Module 2: Causality and DAGs

Econometrics II

Sannah Tijani (stijani@wu.ac.at)
Department of Economics, WU Vienna

Max Heinze (mheinze@wu.ac.at)
Department of Economics, WU Vienna

October 23, 2025

Causality

Identification

Randomization

Practice

Causality

- Causality is when one **cause** leads to some **effect**.
- The **cause** is partly responsible for the **effect**, and the **effect** partly depends on the **cause**.
- A **causal** relationship is useful for making **predictions** about the consequences of changing circumstances or policies; it tells us what would happen in alternative (**counterfactual**) worlds.
- e.g. The effect of colonial institutions on economic growth by Acemoglu, Johnson, and Robinson.

Consider a binary treatment X, and outcome Y. We can think of the **causal effect** τ as the difference in **potential outcomes**:

$$\tau = Y(X = 1) - Y(X = 0)$$

Problem of Causal Inference

In reality, only **one outcome** is realized, the other is **counterfactual**. Thus, we have to estimate this missing outcome to learn about the causal effect.

The potential outcomes framework is called the Neyman-Rubin causal model.

i	X_i	Y_i	$Y_i(1)$	$Y_i(0)$
1	0	1	?	1
2	0	1	?	1
3	1	1	1	?
4	1	0	0	?
...				
N	1	1	1	?

Causality

Identification

Randomization

Practice

Directed Acyclical Graphs

Identification

- We say an effect is **causally identified** if we can interpret it causally in our framework and scope.
- If we want to understand the causal impact of studying on income, $\boldsymbol{y}^{inc} = \boldsymbol{x}^{stu}\beta + \boldsymbol{u}$: Studying → Income
- We likely run into an issue, as you don't get paid for studying but for your skills: Studying → Skills → Income
- Moreover, ability may confound your effect estimates of studying on income, $\boldsymbol{y}^{inc} = \boldsymbol{x}^{stu}\beta + \boldsymbol{u}$, affecting studying, skills, and income.
- You cannot identify the causal effect of studying on income.

Causal Quantities

- **Average Treatment Effect:** the average causal effect is simply the mean of all treatment effects

$$\tau_{ATE} = \mathrm{E}[\tau_i] = \mathrm{E}[Y(1) - Y(0)] = \mathrm{E}[Y(1)] - \mathrm{E}[Y(0)]$$

- **Conditional Average Treatment Effect:** often, we want to control for some third characteristic Z_i:

$$\tau_{CATE} = \mathrm{E}[\tau_i | Z_i = z]$$

- **Average Treatment Effect on the Treated:** we condition on received treatment $Z_i = X_i = 1$.

Average Treatment Effect

- We can use $\mathrm{E}[Y_i(0)] = 0.25$ and $\mathrm{E}[Y_i(1)] = 0.75$ to find that the average treatment effect:

$$\tau_{ATE} = \mathrm{E}[Y(1)] - \mathrm{E}[Y(0)] = 0.5$$

- We can also run an OLS regression and estimate τ_{ATE} :

$$y = x\tau + e$$

Ignorability

- A treatment X is **ignorable** if both potential outcomes are independent of X, the treatment:

$$(Y(1), Y(0)) \perp X$$

- When X is **ignorable**, the treatment is **randomly assigned** and only affects the outcome Y by either realising Y(0) or Y(1):

$$Y = Y(1)X + Y(0)(1 - X)$$

- This condition would be violated in the case of targeted assignment of the subjects, or if the subjects select themselves (survey response)

Conditional Ignorability

A treatment X is **ignorable**, conditional on covariates Z, if:

- $(Y(1), Y(0)) \perp X|Z$
- $P(X = 1) \in (0, 1)$

Potential outcomes are independent of X, conditional on Z, and there are both treated and untreated subjects.

If X is **ignorable**, we can use the sample averages $\mathrm{E}[Y_i(0)]$ and $\mathrm{E}[Y_i(1)]$ as estimates for Y(0) and Y(1), then the estimate of τ_{ATE} will be causally identified.

Causality

Identification

Randomization

Practice

Directed Acyclical Graphs

Application: DAGs in Research

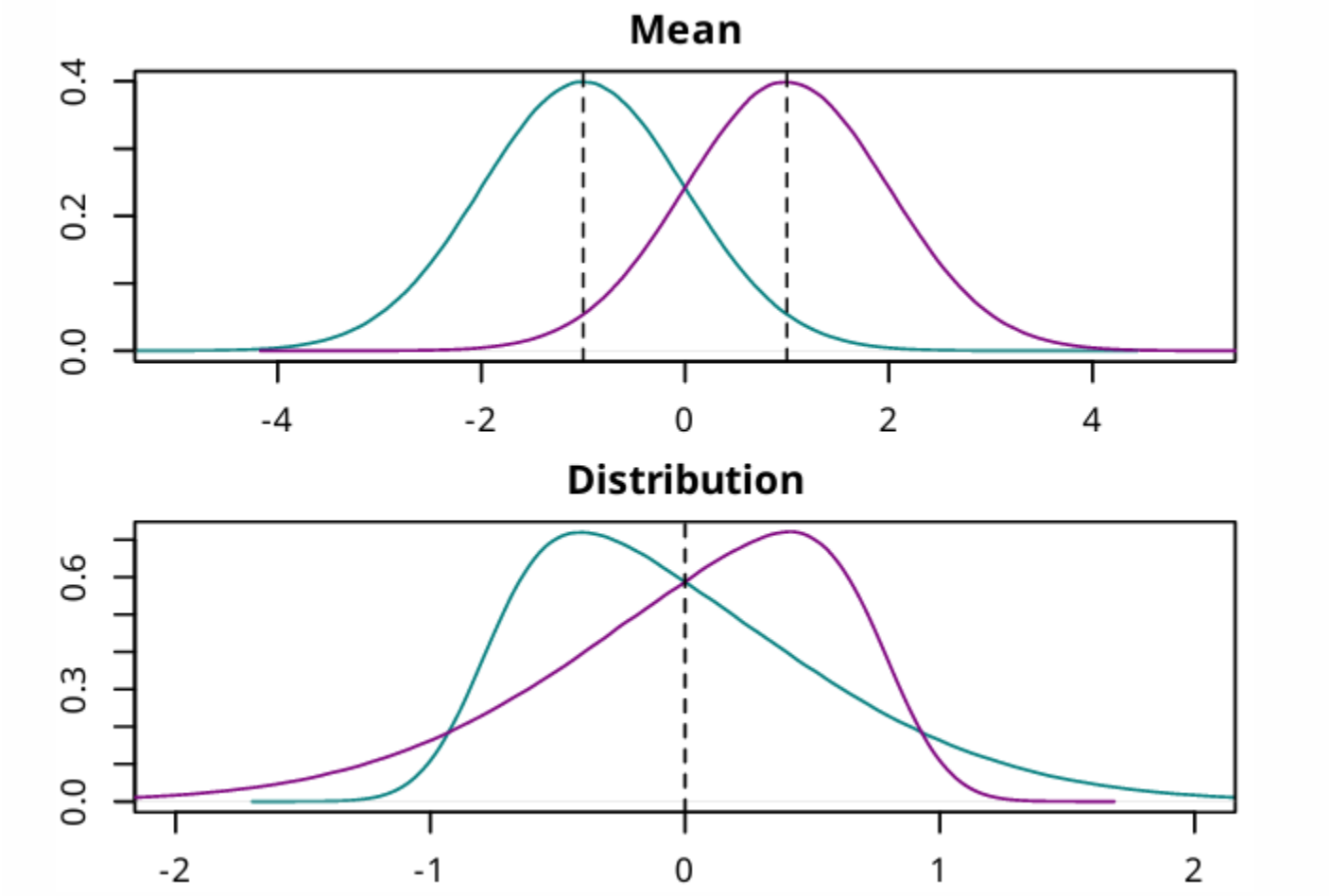
Randomised Experiment

- We have seen that we can estimate a causal effect if we have access to both the **realized outcomes** and its **counterfactual** or if the treatment is **ignorable**.
- Until someone figures out a way to use the first option, experiment with **random assignment** of the treatment is our best option.
- **random assignment** of the treatment solves the selection problem as it makes the treatment independent of potential outcomes.
- However, even in properly randomized experiment there remains **threats** to causal inference.
- The first question to ask is whether the randomization successfully **balanced subjects' characteristics** across the different treatments groups
- The second question to ask is whether there is **sufficient overlap** across the different treatment groups
- For the moment let's focus on two groups: **treated** and **control**.

Imbalance

- An **imbalance** between the treated and control groups occurs when there are differences between the groups i.e, they don't have similar **characteristics**.
- **Imbalance** refers to differences in the distribution of covariates Z between the treatment and the controls. Even if we have **overlap**, the groups might systematically differ on key covariates.
- This is problematic when there are differences in terms of third variables that affect **outcome Y**.
- With enough data, these imbalances should disappear, otherwise we need to account for them before comparing sample means of the groups.

Imbalance (2)



Imbalance Regression Table

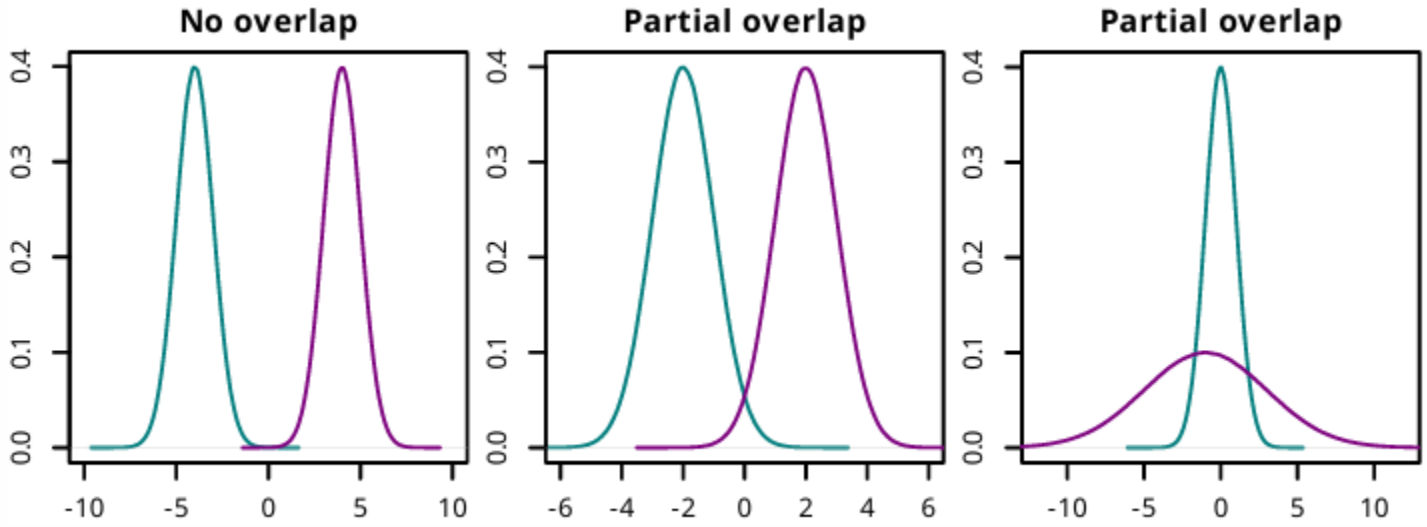
TABLE 2.2.1
Comparison of treatment and control characteristics in the Tennessee STAR experiment

Variable	Class Size			*P*-value for equality across groups
	Small	Regular	Regular/Aide	
Free lunch	.47	.48	.50	.09
White/Asian	.68	.67	.66	.26
Age in 1985	5.44	5.43	5.42	.32
Attrition rate	.49	.52	.53	.02
Class size in kindergarten	15.10	22.40	22.80	.00
Percentile score in kindergarten	54.70	48.90	50.00	.00

Notes: Adapted from Krueger (1999), table I. The table shows means of variables by treatment status for the sample of students who entered STAR in kindergarten. The *P*-value in the last column is for the *F*-test of equality of variable means across all three groups. The free lunch variable is the fraction receiving a free lunch. The percentile score is the average percentile score on three Stanford Achievement Tests. The attrition rate is the proportion lost to follow-up before completing third grade.

Overlap

- **Overlap** describes how similar the range of the data is across groups.
- **Overlap** means that for every combination of observed covariate Z, there is a non-zero probability of observing both treatment and control.
- A lack of **overlap** means that there are no equivalents in the two groups , and we may have to extrapolate beyond the support of the data.



Blocked Experiment

When designing an experiment, we can use **prior information** to get more precise and accurate estimates. Imagine an experiment to test the efficacy of a training program:

- We know that age may be an important factor.
- We could divide the data into different blocks.
- Subjects in a block should have similar ages.
- **Random assignment** of the treatment happens within an age block.
- This helps us minimize issues with imbalance and overlap by running **several** small experiments.

Blocked Experiment(2)

If we conduct an experiment with B blocks:

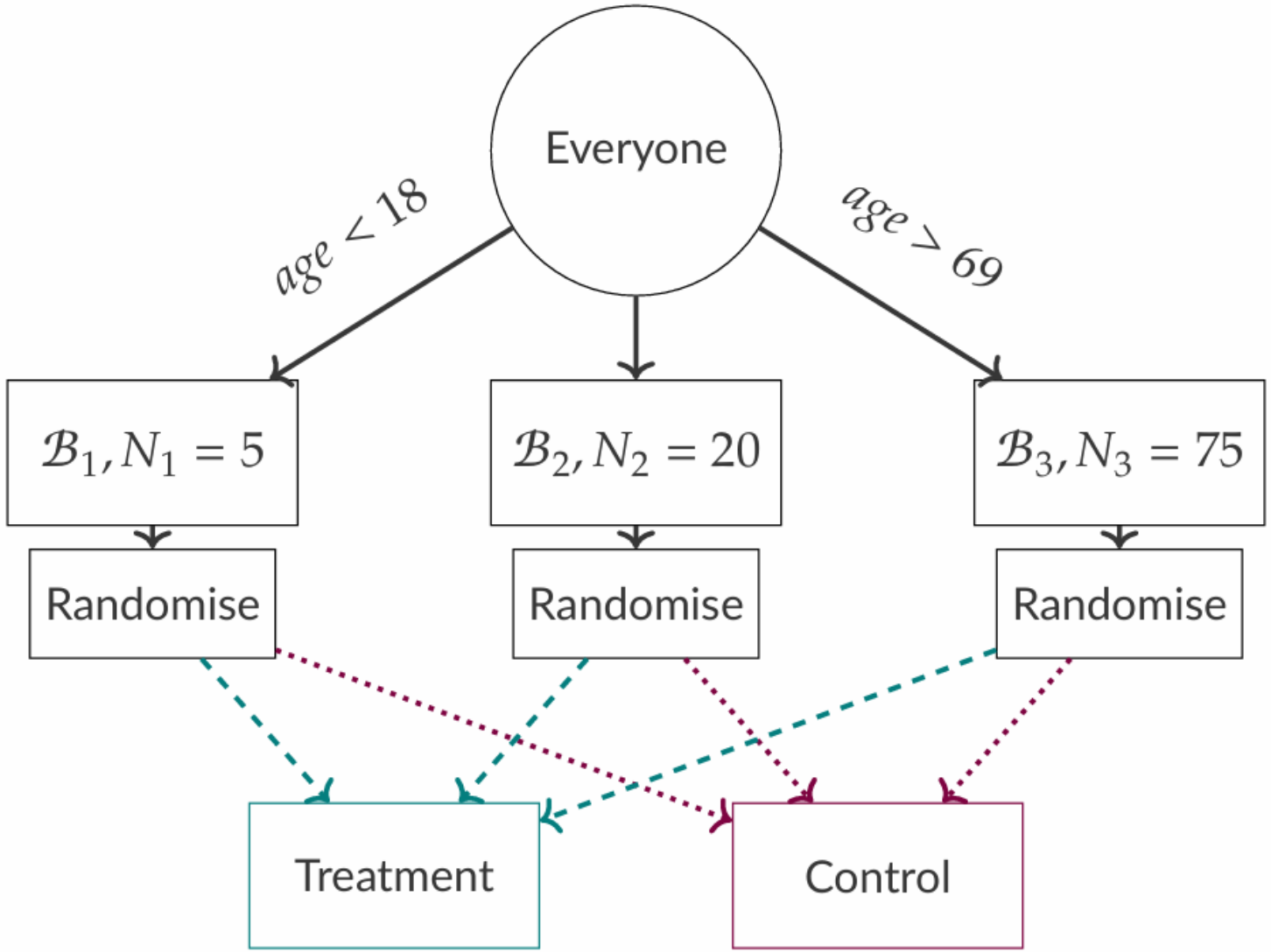
- we can estimate the **ATE** within a block B_b by comparing the sample averages and estimate the overall **ATE** by taking a weighted average:

$$\hat{\tau}^b_{ATE} = \mathrm{E}[Y_j(1)] - \mathrm{E}[Y_j(0)] \text{ where } j \in B_b \hat{\tau}_{ATE} = \frac{\sum_i N_i \hat{\tau}^i}{\sum_i N_i}.$$

- Or by estimating a regression with block indicators

$$y_i = \alpha + x_i \tau_{ATE} + \mathbb{1}(i \in B_1)\gamma_1 + \ldots + \mathbb{1}(i \in B_b)\gamma_b + e_i.$$

Blocked Experiment(3)



Causality

Identification

Randomization

Practice

Directed Acyclical Graphs

Application: DAGs in Research

Exercise

Below are 2 causal questions:

- How would you answer them?
- What model to use?
- What are the issues?
- What solutions can be applied?

Practice task

(1) The effect of media on voting preferences?

(2) The effect of gang presence on income?

Identification

Randomization

Practice

Directed Acyclical Graphs

Application: DAGs in Research

Ways of Viewing Causal Questions

The **Potential Outcomes (PO)** framework, which we covered last week, is **one way** to view causal questions.

- There is a **treatment** X_i that takes on different values for each unit.
- For each possible level of treatment, there is a certain **potential outcome** $Y(x)$.
- Only **one** potential outcome is **observed**, the others are **counterfactuals**.

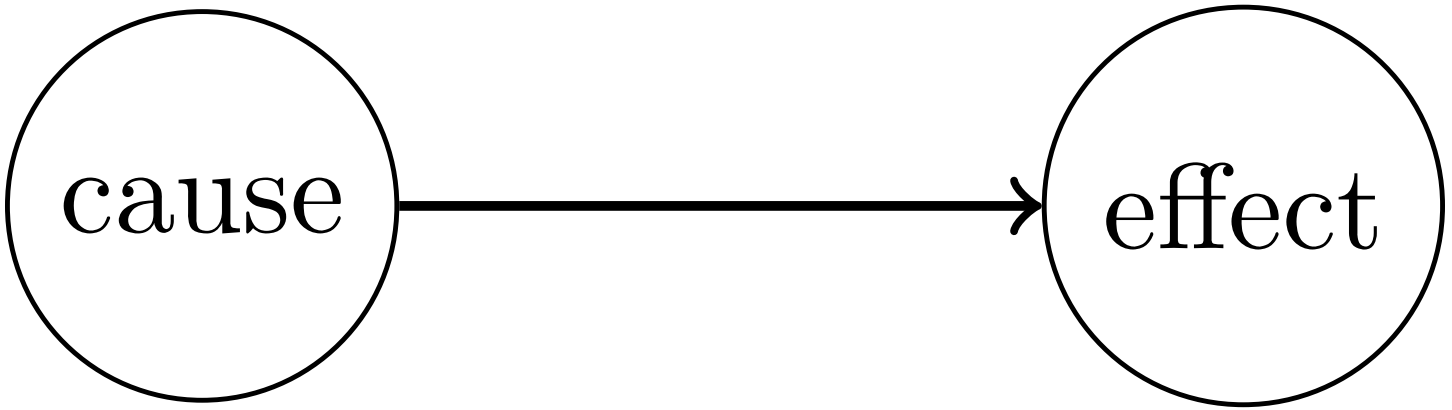
The **potential outcomes** framework relates very clearly to the notion of a **randomized experiment**.

Today, we are discussing a **different framework** that has its strengths **elsewhere**: the **Directed Acyclical Graphs (DAG)** framework.

- It is a **graphical framework** that helps us identify a causal effect in a network of variables.
- It has its **strengths** in a world with a **large number of (observed) variables**, and
- may help people who prefer thinking **graphically** to understand **causal questions**.

A Glorified Flowchart?

To give you an intuition before we start with the theory, a **DAG** looks like this:

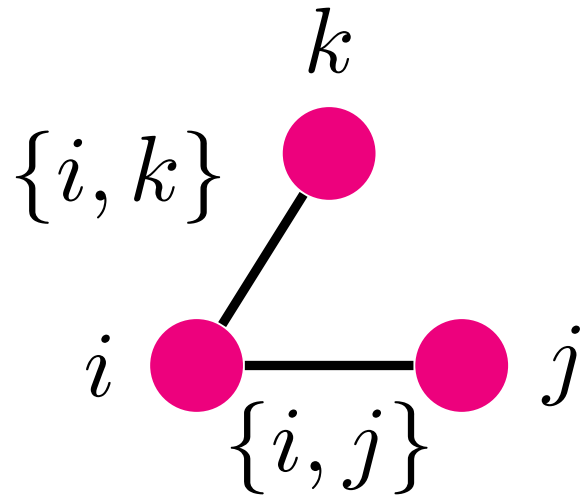


- They are similar in concept to **flowcharts**, but they are not the same.
- You might occasionally have seen them in informal use, e.g. in Econometrics I.
- In this example, the **arrow** represents a causal effect from cause on effect.

A Short Intro to Graph Theory

What you see on the right is what we call a **graph**.

This graph has three **nodes**. They are labeled i, j, and k. Sometimes, we call the **nodes** "vertices," "agents," "points," etc.



Some of the **nodes** in a graph are usually connected to each other, while others are not. We call those connections **edges**. Alternatively, they can be called "links," "connections," "lines," etc.

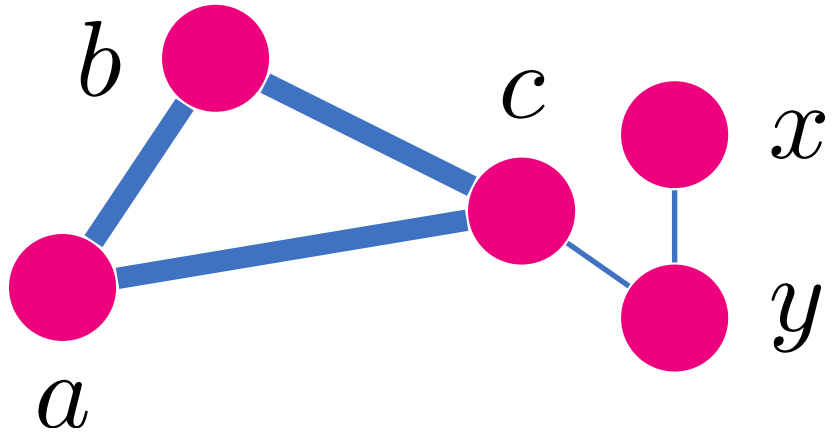
Edges are pairs of two **nodes**. In the second graph, there is one **edge** from i to j. We call this **edge** $\{i, j\}$.

A Short Intro to Graph Theory

This **edge** does not have a direction.

However, we can easily give **edges** a direction. We call an **edge** like this a **directed edge**. When an **edge** is **directed**, the corresponding pair of **nodes** is no longer an **unordered pair**, but an **ordered pair**: $\{j, i\} \neq \{i, j\}$.

A **walk** is a sequence of **edges** that joins a sequence of **nodes**. A **cycle** is a special case of a **walk** where all **edges** are **distinct** and the initial and final **node** are **equal**. In this graph, $\{\{a, b\}, \{b, c\}, \{c, a\}\}$ is a **cycle**.



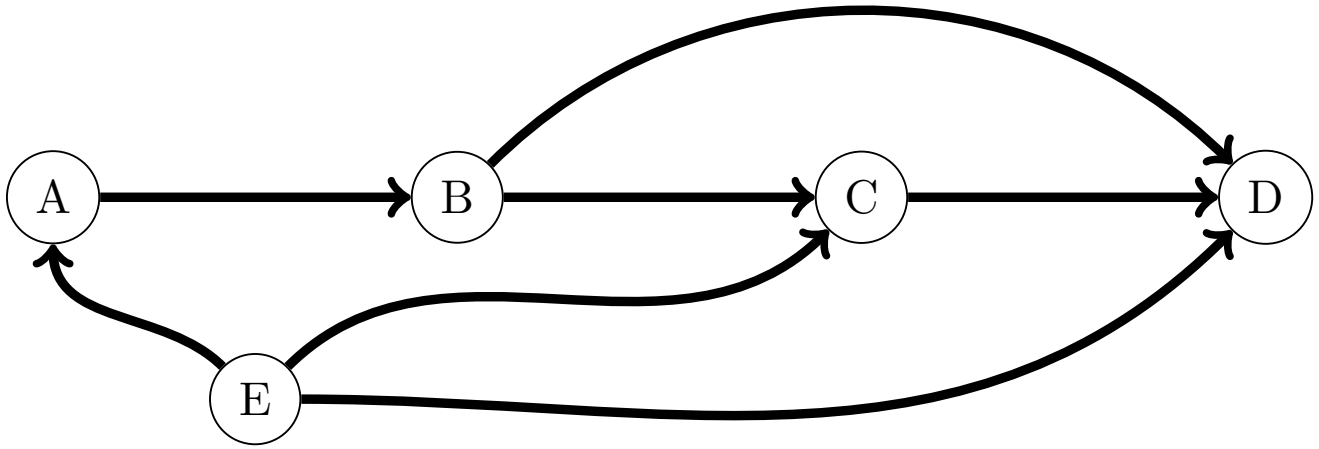
Directed Graphs, Acyclic Graphs

A **graph** that does not contain any **cycles** is called an **acyclic graph**.

If a graph contains only **directed edges**, we call it a **directed graph**.

The following graph is both **directed** and **acyclic**. We therefore call it a

Directed Acyclic Graph (DAG).



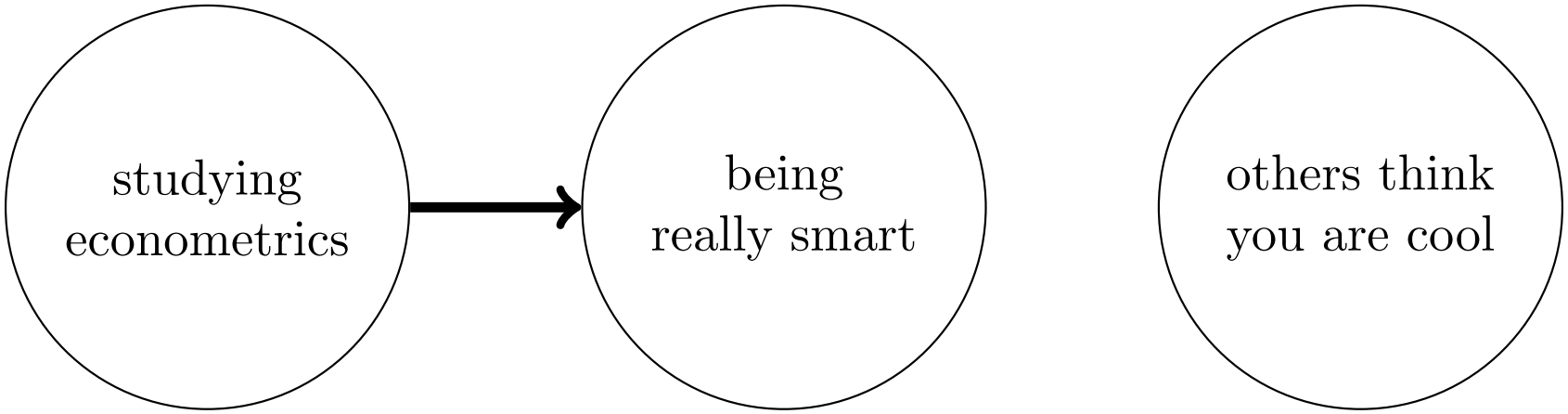
Think

Why is $\{\{A, B\}, \{B, C\}, \{C, E\}, \{E, A\}\}$ not a cycle?

DAGs for Causal Modeling

Why do we talk about **DAGs** in an **Econometrics** class? Because they are really useful for **causal modeling**.

In the following **DAG**, **nodes** represent **(random) variables**, and **edges** represent (hypothesized) **causal effects**.



Missing edges also convey information: the assumption of **no causal effect**.

DAGs and Causal Inference

DAGs are a very useful **framework for causal inference** because

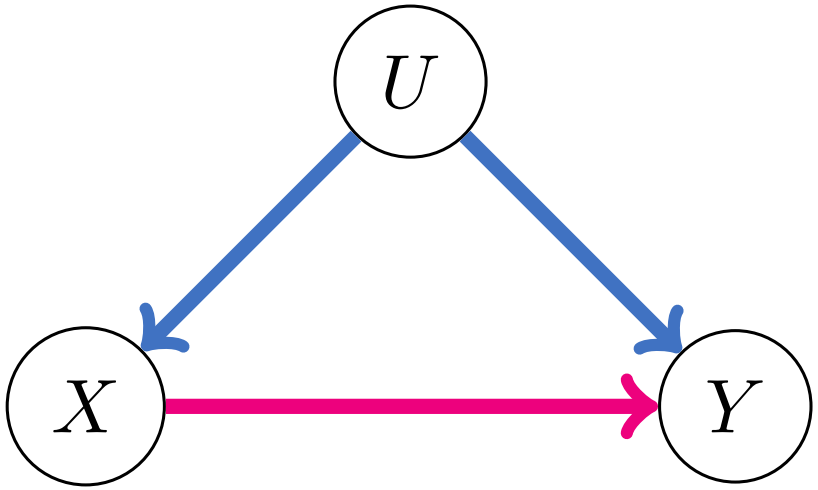
- they **visualize** causal relationships between a number of variables,
    - which allows us to transparently state our **assumptions**,
- and they help us **identify** a causal effect,
    - i.e., they tell us which variables to **control** for to estimate an effect.

The Basics of Causal Inference with DAGs

- Is Y **related** to U?
- Is X **related** to U? Can we **randomize** treatment?
- Are there **other** important variables?

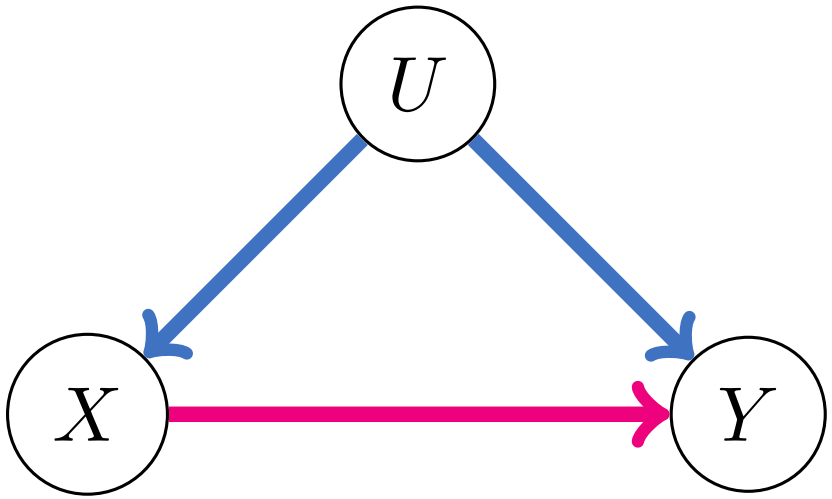
It turns out that there are **two paths** from X to Y,

- one **direct path** $X \to Y$
- and one **backdoor path** $X \leftarrow U \to Y$.



We call it a **backdoor path** because it enters X trough the "back door," via an arrow pointed *at* X.

Confounders



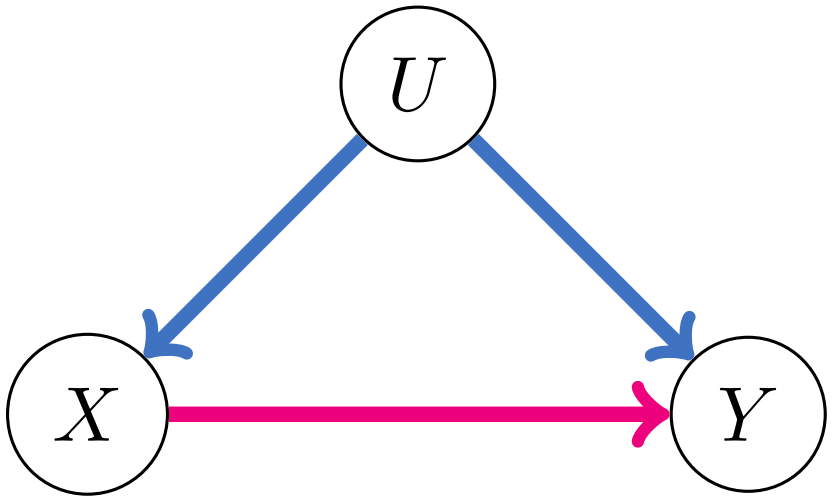
In this DAG, when we want to **isolate** the effect $X \to Y$, there is one **open backdoor path**.

This path **confounds** the causal effect of interest. We therefore call the variable U a **confounder**.

Confounder

A **confounder** is a variable that influences both the dependent and the explanatory variables.

Confounders and Backdoors



If we just look at the connection between X and Y, two effects are mixed together:

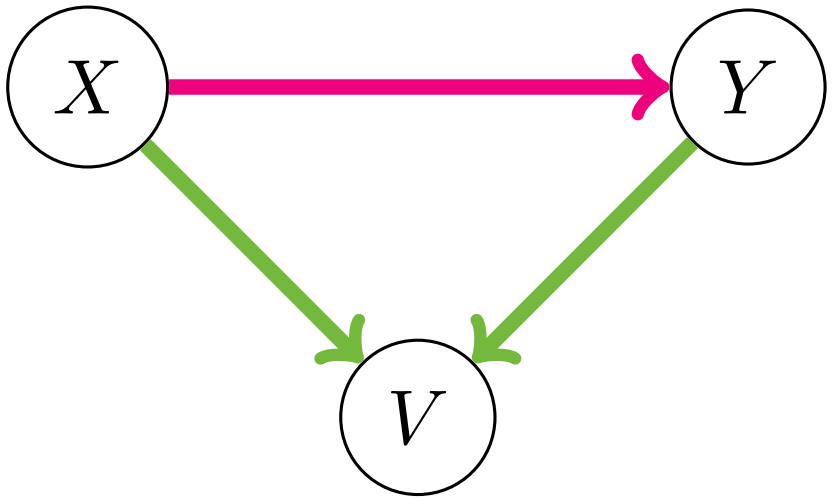
- The effect of X on Y, our **effect of interest**.
- The effect of U on Y via X.

We can **close the backdoor** by **controlling** for the confounder. We only run into problems when we **cannot control** for the confounder.

We would run a **regression** along the lines of:

$$\boldsymbol{y \sim x + u.}$$

Colliders



Now imagine a **different situation**: There is a third variable, V, that is jointly **influenced** by X and Y.

Effects of both variables **collide** at V. We therefore call V a **collider**. There is again one **direct path** and one **backdoor path**, but since the backdoor collides at V, it is **already closed**.

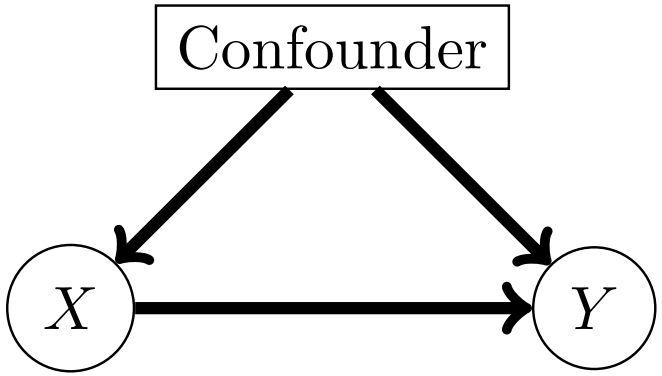
Collider

A **collider** is a variable that is influenced by both the dependent and the explanatory variables.

Closing Backdoors

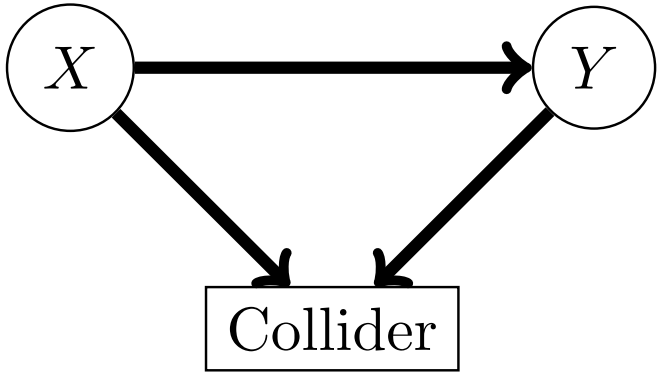
Open **backdoors** between two variables introduce **systematic, non-causal correlation** between them. If we want to estimate a **causal effect**, we need to **close them**. There are **three cases** we have to consider:

Confounders



We **close backdoor paths** by **controlling** for confounders.

Colliders



We can (and **need** to) **leave colliders alone**. The backdoor path is already closed.

Mediators

A **mediator** mediates part of the effect. If we **control for the mediator**, we **remove** the mediated effect and leave only the direct effect.

Randomization

Practice

Directed Acyclical Graphs

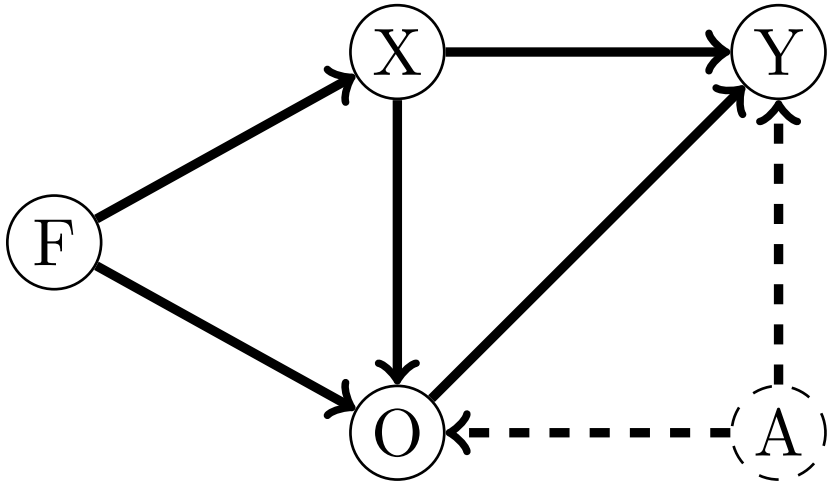
Application: DAGs in Research

Enumerating Paths

How does this framework look like if we apply it to an **example**? Let us look at the following graph on the effect of **gender** (F) based **discrimination** (X) on **earnings** (Y).

We account for **occupation** (O) and **aptitude** (A).

Note that aptitude is **not observed**.



How many paths from X to Y can we **enumerate**?

Enumerating Paths

How many paths between X and Y can we **enumerate**?

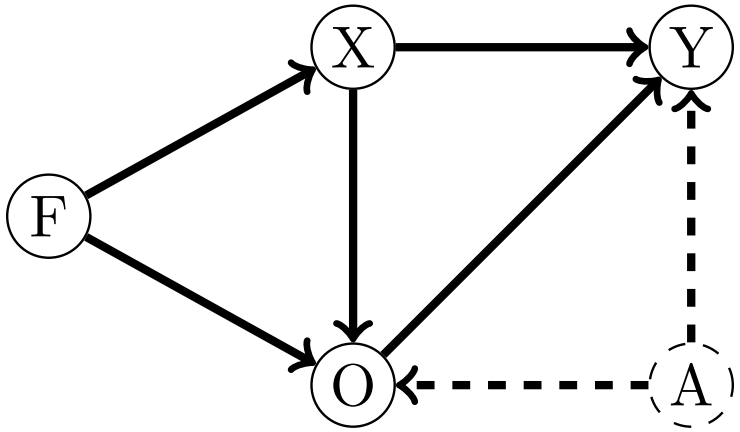
(1) $X \to Y$,

(2) $X \to O \to Y$,

(3) $X \to O \leftarrow A \to Y$,

(4) $X \leftarrow F \to O \to Y$,

(5) $X \leftarrow F \to O \leftarrow A \to Y$.



Which models can we use to isolate the effect of interest?

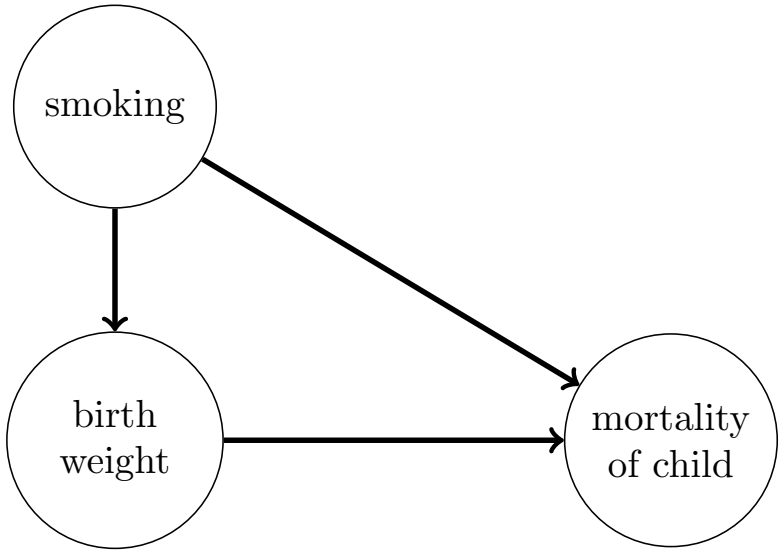
- $Y \sim F$: We get a compound effect of X and O (1, 2, 4).
- $Y \sim X$: We get the effects of X, but they are confounded by F (4).
- $Y \sim X, O$: We get rid of the confounder F and separate the effects of X (1, 2), but they are now confounded by A (3, 5).

Without A, we cannot isolate the causal effect of X on Y in this model. **DAGs** can highlight what **cannot be done**.

Does Smoking During Pregnancy Protect Your Child?

- The debate about whether smoking causes cancer was settled by the 1960s. Its conclusion had been delayed by multiple years because scientists disagreed on the meaning of "to **cause**," and **no ways of discerning causal effects from observational data** were available.
- But even afterwards, one paradox remained. Some researchers argued that **smoking during pregnancy was actually good – *if* the unborn child was underweight**.
- The paradox was **not resolved *until 2006***. Pearl & Mackenzie (2018) argue that it took so long because precise language of causality was not yet available.

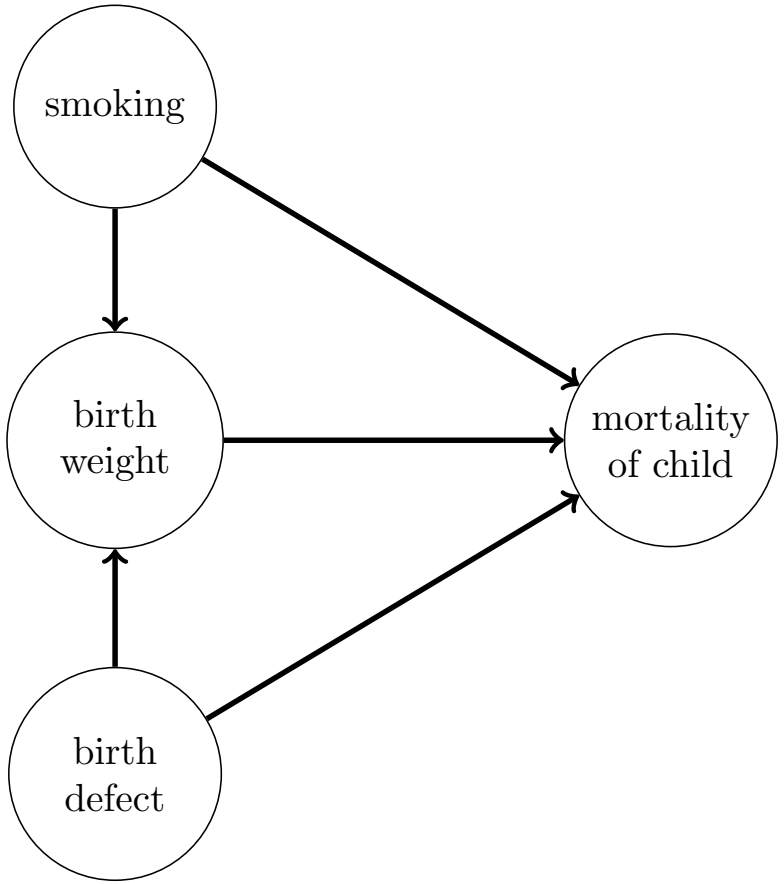
Smoking Mothers



- Underweight infants were found to have a **death rate twenty times higher** than normal-weight newborns.
- Babies of smokers during pregnancy were on average 200 grams lighter than those of non-smokers.
- However, **underweight babies of smoking mothers** had a **higher survival rate** than **underweight babies of non-smoking mothers**.

How come?

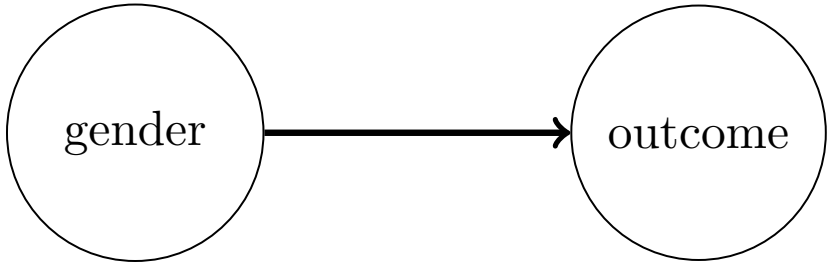
Smoking Mothers



- Scientists at the time cautiously concluded that smoking may not affect the development of the fetus.
- **However**, another explanation makes much more sense:
  - There are (thinking in a blatantly simplified way) **two possible causes** for a low birth weight: **Smoking** and having a **birth defect**.
  - If a mother does not smoke, a low birth weight points much more strongly to a birth defect.
  - **Birth weight** acts as a **collider**.

- The original **paradox** becomes a (literal) textbook case of **collider bias**.

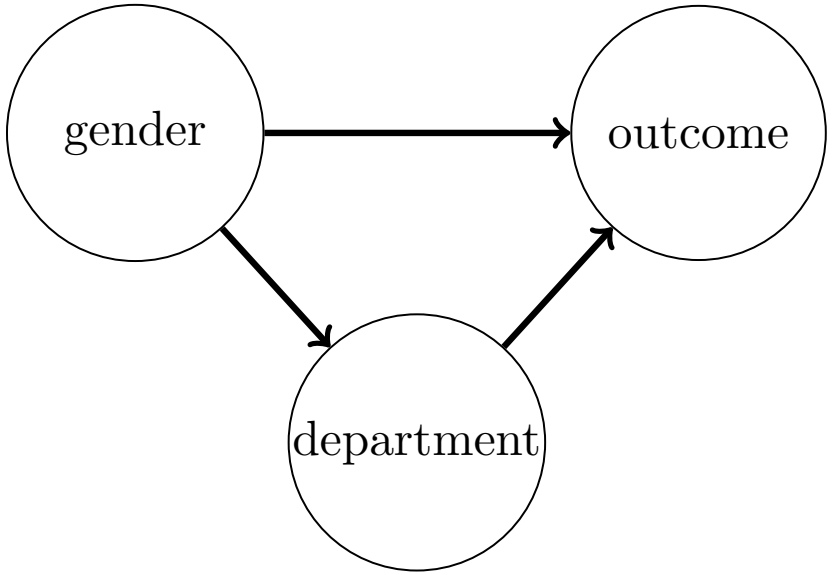
The Berkeley Admissions Paradox



- In 1978, an associate dean at the University of California noticed that **44 percent of applying men** were admitted, but only **35 percent of women**.
- Admission decisions were made by **individual departments**.
- The university surveyed all departments, and found that in every department, **admission decisions were more favorable to women than to men**.

How is this possible?

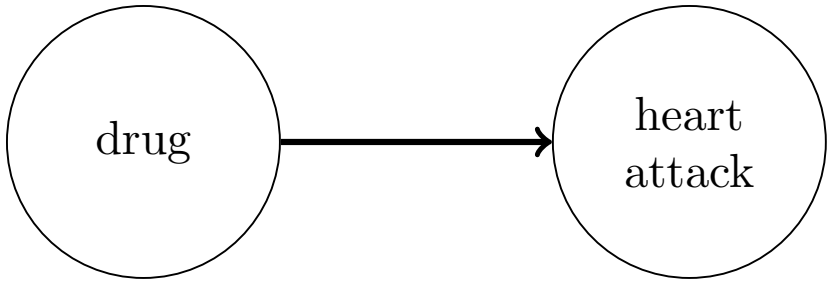
The Berkeley Admissions Paradox



- It turns out that the $\text{gender} \rightarrow \text{outcome}$ relation has an important **mediator**.
- Discrimination is a **causal concept**, and thus a causal graph can help understand the situation.
- Women were applying to **different departments**/majors than men.
  - More **women** applied to humanities departments, which were **harder to get into**.

- The **choice of department** is a mediator. Whether we want to **condition on the mediator** depends on the **specific question** we want to answer.
- In this case, it depends on our **understanding of discrimination** as well as whether we ask about a societal phenomenon or whether the university is at fault.

Simpson's Paradox

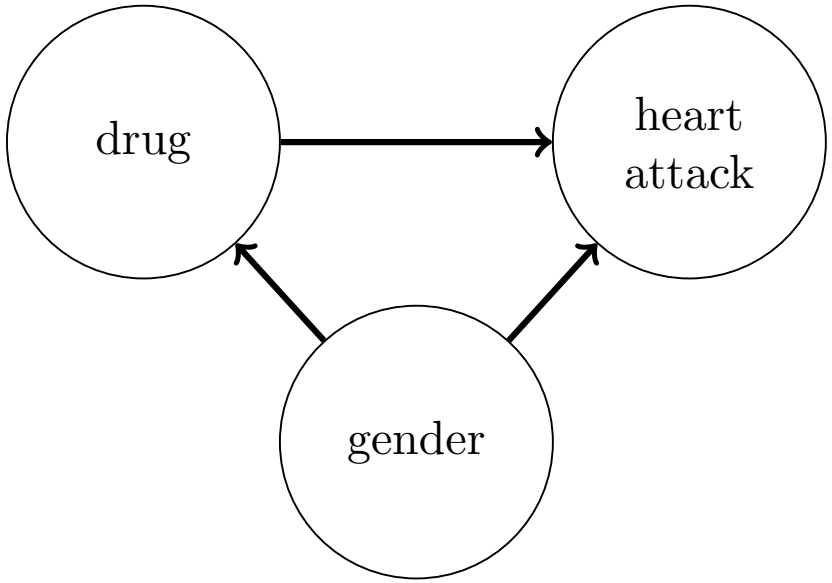


	No Drug		Took Drug	
	Heart Attack	**No Heart Attack**	**Heart Attack**	**No Heart Attack**
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

- Assume a fictional doctor, Dr. Smith, reads about a drug that **reduced** the probability of a **heart attack** among the subjects that took the drug (no randomized trial).
- However, both in **men** and in **women**, the drug seemed to **increase** the propensity to suffer from a heart attack.

How does this "Bad-Bad-Good (BBG)" drug paradox arise?

Simpson's Paradox



	No Drug		Took Drug	
	Heart Attack	**No Heart Attack**	**Heart Attack**	**No Heart Attack**
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

- Actually, the drug is just **plain bad**.
- Since gender is a **confounder** and affects both the propensity to take a drug and the chance of a heart attack, we need to **control for it** when comparing totals.
- In the example, we can do this by **looking at both groups separately** and then averaging. We find out that there is a **negative effect** for **both** women and men, and so the **aggregate effect** is also **negative**.

Summary Slide

- All of the previous three examples are taken from “The Book of Why” by Pearl & Mackenzie (2018). You can read more about them, and about similar examples, in the book.
    - The **smoking mothers** paradox highlighted what can happen if we improperly treat **colliders**.
    - The **Berkeley admissions** paradox was a case of ignoring an important **mediator**
    - **Smith’s Paradox** was a case of ignoring a **confounder**.
- **DAGs** are useful tools, particularly when using observational data, to **visualize** causal networks and make **confounders**, **colliders**, and **mediators** explicit.

References

Cunningham, S. (2021). *Causal inference*. Yale University Press. https://doi.org/10.12987/9780300255881

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning*. Springer US. https://doi.org/10.1007/978-1-0716-1418-1

Pearl, J. (2009). Causality. In *Cambridge Core*. Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic books.