reassignment surgery?
- Do we mean randomly assigning at birth? (a lot of other stuff different)
- one idea: manipulate perceptions–women evaluated differently on paper

No Causation Without Manipulation

Caveats and Implications

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?
- Correlation problem (1) :

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?
- Correlation problem (1) :
  - Regression models can estimate coefficients for immutable characteristics

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?
- Correlation problem (1) :
  - Regression models can estimate coefficients for immutable characteristics
  - But are necessarily imprecise: what do scholars have in mind in models?

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?
- Correlation problem (1) :
  - Regression models can estimate coefficients for immutable characteristics
  - But are necessarily imprecise: what do scholars have in mind in models?
- Design Principle:

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?
- Correlation problem (1) :
  - Regression models can estimate coefficients for immutable characteristics
  - But are necessarily imprecise: what do scholars have in mind in models?
- Design Principle:
  - Pretend you're God designing experiment

Caveats and Implications

- Does not dismiss claims of discrimination on immutable characteristics as legitimate
  - Pervasive effects of racism/sexism in society
  - Suggests: we need a different empirical strategy to evaluate claims
  - What facet of institutionalized racism (or its consequences) causes racial disparities?
- Correlation problem (1) :
  - Regression models can estimate coefficients for immutable characteristics
  - But are necessarily imprecise: what do scholars have in mind in models?
- Design Principle:
  - Pretend you're God designing experiment
  - If that experiment does not exist, be concerned about interpretation

Average Treatment Effects

Average Treatment Effects

Move the goal posts:

Average Treatment Effects

Move the goal posts:
Focus on estimating Average Treatment Effect (ATE)

Average Treatment Effects

Move the goal posts:

Focus on estimating Average Treatment Effect (ATE)

Suppose we have N observations in population ($i = 1, \ldots, N$)

Average Treatment Effects

Move the goal posts:
Focus on estimating Average Treatment Effect (ATE)
Suppose we have N observations in population ($i = 1, \ldots, N$)

$$\text{ATE} = \frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$$

Average Treatment Effects

Move the goal posts:
Focus on estimating Average Treatment Effect (ATE)
Suppose we have N observations in population $(i = 1, \ldots, N)$

$$\begin{aligned} \text{ATE} &= \frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0)) \\ &= E[Y(1) - Y(0)] \text{ Average over population!!!} \end{aligned}$$

Average Treatment Effects

Move the goal posts:
Focus on estimating Average Treatment Effect (ATE)
Suppose we have N observations in population $(i = 1, \ldots, N)$

$$\begin{aligned} \text{ATE} &= \frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0)) \\ &= E[Y(1) - Y(0)] \text{ Average over population!!!} \end{aligned}$$

- Population parameter

Average Treatment Effects

Move the goal posts:
Focus on estimating Average Treatment Effect (ATE)
Suppose we have N observations in population ($i = 1, \dots, N$)

$$\begin{aligned} \text{ATE} &= \frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0)) \\ &= E[Y(1) - Y(0)] \text{ Average over population!!!} \end{aligned}$$

- Population parameter
- It is fixed and unchanging

Estimating ATE under Random Assignment

Estimating ATE under Random Assignment

Estimator for ATE:

Estimating ATE under Random Assignment

Estimator for ATE:

$$\widehat{\text{ATE}} = \text{Average (Treated Units)} - \text{Average (Control Units)}$$

Estimating ATE under Random Assignment

Estimator for ATE:

$$\begin{aligned}\widehat{\text{ATE}} &= \text{Average (Treated Units)} - \text{Average (Control Units)} \\ &= \frac{\sum_{i=1}^{N} Y_i(1) T_i}{\sum_{i=1}^{N} T_i} - \frac{\sum_{i=1}^{N} Y_i(0)(1 - T_i)}{\sum_{i=1}^{N}(1 - T_i)}\end{aligned}$$

Estimating ATE under Random Assignment

Estimator for ATE:

$$
\begin{aligned}
\widehat{\text{ATE}} &= \text{Average (Treated Units)} - \text{Average (Control Units)} \\
&= \frac{\sum_{i=1}^N Y_i(1)T_i}{\sum_{i=1}^N T_i} - \frac{\sum_{i=1}^N Y_i(0)(1-T_i)}{\sum_{i=1}^N (1-T_i)} \\
&= \sum_{i=1}^N [\frac{Y_i(1)T_i}{n_t} - \frac{Y_i(0)(1-T_i)}{n_c}]
\end{aligned}
$$

Estimating ATE under Random Assignment

Estimator for ATE:

$$
\begin{aligned}
\widehat{\text{ATE}} &= \text{Average (Treated Units)} - \text{Average (Control Units)} \\
&= \frac{\sum_{i=1}^{N} Y_i(1) T_i}{\sum_{i=1}^{N} T_i} - \frac{\sum_{i=1}^{N} Y_i(0)(1 - T_i)}{\sum_{i=1}^{N}(1 - T_i)} \\
&= \sum_{i=1}^{N}\left[\frac{Y_i(1) T_i}{n_t} - \frac{Y_i(0)(1 - T_i)}{n_c}\right] \\
&= E[Y(1)|T = 1] - E[Y(0)|T = 0]
\end{aligned}
$$

Average Treatment Effect

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATE?

Average Treatment Effect

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATE?

$$E[Y_{1i} - Y_{0i}] = 1/4 \times (6 + -1 + 0 + 3) = 2$$

Average Treatment Effect

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATE?

$$E[Y_{1i} - Y_{0i}] = 1/4 \times (6 + -1 + 0 + 3) = 2$$

Note: Average effect is positive, but τ_i are negative for some units!

Average Treatment Effect on the Treated

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATT and ATC?

Average Treatment Effect on the Treated

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATT and ATC?

$E[Y_{1i} - Y_{0i} | D_i = 1] = 1/2 \times (6 + -1) = 2.5$

Average Treatment Effect on the Treated

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATT and ATC?

$$E[Y_{1i} - Y_{0i} | D_i = 1] = 1/2 \times (6 + -1) = 2.5$$

$$E[Y_{1i} - Y_{0i} | D_i = 0] = 1/2 \times (0 + 3) = 1.5$$

Naive Comparison: Difference in Means

Naive Comparison: Difference in Means

Comparisons between observed outcomes of treated and control units can often be misleading.

Naive Comparison: Difference in Means

Comparisons between observed outcomes of treated and control units can often be misleading.

- units which select treatment may not be like units which select control.

Naive Comparison: Difference in Means

Comparisons between observed outcomes of treated and control units can often be misleading.

- units which select treatment may not be like units which select control.
- i.e. selection into treatment is often associated with the potential outcomes

Naive Comparison: Difference in Means

Comparisons between observed outcomes of treated and control units can often be misleading.

- units which select treatment may not be like units which select control.
- i.e. selection into treatment is often associated with the potential outcomes
- this means we have violated the assumption of unconfoundness $(Y(1), Y(0)) \perp D$

Selection Bias

Selection Bias

Example: Church Attendance and Political Participation

Selection Bias

Example: Church Attendance and Political Participation

- Church goers likely to differ from non-Church goers on a range of background characteristics (e.g. civic duty)

Selection Bias

Example: Church Attendance and Political Participation

- Church goers likely to differ from non-Church goers on a range of background characteristics (e.g. civic duty)
- Given these differences, turnout for churchgoers could be higher than for non-churchgoers even if church had zero mobilizing effect

Selection Bias

Example: Church Attendance and Political Participation

- Church goers likely to differ from non-Church goers on a range of background characteristics (e.g. civic duty)
- Given these differences, turnout for churchgoers could be higher than for non-churchgoers even if church had zero mobilizing effect

Example: Gender Quotas and Redistribution Towards Women

Selection Bias

Example: Church Attendance and Political Participation

- Church goers likely to differ from non-Church goers on a range of background characteristics (e.g. civic duty)
- Given these differences, turnout for churchgoers could be higher than for non-churchgoers even if church had zero mobilizing effect

Example: Gender Quotas and Redistribution Towards Women

- Countries with gender quotas are likely countries where women are politically mobilized.

Selection Bias

Example: Church Attendance and Political Participation

- Church goers likely to differ from non-Church goers on a range of background characteristics (e.g. civic duty)
- Given these differences, turnout for churchgoers could be higher than for non-churchgoers even if church had zero mobilizing effect

Example: Gender Quotas and Redistribution Towards Women

- Countries with gender quotas are likely countries where women are politically mobilized.
- Given this difference, policies targeted towards women would be more common in quota countries even if these countries had not adopted quotas.

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill in, i.e. **estimate** missing potential outcomes.

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

Types of Assignment Mechanisms

- random assignment

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

Types of Assignment Mechanisms

- random assignment
- selection on observables

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

Types of Assignment Mechanisms

- random assignment
- selection on observables
- selection on unobservables

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

Types of Assignment Mechanisms

- random assignment
- selection on observables
- selection on unobservables

Most statistical models of causal inference attain identification of treatment effects by restricting the assignment mechanism in some way.

$$\boxed{\text{No causation without manipulation?}}$$

$\boxed{\text{No causation without manipulation?}}$

Always ask:
what is the experiment I would run if I had infinite resources and power?

Causal Inference Workflow

Causal Relationship

↓

Ideal Experiment

↓

Quantity of Interest

↓

Identification Strategy

↓

Data → Estimator → Estimate

Summing Up: Neyman-Rubin causal model

Summing Up: Neyman-Rubin causal model

- Useful for studying the "effects of causes", less so for the "causes of effects".

Summing Up: Neyman-Rubin causal model

- Useful for studying the "effects of causes", less so for the "causes of effects".
- No assumption of homogeneity, allows for causal effects to vary unit by unit

Summing Up: Neyman-Rubin causal model

- Useful for studying the "effects of causes", less so for the "causes of effects".
- No assumption of homogeneity, allows for causal effects to vary unit by unit
  - No single "causal effect", thus the need to be precise about the target estimand.

Summing Up: Neyman-Rubin causal model

- Useful for studying the "effects of causes", less so for the "causes of effects".
- No assumption of homogeneity, allows for causal effects to vary unit by unit
  - No single "causal effect", thus the need to be precise about the target estimand.
- Distinguishes between <u>observed</u> outcomes and <u>potential</u> outcomes.

Summing Up: Neyman-Rubin causal model

- Useful for studying the "effects of causes", less so for the "causes of effects".
- No assumption of homogeneity, allows for causal effects to vary unit by unit
  - No single "causal effect", thus the need to be precise about the target estimand.
- Distinguishes between observed outcomes and potential outcomes.
- Causal inference is a missing data problem: we typically make assumptions about the assignment mechanism to go from descriptive inference to causal inference.

Summary: Observational Studies and Causal Inference

Experimental studies:

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status
- Selection: treatment and control groups differ systematically

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status
- Selection: treatment and control groups differ systematically
  - $E[Y(1)|T=1] \neq E[Y(1)|T=0]$, $E[Y(0)|T=0] \neq E[Y(0)|T=1]$

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status
- Selection: treatment and control groups differ systematically
  - $E[Y(1)|T=1] \neq E[Y(1)|T=0]$, $E[Y(0)|T=0] \neq E[Y(0)|T=1]$
  - Observables: things we can see, measure, and use in our study

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status
- Selection: treatment and control groups differ systematically
  - $E[Y(1)|T=1] \neq E[Y(1)|T=0]$, $E[Y(0)|T=0] \neq E[Y(0)|T=1]$
  - Observables: things we can see, measure, and use in our study
  - Unobservables: not observables (big problem)

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status
- Selection: treatment and control groups differ systematically
  - $E[Y(1)|T=1] \neq E[Y(1)|T=0]$, $E[Y(0)|T=0] \neq E[Y(0)|T=1]$
  - Observables: things we can see, measure, and use in our study
  - Unobservables: not observables (big problem)
- Naive difference in means will be biased

Summary: Observational Studies and Causal Inference

Experimental studies:

- Treatment under control of analyst

Observational

- Units (people, countries) control their treatment status
- Selection: treatment and control groups differ systematically
  - $E[Y(1)|T=1] \neq E[Y(1)|T=0]$, $E[Y(0)|T=0] \neq E[Y(0)|T=1]$
  - Observables: things we can see, measure, and use in our study
  - Unobservables: not observables (big problem)
- Naive difference in means will be biased
- Many, many, potential strategies for limiting bias

Avoiding Common Areas of Confusion

Avoiding Common Areas of Confusion

- ‘a causal claim is a statement about what didnt happen’

Avoiding Common Areas of Confusion

- 'a causal claim is a statement about what didnt happen'
- contribution not attribution: we care about a difference which doesnt make it the main reason, nor does it imply a morality claim, it doesnt make T the reason it happened, it doesnt mean that T is responsible for Y

Avoiding Common Areas of Confusion

- 'a causal claim is a statement about what didnt happen'
- contribution not attribution: we care about a difference which doesnt make it the main reason, nor does it imply a morality claim, it doesnt make T the reason it happened, it doesnt mean that T is responsible for Y
- T can 'cause' Y if it is neither necessary nor sufficient

Avoiding Common Areas of Confusion

- 'a causal claim is a statement about what didnt happen'
- contribution not attribution: we care about a difference which doesnt make it the main reason, nor does it imply a morality claim, it doesnt make T the reason it happened, it doesnt mean that T is responsible for Y
- T can 'cause' Y if it is neither necessary nor sufficient
- If you know that on average A causes B and B causes C this doesnt mean you know that A causes C (example A$\rightarrow$B for one subgroup, B$\rightarrow$C for second subgroup, still no A$\rightarrow$C)

Avoiding Common Areas of Confusion

- ‘a causal claim is a statement about what didnt happen’
- contribution not attribution: we care about a difference which doesnt make it the main reason, nor does it imply a morality claim, it doesnt make T the reason it happened, it doesnt mean that T is responsible for Y
- T can ‘cause’ Y if it is neither necessary nor sufficient
- If you know that on average A causes B and B causes C this doesnt mean you know that A causes C (example A$\rightarrow$B for one subgroup, B$\rightarrow$C for second subgroup, still no A$\rightarrow$C)
- estimation of causal effects does not require identical treatment and control groups

Avoiding Common Areas of Confusion

- ‘a causal claim is a statement about what didnt happen’
- contribution not attribution: we care about a difference which doesnt make it the main reason, nor does it imply a morality claim, it doesnt make T the reason it happened, it doesnt mean that T is responsible for Y
- T can ‘cause’ Y if it is neither necessary nor sufficient
- If you know that on average A causes B and B causes C this doesnt mean you know that A causes C (example A$\rightarrow$B for one subgroup, B$\rightarrow$C for second subgroup, still no A$\rightarrow$C)
- estimation of causal effects does not require identical treatment and control groups
- you need a clear counterfactual to have a well-defined causal effect (hence no causation without manipulation). For example of ‘the recession was caused by Wall Street’ may make intuitive sense but is it well-defined?

http://egap.org/methods-guides/10-things-you-need-know-about-causal-inference

1. Making Arguments
2. Potential Outcomes
3. Average Treatment effects
4. Graphical Models
5. Fun With A Bundle of Sticks
6. Visualization

Graphical Models

Graphical Models

- A general framework for representing causal relationships based on directed acyclic graphs (DAG)

Graphical Models

- A general framework for representing causal relationships based on directed acyclic graphs (DAG)
- The work we discuss here comes out of developments by Judea Pearl and others

Graphical Models

- A general framework for representing causal relationships based on directed acyclic graphs (DAG)
- The work we discuss here comes out of developments by Judea Pearl and others
- Particularly useful for thinking through issues of identification.

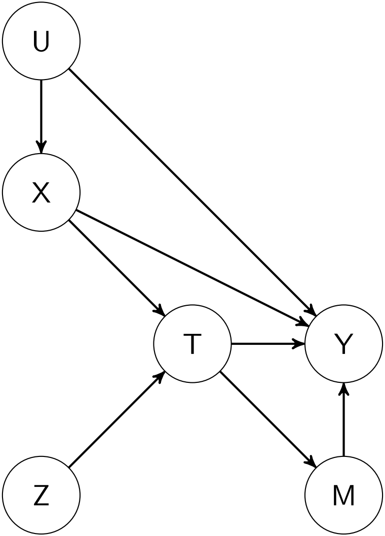
Graphical Models

- A general framework for representing causal relationships based on directed acyclic graphs (DAG)
- The work we discuss here comes out of developments by Judea Pearl and others
- Particularly useful for thinking through issues of identification.
- Provides a graphical representation of the models and a set of rules (do-calculus) for identifying the causal effect.

Graphical Models

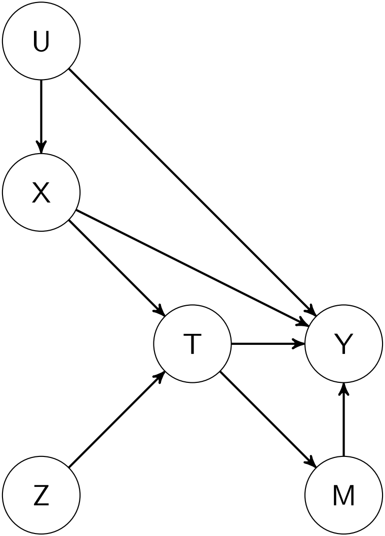
- A general framework for representing causal relationships based on directed acyclic graphs (DAG)
- The work we discuss here comes out of developments by Judea Pearl and others
- Particularly useful for thinking through issues of identification.
- Provides a graphical representation of the models and a set of rules (do-calculus) for identifying the causal effect.
- Nice software that takes the graph and returns an identification strategy: DAGitty at `http://dagitty.net`

Components of a DAG



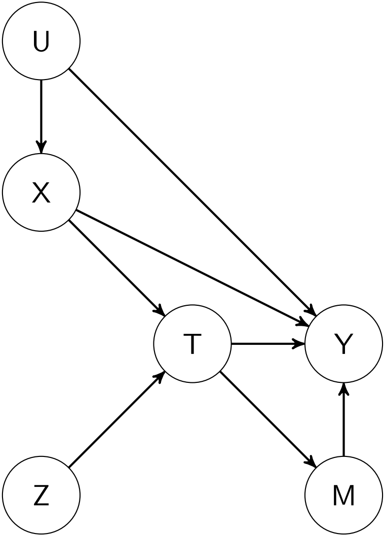
- nodes $\rightarrow$ variables
  (unobserved typically called U or V)

Components of a DAG



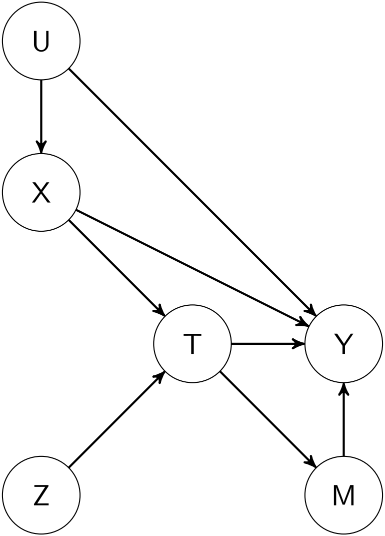
- nodes $\rightarrow$ variables
  (unobserved typically called U or V)
- (directed) arrows $\rightarrow$ causal effects

Components of a DAG



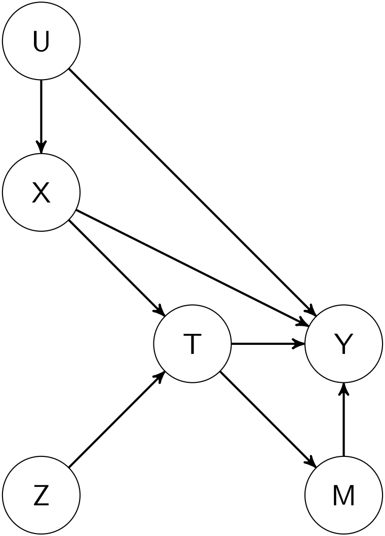
- nodes $\rightarrow$ variables
  (unobserved typically called U or V)
- (directed) arrows $\rightarrow$ causal effects
- absence of nodes $\rightarrow$ no common causes of any pair of variables

Components of a DAG



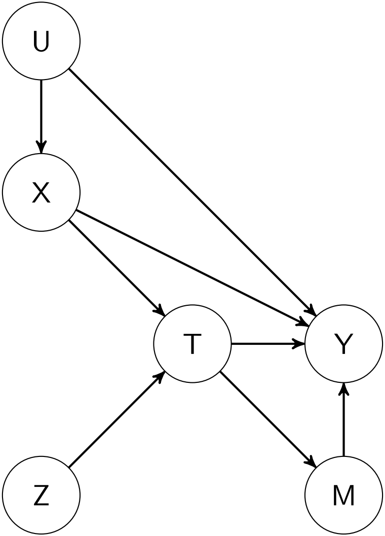
- nodes $\rightarrow$ variables
  (unobserved typically called U or V)
- (directed) arrows $\rightarrow$ causal effects
- absence of nodes $\rightarrow$ no common causes of any pair of variables
- absence of arrows $\rightarrow$ no causal effect

Components of a DAG



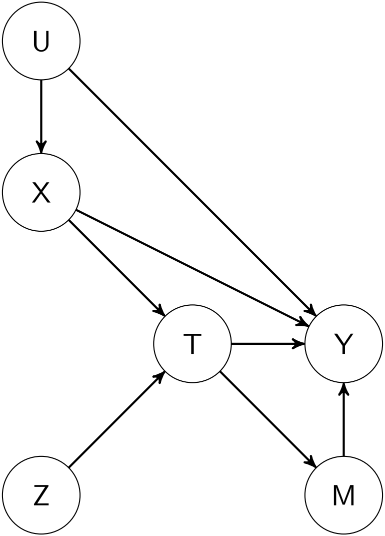
- nodes $\rightarrow$ variables
  (unobserved typically called U or V)
- (directed) arrows $\rightarrow$ causal effects
- absence of nodes $\rightarrow$ no common causes of any pair of variables
- absence of arrows $\rightarrow$ no causal effect
- positioning conveys no mathematical meaning but often is oriented left-to-right with causal ordering for readability.

Components of a DAG



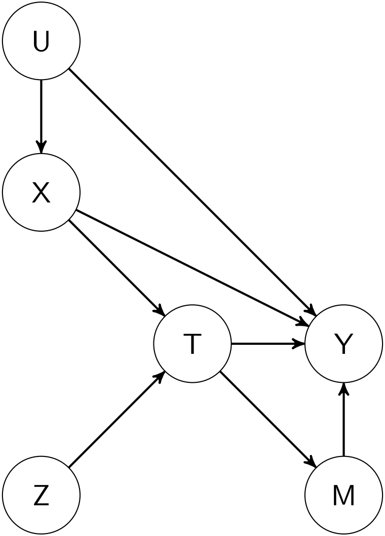
- nodes $\rightarrow$ variables
  (unobserved typically called U or V)
- (directed) arrows $\rightarrow$ causal effects
- absence of nodes $\rightarrow$ no common causes of any pair of variables
- absence of arrows $\rightarrow$ no causal effect
- positioning conveys no mathematical meaning but often is oriented left-to-right with causal ordering for readability.
- dashed lines are used in context dependent ways

Components of a DAG



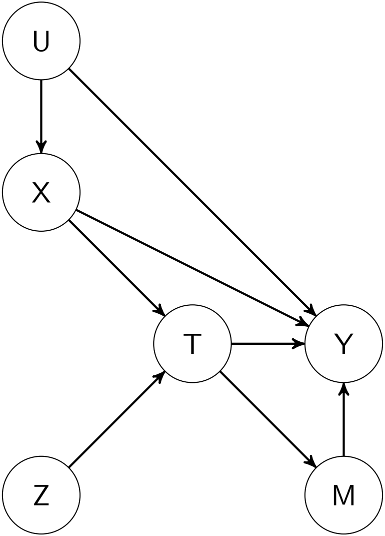
- nodes → variables (unobserved typically called U or V)
- (directed) arrows → causal effects
- absence of nodes → no common causes of any pair of variables
- absence of arrows → no causal effect
- positioning conveys no mathematical meaning but often is oriented left-to-right with causal ordering for readability.
- dashed lines are used in context dependent ways
- all relationships are non-parametric

Relationships in a DAG



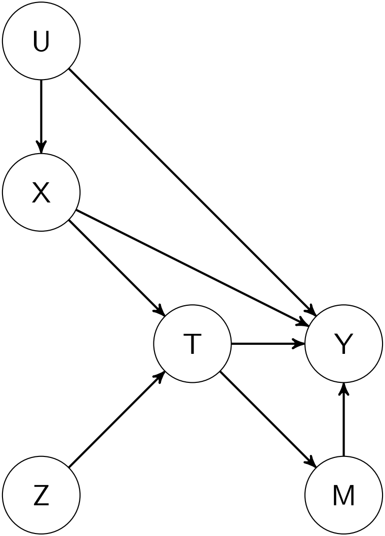
- Parents (Children): directly causing (caused by) a node

Relationships in a DAG



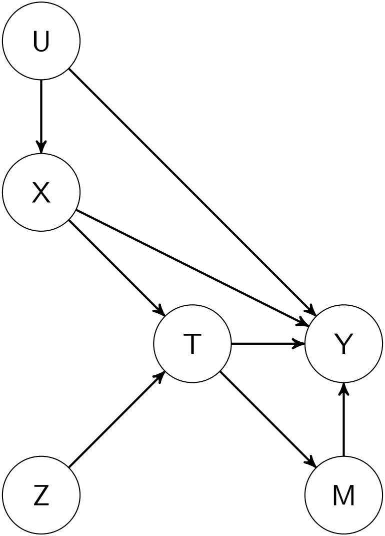
- Parents (Children): directly causing (caused by) a node
- Ancestors (Descendants): directly or indirectly causing (caused by) a node

Relationships in a DAG



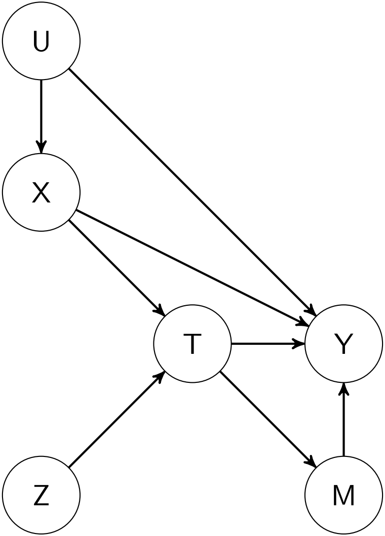
- Parents (Children): directly causing (caused by) a node
- Ancestors (Descendants): directly or indirectly causing (caused by) a node
- Path: a route that connects the variables (path is causal when all arrows point the same way)

Relationships in a DAG



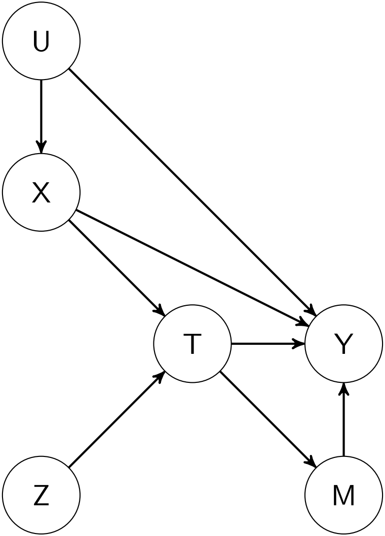
- Parents (Children): directly causing (caused by) a node
- Ancestors (Descendants): directly or indirectly causing (caused by) a node
- Path: a route that connects the variables (path is causal when all arrows point the same way)
- Acyclic implies that there are no cycles and a variable can't cause itself

Relationships in a DAG



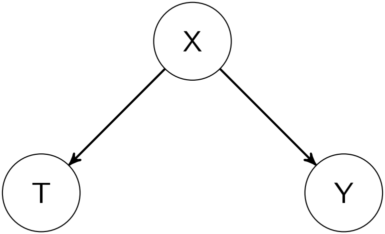
- Parents (Children): directly causing (caused by) a node
- Ancestors (Descendants): directly or indirectly causing (caused by) a node
- Path: a route that connects the variables (path is causal when all arrows point the same way)
- Acyclic implies that there are no cycles and a variable can't cause itself
- Causal Markov assumption: condition on its direct causes, a variable is independent of its non-descendents.

Relationships in a DAG



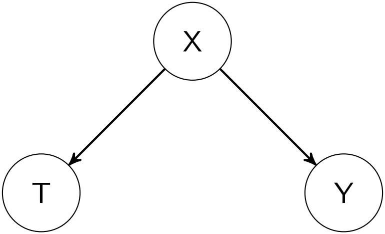
- Parents (Children): directly causing (caused by) a node
- Ancestors (Descendants): directly or indirectly causing (caused by) a node
- Path: a route that connects the variables (path is causal when all arrows point the same way)
- Acyclic implies that there are no cycles and a variable can't cause itself
- Causal Markov assumption: condition on its direct causes, a variable is independent of its non-descendents.
- We will talk in depth about two types of relationships: confounders and colliders

Confounders



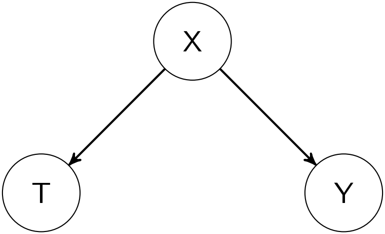
- X is a confounder (or common cause)

Confounders



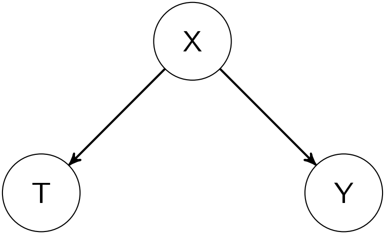
- X is a confounder (or common cause)
- Even without a causal effect or directed edge between T and Y they will have a marginal associational relationship

Confounders



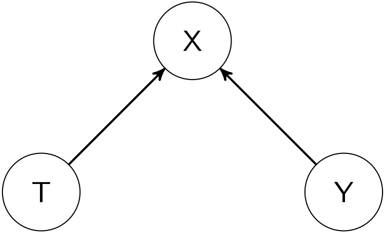
- X is a confounder (or common cause)
- Even without a causal effect or directed edge between T and Y they will have a marginal associational relationship
- Conditional on X, T and Y are unrelated in this graph.

Confounders



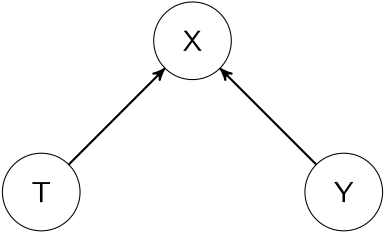
- X is a confounder (or common cause)
- Even without a causal effect or directed edge between T and Y they will have a marginal associational relationship
- Conditional on X, T and Y are unrelated in this graph.
- We can think of conditioning on a confounder as blocking the flow of association.

Colliders



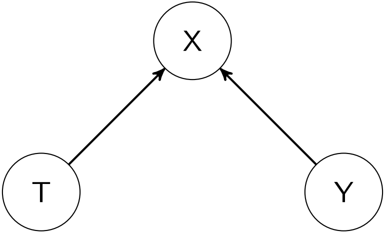
- X is now a collider because two arrows point into it

Colliders



- X is now a collider because two arrows point into it
- In this scenario T and Y are not marginally associated

Colliders



- X is now a collider because two arrows point into it
- In this scenario T and Y are not marginally associated
- If we control for X they become associated and create a connection between T and Y

Colliders are scary because you can induce dependence

Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable

Felix Elwert[1] and Christopher Winship[2]

[1]Department of Sociology, University of Wisconsin, Madison, Wisconsin 53706; email: elwert@wisc.edu

[2]Department of Sociology, Harvard University, Cambridge, Massachusetts 02138; email: cwinship@wjh.harvard.edu

Annu. Rev. Sociol. 2014. 40:31–53

First published online as a Review in Advance on June 2, 2014

The *Annual Review of Sociology* is online at soc.annualreviews.org

This article's doi: 10.1146/annurev-soc-071913-043455

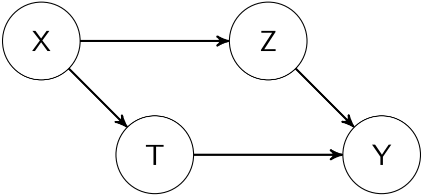
Keywords

causality, directed acyclic graphs, identification, confounding, selection

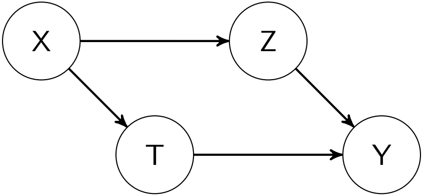
Abstract

Endogenous selection bias is a central problem for causal inference. Recognizing the problem, however, can be difficult in practice. This article introduces a purely graphical way of characterizing endogenous selection bias and of understanding its consequences (Hernán et al. 2004). We use causal graphs (direct acyclic graphs, or DAGs) to highlight that endogenous selection bias stems from conditioning (e.g., controlling, stratifying, or selecting) on a so-called collider variable, i.e., a variable that is itself caused by two other variables, one that is (or is associated with) the treatment and another that is (or is associated with) the outcome. Endogenous selection bias can result from direct conditioning on the outcome variable, a post-outcome variable, a post-treatment variable, and even a pre-treatment variable. We highlight the difference between endogenous selection bias, common-cause confounding, and overcontrol bias and discuss numerous examples from social stratification, cultural sociology, social network analysis, political sociology, social demography, and the sociology of education.

From Confounders to Back-Door Paths

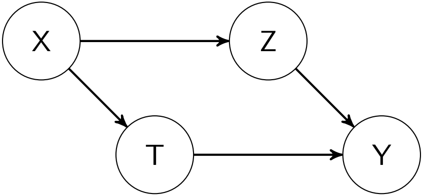


From Confounders to Back-Door Paths



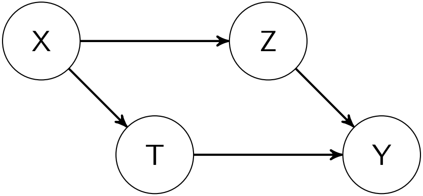
- Identify causal effect of T on Y by conditioning on X, Z or X and Z

From Confounders to Back-Door Paths



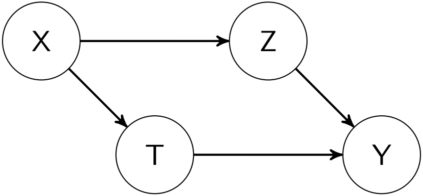
- Identify causal effect of T on Y by conditioning on X, Z or X and Z
- We can formalize this logic with the idea of a back-door path

From Confounders to Back-Door Paths



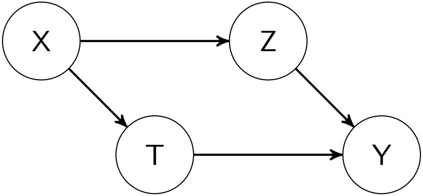
- Identify causal effect of T on Y by conditioning on X, Z or X and Z
- We can formalize this logic with the idea of a back-door path
- A back-door path is "a path between any causally ordered sequence of two variables that begins with a directed edge that points to the first variable." (Morgan and Winship 2013)

From Confounders to Back-Door Paths



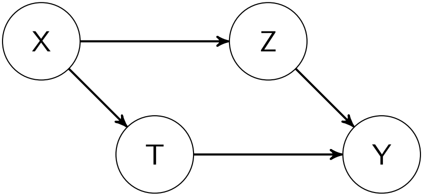
- Identify causal effect of T on Y by conditioning on X, Z or X and Z
- We can formalize this logic with the idea of a back-door path
- A back-door path is "a path between any causally ordered sequence of two variables that begins with a directed edge that points to the first variable." (Morgan and Winship 2013)
- Two paths from T to Y here:
  1. $T \rightarrow Y$ (directed or causal path)
  2. $T \leftarrow X \rightarrow Z \rightarrow Y$ (back-door path)

From Confounders to Back-Door Paths



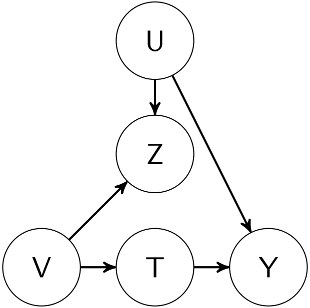
- Identify causal effect of T on Y by conditioning on X, Z or X and Z
- We can formalize this logic with the idea of a back-door path
- A back-door path is "a path between any causally ordered sequence of two variables that begins with a directed edge that points to the first variable." (Morgan and Winship 2013)
- Two paths from T to Y here:
  1. $T \rightarrow Y$ (directed or causal path)
  2. $T \leftarrow X \rightarrow Z \rightarrow Y$ (back-door path)
- Observed marginal association between T and Y is a composite of these two paths and thus does not identify the causal effect of T on Y

From Confounders to Back-Door Paths



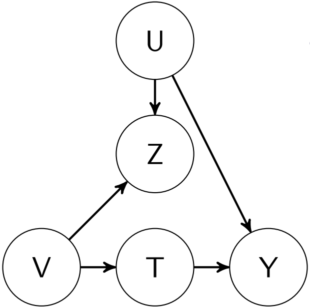
- Identify causal effect of T on Y by conditioning on X, Z or X and Z
- We can formalize this logic with the idea of a back-door path
- A back-door path is "a path between any causally ordered sequence of two variables that begins with a directed edge that points to the first variable." (Morgan and Winship 2013)
- Two paths from T to Y here:
  1. $T \rightarrow Y$ (directed or causal path)
  2. $T \leftarrow X \rightarrow Z \rightarrow Y$ (back-door path)
- Observed marginal association between T and Y is a composite of these two paths and thus does not identify the causal effect of T on Y
- We want to block the back-door path to leave only the causal effect

Colliders and Back-Door Paths



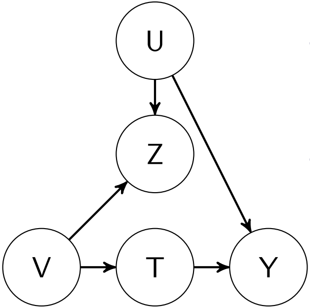
- Z is a collider and it lies along a back-door path from T to Y

Colliders and Back-Door Paths



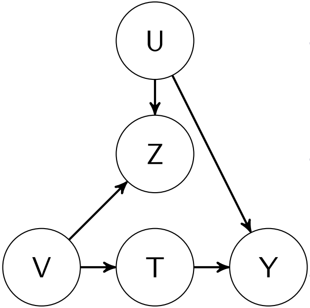
- Z is a collider and it lies along a back-door path from T to Y
- Conditioning on a collider on a back-door path does not help and in fact causes new associations

Colliders and Back-Door Paths



- Z is a collider and it lies along a back-door path from T to Y
- Conditioning on a collider on a back-door path does not help and in fact causes new associations
- Here we are fine unless we condition on Z which opens a path $T \leftarrow V \leftrightarrow U \rightarrow Y$ (this particular case is called M-bias)

Colliders and Back-Door Paths



- Z is a collider and it lies along a back-door path from T to Y
- Conditioning on a collider on a back-door path does not help and in fact causes new associations
- Here we are fine unless we condition on Z which opens a path $T \leftarrow V \leftrightarrow U \rightarrow Y$ (this particular case is called M-bias)
- So how do we know which back-door paths to block?

D-Separation

- Graphs provide us a way to think about conditional independence statements. Consider disjoint subsets of the vertices A, B and C

D-Separation

- Graphs provide us a way to think about conditional independence statements. Consider disjoint subsets of the vertices A, B and C
- A is D-separated from B by C if and only if C blocks every path from a vertex in A to a vertex in B

D-Separation

- Graphs provide us a way to think about conditional independence statements. Consider disjoint subsets of the vertices A, B and C
- A is *D*-separated from B by C if and only if C blocks every path from a vertex in A to a vertex in B
- A path p is said to be blocked by a set of vertices C if and only if at least one of the following conditions holds:

D-Separation

- Graphs provide us a way to think about conditional independence statements. Consider disjoint subsets of the vertices A, B and C
- A is *D*-separated from B by C if and only if C blocks every path from a vertex in A to a vertex in B
- A path p is said to be blocked by a set of vertices C if and only if at least one of the following conditions holds:
  1. p contains a chain structure $a \rightarrow c \rightarrow b$ or a fork structure $a \leftarrow c \rightarrow b$ where the node c is in the set C

D-Separation

- Graphs provide us a way to think about conditional independence statements. Consider disjoint subsets of the vertices A, B and C
- A is D-separated from B by C if and only if C blocks every path from a vertex in A to a vertex in B
- A path p is said to be blocked by a set of vertices C if and only if at least one of the following conditions holds:
  1. p contains a chain structure $a \rightarrow c \rightarrow b$ or a fork structure $a \leftarrow c \rightarrow b$ where the node c is in the set C
  2. p contains a collider structure $a \rightarrow y \leftarrow b$ where neither y nor its descendents are in C

D-Separation

- Graphs provide us a way to think about conditional independence statements. Consider disjoint subsets of the vertices A, B and C
- A is *D*-separated from B by C if and only if C blocks every path from a vertex in A to a vertex in B
- A path p is said to be blocked by a set of vertices C if and only if at least one of the following conditions holds:
    1. p contains a chain structure $a \rightarrow c \rightarrow b$ or a fork structure $a \leftarrow c \rightarrow b$ where the node c is in the set C
    2. p contains a collider structure $a \rightarrow y \leftarrow b$ where neither y nor its descendents are in C
- If A is not D-separated from B by C we say that A is *D*-connected to B by C

Backdoor Criterion

- Generally we want to know if we can nonparametrically identify the average effect of T on Y given a set of possible conditioning variables X

Backdoor Criterion

- Generally we want to know if we can nonparametrically identify the average effect of T on Y given a set of possible conditioning variables X
- Backdoor Criterion for X
  1. No node in X is a descendent of T
     (i.e. don't condition on post-treatment variables!)

Backdoor Criterion

- Generally we want to know if we can nonparametrically identify the average effect of T on Y given a set of possible conditioning variables X
- Backdoor Criterion for X
  1. No node in X is a descendent of T
     (i.e. don't condition on post-treatment variables!)
  2. X D-separates every path between T and Y that has an incoming arrow into T (backdoor path)

Backdoor Criterion

- Generally we want to know if we can nonparametrically identify the average effect of T on Y given a set of possible conditioning variables X
- Backdoor Criterion for X
  1. No node in X is a descendent of T
     (i.e. don't condition on post-treatment variables!)
  2. X D-separates every path between T and Y that has an incoming arrow into T (backdoor path)
- In essence, we are trying to block all non-causal paths, so we can estimate the causal path.

Backdoor Criterion

- Generally we want to know if we can nonparametrically identify the average effect of T on Y given a set of possible conditioning variables X
- Backdoor Criterion for X
  1. No node in X is a descendent of T
     (i.e. don't condition on post-treatment variables!)
  2. X D-separates every path between T and Y that has an incoming arrow into T (backdoor path)
- In essence, we are trying to block all non-causal paths, so we can estimate the causal path.
- Backdoor criterion is just one way to identify the effect: but its the most popular approach in the social sciences and what we are trying to do 99% of the time.

Backdoor Criterion

- Generally we want to know if we can nonparametrically identify the average effect of T on Y given a set of possible conditioning variables X
- Backdoor Criterion for X
  1. No node in X is a descendent of T
     (i.e. don't condition on post-treatment variables!)
  2. X D-separates every path between T and Y that has an incoming arrow into T (backdoor path)
- In essence, we are trying to block all non-causal paths, so we can estimate the causal path.
- Backdoor criterion is just one way to identify the effect: but its the most popular approach in the social sciences and what we are trying to do 99% of the time.
- We will see some other approaches late in the semester.

Thoughts on DAGs and Potential Outcomes

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.
- Potential outcomes is also less of a “foreign language” for most statisticians, but in my experience lumps together a lot of identification assumptions in opaque ignorability conditions.

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.
- Potential outcomes is also less of a "foreign language" for most statisticians, but in my experience lumps together a lot of identification assumptions in opaque ignorability conditions.
- Graphical Models with DAGs are very visually appealing but the operations on the graph can be challenging

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.
- Potential outcomes is also less of a "foreign language" for most statisticians, but in my experience lumps together a lot of identification assumptions in opaque ignorability conditions.
- Graphical Models with DAGs are very visually appealing but the operations on the graph can be challenging
- DAGs very helpful for thinking through identification and the entire causal process

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.
- Potential outcomes is also less of a “foreign language” for most statisticians, but in my experience lumps together a lot of identification assumptions in opaque ignorability conditions.
- Graphical Models with DAGs are very visually appealing but the operations on the graph can be challenging
- DAGs very helpful for thinking through identification and the entire causal process
- Note that both are about non-parametric identification and not estimation. This is good and bad.

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.
- Potential outcomes is also less of a "foreign language" for most statisticians, but in my experience lumps together a lot of identification assumptions in opaque ignorability conditions.
- Graphical Models with DAGs are very visually appealing but the operations on the graph can be challenging
- DAGs very helpful for thinking through identification and the entire causal process
- Note that both are about non-parametric identification and not estimation. This is good and bad.
    - Good: provides a very general framework that applies in non-linear scenarios and interactions

Thoughts on DAGs and Potential Outcomes

- Two very different languages for talking about and thinking about causal inferences
- Potential outcomes is very focused on thinking about the treatment assignment mechanism.
- Potential outcomes is also less of a “foreign language” for most statisticians, but in my experience lumps together a lot of identification assumptions in opaque ignorability conditions.
- Graphical Models with DAGs are very visually appealing but the operations on the graph can be challenging
- DAGs very helpful for thinking through identification and the entire causal process
- Note that both are about non-parametric identification and not estimation. This is good and bad.
  - Good: provides a very general framework that applies in non-linear scenarios and interactions
  - Bad: identification results for identification only holds when variable is completely controlled for (which may be difficult!)

Fun with a Bundle of Sticks

Sen and Wasow (2016) “Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics” *Annual Review of Political Science*.

No Causation Without Manipulation

No Causation Without Manipulation

- One of the difficulties that students have with causal inference is the need for manipulation or an ideal experiment.

No Causation Without Manipulation

- One of the difficulties that students have with causal inference is the need for manipulation or an ideal experiment.
- In many areas the key variables are immutable such as race or gender

No Causation Without Manipulation

- One of the difficulties that students have with causal inference is the need for manipulation or an ideal experiment.
- In many areas the key variables are immutable such as race or gender
- Sen and Wasow argue that we can improve our empirical work on this by seeing race/ethnicity as a composite variable or 'a bundle of sticks' which can be manipulated separately

The Trouble with Race As Treatment

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

1. Race cannot be manipulated

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

1. Race cannot be manipulated
   - without the capacity to manipulate the question is arguably ill-posed and the estimand is unidentified

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

1. Race cannot be manipulated
   - without the capacity to manipulate the question is arguably ill-posed and the estimand is unidentified
2. Everything else is post-treatment

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

1. Race cannot be manipulated
   - without the capacity to manipulate the question is arguably ill-posed and the estimand is unidentified
2. Everything else is post-treatment
   - everything else comes after race which is perhaps unsatisfying

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

1. Race cannot be manipulated
   - without the capacity to manipulate the question is arguably ill-posed and the estimand is unidentified
2. Everything else is post-treatment
   - everything else comes after race which is perhaps unsatisfying
   - this also presumes we are only interested in the total effect

The Trouble with Race As Treatment

There are three problems with race as a treatment in the causal inference sense

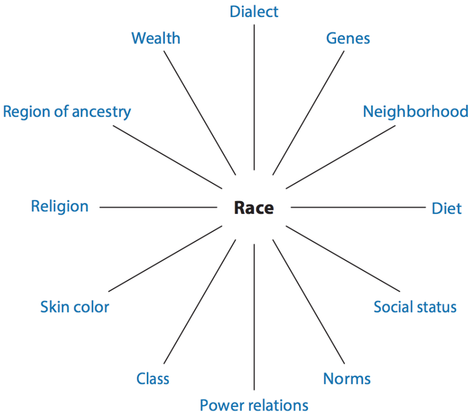
1. Race cannot be manipulated
    - without the capacity to manipulate the question is arguably ill-posed and the estimand is unidentified
2. Everything else is post-treatment
    - everything else comes after race which is perhaps unsatisfying
    - this also presumes we are only interested in the total effect
3. Race is unstable

The Trouble with Race As Treatment

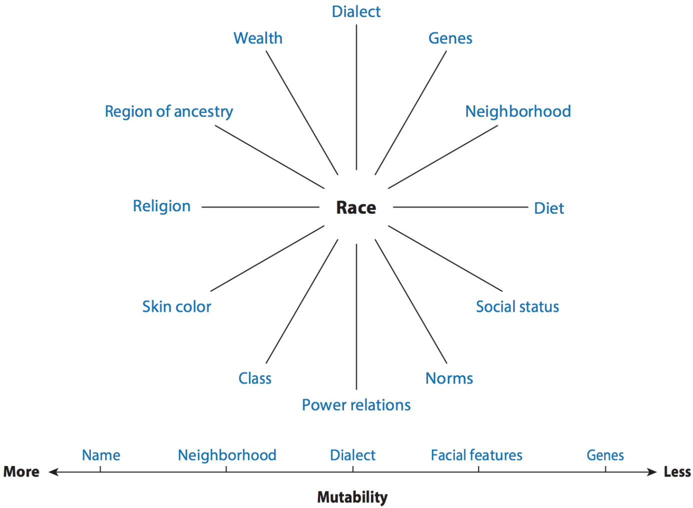
There are three problems with race as a treatment in the causal inference sense

1. Race cannot be manipulated
   - without the capacity to manipulate the question is arguably ill-posed and the estimand is unidentified
2. Everything else is post-treatment
   - everything else comes after race which is perhaps unsatisfying
   - this also presumes we are only interested in the total effect
3. Race is unstable
   - there is substantial variance across treatments which is a SUTVA violation

The Bundle of Sticks



The Bundle of Sticks



Design 1: Exposure Studies

- Approach

Design 1: Exposure Studies

- Approach
  a) "one or more elements of race is identified as a relevant cue"

Design 1: Exposure Studies

- Approach
  a) "one or more elements of race is identified as a relevant cue"
  b) "subjects are treated by exposure to the racial cue"

Design 1: Exposure Studies

- Approach
  a) "one or more elements of race is identified as a relevant cue"
  b) "subjects are treated by exposure to the racial cue"
  c) "unit of analysis is the individual or institution being exposed"

Design 1: Exposure Studies

- Approach
  a) "one or more elements of race is identified as a relevant cue"
  b) "subjects are treated by exposure to the racial cue"
  c) "unit of analysis is the individual or institution being exposed"
- Examples
  - Psychology (Steele 1997 on stereotype threat)

Design 1: Exposure Studies

- Approach
  a) “one or more elements of race is identified as a relevant cue”
  b) “subjects are treated by exposure to the racial cue”
  c) “unit of analysis is the individual or institution being exposed”
- Examples
  - Psychology (Steele 1997 on stereotype threat)
  - Audit/Correspondence Studies (Pager 2003, Bertrand and Mullainathan 2004)

Design 1: Exposure Studies

- Approach
  a) "one or more elements of race is identified as a relevant cue"
  b) "subjects are treated by exposure to the racial cue"
  c) "unit of analysis is the individual or institution being exposed"
- Examples
  - Psychology (Steele 1997 on stereotype threat)
  - Audit/Correspondence Studies (Pager 2003, Bertrand and Mullainathan 2004)
  - Survey Experiments with Racial Cues (Mendelberg 2001)

Design 1: Exposure Studies

- Approach
  a) “one or more elements of race is identified as a relevant cue”
  b) “subjects are treated by exposure to the racial cue”
  c) “unit of analysis is the individual or institution being exposed”
- Examples
  - Psychology (Steele 1997 on stereotype threat)
  - Audit/Correspondence Studies (Pager 2003, Bertrand and Mullainathan 2004)
  - Survey Experiments with Racial Cues (Mendelberg 2001)
  - Field Experiments with Racial Cues (Green 2004, Enos 2011)

Design 1: Exposure Studies

- Approach
  a) "one or more elements of race is identified as a relevant cue"
  b) "subjects are treated by exposure to the racial cue"
  c) "unit of analysis is the individual or institution being exposed"
- Examples
  - Psychology (Steele 1997 on stereotype threat)
  - Audit/Correspondence Studies (Pager 2003, Bertrand and Mullainathan 2004)
  - Survey Experiments with Racial Cues (Mendelberg 2001)
  - Field Experiments with Racial Cues (Green 2004, Enos 2011)
  - Observational Studies (Greiner and Rubin 2010, Wasow 2012)

Design 2: Within-Group Studies

- Approach: identify variation within the racial group along constitutive element.

Design 2: Within-Group Studies

- Approach: identify variation within the racial group along constitutive element.
- Example: Sharkey (2010) exploiting temporal variation in local homicides in Chicago to identify a significant neighborhood effect of proximity to violence on cognitive performance of African-American children

Concluding Thoughts

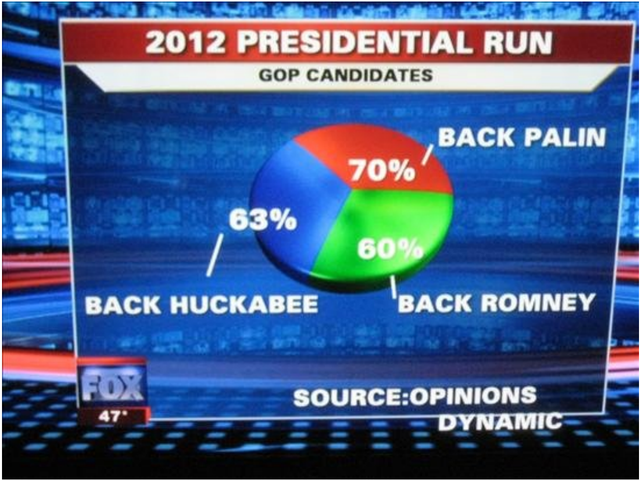
We can study race with causal inference, it just takes very careful design.

Table 2 Overview of exposure and within-group research designs

	Exposure	Within-Group
Unit	Individuals or institutions, potentially from any group	Members of a particular group
Typical treatment	Racial cue or signal (e.g., include distinctively ethnic names on a resume)	Constitutive element of the composite of race (e.g., address anxiety about social belonging in college)
Role of element of race	One "stick" is a proxy for the bundle (e.g., in a phone call with a landlord, dialect signals many traits associated with race)	One "stick" explains part of the bundle (e.g., Middle Passage might partly explain high rates of hypertension among African-Americans)
Examples	Correspondence and audit studies Implicit Association Tests	Experimental manipulation of a constitutive psychological dimension of race Within-race matching

An Intro Motivation

An Intro Motivation



Visualization

Visualization

- Visualization is hard but ultimately extremely important

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking
- Three steps involved

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking
- Three steps involved
  1. clearly define the goal

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking
- Three steps involved
  1. clearly define the goal
  2. estimate quantities of interest

Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking
- Three steps involved
  1. clearly define the goal
  2. estimate quantities of interest
  3. present those quantities in a compelling way

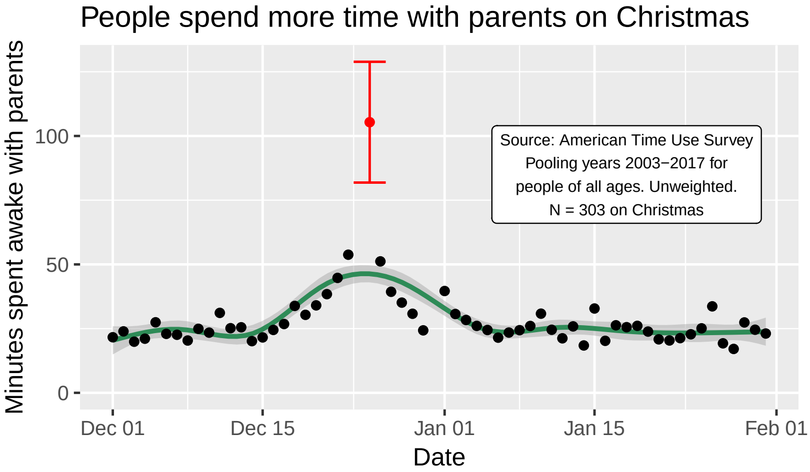
Visualization

- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking
- Three steps involved
  1. clearly define the goal
  2. estimate quantities of interest
  3. present those quantities in a compelling way
- Good design involves thinking carefully about the audience

Visualization

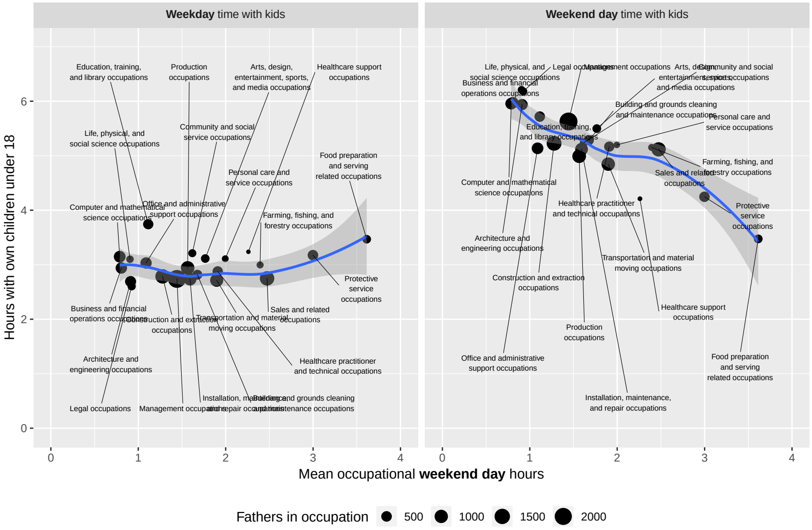
- Visualization is hard but ultimately extremely important
- It is absurd that we spend months collecting data, weeks analyzing it and five minutes slapping it into an unreadable table.
- Visualization can be used for many purposes
  - drawing people into a topic/dataset
  - presenting evidence
  - exploration/model checking
- Three steps involved
  1. clearly define the goal
  2. estimate quantities of interest
  3. present those quantities in a compelling way
- Good design involves thinking carefully about the audience
- I strongly recommend Kieran Healy's new visualization book — great summary of the fundamentals plus `R` code.

Examples



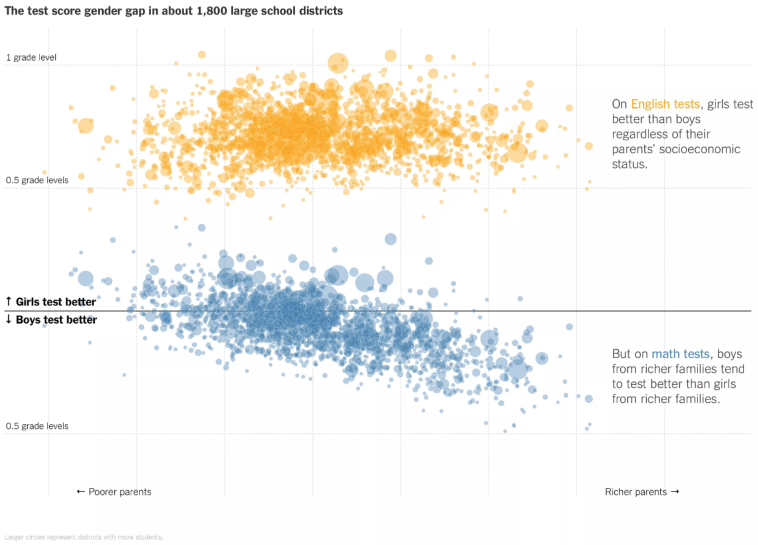
Source: Ian Lundberg

Examples



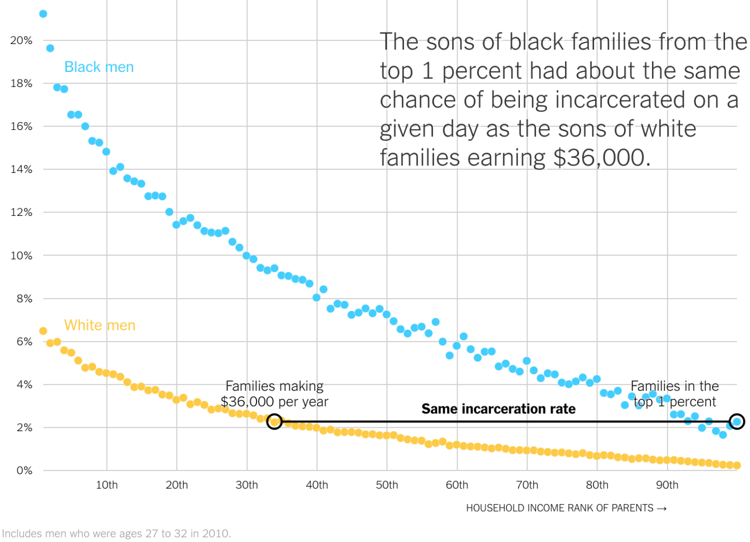
Source: Ian Lundberg

Examples



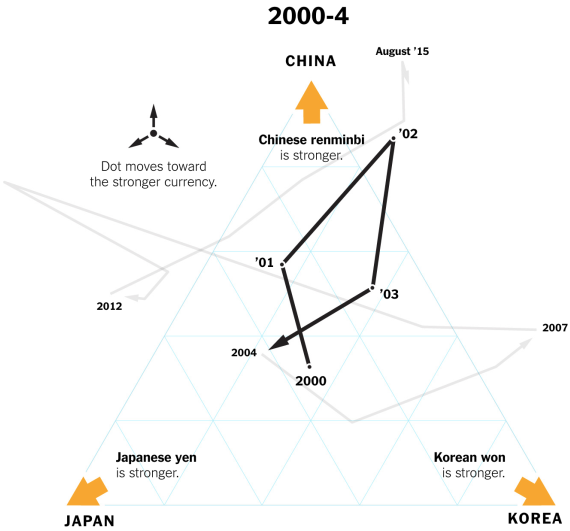
Source: New York Times

Examples



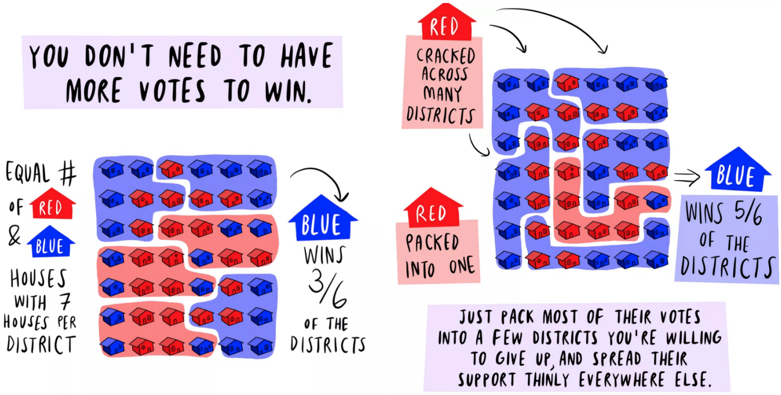
Source: New York Times

Examples



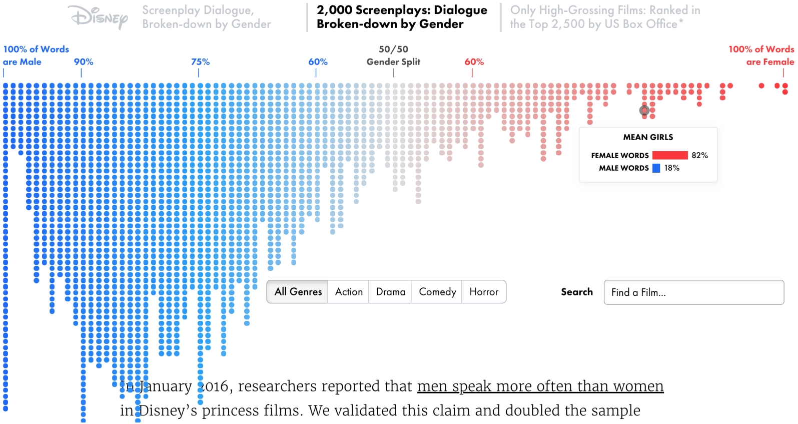
Source: New York Times

Examples



Source: Olivia Walch

Examples



Source: The Pudding

Examples

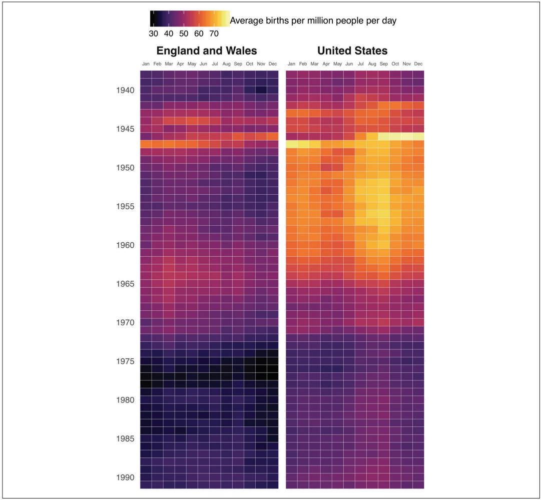
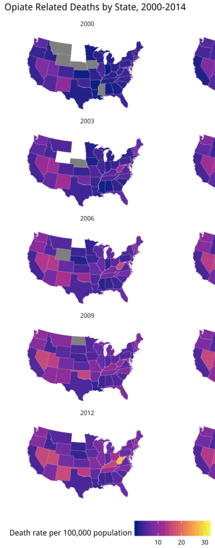


Figure 1. Average births per million people per day, 1938–1991. Each tile represents one month. The underlying count is number of births per month, standardized first by the total population for the period and then by the number of days in that month. Data for the United States are from the U.S. Census Bureau. Data for England and Wales are from the U.K. Office of National Statistics.

Source: Kieran Healy

Examples



Source: Kieran Healy

Reading

- Angrist and Pishke Chapter 2 (The Experimental Ideal) Chapter 3 (Regression and Causality)
- Morgan and Winship Chapters 3-4 (Causal Graphs and Conditioning Estimators)
- Hernan and Robins Chapter 3 Observational Studies
- Optional: Hernan (2018) "The C-word: Scientific euphemisms do not improve causal inference from observational data" *American Journal of Public Health*.
- Optional: Elwert and Winship (2014) "Endogenous selection bias: The problem of conditioning on a collider variable" *Annual Review of Sociology*
- Optional: Morgan and Winship Chapter 11 Repeated Observations and the Estimation of Causal Effects