Section 11: Research Design for Causal Inference [1]

April 17, 2013

[1] Credit to Gary King, Patrick Lam, Maya Sen, and Brandon Stewart

Outline

Announcements

A few pieces of business:

- Last assignment due today

Announcements

A few pieces of business:

- Last assignment due today
- Gary's party on Friday

Announcements

A few pieces of business:

- Last assignment due today
- Gary's party on Friday
- Replications Due May 1st

Announcements

A few pieces of business:

- Last assignment due today
- Gary's party on Friday
- Replications Due May 1st
- Last section next week

What is a Causal Effect?

What is a Causal Effect?

$$Y_i(1) - Y_i(0)$$

What is a Causal Effect?

$$Y_i(1) - Y_i(0)$$

where $Y_i(1) = Y_i(T_i = 1)$ for some treatment variable T.

What is a Causal Effect?

$$Y_i(1) - Y_i(0)$$

where $Y_i(1) = Y_i(T_i = 1)$ for some treatment variable T.

$Y_i(1)$ and $Y_i(0)$ are *potential outcomes*

- represent the outcomes for individual i
- *if* they received treatment or control

What is a Causal Effect?

$$Y_i(1) - Y_i(0)$$

where $Y_i(1) = Y_i(T_i = 1)$ for some treatment variable T.

$Y_i(1)$ and $Y_i(0)$ are *potential outcomes*

- represent the outcomes for individual i
- *if* they received treatment or control

The **fundamental problem of causal inference:**

- only one of $Y_i(1)$ and $Y_i(0)$ is observed

Average Treatment Effects

Average treatment effect (ATE):

Average Treatment Effects

Average treatment effect (ATE):

$$E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$$

Average Treatment Effects

Average treatment effect (ATE):

$$E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$$

Average treatment effect on the treated (ATT):

Average Treatment Effects

Average treatment effect (ATE):

$$E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$$

Average treatment effect on the treated (ATT):

$$E[Y(1|T = 1) - Y(0|T = 1)] = E[Y_t(1) - Y_t(0)]$$

Average Treatment Effects

Average treatment effect (ATE):

$$E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$$

Average treatment effect on the treated (ATT):

$$E[Y(1|T = 1) - Y(0|T = 1)] = E[Y_t(1) - Y_t(0)]$$

ATT & ATE: can't estimate because of unobserved potential outcomes.

Fundamental Problem of Causal Inference

In an ideal world, we would see this:

$Unit_i$	X_i^1	X_i^2	X_i^3	T_i	$Y_i(0)$	$Y_i(1)$	$Y_i(1) - Y_i(0)$
1	2	1	50	0	69	75	6
2	3	1	98	0	111	108	-3
3	2	2	80	1	92	102	10
4	3	1	98	1	112	111	-1

Fundamental Problem of Causal Inference

But in the real world, we see this:

$Unit_i$	X_i^1	X_i^2	X_i^3	T_i	$Y_i(0)$	$Y_i(1)$	$Y_i(1) - Y_i(0)$
1	2	1	50	0	69	?	?
2	3	1	98	0	111	?	?
3	2	2	80	1	?	102	?
4	3	1	98	1	?	111	?

Fundamental Problem of Causal Inference

- For control units, $Y_i(1)$ is the **counterfactual**.

Fundamental Problem of Causal Inference

- For control units, $Y_i(1)$ is the **counterfactual**.
- For treatment units, $Y_i(0)$ is the **counterfactual**.

Fundamental Problem of Causal Inference

- For control units, $Y_i(1)$ is the **counterfactual**.
- For treatment units, $Y_i(0)$ is the **counterfactual**.

Alternative Quantities of Interest

Alternative Quantities of Interest

- Sample average treatment effect:

$$\text{SATE} \equiv \frac{1}{n/2} \sum_{i \in \{I_i = 1\}} \text{TE}_i$$

Alternative Quantities of Interest

- Sample average treatment effect:

$$\text{SATE} \equiv \frac{1}{n/2} \sum_{i \in \{I_i = 1\}} \text{TE}_i$$

- Sample average treatment effect on the treated:

$$\text{SATT} \equiv \frac{1}{n/2} \sum_{i \in \{I_i = 1, T_i = 1\}} \text{TE}_i$$

Alternative Quantities of Interest

Alternative Quantities of Interest

- Population average treatment effect:

$$\text{PATE} \equiv \frac{1}{N}\sum_i \text{TE}_i$$

Alternative Quantities of Interest

- Population average treatment effect:

$$\text{PATE} \equiv \frac{1}{N} \sum_{i} \text{TE}_i$$

- Population average treatment effect on the treated:

$$\text{PATT} \equiv \frac{1}{N^*} \sum_{i \in \{T_i = 1\}} \text{TE}_i$$

where $N^* = \sum_{i=1}^{N} T_i$ is the number of treated units in the population.

Outline

Decomposition of Causal Effect Estimation Error

Decomposition of Causal Effect Estimation Error

- Difference in means estimator: D

Decomposition of Causal Effect Estimation Error

- Difference in means estimator: D
- Estimation Error:

$$\Delta \equiv \text{PATE} - D$$

Decomposition of Causal Effect Estimation Error

- Difference in means estimator: D
- Estimation Error:

$$\Delta \equiv \text{PATE} - D$$

- Pretreatment confounders: X are observed and U are unobserved

Decomposition of Causal Effect Estimation Error

- Difference in means estimator: D
- Estimation Error:

$$\Delta \equiv \text{PATE} - D$$

- Pretreatment confounders: X are observed and U are unobserved
- Decomposition:

$$\begin{aligned} \Delta &= \Delta_S + \Delta_T \\ &= (\Delta_{S_X} + \Delta_{S_U}) + (\Delta_{T_X} + \Delta_{T_U}) \end{aligned}$$

Error due to Δ_S (sample selection), Δ_T (treatment imbalance), and each due to observed (X_i) and unobserved (U_i) covariates

Selection Error

Selection Error

- Definition:

$$\begin{aligned} \Delta_S &\equiv \text{PATE} - \text{SATE} \\ &= \frac{N-n}{N}(\text{NATE} - \text{SATE}), \end{aligned}$$

where NATE is the nonsample average treatment effect.

Selection Error

- Definition:

$$\begin{aligned}\Delta_S &\equiv \text{PATE} - \text{SATE} \\ &= \frac{N-n}{N}(\text{NATE} - \text{SATE}),\end{aligned}$$

  where NATE is the nonsample average treatment effect.
- Δ_S vanishes if:

Selection Error

- Definition:

$$\begin{aligned} \Delta_S &\equiv \text{PATE} - \text{SATE} \\ &= \frac{N-n}{N}(\text{NATE} - \text{SATE}), \end{aligned}$$

  where NATE is the nonsample average treatment effect.
- Δ_S vanishes if:
  1. The sample is a census ($I_i = 1$ for all observations and $n = N$);

Selection Error

- Definition:

$$\Delta_S \equiv \text{PATE} - \text{SATE} = \frac{N-n}{N}(\text{NATE} - \text{SATE}),$$

  where NATE is the nonsample average treatment effect.

- Δ_S vanishes if:
  1. The sample is a census ($I_i = 1$ for all observations and $n = N$);
  2. SATE = NATE; or

Selection Error

- Definition:

$$
\begin{aligned}
\Delta_S &\equiv \text{PATE} - \text{SATE} \\
&= \frac{N-n}{N}(\text{NATE} - \text{SATE}),
\end{aligned}
$$

  where NATE is the nonsample average treatment effect.

- Δ_S vanishes if:
  1. The sample is a census ($I_i = 1$ for all observations and $n = N$);
  2. SATE = NATE; or
  3. Switch quantity of interest from PATE to SATE.

Decomposing Selection Error

Decomposing Selection Error

- Decomposition:

$$\Delta_S = \Delta_{S_X} + \Delta_{S_U}$$

Decomposing Selection Error

- Decomposition:

$$\Delta_S = \Delta_{S_X} + \Delta_{S_U}$$

- $\Delta_{S_X} = 0$ when empirical distribution of (observed) X is identical in population and sample: $\widetilde{F}(X \mid I = 0) = \widetilde{F}(X \mid I = 1)$.

Decomposing Selection Error

- Decomposition:

$$\Delta_S = \Delta_{S_X} + \Delta_{S_U}$$

- $\Delta_{S_X} = 0$ when empirical distribution of (observed) X is identical in population and sample: $\widetilde{F}(X \mid I = 0) = \widetilde{F}(X \mid I = 1)$.
- $\Delta_{S_U} = 0$ when empirical distribution of (unobserved) U is identical in population and sample: $\widetilde{F}(U \mid I = 0) = \widetilde{F}(U \mid I = 1)$.

Decomposing Selection Error

- Decomposition:

$$\Delta_S = \Delta_{S_X} + \Delta_{S_U}$$

- $\Delta_{S_X} = 0$ when empirical distribution of (observed) X is identical in population and sample: $\widetilde{F}(X \mid I = 0) = \widetilde{F}(X \mid I = 1)$.
- $\Delta_{S_U} = 0$ when empirical distribution of (unobserved) U is identical in population and sample: $\widetilde{F}(U \mid I = 0) = \widetilde{F}(U \mid I = 1)$.
- conditions are unverifiable: X is observed only in sample and U is not observed at all.

Decomposing Selection Error

- Decomposition:

$$\Delta_S = \Delta_{S_X} + \Delta_{S_U}$$

- $\Delta_{S_X} = 0$ when empirical distribution of (observed) X is identical in population and sample: $\widetilde{F}(X \mid I = 0) = \widetilde{F}(X \mid I = 1)$.
- $\Delta_{S_U} = 0$ when empirical distribution of (unobserved) U is identical in population and sample: $\widetilde{F}(U \mid I = 0) = \widetilde{F}(U \mid I = 1)$.
- conditions are unverifiable: X is observed only in sample and U is not observed at all.
- Δ_{S_X} vanishes if weighting on X

Decomposing Selection Error

- Decomposition:

$$\Delta_S = \Delta_{S_X} + \Delta_{S_U}$$

- $\Delta_{S_X} = 0$ when empirical distribution of (observed) X is identical in population and sample: $\widetilde{F}(X \mid I = 0) = \widetilde{F}(X \mid I = 1)$.
- $\Delta_{S_U} = 0$ when empirical distribution of (unobserved) U is identical in population and sample: $\widetilde{F}(U \mid I = 0) = \widetilde{F}(U \mid I = 1)$.
- conditions are unverifiable: X is observed only in sample and U is not observed at all.
- Δ_{S_X} vanishes if weighting on X
- Δ_{S_U} cannot be corrected after the fact

Balance in Observed and Unobserved Covariates

Balance in Observed and Unobserved Covariates

- Decomposition:

$$\Delta_T = \Delta_{T_X} + \Delta_{T_U}$$

Balance in Observed and Unobserved Covariates

- Decomposition:

$$\Delta_T = \Delta_{T_X} + \Delta_{T_U}$$

- $\Delta_{T_X} = 0$ when X is balanced between treateds and controls:

$$\widetilde{F}(X \mid T = 1, I = 1) \;\; = \;\; \widetilde{F}(X \mid T = 0, I = 1).$$

  Verifiable from data; can be generated ex ante by blocking or enforced ex post via matching or parametric adjustment

Balance in Observed and Unobserved Covariates

- Decomposition:

$$\Delta_T = \Delta_{T_X} + \Delta_{T_U}$$

- $\Delta_{T_X} = 0$ when X is balanced between treateds and controls:

$$\widetilde{F}(X \mid T=1, I=1) \quad = \quad \widetilde{F}(X \mid T=0, I=1).$$

  Verifiable from data; can be generated ex ante by blocking or enforced ex post via matching or parametric adjustment

- $\Delta_{T_U} = 0$ when U is balanced between treateds and controls:

$$\widetilde{F}(U \mid T=1, I=1) \quad = \quad \widetilde{F}(U \mid T=0, I=1).$$

  Unverifiable. Achieved only by assumption or, on average, by random treatment assignment

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Complete blocking				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\rightarrow ?$	$\rightarrow ?$	$\rightarrow ?$	$\rightarrow ?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\rightarrow ?$	$\rightarrow ?$	$\rightarrow ?$	$\rightarrow ?$
Random treatment assignment			$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				
No selection bias				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				
No selection bias	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				
No selection bias	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Ignorability				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				
No selection bias	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Ignorability				$\stackrel{\text{avg}}{=} 0$

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\stackrel{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to?$	$\to?$	$\to?$	$\to?$
Random treatment assignment			$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				
No selection bias	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$		
Ignorability				$\stackrel{\text{avg}}{=} 0$
No omitted variables				

Effects of Design Components on Estimation Error

Design Choice	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Random sampling	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Stratified Random Sampling	$= 0$	$\overset{\text{avg}}{=} 0$		
Focus on SATE rather than PATE	$= 0$	$= 0$		
Weighting for nonrandom sampling	$= 0$	$=?$		
Large sample size	$\to ?$	$\to ?$	$\to ?$	$\to ?$
Random treatment assignment			$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Complete blocking			$= 0$	$=?$
Exact matching			$= 0$	$=?$
Assumption				
No selection bias	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$		
Ignorability				$\overset{\text{avg}}{=} 0$
No omitted variables				$= 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Ideal experiment	$= 0$	$\rightarrow 0$	$= 0$	$\rightarrow 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Ideal experiment	$= 0$	$\to 0$	$= 0$	$\to 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Ideal experiment	$= 0$	$\to 0$	$= 0$	$\to 0$
Randomized clinicial trials (Limited or no blocking)	$\neq 0$	$\neq 0$	$\stackrel{\text{avg}}{=} 0$	$\stackrel{\text{avg}}{=} 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Ideal experiment	$= 0$	$\to 0$	$= 0$	$\to 0$
Randomized clinicial trials (Limited or no blocking)	$\neq 0$	$\neq 0$	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Randomized clinicial trials (Full blocking)	$\neq 0$	$\neq 0$	$= 0$	$\overset{\text{avg}}{=} 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Ideal experiment	$= 0$	$\rightarrow 0$	$= 0$	$\rightarrow 0$
Randomized clinicial trials (Limited or no blocking)	$\neq 0$	$\neq 0$	$\overset{\text{avg}}{=} 0$	$\overset{\text{avg}}{=} 0$
Randomized clinicial trials (Full blocking)	$\neq 0$	$\neq 0$	$= 0$	$\overset{\text{avg}}{=} 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Social Science Field Experiment (Limited or no blocking)	$\neq 0$	$\neq 0$	$\rightarrow 0$	$\rightarrow 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Social Science Field Experiment (Limited or no blocking)	$\neq 0$	$\neq 0$	$\rightarrow 0$	$\rightarrow 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Social Science Field Experiment (Limited or no blocking)	$\neq 0$	$\neq 0$	$\to 0$	$\to 0$
Survey Experiment (Limited or no blocking)	$\to 0$	$\to 0$	$\to 0$	$\to 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Social Science Field Experiment (Limited or no blocking)	$\neq 0$	$\neq 0$	$\to 0$	$\to 0$
Survey Experiment (Limited or no blocking)	$\to 0$	$\to 0$	$\to 0$	$\to 0$
Observational Study (Representative data set, Well-matched)	≈ 0	≈ 0	≈ 0	$\neq 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Social Science Field Experiment (Limited or no blocking)	$\neq 0$	$\neq 0$	$\to 0$	$\to 0$
Survey Experiment (Limited or no blocking)	$\to 0$	$\to 0$	$\to 0$	$\to 0$
Observational Study (Representative data set, Well-matched)	≈ 0	≈ 0	≈ 0	$\neq 0$
Observational Study (Unrepresentative but partially, correctable data, well-matched)	≈ 0	$\neq 0$	≈ 0	$\neq 0$

The Benefits of Major Research Designs: Overview

	Δ_{S_X}	Δ_{S_U}	Δ_{T_X}	Δ_{T_U}
Social Science Field Experiment (Limited or no blocking)	$\neq 0$	$\neq 0$	$\to 0$	$\to 0$
Survey Experiment (Limited or no blocking)	$\to 0$	$\to 0$	$\to 0$	$\to 0$
Observational Study (Representative data set, Well-matched)	≈ 0	≈ 0	≈ 0	$\neq 0$
Observational Study (Unrepresentative but partially, correctable data, well-matched)	≈ 0	$\neq 0$	≈ 0	$\neq 0$
Observational Study (Unrepresentative data set, Well-matched)	$\neq 0$	$\neq 0$	≈ 0	$\neq 0$

Outline

Research Designs and Estimation Strategies

So here are some possible solutions

Research Designs and Estimation Strategies

So here are some possible solutions

1. Randomized or Experimental Studies

Research Designs and Estimation Strategies

So here are some possible solutions

1. Randomized or Experimental Studies
2. Observational studies

Research Designs and Estimation Strategies

So here are some possible solutions

1. Randomized or Experimental Studies
2. Observational studies
    - Matching

Research Designs and Estimation Strategies

So here are some possible solutions

1. Randomized or Experimental Studies
2. Observational studies
    - Matching
    - Instrumental Variables

Research Designs and Estimation Strategies

So here are some possible solutions

1. Randomized or Experimental Studies
2. Observational studies
    - Matching
    - Instrumental Variables
    - Regression Discontinuity Design

Research Designs and Estimation Strategies

So here are some possible solutions

1. Randomized or Experimental Studies
2. Observational studies
    - Matching
    - Instrumental Variables
    - Regression Discontinuity Design
    - Difference-in-Differences

Classical Randomized Experiment

The gold standard of scientific research.

Classical Randomized Experiment

The gold standard of scientific research.

- The simplest way (conceptually) to compare treated and control units.

Classical Randomized Experiment

The gold standard of scientific research.

- The simplest way (conceptually) to compare treated and control units.
- Units are randomly assigned to receive treatment and control.

Classical Randomized Experiment

The gold standard of scientific research.

- The simplest way (conceptually) to compare treated and control units.
- Units are randomly assigned to receive treatment and control.
- We still can't estimate individual-level causal effects, but we can estimate the population **average treatment effect**, $E[Y_i(1) - Y_i(0)]$.

Classical Randomized Experiment

The gold standard of scientific research.

- The simplest way (conceptually) to compare treated and control units.
- Units are randomly assigned to receive treatment and control.
- We still can't estimate individual-level causal effects, but we can estimate the population **average treatment effect**, $E[Y_i(1) - Y_i(0)]$.
- Principles of Experimentation: Replication, Randomization, Blocking

Classical Randomized Experiments

Note again our assumptions:

Classical Randomized Experiments

Note again our assumptions:

1. **Ignorability.** There are no factors out there that affect both the probability of treatment and the outcome.

Classical Randomized Experiments

Note again our assumptions:

1. **Ignorability.** There are no factors out there that affect both the probability of treatment and the outcome.
    - If the treatment was assigned in a truly random fashion, you are generally ok.

Classical Randomized Experiments

Note again our assumptions:

1. **Ignorability.** There are no factors out there that affect both the probability of treatment and the outcome.
   - If the treatment was assigned in a truly random fashion, you are generally ok.
   - But if it wasn't, then you are in trouble!

Classical Randomized Experiments

Note again our assumptions:

1. **Ignorability.** There are no factors out there that affect both the probability of treatment and the outcome.
    - If the treatment was assigned in a truly random fashion, you are generally ok.
    - But if it wasn't, then you are in trouble!
2. **SUTVA.** We have assumed that assigning treatment to one unit doesn't affect the outcome for another unit.

Classical Randomized Experiments

Note again our assumptions:

1. **Ignorability.** There are no factors out there that affect both the probability of treatment and the outcome.
    - If the treatment was assigned in a truly random fashion, you are generally ok.
    - But if it wasn't, then you are in trouble!
2. **SUTVA.** We have assumed that assigning treatment to one unit doesn't affect the outcome for another unit.
    - Not always reasonable if there are potential peer effects!

.

Classical Randomized Experiments

We also need to think about compliance issues.

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?
  - Never-taker: Unit never takes treatment

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units “take” the treatment as they were supposed to?
  - Never-taker: Unit never takes treatment
  - Always-taker: Unit always takes treatment

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?
  - Never-taker: Unit never takes treatment
  - Always-taker: Unit always takes treatment
  - Complier: Unit takes the treatment like they are supposed to

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?
    - Never-taker: Unit never takes treatment
    - Always-taker: Unit always takes treatment
    - Complier: Unit takes the treatment like they are supposed to
    - Defier: Unit takes treatment when not assigned and control when assigned

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?
  - Never-taker: Unit never takes treatment
  - Always-taker: Unit always takes treatment
  - Complier: Unit takes the treatment like they are supposed to
  - Defier: Unit takes treatment when not assigned and control when assigned
- Naive solution: Focus on "intent to treat" rather than on actual treatment.

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?
  - Never-taker: Unit never takes treatment
  - Always-taker: Unit always takes treatment
  - Complier: Unit takes the treatment like they are supposed to
  - Defier: Unit takes treatment when not assigned and control when assigned
- Naive solution: Focus on "intent to treat" rather than on actual treatment.
- Another solution: Use an IV approach with the intent to treat as an instrument.

Classical Randomized Experiments

We also need to think about compliance issues.

- Did the units "take" the treatment as they were supposed to?
  - Never-taker: Unit never takes treatment
  - Always-taker: Unit always takes treatment
  - Complier: Unit takes the treatment like they are supposed to
  - Defier: Unit takes treatment when not assigned and control when assigned
- Naive solution: Focus on "intent to treat" rather than on actual treatment.
- Another solution: Use an IV approach with the intent to treat as an instrument.

Other issues as well like missing data (measurement error and structural).

Observational Data

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

1. Gather dataset.

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

1. Gather dataset.
2. Estimate ATE or ATT with a model.

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

1. Gather dataset.
2. Estimate ATE or ATT with a model.

SUTVA: assumed (a problematic assumption most of the time)

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

1. Gather dataset.
2. Estimate ATE or ATT with a model.

SUTVA: assumed (a problematic assumption most of the time)

Ignorability: include covariates to get conditional ignorability

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

1. Gather dataset.
2. Estimate ATE or ATT with a model.

SUTVA: assumed (a problematic assumption most of the time)

Ignorability: include covariates to get conditional ignorability

$$(Y(1), Y(0)) \perp T | X$$

Observational Data

We have a dataset where we only observe after the experiment occurred and we have no control over treatment assignment.

This is the case with most of the sciences.

1. Gather dataset.
2. Estimate ATE or ATT with a model.

SUTVA: assumed (a problematic assumption most of the time)

Ignorability: include covariates to get conditional ignorability

$$(Y(1), Y(0)) \perp T | X$$

Treatment assignment is independent of the outcomes (Y) given covariates X.

Problems:

Problems:

- SUTVA assumption

Problems:

- SUTVA assumption
- Omitted variable bias

Problems:

- SUTVA assumption
- Omitted variable bias
    - Don't include all the variables that makes treatment assignment independent of Y.

Problems:

- SUTVA assumption
- Omitted variable bias
  - Don't include all the variables that makes treatment assignment independent of Y.
- Model Dependence

Problems:

- SUTVA assumption
- Omitted variable bias
  - Don't include all the variables that makes treatment assignment independent of Y.
- Model Dependence
  - We try to alleviate the curse of dimensionality and problem of continuous covariates by specifying a model.

Problems:

- SUTVA assumption
- Omitted variable bias
  - Don't include all the variables that makes treatment assignment independent of Y.
- Model Dependence
  - We try to alleviate the curse of dimensionality and problem of continuous covariates by specifying a model.
  - Estimates of ATE or ATT may differ depending on the model you specify.

A Quick Aside for Emphasis

What do we do 99% of the time in the literature

A Quick Aside for Emphasis

What do we do 99% of the time in the literature

We run a regression and interpret every coefficient

A Quick Aside for Emphasis

What do we do 99% of the time in the literature

We run a regression and interpret every coefficient

- What is the treatment? We can only interpret one coefficient causally.

A Quick Aside for Emphasis

What do we do 99% of the time in the literature

We run a regression and interpret every coefficient

- What is the treatment? We can only interpret one coefficient causally.
- Post-Treatment Bias

A Quick Aside for Emphasis

What do we do 99% of the time in the literature

We run a regression and interpret every coefficient

- What is the treatment? We can only interpret one coefficient causally.
- Post-Treatment Bias

We are estimating the conditional mean, not the causal effect.

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
    - perfect **balance**

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
  - perfect **balance**
  - differed only on treatment assignment

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
    - perfect **balance**
    - differed only on treatment assignment
    - $\Rightarrow$ perfect conditional ignorability

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
  - perfect **balance**
  - differed only on treatment assignment
  - $\Rightarrow$ perfect conditional ignorability
- Same treatment effect regardless of the model.

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
    - perfect **balance**
    - differed only on treatment assignment
    - $\Rightarrow$ perfect conditional ignorability
- Same treatment effect regardless of the model.

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
  - perfect **balance**
  - differed only on treatment assignment
  - $\Rightarrow$ perfect conditional ignorability
- Same treatment effect regardless of the model.
- Reduces model dependence

Matching: A way to Ameliorate Model Dependence

- If we had pairs of observations with:
  - perfect **balance**
  - differed only on treatment assignment
  - $\Rightarrow$ perfect conditional ignorability
- Same treatment effect regardless of the model.
- Reduces model dependence
- But remember: Don't match on post-treatment variables!

Instrumental Variables

Instrumental Variables

Goal: Estimate Causal Effects

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Solution so far: Include covariates and match

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Solution so far: Include covariates and match

Another solution: **Instrumental Variables**

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Solution so far: Include covariates and match

Another solution: **Instrumental Variables**

The idea: Find an instrument Z that is randomly assigned (or assignment is ignorable) and that affects Y only through T.

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Solution so far: Include covariates and match

Another solution: **Instrumental Variables**

The idea: Find an instrument Z that is randomly assigned (or assignment is ignorable) and that affects Y only through T.

Example: Y = post-Vietnam War civilian mortality;

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Solution so far: Include covariates and match

Another solution: **Instrumental Variables**

The idea: Find an instrument Z that is randomly assigned (or assignment is ignorable) and that affects Y only through T.

Example: Y = post-Vietnam War civilian mortality; T = serving in the military during Vietnam War;

Instrumental Variables

Goal: Estimate Causal Effects

Problem in Observational Data:
Non-ignorability of treatment assignment (and SUTVA)

Solution so far: Include covariates and match

Another solution: **Instrumental Variables**

The idea: Find an instrument Z that is randomly assigned (or assignment is ignorable) and that affects Y only through T.

Example: $Y =$ post-Vietnam War civilian mortality; $T =$ serving in the military during Vietnam War; $Z =$ draft lottery

The Potential Outcomes Approach

The Potential Outcomes Approach

Assumptions:

The Potential Outcomes Approach

Assumptions:

1. SUTVA:

The Potential Outcomes Approach

Assumptions:

1. SUTVA: Z_i does not affect T_j and Y_j and T_i does not affect Y_j for all $i \neq j$ (non-interference) and there is no variation in the treatment or the instrument.

The Potential Outcomes Approach

Assumptions:

1. SUTVA: Z_i does not affect T_j and Y_j and T_i does not affect Y_j for all $i \neq j$ (non-interference) and there is no variation in the treatment or the instrument.

Figure : SUTVA Assumption implies absence of dotted arrows.

$$\begin{tikzcd}
Z_i \arrow[r] \arrow[dr, dotted] \arrow[drr, dotted] & T_i \arrow[r] \arrow[dr, dotted] & Y_i \\
Z_j \arrow[r] & T_j \arrow[r] & Y_j
\end{tikzcd}$$

The Potential Outcomes Approach

Assumptions:

1. SUTVA: Z_i does not affect T_j and Y_j and T_i does not affect Y_j for all $i \neq j$ (non-interference) and there is no variation in the treatment or the instrument.

Figure : SUTVA Assumption implies absence of dotted arrows.

$$
\begin{tikzcd}
Z_i \arrow[r] \arrow[dr, dotted] \arrow[drr, dotted] & T_i \arrow[r] \arrow[dr, dotted] & Y_i \\
Z_j \arrow[r] & T_j \arrow[r] & Y_j
\end{tikzcd}
$$

Example: The veteran status and civilian mortality of any man was not affected by the draft status of others.

2. Random (Ignorable) Assignment of the Instrument Z

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction:

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

Figure : Exclusion assumption implies absence of dotted arrow.

$$
\begin{tikzcd}
Z_i \arrow[r] \arrow[rr, bend right, dotted] & D_i \arrow[r] & Y_i
\end{tikzcd}
$$

Example: Civilian mortality risk was not affected by draft status once veteran status is taken into account.

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

   Figure : Exclusion assumption implies absence of dotted arrow.

$$
\begin{tikzcd}
Z_i \arrow[r] \arrow[rr, bend right, dotted] & D_i \arrow[r] & Y_i
\end{tikzcd}
$$

   Example: Civilian mortality risk was not affected by draft status once veteran status is taken into account.

4. Nonzero Average Causal Effect of Z on T.

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

   Figure : Exclusion assumption implies absence of dotted arrow.

$$\begin{tikzcd} Z_i \arrow[r] \arrow[rr, bend right, dotted] & D_i \arrow[r] & Y_i \end{tikzcd}$$

   Example: Civilian mortality risk was not affected by draft status once veteran status is taken into account.

4. Nonzero Average Causal Effect of Z on T.
   Example: Having a low lottery number increases the average probability of service.

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

   Figure : Exclusion assumption implies absence of dotted arrow.

$$\begin{tikzcd} Z_i \arrow[r] \arrow[rr, bend right, dotted] & D_i \arrow[r] & Y_i \end{tikzcd}$$

   Example: Civilian mortality risk was not affected by draft status once veteran status is taken into account.

4. Nonzero Average Causal Effect of Z on T.
   Example: Having a low lottery number increases the average probability of service.

5. Monotonicity:

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

   Figure : Exclusion assumption implies absence of dotted arrow.

$$\begin{tikzcd} Z_i \arrow[r] \arrow[rr, bend right, dotted] & D_i \arrow[r] & Y_i \end{tikzcd}$$

   Example: Civilian mortality risk was not affected by draft status once veteran status is taken into account.

4. Nonzero Average Causal Effect of Z on T.
   Example: Having a low lottery number increases the average probability of service.

5. Monotonicity: No Defiers

2. Random (Ignorable) Assignment of the Instrument Z
   Example: Assignment of draft status was random.

3. Exclusion Restriction: Any effect of Z on Y must be via an effect of Z on T.

   Figure : Exclusion assumption implies absence of dotted arrow.

$$\begin{tikzcd} Z_i \arrow[r] \arrow[rr, bend right, dotted] & D_i \arrow[r] & Y_i \end{tikzcd}$$

   Example: Civilian mortality risk was not affected by draft status once veteran status is taken into account.

4. Nonzero Average Causal Effect of Z on T.
   Example: Having a low lottery number increases the average probability of service.

5. Monotonicity: No Defiers
   Example: There is no one who would have served if given a high lottery number, but not if given a low lottery number.

If all the assumptions hold, then the **Local Average Treatment Effect (LATE)** of T on Y is

$$\text{LATE} = \frac{\text{Effect of Z on Y}}{\text{Effect of Z on T}}$$

If all the assumptions hold, then the **Local Average Treatment Effect (LATE)** of T on Y is

$$\text{LATE} = \frac{\text{Effect of Z on Y}}{\text{Effect of Z on T}}$$

It is only a local average treatment effect because it's the effect of T on Y for the subpopulation of compliers, and not the whole population.

If all the assumptions hold, then the **Local Average Treatment Effect (LATE)** of T on Y is

$$\text{LATE} = \frac{\text{Effect of Z on Y}}{\text{Effect of Z on T}}$$

It is only a local average treatment effect because it's the effect of T on Y for the subpopulation of compliers, and not the whole population.

Regression Discontinuity Design

Sometimes called "natural experiments." Here's how it works:

Regression Discontinuity Design

Sometimes called "natural experiments." Here's how it works:

- There is some pre-treatment variable x with a cutoff value – below the cut-off value the unit receives treatment and above the cutoff it receives control.

Regression Discontinuity Design

Sometimes called "natural experiments." Here's how it works:

- There is some pre-treatment variable x with a cutoff value – below the cut-off value the unit receives treatment and above the cutoff it receives control.
- The cutoff should be arbitrary and not related to the other covariates.

Regression Discontinuity Design

Sometimes called "natural experiments." Here's how it works:

- There is some pre-treatment variable x with a cutoff value – below the cut-off value the unit receives treatment and above the cutoff it receives control.
- The cutoff should be arbitrary and not related to the other covariates.
- There should be no real reason to think that the units below the cut-off are substantively that different from the units above the cut-off.

Regression Discontinuity Design

Sometimes called "natural experiments." Here's how it works:

- There is some pre-treatment variable x with a cutoff value – below the cut-off value the unit receives treatment and above the cutoff it receives control.
- The cutoff should be arbitrary and not related to the other covariates.
- There should be no real reason to think that the units below the cut-off are substantively that different from the units above the cut-off.
- We basically take care of ignorability. (Note: you still have to worry about SUTVA.)

Regression Discontinuity

Here are some examples:

Regression Discontinuity

Here are some examples:

- Criminals who are 17 yrs old, 364 days are treated as juveniles, but those who are 18yrs old, 1 day are treated as adults. What is the effect of adult sentencing?

Regression Discontinuity

Here are some examples:

- Criminals who are 17 yrs old, 364 days are treated as juveniles, but those who are 18yrs old, 1 day are treated as adults. What is the effect of adult sentencing?
- European municipalities with populations of 3499 do not need to have dissenting party members sit on their councils, but those with populations of 3500 do. What is the effect of having dissenting party members on council decisions?

Regression Discontinuity

Here are some examples:

- Criminals who are 17 yrs old, 364 days are treated as juveniles, but those who are 18yrs old, 1 day are treated as adults. What is the effect of adult sentencing?
- European municipalities with populations of 3499 do not need to have dissenting party members sit on their councils, but those with populations of 3500 do. What is the effect of having dissenting party members on council decisions?
- HS standardized test takers who get a score of 90.5 get a small scholarship, but those that get 89.5 don't. What is the effect of the scholarship on college success?

Regression Discontinuity

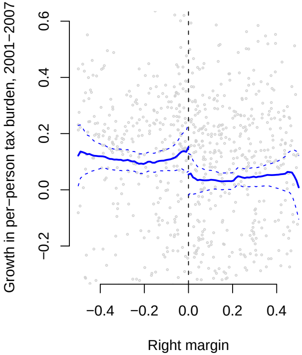
Here are some examples:

- Criminals who are 17 yrs old, 364 days are treated as juveniles, but those who are 18yrs old, 1 day are treated as adults. What is the effect of adult sentencing?
- European municipalities with populations of 3499 do not need to have dissenting party members sit on their councils, but those with populations of 3500 do. What is the effect of having dissenting party members on council decisions?
- HS standardized test takers who get a score of 90.5 get a small scholarship, but those that get 89.5 don't. What is the effect of the scholarship on college success?

It's sometimes easier to see this visually.

Regression Discontinuity (Eggers)

Example: In some French cities, the rightist party wins with 50.5 of the vote%; in other cities, they lose with 49.5% of the vote. What is the effect of rightist governments on tax policies?



Regression Discontinuity

Some thoughts about RDD designs:

Regression Discontinuity

Some thoughts about RDD designs:

- Always pay attention to the bandwidth size...

Regression Discontinuity

Some thoughts about RDD designs:

- Always pay attention to the bandwidth size...
- If you are working on elections: "Are Close Elections Random?" Grimmer, Hersh, Feinstein, Carpenter (2011)

Regression Discontinuity

Some thoughts about RDD designs:

- Always pay attention to the bandwidth size...
- If you are working on elections: "Are Close Elections Random?" Grimmer, Hersh, Feinstein, Carpenter (2011)
- Remember who you are generalizing to!

Difference-in-Differences

The Idea: We can use two units observed at two time points.

Difference-in-Differences

The Idea: We can use two units observed at two time points.

- Imagine two sets of units (treated and control, $T = 1, 0$) each observed at two time points ($D = 1, 2$)

Difference-in-Differences

The Idea: We can use two units observed at two time points.

- Imagine two sets of units (treated and control, $T = 1, 0$) each observed at two time points ($D = 1, 2$)
- There is some intervention on the treated units between the two time points but not the controls

Difference-in-Differences

The Idea: We can use two units observed at two time points.

- Imagine two sets of units (treated and control, $T = 1, 0$) each observed at two time points ($D = 1, 2$)
- There is some intervention on the treated units between the two time points but not the controls
- Take First Differences and Compare

Difference-in-Differences

The Idea: We can use two units observed at two time points.

- Imagine two sets of units (treated and control, $T = 1, 0$) each observed at two time points ($D = 1, 2$)
- There is some intervention on the treated units between the two time points but not the controls
- Take First Differences and Compare

$$(E[Y(D=2)|T=1] - E[Y(D=2)|T=0]) - (E[Y(D=1)|T=1] - E[Y(D=1)|T=0])$$

Similarity is critical! E.g. Card and Krueger on the minimum wage (1994)

References

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. “Identification of Causal Effects Using Instrumental Variables.” *Journal of the American Statistical Association* 91(434):444-455.

Angrist, Joshua David and Pischek, Jörn Steffen Pischke. *Mostly Harmless Econometrics.* Princeton University Press.

Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81(396): 945-960.

Pearl, Judea. *Causality: models, reasoning, and inference*.

Rubin, Donald B. *Matched Sampling for Causal Effects*