Of Causes and Postulates

Macartan Humphreys. Columbia and WZB

18 July 2018

1 Causality is metaphysics

2 No causation without manipulation

3 Ubiquituous causality: You have to care

4 False rumours: of laws, paths, necessity etc

5 Observational data: needs a model

6 Experimental data: still need models

7 The black box: Even more so

8 Qualitative inference: Here too

9 Effects of causes easier to identify than causes of effects

1 Causality is metaphysics

Metaphysics

We will work with the *interventionist* notion of causality

- A cause is a thing that makes a difference (Lewis)

Metaphysics

Difference, with respect to. . . ?

A causal claim is a statement about what *didn't happen*. It is metaphysical (not positivist)

- $Y_i(1) =$ the potential outcome if i receives treatment
- $Y_i(0) =$ the potential outcome if i does not receive treatment
- $Y_i(1) - Y_i(0)$: the individual causal effect.

Metaphysics

Difference, with respect to. . . ?

In Pearl's notation: the distribution of outcomes following an intervention

- $Pr(Y = 1|do(X) = 1, u_i)$
- $Pr(Y = 1|do(X) = 0, u_i)$
- $Pr(Y = 1|do(X) = 1, u_i) - Pr(Y = 1|do(X) = 0, u_i)$

Note: if uncertainty is epistemic then, conditioning on u_i, we expect these probabilities to be ones or zeros.

Metaphysics

We will work with the *interventionist* notion of causality

- But other notions of causality can be handled pretty handily by it. e.g. dependence ideas, "causes of effects" versus "effects of causes"; necessity/sufficiency versus statistical likelihood

$$Y = A * B + C \leftrightarrow Y(A = a, B = b, C = c) = \max(ab, c)$$

- Deterministic / stochastic differences seems something of a red herring: $Y(x)$ generally thought of as fixed, not stochastic
- Question: *are there useful alternative notions?*

Metaphysics and the fundamental problem

We have a fundamental problem

- Causal effect $= Y_i(1) - Y_i(0)$

But we only get to see one of the components of this

Metaphysics

- So: it is metaphysical, *not* positivist.
- Thus it is *itself* a kind of a model.
- But hard to dispense with, short of assuming that nothing that didn't happen could have happened

Inference and measurement

An implication of the fundamental problem is that you cannot *measure* causes; you have to infer them. Or perhaps, imagine them.

2 No causation without manipulation

Causal claims: No causation without manipulation

- Some seemingly causal claims not admissible.
- Did Germany cause the second world war?
- Did Hillary lose because she is a woman?

Causal claims: No causation without manipulation

- To get the definition off the ground, manipulation must be imaginable (whether practical or not)
- This renders thinking about effects of race and gender difficult
- What does it mean to say that Aunt Pat voted for Brexit because she is old?

Causal claims: No causation without manipulation

- To get the definition off the ground, manipulation must be imaginable (whether practical or not)
- This renders thinking about effects of race and gender difficult
- **Compare**: What does it mean to say that Southern counties voted for Brexit because they have many old people?

3 Ubiquituous causality: You have to care

Ubiquituous causality: You have to care

You can **describe**, **represent**, **bear witness**, **appreciate**, without making causal claims.

You can engage in pure **theory**.

But. . .

3 Ubiquituous causality: You have to care

But many seemingly non-causal claims are motivated by causal reasoning, even claims that seem at first to be simple questions of measurement:

- Jack exploited Jill
- It's Jill's fault that bucket fell
- Jack is the most obstructionist member of Congress
- Melania stole Michelle's speech

And:

- Activists need causal claims
- Critics need causal claims

4 False rumours: of laws, paths, necessity etc

Many false rumors

$$ATE = E(Y(1) - Y(0))$$

- No **Determinism** required: Though note the X's are stochastic, not the potential outcomes.
- No **Homogeneity** required: Arbitrary complexity allowed (but not always asked about)
- No **ceteris paribus** required: Estimating average causal effects does not require that treatment and control groups are identical
- No **linearity** required, in either sense

Causality is non transitive

More surprising: No **transitivity** required.

It is possible that A causes B, B causes C but A does not cause C.

- A: Feel depressed; B: Meet with friends; C: Feel happy
- Note: if B was not affected by A then A would have a negative effect on C (the controlled direct effect of A on C is negative)

Causality does not follow a path

More surprising: No **connectedness** (in time and space) required.

X can cause Y even if there is no spatiotemporally continuous "causal path" connecting X and Y.

- Jill is planning on insulting Jack
- Martha sets out to stop Jill
- Martha's bike gets a puncture and she never makes it
- So Jill insults Jack
- Was Martha's puncture a cause of Jill's insult?

An implication is that you cannot expect to be able to "trace a process"

Rumors: "Actual causes" aren't counterfactual causes

Responsibility and cause not coterminous.

Actually doing it does not make it a cause.

Consider:

- A and B throw a stone at the bottle.
- Both are great shots. A is faster and hits the bottle. It breaks.
- B's stone misses but only because A's got there first

Here A's throw is an "actual cause" but not a "counterfactual cause."

- A's throw didn't make a difference.
- Following a tight chain of intermediate steps will not affect that conclusion.
- Should we care about it more than about B's throw?

Rumors: Contribution or attribution?

The counterfactual model focus is on the effects of X or whether X caused Y (given other things, perhaps)

- Focus is on non-rival contributions
- At most it provides a conditional account
- Figuring that X caused Y does not mean that it is *the* cause or the *main* cause.

Consider an outcome Y that might depend on two causes X_1 and X_2:

$$Y(0,0) = 0$$

$$Y(1,0) = 0$$

$$Y(0,1) = 0$$

$$Y(1,1) = 1$$

What caused Y? Which cause was most important?

Rumors: Contribution or attribution?

The counterfactual model focus is on the effects of X or whether X caused Y (given other things, perhaps)

- Focus is on non-rival contributions
- At most it provides a conditional account
- Figuring that X caused Y does not mean that it is *the* cause or the *main* cause.
- This is problem for research programs that define "explanation" in terms of figuring out the things that cause Y
- Real difficulties conceptualizing what it means to say one cause is more important than another cause. What does that mean?

Rumors: Contribution or attribution?

The counterfactual model focus is on the effects of X or whether X caused Y (given other things, perhaps)

- Focus is on non-rival contributions
- At most it provides a conditional account
- Figuring that X caused Y does not mean that it is *the* cause or the *main* cause.
- *Erdogan's increasing authoritarianism was the most important reason for the attempted coup*
  - More important than Turkey's history of coups?
  - What does that mean?

5 Observational data: needs a model

Observation and Models

Much of the time our data comes from observation. From what we see.

We try to make inferences about effects in a population (not "laws").

But we cannot make the inferences without a belief about how the data was generated.

Observation and Models

Should you "control for X":

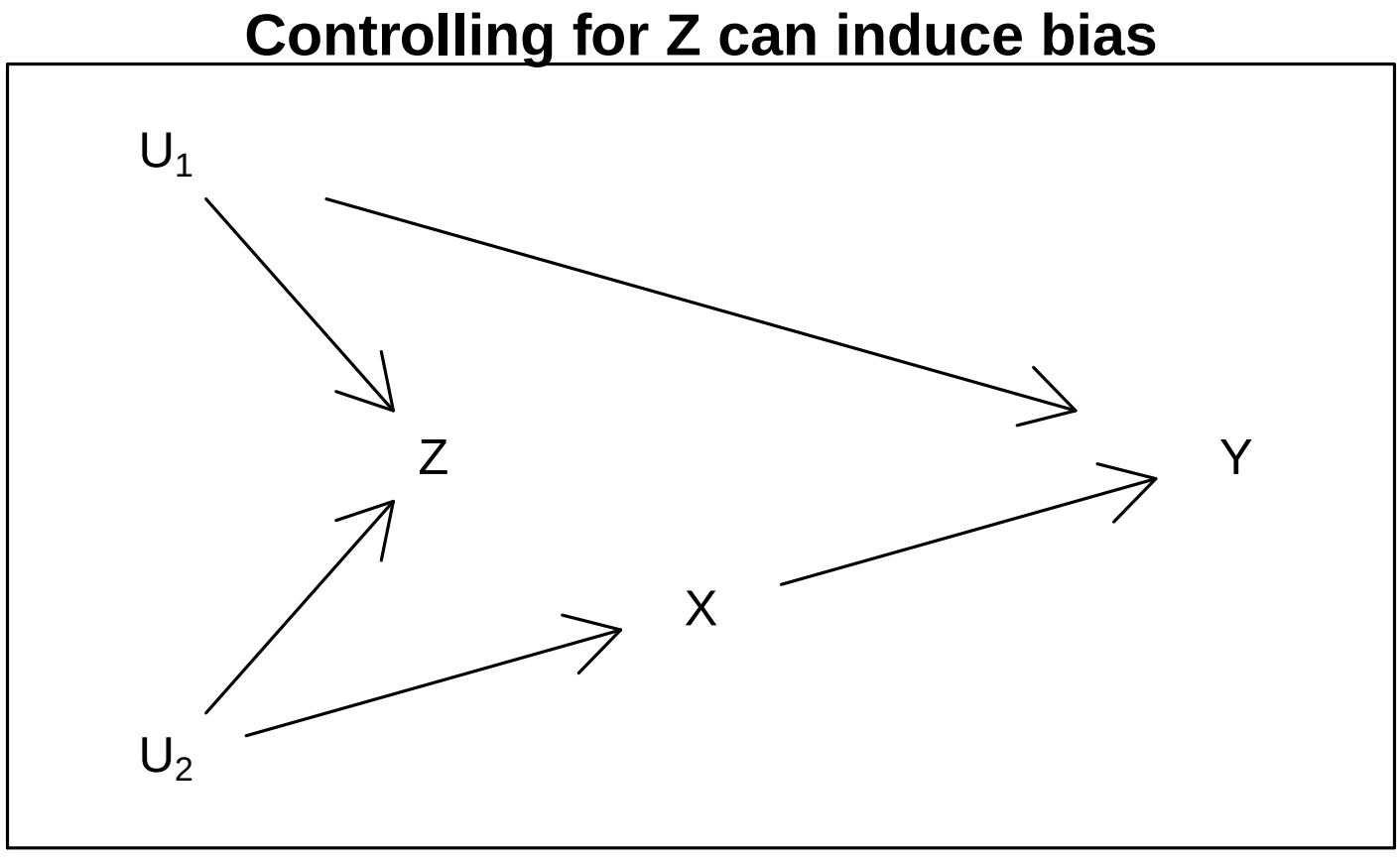
1. $M \to X \to Y$, $M \to Y$
2. $X \to M \to Y$

Challenge for us

- Consider this causal structure:
  - $Z = f_1(U_1, U_2)$
  - $X = f_2(U_2)$
  - $Y = f_3(X, U_1)$

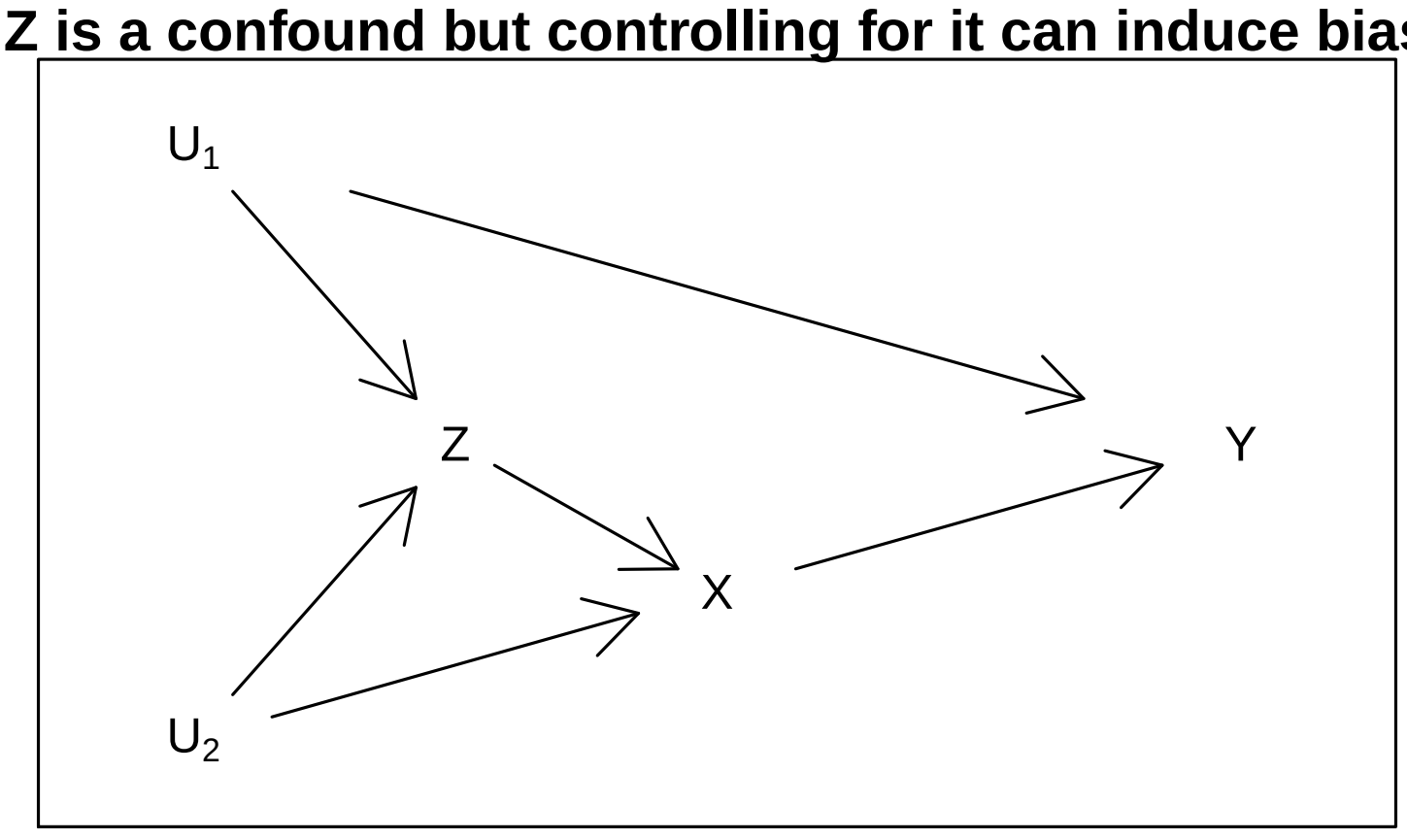
Question: Here Z is prior to X; it is correlated with both Y and X. Would it be useful to "control" for Z when trying to estimate the effect of X on Y?

Challenge for us: Pearl's answer



Collider & Confounder

A bind: from Pearl 1995.



A stronger model can help.

6 Experimental data: still need models

The Magic of the linearity of expectations

Consider the "average causal effect."

$$E(Y_i(1) - Y_i(0))$$

This *can* be estimated since:

$$E(Y_i(1) - Y_i(0)) = E(Y_i(1)) - E(Y_i(0))$$

ie: the expected difference is the difference of expectations

The Magic of the linearity of expectations

This:

$$E(Y_i(1) - Y_i(0)) = E(Y_i(1)) - E(Y_i(0))$$

requires estimating the mean of a population $E(Y_i(1))$.

- Do this using a sample
- Good to go if sample is random

The Magic of the linearity of expectations

- The challenge is to be able to claim that the expected value of an outcome in the treated (or control) units is the same as what the expected value of the treatment (control) outcome for all units

The Magic of the linearity of expectations

A few implications of this:

- You can think of randomization as a sampling problem and apply everything you know about stratification, clustering, weighting and post-stratification. It all goes through
- Units do not have to be assigned with the same probability, but they do need to be assigned with a probability between 0 and 1.
- You have a positive argument for inference – the argument involves no claim around balance on covariates.

Extraordinary power of design based inference

Calculate standard errors and p values using what you *know* from the assignment scheme; not beliefs on distributions you know to be wrong.

- Say you randomized assignment to treatment and your data looked like this.

Unit	1	2	3	4	5	6	7	8	9	10
Treatment	0	0	0	0	0	0	0	1	0	0
Healthy?	3	2	4	6	7	2	4	9	8	2

- Does the treatment improve your health?
- Consider one sided test of null. What is the p value?

But you still need a model

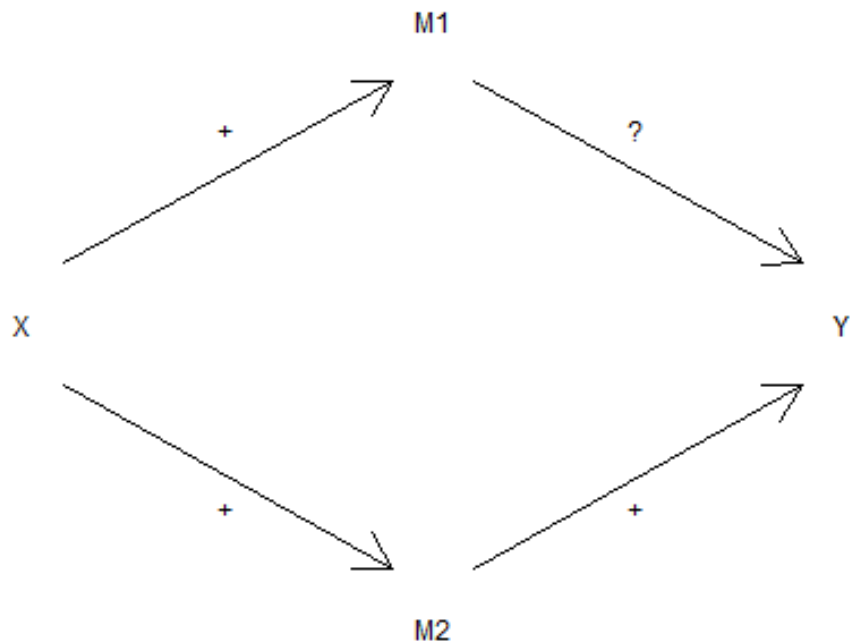
1. At a minimum you generally need to invoke SUTVA. You don;t get causal inference for free.
2. But sometimes the idea of an ATE is itself a model: e.g. the average effect of a $1 increase in income seeks a linear term for a possibly non linear relation

Model 1 is to get off the ground Model 2 is because you want a model of the world

7 The black box: Even more so

The problem of unidentified mediators

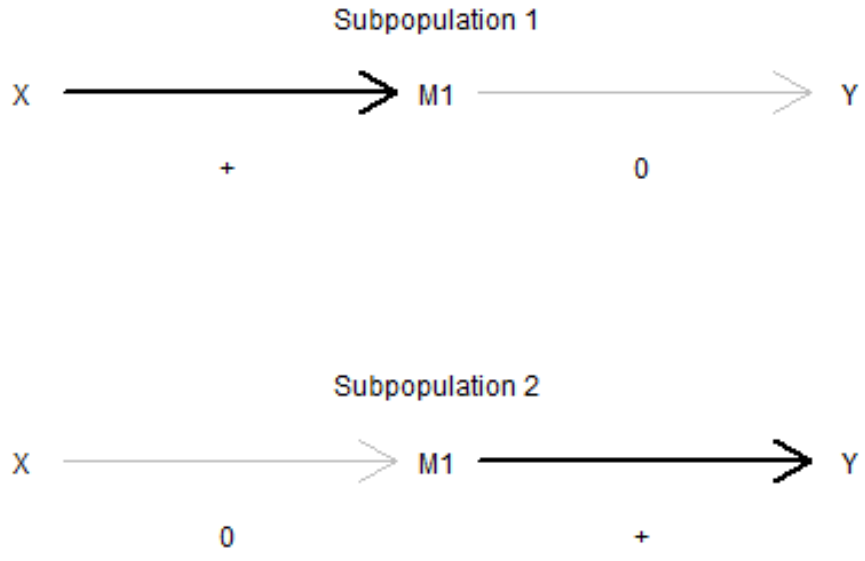
- Consider a causal system like the below.
- The effect of X on M1 and M2 can be measured in the usual way.
- But unfortunately if there are multiple mediators the effect of M1 (or M2) on Y is not identified.
- The 'exclusion restriction' is obviously violated when there are multiple mediators (unless you can account for them all).



The problem of unidentified mediators

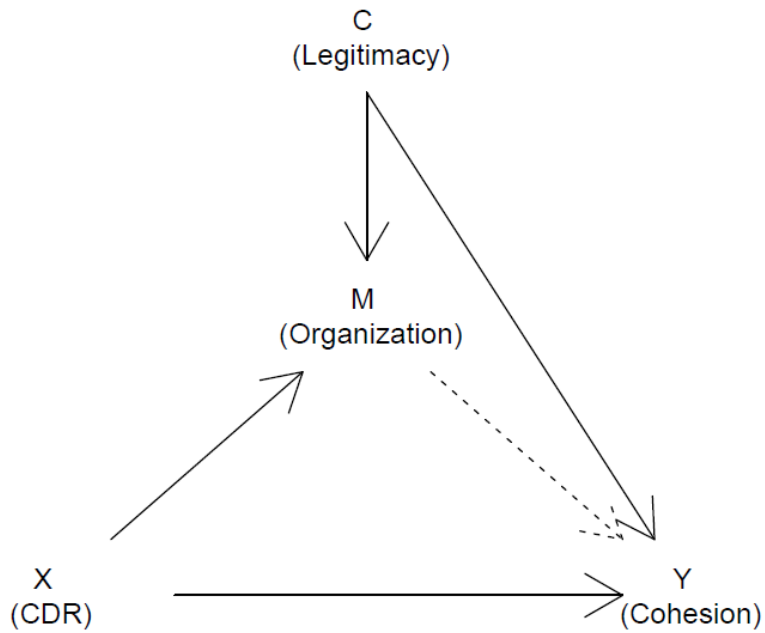
You care about the black box. Is M_1 the mechanism?

- An obvious approach is to first examine the (average) effect of X on M_1 and then use another manipulation to examine the (average) effect of M_1 on Y.
- But **both of these average effects may be positive (for example) even if there is no effect of X on Y through M1**.
- Similarly **both of these average effects may be zero even if X affects on Y through M1 for every unit!**.



The problem of unidentified mediators

- Another somewhat obvious approach is see how the effect of X on Y in a regression is reduced when you control for M. If the effect of X on Y passes through M then surely there should be no effect of X on Y after you control for M.
- But this common strategy is also not guaranteed to produce reliable results
- See Imai on better ways to think about this problem and designs to address it



The problem of unidentified mediators: Quantities

- A **mediation effect** is of the form (see Imai et al):

$$Y_i(0, M_i(1)) - Y_i(0, M_i(0))$$

- This is a well defined counterfactual quantity.
- The bad news is that this is a complex counterfactual and there is no experiment that can recover it.
- You need a model

8 Qualitative inference: Here too

Process tracing has an inference strategy

- Qualitative Inference often depends on *within case* observations
- If there are no mosquitoes then I think it unlikely that the swamp caused malaria
- Inference only possible *conditional* on the model. Not for free.
- No matter how you try to measure it, you cannot observe the causal effect. You need a model.

9 Effects of causes easier to identify than causes of effects

Consider this data

X, Y combinations:

Treatment	Share Healthy (Y=1)
X=0	25%
X=1	75%

Say this is taken from a perfect experiment.

Question: What is the average effect of X on Y?

Consider this data

X, Y combinations:

Treatment	Share Healthy (Y=1)
X=0	25%
X=1	75%

Say this is taken from a perfect experiment.

Question: For a given case with $X = Y = 1$, what are the chances that X caused Y?

Consider this data

X, Y combinations:

Treatment	Share Healthy (Y=1)
X=0	25%
X=1	75%

- It could be that in 75% of cases treatment helps people and in 25% of cases it hurts them. Then in all cases in which $X = 1$ and $Y = 1$ people are healthy *because* they were treated.
- It could also be that in 50% of cases treatment helps people, in 25% of cases they will be healthy no matter what and in 25% of cases they will be sick no matter what. In this case: the treatment helped $50/75 = 2/3$ of cases with $X = Y = 1$.
- *Experimental data cannot discriminate between these situations*

10 So: Models matter. What to do about it?

Models matter. What to do about it?

- Some of these challenges are not solvable without models
- We shouldn't even be trying to dispense with models
- Quit discussions about whether we are after causality or not
- Quit discussions about different notions of causality
- Instead focus on the question of model evaluation
  - Are our models useful or not?
  - When do they lead us astray?

Models matter: Last words

- DAG syntax is a powerful way to express core features of a model
- Models can be usefully described *after* you have conducted research but they can be much more powerful, and credible, if you develop them before hand