Introduction to Causal Inference

Day 2: Causal Graphs and What They Reveal

Chad Hazlett

UCLA

chazlett@ucla.edu

Academia Sinica ISSM, 17 July 2026

Where we are, and today's question

Yesterday: potential outcomes, and two identification scenarios: **experiments** (ignorability) and **conditional ignorability** (CI).

With CI we hope for

$$\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp D_i \mid X_i$$

which we also called "selection on observables", (SOO), and "no unobserved confounders (given X)."

Today we talk about graphical causal models ("DAGs") that help show:

- **Which** X is the "right" X? What really are confounders?
- Is any X good enough?
- What can **go wrong** if I pick certain X?
- More generally: we need the whole causal story, encoded on the graph to properly pose and answer a variety of questions.

The core idea graphs visualize

We all know the mantra “correlation does not imply causation”.

But correlation of D and Y *actually does* imply causation...it just might not be the causal effect of D on Y you are seeing.

If you do some analysis that shows some association of D and Y (possibly after adjusting for X), what explains it?

There are three possibilities:

1. D causes Y (the effect we want to estimate)
2. A **common cause** of D and Y.
3. Analytic errors that create or alter the association of D and Y.

Causal graphs visualize and tell us how to remove unwanted sources of association.

A new motivating example

Suppose people self-select into a new drug. You find:

	Recovered (count / total)	
	Drug	No drug
Mild	81 of 87 (93%)	234 of 270 (87%)
Severe	192 of 263 (73%)	55 of 80 (69%)
Combined	273 of 350 (78%)	289 of 350 (83%)

Pooled, the drug "hurts" (78 vs 83), but within each severity stratum the drug helps (93 vs 87; 73 vs 69).

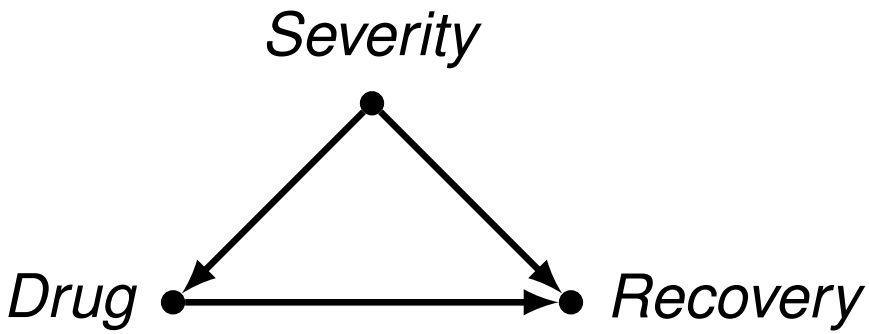
Why? The pooled comparison compares groups that differ in *severity*, not just the drug.

With potential outcomes: $\mathbb{E}[Y(0)|D=1] - \mathbb{E}[Y(0)|D=0] < 0$ (negative bias)

Preview for intuition: Confounding, with a DAG

What we *believe* about the world:

- Severity increases drug-taking
- Severity causes reduced recovery
- The drug might cause recovery.



You *see* the two reasons why *Drug* and *Recovery* are associated:

1. *Drug* $\rightarrow$ *Recovery* (what we want),
2. *Drug* $\leftarrow$ *Severity* $\rightarrow$ *Recovery* (common cause — a spurious association).

Conditioning on *Severity* removes (2), leaving only (1).

Formally now: What's a DAG?

A **Directed Acyclic Graph** (DAG):

- **Nodes** = variables.
- **Arrows, Edges** = direct causal relationships ($A \to B$: A is a direct cause of B).
- **No cycles**: you can't follow arrows and end up back where you started.
- **Missing arrows encode assumptions**: "A is *not* a direct cause of B." These are the assumptions doing the real work.
- **Dashed arrow** / **bidirected arc** = unobserved common cause.

Three things a DAG is *not*:

- Not a list of the variables in your dataset. (Many nodes may be unobserved.)
- Not a statistical model. (No parametric assumptions.)
- Not a complete theory of the world. (Just enough structure for the causal claim you want to make.)

Three path types

A "path" between two nodes is any sequence with edges connecting them (in any direction). Three elementary components of paths:

- **Chain**: $A \rightarrow B \rightarrow C$
- **Fork** (common cause): $A \leftarrow B \rightarrow C$
- **Collider** (common effect): $A \rightarrow B \leftarrow C$

These imply various independence and conditional independence relationships that are useful to us. Let's see how:

Chain

$A \perp\!\!\!\perp C$? No — A affects C through B.

$A \perp\!\!\!\perp C \mid B$? Yes — conditioning on B blocks the chain.

Lesson: a chain transmits association; conditioning on the middle blocks it.

Fork (common cause)

$A \perp\!\!\!\perp C$? No — the common cause B induces association.

$A \perp\!\!\!\perp C \mid B$? Yes — conditioning on the common cause blocks it.

Lesson: a common cause creates spurious association; conditioning on it removes it. **This is classical confounding.**

Collider (common effect)

$A \perp\!\!\!\perp C$? Yes — no causal path, no common cause.

$A \perp\!\!\!\perp C \mid B$? **No!** — conditioning on the collider *opens* a path.

Lesson: colliders are the reverse of the other two. They are *already blocked*, and conditioning on one *creates* an association that wasn't there.

Colliders are real

Pearl's favorite:

Rain $\rightarrow$ Wet lawn $\leftarrow$ Sprinkler was on

Observing a wet lawn, learning it did *not* rain raises the chance the sprinkler ran.

Hollywood:

Looks $\rightarrow$ Fame $\leftarrow$ Talent

Among famous actors, the good-looking ones are less talented (on average).

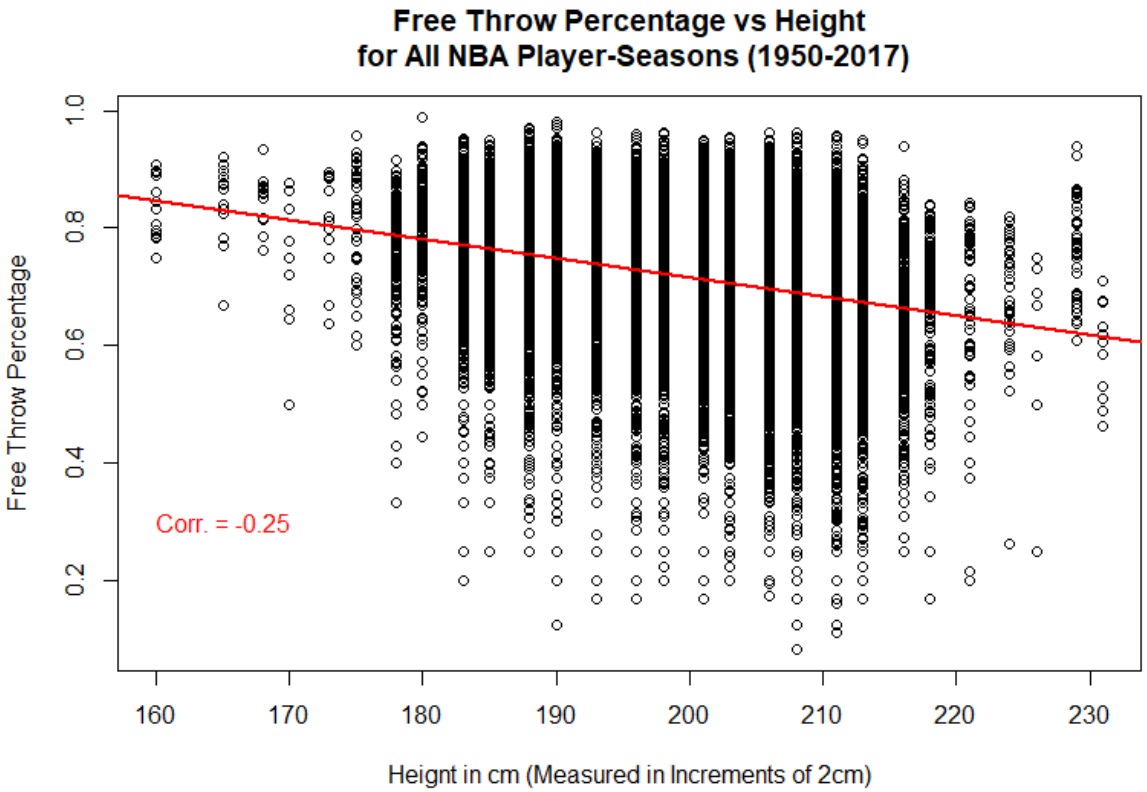
Academic tenure:

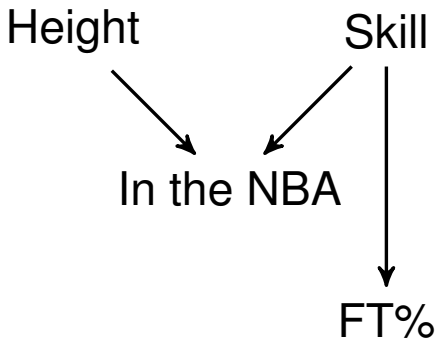
Productivity $\rightarrow$ Tenure $\leftarrow$ Originality

Among tenured faculty, the prolific ones are less original (on average).

Lesson: Conditioning on something can give rise to a spurious association. This produces various selection biases, post-treatment adjustment bias, and other analytic errors.

A collider in the wild





Source: Adam Rohde, via `basketball-reference.com` *and kaggle.com/datasets/drgilermo/nba-players-stats.*

d-separation

Putting the three patterns together:

Rule

***d*-separation.** *A and C are d-separated by a set Z iff every path between them is blocked. A path is blocked when:*

- *It contains a chain or fork at a node in Z,* ***or***
- *It contains a collider such that neither the collider nor any descendant of it is in Z.*

If A and C are d-separated by Z, then

$$A \perp\!\!\!\perp C \mid Z$$

In plain English: When A and C are d-separated by Z in this way, you can condition on Z as a way to remove the association between A and C due to those paths.

Back to the core idea

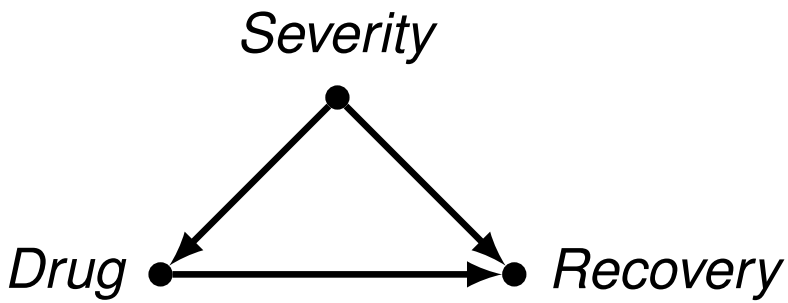
Recall the big idea: You do some analysis and see that D and Y associated. But is this down to $D \rightarrow Y$, or something else?

What we want: a rule for choosing set $\{X\}$ such that, after conditioning, the *only* remaining association between D and Y is the causal arrow $D \rightarrow Y$.

That rule is the **backdoor criterion**.

Front-door and back-door paths

- A **frontdoor path** from D to Y starts with an arrow *out of* D. These carry the effect we want to learn. We want to **leave them open**.
- A **backdoor path** from D to Y starts with an arrow *into* D. These carry spurious association. We want to **block them all**.



Frontdoor: $Drug \rightarrow Recovery$.
Backdoor: $Drug \leftarrow Severity \rightarrow Recovery$. Condition on *Severity* to block it.

The backdoor criterion

Rule

Backdoor criterion. *$\{X\}$ satisfies the backdoor criterion relative to (D, Y) if:*

1. *No node in X is a descendant of D (don't block frontdoor paths),* ***and***
2. *X blocks every backdoor path from D to Y.*

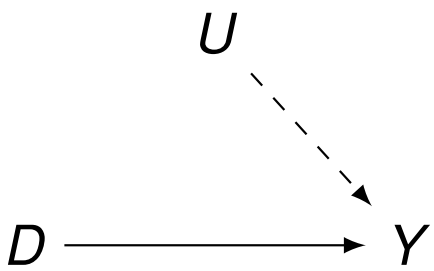
For humans, find a set X that:

1. blocks all *spurious* paths (every backdoor),
2. leaves *causal* paths untouched (no descendants of D),
3. doesn't *open new* spurious paths (no colliders or their descendants unless paired with another blocker downstream).

Backdoor = CI: two languages, one object

Connecting this to yesterday:

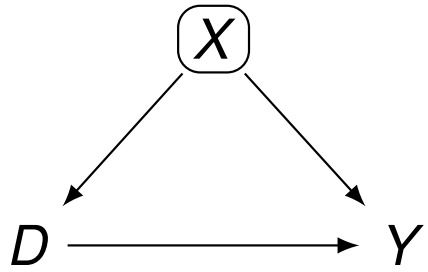
Experiment



Randomization deletes every arrow *into D*.
No backdoor; condition on nothing.

$$\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp D_i$$

Conditional ignorability



X points into D and Y. Condition on X (boxed) to block it.

Assume no other (unmeasured) common cause of D and Y, so

$$\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp D_i \mid X_i$$

Adjustment formula for ATE

I will leave some deeper background Prof. Lai, but in short, to get the ATE based on this we condition on X the same way we did with subclassification:

$$ATE = \sum_x (\mathbb{E}[Y \mid D=1, X=x] - \mathbb{E}[Y \mid D=0, X=x])\ P(X=x)$$

That is: figuring out what to condition on using the DAG and adjusting for it is the same thing we were doing by (i) claiming CI given those variables and (ii) conditioning on them.

Again, because we can't always form discrete strata of X, in practice we rely on other adjustment approaches that effectively move $p(X|D=1)$ and $p(X|D=0)$ to a common $p(X)$ or model away their differences as they affect Y.

So what *is* a confounder?

I cheated yesterday by defining confounders and confounding loosely. Now we can clear that up, and this is where DAGs really help.

The tempting definition: *a variable associated with both D and Y*, which we should therefore control for.

That rule is **wrong**:

- A **mediator** $D \rightarrow M \rightarrow Y$ is associated with both — but controlling for it *removes the very effect* we want.
- A **collider** becomes associated with both once conditioned on — and controlling for it *creates* bias.

"Associated with both" is neither necessary nor sufficient.

So what *is* a confounder? (cont.)

A useful (if circular) definition of confounding is rooted in what adjusting for it does:

- **Confounding** = an open *backdoor* path between D and Y. In other words, a source of "spurious" association.
- A set X **deconfounds** if it *blocks every backdoor* (satisfies the backdoor criterion).
- A "confounder" is a variable worth adjusting for *as part of such a set* to close backdoors.

The well-defined object is the adjustment *set*, not any single variable. The graph tells you which sets work.

Good and bad controls

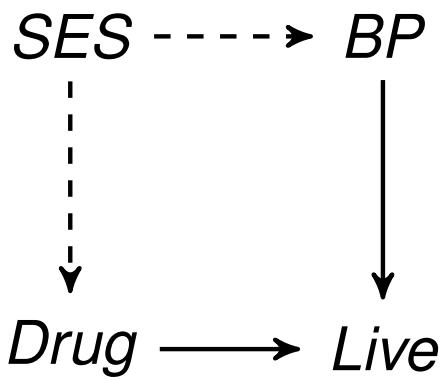
Given a candidate variable Z, should you adjust for it? The *graph* decides.

- Common cause **"confounder"**/ anything on backdoor path: **control** — closes a backdoor.
- **Mediator** $D \to Z \to Y$: **don't** — you would erase part of the effect.
- **Collider** (of D and Y, or descendant thereof): **don't** — opens a spurious path.
- **Descendant of** D (post-treatment): **don't** — same dangers, plus over-control.
- **Cause of** Y **only** (not of D): **optional** — no bias, but can improve *precision*.
- **Instruments** (cause of D but not of Y): **don't**: no bias strictly, but can worsen precision and amplify residual biases.

Lesson: "Controlling for more" is not safer.[1]

[1] *Cinelli, Forney & Pearl 2022, "A Crash Course in Good and Bad Controls."*

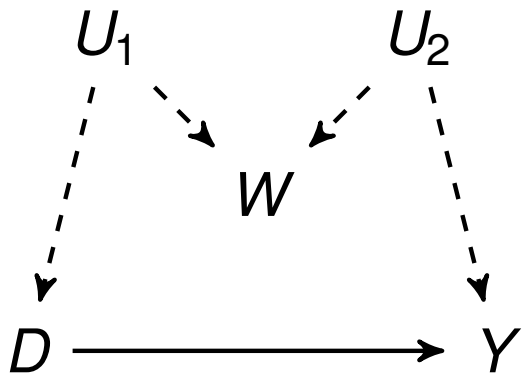
Example 1: a hidden confounder, but a workable adjustment set



SES unobserved (dashed). Can we identify the effect of *Drug* on *Live*?

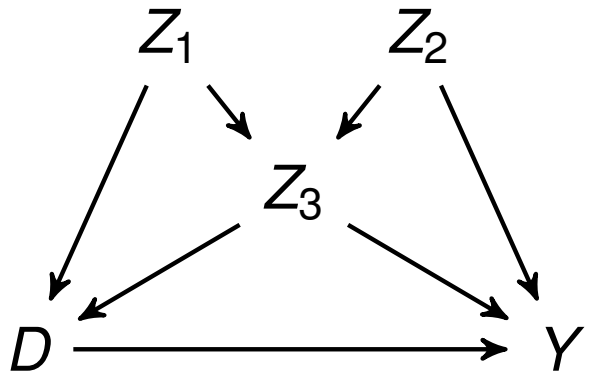
- Backdoor paths? $Drug \leftarrow SES \rightarrow BP \rightarrow Live$
- Observed set that blocks it? $\{BP\}$
- Equivalent CI statement: $\{Live(1), Live(0)\} \perp\!\!\!\perp Drug \mid BP$.

Example 2: conditioning can *break* identification



- Backdoor paths? The only candidate is $D \leftarrow U_1 \rightarrow W \leftarrow U_2 \rightarrow Y$.
- W is a **collider** on it, so the path is already blocked. The empty set identifies the effect!
- A colleague says, "let's play it safe and control for W." You open the path — now D and Y are spuriously associated through U_1 and U_2.

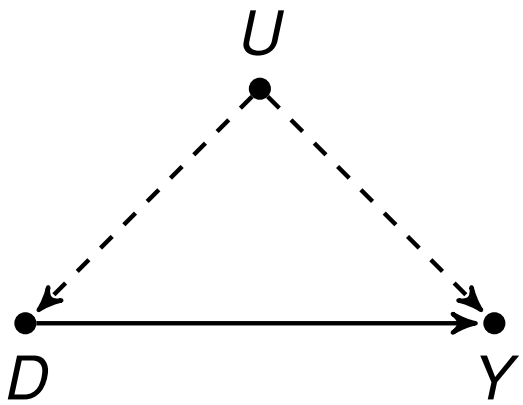
Example 3: The sneaky one



- Z_3 is a **common cause** of D and Y so you *must* adjust for it.
- But Z_3 is *also* a **collider** ($Z_1 \rightarrow Z_3 \leftarrow Z_2$). Conditioning on Z_3 opens that path by associating Z_1 with Z_2.
- Fix: take Z_3 *and* a parent. Valid sets: $\{Z_3, Z_1\}$, $\{Z_3, Z_2\}$, $\{Z_1, Z_2, Z_3\}$ — but *not* $\{Z_3\}$ alone.

This goes way beyond any simple correlational rule about confounders.

Example 4: The worst.



- Backdoor path: $D \leftarrow U \rightarrow Y$. U unobserved.
- No observed variable blocks it. No set satisfies the backdoor criterion.

DAGs tell you when you're stuck — not just when you're fine.

What to do? CI + sensitivity, or find another ID strategy.

Worksheet 1

Reading DAGs

Pen and paper, $\sim$15 min — then we compare notes.

list backdoor paths · test adjustment sets · translate backdoor $\rightarrow$ CI

Don't Condition on Post-Treatment Variables

CI requires conditioning on *pre-treatment* covariates. Conditioning on a variable affected by treatment can **create** bias even if the original assignment was random.

Example: A city randomizes a community policing program across neighborhoods. You want the effect on crime.

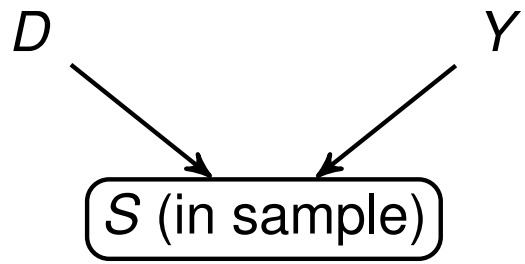
- Some neighborhoods implement the program well.
- You restrict analysis to "neighborhoods that implemented well."
- But implementation is post-treatment and affected by unobserved things (e.g. civic mindedness) that might affect the outcome (crime).
- Restricting to implementing neighborhoods among the treated changes the treated group's composition in a way you don't mimic in the control units.
- This "breaks the randomization", creating a confounder that wasn't there before!

DAGs make this clear with *collider bias*.

Selection bias *is* collider bias

Whenever your sample is *selected* — who ends up in the data depends on other variables — you have **conditioned** on that selection, whether you meant to or not.

If "in the sample" (S) is a common *effect* of D and Y, it is a **collider**. Looking only at $S = 1$ opens a spurious D–Y association.



Example. A job-training study measures employment at follow-up. Both being *treated* and *employment* make people easier to re-contact. Analyze only those you re-contact ($S = 1$) $\Rightarrow$ a training–employment association appears *even if training did nothing*.

Missingness/selection issues are causal issues. Always ask what determined who is in your data.

Example: Blattman & Annan (2009)

"The Consequences of Child Soldiering." Review of Economics and Stats.

Question: What is the effect of abduction by the Lord's Resistance Army (LRA) on education and earnings?

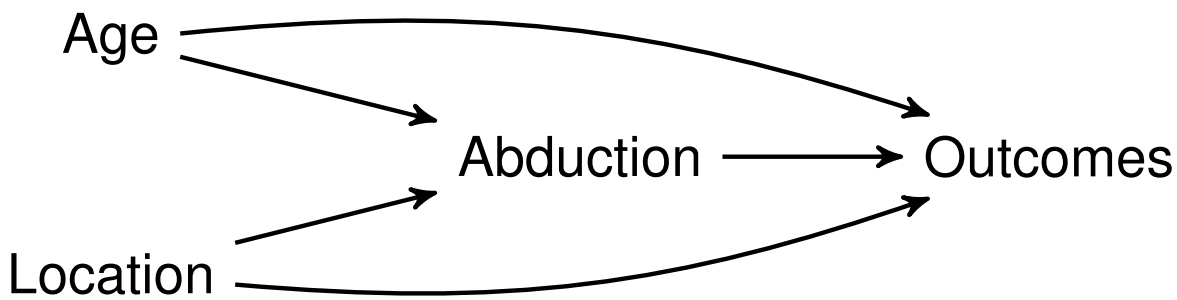
ID strategy: conditional ignorability (CI).

> *"Abduction was large-scale and seemingly indiscriminate; 60,000 to 80,000 youth are estimated to have been abducted. . . Youth were typically taken by roving groups of 10 to 20 rebels during night raids on rural homes. Adolescent males appear to have been the most pliable, reliable and effective forced recruits, and so were disproportionately targeted."*

That's the key – it is arguing that *conditional on age and location*, abduction "as is" random; no other confounders.

Blattman & Annan: the same claim, two languages

The identifying claim as a picture:



Two ways to say the identifying assumption:

- **DAG talk**: age and location are the *only* common causes of abduction and the outcomes — no *other* arrow runs into abduction and also reaches the outcome.
- **CI talk**: abduction is as-good-as-random *within strata of age and location*.

The LRA story (indiscriminate abduction, given age and location) is what *defends* that assumption — not the data.

Blattman & Annan: Evaluating the CI Claim

Data: panel survey of male youth in war-affected regions of Uganda.

- `abd`: abducted by LRA (the treatment)
- `c_ach`, ..., `c_pal`: location indicators
- `age`: age in years
- `fthr_ed`, `mthr_ed`: parental education
- `hh_size96`: household size in 1996
- `educ`, `distress`, `logwage`: outcomes (post-treatment!)

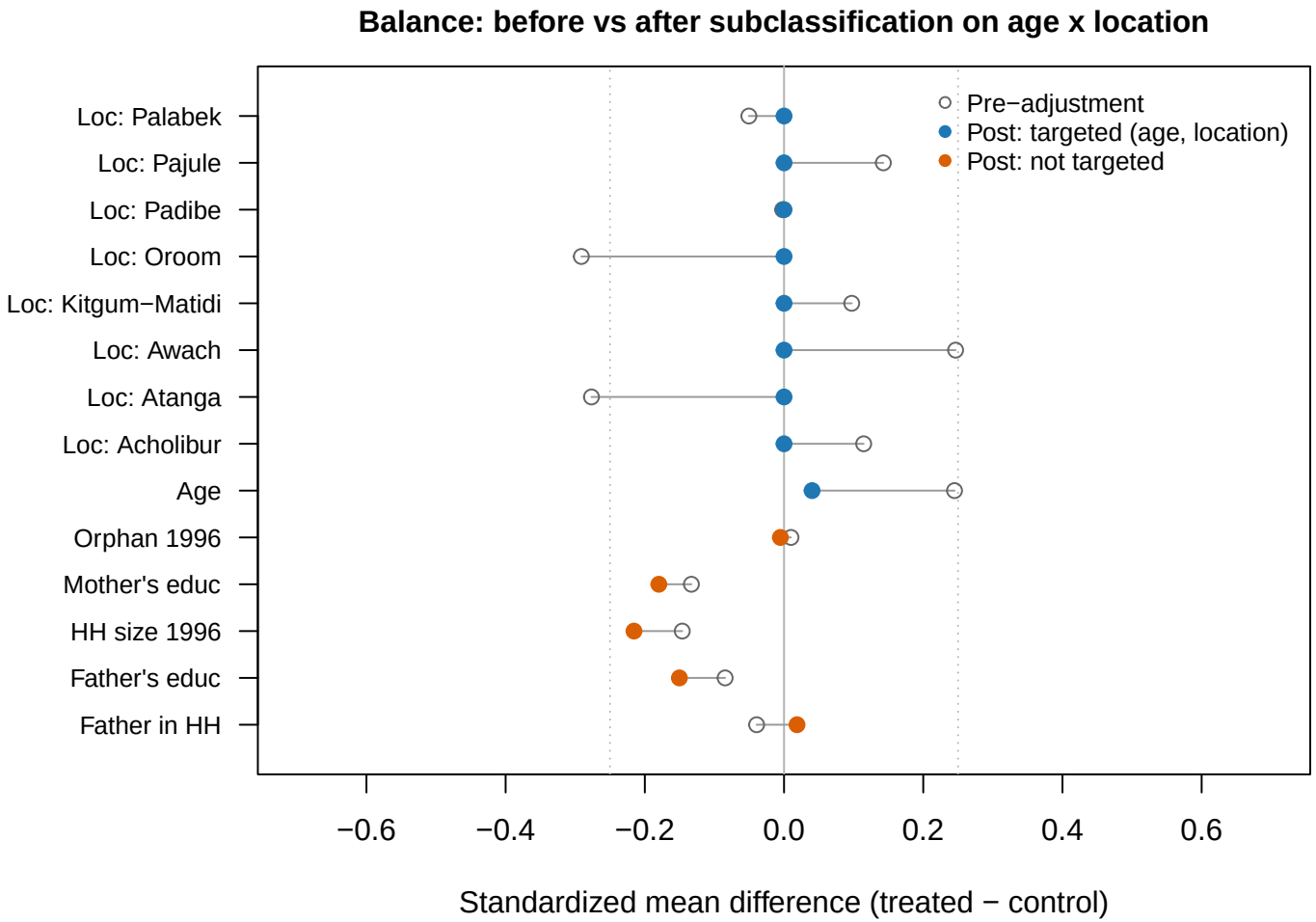
Which variables must you condition on given the CI argument?
Age and location: required by ID strategy.

Should you see balance on other variables, after conditioning on these?
Yes roughly, if CI is right. Gives falsification opportunity, not a test. (Why?)

Why not condition on `educ`, `distress`, `logwage`?

Blattman & Annan: Checking Balance

Apply the "same adjustment on X" procedure to every pre-treatment variable, before vs. after subclassifying on age $\times$ location:



Open circles: pre-adjustment. Filled: post-adjustment. Blue = variables we stratified on; orange = variables we did not.

Reading the Balance Plot

Conditioned-on variables (age, location): post-adjustment SMDs at or very near zero. *Expected* — any variable used to define strata balances by construction. Seeing these at zero only confirms the procedure worked.

Non-targeted variables: these are the *informative* checks.

- Orphan status and father-in-HH: close to balanced before *and* after. Consistent with the CI story.
- Father's education, mother's education, HH size: imbalanced before, and *still imbalanced* (slightly more so) after. Conditional on age and location, abducted youth still come from smaller, less-educated households.

What does this mean?

- Either these variables also drove selection and should be included in X. We can add them to what we adjust for.
- Or (worse) an unobserved thing correlated with these variables is the problem...it is unobserved things we worry about.

Falsification, not a test. Poor balance on (untargeted) observables can warn you there is a problem; cannot tell you there is no problem.

What Makes This a Credible CI Argument?

Remember it is not the procedure that would make these ATEs, it is believing the assumption of no unmeasured confounders beyond age and location.

What is good is that the paper gives reason to believe it:

- A clear argument for why treatment was as-if random conditional on specific, named covariates.
- Grounded in knowledge of how the LRA operated, not just the data.

Contrast with: "We ran a regression with 50 controls and the coefficient was significant." That is not an argument for CI or any other ID strategy.

Worksheet 2

Which X, and What Goes Wrong

evaluate a CI claim · name an omitted confounder · the post-treatment / collider trap

Your job: defend the assumption (not the estimate)

Defending the identification assumption *is* the substantive work. Each **do** comes with a **don't be fooled**:

- **State** the assumption plainly.
  Don't be fooled by: sophistication that hides whether there is an assumption at all.
- **Defend** it — why it is plausible, consider conuterexamples – *name the confounders that keep you up at night*.
  Don't be fooled by: a fancy estimator. Estimation is not identification.
- **Calibrate** your confidence to your defense of the assumption.
  Don't be fooled by: small p-values and tight intervals — precision about the wrong thing.
- **Falsify** where you can. A check is a chance to *find* a problem.
  Don't be fooled by: balance or specification tests as "proof." Nothing in the data rules out an unobserved confounder.
- **(Later) Stress-test** sensitivity to violations of the *core* assumption.
  Don't be fooled by: sensitivity analyses that fuss over lesser problems.

The broader identification landscape

You now have the two foundational cases. Most of the field — and much of the rest of this school — is other *identification strategies*, each resting on a different assumption you must defend:

Strategy	**The assumption you must defend**
Randomized experiment	assignment is random *by design*
Conditional ignorability / SOO	no unobserved confounders given X
Instrumental variables	the instrument affects Y *only* through D
Regression discontinuity	units just above/below the cutoff are comparable
Difference-in-differences	ignorable trend in $Y(0)$
Panel / synthetic control	confounding leaves footprint in pre-tx outcomes

None escapes the need for (demanding) assumptions — they are just different; hopefully you can find one that is defensible in a given setting.

Back to the beginning

Hopefully these four claims now seem more concrete:

1. A causal claim is about a comparison we can *never fully observe*. So no amount of data settles it *by itself*.
2. *"Identification" comes before estimation.* We need assumptions connecting what we *want to know* to what we *can see*. Estimation comes after.
3. There are no "causal methods." Credibility comes from the *assumptions, and the story behind them*, not the technique.
4. Experiments are not a separate *kind* of knowledge. Every design leans on assumptions; with experiments, the assumption holds by design. For others, they need to be argued.

Next: the Safe Inference talk (20 July)

These two days gave you the tools you need for any causal inference machinery, plus a focus on ignorability, conditional ignorability/ backdoor criterion.

In real research we often find we don't have a randomized experiment nor a fully defensible set of confounders to condition on. The usual response: adopt a strong assumption, state a caveat, and hope.

What if we ask a difference question:

- Don't state an answer that holds only when the assumptions are exactly true.
- Instead ask: Given the data we actually have, and assumptions we are genuinely willing to defend, what can we (not) *credibly* conclude?
- We call this safe inference.

All this will be built on the same PO and DAG tools you've just learned, while inviting some creative identification opportunities. See you then!