# Causal Inference I

## Mixtape Session

MIXTAPE SESSIONS

# Roadmap

- Directed Acyclic Graphs
  - Graph notation
  - Backdoor criterion
  - Collider bias
  - Front door criterion
  - Concluding remarks

Judea Pearl, 2011 Turing Award winner, drinking his first IPA

# Judea Pearl and DAGs

- Judea Pearl and colleagues in Artificial Intelligence at UCLA developed DAG modeling to create a formalized causal inference methodology
- They causality concepts extremely clear, they provide a map to the estimation strategy, and maybe best of all, they communicate to others what must be true about the data generating process to recover the causal effect

# Further reading

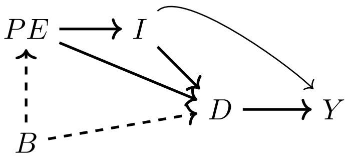
1. Pearl (2018) The Book of Why: The New Science of Cause and Effect, Basic Books (*popular*)
2. Morgan and Winship (2014) Counterfactuals and Causal Inference: Methods and Principles for Social Research, Cambridge University Press, 2nd edition (*excellent*)
3. Pearl, Glymour and Jewell (2016) Causal Inference In Statistics: A Primer, Wiley Books (*accessible*)
4. Pearl (2009) Causality: Models, Reasoning and Inference, Cambridge, 2nd edition (*difficult*)
5. Cunningham (2021) Causal Inference: The Mixtape, Yale, 1st edition (*best choice, no question*)

# Design vs. Model

- DAGs tend to be focused more on the theory of treatment assignment in the world
- As such it's compatible with design-based approaches
- But assumptions in design based approaches tend to emphasize selection into treatment which is not exactly what is meant here

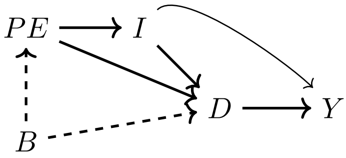
# Causal model

- The causal model is sometimes called the structural model, but for us, I prefer the former as it's less alienating
- Think of this as more connected to the model-based approach discussed earlier
- It's the system of equations describing the relevant aspects of the world
- It necessarily is filled with causal effects associated with some particular comparative statics




- $B$ is a **parent** of $PE$ and $D$
- $PE$ and $D$ are **descendants** of $B$
- There is a **direct (causal) path** from $D$ to $Y$
- There is a **mediated (causal) path** from $B$ to $Y$ through $D$
- There are four **paths** from $PE$ to $Y$ but none are direct, and one is unlike the others

# Colliders




Notice anything different with this DAG? Look closely.

- $D$ is a **collider** along the path $B \to D \leftarrow I$ (i.e., "colliding" at $D$)
- $D$ is a **noncollider** along the path $B \to D \to Y$

# Summarizing Value of DAGs

1. Facilitates the task of designing identification strategy for estimating average causal effects
2. Facilitates the task of testing compatibility of the model with your data
3. Visualizes the identifying assumptions which opens up the model to critical scrutiny

# Creating DAGs

- The DAG is a *relevant* causal relationships describing the relationship between $D$ and $Y$
- It will include:
  - → All direct causal effects among the *relevant* variables in the graph
  - → All common causes of any pair of *relevant* variables in the graph
- No need to model a dinosaur stepping on a bug causing in a million years some evolved created that impacted your decision to go to college
- We get ideas for DAGs from theory, models, observation, experience, prior studies, intuition
- Sometimes called the data generating process.

# Research designs: Selection on observables

- DAGs help us understand the source of problems in our observational (non-experimental) data that make inferring causality hard
- But it also can help us see a way out (this whole class is a way out)
- First way out is the selection on observables research design, which is best described in causal graphs using the **backdoor criterion**
- Selection on observables is not technically difficult, but it does require a reasonable level of confidence in the DAG

# Confounding

- Omitted variable bias has a name in DAGs: "confounding"
- Confounding occurs when when the treatment and the outcomes have a common cause or parent which creates spurious correlation between $D$ and $Y$

$$\begin{tikzcd} D \arrow[rr] & & Y \\ & X \arrow[ul] \arrow[ur] & \end{tikzcd}$$

The *correlation* between $D$ and $Y$ no longer reflects the causal effect of $D$ on $Y$

# Backdoor Paths

- Confounding creates **backdoor paths** between treatment and outcome ($D \leftarrow X \rightarrow Y$) – i.e., spurious correlations
- Not the same as mediation ($D \rightarrow X \rightarrow Y$)
- We can "block" backdoor paths by conditioning on the common cause $X$
- Once we condition on $X$, the correlation between $D$ and $Y$ estimates the causal effect of $D$ on $Y$
- Conditioning means calculating $E[Y|D=1,X] - E[Y|D=0,X]$ for each value of $X$ then combining (e.g., integrating)

$$\begin{tikzcd} D \arrow[rr] & & Y \\ & \boxed{X} \arrow[ul] \arrow[ur] & \end{tikzcd}$$

# Blocked backdoor paths

A backdoor path is blocked if and only if:

- It contains a noncollider that has been conditioned on
- Or it contains a collider that has not been conditioned on

# Examples of blocked paths

Examples:

1. Conditioning on a noncollider blocks a path:

$$\begin{tikzcd} \boxed{X} \arrow[r] \arrow[rr, bend left] & Z & Y \end{tikzcd}$$

2. Conditioning on a collider opens a path (i.e., creates spurious correlations):

$$\begin{tikzcd} Z \arrow[r] & \boxed{X} & Y \arrow[l] \end{tikzcd}$$

3. *Not* conditioning on a collider blocks a path:

$$Z \longrightarrow X \longleftarrow Y$$

# Backdoor criterion

## Backdoor criterion

Conditioning on $X$ satisfies the backdoor criterion with respect to $(D, Y)$ directed path if:

1. All backdoor paths are blocked by $X$
2. No element of $X$ is a collider

In words: If $X$ satisfies the backdoor criterion with respect to $(D, Y)$, then controlling for or matching on $X$ identifies the causal effect of $D$ on $Y$

# What control strategy meets the backdoor criterion?

- List all backdoor paths from $D$ to $Y$. I'll wait.

$$
\begin{tikzcd}
X_1 \arrow[r] \arrow[rr, bend left] & D \arrow[r] & Y \\
& & X_2 \arrow[ul] \arrow[u]
\end{tikzcd}
$$

- What are the necessary and sufficient set of controls which will satisfy the backdoor criterion?

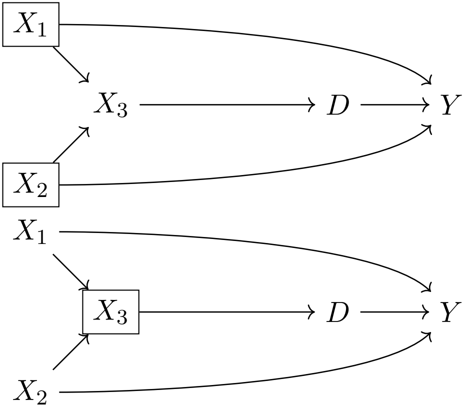
# What if you have an unobservable?

- List all the backdoor paths from $D$ to $Y$.

$$\begin{tikzcd}
 & X_1 \arrow[dr] \arrow[drr] & & \\
U \arrow[r, dashed] \arrow[rrr, dashed, bend right] & X_2 \arrow[r] & D \arrow[r] & Y
\end{tikzcd}$$

- What are the necessary and sufficient set of controls which will satisfy the backdoor criterion?
- What about the unobserved variable, $U$?

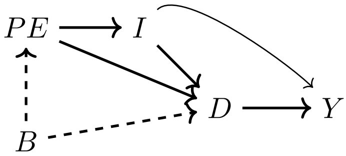
# Multiple strategies




- Conditioning on the common causes, $X_1$ and $X_2$, is sufficient
- …but so is conditioning on $X_3$

# Testing the Validity of the DAG

- The DAG makes testable predictions
- Conditional on $D$ and $I$, parental education ($PE$) should no longer be correlated with $Y$
- Can be hard to figure this out by hand, but software can help (e.g., Daggity.net is browser based, Causal Fusion is more advanced)
- Causal algorithms tend to be DAG based and are becoming popular in industry



# Collider bias

- **Conditioning on a collider introduces spurious correlations; can even mask causal directions**
  - → There is only one backdoor path from $D$ to $Y$

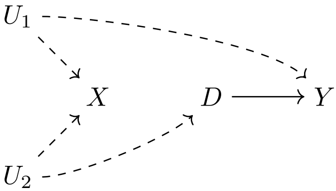
$$
\begin{tikzcd}
\boxed{X_1} \arrow[r] \arrow[rrr, bend left] & D \arrow[rr] \arrow[dr] & & Y \arrow[dl] \\
& & X_2 &
\end{tikzcd}
$$

  - → Conditioning on $X_1$ blocks the backdoor path
  - → But what if we also condition on $X_2$?

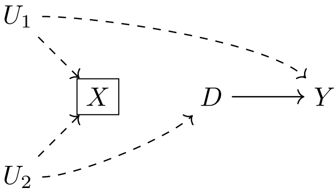
$$
\begin{tikzcd}
\boxed{X_1} \arrow[r] \arrow[rrr, bend left] & D \arrow[rr] \arrow[dr] & & Y \arrow[dl] \\
& & \boxed{X_2} &
\end{tikzcd}
$$

  - → Conditioning on $X_2$ opens up a new path, creating new spurious correlations between $D$ and $Y$

- **Even controlling for pretreatment covariates can create bias**
  - → Name the backdoor paths. Is it open or closed?




  - → But what if we condition on $X$?



# Sample selection example of collider bias

Important: Since unconditioned colliders block back-door paths, what exactly does conditioning on a collider do? Let's illustrate with a fun example and some made-up data

- CNN.com headline: Megan Fox voted worst – but sexiest – actress of 2009 (link)
- Are these two things actually negatively correlated in the world?
- Assume talent and beauty are independent, but each causes someone to become a movie star. What's the correlation between talent and beauty for a sample of movie stars compared to the population as a whole (stars and non-stars)?

- What if the sample consists *only* of movie stars?
- Look at python code

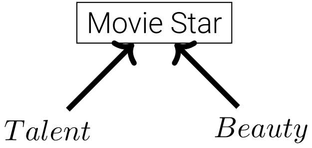
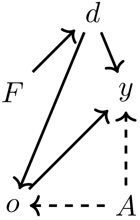



*Figure:* Top left figure: Non-star sample scatter plot of beauty (vertical axis) and talent (horizontal axis). Top right right figure: Star sample scatter plot of beauty and talent. Bottom left figure: Entire (stars and non-stars combined) sample scatter plot of beauty and talent.

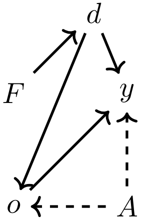
# Occupational sorting and discrimination example of collider bias

- Let's look at another example: very common for think tanks and journalists to say that the gender gap in earnings disappears once you control for occupation.
- But what if occupation is a collider, which it could be in a model with occupational sorting
- Then controlling for occupation in a wage regression searching for discrimination can lead to all kinds of crazy results *even in a simulation where we explicitly design there to be discrimination*

# DAG




$F$ is female, $d$ is discrimination, $o$ is occupation, $y$ is earnings and $A$ is ability. Dashed lines mean the variable cannot be observed. Note, by design, being a female has no effect on earnings or occupation, and has no relationship with ability. So earnings is coming through discrimination, occupation, and ability.




Mediation and Backdoor paths

1. $d \to o \to y$
2. $d \to o \leftarrow A \to y$

*Table:* Regressions illustrating collider bias with simulated gender disparity

| Covariates: | **Unbiased combined effect** | **Biased** | **Unbiased wage effect only** |
|---|---|---|---|
| Female | -3.074*** | 0.601*** | -0.994*** |
| | (0.000) | (0.000) | (0.000) |
| Occupation | | 1.793*** | 0.991*** |
| | | (0.000) | (0.000) |
| Ability | | | 2.017*** |
| | | | (0.000) |
| N | 10,000 | 10,000 | 10,000 |
| Mean of dependent variable | 0.45 | 0.45 | 0.45 |

- Recall we designed there to be a discrimination coefficient of -1
- If we do not control for occupation, then we get the combined effect of $d \rightarrow o \rightarrow y$ and $d \rightarrow y$
- Because it seems intuitive to control for occupation, notice column 2 - the sign flips!
- We are only able to isolate the direct causal effect by conditioning on ability and occupation, but ability is unobserved

# Research design #2: front door criterion

- Confounding creates major issues for us, but if we can observe the confounders, then we can use the backdoor criterion ("selection on observables") to identify causal effects
- What about **unobserved confounding**? It depends on the DAG
- One particular DAG structure that is not widely known outside of Pearl circles is the front door criterion
- Bears some topical resemblance to instrumental variables, but it is nonetheless very different and when available to you the elements can be used to trace out causal effects

# Mechanisms

- Rarely does an intervention operate directly on an outcome; oftentimes it operates on the outcome via a "mechanism"
  - → Example: Parental substance abuse causes foster care removals through child abuse and neglect
- The presence of mechanisms, it turns out, is valuable because of their policy relevance, but also because we can use them *sometimes* for identification

# Frontdoor DAG

$$\begin{tikzcd} D \arrow[r] & M \arrow[r] & Y \\ & U \arrow[ul, dashed] \arrow[ur, dashed] & \end{tikzcd}$$

- We cannot close $D \leftarrow U \rightarrow Y$ because $U$ is not observed and thus simple contrasts are biased estimates of treatment effects
- Pearl (2009) showed that this DAG actually does allow us to recover the effect of $D$ on $Y$, though – just not via the backdoor criterion

# Front door criterion

*If one or more unblocked back door paths connect a causal variable to an outcome variable, the causal effect is identified by conditioning on a set of observed variables $M$ that make up the identifying mechanism if and only if: 1) the variables in $M$ intercept all directed paths from the causal variable to the outcome ("exhaustiveness"); 2) No unblocked back-door paths connecting the causal variable to the variables in the set $M$ and all back door paths from the variables in $M$ to the outcome can be blocked by conditioning on $D$ ("isolation")*

# Exhaustiveness

$$\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}$$

- Exhaustiveness means the variables $M$ are the only paths through which $D$ impacts $Y$.
- In other words, rules out direct effects that bypass $M$ altogether
- "only through $M$" in place of exhaustiveness and you get the idea

# Isolation

$$\begin{tikzcd} D \arrow[r] & M \arrow[r] & Y \\ & U \arrow[ul, dashed] \arrow[ur, dashed] & \end{tikzcd}$$

- Mechanism itself is not confounded with respect to $Y$ (i.e., no unobserved confounding parent linking $M$ and $Y$)
- You are looking for a causal effect contained in a closed but confounded system and the presence of the $M$ mechanism is key
- Pearl and others have suggested smoking ($D$) and lung cancer ($Y$) with $M =$ tar buildup in the lungs might have been candidate but this has been debated)

# Frontdoor three step method

$$
\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}
$$

- Frontdoor criterion is going to take advantage of two things we've seen so far: collider properties and blocking properties
- This is not IV, but as we will see, it bears some similarities to IV
- The final estimator will be the product of two separate calculations (whereas IV is more like the ratio)

# Frontdoor three step method

$$
\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}
$$

1. Estimate the effect of $D$ on $M$. Consider a regression of $M$ on $D$ or simple difference in mean $D$ with respect to $M$

$$D = \alpha_0 + \beta M + \epsilon$$

- $\rightarrow$ $M$ is isolated, so it is not confounded
- $\rightarrow$ $D \leftarrow U \rightarrow Y \leftarrow M$ which is blocked bc $Y$ is a **collider**
- $\rightarrow$ Therefore $\widehat{\beta}$ identifies $\beta$

# Frontdoor three step method

$$\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}$$

2. Estimate the effect of $M$ on $Y$ conditional on $X$
   - → Gets you an unbiased estimate of $M$ effect on $Y$ bc only backdoor path from $M$ to $Y$ is $M \leftarrow D \leftarrow U \rightarrow Y$
   - → So long as we condition on $D$ this path is blocked

$$Y = \alpha_1 + \gamma M + \psi D + \varepsilon$$

# Frontdoor three step method

$$
\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}
$$

3. Multiply $\widehat{\gamma} \times \widehat{\beta}$ and you get the causal effect of $D$ on $Y$

# Examples have been elusive

- Pearl has suggested smoking as an possible example of this but to be valid it requires smoking to not have a direct effect on lung cancer and if it is not the case, it would invalidate the frontdoor design
- Frontdoors requires “closed systems” (as do instruments), and it’s possible that in carefully designed platforms that could either be intentionally designed or happen naturally in a way that is defensible
- Bellemare, et al. (2021) provides a plausible example involving tipping and Uber

# Uber and tipping

- Shared rides could lead to reduced tipping but also increased demand thus creating principal agency issues for Uber and Lyft (drivers versus the firm)
- Harrington (2019): “on average, about 17% of rideshares end up with the driver getting tipped. For strips where a shared trip was authorized, that number is halved to a measly 8.6%.”
- Drivers experiencing such declines probably think it’s caused by sharing rides (e.g., bystander effects, freeriding, etc.) but it also may just be selection (i.e., the marginal rider would’ve tipped that low anyway)
- Bellemare, et al. (2021) suggest Uber’s platform design created FDC DAG that would allow this to be tested

# Assumed Uber Tipping DAG

$$
\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}
$$

- Let $D$ here be authorizing a shared ride (regardless of whether a shared ride occurred), $M$ be a dummy measuring one if sharing did occur, $Y$ be the amount the passengers tipped and $U$ be the unobserved covariates.
- Estimate the effect of authorization $(D)$ on both whether a passenger tips $(Y)$ as well as how much, what they call the extensive and intensive margin of tipping, respectively.
- Data come from Chicago's Department of Business Affairs and Consumer Protection's Transportation Network Providers and is freely available for download from the City of Chicago's website

# Assumptions

$$\begin{tikzcd}
D \arrow[r] & M \arrow[r] & Y \\
& U \arrow[ul, dashed] \arrow[ur, dashed] &
\end{tikzcd}$$

- Key assumption: once the authorization to share a ride is initiated ($D$), then when the ride is shared ($M$)
- No direct effect of authorization on tipping, no unblocked backdoor path from sharing a ride and tipping itself.
- Authors argue that their extensive set of fixed effects will yield plausible conditions for isolation and exhaustiveness are guaranteed.

# Estimation

- Using the logic of the front door criterion, the authors estimate the same two step procedure as shown in the previous simulation with the caveat that they include extensive fixed effects so as to create conditional conditions for isolation and exhaustiveness.
- For illustrative purposes, I will only focus on the effect at the extensive margin (i.e., on whether a passenger tipped at all).

*Table:* Estimation results for tipping at the extensive margin

| Variables: | Naive | Front Door | |
|---|---|---|---|
| | **Tipped** | **Shared Trip** | **Tipped** |
| Sharing authorized $D$ | -0.0628*** | 0.6769*** | -0.0550*** |
| | (0.0001) | (0.0002) | (0.0002) |
| Shared trip $M$ | | | -0.0115*** |
| | | | (0.0002) |
| Full fare | 0.0050*** | -0.0064*** | 0.0049*** |
| | (0.00001) | (0.00001) | (0.00003) |
| Estimated causal effect ($\widehat{\delta}$) | -0.0628*** | | -0.0078** |
| | (0.0001) | | (0.0001) |
| N | 95,670,449 | 95,670,449 | 95,670,449 |

# Interpretation

- Column 1: naive regression simply compares tipping between authorized and non-authorized sharing (6.3pp reduction in tipping)
- Front door criterion: 1pp reduction
- Not surprising drivers don't want ride shares, but authors argue it's caused by selection (i.e., the people using ride shares) not ride share itself
- Unclear if you banned it whether it would increase driver earnings in other words

# Discussion

- DAG front door criterion example. What is the strength of this approach in your opinion (no wrong answer)?
- What is the weakness of this approach in your opinion (no wrong answer)?
- Based on your own background, are there applications you might want to look for these opportunities?

# Summarizing all of this

- Your dataset will not come with a codebook flagging some variables as “confounders”, “mechanisms” and “colliders” because those terms are always context specific
- Except for some unique situations that aren’t generally applicable, you also don’t always know statistically you have an omitted variable bias problem; but both of these are fatal for any application
- You only know to do what you’re doing based on *knowledge about data generating process*.
- All identification must be guided by theory, experience, observation, common sense and knowledge of institutions
- DAGs absorb that information and can be then used to write out the explicit identifying model

# DAGs are not panacea

- DAGs cannot handle, though, reverse causality or simultaneity
- So there are limitations. “All models are wrong but some are useful”
- They are also not everywhere popular (see Twitter ongoing debates which have descended into light hearted jokes as well as aggressive debates)
- But I think they are helpful and while not *necessary*, showcase what is necessary – assumptions
- Heckman (1979) can maybe provide some justification at times