Political Science 15
Introduction to Research in Political Science
Lecture 2a: Introduction to Causality

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett, Paasha Mahdavi, Matto Mildenberger, and Leah Stokes for select slides, used with permission.

Recap from last class

Artwork by @allison_horst

Recall our goals for the course

1. *Learn to think about when a difference is a difference.*
2. *Learn some* `R`.
3. *Learn that correlation is not causation.*

Now we're going to focus on point #3...

...time for Causality!

Why Causality?

Many (most?) interesting questions in political science are about *causes*!

- does democracy improve wealth, or health, or gender equality, or 'Y'?
- do politicians change policies because of public opinion?
- does peacekeeping or development aid reduce violence?
- does education increase political participation?
- ...what does increase political participation?
- ...and many more questions.

Many Analyses Seek to Identify Causal Effects

For example, why would a researcher care to find out whether life spans are longer in democracies than in non-democracies?

Probably, the researcher hoped the research would tell us if democracy *causes* longer life-spans.

Today I want to convince you that observing such a relationship in data often will not tell you about a causal effect.

I will convince you. But you will forget. It's too tempting.

Instead of assuming causation, you will have to be able to do three things:

1. carefully state your conclusion to avoid over-claiming;
2. say *why* a causal claim is not warranted in a given case;
3. understand what *would* allow you to make a causal claim (usually, from an experiment or some other source of *exogenous* variation).

Vitamins lead to a Long Life!

Many people believe that taking vitamins increases lifespan.

Indeed, people who take vitamins live longer!

So: are you going to go back home and take some vitamins this evening?

This is one example of a selection effect:

It looks like there is an "effect" (i.e. that vitamins increase lifespan), but really there is selection bias: a certain kind of person, who also tends to live longer (i.e. is concerned about their health, has higher income, etc.) is more likely to "select into" taking vitamins.

This often happens in observational studies.

Selection effect in practice: $90 vitamins from Goop

goop Wellness

THE MOTHER LOAD

US $90.00 / US $75.00 with subscription

Variables

Dependent variable:

- Outcome we are interested in
- Depends on other variables
- The "effect"
- Also called: response variable, outcome variable

Independent variable:

- Explains our outcome
- The intervention or treatment
- The "cause"
- Also called: explanatory variable, predictor variable

A Classic Example

How could we test the relationship between hospitals and lifespan? What is our IV and what is our DV?

- Independent variable or treatment: going to hospital
- Dependent variable or outcome: five year survival

What are we likely to find? We will likely see that going to the hospital will be associated with lower five-year survival rates.

Does that mean going to the hospital *causes* lower survival?

No!

You should certainly not write a paper or give policy advice that hospitals are bad!

Three Ways To Ponder This

1. Confounding: a confounder is something related to both the IV and the DV. *Higher alcohol consumption* makes people both have higher hospitalization rates and lower survival rates.

2. Selection problem: people go to the hospital when they are sicker.

3. Comparability (in absence of treatment). Ask yourself: how might *those who go to the hospital* and *those who do not* differ on things other than the thing you are looking at (that they are in the hospital)?

Yet more ways to think about it!

You see a relationship between independent variable A and dependent variable, B.

For example, you see countries that are democracies have higher average income.

Maybe this is just due to chance – we will learn some significance tests to deal with that!

But if it is not due to chance, three possibilities remain:

1. A (higher democracy) $\rightarrow$ B (higher income)
2. B (higher income) $\rightarrow$ A (higher democracy)
3. C (unobserved confounder) $\rightarrow$ A & B!

In general, it is easy(ish) to tell if A and B are significantly related, but very difficult to tell the difference between these causal arrangements.

Political Science 15
Introduction to Research in Political Science

Lecture 2b: Experiments and Causality

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett, Paasha Mahdavi, Matto Mildenberger, and Leah Stokes for select slides, used with permission.

Experiments

Example: In a **survey** you find that those who say they watched more political ads were more likely to politically participate. We also find it's a statistically significant relationship ($p < 0.001$, more on this later).

1. A (seeing ads) $\rightarrow$ B (participation)
2. B (participation) $\rightarrow$ A (seeing ads)
3. C (unobserved confounder) $\rightarrow$ A & B!

You *can* decide among these three if you – as the researcher – cause A or B.

Randomization: you randomly assign people to see ads (the IV):

- You could rule out reverse causality (#2)
- You can rule out any confounders (#3)
- You know there is no selection effect because you randomly assigned.
- In terms of comparability: ask yourself, in the absence of the ads, would the treated and control groups look any different? Especially on participation?

Practice:

In a **survey**, you ask people if they saw any candidate election ads and whether they voted in November's election. Among those who saw ads, 80% said they voted. Among those who did not see ads, 60% said they voted in the midterm election.

You run some statistics (a two-sample test for difference in proportions), and find these are statistically significant different groups! ($p < 0.05$). You've got yourself a causal finding...right?

1. What you can and cannot say:
   - You can note the correlation ('reject the null hypothesis') and say "voting rates are significantly higher among those who saw ads than those who did not!"
   - You cannot say "seeing political ads *increased* voting!" (causal assumption)
   - You can say, "seeing ads is *associated with* higher voting rates"

Practice: (continued)

2. Be able to articulate *why* a causal claim is not warranted:
   - e.g. Confounding: Those in more politically competitive areas (e.g. swing states like Ohio) are both more likely to see ads and more likely to participate.
   - or, Comparability: Those who do and don't tend to see ads probably differ on other characteristics, such as where they live, their income and their interest in politics.

3. Know what would be a better design.
   - Run an experiment, randomizing who gets to see political ads and who does not, then look to see if this is related to political participation.

This is Serious: Is HCL really a treatment for COVID-19?

Several observational studies recently showed that Hydroxychloroquine (HCL) *was associated with* faster recovery in COVID-19 patients, especially when taken with azithromycin (Z-pak).

- Among them, a non-randomized study of 36 COVID-19 patients in France published on March 17th found that 100% of the patients taking HCL and Z-pak were cured of the virus compared to 57.1% only treated with HCL and to 12.5% of patients who were not treated. This study, along with a separate study from China, was the basis for the FDA's approval for wide-scale emergency use of HCL as a COVID-19 treatment on March 28th.
- Dr. Rick Bright, BARDA Director at U.S. Department of Health and Human Services:

  "Based upon limited in-vitro and anecdotal clinical data in case series . . . I am authorizing the emergency use of chloroquine phosphate and hydroxychloroquine sulfate for treatment of COVID-19 when clinical trials are not available."

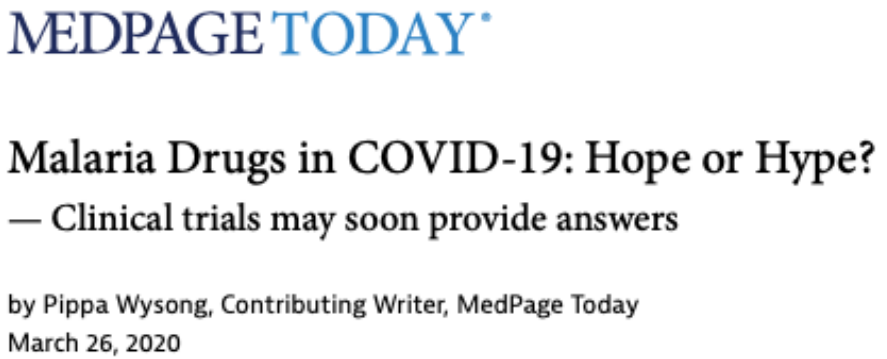
Case Study: Is HCL really a treatment for COVID-19?

Forbes

EDITORS' PICK | 105,899 views | Mar 25, 2020, 04:31am EDT

Hydroxychloroquine Use For COVID-19 Coronavirus Shows No Benefit In First Small—But Limited —Controlled Trial

Tara Haelle Senior Contributor
Healthcare
I offer straight talk on science, medicine, health and vaccines.



MEDPAGE TODAY

Malaria Drugs in COVID-19: Hope or Hype?

— Clinical trials may soon provide answers

by Pippa Wysong, Contributing Writer, MedPage Today
March 26, 2020

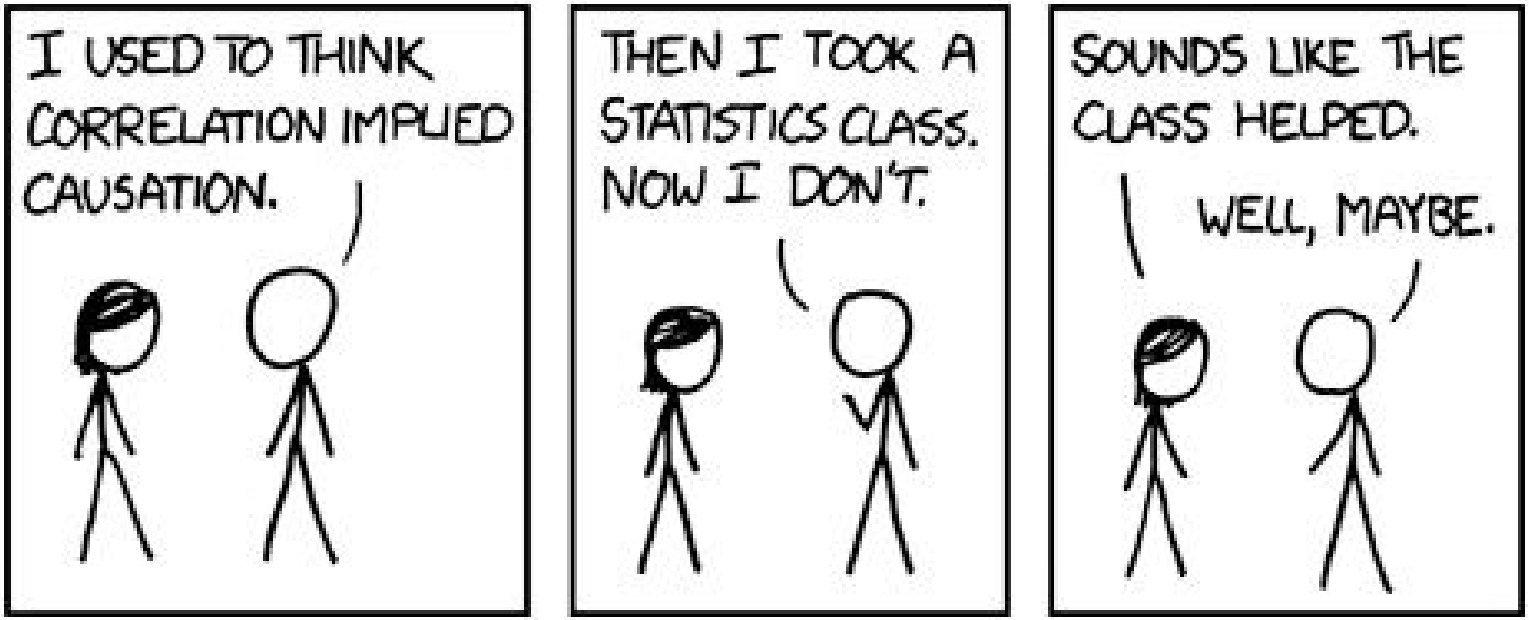
- The first randomized controlled trial (RCT - another word for experiment) of 30 COVID-19 patients in China showed that HCL was no better than the conventional treatment given to patients.
- Experiment showed no statistical difference in viral improvement time for 15 patients who were randomly selected to receive HCL compared to 15 patients who did not.
- The negative results have spurred numerous large RCTs, including two in New York and Minnesota, especially in light of the harmful potential side effects of HCL for certain patients.
- Why do you think the RCT results differ from the observational studies?

Terminology

Treated and Control:

- Treated: those we get some treatment of interest, e.g. those who get HCL, vitamins, democracy, political ads.
- Controls: are those who do not get the "treatment" of interest. e.g. not on HCL, vitamins, autocracy, no political ads.
- Observational study: general term for research where you don't get to *randomize* who gets the treatment. Instead, you just observe some relationship in the world. Check: Typically, can you make causal claims from observational studies?
- Experimental Study and Randomized Control Trial (RCT): common terms for research designs in which you do randomize who gets the treatment. Check: Typically, can you make causal claims from experimental studies?
- Quasi-experimental research: research in which you have observational data, but you find ways to ensure that the treatment was effectively ('as-if') randomly distributed.

Thought for the Day



Political Science 15
Introduction to Research in Political Science
Lecture 2c: Experiments vs. Observational Studies

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett, Paasha Mahdavi, Matto Mildenberger, and Leah Stokes for select slides, used with permission.

Job Training Program Example

Consider a job training program designed to help people get a skilled job and hopefully earn more money.

First, suppose we let people choose whether or not to take the job training program. We then do a survey and compare the incomes of those who take the program to those who do not.

- Get incomes a year later for the treated and for control group and compare means.
- Can you say from this information whether the job training program helps?
- What's a better research design?

Job Training Program: Observational Data

We will use data from LaLonde's 1986 study of a job training program.

First, the **observational** data:

- Some people chose to take a job training program offered to them. These will be our treated.
- Most people in the country did not have access to this program, so our controls come from a national survey.
- Income and other characteristics are measured for both groups.

Does the researcher control the selection process here, or did people self select? People self-selected into the treatment.

Do we think those who take the program and those who did not were comparable? Not comparable.

Can you think of possible confounders that influence both whether people take the program and their incomes? Unemployment, geography, education, motivation.

Let's see the result

You may think "okay, I'm not supposed to give a causal interpretation, but surely the data *suggests* the effect the job training might have on income... It should matter..."

Actual difference in means:

$$\overline{income}_{treated} - \overline{income}_{control} = -\$15,205$$

Why? Do you believe it is causal?

What do you think is biggest confounder or source of non-comparability?

- employment prior to the program is a good guess: 60% of the treated were unemployed before the program, but only 6% of the controls were.
- say we would say "unemployment may be a confounder", or "the treated and controls are not comparable because they differ on mean employment even in the absence of the treatment."
- how can you correctly state the observed result?

Job Training Program: Experiment

Fortunately, there was also **randomization**: those interested in taking the program were randomly assign to either get it or not.

From this experimental sample, we can compare the mean outcome for the treated to the controls.

- Given randomization, are there any confounders?
- Would you say the treated and control groups are now comparable?
- So can you give the result a causal interpretation?

We get:

$$\overline{income}_{treatment} - \overline{income}_{control} = \$1805$$

We would still want to compute statistics (a p-value and confidence interval) to handle the statistical uncertainty...

...but we do think that any difference is *caused* by the job training program.

Key Terms

Internal validity

- Is the experiment well designed? Is it free from confounders or bias? If so, we can say that the effects we find have internal validity.

External validity

- Is the finding *generalizable* to (does it apply to) other populations, situations or cases? Does it apply outside of the context in which the finding was generated? If so, we can say that the effects we find have external validity.

Political Science 15
Introduction to Research in Political Science
Lecture 2d: Limitations of Experiments

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett, Paasha Mahdavi, Matto Mildenberger, and Leah Stokes for select slides, used with permission.

Experiments solve some (but not all!) problems

Internal validity failure example: new depression drug test

- A new drug is being tested to reduce depression. In addition to providing the treated group with the drug, they are also interviewed weekly by a psychiatrist to monitor their progress. The control group is not. This introduces a confounder: if we see a change, is it the medicine, or the interview, that caused the change?

External validity failure example: new depression drug test

- A new drug is being tested to reduce depression. But when the researchers recruit participants, they exclude people who: drink alcohol, have anxiety, take other medication, or only have mild depression because it will be harder to control these factors in the study. Therefore, the findings will be unlikely to generalize to the population at large.

Some more problems with experiments

Experiments (randomization) solve lots of problems. But:

- Not everything can be randomized! (e.g. democracy, gender)
- Not everything *should* be randomized (e.g. wars, medications that may cause birth defects, the right to vote).
- Sometimes it is an ethical dilemma: randomization means denying treatment to some (or exposing some to risk), but not randomizing means we don't really know if something works or is harmful.
- Running experiments is expensive and difficult.
- There are interesting things we can do without making causal inferences, such as prediction tasks and describing things as they exist.

Takeaways

Remember the HCL and Job Training examples when you encounter new observational findings.

We will touch on some of the methods we use to make observational studies a little bit better, usually by making the treated and controls more similar on things not effected by the treatment.

For now, develop the instinct to approach observational results with skepticism, then:

1. interpret correctly and safely
2. be able to say why you are skeptical: potential confounders or reasons for non-comparability
3. in your head, attempt to come up with an 'ideal experiment' (even if you couldn't implement it)

Political Science 15
Introduction to Research in Political Science
Lecture 2e: Formalizing our Model

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett, Paasha Mahdavi, Matto Mildenberger, and Leah Stokes for select slides, used with permission.

Recap from Week 1

Your to-do list for Week 1

- ✓ Read the syllabus (yes, the whole thing)
- ✓ Attend section
- ☐ Read Chapter 1 of *Real Stats*
- ☐ Watch video modules (watch 90% by Sunday before midnight for participation points)
- ☐ Fill out Week 1 survey (do this or you risk being dropped)
- ☐ Fill out "Find a study buddy system" on GauchoSpace (if you'd like)
- ☐ Hand in Problem Set 1 (due Friday before midnight)

Recap from Week 1

What you should know (or be on your way to knowing)

- Correlation is not causation (course mantra!)
- Why an observed relationship might not be causal
  - Confounding
  - Selection bias
  - Non-comparability
- How experiments work
  - Random assignment of treatment
  - Treated and control groups
- Experiments versus observational studies
- Limitations of experiments
- Strategy for analyzing results
  1. Carefully state your results without over-claiming
  2. Articulate why or why not a causal claim is warranted
  3. Say what *would* allow you to make a causal claim otherwise

The most important equation of the course

Bivariate regression equation

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

This is the most important equation of the course. The equation above is sometimes referred to as the regression model, the regression equation, or the bivariate regression equation.

It is a **stylized representation of the world** where X affects Y in linear manner.

Famous aphorism in statistics

"All models are wrong, but some are useful" – George Box

Bivariate regression equation

Y_i is the **dependent variable**, or outcome of interest

X_i is the **independent variable**, i.e. a treatment or explanatory factor

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

ϵ_i is the **error term**, or everything that we haven't captured in our model. We cannot observe this!

β_0 is the **intercept/constant**, or the value of Y when X is 0

β_1 is the **slope**, or how much change in Y is associated with a one-unit change in X

Applied to our section example

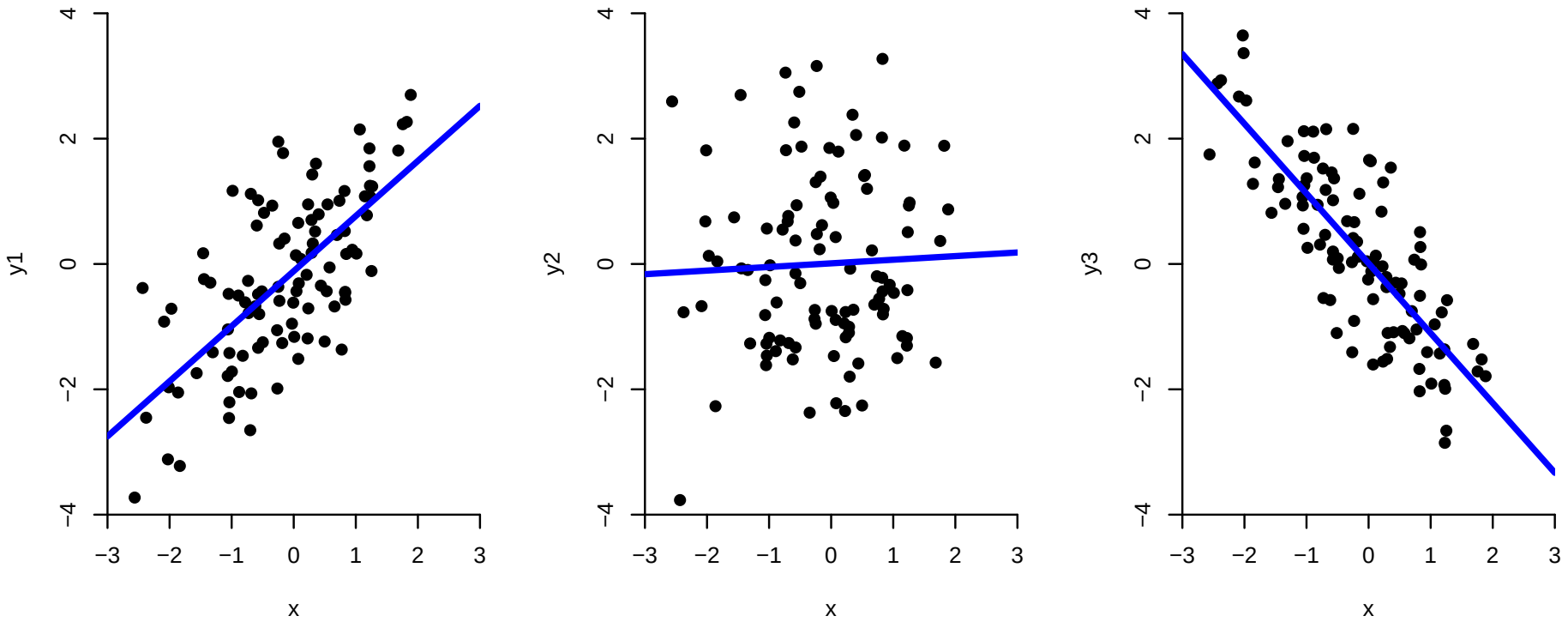
In our first section, the bivariate regression model could be written as:

$$MathScore_i = \beta_0 + \beta_1 Absences_i + \epsilon_i$$

	school	gender	math_score	language_score	absences
1	1	M	56.4	44.1	4
2	1	M	36.3	48.2	4
3	1	F	31.9	42.7	2
4	1	F	28.9	42.2	3
5	1	F	6.3	28.6	3
6	1	F	62.6	48.5	13
7	1	F	58.3	38.6	11
8	1	M	9.3	16.2	7
9	1	M	51.2	52.0	10
10	1	M	48.3	43.8	9
11	1	F	40.0	36.7	4
12	1	M	32.6	12.8	5
13	1	F	62.9	63.3	5
14	1	F	67.3	45.7	3
15	1	M	23.9	26.7	1

Visualizing this model

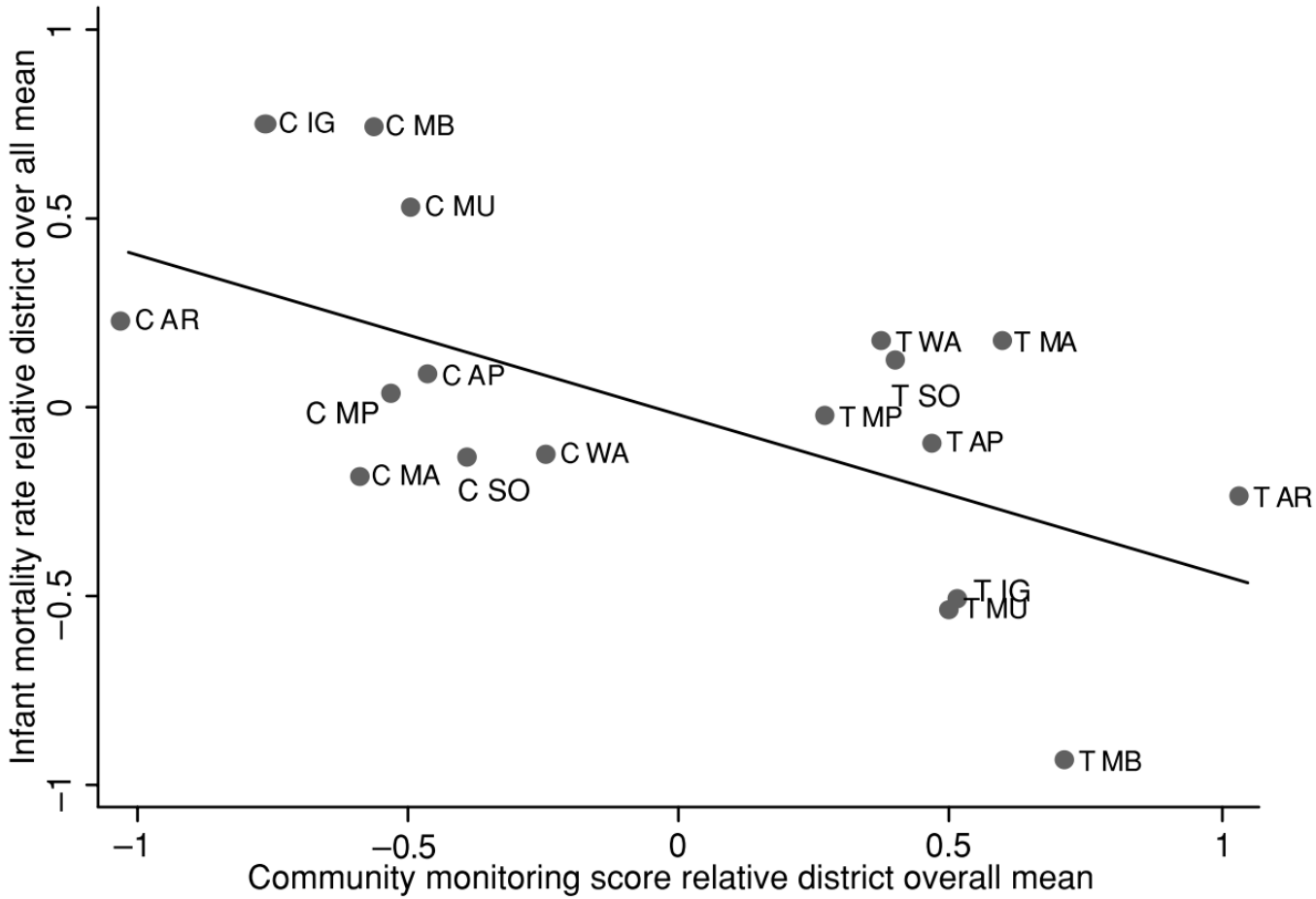
Scatterplots with regression lines:



- The (estimated) β_0 will determine the value of Y when X is 0
- The (estimated) β_1 will determine how shallow/steep the blue line is

Visualizing this model with real data

Scatterplot with regression line (real data):



Source: Bjōrkman, M., & Svensson, J. (2009). Power to the People: Evidence From a Randomized Field Experiment on Community-Based Monitoring in Uganda.

Endogeneity

Key point

Endogeneity is a fundamental problem in statistics. This is another way of saying correlation $\neq$ causation.

- An independent variable is endogenous if it is correlated with the error term in the model.
- That is, X is endogenous if changes in it are associated to factors captured by ϵ.
- An independent variable is exogenous if it is not associated with factors captured in the error term.

Endogeneity and confounders

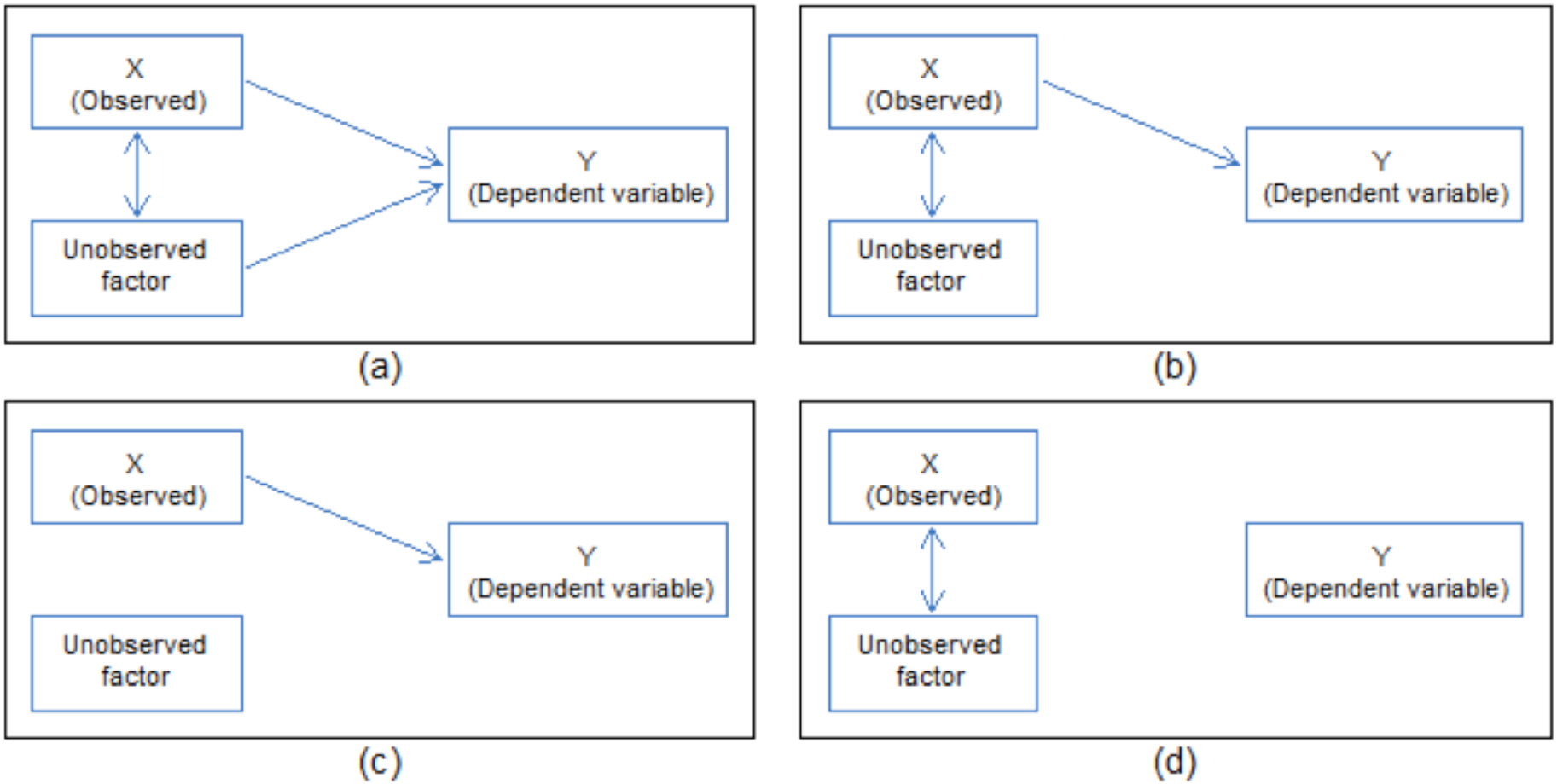
An example of endogeneity is a confounder, i.e. a variable that affects both the IV and the DV.

- If you don't include the confounder in your model, then it will live in the error term ϵ.
- And since the confounder is correlated with the IV (by definition), then the IV will be correlated with the error term $\Rightarrow$ endogeneity!
- Consequently, if you called the observed correlation between the IV and the DV "causal", this might be spurious, because it is really the confounder that is driving the results.

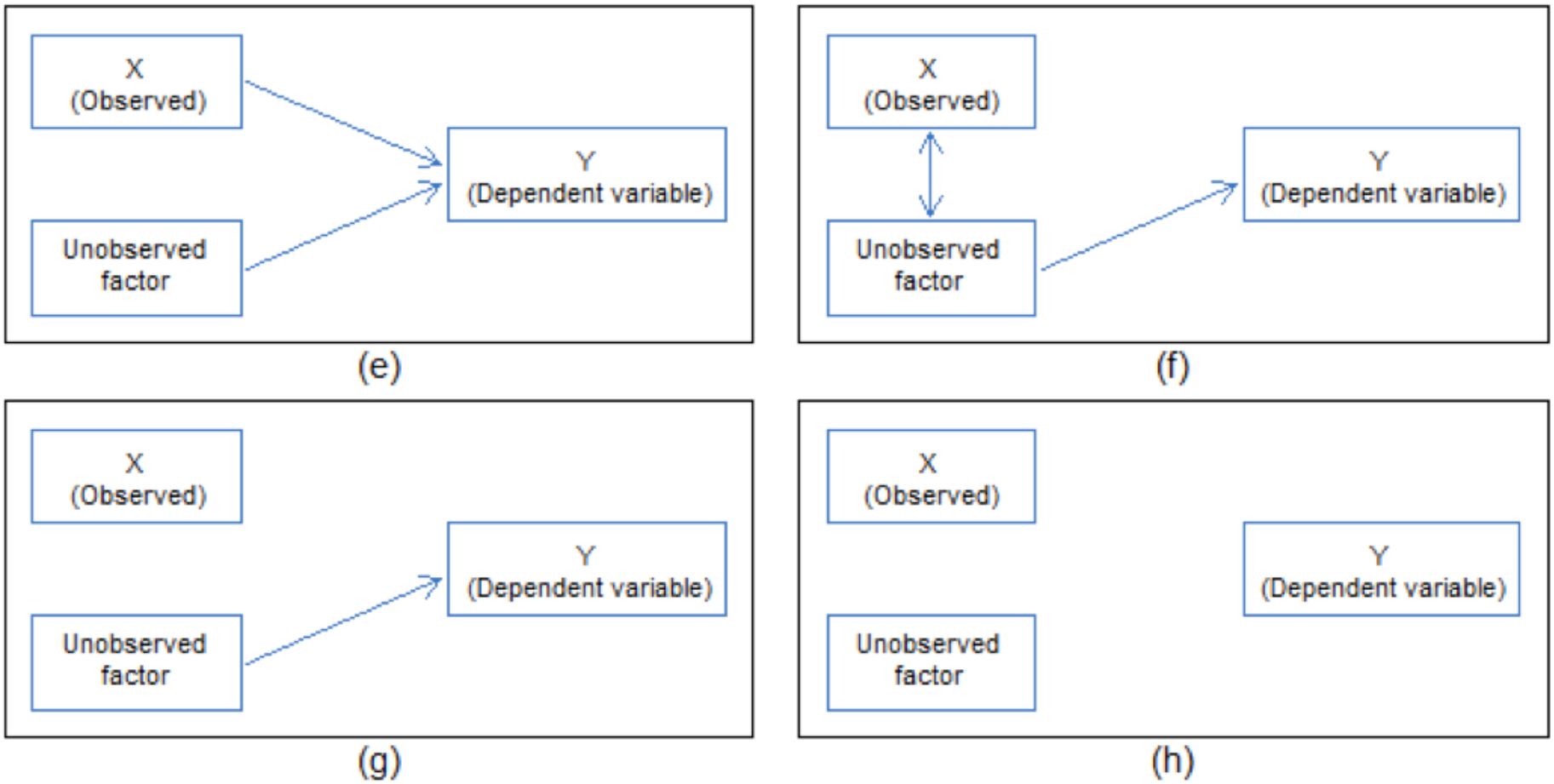
How to create exogeneity

- Remember that **the error term is unobservable**, so it is hard to know for sure if an independent variable is endogenous or exogenous.
- Why is this such a problem? Because it is difficult (but not impossible) to assess causality for endogenous independent variables.
- One way to construct an exogenous IV is using random assignment.

Different configurations of endogeneity/exogeneity



Different configurations of endogeneity/exogeneity



Randomness vs. Randomization

- The textbook uses both terms. It's a bit confusing.
- Important: These two terms do not mean the same thing.
  - **Randomness**: noise in the data (i.e. effects from sampling on a given day) that could go away with larger sample sizes. We will address some of these concerns with t-tests, p-values, and confidence intervals.
  - **Randomization** (aka random assignment): using a coin toss to create treatment and control groups; creates exogeneity.

Next Week

- Intro to probability
  - Read *Real Stats* Appendix A-G
- Intro to bivariate regression (OLS)
  - Read *Real Stats* Chapter 3
- Why exploring your data and having replicable analysis matters!
  - Read a cautionary tale in *Real Stats* Chapter 2