Causal Inference: An Introduction

Qingyuan Zhao

Statistical Laboratory, University of Cambridge

4th March, 2020 @ Social Sciences Research Methods Programme (SSRMP), University of Cambridge

Slides and more information are available at
`http://www.statslab.cam.ac.uk/~qz280/`.

About this lecture

About me

2019 –	University Lecturer in the Statistical Laboratory (in Centre for Mathematical Sciences, West Cambridge).
2016 – 2019	Postdoc: Wharton School, University of Pennsylvania.
2011 – 2016	PhD in Statistics: Stanford University.

Disclaimer

- I am a statistician who work on causal inference, but not a social scientist.
- **Bad news:** What's in this lecture may not reflect the current practice of causal inference in social sciences.
- **Good news (hopefully):** What's in this lecture will provide you an up-to-date view on the **design**, **methodology**, and **interpretation** of causal inference (especially observational studies).
- I tried to make the materials as accessible as possible, but some amount of maths seemed inevitable. Please bear with me and don't hesitate to ask questions.

Growing interest in causal inference

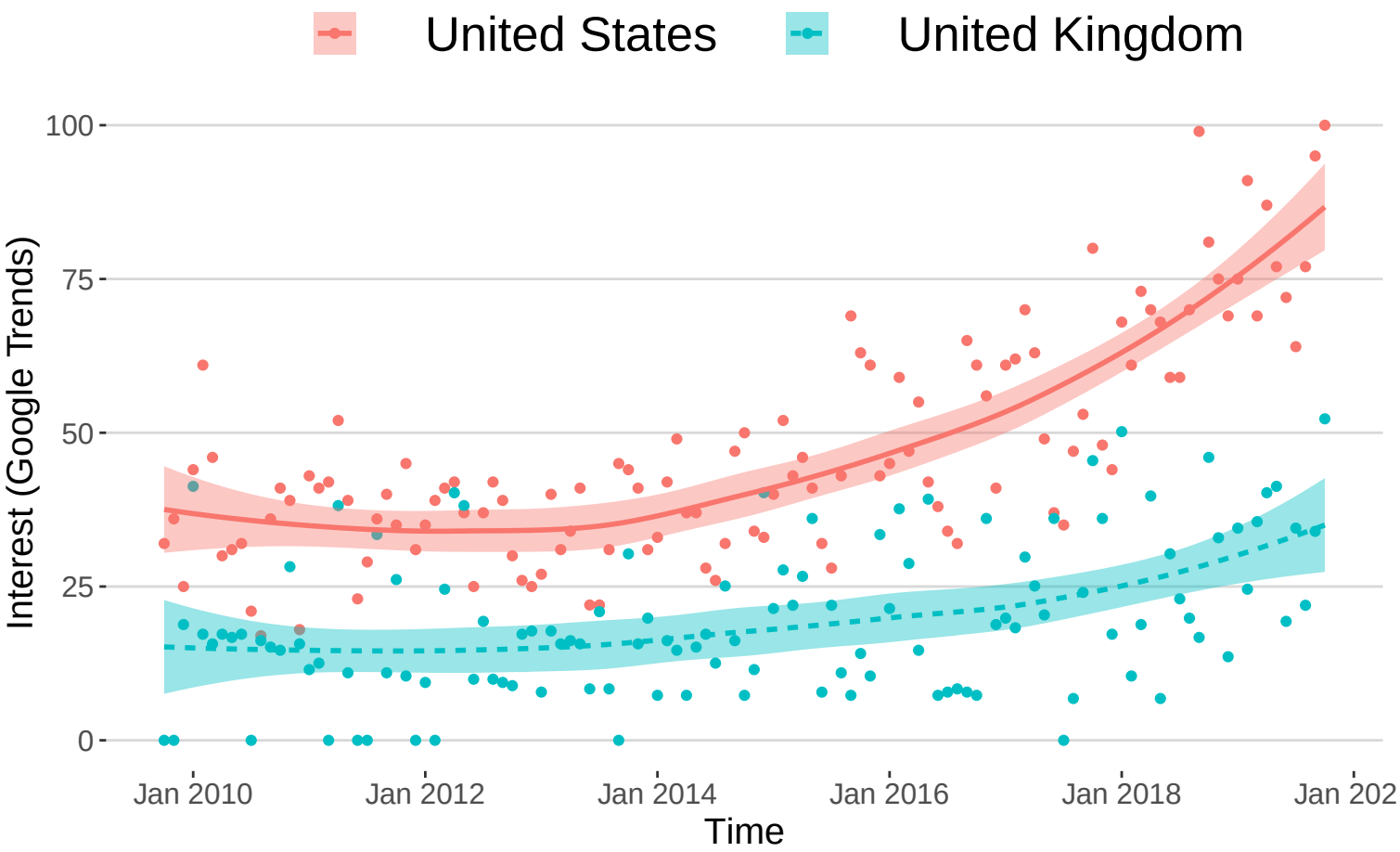


Figure: Data from Google Trends.

A diverse field

Causal inference is **driven by applications** and is **at the core of statistics** (*the science of using information discovered from collecting, organising, and studying numbers*—Cambridge Dictionary).

Many origins of causal inference

- Biology and genetics;
- Agriculture;
- Epidemiology, public health, and medicine;
- Economics, education, psychology, and other social sciences;
- Artificial intelligence and computer science;
- Management and business.

In the last decade, independent developments in these disciplines have been merging into a single field called "Causal Inference".

Examples in social sciences

1. Economics: How does supply and demand (causally) depend on price?
2. Policy: Are job training programmes actually effective?
3. Education: Does learning "mindset" affect academic achievements?
4. Law: Is it justifiable to sue the factory over injuries due to poor working conditions?
5. Psychology: What is the effect of family structure on children's outcome?

Outline for this lecture

To study causal relationships, empirical studies can be categorised into

Randomised Experiments (Part I)

1. Completely randomised;
2. Stratified (pairs or blocks);
3. With regression adjustment (also called covariance adjustment)?
4. More sophisticated designs (e.g. sequential experiments).

$\downarrow\downarrow$ Question: How to define causality? (Part II) $\downarrow\downarrow$

Observational Studies (Part III)

Also called quasi-experiments in social sciences (I think it's a poor name).

1. Controlling for confounders;
2. Instrumental variables;
3. Regression discontinuity design;
4. Negative control (e.g. difference in differences).

Part I: Randomised experiments

The breakthrough

- The idea of randomised experiments dates back to the early development of experimental psychology in the late 1800s by Charles Sanders Peirce (American philosopher).
- In 1920s, Sir Ronald Fisher established randomisation as a principled way for causal inference in scientific research (*The Design of Experiments*, 1935).

Fundamental logic*

1. Suppose we let half of the participants to receive the treatment **at random**,
2. If **significantly more** treated participants have better outcome,
3. Then the treatment **must be** beneficial.

Randomisation (1) $\Longrightarrow$ a choice of statistical error (2) vs. causality (3).
(because there can be no other logical explanations)

*We will revisit this logic when moving to observational studies.

Randomisation

Some notations

- A is treatment (e.g. job training), for now let A be binary (0=control, 1=treated);
- Y is outcome (e.g. employment status 6 months after job training).
- X is a vector of covariates **measured before the treatment** (e.g. gender, education, income, ...).
- Subscript $i = 1, \ldots, n$ indexes the study participants.

Different designs of randomised experiments

- **Bernoulli trial**: $A_1, \ldots, A_n$ independent and $\mathbb{P}(A_i = 1) = 0.2$.
- **Completely randomised**:

$$\mathbb{P}(A_1 = a_1, \ldots, A_n = a_n) = \binom{n}{n/2}^{-1} \text{ if } a_1 + \cdots + a_n = n/2.$$

- **Stratified**: $A_1, \ldots, A_n$ independent, $\mathbb{P}(A_i = 1 \mid X_i) = \pi(X_i)$ where $\pi(\cdot)$ is a given function. For example:

$$\mathbb{P}(A_i = 1 \mid X_{i1} = \text{male}) = 0.5 \text{ and } \mathbb{P}(A_i = 1 \mid X_{i1} = \text{female}) = 0.75.$$

- **Blocked**: Completely randomised within each block of participants similar in X.

Statistical inference: Approach 1

Randomisation inference (permutation test)

Test the hypothesis $H_0 : A \perp\!\!\!\perp Y \mid X$ (or $H_0 : A \perp\!\!\!\perp Y$ if randomisation does not depend on X).

1. Choose a **test statistic** $T(X, A, Y)$ (e.g. in a blocked experiment with matched pairs, the average pairwise treated-minus-control difference in Y).
2. Obtain the **randomisation distribution** of $T(X, A, Y)$ by permuting A, according to how it was randomised.
3. Compute the **p-value**:

$$\mathbb{P}_{A \sim \pi}\Big(T(X, A, Y) \geq T(X, A^{\text{obs}}, Y) \mid X, Y \Big).$$

Note that the randomisation inference treats X and Y as given and **only considers randomness in the treatment** $A \sim \pi$ (which is exactly the randomness introduced by the experimenter).

Statistical inference: Approach 2

Regression analysis

- Simplest form:
$$\mathbb{E}[Y|A] = \alpha + \beta A.$$
- Regression adjustment (also called covariance adjustment):
$$\mathbb{E}[Y|A, X] = \alpha + \beta A + \gamma X + \delta AX.$$
- More complex mixed-effect models, to account for heterogeneity of the participants.

Interpretation of regression analysis

- Slope coefficient β of the treatment A in these regression models is usually interpreted as the **average treatment effect**, although this becomes difficult to justify in complex designs/regression models.
- To differentiate from **structural equation models**, regression models were written in the form of $\mathbb{E}[Y|A] = \alpha + \beta A$ instead of the "traditional" form $Y = \alpha + \beta A + \epsilon$. We will explain their differences later.

Comparison of the two approaches

Randomisation inference

Advantages:

1. Only uses randomness in the design.
2. Distribution-free and exact finite-sample test.

Disadvantages:

1. Only gives a hypothesis test for "no treatment effect whatsoever" (can be extended to constant treatment effect).

Regression analysis

Advantages:

1. Account for treatment effect heterogeneity.
2. Well-developed extensions: mixed-effect models, generalised linear models, Cox proportional-hazards models, etc.

Disadvantages:

1. Inference usually relies on normality or large-sample approximations.
2. Causal interpretation is model-dependent!

Internal vs. external validity

Internal validity

- Campbell and Stanley (1963): "Whether the experimental treatments make a difference in this specific experimental instance".
- Exactly what randomisation inference tries to do.

External validity

- Shadish, Cook and Campbell (2002): "Whether the cause-effect relationship holds over variation in persons, settings, treatment variables, and measurement variables".

Related concepts

- Another important concept in social sciences is **construct validity**: "the validity if inferences about the higher order constructs that represent sampling particulars". See Shadish et al. (2002) for more discussion.
- Perice's three kinds of inferences: deduction, induction, abduction.

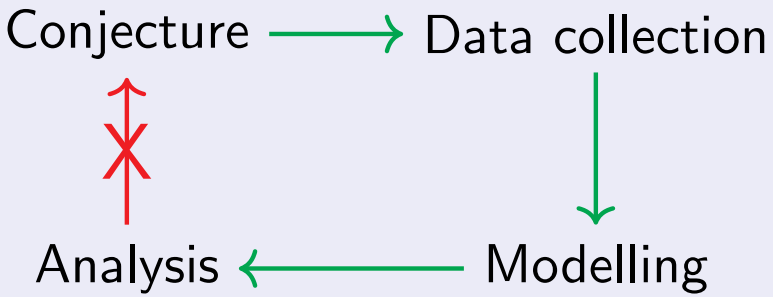
How causal inference became irrelevant

The narrow-minded view of causality

- **"Correlation does not imply causation"**
- $\Longrightarrow$ **Causality can only be established by randomised experiments**
- $\Longrightarrow$ Causal inference became absent in statistics until 1980s.
- Example: "Use of Causal Language" in the author guidelines of *JAMA*:

> **Causal language** *(including use of terms such as effect and efficacy)* **should be used only for randomised clinical trials**. *For all other study designs, methods and results should be described in terms of association or correlation and should avoid cause-and-effect wording.*

Broken cycle of statistical research



"Clouds" over randomised experiments

(Borrowing the metaphor from the famous 1900 speech by Kelvin.)

Smoking and Lung cancer (1950s)

- Hill, Doll and others: **Overwhelming association** between smoking and lung cancer, **in many populations**, and **after conditioning on many variables**.
- Fisher and other statisticians: **But correlation is not causation.**

Infeasibility of randomised experiments

- Ethical problems, high cost, and many other reasons.

Non-compliance

- People may not comply with assigned treatment or drop out during the study.

$\Longrightarrow$ **Need for causal inference from observational data.**

Part II: How to define causality?

Definition 0: Implicitly from randomisation

Recall the logic of randomised experiment:

1. Suppose we let half of the participants to receive the treatment **at random**,
2. If **significantly more** treated participants have better outcome,
3. Then the treatment **must be** beneficial (because there can be no other logical explanation).

Randomisation (1) $\Longrightarrow$ a choice of statistical error (2) vs. causality (3). (because there can be no other logical explanations)

For observational studies, we need a definition of causality that **does not hinge on (explicit) randomisation**.

Pioneers in causal inference have come up with three definitions/languages:

1. Counterfactual (also called potential outcome);
2. Causal graphical model;
3. Structural equation model.

Part II: How to define causality?

Definition 1: Counterfactuals (Neyman, 1923; Rubin, 1974)

- Participants have two **counterfactuals**, $Y(0)$ and $Y(1)$.
- We **only observe one counterfactual** (in any study, randomised or not),

$$Y = Y(A) = \begin{cases} Y(1), & \text{if } A = 1, \\ Y(0), & \text{if } A = 0. \end{cases}$$

i	$Y_i(0)$	$Y_i(1)$	A_i	Y_i
1	-3.7	?	0	-3.7
2	2.3	?	0	2.3
3	?	7.4	1	7.4
4	0.8	?	0	0.8
⋮	⋮	⋮	⋮	⋮

- Rubin calls this the **"science table"** (I didn't find this terminology useful).
- The goal of causal inference is to infer the difference

$$\text{Distribution of } Y(0) \text{ vs. Distribution of } Y(1).$$

- Example: Average treatment effect is defined as $\mathbb{E}[Y(1) - Y(0)]$.

Part II: How to define causality?

Definition 1: Counterfactuals (Neyman, 1923; Rubin, 1974)

- We would like to infer about the difference between

$$\text{Distribution of } Y(0) \text{ vs. Distribution of } Y(1).$$

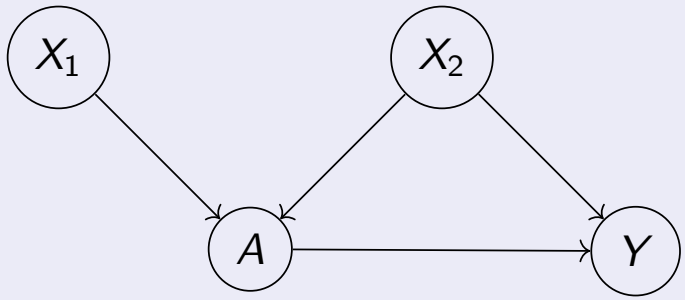
- How is this possible? If we know $A \perp\!\!\!\perp Y(0) \mid X$, then

$$\begin{aligned}\mathbb{P}(Y(0)=y) &= \mathbb{E}[\mathbb{P}(Y(0)=y \mid X)] \\ &= \mathbb{E}[\mathbb{P}(Y(0)=y \mid A=0, X)] \\ &= \mathbb{E}[\mathbb{P}(Y=y \mid A=0, X)]\end{aligned}$$

- **Remark 1:** The above derivation is called **causal identification**.
- **Remark 2:** In the literature, the key assumption $A \perp\!\!\!\perp Y(0) \mid X$ is called **"randomisation"**, **"ignorability"**, or **"no unmeasured confounders"**.
- **Remark 3:** An synonym for counterfactual is **potential outcome**. I like to use **potential outcome** for randomised experiments (looking forward) and **counterfactual** for observational studies (looking backward).

Part II: How to define causality?

Definition 2: Graphical models



- **Probabilistic graphical models**/Bayesian networks (Pearl, 1985; Lauritzen, 1996): Joint distribution factorises according to the graph:

$$\begin{aligned}&\mathbb{P}(X_1 = x, X_2 = x, A = a, Y = y)\\ =&\mathbb{P}(X_1 = x_1, X_2 = x_2)\,\mathbb{P}(A = a \mid X_1 = x_1, X_2 = x_2)\,\mathbb{P}(Y = y \mid X_2 = x_2, A = a).\end{aligned}$$

- We can obtain **conditional independence** between the variables by applying the **d-separation criterion** (details omitted; imagine information flowing like water).
- Examples: $Y \perp\!\!\!\perp X_1 \mid A$; $X_1 \perp\!\!\!\perp X_2$ but $X_1 \not\perp\!\!\!\perp X_2 \mid A$ (this is called **collider bias**).

How to define causality?

Definition 2: Graphical models

- **Causal graphical models** (Robins, 1986; Spirtes et al., 1993; Pearl, 2000): Joint distribution in interventional settings also described by the graph:

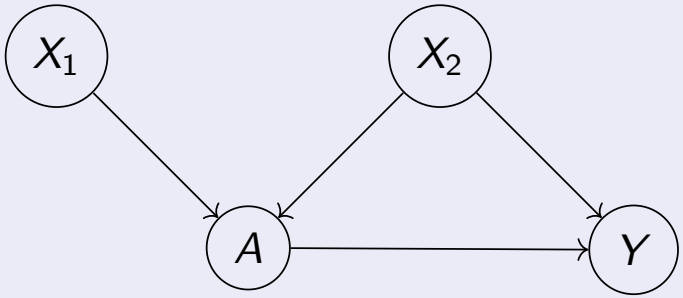
$$\begin{aligned}&\mathbb{P}(X_1 = x_1, X_2 = x_2, A = a, Y(a) = y)\\ =&\mathbb{P}(X_1 = x_1, X_2 = x_2)\,\mathbb{P}(A = a \mid X_1 = x_1, X_2 = x_2)\,\mathbb{P}(Y(a) = y \mid X_2 = x_2).\end{aligned}$$

- **Remark:** Computer scientists use the **do** notation introduced by Pearl:

$$\mathbb{P}(Y = y \mid \mathbf{do}(A = a)) = \mathbb{P}(Y(a) = y).$$

How to define causality?

Definition 3: Structural equations (Wright, 1920s; Haavelmo, 1940s)



- From the graph we may define a set of structural equations:

$$\begin{aligned} X_1 &= f_{X_1}(\epsilon_{X_1}), \\ X_2 &= f_{X_2}(\epsilon_{X_2}), \\ A &= f_A(X_1, X_2, \epsilon_A), \\ Y &= f_Y(A, X_2, \epsilon_Y). \end{aligned}$$

- Parameters in the structural equations are **causal effects**. For example, if $f_Y(A, X_2, \epsilon_Y) = \beta_{AY} A + \beta_{XY} X_2 + \epsilon_Y$, then β_{AY} is the causal effect of A on Y.
- **Remark:** Structural equations are different from regressions that only model the conditional expectation $\mathbb{E}[Y \mid A, X]$.

Unification of the definitions

Define counterfactual from graphs

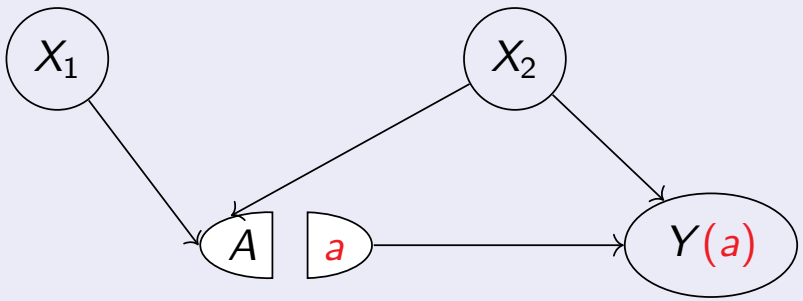
- Structural equations are **structural** instead of **regression** because they also govern the interventional settings (Pearl, 2000):

$$Y(a) = F_Y(a, X, \epsilon_Y).$$

- That is, $Y(0) = F_Y(0, X, \epsilon_Y)$ and $Y(1) = F_Y(1, X, \epsilon_Y)$ share the randomness in X and ϵ_Y.

Single-world intervention graphs (Richardson and Robins, 2013)

- Distribution of counterfactuals factorises according to an extended graph (obtained by splitting and relabelling the nodes).



- Apply the d-separation, we get $Y(a) \perp\!\!\!\perp A \mid X_2$ (and also $Y(a) \perp\!\!\!\perp A \mid X_1, X_2$).

Recap

"Equivalence" of the definitions of causality

Graphical models

$\rightarrow$ Define structural equations

$\rightarrow$ Define counterfactuals

$\rightarrow$ Embed in extended graph.

Strengths of the different approaches

- Graphical model: Good for understanding the scientific problems.
- Structural equations: Good for fitting simultaneous models for the variables (especially for abstract constructs in social sciences).
- Counterfactuals: Good for articulating the inference for a small number of causes and effects.

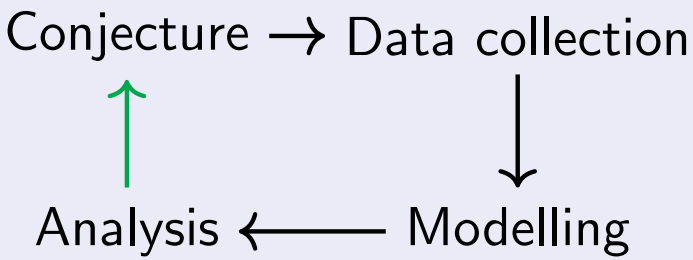
Modern causal inference

Logic of randomised experiment

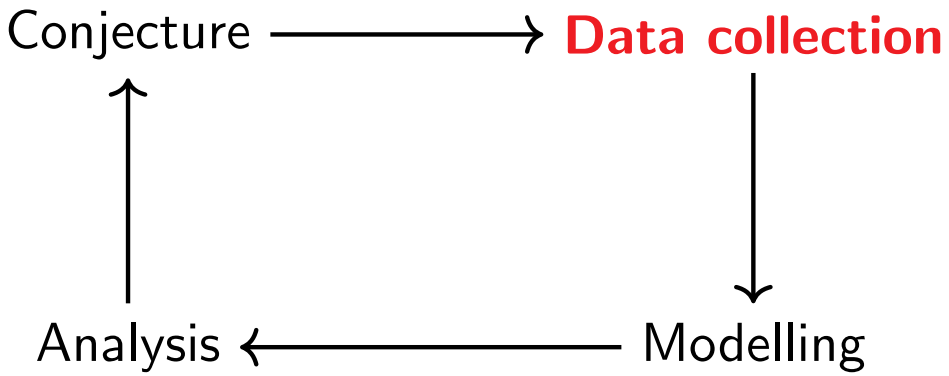
Randomisation (1) $\Longrightarrow$ a choice of statistical error (2) vs. causality (3).

Logic of observational studies

- View randomisation as **a breakable identification assumption**.
  - Examples: need to use pseudo-RNGs; non-compliance and missing data.
- Causal inference from observational studies becomes a choice between
  1. Identification and modelling assumptions being violated;
  2. Statistical error;
  3. True causality.
- Causal inference is **abductive** (inference to the best explanation).
  - Strength of causal inference = credibility of the assumptions.
- Cycle of statistical research is restored:



Part III: Designing observational studies



Study design = How data are collected in a study.

- This is slightly different from the traditional notion of **experimental design** (often about how to minimise the statistical error in **a regression analysis**).
- In modern causal inference, study design refers to **how data are collected to meet the identification assumption (independent of analysis)**.
  - Common designs in observational studies: controlling for confounders, instrumental variables, regression discontinuity, difference-in-differences.

Design trumps analysis (Rubin, 2008)

Logic of observational studies

Causal inference from observational studies becomes a choice between

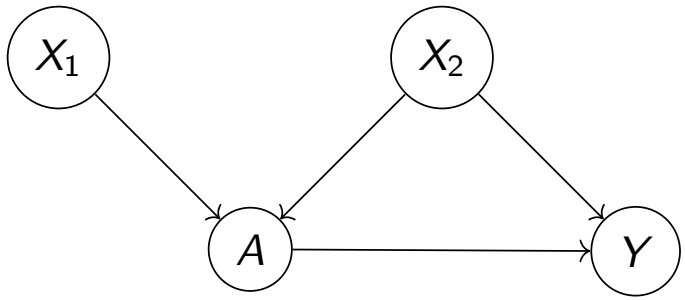
1. Identification and modelling assumptions being violated;
2. Statistical error;
3. True causality.

A decomposition of estimation error (Zhao, Keele, and Small, 2019)

$$\begin{aligned}&\text{Causal estimator} - \text{True causal effect}\\ =\ &\textbf{Design bias} + \text{Modelling bias} + \text{Statistical noise}.\end{aligned}$$

- The first term (**Design bias**) is fixed once we decide how to collect data.
- The last two terms resemble the familiar bias-variance trade-off in statistics. We can hope to make it small by using better statistical methods and or having a large sample.
- $\Longrightarrow$ **Design** $\gg$ **Modelling** $>$ **Analysis**.

Design 1: Controlling for confounders



- Loosely speaking, confounders are common causal ancestors of the treatment and the outcome (for example, X_2 in the above graph).

Identifying assumption: No unmeasured confounders

In counterfactual terms: $Y(0) \perp\!\!\!\perp A \mid X$ and $Y(1) \perp\!\!\!\perp A \mid X$ for measured X.

- In the above example, this would hold if $X = X_2$ or $X = (X_1, X_2)$. It would not hold if $X = X_2$ and there is another U_3 affecting both A and Y directly.
- This can be checked using the **single-world intervention graphs**.
- This assumption is also called **ignorability**, **exogeneity**, **unconfoundedness**, **selection on observables**, etc.

Which covariates should be controlled for?

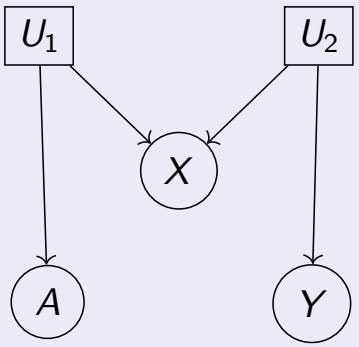
Counterfactualists: Measuring pre-treatment covariate always helps

- Rubin (2009), replying to Pearl and others:

> *I* **cannot think of a credible real-life situation** *where I would intentionally allow substantially different observed distributions of a true covariate in the treatment and control groups.*

- Logic: observational studies should try to **mimic** randomised experiments.

Graphists: Counterexample (M-bias)



- X is measured, U_1 and U_2 are unmeasured, all temporally precede A.
- Conditioning on X introduces **spurious association** between A and Y.

This debate is still ongoing. **My take: measure as many covariates as possible, but think about if any would introduce bias via the M-structure.**

Statistical methods: Approach 1

Create a **pseudo-population** to mimic randomised experiment

- **Matching:** Create pairs of treated and control participants with similar pre-treatment characteristics (in terms of the covariates X).
  - Many algorithms: nearest-neighbour matching, Mahalanobis distance matching, optimal matching, etc.
- **Propensity-score matching:** Match on the (estimated) propensity score $\pi(X) = \mathbb{P}(A = 1 \mid X)$ to reduce the dimensionality.
- **Stratification:** Create strata/blocks in terms of X or $\pi(X)$. Treat participants within a stratum/block as randomised.
- **Weighting:** Weight the participants by the inverse of the probability of receiving the observed treatment.
  - That is, weight participant i by $\frac{1}{\pi(X_i)}$ if $A_i = 1$ (treated) and by $\frac{1}{1 - \pi(X_i)}$ if $A_i = 0$ (control).

Randomisation inference or regression analysis (for randomised experiments) can then be applied to the pseudo-population.

Statistical methods: Approach 2

Outcome regression (also called standardisation)

Recall that if $A \perp\!\!\!\perp Y(0) \mid X$, then

$$\mathbb{E}[Y(0)] = \mathbb{E}[\mathbb{E}(Y(0) \mid X)] = \mathbb{E}[\mathbb{E}[Y(0) \mid A = 0, X]] = \mathbb{E}[\mathbb{E}[Y \mid A = 0, X]].$$

Two steps to estimate $\mathbb{E}[Y(0)]$ (average counterfactual under control):

- Estimate $\mathbb{E}[Y \mid A = 0, X]$ by regression **using control participants**.
- **Average** the predicted $\mathbb{E}[Y \mid A = 0, X]$ **over all participants**.

We can do the same thing to estimate $\mathbb{E}[Y(1)]$ and take the difference to estimate $\mathbb{E}[Y(1) - Y(0)]$ (average treatment effect).

Statistical methods: Which one to use?

- Both approaches are better than the "standard" regression (e.g. $Y = \alpha + \beta A + \gamma X + \epsilon$), because interpreting the results of the "standard" regression **requires that we correctly specify the structural equation**.
- Both approaches are **semiparametric** in the sense that the "nuisance parameters" $\pi(X)$ and $\mathbb{E}[Y \mid A = 0, X]$ can be estimated nonparametrically.

More complicated methods

- State-of-the-art: estimate $\pi(X)$ and $\mathbb{E}[Y \mid A = 0, X]$ using **machine learning** and then combine them in a **"doubly robust"** estimator.
- What they are trying to do is to minimise the **"Modelling bias"**:

$$\begin{aligned} &\text{Causal estimator} - \text{True causal effect} \\ &= \text{Design bias} + \textbf{Modelling bias} + \text{Statistical noise}. \end{aligned}$$

- My take: Too much sophistication not really necessary in "normal" applications. Save your time for study design and data collection. Choose the method you are most comfortable with.

Another key assumption

Overlap assumption (also called positivity)

- A key assumption that was implicit in the above discussion is:

$$0 < \pi(x) = \mathbb{P}(A = 1 \mid X = x) < 1, \text{ for all } x.$$

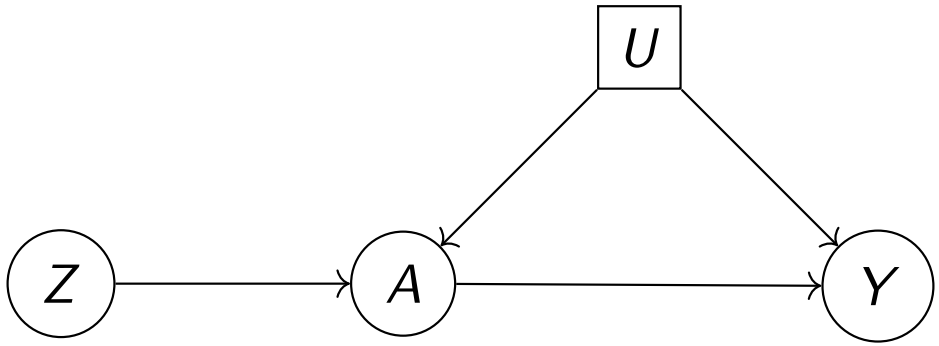
- This means that the treated participants and control participants have **overlapping X distributions**.
- In other words, any study participant have at least some chance of receiving treatment (or control).

- You should always check the overlap assumption and define your study population accordingly (e.g. by comparing histograms).
- Matching methods are helpful in this regard, because you can examine whether the matched participants are indeed similar.

Recap

- Study designs discussed so far assume **no unmeasured confounders**
  - Either by randomisation in **randomised experiments**;
  - Or by treating it as an explcit assumption in **observational studies**.
- Next: Other observational study designs that try to remove or reduce bias due to unmeasured confounders.

Design 2: Instrumental variables

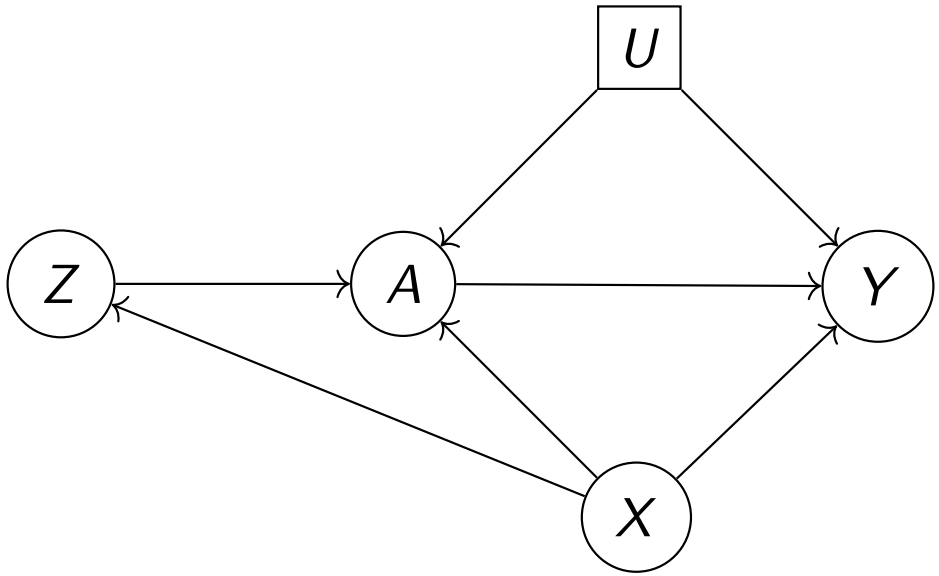


- Z is an instrumental variable (IV); U is unmeasured confounder.
- Idea: use exogenous (or unconfounded) randomness in A.

Examples of IV

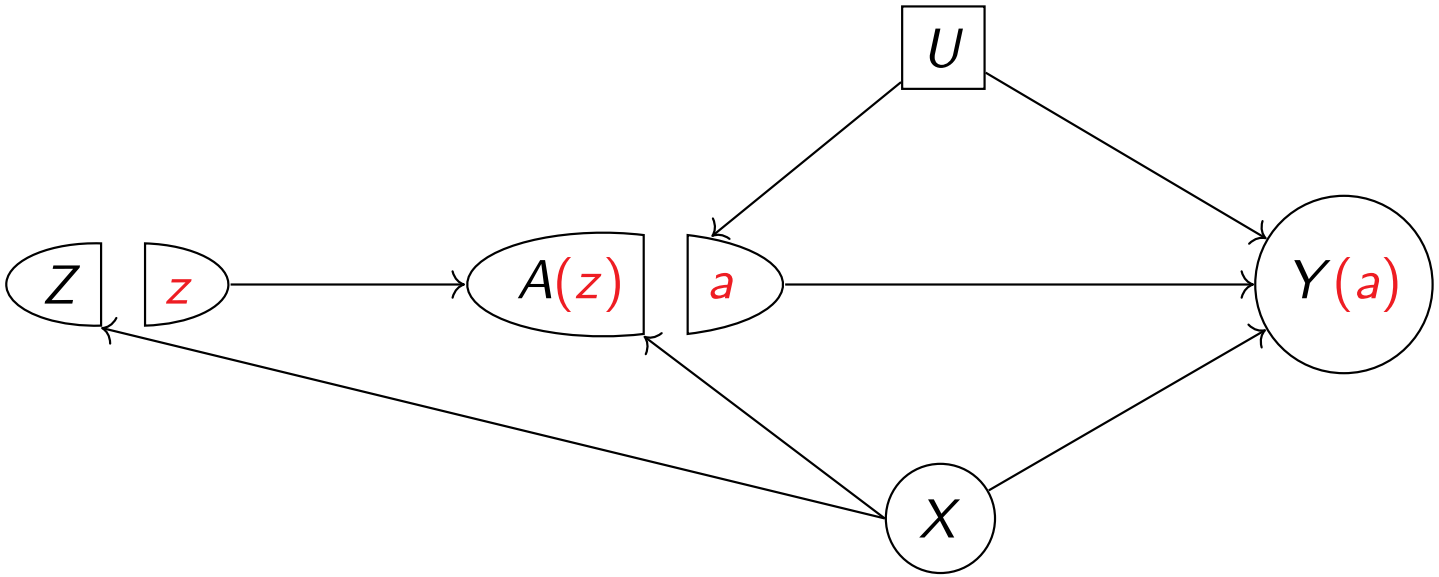
- Draft lottery for Vietnam war (treatment: military service).
- Distance to closest college (treatment: college education).
- Favourable growing condition for crops (treatment: market price, outcome: market demand).
- Randomised cash incentive to quit smoking (treatment: quit smoking).
- Randomised treatment assignment (treatment: actual treatment received, could be different to the IV due to non-compliance).

Assumptions for instrumental variables



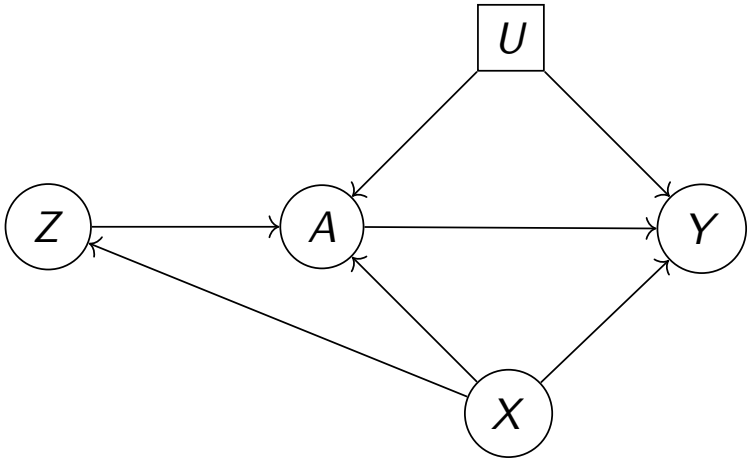
1. Z must affect A.
2. There is no unmeasured Z-Y confounders.
3. There is no direct effect from Z to Y.

Assumptions for instrumental variables



1. Z must affect A: $A(z)$ **depends on** z.
2. There is no unmeasured Z-Y confounders: $Y(a) \perp\!\!\!\perp Z \mid X$.
3. There is no direct effect from Z to Y: $\mathbf{Y(a,z) = Y(a)}$.

Statistical methods for instrumental variables



Two-stage least squares (most widely used)

- **Stage 1:** Regress A on Z and X.
- **Stage 2:** Regress Y on **predicted A from stage 1** and X.
- **Special case:** when there is no X, this is equivalent to the **Wald estimator**:

$$\frac{\text{Slope of } Y \sim Z \text{ regression}}{\text{Slope of } A \sim Z \text{ regression}}$$

.

Remark: Can also use randomisation inference (Imbens and Rosenbaum, 2005).

How to interpret instrumental variable studies

Appropriateness of the assumptions

1. IV must affect treatment.
2. There is no unmeasured IV-outcome confounders.
3. There is no direct effect from IV to outcome.

Additional assumptions

Instrumental variable design often makes additional assumptions. Examples:

- **Homogeneity:** $Y(A=1) - Y(A=0)$ is constant.
- **Monotonicity:** $A(Z=1) \geq A(Z=0)$ (e.g. IV is random encouragement).

Complier average treatment effect

Under monotonicity (and binary IV and treatment), it is well known that

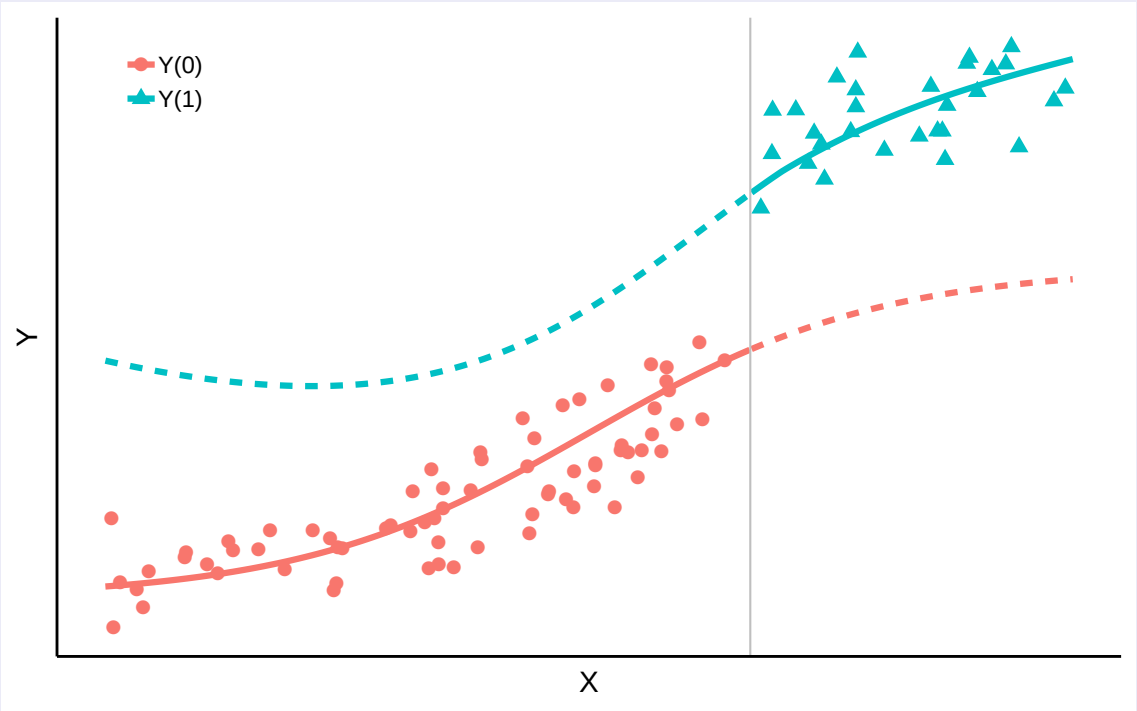
$$\text{The Wald estimator} \to \mathbb{E}[Y(1) - Y(0) \mid A(1)=1, A(0)=0]$$

The condition $\{A(1)=1, A(0)=0\}$ corresponds to the participants who would **comply** with treatment encouragement.

Design 3: Regression discontinuity

Natural experiment: Sharp discontinuity

- Covariate X: Test score.
- Treatment A: Scholarship determined by test score $A = I(X \geq c)$.
- Outcome Y: Future test score.



- Regression discontinuity tries to estimate $\mathbb{E}[Y(1) - Y(0) \mid X = c]$.

Sharp regression discontinuity design

Assumptions

1. X has positive density around the discontinuity c.
2. $\mathbb{E}[Y(0) \mid X]$ and $\mathbb{E}[Y(1) \mid X]$ are **continuous** in x.

- **Remark:** $A = I(X \geq c)$ satisfies the no unmeasured confounders assumption $Y(0) \perp\!\!\!\perp A \mid X$ but not the overlap assumption $0 < \mathbb{P}(A = 1 \mid X = x) < 1$.

Statistical methods

- **Broken line regression:** assume

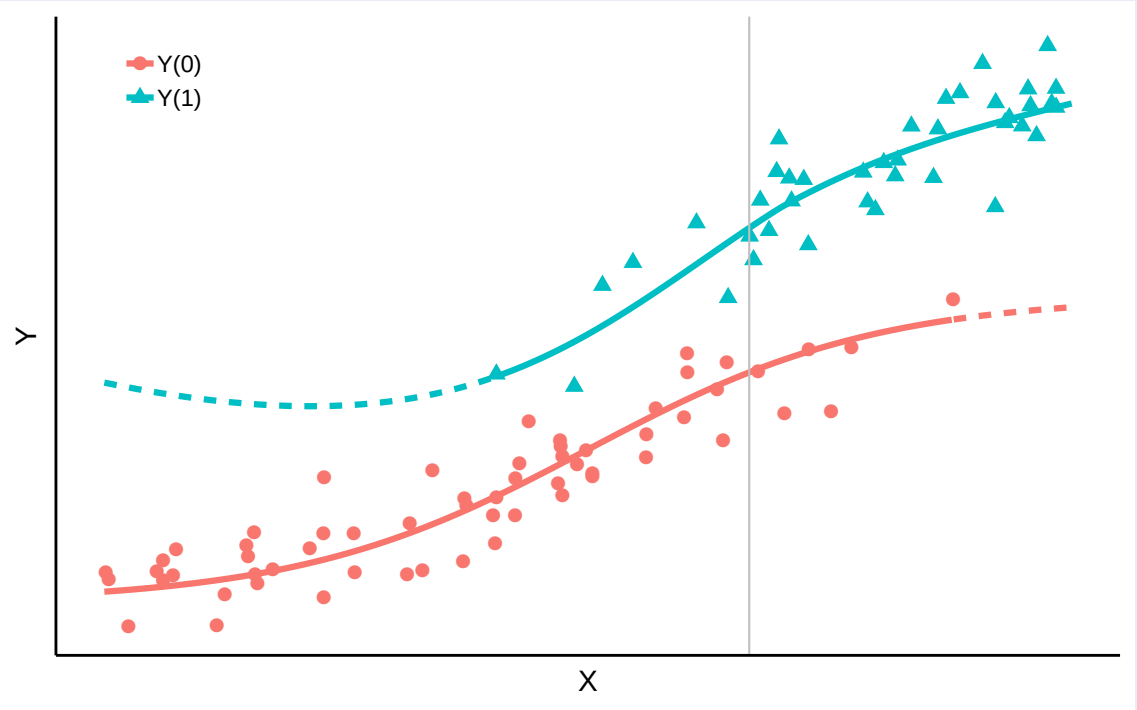
$$\mathbb{E}[Y \mid X] = \begin{cases} \alpha_0 + \gamma_0 x, & \text{if } x < c, \\ \alpha_1 + \gamma_1 x, & \text{if } x \geq c, \end{cases}$$

  Jump can be estimated by $(\hat{\alpha}_1 - \hat{\alpha}_0) + c(\hat{\gamma}_1 - \hat{\gamma}_0)$.
- More robust: **local linear regression** using participants close to the discontinuity.
- Can also use randomisation inference (use randomness in X near c).

Extension

Fuzzy regression discontinuity design

- A is not a deterministic function of X, but $\mathbb{P}(A = 1 \mid X = x)$ has a discontinuity at $x = c$ (jump size < 1).

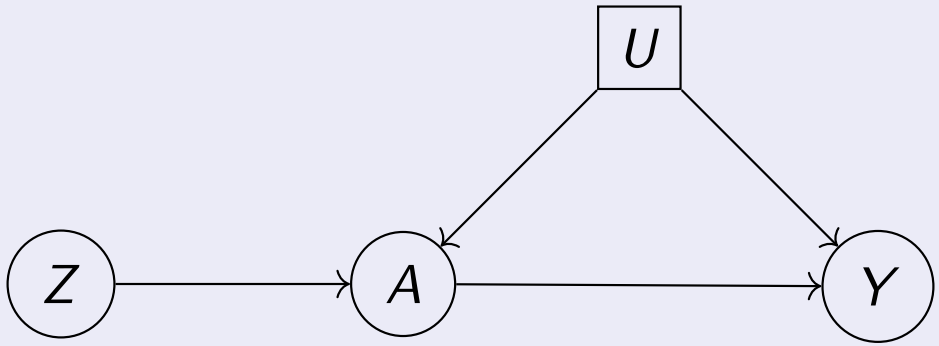


- Can be similarly analysed (broken-line regression, local linear regression, randomisation inference, . . .).

Design 4: Negative controls

- Negative control is a general class of designs that utilise **lack of direct causal effect or association**.
- In other words, these designs utilise **specificity of causal effect**.
- This approach is still under active development. It usually requires additional assumptions beyond specificity.

Example: Instrumental variables

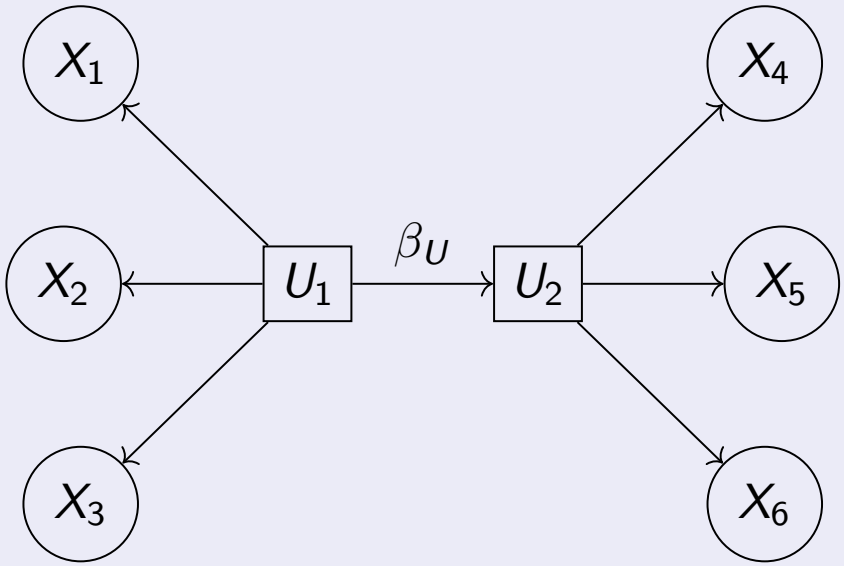


Key assumptions (specificity):

1. IV is independent of unmeasured confounder.
2. IV has no direct effect on outcome.

Design 4: Negative control

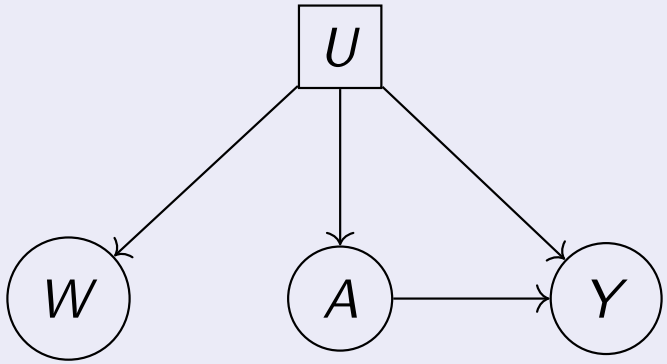
Confirmatory factor analysis and latent variable models



- U_1 and U_2: Latent abstract constructs (e.g. confidence, reading ability, personality, ...).
- X_1 to X_6: Measurements of the latent variables.
- Key assumption (specificity): **lack of association between the measurements** (except those explained by the causal effect of U_1 on U_2).
- **Remark:** Analysis of these designs usually relies on strong parametric assumptions.

Design 4: Negative control

Example: Difference-in-differences (DID)



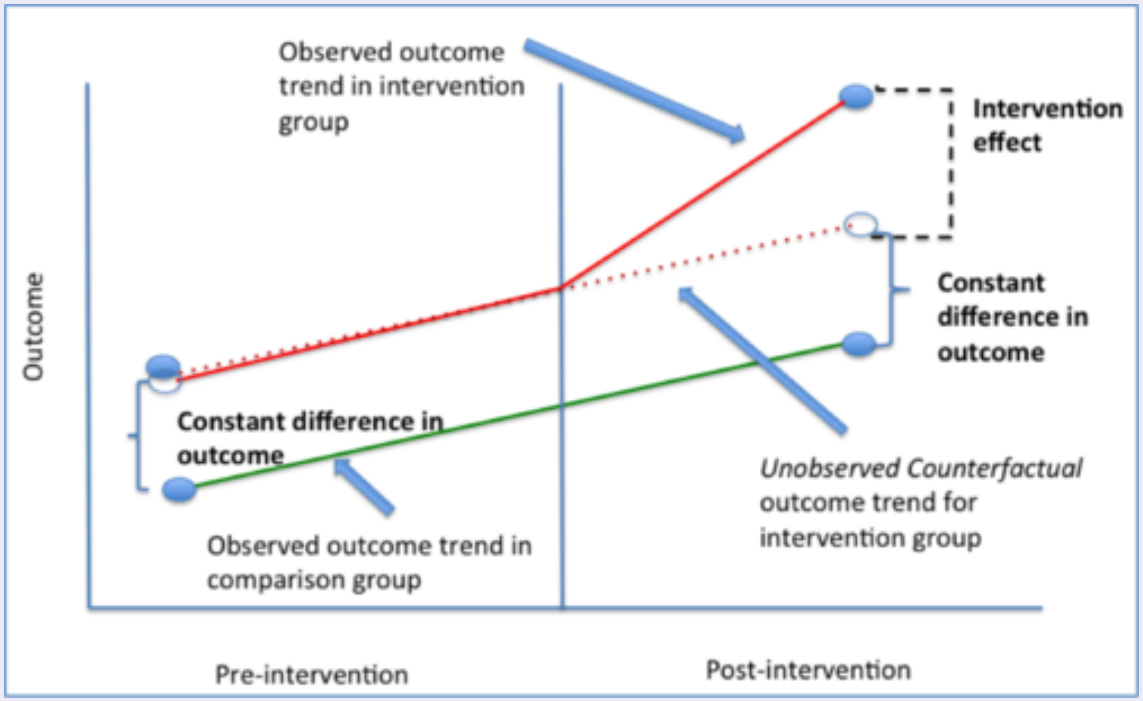
- W and Y are repeated measurements before and after the intervention.
- Example: A is change in minimum wage. W and Y are unemployment rates before and after the change.
- Key assumption (specificity): **Lack of direct effect of A on W**.

Design 4: Negative control

Example: Difference-in-differences (DID)

- DID requires an stronger assumption (than just specificity) called **parallel trends**:

$$\mathbb{E}[Y(0) - W \mid A = 1] = \mathbb{E}[Y(0) - W \mid A = 0].$$



- Estimator: "difference in differences" as illustrated in the figure.

Summary

Part I: Randomised experiments

- Randomisation $\Longrightarrow$ choose between 1. Statistical error and 2. Causality.
- Statistical methods: randomisation inference and regression analysis.

Part II: How to define causality

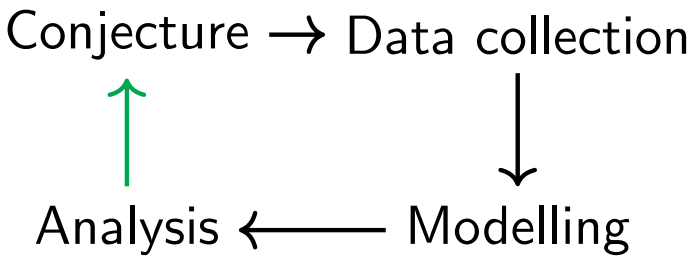
- 1. Counterfactuals; 2. Graphical models; 3. Structural equations.
- "Equivalence" of the definitions and their relative strengths.
- Logic of observational studies: Choose between 1. False assumptions; 2. Statistical error; 3. Causality.

Part III: Designing observational studies

- Design 1: Controlling for confounders;
- Design 2: Instrumental variables;
- Design 3: Regression discontinuity;
- Design 4: Negative controls.

Principles of causal inference

- Observation (seeing) is not intervention (doing).
- Randomised experiment is the gold standard of causal inference.
- Causal inference is abductive (inference to the best explanation).
- Internal, external, and construct validities.
- Design trumps analysis.
- Cycle of statistical research.



Further readings

Book-long treatments (from less mathematical to most mathematical):

- Mackenzie and Pearl (2018) *The Book of Why: The New Science of Cause and Effect*. [General]
- Rosenbaum (2017) *Observation and Experiment: An Introduction to Causal Inference*. [General]
- Freedman (2009) *Statistical Models: Theory and Practice*. [Undergraduate]
- Shadish, Cook, and Campbell (2002) *Experimental and Quasi-Experimental Designs*. [Undergraduate/Postgraduate]
- Angrist and Pischke (2008) *Mostly Harmless Econometrics: An Empiricists Companion*. [Undergraduate/Postgraduate]
- Hernán and Robins (2020) *Causal Inference: What If*. [Part I: Undergraduate; Part II & III: Postgraduate]
- Imbens and Rubin (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences*. [Postgraduate]
- Pearl (2009) *Causality: Models, Reasoning, and Inference*. [Postgraduate]
- Rosenbaum (2010) *Design of Observational Studies*. [Postgraduate]
- Zhao (2019) *Causal Inference Lecture Notes*. [Postgraduate; unpublished and available upon request].

Further readings

Randomised experiments

- **Experimental design:** Box (1978) *Statistics for Experimenters: Design, Innovation, and Discovery*.
- **Randomisation inference:** Rosenbaum (2002) *Observational Studies*. Imbens and Rubin (2015, Chapter 5)
- **Regression adjustment:** Imbens and Rubin (2015, Chapter 7).

Languages of causal inference

- **Counterfactuals:** Imbens and Rubin (2015, Chapters 1–2); Hernán and Robins (2020, Chapters 1–3).
- **Graphical models:** Lauritzen (1996) *Graphical Models* [probabilistic graphical models only]; Pearl (2009); Spirtes, Glymour, and Scheines (2000) *Causation, Prediction, and Search*.
- **Structural equations:** Bollen (1989) *Structural Equations with Latent Variables*; Peters, Janzing, and Schölkopf (2017) *Elements of Causal Inference: Foundations and Learning Algorithms*.

Further readings

Observational studies

- **Controlling for confounders (randomisation inference):** Rosenbaum (2002, 2010);
- **Controlling for confounders (pseudo-population):** Imbens and Rubin (2015); Stuart (2010) Matching Methods for Causal Inference: A Review and a Look Forward (in *Statistical Science*).
- **Controlling for confounders (regression and semiparametric inference):** Hernán and Robins (2020).
- **Instrumental variables:** Angrist and Pischke (2008); Baiocchi, Cheng, Small (2015) Tutorial in Biostatistics: Instrumental Variable Methods for Causal Inference (in *Statistics in Medicine*).
- **Regression discontinuity:** Shadish, Cook, and Campbell (2002); Imbens and Lemieux (2008) Regression discontinuity designs: A guide to practice (in *Journal of Econometrics*).
- **Structural equations with latent variables:** Bollen (1989).
- **Difference in differences:** Angrist and Pischke (2008).

Further readings

Topics not covered in this lecture

- **Sequentially randomised experiments:** Multiple treatments at different time. See Hernán and Robins (2020).
- **Effect modification (treatment effect heterogeneity):** Estimate $\mathbb{E}[Y(1) - Y(0) \mid X = x]$ as a function of x. See the results from a recent data challenge in the journal *Observational Studies*.
- **Dynamic treatment regimes:** How to optimally make sequential interventions? See Kosorok and Laber (2019) Precision Medicine (in *Annual Review of Statistics and Its Application*).
- **Sensitivity analysis:** What if the identification assumptions are violated to a limited degree? See Rosenbaum (2002, 2010).
- **Causal mediation analysis:** Seperate direct and indirect causal effects. See Vanderweele (2015) *Explanation in Causal Inference: Methods for Mediation and Interaction*.
- **Corroboration of evidence (research synthesis):** How to combine evidence from different studies (possibily with different designs)? Often done in a qualitative way, more quantitative developments needed. Classical book: Hedges and Olkin (1985) *Statistical Methods for Meta-Analysis*.

Resources in Cambridge

- The Statistical Laboratory has a free consulting service called Statistics Clinic (`http://www.talks.cam.ac.uk/show/index/21850`).
- I run a reading group in causal inference (`http://talks.cam.ac.uk/show/index/105688`).
- I run a Part III course in causal inference for maths students (`http://www.statslab.cam.ac.uk/~qz280/teaching/Causal_Inference_2019.html`).
- There are several causal inference researchers in MRC Biostatistics Unit, Cambridge social sciences and other subjects.
- Best way to reach me: email me (qz280@cam) about my availability in the Statistics Clinic.

That's all! Questions?