DAGs and potential outcomes

Session 5

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

do()ing observational causal inference

Potential outcomes

do()ing observational causal inference

Structural models

$$\mathrm{Loc} = f_{\mathrm{Loc}}(\mathrm{U1})$$
$$\mathrm{Bkgd} = f_{\mathrm{Bkgd}}(\mathrm{U1})$$
$$\mathrm{JobCx} = f_{\mathrm{JobCx}}(\mathrm{Edu})$$
$$\mathrm{Edu} = f_{\mathrm{Edu}}(\mathrm{Req}, \mathrm{Loc}, \mathrm{Year})$$
$$\mathrm{Earn} = f_{\mathrm{Earn}}(\mathrm{Edu}, \mathrm{Year}, \mathrm{Bkgd},$$
$$\mathrm{Loc}, \mathrm{JobCx})$$

Structural models

dagify() in ggdag forces you to think this way

$$\text{Earn} = f_{\text{Earn}}(\text{Edu}, \text{Year}, \text{Bkgd},$$
$$\text{Loc}, \text{JobCx})$$
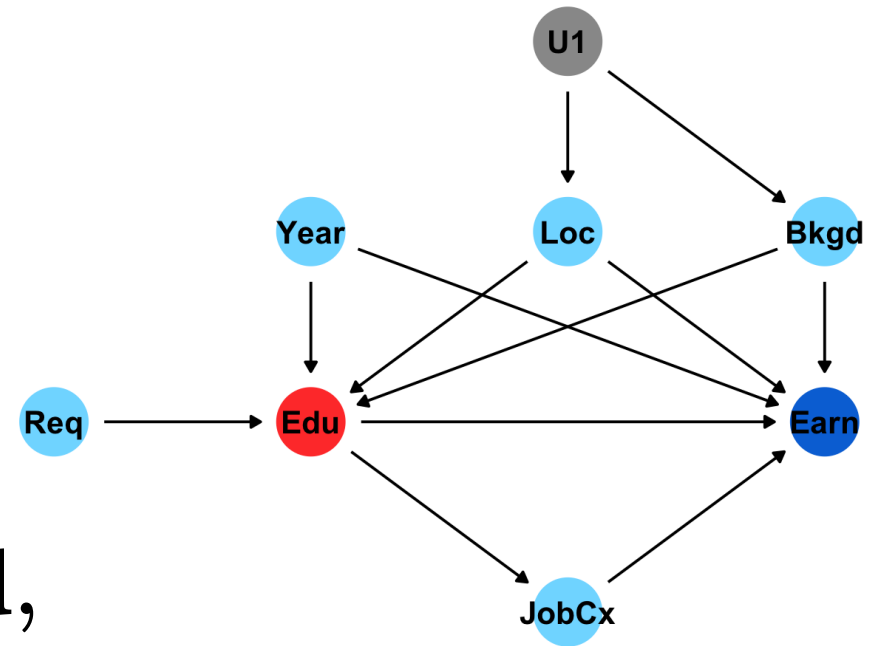$$\text{Edu} = f_{\text{Edu}}(\text{Req}, \text{Loc}, \text{Year})$$
$$\text{JobCx} = f_{\text{JobCx}}(\text{Edu})$$
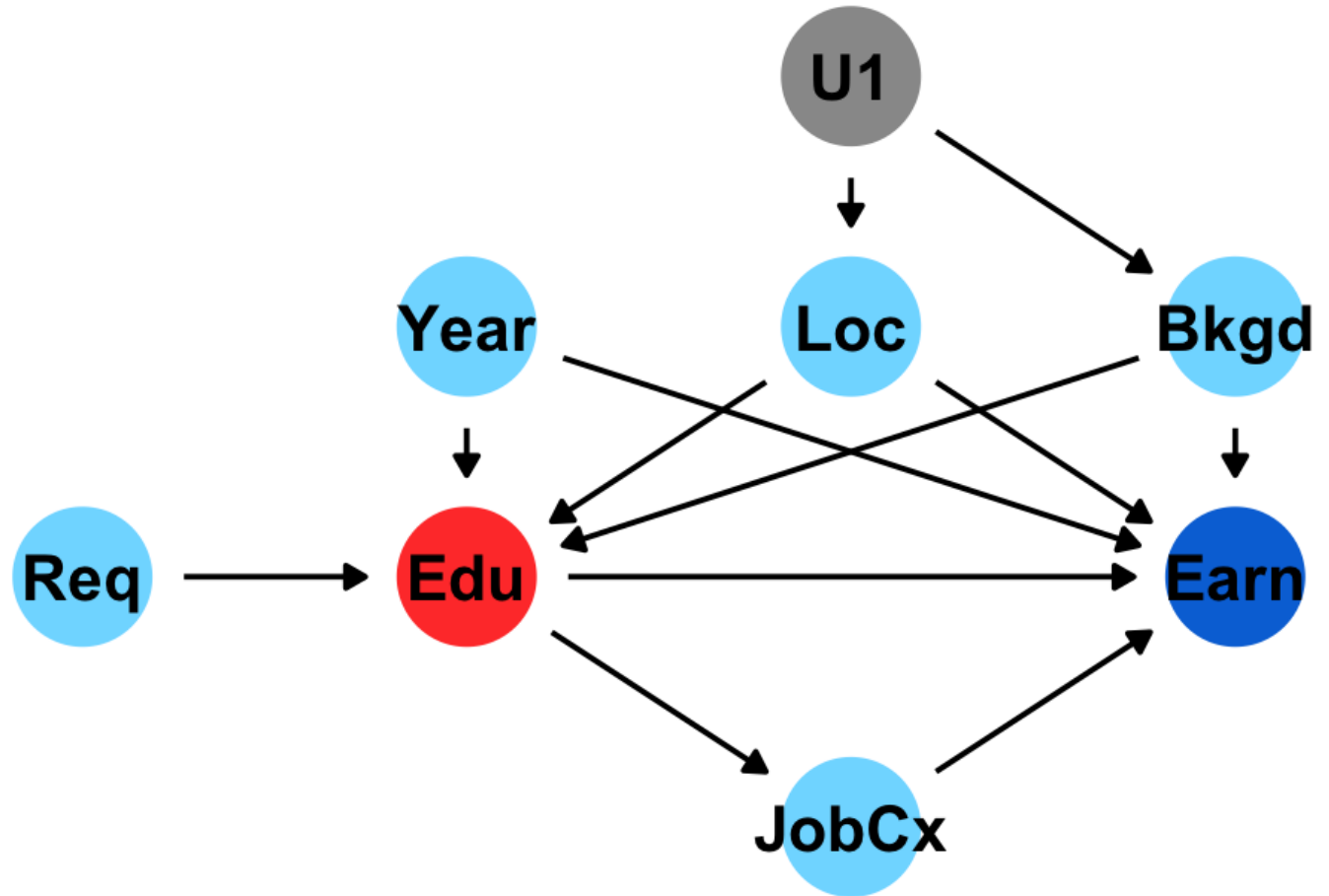$$\text{Bkgd} = f_{\text{Bkgd}}(\text{U1})$$
$$\text{Loc} = f_{\text{Loc}}(\text{U1})$$

```
dagify(
  Earn ~ Edu + Year + Bkgd + Loc + JobCx,
  Edu ~ Req + Loc + Bkgd + Year,
  JobCx ~ Edu,
  Bkgd ~ U1,
  Loc ~ U1
)
```

All these nodes are related; there's correlation between them all

We care about **Edu** → **Earn**, but what do we do about all the other nodes?

Causal identification

A causal effect is *identified* if the association between treatment and outcome is properly stripped and isolated
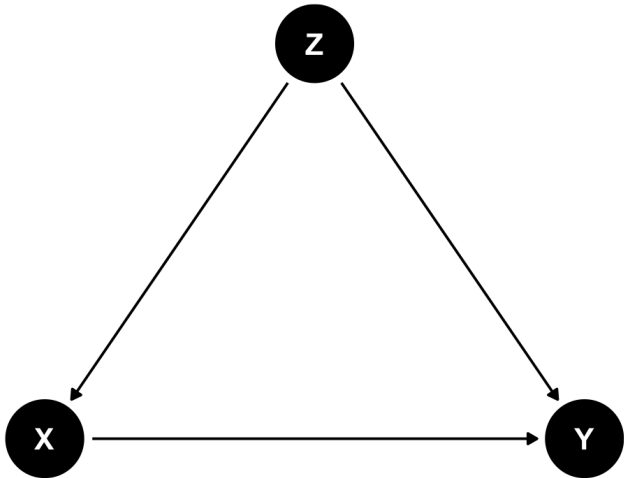
Paths and associations

Arrows in a DAG transmit associations

You can redirect and control those paths by "adjusting" or "conditioning"

Three types of associations

Confounding

Common cause

Causation
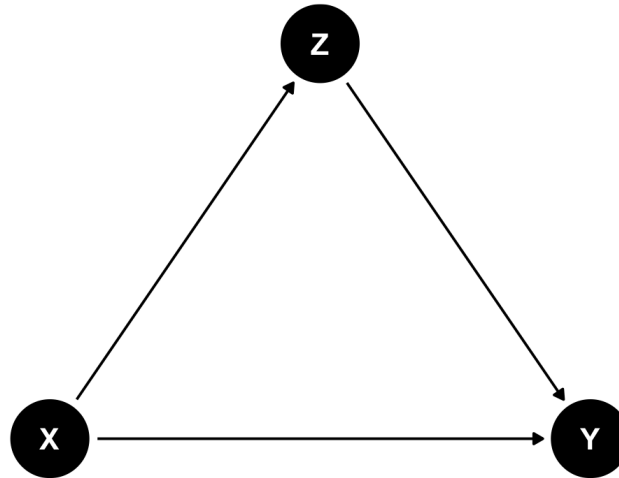
Mediation

Collision
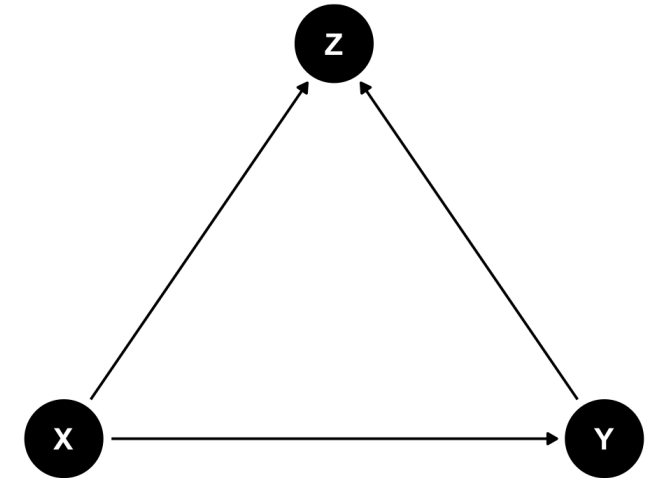
Selection / endogeneity

Interventions

do-operator

Making an intervention in a DAG

$$P[Y \mid do(X = x)] \quad \text{or} \quad E[Y \mid do(X = x)]$$

P = probability distribution, or E = expectation/expected value

**Y = outcome, X = treatment;
x = specific value of treatment**

Interventions

$$E[Y \mid do(X = x)]$$

E[Earnings | *do*(One year of college)]

E[Firm growth | *do*(Government R&D funding)]

E[Air quality | *do*(Carbon tax)]

E[Juvenile delinquency | *do*(Truancy program)]

E[Malaria infection rate | *do*(Mosquito net)]

Interventions

When you *do*() X, delete all arrows into it

Observational DAG

A

X → Y

A → X

A → Y

Experimental DAG

A

X → Y

A → Y
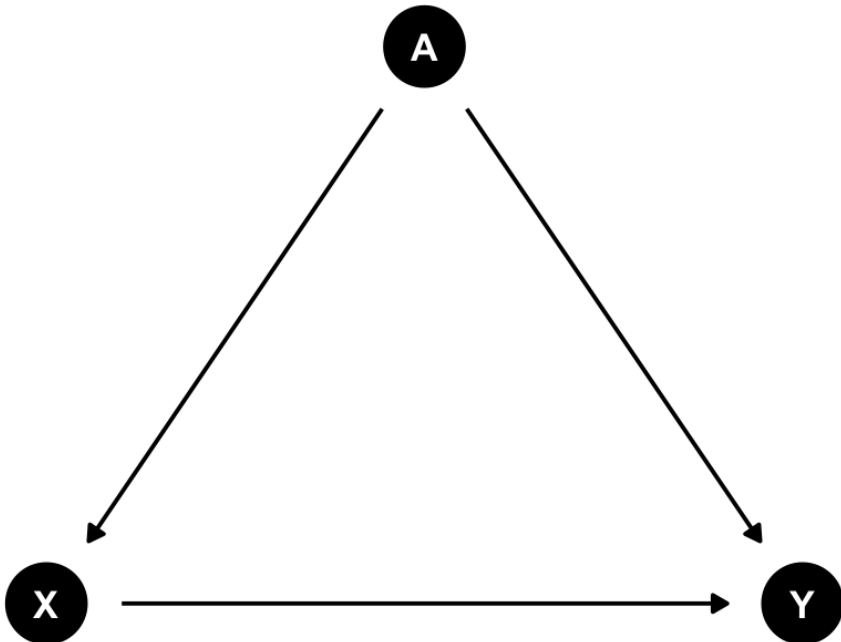
Interventions

$$E[\text{Earnings} \mid do(\text{College education})]$$

Observational DAG
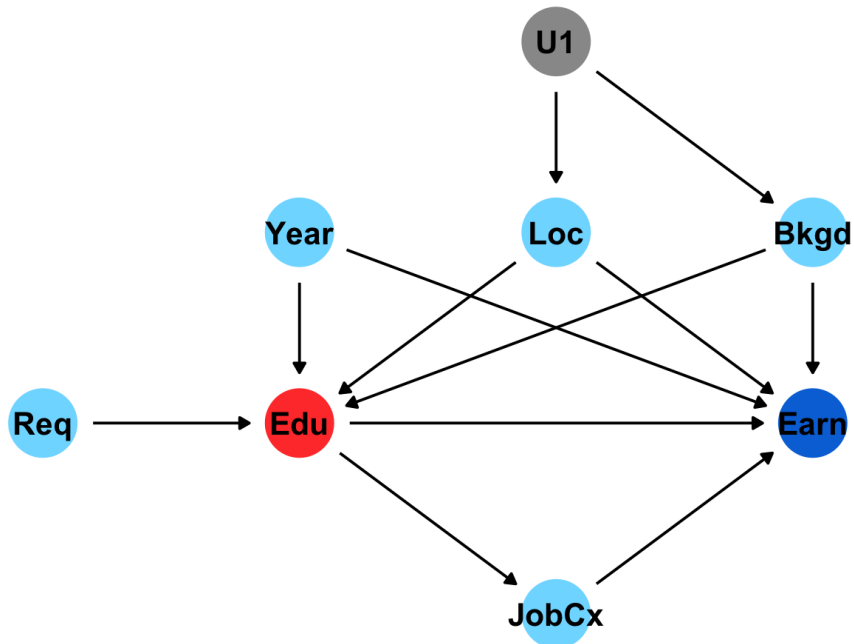
Experimental DAG
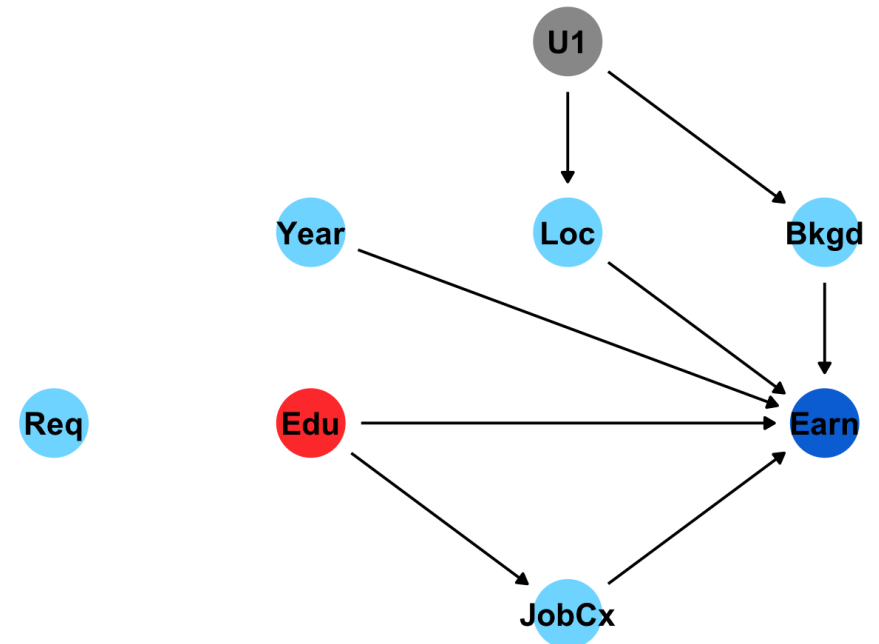
Un*do*()ing things

We want to know **P[Y | *do*(X)]** but all we have is observational data X, Y, and Z

$$P[Y \mid do(X)] \neq P(Y \mid X)$$

Correlation isn't causation!

Un*do*()ing things

Our goal with observational data:
Rewrite P[Y | *do*(X)] so that it doesn't have a
***do*() anymore (is "*do*-free")**

do-calculus

A set of three rules that let you manipulate a DAG in special ways to remove *do*() expressions

The do-calculus Let G be a CGM, $G_{\overline{T}}$ represent G post-intervention (i.e with all links into T removed) and $G_{\underline{T}}$ represent G with all links *out of* T removed. Let $do(t)$ represent intervening to set a single variable T to t,

Rule 1: $\mathbb{P}(y|do(t), z, w) = \mathbb{P}(y|do(t), z)$ if $Y \perp\!\!\!\perp W|(Z, T)$ in $G_{\overline{T}}$

Rule 2: $\mathbb{P}(y|do(t), z) = \mathbb{P}(y|t, z)$ if $Y \perp\!\!\!\perp T|Z$ in $G_{\underline{T}}$

Rule 3: $\mathbb{P}(y|do(t), z) = \mathbb{P}(y|z)$ if $Y \perp\!\!\!\perp T|Z$ in $G_{\overline{T}}$, and Z is not a decedent of T.

WAAAAAY beyond the score of this class!
Just know it exists and computer algorithms can do it for you!

Special cases of *do*-calculus

Backdoor adjustment

Frontdoor adjustment

Backdoor adjustment

$$P[Y \mid do(X)] = \sum_Z P(Y \mid X, Z) \times P(Z)$$
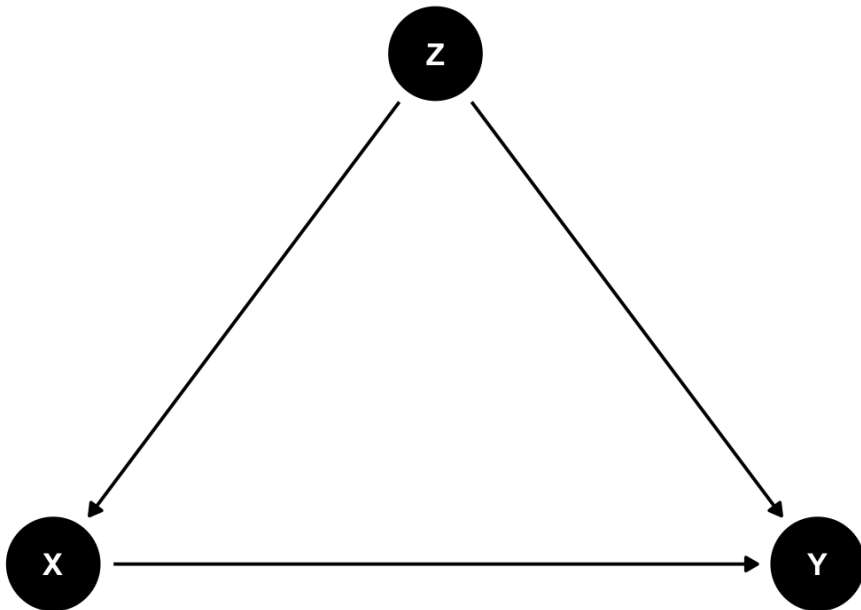
↑ **That's complicated!**

The right-hand side of the equation means "the effect of X on Y after adjusting for Z"

There's no *do()* on that side!

Frontdoor adjustment
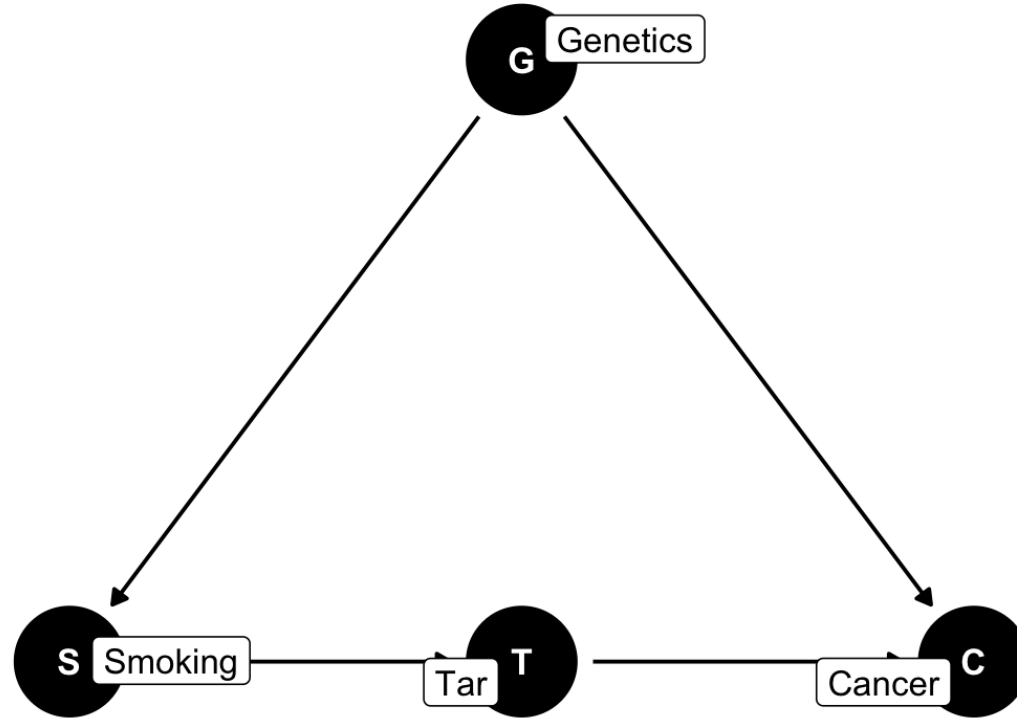
**S → T is *d*-separated; T → C is *d*-separated
combine the effects to find S → C**

Moral of the story

If you can transform *do*() expressions to *do*-free versions, you can legally make causal inferences from observational data

Backdoor adjustment is easiest to see + dagitty and **ggdag** do this for you!

Fancy algorithms (found in the **causaleffect** package) can do the official *do*-calculus for you too

Potential outcomes

Program effect

Y — Outcome variable

Post-program **outcome level**

Outcome with program

Program effect

δ

Outcome
without program

Outcome change

Pre-program
outcome level

X

Before program **During program** **After program**
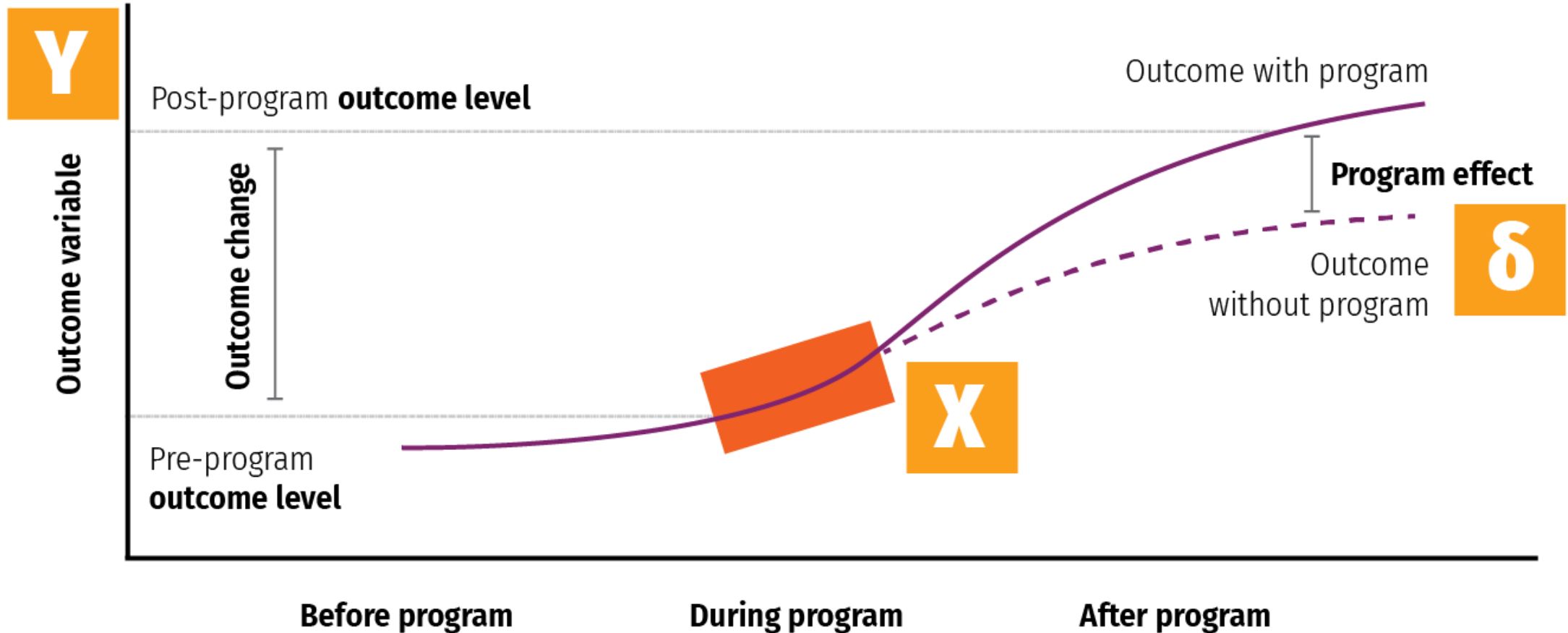
Some equation translations

Causal effect = δ (delta)

$$\delta = P[Y \mid do(X)]$$

$$\delta = E[Y \mid do(X)] - E[Y \mid \hat{do}(X)]$$

$$\delta = (Y \mid X = 1) - (Y \mid X = 0)$$

$$\delta = Y_1 - Y_0$$

Fundamental problem of causal inference

$$\delta_i = Y_i^1 - Y_i^0 \quad \text{in real life is} \quad \delta_i = Y_i^1 - ???$$

Individual-level effects are impossible to observe!

There are no individual counterfactuals!

Average treatment effect (ATE)

Solution: Use averages instead

$$\text{ATE} = E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

Difference between average/expected value when program is on vs. expected value when program is off

$$\delta = (\bar{Y} \mid P = 1) - (\bar{Y} \mid P = 0)$$

Person	Age	Treated	Outcome with program	Outcome without program	Effect
1	Old	TRUE	**80**	60	**20**
2	Old	TRUE	**75**	70	**5**
3	Old	TRUE	**85**	80	**5**
4	Old	FALSE	70	**60**	**10**
5	Young	TRUE	**75**	70	**5**
6	Young	FALSE	80	**80**	**0**
7	Young	FALSE	90	**100**	**-10**
8	Young	FALSE	85	**80**	**5**

Person	Age	Treated	Outcome with program	Outcome without program	Effect
1	Old	TRUE	**80**	60	**20**
2	Old	TRUE	**75**	70	**5**
3	Old	TRUE	**85**	80	**5**
4	Old	FALSE	70	**60**	**10**
5	Young	TRUE	**75**	70	**5**
6	Young	FALSE	80	**80**	**0**
7	Young	FALSE	90	**100**	**-10**
8	Young	FALSE	85	**80**	**5**

$$\delta = (\bar{Y} \mid P = 1) - (\bar{Y} \mid P = 0) \qquad \text{ATE} = \frac{20+5+5+5+10+0+-10+5}{8} = 5$$

CATE

ATE in subgroups

Is the program more effective for specific age groups?

Person	Age	Treated	Outcome with program	Outcome without program	Effect
1	Old	TRUE	**80**	60	**20**
2	Old	TRUE	**75**	70	**5**
3	Old	TRUE	**85**	80	**5**
4	Old	FALSE	70	**60**	**10**
5	Young	TRUE	**75**	70	**5**
6	Young	FALSE	80	**80**	**0**
7	Young	FALSE	90	**100**	**-10**
8	Young	FALSE	85	**80**	**5**

$$\delta = (\bar{Y}_{\mathrm{O}} \mid P = 1) - (\bar{Y}_{\mathrm{O}} \mid P = 0)$$

$$\mathrm{CATE}_{\mathrm{Old}} = \frac{20 + 5 + 5 + 10}{4} = 10$$

$$\delta = (\bar{Y}_{\mathrm{Y}} \mid P = 1) - (\bar{Y}_{\mathrm{Y}} \mid P = 0)$$

$$\mathrm{CATE}_{\mathrm{Young}} = \frac{5 + 0 - 10 + 5}{4} = 0$$

ATT and ATU

Average treatment on the treated

ATT / TOT

Effect for those with treatment

Average treatment on the untreated

ATU / TUT

Effect for those without treatment

Person	Age	Treated	Outcome with program	Outcome without program	Effect
1	Old	TRUE	**80**	60	**20**
2	Old	TRUE	**75**	70	**5**
3	Old	TRUE	**85**	80	**5**
4	Old	FALSE	70	**60**	**10**
5	Young	TRUE	**75**	70	**5**
6	Young	FALSE	80	**80**	**0**
7	Young	FALSE	90	**100**	**-10**
8	Young	FALSE	85	**80**	**5**

$$\delta = (\bar{Y}_{\text{T}} \mid P = 1) - (\bar{Y}_{\text{T}} \mid P = 0) \qquad \text{CATE}_{\text{Treated}} = \frac{20+5+5+5}{4} = 8.75$$

$$\delta = (\bar{Y}_{\text{U}} \mid P = 1) - (\bar{Y}_{\text{U}} \mid P = 0) \qquad \text{CATE}_{\text{Untreated}} = \frac{10+0-10+5}{4} = 1.25$$

The ATE is the weighted average of the ATT and ATU

$$\mathrm{ATE} = (\pi_{\mathrm{Treated}} \times \mathrm{ATT}) + (\pi_{\mathrm{Untreated}} \times \mathrm{ATU})$$

$$\left(\frac{4}{8} \times 8.75\right) + \left(\frac{4}{8} \times 1.25\right)$$

$$4.375 + 0.625 = 5$$

π here means "proportion," not 3.1415

Selection bias

ATE and ATT aren't always the same

ATE = ATT + Selection bias

$$5 = 8.75 + x$$
$$x = -3.75$$

Randomization fixes this, makes x = 0

Actual data

Person	Age	Treated	Actual outcome
1	Old	TRUE	80
2	Old	TRUE	75
3	Old	TRUE	85
4	Old	FALSE	60
5	Young	TRUE	75
6	Young	FALSE	80
7	Young	FALSE	100
8	Young	FALSE	80

Treatment not
randomly assigned

We can't see
unit-level causal effects

What do we do?!

Actual data

Person	Age	Treated	Actual outcome
1	Old	TRUE	80
2	Old	TRUE	75
3	Old	TRUE	85
4	Old	FALSE	60
5	Young	TRUE	75
6	Young	FALSE	80
7	Young	FALSE	100
8	Young	FALSE	80

Treatment seems to be correlated with age
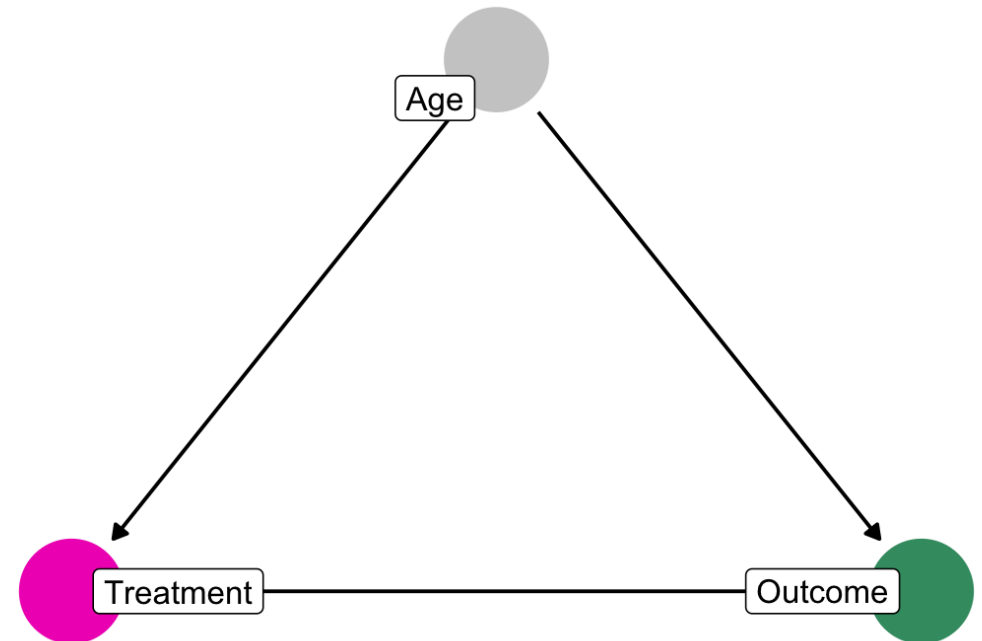
Actual data

Person	Age	Treated	Actual outcome
1	Old	TRUE	80
2	Old	TRUE	75
3	Old	TRUE	85
4	Old	FALSE	60
5	Young	TRUE	75
6	Young	FALSE	80
7	Young	FALSE	100
8	Young	FALSE	80

We can estimate the ATE by finding the weighted average of age-based CATEs

As long as we assume/pretend treatment was randomly assigned within each age = unconfoundedness

$$\widehat{\text{ATE}} = \pi_{\text{Old}} \widehat{\text{CATE}}_{\text{Old}} + \pi_{\text{Young}} \widehat{\text{CATE}}_{\text{Young}}$$

Actual data

$$\widehat{\text{ATE}} = \pi_{\text{Old}}\widehat{\text{CATE}}_{\text{Old}} + \pi_{\text{Young}}\widehat{\text{CATE}}_{\text{Young}}$$

Person	Age	Treated	Actual outcome
1	Old	TRUE	80
2	Old	TRUE	75
3	Old	TRUE	85
4	Old	FALSE	60
5	Young	TRUE	75
6	Young	FALSE	80
7	Young	FALSE	100
8	Young	FALSE	80

$$\widehat{\text{CATE}}_{\text{Old}} = \frac{80+75+85}{3} - \frac{60}{1} = 20$$

$$\widehat{\text{CATE}}_{\text{Young}} = \frac{75}{1} - \frac{80+100+80}{3} = -11.667$$

$$\widehat{\text{ATE}} = \left(\frac{4}{8} \times 20\right) + \left(\frac{4}{8} \times -11.667\right) = 4.1667$$

¡¡¡DON'T DO THIS!!!

$$\widehat{\text{ATE}} = \widehat{\text{CATE}}_{\text{Treated}} - \widehat{\text{CATE}}_{\text{Untreated}}$$

Person	Age	Treated	Actual outcome
1	Old	TRUE	80
2	Old	TRUE	75
3	Old	TRUE	85
4	Old	FALSE	60
5	Young	TRUE	75
6	Young	FALSE	80
7	Young	FALSE	100
8	Young	FALSE	80

$$\widehat{\text{CATE}}_{\text{Treated}} = \frac{80+75+85+75}{4} = 78.75$$

$$\widehat{\text{CATE}}_{\text{Untreated}} = \frac{60+80+100+80}{4} = 80$$

$$\widehat{\text{ATE}} = 78.75 - 80 = -1.25$$

You can only do this if treatment is random!

$$\widehat{\text{ATE}} = \pi_{\text{Old}}\widehat{\text{CATE}}_{\text{Old}} + \pi_{\text{Young}}\widehat{\text{CATE}}_{\text{Young}}$$

We used age here because it correlates with (and confounds) the outcome

And we assumed unconfoundedness; that treatment is randomly assigned within the groups
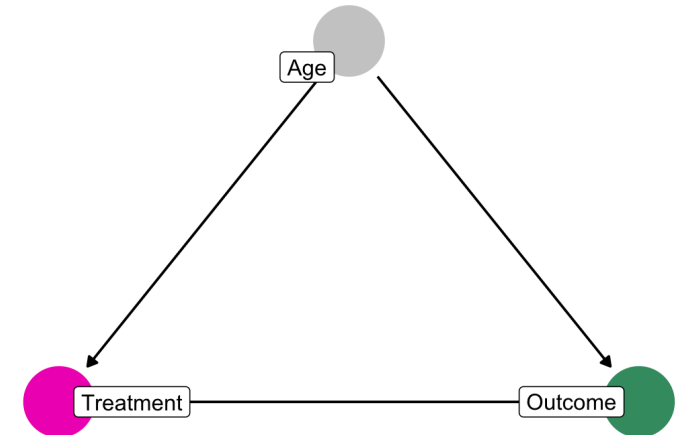
Does attending a private university cause an increase in earnings?

TABLE 2.1
The college matching matrix

| Applicant group | Student | Private | | | Public | | | 1996 earnings |
		Ivy	Leafy	Smart	All State	Tall State	Altered State	
A	1		Reject	Admit		Admit		110,000
	2		Reject	Admit		Admit		100,000
	3		Reject	Admit		Admit		110,000
B	4	Admit			Admit		Admit	60,000
	5	Admit			Admit		Admit	30,000
C	6		Admit					115,000
	7		Admit					75,000
D	8	Reject			Admit	Admit		90,000
	9	Reject			Admit	Admit		60,000

Note: Enrollment decisions are highlighted in gray.

**Average private –
Average public**

TABLE 2.1
The college matching matrix

Applicant group	Student	Private			Public			1996 earnings
		Ivy	Leafy	Smart	All State	Tall State	Altered State	
A	1		Reject	Admit		Admit		110,000
	2		Reject	Admit		Admit		100,000
	3		Reject	Admit			Admit	110,000
B	4	Admit			Admit		Admit	60,000
	5	Admit			Admit		Admit	30,000
C	6		Admit					115,000
	7		Admit					75,000
D	8	Reject			Admit	Admit		90,000
	9	Reject			Admit	Admit		60,000

Note: Enrollment decisions are highlighted in gray.

$$\frac{110 + 100 + 60 + 115 + 75}{5} = 92$$

$$\frac{110 + 30 + 90 + 60}{4} = 72.5$$

$$\left(92 \times \frac{5}{9}\right) - \left(72.5 \times \frac{4}{9}\right) = 18,888$$

This is wrong!

$$\widehat{\text{ATE}} = \pi_{\text{Private}} \widehat{\text{CATE}}_{\text{Private}} + \pi_{\text{Public}} \widehat{\text{CATE}}_{\text{Public}}$$

Table 2.1
The college matching matrix

| Applicant group | Student | Private | | | Public | | | 1996 earnings |
		Ivy	Leafy	Smart	All State	Tall State	Altered State	
A	1		Reject	Admit		Admit		110,000
	2		Reject	Admit		Admit		100,000
	3		Reject	Admit			Admit	110,000
B	4	Admit			Admit		Admit	60,000
	5	Admit			Admit		Admit	30,000
C	6		Admit					115,000
	7		Admit					75,000
D	8	Reject			Admit	Admit		90,000
	9	Reject			Admit	Admit		60,000

Note: Enrollment decisions are highlighted in gray.

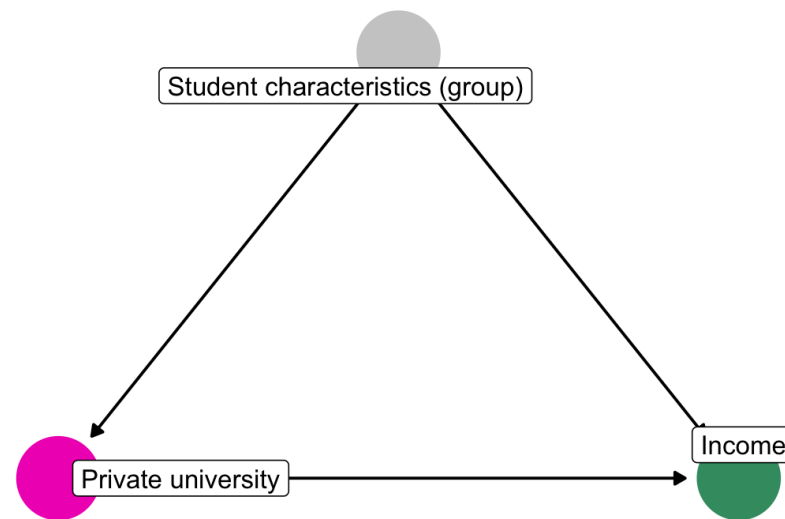**These groups look like they have similar characteristics**

Unconfoundedness?

43 / 45

TABLE 2.1
The college matching matrix

| Applicant group | Student | Private | | | Public | | | 1996 earnings |
		Ivy	Leafy	Smart	All State	Tall State	Altered State	
A	1		Reject	Admit		Admit		110,000
	2		Reject	Admit		Admit		100,000
	3		Reject	Admit			Admit	110,000
B	4	Admit			Admit		Admit	60,000
	5	Admit			Admit		Admit	30,000
C	6		Admit					115,000
	7		Admit					75,000
D	8	Reject			Admit	Admit		90,000
	9	Reject			Admit	Admit		60,000

Note: Enrollment decisions are highlighted in gray.

CATE Group A + CATE Group B

$$\frac{110 + 100}{2} - 110 = -5,000$$

$$60 - 30 = 30,000$$

$$(-5 \times \frac{3}{5}) + (30 \times \frac{2}{5}) = 9,000$$

This is less wrong!

$$\widehat{\text{ATE}} = \pi_{\text{Group A}} \widehat{\text{CATE}}_{\text{Group A}} + \pi_{\text{Group B}} \widehat{\text{CATE}}_{\text{Group B}}$$

Matching with regression

$$\text{Earnings} = \alpha + \beta_1 \text{Private} + \beta_2 \text{Group} + \epsilon$$

```
model_earnings <- lm(earnings ~ private + group_A, data = schools_small)
```

term	estimate	std.error	statistic	p.value
(Intercept)	40000	11952.29	3.35	0.08
privateTRUE	10000	13093.07	0.76	0.52
group_ATRUE	60000	13093.07	4.58	0.04

β₁ = $10,000 **This is less wrong!** **Significance details!**