Lecture 6: Notions of Bias II

POL-GA 1251
Quantitative Political Analysis II
Prof. Cyrus Samii
NYU Politics

February 8, 2018


DANGER
MINES

Sources of bias in estimating causal effects

Nature, research design, & their interaction (biased data)	▶ Sample selection & missing data ▶ Confounding & non-random assignment ▶ Measurement error
Data analysis (biased methods)	▶ Improper averaging of heterogeneous effects ▶ Controlling for post-treatment variables ▶ Controlling for instruments ▶ Misspecification

Biased methods: Heterogenous treatment effects

Heterogenous treatment effects

Recall from last week: stratified (a.k.a. "matching") estimators, $\hat{\rho}_{ATT}$ and $\hat{\rho}$, vs. saturated OLS estimator, $\hat{\delta}_R$:

$$E[\hat{\rho}_{ATT}] = \frac{\sum_x \delta_X \Pr[D_i = 1|X_i = x] \Pr[X_i = x]}{\sum_x \Pr[D_i = 1|X_i = x] \Pr[X_i = x]}$$

and

$$E[\hat{\rho}] = \sum_x \delta_X \Pr[X_i = x],$$

versus

$$\begin{aligned}\hat{\delta}_R &\xrightarrow{a} \frac{\sum_x \delta_X [\Pr[D_i = 1|X_i = x](1 - \Pr[D_i = 1|X_i = x])] \Pr[X_i = x]}{\sum_x [\Pr[D_i = 1|X_i = x](1 - \Pr[D_i = 1|X_i = x])] \Pr[X_i = x]} \\ &= \frac{\sum_x \delta_X \mathrm{Var}(D_i|X_i = x) \Pr[X_i = x]}{\sum_x \mathrm{Var}(D_i|X_i = x) \Pr[X_i = x]},\end{aligned}$$

and then similar for the more generalized regression setting (Aronow & Samii 2016).

Heterogenous treatment effects

- Under CIA, OLS estimator with controls is biased for ATE, ATT.
- Size of bias depends on how $\mathrm{E}[\rho_i|X_i]$ correlates with $\mathrm{Var}(D_i|X_i)$.

Heterogenous treatment effects

Table: Karlan and Zinman (2008) treatment effect weighting

		Weight	
Risk group	Effect	FE	Sample
Low	-32.4	0.044	0.125
Medium	-9.9	0.058	0.092
High	-2.7	0.898	0.783
Average		-4.393	-7.047
Std. error		(1,129)	(1.917)

From Gibbons et al. 2018.

Notes: risk groups are experimental strata. Effect shows stratum-specific treatment effects. FE and Sample show the conditional variance and population weights, respectively. Average shows the resulting conditional-variance and population weighted averages.

Biased methods: Post-treatment bias

Post-treatment bias

Rosenbaum (1984) provides the classic analysis. First, suppose,

- as usual, $\mathscr{D} = \{0, 1\}$, SUTVA, (Y_{1i}, Y_{0i}), X_i.
- target estimand is $\rho = \mathrm{E}[Y_{1i} - Y_{0i}]$.
- possible identifying assumptions are random assignment or CIA.

Post-treatment bias

Rosenbaum (1984) provides the classic analysis. First, suppose,

- as usual, $\mathscr{D} = \{0, 1\}$, SUTVA, (Y_{1i}, Y_{0i}), X_i.
- target estimand is $\rho = \mathrm{E}[Y_{1i} - Y_{0i}]$.
- possible identifying assumptions are random assignment or CIA.
- Add to the mix (S_1, S_0), post-treatment variables, where

$$S_i = D_i S_{1i} + (1 - D_i) S_{0i}$$

- Suppose we control for both S_i and X_i, yielding in expectation:

$$\begin{aligned}\tilde{\rho} &= \mathrm{E}_{S,X}\{\mathrm{E}[Y_i | D_i = 1, S_i = s, X_i = x] - \mathrm{E}[Y_i | D_i = 0, S_i = s, X_i = x]\} \\ &= \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i} | D_i = 1, S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i} | D_i = 0, S_{0i} = s, X_i = x]\}\end{aligned}$$

- Alarm bells should be ringing. Do you see why?

Post-treatment bias

To characterize the bias, let's first define,

$$\tilde{\nu} = \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i}|S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i}|S_{0i} = s, X_i = x]\},$$

the "net treatment difference." It is not a proper causal quantity.

Post-treatment bias

An expression for post-treatment bias:

$$\tilde{\rho} - \rho = \underbrace{(\tilde{\rho} - \tilde{\nu})}_{A} + \underbrace{(\tilde{\nu} - \rho)}_{B}$$

Post-treatment bias

An expression for post-treatment bias:

$$\tilde{\rho} - \rho = \underbrace{(\tilde{\rho} - \tilde{\nu})}_{A} + \underbrace{(\tilde{\nu} - \rho)}_{B}$$

- A assured to be zero only if random assignment or CIA holds for both (Y_{1i}, Y_{0i}) and (S_{1i}, S_{0i}):

$$(Y_{1i}, S_{1i}, Y_{0i}, S_{0i}) \perp\!\!\!\perp D_i | X_i \Rightarrow Y_{di} \perp\!\!\!\perp D_i | X_i, S_{di}$$

$$\begin{aligned}\tilde{\rho} - \tilde{\nu} = {} & \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i}|D_i = 1, S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i}|D_i = 0, S_{0i} = s, X_i = x]\} \\ & - \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i}|S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i}|S_{0i} = s, X_i = x]\}.\end{aligned}$$

Post-treatment bias

An expression for post-treatment bias:

$$\tilde{\rho} - \rho = \underbrace{(\tilde{\rho} - \tilde{\nu})}_{A} + \underbrace{(\tilde{\nu} - \rho)}_{B}$$

- A assured to be zero only if random assignment or CIA holds for both (Y_{1i}, Y_{0i}) and (S_{1i}, S_{0i}):

$$(Y_{1i}, S_{1i}, Y_{0i}, S_{0i}) \perp\!\!\!\perp D_i | X_i \Rightarrow Y_{di} \perp\!\!\!\perp D_i | X_i, S_{di}$$

$$\begin{aligned}\tilde{\rho} - \tilde{\nu} = {} & \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i}|D_i = 1, S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i}|D_i = 0, S_{0i} = s, X_i = x]\} \\ & - \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i}|S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i}|S_{0i} = s, X_i = x]\}.\end{aligned}$$

- B assured to be zero only if $S_{1i} = S_{0i}$ (no effect) for all i:

$$\tilde{\nu} - \rho = \mathrm{E}_{S,X}\{\mathrm{E}[Y_{1i}|S_{1i} = s, X_i = x] - \mathrm{E}[Y_{0i}|S_{0i} = s, X_i = x]\} - \mathrm{E}[Y_{1i} - Y_{0i}]$$

.

Post-treatment bias

Implications:

- Post-treatment bias assured to be zero only if (i) CIA or random assignment holds with respect to X_i for both Y and S, *and* (ii) treatment has no effect on S.
- If CIA or random assignment holds with respect to X_i, controlling for S may induce bias that could have been avoided by leaving S out.

Post-treatment bias

Intuitions about post-treatment bias:

- Suppose $\mathrm{E}[Y_{di}|S_{di}=s, X_i=x] = \alpha_d + x\beta + s\gamma$.
- Then the true treatment effect is given by,

$$\begin{aligned}\rho &= \mathrm{E}[(\alpha_1 + X_i\beta + S_{1i}\gamma) - (\alpha_0 + X_i\beta + S_{0i}\gamma)] \\ &= (\alpha_1 - \alpha_0) + \mathrm{E}[S_{1i} - S_{0i}]\gamma.\end{aligned}$$

- The estimate that conditions on S is given by,

$$\begin{aligned}\tilde{\rho} &= \mathrm{E}[(\alpha_1 + X_i\beta + S_i\gamma) - (\alpha_0 + X_i\beta + S_i\gamma)] \\ &= (\alpha_1 - \alpha_0).\end{aligned}$$

- The bias, $\tilde{\rho} - \rho = -\mathrm{E}[S_{1i} - S_{0i}]\gamma$ amounts to the portion of ρ that has been "stolen away" by conditioning on S.

Post-treatment bias

Sometimes we feel that we need to control for S because it proxies for some pre-treatment variable, U, which was not measured.

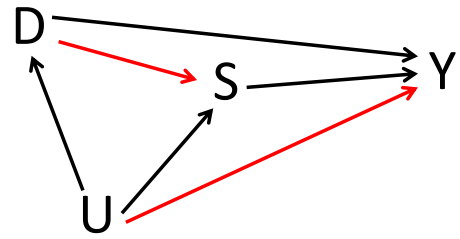
- Suppose that CIA holds when conditioning on X and U.
- We call S a "surrogate" for U when,

$$Y_{di} \perp\!\!\!\perp U_i | (S_{di}, X_i) \text{ for } d = 0, 1$$

  Then, controlling for S_{di} breaks dependency between Y_{di} and U_i.
- If so, bias component $A = 0$ (D_i's drop out of conditioning sets).
- Then, if $S_{1i} = S_{0i}$ for all i as well, post-treatment bias is zero.
- These three conditions imply that controlling for S is justified for identifying ρ.

Post-treatment bias

- Suppose *conditional on X* the following holds, where the red lines are causal relations that we are unsure about.



- If D affects S, conditioning on S *induces* new dependencies between U and D, which confound effect of D on Y. So that's why we need $S_{1i} = S_{0i}$.
- If U affects Y in a manner that is not channeled through S, then conditioning on S does not block the confounding that U introduces to the effect of D on Y. So that's we we need the conditional independence (exclusion restriction) assumption from the previous slide.

Post-treatment bias

Generally speaking,

- Controlling for post-treatment variables is a *bad idea* for experiments or well-identified natural experiments.
- In observational studies where lots of controls are needed for CIA, it may be justified, but *only if the necessary surrogacy and zero-effect assumptions* are plausible.
- At the very least, if you feel that you have to control for a post-treatment variable, conduct a sensitivity analysis.
- We will look into questions of mediation and "direct" vs. "indirect" effects later in the semester.

Biased methods: Bias amplification

Bias amplification

- Recall the homework problem where we had an instrument, $Z_i \perp\!\!\!\perp (Y_{1i}, Y_{0i})$, an endogenous binary treatment, $D_i(Z_i)$, and observed outcomes, $Y_i = D_i(Z_i)Y_{1i} + [1 - D_i(Z_i)]Y_{0i}$. The population was divided into three principal strata:

Principal stratum	$Z_i = 0$ $D_i(0)$	$Z_i = 1$ $D_i(1)$
Always takers	1	1
Compliers	0	1
Never takers	0	0

(We assume no "defiers.")

Bias amplification

- Conditioning on $Z_i = 1$ yields an estimate,

$$\mathrm{E}[Y_i|D_i(1) = 1, Z_i = 1] - \mathrm{E}[Y_i|D_i(1) = 0, Z_i = 1],$$

which is a comparison of average outcomes among "always takers" and "compliers" against average outcomes among "never takers," and therefore not a valid causal effect.

Recall, "CEF is causal when it describes differences in average potential outcomes for a fixed reference population." (Angrist & Pischke, p.52)

As such this difference does not define a causal effect.

Bias amplification

- Similar for regression.
- Suppose

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i, \qquad (1)$$

but ε_i correlated with X_{i2}, so we have endogeneity.

- Suppose and instrument Z_i that is correlated with X_i but not ε_i, and that we try to fit the following model with OLS,

$$Y_i = \beta_{1,Z} + \beta_{2,Z} X_i + \beta_3 Z_i + \varepsilon_{Z,i},$$

yielding,

$$\hat{\beta}_{2,Z} = \frac{\text{Cov}\left[\tilde{Y}, \tilde{X}_i\right]}{\text{Var}\left[\tilde{X}_i\right]}.$$

- $\tilde{X}_i$ is purged of variation due to Z_i, which was uncorrelated with ε_i, in which case the endogenous variation correlated with ε_i has been "concentrated"!
- Bias from controlling for Z_i may be worse that the bias from estimation of (1) via OLS.

Bias amplification

- Another look (Clarke 2005, 2009): suppose true model:

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \varepsilon_i,$$

with ε_i independent. Want β_2, but have two misspecified models:

$$(A) \qquad Y_i = \beta_1^a + \beta_2^a X_{i2} + \varepsilon_i^a,$$
$$(B) \qquad Y_i = \beta_1^b + \beta_2^b X_{i2} + \beta_3^b X_{i3} + \varepsilon_i^b$$

By FWL and OVB formula,

$$\mathrm{E}[\hat{\beta}_2^a] = \beta_2 + \beta_3 \delta_{3,2} + \beta_4 \delta_{4,2}$$
$$\mathrm{E}[\hat{\beta}_2^b] = \beta_2 + \beta_4 \tilde{\delta}_{4,2(3)}$$
$$\left|\mathrm{E}[\hat{\beta}_2^a] - \mathrm{E}[\hat{\beta}_2^b]\right| = \left|\beta_3 \delta_{3,2} + \beta_4 (\delta_{4,2} - \tilde{\delta}_{4,2(3)})\right|$$

- Possible that (B) is more biased than (A). E.g., sps all positive biases, but β_3 near zero, β_4 large, and $\delta_{4,2} << \tilde{\delta}_{4,2(3)}$. Note last condition implies substantial correlation between X_{i2} and X_{i3}.

Bias amplification more generally

Pearl (2010):

[B]eing a good predictor of treatment assignment [may compromise] the bias-reducing potential of a covariate, for it tends to amplify bias due to other, uncontrolled confounders. One would do better therefore to rank order covariates based on their importance with respect to the outcome variable.

This may be overstating it, but the general recommendation to pay attention to relations to outcomes as well as assignment is warranted when trying to meet the CIA .

Biased methods: Model dependence, misspecification & “extreme counterfactuals”

Model dependence, misspecification & "extreme counterfactuals"

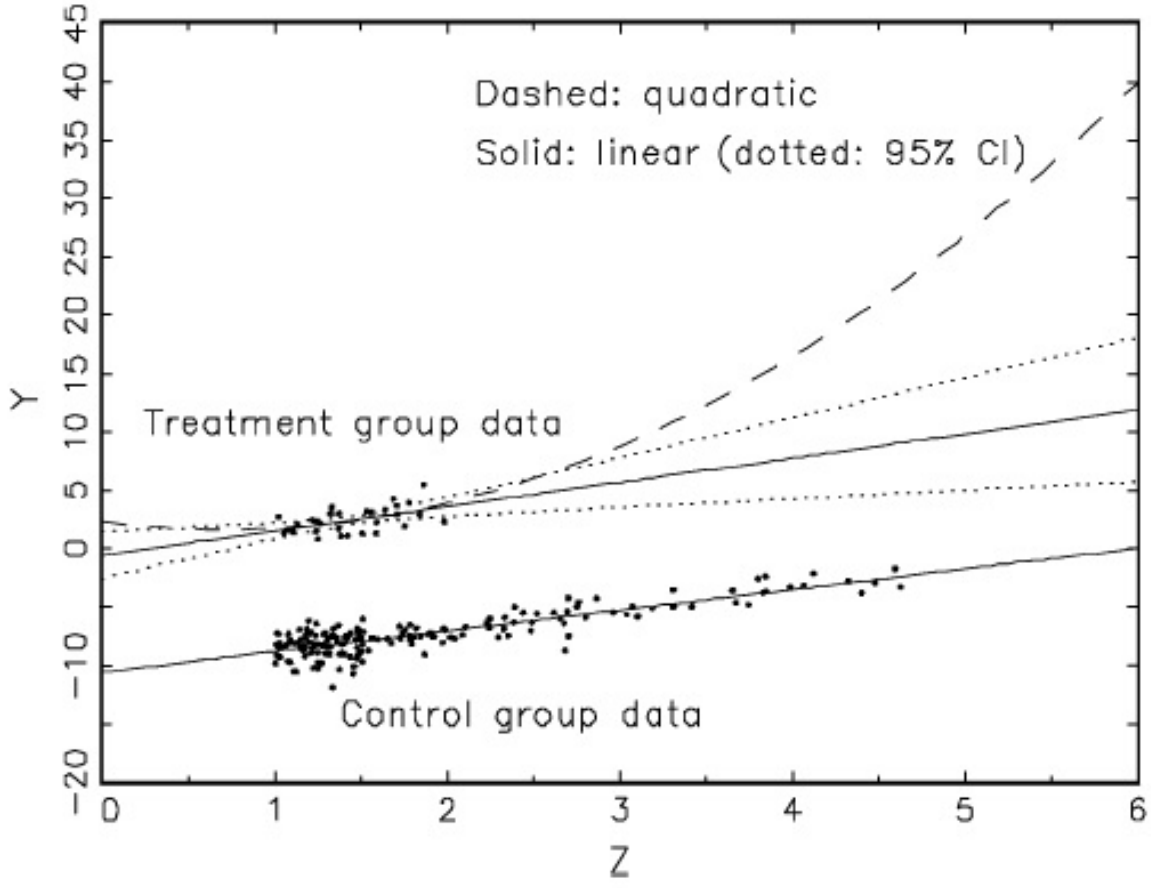


Fig. 4 An illustration of how the degree of extrapolation bias is more severe (and model dependent) than interpolation bias.

(King and Zeng, 2006)

Model dependence, misspecification & "extreme counterfactuals"

This is the problem of extreme counterfactuals—predictions, what-if questions, and causal inferences that are so far from the data that inferences wind up being drawn on the basis of minor model specification choices no one would like to defend, rather than empirical evidence....Our confidence interval for counterfactuals farther from the data are wider, but the inference may be considerably more uncertain than the confidence interval indicates.
(King and Zeng, 2006, p. 132)

Model dependence, misspecification & "extreme counterfactuals"

- Recall under CIA, we get ATE as,

$$\rho = \mathrm{E}_X \{\mathrm{E}[Y_{1i}|D_i = 1, X_i = x] - \mathrm{E}[Y_{0i}|D_i = 0, X_i = x]\}$$

- Potential for bias when we cannot implement this directly – that is, when we cannot really line up treated and control observations for all values of x.
- *To make causal inferences in situations with nonoverlapping densities, we must therefore either eliminate the region outside of common support or attempt to extrapolate to the needed data, such as by using a parametric model.* (p. 149)
- Extrapolation to areas with no overlap can be perilous.
- Motivates strategy of "changing the goal posts" by focusing on effects in areas of common support and non-parametric methods like matching and weighting.

Model dependence, misspecification & "extreme counterfactuals"

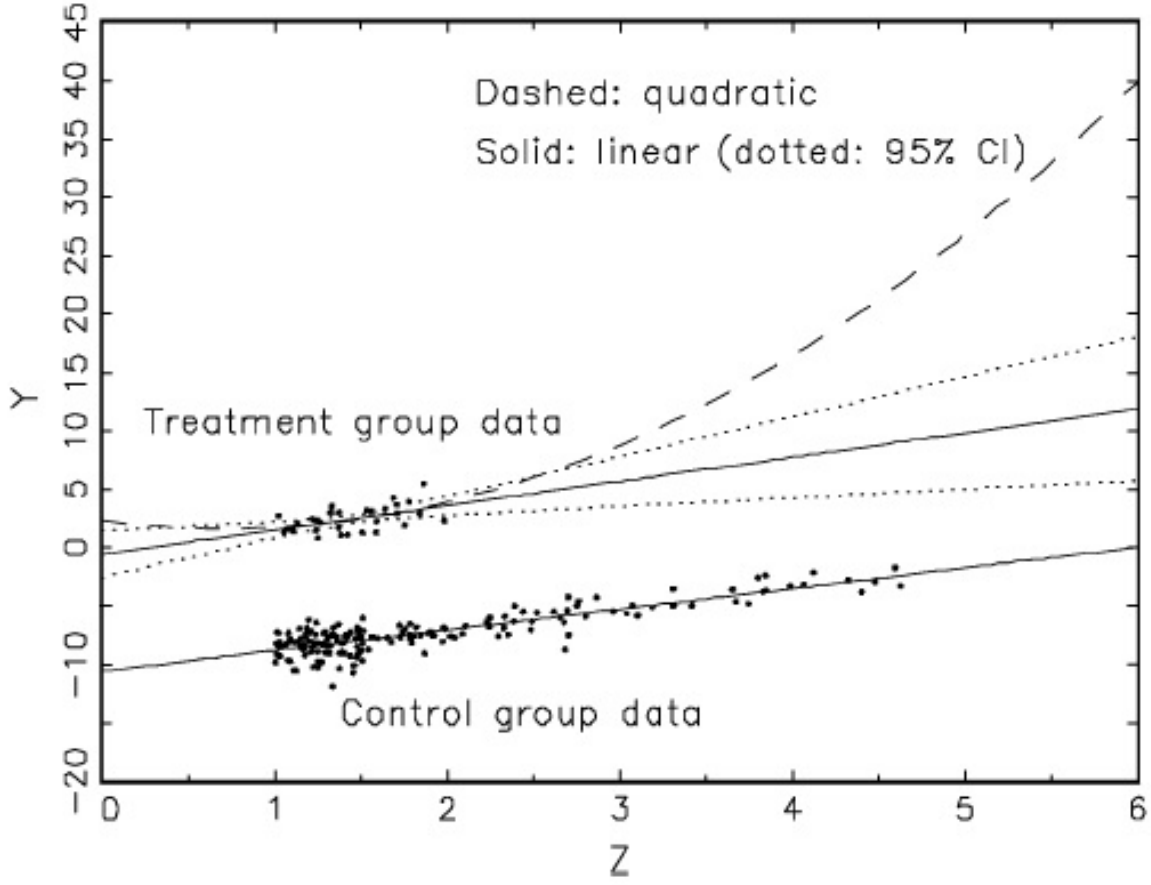


Fig. 4 An illustration of how the degree of extrapolation bias is more severe (and model dependent) than interpolation bias.

Remarks on Biased Methods

- These biases make CIA studies especially vulnerable to hidden confounding.
- E.g., suppose in a regression study someone adds a control variable to the specification and the effect estimate changes.

Remarks on Biased Methods

- These biases make CIA studies especially vulnerable to hidden confounding.
- E.g., suppose in a regression study someone adds a control variable to the specification and the effect estimate changes.
- Is it because you are actually reducing confounding?

Remarks on Biased Methods

- These biases make CIA studies especially vulnerable to hidden confounding.
- E.g., suppose in a regression study someone adds a control variable to the specification and the effect estimate changes.
- Is it because you are actually reducing confounding?
- Or is it because you are shifting the effective sample?

Remarks on Biased Methods

- These biases make CIA studies especially vulnerable to hidden confounding.
- E.g., suppose in a regression study someone adds a control variable to the specification and the effect estimate changes.
- Is it because you are actually reducing confounding?
- Or is it because you are shifting the effective sample?
- Or is it because you are introducing bias amplification?

Remarks on Biased Methods

- These biases make CIA studies especially vulnerable to hidden confounding.
- E.g., suppose in a regression study someone adds a control variable to the specification and the effect estimate changes.
- Is it because you are actually reducing confounding?
- Or is it because you are shifting the effective sample?
- Or is it because you are introducing bias amplification?
- Or is it because you are compounding misspecification?

Remarks on Biased Methods

- These biases make CIA studies especially vulnerable to hidden confounding.
- E.g., suppose in a regression study someone adds a control variable to the specification and the effect estimate changes.
- Is it because you are actually reducing confounding?
- Or is it because you are shifting the effective sample?
- Or is it because you are introducing bias amplification?
- Or is it because you are compounding misspecification?
- Or is it another form of bad control, like post-treatment control?

Remarks on Biased Methods

- If one can minimize reliance on CIA, control strategies, and modeling, then one can reduce the potential for these "errors of commission."
- When control strategies are *unavoidable*, the general recommendation is sensitivity analysis, to check for sensitivity to violations of CIA, assumptions about post-treatment bias, bias amplification, etc.