

# STATISTICAL MODELING AND CAUSAL INFERENCE WITH R

## Week 4: Causal Graphs

---

Manuel Bosancianu

Max Schaub

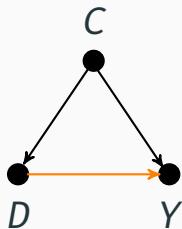
September 28, 2020

Hertie School

# d-separation

---

# Canonical configurations: confounders



$D \leftarrow C \rightarrow Y$  is an **open** back-door path ( $C$  is not a collider).

**Solution:** condition on  $C$  to close the path.

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 C_i + \epsilon_i \quad (1)$$

# Canonical configurations: mediators



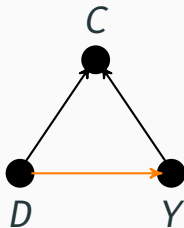
$D \rightarrow C \rightarrow Y$ : causal path between treatment and outcome.

**Solution:** do not condition on  $C$ , to keep the path open.

Rules:

- ✓ close all non-causal paths linking  $D$  to  $Y$
- ✓ do not close any causal path between  $D$  and  $Y$

# Canonical configurations: colliders



$D \rightarrow C \leftarrow Y$ : noncausal path between treatment and outcome.

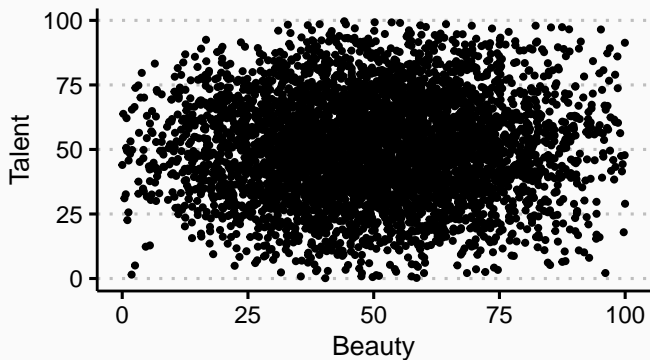
**Solution:** do not condition on C, as path is closed already.

Don't condition on a descendant of a collider, either (Pearl, Glymour, & Jewell, 2016, p. 44-45).

# Truncation in data

Is there a negative relationship between talent and beauty among actors?

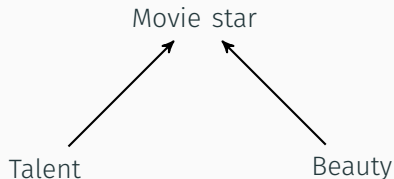
Strong argument for "no"; spurious relationship.



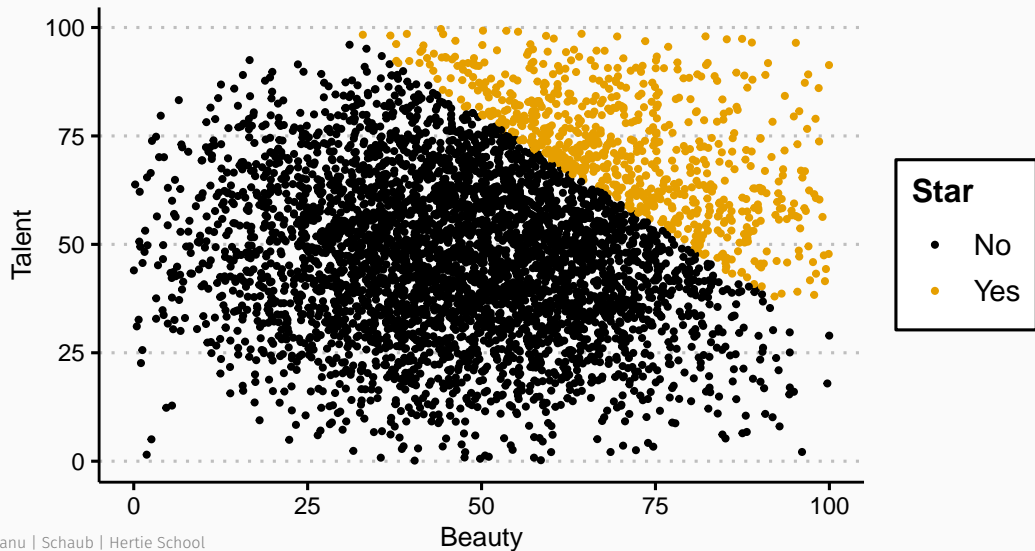
# Selection by cutoff

$$\text{Star potential}_i = \text{Beauty}_i + \text{Talent}_i \quad (2)$$

Top 15% succeed as actors.

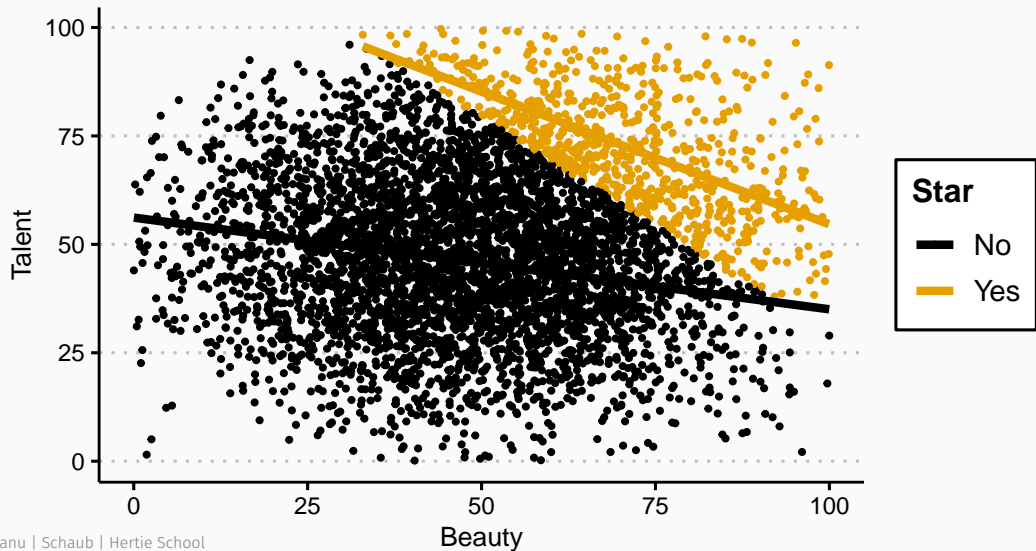


# Effects of conditioning on collider





# Effects of conditioning on collider



# Strategy for $d$ -separation

A few simple steps (Cunningham, 2021, p. 73):

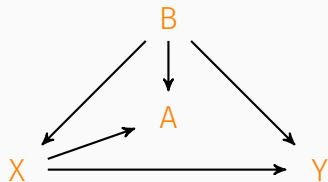
1. write down all paths between  $D$  and  $Y$
2. identify open/closed back-door paths (any confounders or colliders?)
3. find conditioning strategy that closes all open back-doors

Last step is not always possible.

# Practice

---

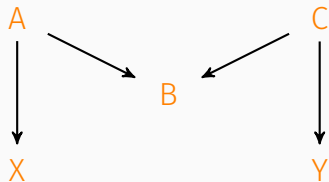
# Practice #1



Estimate  $X \rightarrow Y$

One back-door path:  
 $X \leftarrow B \rightarrow Y$ .

## Practice #2



Estimate  $X \rightarrow Y$

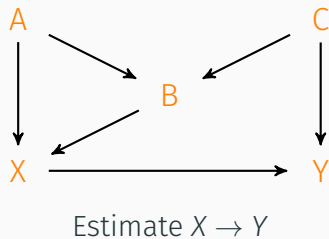
"M-bias."

One back-door path:

$$X \leftarrow A \rightarrow B \leftarrow C \rightarrow Y$$

No need to control for any variable.

## Practice #3



Two back-door paths:

$$X \leftarrow A \rightarrow B \leftarrow C \rightarrow Y$$

$$X \leftarrow B \leftarrow C \rightarrow Y$$

Either control for  $B$  & either  $A$  or  $C$ ,  
or just for  $C$ .

## Racially-biased policing

---

# The setup

**Question:** Can commonly-used administrative data on police stops reveal racially-biased policing?

**Fundamental** problem: Impossible to detect if administrative records suffer from bias.

Race also influences whether police decide to stop a person (only a stop triggers a report being filled in).



# The setup

Studying racial bias in policing: "inherently causal inquiry".

**Counterfactual:** What would be the outcome if an individual of different race would encounter police in same location, time, for same criminal conduct...

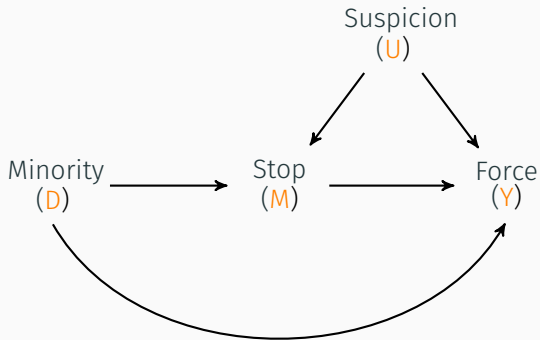
Multiple estimands:  $ATE$  or  $ATE_{M=1}$ .

# The setup

Studying racial bias in policing: "inherently causal inquiry".

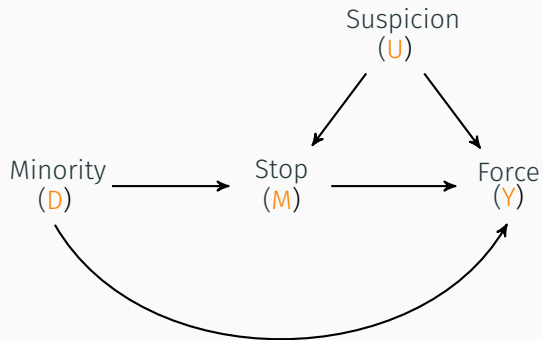
**Counterfactual:** What would be the outcome if an individual of different race would encounter police in same location, time, for same criminal conduct...

Multiple estimands:  $ATE$  or  $ATE_{M=1}$ .



$D \in \{0; 1\}$ , with 1 designating "minority".

Unobservables include: sense of threat, suspicion.



$$ITE_i = Y_i^1(M_i^1) - Y_i^0(M_i^0) \quad (3)$$

# Difficulty with ATE

$$ATE = E[Y_i^1(M_i^1)] - E[Y_i^0(M_i^0)] \quad (4)$$

Quantity depends on:

- ✓ minorities are stopped at differential rates
- ✓ minorities are subjected to violence at differential rates

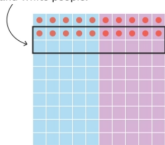
# Partial illustration (Bronner, 2020)

## How numbers that appear equitable can obscure bias

Let's say a police officer is patrolling the street, looking for people with contraband. The officer sees 100 people, some of whom have ● contraband on their person. Say the crowd is evenly split between Black and white people.

### SCENARIO 1

The police officer stops 20 people, pulling aside equal numbers of Black and white people.



Of the 20 people stopped, the officer uses ○ force against 8 of them.

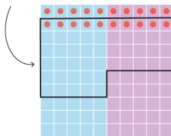


The police officer used force against stopped white people and stopped Black people at the same rate: 40%.

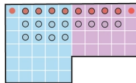
But that's not the only scenario that can lead to that 40% number.

### SCENARIO 2

This time, of the 100 people the officer sees, he stops 50. But this time he is biased in whom he pulls aside.



The officer uses force against 20 people this time.



This time, like last time, the police officer used force against stopped white people and stopped Black people at the same rate: 40%.

### ANALYSIS

Things might appear equal, but in the second scenario, more Black people were stopped by the police than white people.

While use of force among **stopped** people is equal, use of force among all **observed** people is not:

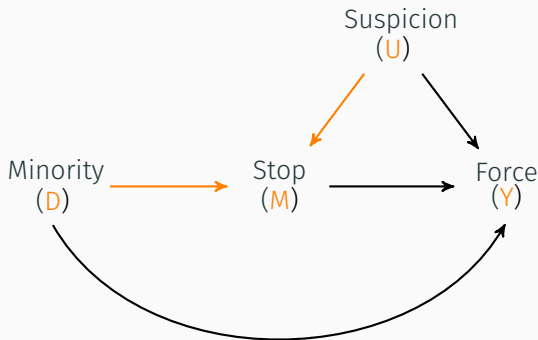
$$\frac{12}{50} = 24\% \text{ of Black people have force used against them}$$

$$\frac{8}{50} = 16\% \text{ of white people have force used against them}$$

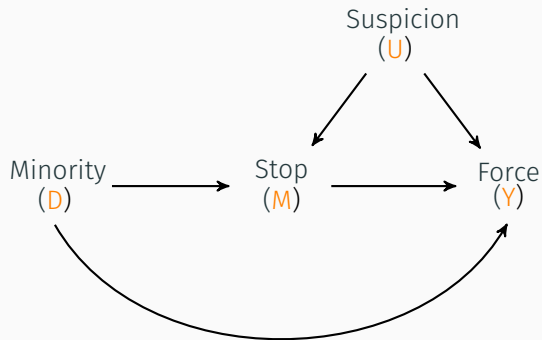
### CONCLUSION

This is why knowing how often police use force against people they've stopped is **not enough information** to know whether use of force is racially biased. In real life, we don't have data on everyone who was observed but not stopped, but we need that to know whether use of force is biased overall.

# What is this about?



**Collider bias:** police records condition on a collider, by recording only stops that are impacted themselves by race.

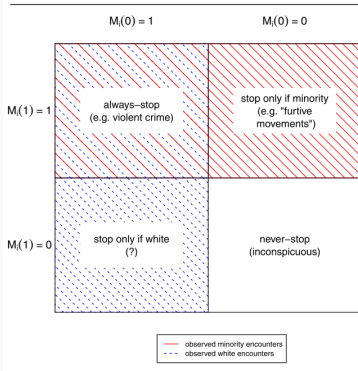


$$Naive = \hat{\delta} = E[Y_i^1 | D_i = 1, M_i = 1] - E[Y_i^0 | D_i = 0, M_i = 1] \quad (5)$$



# Strata of contexts

**FIGURE 2. Principal Strata and Observed Police-Civilian Encounters**



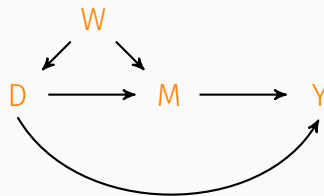
*Notes:* The figure displays the four principal strata that comprise police–civilian encounters based on how the mediator  $M$  (whether a civilian is stopped by police) responds to treatment  $D$  (whether the civilian is a racial minority). Minorities in the “always stop” and anti-minority racial stop strata, highlighted in red, are stopped by police and, thus, appear in police administrative data. Likewise, white civilians in the “always-stop” and anti-white racial stop strata, highlighted in blue, appear in police data. “Never stop” encounters are unobserved. Because white and nonwhite encounters are drawn from different principal strata, the two groups are incomparable and estimates of causal quantities using observed encounters will be statistically biased absent additional assumptions.

# Assumptions needed for $NATE = ATE$

$$M_i(d) \perp\!\!\!\perp D_i | X_i \quad (6)$$

After controlling for contextual factors (neighborhoods, police practices), race is independent of encounters.

Can't be tested without data from "non-stops".



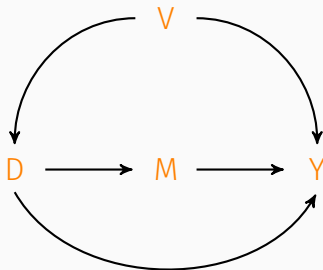
Assumption violation

# Assumptions needed for $NATE = ATE$

$$Y_i(d, m) \perp\!\!\!\perp D_i | M_{0i} = m', M_{1i} = m'', X_i \quad (7)$$

A contextual factor shaped by race, and which influences violence, needs to be controlled for.

The 2 assumptions: "treatment ignorability".



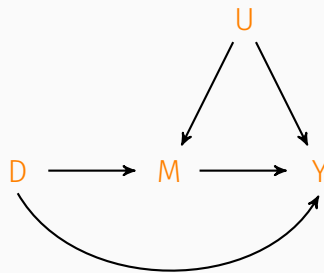
Assumption violation

# Assumptions needed for $NATE = ATE$

$$Y_i(d, m) \perp\!\!\!\perp M_{0i} | D_i = d, M_{1i} = 1, X_i \quad (8)$$

Violence rates in "always-stop" encounters similar to those in "racial stops".

Many factors unrecorded in reports make this a strong assumption ("mediator ignorability")



Assumption violation

Thank **you** for the kind attention!

# References

- Bronner, L. (2020, June). *Why Statistics Don't Capture The Full Extent Of The Systemic Bias In Policing*. Retrieved from <https://fivethirtyeight.com/features/why-statistics-dont-capture-the-full-extent-of-the-systemic-bias-in-policing/>
- Cunningham, S. (2021). *Causal Inference: The Mixtape*. New Haven, CT: Yale University Press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. Chichester, UK: Wiley.