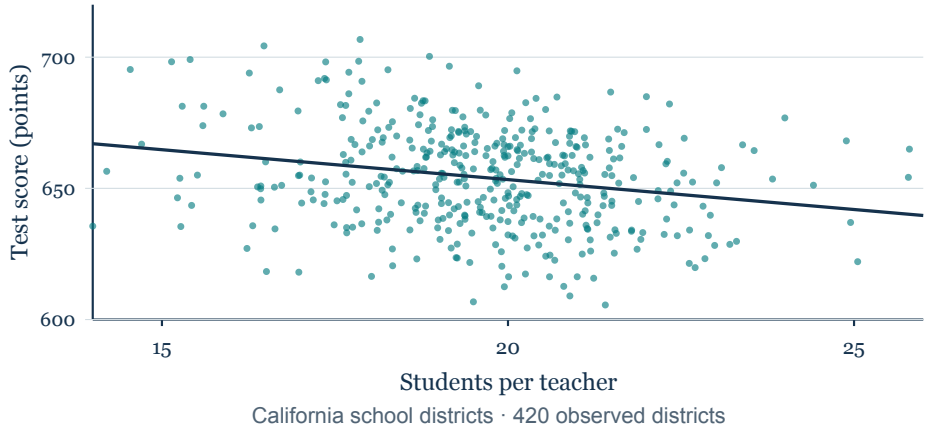
Do smaller classes improve learning?

Compare districts. Then ask what else differs.



California school districts · 420 observed districts

Practice

Do you expect the line to slope upward or downward? Name one other district characteristic that could help explain the pattern.

Explanation: A downward slope is plausible: districts with more students per teacher may have lower scores. Family resources could also matter. For example, districts serving richer families may have both more teachers and higher scores. The graph lets us check the observed pattern; it does not yet isolate the effect of hiring teachers.

Which question are we answering?

The same data can describe a pattern, predict an outcome or help study a policy effect.

> Description: how do observed districts differ?

> Prediction: how well can we forecast a new outcome?

> Causation: what changes when the policy changes?

Define the observation before the equation

One observation is one district; define the variables and their units.

> i: school district

> Y_i: average score, in points

> X_i: students per teacher

> $n = 420$ districts

Practice

Would ten years of the same 420 districts form a panel? Would ten different districts observed each year necessarily do so?

Explanation: The first is a panel if district identities are tracked. The second may be repeated cross sections; the calendar alone does not create within-unit observations.

Different data structures support different comparisons

Identify the unit and whether the same units return over time.

> Cross section: different units at a date; time series: observations ordered by date.

> Pooled cross sections: samples at several dates; panel: identified units followed over time.

> Data structure and sampling design are separate questions.

Random sampling is a claim about how observations are drawn

Independence and a common distribution need a substantive argument.

> Random sample: (Y_i, X_i), $i = 1, \ldots, n$, are independent draws from one joint distribution.

> iid: independent and identically distributed.

> Ask how units were selected and whether they share shocks.

Start with a rule for how the outcome is produced

A causal model states mechanisms before asking what would change.

> Staffing: $X = 22 - C + V$

> Learning: $Y = 700 - 2X + 5C + U$

> $C = 2, V = 2, U = 3 \Rightarrow X = 22, Y = 669$

Define the intervention before its counterfactual

Change the staffing rule; retain the stated learning mechanism.

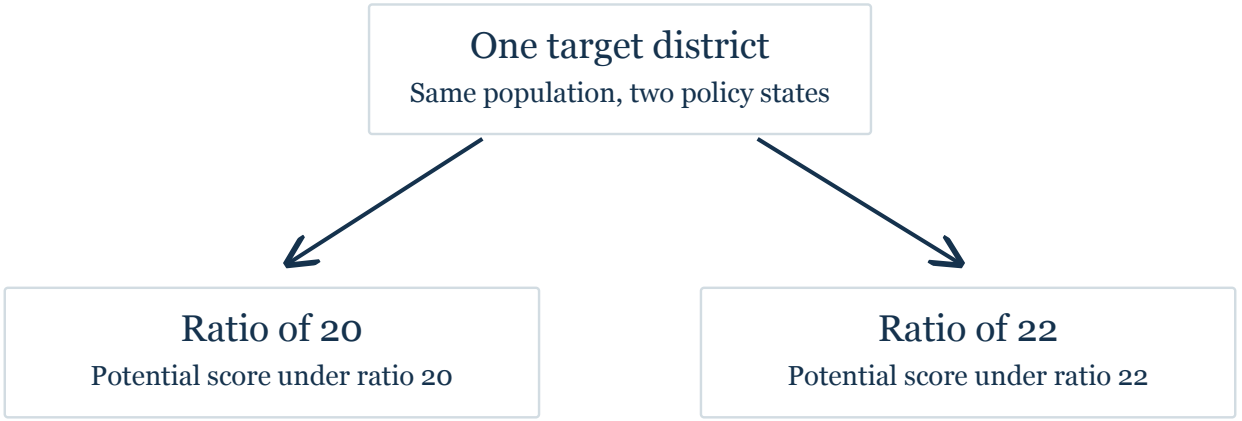
> Intervene: replace the assignment rule by $X = 20$

> Alternative score: $700 - 2(20) + 5(2) + 3 = 673$

> Same-district difference: $673 - 669 = 4$ points

The comparison we want is partly missing

A policy effect compares potential outcomes under two staffing choices.



- Potential outcomes: $Y_i(20), Y_i(22)$
- Causal contrast: $Y_i(20) - Y_i(22)$
- Observed groups do not automatically supply the missing outcome.

Average effects require a population and an identification argument

Define the policy states, population and averaging rule.

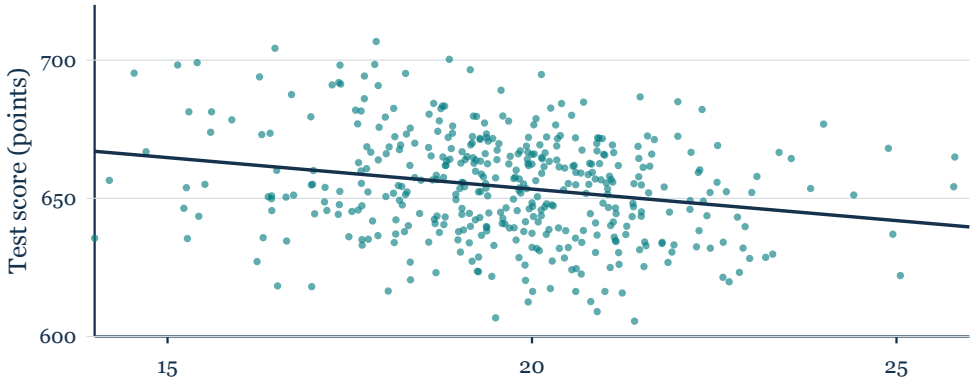
> $ATE = E[Y(1) - Y(0)]$; $ATT = E[Y(1) - Y(0) \mid D = 1]$.

> Exchangeability: $(Y(1), Y(0)) \perp D \mid C$; overlap: $0 < P(D = 1 \mid C) < 1$.

> Under the stated conditions, $ATE = E_C\{E[Y \mid D = 1, C] - E[Y \mid D = 0, C]\}$.

Read the data before fitting a line

Start with axes, units, range and unusual observations.



California school districts · 420 observed districts

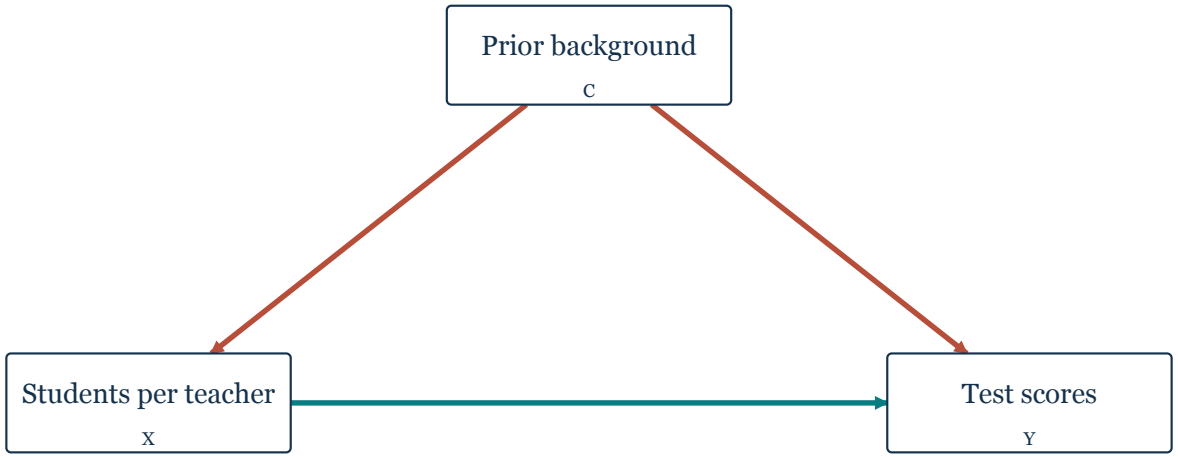
Practice

Districts at about the same student–teacher ratio have different scores. Does this prevent a line from describing their average pattern?

Explanation: No. A line can summarise the average pattern even when individual districts lie above and below it. Those differences tell us that staffing alone does not fully explain scores.

Use a DAG to make the causal story visible

Begin with assumptions about the arrows and mechanisms.



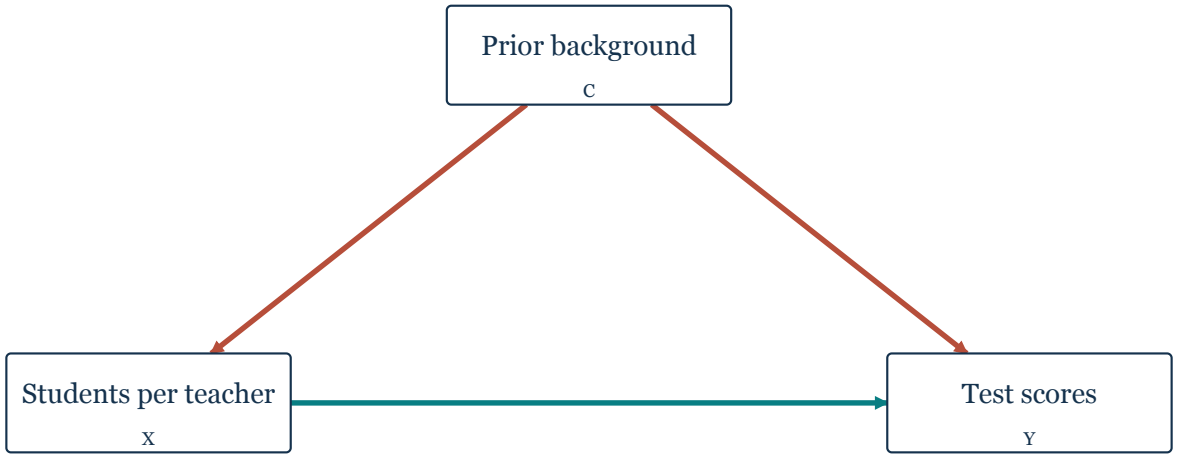
1 open directed route · 1 open alternative · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

- Node: a variable; arrow: a proposed direct causal mechanism.
- Directed: arrows have a direction. Acyclic: no directed loop.
- Missing arrows and independent background inputs are assumptions.

Trace a path one connection at a time

A causal route follows arrows; a general path need not.



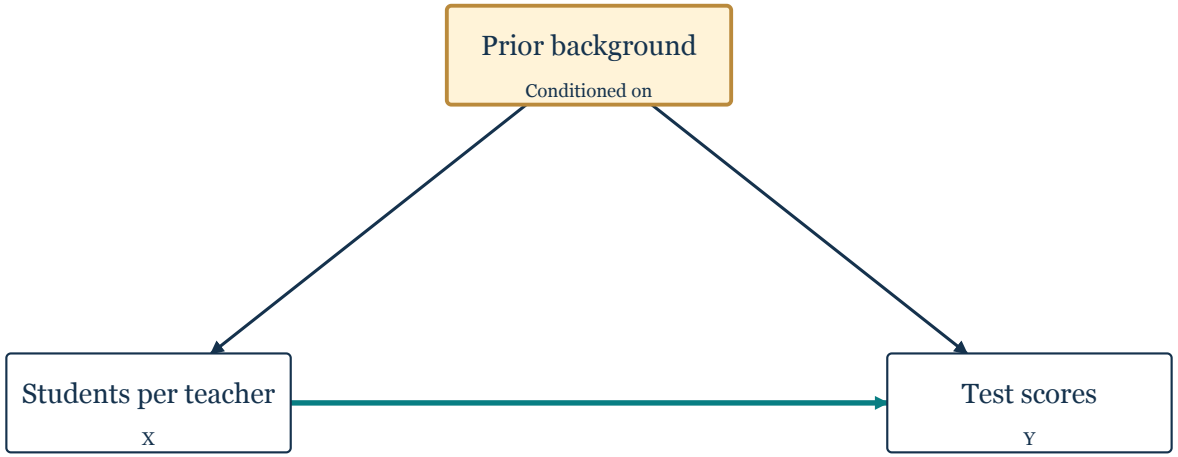
1 open directed route · 1 open alternative · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

- Causal route: $X \to Y$
- Backdoor route: $X \leftarrow C \to Y$

A common cause creates a backdoor comparison

Compare districts with similar prior background.



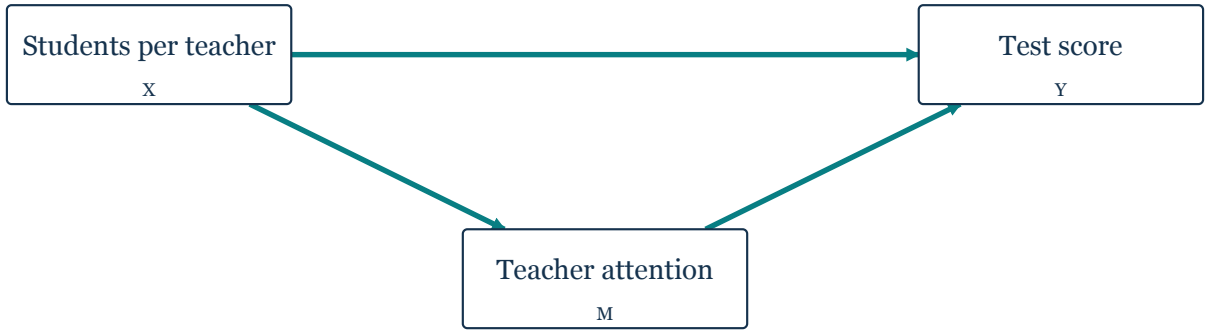
1 open directed route · 0 open alternatives · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

- Compare at fixed C: block $X \leftarrow C \rightarrow Y$
- Conditioning selects comparisons; intervention changes a mechanism.

An intermediate mechanism is a mediator

A variable caused by staffing can carry part of its effect.



2 open directed routes · 0 open alternatives · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

- Mediated route: $X \to M \to Y$
- Total effect allows the mediator to respond.

Selection can create a relationship

A four-case scholarship example makes the mechanism visible.

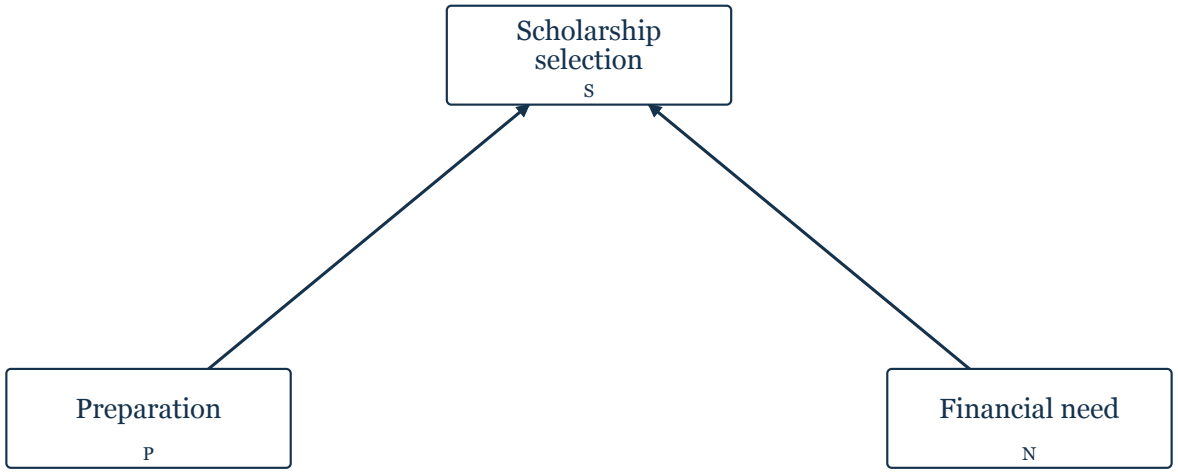
> Equally likely (P, N): $(0,0), (0,1), (1,0), (1,1)$

> Award if $P = 1$ or $N = 1$; exclude only $(0,0)$

> Among recipients: $\Pr(N = 1 \mid P = 0) = 1$, $\Pr(N = 1 \mid P = 1) = 1/2$

A shared consequence behaves differently

Conditioning can open a route that was blocked.



0 open directed routes · 0 open alternatives · gold: conditioned on

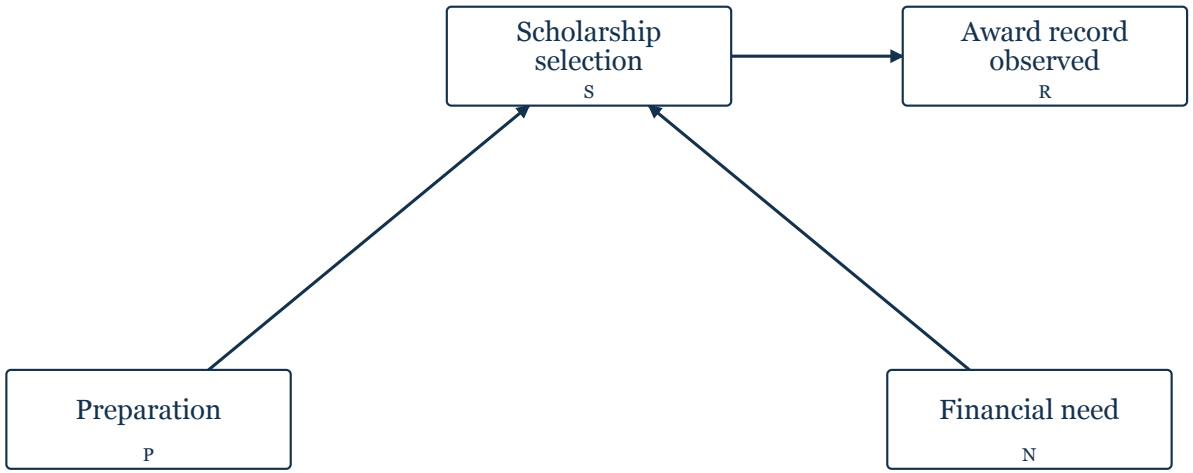
Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

> Collider: $P \rightarrow S \leftarrow N$

> Unconditioned collider blocks; selecting recipients opens the path.

Selecting a consequence can also open the route

A descendant is reached by following arrows forward.



0 open directed routes · 0 open alternatives · gold: conditioned on

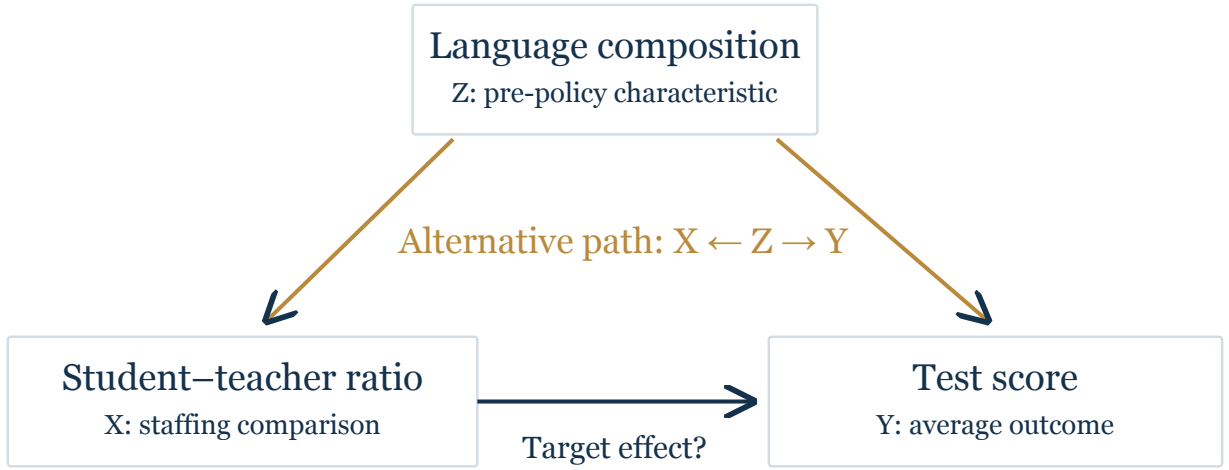
Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

> $P \to S \leftarrow N$, with $S \to R$

> Conditioning on a collider descendant can open its path.

Why the observed groups might differ

A confounder is another factor related to both staffing and scores.



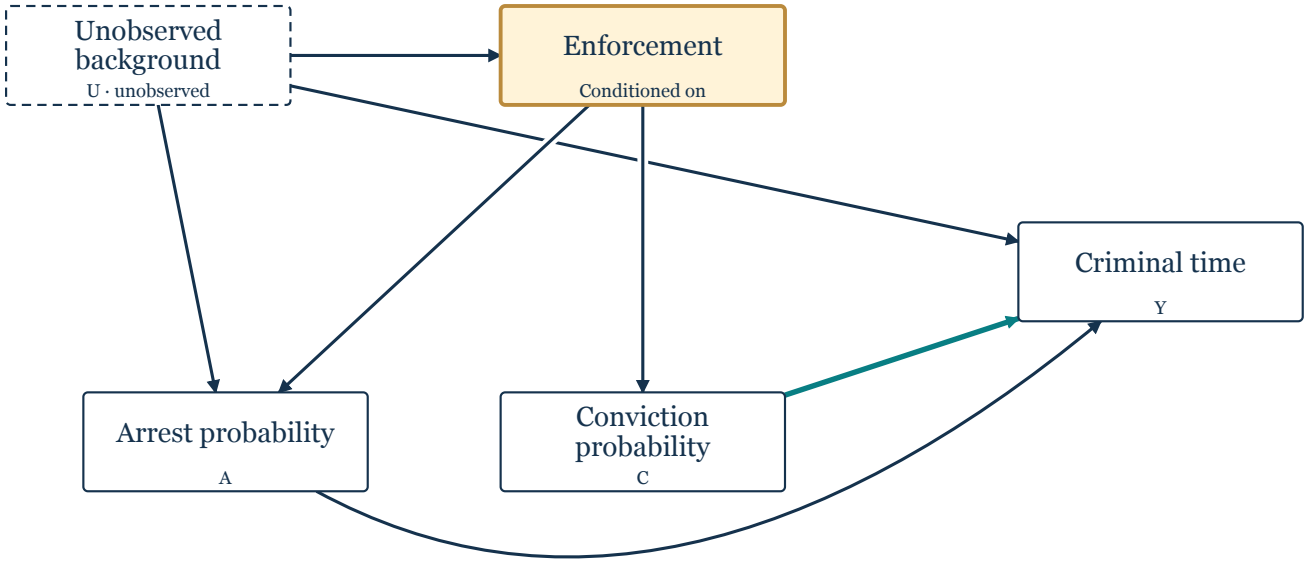
Proposed causal structure, not a relationship proved by the scatterplot.

Arrows show the proposed explanation. The scatterplot alone does not establish it.

- Z: English-learner share, in percent
- Possible confounder: related to staffing and scores
- $X \leftarrow Z \rightarrow Y$: an alternative explanation

Law and order: identify the conviction comparison

A useful adjustment closes every drawn backdoor route.



1 open directed route · 0 open alternatives · gold: conditioned on

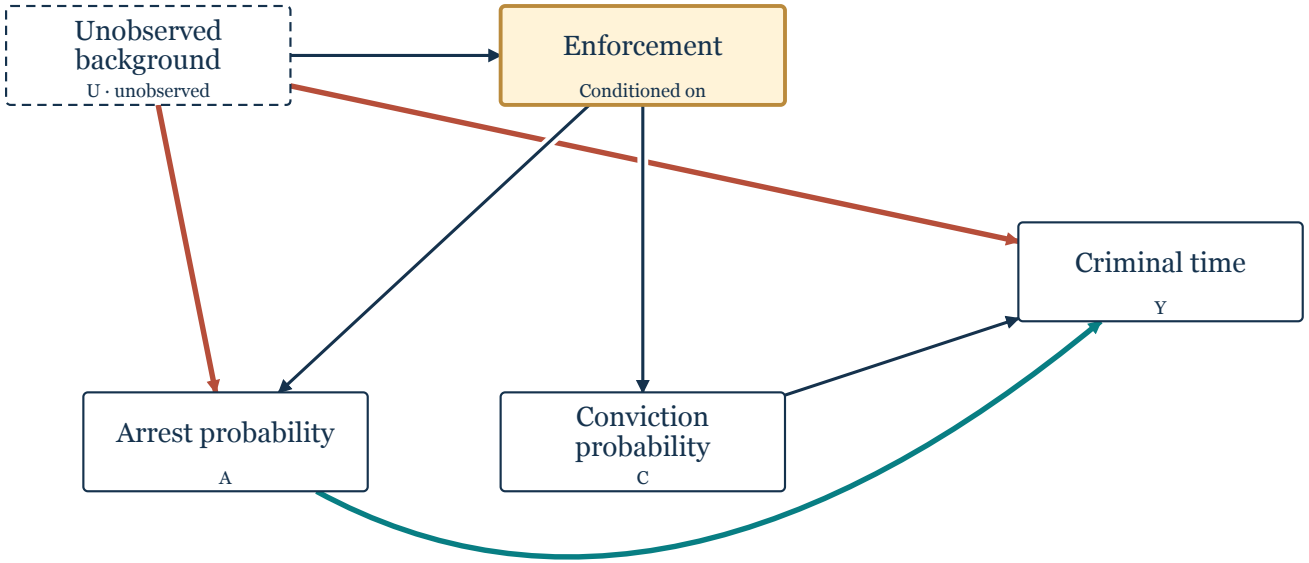
Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

> Target: conviction probability → criminal time

> Ask which variable closes every backdoor route.

Changing the target can change the required controls

The arrest comparison still has an unobserved route.



1 open directed route · 1 open alternative · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

> Remaining route: $A \leftarrow U \rightarrow Y$

> A control plan is specific to the chosen explanatory variable.

Before estimating, state what the data can answer

> Before estimating, state what the data can answer

Practice

Write four sentences naming the population, staffing intervention and outcome, one alternative explanation, and evidence that could address it.

Explanation: One observation is a school district, and staffing is measured as students per teacher. A reduction of two students per teacher gives about +4.56 fitted score points in the simple model. The difference is about +2.20 points after holding English-learner share fixed. Other district differences or staffing responses to learning needs could still prevent either number from measuring the causal gain from hiring teachers.

Audit the school observations before estimating

Try the comparison before opening the explanation.

> Read schools.csv with import delimited; inspect score, str, el and income with summarize.

> count if el>20 counts districts with English-learner share above 20 percent.

> Save your commands in a do-file and their printed results in a log.

Practice

Use schools.csv in Stata. Verify the number of districts, inspect score, str, el and income, and count districts with el above 20. Save your commands and output. Explain which sampling and causal claims cannot be verified merely by these counts.

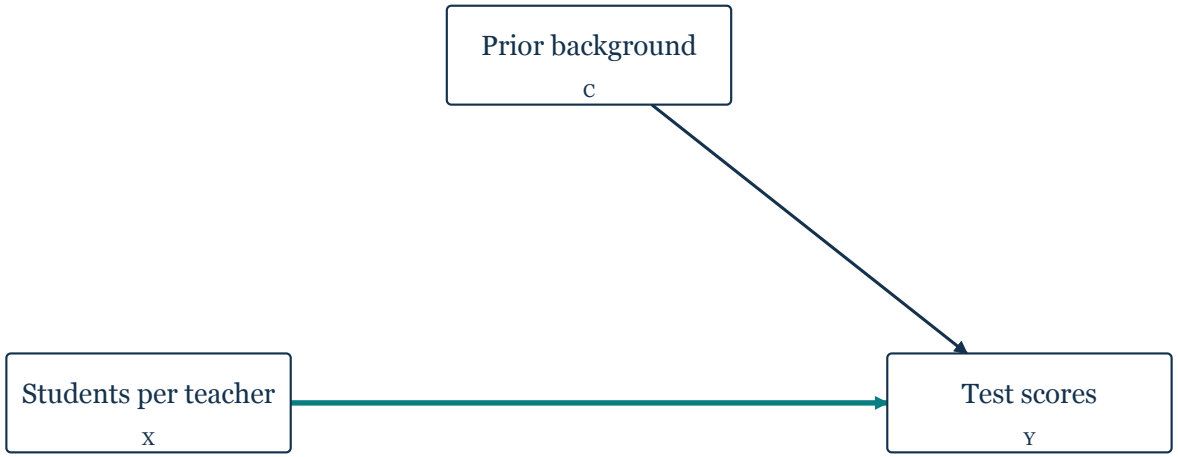
Audit the school observations before estimating: explanation

Follow the calculation and interpret its units.

> There are 420 district observations. The variables are district test score, students per teacher, English-learner percentage and district income in thousands of dollars. There are 118 districts with English-learner share above 20 percent; mean score is about 654.157 and mean staffing ratio about 19.640. The overleaf_technical.do reference prints the summaries. Correct execution, documented units and an explicit population statement are required. Counts do not establish random sampling, independent errors or a causal staffing effect.

An intervention graph changes a different object

The policy removes incoming staffing arrows rather than selecting a subgroup.



1 open directed route · 0 open alternatives · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

> Intervention: replace the assignment mechanism.

> Conditioning: compare within an observed subgroup.

Would the result apply in a new setting?

Check credibility in the original setting and relevance to the new one.

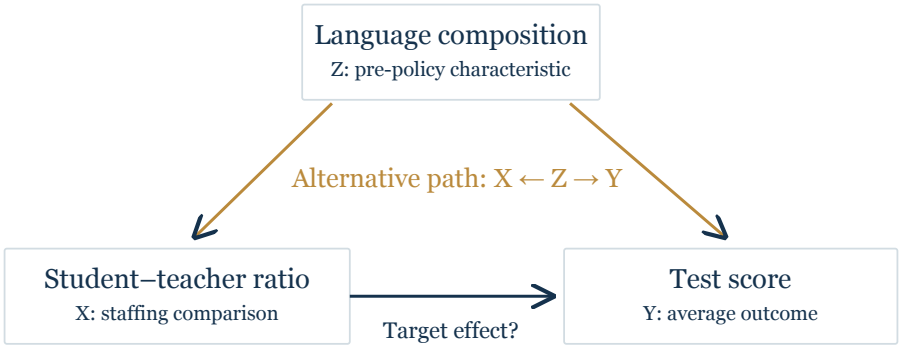
Practice

Propose a local staffing study. Name the schools or districts, policy change, learning measure, comparison and main concern.

Explanation: For example, study school-level learning gains after a funded reduction in the staffing ratio. Compare funded schools with schools whose lack of funding gives a credible estimate of what would otherwise have happened. If struggling schools are selected for funding, the comparison must address their different expected progress.

Diagnose the claim

A test of the slope and a credible policy comparison answer different questions.



Proposed causal structure, not a relationship proved by the scatterplot.

Arrows show the proposed explanation. The scatterplot alone does not establish it.

Practice

Rewrite: 'Significant slope proves that smaller classes cause higher scores.'

Explanation: Districts with lower student–teacher ratios have higher fitted scores in these data. A causal staffing effect would require a defensible comparison addressing factors such as family resources and staffing responses to student needs.

What would convince you next?

Choose new data or a new comparison to address a specific concern.

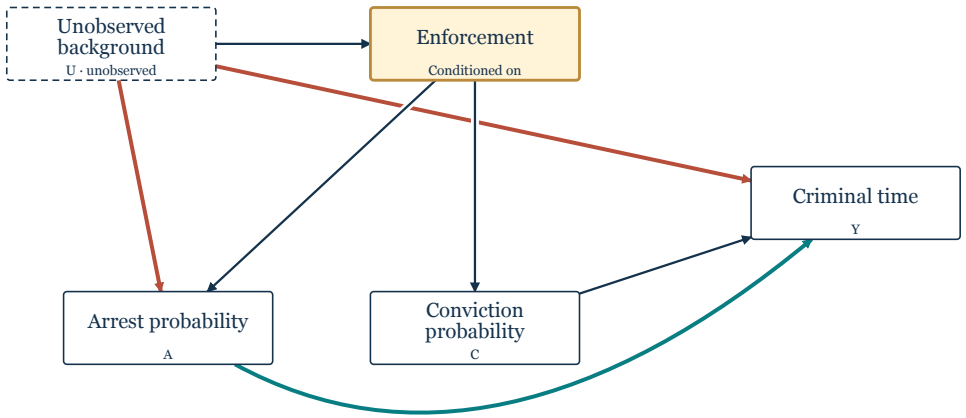
Practice

Compare adding more districts with allocating a staffing offer by chance. What can each improve?

Explanation: More districts can improve precision if the sampling conditions are suitable. Chance allocation can support a causal comparison by making the offer unrelated to other factors before the policy. We must still check whether the offer was implemented and whether the groups were affected by other differences.

DAG practice: conviction and arrest

Trace the routes before choosing a control.



1 open directed route · 1 open alternative · gold: conditioned on

Assumed causal structure · dashed boxes: unobserved · gold boxes: conditioned on

Practice

In the law-and-order model, hold enforcement fixed. Which backdoor route from arrest probability to criminal time remains open? Explain why the same enforcement adjustment can be enough for the drawn conviction comparison but insufficient for arrest.

DAG practice: conviction and arrest: explanation

Explain what conditioning changes in this stated graph.

> For arrest, the route $A \leftarrow U \rightarrow Y$ remains open because unobserved background affects both arrest probability and criminal time. Holding enforcement fixed does not block that route. For conviction, every drawn backdoor starts through enforcement, so conditioning on enforcement blocks those routes in this stated graph. The result is specific to the chosen explanatory variable and assumed arrows. If the real conviction process had another unmeasured parent affecting the outcome, that control set could also be insufficient.