Causal inference lightning recap 1

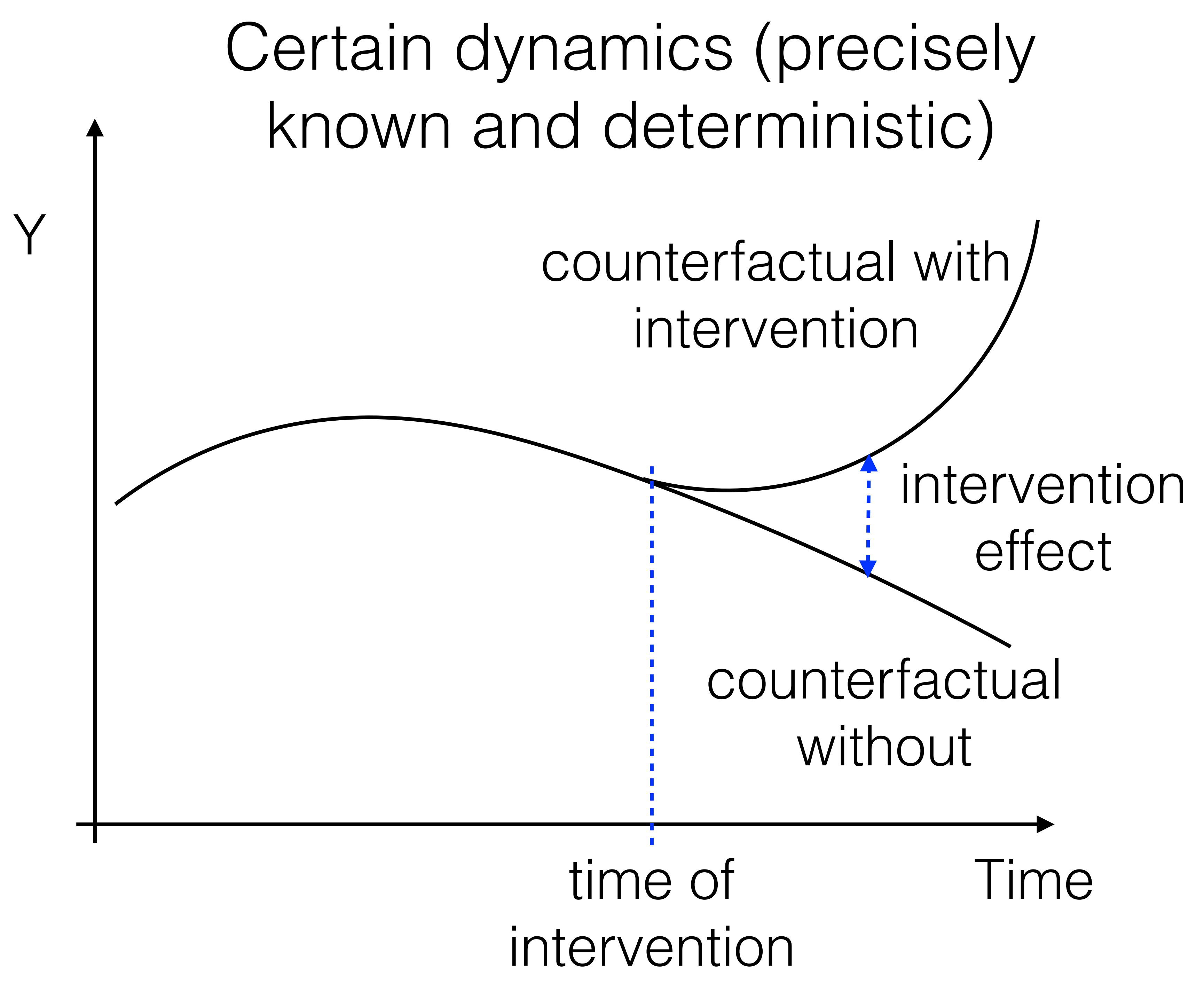
Counterfactuals/scenarios: things that could have happened (past) or could happen (future).

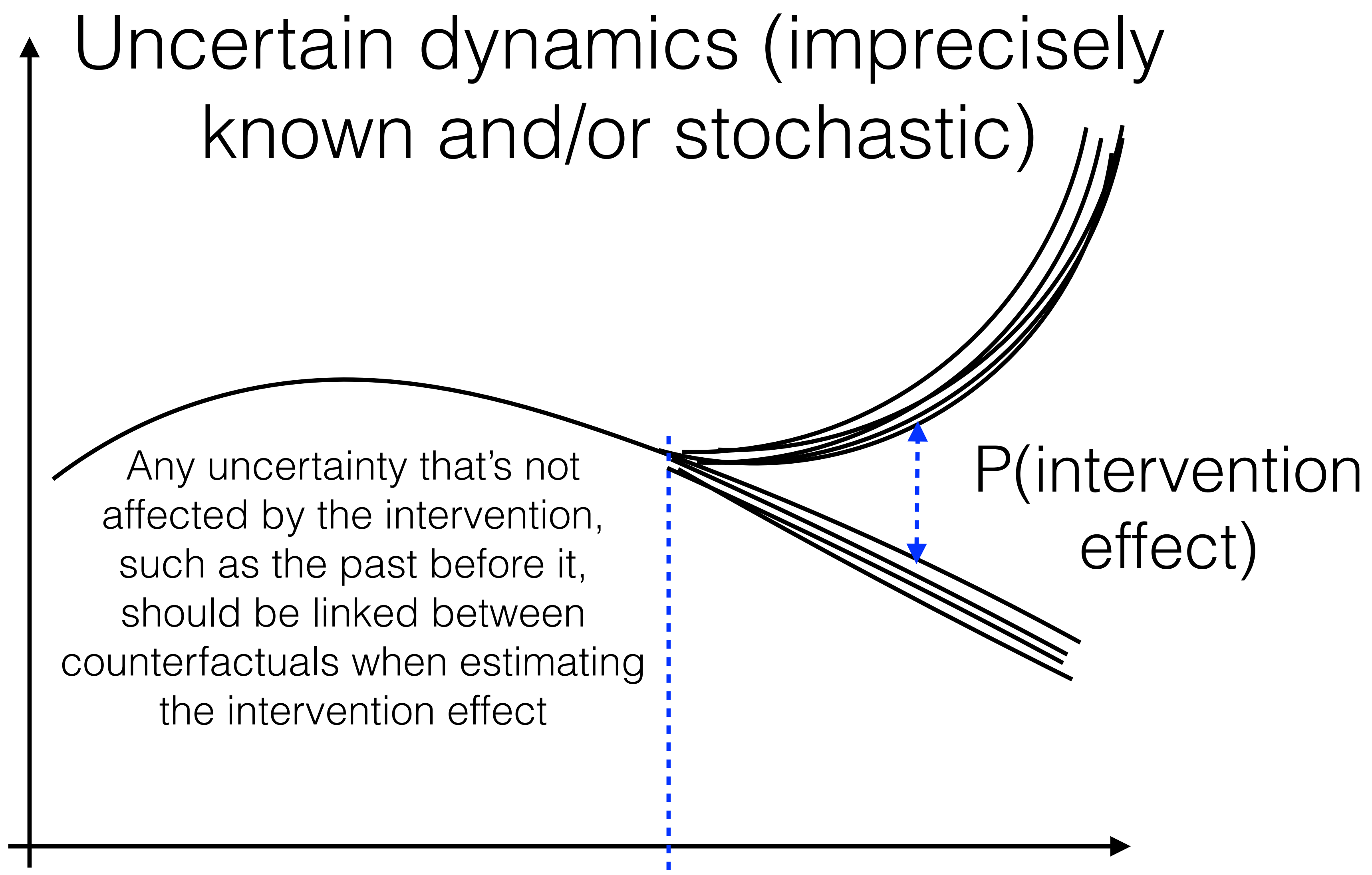
An intervention: an action that causes two counterfactuals to begin diverging from that point in time.

A comparison of precisely defined counterfactuals is necessary to:

- define causality and related words like “affects”, “because of” (i.e. attribution). e.g. “obesity causes poor health” is an ill-defined statement because the intervention is not specified. Reducing obesity by chopping off arms would not improve health.

- define the meaning of some common but loaded words like *should*: “we should do X” rests upon defining some measure of value and finding this is higher when doing X than when doing some precisely defined alternative to X.





Obesity example credit: Christophe, I think

Causal inference lightning recap 2

Learning that (un)desirable outcome Y is correlated with actionable variable X only implies that we should change X to get more (or less) Y if the correlation is causal.

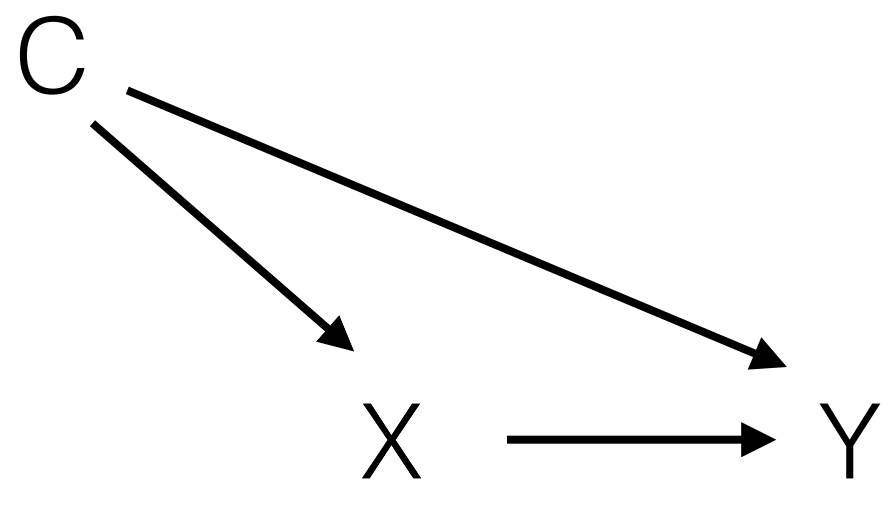
Correlation does not imply causation. But sometimes it does. So when does it?

“Draw your assumptions before your conclusions” - Miguel Hernán

DAG = directed acyclic graph

DAG for causality: a model for a set of causal effects, visualised as a graph.

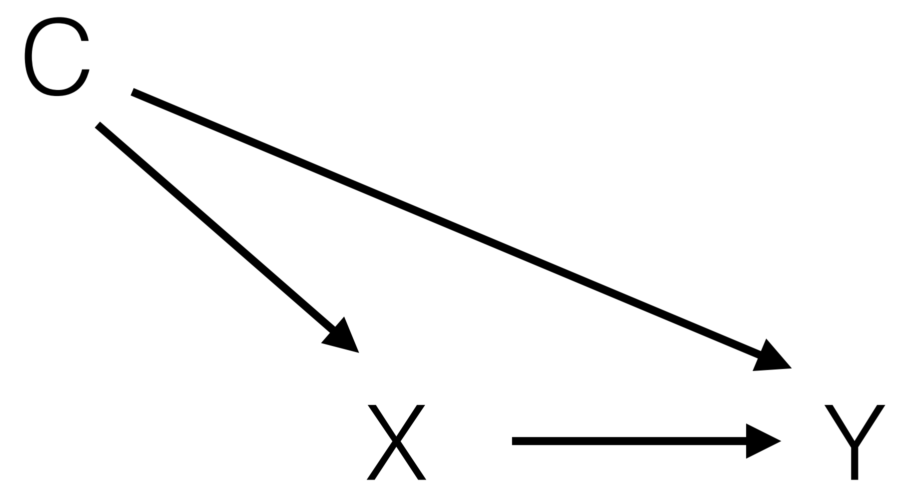
Sometimes (like here) time flows left to right, so all arrows go left to right.



This DAG means:

- C (for confounder) causally influences X. i.e. in a (perhaps necessarily hypothetical) experiment in which C were manipulated, keeping other things constant, X would change. Or if there is uncertainty present, the distribution for P(X) would change.
- C causally influences Y
- X causally influences Y but not C
- Y does not causally influence X or C
- Other omitted variables may causally affect X or Y or C, but these do not result in ‘open pathways’ between any two of X or Y or C…

Causal inference lightning recap 3



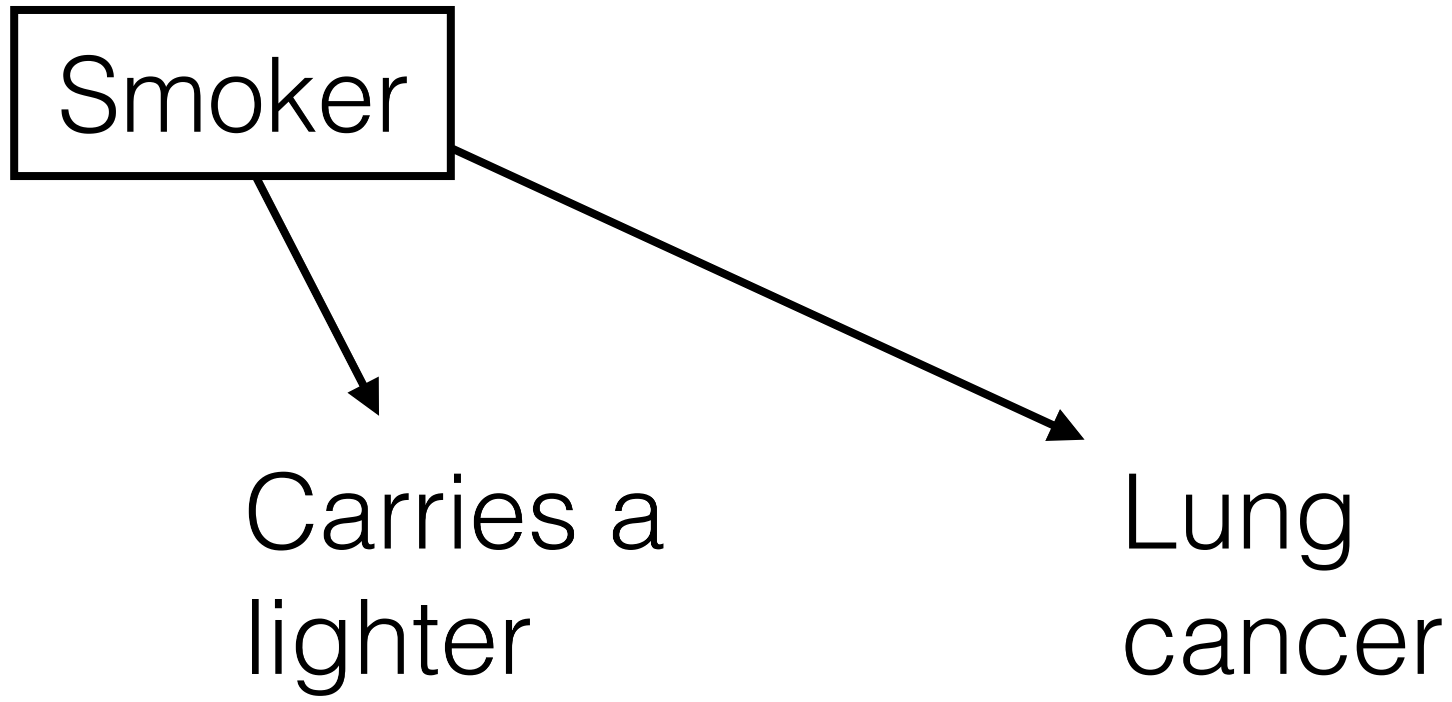
This DAG says C and X affect Y but not *how*. P(Y) could be any function of X and C: different models, same DAG.

Knee-jerk way of 'adjusting for confounding' when trying to interpret correlation between X and Y as causal: `lm(Y ~ X + C)`

What that means: $P(Y \mid X, C, \beta_X, \beta_C, \alpha, \sigma) = N(Y \mid \beta_X X + \beta_C C + \alpha, \sigma^2)$ with model parameters β_X, β_C, α, σ to be estimated. Be aware:

1. Other functional forms and distributions may be better
2. This is not estimating the effect of C on X
3. This is not estimating the *total* effect of C on Y, which is the sum of two pathways (tangent: Bayesian estimators may be better?)
4. Collinear predictors may be a problem: in the limit that X is proportional to C, $X = mC$, can re-write that as $N(Y \mid (\beta_X + m\beta_C)X + \alpha, \sigma^2)$. With this model and data, can only estimate $\beta_X + m\beta_C$, not β_X and β_C separately. Is not a problem if you only want to use the model to predict Y for more data with similarly correlated X & C. Is a problem for causal inference.

Causal inference lightning recap 4

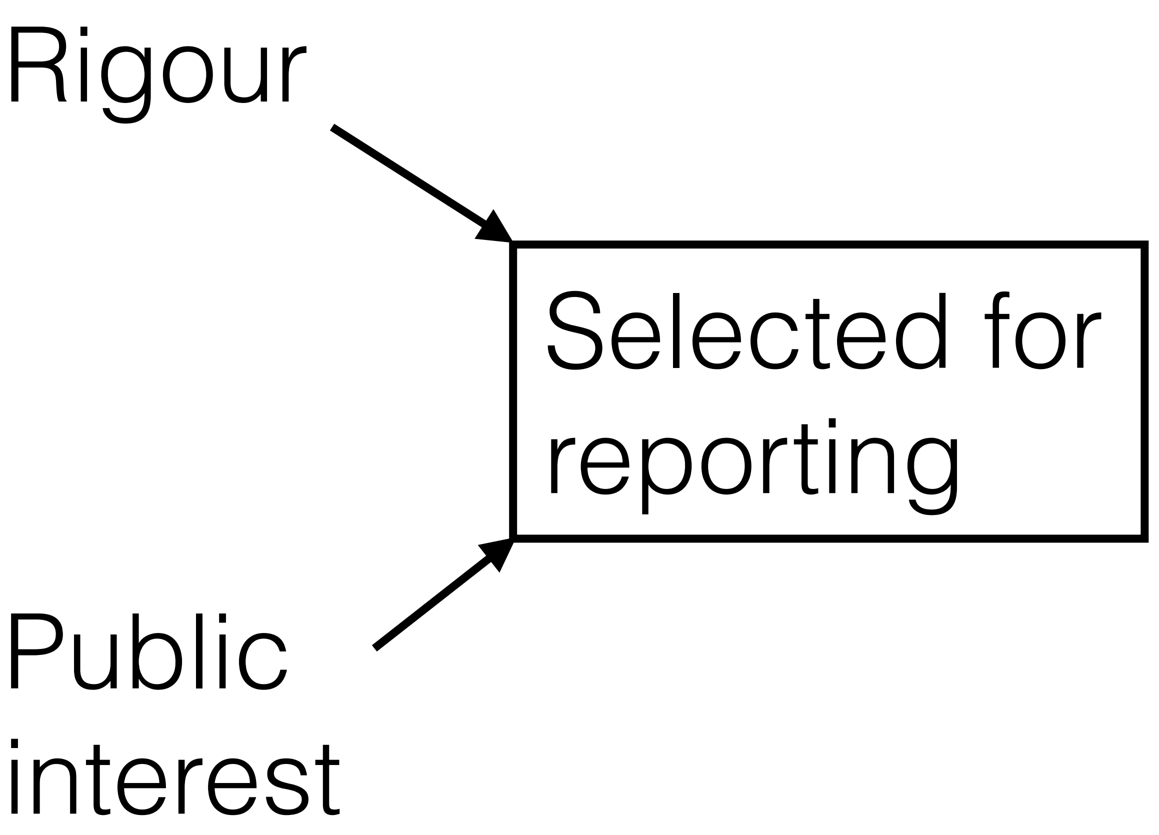


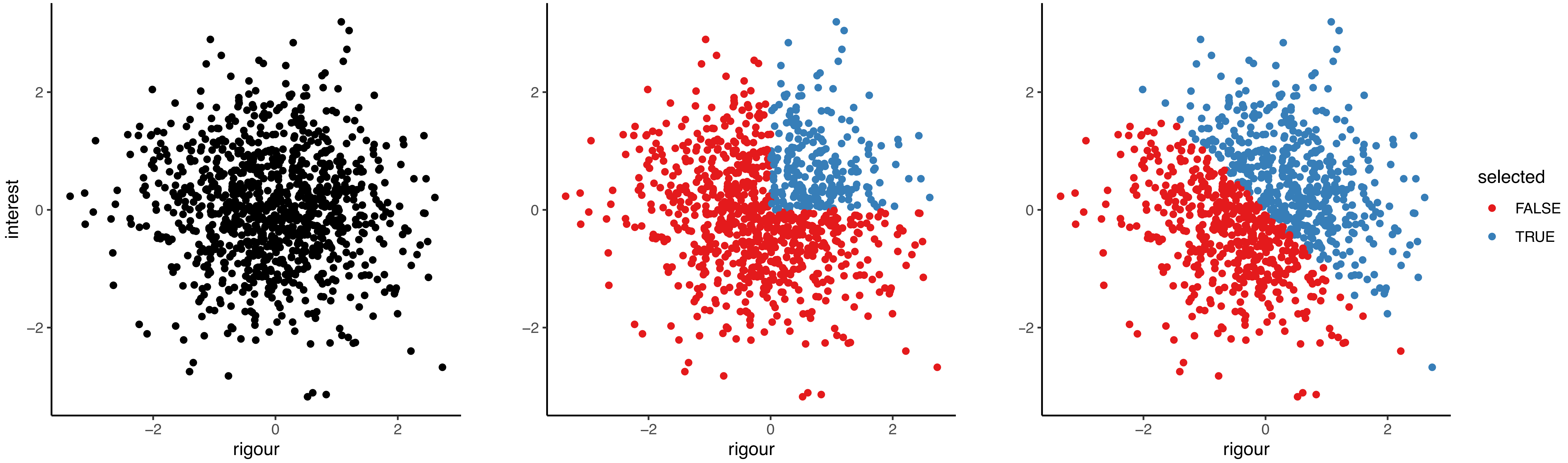
Alternative way of adjusting for the confounder: stratify by it

‘Box’ a variable to indicate adjusting for it somehow

Carrying a lighter is associated with an outcome of lung cancer. The association disappears when you condition on being a smoker, which is a common cause.

Conditioning on a study being covered in the popular press may induce a non-causal correlation between its rigour and public interest, depending on how those are combined for selection.





Examples credit: *Causal Inference: What If* and *Statistical Rethinking*