

Post-treatment conditioning

Han Zhang

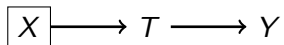
- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

Two representations of causal analysis

- Counterfactual: y^1 and y^0
- DAG: graphical representations of the *theorized* data-generating process.

- Node: random variables.
 - Can be observed or unobserved. Just variables that the theory believe to be relevant.
 - rectangle means unobserved.
- Edge (arrow): directly causal effect between two variables.
 - Only direct relations. E.g., the effect of X on Y is through T .
 - Missing edges between nodes encode exclusion restrictions.

$$X \longrightarrow T \longrightarrow Y$$



Elements

- Path: a sequence of edges connecting two variables (Do not consider direction)
- Causal path: following the directions and from treatment to outcome.
 $T \rightarrow Y$ and $T \rightarrow C \rightarrow Y$.
- Noncausal path: E.g., $T \rightarrow C \leftarrow U \rightarrow Y$. Brings spurious associations between treatment and outcomes.

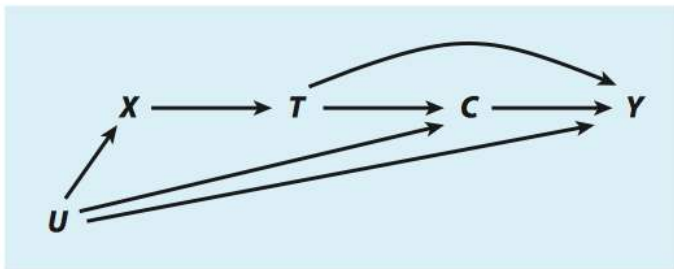


Figure 1

- Path analysis
- Often assume linear relationships
- Hard to represent interaction

DAG (Directed Acyclic Graph)

- DAG encodes all marginal (unconditional) and conditional independence relations
- There is no parametric assumptions now.

- DAG helps to make sure whether the causal relations can be identified.
- identification: given ideal data (non response errors etc.), whether the causal relationship can be recovered in a DAG; if so, under what condition.
- Estimation: choose real statistical models (regression, matching ...).

Simple case

Examine whether $T \perp\!\!\!\perp Y$

$$T \longrightarrow Y$$

Mutual Dependence

- Confounding bias
- Condition on C eliminates confounding bias
- E.g., $A \not\perp B$, but $A \perp B \mid C$.

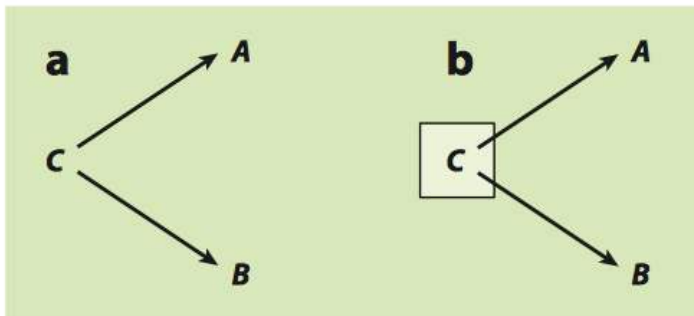
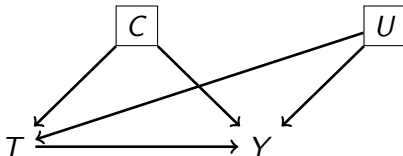


Figure 3

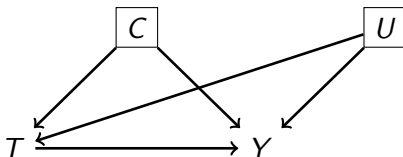
Mutual Dependence

When you have unobserved variables U , $T \rightarrow Y$ remain unidentified.



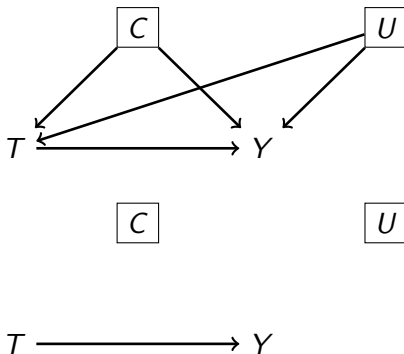
What does experiments do

- Randomization breaks the links from other variables to treatment/outcomes.



What does experiments do

- Randomization breaks the links from other variables to treatment/outcomes.



Mutual Causation

- Endogenous selection bias
- Condition on C **create** endogenous selection bias.

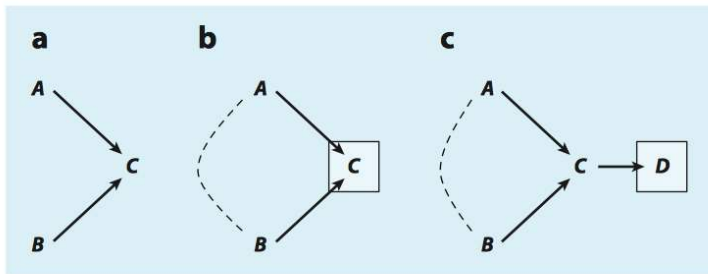


Figure 4

Mutual Causation

- $A \perp\!\!\!\perp B, p(A, B) = p(A)p(B).$
- $A \not\perp\!\!\!\perp B \mid C \iff p(A, B \mid C) \neq p(A \mid C)p(B \mid C),$ or
- $p(A \mid B, C) \neq p(A \mid C)$ (giving B as information influence conditional probability of A given C)
- Conditioning on C gives wrong implication that A, B are causally related.

Mutual Causation

- $A \perp\!\!\!\perp B, p(A, B) = p(A)p(B).$
- $A \not\perp\!\!\!\perp B \mid C \iff p(A, B \mid C) \neq p(A \mid C)p(B \mid C),$ or
- $p(A \mid B, C) \neq p(A \mid C)$ (giving B as information influence conditional probability of A given C)
- Conditioning on C gives wrong implication that A, B are causally related.

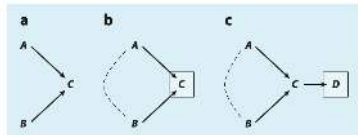


Figure 4

- Example: A is exam difficulty; B is student intelligence; C is test scores.
- Knowing low test score, we can infer A from $p(A \mid C)$ that the exam will possibly be hard; but with further information of student intelligence, $p(A \mid B, C) \neq p(A \mid C)$ typically.

Mediation

- Overcontrol bias
 - $A \not\perp\!\!\!\perp B$, $p(A, B) \neq p(A)p(B)$.
 - $A \perp\!\!\!\perp B \mid C$, or
 - But this only means that the direct effect of A on B is zero; the total effect is not zero.
- Condition on C **create** overcontrol bias.
- More about interpretation of total/direct effect.



Figure 2

(a) A and B are associated by causation. The marginal association between A and B identifies the causal effect of A on B . (b) A and B are conditionally independent given C . The conditional association between A and B given C does not identify the causal effect of A on B (overcontrol bias).

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

Identification in General DAGs

A path between two variables, A and B , does not transmit association and is said to be blocked (closed, or d-separated) if

- the path contains a noncollider, C , that has been conditioned on
 - $A \rightarrow \boxed{C} \rightarrow B$ (mediation)
 - $A \leftarrow \boxed{C} \rightarrow B$ (confounding)

or if

- the path contains a collider, C , and neither the collider nor any of its descendants have been conditioned on.

Goal: try to find an observed set of variables as conditions that:

- block all noncausal paths between treatment and outcome.
- do not block any causal paths between treatment and outcome.

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)**
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

Sample truncation bias

Effect of education on income (truncated to only contain low earners); U are other observed errors (embedded in conventional control variables).

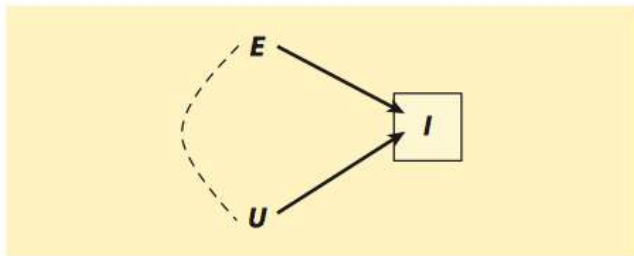


Figure 5

Other examples

- social movement: only sample protests that have 1,000 participants.

Nonresponse bias

A divorced father's income, I , and the amount of child support he pays, P , both influence whether a father responds to the study, R .

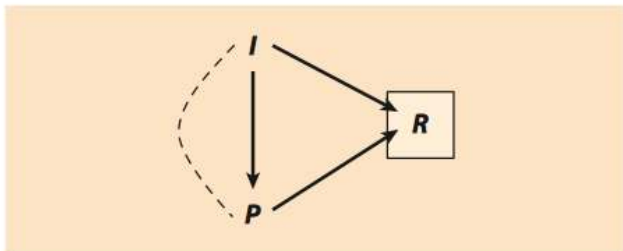


Figure 6

- Nonresponse is causally determined by both treatment and outcome.
- List deletions of non-responding subjects implies conditioning on R .
- Does multiple imputation also introduce nonresponse bias?

Sample selection bias

B , topping the Billboard charts (treatment); R , inclusion in the Rolling Stone 500 (outcome); S , sample selection]

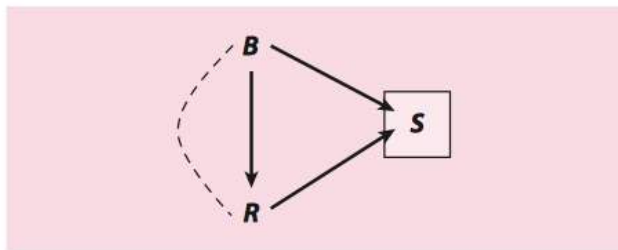


Figure 7

Heckman selection bias

M , motherhood (treatment); W_R , unobserved reservation wage; W_O , offer wage (outcome); E , employment; ϵ , error term on offer wage]

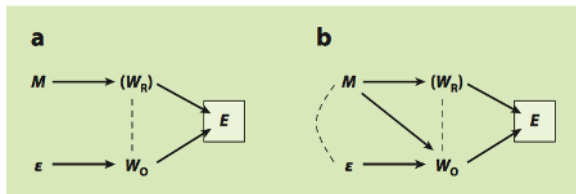


Figure 8

- (a): many dataset restricted attention to employed women (conditioning on colliders): an association between motherhood and wages even if the causal effect of motherhood on wages is in fact zero.
- (b): motherhood may indeed have an effect on offer wages (e.g., because of mothers' differential productivity compared with childless women or because of employer discrimination)
- Even if W_R is measured, there is still a collider W_O .

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 **Overcontrol bias**
- 5 Condition on pre-treatment colliders

- Censoring: only dependent variables are observed
- (compare with) Truncated: no information at all is available for the non-selected observations.
- Choices
 - handle censoring using duration models (survival analysis, event history models).
 - discard

- P : poverty (treatment)
- D : divorce rates (outcome)
- C : censoring/attrition.
- U : unmeasured factors, such as marital distress.

- P : poverty (treatment)
- D : divorce rates (outcome)
- C : censoring/attrition.
- U : unmeasured factors, such as marital distress.

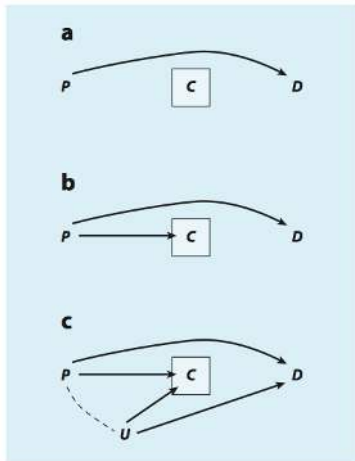
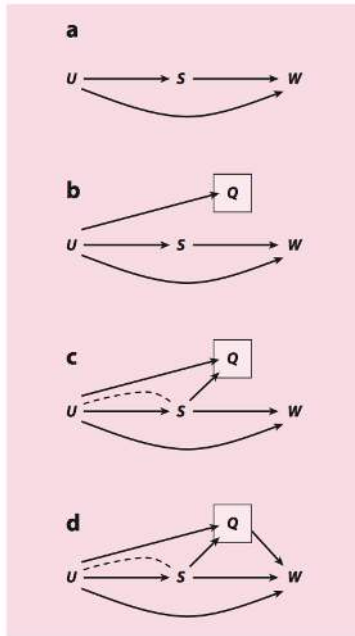


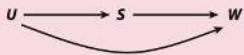
Figure 9

- S : schooling (treatment)
- W : wages (outcome)
- U : ability (unobserved)
- Q : measure test score of ability (such as IQ), as a proxy for U .

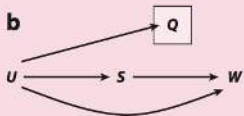
- S : schooling (treatment)
- W : wages (outcome)
- U : ability (unobserved)
- Q : measure test score of ability (such as IQ), as a proxy for U .



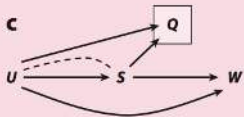
a



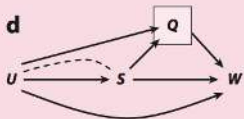
b



c

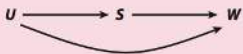


d

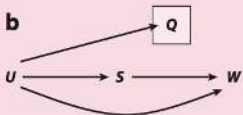


- (a): U is unobserved; the causal path is not blocked and the effect of S on W is unidentified.

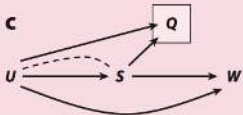
a



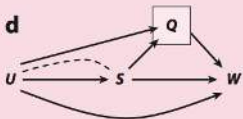
b



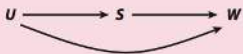
c



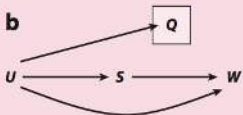
d



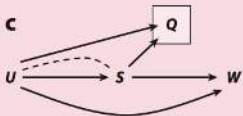
a



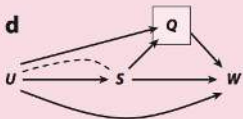
b



c

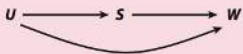


d

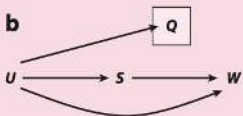


- (a): U is unobserved; the causal path is not blocked and the effect of S on W is unidentified.
- (b): Q is an imperfect measure;

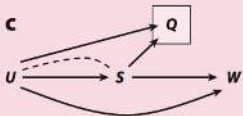
a



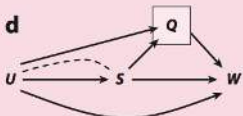
b



c

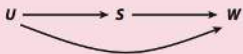


d

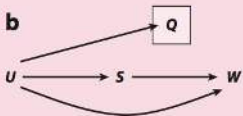


- (a): U is unobserved; the causal path is not blocked and the effect of S on W is unidentified.
- (b): Q is an imperfect measure;
- (c): Q is a collider; exogenous selection bias

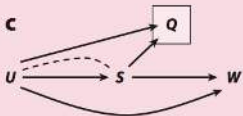
a



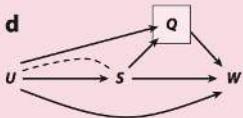
b



c



d



- (a): U is unobserved; the causal path is not blocked and the effect of S on W is unidentified.
- (b): Q is an imperfect measure;
- (c): Q is a collider; exogenous selection bias
- (d): Q has a causal impact on outcome W . Say employers use Q as hiring criteria; overcontrol bias.

Mediation analysis

Randomized class-size experiment:

- T : class size in first grade
- Y : high school graduation
- M : boosting student achievement in third grade
- U : unobserved causes of mediator variables

Treatment is randomized but the mediation is not.

Mediation analysis

Randomized class-size experiment:

- T : class size in first grade
- Y : high school graduation
- M : boosting student achievement in third grade
- U : unobserved causes of mediator variables

Treatment is randomized but the mediation is not.

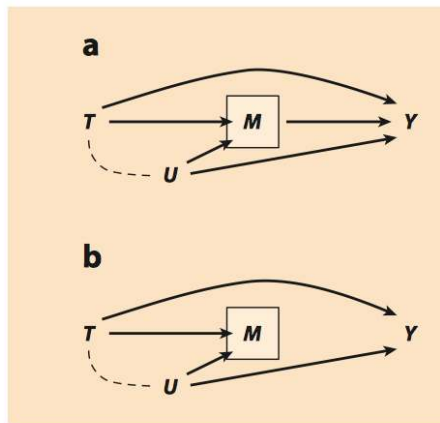


Figure 11

Mediation analysis

Randomized class-size experiment:

- (a): total effect between T and Y is unidentified if condition on M .
- (b):
 - total effect is identified
 - total effect = direct + indirect effects
 - directed effect (effect of T on Y net of other factors) is unidentified (due to collider).
 - Wrongly claim that there is indirect effect.

Treatment is randomized but the mediation is not.

Mediation analysis

Randomized class-size experiment:

- (a): total effect between T and Y is unidentified if condition on M .
- (b):
 - total effect is identified
 - total effect = direct + indirect effects
 - directed effect (effect of T on Y net of other factors) is unidentified (due to collider).
 - Wrongly claim that there is indirect effect.

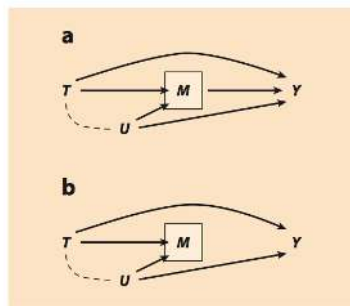


Figure 11

Treatment is randomized but the mediation is not.

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

- 1 Intro to DAG
- 2 Identification in DAG
- 3 Endogenous selection bias ((Post-) Outcome Collider)
- 4 Overcontrol bias
- 5 Condition on pre-treatment colliders

Pre-treatment Collider

- Spread of behaviors
- $Y_{i,t}$ is civic engagement of individual i at time t . Can be extended to other things (say, musical taste, smoking behaviors..)
- pre-treatment collider $F_{i,j}$: existence of friendship ties.
- U_i, U_j are individual attributes that shapes friendship ties (homophily) as well as outcome.

Pre-treatment Collider

- Spread of behaviors
- $Y_{i,t}$ is civic engagement of individual i at time t . Can be extended to other things (say, musical taste, smoking behaviors..)
- pre-treatment collider $F_{i,j}$: existence of friendship ties.
- U_i, U_j are individual attributes that shapes friendship ties (homophily) as well as outcome.

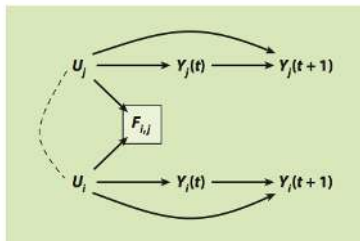
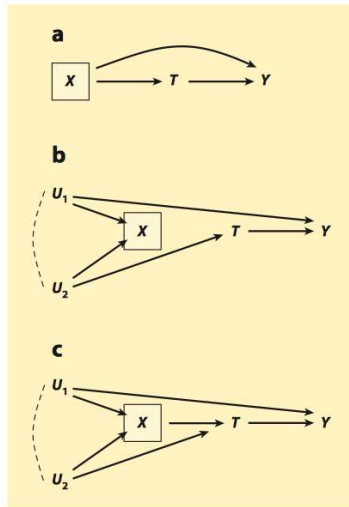


Figure 12

Pre-treatment Collider

- Without a priori theory, hard to distinguish colliders from confounding variables.
- X : 1) pre=treatment and 2) associated with both treatment and outcome.



- Do not condition on post-treatment variables (colliders and intermediates)
- pre-treatment colliders
 - do not condition when it is not a confounder.