Causality

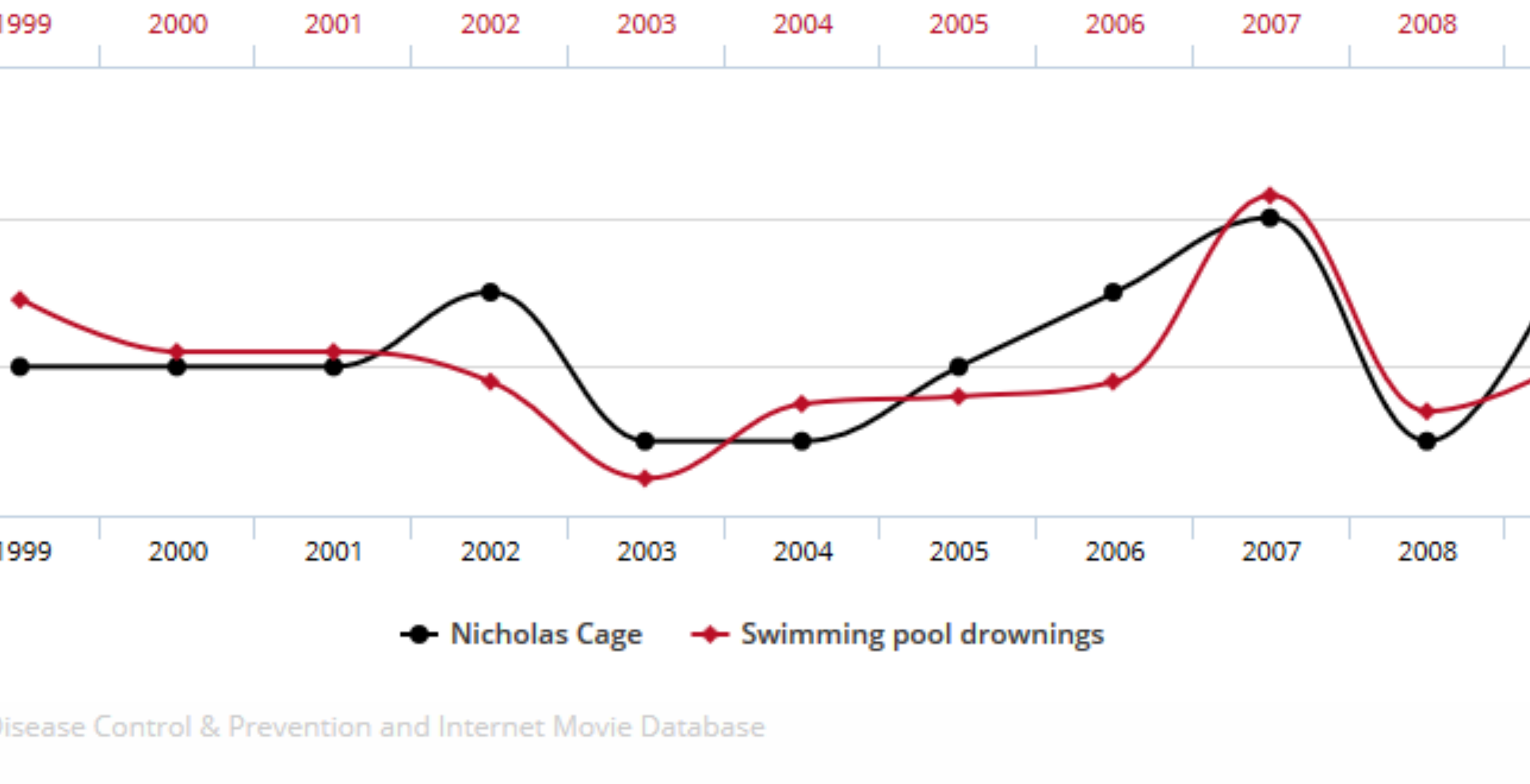
Julian Faraway

Number of people who drowned by falling into a pool

correlates wtih

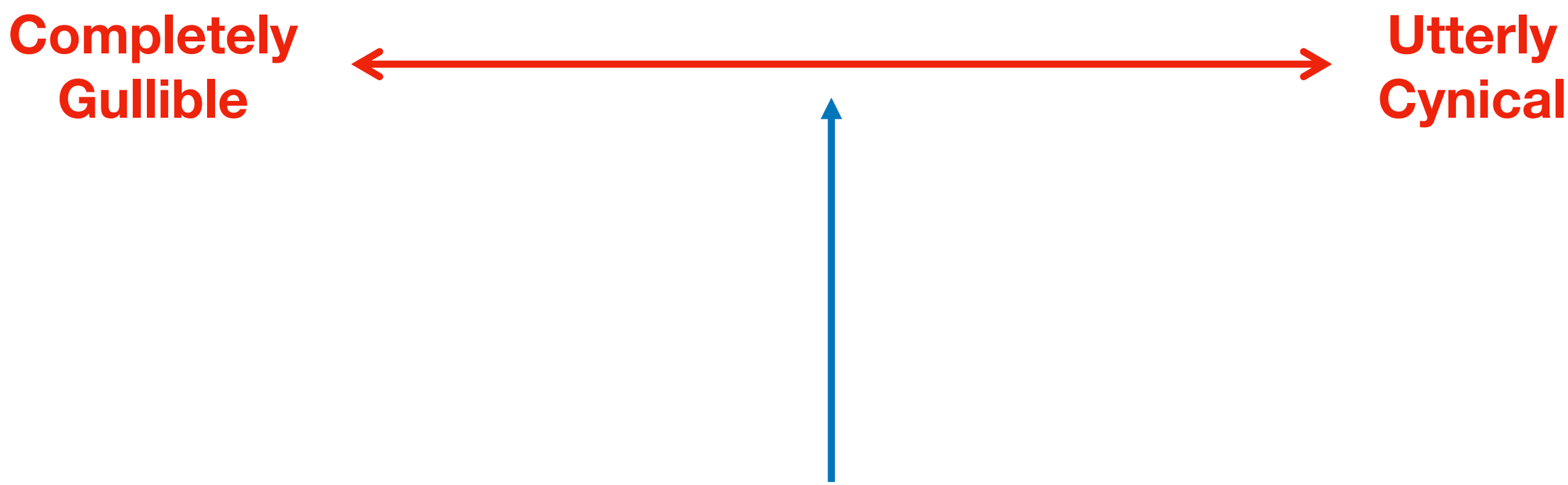
Films Nicolas Cage appeared in

Correlation: 66.6% (r=0.666004, p>0.05)



Disease Control & Prevention and Internet Movie Database

Correlation is not Causation



David Hume

Enquiry Concerning
Human Understanding

1748

introduces the idea of the
counterfactual

An object is the cause of another..

“if the first object had not been,
the second never had existed”

Action → **Outcome**

Counterfactual Action → **Counterfactual Outcome**

Difference is the effect of the action

High Air Pollution in Ulaanbaatar

Poor Health outcomes

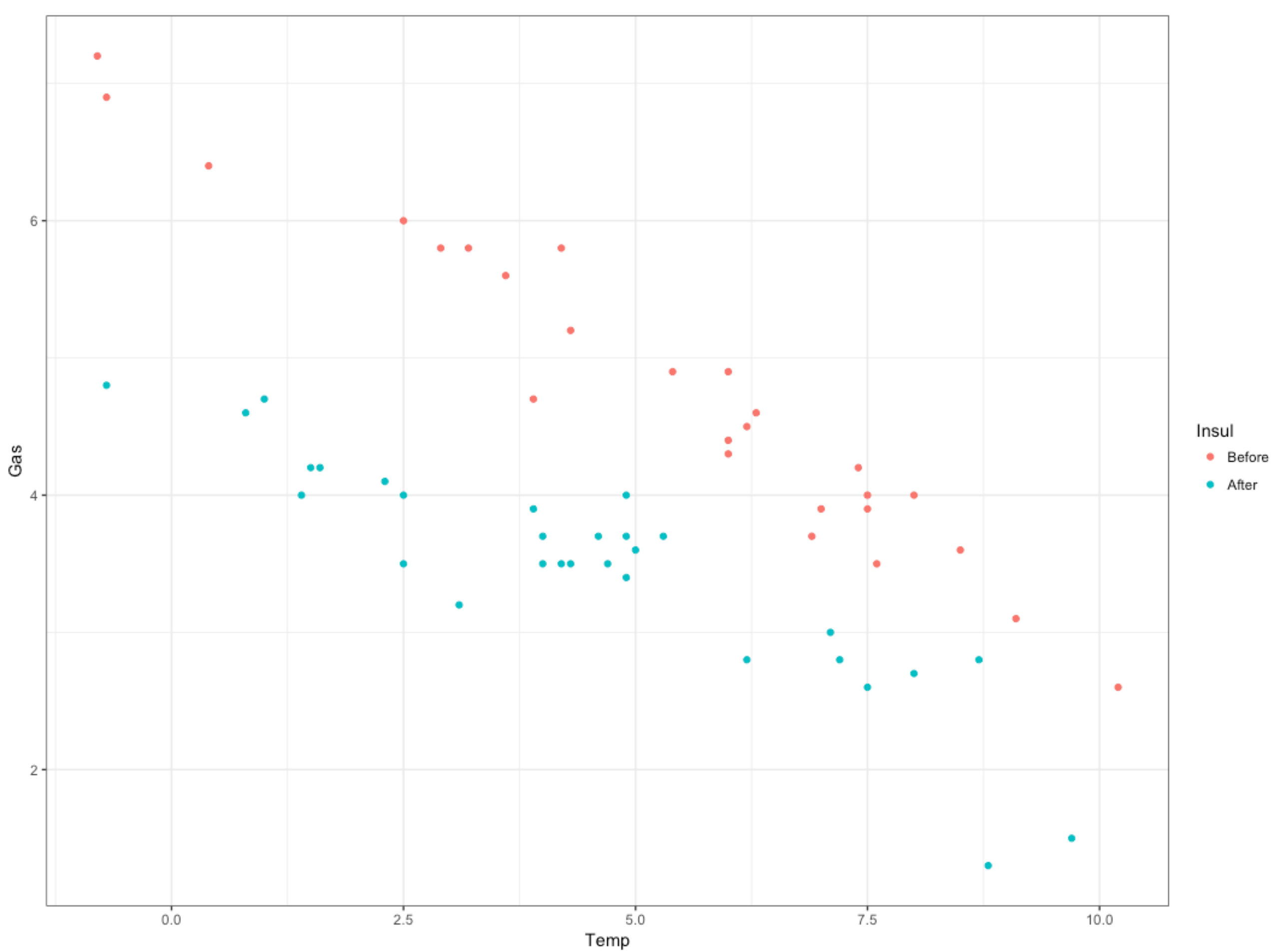
Difference?

Lower Air Pollution in Ulaanbaatar

Better Health Outcomes?

Problems with Counterfactuals

- Never get to observe
- Can the counterfactual action/decision/assignment potentially occur? What if I bought a new car instead of repairing the old one? What if I were the other gender?
- Counterfactual approximations in space and time.
  - Air Pollution/Health in Canadian Cities?
  - Air Pollution/Health in Ulaanbaatar in the past?


Gas
6
4
2
0.0
2.5
5.0
7.5
10.0
Temp
Insul
Before
After

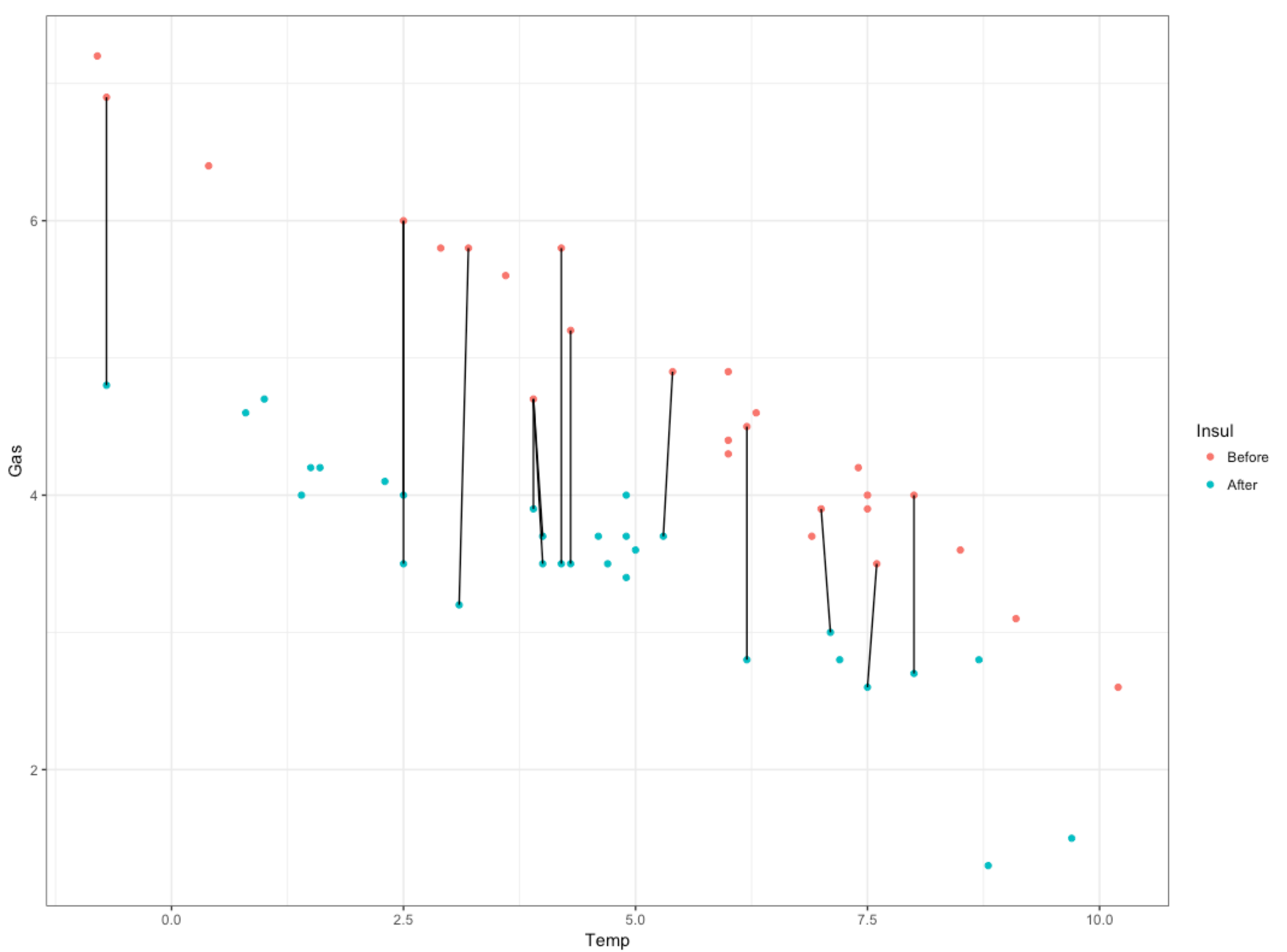
Matching

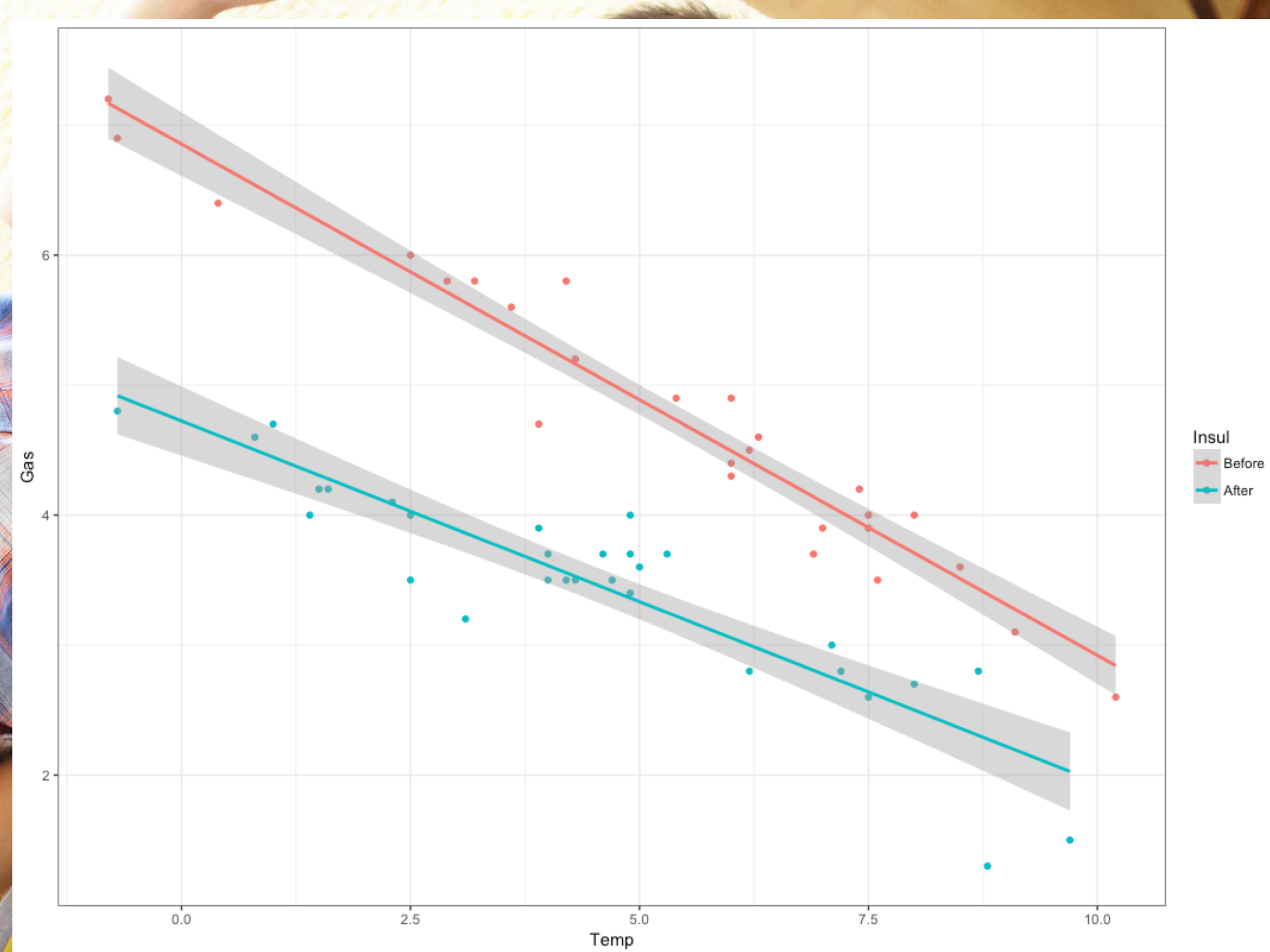
- Identify pairs of observations that are identical/similar in the covariates but have different treatments.

- Attempts to approximate the counterfactual outcome

- Exact matching can be difficult where observations are multivariate – *propensity score* matching can help

- Unobserved characteristics of the pair may be different in a consistent manner thus voiding the causal conclusion


Gas
6
4
2
0.0
2.5
5.0
7.5
10.0
Temp
Insul
Before
After


Gas
6
4
2
0.0
2.5
5.0
7.5
10.0
Temp
Insul
Before
After

(Linear) Modelling

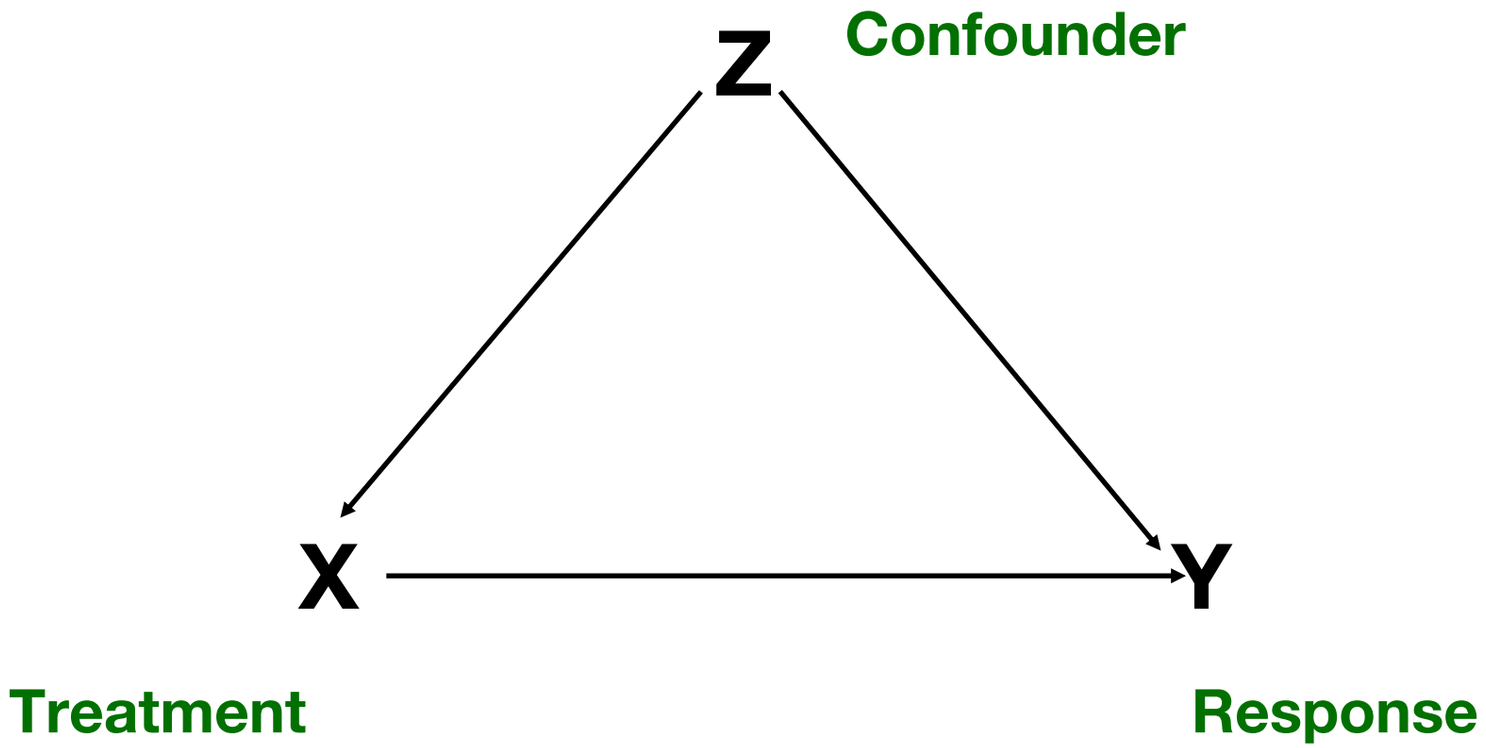
- Difference between the treatments is expressed as parameter(s) in the model.

- Avoids the need for matching

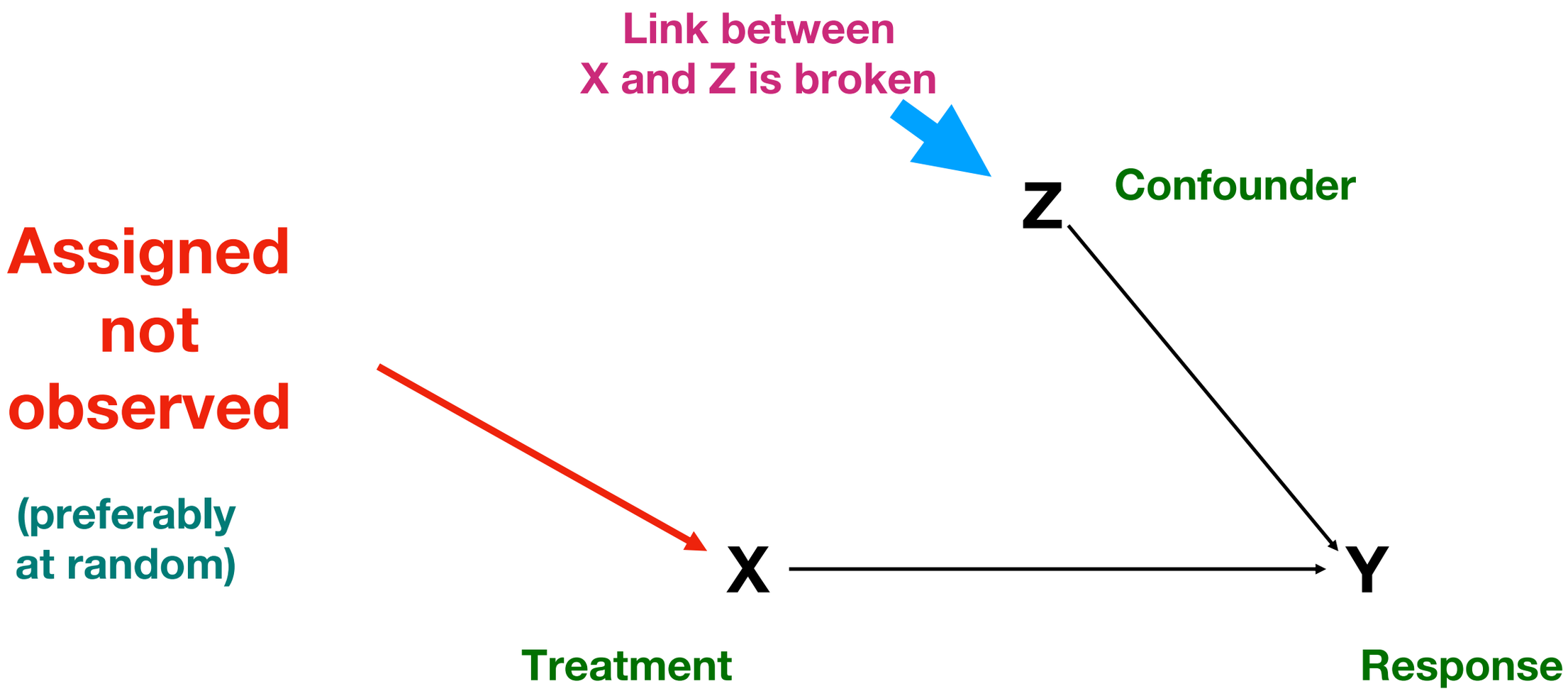
- Requires additional assumptions

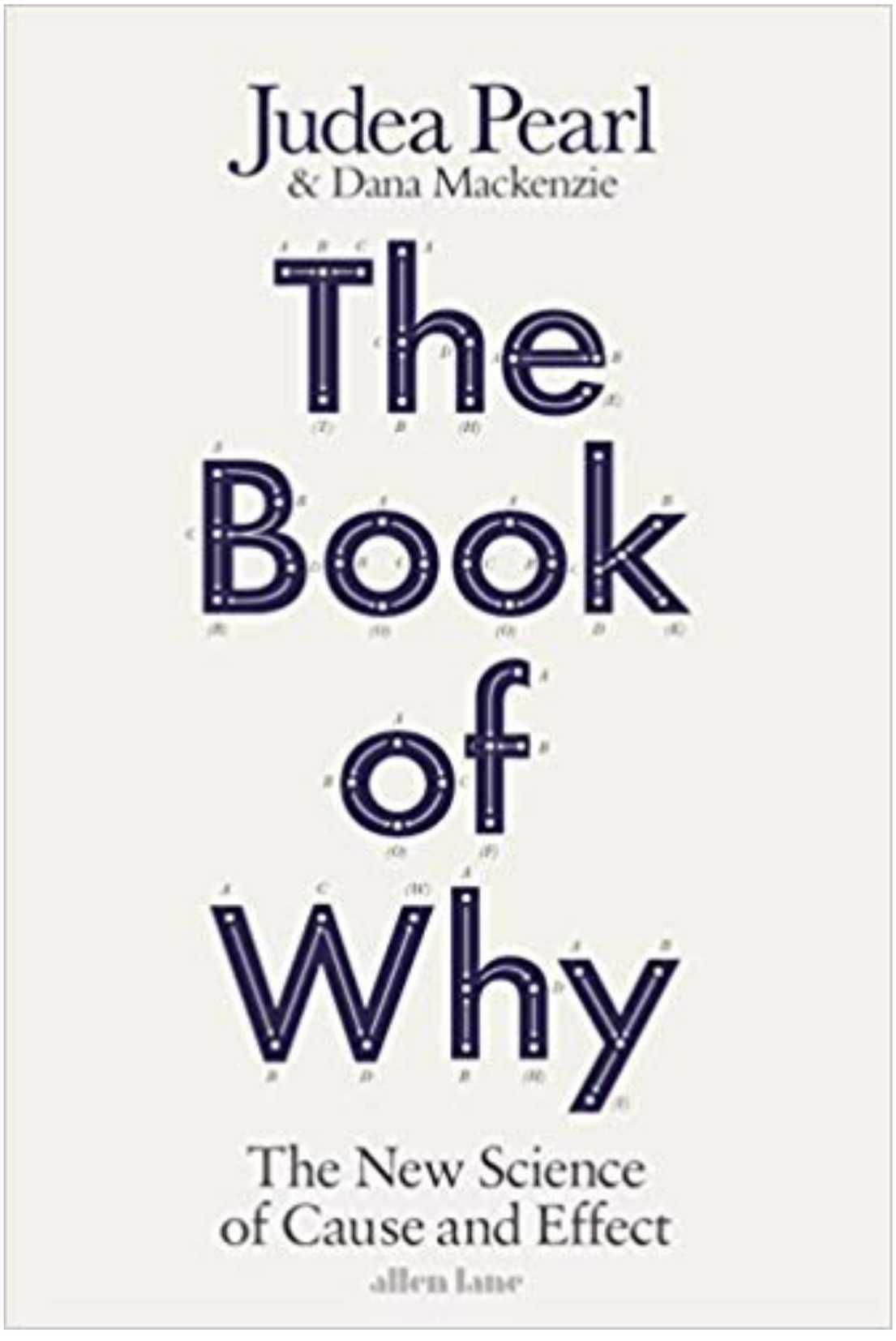
- Non-treatment predictors are called *confounders.* Including such variables in the model is called *adjusting for the confounder.*

Graphical Models



Experiment


Link between
X and Z is broken
Z
Confounder
Assigned
not
observed
(preferably
at random)
X
Y
Treatment
Response


Judea Pearl
& Dana Mackenzie
The Book of Why
The New Science of Cause and Effect
allen lane

Statistical/Probabilistic reasoning alone cannot support causal inference

Determining the joint probability distribution of variables alone says nothing about causation

P(Disease | Symptom)

Pearl promises to determine the necessary set of non-data assumption that are sufficient to make a causal conclusion

Bradford Hill Criteria

Strength

Consistency

Specificity

Temporality

Biological Gradient

Plausibility

Coherence

Experiment

Analogy

Qualitative and somewhat specific to epidemiology