500 Class 08 (Recording)

https://thomaselove.github.io/500-2025/

2025-03-06

What's in these Slides?

- Instrumental Variables
  - Is An Economist Qualified to Solve Puzzle of Autism? (Whitehouse 2007)
  - The Exclusion Restriction
  - Propensity Scores and Instruments, Together
  - The Formalized Instrumental Variables Assumptions
- Landrum and Ayanian 2001 comparing methods

Is An Economist Qualified to Solve Puzzle of Autism?

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

> *In the spring of 2005, Cornell University economist Michael Waldman noticed a strange correlation in Washington, Oregon and California. The more it rained or snowed, the more likely children were to be diagnosed with autism.*

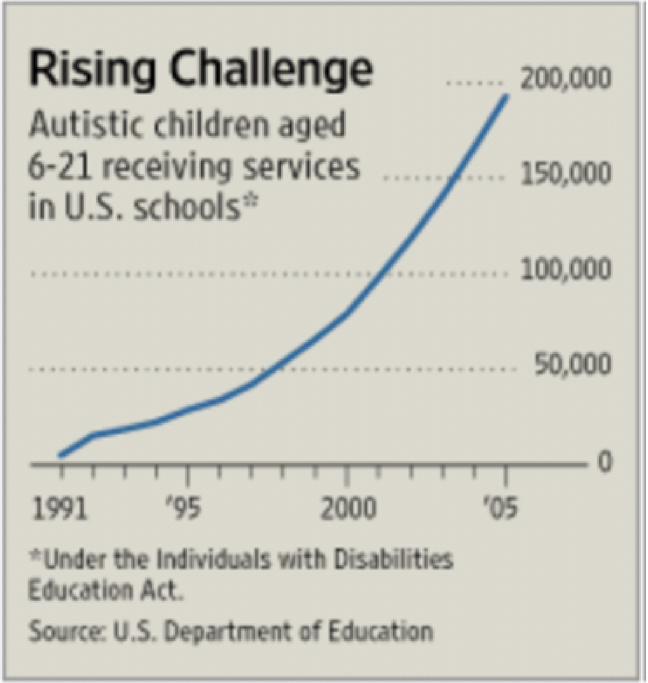
- What do kids do more during rain or snow that influences autism?

Is An Economist Qualified to Solve Puzzle of Autism?

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

> *In the spring of 2005, Cornell University economist Michael Waldman noticed a strange correlation in Washington, Oregon and California. The more it rained or snowed, the more likely children were to be diagnosed with autism.*

- What do kids do more during rain or snow that influences autism?
- Watch TV?


Rising Challenge
Autistic children aged 6-21 receiving services in U.S. schools*
200,000
150,000
100,000
50,000
0
1991
'95
2000
'05
*Under the Individuals with Disabilities Education Act.
Source: U.S. Department of Education

Autism on the Rise

Studies in recent decades have shown the proportion of children with autism growing, though researchers aren't sure the disorder has actually become more prevalent. Greater awareness, broadening definitions of the disorder and the availability of special-education programs may have made parents more likely to get their children diagnosed.

Most researchers now recognize that heredity plays a central role in autism, and they are making progress in identifying the genes responsible. They're also looking into the possibility of interaction with environmental factors, both in the womb and after birth.

Professor 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*, for instance, provide popular treatments of some key ideas.

What is an Instrumental Variable?

- An instrumental variable is 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.
- An ideal instrumental variable to test the link between an exposure and an outcome has...
  - a strong correlation with receipt of the exposure
  - no direct effect on the outcome or on other factors that cause the outcome (exclusion restriction)

Then, if data links the IV to the outcome, it suggests that the *exposure* must be contributing to the outcome.

Back to Autism for a Moment (from Whitehouse, WSJ)

In principle, the best way to figure out whether television triggers autism would be to do what medical researchers do: randomly select a group of susceptible babies at birth to refrain from television, then compare their autism rate to a similar control group that watched normal amounts of TV. If the abstaining group proved less likely to develop autism, that would point to TV as a culprit.

Economists usually have neither the money nor the access to children needed to perform that kind of experiment… Instead, economists look for instruments – natural forces or government policies that do the random selection for them.

On Instruments (from Whitehouse, WSJ)

First developed in the 1920s, the technique helps them separate cause and effect.

Establishing whether A causes B can be difficult, because often it could go either way. If television watching were shown to be unusually prevalent among autistic children, it could mean either that television makes them autistic or that something about being autistic makes them more interested in TV.

The ideal instrument is a variable that is correlated with A but has no direct effect of its own on B. It should also have no connection to other factors that might cause B. If data in a study nonetheless show that the instrumental variable is linked to B, it suggests that A must be contributing to B.

A Simple Example

This example largely comes from https://www.statology.org/instrumental-variables/.

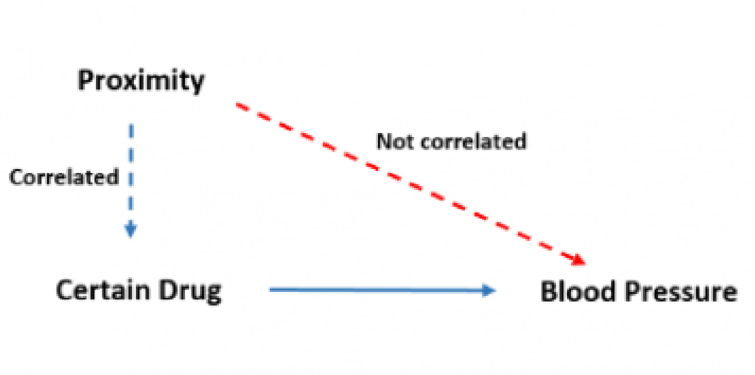
1. Suppose we want to estimate the effect that a certain drug has on systolic blood pressure.
2. Other variables like time spent exercising, overall diet, and stress levels also affect systolic blood pressure.
3. So if we run a simple linear regression to predict SBP based on using the drug or not using the drug alone, we can't be sure that we have accurately captured the effect of the drug on SBP.

Identifying an Instrument

An instrumental variable is a third variable introduced into regression analysis that is correlated with the predictor variable, but uncorrelated with the outcome variable. Our goal is to estimate the true causal effect that our exposure has on our outcome.

An example of an instrumental variable that we may use in this regression analysis is an individual's proximity to a pharmacy.

- "proximity" would likely be highly correlated with whether or not the individual takes the certain drug because an individual wouldn't be able to obtain it if they don't live near a pharmacy.
- But "proximity" isn't expected to have any correlation with SBP (our outcome) except through the exposure (taking the drug.)


Proximity
Correlated
Not correlated
Certain Drug
Blood Pressure

Instrumental Variables Regression (Two-Stage Least Squares)

1. Fit a regression using the instrument as the predictor of our exposure.

Certain Drug $= \beta_{01} + \beta_{11}$ (Proximity)

Predicted Values are called $\hat{cd}$.

Stage 2 of the IV Regression

2. Fit a regression using the predicted values of our exposure from model 1 as the predictor for our outcome.

SBP $= \beta_{02} + \beta_{12} \hat{cd}$

Because we used "proximity" alone to come up with $\hat{cd}$ and since we know that proximity should not be correlated with SBP, the effect seen in the second stage regression can be attributed to the drug.

So the effect captured in β_{12} then tells us about a causal effect of the drug on SBP.

Key Assumptions in this Simple Example

1. Our instrument (proximity) is highly correlated with our exposure (use of the drug.)
2. Our instrument (proximity) is not correlated with our outcome (systolic blood pressure.)
3. Our instrument (proximity) is also not correlated with the other variables that are left out of the model (like exercise, diet, stress.)

If the instrument does not meet these criteria, it will likely produce unreliable and biased results.

Instruments have changed the world

> *Joshua Angrist (and others) have studied the cost of war. Steven Levitt has examined the effect of adding police on crime (for example.) Their work has played an important role in public-policy debates.*

Angrist's work (with Imbens and Rubin in 1990) is perhaps the most cited.

> *Did service during the Vietnam War have a negative effect on people's future earnings?*
>
> *It wouldn't be enough to say that people who served ended up poorer. Perhaps a lack of opportunities in the civilian world made them more likely to enlist in the first place.*

Angrist (1990)

As an instrumental variable, Professor Angrist chose the draft lottery, which made some people more likely than others to serve in the Vietnam-era military, but didn't have any connection to their initial circumstances.

- Data: On average, white men with draft-eligible lottery numbers had much lower earnings many years later.
- Data on non-whites were inconclusive.

  Prof. Angrist concluded that conscription had a detrimental effect on future earnings.

Also see Angrist, Imbens and Rubin (1996, JASA)

Levitt (1997)

Chicago's Professor Levitt tackled police staffing and crime. That's an issue where cause and effect are hard to disentangle because cities with many criminals are likely to have more police, but that doesn't mean an excess of officers causes crime.

Prof. Levitt took advantage of the fact that mayors and governors tend to put more police on the streets in election years. Using election cycles, he concluded in a 1997 paper that adding police reduces violent crime.

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.

See Rosenbaum (2010) or Greevy et al (2004) or Holland (1998)

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.

A second instrument

- Communities with larger rates 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.

The Instrumental Variable 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.

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

Section 1

Landrum Mary Beth and Ayanian John Z 2001 Causal Effect of Ambulatory Specialty Care on Mortality Following Myocardial Infarction: A Comparison of Propensity Score and Instrumental Variable Analyses

Landrum and Ayanian (2001)

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

- Landrum and Ayanian (2001) studied 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: focus here is on NY fee-for-service patients.

Landrum and Ayanian (2001) Design

- 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...

Landrum and Ayanian (2001) 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

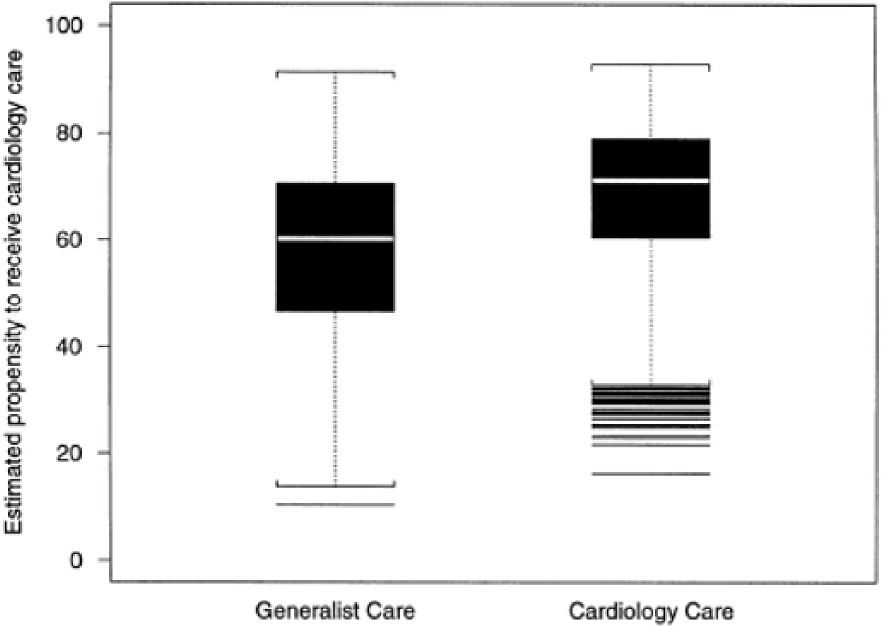
- Logistic regression predicting receipt of cardiology care in 90 days following discharge.
- 41 clinical and provider characteristics included...
  - All those covariates described previously
  - Patient demographic and clinical characteristics
  - Treatments received during hospitalization
  - Inpatient provider and hospital characteristics

The propensity model did not include the "instrument" we will discuss shortly.

Overlap in Propensity Score shown in next slide:

- 34 patients got cardiology care despite a low PS.
- Did we not observe all predictors of cardiology care?

Overlap in Propensity for Cardiology Care



Step 1. Propensity Score 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 using 3 or 10 strata instead of 5

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

Among matched patients,

- 18 m mortality for cardiology patients was 11.7%
- 18 m mortality for generalist patients was 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+)

We'll obtain an estimate of the 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

Formalized Instrumental Variables Assumptions

1. SUTVA (Stable Unit Treatment Value Assumption) - unaffected by other subjects
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

The SUTVA Assumption

- 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

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 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.

Ignorable Assignment of the Instrument

- 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)

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 pretty strongly 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.

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.

Analytic Decisions in Landrum and Ayanian

- 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 strongly correlated with density, so estimated treatment effects at fixed levels of those characteristics as well.

Landrum and Ayanian IV Results

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

- IV Model with No Covariates: LATE = -9.5%, SE = 7.9
- IV Model with Covariates: LATE = -1.0%, SE = 8.4

Covariates controlled for teaching hospital and inpatient treatments.

- 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.

When are IV methods especially attractive?

1. An instrument is available, and...
2. Assignment to a treatment is ignorable, but compliance with the assignment is not perfect so that the dose of treatment received is non-ignorable.
3. 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.

- An interesting perspective on the roles of IV and PS, (and a nice application of PS with subclassification) is provided in Coyte et al.'s joint replacement study (2000) J of Health Econ.

Propensity Scores vs. Instrumental Variables?

- Both propensity score methods and instrumental variables (IV) can be used to adjust for unobserved covariates that affect treatment assignment when treatment assignment and treatment outcome are confounded.
- Some questions call for PS adjustment, others for IV models of Rx effect.
- Each 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).

Agenda for Zoom Call at 10 AM 2025-03-06

- How did Lab 3 go?
- Tanenbaum (2017) paper
- Rosenbaum Chapter 7 discussion