Lecture 16
Assessing unconfoundedness

Outline

- Assessing unconfoundedness
  - Negative control outcome
  - Negative control treatment
- Issues with over-adjustment
  - Adjust for post-treatment covariates
  - M-bias
- Suggested reading: Imbens and Rubin book Chapter 21.1-21.4, Peng's book Chapter 16

Unconfoundedness and balance

- Unconfoundedness property: $W_i \perp \big(Y_i(0), Y_i(1)\big) \mid \boldsymbol{X}_i$
- This is an untestable assumption: we can never test for the unconfoundedness property as it is an assumption on the partially unmeasured potential outcomes

- We assess balancing of covariates and test for $W_i \perp \boldsymbol{X}_i \mid e(\boldsymbol{X}_i)$
- What we really care about is the balance of potential outcomes:

$$W_i \perp \big(Y_i(0), Y_i(1)\big) \mid e(\boldsymbol{X}_i)$$

  within strata of observed covariates, potential outcomes corresponding to both treatment conditions need to be balanced between groups

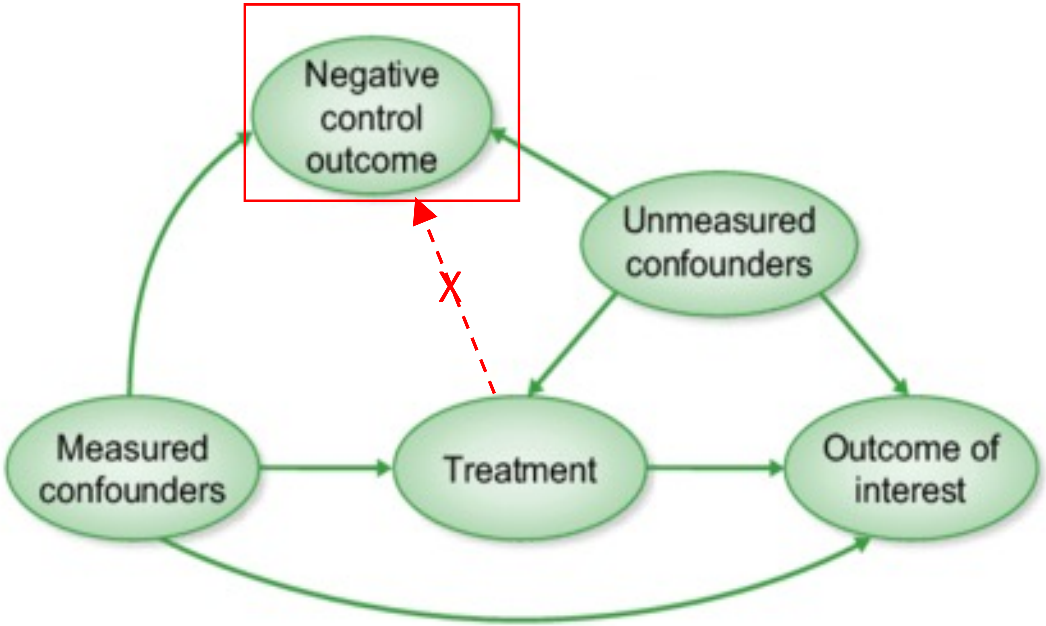
- Covariate balancing is a necessary, but not sufficient condition, especially when there are unmeasured confounding pre-treatment covariates

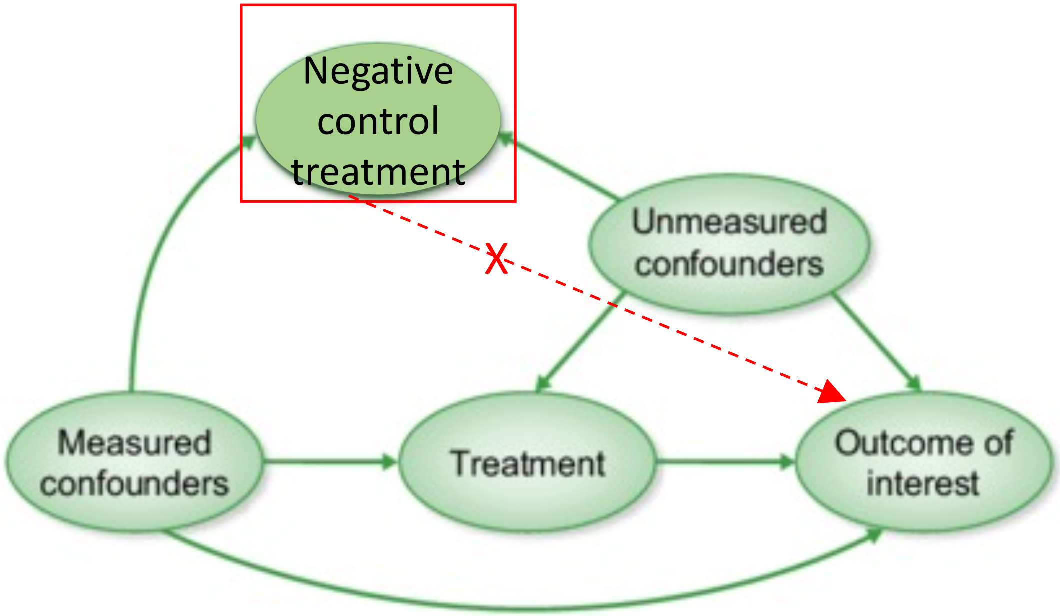
Assessing unconfoundedness

- We can not test for unconfoundedness but we can assess the credibility of the unconfoundedness assumption indirectly

- Three approaches
  - **Negative control outcome**: choose proxy of the real outcome that
    1. Share a similar set of possible unmeasured confounding variables with the real outcome
    2. We know a priori that the treatment have zero causal effect on the proxy
  - **Negative control treatment**: choose new “treatment” that
    1. Share a similar set of possible unmeasured confounding variables with the real treatment
    2. We know a priori that the new “treatment” has zero causal effect on the outcome
  - **Assess robustness of the ATE estimation** given different sets of pre-treatment covariates

Negative control treatments and negative control outcomes





Negative control outcome (pseudo-outcome)

- One common way to find a good proxy of the outcome is the lagged outcome
  - E.x., outcome is the earning 1 year after treatment, lagged outcome is the earning 1 year before treatment

- The idea: the lagged outcome Y_i^{lag}, can be considered a proxy for $Y_i(0)$ and, given it is observed before the treatment, it is unaffected by the treatment

- By definition, the lagged outcome is also a pre-treatment covariate
  - Define $\boldsymbol{X}_i^r = \boldsymbol{X}_i \backslash Y_i^{lag}$, we test for the independence

$$H_0\text{: } W_i \perp Y_i^{lag} \mid \boldsymbol{X}_i^r$$

- In general, negative control outcome satisfies that $Y_i^{lag}(0) \equiv Y_i^{lag}(1)$, so we always observe its potential outcomes

- If we do not reject H_0, it suggests that the unconfoundedness assumption is plausible.

The Imbens-Rubin-Sacerdote lottery data

Table 21.1. ***Summary Statistics for Selected Lottery Sample for the IRS Lottery Data***

Variable	Label	All ($N = 496$) Mean	All ($N = 496$) (S.D.)	Non-Winners ($N_t = 259$) Mean	Winners ($N_c = 237$) Mean	[t-Stat]	Nor Dif
Year Won	(X_1)	6.23	(1.18)	6.38	6.06	−3.0	−0.27
Tickets Bought	(X_2)	3.33	(2.86)	2.19	4.57	9.9	0.90
Age	(X_3)	50.22	(13.68)	53.21	46.95	−5.2	−0.47
Male	(X_4)	0.63	(0.48)	0.67	0.58	−2.1	−0.19
Years of Schooling	(X_5)	13.73	(2.20)	14.43	12.97	−7.8	−0.70
Working Then	(X_6)	0.78	(0.41)	0.77	0.80	0.9	0.08
Earnings Year -6	(Y_{-6})	13.84	(13.36)	15.56	11.97	−3.0	−0.27
Earnings Year -5	(Y_{-5})	14.12	(13.76)	15.96	12.12	−3.2	−0.28
Earnings Year -4	(Y_{-4})	14.21	(14.06)	16.20	12.04	−3.4	−0.30
Earnings Year -3	(Y_{-3})	14.80	(14.77)	16.62	12.82	−2.9	−0.26
Earnings Year -2	(Y_{-2})	15.62	(15.27)	17.58	13.48	−3.0	−0.27
Earnings Year -1	(Y_{-1})	16.31	(15.70)	18.00	14.47	−2.5	−0.23
Pos Earnings Year -6	$(Y_{-6} > 0)$	0.69	(0.46)	0.69	0.70	0.3	0.03
Pos Earnings Year -5	$(Y_{-5} > 0)$	0.71	(0.45)	0.68	0.74	1.6	0.14
Pos Earnings Year -4	$(Y_{-4} > 0)$	0.71	(0.45)	0.69	0.73	1.1	0.10
Pos Earnings Year -3	$(Y_{-3} > 0)$	0.70	(0.46)	0.68	0.73	1.4	0.13
Pos Earnings Year -2	$(Y_{-2} > 0)$	0.71	(0.46)	0.68	0.74	1.6	0.15
Pos Earnings Year -1	$(Y_{-1} > 0)$	0.71	(0.45)	0.69	0.74	1.2	0.10

The Imbens-Rubin-Sacerdote lottery data

Pseudo-Outcome	Remaining Covariates	Selected Covariates	Est	$(\widehat{\text{s. e.}})$
Y_{-1}	$X_1, \ldots, X_6, Y_{-6}, \ldots, Y_{-2}, Y_{-6} > 0, \ldots, Y_{-2} > 0$	X_2, X_5, X_6, Y_{-2}	-0.53	(0.58)
$\frac{Y_{-1}+Y_{-2}}{2}$	$X_1, \ldots, X_6, Y_{-6}, \ldots, Y_{-3}, Y_{-6} > 0, \ldots, Y_{-3} > 0$	X_2, X_5, X_6, Y_{-3}	-1.16	(0.71)
$\frac{Y_{-1}+Y_{-2}+Y_{-3}}{3}$	$X_1, \ldots, X_6, Y_{-6}, Y_{-5}, Y_{-4}, Y_{-6} > 0, Y_{-5} > 0, Y_{-4} > 0$	X_2, X_5, X_6, Y_{-4}	-0.39	(0.77)
$\frac{Y_{-1}+\ldots+Y_{-4}}{4}$	$X_1, \ldots, X_6, Y_{-6}, Y_{-5}, Y_{-6} > 0, Y_{-5} > 0$	X_2, X_5, X_6, Y_{-5}	-0.56	(0.89)
$\frac{Y_{-1}+\ldots+Y_{-5}}{5}$	$X_1, \ldots, X_6, Y_{-6}, Y_{-6} > 0$	X_2, X_5, X_6, Y_{-6}	-0.49	(0.87)
$\frac{Y_{-1}+\ldots+Y_{-6}}{6}$	$X_1, \ldots, X_6$	X_2, X_5, X_6	-2.56	(1.55)
Actual outcome Y	$X_1, \ldots, X_6, Y_{-6}, \ldots, Y_{-1}, Y_{-6} > 0, \ldots, Y_{-1} > 0$	X_2, X_5, X_6, Y_{-1}	-5.74	(1.14)

Worse balance as no previous earnings are controlled

Negative control treatment (pseudo-treatment)

- One common case of negative control treatment is when there are multiple control groups
- Suppose we have two control groups and one treatment group $G_i \in \{c_1, c_2, t\}$ [e.g., ineligibles, eligible nonparticipants and participants]

$$W_i = \begin{cases} 0 & \text{if } G_i = c_1, c_2, \\ 1 & \text{if } G_i = t. \end{cases}$$

- We test for

$$G_i \perp\!\!\!\perp Y_i(0) \,\Big|\, X_i, G_i \in \{c_1, c_2\}$$

which is equivalent to

$$G_i \perp\!\!\!\perp Y_i^{\text{obs}} \,\Big|\, X_i, G_i \in \{c_1, c_2\},$$

Define pseudo-treatment for the lottery data

- One option is to have a comparison control group, of individuals who did not play the lottery at all
- Then we can compare between the “losers” and non-lottery players
- This comparison group is good because “losers” and non-lottery players can be substantially different due to various reasons (so they may share the same unmeasured confounders with that between “losers” and “winners”)

- However, we do not have such data

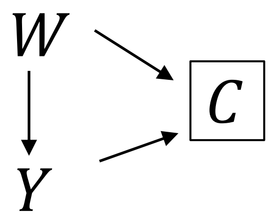
- Here, we split the winners into two subgroups
    - Median yearly prize for the winners is $31,800
    - We treat the winners with yearly prize less than $30,000 as the other group of control
    - Treat the winners with yearly prize larger than $30,000 as the treated group

Pseudo-treatment analysis for the lottery data

Table 21.4. ***Estimates of Average Difference in Outcomes for Controls and Small Winners (less than $30,000) for the IRS Lottery Data***

Outcome	Subpopulation	Est	$(\widehat{\text{s.e.}})$
Y_i	All	−0.82	(1.37)
$\mathbf{1}_{Y_i=0}$	$Y_{i,-1} = 0$	−0.02	(0.05)
$\mathbf{1}_{Y_i=0}$	$Y_{i,-1} > 0$	0.07	(0.05)
Y_i	$Y_{i,-1} = 0$	−1.18	(1.10)
Y_i	$Y_{i,-1} > 0$	−0.16	(0.69)

Adjust for post-treatment covariates

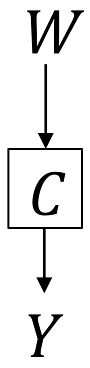


- Collider bias:
  Conditional on a collider C creates non-causal association between A and Y
- Example:

  W: Give folic acid supplements to pregnant women shortly after conception

  Y: fetus's risk of developing a cardiac malformation

  C: survival at birth

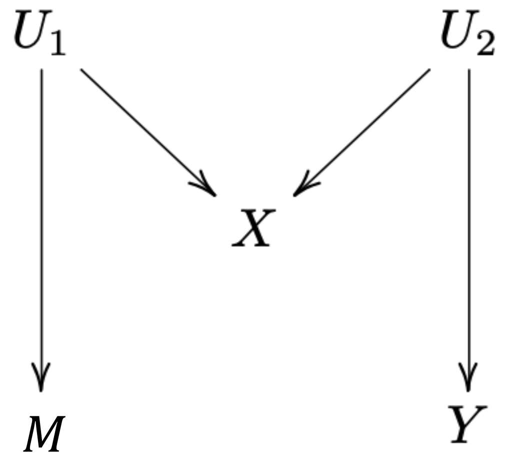


- This is also commonly known as selection bias exists as C can be a selection condition which is unavoidable

- We should avoid adjusting for post-treatment covariates

M bias

A simulation example (code from Chatper 16.3.1 of Peng's book)



- Adjust for X introduces more confounders

```
> ## M bias with large sample size
> n  = 10^6
> U1 = rnorm(n)
> U2 = rnorm(n)
> X  = U1 + U2 + rnorm(n)
> Y  = U2 + rnorm(n)
> ## with a continuous treatment Z
> Z  = U1 + rnorm(n)
> round(summary(lm(Y ~ Z))$coef[2, 1], 3)
[1] 0
> round(summary(lm(Y ~ Z + X))$coef[2, 1], 3)
[1] -0.2
>
> ## with a binary treatment Z
> Z  = (Z >= 0)
> round(summary(lm(Y ~ Z))$coef[2, 1], 3)
[1] 0.002
> round(summary(lm(Y ~ Z + X))$coef[2, 1], 3)
[1] -0.42
```