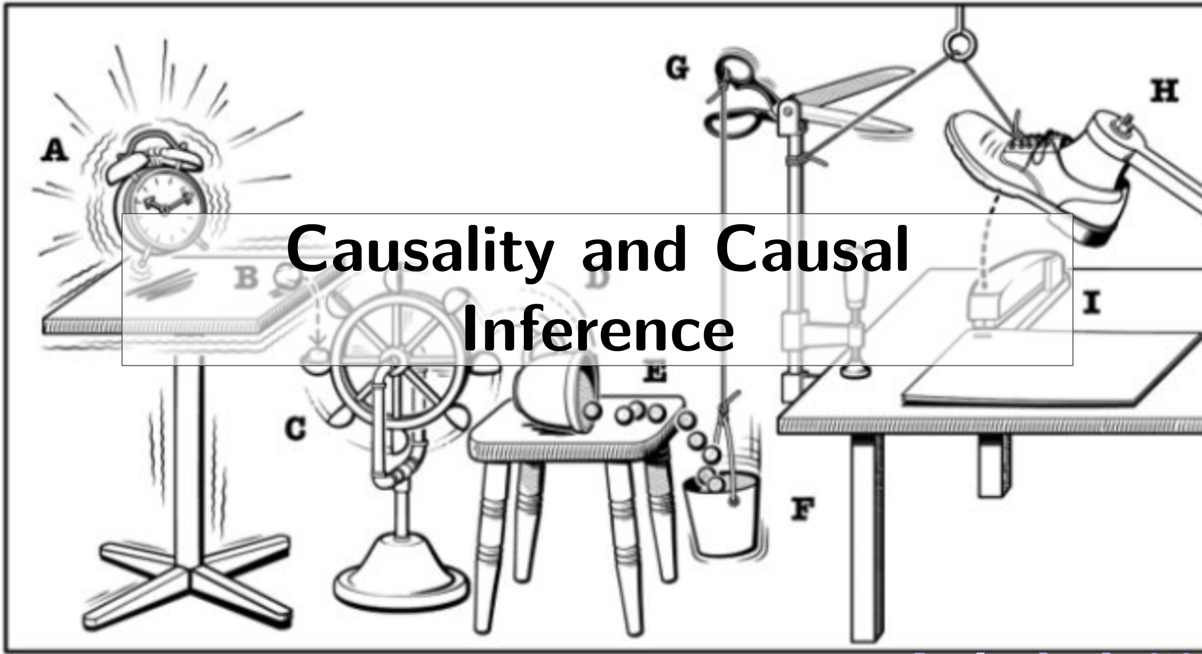
Causality and Causal Inference



Today:

- Introduce the Rubin Causal Model;
- Discuss the fundamental problem of causal inference;
- Use descriptive statistics to get around the fundamental problem of causal inference.

What is causality?!

- In this class we want to use political numbers appropriately – often this will look like does variable X cause a change in variable y:
  - Does education level change the gender wage gap?
  - Does incumbency confer an advantage?
  - Do nuclear weapons decrease the chance of conflict?
  - Do social safety nets decrease labor?
- Many, many definitions – here are two to think about:
  - Counterfactual causality:
    - The difference between worlds – we'll talk about this;
    - In the version of this we will discuss today causality is deterministic;
  - Probabilistic causation:
    - A change in a variable will on average cause a change in another variable;
    - This is probabilistic – note the 'on average' part.

The Rubin Causal Model

...the causal effect of one treatment, T, over another, C, for a particular unit is the difference between what would have happened if the unit had been exposed to T and what would have happened if the unit had been exposed to C...

The Rubin Causal Model

- **Example**: a political campaign is seeking money for its 'war chest' and wants to know if going door to door to visit party members will increase contributions;
- Let's think about a party member – the campaign wants to know:
  - The **donation** this person will make **if they get a visit** from the campaign;
  - The **donation** this person will make **if they don't get a visit** from the campaign;
- Some useful terms in the Rubin Causal Model:
  - The party member is a **unit**;
  - The visit is a **treatment**;
  - The group that gets no visit is the **control**;
  - The two possible donations are **potential outcomes**.

The Rubin Causal Model

Unit	Y_i(visit)	Y_i(none)	
1	\$675	\$150	

The Rubin Causal Model

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	
1	\$675	\$150	
2	\$3600	\$2500	
3	\$1900	\$3300	
4	\$2300	\$1000	
5	\$2600	\$2000	
6	\$3000	\$0	
7	\$1950	\$2500	

The Rubin Causal Model

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	$Y_i(\text{visit}) - Y_i(\text{none})$
1	\$675	\$150	\$525
2	\$3600	\$2500	\$1100
3	\$1900	\$3300	-\$1400
4	\$2300	\$1000	\$1300
5	\$2600	\$2000	\$600
6	\$3000	\$0	\$3000
7	\$1950	\$2500	-\$550

The Rubin Causal Model

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	$Y_i(\text{visit}) - Y_i(\text{none})$
1	\$675	?	\$525
2	\$3600	?	\$1100
3	?	\$3300	-\$1400
4	\$2300	?	\$1300
5	?	\$2000	\$600
6	?	\$0	\$3000
7	\$1950	?	-\$550

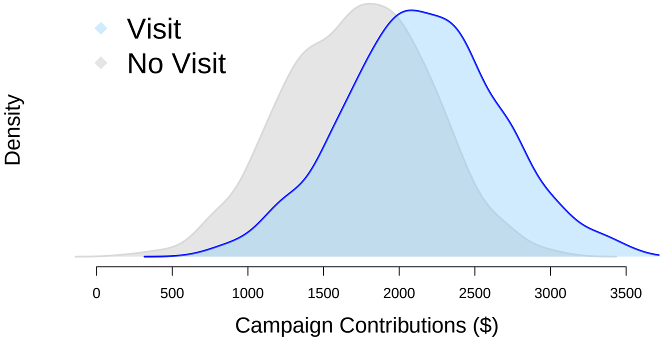
The Rubin Causal Model

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	$Y_i(\text{visit}) - Y_i(\text{none})$
1	\$675	?	?
2	\$3600	?	?
3	?	\$3300	?
4	\$2300	?	?
5	?	\$2000	?
6	?	\$0	?
7	\$1950	?	?

The **fundamental problem of causal inference**: one of the potential outcomes is ALWAYS missing.

Estimating Effects: Graphically via Density Plots

- Can't observe unit level casual effects directly;



Estimating Effects: Average Causal Effect

- Can't observe unit level casual effects directly;

- Can compute an average causal effect – for n_T units in the treatment group and n_C units in the control group:

$$\frac{1}{n_T}\sum_i Y_i(\mathrm{T}) - \frac{1}{n_C}\sum_i Y_i(\mathrm{C});$$

- For our example:

$$\frac{\$675 + \$3600 + \$2300 + \$1950}{4} - \frac{\$3300 + \$2000 + \$0}{3} \approx \$365.$$

Estimating Effects: Average Causal Effect

- Can't observe unit level casual effects directly;

- Can compute an average causal effect – for n_T party members in the visit group and n_C party members in the no visit group:

$$\frac{1}{n_T}\sum_i Y_i(\text{visit}) - \frac{1}{n_C}\sum_i Y_i(\text{none});$$

- For our example:

$$\frac{\$675 + \$3600 + \$2300 + \$1950}{4} - \frac{\$3300 + \$2000 + \$0}{3} \approx \$365.$$

Estimating Effects: Inferring Unit Level Outcomes

- Can't observe unit level casual effects directly;

- So what can we do? We can **infer** unit level effects by making an assumption – consider **a unit in the treatment group**:

$$\underbrace{Y_i(\mathrm{T})}_{\text{Observed!}} - Y_i(\mathrm{C}) = Y_i^*$$

Estimating Effects: Inferring Unit Level Outcomes

- Can't observe unit level casual effects directly;

- So what can we do? We can **infer** unit level effects by making an assumption – consider **a unit in the treatment group**:

$$\underbrace{Y_i(\mathrm{T})}_{\text{Observed!}} - Y_i(\mathrm{C}) = Y_i^*$$

$$\underbrace{Y_i(\mathrm{T})}_{\text{Observed!}} - Y_i(\mathrm{C}) = \underbrace{\frac{1}{n_T}\sum_i Y_i(\mathrm{T}) - \frac{1}{n_C}\sum_i Y_i(\mathrm{C})}_{\text{Observed!}}$$

$$Y_i(\mathrm{C}) = \underbrace{Y_i(\mathrm{T})}_{\text{Observed!}} - \underbrace{\left(\frac{1}{n_T}\sum_i Y_i(\mathrm{T}) - \frac{1}{n_C}\sum_i Y_i(\mathrm{C})\right)}_{\text{Observed!}}$$

- This is called the **assumption of constant effect**.

Estimating Effects: Inferring Unit Level Outcomes

- Can't observe unit level casual effects directly;

- So what can we do? We can **infer** unit level effects by making an assumption – consider **a unit in the control group**:

$$Y_i(\mathrm{T}) - \underbrace{Y_i(\mathrm{C})}_{\text{Observed!}} = Y_i^*$$

$$Y_i(\mathrm{T}) - \underbrace{Y_i(\mathrm{C})}_{\text{Observed!}} = \underbrace{\frac{1}{n_T}\sum_i Y_i(\mathrm{T}) - \frac{1}{n_C}\sum_i Y_i(\mathrm{C})}_{\text{Observed!}}$$

$$Y_i(\mathrm{T}) = \underbrace{Y_i(\mathrm{C})}_{\text{Observed!}} + \underbrace{\left(\frac{1}{n_T}\sum_i Y_i(\mathrm{T}) - \frac{1}{n_C}\sum_i Y_i(\mathrm{C})\right)}_{\text{Observed!}}$$

- This is called the **assumption of constant effect**.

Estimating Effects: Inferring Unit Level Outcomes

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	$Y_i(\text{visit}) - Y_i(\text{none})$
1	\$675	?	?
2	\$3600	?	?
3	?	\$3300	?
4	\$2300	?	?
5	?	\$2000	?
6	?	\$0	?
7	\$1950	?	?

Estimating Effects: Inferring Unit Level Outcomes

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	$Y_i(\text{visit}) - Y_i(\text{none})$
1	\$675	?	\$365
2	\$3600	?	\$365
3	?	\$3300	\$365
4	\$2300	?	\$365
5	?	\$2000	\$365
6	?	\$0	\$365
7	\$1950	?	\$365

Estimating Effects: Inferring Unit Level Outcomes

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	$Y_i(\text{visit}) - Y_i(\text{none})$
1	\$675	**\$310**	\$365
2	\$3600	**\$3235**	\$365
3	**\$3665**	\$3300	\$365
4	\$2300	**\$1935**	\$365
5	**\$2365**	\$2000	\$365
6	**\$365**	\$0	\$365
7	\$1950	**\$1585**	\$365

Assignment Effects

- Fair warning 1: the assignment of units to the treatment and the control groups is **very consequential**;

- Question: how many ways are there to assign our 7 party members?

- Question: how many ways would there be to assign 1000 party members?

Assignment Effects

- Fair warning 1: the assignment of units to the treatment and the control groups is **very consequential**;

- Question: how many ways are there to assign our 7 party members?

$$\text{Answer: } 2^7 - 2 = 126;$$

- Question: how many ways would there be to assign 1000 party members?

$$\text{Answer: } 2^{1000} - 2 \approx 10^{301}.$$

Assignment Effects

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	
1	**\$675**	\$150	
2	**\$3600**	\$2500	
3	\$1900	**\$3300**	Average CE was:
4	**\$2300**	\$1000	\$365
5	\$2600	**\$2000**	
6	\$3000	**\$0**	
7	**\$1950**	\$2500	

Assignment Effects

Unit	Y_i(visit)	Y_i(none)	
1	**$675**	$150	
2	$3600	**$2500**	
3	$1900	**$3300**	Average CE is:
4	$2300	**$1000**	-$1208
5	$2600	**$2000**	
6	$3000	**$0**	
7	$1950	**$2500**	

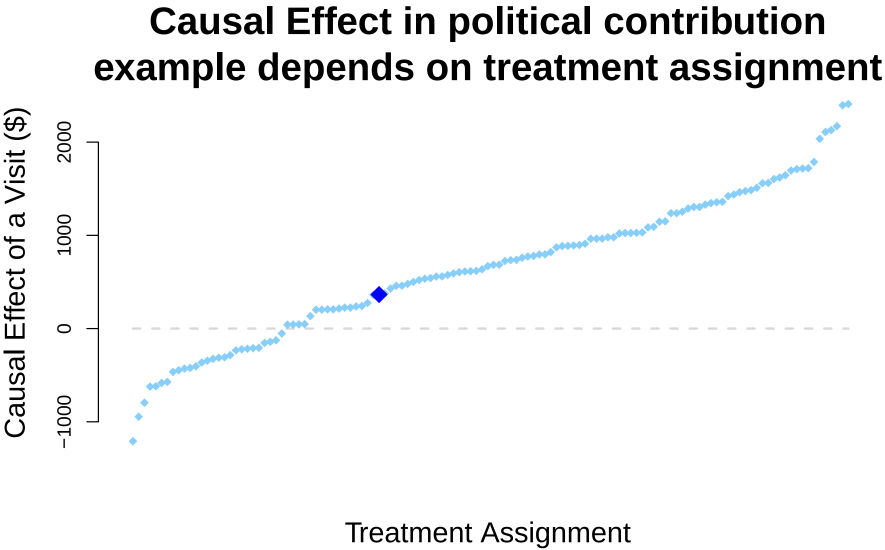
Assignment Effects: a perfect campaign manager

Unit	Y_i(visit)	Y_i(none)	
1	**\$675**	\$150	
2	**\$3600**	\$2500	
3	\$1900	**\$3300**	Average CE is:
4	**\$2300**	\$1000	-\$465
5	**\$2600**	\$2000	
6	**\$3000**	\$0	
7	\$1950	**\$2500**	

Assignment Effects: a perfect visit advocate

Unit	$Y_i(\text{visit})$	$Y_i(\text{none})$	
1	\$675	**\$150**	
2	**\$3600**	\$2500	
3	**\$1900**	\$3300	Average CE is:
4	\$2300	**\$1000**	\$2129
5	**\$2600**	\$2000	
6	\$3000	**\$0**	
7	**\$1950**	\$2500	

Assignment Effects



Why should we care?

Causal inference is currently one of the most popular research methods in political science and is a simple way to gauge the effect of a binary variable.