Causality

Professor Ji-Woong Chung
Korea University

This lecture note is based on Todd Gormley's.

Outline

Motivate why we care about causality

Three possible biases

- Omitted Variable Bias
- Measurement error bias
- Simultaneity bias

Outline

Motivate why we care about causality

Three possible biases

- Omitted Variable Bias
- Measurement error bias
- Simultaneity bias

Motivation

- As researchers, we are interested in making causal statements.
- Example #1: What is the effect of a change in corporate taxes on firms' leverage choice?
- Example #2: What is the effect of giving a CEO more stock ownership in the firm on the CEO's desire to take on risky investments?
- We don't like to just say variables are 'associated' or 'correlated' with each other.

What do we mean by causality?

Recall from earlier lecture that if our linear model is the following:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u$$

And we want to infer β_1 as the causal effect of x_1 on y holding all else equal, then we need to make the following assumptions...

The Basic Assumptions

- Assumption #1: $E(u) = 0$
- Assumption #2: $E(u|x_1 \dots x_k) = E(u)$
  - In words, the average of u (i.e., unexplained portion of y) does not depend on the value of x.
  - "Conditional mean independence" (CMI).

Outline

Motivate why we care about causality

Three possible biases

- Omitted Variable Bias
- Measurement error bias
- Simultaneity bias

Three Main Ways This Will Be Violated

- Omitted variable bias
- Measurement error bias
- Simultaneity bias

Now let's go through each in turn...

Outline

Motivate why we care about causality

Three possible biases

- Omitted Variable Bias
- Measurement error bias
- Simultaneity bias

Omitted Variable Bias (OVB)

- Probably the most common concern you will hear researchers worry about.
- Basic idea = the estimation error u contains another variable, e.g., z that affects y **and** is correlated with an x.
- Please note! The omitted variable is only problematic if correlated with an x.

OVB More Formally with One Variable

You estimate:

$$y = \beta_0 + \beta_1 x + u$$

But the true model is:

$$y = \beta_0 + \beta_1 x + \beta_2 z + v$$

Then

$$\hat{\beta}_1 = \beta_1 + \delta_{xz}\beta_2$$

where δ_{xz} is the coefficient you'd get from regressing the omitted variable z on x.

$$\delta_{xz} = \frac{\text{cov}(x, z)}{\text{var}(x)}$$

Interpreting the OVB Formula

$$\hat{\beta}_1 = \underbrace{\beta_1}_{\text{Effect of } x \text{ on } y} + \underbrace{\underbrace{\frac{\text{cov}(x,z)}{\text{var}(x)}}_{\text{Regression of } z \text{ on } x} \underbrace{\beta_2}_{\text{Effect of } z \text{ on } y}}_{\text{Bias}}$$

- Easy to see the estimated coefficient is only unbiased if $\text{cov}(x,z) = 0$ [i.e., x and z are uncorrelated] or z has no effect on y [i.e., $\beta_2 = 0$].

Direction and Magnitude of the Bias

$$\hat{\beta}_1 = \beta_1 + \frac{\text{cov}(x, z)}{\text{var}(x)}\beta_2$$

- Direction of bias given by signs of β_2, $\text{cov}(x, z)$.
  - E.g., if we know z has a positive effect on y [i.e., $\beta_2 > 0$] and x and z are positively correlated [$\text{cov}(x, z) > 0$], then the bias will be positive.
- Magnitude of the bias will be given by magnitudes of β_2, $\frac{\text{cov}(x,z)}{\text{var}(x)}$.

Example – One Variable Case

Suppose we estimate:

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + w$$

But the true model is:

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{ability} + u$$

What is likely bias on β_1? Recall

$$\hat{\beta}_1 = \beta_1 + \beta_2 \frac{\text{cov}(\text{educ, ability})}{\text{var}(\text{educ})}$$

.

Example – Answer

- Ability & wages are likely positively correlated so $\beta_2 > 0$.
- Ability & education are likely positively correlated so cov(education, ability) > 0.
- Thus, the bias is likely to be positive! $\hat{\beta}_1$ is too big!

OVB – General Form

- Once you move away from the simple case of just one omitted variable, determining the sign (and magnitude) of bias will be a lot harder.
  - Let β be the vector of coefficients on k included variables.
  - Let γ be the vector of coefficients on l excluded variables .
  - Let $\boldsymbol{X}$ be the matrix of observations of included variables.
  - Let $\boldsymbol{Z}$ be the matrix of observations of excluded variables.

OVB – General Form Intuition

$$\hat{\beta} = \beta + \underbrace{\frac{E[\boldsymbol{X}'\boldsymbol{Z}]}{E[\boldsymbol{X}'\boldsymbol{X}]}}_{\substack{\text{Vector of regression} \\ \text{coefficient}}} \underbrace{\gamma}_{\substack{\text{Vector of partial effects} \\ \text{of excluded variables}}}$$

- Same idea as before but more complicated.
- This can be a real mess!

Eliminating Omitted Variable Bias

How we try to get rid of this bias will depend on the type of omitted variable:

- Observable omitted variable
- Unobservable omitted variable

How can we deal with an observable omitted variable?

Observable Omitted Variables

- This is easy! Just add them as controls.
  - E.g., if the omitted variable z in my simple case was 'leverage,' then add leverage to regression.
- A functional form misspecification is a special case of an observable omitted variable.
- Let's now talk about this...

Functional Form Misspecification

Assume true model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + u$$

However, we omit the squared term x_2^2.

- Just like any OVB, bias on $(\beta_0, \beta_1, \beta_2)$ will depend on β_3 and correlations among (x_1, x_2, x_2^2).
- You get the same type of problem if you have an incorrect functional form for y [e.g., it should be $\ln(y)$ not y].

In some sense, this is a minor problem... Why?

Tests for Correction Functional Form

- You could add additional squared and cubed terms and look to see whether they make a difference and/or have non-zero coefficients.
- This isn't as easy when the possible models are not nested...

Non-Nested Functional Form Issues

Two non-nested examples

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$
$$\text{versus} \quad y = \beta_0 + \beta_1 \ln(x_1) + \beta_2 \ln(x_2) + u$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$
$$\text{versus} \quad y = \beta_0 + \beta_1 x_1 + \beta_2 z + u$$

Let's use the first example and see how we can try to figure out which is right

Nonnested models: neither equation is a special case of the other.

Davidson-MacKinnon Test, 1981 [Part 1]

To test which is correct you can try this...

- Take fitted values $\hat{y}_1$ from 1st model and add them as a control in 2nd model.

$$y = \beta_0 + \beta_1 \ln(x_1) + \beta_2 \ln(x_2) + \theta_1 \hat{y} + u$$

- If 2nd model is correct, then θ_1 should be insignificant. If significant, rejects 2nd model!
- Then do the reverse and look at t-stat on θ_2 in:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \theta_2 \tilde{y} + v$$

where $\tilde{y}$ is predicted value from 2nd model... if significant, then 1st model is also rejected.

Davidson-MacKinnon Test, 1981 [Part 2]

Number of weaknesses to this test:

- A clear winner may not emerge.
  - Both might be rejected.
  - Both might be accepted [If this happens you can use the R^2 to choose which model is a better fit].
- And rejecting one model does NOT imply that the other model is correct.

Bottom Line Advice on Functional Form

Practically speaking, you hope that changes in functional form won't affect coefficients on key variables very much...

- But if it does... You need to think hard about why this is and what the correct form should be.
- The prior test might help with that...

But, if the effects of key independent variables on y are not very different, then it does not really matter which model is used.

Eliminating Omitted Variable Bias

How we try to get rid of this bias will depend on the type of omitted variable:

- Observable omitted variable
- Unobservable omitted variable

Unobservables are much harder to deal with but one possibility is to find a proxy variable.

Unobserved Omitted Variables Example

- Consider the estimation:

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{ability} + u$$

Problem: we don't observe & can't measure ability.

What can we do?
Answer: Find a proxy variable that is correlated with the unobserved variable, e.g., IQ.

Proxy Variables [Part 1]

Consider the following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3^* + u$$

where x_3^* is unobserved but we have a proxy, x_3. Then suppose:

$$x_3^* = \delta_0 + \delta_3 x_3 + v_3$$

- v_3 is the error associated with the proxy's imperfect representation of unobservable x_3^*.
- Intercept just accounts for different scales [e.g., ability has different average value than IQ]

Proxy Variables [Part 2]

- If we are only interested in β_1 or β_2, we can just replace x_3 with x_3^* and we run the regression of y on x_1, x_2, and x_3.
- But for this to give us consistent estimates of β_1 and β_2, we need to make some assumptions.

  #1 – We've got the right model.

  #2 – Other variables don't explain our unobserved variable after we've accounted for our proxy.

Proxy Variables – Assumptions

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3^* + u$$

$$x_3^* = \delta_0 + \delta_3 x_3 + v_3$$

#1 – $E(u|x_1, x_2, x_3^*) = 0$; i.e., we have the right model and x_3 would be irrelevant if we control for x_1, x_2, x_3^* such that:
$E(u|x_1, x_2, x_3^*, x_3) = E(u|x_1, x_2, x_3^*)(= 0)$

- This is a common (and important) assumption.

#2 – $E(v_3|x_1, x_2, x_3) = 0$; i.e., x_3 is a good proxy for x_3^* such that after controlling for x_3, x_3^* does not depend on x_1 or x_2.

- I.e., $E(x_3^*|x_1, x_2, x_3) = E(x_3^*|x_3) = \delta_0 + \delta_3 x_3$

Why the Proxy Works...

Recall the true model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3^* + u$$

Now plug-in for x_3^* using $x_3^* = \delta_0 + \delta_3 x_3 + v_3$

$$y = \underbrace{(\beta_0 + \beta_3 \delta_0)}_{\alpha_0} + \beta_1 x_1 + \beta_2 x_2 + \underbrace{(\beta_3 \delta_3)}_{\alpha_1} x_3 + \underbrace{(u + \beta_3 v_3)}_{e}$$

Prior assumptions ensure that $E(e|x_1, x_2, x_3) = 0$ such that the estimates of $\alpha_0, \beta_1, \beta_2, \alpha_1$ are consistent.

Note: β_0 and β_3 are not identified.

Proxy Assumptions are Key [Part 1]

Suppose assumption #2 is wrong such that x_3^* is correlated to all of the observed variables:

$$x_3^* = \delta_0 + \delta_3 x_3 + \underbrace{\gamma_1 x_1 + \gamma_2 x_2 + w}_{v}$$

where $E(w|x_1, x_2, x_3) = 0$

If the above is true, $E(v|x_1, x_2, x_3) \neq 0$, and if you substitute into the model of y you'd get...

Proxy Assumptions are Key [Part 2]

Plugging in for x_3^* you'd get:

$$y = \underbrace{(\beta_0 + \beta_3\delta_0)}_{\alpha_0} + \underbrace{(\beta_1 + \beta_3\gamma_1)}_{\alpha_1} x_1 + \underbrace{(\beta_2 + \beta_3\gamma_2)}_{\alpha_2} x_2 + \underbrace{(\beta_3\delta_3)}_{\alpha_3} x_3 + e$$

- E.g., α_1 captures effect of x_1 on y (β_1) but also its correlation with unobserved variable.
- We'd get consistent estimates of $\alpha_0, \alpha_1, \alpha_2, \alpha_3$.
- But that isn't what we want!

Proxy Variables – Example #1

Consider the earlier wage estimation:

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{ability} + u$$

- If we use IQ as a proxy for unobserved *ability*, what assumption must we make? Is it plausible?
- Answer: We assume $E(ability|educ, IQ) = E(ability|IQ)$,

  i.e., average ability does not change with education after accounting for IQ... Could be a questionable assumption!

Proxy Variables – Example #2

Consider the Q-theory of investment:

$$\text{Investment} = \beta_0 + \beta_1 Q + u$$

- Can we estimate β_1 using a firm's market-to-book ratio (MTB) as a proxy for Q?
- Answer: Even if we believe this is the correct model (Assumption #1) or that Q only depends on MTB (Assumption #2), e.g., $Q = \delta_0 + \delta_1 MTB$, we are still not getting an estimate of β_1... see next slide for the math.

Proxy Variables – Example #2 [Part 2]

Even if assumptions held, we'd only be getting consistent estimates of:

$$\text{Investment} = \alpha_0 + \alpha_1 MTB + e$$

where $\alpha_0 = \beta_0 + \beta_1\delta_0$ and $\alpha_1 = \beta_1\delta_1$.

- While we can't get β_1, is there something we can get if we make assumptions about the sign of δ_1?
- Answer: Yes, the sign of β_1.

Proxy Variables – Summary

- If the coefficient on the unobserved variable isn't what we are interested in, then a proxy for it can be used to identify and remove/mitigate OVB from the other parameters.
- A proxy can also be used to determine the sign of the coefficient on an unobserved variable.

Random Coefficient Model

So far we've assumed that the effect of x on y (i.e., β) was the same for all observations.

- In reality, this is unlikely true; the model might look more like:

$$y_i = a_i + b_i x_i$$

  where $a_i = \alpha + c_i$, $b_i = \beta + d_i$

- α is the average intercept, $E(a_i)$, and β is what we call the "average partial effect" (APE) , $E(b_i)$.

Random Coefficient Model [Part 2]

Regression would seem to be incorrectly specified but if willing to make assumptions, we can identify the APE.

- Plug in for a_i and b_i:

$$\begin{aligned} y_i &= (\alpha + c_i) + (\beta + d_i)x_i \\ &= \alpha + \beta x_i + \underbrace{(c_i + d_i x_i)}_{u_i} \end{aligned}$$

- The error term contains an interaction between an unobservable, d_i, and the observed explanatory variable, x_i.
- Identification requires

$$E(u|x_i) = E(c_i + d_i x_i | x_i) = 0$$

1

What does this imply?

[1] The error term is heteroskedastic: $Var(u_i|x_i) = \sigma_c^2 + \sigma_d^2 x_i^2$

Random Coefficient Model [Part 3]

This amounts to requiring:

$$E(c_i|x_i) = E(c_i) = 0 \Rightarrow E(a_i|x_i) = E(a_i)$$
$$E(d_i|x_i) = E(d_i) = 0 \Rightarrow E(b_i|x_i) = E(b_i)$$
$$\because a_i = \alpha + c_i, \quad b_i = \beta + d_i$$

- We must assume that the individual intercepts (a_i) and slopes (b_i) are mean independent (i.e., uncorrelated with the value of x) in order to estimate the APE.
- I.e., knowing x does not help us predict the individual's partial effect.

Random Coefficient Model [Part 4]

Implications of APE

- Be careful interpreting coefficients when you are implicitly arguing elsewhere in the paper that the effect of x varies across observations.
- Keep in mind the assumption this requires.
- And describe results using something like... "we find that on average an increase in x causes a β change in y."

Outline

Motivate why we care about causality

Three possible biases

- Omitted Variable Bias
- Measurement error bias
- Simultaneity bias

Measurement Error (ME) Bias

Estimation will have measurement error whenever we measure the variable of interest imprecisely.

- Example #1: Altman-Z-score is a noisy measure of default risk.
- Example #2: Average tax rate is a noisy measure of marginal tax rate.

Such measurement error can cause bias, and the bias can be quite complicated.

Measurement Error vs. Proxies

Measurement error is like a proxy variable but very different conceptually.

- A proxy is used for something that is entirely unobservable or unmeasurable (e.g., ability).
- With measurement error, the variable we don't observe is well-defined and can be quantified... it's just that our measure of it contains error.

ME of Dependent Variable [Part 1]

It could not a big issue (in terms of bias); sometimes just causes our standard errors to be larger.

- Example:

  $$y^* = \beta_0 + \beta_1 x_1 + u$$

  But we measure y^* with error $e = y - y^*$ Because we only observe y we estimate:

  $$y = \beta_0 + \beta_1 x_1 + (u + e)$$

ME of Dependent Variable [Part 2]

As long as $E(e|x) = 0$ the OLS estimates $\hat{\beta}_1$ are consistent and unbiased.

- I.e., as long as the measurement error of y is uncorrelated with the x's we're okay.
- Only issue is that we get larger standard errors when e and u are uncorrelated [which is what we typically assume] because $\text{Var}(u + e) > \text{Var}(u)$.

ME of Dependent Variable [Part 3]

What are some common examples of ME?

- Market leverage – typically use book value of debt because market value is hard to observe.
- Firm value – again hard to observe market value of debt so we use book value.
- CEO compensation – value of options are approximated using Black-Scholes.

Is assuming e and x are uncorrelated plausible?

ME of Dependent Variable [Part 4]

Answer = Maybe... maybe not.

- Example: Firm leverage is measured with error; hard to observe the market value of debt so we use book value.
- But the measurement error is likely to be larger when firms are in distress... Market value of debt falls; book value does not.
- This error could be correlated with x's if it includes things like profitability (i.e., ME larger for low-profit firms).
- This type of ME will cause inconsistent estimates.

ME of Independent Variable [Part 1]

- Let's assume the model is:

$$y = \beta_0 + \beta_1 x^* + u$$

  But we observe x^* with error $e = x - x^*$
- We assume that $E(y|x^*, x) = E(y|x^*)$ [i.e., x doesn't affect y after controlling for x^*; this is standard and uncontroversial because it is just stating that we have written the correct model].

ME of Independent Variable [Part 2]

There are lots of examples!

- Average Q measures marginal Q with error.
- Altman-Z score measures default probability with error.

Will this measurement error cause bias?

ME of Independent Variable [Part 3]

Answer depends crucially on what we assume about the measurement error e.

- Literature focuses on two extreme assumptions:

#1 Measurement error e is uncorrelated with the observed measure x.

#2 Measurement error e is uncorrelated with the unobserved measure x^*.

Assumption #1: e Uncorrelated with x

Substituting x^* with what we actually observe $x^* = x - e$ into the true model we have:

$$y = \beta_0 + \beta_1 x^* + u = \beta_0 + \beta_1(x - e) + u = \beta_0 + \beta_1 x + (u - \beta_1 e)$$

- Is there a bias?
- Answer = No. x is uncorrelated with e by assumption and x is uncorrelated with u by earlier assumptions.
- What happens to our standard errors?
- Answer = They get larger; error variance is now $\text{Var}(u) + \beta_1^2 \text{Var}(e)$.

Assumption #2: e Uncorrelated with x^*

We are still estimating: $y = \beta_0 + \beta_1 x^* + u$, but now x is correlated with e.

- e uncorrelated with x^* guarantees e is correlated with x: $Cov(x, e) = E(xe) = E(x^* e) + E(e^2) = \sigma_e^2$
- I.e., an independent variable will be correlated with the error... we will get biased estimates!

This is what people call the Classical Error-in-Variables (CEV) assumption.

CEV with 1 variable = Attenuation Bias

If you work out the math, you can show that the estimate of β_1 in the prior example (which had just one independent variable) is:[2]

$$plim(\hat{\beta}_1) = \beta_1 \frac{\text{var}(x^*)}{\text{var}(x^*) + \text{var}(e)}$$

- The estimate is always biased towards zero; i.e., it is an attenuation bias.
- And if the variance of error Var(e) is small, then attenuation bias won't be that bad.

[2] $plim(\hat{\beta}_1) = \beta_1 + \frac{Cov(x, u - \beta_1 e)}{Var(x)} = \beta_1 - \frac{\beta_1 \text{var}(e)}{\text{var}(x^*) + \text{var}(e)}$

Measurement Error... Not So Bad?

Under the current setup, measurement error doesn't seem so bad...

- If the error is uncorrelated with the observed x, no bias.
- If the error is uncorrelated with the unobserved x^* we get an attenuation bias... so at least the sign on our coefficient of interest is still correct.

Why is this misleading?

Nope, Measurement Error is Bad News

Truth is, measurement error is probably correlated a bit with both the observed x and unobserved x^*.

- I.e., some attenuation bias is likely.

Moreover, even in the CEV case, if there is more than one independent variable, the bias gets horribly complicated. . .

ME with more than one variable

If estimating:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

and just one of the x's is mismeasured, then...

- ALL the β's will be biased if the mismeasured variable is correlated with any other x [which presumably is true since it was included!].
- Sign and magnitude of biases will depend on all the correlations between x's; i.e., big mess!

ME Example

Fazzari, Hubbard, and Petersen (1988) is a classic example of a paper with ME problem.

- Regresses investment on Tobin's Q (it's a measure of investment opportunities) and cash.
- Finds a positive coefficient on cash; argues there must be financial constraints present.
- But Q is a noisy measure; all coefficients are biased!

Erickson and Whited (2000) argue the positive coefficient disappears if you correct the ME.

Outline

Motivate why we care about causality

Three possible biases

- Omitted Variable Bias
- Measurement error bias
- Simultaneity bias

Simultaneity Bias

This will occur whenever any of the supposedly independent variables (i.e., the x's) can be affected by changes in the y variable; e.g.:

$$y = \beta_0 + \beta_1 x + u$$
$$x = \delta_0 + \delta_1 y + \nu$$

I.e., changes in x affect y and changes in y affect x; this is the simplest case of reverse causality.

An estimate of β_1 will be biased...

Simultaneity Bias Continued...

To see why estimating $y = \beta_0 + \beta_1 x + u$ won't reveal the true β_1, solve for x:

$$\begin{aligned} x &= \delta_0 + \delta_1 y + \nu \\ &= \delta_0 + \delta_1(\beta_0 + \beta_1 x + u) + \nu \\ x &= \left(\frac{\delta_0 + \delta_1\beta_0}{1 - \delta_1\beta_1}\right) + \left(\frac{\nu}{1 - \delta_1\beta_1}\right) + \left(\frac{\delta_1}{1 - \delta_1\beta_1}\right) u \end{aligned}$$

x is correlated with u! I.e., bias!

Simultaneity Bias in Other Regressors

- Prior example is a case of reverse causality (the variable of interest is affected by y).
- But if y affects any x, there will be a bias; e.g.:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

and:

$$x_2 = \gamma_0 + \gamma_1 y + w$$

- Easy to show that x_2 is correlated with u; and there will be a bias on all coefficients.
- This is why people use lagged x's.

Simultaneity Bias – Summary

- If your x might also be affected by the y (i.e., reverse causality), you won't be able to make causal inferences using OLS.
- Instrumental variables or natural experiments will be helpful with this problem.
- Also, you can't get causal estimates with OLS if controls are affected by the y.

"Bad Controls"

- Like simultaneity bias... but this is when one x is affected by another x; e.g.:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

where:

$$x_2 = \gamma_0 + \gamma_1 x_1 + \nu$$

- Angrist-Pischke call this a "bad control," and it can introduce a subtle selection bias when working with natural experiments [we will come back to this in later lecture]

"Bad Controls" continued

- But just to preview it... If you have an x that is truly exogenous (i.e., random) [as you might have in a natural experiment], do not put in controls that are also affected by x!
- Only add controls unaffected by x or just regress your various y's on x and x alone!
- We will revisit this in a later lecture...

Summary of Today [Part 1]

- We need conditional mean independence (CMI) to make causal statements.
- CMI is violated whenever an independent variable x is correlated with the error u.
- Three main ways this can be violated:
  - Omitted variable bias
  - Measurement error bias
  - Simultaneity bias

Summary of Today [Part 2]

The biases can be very complex.

- If more than one omitted variable or the omitted variable is correlated with more than one regressor, the sign of bias is hard to determine.
- Measurement error of an independent variable can (and likely does) bias all coefficients in ways that are hard to determine.
- Simultaneity bias can also be complicated.

Summary of Today [Part 3]

To deal with these problems there are some tools we can use.

- E.g., Proxy variables [discussed today].
- We will talk about other tools later, e.g.:
  - Instrumental variables
  - Natural experiments
  - Regression discontinuity