FLS 6441 - Methods III: Explanation and Causation

Week 2 - A Framework for Explanation

Jonathan Phillips

March 2019

Section 1

Explanation

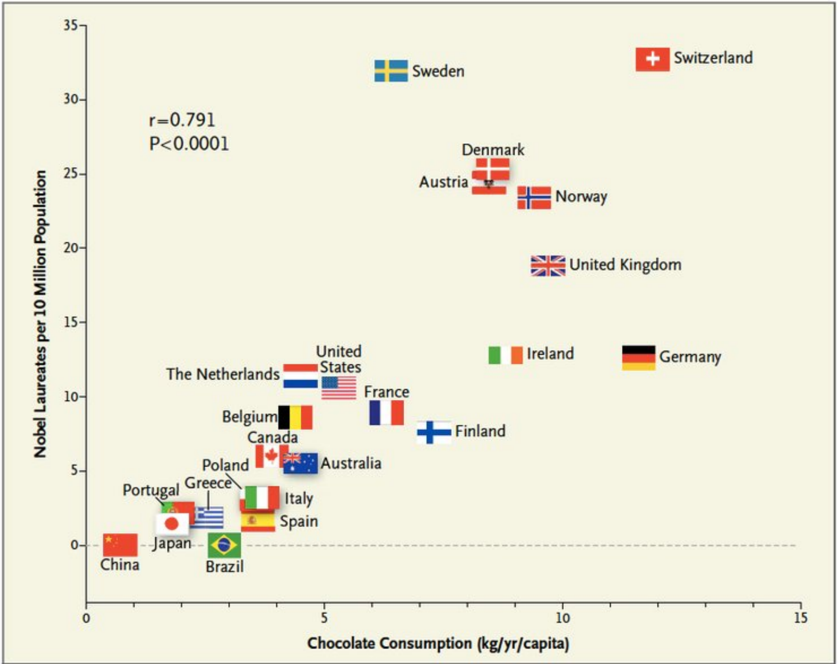


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel

Explanation

- Why isn't correlation enough?

Explanation

- Why isn't correlation enough?
  - For **prediction**, correlation is fine: If we know a country has chocolate consumption of 10kg/yr/capita we can reasonably predict it will have about 25 Nobel Laureates

Explanation

- Why isn't correlation enough?
  - For **prediction**, correlation is fine: If we know a country has chocolate consumption of 10kg/yr/capita we can reasonably predict it will have about 25 Nobel Laureates
  - But for **intervention**, correlation does not help: forcing people to eat more chocolate does nothing on its own to produce more Nobel Laureates

Explanation

- Why isn't correlation enough?
  - For **prediction**, correlation is fine: If we know a country has chocolate consumption of 10kg/yr/capita we can reasonably predict it will have about 25 Nobel Laureates
  - But for **intervention**, correlation does not help: forcing people to eat more chocolate does nothing on its own to produce more Nobel Laureates
  - For **explanation**, correlation also fails - it is no *explanation* to say that Switzerland has the most Nobel Laureates because it has the highest chocolate consumption

Explanation

- What does it mean to explain something?

Explanation

- What does it mean to explain something?
- To give an account of what happens, *and why*
  - The ’chain of causation’

Explanation

- What does it mean to explain something?
- To give an account of what happens, *and why*
  - The ’chain of causation’
- If D explains y, we are saying that the *absence* of D would have led to a different value of y

Explanation

- What does it mean to explain something?
- To give an account of what happens, *and why*
  - The ’chain of causation’
- If D explains y, we are saying that the *absence* of D would have led to a different value of y
- There exists a ’counterfactual’ possibility that did not happen

Explanation

Deterministic Explanation

Explanation

Deterministic Explanation

- **Sufficient conditions:** Every time D happens, Y happens

Explanation

Deterministic Explanation

- **Sufficient conditions:** Every time D happens, Y happens
- **Necessary conditions:** Y does not happen if D does not happen ('*but for*')

Explanation

Deterministic Explanation

- **Sufficient conditions:** Every time D happens, Y happens
- **Necessary conditions:** Y does not happen if D does not happen (’*but for*’)

Probabilistic Explanation

Explanation

Deterministic Explanation

- **Sufficient conditions:** Every time D happens, Y happens
- **Necessary conditions:** Y does not happen if D does not happen ('*but for*')

Probabilistic Explanation

- If D happens, the **probability** of Y increases

Explanation

Deterministic Explanation

- **Sufficient conditions:** Every time D happens, Y happens
- **Necessary conditions:** Y does not happen if D does not happen ('*but for*')

Probabilistic Explanation

- If D happens, the **probability** of Y increases
- Treatment effects are a distribution, not a single value

Explanation

- Two perspectives on explanation:

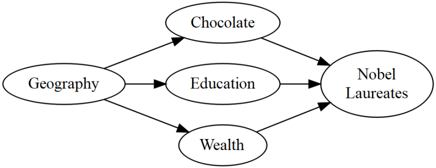
Explanation

- Two perspectives on explanation:

Causes of Effects	**Effects of Causes**
What caused Y?	Does D cause Y?
Why does Switzerland have so many Nobel laureates?	Does chocolate cause more Nobel laureates?
Backward-looking	Forward-looking

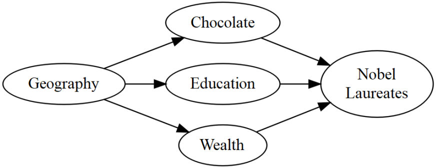
Explanation

- Two perspectives on explanation:



Explanation

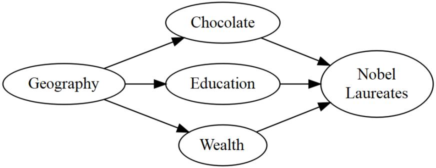
- Two perspectives on explanation:



- Identifying the source of **ALL** of the variation in Nobel Laureates

Explanation

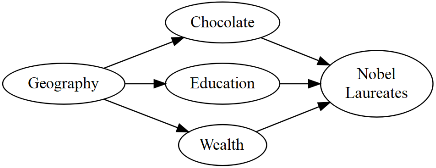
- Two perspectives on explanation:

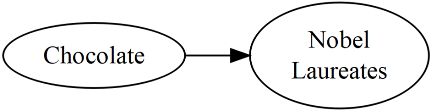


- Identifying the source of **ALL** of the variation in Nobel Laureates
- An infinite task!

Explanation

- Two perspectives on explanation:

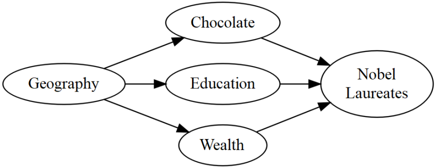


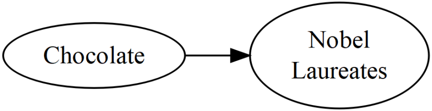


- Identifying the source of **ALL** of the variation in Nobel Laureates
- An infinite task!

Explanation

- Two perspectives on explanation:



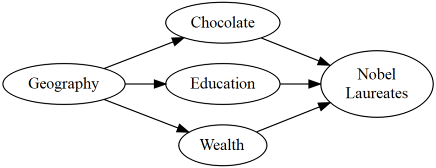


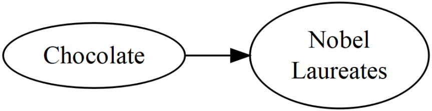
- Identifying the source of **ALL** of the variation in Nobel Laureates
- An infinite task!

- Identifying how much **ONE** variable causes variation in Nobel Laureates

Explanation

- Two perspectives on explanation:



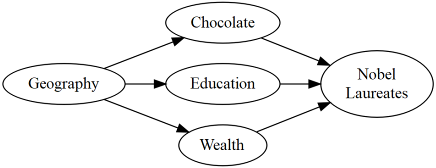


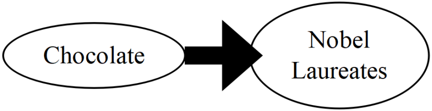
- Identifying the source of **ALL** of the variation in Nobel Laureates
- An infinite task!

- Identifying how much **ONE** variable causes variation in Nobel Laureates
- This we can do!

Explanation

- Two perspectives on explanation:



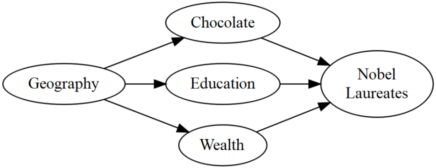


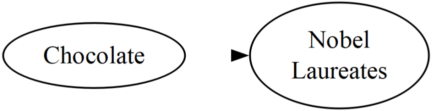
- Identifying the source of **ALL** of the variation in Nobel Laureates
- An infinite task!

- Identifying how much **ONE** variable causes variation in Nobel Laureates
- This we can do!

Explanation

- Two perspectives on explanation:





- Identifying the source of **ALL** of the variation in Nobel Laureates
- An infinite task!

- Identifying how much **ONE** variable causes variation in Nobel Laureates
- This we can do!

Explanation

- A focus on a single explanatory variable D requires a clear definition of **'Treatment'**

Explanation

- A focus on a single explanatory variable D requires a clear definition of **’Treatment’**
- AND to clearly define a **’Control’**

Explanation

- A focus on a single explanatory variable D requires a clear definition of **’Treatment’**
- AND to clearly define a **’Control’**
  - What is the opposite of investing $1bn in education?

Explanation

- A focus on a single explanatory variable *D* requires a clear definition of **'Treatment'**
- AND to clearly define a **'Control'**
  - What is the opposite of investing $1bn in education?
  - No investment, or investing it elsewhere?

Explanation

- A focus on a single explanatory variable D requires a clear definition of **'Treatment'**
- AND to clearly define a **'Control'**
  - What is the opposite of investing \$1bn in education?
  - No investment, or investing it elsewhere?
- Define treatment:

$$D_i = \begin{cases} 1, \text{ if treated} \\ 0, \text{ if not treated} \end{cases}$$

Explanation

- Defining our outcome variable:

Explanation

- Defining our outcome variable:
  - Is it the outcome we really care about? Or just what's easy to measure?

Explanation

- Defining our outcome variable:
  - Is it the outcome we really care about? Or just what's easy to measure?
  - What theory are we testing?

Explanation

- Defining our outcome variable:
  - Is it the outcome we really care about? Or just what's easy to measure?
  - What theory are we testing?
- Tempting to look at many outcomes, but the risk of 'cherry-picking'

Explanation

- Defining our outcome variable:
  - Is it the outcome we really care about? Or just what's easy to measure?
  - What theory are we testing?
- Tempting to look at many outcomes, but the risk of 'cherry-picking'
  - All outcomes are **probabilistic** (due to all the other factors we haven't accounted for)

Explanation

- Defining our outcome variable:
  - Is it the outcome we really care about? Or just what's easy to measure?
  - What theory are we testing?
- Tempting to look at many outcomes, but the risk of 'cherry-picking'
  - All outcomes are **probabilistic** (due to all the other factors we haven't accounted for)
  - If we study 20 outcomes, on average one will show a significant effect even with no real causal effect

Explanation

- Defining our outcome variable:
  - Is it the outcome we really care about? Or just what's easy to measure?
  - What theory are we testing?
- Tempting to look at many outcomes, but the risk of 'cherry-picking'
  - All outcomes are **probabilistic** (due to all the other factors we haven't accounted for)
  - If we study 20 outcomes, on average one will show a significant effect even with no real causal effect
- So we usually want to study a **single outcome**

Explanation

- What are the **units** of our analysis?

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?
- eg. How does the electoral system affect attitudes to redistribution?

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?
- eg. How does the electoral system affect attitudes to redistribution?
  - Treatment at the national level

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?
- eg. How does the electoral system affect attitudes to redistribution?
  - Treatment at the national level
  - Outcome varies at the individual level

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?
- eg. How does the electoral system affect attitudes to redistribution?
  - Treatment at the national level
  - Outcome varies at the individual level
  - Measurement needed at the lowest (individual) level

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?
- eg. How does the electoral system affect attitudes to redistribution?
  - Treatment at the national level
  - Outcome varies at the individual level
  - Measurement needed at the lowest (individual) level
  - But our analysis needs to take account of the ’clustered’ treatment

Explanation

- What are the **units** of our analysis?
  - Countries? Political Parties? Individuals?
  - At what level does causality operate?
- eg. How does the electoral system affect attitudes to redistribution?
  - Treatment at the national level
  - Outcome varies at the individual level
  - Measurement needed at the lowest (individual) level
  - But our analysis needs to take account of the ’clustered’ treatment
- Units are **time-specific**: the same person 10 minutes later is a different unit

Section 2

Causal Inference

Causal Inference

- The **causal effect** of treatment is how each unit's outcome differs when it is treated and not treated

Causal Inference

- The **causal effect** of treatment is how each unit's outcome differs when it is treated and not treated
- This means comparing the **Potential Outcomes** for unit i:

$$Y_{Di} = \begin{cases} Y_{1i} \text{ Potential Outcome if unit i treated} \\ Y_{0i} \text{ Potential Outcome if unit i NOT treated} \end{cases}$$

Causal Inference

- The **causal effect** of treatment is how each unit's outcome differs when it is treated and not treated
- This means comparing the **Potential Outcomes** for unit i:

$$Y_{Di} = \begin{cases} Y_{1i} \text{ Potential Outcome if unit i treated} \\ Y_{0i} \text{ Potential Outcome if unit i NOT treated} \end{cases}$$

- Individual Treatment Effect for unit i: $\alpha_i = Y_{1i} - Y_{0i}$

Causal Inference

- The **causal effect** of treatment is how each unit's outcome differs when it is treated and not treated
- This means comparing the **Potential Outcomes** for unit i:

$$Y_{Di} = \begin{cases} Y_{1i} \text{ GDP Growth of Brazil in 2010 if a Democracy} \\ Y_{0i} \text{ GDP Growth of Brazil in 2010 if NOT a Democracy} \end{cases}$$

- Individual Treatment Effect for unit i: $\alpha_i = Y_{1i} - Y_{0i}$

Causal Inference

- We are explicitly thinking about **counterfactuals**

Causal Inference

- We are explicitly thinking about **counterfactuals**
  - What would have happened to the same unit if the treatment had not happened?

Causal Inference

- We are explicitly thinking about **counterfactuals**
  - What would have happened to the same unit if the treatment had not happened?
  - What would Brazil's GDP growth rate be if we lived in a dictatorship?

Causal Inference

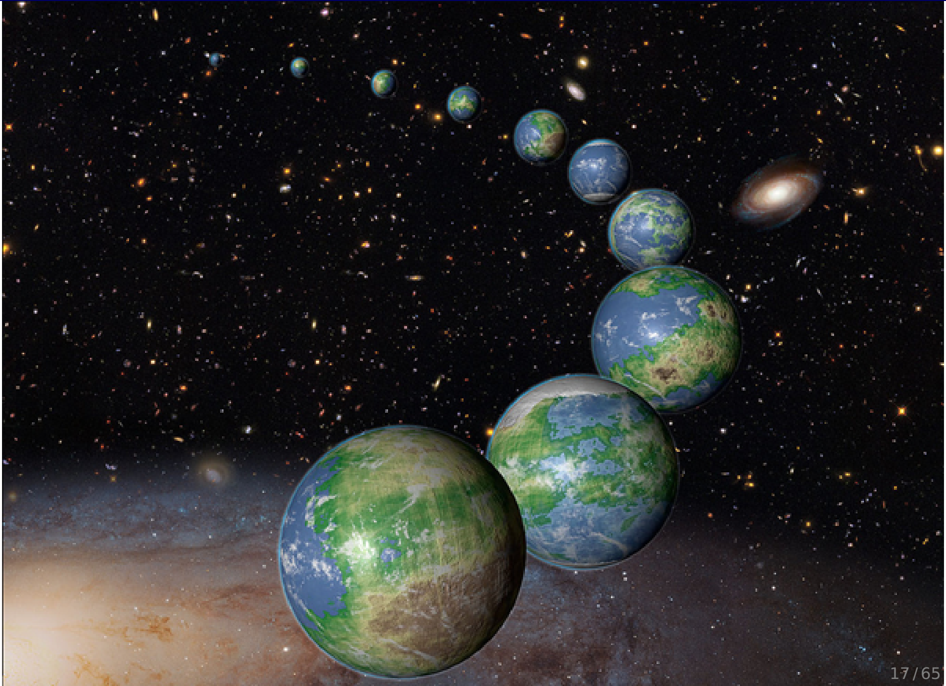
- We are explicitly thinking about **counterfactuals**
  - What would have happened to the same unit if the treatment had not happened?
  - What would Brazil's GDP growth rate be if we lived in a dictatorship?
  - Would World War I still have happened if Archduke Franz Ferdinand had not been assassinated in 1914?

Causal Inference

- We are explicitly thinking about **counterfactuals**
  - What would have happened to the same unit if the treatment had not happened?
  - What would Brazil's GDP growth rate be if we lived in a dictatorship?
  - Would World War I still have happened if Archduke Franz Ferdinand had not been assassinated in 1914?
  - Would Brazil have won the 2014 World Cup if Neymar had not been injured?

Causal Inference

- We are explicitly thinking about **counterfactuals**
  - What would have happened to the same unit if the treatment had not happened?
  - What would Brazil's GDP growth rate be if we lived in a dictatorship?
  - Would World War I still have happened if Archduke Franz Ferdinand had not been assassinated in 1914?
  - Would Brazil have won the 2014 World Cup if Neymar had not been injured?



Causal Inference

Potential Outcomes are just another Variable for each Unit

	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	Y_1	Y_0	$Y_1 - Y_0$
Brasil	4	1	3
Argentina	7	4	3
Bolivia	2	4	-2
Colombia	7	7	0
Peru	5	4	1

Causal Inference

- Political Science is not about explaining individual events

Causal Inference

- Political Science is not about explaining individual events
- We ideally want general theories that apply to *all our units*

Causal Inference

- Political Science is not about explaining individual events
- We ideally want general theories that apply to *all our units*
- To explain a systematic treatment - not a single event - we need **multiple counterfactual comparisons**

Causal Inference

- Political Science is not about explaining individual events
- We ideally want general theories that apply to *all our units*
- To explain a systematic treatment - not a single event - we need **multiple counterfactual comparisons**
- We know how democracy works in Europe; the question is what will happen if it becomes more common in the whole world?

Causal Inference

- Political Science is not about explaining individual events
- We ideally want general theories that apply to *all our units*
- To explain a systematic treatment - not a single event - we need **multiple counterfactual comparisons**
- We know how democracy works in Europe; the question is what will happen if it becomes more common in the whole world?

Average Treatment Effect

We want to calculate an **Average Treatment Effect**

Causal Inference

- Political Science is not about explaining individual events
- We ideally want general theories that apply to *all our units*
- To explain a systematic treatment - not a single event - we need **multiple counterfactual comparisons**
- We know how democracy works in Europe; the question is what will happen if it becomes more common in the whole world?

Average Treatment Effect

We want to calculate an **Average Treatment Effect**

$$ATE = E(\alpha_i) = E(Y_1 - Y_0) = E(Y_1) - E(Y_0) = \frac{\sum_i (Y_{1i} - Y_{0i})}{N}$$

Causal Inference

Potential Outcomes are just another Variable for each Unit

	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	Y_1	Y_0	$Y_1 - Y_0$
Brasil	4	1	3
Argentina	7	4	3
Bolivia	2	4	-2
Colombia	7	7	0
Peru	5	4	1
Average Treatment Effect	**5**	**4**	**1**

Causal Inference

- In reality, some units are **actually treated** ($D = 1$), others are **actually control** ($D = 0$)

Causal Inference

- In reality, some units are **actually treated** ($D = 1$), others are **actually control** ($D = 0$)

Average Treatment Effect on the Treated

$$\text{ATT}=\text{E}(\alpha_i|D = 1) = E(Y_1 - Y_0|D = 1) = \frac{\sum_i (Y_{1i} - Y_{0i}|D=1)}{N_{Treated}} \quad (1)$$

Causal Inference

- In reality, some units are **actually treated** ($D = 1$), others are **actually control** ($D = 0$)

Average Treatment Effect on the Treated

$$\text{ATT}=\text{E}(\alpha_i|D = 1) = E(Y_1 - Y_0|D = 1) = \frac{\sum_i (Y_{1i} - Y_{0i}|D=1)}{N_{Treated}} \quad (1)$$

Average Treatment Effect on the Untreated (Control)

$$\text{ATU}=\text{E}(\alpha_i|D = 0) = E(Y_1 - Y_0|D = 0) = \frac{\sum_i (Y_{1i} - Y_{0i}|D=0)}{N_{Control}} \quad (2)$$

Causal Inference

- In reality, some units are **actually treated** ($D = 1$), others are **actually control** ($D = 0$)

Average Treatment Effect on the Treated

$$\text{ATT}=\text{E}(\alpha_i|D = 1) = E(Y_1 - Y_0|D = 1) = \frac{\sum_i (Y_{1i} - Y_{0i}|D=1)}{N_{Treated}} \quad (1)$$

Average Treatment Effect on the Untreated (Control)

$$\text{ATU}=\text{E}(\alpha_i|D = 0) = E(Y_1 - Y_0|D = 0) = \frac{\sum_i (Y_{1i} - Y_{0i}|D=0)}{N_{Control}} \quad (2)$$

- The three effect estimates are usually different
  - The effect democracy has had in Europe is different to the effect if it were introduced in Africa

Causal Inference

Potential Outcomes Example

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Brasil	1	4	1	3
Argentina	0	7	4	3
Bolivia	1	2	4	-2
Colombia	0	7	7	0
Peru	0	5	4	1

Causal Inference

Potential Outcomes Example

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Brasil	1	4	1	3
Bolivia	1	2	4	-2
ATT	**1**	**3**	**2.5**	**0.5**

Causal Inference

Potential Outcomes Example

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Argentina	0	7	4	3
Colombia	0	7	7	0
Peru	0	5	4	1
ATU	**0**	**6.3**	**5**	**1.3**

Causal Inference

The Fundamental Problem of Causal Inference

- No units can receive **both** treatment and control

Causal Inference

The Fundamental Problem of Causal Inference

- No units can receive **both** treatment and control
- So we can never observe both Y_1 and Y_0 **for the same unit**

Causal Inference

The Fundamental Problem of Causal Inference

- No units can receive **both** treatment and control
- So we can never observe both Y_1 and Y_0 **for the same unit**
- *Individual* Treatment Effects are **Impossible to Estimate**

Causal Inference

The Fundamental Problem of Causal Inference

- No units can receive **both** treatment and control
- So we can never observe both Y_1 and Y_0 **for the same unit**
- *Individual* Treatment Effects are **Impossible to Estimate**

$$Y_i^{obs} = \begin{cases} Y_{1i} \text{ if } D_i = 1 \\ Y_{0i} \text{ if } D_i = 0 \end{cases}$$

Causal Inference

The Fundamental Problem of Causal Inference

- No units can receive **both** treatment and control
- So we can never observe both Y_1 and Y_0 **for the same unit**
- *Individual* Treatment Effects are **Impossible to Estimate**

$$Y_i^{obs} = \begin{cases} Y_{1i} \text{ if } D_i = 1 \\ Y_{0i} \text{ if } D_i = 0 \end{cases}$$

$$Y_i^{obs} = D_i \cdot Y_{1i} + (1 - D_i) \cdot Y_{0i}$$

Causal Inference

Potential Outcomes Example

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Brasil	1	4	1	3
Argentina	0	7	4	3
Bolivia	1	2	4	-2
Colombia	0	7	7	0
Peru	0	5	4	1

Causal Inference

Potential Outcomes Example

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	Treatment Effect
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Brasil	1	4	?	?
Argentina	0	?	4	?
Bolivia	1	2	?	?
Colombia	0	?	7	?
Peru	0	?	4	?

Causal Inference

Potential Outcomes Example

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	**Observed** GDP Growth
	D_i	Y_1	Y_0	Y^{obs}
Brasil	1	4	?	4
Argentina	0	?	4	4
Bolivia	1	2	?	2
Colombia	0	?	7	7
Peru	0	?	4	4

Causal Inference

Potential Outcomes Example

	Democracy?	**Observed** GDP Growth
	D_i	Y^{obs}
Brasil	1	4
Argentina	0	4
Bolivia	1	2
Colombia	0	7
Peru	0	4

Causal Inference

- Actually, nothing stops us calculating the **Average Treatment Effect**

Causal Inference

- Actually, nothing stops us calculating the **Average Treatment Effect**
- The question is, is the ATE accurate?

Causal Inference

- Actually, nothing stops us calculating the **Average Treatment Effect**
- The question is, is the ATE accurate?

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	**Treatment Effect**
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Brasil	1	4	1	3
Argentina	0	7	4	3
Bolivia	1	2	4	-2
Colombia	0	7	7	0
Peru	0	5	4	1
Average Treatment Effect		**5**	**4**	**1**

Causal Inference

- Actually, nothing stops us calculating the **Average Treatment Effect**
- The question is, is the ATE accurate?

	Democracy?	GDP Growth if Democracy	GDP Growth if NOT Democracy	**Treatment Effect**
	D_i	Y_1	Y_0	$Y_1 - Y_0$
Brasil	1	4	?	?
Argentina	0	?	4	?
Bolivia	1	2	?	?
Colombia	0	?	7	?
Peru	0	?	4	?
Average Treatment Effect		**3**	**5**	**-2**

Causal Inference

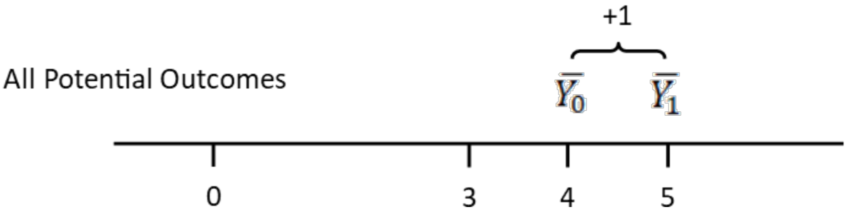
- **So what went wrong?**

Causal Inference

- **So what went wrong?**
- The potential outcomes we **observe** are a **biased representation** of the potential outcomes of all the units

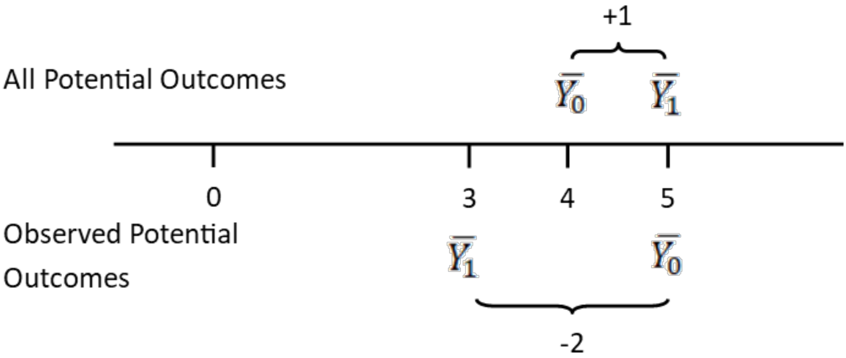
Causal Inference

- **So what went wrong?**
- The potential outcomes we **observe** are a **biased representation** of the potential outcomes of all the units



Causal Inference

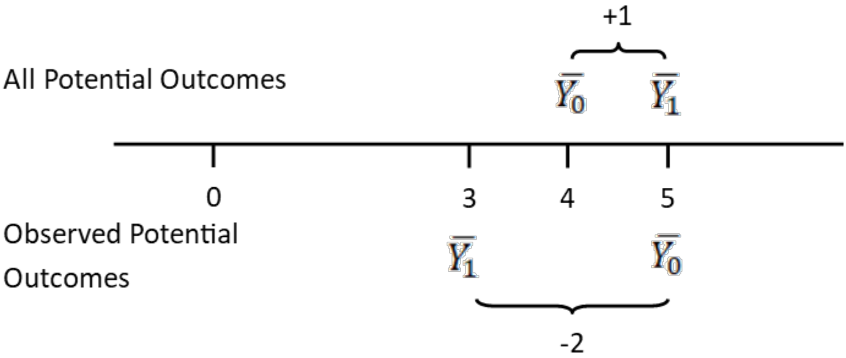
- **So what went wrong?**
- The potential outcomes we **observe** are a **biased representation** of the potential outcomes of all the units



- $E(Y_1)$ values are **biased lower** in the observed data

Causal Inference

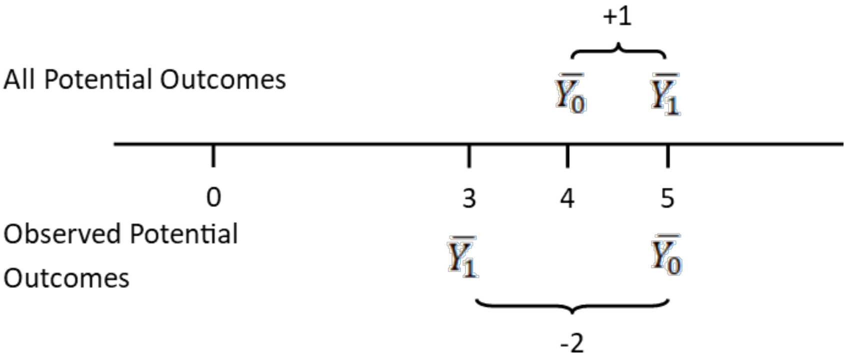
- **So what went wrong?**
- The potential outcomes we **observe** are a **biased representation** of the potential outcomes of all the units



- $E(Y_1)$ values are **biased lower** in the observed data
- $E(Y_0)$ values are **biased higher** in the observed data

Causal Inference

- **So what went wrong?**
- The potential outcomes we **observe** are a **biased representation** of the potential outcomes of all the units



- $E(Y_1)$ values are **biased lower** in the observed data
- $E(Y_0)$ values are **biased higher** in the observed data
- So $E(Y_1) - E(Y_0)$ is **biased**

Causal Inference

- The Fundamental Problem of Causal Inference means we can only discover causal relationships by comparing **across units**

Causal Inference

- The Fundamental Problem of Causal Inference means we can only discover causal relationships by comparing **across units**
- Comparing treated i and control j units

Causal Inference

- The Fundamental Problem of Causal Inference means we can only discover causal relationships by comparing **across units**
- Comparing treated i and control j units
- If potential outcomes are biased in our observed data:

Causal Inference

- The Fundamental Problem of Causal Inference means we can only discover causal relationships by comparing **across units**
- Comparing treated i and control j units
- If potential outcomes are biased in our observed data:
  - Our **counterfactual case** j does not represent what would have happened to i in the absence of treatment

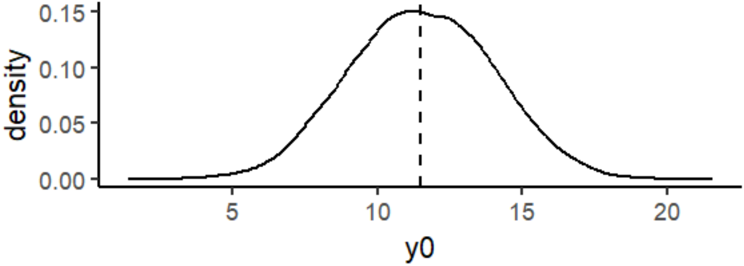
Causal Inference

- The Fundamental Problem of Causal Inference means we can only discover causal relationships by comparing **across units**
- Comparing treated i and control j units
- If potential outcomes are biased in our observed data:
  - Our **counterfactual case** j does not represent what would have happened to i in the absence of treatment
  - Counterfactuals are not **plausible**

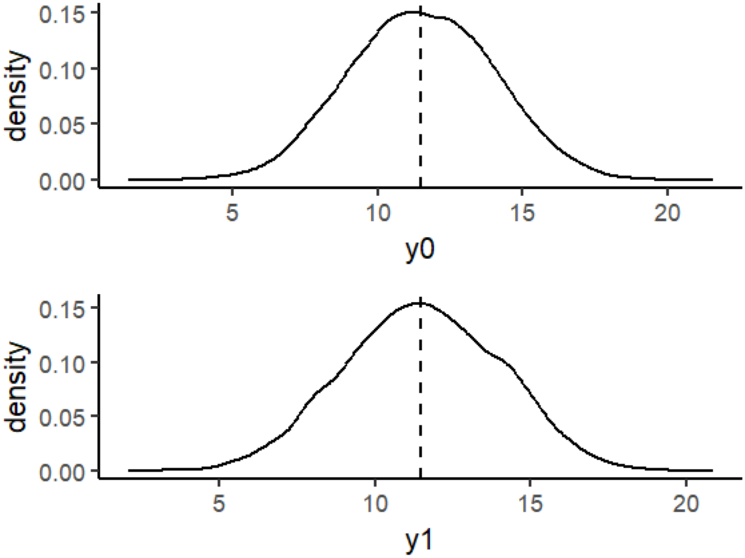
Causal Inference

- The Fundamental Problem of Causal Inference means we can only discover causal relationships by comparing **across units**
- Comparing treated *i* and control *j* units
- If potential outcomes are biased in our observed data:
  - Our **counterfactual case** *j* does not represent what would have happened to *i* in the absence of treatment
  - Counterfactuals are not **plausible**
  - Causal effects are **biased**

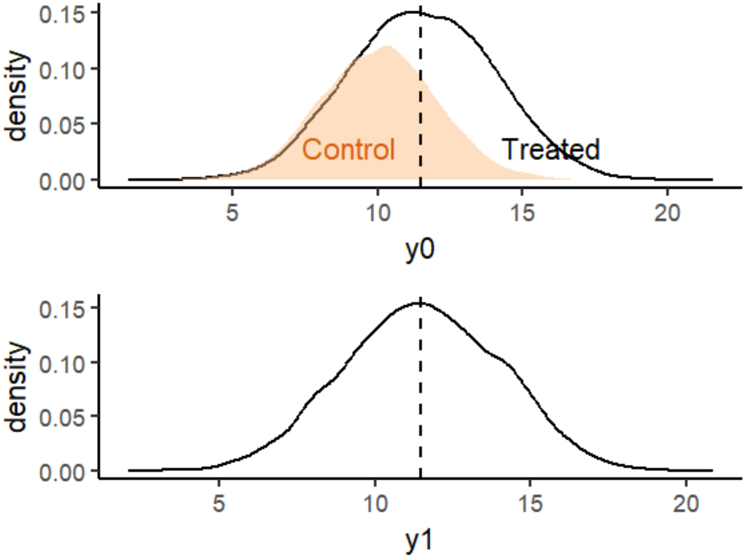
Causal Inference



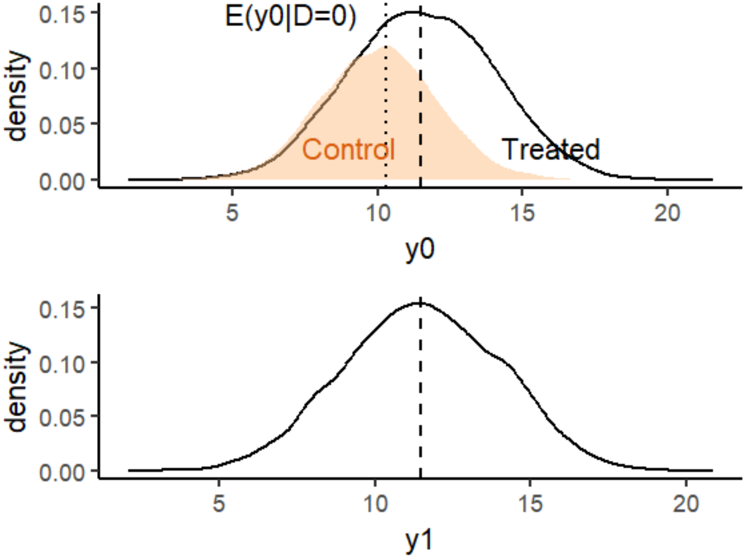
Causal Inference



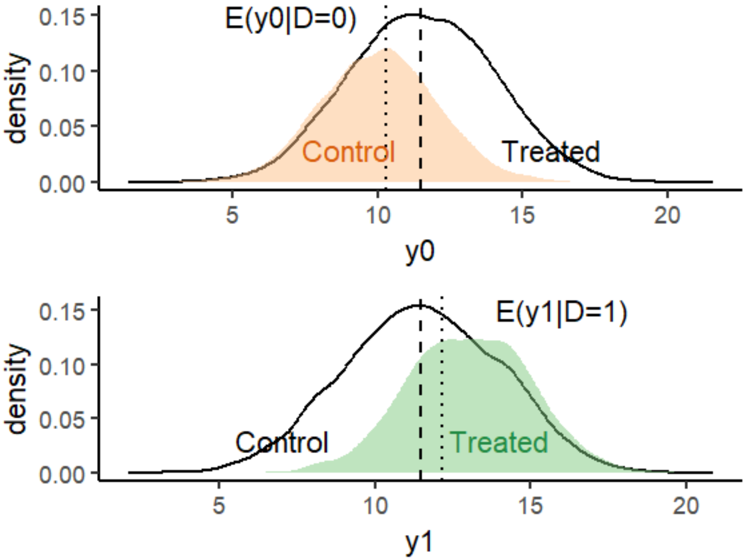
Causal Inference



Causal Inference



Causal Inference



Causal Inference

- Lots of averages:

		Hypothetical outcome	
		$Y0$	$Y1$
Actual Treatment	$D=0$	$E(Y_{0i}\|D=0)$	$E(Y_{1i}\|D=0)$
	$D=1$	$E(Y_{0i}\|D=1)$	$E(Y_{1i}\|D=1)$

Causal Inference

- Lots of averages:

		Hypothetical outcome	
		$Y0$	$Y1$
Actual Treatment	$D = 0$	$E(Y_{0i} \mid D = 0)$	$E(Y_{1i} \mid D = 0)$
	$D = 1$	$E(Y_{0i} \mid D = 1)$	$E(Y_{1i} \mid D = 1)$

Causal Inference

- All our causal estimates are **averages**

Causal Inference

- All our causal estimates are **averages**
  - We cannot distinguish the null hypothesis of no average effect from the sharp null hypothesis of no individual effects

Causal Inference

- All our causal estimates are **averages**
  - We cannot distinguish the null hypothesis of no average effect from the sharp null hypothesis of no individual effects

	No Average Effect $E(Y_1 - Y_0) = 0$	"Sharp null": No individual effects $(Y_{1i} - Y_{0i} = 0)$
Brasil	2	0
Argentina	-1	0
Bolivia	1	0
Colombia	0	0
Peru	-2	0
Average	**0**	**0**

Section 3

Why Observational Data is Biased

Bias

- Why are potential outcomes biased in our data?

Bias

- Why are potential outcomes biased in our data?
  1. Omitted Variables

Bias

- Why are potential outcomes biased in our data?
    1. Omitted Variables
    2. Reverse Causation

Bias

- Why are potential outcomes biased in our data?
  1. Omitted Variables
  2. Reverse Causation
  3. Selection Bias

Bias

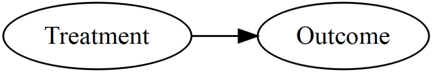
- Why are potential outcomes biased in our data?
  1. Omitted Variables
  2. Reverse Causation
  3. Selection Bias
- In all of these cases **the potential outcomes are distorted**

Bias

- Why are potential outcomes biased in our data?
  1. Omitted Variables
  2. Reverse Causation
  3. Selection Bias
- In all of these cases **the potential outcomes are distorted**
- So basic regression is **biased**

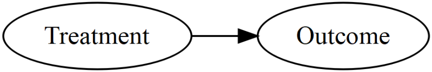
Omitted Variable Bias

A real causal relationship:

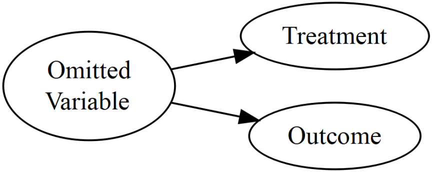


Omitted Variable Bias

A real causal relationship:

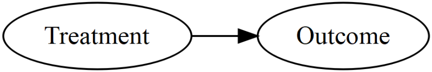


Being misled by omitted variable bias:

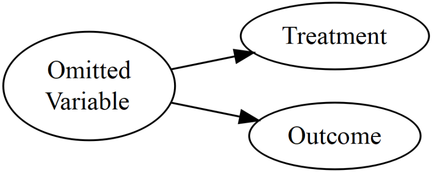


Omitted Variable Bias

A real causal relationship:



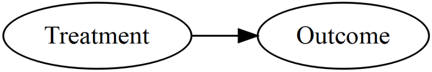
Being misled by omitted variable bias:



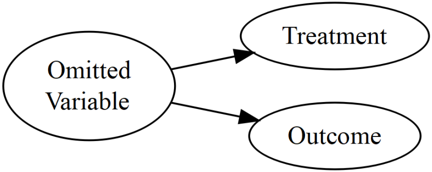
- A third variable causes some units to have **different values of potential outcomes**, AND for those **same units to be treated**

Omitted Variable Bias

A real causal relationship:



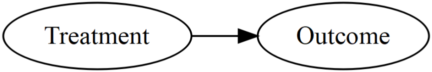
Being misled by omitted variable bias:



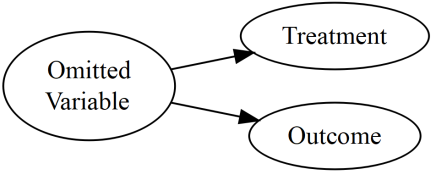
- A third variable causes some units to have **different values of potential outcomes**, AND for those **same units to be treated**
- So treated units have non-representative Y_1

Omitted Variable Bias

A real causal relationship:



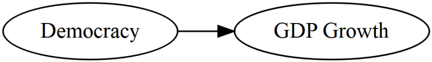
Being misled by omitted variable bias:



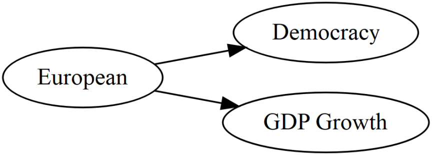
- A third variable causes some units to have **different values of potential outcomes**, AND for those **same units to be treated**
- So treated units have non-representative Y_1
- And control units have non-representative Y_0

Omitted Variable Bias

A real causal relationship:

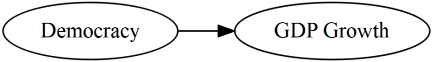


Being misled by omitted variable bias:

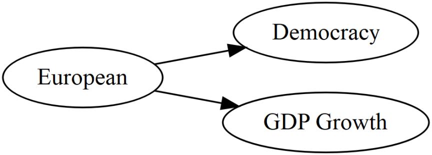


Omitted Variable Bias

A real causal relationship:

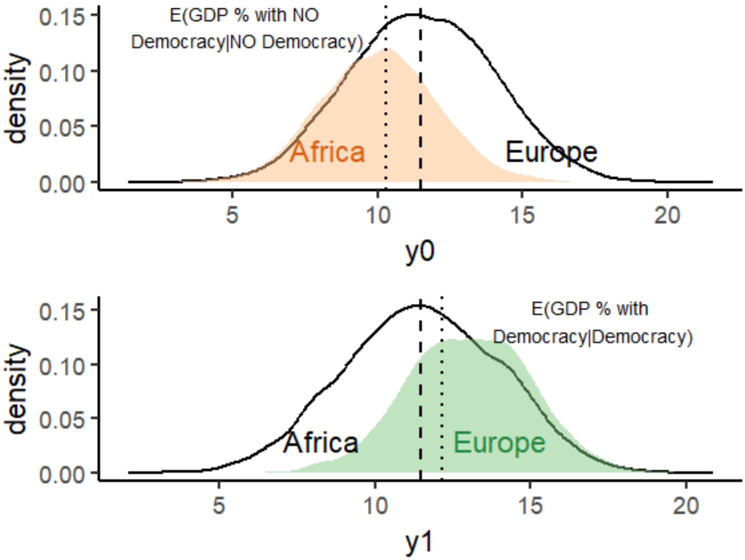


Being misled by omitted variable bias:



- European countries faced conditions that encouraged both democracy and rapid GDP growth

Omitted Variable Bias



Omitted Variable Bias

- Let's say that $Y_{1i} = Y_{0i} + \alpha$, where α is the real constant treatment effect

$$A\hat{T}E = E(Y_1|D=1) - E(Y_0|D=0)$$

Omitted Variable Bias

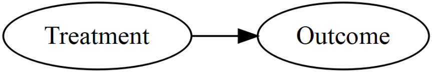
- Let's say that $Y_{1i} = Y_{0i} + \alpha$, where α is the real constant treatment effect

$$A\hat{T}E = E(Y_1|D=1) - E(Y_0|D=0)$$

$$A\hat{T}E = \underbrace{\alpha}_{\text{Real ATE}} + \underbrace{E(Y_0|D=1) - E(Y_0|D=0)}_{\text{Bias}}$$

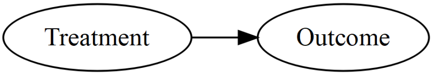
Reverse Causation

A real causal relationship:

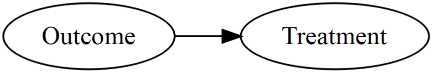


Reverse Causation

A real causal relationship:

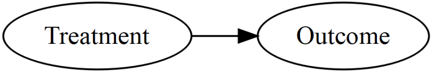


Being misled by reverse causation:

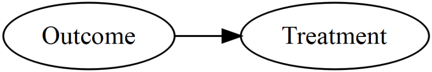


Reverse Causation

A real causal relationship:



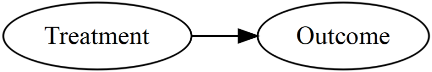
Being misled by reverse causation:



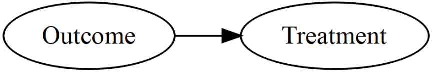
- D does not affect Y, but higher Y makes treatment (D) more likely

Reverse Causation

A real causal relationship:



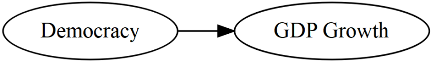
Being misled by reverse causation:



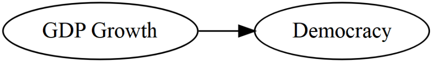
- D does not affect Y, but higher Y makes treatment (D) more likely
- So the two variables are **correlated**

Reverse Causation

A real causal relationship:

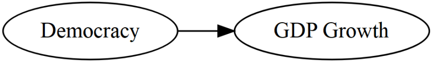


Being misled by reverse causation:

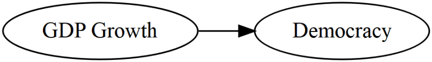


Reverse Causation

A real causal relationship:



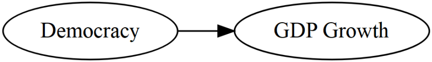
Being misled by reverse causation:



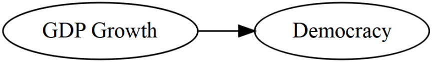
- GDP Growth encourages democratization

Reverse Causation

A real causal relationship:

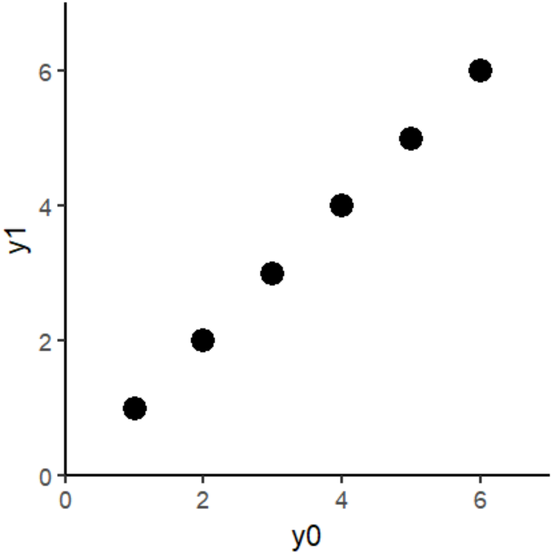


Being misled by reverse causation:



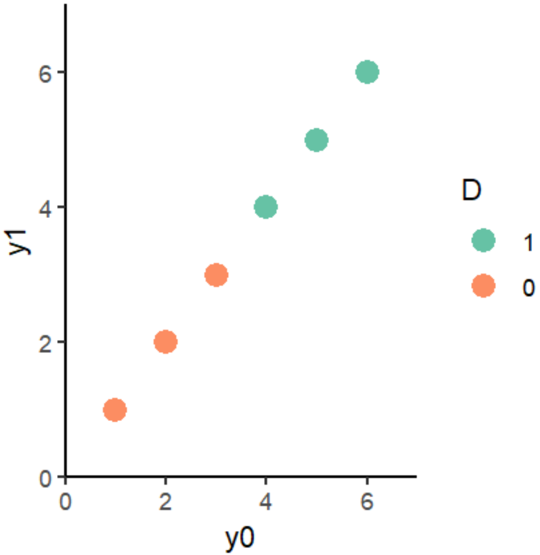
- GDP Growth encourages democratization
- So democracies are more likely to have experienced high growth rates

Reverse Causation

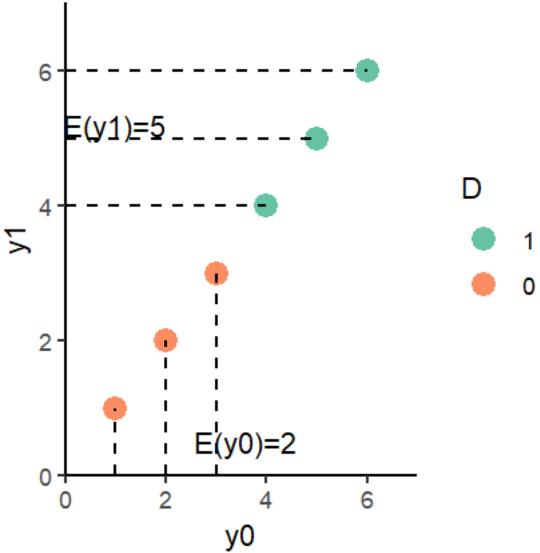


- $E(Y_1 - Y_0) = 0$

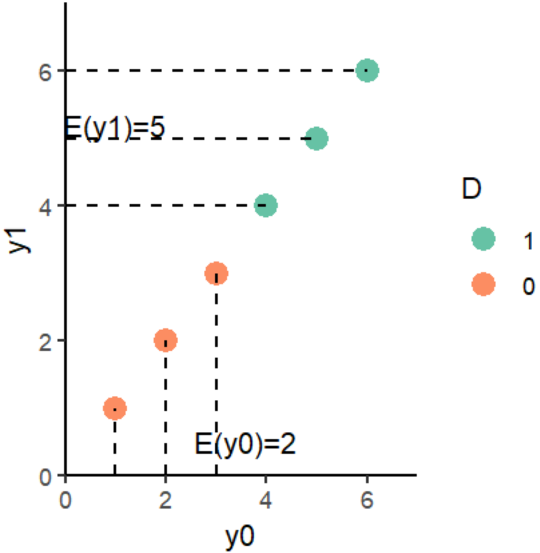
Reverse Causation



Reverse Causation



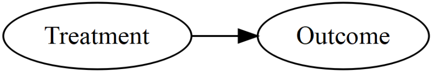
Reverse Causation



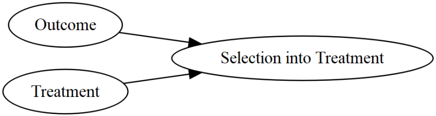
- $E(Y_1|D=1) - E(Y_0|D=0) = 5 - 2 = 3$

Selection Bias

A real causal relationship:

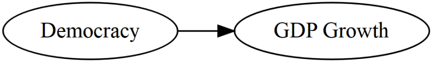


Being misled by Selection Bias:

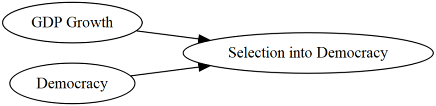


Selection Bias

A real causal relationship:

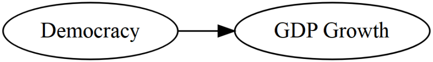


Being misled by Selection Bias:

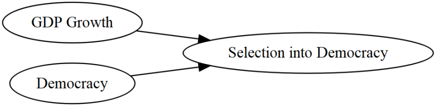


Selection Bias

A real causal relationship:



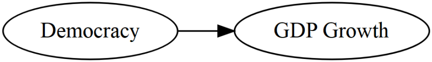
Being misled by Selection Bias:



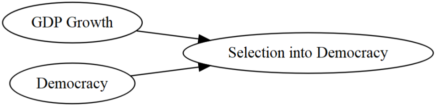
- The units which benefit most from treatment (largest $y_1 - y_0$) **choose treatment**

Selection Bias

A real causal relationship:



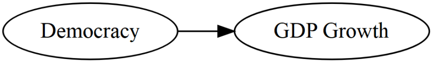
Being misled by Selection Bias:



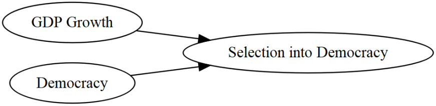
- The units which benefit most from treatment (largest $y_1 - y_0$) **choose treatment**
- We don't see any of the low y_1's of units which avoid treatment

Selection Bias

A real causal relationship:



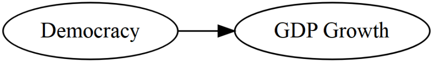
Being misled by Selection Bias:



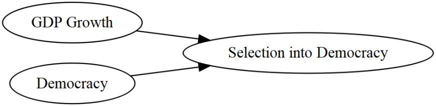
- The units which benefit most from treatment (largest $y_1 - y_0$) **choose treatment**
- We don't see any of the low y_1's of units which avoid treatment
  - Countries which can boost their GDP growth by becoming a democracy choose to democratize

Selection Bias

A real causal relationship:

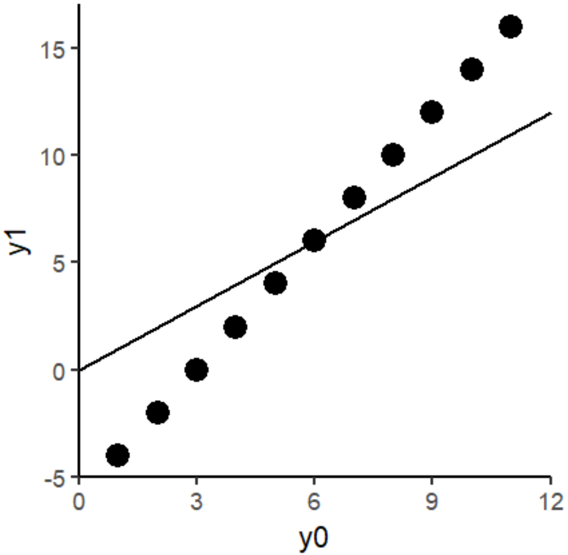


Being misled by Selection Bias:

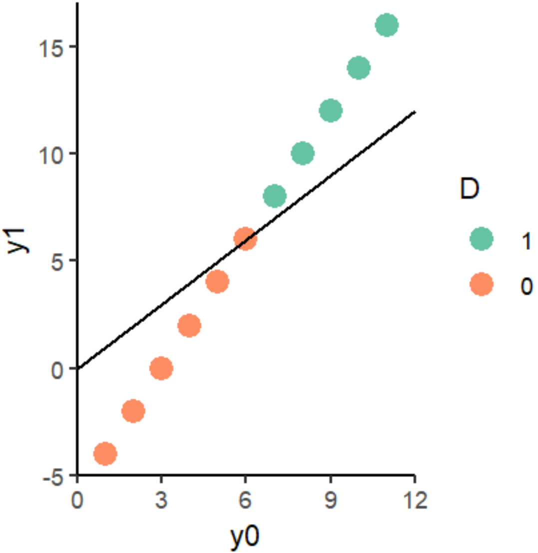


- The units which benefit most from treatment (largest $y_1 - y_0$) **choose treatment**
- We don't see any of the low y_1's of units which avoid treatment
  - Countries which can boost their GDP growth by becoming a democracy choose to democratize
  - Ex. Mexico? Myanmar?

Self-Selection Bias

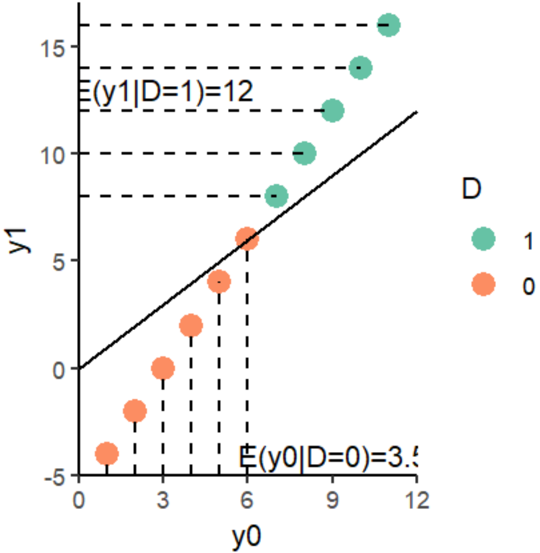


Self-Selection Bias



- $E(y_1) - E(y_0) = 0$

Self-Selection Bias



- $E(y_1|D=1) - E(y_0|D=0) = 8.5$

Self-Selection Bias

- Allow treatment effects to vary across individuals, so $Y_{1i} = Y_{0i} + \alpha_i$

Self-Selection Bias

- Allow treatment effects to vary across individuals, so $Y_{1i} = Y_{0i} + \alpha_i$

$$\underbrace{E(Y_i|D=1) - E(Y_i|D=0)}_{\text{Observed Effect}} = \underbrace{E(Y_{1i} - Y_{0i})}_{\text{Real ATE}}$$
$$+ \underbrace{\frac{1}{2}\Big[E(Y_{1i}|D=1) - E(Y_{1i}|D=0)\Big]}_{\text{Imbalance on } Y_1} + \underbrace{\frac{1}{2}\Big[E(Y_{0i}|D=1) - E(Y_{0i}|D=0)\Big]}_{\text{Imbalance on } Y_0} \tag{3}$$

NB: For equal-sized treatment and control groups

Self-Selection Bias

- Allow treatment effects to vary across individuals, so $Y_{1i} = Y_{0i} + \alpha_i$

$$\underbrace{E(Y_i|D=1) - E(Y_i|D=0)}_{\text{Observed Effect}} = \underbrace{E(Y_{1i} - Y_{0i})}_{\text{Real ATE}}$$
$$+ \underbrace{\frac{1}{2}\Big[E(Y_{1i}|D=1) - E(Y_{1i}|D=0)\Big]}_{\text{Imbalance on } Y_1} + \underbrace{\frac{1}{2}\Big[E(Y_{0i}|D=1) - E(Y_{0i}|D=0)\Big]}_{\text{Imbalance on } Y_0} \quad (3)$$

NB: For equal-sized treatment and control groups

Treatment Assignment Mechanism

- In all of these cases, **which units receive 'treatment' ($D_i = 1$)**, and why, affect our estimate of the relationship between D and Y

Treatment Assignment Mechanism

- In all of these cases, **which units receive 'treatment' ($D_i = 1$)**, and why, affect our estimate of the relationship between D and Y
  - This is the **Treatment Assignment Mechanism**

Treatment Assignment Mechanism

- In all of these cases, **which units receive 'treatment' ($D_i = 1$)**, and why, affect our estimate of the relationship between D and Y
  - This is the **Treatment Assignment Mechanism**
- Messy treatment assignment mechanisms are why basic regression is no use for explanation

Treatment Assignment Mechanism

- In all of these cases, **which units receive 'treatment' ($D_i = 1$)**, and why, affect our estimate of the relationship between D and Y
  - This is the **Treatment Assignment Mechanism**
- Messy treatment assignment mechanisms are why basic regression is no use for explanation
  - It means our comparison control cases are really misleading

Treatment Assignment Mechanism

- In all of these cases, **which units receive 'treatment' ($D_i = 1$)**, and why, affect our estimate of the relationship between D and Y
  - This is the **Treatment Assignment Mechanism**
- Messy treatment assignment mechanisms are why basic regression is no use for explanation
  - It means our comparison control cases are really misleading
  - Y_0 for North Korea is not a good guide to the Y_0 for Sweden

Treatment Assignment Mechanism

- In all of these cases, **which units receive 'treatment' ($D_i = 1$)**, and why, affect our estimate of the relationship between D and Y
  - This is the **Treatment Assignment Mechanism**
- Messy treatment assignment mechanisms are why basic regression is no use for explanation
  - It means our comparison control cases are really misleading
  - Y_0 for North Korea is not a good guide to the Y_0 for Sweden
  - What would happen if the control units got treated?

Treatment Assignment Mechanism

- The comparability of treatment and control units depends on how they got to be treated

Treatment Assignment Mechanism

- The comparability of treatment and control units depends on how they got to be treated

Treatment Assignment Mechanism

The set of factors that determine why some units have $D = 0$ and others have $D = 1$

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**
  - Independent means the values of the potential outcomes give us no information about whether that unit was treated

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**
  - Independent means the values of the potential outcomes give us no information about whether that unit was treated
  - Potential outcomes are 'balanced' across control and treatment groups

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**
  - Independent means the values of the potential outcomes give us no information about whether that unit was treated
  - Potential outcomes are 'balanced' across control and treatment groups

Independence of Treatment Assignment

Treatment Assignment does NOT depend on the values of units' Potential Outcomes

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**
  - Independent means the values of the potential outcomes give us no information about whether that unit was treated
  - Potential outcomes are 'balanced' across control and treatment groups

Independence of Treatment Assignment

Treatment Assignment does NOT depend on the values of units' Potential Outcomes

$(Y_1, Y_0) \perp D$

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**
  - Independent means the values of the potential outcomes give us no information about whether that unit was treated
  - Potential outcomes are 'balanced' across control and treatment groups

Independence of Treatment Assignment

Treatment Assignment does NOT depend on the values of units' Potential Outcomes

$$(Y_1, Y_0) \perp D$$

$$Pr(D|(Y_1, Y_0)) = Pr(D)$$

Treatment Assignment Mechanism

- Explanation is more reliable where the **Treatment Assignment Mechanism is Independent of Potential Outcomes**
  - Independent means the values of the potential outcomes give us no information about whether that unit was treated
  - Potential outcomes are 'balanced' across control and treatment groups

Independence of Treatment Assignment

Treatment Assignment does NOT depend on the values of units' Potential Outcomes

$(Y_1, Y_0) \perp D$

$Pr(D|(Y_1, Y_0)) = Pr(D)$

$E(Y|D=1) = E(Y|D=0) = E(Y)$

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?
  5. What is the Treatment Assignment Mechanism?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?
  5. What is the Treatment Assignment Mechanism?
  6. Draw a causal diagram of the variables in the study, including the treatment assignment mechanism

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?
  5. What is the Treatment Assignment Mechanism?
  6. Draw a causal diagram of the variables in the study, including the treatment assignment mechanism
  7. Is Treatment Assignment independent of Potential Outcomes?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?
  5. What is the Treatment Assignment Mechanism?
  6. Draw a causal diagram of the variables in the study, including the treatment assignment mechanism
  7. Is Treatment Assignment independent of Potential Outcomes?
  8. Describe the risk of:
     - Omitted Variable Bias?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?
  5. What is the Treatment Assignment Mechanism?
  6. Draw a causal diagram of the variables in the study, including the treatment assignment mechanism
  7. Is Treatment Assignment independent of Potential Outcomes?
  8. Describe the risk of:
     - Omitted Variable Bias?
     - Reverse Causation?

Summary

- Template to analyze a paper:
  1. What are the treatment and outcome variables?
  2. What are the Potential Outcomes?
  3. What is the Fundamental Problem of Causal Inference in this case?
  4. How do we define the Average Treatment Effect (ATE) in this case? ATT? ATU?
  5. What is the Treatment Assignment Mechanism?
  6. Draw a causal diagram of the variables in the study, including the treatment assignment mechanism
  7. Is Treatment Assignment independent of Potential Outcomes?
  8. Describe the risk of:
     - Omitted Variable Bias?
     - Reverse Causation?
     - Self-Selection?

DOES OIL HINDER DEMOCRACY?

By MICHAEL L. ROSS*

INTRODUCTION

POLITICAL scientists believe that oil has some very odd properties. Many studies show that when incomes rise, governments tend to become more democratic. Yet some scholars imply there is an exception to this rule: if rising incomes can be traced to a country's oil wealth, they suggest, this democratizing effect will shrink or disappear. Does oil really have antidemocratic properties? What about other minerals and other commodities? What might explain these effects?

The claim that oil and democracy do not mix is often used by area specialists to explain why the high-income states of the Arab Middle East have not become democratic. If oil is truly at fault, this insight could help explain—and perhaps, predict—the political problems of oil exporters around the world, such as Nigeria, Indonesia, Venezuela, and the oil-rich states of Central Asia. If other minerals have similar properties, this effect might help account for the absence or weakness of democracy in dozens of additional states in sub-Saharan Africa, Latin America, and Southeast Asia. Yet the "oil impedes democracy" claim has received little attention outside the circle of Mideast scholars; moreover, it has not been carefully tested with regression analysis, either within or beyond the Middle East.

I use pooled time-series cross-national data from 113 states between 1971 and 1997 to explore three aspects of the oil-impedes-democracy claim. The first is the claim's validity: is it true? Although the claim has been championed by Mideast specialists, it is difficult to test by examining only cases from the Middle East because the region provides scholars with

* Previous versions of this article were presented to seminars at Princeton University, Yale University, and the University of California, Los Angeles, and at the September 2000 annual meeting of the American Political Science Association in Washington, D.C. For their thoughtful comments on earlier drafts, I am grateful to Pradeep Chhibber, Indra de Soysa, Geoffrey Garrett, Phil Keefer, Steve Knack, Miriam Lowi, Ellen Lust-Okar, Lant Pritchett, Nicholas Sambanis, Jennifer Widner, Michael Woolcock, and three anonymous reviewers. I owe special thanks to Irfan Nooruddin for his research assistance and advice and to Colin Xu for his help with the Stata. I wrote this article while I was a visiting scholar at The World Bank in Washington, D.C. The views I express in this article, and all remaining errors, are mine alone.

World Politics 53 (April 2001), 325–61

Summary

- Try experimenting with the Causal Relationships App here

Summary

- Try experimenting with the Causal Relationships App here
- Can you create an artificial effect between D and Y even when there is no direct causal effect?

Summary

- Try experimenting with the Causal Relationships App here
- Can you create an artificial effect between D and Y even when there is no direct causal effect?
- Under what conditions can you recover the real treatment effect?

Section 4

Rest of the Course

Rest of the Course

- The rest of the course is mostly about:
  - **Design-Based Solutions** to the Fundamental Problem of Causal Inference: Which treatment assignment mechanisms **avoid these biases** and provide plausible counterfactuals

Rest of the Course

- The rest of the course is mostly about:
  - **Design-Based Solutions** to the Fundamental Problem of Causal Inference: Which treatment assignment mechanisms **avoid these biases** and provide plausible counterfactuals
  - How much can we learn with better research design?

Rest of the Course

- The rest of the course is mostly about:
  - **Design-Based Solutions** to the Fundamental Problem of Causal Inference: Which treatment assignment mechanisms **avoid these biases** and provide plausible counterfactuals
  - How much can we learn with better research design?
  - **Model-Based Solutions:** Not so much.

Rest of the Course

		Independence of Treatment Assignment	**Researcher Controls Treatment Assignment?**
Controlled Experiments	Field Experiments	✓	✓
	Survey and Lab Experiments	✓	✓
Natural Experiments	Randomized Natural Experiments	✓	
	Instrumental Variables	✓	
	Discontinuities	✓	
Observational Studies	Difference-in-Differences		
	Controlling for Confounding		
	Matching		
	Comparative Cases and Process Tracing		