

Statistical Methods of Causal Inference

Lecture 1: The Potential Outcomes Framework

Dr. Sanchayan Banerjee

Institute for Environmental Studies
Vrije Universiteit Amsterdam

July 17, 2023

S.Banerjee@vu.nl

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

Who am I?



Call me **Dr. Sunshine**. An economist by training, I completed my PhD in Behavioural, Experimental, and Environmental Economics at the **London School of Economics** and Political Science. I joined VU Amsterdam as an Assistant Professor in Environmental and Behavioural Economics in September 2022. My research interests lie in developing and testing citizen-oriented behavioural strategies (“Nudge+”) for improving pro-social behaviours (food, energy, public health, education, donation, ...). In my free time, you will find me cooking, enjoying a pint (s), running, or travelling (doing weird adventures)!

The teaching team:

- Dr. Sanchayan Banerjee, Coordinator & Primary Lecturer on Experimental Design (Days 1-2, 7-8, 9).
- Mr. Jack Fitzgerald, Co-coordinator & Primary Lecturer and Instructor on Observational Methods (Days 3-6, 9)
- Mr. Marjan Nikoloski, Instructor on Experimental Design (Days 1-2, 7-8)
- Dr. Katharina Brütt, Guest Teacher (Day 2)
- Prof. Dr. Hans Koster, Guest Teacher (Day 5)

Course Roadmap:

- ① 17 July: Potential Outcomes Framework
- ② 18 July: RCTs and Matching
- ③ 19 July: Difference in Differences
- ④ 20 July: Instrumental Variable Estimation
- ⑤ 21 July: Regression Discontinuity Design
- ⑥ 24 July: Synthetic Control
- ⑦ 25 July: Power Analysis and Advanced Experimental Design
- ⑧ 26 July: Practical Issues in Experiments
- ⑨ 27 July: How to set up your (quasi)-experiment: Get, set, and go!
- ⑩ 28 July: Final Matters (exam in am; Q&A w Dr. John List in pm)

Course Structure:

- Nine 2.5-3 hour lectures (NB: Two lecturers)
- Nine 1.5-2 hour computer practicals (choice of R/STATA)
- Office Hours with me (on demand, shoot me an email).

Course Assessment:

- Daily Quizzes: Best 6 out of 8 will contribute to 40% of the grade.
- Quizzes to be released at the end of the lecture (beginning Day 2) – finish before seminar on same day.
- Final Written Examination: 60% of the grade.
- 80% attendance is compulsory in all lectures and computer practicals to be eligible for examination or attendance certificates.

Course Readings:

- ① Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press, 2009.
- ② Cunningham, Scott. *Causal Inference: The Mixtape*. New Haven: Yale University Press, 2021.
- ③ Moffatt, Peter G. "Experimetrics: A Survey." *Foundations and Trends in Econometrics* 11, no. 1-2 (2021): 1-152.

Note: Individual lectures can be accompanied by a few journal articles. Where ever so, this will be indicated in the lecture slides. For every lecture, indicative chapters from the core readings will also be mentioned.

Course Logistics:

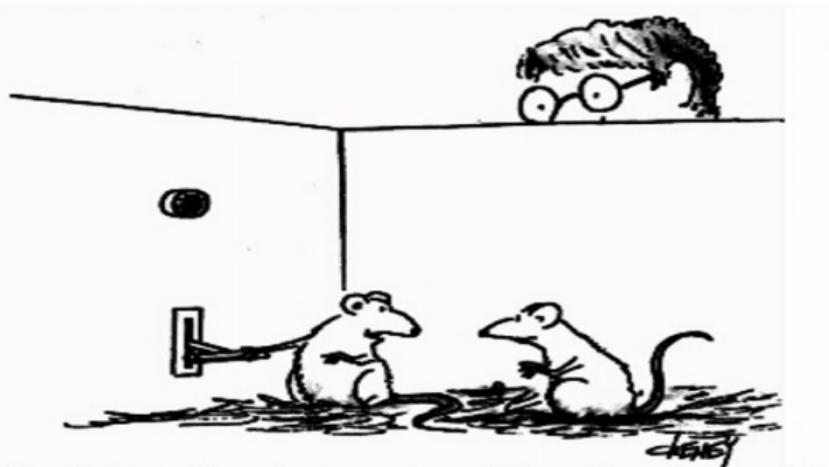
- Lectures will take place in NU-3B07 (Nieuwe Building, Wing B, Room 07), unless otherwise mentioned.
- Lectures will run between 9 am and 12 pm (with lots of breaks included), unless otherwise mentioned.
- Computer practicals will take place in NU-6B15 (Week 1) and NU-6B20 (Week 2), unless otherwise mentioned.
- Computer practicals will run between 1.30 pm and 3.30 pm, unless otherwise mentioned.

When in doubt, your one stop solution is [Canvas](#) (please refer to *Modules/Files* pages.)

If nothing works out, write to me at: S.Banerjee@vu.nl. I will respond to you within 24 hours.

Why are we here?

To learn about causality. Exhibit A – Causal enough?



It's a rather interesting phenomenon. Every time I press this lever, that post-graduate student breathes a sigh of relief.

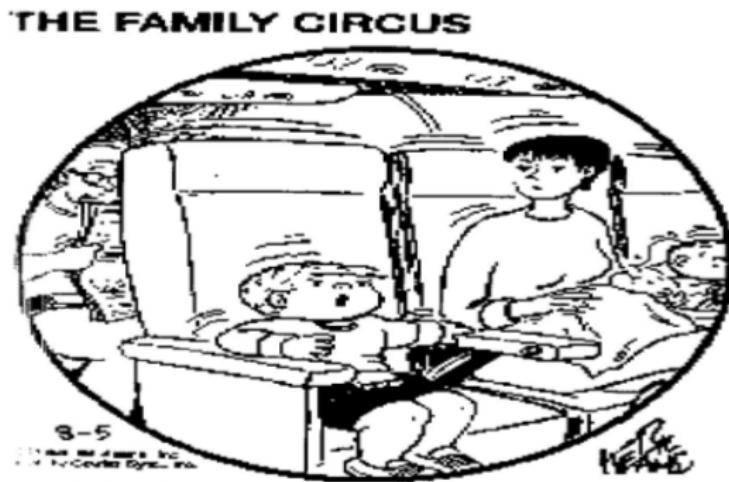
Why are we here?

To learn about causality. Exhibit B – Causal enough?



Why are we here?

To learn about causality. Exhibit C – Causal enough?



"I wish they didn't turn on that seatbelt sign so much! Every time they do, it gets bumpy."

Why are we here?

To learn about causality. But first beware!!!

We are surrounded by many different patterns of events in our daily lives. They all might co-vary ("correlate") with one another, often due to pure chance errors. However, we should be careful to not think of them as cause-and-effect ("causal") relationships.

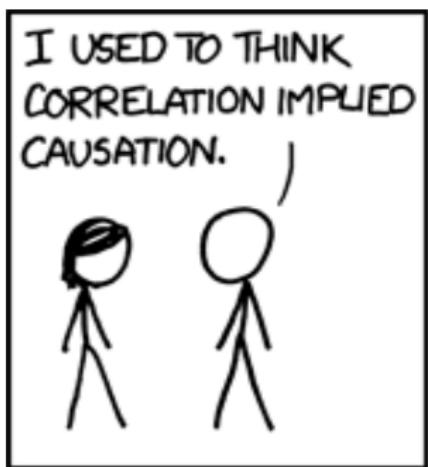
Such seemingly odd associations are often called "spurious correlations"¹. Don't let them fool you.

¹ Interested in Spurious Correlations, [Click Here](#)

Mantra: Correlation is not causation

Course Objective:

By the end of this course, we want you to be smarter in inferring causal relationships...



What is this course about?

- How to study causal effects?
 - ① Concepts, definitions, and theoretical foundations.
 - ② Research designs
 - ③ Implementation of these designs Hands-on examples and exercises
- Formal technical presentation in many places, but that's not the main point. The math is only to back-up your intuitive understanding.
- Only cases where measurement is quantifiable and numerical (no qualitative exposure)
- What this course is **NOT** about?
 - ① Proofs and derivations (It is not an econometrics course)
 - ② Description & Prediction (It is not a simple regression course)

What kind of causal research questions do we consider?

- The course is primarily about **effects of causes**.
 - $D \rightarrow Y$
 - How do the values of an **outcome variable** (Y) change when a **treatment variable** (D) takes on different values?
 - For simplicity, for most of the times, we will focus on a simple case where the treatment has only two values, “active treatment” (1) and “no treatment (control)” (0), simply put as “treatment” and “control”.
- What we do not consider in this course?
 - **Causes of Effects**
 - Substantive explanations of causal mechanisms: *Why* a treatment has the effect that it does?

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

Background

Asking causal questions.

Descriptive and predictive exercises can be very important and interesting², but in modern day applied econometrics, causality trumps it all.

Why?

- Causal relationships directly test theories of how the world works
- Causal relationships provide us with counterfactuals – how the world would have looked with different sets of policies/ circumstances.

²Growth of data science and machine learning has largely centred around prediction questions. Not our core focus but regression methods are a really valuable foundation for machine learning and so we may return to this later in the course.

Background

Some important causal questions today!

- What is the effect of minimum wage on employment?
- What is the effect of increasing emissions on climate change?
- What is the effect of vaccination on health outcomes?
- What is the effect of universal basic income on productivity?
- What is the effect of technology on our emotional well-being?
- What is the effect of number of children on divorce rates?

and so on...

Example – I

Example I – Seats belt mandates and Road Accidents in the US³

³for more details, see Cohen & Einav (2003)

Example – I

What do you expect the use of seat belts mandates to do to traffic fatalities?

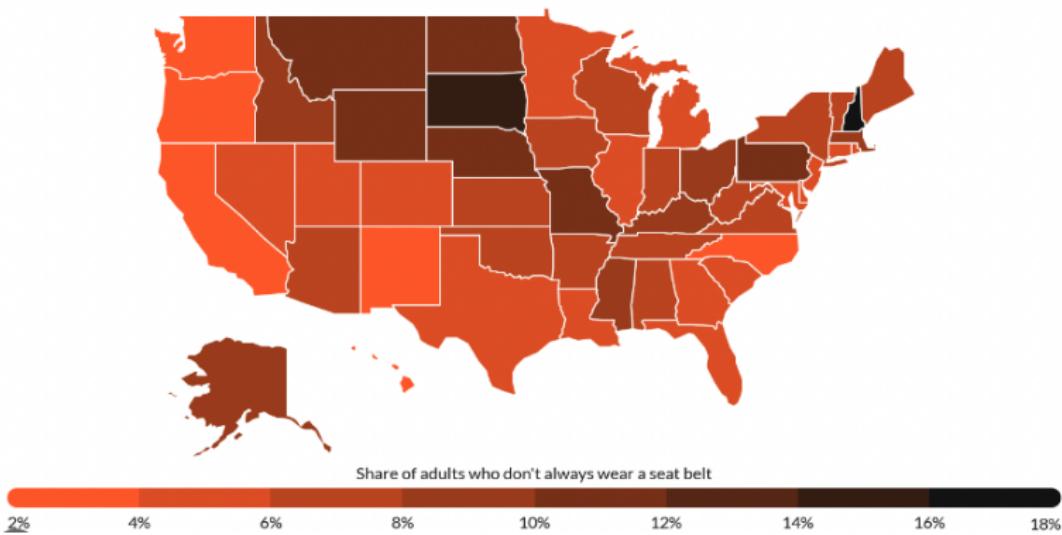
Take the quiz. Scan the QR code below Let's see what you think.



Example – I

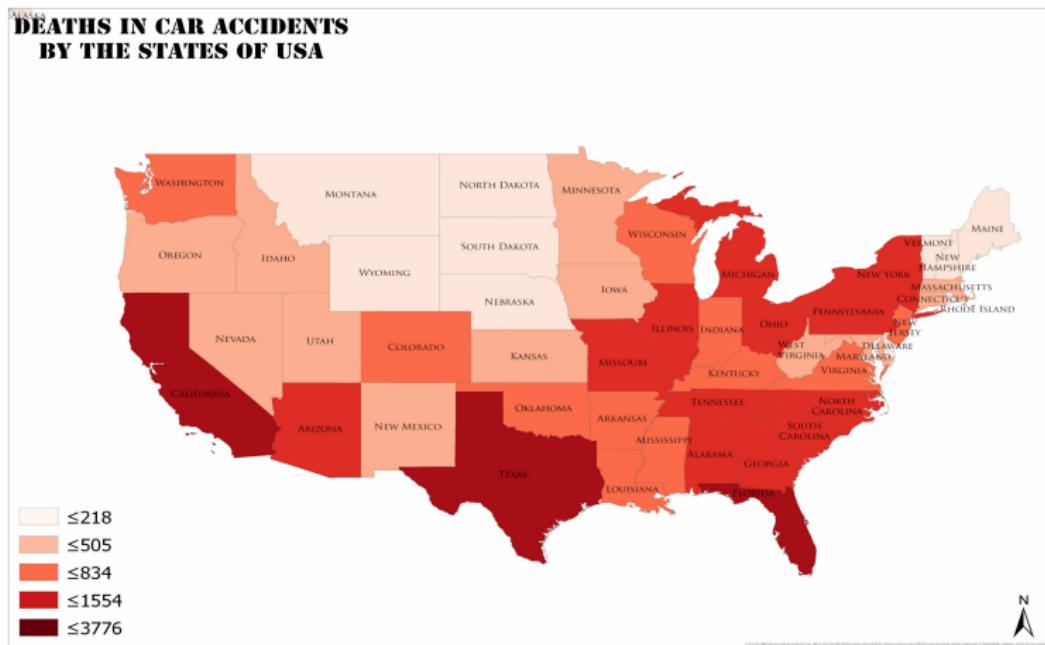
Take a look at the seat belt usage (by states) in the US in 2017.

Residents of rural states are more likely to report not wearing a seat belt



Example – I

Take a look at the traffic accidents (by states) in the US in 2017.



Example – I

So what do you conclude based on these two figures?

- We observe that “states with high reports of low compliance” have “low rates of traffic fatalities”.
- A naive response would be to conclude seat belt mandates **are not effective – in fact they do the opposite**

But why is it naive to claim this? Simply because of the role of “confounders” in this relationship.

Internal Validity

What is a confounder?

Let's assume X (seat belt use) $\rightarrow Y$ (traffic fatalities). This relationship between X & Y is confounded...

- *iff* there is another variable Z (let's say, traffic density) that is correlated with both X and Y , and we have not accounted for it.

A failure to account for Z or rule out Z will lead to a “bias” (in this case, **omitted variable bias**) in the relationship between X & Y .

But why is that so?

Internal Validity

Recap: Ordinary Least Squares and Gauss-Markov Assumptions

Ordinary Least Squares (OLS) estimators perform well under a quite broad variety of different circumstances (the so-called “Gauss Markov Assumptions”), which state:

- ① Error term has conditional mean zero (Exogeneity) $\rightarrow E(u_i|X_i) = 0$
- ② $(X_i, Y_i), \forall i = [1, \dots, n]$ are independent and identically distributed (i.i.d.) draws from their joint distribution.
- ③ Large outliers are unlikely i.e., X_i & Y_i have nonzero finite fourth moments.

In the example 1 above, assumption ① is violated. How?

Internal Validity

Threats to Exogeneity – Omitted Variable Bias

We are interested in estimating the linear relationship between X (seat belt usage) and Y (traffic fatalities). We assume the true relationship between X & Y is given by the relationship below:

$$Y = \alpha + \beta X + u$$

Our OLS estimator $\hat{\beta}$ is an unbiased estimator iff assumptions 1-3 holds. But in this case, it does not. We have **wrongly** omitted a variable Z (traffic density) which has been subsumed in the error term, u , s.t.,

$$u_i = \epsilon_i + Z_i, \text{ and}$$

$$[E(u_i|X_i) = E(\epsilon_i, Z_i|X_i)] \neq 0 \text{ since } \text{corr}(Z_i, X_i) \neq 0$$

Internal Validity

Threats to Exogeneity – Omitted Variable Bias

Under such circumstances, the OLS estimator of X , $\hat{\beta}_x$, turns out to be:

$$\hat{\beta}_x = \beta_x + \beta_z * [\text{corr}(X_i, Z_i) / \text{var}(x)] \neq \beta_x$$

Omitted variable bias is the bias in the OLS estimator that arises when the regressor, X , is correlated with an omitted variable, Z . For omitted variable bias to occur, two conditions must be fulfilled:

- X is correlated with the omitted variable, Z .
- The omitted variable is a determinant of the dependent variable, Y .

Internal Validity

Threats to Exogeneity – Omitted Variable Bias

Omitted Variable Bias is a problem that cannot be solved by increasing the number of observations used to estimate is inconsistent: OVB prevents the estimator from converging in probability to the true parameter value. Strength and direction of the bias are determined by the correlation between the error term and the regressor. For example,

	Corr (X,Z) >0	Corr (X,Z) <0
$\beta_z > 0$	Positive bias	Negative bias
$\beta_z < 0$	Negative bias	Positive bias

Internal Validity

Other Threats to Exogeneity (t.b.c.)

Answering causal questions can be pretty difficult—both practically and econometrically – A “kitchen sink” regression approach is tricky...

Practical challenges	Econometric challenges
Which variables?	Omitted-variable bias
Which functional form(s)?	Reverse Causality (lecture on IV)
Is the sample representative?	Measurement error
Does data exist? How much?	How precise should we be? (specification error)

Many of these challenges relate to the need to hold all else constant (*ceterus paribus*), which is often talked about in terms of exogeneity, i.e., assumption 1.

A general violation of assumption 1 is referred to as “endogeneity” → threat to internal validity

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

Statistics & Causality

Paul W. Holland: “Statistics and Causal Inference”

“The emphasis here will be on measuring the effects of causes because this seems to be a place where statistics...has contributions to make. It is my opinion that an emphasis on the **effects of causes** rather than on the causes of effects is, in itself, an important consequence of bringing statistical reasoning to bear on the analysis of causation...”

- This, in turn, leads naturally to the question of the conditions, assumptions, etc., that are required to use statistics to make causal inferences.
- Once again, remember the mantra: **Correlation is not causation**

The Potential Outcomes Framework

The Potential Outcomes Framework

A general statistical framework for the analysis of cause and effect.

- Underlying ideas are actually much older than the modern field of statistics

Consider the following quote:

“If a person eats of a particular dish, and dies in consequence, that is, would not have died if he had not eaten of it, people would be apt to say that eating of that dish was the cause of his death.” (John Stuart Mill: *A system of logic*)

The Potential Outcomes Framework

A general statistical framework for the analysis of cause and effect.

- Underlying ideas are actually much older than the modern field of statistics

Consider the following quote:

“If a person eats of a particular dish, and dies in consequence, that is, would not have died if he had not eaten of it, people would be apt to say that eating of that dish was the cause of his death.” (John Stuart Mill: *A system of logic*)

The Potential Outcomes Framework

A general statistical framework for the analysis of cause and effect.

- Underlying ideas are actually much older than the modern field of statistics

Consider the following quote:

"If a person eats of a particular dish, and dies in consequence, that is, **would not have died if he had not eaten of it**, people would be apt to say that eating of that dish was the cause of his death." (John Stuart Mill: *A system of logic*)

The Potential Outcomes Framework

A general statistical framework for the analysis of cause and effect.

- Underlying ideas are actually much older than the modern field of statistics

Consider the following quote:

"If a person eats of a particular dish, and dies in consequence, that is, would not have died if he had not eaten of it, **people would be apt to say that eating of that dish was the cause of his death.**" (John Stuart Mill: A *system of logic*)

A relationship between X & Y is causal iff

- ① if X, then Y
- ② if not X, then not Y

The Potential Outcomes Framework

Notable antecedents:

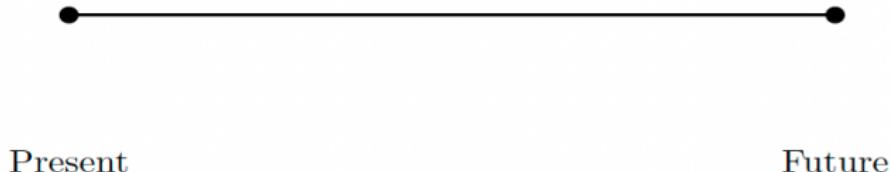
- ① Jerzy Neyman (1923) introduced the potential outcomes notation for experiments
- ② R.A. Fisher (1925) proposed actually randomising treatments to units
- ③ Donald Rubin (1974) extended the potential outcomes framework [a.k.a. the “Rubin Causal Model” (Holland 1986)] to observational studies

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

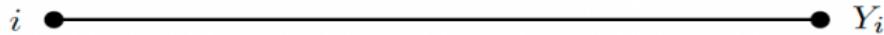
The World of Potential Outcomes

Potential Outcomes → Potential Future Worlds



The World of Potential Outcomes

Potential Outcomes \longrightarrow Potential Future Worlds

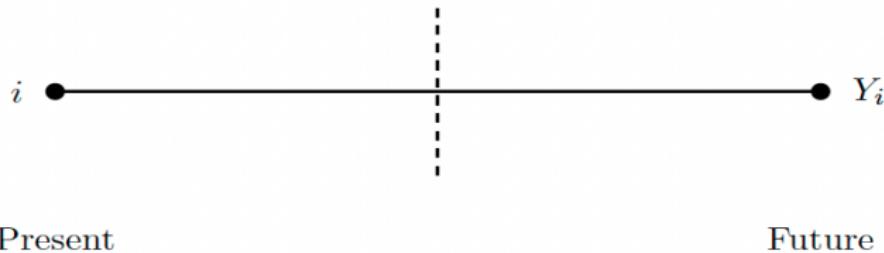


Present

Future

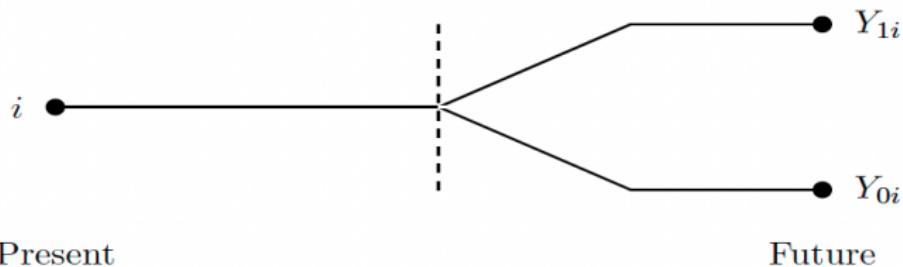
The World of Potential Outcomes

Potential Outcomes → Potential Future Worlds



The World of Potential Outcomes

Potential Outcomes \rightarrow Potential Future Worlds



The World of Potential Outcomes

Example II – World Bank Poverty Reduction Program

The World of Potential Outcomes

Example II – World Bank Poverty Reduction Program

Consider the Future GDP Per Capita of different countries in two different:

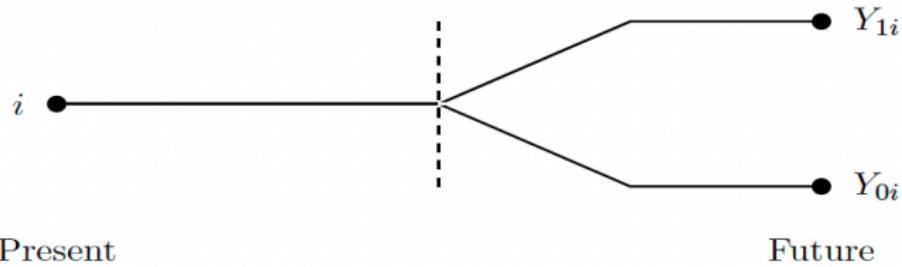
Y_{0i} → potential outcome when the country i does not receive aid; and

Y_{1i} → potential outcome when the country i receives aid

Unit	Y_{0i}	Y_{1i}	$Y_{1i} - Y_{0i}$
Denmark	\$41,932	\$43,445	\$1,513
Liberia	\$806	\$878	\$72
Australia	\$42,063	\$43,544	\$1,481
Afghanistan	\$988	\$1,946	\$958
South Korea	\$31,079	\$33,140	\$2,061
Haiti	\$1,184	\$1,703	\$519
Canada	\$40,885	\$43,247	\$2,362
Papua New Guinea	\$1,972	\$2,539	\$567
Saudi Arabia	\$50,792	\$53,644	\$2,852
Honduras	\$3,634	\$4,593	\$959

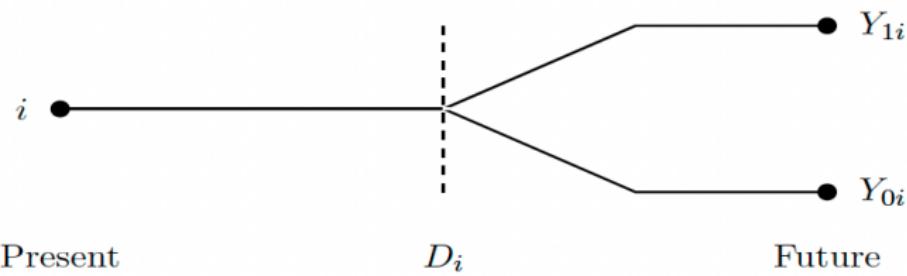
The World of Potential Outcomes

Potential Outcomes \rightarrow Potential Future Worlds



The World of Potential Outcomes

Potential Outcomes \rightarrow Potential Future Worlds



The World of Potential Outcomes

Definition (Outcome)

Y_i : Observed outcome variable of interest for unit i

Definition (Potential Outcome)

Y_{0i} and Y_{1i} : Potential outcomes for unit i

$$Y_{.i} = \begin{cases} Y_{1i} & \text{Potential outcome for unit } i \text{ with treatment} \\ Y_{0i} & \text{Potential outcome for unit } i \text{ without treatment} \end{cases}$$

Definition (Treatment)

D_i : Indicator of treatment intake for $unit i$

$$D_i = \begin{cases} 1 & \text{if unit } i \text{ received the treatment} \\ 0 & \text{otherwise.} \end{cases}$$

The World of Potential Outcomes

$$\text{Potential Outcome} = \begin{cases} Y_{1i}, & \text{if } D_i = 1 \\ Y_{0i}, & \text{if } D_i = 0 \end{cases}$$

The World of Potential Outcomes

$$\text{Potential Outcome} = \begin{cases} Y_{1i}, & \text{if } D_i = 1 \\ Y_{0i}, & \text{if } D_i = 0 \end{cases}$$

Y_{1i} : Unit i 's measurement on Y if we had ended up in the potential future where $D_i = 1$

- Regardless of whether we actually end up in that world
- Our example: *GDP per capita for a country, had it been a part of the development program, irrespective of whether it actually was part of the program*

Y_{0i} : Unit i 's measurement on Y if we had ended up in the potential future where $D_i = 0$

- Regardless of whether we actually end up in that world
- Our example: *GDP per capita for a country, had it not been a part of the development program, irrespective of whether it actually was part of the program*

The World of Potential Outcomes

$$\text{Potential Outcome} = \begin{cases} Y_{1i}, & \text{if } D_i = 1 \\ Y_{0i}, & \text{if } D_i = 0 \end{cases}$$

The *causal effect* of an intervention or treatment on the outcome for unit i is the difference between its two potential outcomes:

$$Y_{1i} - Y_{0i}$$

Fundamental Problem of Causal Inference

However, there is a fundamental problem of “identification” here.

- It is impossible to observe the values of Y_{0i} and Y_{1i} on the same unit i at the same time period.
- Therefore, it is impossible to observe the effect of D on unit i .
- *A priori* each potential outcome could be observed.
- After assignment of D_i , the treatment, one outcome is observed, the other becomes the counterfactual.

This identification problem is often referred to as the Fundamental Problem of Causal Inference (FPCI).

Fundamental Problem of Causal Inference

Example II – World Bank Poverty Reduction Program

Consider example II once again. Let's say we observe now the treatment status, D_i of the country i .

Unit	Y_{0i}	Y_{1i}	$Y_{1i} - Y_{0i}$	D_i
Denmark	\$41,932	\$43,445	\$1,513	
Liberia	\$806	\$878	\$72	
Australia	\$42,063	\$43,544	\$1,481	
Afghanistan	\$988	\$1,946	\$958	
South Korea	\$31,079	\$33,140	\$2,061	
Haiti	\$1,184	\$1,703	\$519	
Canada	\$40,885	\$43,247	\$2,362	
Papua New Guinea	\$1,972	\$2,539	\$567	
Saudi Arabia	\$50,792	\$53,644	\$2,852	
Honduras	\$3,634	\$4,593	\$959	

Fundamental Problem of Causal Inference

Example II – World Bank Poverty Reduction Program

Some countries received the aid ($D_i = 1$) whereas others did not ($D_i = 0$).

Unit	Y_{0i}	Y_{1i}	$Y_{1i} - Y_{0i}$	D_i
Denmark	\$41,932	\$43,445	\$1,513	0
Liberia	\$806	\$878	\$72	1
Australia	\$42,063	\$43,544	\$1,481	0
Afghanistan	\$988	\$1,946	\$958	1
South Korea	\$31,079	\$33,140	\$2,061	0
Haiti	\$1,184	\$1,703	\$519	1
Canada	\$40,885	\$43,247	\$2,362	0
Papua New Guinea	\$1,972	\$2,539	\$567	1
Saudi Arabia	\$50,792	\$53,644	\$2,852	0
Honduras	\$3,634	\$4,593	\$959	1

Fundamental Problem of Causal Inference

Example II – World Bank Poverty Reduction Program

Consequently, we observe only those outcomes that have realised.

Unit	Y_{0i}	Y_{1i}	$Y_{1i} - Y_{0i}$	D_i
Denmark	\$41,932	\$43,445	\$1,513	0
Liberia	\$806	\$878	\$72	1
Australia	\$42,063	\$43,544	\$1,481	0
Afghanistan	\$988	\$1,946	\$958	1
South Korea	\$31,079	\$33,140	\$2,061	0
Haiti	\$1,184	\$1,703	\$519	1
Canada	\$40,885	\$43,247	\$2,362	0
Papua New Guinea	\$1,972	\$2,539	\$567	1
Saudi Arabia	\$50,792	\$53,644	\$2,852	0
Honduras	\$3,634	\$4,593	\$959	1

Most common estimands in this framework

Wait, Esti-what?

Glossary (some terms to remember):

- ① **Estimand** → Unobserved population parameter or function.
- ② **Estimator** → A function that can be applied to the observed data.
- ③ **Estimate** → A specific output of the said function.

OK. Let's resume back!

Individual Treatment Effect

Definition (Individual Treatment Effect)

Causal effect of the treatment on the outcome for unit i , defined by the comparison of two potential outcomes:

$$\tau_i = Y_{1i} - Y_{0i}$$

This **cannot be observed**, and is also very hard to estimate:

- We cannot observe both potential outcomes Y_{1i} and Y_{0i} for the same unit i .
- Hard to reliably fill in the missing potential outcome for any one unit i .

Average Treatment Effect (ATE)

- Consider a fixed group (**population**) of units $i = 1, \dots, N$
- Values of the potential outcomes for this population can be represented as two vectors:

$$\mathbf{Y}_1 = (Y_{11}, Y_{12}, \dots, Y_{1N})$$

$$\mathbf{Y}_0 = (Y_{01}, Y_{02}, \dots, Y_{0N})$$

- A population causal estimand is a comparison of \mathbf{Y}_1 and \mathbf{Y}_0
- A common choice is a difference of their expected values (means).

Average Treatment Effect (ATE)

Definition (Average treatment effect, ATE)

$$\tau_{ATE} = \frac{1}{N} \sum_{i=1}^N (Y_{1i} - Y_{0i})$$

or equivalently

$$\tau_{ATE} = \mathbb{E}[Y_{1i} - Y_{0i}]$$

- In the rest of this course, we will consider various assumptions under which τ_{ATE} can be **identified** from observed information
- Note on notation: We represent the **estimand** as a greek letter (in this case τ , but could be anything). We typically represent an **estimator** for that estimand as a greek letter with something on top (e.g. $\tilde{\tau}$ or $\hat{\tau}$). An **estimate** will be a realised number (interval, etc.).

Average Treatment Effect on the Treated (ATT)

Definition (Average treatment effect on the treated, ATT)

$$\tau_{ATT} = \frac{1}{N_1} \sum_{i=1}^N D_i(Y_{1i} - Y_{0i}) \quad \text{where} \quad N_1 = \sum_{i=1}^N D_i$$

or equivalently $\tau_{ATT} = \mathbb{E}[Y_{1i} - Y_{0i}|D_i = 1]$

- In words, N_1 equals the number of treated units.
- When would $\tau_{ATT} \neq \tau_{ATE}$? When D_i and Y_{di} are associated.
- Exercise: Define τ_{ATC} , ATE on the untreated (control) units.

$$\tau_{ATC} = \mathbb{E}[Y_{1i} - Y_{0i}|D_i = 0]$$

Conditional Average Treatment Effect (CATE) or Local Average Treatment Effect (LATE)

Definition (Conditional average treatment effects, CATE)

$$\tau_{CATE}(x) = \mathbb{E}[Y_{1i} - Y_{0i} | X_i = x]$$

where X_i is a **pre-treatment covariate** for unit i

- In words, $\tau_{CATE}(x)$ is a **subgroup effect**, treatment effect on units who have particular characteristics x .
- This estimand sometimes goes by other names (e.g. local average treatment effect or LATE).
- This is an increasing important area for causal inference (e.g. optimal policy targeting), and we will return to it later!



Internal Validity

- The term 'population' is used here in an unfamiliar-seeming way: It refers *only* to the N units for whom we have observed data (i.e. what we would typically call the 'sample')
- The estimands considered on this course are defined and estimated for this population
- Internal validity refers to the validity of our estimates of these effects. This class is focused only on internal validity.

External Validity

- External validity refers to the validity generalising our estimates of causal effects from the 'population' of N units to any other population (note, this could include generalising from a realised sample to a population)
- If claimed, external validity has to be justified by different kinds of arguments, e.g.
 - Representative sampling (ideally probability sampling) of the N units from a larger population. This is a population inference task, as in survey research
 - Substantive theory / assumptions / wishful thinking about why a causal effect for these N units would also apply elsewhere

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

What is identification?

Identification

- Definition: Sufficient condition(s) for drawing a conclusion given the data generating process.
- A statistical model is said to be **identifiable** if it is theoretically possible to learn the true values of its parameters if an infinite number of observations were obtained.
- A parameter of a statistical model is said to be **identified** if the data require that the parameter converges with probability one to the true value as the sample size goes to infinity.
- In this class, we focus on the specific subset of identification issues known as **causal identification**.

What is identification strategy?

An identification strategy

- Definition: Sufficient condition(s) for drawing a conclusion about a causal effect given the data generating process.
- An **identification strategy** is the combination of a description of the source of variation in a causal variable and the application of a statistical technique to make use of that information.
 - where “description of the source of variation in a causal variable” implies knowing about the assignment process that led to an observed distribution of treatment(s).

Recall: Our objective is to be able to identify $\tau_{ATE} = E[Y_{1i} - Y_{0i}]$ from the observed information in the data.

Identification Assumption for the Potential Outcomes Framework

SUTVA

Underlying all methods that make causal claims in the potential outcomes framework is:

The Stable Unit Treatment Value Assumption (a.k.a. **SUTVA**)

SUTVA implies **no unmodelled spillovers**. In other words, potential outcomes for a given individual responds to only its **own** treatment status; potential outcomes are invariant to random assignment of others.

Identification Assumption for the Potential Outcomes Framework

What is SUTVA?

SUTVA comes into three parts:

- ① There is only ONE version of the treatment i.e. the treatment is exhaustive and there are no hidden assignments
 - Any units that take the same value of D experienced the same regime
- ② Units cannot interfere with one another in ways that affect treatments or measurement of outcomes i.e. there are no interaction effects
 - For example, units in the control group cannot learn about the treatment by interacting with peers in the treatment group.
- ③ There is no measurement error i.e. neither the treatment nor the observed outcomes are measured with error.

Identification Assumption for the Potential Outcomes Framework

If SUTVA holds, we can link the potential outcomes to a unit's observed outcome using the following switching equation:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$$

Implications:

- If $D_i = 1$, then $Y_i \rightarrow Y_{1i}$, and
- If $D_i = 0$, then $Y_i \rightarrow Y_{0i}$

Recap: Fundamental problem of causal inference

Recall, FPCI tells us that half the potential outcomes are missing.

So how can we estimate: $\tau_{ATE} = E[Y_{1i} - Y_{0i}]$.

- A large amount of homogeneity would solve this problem, that is if
 - (Y_{1i}, Y_{0i}) constant across individuals
 - (Y_{1i}, Y_{0i}) constant across time
- Unfortunately, however, there is a large heterogeneity in the individual responses.

The Statistical Solution to FPCI

The *Statistical Solution* to the *Fundamental Problem of Causal Inference*

- Make use of the population from which the units are drawn
- Instead of calculating

$$Y_{1i} - Y_{0i} \text{ for each } i$$

Estimate

$$E(Y_1 - Y_0) = E(Y_1) - E(Y_0) \text{ over all } i \text{ in the population}$$

- Use units with $D_i = 1$ to obtain information about Y_1
- Use units with $D_i = 0$ to obtain information about Y_0

The Statistical Solution to FPCI

Observed Quantities

$E(Y_{1i}|D_i = 1)$: Average actual outcomes for units with
 $D = 1$

$E(Y_{0i}|D_i = 0)$: Average actual outcomes for units with
 $D = 0$

Unobserved Quantities

$E(Y_{1i}|D_i = 0)$: Average counterfactual outcomes for units
with $D = 0$

$E(Y_{0i}|D_i = 1)$: Average counterfactual outcomes for units
with $D = 1$

The Statistical Solution to FPCI

Can we ever say that the following equalities hold?

$$E(Y_0) = E(Y_0|D_i = 0)$$

$$E(Y_1) = E(Y_1|D_i = 1)$$

Yes, if we are willing to make the assumption of *independence*

- The average values of Y_0 or Y_1 do not depend on whether we focus on observed values or the whole population of observed and unobserved values
- In other words, potential outcomes are uncorrelated with treatment assignment
- In formal notation: $(Y_0, Y_1) \perp\!\!\!\perp D$

Back to Example II

To measure average causal effects, we want to be able to calculate

$$E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

But we have that

$$\begin{aligned} E(Y_0) &= E(Y_0|D_i = 0) \Pr(D_i = 0) \\ &\quad + E(Y_0|D_i = 1) \Pr(D_i = 1) \end{aligned}$$

$$\begin{aligned} E(Y_1) &= E(Y_1|D_i = 0) \Pr(D_i = 0) \\ &\quad + E(Y_1|D_i = 1) \Pr(D_i = 1) \end{aligned}$$

Unit	Y_{0i}	Y_{1i}	$Y_{1i} - Y_{0i}$	D_i
Denmark	\$41,932	\$43,445	\$1,513	0
Liberia	\$806	\$878	\$72	1
Australia	\$42,063	\$43,544	\$1,481	0
Afghanistan	\$988	\$1,946	\$958	1
South Korea	\$31,079	\$33,140	\$2,061	0
Haiti	\$1,184	\$1,703	\$519	1
Canada	\$40,885	\$43,247	\$2,362	0
Papua N.G.	\$1,972	\$2,539	\$567	1
Saudi Arabia	\$50,792	\$53,644	\$2,852	0
Honduras	\$3,634	\$4,593	\$959	1

We want to be able to use

$$\begin{aligned} E(Y_0) &= E(Y_0|D_i = 0) \\ E(Y_1) &= E(Y_1|D_i = 1) \end{aligned}$$

Back to Example II

<i>Unit</i>	Y_{0i}	Y_{1i}	$Y_{1i} - Y_{0i}$	D_i
Denmark	\$41,932	\$43,445	\$1,513	0
Liberia	\$806	\$878	\$72	1
Australia	\$42,063	\$43,544	\$1,481	0
Afghanistan	\$988	\$1,946	\$958	1
South Korea	\$31,079	\$33,140	\$2,061	0
Haiti	\$1,184	\$1,703	\$519	1
Canada	\$40,885	\$43,247	\$2,362	0
Papua New Guinea	\$1,972	\$2,539	\$567	1
Saudi Arabia	\$50,792	\$53,644	\$2,852	0
Honduras	\$3,634	\$4,593	\$959	1

Can we just use $E(Y_0|D_i = 0)$ in place of $E(Y_0)$ and $E(Y_1|D_i = 1)$ in place of $E(Y_1)$?

- Let's look at the sample mean analogues for these expectations:

$$\bar{Y}_0 = \frac{\$41,932 + \$806 + \$42,063 + \$988 + \$31,079 + \$1,184 + \$40,885 + \$1,972 + \$50,792 + \$3,634}{10} \\ = \$21,533.50$$

$$\bar{Y}_0|D=0 = \frac{\$41,932 + \$42,063 + \$31,079 + \$40,885 + \$50,792}{5} = \$41,350.20$$

$$\bar{Y}_1 = \frac{\$43,445 + \$878 + \$43,544 + \$1,946 + \$33,140 + \$1,703 + \$43,247 + \$2,539 + \$53,644 + \$4,593}{10} \\ = \$22,867.90$$

$$\bar{Y}_1|D=1 = \frac{\$878 + \$1,946 + \$1,703 + \$2,539 + \$4,593}{5} = \$2,331.80$$

- In both cases, it looks like we are about \$20,000 off the mark!

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

Random Assignment

Randomised experiments are the gold standard

The reason that randomised experiments are the gold standard for causal inference...

is that random assignment achieves independence

(“in expectation” or “asymptotically”)

Random Assignment

If we do not have random assignment, we have **selection bias**.

- This means there is some non-random process that *selects* units into the treatment and control groups
- If we cannot account for this process, we do not know how to appropriate estimate the *causal* relationship between a predictor and the outcome variable of interest.

Selection Bias

Let's say we don't know the data generating process that led to the observed distribution of D_i

Selection Bias

Let's say we don't know the data generating process that led to the observed distribution of D_i

Naïve comparison of average poverty change

$$E(Y_i|D_i = 1) - E(Y_i|D_i = 0) = E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0)$$

Selection Bias

Let's say we don't know the data generating process that led to the observed distribution of D_i

Naïve comparison of average poverty change

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}_{\text{Observed difference in average outcome measures}} = E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0)$$

Selection Bias

Let's say we don't know the data generating process that led to the observed distribution of D_i

Naïve comparison of average poverty change

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}_{\text{Observed difference in average outcome measures}} = E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0)$$
$$= E(Y_{1i}|D_i = 1) - \cancel{E(Y_{0i}|D_i = 1)} + \cancel{E(Y_{0i}|D_i = 1)} - E(Y_{0i}|D_i = 0)$$

Selection Bias

Let's say we don't know the data generating process that led to the observed distribution of D_i

Naïve comparison of average poverty change

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}_{\text{Observed difference in average outcome measures}} = E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0)$$

$$= \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}_{\text{Average treatment effect on the treated}} + \underbrace{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Selection bias}}$$

Selection Bias

Let's say we don't know the data generating process that led to the observed distribution of D_i

Naïve comparison of average poverty change

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}_{\text{Observed difference in average outcome measures}} = \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Average treatment effect on the treated}} + \underbrace{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Selection bias}}$$

- If *selection bias* is large and in the same direction as the average treatment effect on the treated, the treatment effect will seem larger than it really is
- If *selection bias* is large and in the opposite direction as the average treatment effect on the treated, the treatment effect will seem smaller than it really is, or perhaps even reversed

Random Assignment and Causal Identification

Random assignment solves the selection problem

If D_i is independent of Y_i

$$\begin{aligned} E(Y_{1i}|D_i = 1) &= E(Y_{1i}|D_i = 0) \\ E(Y_{0i}|D_i = 0) &= E(Y_{0i}|D_i = 1) \end{aligned}$$

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}$$

Observed difference in average
outcome measures

$$= \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}_{\text{Average treatment effect on the treated}} + \underbrace{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Selection bias}}$$

Random Assignment and Causal Identification

Random assignment solves the selection problem

If D_i is independent of Y_i

$$\begin{aligned} E(Y_{1i}|D_i = 1) &= E(Y_{1i}|D_i = 0) \\ E(Y_{0i}|D_i = 0) &= E(Y_{0i}|D_i = 1) \end{aligned}$$

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}$$

Observed difference in average
outcome measures

$$= \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}_{\text{Average treatment effect on the treated}} + \underbrace{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Selection bias}}$$

Random Assignment and Causal Identification

Random assignment solves the selection problem

If D_i is independent of Y_i

$$\begin{aligned} E(Y_{1i}|D_i = 1) &= E(Y_{1i}|D_i = 0) \\ E(Y_{0i}|D_i = 0) &= E(Y_{0i}|D_i = 1) \end{aligned}$$

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}$$

Observed difference in average
outcome measures

$$= \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}_{\text{Average treatment effect on the treated}} + \underbrace{0}_{\text{Selection bias}}$$

Random Assignment and Causal Identification

Random assignment solves the selection problem

If D_i is independent of Y_i

$$\begin{aligned} E(Y_{1i}|D_i = 1) &= E(Y_{1i}|D_i = 0) \\ E(Y_{0i}|D_i = 0) &= E(Y_{0i}|D_i = 1) \end{aligned}$$

$$\underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}$$

Observed difference in average
outcome measures

$$= \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}$$

Average treatment effect on the
treated

Random Assignment and Causal Identification

“Correlation is not causation” revisited

$$\begin{aligned}
 & \underbrace{E(Y_i|D_i = 1) - E(Y_i|D_i = 0)}_{\text{Observed difference in average outcome measures}} \\
 &= \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}_{\text{Average treatment effect on the treated}} + \underbrace{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Selection bias}}
 \end{aligned}$$

On the left-hand-side, we have a simple difference of means between units in the $D = 1$ group and units in the $D = 0$ group

- We can always estimate this quantity
- But if we do not know about the process that produced the observed distribution of D values across units, we do not know how much selection bias is contributing to the estimate

Random Assignment and Causal Identification

- Naïvely examining a difference of means is unappealing unless we know something about the *assignment mechanism*

Definition (Assignment Mechanism)

Assignment mechanism is the procedure that determines which units get which treatment. Examples include:

- random assignment
- selection on observables
- selection on unobservables
- Most statistical models of causal inference attain identification of treatment effects by restricting the assignment mechanism in some way

Assignment Mechanism

Things we want to know about the assignment mechanism

- ① Is it individualistic?
 - The probability that a unit is assigned to a particular treatment does not depend on the characteristics (covariates or potential outcomes) of any other units.
- ② Is it probabilistic?
 - Every unit has some chance of being in any treatment group, depending on their values on covariates.
- ③ Is it unconfounded?
 - Assignment mechanism does not depend on units' potential outcomes
- ④ Is it known?
 - In randomised experiments, we know it because it is controlled by the researcher.
 - In observational (quasi-experimental) research, we have to convince ourselves that we know it by making a lot of assumptions.

Outline

- 1 Housekeeping & General Information
- 2 Correlation versus Causation
- 3 Potential Outcomes Background
- 4 The Logic of Potential Outcomes Framework
- 5 The identification problem for causal inference
- 6 Random Assignment and Causal Inference
- 7 Summary

Summary

Design Trumps Analysis⁴

- ① Randomised Experiments (Gold Standard)
- ② Observational Studies
 - Selection on Observables
 - Regression (Kitchen-Sink)
 - Matching
 - Weighting
 - Selection on Unobservables
 - Difference-in-Differences
 - Instrumental Variables
 - Regression Discontinuity Design
 - Synthetic Control

⁴see Rubin, 2008

Summary

Key Takeaways

- The fundamental problem of causal inference presents us with a missing data problem.
- To overcome this missing data problem, we use units in each category to go from individual-level measurements of causal effects to population-level estimates of average causal effects.
- This introduces the possibility of selection bias, which we have to use identification assumptions to overcome.
- One such identification assumption is independence.
- Randomisation achieves independence in expectation.

Summary

Key Ideas

- Causality is defined by potential outcomes, not by realised (observed) outcomes.
- Observed association is neither necessary nor sufficient for causality.
- Estimation of causal effects of a treatment (usually) starts with studying the assignment mechanism.
- We want to achieve the standards of a randomised experiment even if we do not have one.
- When we do not have one, our ability to make causal inferences often relies on making untestable assumptions about the assignment mechanism based on our knowledge of the situation.