Causal Inference Workshop

Week 2 - Potential Outcomes Framework

Causal Inference Workshop

January 30, 2024

Workshop outline

A. Causal inference fundamentals
   - Modeling assumptions matter too
   - Conceptual framework (potential outcomes framework)
B. Design stage: common identification strategies
   - IV + RDD [coding]
   - DiD, DiDiD, Event Studies, New TWFE Lit [coding]
   - Synthetic Control / Synthetic DiD [coding]
C. Analysis stage: strengthening inferences
   - Limitations of identification strategies, pre-estimation steps
   - Estimation [controls] and post-estimation steps [supporting assumptions]
D. Other topics in causal inference and sustainable development
   - Inference (randomization inference, bootstrapping)
   - Weather data regressions, other common/fun SDev topics [coding]
   - Remote sensing data, other common/fun SDev topics

Causal inference roadmap

- *Potential outcomes* [framework] [today]
    - Causal effect is difference between two potential outcomes
- *Identification* [application/implementation]
    - Identifying assumptions needed for a statistical estimate to have causal interpretation
    - Removing selection bias in regressions
    - E.g., RD, IV, ...
- *Estimation* [application/implementation] [last week]
    - (Usually) use linear regression model

Outline

Workshop outline

Potential outcomes framework

An alternative framework, the DAG

Summary

Causal inference roadmap

- *Potential outcomes (PO)* [framework]
  - Causal effect is difference between two potential outcomes
  - Lingua franca for expressing causal statements in economics / social sciences
  - This is one “approach to causality” (Imbens 2020)
    - Builds on Neyman 1923
    - Extended to observational studies by Rubin 1974
  - PO framework is not the *only* approach
    - Directed Acyclic Graph (DAG) approach is another alternative

- *Identification* [application/implementation]

- *Estimation* [application/implementation]

Outline

Workshop outline

Potential outcomes framework

- The original selection bias problem
- Treatment effects as a linear regression
- When does IA/CIA not hold?

An alternative framework, the DAG

- A very basic overview of DAGs
- Comparative strengths and weaknesses of the PO and DAG approaches

Summary

Potential outcomes framework and treatment effects

- We have:
  - A population, of which we observe sample of units $i = 1, ..., N$
  - A binary treatment of interest $D_i \in \{0, 1\} \rightarrow$ want to estimate the causal effect of D on Y
  - Let unit i's potential outcomes be: Y_i^1 if received treatment, Y_i^0 otherwise
  - Let unit i's observable outcome be: Y_i

Potential outcomes framework and treatment effects

- We have:
  - A population, of which we observe sample of units $i = 1, ..., N$
  - A binary treatment of interest $D_i \in \{0, 1\} \rightarrow$ want to estimate the causal effect of D on Y
  - Let unit i's potential outcomes be: Y_i^1 if received treatment, Y_i^0 otherwise
  - Let unit i's observable outcome be: Y_i

- Note the difference between *potential* outcomes (Y_i^1, Y_i^0) and *observable* or "actual" outcomes (Y_i); can relate them according to: $Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$

Potential outcomes framework and treatment effects

- We have:
  - A population, of which we observe sample of units $i = 1, \ldots, N$
  - A binary treatment of interest $D_i \in \{0, 1\} \rightarrow$ want to estimate the causal effect of D on Y
  - Let unit i's potential outcomes be: Y_i^1 if received treatment, Y_i^0 otherwise
  - Let unit i's observable outcome be: Y_i

- Note the difference between *potential* outcomes (Y_i^1, Y_i^0) and *observable* or "actual" outcomes (Y_i); can relate them according to: $Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$

- Define:

individual treatment effects (TEs)	$Y_i^1 - Y_i^0 \; \forall i$	*ideally estimate; unknowable*
average treatment effect (ATE)	$\mathbb{E}[Y_i^1 - Y_i^0]$	*reasonably estimate; unknowable, but might be identifiable*
average treatment effect on the treated (ATT)	$\mathbb{E}[Y_i^1 - Y_i^0 \mid D_i = 1]$	*reasonably estimate; unknowable, but might be identifiable*
difference in average observed outcomes	$\mathbb{E}[Y_i \mid D_i = 1] - \mathbb{E}[Y_i \mid D_i = 0]$	*directly identifiable, straightforward estimator*

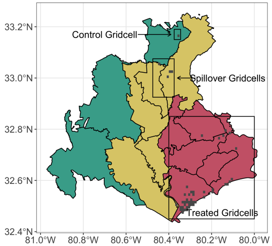
Potential outcomes framework and treatment effects: assumptions

- Assume additive treatment effects and no interference between units
- **Stable Unit Treatment Value Assumption (SUTVA)**: treatment received by one unit does not affect potential outcomes for other units
  - Each unit has only two possible potential outcomes Y_i^1, Y_i^0, which implies:
    - No spillovers
    - No general equilibrium effects

Potential outcomes framework and treatment effects: assumptions

Example of possible SUTVA violation:

- What are the effects of plastic bag laws on plastic litter in the environment?
- Use data on $\sim$100k shoreline cleanups
- Aggregate outcome data to 0.01 lat/lon gridcells
- Treatment at the zip code level (is there a policy in zip code?)
- Why may SUTVA be violated?



Potential outcomes framework and treatment effects: assumptions

- Assume additive treatment effects and no interference between units
- **Stable Unit Treatment Value Assumption (SUTVA)**: treatment received by one unit does not affect potential outcomes for other units
  - Each unit has only two possible potential outcomes Y_i^1, Y_i^0, which implies:
    - No spillovers
    - No general equilibrium effects
  - Often not realistic in economics studies
  - Many papers on SUTVA as nuisance
    - Can change how treatment is defined (e.g., within-household spillover)
    - Change level at which you interpret results
  - Some papers on SUTVA as substance (modeling the impact of the interference between units), e.g., spillovers:
    - (Hong and Raudenbush 2006; Hudgens and Halloran 2008; Aronow and Samii 2017; Rosenbaum 2007)

Potential outcomes framework and the selection bias problem

- Back to the various parameters:

individual treatment effects (TEs)	$Y_i^1 - Y_i^0 \ \forall i$	*ideally estimate; unknowable*
average treatment effect (ATE)	$\mathbb{E}[Y_i^1 - Y_i^0]$	*reasonably estimate; unknowable, but might be identifiable*
average treatment effect on the treated (ATT)	$\mathbb{E}[Y_i^1 - Y_i^0 \vert D_i = 1]$	*reasonably estimate; unknowable, but might be identifiable*
difference in average observed outcomes	$\mathbb{E}[Y_i \vert D_i = 1] - \mathbb{E}[Y_i \vert D_i = 0]$	*directly identifiable, straightforward estimator*

- We never observe causal effects

- What we can do is compute the difference in average observed outcomes:

$$\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \dots$$

$$= \underbrace{\mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]}_{\text{ATT}} + \underbrace{\mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]}_{\text{selection bias}}$$

▸ math details

Potential outcomes framework and the selection bias problem

- What we can do is compute the difference in average observed outcomes:

$$\begin{aligned}\mathbb{E}[Y_i|D_i=1]-\mathbb{E}[Y_i|D_i=0] &= \ldots \\ &= \underbrace{\mathbb{E}[Y_i^1-Y_i^0|D_i=1]}_{\text{ATT}}+\underbrace{\mathbb{E}[Y_i^0|D_i=1]-\mathbb{E}[Y_i^0|D_i=0]}_{\text{selection bias}}\end{aligned}$$

- Selection bias is the average difference in Y_i^0 between the treated and untreated
- Necessary and sufficient condition to remove selection bias and thus identify ATT: $\mathbb{E}[Y_i^0|D_i=1]=\mathbb{E}[Y_i^0|D_i=0]$
- $(Y_i^0, Y_i^1)\perp\!\!\!\perp D_i \implies Y_i^0 \perp\!\!\!\perp D_i \implies \mathbb{E}[Y_i^0|D_i=1]=\mathbb{E}[Y_i^0|D_i=0]$

▸ math details

Independence assumption and selection bias

- When treatment is independent of POs $\rightarrow$ no selection bias *in expectation*
  - $(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i$; **independence assumption** (IA)
  - Selection bias is eliminated and $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]$ or difference in average observed outcomes equals the ATT. Note that IA is sufficient but not necessary.
  - Holds in expectation for experiments, not for (virtually any) observational study

Independence assumption and selection bias

- When treatment is independent of POs $\rightarrow$ no selection bias *in expectation*
  - $(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i$; **independence assumption** (IA)
  - Selection bias is eliminated and $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]$ or difference in average observed outcomes equals the ATT. Note that IA is sufficient but not necessary.
  - Holds in expectation for experiments, not for (virtually any) observational study

- Let's consider various assignment mechanisms:
  - Random assignment (e.g., experiments)
  - Selection on observables
  - Selection on unobservables

Independence assumption and selection bias

- When treatment is independent of POs $\rightarrow$ no selection bias *in expectation*
  - $(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i$; **independence assumption** (IA)
  - Selection bias is eliminated and $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]$ or difference in average observed outcomes equals the ATT. Note that IA is sufficient but not necessary.
  - Holds in expectation for experiments, not for (virtually any) observational study

- Let's consider various assignment mechanisms:
  - Random assignment (e.g., experiments)
    - If treatment is randomly assigned, IA holds and identifies ATT (no selection bias in expectation, **NOT** for any single trial)
  - Selection on observables
  - Selection on unobservables

Independence assumption and selection bias

- When treatment is independent of POs $\rightarrow$ no selection bias *in expectation*
  - $(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i$; **independence assumption** (IA)
  - Selection bias is eliminated and $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]$ or difference in average observed outcomes equals the ATT. Note that IA is sufficient but not necessary.
  - Holds in expectation for experiments, not for (virtually any) observational study

- Let's consider various assignment mechanisms:
  - Random assignment (e.g., experiments)
  - Selection on observables
    - If conditional on some pre-treatment characteristic X_i, we have $(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i|X_i$, we can once again eliminate selection bias in expectation (**conditional independence assumption**, CIA)
    - Compare outcomes within each stratum of X_i
  - Selection on unobservables

Independence assumption and selection bias

- When treatment is independent of POs → no selection bias *in expectation*
  - $(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i$; **independence assumption** (IA)
  - Selection bias is eliminated and $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]$ or difference in average observed outcomes equals the ATT. Note that IA is sufficient but not necessary.
  - Holds in expectation for experiments, not for (virtually any) observational study

- Let's consider various assignment mechanisms:
  - Random assignment (e.g., experiments)
  - Selection on observables
  - Selection on unobservables
    - Will need other identification strategies to eliminate selection bias

Identifying assumptions

- We can recover an **unbiased** estimator of a causal effect if **(conditional) independence assumption** holds:
  - if IA holds ($(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i$) $\rightarrow$ estimate ATT and ATE
  - if $\cancel{IA}$, but CIA ($(Y_i^0, Y_i^1) \perp\!\!\!\perp D_i | X_i$) $\rightarrow$ can estimate ATT and ATE in each stratum (and then combine)
  - if $\cancel{CIA}$, need relevant exogenous source of variation in D_i (e.g., $(Y_i^0, Y_i^1) \perp\!\!\!\perp Z_i,\ Z_i \not\!\perp\!\!\!\perp D_i$) $\rightarrow$ estimate a LATE

- Need an **identification strategy** that convinces us that IA holds

- Bottom-line:
  - Econometrics / regression controls won't bring causality $\rightarrow$ need identification strategy
  - BUT, even with good identification strategy, no reason to expect balance for all relevant pre-treatment characteristics $\rightarrow$ control for relevant pre-treatment variables

Outline

Treatment effects as a linear regression

- *Potential outcomes (PO)* [framework] [just now]
- *Identification* [application/implementation]
- *Estimation* [application/implementation] [last week]

Treatment effects as a linear regression

- Suppose heterogeneous TE: $Y_i^1 - Y_i^0 = \beta_i$
    - Denote ATT by β; that is, $\beta = \mathbb{E}[\beta_i | D_i = 1] = \mathbb{E}[Y_i^1 - Y_i^0 | D_i = 1]$

- Then we can write

$$
\begin{aligned}
Y_i &= Y_i^0 + (Y_i^1 - Y_i^0) D_i \\
&= Y_i^0 + \beta_i D_i \\
&= Y_i^0 + (\beta_i - \beta + \beta) D_i \\
&= \mathbb{E}[Y_i^0] + \beta D_i + Y_i^0 - \mathbb{E}[Y_i^0] + (\beta_i - \beta) D_i \\
&= \alpha + \beta D_i + e_i
\end{aligned}
$$

where $\alpha \equiv Y_i^0$ and $e_i \equiv Y_i^0 - \mathbb{E}[Y_i^0] + (\beta_i - \beta) D_i$

Treatment effects as a linear regression

- Think about OLS regression of Y_i on D_i and a constant:

$$\hat{\beta}_{ols} = \bar{Y}_1 - \bar{Y}_0$$

  - $\mathbb{E}[\hat{\beta}_{ols}] = \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0]$

- By $Y = \alpha + \beta D_i + e_i$ and the definition of e_i, we have
  - $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \beta + \mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0]$
  - $\mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0] = \mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]$

math details, part 1 | math details, part 2 | math details, part 3

Treatment effects as a linear regression

- Think about OLS regression of Y_i on D_i and a constant:

$$\hat{\beta}_{ols} = \bar{Y}_1 - \bar{Y}_0$$

  - $\mathbb{E}[\hat{\beta}_{ols}] = \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0]$

- By $Y = \alpha + \beta D_i + e_i$ and the definition of e_i, we have
  - $\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \beta + \mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0]$
  - $\mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0] = \mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]$

- Thus:

$$\begin{aligned}\mathbb{E}[\hat{\beta}_{ols}] &= \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] \\ &= \beta + \mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0] \\ &= \underbrace{\beta}_{\text{ATT}} + \underbrace{\mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]}_{\text{selection bias}}\end{aligned}$$

math details, part 1 | math details, part 2 | math details, part 3

Treatment effects as a linear regression

- Thus:

$$\begin{aligned}
\mathbb{E}[\hat{\beta}_{ols}] &= \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] \\
&= \beta + \mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0] \\
&= \underbrace{\beta}_{\text{ATT}} + \underbrace{\mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]}_{\text{selection bias}}
\end{aligned}$$

- Which means that:
  - $\mathbb{E}[\hat{\beta}_{ols}]$ is equal to ATT iff:
    - there is no selection bias (identification problem; independence)

Outline

Endogeneity

- In simple linear regression model $y_i = \alpha + \beta x_i + e_i$, variable x_i is:
  - **endogenous** if it is correlated with the error term, or $cov[x_i, e_i] \neq 0$
  - **exogenous** otherwise, if $cov[x_i, e_i] = 0$ (A3. of CLRM)
- If x is endogenous, then OLS estimator of β will be biased and inconsistent for β
- In our setting (potential outcomes framework), if treatment D_i is endogenous ($cov[D_i, e_i] \neq 0$), there is imbalance in potential outcomes across treatment groups
  - $\rightarrow$ CIA doesn't hold (again, identification problem $\leftrightarrow$ regression problem)

Sources of endogeneity

- **Reverse causality or simultaneity**
  - If y also affects D, this is captured by e, making e correlated with D
- **Measurement error** in D that is **correlated** with y
- **Omitted variable bias (OVB)**
  - If omitted variable w is correlated with D, e is correlated with D (w is a "confounding variable")

→ in observational studies, excluding confounder creates bias, so must adjust for all confounders; but we can rarely be certain to have measured all confounders, which is why we turn to alternative **"identification" strategies**

Outline

Outline

Directed acyclic graphs (DAGs)

- An alternative to the potential outcomes framework is the causal graph framework or work on directed acyclic graphs (DAGs) (Pearl 2009)
  - PO and DAG frameworks are *not* contradicting; both define causality using counterfactuals
  - Each framework has its own benefits (see Imbens 2020 for a review of these) and are therefore complementary perspectives

Directed acyclic graphs (DAGs)

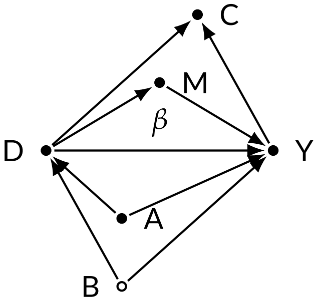
- Relationships between random variables are encoded with nodes and directed edges
  - Nodes are random variables (solid for observed variables, hollow for unobserved)
  - Arrows represent possible direct causal relationships
  - Paths are sequences of edges
  - DAG is a *complete* encoding of assumptions about causal relationships

Directed acyclic graphs (DAGs)

- Relationships between random variables are encoded with nodes and directed edges
  - Nodes are random variables (solid for observed variables, hollow for unobserved)
  - Arrows represent possible direct causal relationships
  - Paths are sequences of edges
  - DAG is a *complete* encoding of assumptions about causal relationships

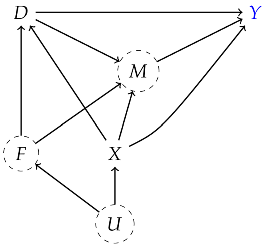
  - Types of elementary paths:
    - **Mediating path:** $D \rightarrow M \rightarrow Y$
    - **Confounding paths:** $D \leftarrow A \rightarrow Y$ (closed); $D \leftarrow\!-\!- B -\!-\!\rightarrow Y$ (open)
    - **Colliding path:** $D \rightarrow C \leftarrow Y$

  - Identification strategies:
    - Blocking back-door paths (adjusting for all confounders)
    - Instruments (alternative identification strategies)
      $\rightarrow$ same conclusion as with potential outcomes framework



401(k) eligibility and financial wealth example (Chernozhukov et al. 2025)

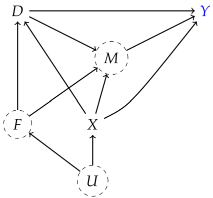
- **Question:** What is the causal effect of 401(k) *eligibility* on household financial wealth?
- **Variables:**
  - D: eligibility for a 401(k) plan
  - Y: net financial assets
  - X: observed worker characteristics (income, age)
  - F: unobserved firm characteristics
  - M: employer match
  - U: latent factors



- Eligibility may affect wealth:
  - directly (saving incentives, tax advantages)
  - indirectly through employer matching
- Selection into eligible jobs depends on X and unobserved F

Backdoor blocking and identification

- **Goal:** identify the causal effect of D on Y.
- **Backdoor criterion:** choose a set Z such that
  1. no variable in Z is a **descendant** of D, and
  2. Z **blocks every backdoor path** from D to Y (a backdoor path starts with an arrow into D).



- **Why?**
  - Rule (1) prevents blocking causal paths (e.g. $D \to M \to Y$).
  - Rule (2) removes non-causal association due to confounding.
- **In the 401(k) DAG, backdoor paths include:**
  - $D \leftarrow X \to Y$
  - $D \leftarrow F \to M \to Y$
  - $D \leftarrow F \leftarrow U \to X \to Y$
- **Valid adjustment set:** $Z = \{F, X\}$ blocks all backdoor paths and implies

$$Y(d) \perp\!\!\!\perp D \mid F, X.$$

Outline

Strengths and weaknesses of the PO and DAG approaches

- See Imbens 2020 for a review of the relevance of DAGs for empirical economics
  - **Experiments and manipulability:**
    $\rightarrow$ PO framework elevates randomized experiments as “gold standard”, while DAG doesn’t deem experiments special ($\sim$ notion of manipulability)
  - **Parts of causal analysis addressed: (pre-identification, identification, post-identification)**
    $\rightarrow$ DAGs only consider step 2, while steps 2 and 3 are considered jointly in PO
  - **Representation of identifying assumptions and identification strategies**
    $\rightarrow$ Identifying assumptions explicit in graphical versions and often much clearer than algebraic versions, BUT many other assumptions not easily captures in DAG framework; accounting for treatment heterogeneity difficult with DAGs

- Bottom-line:
  - Can be very helpful for thinking about or communicating research designs
  - May be helpful to know how to represent your analysis in both frameworks

Outline

Causal inference roadmap

- *Potential outcomes* [framework] [today]
  - Causal effect is the difference between two potential outcomes
  - We can't observe this difference, but can see differences in average observed outcomes
  - If **(conditional) independence assumption** holds, can estimate unbiased ATT
- *Identification* [application/implementation] [up next!]
  - In most empirical settings, IA and CIA do not hold, which is why we need an **identification strategy**
  - Want to eliminate selection bias (identification problem)
- *Estimation* [application/implementation] [last week]
  - (Usually) use linear regression model
  - $\hat{\beta}_{OLS}$ unbiased estimator for ATT if e is uncorrelated with treatment (regression problem)

Questions? Comments?

Thank you!

References

Heavily based on Claire Palandri's 2022 version and Anna Papp's 2024 version of the Causal Inference Workshop.

Aronow, Peter M., and Cyrus Samii. 2017. "ESTIMATING AVERAGE CAUSAL EFFECTS UNDER GENERAL INTERFERENCE, WITH APPLICATION TO A SOCIAL NETWORK EXPERIMENT." *The Annals of Applied Statistics* 11 (4): 1912–1947. ISSN: 19326157, accessed January 24, 2024. http://www.jstor.org/stable/26362172.

Chernozhukov, Victor, Christian Hansen, Nathan Kallus, Martin Spindler, and Vasilis Syrgkanis. 2025. "Causal Inference via Directed Acyclic Graphs and Nonlinear Structural Equation Models." Chap. 7 in *Applied Causal Inference Powered by ML and AI*. Working manuscript. Online. https://causalml-book.org.

Hong, Guanglei, and Stephen W. Raudenbush. 2006. "Evaluating Kindergarten Retention Policy: A Case Study of Causal Inference for Multilevel Observational Data." *Journal of the American Statistical Association* 101 (475): 901–910. ISSN: 01621459, accessed January 24, 2024. http://www.jstor.org/stable/27590770.

Hudgens, Michael G., and M. Elizabeth Halloran. 2008. "Toward Causal Inference with Interference." *Journal of the American Statistical Association* 103 (482): 832–842. ISSN: 01621459, accessed January 24, 2024. http://www.jstor.org/stable/27640105.

Imbens, Guido W. 2020. "Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics." *Journal of Economic Literature* 58 (4): 1129–79. https://doi.org/10.1257/jel.20191597. https://www.aeaweb.org/articles?id=10.1257/jel.20191597.

Pearl, Judea. 2009. *Causality: models, reasoning, and inference*. New York: Cambridge University Press.

Rosenbaum, Paul R. 2007. "Interference between Units in Randomized Experiments." *Journal of the American Statistical Association* 102 (477): 191–200. ISSN: 01621459, accessed January 24, 2024. http://www.jstor.org/stable/27639831.

Rubin, Donald B. 1974. "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of*

Appendix

$$
\begin{aligned}
\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] \\
&= \mathbb{E}[Y_i^1|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0] \\
&= \mathbb{E}[Y_i^1|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 1] + \mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0] \\
&= \underbrace{\mathbb{E}[Y_i^1 - Y_i^0|D_i = 1]}_{\text{ATT}} + \underbrace{\mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]}_{\text{selection bias}}
\end{aligned}
$$

back

Appendix

$$Y_i = \lambda + \gamma D_i + \varepsilon_i, \quad \mathbb{E}[D_i \varepsilon_i] = 0$$

Projection ceofficient of this projection model:

$$\gamma = \frac{cov[Y_i, D_i]}{Var[D_i]} = \frac{\mathbb{E}[Y_i D_i] - \mathbb{E}[Y_i]\mathbb{E}[D_i]}{\mathbb{E}[D_i^2] - \mathbb{E}[D_i]^2} =$$

$$= \frac{\mathbb{E}[Y_i | D_i = 1] P(D_i = 1) - \big(\mathbb{E}[Y_i | D_i = 0] P(D_i = 0) + \mathbb{E}[Y_i | D_i = 1] P(D_i = 1)\big) P(D_i = 1)}{P(D_i = 1) - P(D_i = 1)^2}$$

$$= \frac{\big(\mathbb{E}[Y_i | D_i = 1] - \mathbb{E}[Y_i | D_i = 0]\big)\big(P(D_i = 1) - P(D_i = 1)^2\big)}{P(D_i = 1) - P(D_i = 1)^2}$$

$$= \mathbb{E}[Y_i | D_i = 1] - \mathbb{E}[Y_i | D_i = 0]$$

back

OLS with a binary regressor (1/2)

$$\hat{\beta}_{\text{OLS}} = \frac{\sum_{i=1}^{n} Y_i(D_i - \bar{D})}{\sum_{i=1}^{n} D_i(D_i - \bar{D})}, \qquad \bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i.$$

Define

$$N_1 = \sum_{i=1}^{n} D_i, \quad N_0 = n - N_1 \quad \left(\text{and thus } \bar{D} = \frac{N_1}{n}\right)$$

$$\bar{Y}_1 = \frac{1}{N_1}\sum_{i:D_i=1} Y_i, \quad \bar{Y}_0 = \frac{1}{N_0}\sum_{i:D_i=0} Y_i.$$

Denominator:

$$\sum_{i=1}^{n} D_i(D_i - \bar{D}) = \sum_{i=1}^{n}(D_i - \bar{D}D_i) = (1 - \bar{D})\sum_{i=1}^{n} D_i = \frac{N_0 N_1}{n}.$$

OLS with a binary regressor (2/2)

Numerator:

$$\sum_{i=1}^{n} Y_i(D_i - \bar{D}) = \sum_{i=1}^{n} Y_i D_i - \bar{D} \sum_{i=1}^{n} Y_i.$$

$$= N_1 \bar{Y}_1 - \bar{D}(N_1 \bar{Y}_1 + N_0 \bar{Y}_0) = \frac{N_1 N_0}{n}(\bar{Y}_1 - \bar{Y}_0).$$

Combining numerator and denominator:

$$\hat{\beta}_{\text{OLS}} = \frac{\frac{N_1 N_0}{n}(\bar{Y}_1 - \bar{Y}_0)}{\frac{N_1 N_0}{n}} = \bar{Y}_1 - \bar{Y}_0.$$

$$\boxed{\hat{\beta}_{\text{OLS}} = \bar{Y}_1 - \bar{Y}_0.}$$

back

Appendix

Since:

- $\mathbb{E}[Y_i|D_i = 1] = \alpha + \beta + \mathbb{E}[e_i|D_i = 1]$ and
- $\mathbb{E}[Y_i|D_i = 0] = \alpha + \mathbb{E}[e_i|D_i = 0]$

We then have:

$$\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = \beta + \mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0]$$

back

Appendix

Since:

$$e_i = Y_i^0 - \mathbb{E}[Y_i^0] + (\beta_i - \beta)D_i$$

We have:

$$\begin{aligned}
\mathbb{E}[e_i|D_i = 1] - \mathbb{E}[e_i|D_i = 0] \\
&= \mathbb{E}[\beta_i - \beta|D_i = 1] + e\mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0] - \mathbb{E}[Y_i^0|D_i = 0] + \mathbb{E}[Y_i^0] \\
&= \mathbb{E}[\beta_i|D_i = 1] - \beta + \mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0] \\
&= \mathbb{E}[Y_i^0|D_i = 1] - \mathbb{E}[Y_i^0|D_i = 0]
\end{aligned}$$

back