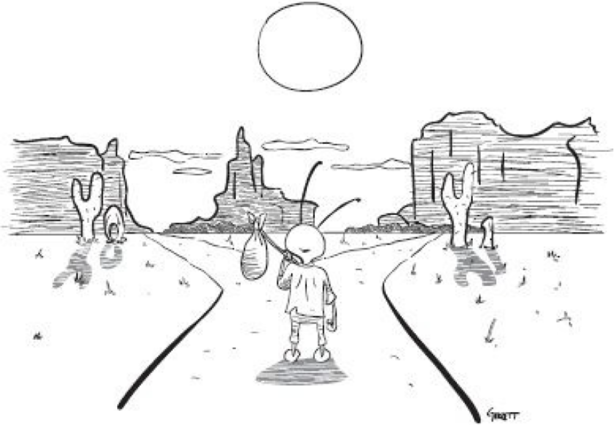
Recitation: Review

Felipe Balcazar

NYU

October, 2022

All else equal



- What is the impact of a choice/event (D) on an outcome (Y)?

Fundamental Problem of Causal Inference

Definition (Fundamental Problem of Causal Inference)

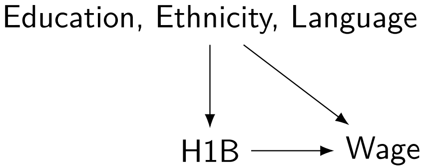
We cannot observe both potential outcomes. How can we calculate the avg. causal effect?

Naive comparisons have bias: difference in means

$$
\begin{aligned}
E[Y_{1,i}|D_i = 1] - E[Y_{0,i}|D_i = 0] &= \underbrace{E[Y_{1,i} - Y_{0,i}|D_i = 1]}_{ATT} + \underbrace{\{E[Y_{0,i}|D_i = 1] - E[Y_{0,i}|D_i = 0]\}}_{\text{Selection bias}} \\
&= \kappa + \underbrace{\{E[Y_{0,i}|D_i = 1] - E[Y_{0,i}|D_i = 0]\}}_{\text{Selection bias}} \\
&= \text{avg. causal effect} + \text{Selection bias}
\end{aligned}
$$

Thus, most naive comparisons are not "other things equal" *(ceteris paribus)* comparisons.

Effect of being migrant on an H1B Visa on wages



Effect of being migrant on an H1B Visa on wages

$$
\begin{tikzcd}
\text{Education, Ethnicity, Language} \arrow[d] \arrow[dr] & \\
\text{H1B} \arrow[r] & \text{Wage}
\end{tikzcd}
$$

$$E[Y_{1,i}|D_i = 1] - E[Y_{0,i}|D_i = 0] = \kappa + E[Y_{0i}|Di = 1] - E[Y_{0i}|Di = 0]$$

The method for causal inference: regression

$$Y_i = \alpha + \beta D_i + \varepsilon_i.$$

The method for causal inference: regression

$$Y_i = \alpha + \beta D_i + \varepsilon_i.$$

Remember

- $\widehat{\alpha} = E(Y_i | D_i = 0)$.
- $\widehat{\beta} = E(Y_i | D_i = 1) - E(Y_i | D_i = 0)$.

The method for causal inference: regression

$$Y_i = \alpha + \beta D_i + \varepsilon_i.$$

Remember

- $\widehat{\alpha} = E(Y_i | D_i = 0)$.
- $\widehat{\beta} = E(Y_i | D_i = 1) - E(Y_i | D_i = 0)$.

- Is β the average treatment effect?

The method for causal inference: regression

$$Y_i = \alpha + \beta D_i + \delta X_i + \epsilon_i.$$

The method for causal inference: regression

$$Y_i = \alpha + \beta D_i + \delta X_i + \epsilon_i.$$

Remember

- $\widehat{\alpha} = E(Y_i|D_i = 0, X_i)$.
- $\widehat{\beta} = E(Y_i|D_i = 1, X_i) - E(Y_i|D_i = 0, X_i)$.

The method for causal inference: regression

$$Y_i = \alpha + \beta D_i + \delta X_i + \epsilon_i.$$

Remember

- $\widehat{\alpha} = E(Y_i | D_i = 0, X_i)$.
- $\widehat{\beta} = E(Y_i | D_i = 1, X_i) - E(Y_i | D_i = 0, X_i)$.

- Is β the average treatment effect? If not, when?

Effect of being migrant on an H1B Visa on wages

$$Wage_i = \alpha + \beta H1B_i + \underbrace{\delta_1 Edu_i + \delta_2 Eth_i + \delta_3 Lang_i}_{\delta X_i} + \epsilon_i$$

Recall:

- $\widehat{\alpha} = E(Wage_i | H1B_i = 0, X_i)$.
- $\widehat{\beta} = E(Wage_i | H1B_i = 1, X_i) - E(Wage_i | H1B_i = 0, X_i)$.

Effect of being migrant on an H1B Visa on wages

$$Wage_i = \alpha + \beta H1B_i + \underbrace{\delta_1 Edu_i + \delta_2 Eth_i + \delta_3 Lang_i}_{\delta X_i} + \epsilon_i$$

Recall:

- $\widehat{\alpha} = E(Wage_i | H1B_i = 0, X_i)$.
- $\widehat{\beta} = E(Wage_i | H1B_i = 1, X_i) - E(Wage_i | H1B_i = 0, X_i)$.

- Is β the average treatment effect? If not when?

Random assignment eliminates selection bias

- When D_i is randomly assigned:

$$E[Y_{0i}|Di = 1] = E[Y_{0i}|Di = 0].$$

- If this occurs the difference in means by treatment status captures the causal effect of treatment.

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = ATE + \underbrace{E[Y_{0i}|Di = 1] - E[Y_{0i}|Di = 0]}_{0}$$

$$= ATE$$

Random assignment eliminates selection bias

- When D_i is randomly assigned:

$$E[Y_{0i}|Di = 1] = E[Y_{0i}|Di = 0].$$

- If this occurs the difference in means by treatment status captures the causal effect of treatment.

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = ATE + \underbrace{E[Y_{0i}|Di = 1] - E[Y_{0i}|Di = 0]}_{0}$$
$$= ATE$$

Can you think about a simple experiment for the effect of H1B on wages?

Random assignment eliminates selection bias

- When D_i is randomly assigned:

$$E[Y_{0i}|Di = 1] = E[Y_{0i}|Di = 0].$$

- If this occurs the difference in means by treatment status captures the causal effect of treatment.

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = ATE + \underbrace{E[Y_{0i}|Di = 1] - E[Y_{0i}|Di = 0]}_{0}$$
$$= ATE$$

Can you think about a simple experiment for the effect of H1B on wages? Can the experiment fail?

Random assignment eliminates selection bias

- When D_i is randomly assigned:

$$E[Y_{0i}|Di = 1] = E[Y_{0i}|Di = 0].$$

- If this occurs the difference in means by treatment status captures the causal effect of treatment.

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = ATE + \underbrace{E[Y_{0i}|Di = 1] - E[Y_{0i}|Di = 0]}_{0}$$

$$= ATE$$

Can you think about a simple experiment for the effect of H1B on wages? Can the experiment fail? How do we now?

Are they perfect?

- Experiments can be expensive to carry out.
  - Often you can run them on (small) random samples.
  - This is contingent on the treatment and outcome of interest.
- Often you need cooperation from local political actors.
- They can be time consuming (perhaps it is not wise to do one if you are graduating soon).

Are they perfect?

- Experiments can be expensive to carry out.
  - Often you can run them on (small) random samples.
  - This is contingent on the treatment and outcome of interest.
- Often you need cooperation from local political actors.
- They can be time consuming (perhaps it is not wise to do one if you are graduating soon).
- They may lack external validity.
  - But they are replicable and scalable, addressing this issue.

Are they perfect?

- Experiments can be expensive to carry out.
  - Often you can run them on (small) random samples.
  - This is contingent on the treatment and outcome of interest.
- Often you need cooperation from local political actors.
- They can be time consuming (perhaps it is not wise to do one if you are graduating soon).
- They may lack external validity.
  - But they are replicable and scalable, addressing this issue.
- You cannot use them to study any type of question.
- Despite these shortcomings, they are the *gold standard*.
  - They are internally valid!
  - They solve the problem of selection bias!
  - We can estimate the direction and magnitude of the treatment!

Ethical Concerns

- Some people object to field experiments on ethical grounds.
- Some units are denied potentially helpful treatments just for purposes of evaluation.
- Participant consent is important, but may some times influence the outcomes (*Hawthorne* Effects).
- They are often justified on the grounds of:
    - Scarce resources/staggered implementation.
    - Potential gains from learning about impact of treatment compensates for ethical considerations.

Interaction effects: regression form

$$Y_i = \alpha + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i \times X_i + \varepsilon_i$$

- β_1 is the effect of the treatment conditional on $X = 0$:

$$E(Y_i|D_i = 1, X_i = 0) - E(Y_i|D_i = 0, X_i = 0)$$

- $\beta_1 + \beta_3$ is the effect of the treatment conditional on $X = 1$:

$$E(Y_i|D_i = 1, X_i = 1) - E(Y_i|D_i = 0, X_i = 1)$$

- β_3 is the additional effect of the treatment for $X = 1$.

Interaction effects: regression form

$$Y_i = \alpha + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i \times X_i + \varepsilon_i$$

- β_1 is the effect of the treatment conditional on $X = 0$:

$$E(Y_i|D_i = 1, X_i = 0) - E(Y_i|D_i = 0, X_i = 0)$$

- $\beta_1 + \beta_3$ is the effect of the treatment conditional on $X = 1$:

$$E(Y_i|D_i = 1, X_i = 1) - E(Y_i|D_i = 0, X_i = 1)$$

- β_3 is the additional effect of the treatment for $X = 1$.
- Neither D nor X need to be binary!

$$Y_i = \alpha + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i \times X_i + \varepsilon_i$$

- β_1 is the effect of the treatment conditional on $X = 0$:

$$E(Y_i|D_i = 1, X_i = 0) - E(Y_i|D_i = 0, X_i = 0)$$

- $\beta_1 + \beta_3$ is the effect of the treatment conditional on $X = 1$:

$$E(Y_i|D_i = 1, X_i = 1) - E(Y_i|D_i = 0, X_i = 1)$$

- β_3 is the additional effect of the treatment for $X = 1$.
- Neither D nor X need to be binary!

Interaction effects: regression form

$$Y_i = \alpha + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i \times X_i + \varepsilon_i$$

- β_1 is the effect of the treatment conditional on $X = 0$:

$$E(Y_i|D_i = 1, X_i = 0) - E(Y_i|D_i = 0, X_i = 0)$$

- $\beta_1 + \beta_3$ is the effect of the treatment conditional on $X = 1$:

$$E(Y_i|D_i = 1, X_i = 1) - E(Y_i|D_i = 0, X_i = 1)$$

- β_3 is the additional effect of the treatment for $X = 1$.
- Neither D nor X need to be binary!

Interaction effects: regression form

$$Y_i = \alpha + \beta_1 D_i + \beta_2 X_i + \beta_3 D_i \times X_i + \varepsilon_i$$

- β_1 is the effect of the treatment conditional on $X = 0$:

$$E(Y_i | D_i = 1, X_i = 0) - E(Y_i | D_i = 0, X_i = 0)$$

- $\beta_1 + \beta_3$ is the effect of the treatment conditional on $X = 1$:

$$E(Y_i | D_i = 1, X_i = 1) - E(Y_i | D_i = 0, X_i = 1)$$

- β_3 is the additional effect of the treatment for $X = 1$.
- Neither D nor X need to be binary!

$$Wage_i = \alpha + \beta_1 H1B_i + \beta_2 H1B_i \times Eth_i + \delta X_i + \epsilon$$

So

- $\widehat{\beta}_1 = E(Wage_i | H1B_i = 1, Eth_i = 0, X_i) - E(Wage_i | H1B_i = 0, Eth_i = 0, X_i)$
- $\widehat{\beta}_1 + \widehat{\beta}_2 = E(Wage_i | H1B_i = 1, Eth_i = 1, X_i) - E(Wage_i | H1B_i = 0, Eth_i = 1, X_i)$

Practice Questions

1. What is the unit of analysis (i)?
2. What is the treatment? Which are the treated units ($D_i = 1$) and which are control units ($D_i = 0$)?
3. What is the outcome variable Y_i? How would you measure it?
4. What is the ideal counterfactual to assess the causal relationship?
5. What is the naive comparison (simple correlation)? What do you expect to be the sign of the naive comparison?
6. What do Y_{0i} and Y_{1i} correspond to in this context?
7. What sign do you expect selection bias to have in this context? Mention explicitly the main observable and non-observable confounders that may generate this selection bias. (Use selection bias formula and a DAG).
8. Does the naive comparison tend to over-estimate or under-estimate the causal effect?