500 Class 10 Slides

github.com/THOMASELOVE/2020-500

2020-04-09

Agenda

- Some details on Sensitivity Analysis
- Instrumental Variables: An Introduction

Sensitivity Analysis

Lots of things can be described as part of a sensitivity analysis. We are focusing on one issue: quantifying departures from randomized (i.e. ignorable) treatment assignment.

Ignorable treatment assignment means that if two people have the same values of the observed covariates (and thus, for example, the same propensity score) then they have the same probability of treatment.

- Rosenbaum's bounds on Γ are just one possibility.
- Γ and Θ_p and Λ and Δ are just different methods of describing departure from ignorable treatment assignment in matched pairs, although only Γ applies outside of matched pairs.

Γ and Θ_p

We can express this in terms of Γ or Θ_p pretty easily in the matched pairs setting.

$$\frac{1}{1+\Gamma} \leq \Theta_p \leq \frac{\Gamma}{1+\Gamma}$$

$\Gamma = 2$ is the same magnitude of departure from ignorable treatment assignment as the interval from 0.33 to 0.67 for Θ_p.

If $\Gamma = 2$, then Harry might be twice as likely as Sally to receive the treatment (so Harry's probability Θ_H is $2/3$ and Sally's is $1/3$) or Sally might be twice as likely as Harry (so Harry's probability could be as low as $1/3$) to receive the treatment.

Amplifying the Γ value with Λ and Δ

This approach, like Θ_p bounds, applies only in the case of matched pairs.

- Λ tells you about the relationship of an unobserved covariate with treatment assignment.
- Δ tells you about the relationship of an unobserved covariate with the outcome.

$$\Gamma = (\Lambda\Delta + 1)/(\Lambda + \Delta)$$

Table 9.1

Table 9.1. Understanding the sensitivity parameter Γ

Γ	*Range of possible values of* θ_p		Λ	Δ
1	0.50	0.50	1	1
1.05	0.49	0.51	1.37	1.37
1.1	0.48	0.52	1.40	1.80
1.25	0.44	0.56	2	2
1.5	0.40	0.60	2	4
2	0.33	0.67	3	5
2.5	0.29	0.71	4	6
3	0.25	0.75	5	7
3.5	0.22	0.78	6	8
4	0.20	0.80	7	9
4.5	0.18	0.82	8	10
5	0.17	0.83	9	11
6	0.14	0.86	11	13
7	0.12	0.88	13	15
8	0.11	0.89	15	17
9	0.10	0.90	17	19

Using the Amplification

If $\Gamma = 1.5$ then, for example we could use

- a bound on Θ_p from 0.40 to 0.60
- or a combination of $\Lambda = 2$ and $\Delta = 4$
- or a combination of $\Lambda = 4$ and $\Delta = 2$
- or a requirement that $\Lambda = 1.5$ and that the unobserved covariate be a perfect predictor of the outcome.
- or a requirement that $\Delta = 1.5$ and that the unobserved covariate be a perfect predictor of treatment assignment.

Some Thoughts on Instrumental Variables

Slides for Class 10

2020-04-09

PQHS 500 / CRSP 500

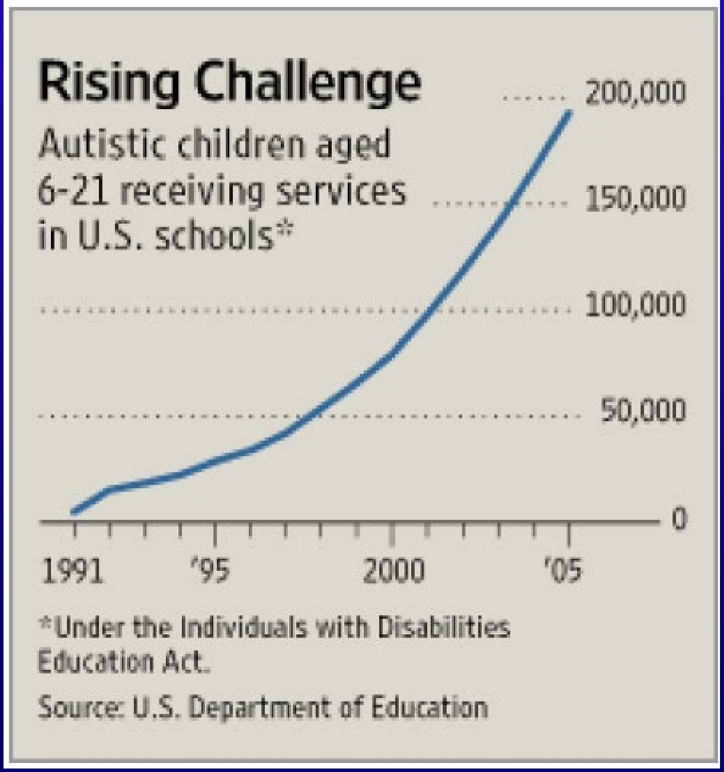
Thomas E. Love, Ph.D.
Thomas.Love@case.edu

https://github.com/THOMASELOVE/2020-500

Is An Economist Qualified to Solve Puzzle of Autism?

By Mark Whitehouse, *Wall Street Journal*, page A1, Feb 27 2007

- Michael Waldman (Cornell)
  - The more it rained or snowed in WA, OR and CA, the more likely kids were to be diagnosed with autism
  - What do kids do more during rain or snow that influences autism?
  - Watch TV?



Waldman’s Interest in Autism

- Professor Waldman’s 2-year-old son was identified with an autism-spectrum disorder in 2003.
- Hoping to eliminate potential triggers, Professor Waldman supplemented recommended therapy with a sharp reduction in TV watching. His son had started watching more TV the previous summer, after a baby sister was born.
- Waldman’s son improved within six months and today has fully recovered -- "When I saw the rapid progress, which was certainly not what anyone had been predicting, I became very curious as to whether television watching might have played a role in the onset of the disorder."

Does TV Trigger Autism?

- Ideal Study: randomly select a group of susceptible babies at birth to refrain from TV
  - Compare their autism rates to a control group
- Economists look for "natural experiments" and use **instrumental variable** methods to hopefully approximate the rigor of a randomized trial.
  - Think of an instrument as a randomized "nudge" towards a treatment…

Levitt and Dubner, *Freakonomics* and *SuperFreakonomics* and NY Times blog, etc.

Instrumental Variables

- IV: a “random”, “policy” or “natural” nudge or encouragement towards a particular exposure that affects the outcome only to the extent that it affects acceptance of the exposure.
- Ideal IV to test exposure-outcome linkage is …
  - Strongly correlated with exposure
  - No direct effect on the outcome or on other factors that cause the outcome. (Exclusion Restriction)
  - Then if data links IV to outcome, it suggests that exposure must be contributing to the outcome.

Understanding the Exclusion Restriction for an Instrument

- Randomized encouragement to either an active drug or a double-blind placebo is the experimental design that most closely approximates an instrument.
  - Encouragement is actually randomized.
  - Neither subject nor investigator knows what treatment subject is being encouraged to do.
  - So there are few opportunities for encouragement to affect a clinical outcome without shifting the amount of active drug that is consumed.

Rosenbaum (2010), Greevy, Silber, Cnaan & Rosenbaum (2004), Holland (1988)

Did Vietnam War service have a negative effect on people's future earnings?

- Not enough to see that vets ended up poorer.
  - Perhaps a lack of opportunities in the civilian world made them more likely to enlist.
- Angrist (@ MIT) chose draft lottery as an IV
  - Some people were more likely than others to serve in the military but this had no connection to their initial circumstances.
  - Data: On average, white men with draft-eligible lottery #s had much lower earnings many years later. (Data on non-whites were inconclusive.)

Angrist, Imbens and Rubin (1996, JASA)

Instrumental Variable for Looking at TV → Autism relationship

- Waldman et al. selected precipitation
  - Kids tended to spend more time in front of TV when it rained or snowed than when it didn't.
  - IV argument: Precipitation "randomly selects" some kids to watch more TV than others.
  - Study conducted in WA, OR and CA where rain and snowfall vary a lot.
- Kids growing up in periods of unusually high precipitation were in fact more likely to be diagnosed with autism.

Going Further

- A second instrument produced a similar result
  - Communities with larger % of households subscribing to Cable also had higher autism rates.
- Conclusions
  - TV watching could be a cause of autism.
  - Precipitation could be linked to potential triggers other than TV watching (household mold?)
  - Marginal Effect: data reflect TV effect on the kids who changed their habits because of rain or snow.
  - Does nothing to explain the mechanism by which TV would influence autism, as in all IV studies.

What About Instrumental Variables?

- Idea: Find a variable (the instrument)
  - strongly correlated with the treatment choice
  - but having no direct effect on the outcome (outside of the instrument's influence on treatment selection)
- If these two conditions are not met, then IV is not a useful approach.
- In health care, treatment selection is usually closely linked to outcome.

Earle et al (2001), Landrum and Ayanian (2001), Posner et al (2001)

Propensity Scores and Instrumental Variables Together: Ambulatory Specialty Care following an Acute MI

- Studying the effect of specialist (cardiologist) vs. generalist outpatient care for acute MI pts.
  - Patients getting specialty care for AMI were younger, less likely to have chronic illnesses, and more likely to have prior cardiac disease.
- Data from Cooperative Cardiovascular Project
  - 200,000 Medicare AMI patients treated in 1994-5.
  - Especially detailed clinical data available in several states – here, we focus on NY fee-for-service pts.

Landrum and Ayanian (2001), Stukel et al. (2007)

Design Specifications for the Study

- Outcome: 18m mortality after AMI discharge
- Treatment: Cardiology vs. Generalist Care
  - 3,551 (65%) **cardiology** care: had at least one office visit with cardiologist in 90d post-discharge
  - 1,916 (35%) **generalist** care: had at least one PCP office visit without any cardiologist visits in 90d
- Unadjusted mortality was substantially lower (9.3% vs. 15.8% at 18m) in cardiology group
  - But the two groups were very different in terms of important baseline covariates linked to mortality…

Treatment Groups are Very Different

- Cardiology patients were:
  - Younger (mean 73.4 vs 74.5, stdzd diff = 20%)
  - More often male, white and to have a prior history of AMI or angina than generalist patients
  - Less likely to also have stroke, COPD or diabetes or prior CHF than generalist patients
  - More likely to be discharged from:
    - A teaching hospital and also an urban hospital
    - A hospital with invasive cardiac services
    - A hospital with cardiology care facilities, or angiography, or angioplasty, or bypass surgery facilities

Standard Logistic Regression Analysis

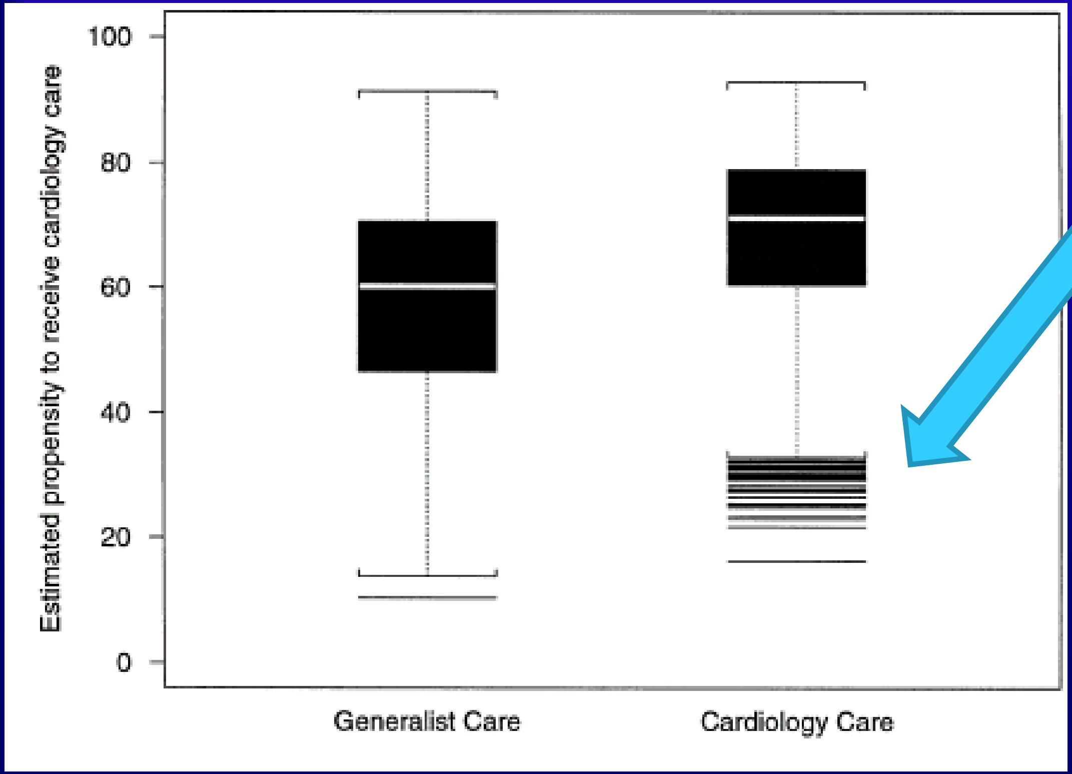
- Adjusting for these (and more) observed differences in a logistic regression model reduced the unadjusted absolute differences in 18-month mortality from 6.5% to 2.9%.
- But given the substantial differences in observed characteristics, and the likelihood that patients differ in terms of unobserved characteristics related to outcome as well…

Propensity Scores for the Ambulatory Care Study

- Logistic regression predicting receipt of cardiology care in 90 days following discharge.
- 41 clinical and provider characteristics…
  - All those covariates described previously
  - Patient demographic and clinical characteristics
  - Treatments received during hospitalization
  - Inpatient provider and hospital characteristics
  - Did not include the “instrument” to be discussed.

Estimated Propensity for Cardiology Care (using 41 clinical and provider covariates) for the Two Groups

- Mean PS is 0.69 in cardiology group, 0.58 in generalist group with substantial overlap…



- 34 patients got cardiology care despite low PS – did we not observe all predictors of cardiology care?

Landrum & Ayanian, Fig 1.

Step 1: Propensity Score Stratification

- Key Results from PS Stratification
  - Balance achieved on most covariates – exceptions are % rural, history of HTN, in-hosp. cardiac arrest
  - In patients least likely to receive cardiology care (by PS quintile) cardiology care estimated to reduce 18m mortality by 11.5 points (12.3% vs 23.8%)
  - Other quintiles show much smaller differences
  - Average of differences = 3.1 points (est. mortality reduction if all pts. in cohort got cardiology care)
  - Differences weighted by cardiology pts in quintile = 2.3 points (average causal effect among the treated)
  - Results insensitive to # of strata = 3 or 10 instead.

Step 2: Propensity Score Matching

- 1775 of 1916 generalist patients (93%) matched to a cardiology patient with closest estimated PS [inside 0.6 SD of logit PS caliper]
  - 1776 unmatched cardiology patients were those with highest propensity for cardiology care
  - Covariate Balance is excellent among matches
  - 18 m mortality for matched cardiology pts. 11.7%
  - 18 m mortality for matched generalist pts. 14.7%
  - 3.0 point absolute reduction in mortality for cardiology care (standard error = 1.1)

Step 3: Instrumental Variable

- Tough part: identifying the instrument
  - Related to treatment but not outcomes
  - Selection: density of cardiologists in patient's county of residence, at two levels (above or below 6.7 cardiologists per 100,000 population age 65+)
- Local Average Treatment Effect for all "Marginal" Patients (also called "Compliers")
  - Would get cardiology if they lived in a high density area but not if they lived in a low density area
  - Cannot identify "compliers" from observed data

More Formalized Instrumental Variables Assumptions

1. Stable Unit Treatment Value Assumption (SUTVA) – unaffected by other pts.
2. Non-zero causal effect of instrument on treatment (IV predicts treatment status)
3. Ignorable assignment of the instrument
4. Exclusion Restriction (IV has no effect on outcomes other than through the treatment)
5. Monotonicity of IV’s effect on treatment

A1. Stable Unit Treatment Value Assumption (SUTVA)

- A patient's potential treatments / outcomes are assumed unrelated to treatment status of other patients
  - Treatment Status (cardiology / generalist) and Mortality assumed unaffected by the care received by other patients
  - Access to care does vary across geographical areas
  - Patients in high-density areas may have increased access to all kinds of specialists…

A2. Non-Zero Causal Effect of Instrument on Treatment

- The IV must predict treatment status
  - We can check this (to some degree) in data.
  - Likelihood of receiving cardiology care was significantly and positively associated with quintile of cardiology density ($p < 0.001$) in New York
  - This wasn’t true in other states (TX, CA, MA) – so this instrument wouldn’t be appropriate there.

A3. Ignorable Assignment of IV

- Patients from different density areas must be similar (in both observed and unobserved characteristics) to what they would have been had density been randomly assigned.
  - Can't verify directly, but if patients are similar in terms of observed characteristics, that provides some evidence of the validity of the assumption.
  - In fact, observed data in NY looks balanced when we compare low density to high density patients.
  - Not true in other states (FL high density areas had older patients, for instance)

A4. Exclusion Restriction

- IV assumed to have no effect on outcomes other than through its effect on the treatment
  - Can’t verify this directly, either, but …
  - Density is actually correlated with hospital characteristics and with in hospital treatments.
  - Instrument could have an effect on mortality through these other characteristics / treatments.
  - High density areas were more urban, and care processes differ between urban and rural areas.

A5. Monotonicity

- IV assumed to affect treatment monotonically
  - Can’t verify this directly, either.
  - If a patient in a low-density area received cardiology care, have to believe (s)he would also have received cardiology care if (s)he lived in a high density area.
  - Often seems pretty reasonable with this sort of instrument.

IV Analytic Decisions in this Study

- Divided patients into quintiles according to density of cardiologists in their county
- Non-parametric approaches to estimating treatment effects: avoid further assumptions
  - Density strongly related to urban/rural location, so also estimated these effects separately
  - Hospital (teaching or not) and inpatient treatment (coronary angiography or not) also correlated with density, so estimated treatment effects at fixed levels of those characteristics as well.

Landrum and Ayanian (2001)

Instrumental Variable Results

- LATE estimate = difference in 18m mortality (cardiology – generalist) among pts. for whom cardiologist supply determined treatment.

	LATE	Standard Error
IV model with no covariates	-9.5 %	7.9
Controlling for teaching hospital and inpatient treatment as well	-1.0 %	8.4

- SEs are large because treatment determined by supply for only about 15% of population.
  - Lowest Quintile of Density: 57% got cardiology
  - Highest Quintile of Density: 72% got cardiology

Comparison of PS and IV Approaches

- PS analyses found a small but significant benefit, concentrated among patients with lowest propensity to receive cardiology care.
- IV point estimates were consistent with a small benefit to ambulatory cardiology care, but were not precisely estimated, so the differences between groups were not statistically significant.

Issues to Consider

- Both PS and IV approaches rely on critical and untestable assumptions.
- They are looking at different things - if there is heterogeneity in the impact of cardiology care across strata of patients, PS and IV estimates of causal effects may differ even if both sets of assumptions hold up.
- Methods estimate effects for different people.

Importance of the Policy Question

- PS analysis lets us identify characteristics of the population to make recommendations for individual patients (subgroup analyses, etc.)
- IV analysis more applicable if we want to look at, say, the impact of increasing the supply of cardiologists, because IV demonstrates the marginal effect of such changes.

Surgery vs. Medical Therapy for Native Valve Endocarditis [PS and IV]

Analysis of the Impact of Early Surgery on In-Hospital Mortality of Native Valve Endocarditis

Use of Propensity Score and Instrumental Variable Methods to Adjust for Treatment-Selection Bias

Tahaniyat Lalani, MBBS, MHS; Christopher H. Cabell, MD, MHS; Daniel K. Benjamin, PhD; Ovidiu Lasca, BA; Christoph Naber, MD; Vance G. Fowler, Jr, MD, MHS; G. Ralph Corey, MD; Vivian H. Chu, MD, MHS; Michael Fenely, MD; Orathai Pachirat, MD; Ru-San Tan, MBBS, MRCP; Richard Watkin, MRCP; Adina Ionac, MD, PhD; Asuncion Moreno, MD; Carlos A. Mestres, MD, PhD; José Casabé, MD, PhD; Natalia Chipigina, PhD; Damon P. Eisen, MBBS, MD, FRACP; Denis Spelman, MBBS, FRACP, FRCPA, MPH; Francois Delahaye, MD, PhD; Gail Peterson, MD; Lars Olaison, MD, PhD; Andrew Wang, MD; for the International Collaboration on Endocarditis–Prospective Cohort Study (ICE-PCS) Investigators

Background—The impact of early surgery on mortality in patients with native valve endocarditis (NVE) is unresolved. This study sought to evaluate valve surgery compared with medical therapy for NVE and to identify characteristics of patients who are most likely to benefit from early surgery.

Methods and Results—Using a prospective, multinational cohort of patients with definite NVE, the effect of early surgery on in-hospital mortality was assessed by propensity-based matching adjustment for survivor bias and by instrumental variable analysis. Patients were stratified by propensity quintile, paravalvular complications, valve perforation, systemic embolization, stroke, *Staphylococcus aureus* infection, and congestive heart failure. Of the 1552 patients with NVE, 720 (46%) underwent early surgery and 832 (54%) were treated with medical therapy. Compared with medical therapy, early surgery was associated with a significant reduction in mortality in the overall cohort (12.1% [87/720] versus 20.7% [172/832]) and after propensity-based matching and adjustment for survivor bias (absolute risk reduction [ARR] −5.9%, $P<0.001$). With a combined instrument, the instrumental-variable–adjusted ARR in mortality associated with early surgery was −11.2% ($P<0.001$). In subgroup analysis, surgery was found to confer a survival benefit compared with medical therapy among patients with a higher propensity for surgery (ARR −10.9% for quintiles 4 and 5, $P=0.002$) and those with paravalvular complications (ARR −17.3%, $P<0.001$), systemic embolization (ARR −12.9%, $P=0.002$), *S aureus* NVE (ARR −20.1%, $P<0.001$), and stroke (ARR −13%, $P=0.02$) but not those with valve perforation or congestive heart failure.

Conclusions—Early surgery for NVE is associated with an in-hospital mortality benefit compared with medical therapy alone. (***Circulation*. 2010;121:1005-1013.**)

Key Words: surgery ■ endocarditis ■ drug therapy ■ hospital mortality

- Conclusion: Early surgery for NVE is associated with lower rates of in-hospital mortality.
- Hard to attribute mortality difference to selection bias - either observed or “hidden.”

Lalani et al. (2010)

Lalani et al. approach: Propensity Scores + Instrumental Variables

- Cohort study: Estimated PS(early surgery), did 1-1 matching with replacement (among medical patients) within a 0.05 caliper
  - Required each medical patient to survive at least as long as the time to surgery in the matched surgical patient so as to match on follow-up times
- Constructed an instrument (using BIPROBIT routine in Stata) - combination of covariates.
  - High correlation with early surgery, no direct relationship with mortality.

When Are Instrumental Variables Methods Especially Attractive?

An instrument is available, and …

- Assignment to a treatment is ignorable, but compliance with the assignment is not perfect so that the dose of treatment received is non-ignorable.
- Data are weak, in the sense that observed covariates provide insufficient insight into the background to allow estimated effects (adjusting for covariates) to be due to treatment.

Propensity Scores vs. Instrumental Variables?

- Some questions call for PS adjustment, others for IV models of Rx effect.
- Both have unverifiable assumptions:
  - PS adjusts for selection bias in terms of identified covariates – we must presume this is sufficient to also adjust for unobserved covariates. Sensitivity analysis can help.
  - IV presumes we can and do identify appropriate instrument(s).