

Workshop “Causal Graphs”

InFER, Goethe-Universität Frankfurt am Main
Julian Schuessler

Post-Doc, Centre for the Experimental-Philosophical Study of Discrimination,
Department of Political Science, Aarhus Universitet

December 18, 2024

Section 1

Intro

Modus operandi

- ▶ I'm Julian, pol sci post-doc @ Aarhus U
 - ▶ Causal inference methods, experiments, discrimination, public opinion
- ▶ Pen & paper & (some) R/Stata
- ▶ *Ask questions.*
- ▶ *There are no stupid questions.*

Course Overview

- ▶ DAGs, DAGs, DAGs
- ▶ Applied to problems of relevance to social scientists
- ▶ You will relearn most of what you know about statistics
- ▶ Mostly “qualitative”, theoretical understanding
- ▶ I.e., you will be able to solve problems with data by *thinking*, not by *tinkering*
- ▶ Some implementation in R, STATA, esp. wrt. sensitivity analysis

What is this course is not

- ▶ Ready-made, but ultimately black-box models
- ▶ “Here are 5 variations of a multilevel structural equation model with triple cross-level interactions”

Section 2

Introductions

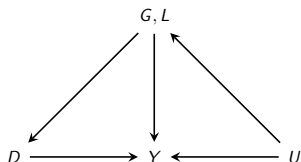
Section 3

A Running Example

Running Example

- ▶ Darfur (Sudan): Mass violence against civilians in 2003/04, killing an estimated 200,000
- ▶ Indictments of genocide and other crimes in the International Criminal Court
- ▶ Does “violence beget violence”? Or do individual experiences of violence lead people to demand peace?
- ▶ *What is the causal effect of experiencing violence on peace attitudes?* (Hazlett 2020)
- ▶ Suppose we study individuals. Draw a DAG for this research situation, with variables: exposure to violence D , gender G , village/location L , attitudes towards peace Y , other unobserved causes of attitudes towards peace U

A Possible Causal Graph for Hazlett 2020



- ▶ What is the research question?
- ▶ What are the crucial assumptions in this DAG?
- ▶ Can I test my assumptions?
- ▶ Given a research question, can I answer it? What kind of estimation strategy and which control variables do I need to use?

Section 4

Causal and Non-Causal Questions

Causal & Non-Causal Questions

- ▶ Causal question: Effect of violence on attitudes

Hazlett 2020: Basic Result

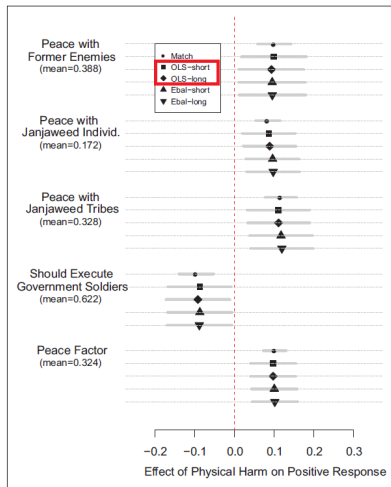
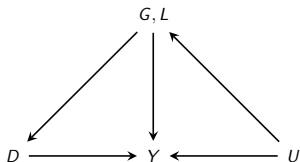


Figure: Fig. 2 from Hazlett 2020.

Causal & Non-Causal Questions

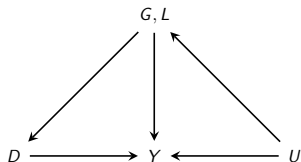


- ▶ Causal question: Effect of violence on attitudes
- ▶ Exercise: Based on the three variables gender, violence, attitudes, what other questions could one ask? Are they non-causal or causal?

Causal & Non-Causal Questions

- ▶ Causal question: Effect of violence on attitudes
- ▶ Non-causal question: Correlation/dependence between violence and attitudes
- ▶ Non-causal question: Dependence between violence and attitudes, controlling for (conditioning on, given the same) gender
- ▶ Causal question: Effect of violence on attitudes among women (conditional/subgroup effect/effect heterogeneity)
- ▶ Causal question: Joint effect of gender & violence on attitudes (causal interaction)
- ▶ Causal question: Effect of gender on attitudes through violence (mediation)

Roadmap



- ▶ What is the research question? ✓
- ▶ What are the crucial assumptions in this DAG?
- ▶ Can I test my assumptions?
- ▶ Given a research question, can I answer it? What kind of estimation strategy and which control variables do I need to use?

Section 5

DAG Basics

Directed Acyclic Graphs

- ▶ Directed acyclic graphs (DAGs)
- ▶ Directed: Every connection has a direction (no simple lines, no arrows going both ways)
- ▶ Acyclic: No cycles in the graph - no “mutual causality”, “feedback loops”
- ▶ First used by biologist Sewall Wright in the 1920s, important for traditional structural equation modelling in the 60s–80s, resurgence due to work by computer scientist Judea Pearl in 1995
- ▶ Popular framework for “working with” causality in machine learning/AI, statistics, political science/sociology...

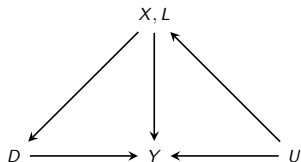
Causal DAGs

- ▶ Causal DAGs visualize causal assumptions
- ▶ Important assumptions are the *arrows left out*
 - ▶ No arrow = assume that no such causal effect exists, period
- ▶ Drawing an arrow just implies that there *might* be a causal effect
- ▶ Bad assumptions in, bad results out
- ▶ Good assumptions in, good results out
- ▶ Causal inference without assumptions is impossible

Cycles

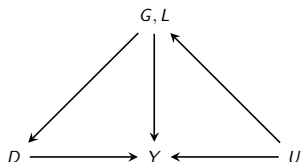
- ▶ We usually think that causality takes time
- ▶ A cyclic graph $D \leftrightarrow Y$ is short-hand for $D_1 \rightarrow Y_1 \rightarrow D_2 \rightarrow Y_2 \dots$
- ▶ When you think you have cycles, the issue is usually that you don't have sufficiently fine-grained data over time
- ▶ There is not much you can learn about causality when you have cycles
- ▶ This course assumes you do not have cycles

Reading Assumptions from a DAG



- ▶ What are the assumptions in this DAG?
 - ▶ No edge from U to D
 - ▶ No edge from Y to U , D to X , D to U ... (cycles)
- ▶ Indeed, all variables save D and U are already directly connected

Roadmap



- ▶ What is the research question? ✓
- ▶ What are the crucial assumptions in this DAG? ✓
- ▶ Can I test my assumptions?
- ▶ Given a research question, can I answer it? What kind of estimation strategy and which control variables do I need to use?

Section 6

Testing assumptions: d-separation

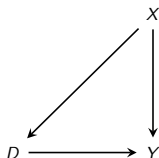
Introduction

- ▶ Any observed correlation between variables must be due to *some* causal connection
- ▶ In graph terms, there must be an *open path* between the variables that creates the correlation
- ▶ But depending on the graph, there may also be situations where there is no such open path
- ▶ Then, we say that the variables are “d-separated”, and hence uncorrelated
- ▶ This zero correlation would be something we could check in the data \implies test of assumptions/DAG
- ▶ To check this, we need to understand when a path is “open” or “blocked”
- ▶ This is also important for understanding graphs more broadly, e.g. for causal inference

Strategy

- ▶ Any DAG can be split up into direct effects $A \rightarrow B$, which (probably) create a dependency between A and B
- ▶ As well as three different three-variable combinations that are chained together
- ▶ If one understands how these three combinations create correlations, one can analyze any larger graph based on this

Splitting up a DAG into Paths



- ▶ A path is any connection between two “nodes” on the graph, irrespective of the direction of arrows, without going over the same node twice
- ▶ Which paths are in this graph that start from D or X ?
- ▶ $D \rightarrow Y$ and $D \leftarrow X$
- ▶ $D \leftarrow X \rightarrow Y$ and $D \rightarrow Y \leftarrow X$
- ▶ $X \rightarrow D$ and $X \rightarrow Y$
- ▶ $X \rightarrow D \rightarrow Y$ and $X \rightarrow Y \leftarrow D$
- ▶ (Plus others that start in Y)

Path Type I

$$D \longrightarrow M \longrightarrow Y$$

- ▶ In this graph, do D and Y correlate?
 - ▶ Yes
- ▶ Do D and Y correlate when I control for / condition on M ?
 - ▶ No
- ▶ The path is *open*. Conditional on M , it is *blocked*
- ▶ Simulation in R:

```
D <- rnorm(1000)
M <- 0.4*D + rnorm(1000)
Y <- -0.6*M + rnorm(1000)
lm(Y ~ D)
lm(Y ~ D + M)
```

Path Type I

$$D \longrightarrow M \longrightarrow Y$$

- ▶ In this graph, do D and Y correlate?
 - ▶ Yes
- ▶ Do D and Y correlate when I control for / condition on M ?
 - ▶ No
- ▶ Examples:
 - ▶ Kick football \rightarrow broken window \rightarrow inhabitant angry
 - ▶ Settler mortality \rightarrow Property rights \rightarrow GDP (AJR)
 - ▶ Drug \rightarrow activity of neurotransmitters \rightarrow Subjective experience

Path Type II

$$D \longleftarrow M \longrightarrow Y$$

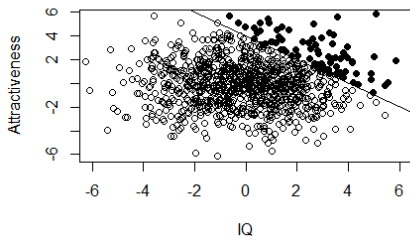
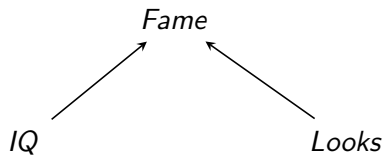
- ▶ In this graph, do D and Y correlate?
 - ▶ Yes
- ▶ Do D and Y correlate when I control for / condition on M ?
 - ▶ No
- ▶ The path is *open*. Conditional on M , it is *blocked*
- ▶ Example: Ice cream sales \leftarrow Season \rightarrow pool drownings

Path Type III

$$D \longrightarrow M \longleftarrow Y$$

- ▶ In this graph, do D and Y correlate?
 - ▶ No
- ▶ Do D and Y correlate when I control for / condition on M ?
 - ▶ Yes
- ▶ The path is *blocked*. Conditional on M , it is *open*
- ▶ M acts as a *collider*

Collider: Example 1



Collider: Example 2

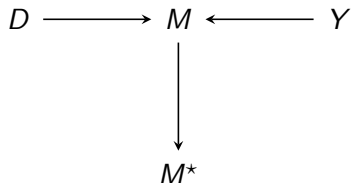


- ▶ If study is newsworthy and published in *Science*...
- ▶ ... it is probably less rigorous

Colliders: More examples

- ▶ Eye color of father \rightarrow eye color of child \leftarrow eye color of mother
- ▶ Lawn sprinkler with timer \rightarrow Wet grass \leftarrow Rain

Descendants of Colliders



- ▶ M^* is a “descendant” of M
- ▶ Acts as a “proxy”, therefore, conditioning on M^* has same qualitative consequences as conditioning on M - opens up path
- ▶ Especially clear when $M^* = M$ - then they contain the exact same information

Blocking paths: Summary

- ▶ Chain of mediation: Path is open unconditionally, but blocked conditional on the middle node. $D \not\perp\!\!\!\perp Y$ but $D \perp\!\!\!\perp Y|M$.
- ▶ Common cause/fork: Path is open unconditionally, but blocked conditional on the middle node. $D \not\perp\!\!\!\perp Y$ but $D \perp\!\!\!\perp Y|M$.
- ▶ Collider: Path is blocked unconditionally, but open conditional on the middle node or one of its descendants. $D \perp\!\!\!\perp Y$ but $D \not\perp\!\!\!\perp Y|M$.

d-separation: Definition

- ▶ If M blocks every path between two nodes D and Y , then D and Y are **d-separated**, conditional on M , and thus are independent conditional on M
- ▶ **testable implication** of the graph
- ▶ “d-separation” = “directional separation” (in directed graphs)
- ▶ Path p may be very long, but as long as you block sub-path, you block the whole path
- ▶ If testable implication does not hold, something about the graph is wrong
- ▶ Note: Statement about graph \implies statement about data

Blocking long paths

- ▶ $A \rightarrow B \leftarrow C \rightarrow D$ is blocked since B acts as a collider
- ▶ Conditional on B , it is open
- ▶ Conditional on B and C , it is blocked (because C acts as a “mediator”)

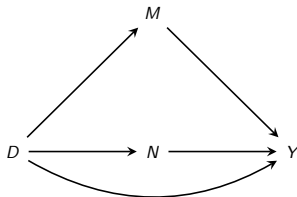
d-separation: Recipe

- ▶ Write down all paths between two variables of interest D , Y
- ▶ Check whether they are open or blocked, perhaps conditional on some control variables M
- ▶ If all are blocked: D and Y are d-separated (and therefore independent/uncorrelated), when controlling for M
- ▶ If at least one path is open: D and Y will probably correlate (controlling for M)

d-separation: Practice & DAGitty

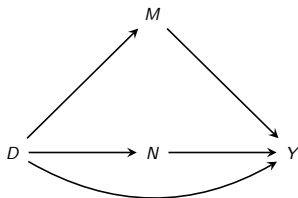
- ▶ Needs practice
- ▶ Automated: <http://dagitty.net/>
- ▶ Also R package dagitty

Exercise: d-separation



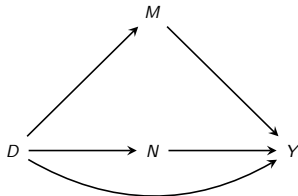
- ▶ This graph makes many strong assumptions. Are they testable?
- ▶ That is, is there a regression you could run using some (or all) of the variables to show that this graph is wrong?
- ▶ Yes: There is (exactly) one pair of variables that is d-separated conditional on other variable(s)
- ▶ To show d-separation, enumerate all paths between two variables and specify how they would be blocked
- ▶ Take a sheet of paper & 5–10 minutes

Exercise: d-separation



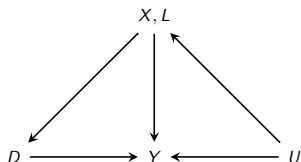
- ▶ Almost all pairs of variables are directly connected \implies will likely correlate
- ▶ Except for M and N
- ▶ Connected through D and Y
- ▶ Y is always a collider on these paths, e.g. $M \rightarrow Y \leftarrow N$
- ▶ Only open path is $M \leftarrow D \rightarrow N$
- ▶ D is “confounder”, controlling for D blocks the path, does not open other paths
- ▶ So M and N should be independent, given D
- ▶ E.g., $1m(M \sim N + D)$

Exercise: d-separation



- Exercise: Confirm this in DAGitty

Roadmap



- ▶ What is the research question? ✓
- ▶ What are the crucial assumptions in this DAG? ✓
- ▶ Can I test my assumptions? ✓
- ▶ Given a research question, can I answer it? What kind of estimation strategy and which control variables do I need to use?

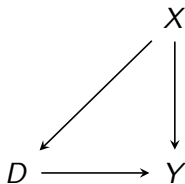
Section 7

Back-Door Criterion

From d-separation to Identification

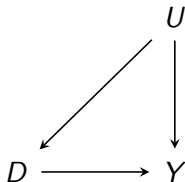
- ▶ We now know which graphs create (non-)correlations/dependencies
- ▶ Can we use our “path”-knowledge to determine whether we can answer causal and non-causal research questions?
- ▶ How can we use this knowledge to determine valid control variables?

Intuition



- ▶ Which paths does the association between D and Y consist of?
- ▶ 1) causal effect of D on Y and 2) confounding due to X
- ▶ If we could randomize D , $X \rightarrow D$ would be deleted
- ▶ If you cannot do this, find control variables such that
 - ▶ “Bad”, “spurious”, “non-causal” paths between D and Y are blocked
 - ▶ All “causal” paths are left open
 - ▶ No new “non-causal” paths are opened up (colliders...)

Intuition



- ▶ If you cannot randomize, find control variables such that
 - ▶ “Bad”, “spurious”, “non-causal” paths between D and Y are blocked
 - ▶ All “causal” paths are left open
 - ▶ No new “non-causal” paths are opened up (colliders...)
- ▶ This is *unrelated* to d-separation: d-separation is for testing graphs; and if two variables are d-separated, by definition all paths between them are blocked
- ▶ But for identifying causal effects, we certainly want to leave certain paths open (although we also want to block *some*)

The Back-Door Criterion

- ▶ We are interested in the effect of D on Y . In a DAG, a set of variables X satisfies the backdoor criterion for this effect if
 - 1) no node in X is influenced by D , and
 - 2) X blocks every path between D and Y that contains an arrow into D
- ▶ A path that starts with an arrow into D is called a **back-door path**
- ▶ Blocking back-door paths makes sure we block “bad” paths
- ▶ Not conditioning on effects/descendants of D makes sure we leave all “good” causal paths open and that we do not open up new bad paths
- ▶ Holds for any DAG \implies non-parametric, distribution-free

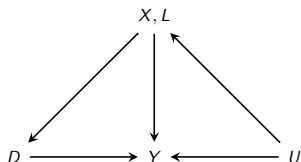
BDC: Notes

- ▶ If observed controls fulfill BDC for effect of interest, we say effect of interest is *identified*
- ▶ That is, causal effect is measurable in principle given our assumptions
- ▶ If X fulfills BDC wrt effect of D on Y , this implies counterfactual ignorability: $Y(d) \perp\!\!\!\perp D | X$

BDC: Recipe

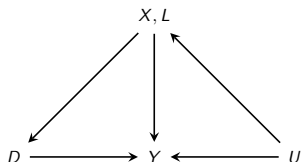
- ▶ Write down all back-door paths between D and Y
- ▶ Check whether they can be blocked by control variables X
- ▶ Make sure no control variable in X is influenced by D

Roadmap



- ▶ What is the research question? ✓
- ▶ What are the crucial assumptions in this DAG? ✓
- ▶ Can I test my assumptions? ✓
- ▶ Given a research question, can I answer it? What kind of estimation strategy and which control variables do I need to use? ✓
 - ▶ Answered for standard adjustment strategies

Roadmap



- ▶ More questions:
 - ▶ What is post-treatment bias?
 - ▶ How can I quantitatively assess the sensitivity of causal effect estimates?
 - ▶ What is the difference between effect heterogeneity and causal interaction?
 - ▶ What can DAGs tell us about causal mediation?
 - ▶ Why are social network studies generically confounded?

Section 8

Post-Treatment Bias

Survey

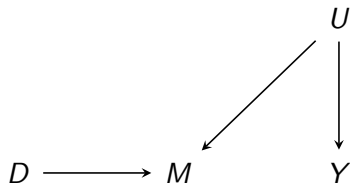
- ▶ “When I am interested in estimating a causal effect of a treatment on an outcome, I should control for variables correlated with the treatment and outcome. Otherwise, I get omitted variable-bias.”

Post-Treatment Variables: Problem 1

$$D \longrightarrow M \longrightarrow Y$$

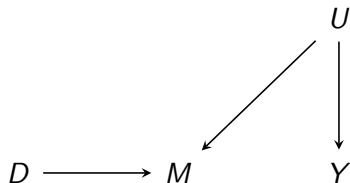
- ▶ Which set of variables in this graph satisfy the BDC wrt effect of D on Y ?
- ▶ The empty set \emptyset - no controls necessary
- ▶ Here, correlation is causation
- ▶ No paths into D - as if we intervened on it
- ▶ Does M correlate with D and Y ?
- ▶ “ M correlates with D and Y . I’ve learned in stats that I need to control for it. Otherwise, I have omitted-variable bias”
- ▶ Bad idea: Conditional on M , D and Y are d-separated! Even though D may have an effect on Y
- ▶ Montgomery et al. 2018 AJPS estimate that 50 % of political science experiments do this. Huge problem.

Post-Treatment Variables: Problem 2



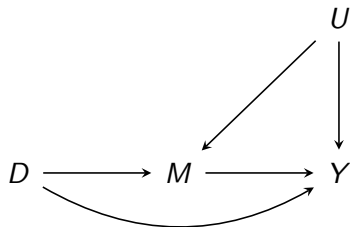
- ▶ It gets worse. Which set of variables in this graph satisfy the BDC wrt effect of D on Y ?
- ▶ The empty set - no controls necessary
- ▶ Also, no causal effect of D on Y !

Post-Treatment Variables: Problem 2



- ▶ “ M correlates with D and Y . I’ve learned in stats that I need to control for it. Otherwise, I have omitted-variable bias”
- ▶ Bad idea: Conditional on M , D and Y are d-connected! Collider!

Post-Treatment Variables: General Case



- ▶ This graph applies to situations where there are no back-door paths into D . Perhaps via randomization, or you block them by conditioning on X (not shown).
- ▶ Conditioning on M is forbidden by the BDC and will have two consequences:
 - ▶ 1. You block a causal path, which you do not want
 - ▶ 2. You open up a non-causal path, which you do not want
- ▶ This introduces bias, and it can go in any direction

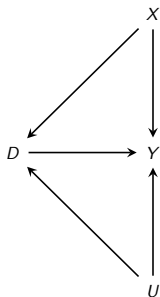
Post-Treatment Variables: Remarks

- ▶ Although clear using causal graphs, the fact that conditioning on the descendants of the treatment may actually introduce bias is not well-known
- ▶ Usually not mentioned in textbooks that do not use causal graphs
- ▶ Even if mentioned, not really explained (see for example “Mostly Harmless Econometrics”, section on “Bad Control”)

Section 9

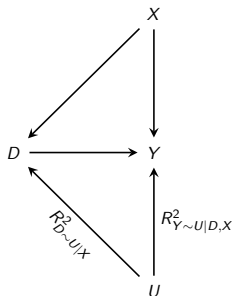
Sensitivity Analysis for Unobserved
Confounding: sensemakr

Another Possible Causal Graph for Hazlett 2020



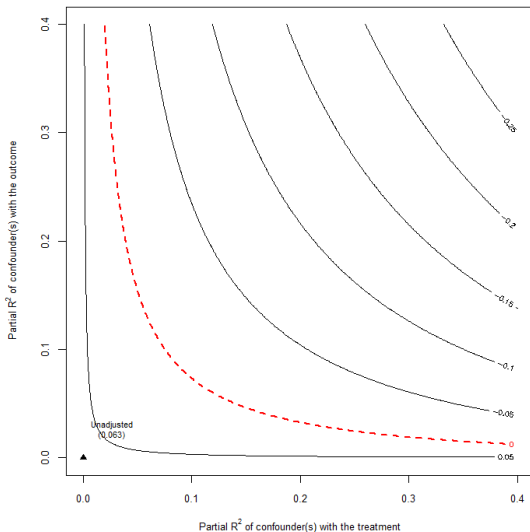
- ▶ violence D
- ▶ attitudes Y
- ▶ gender, village X
- ▶ Unobserved U ?

Another Possible Causal Graph for Hazlett 2002



- ▶ violence D
- ▶ attitudes Y
- ▶ gender, village X
- ▶ Unobserved U ?

Sensitivity Analysis via sensemakr (Cinelli/Hazlett)

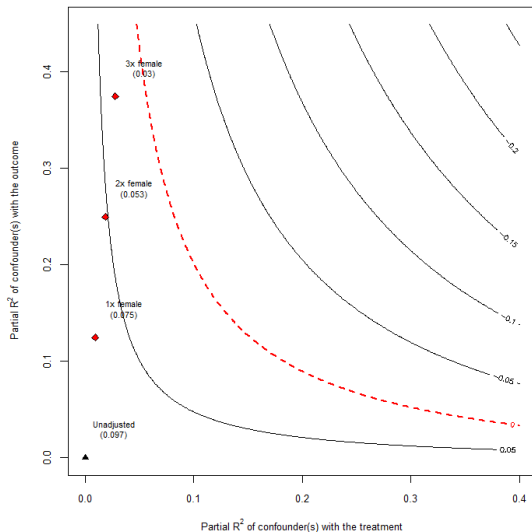


sensemakr: Exercise

- ▶ Add female and village as control variables
- ▶ Use female for “benchmarking”
- ▶ Plot the results

```
sensitivity.2 <- sensemakr(model.2,  
  treatment = "d", benchmark_covariates = "x", kd =  
  1:3)  
plot(sensitivity.2)
```
- ▶ What does the plot tell us?

sensemakr: Benchmarking



sensemakr: Comments

- ▶ Approach by `sensemakr` relies on assumption of linear “target” regression that includes U
- ▶ But, U may contain many, many variables that impact on D and Y in complicated ways
- ▶ Sensitivity of significance tests (t-values) by using `plot(sensitivity.1, sensitivity.of = "t-value")`
- ▶ Exercise: Plot sensitivity of t-values, including benchmarks, and interpret

Section 10

Interim Summary

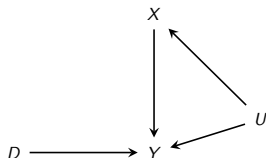
Summary

- ▶ Statistical control for additional variables may make a huge difference
- ▶ W/o a clearly articulated question and assumptions impossible to justify whether additional control is “good” or “bad”
- ▶ Causal graphs visualize causal assumptions
- ▶ Causal assumptions imply certain (non-)correlations via d-separation
- ▶ Given a causal question & DAG, can tell what kind of control is (not) necessary via back-door criterion
- ▶ Danger of post-treatment bias
- ▶ If unobserved confounding suspected: Sensitivity analysis

Section 11

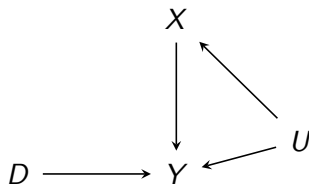
Effect Heterogeneity & Causal Interaction

Effect Heterogeneity



- ▶ Messages D , socio-economic characteristics X , turnout Y (Imai/Strauss 2011)
- ▶ Limited budget for messages D , to which people (X) should you send it as to maximize turnout?
- ▶ Depends on effect heterogeneity: X “determines” size of effect of D on Y
- ▶ Can we estimate this without bias?
- ▶ Yes: If X fulfills BDC for effect of D on Y , X -specific effect also identified

Causal Interaction



- ▶ Messages D , socio-economic characteristics X , turnout Y (Imai/Strauss 2011)
- ▶ Different Q: How do D and X causally interact in determining turnout?
- ▶ Can we estimate this without bias?
- ▶ BDC for “joint effect” of D, X on Y not fulfilled: Back-door path $X \leftarrow U \rightarrow Y$ cannot be blocked

Effect Heterogeneity vs. Causal Interaction

- ▶ Two different causal questions
- ▶ The statistical model may be the same for both questions, e.g. $Y = \beta_1 D + \beta_2 X + \beta_3 DX + \epsilon$
- ▶ But it may only be able to answer one of the questions in any given situation
- ▶ Given graph, which question would $Y = \beta_1 D + \beta_2 X + \beta_3 DX + \beta_4 U + \epsilon$ answer?
- ▶ Causal interaction – only! Control for U would generally change estimate of D – X interaction!
- ▶ Just like a joint experiment on D and X would answer causal interaction question, not effect heterogeneity question
- ▶ Intro survey: “If we could, we would always want to randomize variables in order to answer scientific and policy questions”
- ▶ No...

Section 12

Causal Mechanisms: Mediation Analysis

Survey

- ▶ “The best way to analyse direct and indirect effects of a variable is to randomize treatment and mediator.”
- ▶ “When I measure all mediators and all variables that impact on two or more of them (“confounders”), I can statistically disentangle the effect of an independent variable D on a dependent variable Y into different indirect effects”

Substantive Examples for (in)direct effects

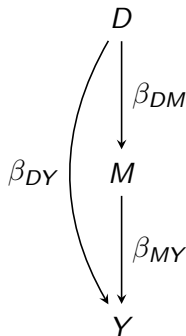
- ▶ Do macroeconomic conditions affect the vote for the incumbent mostly through individual evaluations of the economy?
- ▶ Does the incumbency effect exist because strong incumbents scare off high-quality challengers?
- ▶ Do PR systems redistribute more because of different coalition dynamics?
- ▶ Are hiring processes discriminatory; i.e., is there a direct effect of socio-economic background/gender/race...on the probability to receive a job?
- ▶ Do some genes cause lung cancer only through their effect on smoking behaviour?
- ▶ Does Cognitive Behavioral Therapy only work because it leads people to use anti-depressants more often?

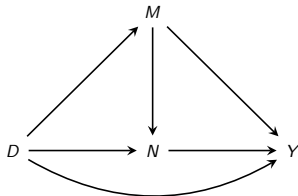
History of (in)direct effects

- ▶ Pretty clear in linear structural models
- ▶ Generalization of direct and indirect effects in Pearl 2001
- ▶ Followed by increased interest in statistics, epidemiology, sociology, political science
- ▶ E.g., Imai et al. 2010ff. implementation in `mediation` package

Direct and indirect effects in linear models

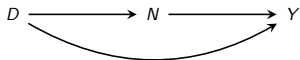
- ▶ What is the direct, what is the indirect effect of D on Y in this model?
- ▶ Direct: β_{DY} , indirect: $\beta_{DM}\beta_{MY}$
- ▶ Linear models allow for easy estimation strategies using series of linear regressions
- ▶ But many things are nonlinear...





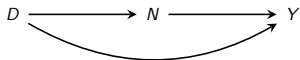
- ▶ randomized treatment D
- ▶ cost/benefit beliefs M
- ▶ anxiety N
- ▶ opposition to immigration Y

Simplified Mediation Graph



- ▶ randomized treatment D
- ▶ (cost/benefit beliefs M)
- ▶ anxiety N
- ▶ opposition to immigration Y

Estimation in the Easiest Case



- ▶ **If this graph is correct and you stick to linear regressions...**
- ▶ Regression of Y on D : ATE of D
- ▶ Regression of Y on D and N : Direct effect of D
- ▶ ATE - Direct Effect = Indirect effect
- ▶ Or: Regression of N on D \times Regression of Y on N and D = IE
- ▶ “Kenny/Baron” approach

Differencing Approach

Table:

	<i>Dependent variable:</i>	
	Opposition to Immigration (1-4)	
	(1)	(2)
tone_eth	0.439*** (0.134)	0.161 (0.116)
emo		0.188*** (0.018)
Constant	2.914*** (0.068)	1.674*** (0.134)
Observations	265	265
R ²	0.040	0.315
Adjusted R ²	0.036	0.309
Residual Std. Error	0.949 (df = 263)	0.803 (df = 262)
F Statistic	10.820*** (df = 1; 263)	60.162*** (df = 2; 262)

Note:

*p<0.1; **p<0.05; ***p<0.01

$$IE = 0.439 - 0.161 = 0.278$$

Product Approach

Table:

	<i>Dependent variable:</i>	
	emo (1)	immigr (2)
tone_eth	1.480*** (0.380)	0.161 (0.116)
emo		0.188*** (0.018)
Constant	6.594*** (0.193)	1.674*** (0.134)
Observations	265	265
R ²	0.054	0.315
Adjusted R ²	0.051	0.309
Residual Std. Error	2.703 (df = 263)	0.803 (df = 262)
F Statistic	15.143*** (df = 1; 263)	60.162*** (df = 2; 262)

Note:

*p<0.1; **p<0.05; ***p<0.01

$$IE = 1.480 * 0.188 = 0.278$$

Problems with the Classic Linear Approach

- ▶ What if there is unobserved confounding? Other problems?
- ▶ Sensitivity analysis?
- ▶ Unclear how to implement this for nonlinear models (e.g., logit)
- ▶ No standard errors for indirect effect

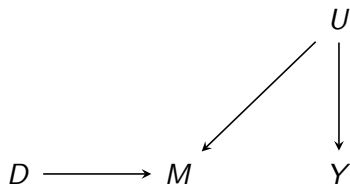
Section 13

Mediation: Identification & Post-Treatment Confounding

Natural Effects: Identification

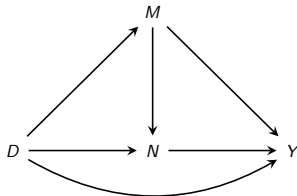
- ▶ I am skipping the (intricate) general definition of direct and indirect effects
- ▶ Instead, focus again on identification
- ▶ I.e., when can we estimate direct/indirect effects from data?
How can we think about choosing control variables?

Recap: Post-Treatment Variables: Problem 2



- ▶ Conditional on M , D and Y are d-connected! Collider!
- ▶ Control for U necessary to correctly infer zero direct effect of D on Y

A Possible Causal Graph for the Brader et al. Study

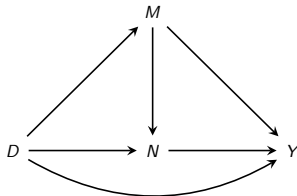


- ▶ randomized treatment D
- ▶ cost/benefit beliefs M
- ▶ anxiety N
- ▶ opposition to immigration Y
- ▶ If we want “direct effect that does not go through N ”, control for M ?
- ▶ Poll

Identification of Natural Direct Effects

- ▶ Graphical version of Sequential Ignorability (Imai et al. 2010) due to Pearl 2014:
- ▶ There are covariates X such that
- ▶ 1. X and D block all D -avoiding back-door paths from N to Y
- ▶ 2. X blocks all back-door paths from D to N and from D to Y , and no member of X is descendant of D
- ▶ In essence: Control for confounders of D and Y , D and N , and N and Y
- ▶ And: “no member of X is descendant of D ”

A Possible Causal Graph for the Brader et al. Study



- ▶ If we want “direct effect not through N ”, control for M ?
- ▶ This effect is fundamentally unidentifiable! (w/o further assumptions)
 - ▶ Control for M : Block part of effect that goes through N via M
 - ▶ Do not control: Open confounding path $N \leftarrow M \rightarrow Y$
- ▶ M acts as a *post-treatment confounder*
- ▶ This is an issue even if M is measured!

Interim Summary

- ▶ “When I measure all mediators and all other relevant variables (“confounders”), I can disentangle the effect of the treatment into different indirect effects”
 - ▶ No, because other mediators may act as post-treatment confounders

Section 14

Mediation Analysis: Estimation & Sensitivity Analysis

Imai et al.: “mediation” package

- ▶ R: `install.packages("mediation")`
- ▶ General idea:
 - ▶ Fit a regression of Y on D and M (plus controls) (outcome model)
 - ▶ Fit a regression of M on D (plus controls) (mediator model)
 - ▶ Package calculates Total, Direct, Indirect effect from that
- ▶ Supports many, many models in R

Basic Usage

```
model.m <- lm(m ~ d, data=df)
```

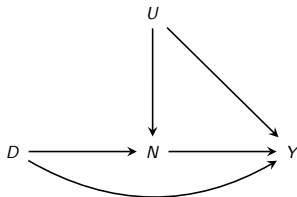
```
model.y <- lm(y ~ d + m, data=df)
```

```
out.1 <- mediate(model.m, model.y,  
sims=1000, treat="d",  
mediator="m",  
boot=FALSE)
```

```
plot(out.1)
```

```
summary(out.1)
```

Unobserved Mediator-Outcome Confounding



- ▶ randomized treatment D
- ▶ (cost/benefit beliefs M)
- ▶ anxiety N
- ▶ opposition to immigration Y
- ▶ U unobserved confounder

Sensitivity Analysis in mediation Package

```
sensout.1 <- medsens(  
  out.1, sims=10000, rho.by=.01)  
  
summary(sensout.1)  
plot(sensout.1, sens.par="R2")  
plot(sensout.1)
```

Sensitivity Results

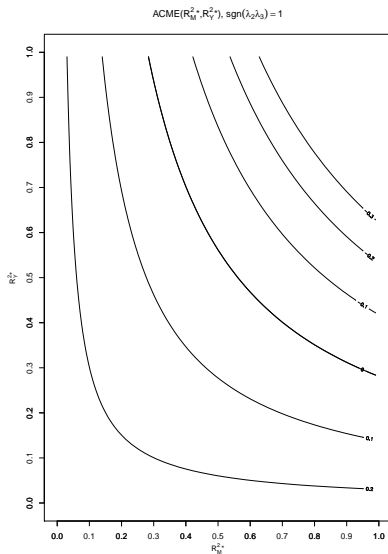


Figure: Output from `plot(sensout.1, sens.par="R2")`

Sensitivity Results

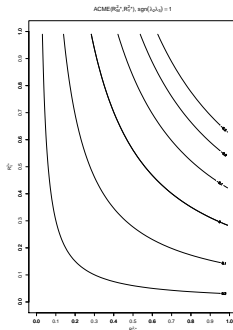


Figure: Output from `plot(sensout.1, sens.par="R2")`

- ▶ Original point estimate not shown at $(0, 0)$
- ▶ As in `sensmakr`, contour lines point estimates for varying combinations of R^2 that unobserved confounder explains (in M and in Y)

Alternative Sensitivity Results

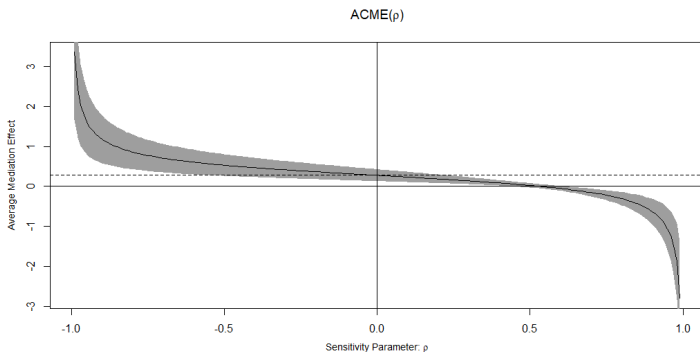


Figure: Output from `plot(sensout.1)`

- ▶ Dashed horizontal line is original point estimate
- ▶ Solid black line point estimate for varying ρ
- ▶ Shaded area are confidence intervals

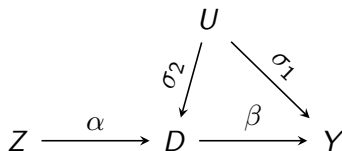
Basic Mediation: Summary

- ▶ In “simple” studies where aim is to estimate the total causal effect, control for post-treatment variables/mediators may create bias
 - ▶ Control away part of the causal effect of interest and/or
 - ▶ Open up non-causal paths (colliders)
- ▶ If aim is to estimate direct/indirect effects, control for mediators seems to make sense
- ▶ Unobserved confounders of M and Y still create problems; can use sensitivity analysis
- ▶ New (and unique) problem: Post-treatment confounding
- ▶ Not solvable without stronger assumptions
- ▶ Sensitivity analysis under stronger assumptions possible: `multimed`

Section 15

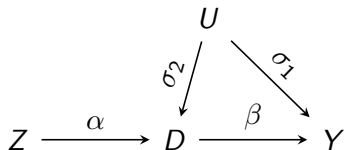
Instrumental Variables via DAGs

IV in Linear Case



- ▶ Regard graph as depiction of *structural* or *causal* equations:
 $D = \alpha Z + U$
 $Y = \beta D + U$
- ▶ These are not regressions, but “nature” or “society”
- ▶ E.g., in our simulations, we played nature
- ▶ β is a causal effect that may or may not be *identified* via the regression of Y on D

IV in Linear Case



- ▶ $\text{cov}(Z, D) = \alpha$
- ▶ $\text{cov}(Z, Y) = \alpha \cdot \beta$
- ▶ We want β . We can estimate $\text{cov}(Z, D)$ and $\text{cov}(Z, Y)$
- ▶ So: $\beta = \frac{\text{cov}(Z, Y)}{\text{cov}(Z, D)}$
- ▶ Here, Z acts as an **instrumental variable for the effect of D on Y**

Cyrus Samii's IV Greatest Hits Collection

- ▶ Draft lottery numbers → military service → income (Angrist 1990)
- ▶ Quarter of birth → schooling → income (Angrist & Krueger 1991)
- ▶ Election year → number of police → crime (Levitt 1997)
- ▶ Sibling sex composition → number of children → labor supply (Angrist & Evans 1998)
- ▶ Settler mortality → investment in institutions → avg. income (Acemoglu et al. 2001)
- ▶ Rain → avg. income → civil war (Miguel et al. 2004)
- ▶ Density of railroads → segregation → inequality (Ananat 2011)

Alternative Derivation

- ▶ In fact, only structural model for Y needs to be linear:
- ▶ $Y = \beta D + U$, where U and D correlate (back-door path)
- ▶ Using this equation, $\text{cov}(Z, Y)$ is
- ▶ $= \text{cov}(Z, \beta D + U) = \beta \text{cov}(Z, D) + \text{cov}(Z, U)$
- ▶ The graph says $Z \perp\!\!\!\perp U$, so $\text{cov}(Z, U) = 0$
- ▶ Solve for $\beta = \frac{\text{cov}(Z, Y)}{\text{cov}(Z, D)}$
- ▶ We have made no assumption on structural model for D !

2nd Alternative/Two-Stage Least Squares

- ▶ Let $D = \mu + \delta_1 Z + \epsilon$ be the **linear projection** of D on Z
- ▶ This is not structural, nor a regression, but a **linear approximation** to $E[D|Z]$ that almost always exists (OLS in the population)
- ▶ $cov(Z, \epsilon) = 0$ by construction (as for regression error when indep. vars are discrete)
- ▶ Insert into structural model for Y :
- ▶ $Y = \beta(\mu + \delta_1 Z + \epsilon) + U$
- ▶ $= \beta\mu + \beta\delta_1 Z + \beta\epsilon + U$
- ▶ This is a mix of structural and linear-projection coefficients
- ▶ Could be estimated via OLS if $cov(Z, \beta\epsilon + U) = \beta cov(Z, \epsilon) + cov(Z, U) = 0$, which is true by construction (ϵ)/by assumption (U)

Two-Stage Least Squares

- ▶ $\beta\mu + \beta\delta 1Z + \beta\epsilon + U$
- ▶ $= \beta\mu + \beta\hat{D} + \beta\epsilon + U$
- ▶ Where $\hat{D} = \delta 1Z$ are fitted values from first-stage linear projection
- ▶ This suggests:
- ▶ OLS of D on Z , regardless of what kind of variables D and Z are. Generate \hat{D} .
- ▶ OLS of Y on \hat{D} . Coefficient is consistent estimate of causal effect β
- ▶ **Two-Stage Least Squares**
- ▶ Implemented in standard statistical software (which also gives correct standard errors)

Implementation of 2SLS

```
library(estimatr)
```

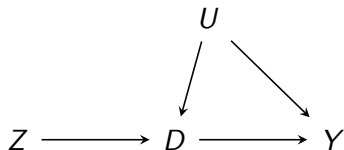
```
iv_robust(Y ~ D + X | Z + X, data = dat)
```

- ▶ Y is outcome
- ▶ D is treatment
- ▶ Z is instrument
- ▶ X are controls

Section 16

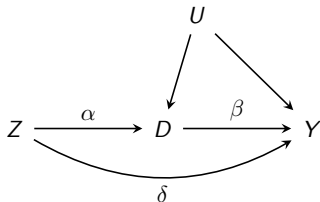
IV Assumptions & Covariates

Basic IV



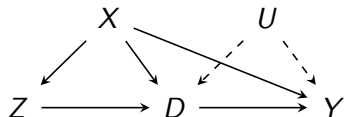
- ▶ Z does not directly affect Y (“exclusion restriction”, “no direct effect”)
- ▶ No variables impacting Z and Y or Z and $D \implies$ no back-door paths from instrument to treatment or outcome
- ▶ E.g. Z US Vietnam war draft lottery, D actually serving in Vietnam war, Y wages after return, U unobserved ability (Angrist 1990)

IV with Direct Effect of Instrument



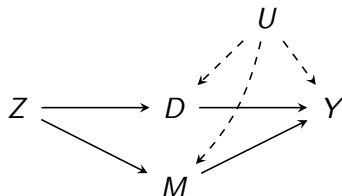
- ▶ What is the IV estimator $\frac{\text{cov}(Z, Y)}{\text{cov}(Z, D)}$ in this case?
- ▶ $\text{cov}(Z, D) = \alpha$
- ▶ $\text{cov}(Z, Y) = \alpha\beta + \delta$
- ▶ $\hat{\beta} = \frac{\alpha\beta + \delta}{\alpha} = \beta + \frac{\delta}{\alpha}$
- ▶ Asymptotic bias $\frac{\delta}{\alpha}$. Z not a valid IV
- ▶ Larger if instrument is “weaker” (smaller α)

IV with Covariates



- ▶ X : birth year – demand of military varied from year to year
- ▶ Control for X blocks back-door path – all good

Post-Instrument Covariates



- ▶ M attending college to defer the draft
- ▶ Is Z still a valid IV? Must we control for M ? Poll!
- ▶ If no control for M : Violation of exclusion restriction
- ▶ But conditioning on M creates non-causal association between Z and Y (collider!)
- ▶ $\implies Z$ not a valid instrument
- ▶ See Glynn et al. forthcoming, Schuessler et al. 2021 for deeper discussion

Section 17

Instrumental Variables: (Almost) Nonparametric Case

From Linear to Nonparametric Case

- ▶ IV in linear case is very easy
- ▶ In nonparametric case, complications occur
- ▶ The problem comes from heterogeneity in causal effects
- ▶ In linear causal models, everyone has the same individual causal effect $Y_1(u) - Y_0(u) = \beta$
- ▶ In more realistic nonparametric case, $Y_1(u) - Y_0(u)$ varies across u /across individuals i

Nonparametric Case

- ▶ With binary Z and D , we can write the structural equations of the simple IV model as
 - ▶ $Y_i = \mu_1 + \beta_i D_i + \epsilon_i$
 $D_i = \mu_2 + \alpha_i Z_i + \epsilon_i$
 $Z_i \perp\!\!\!\perp \epsilon_i$
- ▶ Where $\beta_i = Y_{D=1}(u) - Y_{D=0}(u)$ and $\alpha_i = D_{Z=1}(u) - D_{Z=0}(u)$ are unit-level causal effects
- ▶ Very easy to show that $E[Y|Z = 1] - E[Y|Z = 0] = E[\alpha_i \beta_i]$ by structural definition of counterfactuals & BDC
- ▶ Also clear that $E[D|Z = 1] - E[D|Z = 0] = E[\alpha_i]$

Nonparametric Case

- ▶ $Y_i = \mu_1 + \beta_i D_i + \epsilon_i$
 $D_i = \mu_2 + \alpha_i Z_i + \epsilon_i$
- ▶ In this model, $\alpha_i, \beta_i, \epsilon_i$ are all part of U_i , the unobserved confounders that influence
 - ▶ D and Y and
 - ▶ how D reacts to Z and Y reacts to D (interactions!)
- ▶ So α_i and β_i correlate
- ▶ With binary D , $\alpha_i = D_{Z=1}(u) - D_{Z=0}(u)$ can only take on three values: 1, 0, -1
- ▶ It turns out that now, bad things can happen with our usual IV estimator

Nonparametric Case

- ▶ $\frac{\text{cov}(Z, Y)}{\text{cov}(Z, D)} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]} =$
- ▶ $\frac{E[\alpha_i \beta_i]}{E[\alpha_i]}$ is the IV estimator if our graph is correct
- ▶ Now let's say for people with $\alpha_i = 1$ and $\alpha_i = -1$, $\beta_i = 1$; for $\alpha_i = 0$ units, $\beta_i = 0$. All α_i equally likely ($\frac{1}{3}$)
- ▶ Then by LIE, $ATE = E[\beta_i] = E[\beta_i | \alpha_i = 1]P(\alpha_i = 1) + E[\beta_i | \alpha_i = -1]P(\alpha_i = -1) + E[\beta_i | \alpha_i = 0]P(\alpha_i = 0) = \frac{2}{3}$
- ▶ But also by LIE: $E[\alpha_i \beta_i] = E[\alpha_i \beta_i | \alpha_i = 1]P(\alpha_i = 1) + E[\alpha_i \beta_i | \alpha_i = -1]P(\alpha_i = -1) + E[\alpha_i \beta_i | \alpha_i = 0]P(\alpha_i = 0)$
- ▶ $= E[\beta_i | \alpha_i = 1]P(\alpha_i = 1) + E[-\beta_i | \alpha_i = -1]P(\alpha_i = -1)$
- ▶ $= 1 \cdot \frac{1}{3} - 1 \cdot \frac{1}{3} = 0!$
- ▶ IV estimator will be 0 even though ATE is $\frac{2}{3}$!

Parametric Solutions

- ▶ Solutions by making stronger assumptions:
- ▶ $E[\alpha_i\beta_i] = E[\alpha_i]E[\beta_i]$. This is almost like assuming away confounding (uncorrelated effect heterogeneity)
- ▶ $\beta_i = \beta$, a constant, so $E[\alpha_i\beta_i] = E[\alpha_i]\beta$. This is similar to linearity (constant causal effects)
- ▶ Most common in political science/econ: Assume away that people exist with $\alpha_i = -1$
- ▶ Since their choice of D reacts to Z “in the opposite way”, $\alpha_i = -1$ units are also called **defiers**
- ▶ Assumption also sometimes called **monotonicity**, because Z may not have positive AND negative impact on D
- ▶ Since this restricts the structural function for D , **it is a parametric assumption**

The LATE Model

- ▶ $Z \perp\!\!\!\perp Y_{D=1}, Y_{D=0}, D_{Z=1}, D_{Z=0} \implies Z \perp\!\!\!\perp \beta_i, \alpha_i$
- ▶ This is the **old** instrumental assumption from linear case (no back-door paths to or direct effect on Y), plus BDC for $Z \rightarrow D$ (**new!**)
- ▶ No defiers: $P(\alpha_i = -1) = 0$ (**new!**)
- ▶ Relevance/first-stage: $E[D|Z = 1] - E[D|Z = 0] \neq 0$ (**old**)
- ▶ Then the IV estimator by above reasoning evaluates to
- ▶
$$\frac{E[\alpha_i \beta_i | \alpha_i = 1] P(\alpha_i = 1)}{E[\alpha_i]} =$$
- ▶
$$\frac{E[\beta_i | \alpha_i = 1] P(\alpha_i = 1)}{P(\alpha_i = 1)} =$$
- ▶ $E[\beta_i | \alpha_i = 1]$
- ▶ $= E[Y_{D=1} - Y_{D=0} | D_{Z=1} - D_{Z=0} = 1]$

The LATE Model

- ▶ $E[\beta_i | \alpha_i = 1] = E[Y_{D=1} - Y_{D=0} | D_{Z=1} - D_{Z=0} = 1]$
- ▶ The average effect of D on Y for those units whose choice of D reacts to Z
- ▶ **Local** Average Treatment Effect (LATE), Complier Average Causal Effect (CACE)

LATE is not ATE

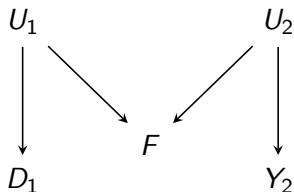
- ▶ In our example, $\alpha_i = 1 \implies \beta_i = 1$ and $\alpha_i = 0 \implies \beta_i = 0$
- ▶ If $P(\alpha_i = 1) = 0.5$ and no defiers, this would mean $ATE = 0.5 \neq LATE = 1$
- ▶ ATE is usually more relevant for policy and science
- ▶ Compliers may be small part of overall population. Fortunately, first-stage is share of compliers: $E[\alpha_i] = P(\alpha_i = 1)$, so we can check this
- ▶ Plus, we cannot directly observe who a complier is, because we cannot observe α_i ; so LATE is not really a covariate-specific effect
- ▶ In general, people debate whether LATE is useful to know or whether we should care about ATE (e.g. Heckman)
- ▶ It turns out that using an instrument, even without monotonicity, one can at least **partially identify/bound** the ATE (Balke/Pearl 1997)

Section 18

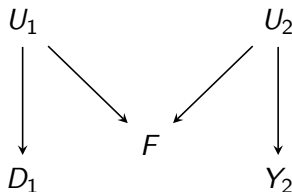
Network Studies: The Shalizi/Thomas Critique

Network Studies

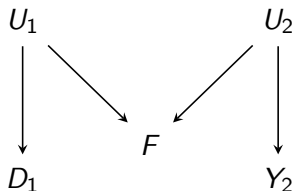
- ▶ One foundational question for network studies: Do features of one unit affect connected ones?
- ▶ E.g., if one student gets bullied by some (D), will his/her friends like him less (Y)?
- ▶ Such questions should also be of interest to qualitative researchers



- ▶ D_1 indicates whether student 1 gets bullied. U_1 are 1's traits
- ▶ Y_2 is what student 2 thinks about student 1. U_2 are 2's traits.
- ▶ F is a dummy that indicates whether 1 and 2 are friends.
- ▶ For each student, regress friends' average opinions on bullying variable D . What happens?



- ▶ Let's say U just contains age: Students of the same age flock together (homophily)
- ▶ Let's say younger students are more likely to be bullied than older ones
- ▶ And older ones are generally friendlier about others than younger ones
- ▶ But friends do not change their opinion about friends when they see them getting bullied (no effect)



- ▶ Those who are bullied (D_1) are more likely to be young (U_1), more likely to have younger friends (those with $F = 1$ are younger, U_2), therefore less liked (Y_2)
- ▶ Those who are not bullied (D_1) are more likely to be older (U_1), more likely to have older friends (those with $F = 1$ are older, U_2), therefore more liked (Y_2)
- ▶ Systematic relationship between bullying and friends' opinion, even though no causal effect
- ▶ Issue is *implicit conditioning on F* in the analysis
- ▶ Shalizi/Thomas 2011, Sociological Research & Methods.

Section 19

Missing Data & Selection Bias from a DAG View

Poll

- ▶ Have you read the paper by Peter Selb and me?

Breakout Rooms

- ▶ Discuss & Collect in Google Doc: Why is missing data a problem?
 - ▶ Which different problems can occur due to missing data?
 - ▶ On what do biases due to missingness depend?
 - ▶ How could we solve these problems?
 - ▶ How is this related to external validity?

Two Kinds of Problems

- ▶ 1. Your data suffers from non-response / non-random sampling
- ▶ This is the problem of **(Sample) Selection Bias**
- ▶ 2. You have an experimental causal effect estimate for one population. Can you infer the effect in a different population?
- ▶ This is the problem of **External Validity / Generalizability / Transportability**

Basic Formalization

- ▶ Selection variable S
- ▶ Selection Bias: Data conditional on $S = 1$, but we want quantity for everyone
 - ▶ S describes subsets of population of interest
 - ▶ Forced to condition on $S = 1$
 - ▶ Outcome Y missing if $S = 0$
- ▶ Special case of “missing data” (imputation etc.)
 - ▶ X_1, X_2, \dots, Y all may contain some pattern of missingness
 - ▶ Specific missingness indicators possible (but stronger assumption)

Not all selection biases are the same

- ▶ Especially in econometrics, selection bias from “selection into treatment” = back door paths
- ▶ Different problem from “sample” selection bias = selection bias
- ▶ The problems interact

Section 20

Schuessler/Selb: Prototypical Selection Scenarios

- ▶ Interest is in marginal distribution / mean of Y ($P(Y)$ / $E[Y]$)
- ▶ Usually needed: Population distribution of adjustment variables $P(X)$ from census (“external” or “auxiliary” data)
- ▶ Or, depending on concrete estimation method, X for everyone in the sampling frame (not only respondents)

Prototypical Selection Scenario I

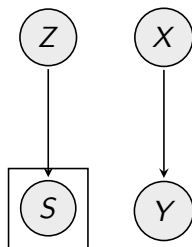


Figure: Y is missing completely at random (MCAR)

- ▶ d-separation gives $Y \perp\!\!\!\perp S$
- ▶ $P(Y|S = 1) = P(Y)$: Sampled data is representative
- ▶ Extreme case: Randomization device Z is only cause of S
- ▶ But random sampling not necessary

Prototypical Selection Scenario II

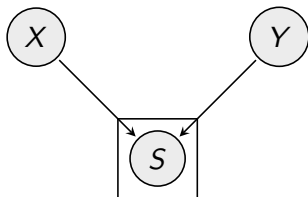


Figure: Y is missing not at random (MNAR)

- ▶ Forced to condition on collider
- ▶ Creates association between X and Y where none exists in population!
- ▶ Explanation for Lahtinen et al. 2019: From validation data, infer that nonresponse leads to underestimation of correlation between SES X and turnout Y

```
set.seed(923)
n <- 1000
u <- rnorm(n)
x <- u + rnorm(n)
y <- u + rnorm(n)
s <- (x + y) > 1
mean(s)
[1] 0.349
cor(x, y)
[1] 0.4895506
cor(x[s==1], y[s==1])
[1] -0.1328959
```


Prototypical Selection Scenario III

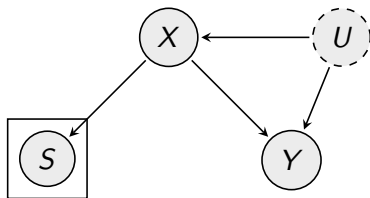


Figure: Y is missing at random (MAR), conditional on X

- ▶ X d-separates S and Y
- ▶ $P(Y) = \sum_x P(Y|x, S = 1)P(x)$
- ▶ e.g. via IPW = Horvitz-Thompson, or MRP
- ▶ We can “recover” or “identify” population distribution of Y

(Simple) MNAR cannot be solved nonparametrically

- ▶ If Y directly affects S , no way to solve this problem without additional data / parametric assumptions
- ▶ E.g. Heckman selection model

Summary: Graphical Criterion and Estimation

- ▶ For recovery of $P(Y)$, need X that d-separate Y from $S \rightarrow Y \perp\!\!\!\perp S | X$
- ▶ Need $P(X)$ from census, e.g. joint distribution of age/gender/region
- ▶ Then use weighting, MRP...

Section 21

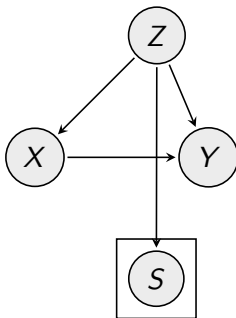
Schuessler/Selb: Inference on Conditional Distributions

- ▶ Interest now in $P(Y|X)$, $E[Y|X]$, or regression of Y on X
- ▶ Usually needed: Population distribution $P(Z, X)$ from census (“external” or “auxiliary” data)

Textbook Case

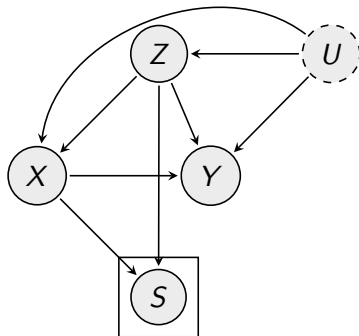
- ▶ The textbook case is $Y \perp\!\!\!\perp S | X$
- ▶ Not practical: Usually regressors of interest X are small subset of those available/needed for non-response adjustment (X, Z)
- ▶ And inclusion of Z in regression model will change regression coefficients of X if correlated

Use of Auxiliary Variables for Inferring Cond. Distributions



- ▶ $Y \not\perp\!\!\!\perp S|X$, but $Y \perp\!\!\!\perp S|X, Z$
- ▶ $P(Y|X) = \sum_z P(Y|X, z, S = 1)P(z|X)$ (LoTP and d-sep.)

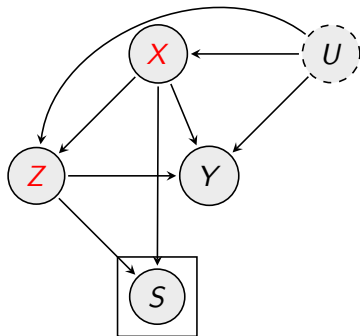
Use of Auxiliary Variables for Inferring Cond. Distributions II



- ▶ Unobserved confounders do not change story
- ▶ As long as $U \rightarrow S$ and $Y \rightarrow S$ do not exist

Use of Auxiliary Variables for Inferring Cond. Distributions

III



- ▶ Causal ordering of X and Z does not matter!
- ▶ Unlike back-door criterion

Graphical Criterion

- ▶ Find Z that together with regressors X d-separate Y from S
 $\implies Y \perp\!\!\!\perp S | Z, X$
- ▶ Using population data / sampling frame data on Z, X , estimate $P(S = 1 | X, Z)$
- ▶ Predict $P(S = 1 | X, Z)$ for each observation, use inverse as weight in estimation (works for OLS, 2SLS, MLE, and more)

Section 22

External Validity, Transportability, Etc.

External Validity

- ▶ “I believe the causal effect of X on Y is 3.4 in your sample, but is it externally valid?”
- ▶ What is “external”?
- ▶ Need to specify the other population
- ▶ Need to measure effect moderators in other population

Conceptualization of Transportability

- ▶ Suppose we have a valid causal effect estimate for one population $S = 1$
- ▶ Want to “transport” it to other population $S = 0$
- ▶ S now of intrinsic interest, not nuisance
- ▶ Not sensible to conceptualize S as caused by something else
 - ▶ Researcher chooses S , not units
- ▶ Instead, S causes dissimilarities between populations
- ▶ S as selection versus S as (dis)similarity
- ▶ Corollary: Heckman-style models do not make sense for transportability!

Selection Bias vs. Transportability, Data Structure

X	Y	S
1	NA	0
4	2	1
3	3	1

X	Y	S
1	3	0
4	2	1
3	3	1

Formalization

- ▶ Pearl/Bareinboim 2014: $P(Y|do(x))$ known, want $P^*(Y|do(x))$
- ▶ Equivalent (ibid., fn. 16): $P(Y|do(x), S = 1)$ known, want $P(Y|do(x), S = 0)$

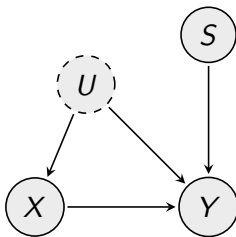
Effect Heterogeneity

- ▶ Issue in transportability is effect heterogeneity – only
- ▶ Critical question: Under what assumptions can we capture “relevant” effect heterogeneity using observed Z only, and what does the estimand look like?

Selection Diagrams

- ▶ Gives rise to selection diagrams (Pearl/Bairnboim)
- ▶ S points into variables V where
 - ▶ Structural function f for V and/or
 - ▶ Unobserved causes U for V differ between populations
- ▶ In nonparametric graphs, never specify f , so might as well go with “ U differ”-interpretation

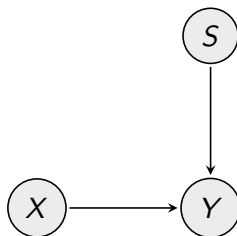
Simple Selection Diagram



Graphical Criterion

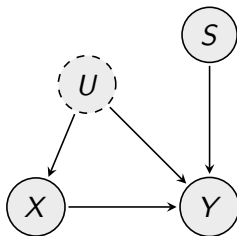
- ▶ Graphical criterion from before can still be used:
 $(Y \perp\!\!\!\perp S | Z, X)_{G_{\overline{X}}}$ (“S-admissability”)
- ▶ If Z not influenced by X , then estimate
 $\sum_z P(Y|do(X), z, S = 1)P(z|S = 0)$ (Pearl/Bairnboim 2014, Corollary 1)
- ▶ Similar to causal-inference-selection-bias case
- ▶ $P(Y|do(X), z, S = 1)$ directly measured in experiment
- ▶ $P(z|S = 0)$ distribution of relevant moderators in other population
- ▶ Estimation via interactive models, averaging over target population distribution of moderators

Trivial Transportability



- ▶ No confounding problem to solve, experiment not necessary
- ▶ $P(Y|X, S = s)$ equals $P(Y|do(x), S = s)$
- ▶ Simply measure X and Y observationally in population of interest

Transportability Failure

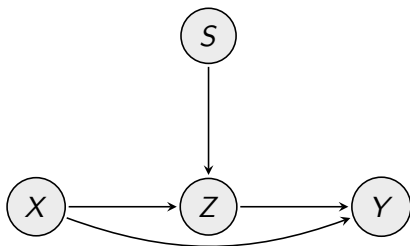


- ▶ $P(Y|do(x), S = 1)$ given from one population
- ▶ Other population somehow differs in $P(U)$
- ▶ Transportation not possible
- ▶ If U measured, effect of interest trivially transportable as before

Non-DAG Approach

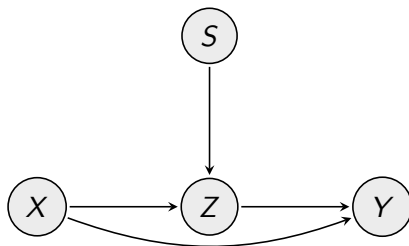
- ▶ Stuart et al. 2011:
 - ▶ $S \perp\!\!\!\perp Y(x) | Z$
 - ▶ “all potential moderators of the treatment effect measured”
 - ▶ Interpretation wrong – every cause of Y is a potential moderator
 - ▶ Assumption fails if Z influenced by X , even if no unobserved confounding

Post-Treatment Moderators



- ▶ $(Y \perp\!\!\!\perp S | Z, X)_{G_{\overline{X}}}$ holds
- ▶ But correct adjustment formula weights with $P(z|x, S = 0)$, as X influences Z

Post-Treatment Moderators



- ▶ This is VERY likely in the social sciences
- ▶ E.g. Z updated beliefs, depend on priors, vary across contexts, influence behavior
- ▶ Adjustment for post-treatment moderator Z would be necessary (in two-stage style model)

Section 23

Summary

Checklist for Missing Data / External Validity

1. What is the population of interest, how does it relate to the sample/the experiment?
 - 1.1 Population of interest sampled, but nonresponse: Selection bias
 - 1.2 Population of interest sampled, but experiment only in other population, full response: Transportability
2. If selection bias, what is the quantity of interest: Mean of Y , regression of Y on X , causal effect of X on Y ?
 - 2.1 If mean/regression, selection bias: Need $Y \perp\!\!\!\perp S | X, Z$, and population data on X, Z . Weighting estimator with inverse of $P(S = 1 | X, Z)$
 - 2.2 If causal effect, selection bias: Generalized adj. criterion, population data on Z . Weights are inverse of $P(S = 1 | Z)$
3. If transportability, look for Z such that $(Y \perp\!\!\!\perp S | Z, X)_{G_{\bar{X}}}$
 - 3.1 Estimation via interactive models, averaging over target population distribution of moderators

Section 24

Further Literature

Books

Hernán MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020.

<https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>

Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. Causal inference in statistics: A primer. John Wiley & Sons, 2016.

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. The MIT Press, 2017.

Shalizi, Cosma. "Advanced data analysis from an elementary point of view." (2021) <http://www.stat.cmu.edu/cshalizi/ADAfaEPoV/>

Literature

Pearl, Judea. "Direct and indirect effects." Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence. 2001.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. "Identification, inference and sensitivity analysis for causal mediation effects." Statistical science (2010): 51-71.

Imai, Kosuke, et al. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." American Political Science Review (2011): 765-789.

Imai, Kosuke, and Teppei Yamamoto. "Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments." Political Analysis (2013): 141-171.

Tingley, Dustin, et al. "Mediation: R package for causal mediation analysis." (2014).

Literature

Pearl, Judea. "Interpretation and identification of causal mediation." *Psychological methods* 19.4 (2014): 459.

Pearl, Judea. *Causality*. Cambridge university press, 2009.

Goldin, Claudia, and Cecilia Rouse. "Orchestrating impartiality: The impact of 'blind' auditions on female musicians." *American economic review* 90.4 (2000): 715-741.

Yamamoto, Teppei. "Identification and estimation of causal mediation effects with treatment noncompliance." Unpublished manuscript (2013).

Literature

Brader, Ted, Nicholas A. Valentino, and Elizabeth Suhay. "What triggers public opposition to immigration? Anxiety, group cues, and immigration threat." *American Journal of Political Science* 52.4 (2008): 959-978.

Frölich, Markus, and Martin Huber. "Direct and indirect treatment effects—causal chains and mediation analysis with instrumental variables." *Journal of the Royal Statistical Society Series B* 79.5 (2017): 1645-1666.

Hazlett, Chad. "Angry or Weary? How Violence Impacts Attitudes toward Peace among Darfurian Refugees." *Journal of Conflict Resolution* 64.5 (2020): 844-870.

Cinelli, Carlos, and Chad Hazlett. "Making sense of sensitivity: Extending omitted variable bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82.1 (2020): 39-67.

Literature

Schuessler, Julian, Glynn, Adam N., and Rueda, Miguel. R.
"Post-Instrument Bias". Working paper. (2021)

Glynn, Adam N., Rueda, Miguel. R., Schuessler, Julian
"Post-Instrument Bias in Linear Models". Forthcoming in
Sociological Methods & Research.

Schuessler, Julian, and Peter Selb. "Graphical causal models for
survey inference." Working Paper (2021).

Elwert, Felix, and Christopher Winship. "Endogenous selection
bias: The problem of conditioning on a collider variable." Annual
review of sociology 40 (2014): 31-53.

Knox, Dean, Will Lowe, and Jonathan Mummolo. "Administrative
records mask racially biased policing." American Political Science
Review 114.3 (2020): 619-637.