HETEROGENEITY OF TREATMENT EFFECTS & INSTRUMENTAL VARIABLES

MATT MACIEJEWSKI, PHD

DURHAM COIN

9-21-2022 HERC PRESENTATION

DISCLOSURES

- VA HSR&D Funding
  - RCS 10-391, IIR 20-114, IIR 18-019, C19 21-278

- Own Amgen stock due to spouse's employment

AGENDA

- General research question we want to answer and why it is biased in non-randomized studies. Implication of bias for interpretation
- How instrumental variable (IV) analysis can help address this bias and required assumptions
- Interpretation of IV results: the marginal patient
- Werner analysis (2019 JAMA Intern Med)

HEALTH SYSTEMS NEED ANSWERS TO PRESSING QUESTIONS

- Due to feasibility, ethics or other reasons, RCTs won't answer all questions

- In non-randomized study, want to know if treatment works as intended…
  - In those who choose treatment (average treatment effect on the treated, ATT)
  - In eligible patients who are randomly chosen for treatment (ATE)

- What we estimate to answer this question:

$$Y_i = \beta_0 + \beta_1 \cdot Tx_i + \beta_2 \cdot X_i + U_i$$

Y = Outcome, Tx = treatment, X = covariates, U = residual (unobs confounders + random error)

BIAS IN USUAL OUTCOME MODEL

- $Y_i = \beta_0 + \beta_1 \cdot Tx_i + \beta_2 \cdot X_i + U_i$

- Problem: In non-randomized studies, U = random error AND unobserved confounders (e.g., ability, compulsiveness, perceived risk of getting COVID)
  - If measurement of covariates (X's) was exhaustive, then U may be close to random and Cov(Tx,U) → 0
  - If treatment is randomly allocated, then (by assumption) unobserved confounders are balance between arms

- Since treatment is not allocated randomly and measurement of covariates (X) is often incomplete, unobserved confounding cannot be ruled out
  - Cov(Tx,U) ≠ 0 because patients may have private information about returns to treatment to decide whether or not to choose treatment

INTERPRETATION CAN BE FRAUGHT IF THERE IS BIAS

- $Y_i = \beta_0 + \beta_1 \cdot Tx_i + \beta_2 \cdot X_i + U_i$

- Treatment effect estimate will be biased by unobserved confounders (e.g., ability broadly defined, prior treatment attempts, disease severity)

- If treatment effect is not null, is it due to…
  - Treatment effect?
  - Selection effects?
  - Both?

TREATMENT VS. SELECTION EFFECTS IN RCT

- Randomization simplifies interpretation because selection is controlled
  - Interpretation follows from 3 possible outcome differences

	Treatment Effect is…		
	Harmful	**Null**	**Protective**
No selection	Tx outcome worse than Control	Treatment = Control	Tx outcome better than Control

TREATMENT VS SELECTION EFFECTS IN NON-RANDOMIZED STUDIES

- Lack of randomization creates possibility of selection bias
  - Now there are 9 possible outcomes with multiple interpretations

	Treatment Effect is…		
	Harmful	**Null**	**Protective**
Healthier pts get Tx	Treatment > < Control (depends)	Treatment > Control	Treatment > Control
No selection	Treatment < Control	Treatment = Control	Treatment > Control
Sicker pts get Tx	Treatment < Control	Treatment < Control	Treatment > < Control (depends)

IF OUTCOMES BETTER IN TREATMENT GROUP THAN CONTROL GROUP, WHICH INTERPRETATION IS CORRECT?

- 5 possible outcomes with multiple interpretations!!

	Treatment Effect is…		
	Harmful	**Null**	**Protective**
Healthier pts get Tx	Treatment > < Control (depends)	Treatment > Control	Treatment > Control
No selection			Treatment > Control
Sicker pts get Tx			Treatment > < Control (depends)

HOW TO PROCEED?

- Shouldn't just estimate treatment equation
  - Identify proxy measures for key unobserved confounders
- Match or weight based on propensity scores
  - PS doesn't address unobserved confounding
- Estimate instrumental variable model
  - In an RCT with imperfect compliance, valid instrument is randomization assignment

IV ANALYSIS REQUIRES ESTIMATION OF TREATMENT SELECTION

- $Tx_i = \alpha_0 + \alpha_1 \cdot X_i + \alpha_2 \cdot Z_i + U_i$

- Tx = observed treatment, X = covariates, Z = instrument
  - Propensity scores have a similar specification (but no Z)

- How observed treatment status is realized in population:

  $Tx = p \cdot Tx_1 + (1 - p) \cdot Tx_0$

  where Tx_1 = choose treatment, Tx_0 = choose not treatment (e.g., control)

- With perfect compliance, $Pr(Tx_1) = Pr(Tx_0) = 1$

ASSUMPTIONS OF IV ANALYSIS

- $Tx_i = \alpha_0 + \alpha_1 \cdot X_i + \alpha_2 \cdot Z_i + U_i$

- Assumption #1: Instrument is predictive of treatment assignment (the way a coin flip correlates with treatment assignment in RCT)
  - Cov(Tx, Z) ≠ 0; instrument considered strong if F-test on $\alpha_2 > 10$

- Assumption #2: Instrument is uncorrelated with outcome or residual of outcome equation (not strictly testable)
  - Referred to as exclusion restriction

- Other assumptions
  - Stable unit treatment value assumption (SUTVA): treatment affects only person treated (no herd effects)

TWO WAYS TO DO IV ANALYSIS

- $Tx_i = \alpha_0 + \alpha_1 \cdot X_i + \alpha_2 \cdot Z_i + U_i$

- From treatment equation, you get 2 terms:
  - Predicted treatment status: $Tx_i = \widehat{Tx_i}$
  - Residual: $\in_i = Tx_i - \widehat{Tx_i}$

- Two-stage least squares: replace observed treatment (Tx_i) with predicted treatment ($\widehat{Tx_i}$) and run outcome equation since Cov($\widehat{Tx_i}$,U) = 0

$$Y_i = \beta_0 + \beta_1 \cdot \widehat{Tx_i} + \beta_2 \cdot X_i + U_i$$

- Two-stage residual inclusion: add residual to original outcome equation

$$Y_i = \beta_0 + \beta_1 \cdot Tx_i + \beta_2 \cdot X_i + \beta_3 \cdot \in_i + U_i$$ (Terza & Basu 2008 JHE)

IV ANALYSIS UNDER IMPERFECT COMPLIANCE

- $Tx_i = \alpha_0 + \alpha_1 \cdot X_i + \alpha_2 \cdot Z_i + U_i$

  Remember, $Tx = p \cdot Tx_1 + (1 - p) \cdot Tx_0$

- Unlikely that perfect compliance ($Pr(Tx_1) = Pr(Tx_0) = 1$) holds, which results in…
  - 2 subgroups in treatment arm: compliers ($Tx_1 > Tx_0$) + always-takers ($Tx_1 = Tx_0 = 1$)
  - 2 groups in control arm: compliers ($Tx_0 > Tx_1$) + never-takers ($Tx_1 = Tx_0 = 0$)

- Note: instrument (Z_i) does not inform treatment selection of always-takers or never-takers

IMPLICATION OF IMPERFECT COMPLIANCE FOR INTERPRETATION

- If there is imperfect compliance and instrument (Z_i) does not inform treatment selection of always-takers or never-takers, then treatment effect estimate generalizes only to compliers
  - These are referred to as "marginal patients"

- Instead of estimating the average treatment effect, IV yields what is called the Local Average Treatment Effect (LATE)
  - LATE represents ATE for patients that change treatment status according to their treatment assignment
  - LATE also referred to as complier average causal effect (CACE)

- Fundamental challenge of LATE: not entirely sure who LATE generalizes to
  - Can't easily identify marginal patients in analytic cohort

CHALLENGE OF IV INTERPRETATION OF LATE

- Treatment effect estimate generalizes only to compliers (aka marginal patients)

- Fundamental challenge of LATE: not entirely sure who LATE generalizes to
  - Can't easily identify marginal patients in analytic cohort

- Baiocchi, Cheng & Small 2014 Stats in Medicine outline a method for characterizing compliers
  - Prevalence ratio for binary covariates $= \frac{P(X=1|C=complier)}{P(X=1)}$
  - Can compare estimated prevalence ratio to prevalence of binary covariates in overall cohort (see sections 5.2-5.3 of 2014 paper)

WERNER ANALYSIS OF HOSPITAL DISCHARGE VS HOME WITH HOME HEALTH CARE

- Summarize RQ: Are hospital readmission and costs similar between patients discharged home with home health (HH) and patients discharged to a skilled nursing facility (SNF)?
  - 2010-2016 FFS Medicare data (n=17,235,854)

- Why might non-randomized study of outcome differences be biased?
  - Confounding by indication (healthier patients discharged home)
  - Management of complications and functional improvements may differ by setting

WERNER ANALYSIS OF HOSPITAL DISCHARGE VS HOME WITH HOME HEALTH CARE

- Summarize RQ: Are hospital readmission and costs similar between patients discharged home with home health (HH) and patients discharged to a skilled nursing facility (SNF)?
  - 2010-2016 FFS Medicare data (n=17,235,854)

- Outcome eq: $Readmit_i = \beta_0 + \beta_1 \cdot Home\ w/HH_i + \beta_2 \cdot X_i + U_i$

- Treatment eq: $Home\ w/HH_i = \alpha_0 + \alpha_1 \cdot X_i + \alpha_2 \cdot DiffDist_i + U_i$
  - Instrument = differential distance from home to home health and from home to SNF
  - DiffDist = (miles from home to nearest HH agency) – (miles from home to nearest SNF)

DID INSTRUMENT SATISFY IV CONDITIONS IN WERNER?

- Treatment eq: $Home\ w/HH_i = \alpha_0 + \alpha_1 \cdot X_i + \alpha_2 \cdot DiffDist_i + U_i$
  - Instrument = differential distance from home to home health and from home to SNF
  - DiffDist = (miles from home to nearest HH agency) – (miles from home to nearest SNF)

- Assumption #1: Differential distance predicts patient selection into HH or SNF
  - F-test from logistic regression of treatment = 263.4 (F > 10 is strong per Staiger & Stock 1997)

- Assumption #2: Instrument is uncorrelated with outcome or residual of outcome equation (not strictly testable)
  - Satisfied if patients don't choose where to live based on distance to HH and SNF?
  - Indirect tests
    - Median split: are patient characteristics balanced based on value of instrument?
    - Falsification test: is Z uncorrelated with outcome not relevant to treatment (e.g., admission far from home)?

READMIT DIFFERENCES FROM WERNER 2019

Method	Result
Unadjusted	2.0% lower readmit rate if HH
Covariate adjusted	1.6% lower readmit rate if HH
IV results	5.6% higher readmit rate if HH

WERNER STATEMENT ABOUT INFERENCE

- "When interpreting these results, it is important to understand the population to whom they apply….the results of instrumental variable analyses apply to the so-called marginal patients.[23] *The marginal patients in this study are those discharged to home with home health care solely because of their closer proximity to a home health agency than to an SNF*, conditional on health characteristics. In this context, these marginal patients may be interpreted as those whose need for home health vs SNF is borderline and either setting would be reasonable…."

SUMMARY

- IV methods can help address unobserved confounding

- Requires an instrument that meets IV assumptions
  - Quasi-randomizer

- Tradeoff is that inference changes from ATE to LATE
  - Generalizability changes from entire cohort to compliers
  - Compliers aren't necessarily identifiable

READINGS

- Abadie & Cattaneo, 2018 Annual Review of Economics, 10: 465-503
- Baiocchi, Cheng & Small, 2014 Stats in Med, 33: 2297-2340
- Basu, Coe & Chapman, 2018 Health Econ, 27(6): 937-955
- Goldman, Bhattacharya, et al., 2001 JASA, 96(455): 883-894
  - **Great** appendix walking through IV assumptions intuitively
- Harris & Remler, 1998 HSR, 33(5 Pt 1): 1337-1360
- Maciejewski, Dowd & Norton, 2022 JAMA, 327(12): 1177-1178
- Staiger and Stock, 1997 Econometrica 65(3): 557-586
- Terza, Basu & Rathouz, 2008 J Health Econ, 27(3): 531-543
- Werner et al., 2019 JAMA Internal Med, 179(5): 617-623

THANKS!

QUESTIONS?