Introduction to causal inference

Session 11

MATH 80667A: Experimental Design and Statistical Methods
for Quantitative Research in Management
HEC Montréal

Outline

Directed acyclic graphs

Causal mediation

Directed acyclic graphs

Types of data

Experimental

You have control over which units get treatment

Observational

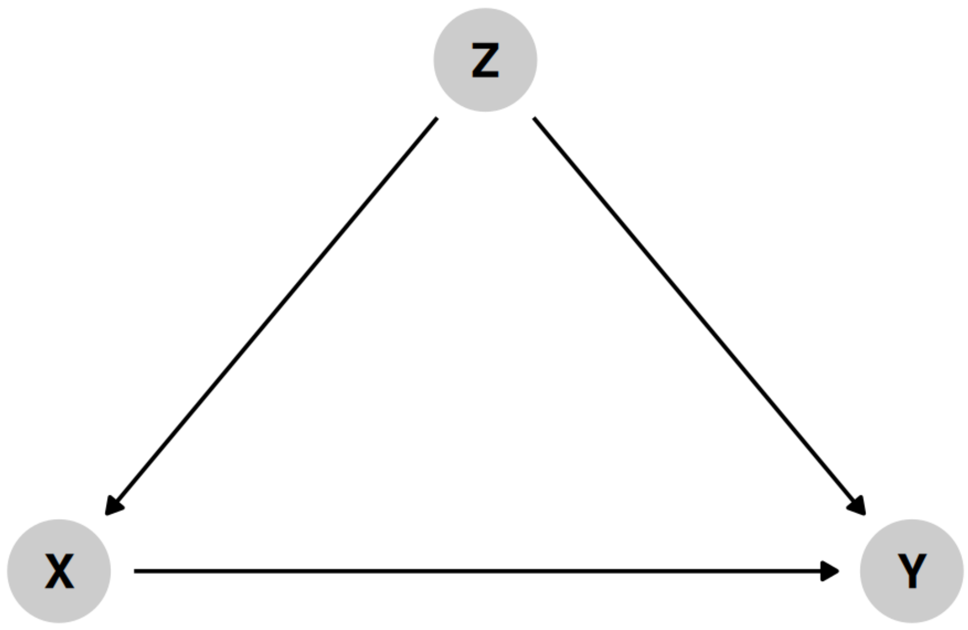
You don't have control over which units get treatment

Causal diagrams

Directed acyclic graphs (DAGs)

Directed: Each node has an arrow that points to another node

Acyclic: You can't cycle back to a node (and arrows only have one direction)



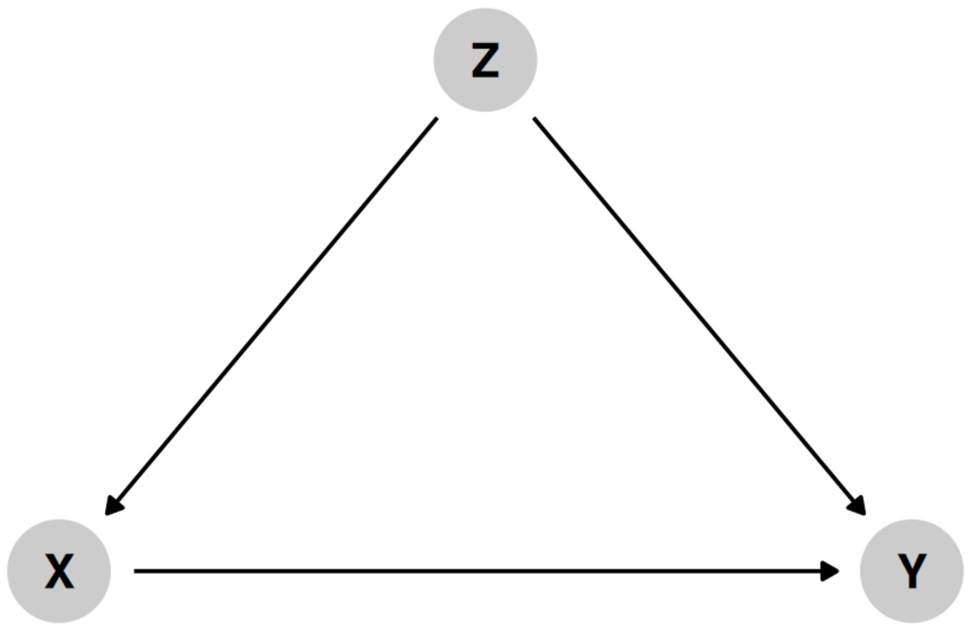
Causal diagrams

Directed acyclic graphs (DAGs)

Graphical model of the process that generates the data

Maps your philosophical model

Fancy math ("*do*-calculus") tells you what to control for to isolate and identify causation



How to draw a DAG

What is the causal effect of an additional year of education on earnings?

Step 1: List variables

Step 2: Simplify

Step 3: Connect arrows

Step 4: Use logic and math to determine which nodes and arrows to measure

1. List variables

Education (treatment) $\longrightarrow$ Earnings (outcome)

Location

Ability

Demographics

Socioeconomic status

Year of birth

Compulsory schooling laws

Job connections

2. Simplify

Education (treatment) → Earnings (outcome)

Location

Ability

Demographics

Socioeonomic status

Year of birth

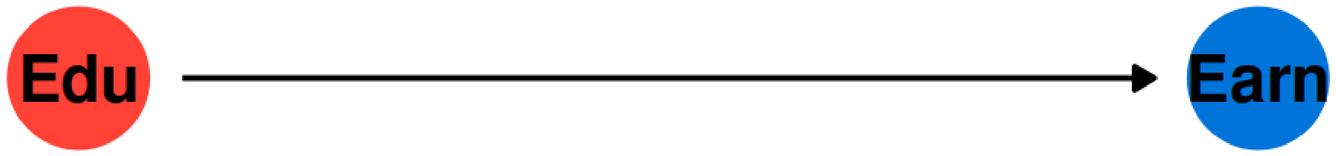
Compulsory schooling laws

Job connections

Background

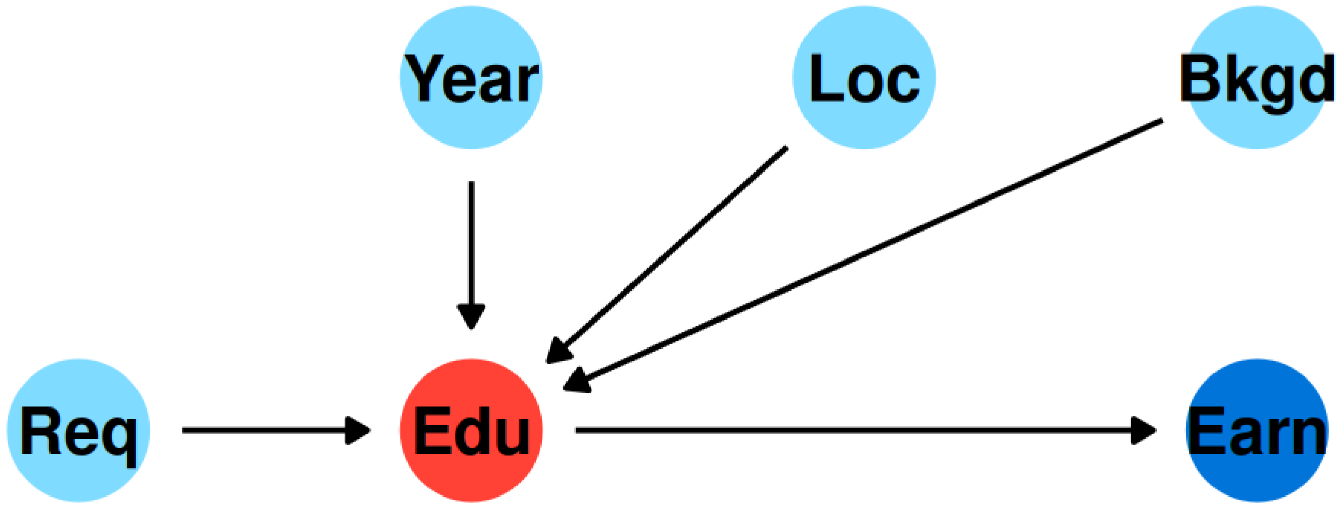
3. Draw arrows

Education causes earnings



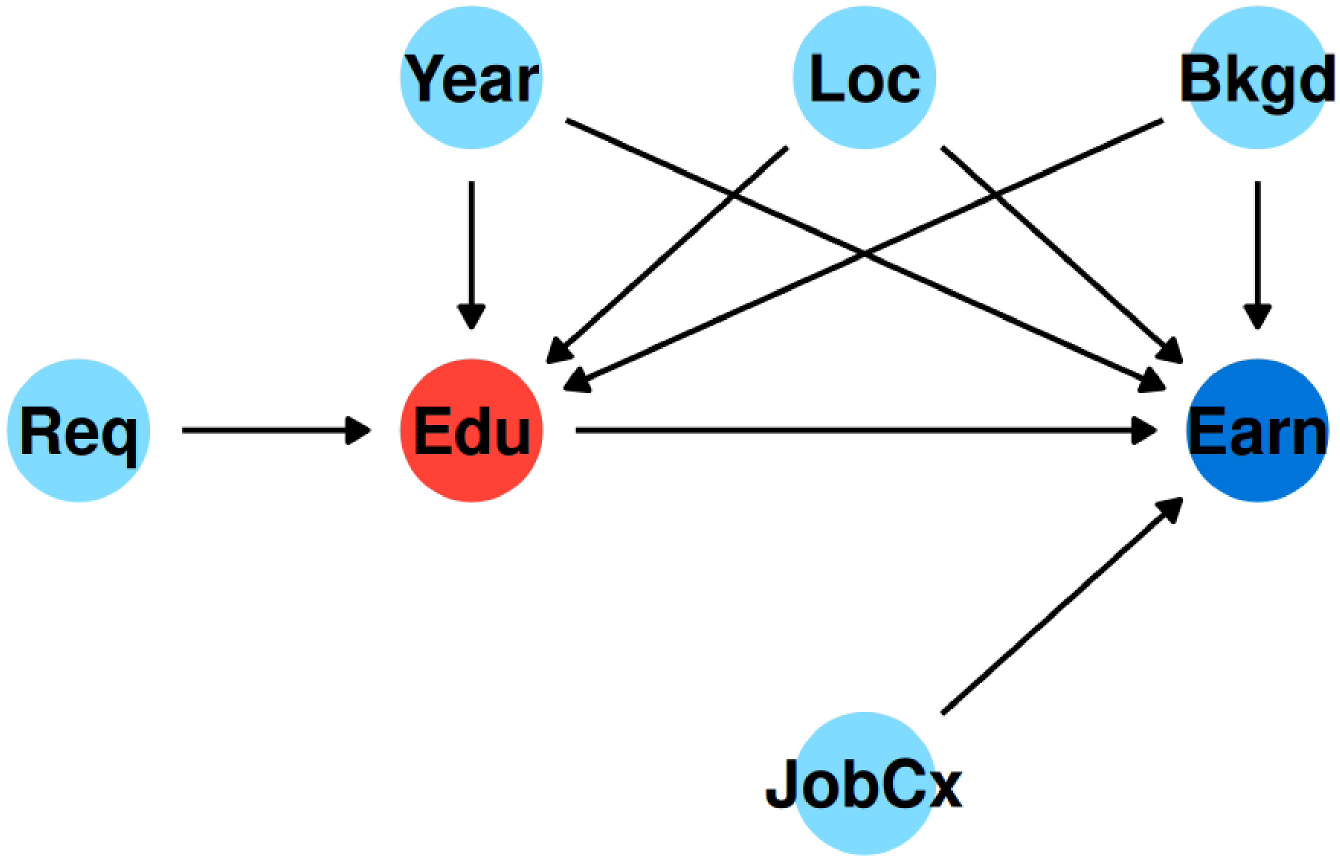
3. Draw arrows

Background, year of birth, location, job connections, and school requirements all cause education



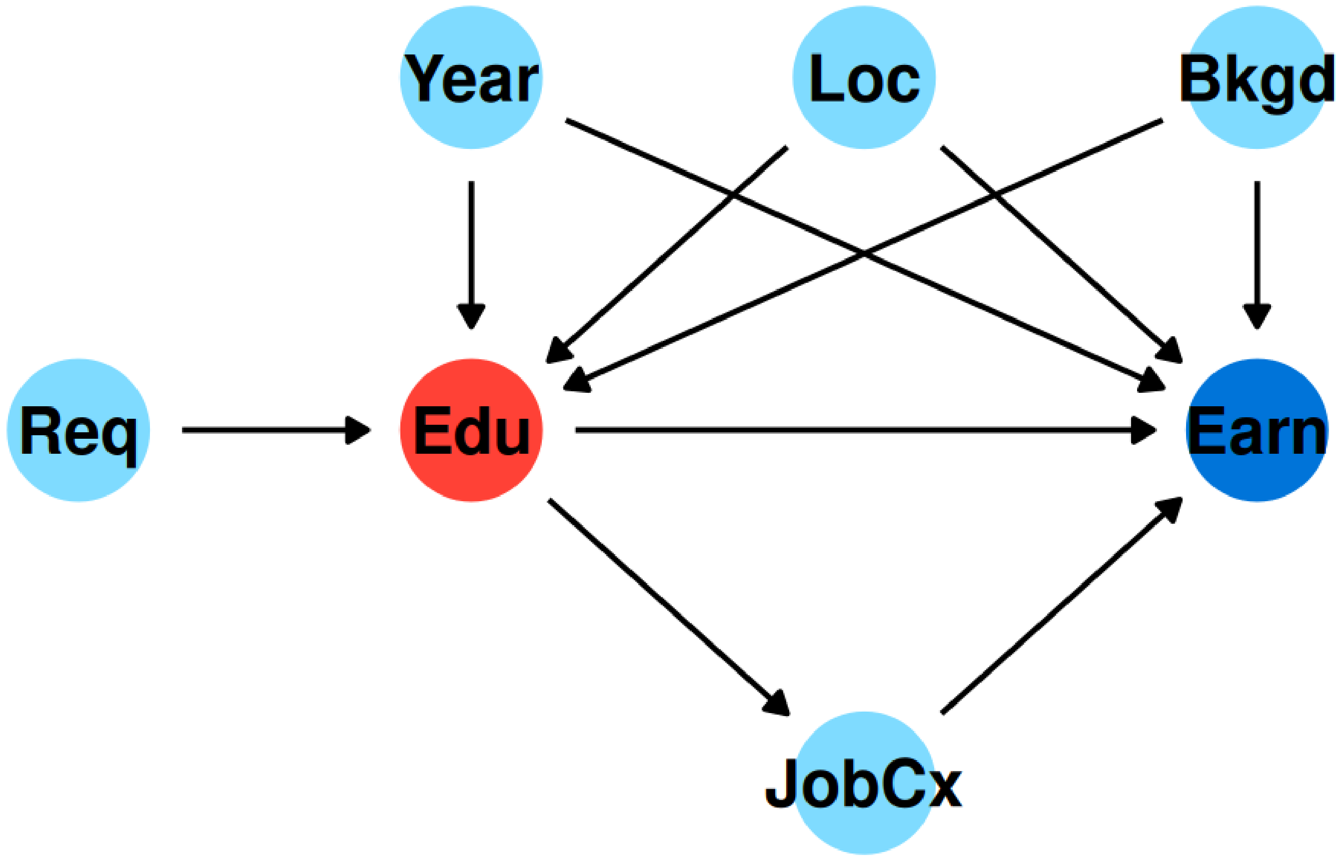
3. Draw arrows

Background, year of birth, and location all cause earnings too



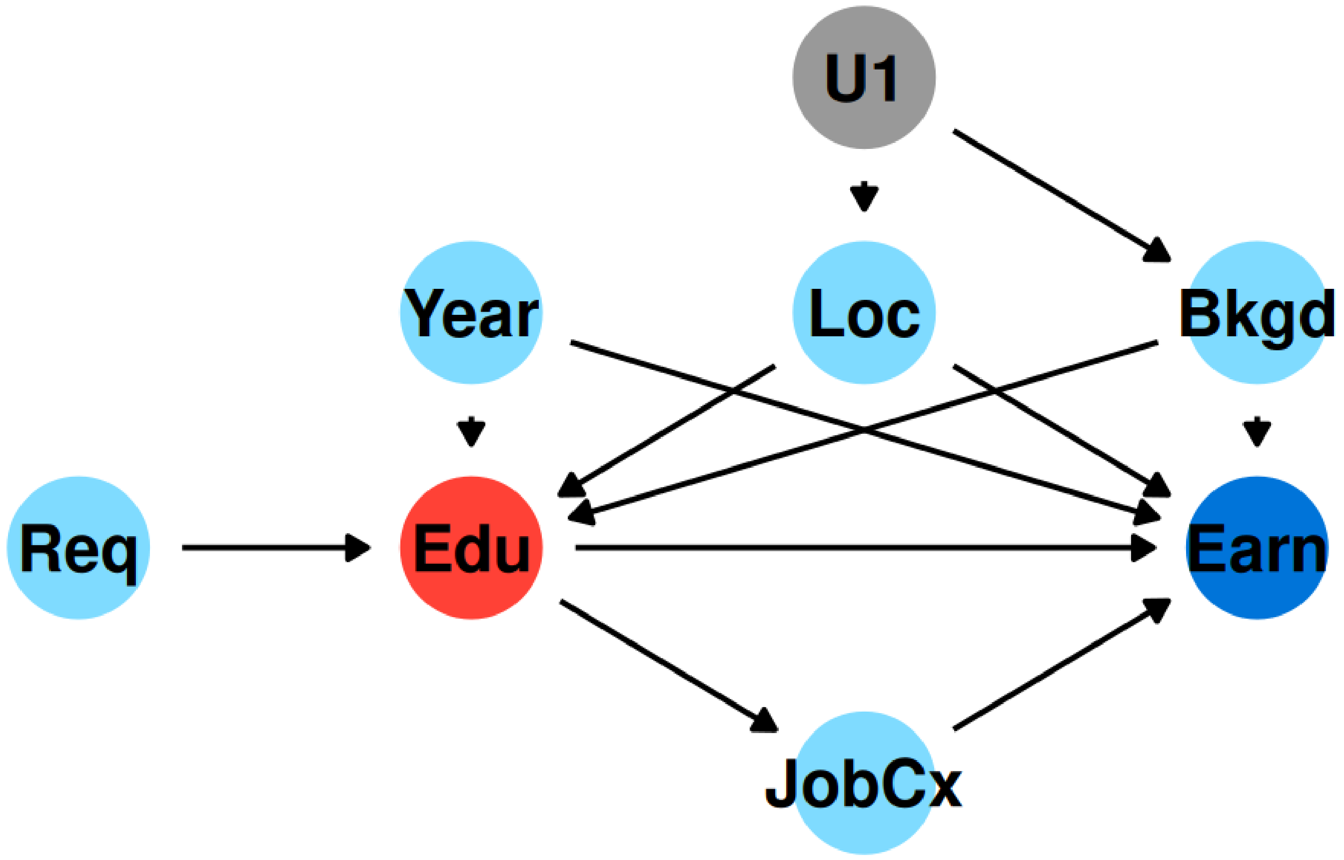
3. Draw arrows

Education causes job earnings



3. Draw arrows

Location and background are probably related, but neither causes the other. Something unobservable (U1) does that.



Causal identification

A causal effect is *identified* if the association between treatment and outcome is propertly stripped and isolated

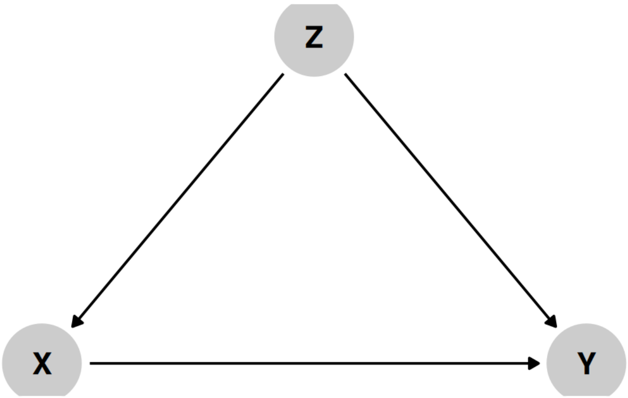
Paths and associations

Arrows in a DAG transmit associations

You can redirect and control those paths by "adjusting" or "conditioning"

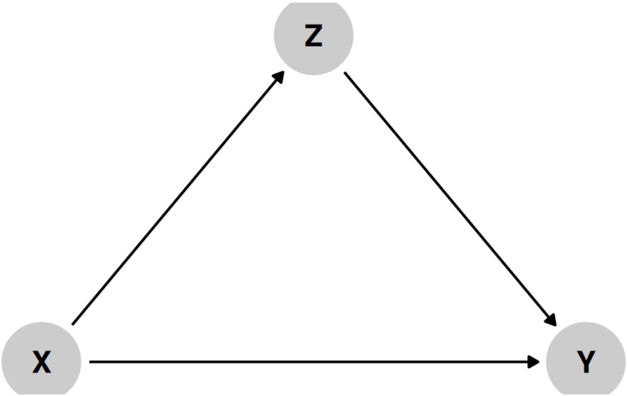
Three types of associations

Confounding



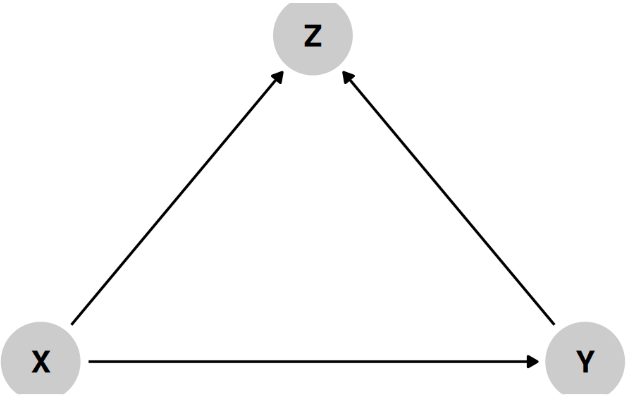
Common cause

Causation



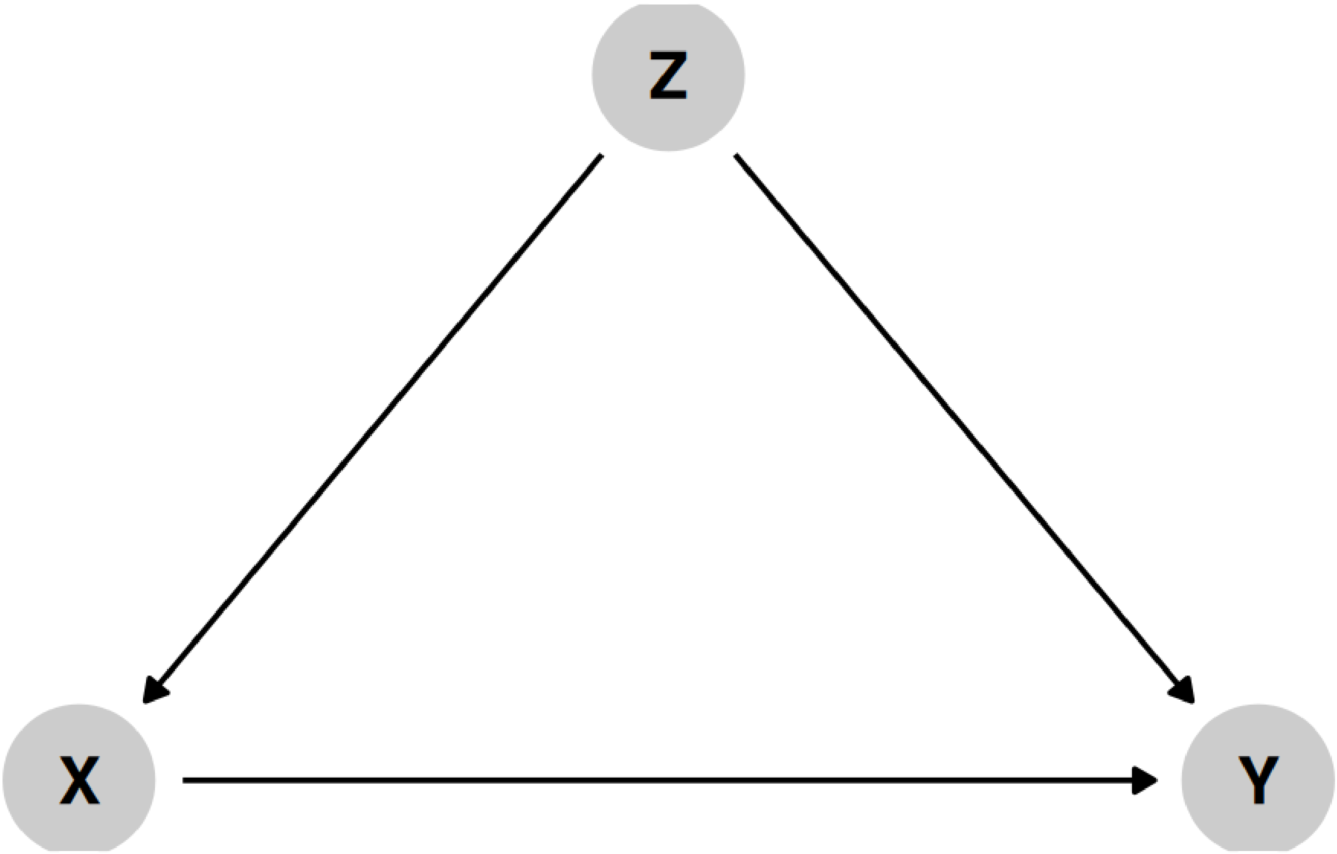
Mediation

Collision



Selection / endogeneity

Confounding

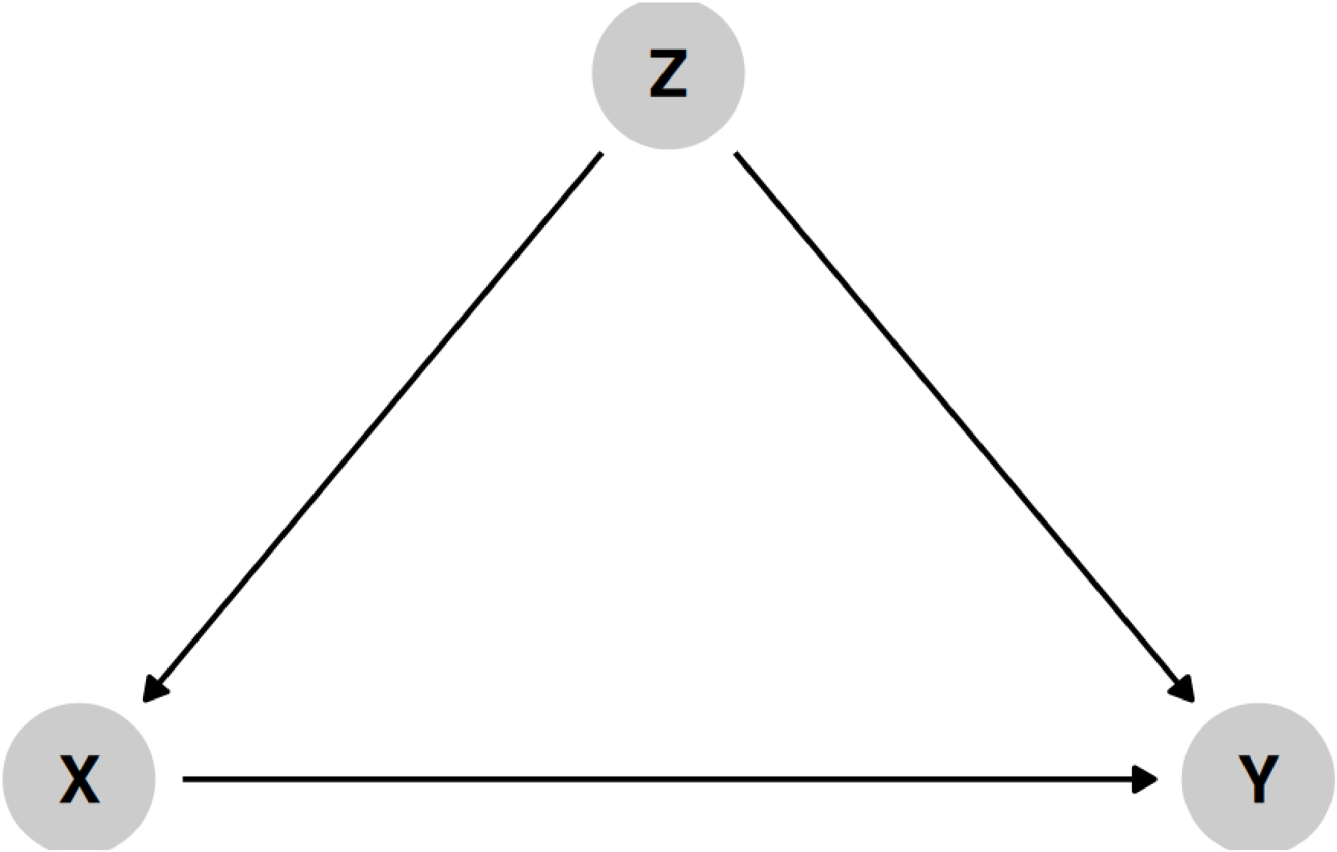


X causes Y

But Z causes both X and Y

Z *confounds* the X $\rightarrow$ Y association

Paths



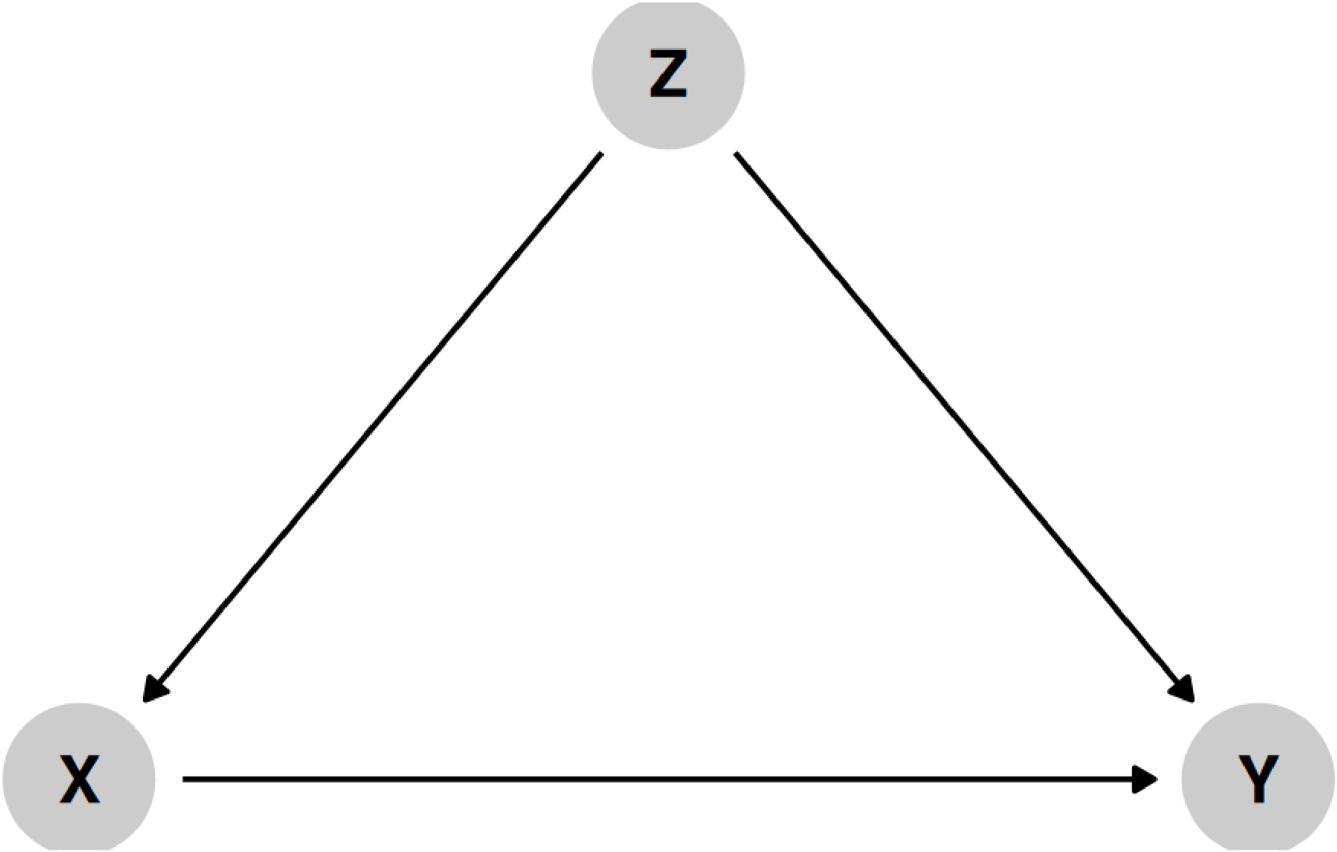
Paths between X and Y?

X → Y

X ← Z → Y

Z is a *backdoor*

d-connection

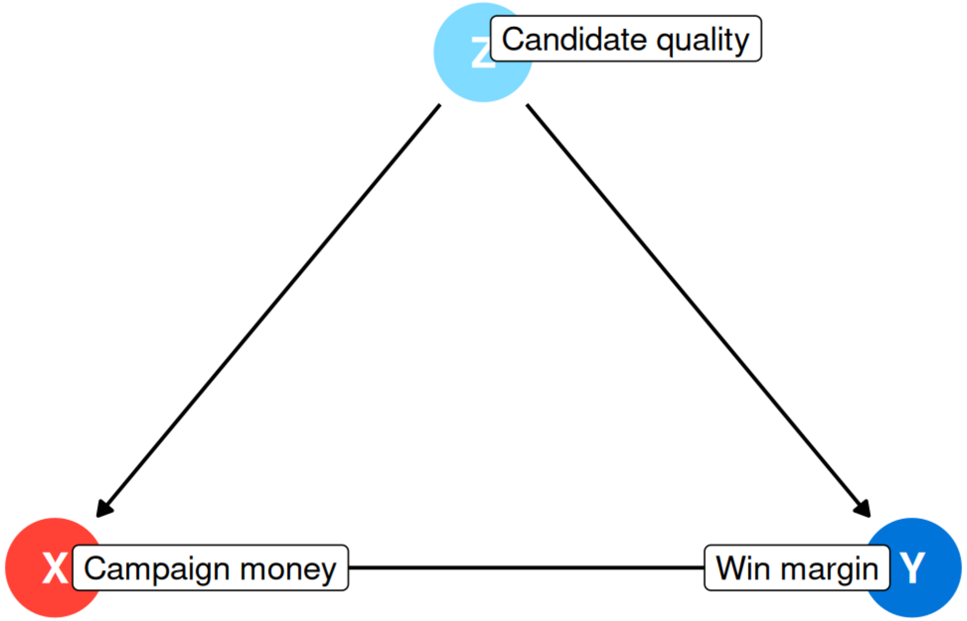


X and Y are "*d*-connected" because associations can pass through Z

The relationship between X and Y is not identified / isolated

Effect of money on elections

What are the paths between money and win margin?

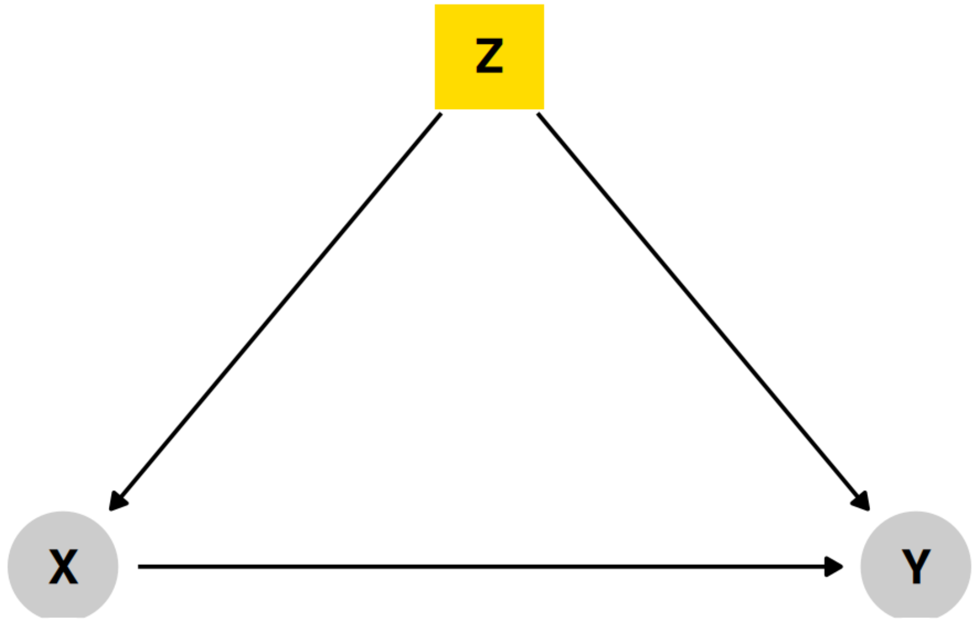


Money → Margin

Money ← Quality → Margin

Quality is a *backdoor*

Closing doors



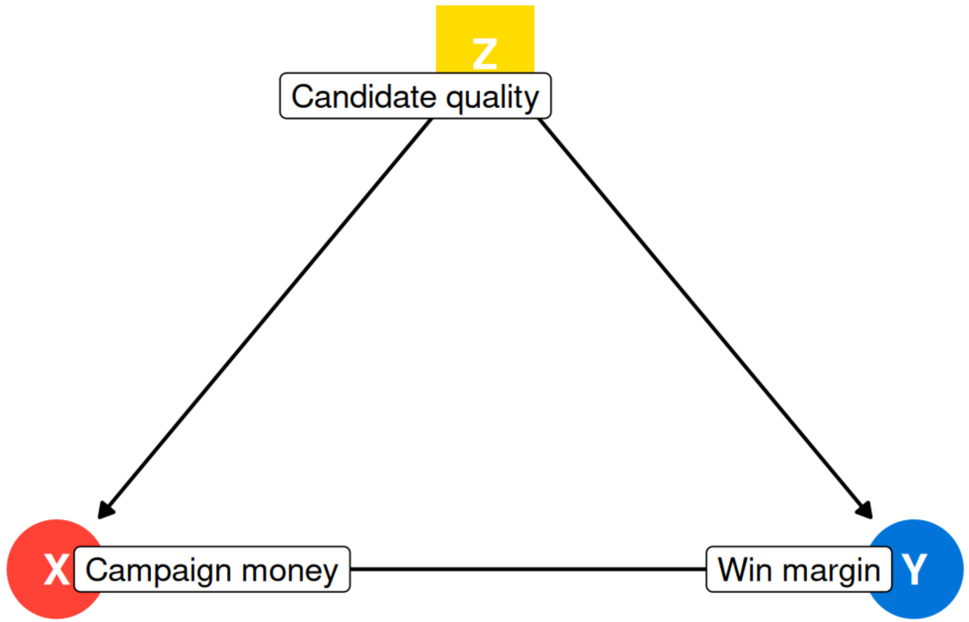
Close the backdoor by adjusting for Z

Closing doors

Find the part of campaign money that is explained by quality, remove it.
This is the residual part of money.

Find the part of win margin that is explained by quality, remove it. This is the residual part of win margin.

Find the relationship between the residual part of money and residual part of win margin.
This is the causal effect.

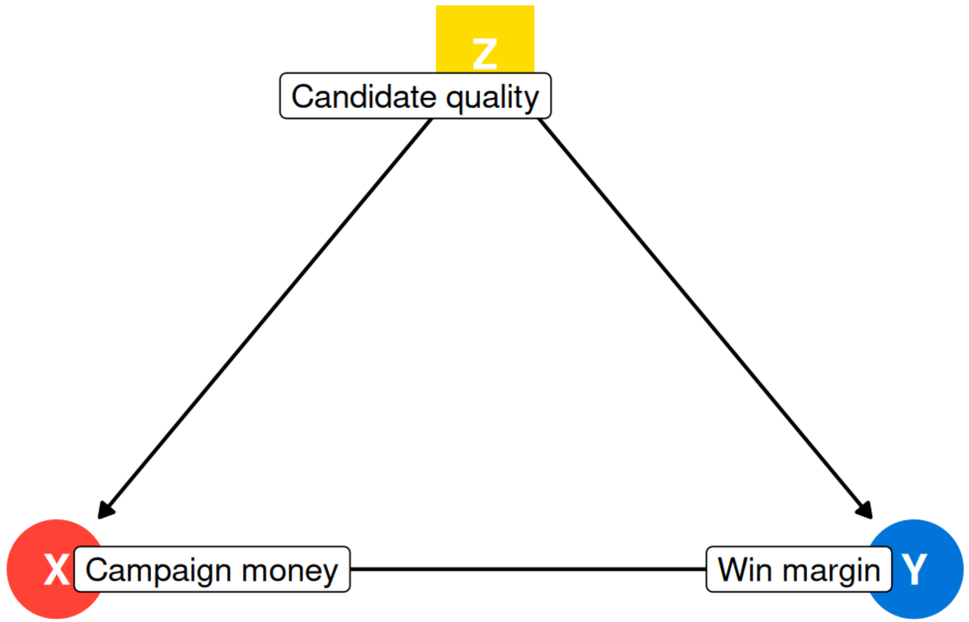


Closing doors

Compare candidates as if they had the same quality

Remove differences that are predicted by quality

Hold quality constant



How to adjust

Include term in regression

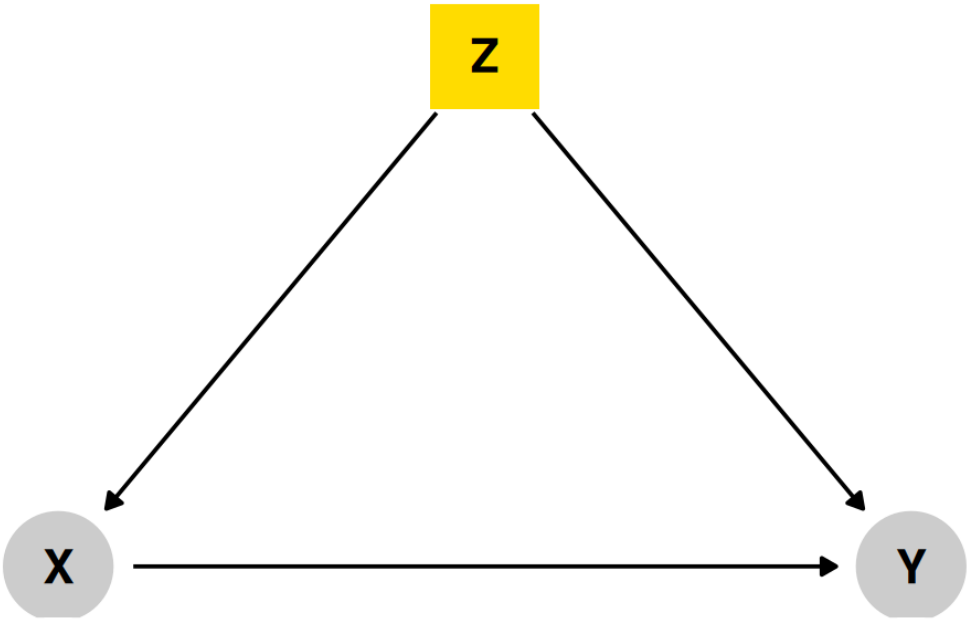
$$\text{Win margin} = \beta_0 + \beta_1 \text{Campaign money} + \beta_2 \text{Candidate quality} + \varepsilon$$

Matching

Stratifying

Inverse probability weighting

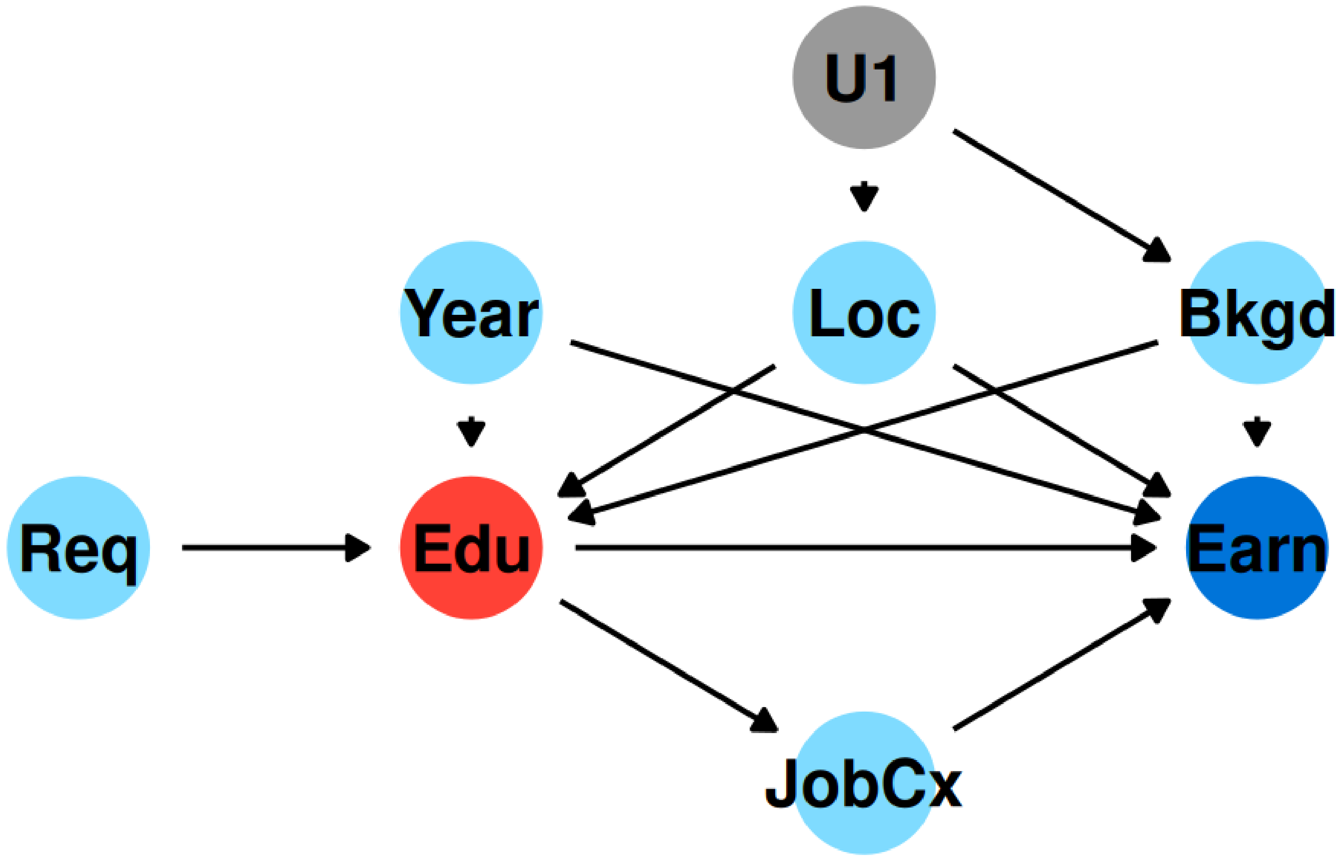
d-separation



If we control for Z, X and Y are now "*d*-separated" and the association is isolated!

Closing backdoors

Block all backdoor paths to identify the main pathway you care about



All paths

Education → Earnings

Education → Job connections → Earnings

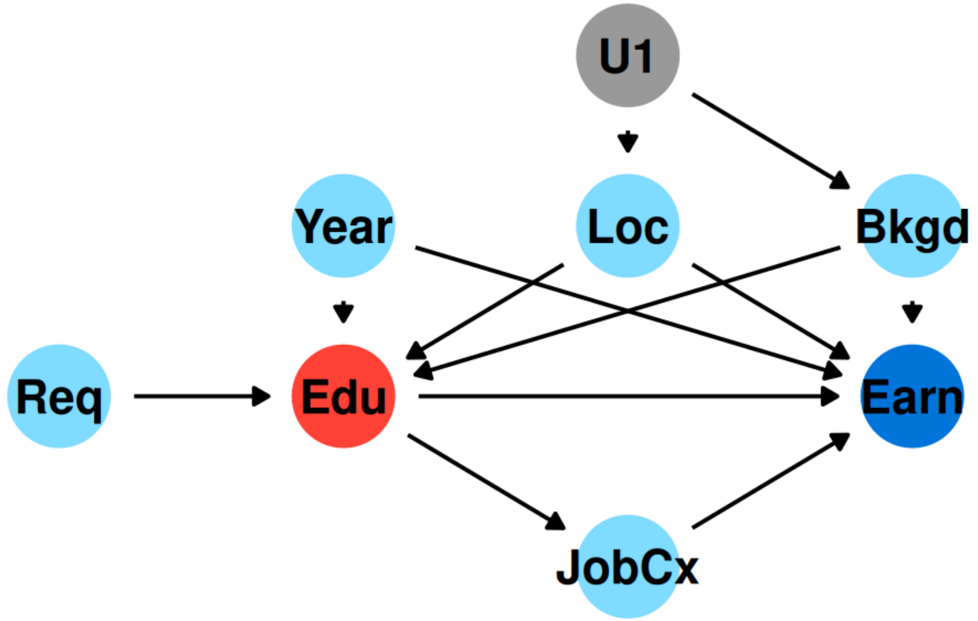
Education ← Background → Earnings

Education ← Background ← U1 → Location → Earnings

Education ← Location → Earnings

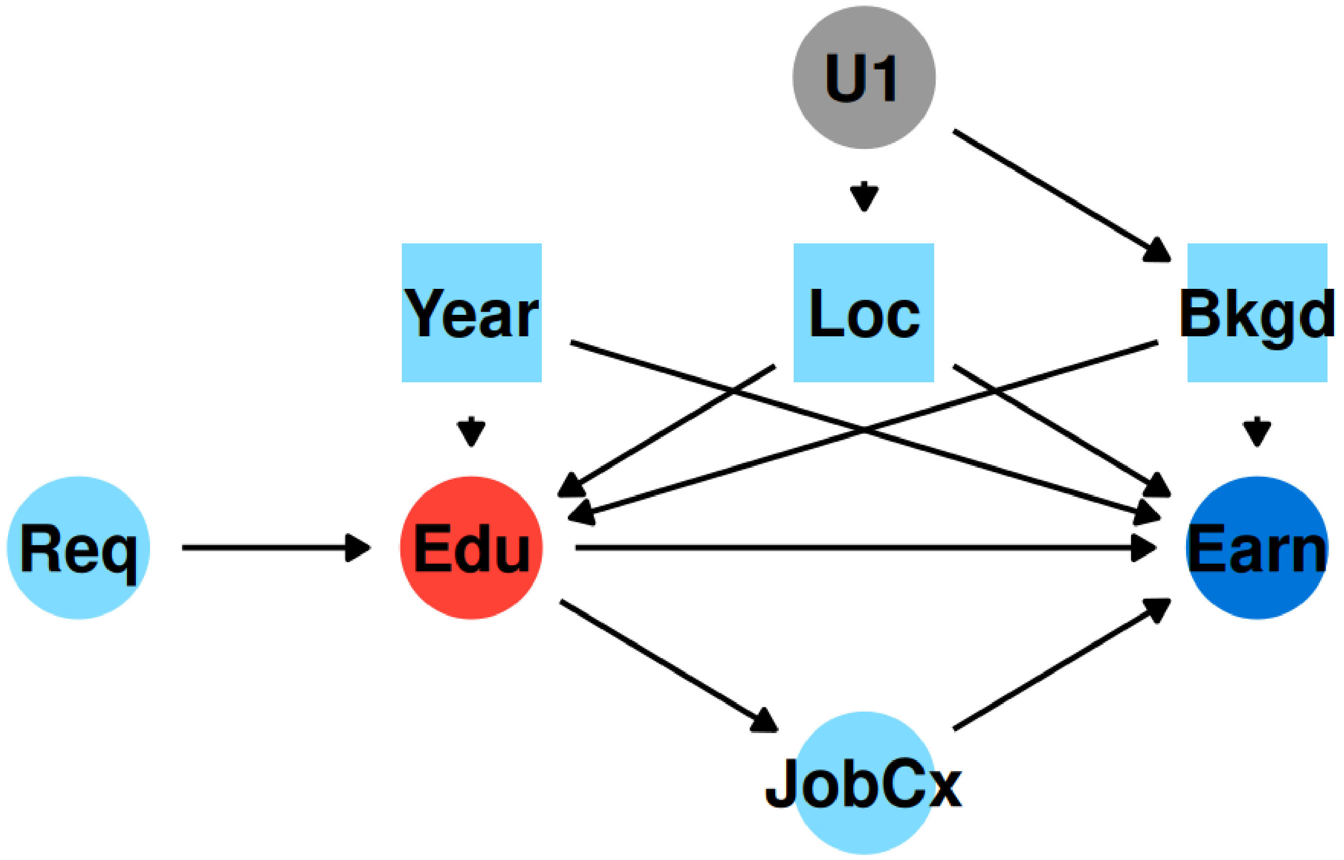
Education ← Location ← U1 → Background → Earnings

Education ← Year → Earnings



All paths

Adjust for Location, Background and Year to isolate the Education → Earnings causal effect

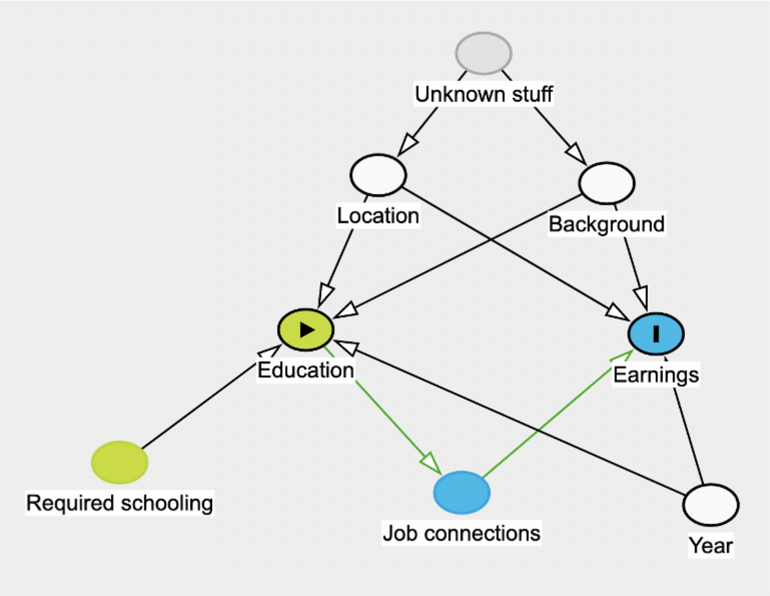


How do you know if this is right?

You can test the implications of the model to see if they're right in your data

$$X \perp Y \mid Z$$

X is independent of Y, given Z

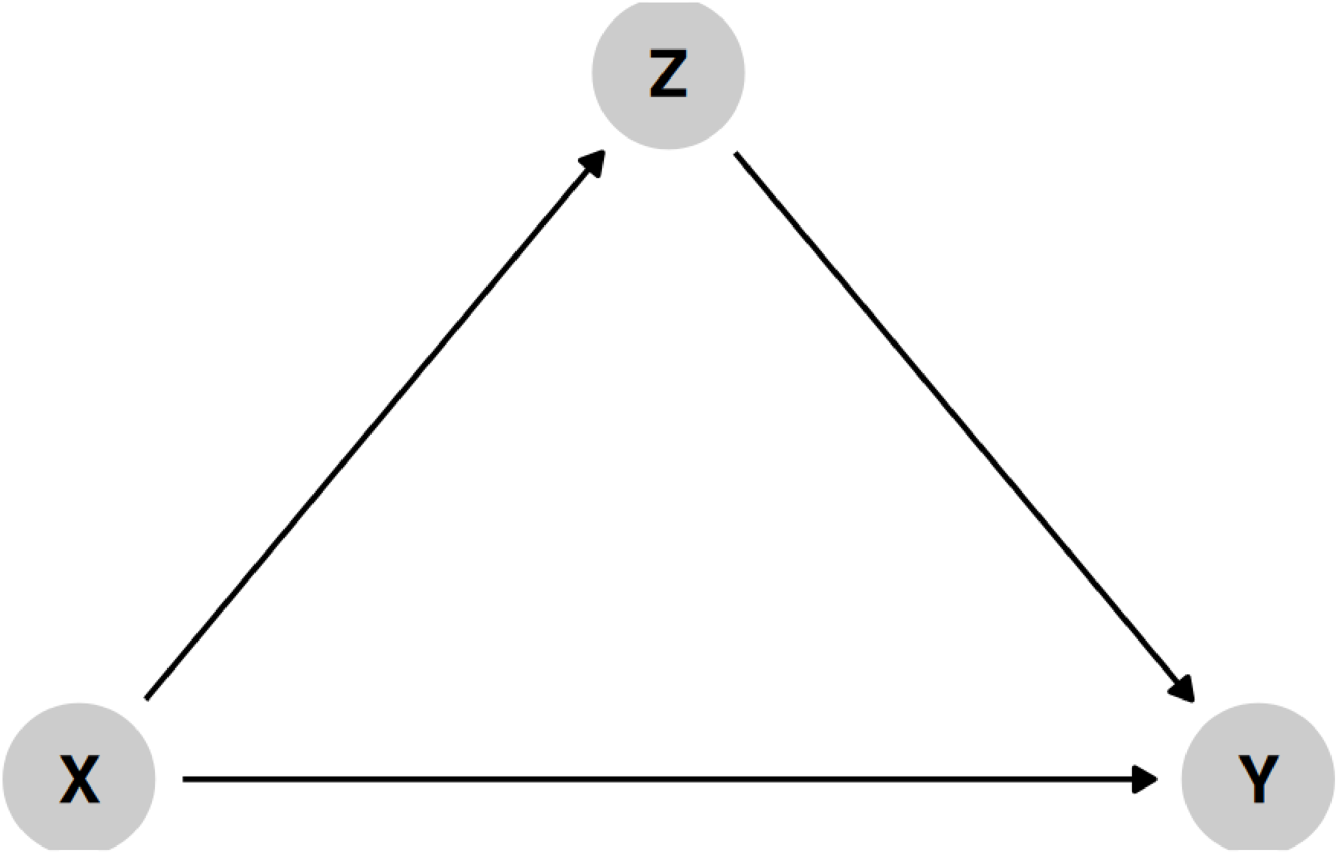


Testable implications

The model implies the following conditional independences:

- Education ⊥ Earnings | Background, Job connections, Location, Year
- Required schooling ⊥ Job connections | Education
- Required schooling ⊥ Year
- Required schooling ⊥ Earnings | Background, Job connections, Location, Year
- Required schooling ⊥ Earnings | Background, Education, Location, Year
- Required schooling ⊥ Background
- Required schooling ⊥ Location
- Job connections ⊥ Year | Education
- Job connections ⊥ Background | Education
- Job connections ⊥ Location | Education
- Year ⊥ Background
- Year ⊥ Location

Causation

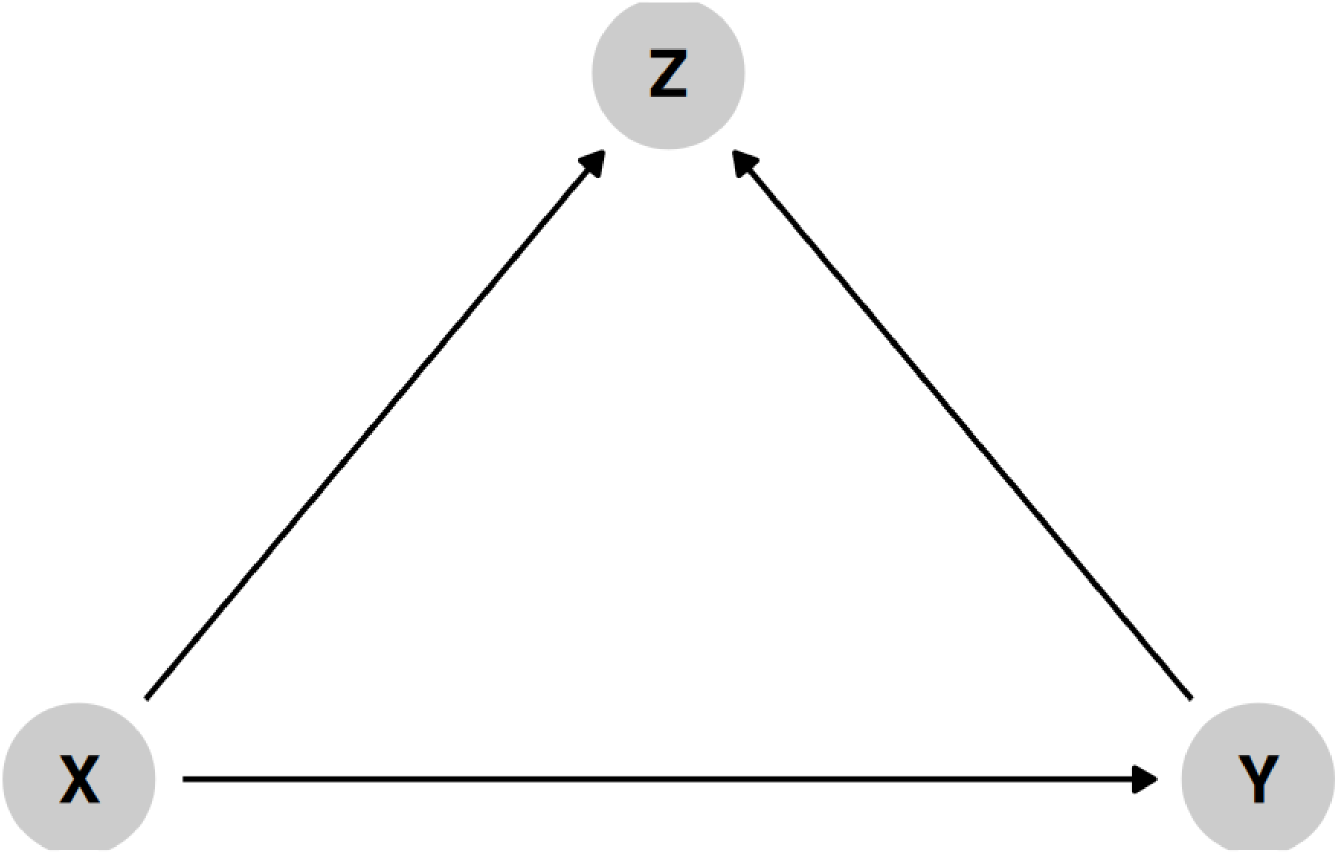


X causes Y

X causes Z which causes Y

Z is a mediator

Colliders



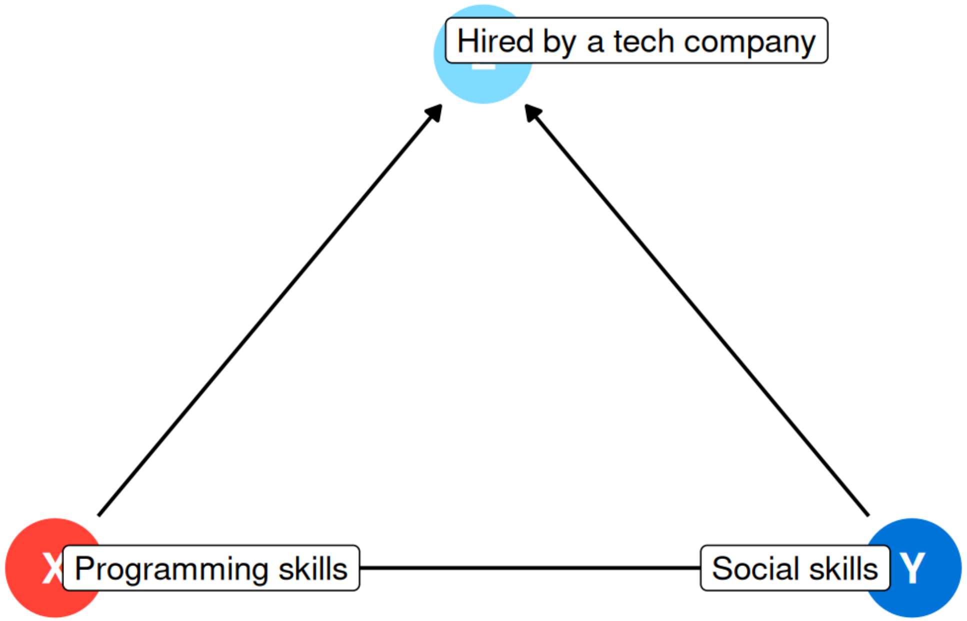
X causes Z

Y causes Z

Should you control for Z?

Programming and social skills

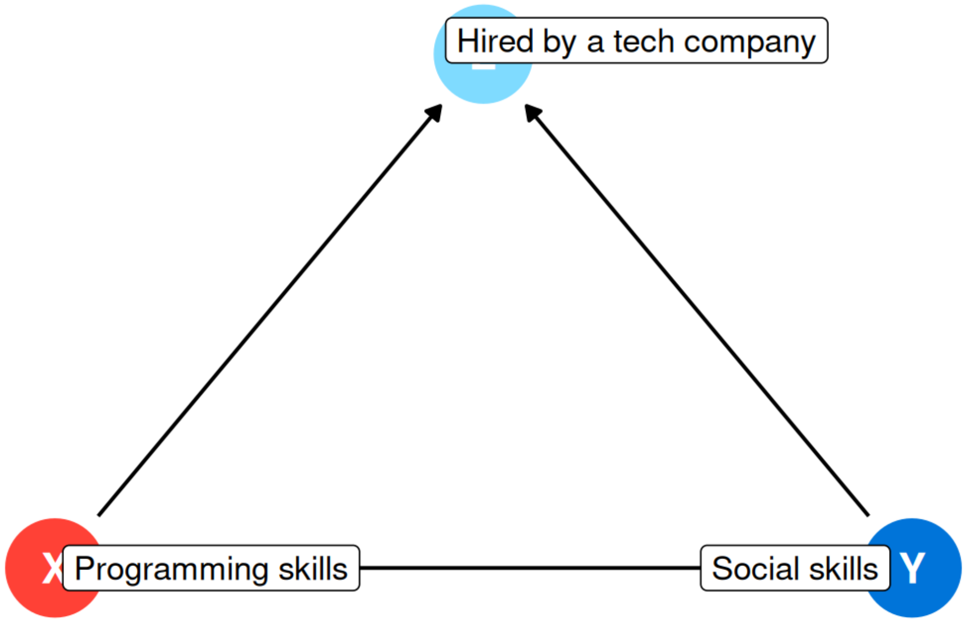
Do programming skills reduce social skills?



You go to a tech company and conduct a survey. You find a negative relationship!
Is it real?

Programming and social skills

Do programming skills reduce social skills?

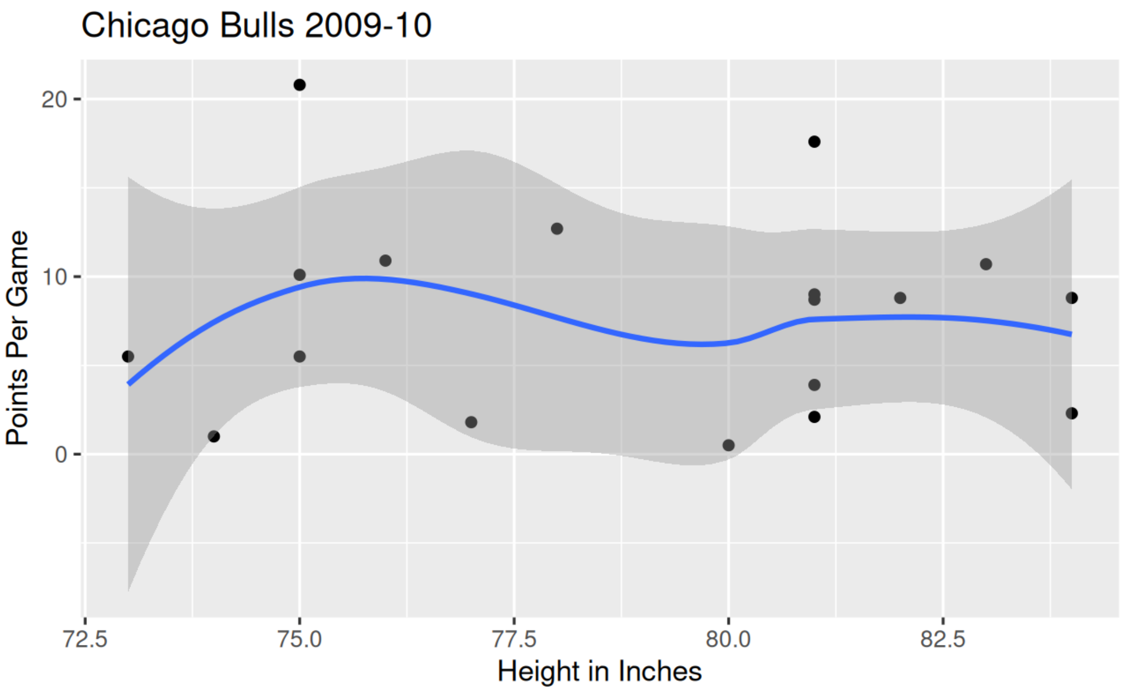


No! Hired by a tech company is a collider and we controlled for it.

This inadvertently connected the two.

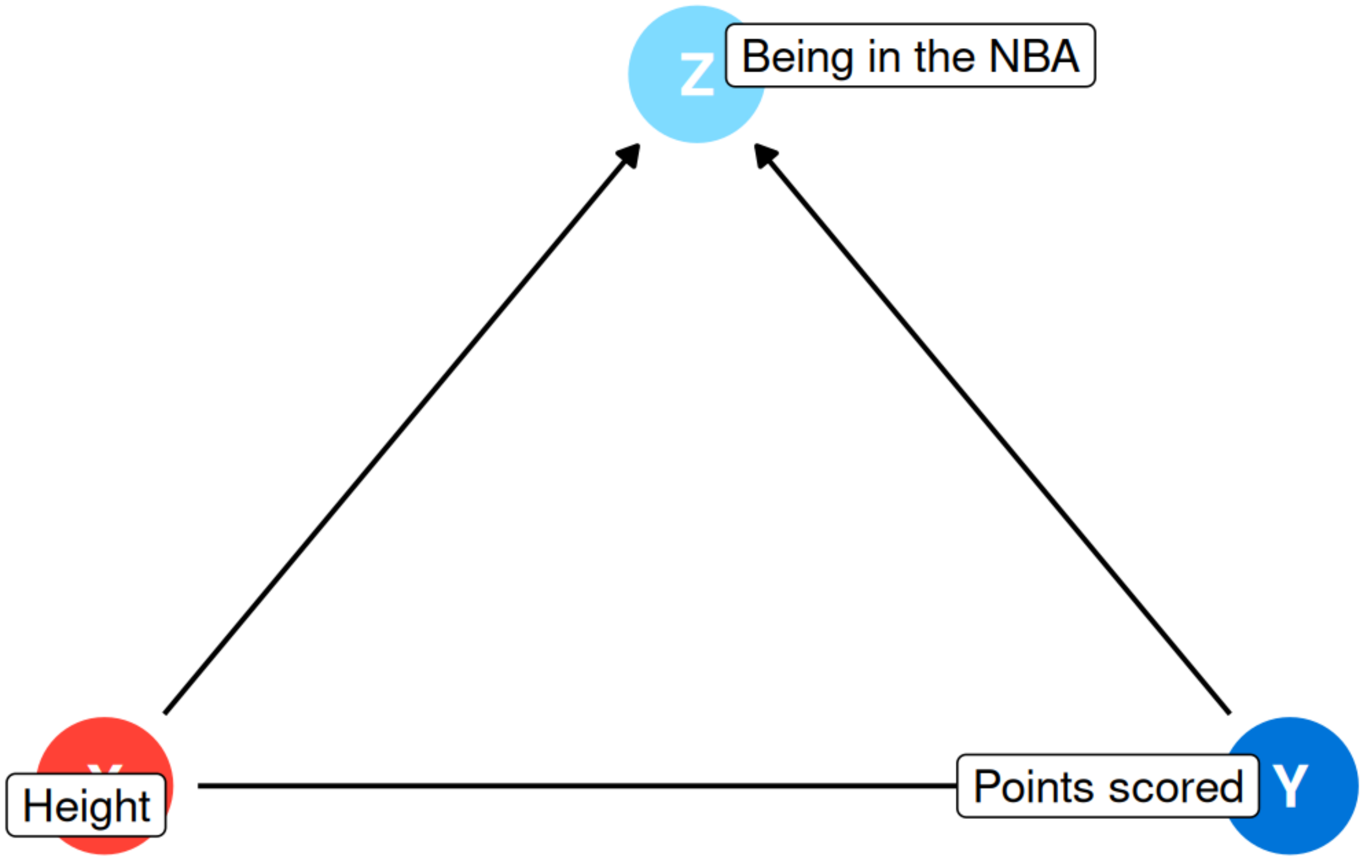
Colliders can create fake causal effects

Colliders can hide real causal effects

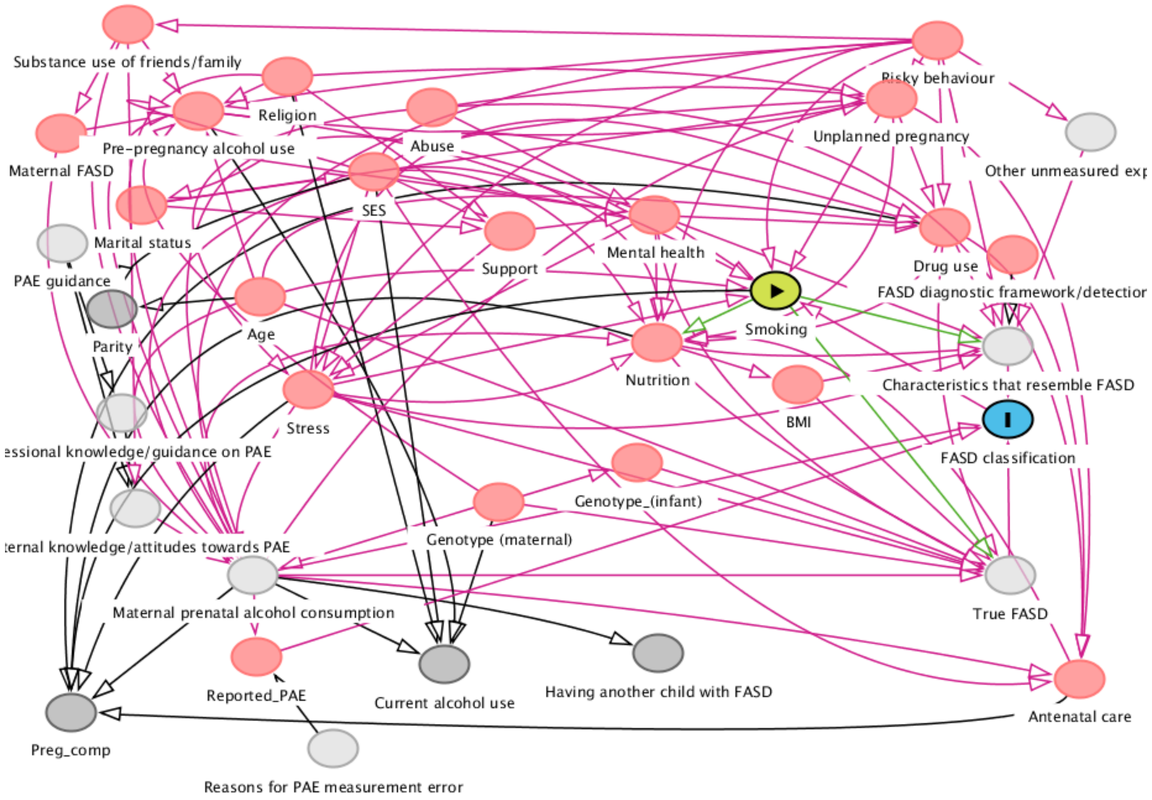


Height is unrelated to basketball skill… among NBA players

Colliders and selection bias



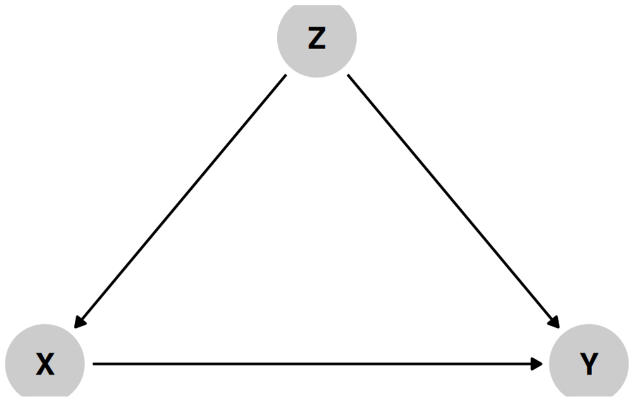
Life is inherently complex



Postulated DAG for the effect of smoking on fetal alcohol spectrum disorders (FASD)

Three types of associations

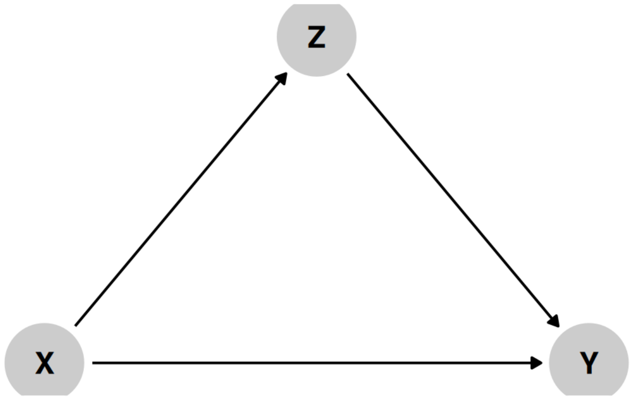
Confounding



Common cause

Causal forks **X** ← **Z** → **Y**

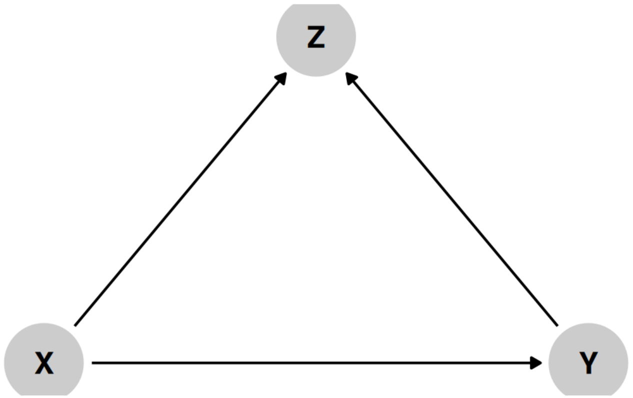
Causation



Mediation

Causal chain **X** → **Z** → **Y**

Collision



Selection / endogeneity

inverted fork **X** → **Z** ← **Y**

Causal mediation

Key references

- Imai, Keele and Tingley (2010), A General Approach to Causal Mediation Analysis, *Psychological Methods*.
- Pearl (2014), Interpretation and Identification of Causal Mediation, *Psychological Methods*.
- Baron and Kenny (1986), The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations, *Journal of Personality and Social Psychology*

Limitations:

- Bullock, Green, and Ha (2010), Yes, but what's the mechanism? (don't expect an easy answer)
- Uri Simonsohn (2022) Mediation Analysis is Counterintuitively Invalid

Fundamental problem of causal inference

Observe outcome for a single treatment

With binary treatment X_i, I observed either $Y_i \mid \mathrm{do}(X_i = 1)$ or $Y_i \mid \mathrm{do}(X_i = 0)$ given intervention **do**.

Thus define causal effect as an **average** treatment

$$\mathsf{E}[Y_i \mid \mathrm{do}(X_i = 1)] - \mathsf{E}[Y_i \mid \mathrm{do}(X_i = 0)]$$

Effect cannot be estimated directly in general, even with randomized experiments.

Sequential ignorability assumption

Define

- treatment of individual i as X_i,
- potential mediation given treatment x as $M_i(x)$ and
- potential outcome for treatment x and mediator m as $Y_i(x, m)$.

Given pre-treatment covariates W, potential outcomes for mediation and treatment are conditionally independent of treatment assignment.

$$Y_i(x', m), M_i(x) \perp\!\!\!\perp X_i \mid W_i = w$$

Given pre-treatment covariates and observed treatment, potential outcomes are independent of mediation.

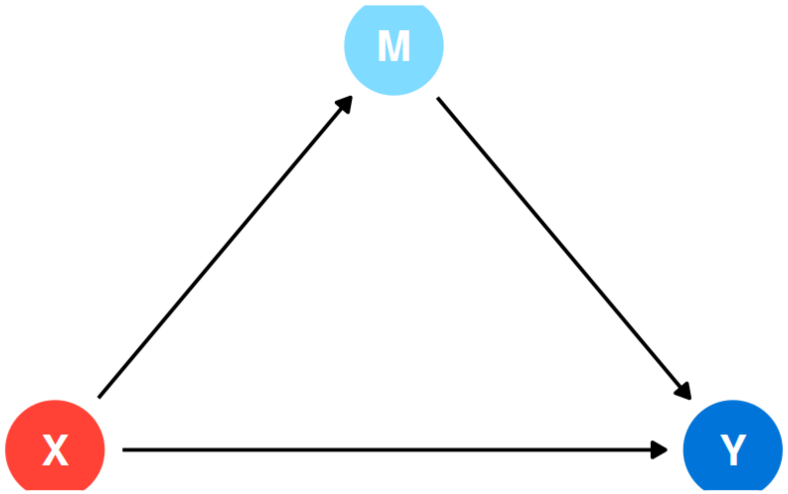
Total effect

Total effect: overall impact of X (both through M and directly)

$$\mathsf{TE}(x, x^*) = \mathsf{E}[Y \mid \mathrm{do}(X = x)] - \mathsf{E}[Y \mid \mathrm{do}(X = x^*)]$$

This can be generalized for continuous X to any pair of values (x_1, x_2).

X → M → Y
plus
X → Y



Average controlled direct effect

$$\begin{aligned}\mathsf{CDE}(m, x, x^*) &= \mathsf{E}[Y \mid \mathrm{do}(X = x, m = m)] - \mathsf{E}[Y \mid \mathrm{do}(X = x^*, m = m)\\ &= \mathsf{E}\{Y(x, m) - Y(x^*, m)\}\end{aligned}$$

Expected population change in response when the experimental factor changes from x to x^* and the mediator is set to a fixed value m.

Direct and indirect effects

Natural direct effect: $\text{NDE}(x, x^*) = \text{E}[Y\{x, M(x^*)\} - Y\{x^*, M(x^*)\}]$

- expected change in Y under treatment x if M is set to whatever value it would take under control x^*

Natural indirect effect: $\text{NIE}(x, x^*) = \text{E}[Y\{x^*, M(x)\} - Y\{x^*, M(x^*)\}]$

- expected change in Y if we set X to its control value and change the mediator value which it would attain under x

Counterfactual conditioning reflects a physical intervention, not mere (probabilistic) conditioning.

Total effect is $\text{TE}(x, x^*) = \text{NDE}(x, x^*) - \text{NIE}(x^*, x)$