Segment 6: Quasi Experiments

Section 01: Instrumental Variables

Operating Question: Segment 6

Are there any non-experimental circumstances when I can make causal inferences without conditional ignorability?

Conditional Ignorability!

$$Y^0, Y^1 \perp\!\!\!\perp Z|X$$

- In an observational study, no guarantee that we thought of (or measured) all the required X needed
  - Probably the most cited threat to validity of an observational study
- Can there be study circumstances that offer a similar guarantee purely by feature of the study design?
  - That is, can we somehow recreate the conditions of a randomized study that "automatically" balances observed and unobserved confounders
  - *Without* having to explicitly condition on X
  - As would be the case in a completely randomized study

From "Restructuring" to "Getting Lucky"

- Most of the methods we've discussed for analyzing observational studies rely in large part on:
  - Knowing (and observing) the "right" X to believe the conditional ignorability assumption
  - Using those X to "restructure" the data to recreate the conditions of a randomized experiment

- A different class of methods aims to directly leverage circumstances of the study to *avoid* conditional ignorability:
  - If the circumstances are right, we may not need to "restructure" with X
  - We may be lucky that we can otherwise leverage the conditions of an "as if randomized" study design
  - Even without conditional ignorability
  - "Automatically" adjust for **unobserved** confounders

Quasi-Experimental Methods

- Maybe we can't enumerate/measure all X
- Are there certain study circumstances that can "naturally" recreate the conditions of a randomized study?
  - "Natural experiment": nature randomly assigns some treatment variable
  - Natural grouping of units
  - Treatment is "quasi-randomized"
- **Very** specialized circumstances that offer the potential to control for unmeasured confounders
- Methods to take advantage of these specialized circumstances are sometimes described as "quasi-experimental" methods
  - Many have a long history in econometrics, but have more recently gained popularity in other disciplines

Instrumental Variables (IVs)

Setting: We are interested in the effect of a treatment (T) on an outcome Y, but we are not confident that we can satisfy conditional ignorability ($|X$)....

...But the data have a particular type of variable, called an *instrument*, Z, that is:

1. Essentially randomly assigned (unrelated to potential outcomes)
2. Predictive of T
3. Otherwise unrelated to Y

This could arise because of a "broken" randomized study or because of lucky "natural" circumstances

Example: Randomized Encouragement Design

Sesame Street example from Gelman, Hill, Vehtari

- Units: preschool children
- T: watching sesame street ($T = 1$) vs. not ($T = 0$)
- Y: letter recognition

...but even if we *randomize* children to watch, we cannot *force* them to watch...

- Z, random *encouragement* to watch ($Z = 1$) vs. not ($Z = 0$).
- Z is an *instrument* that induces variation in T, but is otherwise unrelated to Y

Key Idea: Leverage the randomized nature of Z to estimate a causal effect of T *in a certain subpopulation of units*

Example: Treatment Noncompliance

E.g., Shetty et al. (2008)

- Units: facial surgery patients
- Z: randomized treatment assignment
- T: receiving RIF surgery ($T = 1$) vs. MMF ($T = 0$)
- Y: Post-treatment complications

...but even though we *randomized* patients to receive RIF or MMF, the operating surgeon could *overrule* the randomization...

- Noncompliance $\rightarrow Z$ is an *instrument* that induces variation in T, but is otherwise unrelated to Y
- Intention to Treat Effect: $\bar{Y}|Z = 1 - \bar{Y}|Z = 0$, effect of *randomization* not the effect of *treatment*
- Actually possible to estimate the causal effect of *treatment*, T, in a certain subpopulation of units

Example: Hypothetical Dietary Experiment

With treatment noncompliance

Table: Observed Data from the Hypothetical Dietary Experiment, **Idealized Assignment**

Unit, i	Female, x_{1i}	Age, x_{2i}	Treatment Z_i	Potential T_i^0	Potential T_i^1	Potential Y_i^0	Potential Y_i^1
Audrey	1	40	0	0	1	140	135
Anna	1	40	1	0	0	140	135
Bob	0	50	0	0	0	150	140
Bill	0	50	1	0	0	150	140
Caitlin	1	60	0	0	1	160	155
Cara	1	60	1	0	1	160	155
Dave	0	70	0	1	0	170	160
Doug	0	70	1	0	1	170	160

General Formulation

with binary Z, T

The sample can be stratified into four strata based on the values of T^0, T^1

T^0	T^1	
0	0	"Never-Takers"
0	1	"Compliers"
1	0	"Defiers"
1	1	"Always-Takers"

This language is often used to frame IV methods that are not about "noncompliance" per se.

Local Average Treatment Effects

Instrumental variables can permit estimation of causal treatment effects (of T) in a *certain subpopulation of units*. But what subpopulation?

- Only *some* units have their T affected by Z
  - E.g., T_i^0 =rarely watched, T_i^1 = frequently watched
  - E.g., T_i^0 =MMF, T_i^1 =RIF
- IV methods can estimate causal effects of T *only in the units whose* T *could be altered by* Z
  - $\{\text{i}; T_i^0 \neq T_i^1\}$
  - *Local Average Treatment Effect* (LATE)
- IV typically focuses on a particular LATE called the *Complier Average Causal Effect* (CACE)
  - $\{i; T_i^0 = 0, T_i^1 = 1\}$: "Compliers"
- Other groups of units don't have a well-defined effect of T
  - $\{i; T_i^0 = 0, T_i^1 = 0\}$: "Never-takers"
  - $\{i; T_i^0 = 1, T_i^1 = 1\}$: "Always-takers"

Key IV Assumptions

1. Ignorability of the instrument
2. Monotonicity
3. Nonzero association between instrument and treatment
4. Exclusion restriction

The viability of these assumptions will be based on specialized circumstances of the study, and can be very difficult to justify in practice...but a valid instrumental variable can be *very* powerful for causal inference

IV Assumption 1: Ignorability of instrument

$$Y^0, Y^1 \perp\!\!\!\perp Z$$

- Trivial in a "broken" randomized experiment
  - Random encouragement design
  - Treatment noncompliance
- May also occur as some form of *natural experiment*
- May be difficult to justify in an observational study

IV Assumption 2: Monotonicity

Having $Z = 1$ (vs. $Z = 0$) can only *increase* the value of T

AKA, "no defiers:

$$\{i; T_i^0 = 1, T_i^1 = 0\} = \emptyset$$

IV Assumption 3: Nonzero $Z \leftrightarrow T$ Association

The instrument (Z) needs to have at least some impact on which units receive treatment (T)

If the instrument doesn't induce any unit to adopt treatment, then there is no subpopulation in which we can learn the effect of T

IV Assumption 4: Exclusion Restriction

The instrument has exactly zero effect on the outcome among units for whom the instrument does not affect the treatment

The instrument can only affect the outcome "through" changing the treatment

No causal effect of the instrument on the outcome in "always takers" and "never takers"

Note: In practice, violations of the exclusion restriction are the most common threat to the validity of IV analyses

Example: Randomized Encouragement Design

Sesame Street example from Gelman, Hill, Vehtari

1. Ignorability of the instrument
   - Randomized study
2. Monotonicity
   - Encouraging to watch sesame street doesn't cause children to *not* watch
   - No children would *not watch* if encouraged to do so, but *watch* if not encouraged ("no defiers")
3. Nonzero association between instrument and treatment
   - At least some children watched *because* of the encouragement
4. Exclusion restriction
   - Simply encouraging to watch cannot change in a child's letter recognition without actually inducing them to watch
   - **Threat:** What if some parents never let their child watch TV, but the encouragement materials prompted *other* educational activities that improved letter recognition?

Identification of Causal Effects Using Instrumental Variables

Joshua D. ANGRIST, Guido W. IMBENS, and Donald B. RUBIN

We outline a framework for causal inference in settings where assignment to a binary treatment is ignorable, but compliance with the assignment is not perfect so that the receipt of treatment is nonignorable. To address the problems associated with comparing subjects by the ignorable assignment—an "intention-to-treat analysis"—we make use of instrumental variables, which have long been used by economists in the context of regression models with constant treatment effects. We show that the instrumental variables (IV) estimand can be embedded within the Rubin Causal Model (RCM) and that under some simple and easily interpretable assumptions, the IV estimand is the average causal effect for a subgroup of units, the compliers. Without these assumptions, the IV estimand is simply the ratio of intention-to-treat causal estimands with no interpretation as an average causal effect. The advantages of embedding the IV approach in the RCM are that it clarifies the nature of critical assumptions needed for a causal interpretation, and moreover allows us to consider sensitivity of the results to deviations from key assumptions in a straightforward manner. We apply our analysis to estimate the effect of veteran status in the Vietnam era on mortality, using the lottery number that assigned priority for the draft as an instrument, and we use our results to investigate the sensitivity of the conclusions to critical assumptions.

KEY WORDS: Compliers; Intention-to-treat analysis; Local average treatment effect; Noncompliance; Nonignorable treatment assignment; Rubin-Causal-Model; Structural equation models.

Example: Military Service

From Angrist, Imbens, and Rubin (1996)

- Units: US adults eligible for military service during the Vietnam era
- T: Serving in the military ($T = 1$) vs. not ($T = 0$)
- Y: Civilian (non-combat) Mortality
- Z: Low draft lottery number ($Z = 1$) vs. high draft lottery number ($Z = 0$)
- Some people served ($T = 1$) even with high draft numbers ($Z = 0$) and *vice versa* $\rightarrow$ Noncompliance

Example: Military Service

From Angrist, Imbens, and Rubin (1996)

1. Random instrument?
   - Draft was a lottery
2. Monotonicity?
   - Would anyone have enlisted in the military with a high draft number but avoided service with a low draft number?
3. Nonzero $Z \leftrightarrow T$ association?
   - Were there at least *some* people who enlisted *because* of the draft?
4. Exclusion restriction
   - Having a low draft number does not impact mortality without also impacting military service. Draft number has no "direct effect" on civilian mortality
   - **Threat:** People with low lottery numbers changed educational plans to defer their service $\rightarrow$ impact on civilian mortality

Evaluating Short-Term Drug Effects Using a Physician-Specific Prescribing Preference as an Instrumental Variable

M. Alan Brookhart, Philip S. Wang, Daniel H. Solomon, and Sebastian Schneeweiss

Background: Postmarketing observational studies of the safety and effectiveness of prescription medications are critically important but fraught with methodological problems. The data sources available for such research often lack information on indications and other important confounders for the drug exposure under study. Instrumental variable methods have been proposed as a potential approach to control confounding by indication in nonexperimental studies of treatment effects; however, good instruments are hard to find.

Methods: We propose an instrument for use in pharmacoepidemiology that is based on a time-varying estimate of the prescribing physician's preference for one drug relative to a competing therapy. The use of this instrument is illustrated in a study comparing the effect of exposure to COX-2 inhibitors with nonselective, nonsteroidal anti-inflammatory medications on gastrointestinal complications.

Results: Using conventional multivariable regression adjusting for 17 potential confounders, we found no protective effect due to COX-2 use within 120 days from the initial exposure (risk difference = −0.06 per 100 patients; 95% confidence interval = −0.26 to 0.14). However, the proposed instrumental variable method attributed a protective effect to COX-2 exposure (−1.31 per 100 patients; −2.42 to −0.20) compatible with randomized trial results (−0.65 per 100 patients; −1.08 to −0.22).

Conclusions: The instrumental variable method that we have proposed appears to have substantially reduced the bias due to unobserved confounding. However, more work needs to be done to understand the sensitivity of this approach to possible violations of the instrumental variable assumptions.

trials are underpowered to detect uncommon adverse events, nonexperimental postmarketing studies are needed to evaluate the safety of approved drugs.[2,3] This function of postmarketing research has come under increased scrutiny with the recent safety concerns pertaining to several widely used therapeutic agents, including selective COX-2 inhibitors, hormone replacement therapy, and selective serotonin reuptake inhibitors.

Studies of outcomes associated with exposure to pharmaceutical products as they are used in routine practice are inherently nonexperimental. Often such studies are based on healthcare claims data containing longitudinal information on pharmacy dispensing, healthcare encounters, procedures, and International Classification of Diseases-coded diagnoses.[4] Although these files contain data on large populations followed over extended periods of time, they often lack detailed information on clinical indications for specific therapies. This problem is thought to be particularly acute in studies of intended drug effects because of the difficulty in adjusting for confounding by indication[5]; ie, patients who are thought to benefit most from a drug are more likely to receive therapy.[6–8] Although epidemiologists have a variety of design options and analytic tools to adjust for measured confounders, pharmacoepidemiologic studies have consistently been criticized for having incomplete information on many potential predictors of study outcomes that might lead to selective prescribing.[9–13] Some authors argue that it is impossible for current epidemiologic methods to fully adjust confounding by indication in studies

Example: Prescribing Preferences

From Brookhart et al. (2006)

- Units: 65+ year old new oral NSAID users
- T: COX-2 NSAID ($T = 1$) vs. non-selective NSAID ($T = 0$)
- Y: Short-term GI complications
- Z: Prescribing physician's preference for COX-2 vs. nonselective NSAID based on the *last* written prescription
- Receipt of COX-2 vs. non-selective NSAID is definitely not random....but maybe physician preference is random with respect to patient's outcomes

Example: Prescribing Preferences

From Brookhart et al. (2006)

1. Random instrument?
   - Physician's previous prescription is unrelated to GI risk for current patient
   - **Threat:** But maybe physicians who frequently prescribe COX-2 tend to see higher risk patients....
2. Monotonicity?
   - Would a $Z = 1$ preferring physician always prescribe $T = 0$ and a $Z = 0$ physician always prescribe $T = 1$?
3. Nonzero $Z \leftrightarrow T$ association?
4. Exclusion restriction
   - Physician's prescribing preference cannot directly impact GI outcomes, it can only impact "through" prescription choice
   - **Threat:** Physicians who frequently proscribe COX-2 may also be more likely to co-prescribe *other* therapies that impact GI outcomes

Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?

Analysis Using Instrumental Variables

Mark McClellan, MD, PhD; Barbara J. McNeil, MD, PhD; Joseph P. Newhouse, PhD

Objective.—To determine the effect of more intensive treatments on mortality in elderly patients with acute myocardial infarction (AMI).

Design.—Analysis of incremental treatment effects using differential distances as instrumental variables to account for unobserved case-mix variation (selection bias) in observational Medicare claims data (1987 through 1991).

Main Outcome Measure.—Survival to 4 years after AMI

and comprehensive data se
able to evaluate medical tre
providers,[4] outcomes resea
the fastest-growing fields i
search today. Its results ca
determine not only what ef

Example: Distance to Hospital

From McClellan et al. (1994)

- Units: 65+ year old US Medicare beneficiaries who had a heart attack (AMI)
- T: Intensive treatment ($T = 1$) vs. non-intensive ($T = 0$)
- Y: Survival
- Z: Distance from patient's residential location to hospital with intensive treatment capabilities
- Many patients have hospital (or procedure) selected based on risk
- But some patients end up in an intensive vs. non-intensive hospital based on which type of hospital is closest
  - A "complier" is a patient who would receive $T = 1$ if living close to a treatment-intensive hospital, but $T = 0$ if living far from a treatment-intensive hospital

General Analytic Framework for IVs

T^0	T^1	c
0	0	"Never-Takers"
0	1	"Compliers"
1	0	"Defiers"
1	1	"Always-Takers"

Considerations for IV Estimation

- Have to get pretty lucky...
  - Very hard to satisfy the IV assumptions outside of "broken" randomized experiments
- CACE tells us about an average affect in a well-defined subpopulation, but we still don't know exactly which units are compliers
- ITT may be more/less relevant than a LATE
  - E.g., If we can't control who actually *takes* treatment, maybe ITT corresponds to the more practical policy where all we can do is *encourage*
- Can be formulated via regression equations and two-stage least squares
  - More common in econometrics literature
  - C.f. Angrist, Imbens, and Rubin (1996)