Potential Outcomes Causal Model

Rui Wang

Do public assistance increase the number of paupers (Yule, 1899)

Yule wants to know if public assistance increase the number of paupers in England

$$Pauper = \alpha + \delta Outrelief + \beta_1 Old + \beta_2 Pop + u$$

- Data from 1871 and 1881 census
- $Pauper$ the growth rate of pauper
- $Outrelief$ the growth rate of out-relief ratio (Measures on the strength of public assistance)
- Old and Pop are the growth rates of people aged above 65 and population

Results

Source	SS	df	MS
Model	5875.32014	3	1958.44005
Residual	2551.89861	28	91.1392359
Total	8427.21875	31	271.845766

Number of obs	=	32
F(3, 28)	=	21.49
Prob > F	=	0.0000
R-squared	=	0.6972
Adj R-squared	=	0.6647
Root MSE	=	9.5467

paup	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
outrelief	.7520945	.1349873	5.57	0.000	.4755856	1.028603
old	.055602	.2233568	0.25	0.805	-.4019236	.5131276
pop	-.3107383	.0668514	-4.65	0.000	-.4476771	-.1737995
_cons	63.18774	27.14388	2.33	0.027	7.58602	118.7895

A 10-percetage-point change in the out-relief growth rate is associated with a 7.5-percentage-point increase in the pauperism growth rate.

Does public assistance increase pauperism?

Correlation and causality

- 1. Correlation does not mean causality
  - Mortality rate due to drowning increases, when ice cream consumption increase
    - Did the ice cream cause death due to drowning?
  - Every morning, when the rooster crows, the sun rises
    - Did the rooster crowing cause the sun to rise?

Correlation and causality

- 2. No observed correlation does not mean no causality
- Examples from Scott Cunningham
  - A sailor is sailing her sailboat across a lake
  - Wind blows, and she perfectly counters by turning the rudder
  - The same aliens observe from space and say “Look at the way she’s moving that rudder back and forth but going in a straight line. That rudder is broken.”
  - So they send her a new rudder

No correlation doesn’t mean no causality. Artwork by Seth Hahne.

ED and Causality

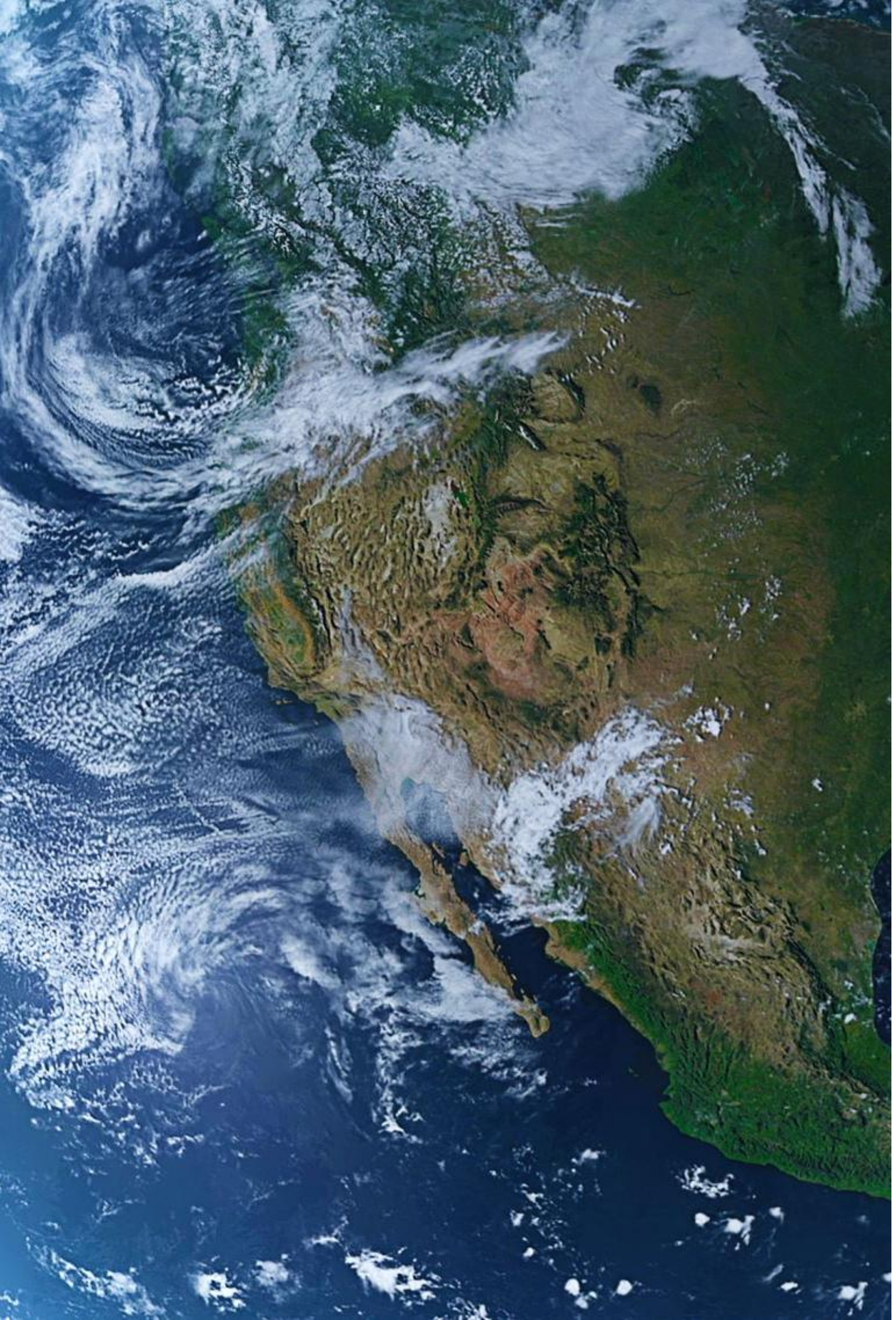
- 1. Many ED questions are policy and causal questions.
  - We want to know the causal effect of policy changes on outcomes
  - Does education increase individuals’ earnings?
  - Does class size affect academic performance?
  - Does attending college increase social mobility?
  - Does monetary incentive increase teachers’ performance?
  - Does health insurance increase individuals’ health status?
  - Does an increase in administrative burden reduce citizens’ satisfaction?
  - ….

Public policy and causality

- 2. Policy recommendations that ignore causality can backfire
  - If Yule's (1899) estimated relationship between public assistance and pauper is causal, we want to cut public assistance to combat pauperism.
  - Another example is hospital and health (from Angrist and Pischke, 2009)
    - 2005 NHIS, 1 poor health-5 excellent health

Groupz	Sample Size	Mean Health Status	Std. Error
Attend hospital	7,774	3.21	0.014
Not attend hospital	90,049	3.93	0.003

- Does going to hospital make people sicker?
- Can we improve peoples' health by closing hospitals?



Causality intuition

- A causal effect is defined as a comparison between the actual state of the world and the counterfactual state of the world.
  - Example: treatment in hospital

Does ACT preparation-class increase ACT scores

Student names	ACT score	Attend preparation-class?
1	14	Yes
2	22	No
3	18	Yes
4	29	No
5	14	Yes
6	24	No
7	36	No
8	22	No
9	25	No
10	32	Yes

- In 2020, Lakeview High School offered an ACT Pre-class to help their students prepare for ACT (American College Test).
- In the same year, 10 students took the ACT. The table on the left shows the ACT scores (Y_i) for those students as well as whether they had attended the ACT pre-class.
- Do you think the ACT preparation-class

Introduction to Neyman causal model (1923)

- Rubin 1974
- Let the treatment (college attendance) be a binary (dummy) variable:

D_i=

D_i=1 if i attended college

D_i=0 if i did not attend college

Where i indexes an individual observation, such as a person

- Potential outcomes (Inco

Y_{id}=

Y_{i1} if i attended college

Y_{i0} if i do not attend college

$$Y_{id} = Y_{i0} - (Y_{i0} - Y_{i1})D_i$$

Concepts

- Individual treatment effects or causal effects

$$\delta_i = Y_{i1} - Y_{i0}$$

- Average treatment effect (ATE) 平均干预效应
  - The average treatment effect is the population average of all i individual treatment effects

$$E[\delta_i] = E[Y_{i1} - Y_{i0}] = E[Y_{i1}] - E[Y_{i0}]$$

  - Policymakers care about average treatment effects

The Fundamental problem of causal inference

- It is impossible to observe both Y_{i1} and Y_{i0} for the same individual and so individual causal effects, δ_i , are unknowable.
- ATE is a quantity that cannot be calculated but need to be estimated.

Other treatment effects of interests

- Average treatment effect on the treated (ATT) 干预组干预效应
  - The average treatment effect on the treatment group is equal to the average treatment effect conditional on being a treatment group member

$$ATT = E[Y_{i1} - Y_{i0}|D_i = 1] = E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 1]$$

  - When do policymakers want to know ATT?
    - Free and reduced-price lunch
    - Need based scholarship
    - Anti-poverty program
    - … …

Other treatment effects of interests

- Average treatment effect on the untreated (ATU)未干预组干预效应
  - The average treatment effect on the untreated group is equal to the average treatment effect conditional on being untreated:

$$ATU = E[Y_{i1} - Y_{i0}|D_i = 0] = E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0]$$

  - Do $ATT = ATU$? Why?

  - Among ATE, ATT, and ATU, ATE and ATT are the two most common ones of interest.

Identifying Y_1 and Y_0

- In the previous example, we do not observe Y_1 and Y_0 for the same students at the same time.
- Thus, we can not estimate the causal impact of attending preparation class and students' performance

Student names	ACT score	Attend preparation-class? (D_i)
1	14 (Y_1 for student 1)	Yes (1)
2	22 (Y_0 for student 2))	No (0)
3	18(Y_1 for student 3))	Yes (1)
4	29(Y_0)	No (0)
5	14(Y_1)	Yes (1)
6	24(Y_0)	No (0)
7	36(Y_0)	No (0)
8	22(Y_0)	No (0)
9	25(Y_0)	No (0)
10	32(Y_1)	Yes (1)

ACT preparation-class and ACT scores with full knowledge

Students	Y_{i1}	Y_{i0}	$\delta_i = Y_{i1} - Y_{i0}$	D_i
1	14	4	10	1
2	18	22	-4	0
3	18	4	14	1
4	25	29	-4	0
5	14	7	7	1
6	36	24	12	0
7	24	36	-12	0
8	18	22	-4	0
9	11	25	-14	0
10	32	29	3	1

- In 2020, Lakeview High School offered an ACT Pre-class to help their students prepare for ACT (American College Test).

- In the same year, 10 students took the ACT. The table on the left shows the ACT scores (Y_i) for those students as well as whether they had attended the ACT pre-class.

- In this example, what are the ATE, ATT, and ATU of the program?
- What are the differences in ACT between the treated and the untreated group without the program?

ACT preparation-class and ACT scores with full knowledge (2)

Students	Y_{i1}	Y_{i0}	$\delta_i = Y_{i1} - Y_{i0}$	D_i
1	14	4	10	1
2	18	22	-4	0
3	18	4	14	1
4	25	29	-4	0
5	14	7	7	1
6	36	24	12	0
7	24	36	-12	0
8	18	22	-4	0
9	11	25	-14	0
10	32	29	3	1

$$ATE = E(\delta_i) = 0.8$$

$$ATT = E(\delta_i | D_i = 1) = 8.5$$

$$ATU = E(\delta_i | D_i = 0) = -\frac{26}{6} \approx -4.33$$

Note that

$$ATE = p * ATT + (1-p) * ATU$$
$$0.8 = 0.4 * 8.5 + 0.6 * -4.33$$

Difference without the program

Treated-Untreated=11- $\frac{158}{6} \approx -15.33$

ACT preparation-class and ACT scores with full knowledge (3)

Students	Y_{i1}	Y_{i0}	$\delta_i = Y_{i1} - Y_{i0}$	D_i
1	14			1
2		22		0
3	18			1
4		29		0
5	14			1
6		24		0
7		36		0
8		22		0
9		25		0
10	32			1

$$ATE = E(\delta_i) = 0.8$$

$$ATT = E(\delta_i | D_i = 1) = 8.5$$

$$ATU = E(\delta_i | D_i = 0) = -\frac{26}{6} \approx -4.33$$

- However, we do not know Y_{i0} for those that participated in the ACT program. We also do not know Y_{i1} for those that did not attend the program.

Example: ACT preparation-class and ACT scores （4）

Students	Y	D_i
1	14	1
2	22	0
3	18	1
4	29	0
5	14	1
6	24	0
7	36	0
8	22	0
9	25	0
10	32	1

$ATE = E(\delta_i)$ =0.8

$$ATT = E(\delta_i|D_i = 1) = 8.5$$

$$ATU = E(\delta_i|D_i = 0) \approx -4.33$$

Difference without the program=11- $\frac{158}{6} \approx -15.33$

- The differences in ACT scores between the treated and untreated groups

$$E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$$

Example: ACT preparation-class and ACT scores （5）

Students	Y	D_i
1	14	1
2	22	0
3	18	1
4	29	0
5	14	1
6	24	0
7	36	0
8	22	0
9	25	0
10	32	1

$ATE = E(\delta_i)$ =0.8

$ATT = E(\delta_i|D_i = 1) = 8.5$

$ATU = E(\delta_i|D_i = 0) \approx -4.33$

Difference without the program=11- $\frac{158}{6} \approx -15.33$

- The differences in ACT scores between the treated and untreated groups

$E(Y_i|D_i = 1) - E(Y_i|D_i = 0) \approx -6.833$

- Substantial biases when relying on the comparison of means

Simple difference in mean outcomes (SDO)

- SDO is the difference between the population average outcome for the treatment and control groups, and can be approximated by the sample averages:

$$SDO = E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 0]$$

- SDO is an estimate of the ATE but SDO may be biased

Decompose SDO

- The simple difference in mean outcomes can be decomposed into three parts:

$$\text{SDO} = E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 0]$$

$$= ATE$$ Parameter of interest

$$+ E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]$$ Selection bias

$$+ (1 - \pi)(ATT - ATU)$$ Heterogeneous treatment effect weighted by the share of the population in the control group (π is the percent in the treatment group)

Proof

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1 - \pi)(ATT - ATU)$$

The expression of the previous equation could be transformed as below

Proof

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1 - \pi)(ATT - ATU)$$

$$ATE = \pi ATT + (1 - \pi)ATU$$

$$ATE = \pi(E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 1]) + (1 - \pi)(E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0])$$

We start by rearranging the expression of ATE and separate different expressions

Proof

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1 - \pi)(ATT - ATU)$$

We start by rearranging the expression of ATE and separate different expressions

$$ATE = \pi ATT + (1 - \pi)ATU$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$

Proof

We need selection, which equals to $E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]$ So we introduce those two terms in our equation

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1 - \pi)(ATT - ATU)$$

$$ATE = \pi ATT + (1 - \pi)ATU$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$
$$- E[Y_{i0}|D_i = 1] + E[Y_{i0}|D_i = 0] + E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]$$

Proof

We need selection, which equals to $E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]$ So we introduce those two terms in our equation

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1-\pi)(ATT - ATU)$$

$$ATE = \pi ATT + (1-\pi)ATU$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$
$$+ E[Y_{i0}|D_i = 1] - Selection \quad - E[Y_{i0}|D_i = 0]$$

Proof

We need $\pi - 1$, so we combine different terms and introduce $E[Y_{i1}|D_i = 1]$ to get it.

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1-\pi)(ATT - ATU)$$

$$ATE = \pi ATT + (1-\pi)ATU$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$
$$+ E[Y_{i0}|D_i = 1] - Selection \quad - E[Y_{i0}|D_i = 0]$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$
$$+ E[Y_{i0}|D_i = 1] - Selection \quad - E[Y_{i0}|D_i = 0] + E[Y_{i1}|D_i = 1] - E[Y_{i1}|D_i = 1]$$

Proof

We need $\pi - 1$, so we combine different terms and introduce $E[Y_{i1}|D_i = 1]$ to get it.

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1 - \pi)(ATT - ATU)$$

$$ATE = \pi ATT + (1 - \pi)ATU$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$
$$-Selection \quad + E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]$$

$$ATE = -(1 - \pi)E[Y_{i1}|D_i = 1] + (1 - \pi)E[Y_{i0}|D_i = 1] + (1 - \pi)E[Y_{i1}|D_i = 0] - \boldsymbol{E[Y_{i0}|D_i = 0]} - (1 - \pi)E[Y_{i0}|D_i = 0]$$
$$-Selection \quad + \boldsymbol{E[Y_{i1}|D_i = 1]}$$

Proof

We need $\pi - 1$, so we combine different terms and introduce $E[Y_{i1}|D_i = 1]$ to get it.

$$ATE = \text{SDO} - (E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]) - (1-\pi)(ATT - ATU)$$

$$ATE = \pi ATT + (1-\pi)ATU$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$

$$ATE = \pi E[Y_{i1}|D_i = 1] - \pi E[Y_{i0}|D_i = 1] + E[Y_{i1}|D_i = 0] - E[Y_{i0}|D_i = 0] - \pi E[Y_{i1}|D_i = 0] + \pi E[Y_{i0}|D_i = 0]$$
$$-Selection \quad + E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0]$$

$$ATE = -(1-\pi)E[Y_{i1}|D_i = 1] + (1-\pi)E[Y_{i0}|D_i = 1] + (1-\pi)E[Y_{i1}|D_i = 0] - \boldsymbol{E[Y_{i0}|D_i = 0]} - (1-\pi)E[Y_{i0}|D_i = 0]$$
$$-Selection \quad + \boldsymbol{E[Y_{i1}|D_i = 1]}$$

$$ATE = -(1-\pi)(E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 1] - E[Y_{i1}|D_i = 0] + E[Y_{i0}|D_i = 0])$$
$$-Selection \quad + SDO$$

Proof

Rearrange the previous equation,

$$\boldsymbol{E[Y_{i1}|D_i=1]-E[Y_{i0}|D_i=0]=SDO}$$

$$E[Y_{i0}|D_i=1]\ -(1-\pi)E[Y_{i1}|D_i=1]=-\mathrm{ATT}$$

$$E[Y_{i1}|D_i=0]-(1-\pi)E[Y_{i1}|D_i=0]=\text{ATU}$$

$$ATE=\text{SDO}-(E[Y_{i0}|D_i=1]-E[Y_{i0}|D_i=0])-(1-\pi)(ATT-ATU)$$

$$ATE=\pi ATT+(1-\pi)ATU$$

$$ATE=\pi E[Y_{i1}|D_i=1]-\pi E[Y_{i0}|D_i=1]+E[Y_{i1}|D_i=0]-E[Y_{i0}|D_i=0]-\pi E[Y_{i1}|D_i=0]+\pi E[Y_{i0}|D_i=0]$$

$$ATE=\pi E[Y_{i1}|D_i=1]-\pi E[Y_{i0}|D_i=1]+E[Y_{i1}|D_i=0]-E[Y_{i0}|D_i=0]-\pi E[Y_{i1}|D_i=0]+\pi E[Y_{i0}|D_i=0]$$

$$-Selection\quad +E[Y_{i0}|D_i=1]-E[Y_{i0}|D_i=0]$$

$$ATE=-(1-\pi)E[Y_{i1}|D_i=1]+(1-\pi)E[Y_{i0}|D_i=1]+(1-\pi)E[Y_{i1}|D_i=0]-\boldsymbol{E[Y_{i0}|D_i=0]}-(1-\pi)E[Y_{i0}|D_i=0]$$

$$-Selection\quad +\boldsymbol{E[Y_{i1}|D_i=1]}$$

$$ATE=-(1-\pi)(ATT-ATU)$$

$$-Selection\quad +SDO$$

Back to our example

- $ATE = E(\delta_i)$ =0.8
- $ATT = E(\delta_i|D_i = 1) = 8.5$
- $ATU = E(\delta_i|D_i = 0) =- \frac{26}{6} \approx- 4.33$
- Difference without the program=11- $\frac{158}{6} =- \frac{92}{6} \approx- 15.33$

$$\text{SDO} = E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 0]$$

$$= ATE + E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0] + (1 - \pi)(ATT - ATU)$$

$$= 0.8 + (-15.33) + (1 - 0.4)\big(8.5 - (-4.33)\big) =- 6.83$$

When SDO is not biased

$$SDO = ATE + E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0] + (1-\pi)(ATT - ATU)$$

- SDO is not necessarily biased in all situations. We can use SDO when

$$(Y_{i1}, Y_{i0}) \perp D_i$$

In this simple case, we do not consider other factors.

Independence assumption: The treatment is independent of the potential outcomes.

Be achieved through randomization of the treatment assignment.

Coin flips, lottery, sorted by birthday… …

- Random assignment means

$$\mathrm{E}[\mathrm{Y}_{\mathrm{i1}}|\mathrm{D}_{\mathrm{i}} = 1] - \mathrm{E}[\mathrm{Y}_{\mathrm{i1}}|\mathrm{D}_{\mathrm{i}} = 0] = 0$$

$$\mathrm{E}[\mathrm{Y}_{\mathrm{i0}}|\mathrm{D}_{\mathrm{i}} = 1] - \mathrm{E}[\mathrm{Y}_{\mathrm{i0}}|\mathrm{D}_{\mathrm{i}} = 0] = 0$$

$$SDO = ATE + E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0] + (1 - \pi)(ATT - ATU)$$

Eliminating the selection and heterogenous treatment effect bias

$$E[Y_{i1}|D_i = 1] - E[Y_{i1}|D_i = 0] = 0$$
$$E[Y_{i0}|D_i = 1] - E[Y_{i0}|D_i = 0] = 0$$

- Selection Bias

$$E[Y_{i1}|D_i = 1] - E[Y_{i1}|D_i = 0] = 0$$

- Heterogenous treatment effect bias

$$(1 - \pi)(ATT - ATU)$$

We just need $ATT - ATU = 0$

$$ATT - ATU = E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 1] - E[Y_{i1}|D_i = 0] + E[Y_{i0}|D_i = 0] = 0$$

- However, most times random assignment are not possible!

The SUTVA in the estimates of ATE, ATT, & ATU

- Potential outcomes model places a limit on what we can measure: ”the stable unit-treatment value assumption” (SUTVA) (Rubin, 1978, 1980, 1990)
- SUTVA means that average treatment effects are parameters that assume
  - (1) homogenous dosage and
  - (2) potential outcomes are invariant to who else (and how many) is treated.

SUTVA: Homogenous dose

- The treatment is received in homogeneous does to all units.
  - Teacher partiality?
  - Assigned to different classes?
  - Exposure to different peers(classmates)?

- Easy to imagine violations of the homogenous does assumptions. But, not a problem, you will be fine if you carefully define the treatment.

SUTVA: Does not affect other units

- The treatment should not affect the untreated units in any ways.
- However, it is common to have spillover effects. In our cases, potential spillover happens when participants
  - share lecture materials, note, slides, and etc.
  - offer tutoring to their classmates
  - rise expectation of their classmates
  - …
- Only change partial equilibrium (局部均衡) only
  - The treatment only change individuals' outcome, and does not affect general equilibrium
  - E.g., Expanding college lead to increases the supply of college educated workers in the market, and individuals become more difficult in finding jobs

Another slide on spillover (1)

1. Contagion: The effect of being vaccinated on one’s probability of contracting a disease depends on whether others have been vaccinated.
2. Displacement: Police interventions designed to suppress crime in one location may displace criminal activity to nearby locations.
3. Communication: Interventions that convey information about commercial products, entertainment, or political causes may spread from individuals who receive the treatment to others who are nominally untreated.

Another slide on spillover (2)

4. Social comparison: An intervention that offers housing assistance to a treatment group may change the way in which those in the control group evaluate their own housing conditions.
5. Signaling: Policy interventions are sometimes designed to "send a message" to other units about what the government intends to do or what it has the capacity to do.
6. Persistence and memory: Within-subjects experiments, in which outcomes for a given unit are tracked over time, may involve "carryover" or "anticipation"

Questions?