DAGs and “What Is Your Estimand?”

Richard Patti and Michael Lachanski

A Crash Course in Directed Acyclic Graphs (DAGs)

- No clear description of “rules of the game” (Cunningham 2021: chap. 3)
- Not great for studying interactions/equalization
- Indispensable for reading contemporary articles in:
  - Education, epidemiology, political science, sociology
- Plan: 45 minutes on DAGs

The Plan

- Discussion meant to empower you to read “What Is Your Estimand?”
- Presentation is condensed version of Elwert (2021)
  - Will not cover, e.g. “d-separation”
  - https://www.youtube.com/watch?v=DK3sIDzHdMg
- What is a DAG? Why DAGs? How to use DAGs?
- Using DAGs - 3 fundamental sources of association
- Toy exercises

Data Generating Process

- How were your data generated?
- Could be equation (simulate it): $Y_i = X_i b + e_i$

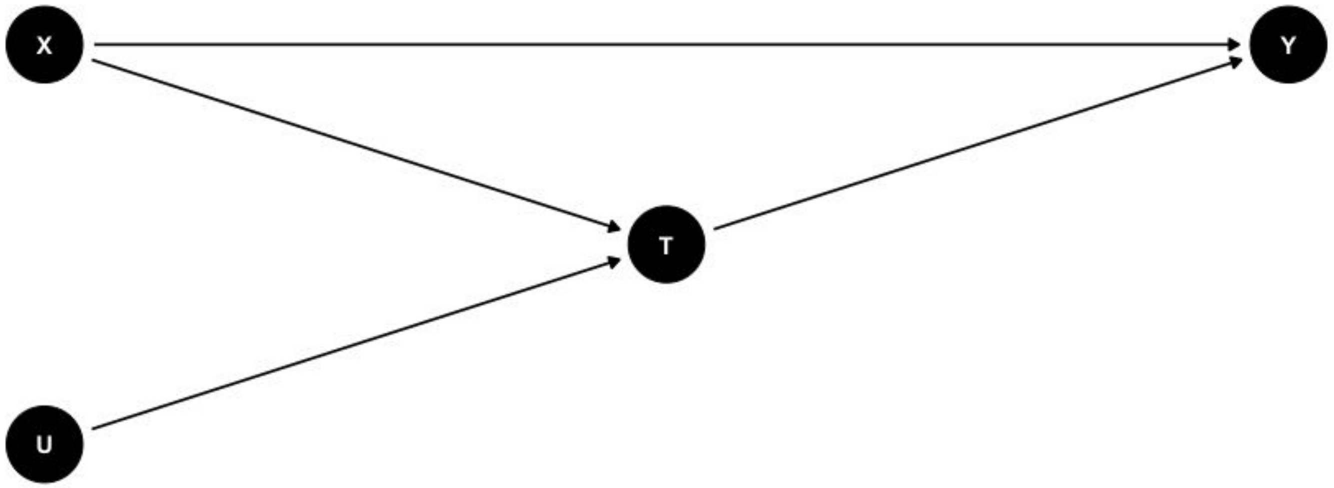
- Could be description (Massey 2007: 3):

"The most common form of stratification in foraging societies occurs on the basis of gender. Stratification between males and females derives primarily from the amount of time that men spend alone together, typically on a hunt, and is thus determined by local ecology (Massey 2005a). Societies where men spend large amounts of time away from women hunting large game tend to be more gender-stratified. During the time they are away on their own, males reinforce male predispositions and tendencies to become more aggressive and domineering (Macoby 1998). At the same time, females left by themselves reinforce female predispositions and tendencies to become more caring and nurturing. The end result is a divergence in gender-specific attitudes and behaviors that works to the detriment of females once the two sexes reunite (Macoby 1998; Sanday 1981). Compared with foraging societies built around the hunting of large mammals, societies that rely on aquatic resources, gathering and scavenging, or the pursuit of small game tend to be much lower in gender inequality."

What is a DAG? A “Directed Acyclic Graph”

- Almost all social science DGPs can be rendered using DAGs.
- In sociology, convention is that time flows from left to right on the graph.

1. DAGs Annotate the Data-Generating Process

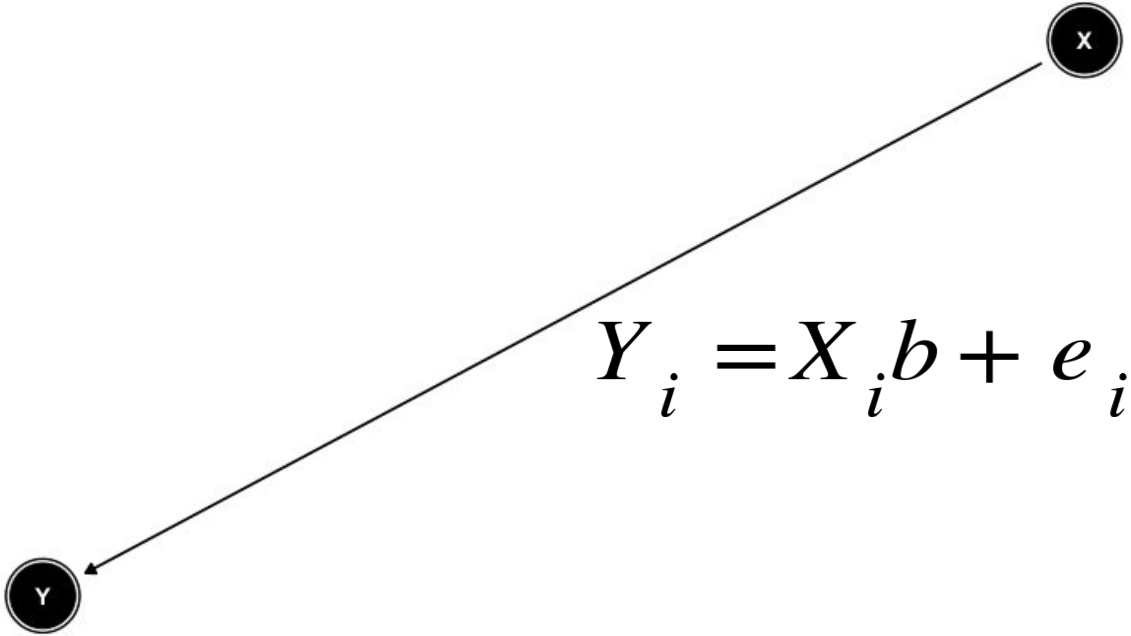


Components of a DAG

1. Node is a variable (“graph”)
2. A arrow is a causal relationship (“directed”)
3. MISSING ARROWS represent absent direct causal effects
    - Missing arrows are equivalent to “exclusion restrictions”.
    - Without exclusions, one cannot point-identify causal effects.
    - When you drop an arrow from node X to node Y, you are asserting that there is no direct causal path from X to Y.

- In the vast majority of social science research, DAGs are causal DAGs. We will only discuss causal DAGs.
- An arrow from X to Y means that X directly causes Y. This is the key feature of a causal DAG.
- Rarely, papers use this notation to mean that X is associated with Y. Careful!

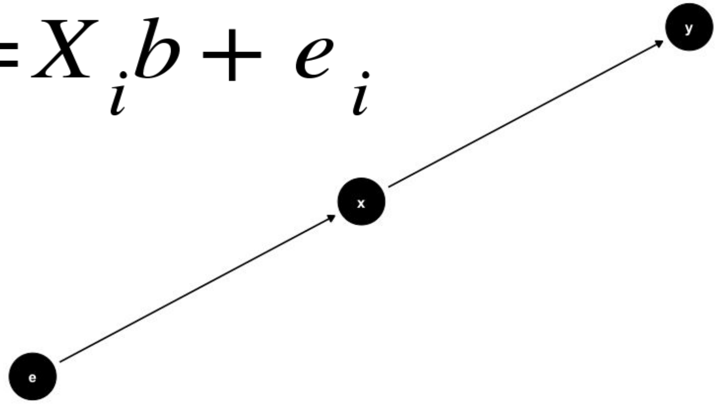
Example: Equation (X causes Y)

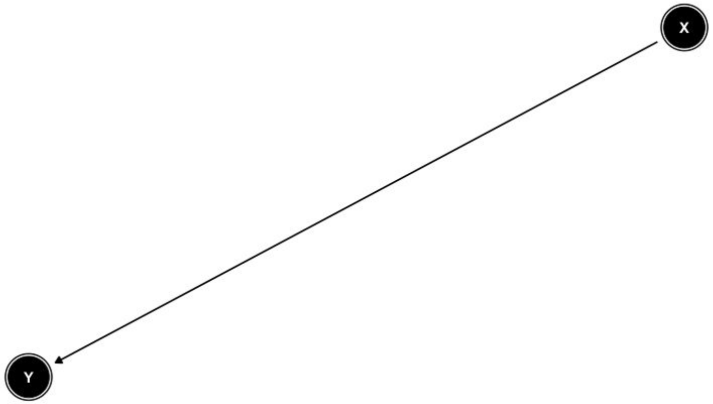


$$Y_i = X_i b + e_i$$

Example: Equation (X causes Y)

$$Y_i = X_i b + e_i$$

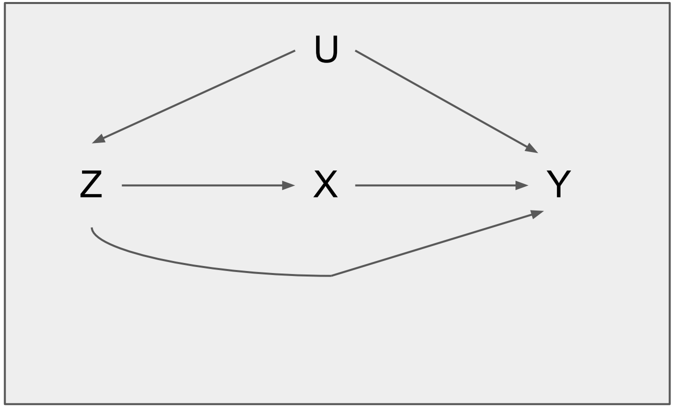


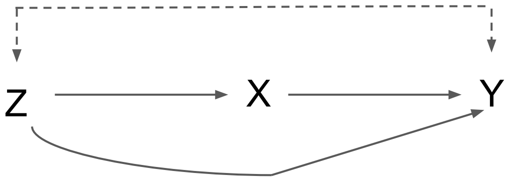


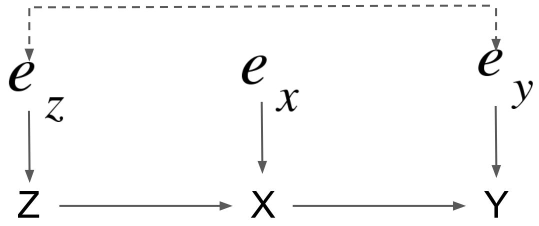
Most variables have lots of causes. The unmeasured causes affecting each variable are called the “structural error term” of the variable. An “idiosyncratic” structural error term includes all unobserved variables that only affect one measured variable in your graph. By convention, the idiosyncratic structural error terms are not drawn in the DAG. Here are two equivalent graphs. They represent the same equation and the same DGP.

Correlated Error Terms

Sometimes, some *unmeasured* variable, *U*, may affect two or more variables in your DGP, e.g. Z and Y. Since *U* is thus part of the error term of both variables, we say that *Z* and *Y* have "correlated errors".
Correlated errors, UNLIKE IDIOSYNCRATIC ERRORS, always need to be drawn explicitly!







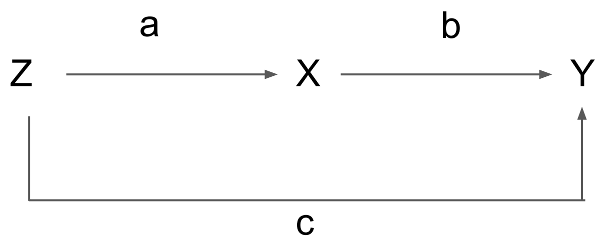
Here are three equivalent ways of drawing correlated error terms between *X* and *Y*. These three graphs represent the same DGP.

DAG vs. Path Models

Path models are a (very) special case of DAGs. The difference is that path models assume that the data are generated by a linear model, in which the causal effect of one variable on another is the exact same for every individual (“homogeneity”).

In a path model, the strength of the causal effect of each arrow is summarized by a single number, the so-called path coefficient.

Path coefficients represent causal effects in the regression metric.

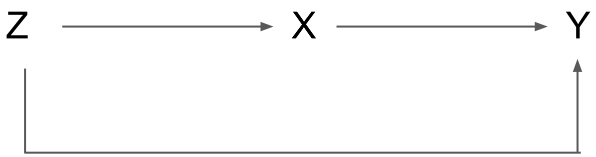


For example, this graph says that changing *Z* by one unit causes an increase in *X* by *a* units and that this effect is the same (on average) for all individuals. Furthermore, path analysis typically assumes that the nodes are jointly normal. The “path coefficients” are *parameters*, i.e., the structural/causal effects in the DGP.

DAGs are Non-parametric

In contrast to path models, DAGs are completely non-parametric objects.

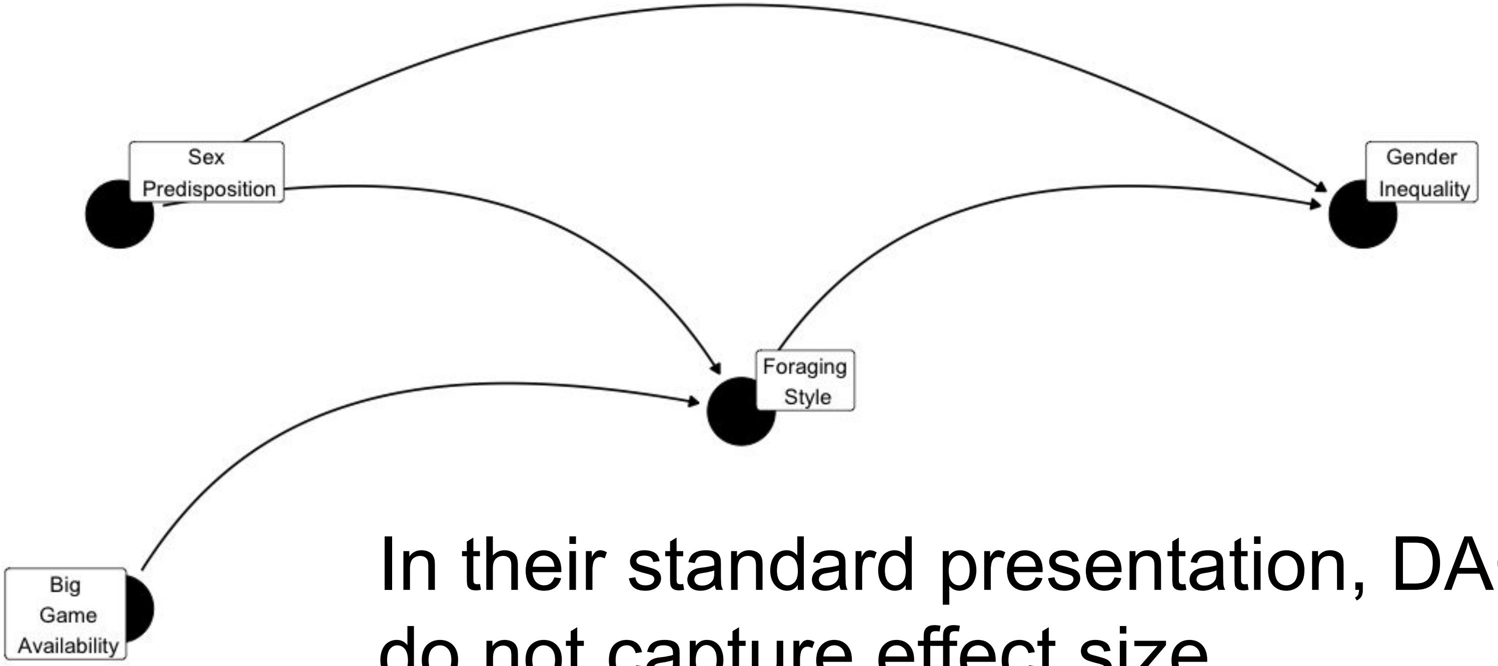
1. No distributional assumptions (nodes can have any distribution)
2. No functional form assumptions (arrows can be linear or non-linear, effects may be homogeneous or heterogeneous)



This makes DAGs a very general notation for the *qualitative* causal structure of the assumed DGP.
This generality guards against fallacious conclusions when distributional or functional form assumptions are ill justified (as they often are).

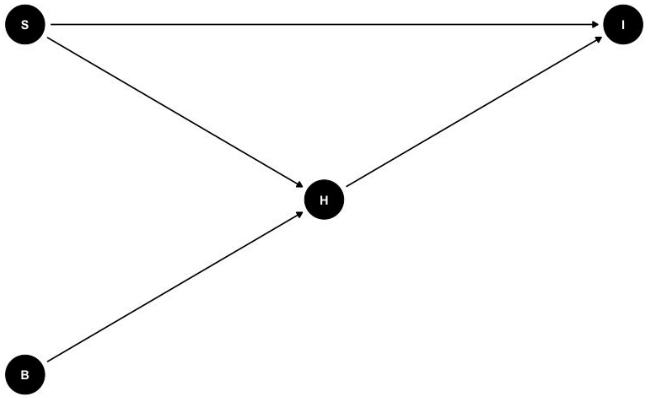
DAGs Capture Theories

Gender Inequalities in Foraging Societies (Massey 2007)



In their standard presentation, DAGs do not capture effect size.

Illustration: A Model of Gender Inequality in Foraging Societies



Remember: All causal claims made from data are relative to the assumed DGP. If you do not believe this DGP, you should change it.

This DAG captures the qualitative DGP summarized in Massey (2007: 3):

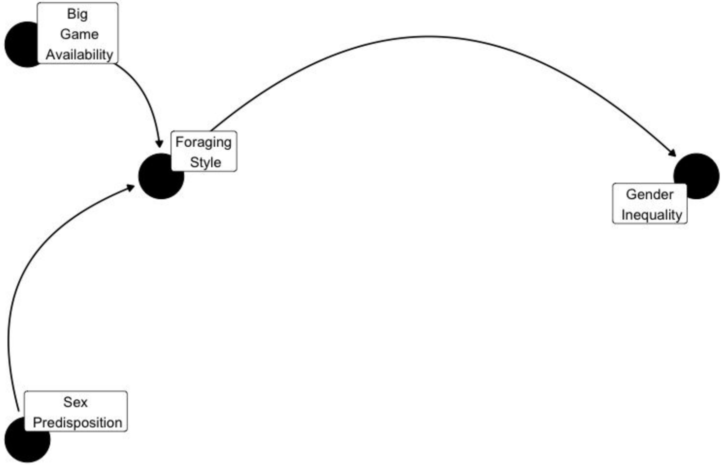
Sex predispositions, S, cause gender inequality, I:	S → I
Hunting big game, H, cause gender inequality, I:	H → I
Big game availability, B, cause men to hunt big game, H:	B → H
Sex predispositions, S, cause men to hunt big game, H:	S → H

This DAG also includes several exclusion restrictions:

Big game availability causes gender inequality only via hunting big game:	No arrow B → I
Sex predispositions do not cause big game availability:	No arrow S → B

DAGs Capture Theories

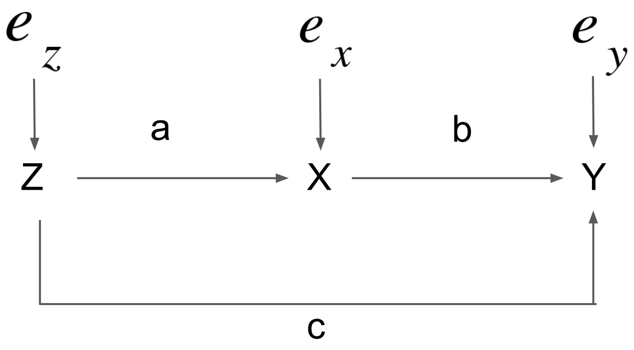
Gender Inequalities in Foraging Societies (Massey 2007)



If you don’t believe a DAG, you can change it.

Graphs Represent Structural Equations

Linear path model

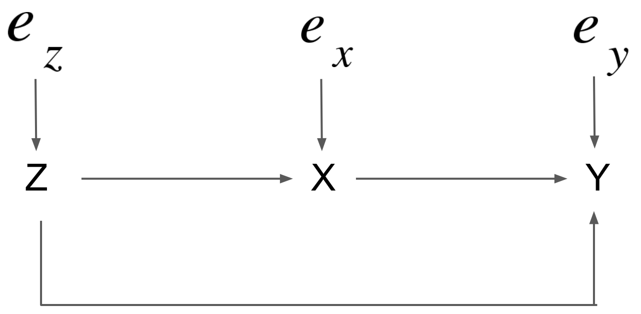


$$Z_i = e_{iZ}$$

$$X_i = p + aZ_i + e_{iX}$$

$$Y_i = g + cZ_i + bX_i + e_{iY}$$

Non-parametric structural equation model



$$Z_i = e_{iZ}$$

$$X_i = f_{iX}\left(Z_i,\ e_{iX}\right)$$

$$Y_i = f_{iY}\left(Z_i, X_i, e_{iY}\right)$$

DAGs Contain No Directed Cycles ("acyclic")

With all that, we have defined path models/DAGs as visual representations of structural equations that are assumed to have generated the data.

NB: $f_v(pa_v)$ says that V is generated (caused) by some (likely unknown) function, $f_v(\;)$, that takes the parents of V (pa_v, all variables with a direct arrow into V) as inputs; $f_v(\;)$ places no constraints on the functional form and may include interactions.

This non-parametric structural equation modeling REQUIRES that DAGs are "acyclic" in the sense that they don't contain directed cycles of arrows because Causes must occur before their effects. A directed cycle would say that the future causes the past => nonsense.

Apparent counterexamples are almost always resolved by articulating the DAG in time (e.g., by adding time subscripts to multiple states of a variable). There is also theory for cyclic, partially direct, and even undirected graphs. Naturally, such graphs support weaker conclusions, because they make weaker assumptions. See Pearl (2009) and Maathuis et al. (2018).

2. Inferring Association from Causation

- DAGs allow you to distinguish causal correlations from associations.
- DAGs represent an *assumed solution* to the “identification problem”. Problem of “causal discovery” has no general solution (Robins and Wasserman 1999).
- Having notated the causal structure of a DGP graphically, we can now infer *what associations should exist* (be observable) in the data *if the graph represented the true DGP*.
- This is known as “deriving the associational implications of a model”.

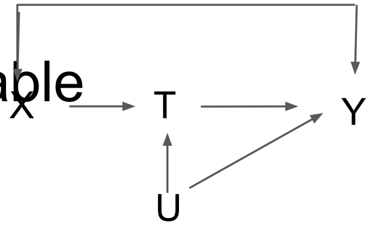
2. Inferring Association from Causation

- In other words, we ask: if this graph captures the process that generated the data, what variables should be associated with each other, and which ones should be independent?
- Problem of identifying **all** associations from DAG is completely solved: “d-separation” and taught in a full course on this subject (Pearl 1988). Takes too much time to cover seriously.
- Instead, we give some heuristics for “deriving the associational implications of a model”.

Outlook

- Simple rules connect the causal assumptions in the DAG to statistical associations in the data.

1. The causal effects in a DGP give rise to observable associations in data.
2. We also cover some heuristics.

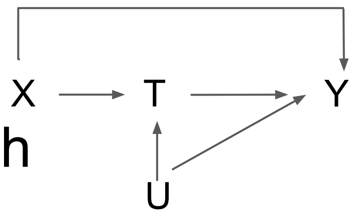


- In the following slides, we define the term path. See Pearl (2009) for details.

Paths: Causal and Non-causal

Definition: A path between two variables is a non-self-intersecting sequence of adjacent edges.

Meaning: the direction of the arrows does not matter; a path may not contain cycles (i.e., touch any variable more than once).



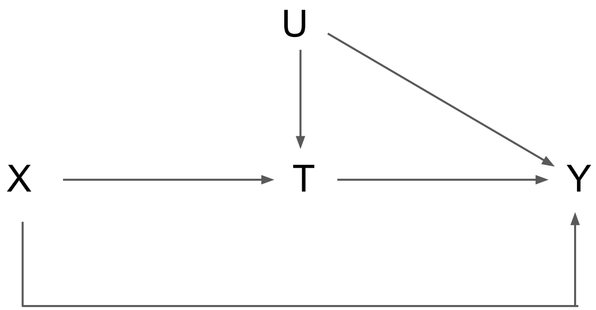
Fact: Every path between a treatment, *T*, and an outcome, *Y*, is either causal or non-causal.

Definition: A causal path between two variables is a path in which all arrows point away from *T* and towards *Y*. The union of causal paths comprise the total causal effect.

Paths: Causal and Non-causal

Definition: A non-causal path is a path from T to Y in which at least one arrow points against the flow of time. Non-causal paths may carry spurious association.

Definition: A path between treatment, T, and outcome, Y, that begins with an arrow into T, T←, is called a backdoor path. Backdoor paths are non-causal.



Exercise: List the causal and non-causal paths from T to Y.

Causal: T→Y

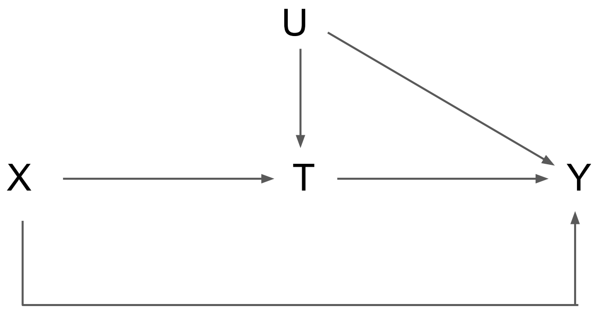
Non-causal: T←U→Y and T←X→Y

Collider Variables Are Key

Collider variables are key for working with DAGs.

Definition: When two arrows on a given path point directly into a variable on the path, then that variable is a collider variable on the path.

Fact: On any given path, any one variable is either a collider or a non-collider. Definition: A variable that is not a collider on the path is a non-collider on the path.



Ex: T is a collider on the path: X→T←U→Y

Ex: T is not a collider on the path: X→T→Y ⇒ Colliders are path-specific!

: non-causal (spurious) association

: conditioning

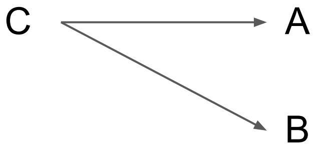
Three Rules of Association

Now let's ask which paths transmit, and which don't transmit, association. We start with the primitives. It turns out that all marginal and conditional associations originate from only 3 causal structures.

(1) Direct and indirect causation

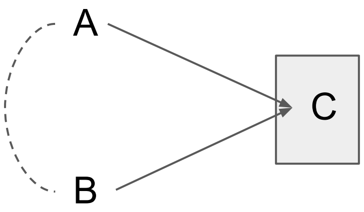
$A \not\perp\!\!\!\perp B$ and $A \perp\!\!\!\perp B \mid C$

(2) Common cause confounding



$A \not\perp\!\!\!\perp B$ and $A \perp\!\!\!\perp B \mid C$

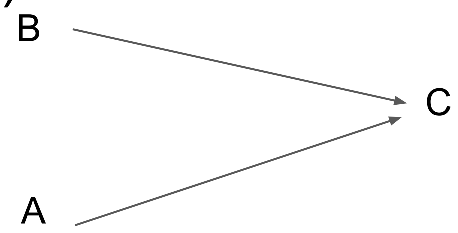
(3) Conditioning on a common effect ("collider"): Selection $A \perp\!\!\!\perp B$ and $A \not\perp\!\!\!\perp B \mid C$



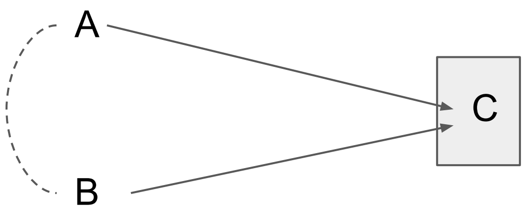
Conditioning on a Collider

Notice: No causal effect of A on B (or vice versa) and no confounding.

$A \perp\!\!\!\perp B$: Marginally independent



$A \not\perp\!\!\!\perp B \mid C$: Conditionally dependent



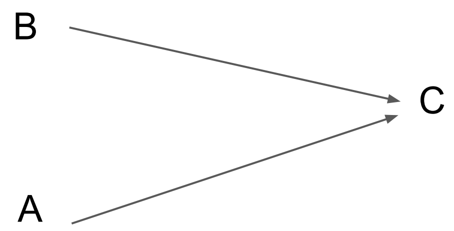
Conditioning on a Collider: Examples

<u>Pearl's Sprinkler Example</u>

A: It rains

B: The sprinkler is on

C: The lawn is wet

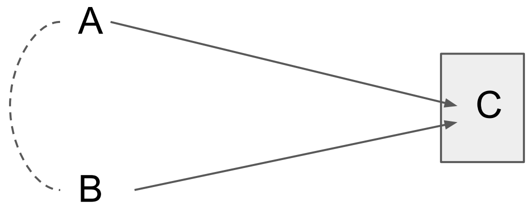


<u>Academic Tenure Example</u>

A: Productivity

B: Originality

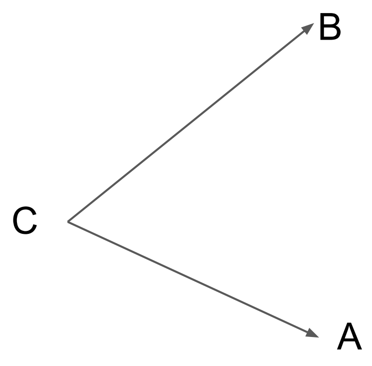
C: Tenure



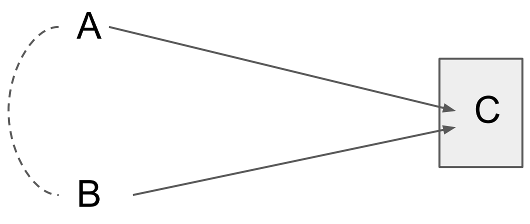
Confounding vs. Selection Bias

Confounding bias and selection bias are different.

Suppose that *A* is treatment and *B* is the outcome.



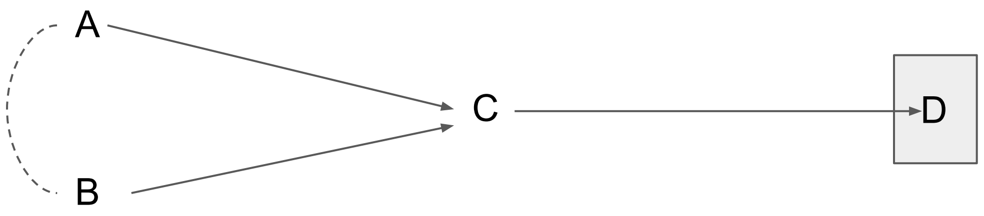
Confounding Bias: failure to condition on a common cause.



Selection bias: mistakenly conditioning on a common effect.

There's More: Conditioning on a Descendant of a Collider

Conditioning on the descendant, D, of a collider, C, results in the same problem as conditioning on C itself: It induces an association between the collider's parents.



What Is Your Estimand?

Proposed Framework

Step 1: Set a theoretical estimand

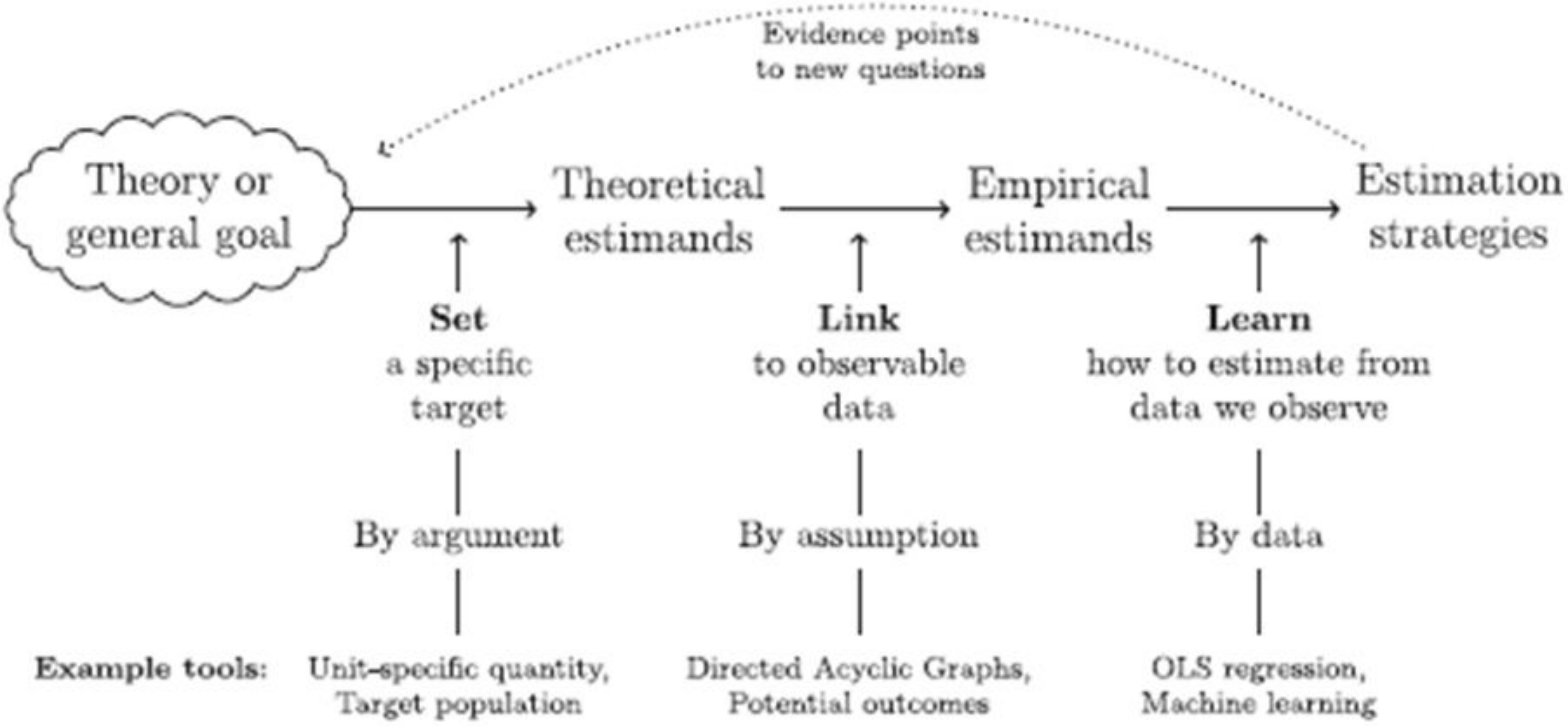
- A goal outside the model

Step 2: Link to an empirical estimand

- Observable quantity (use DAGs)

Step 3: Estimation

- Learn the empirical estimand from data


Evidence points
to new questions
Theory or
general goal
Theoretical
estimands
Empirical
estimands
Estimation
strategies
Set
a specific
target
Link
to observable
data
Learn
how to estimate from
data we observe
By argument
By assumption
By data
Example tools:
Unit-specific quantity,
Target population
Directed Acyclic Graphs,
Potential outcomes
OLS regression,
Machine learning

Theoretical Estimand (Set the Target)

What are you trying to estimate?

- Unit-specific quantity
- Target Population

Theoretical Estimand (Set the Target)

What are you trying to estimate?

$$\frac{1}{n}\sum_{i=1}^{n} Y_i$$

↑ Mean over every i among U.S. adults (target population)

↑ Whether each i is employed (unit-specific quantity)

Theoretical Estimand (Set the Target)

What are you trying to estimate?

- Unit-specific quantity
- Target Population
- Exists outside Statistical Model

Theoretical Estimand (Set the Target)

What are you trying to estimate?

$$\frac{1}{n}\sum_{i=1}^{n}\Big(Y_i(1) - Y_i(0)\Big) \qquad (2)$$

↗	↑	↖
Mean over every i among U.S. adults	Employment if enrolled in job training	Employment if not enrolled in job training
(target population)	(unit-specific quantity)	

Unit-Specific Quantity

Defined for each unit in the population

- Descriptive
- Causal

Target Population

Empirically Tractable

Target Population

Empirically Tractable

Theoretical Interest in itself OR

Informative about a broader population

Target Population

Empirically Tractable

Theoretical Interest in itself OR

Informative about a broader population

Address the tension!

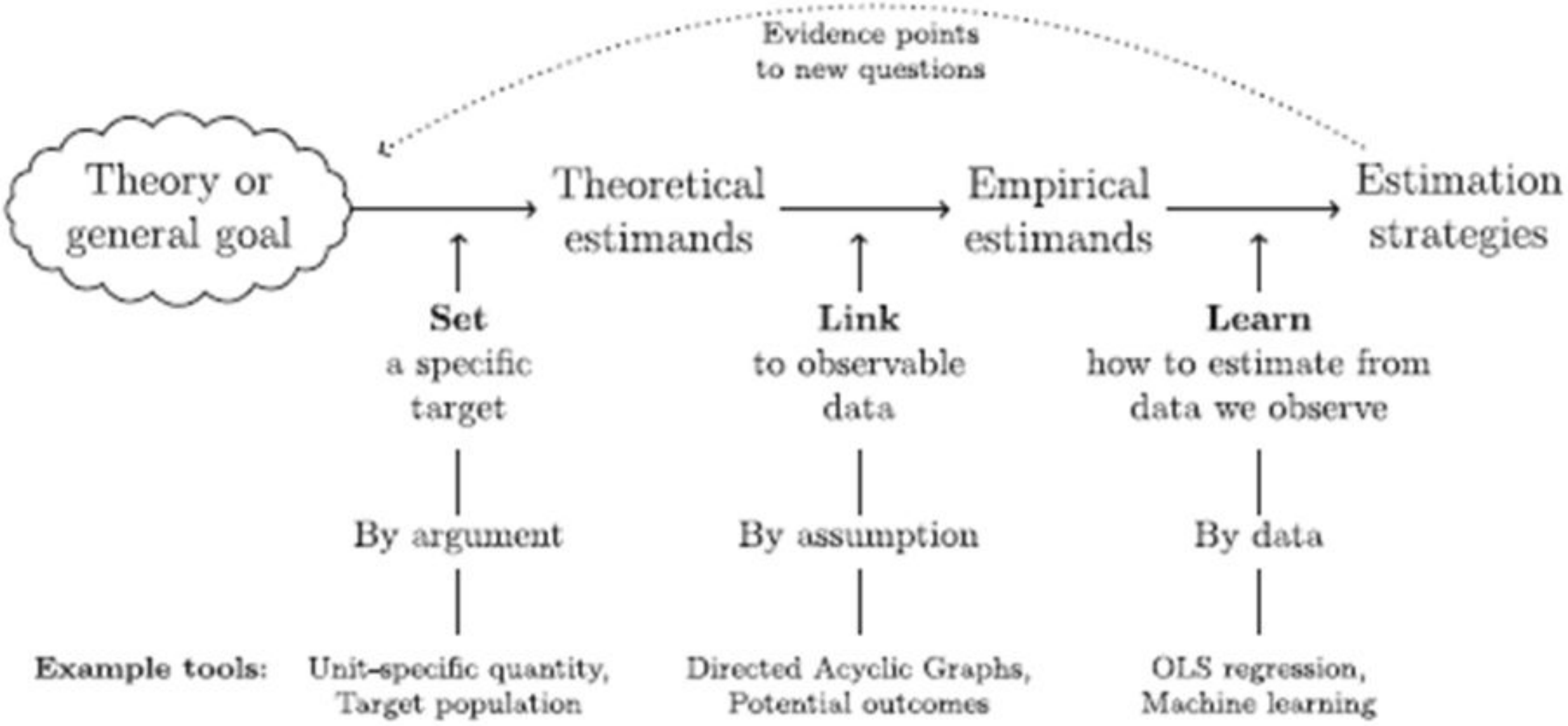
Ex. Angrist and Evans (1998) birth of third child

Theoretical Estimand

Descriptive estimand

Causal estimand- counterfactuals

Formalize quantity relevant to theory


Evidence points
to new questions
Theory or
general goal
Theoretical
estimands
Empirical
estimands
Estimation
strategies
Set
a specific
target
Link
to observable
data
Learn
how to estimate from
data we observe
By argument
By assumption
By data
Example tools:
Unit-specific quantity,
Target population
Directed Acyclic Graphs,
Potential outcomes
OLS regression,
Machine learning

Identification: Link to an Empirical Estimand

Observable Data (Formalized with DAGs)

Identification: Link to an Empirical Estimand

$$\frac{1}{n}\sum_{i=1}^{n}\Big(Y_i(1) - Y_i(0)\Big) \qquad (2)$$

↗ Mean over every i among U.S. adults

↑ Employment if enrolled in job training

↑ Employment if not enrolled in job training

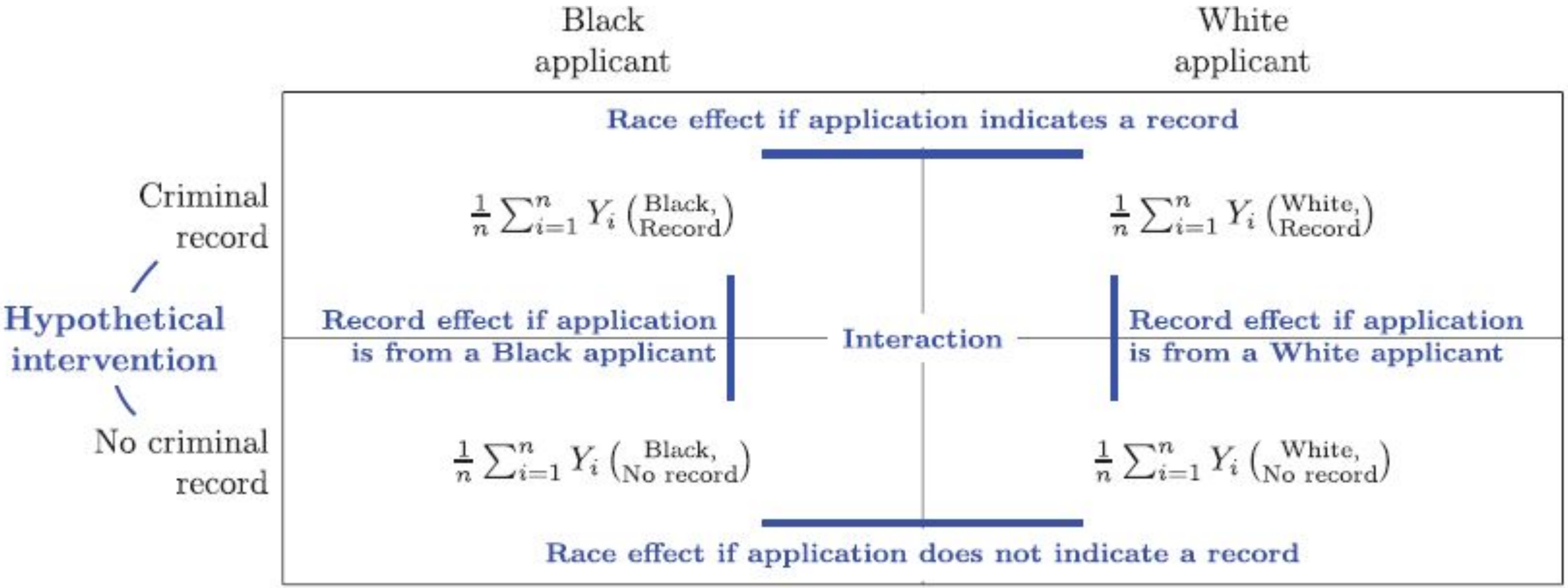
(target population)

(unit-specific quantity)

Table 1. Unit-Specific Quantities Defined in Potential Outcomes Unlock Many Causal Estimands for Inquiry

Estimand name	Mathematical statement	DAG	Reference	Colloquial terms
Average treatment effect	$\frac{1}{n}\sum_{i}\left(Y_i(d') - Y_i(d)\right)$	$D \longrightarrow Y$	Morgan and Winship (2015)	Effect
Conditional average treatment effect	$\frac{1}{n_x}\sum_{i:X_i=x}\left(Y_i(d') - Y_i(d)\right)$	$X \rightarrow D \rightarrow Y$; $X \rightarrow Y$	Athey and Imbens (2016)	Effect heterogeneity or moderation
Causal interaction	$\frac{1}{n}\sum_{i}\left(\left(Y_i(a',d') - Y_i(a',d)\right) - \left(Y_i(a,d') - Y_i(a,d)\right)\right)$	$A \rightarrow Y$; $D \rightarrow Y$	Vanderweele (2015)	Joint treatment effect
Controlled direct effect	$\frac{1}{n}\sum_{i}\left(Y_i(d',m) - Y_i(d,m)\right)$	$D \rightarrow M \rightarrow Y$; $D \longrightarrow Y$	Acharya et al. (2016)	Mediation (Illustrations: Example 2)
Natural direct effect	$\frac{1}{n}\sum_{i}\left(Y_i(d', M_i(d)) - Y_i(d, M_i(d))\right)$	$D \rightarrow M \rightarrow Y$; $D \longrightarrow Y$	Imai et al. (2011)	Mediation (Part B of the Online Supplement)
Effect of time-varying treatment	$\frac{1}{n}\sum_{i}\left(Y_i(d'_1, d'_2) - Y_i(d_1, d_2)\right)$	$D_1 \rightarrow D_2 \rightarrow Y$; $D_1 \rightarrow Y$	Wodtke et al. (2011)	Cumulative effect

Identification: Link to an Empirical Estimand



Identification: Link to an Empirical Estimand

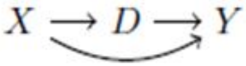
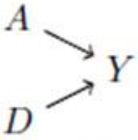
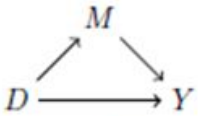
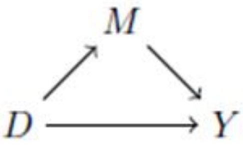
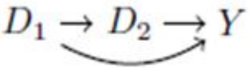
$$\tau = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i\binom{\text{White,}}{\text{Record}} - Y_i\binom{\text{Black,}}{\text{No record}}\right) \tag{3}$$

Mean over **all** applications

Potential outcome under one condition

Potential outcome under another condition

Table 1. Unit-Specific Quantities Defined in Potential Outcomes Unlock Many Causal Estimands for Inquiry

Estimand name	Mathematical statement	DAG	Reference	Colloquial terms
Average treatment effect	$\frac{1}{n}\sum_i \left(Y_i(d') - Y_i(d)\right)$		Morgan and Winship (2015)	Effect
Conditional average treatment effect	$\frac{1}{n_x}\sum_{i:X_i=x} \left(Y_i(d') - Y_i(d)\right)$		Athey and Imbens (2016)	Effect heterogeneity or moderation
Causal interaction	$\frac{1}{n}\sum_i \left(\left(Y_i(a', d') - Y_i(a', d)\right) - \left(Y_i(a, d') - Y_i(a, d)\right)\right)$		Vanderweele (2015)	Joint treatment effect
Controlled direct effect	$\frac{1}{n}\sum_i \left(Y_i(d', m) - Y_i(d, m)\right)$		Acharya et al. (2016)	Mediation (Illustrations: Example 2)
Natural direct effect	$\frac{1}{n}\sum_i \left(Y_i(d', M_i(d)) - Y_i(d, M_i(d))\right)$		Imai et al. (2011)	Mediation (Part B of the Online Supplement)
Effect of time-varying treatment	$\frac{1}{n}\sum_i \left(Y_i(d_1', d_2') - Y_i(d_1, d_2)\right)$		Wodtke et al. (2011)	Cumulative effect

Identification: Link to an Empirical Estimand

$$\tau = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i\binom{\text{White,}}{\text{Record}} - Y_i\binom{\text{Black,}}{\text{No record}}\right) \quad (3)$$

Mean over **all** applications

Potential outcome under one condition

Potential outcome under another condition

Identification: Link to an Empirical Estimand

Factual outcomes

$$\theta = \frac{1}{n_{\text{WR}}} \sum_{i \in \mathcal{S}_{\text{WR}}} Y_i - \frac{1}{n_{\text{BN}}} \sum_{i \in \mathcal{S}_{\text{BN}}} Y_i$$

Mean among applications assigned to the condition (**W**hite, **R**ecord)

Mean among applications assigned to the condition (**B**lack, **N**o record)

Estimation- Learn the Empirical Estimand From the Data

Estimation as a tool to estimate unknown components (not just parameters in a model)

Estimation- Learn the Empirical Estimand From the Data

Estimation as a tool to estimate unknown components (not just parameters in a model)

Mean over entire sample

Expected outcome among cases with the covariate values $\vec{x}_i$ of unit i who are factually treated ($D = 1$)

$$\theta = \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}(Y \mid \vec{X} = \vec{x}_i, D = 1) - \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}(Y \mid \vec{X} = \vec{x}_i, D = 0) \quad (5)$$

Mean over entire sample

Expected outcome among cases with the covariate values $\vec{x}_i$ of unit i who are factually untreated ($D = 0$)

Estimation- Learn the Empirical Estimand From the Data

Estimation as a tool to estimate unknown components (not just parameters in a model)

Mean over entire sample ↘

Regression prediction at observed covariate values $\vec{x}_i$ with treatment set to $D = 1$ ↙

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, D = 1) - \frac{1}{n}\sum_{i=1}^{n} \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, D = 0)$$

↗ Mean over entire sample

↖ Regression prediction at observed covariate values $\vec{x}_i$ with treatment set to $D = 0$

Estimation- Learn the Empirical Estimand From the Data

Estimation as a tool to estimate unknown components (not just parameters in a model)

Which θ minimizes mean squared error?

- Performance in treatment and control group

Estimation to predict conditional means rather than estimating coefficients

Examples:
How Estimands Improve Practice

College Completion

Probability of college completion Y ↘ | Among those with ↓ | Gender g, birth cohort c, and parent characteristics p ↙

$$\tau(g,p,c) = \mathbf{P}\left(Y = 1 \;\middle|\; \begin{matrix} G = g \\ C = c \\ P = p \end{matrix}\right)$$

College Completion

Probability of college completion Y ↘ Among those with ↓ Gender g, birth cohort c, and parent characteristics p ↙

$$\tau(g,p,c) = \mathrm{P}\left(Y = 1 \;\middle|\; \begin{matrix} G = g \\ C = c \\ P = p \end{matrix} \right)$$

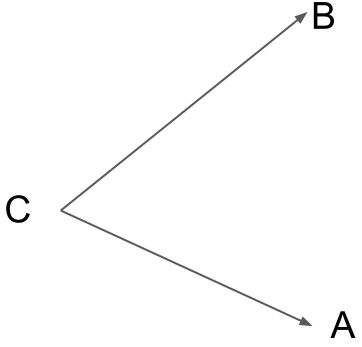
Probability of college completion Y ↘ Among those with ↓ Gender g, birth cohort c, and parent characteristics p ↙

$$\theta(g,p,c) = \mathrm{P}\left(Y = 1 \;\middle|\; \begin{matrix} G = g \\ C = c \\ P = p \\ R = 1 \end{matrix} \right)$$

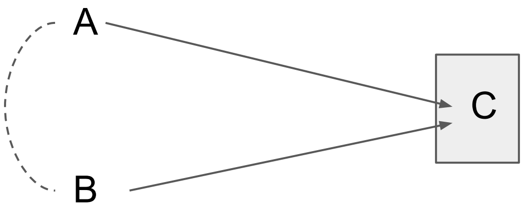
↑ and who are **alive and willing to respond** $(R = 1)$

Confounding vs. Selection Bias

Confounding bias and selection bias are different. Suppose that *A* is treatment and *B* is the outcome.

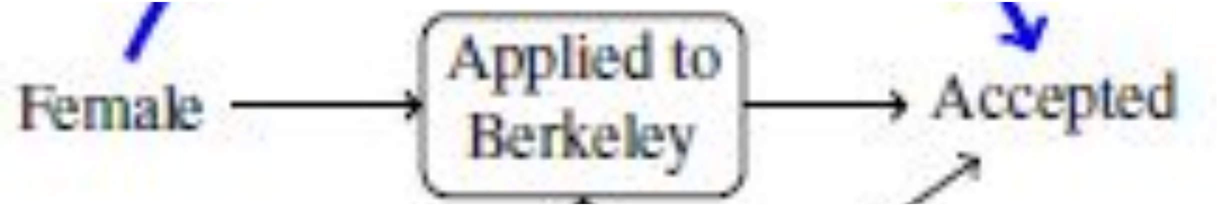


Confounding Bias: failure to condition on a common cause.

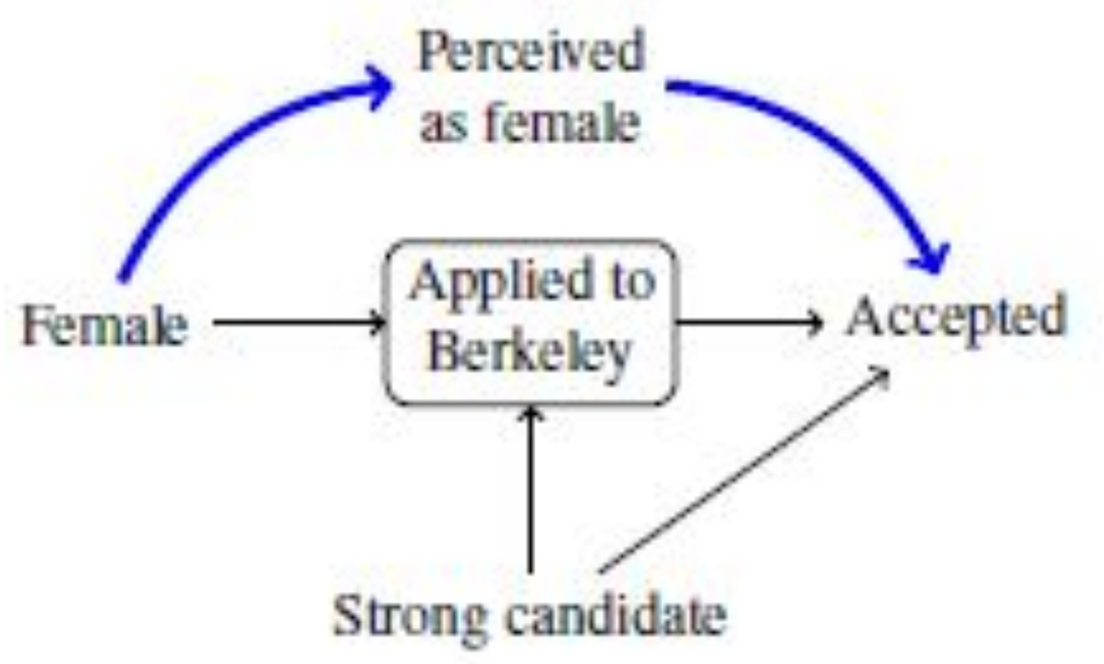


Selection bias: mistakenly conditioning on a common effect.

Berkeley Admissions



Berkeley Admissions



Identification: Link to an Empirical Estimand

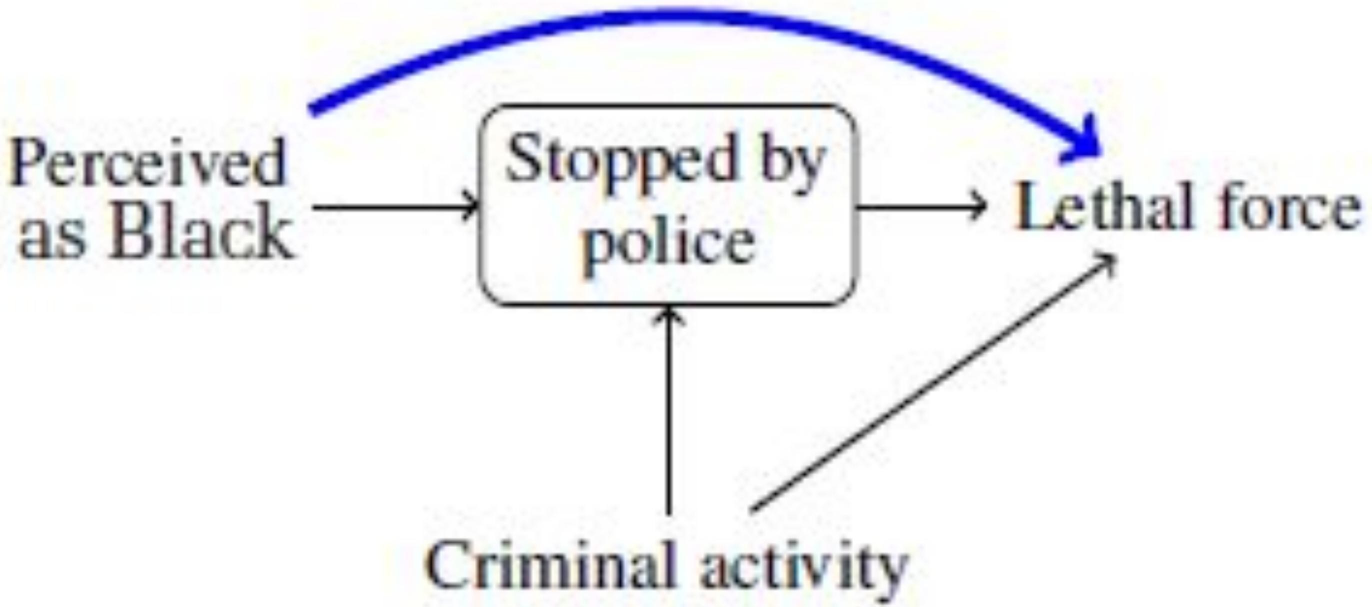
Factual outcomes

$$\theta = \frac{1}{n_{\text{WR}}} \sum_{i \in \mathcal{S}_{\text{WR}}} Y_i - \frac{1}{n_{\text{BN}}} \sum_{i \in \mathcal{S}_{\text{BN}}} Y_i$$

Mean among applications assigned to the condition (**W**hite, **R**ecord)

Mean among applications assigned to the condition (**B**lack, **N**o record)

Example


Perceived as Black
Stopped by police
Lethal force
Criminal activity

Identification: Link to an Empirical Estimand

Factual outcomes

$$\theta = \frac{1}{n_{\text{WR}}} \sum_{i \in \mathcal{S}_{\text{WR}}} Y_i - \frac{1}{n_{\text{BN}}} \sum_{i \in \mathcal{S}_{\text{BN}}} Y_i$$

Mean among applications assigned to the condition (**W**hite, **R**ecord)

Mean among applications assigned to the condition (**B**lack, **N**o record)

~FIN~

References

Angrist, Joshua D., and William N. Evans. 1998. “Children and Their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size.” The American Economic Review 88(3):450–77.

Bickel, Peter J., Eugene A. Hammel, and J. William O’Connell. 1975. “Sex Bias in Graduate Admissions: Data from Berkeley.” Science 187(4175): 398–404.

Cunningham, S. 2021. Causal inference. In *Causal Inference*. Yale University Press.

Fryer, Roland G. 2019. “An Empirical Analysis of Racial Differences in Police Use of Force.” Journal of Political Economy 127(3):1210–61.

Fryer, Roland G. 2020. “An Empirical Analysis of Racial Differences in Police Use of Force: A Response.” Journal of Political Economy 128(10):4003–4008.

Lundberg, I., Johnson, R., and Stewart, B. M. 2021. What is your estimand? Defining the target quantity connects statistical evidence to theory. American Sociological Review, *86*(3):532-565.

Massey, D.S. 2007. *Categorically unequal: The American stratification system*. Russell Sage Foundation.

Pager, Devah. 2003. “The Mark of a Criminal Record.” American Journal of Sociology 108(5):937–75.

Pearl, Judea. 2009. *Causality*. Cambridge University Press.