Causal Inference

Justin Grimmer

Associate Professor
Department of Political Science
University of Chicago

March 28th, 2018

Purpose, Scope, and Examples

Goal in causal inference is to assess the causal effect of some potential cause (e.g. an institution, intervention, policy, or event) on some outcome.

Examples of such research questions include...

What is the effect of:

- political institutions on corruption?
- voting technology on voting fraud?
- incumbency status on vote shares?
- peacekeeping missions on peace?
- mass media on voter preferences?
- church attendance on turnout?

What Do We Mean by Causal Inference?

As in all statistics, we must begin with a model of the reality we are interested in studying, such as:

$$y_i = \alpha + \tau D_i + X_i\beta + \epsilon_i$$

Key problems with regression:

- Endogeneity and omitted variable bias
- Misspecified functional form
- Heterogenous treatment effects

What Do We Mean by Causal Inference?

As in all statistics, we must begin with a model of the reality we are interested in studying, such as:

$$y_i = \alpha + \tau D_i + X_i\beta + \epsilon_i$$

Key problems with regression:

- Endogeneity and omitted variable bias
- Misspecified functional form
- Heterogenous treatment effects

What Do We Mean by Causal Inference?

As in all statistics, we must begin with a model of the reality we are interested in studying, such as:

$$y_i = \alpha + \tau D_i + X_i \beta + \epsilon_i$$

Key problems with regression:

- Endogeneity and omitted variable bias
- Misspecified functional form
- Heterogenous treatment effects

Neyman-Rubin Potential Outcomes Model

Figure 1: Neyman

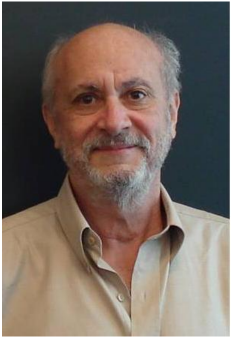
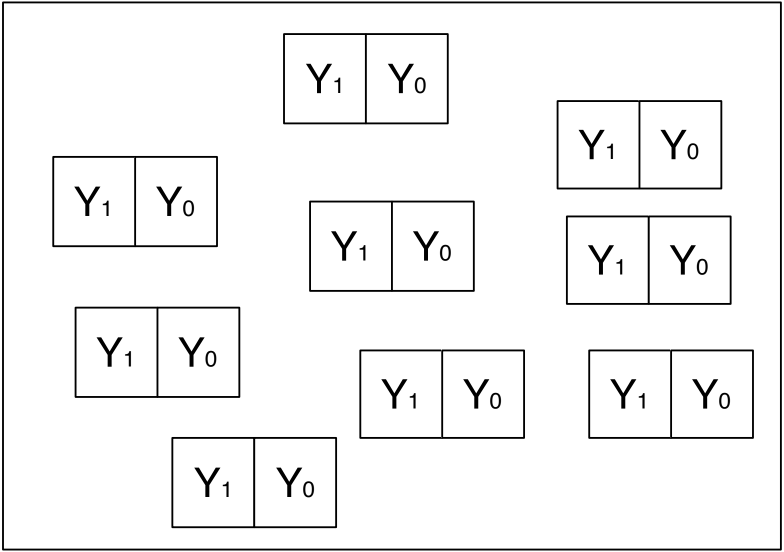
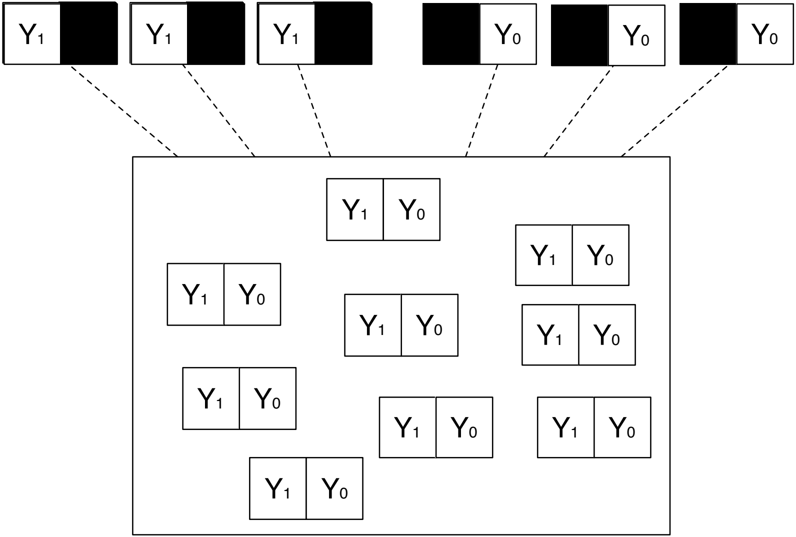


Figure 2: Rubin

Neyman Urn Model



Neyman Urn Model



Causality with Potential Outcomes

Definition (Treatment)

D_i: Indicator of treatment intake for *unit i*

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

Definition (Outcome)

Y_i: Observed outcome variable of interest for unit i. The treatment occurs temporally before the outcome.

Definition (Potential Outcomes)

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

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

Causality with Potential Outcomes

Definition (Causal Effect)

Causal effect of the treatment on the outcome for unit i is the difference between its two potential outcomes:

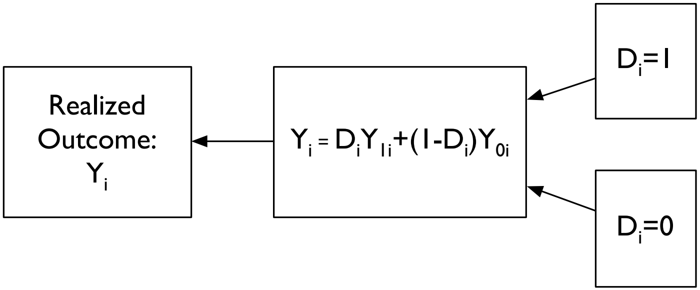
$$\tau_i = Y_{1i} - Y_{0i}$$

Assumption

Observed outcomes are realized as

$$Y_i = D_i \cdot Y_{1i} + (1 - D_i) \cdot Y_{0i} \text{ so } Y_i = \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases}$$

Causal Inference as a Missing Data Problem



Definition (Fundamental Problem of Causal Inference)

We cannot observe both potential outcomes. So how can we calculate $\tau_i = Y_{1i} - Y_{0i}$?

Fundamental Problem of Causal Inference

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What do we observe?

i	D_i	Y_{1i}	Y_{0i}	τ_i	Y_i
1	1	10	?	?	10
2	1	1	?	?	1
3	0	?	3	?	3
4	0	?	2	?	2

Causal inference is difficult because it involves missing data.

Fundamental Problem of Causal Inference

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What do we observe?

i	D_i	Y_{1i}	Y_{0i}	τ_i	Y_i
1	1	10	?	?	10
2	1	1	?	?	1
3	0	?	3	?	3
4	0	?	2	?	2

Causal inference is difficult because it involves missing data.

Causal Inference as a Missing Data Problem

How can we calculate $\tau_i = Y_{1i} - Y_{0i}$?

- Homogeneity is one solution:
  - If $\{Y_{1i}, Y_{0i}\}$ is constant across individuals, then cross-sectional comparisons will recover τ_i
  - If $\{Y_{1i}, Y_{0i}\}$ is constant across time, then before and after comparisons will recover τ_i

In social phenomena, unfortunately, homogeneity is very rare.

Causal Inference as a Missing Data Problem

How can we calculate $\tau_i = Y_{1i} - Y_{0i}$?

- Homogeneity is one solution:
  - If $\{Y_{1i}, Y_{0i}\}$ is constant across individuals, then cross-sectional comparisons will recover τ_i
  - If $\{Y_{1i}, Y_{0i}\}$ is constant across time, then before and after comparisons will recover τ_i

In social phenomena, unfortunately, homogeneity is very rare.

Other Assumptions

Assumption

Observed outcomes are realized as

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

- Embedded in this formulation is the assumption that potential outcomes for unit i are unaffected by treatment assignment for unit j.

- Assumption known by several names:
    - **S**table **U**nit **T**reatment **V**alue **A**ssumption (SUTVA)
    - No interference
    - Individualized Treatment Response

- Examples: vaccination, fertilizer on plot yield, communication

Other Assumptions

Assumption

Observed outcomes are realized as

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

- Embedded in this formulation is the assumption that potential outcomes for unit i are unaffected by treatment assignment for unit j.
- Assumption known by several names:
  - **S**table **U**nit **T**reatment **V**alue **A**ssumption (SUTVA)
  - No interference
  - Individualized Treatment Response
- Examples: vaccination, fertilizer on plot yield, communication

Potential Outcomes with Interference

Let $\mathbf{D} = \{D_i, D_j\}$ be the set of vectors of treatment assignments for two units i (me) and j (you).

How many elements in $\mathbf{D}$?

$$\mathbf{D} = \{(D_i = 0, D_j = 0), (D_i = 1, D_j = 0), (D_i = 0, D_j = 1), (D_i = 1, D_j = 1)\}$$

How many potential outcomes for unit i?

$$Y_{1i}(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) \\ Y_{1i}(1,0) \end{cases} \quad Y_{0i}(\mathbf{D}) = \begin{cases} Y_{0i}(0,1) \\ Y_{0i}(0,0) \end{cases}$$

Potential Outcomes with Interference

Let $\mathbf{D} = \{D_i, D_j\}$ be the set of vectors of treatment assignments for two units i (me) and j (you).

How many elements in $\mathbf{D}$?

$$\mathbf{D} = \{(D_i = 0, D_j = 0), (D_i = 1, D_j = 0), (D_i = 0, D_j = 1), (D_i = 1, D_j = 1)\}$$

How many potential outcomes for unit i?

$$Y_{1i}(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) \\ Y_{1i}(1,0) \end{cases} \quad Y_{0i}(\mathbf{D}) = \begin{cases} Y_{0i}(0,1) \\ Y_{0i}(0,0) \end{cases}$$

Potential Outcomes with Interference

Let $\mathbf{D} = \{D_i, D_j\}$ be the set of vectors of treatment assignments for two units i (me) and j (you).

How many elements in $\mathbf{D}$?

$$\mathbf{D} = \{(D_i = 0, D_j = 0), (D_i = 1, D_j = 0), (D_i = 0, D_j = 1), (D_i = 1, D_j = 1)\}$$

How many potential outcomes for unit i?

$$Y_{1i}(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) \\ Y_{1i}(1,0) \end{cases} \quad Y_{0i}(\mathbf{D}) = \begin{cases} Y_{0i}(0,1) \\ Y_{0i}(0,0) \end{cases}$$

Potential Outcomes with Interference

Let $\mathbf{D} = \{D_i, D_j\}$ be the set of vectors of treatment assignments for two units i (me) and j (you).

How many elements in $\mathbf{D}$?

$$\mathbf{D} = \{(D_i = 0, D_j = 0), (D_i = 1, D_j = 0), (D_i = 0, D_j = 1), (D_i = 1, D_j = 1)\}$$

How many potential outcomes for unit i?

$$Y_{1i}(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) \\ Y_{1i}(1,0) \end{cases} \quad Y_{0i}(\mathbf{D}) = \begin{cases} Y_{0i}(0,1) \\ Y_{0i}(0,0) \end{cases}$$

Potential Outcomes with Interference

How many causal effects for unit i?

$$\tau_i(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) - Y_{0i}(0,0) \\ Y_{1i}(1,1) - Y_{0i}(0,1) \\ Y_{1i}(1,0) - Y_{0i}(0,0) \\ Y_{1i}(1,0) - Y_{0i}(0,1) \\ Y_{1i}(1,1) - Y_{1i}(1,0) \\ Y_{0i}(0,1) - Y_{0i}(0,0) \end{cases}$$

How many potential outcomes are observed for unit i?

Since we only observe one of the four potential outcomes, the missing data problem for causal inference is even more severe.

Potential Outcomes with Interference

How many causal effects for unit i?

$$\tau_i(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) - Y_{0i}(0,0) \\ Y_{1i}(1,1) - Y_{0i}(0,1) \\ Y_{1i}(1,0) - Y_{0i}(0,0) \\ Y_{1i}(1,0) - Y_{0i}(0,1) \\ Y_{1i}(1,1) - Y_{1i}(1,0) \\ Y_{0i}(0,1) - Y_{0i}(0,0) \end{cases}$$

How many potential outcomes are observed for unit i?

Since we only observe one of the four potential outcomes, the missing data problem for causal inference is even more severe.

Potential Outcomes with Interference

How many causal effects for unit i?

$$\tau_i(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) - Y_{0i}(0,0) \\ Y_{1i}(1,1) - Y_{0i}(0,1) \\ Y_{1i}(1,0) - Y_{0i}(0,0) \\ Y_{1i}(1,0) - Y_{0i}(0,1) \\ Y_{1i}(1,1) - Y_{1i}(1,0) \\ Y_{0i}(0,1) - Y_{0i}(0,0) \end{cases}$$

How many potential outcomes are observed for unit i?

Since we only observe one of the four potential outcomes, the missing data problem for causal inference is even more severe.

Potential Outcomes with Interference

How many causal effects for unit i?

$$\tau_i(\mathbf{D}) = \begin{cases} Y_{1i}(1,1) - Y_{0i}(0,0) \\ Y_{1i}(1,1) - Y_{0i}(0,1) \\ Y_{1i}(1,0) - Y_{0i}(0,0) \\ Y_{1i}(1,0) - Y_{0i}(0,1) \\ Y_{1i}(1,1) - Y_{1i}(1,0) \\ Y_{0i}(0,1) - Y_{0i}(0,0) \end{cases}$$

How many potential outcomes are observed for unit i?

Since we only observe one of the four potential outcomes, the missing data problem for causal inference is even more severe.

Potential Outcomes with Interference

The No Interference assumption states that unit i's potential outcomes depend on D_i, not $\mathbf{D}$:

$Y_{1i}(1,1) = Y_{1i}(1,0)$ and $Y_{0i}(0,1) = Y_{0i}(0,0)$

This assumption furthermore allows us to define the effect for unit i as $\tau_i = Y_{1i} - Y_{0i}$.

No interference is an example of an **exclusion restriction**. We rely on outside information to rule out the possibility of certain causal effects (e.g. you taking the treatment has no effect on my potential outcomes).

Note that traditional models like regression also involve an implicit SUTVA assumption (Y_i depends on X_i)

Potential Outcomes with Interference

The No Interference assumption states that unit i's potential outcomes depend on D_i, not $\mathbf{D}$:

$Y_{1i}(1,1) = Y_{1i}(1,0)$ and $Y_{0i}(0,1) = Y_{0i}(0,0)$

This assumption furthermore allows us to define the effect for unit i as $\tau_i = Y_{1i} - Y_{0i}$.

No interference is an example of an **exclusion restriction**. We rely on outside information to rule out the possibility of certain causal effects (e.g. you taking the treatment has no effect on my potential outcomes).

Note that traditional models like regression also involve an implicit SUTVA assumption (Y_i depends on X_i)

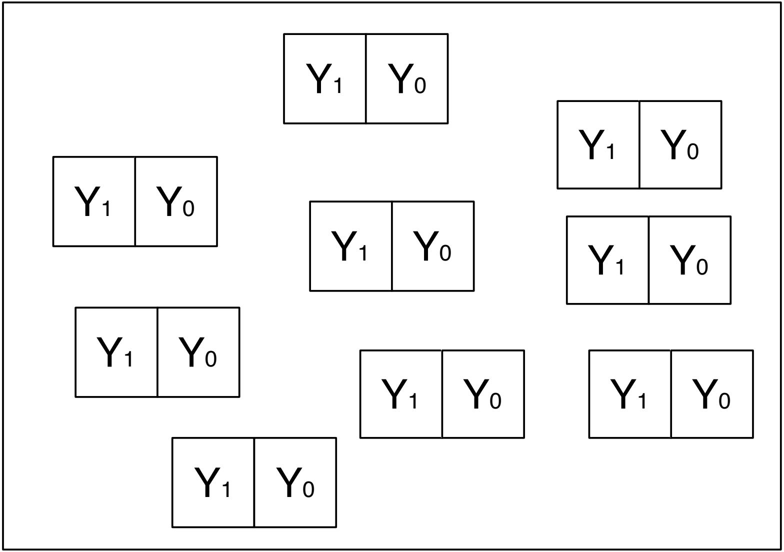
Potential Outcomes with Interference

Some Examples of Interference:

- Contagion
- Displacement
- Communication
- Deterrence

Causal inference in the presence of interference between subjects is an area of active research. Specially tailored experimental designs have been developed to study these interactions, e.g. Miguel and Kremer (2004) and Sinclair, McConnell, and Green (2012).

Back to the Neyman Urn Model



Estimands

Because τ_i are unobservable, we shift what we are interested in to:

Definition (Average Treatment Effect (ATE))

$$\tau_{ATE} = \text{Average of all treatment potential outcomes} - \text{Average of all control potential outcomes}$$

or

$$\tau_{ATE} = \frac{\sum_i^N Y_{1i}}{N} - \frac{\sum_i^N Y_{0i}}{N}$$

or

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

or

$$\tau_{ATE} = E[\tau_i]$$

Other Estimands

Definition (Average treatment effect on the treated (ATT))

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

Definition (Average treatment effect on the controls (ATC))

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

Definition (Average treatment effects for subgroups)

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

or

$$\tau_{ATT(X)} = E[Y_{1i} - Y_{0i} | D_i = 1, X_i = x]$$

Other Estimands

Definition (Average treatment effect on the treated (ATT))

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

Definition (Average treatment effect on the controls (ATC))

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

Definition (Average treatment effects for subgroups)

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

or

$$\tau_{ATT(X)} = E[Y_{1i} - Y_{0i} | D_i = 1, X_i = x]$$

Other Estimands

Definition (Average treatment effect on the treated (ATT))

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

Definition (Average treatment effect on the controls (ATC))

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

Definition (Average treatment effects for subgroups)

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

or

$$\tau_{ATT(X)} = E[Y_{1i} - Y_{0i} | D_i = 1, X_i = x]$$

Average Treatment Effect

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATE?

$E[Y_{1i} - Y_{0i}] = 1/4 \times (6 + -1 + 0 + 3) = 2$

Note: Average effect is positive, but τ_i are negative for some units!

Average Treatment Effect

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATE?

$E[Y_{1i} - Y_{0i}] = 1/4 \times (6 + -1 + 0 + 3) = 2$

Note: Average effect is positive, but τ_i are negative for some units!

Average Treatment Effect

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATE?

$E[Y_{1i} - Y_{0i}] = 1/4 \times (6 + -1 + 0 + 3) = 2$

Note: Average effect is positive, but τ_i are negative for some units!

Average Treatment Effect on the Treated

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATT and ATC?

$E[Y_{1i} - Y_{0i}|D_i = 1] = 1/2 \times (6 + -1) = 2.5$

$E[Y_{1i} - Y_{0i}|D_i = 0] = 1/2 \times (0 + 3) = 1.5$

Average Treatment Effect on the Treated

Imagine a study population with 4 units:

i	D_i	Y_{1i}	Y_{0i}	τ_i
1	1	10	4	6
2	1	1	2	-1
3	0	3	3	0
4	0	5	2	3

What is the ATT and ATC?

$$E[Y_{1i} - Y_{0i} | D_i = 1] = 1/2 \times (6 + -1) = 2.5$$

$$E[Y_{1i} - Y_{0i} | D_i = 0] = 1/2 \times (0 + 3) = 1.5$$

Naive Comparison: Difference in Means

Comparisons between *observed* outcomes of treated and control units can often be misleading.

$$
\begin{aligned}
E[Y_i|D=1] - E[Y_i|D_i=0} \\
&= E[Y_{1i}|D_i=1] - E[Y_{0i}|D_i=0] \\
&= \underbrace{E[Y_{1i} - Y_{0i}|D_i=1]}_{\text{ATT}} + \underbrace{\{E[Y_{i0}|D_i=1] - E[Y_{0i}|D_i=0]\}}_{\text{BIAS}}
\end{aligned}
$$

- Bias term unlikely to be 0 in most applications.
- Selection into treatment is often associated with the potential outcomes.

Selection Bias

$$
\begin{aligned}
E[Y_i|D_i = 1] - E[Y_i|D_i = 0]& \\
&= E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0] \\
&= \underbrace{E[Y_{1i} - Y_{0i}|D_i = 1]}_{\text{ATT}} + \underbrace{\{E[Y_{i0}|D_i = 1] - E[Y_{0i}|D_i = 0]\}}_{\text{BIAS}}
\end{aligned}
$$

Example: Church Attendance and Political Participation

- Churchgoers are likely to differ from non-churchgoers on a range of background characteristics (e.g. civic duty).
- Given these differences, turnout for churchgoers would be higher than for non-churchgoers even if churchgoers never attended church or church had zero mobilizing effect ($E[Y_0|D = 1] - E[Y_0|D = 0] > 0$).

Selection Bias

$$\begin{aligned}
E[Y_i|D_i = 1] - E[Y_i|D_i = 0] & \\
&= E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0] \\
&= \underbrace{E[Y_{1i} - Y_{0i}|D_i = 1]}_{\text{ATT}} + \underbrace{\{E[Y_{i0}|D_i = 1] - E[Y_{0i}|D_i = 0]\}}_{\text{BIAS}}
\end{aligned}$$

Example: Gender Quotas and Redistribution Towards Women

- Countries with gender quotas are likely countries where women are politically mobilized.
- Given this difference, policies targeted towards women are more common in quota countries even if these countries had not adopted quotas ($E[Y_0|D = 1] - E[Y_0|D = 0] > 0$).

Regression to Estimate the Average Treatment Effect

What happens when you run a regression of the observed outcome on the treatment indicator to estimate the ATE?

The ATE can be expressed as a regression equation:

$$
\begin{aligned}
Y_i &= D_i\, Y_{1i} + (1 - D_i)\, Y_{0i} \\
&= Y_{0i} + (Y_{1i} - Y_{0i})\, D_i \\
&= \underbrace{\bar{Y}_0}_{\alpha} + \underbrace{(\bar{Y}_1 - \bar{Y}_0)}_{\tau_{Reg}}\, D_i + \underbrace{\{(Y_{i0} - \bar{Y}_0) + D_i \cdot [(Y_{i1} - \bar{Y}_1) - (Y_{i0} - \bar{Y}_0)]\}}_{\epsilon} \\
&= \alpha + \tau_{Reg}\, D_i + \epsilon_i
\end{aligned}
$$

- τ_{Reg} could be biased and inconsistent for τ_{ATE} in two ways:
    - Baseline difference in potential outcomes under control that is correlated with D_i.
    - Individual treatment effects τ_i are correlated with D_i
- Effect heterogeneity implies "heteroskedasticity", i.e. error variance differs by values of D_i.
    - Neyman model imples "robust" standard errors.

Regression to Estimate the Average Treatment Effect

What happens when you run a regression of the observed outcome on the treatment indicator to estimate the ATE?

The ATE can be expressed as a regression equation:

$$
\begin{aligned}
Y_i &= D_i\, Y_{1i} + (1 - D_i)\, Y_{0i} \\
&= Y_{0i} + (Y_{1i} - Y_{0i})\, D_i \\
&= \underbrace{\bar{Y}_0}_{\alpha} + \underbrace{(\bar{Y}_1 - \bar{Y}_0)}_{\tau_{Reg}} D_i + \underbrace{\{(Y_{i0} - \bar{Y}_0) + D_i \cdot [(Y_{i1} - \bar{Y}_1) - (Y_{i0} - \bar{Y}_0)]\}}_{\epsilon} \\
&= \alpha + \tau_{Reg} D_i + \epsilon_i
\end{aligned}
$$

- τ_{Reg} could be biased and inconsistent for τ_{ATE} in two ways:
  - Baseline difference in potential outcomes under control that is correlated with D_i.
  - Individual treatment effects τ_i are correlated with D_i
- Effect heterogeneity implies "heteroskedasticity", i.e. error variance differs by values of D_i.
  - Neyman model imples "robust" standard errors.

Regression to Estimate the Average Treatment Effect

What happens when you run a regression of the observed outcome on the treatment indicator to estimate the ATE?

The ATE can be expressed as a regression equation:

$$
\begin{aligned}
Y_i &= D_i\, Y_{1i} + (1 - D_i)\, Y_{0i} \\
&= Y_{0i} + (Y_{1i} - Y_{0i})\, D_i \\
&= \underbrace{\bar{Y}_0}_{\alpha} + \underbrace{(\bar{Y}_1 - \bar{Y}_0)}_{\tau_{Reg}} D_i + \underbrace{\{(Y_{i0} - \bar{Y}_0) + D_i \cdot [(Y_{i1} - \bar{Y}_1) - (Y_{i0} - \bar{Y}_0)]\}}_{\epsilon} \\
&= \alpha + \tau_{Reg} D_i + \epsilon_i
\end{aligned}
$$

- τ_{Reg} could be biased and inconsistent for τ_{ATE} in two ways:
  - Baseline difference in potential outcomes under control that is correlated with D_i.
  - Individual treatment effects τ_i are correlated with D_i
- Effect heterogeneity implies "heteroskedasticity", i.e. error variance differs by values of D_i.
  - Neyman model imples "robust" standard errors.

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill them in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

Definition (Assignment Mechanism)

Assignment mechanism is the procedure that determines which units are selected for 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.

The Assignment Mechanism

- Since missing potential outcomes are unobservable we must make assumptions to fill them in, i.e. **estimate** missing potential outcomes.
- In the causal inference literature, we typically make assumptions about the **assignment mechanism** to do so.

Definition (Assignment Mechanism)

Assignment mechanism is the procedure that determines which units are selected for 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

Imagine a study population with 4 units:

i	$\Pr(D_i = 1)$	D_i	Y_{1i}	Y_{0i}	τ_i
1	?	1	10	4	6
2	?	1	1	2	-1
3	?	0	3	3	0
4	?	0	5	2	3

$\boxed{\text{No causation without manipulation?}}$

Always ask:

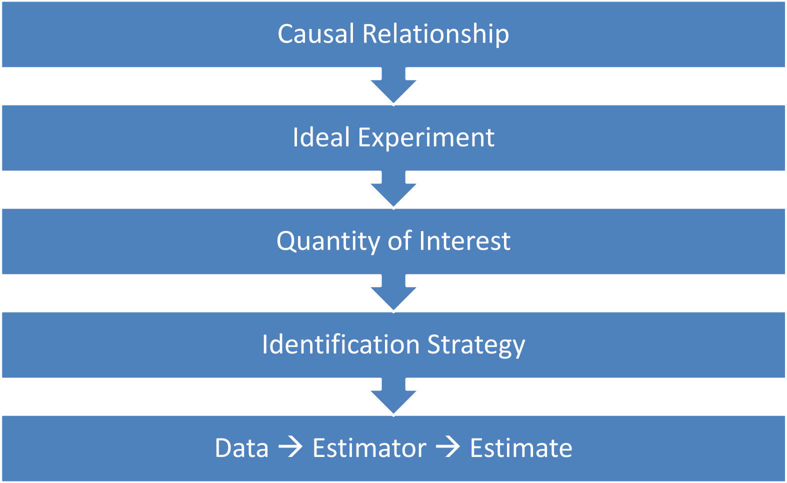
What is the ideal experiment you would run if you had infinite resources and power?

$\boxed{\text{No causation without manipulation?}}$

Always ask:

What is the ideal experiment you would run if you had infinite resources and power?

Causal Inference Workflow

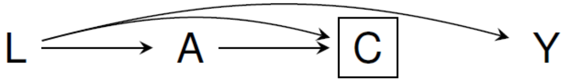


Summing Up: Neyman-Rubin causal model

- Useful for studying the "effects of causes," less so for the "causes of effects."
- No assumption of homogeneity, allows for causal effects to vary unit by unit.
    - No single "causal effect," thus the need to be precise about the target estimand.
- Distinguishes between *observed* outcomes and *potential* outcomes.
- Causal inference is a missing data problem: we typically make assumptions about the assignment mechanism to go from descriptive inference to causal inference.

Alternative Causal Models

The Neyman-Rubin causal model is popular in the social and health sciences, but alternatives exist:



- Structural Equation Modeling:
  - Write down causal model using Directed Acyclic Graphs (DAG)
  - Causal effects are defined by interventions that set variables to specified values in the causal model.
  - Set of axioms ("Do Calculus") that establish identifiablity of causal parameters given structure of the causal graph.
  - Can be re-expressed in potential outcome notation (though sometimes difficult!)
- Causality without Counterfactuals (Dawid 2000)