Week 11: Causality with Unmeasured Confounding

Brandon Stewart[1]

Princeton

November 28–30, 2022

[1] Many slides are from Matt Blackwell, Adam Glynn, Jens Hainmueller and Ian Lundberg.

Where We've Been and Where We're Going...

- Last Week
  - selection on observables and measured confounding
- This Week
  - natural experiments
  - instrumental variables
  - other approaches
- The Following Week
  - repeated observations and wrap up
- Long Run
  - probability $\rightarrow$ inference $\rightarrow$ regression $\rightarrow$ causal inference

1 Natural Experiments

2 Instrumental Variables

3 Regression Discontinuity

4 Interrupted Time-Series

Unmeasured Confounding

- Last week we considered cases of measured confounding

$$\begin{tikzcd} & X \arrow[dl] \arrow[dr] & \\ A \arrow[rr] & & Y \end{tikzcd}$$

- In this case we block the backdoor path $A \leftarrow X \rightarrow Y$ by conditioning on X.
- What happens in the general case where X is unobserved?
- Under selection on unobservables we are going to need a different approach. We will talk about several over the next two weeks.
- There is No Free Lunch $\rightsquigarrow$ we can't get something for nothing, we will need new variables, new assumptions and/or new approaches.
- Goal: give you a feel for what is possible, but note that you will need to do work beyond class if you want to use one of these techniques. Think "on ramp" more than "comprehensive tutorial."

Approaches to Unmeasured Confounding

- Natural Experiments
- Instrumental Variables
- Regression Discontinuity
- Interrupted Time-Series/Diff-in-Diff/Synthetic Control
- Bounding
- Sensitivity Analysis
- Front Door Adjustment

Natural Experiments

- Natural Experiment: broadly speaking an "experiment" where the treatment is randomized but the randomization was not controlled by the researcher.
- Leverages an exogenous (external to the system) event to measure the effect of an otherwise endogenous (internal to the system) phenomenon.
- Trickier to analyze than regular experiments
  - we ought to be suspicious of randomization we don't control
  - nature may not choose exactly the treatment we want
  - not immediately obvious which groups are comparable
  - valid comparison may not estimate the causal effect of interest
- When available, a useful way to capitalize on randomness in the world to make causal inferences.
- See Dunning (2012) *Natural Experiments in the Social Sciences*

Caution on terminology

It is worth noting that the label "natural experiment" is perhaps unfortunate. As we shall see, the social and political forces that give rise to as-if random assignment of interventions are not generally "natural" in the ordinary sense of that term. Second, natural experiments are observational studies, not true experiments, again, because they lack an experimental manipulation. In sum, natural experiments are neither natural nor experiments.

—Dunning (2012) pg 16

Natural Experiment Examples (True Randomization)

Randomness	Focus	Citation
Vietnam draft	labor market	Angrist 1990
randomized quotas	female leadership in Indian village council presidencies	Chattopadhyay & Duflo 2004
randomized term lengths	tenure in office on legislative performance	Dal Bo & Rossi 2010
lottery	effect of winnings on political attitudes	Doherty, Green & Gerber 2006
randomized ballot order	ballot order effects in CA	Ho & Imai 2008

Natural Experiment Examples (As If Randomization)

Randomness	Focus	Citation
election monitor assignment	international election monitoring on fraud	Hyde 2007
random shelling by drunk soldiers	indiscriminate violence on rebellion	Lyall 2009
hurricane	study of friendship formulation	Phan and Airoldi 2015
2006 Israel-Hezbollah war	stress on unborn babies	Torche and Shwed 2015
Snowden revelations	reading behavior on wikipedia	Penney 2016
terrorist attacks	perception of immigrants	Legewie 2013

Questions to Ask Yourself

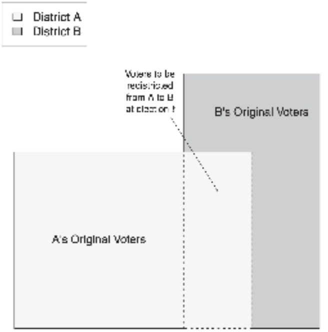
From Sekhon and Titiunik (2012) "When Natural Experiments Are Neither Natural nor Experiments" *American Political Science Review*

1. "is the proposed treatment-control comparison guaranteed to be valid by the assumed randomization?"
2. "if not, what is the comparison that is guaranteed by the randomization, and how does this comparison relate to the comparison the researcher wishes to make?"

Example: Redistricting

Sekhon and Titiunik 2012 discussion of Ansolabehere et al. 2000

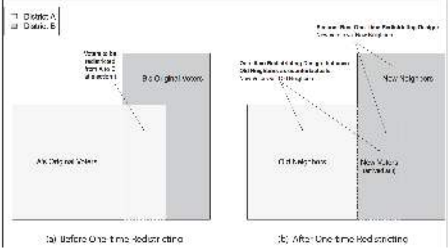
- problem: difficult to estimate incumbency advantage.
- proposed solution: redistricting as a natural experiment
- two types of variation: temporal (voting before and after) and cross-sectional (some voters move)
- key claim: if district is randomly redrawn, we can attribute voting differences between new voters in a district and their new neighbors to a lack of incumbency advantage
- idea is that two groups have same incumbent, same challenger, same campaign environment, but different histories with incumbent



(a) Before One-time Redistricting

Example: Redistricting

FIGURE 1. Illustration of Redistricting Research Designs



1) Is the comparison of B's old voters and B's new voters guaranteed to be valid if we assume that voters are redistricted randomly?
   - No! only voters in A were subject to randomization
   - Natural experiment creates three distinct groups: new voters, new neighbors, old neighbors and not all comparisons are valid.
2) What comparison is guaranteed to be valid if redistricting is done at random?
   - random redistricting guarantees that old neighbors and new voters are comparable.
   - need to find a new design (see Sekhon and Titiunik 2012 for more)

Pitfalls

- We still really need theory to guide our thinking
- Understanding the assignment and causal process is extremely important (was this really random? is the treatment really what we care about?)
- The result only applies to a limited population — is it significant?
- Be sure to verify that your "as-if-random" assignment is really random (e.g. placebo tests, balance tests)
- If possible, use sensitivity tests to evaluate susceptibility to unobserved confounding (e.g. Cinelli and Hazlett 2020)
- Convincingly analyzing a natural experiment takes at least as much careful thought not less!

Reasons to Be Excited

- Now that you know what to look for you may see more natural experiments out there
- Exogenous randomization can help us make credible causal inferences in places where we never could have run an experiment
- It is often pretty easy to communicate these kinds of methods to non-experts
- Salganik (2018) argues that with always-on digital data collection we will be in better shape moving forward to leverage natural experiments as the opportunities arise.

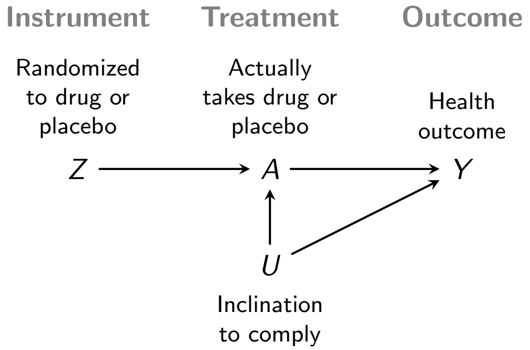
1 Natural Experiments

2 Instrumental Variables

3 Regression Discontinuity

4 Interrupted Time-Series

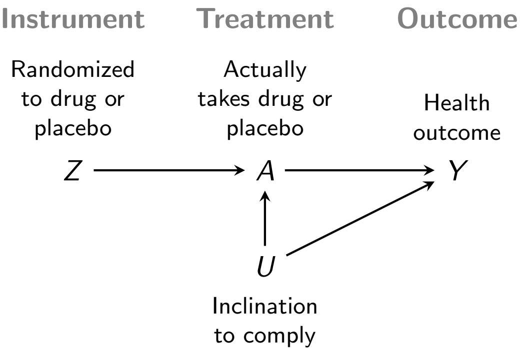
Instrumental variables: Experiment with noncompliance



Two ideas

1. Intent to treat effect
2. Average effect among compliers

Instrumental variables: 1) Intent to treat effect



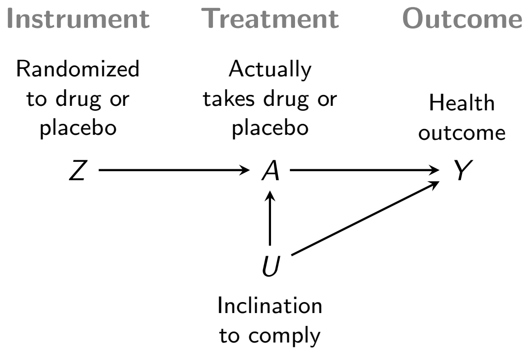
Ignore A. What is the effect of Z on Y?

$$\underbrace{E(Y^1 - Y^0)}_{\substack{\text{Outcome under } Z=1 \quad \text{Outcome under } Z=0}} \underset{\substack{\text{By Positivity} \\ \text{Consistency} \\ \text{Exchangeability for } Z}}{=} \underbrace{E(Y \mid Z = 1) - E(Y \mid Z = 0)}_{\text{Mean difference in observable outcomes}}$$

Instrumental variables: 2) Average effect among compliers

Imbens, G., & Angrist, J. (1994). Identification and estimation of local average treatment effects. Econometrica, 62(2), 467-475.

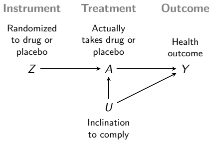
Instrumental variables: 2) Average effect among compliers



Key insight: The effect of Z on Y operates entirely through A

1. Study the effect of $Z \rightarrow Y$ (we just did)
2. Study the effect of $Z \rightarrow A$
3. Learn about $A \rightarrow Y$ since $Z \rightarrow Y$ is $Z \rightarrow A \rightarrow Y$

Instrumental variables: 2) Average effect among compliers

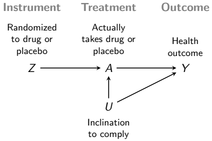


The effect $Z \to A$ has four **principal strata**:
latent sets of people who respond to Z a particular way

Compliers	$A^{Z=0} = 0$	$A^{Z=1} = 1$	(follow assignment)
Always takers	$A^{Z=0} = 1$	$A^{Z=1} = 1$	(always take treatment)
Never takers	$A^{Z=0} = 0$	$A^{Z=1} = 0$	(never take treatment)
Defiers	$A^{Z=0} = 1$	$A^{Z=1} = 0$	(defy assignment)

Discuss: In which strata is the effect $Z \to Y$ zero?

Instrumental variables: 2) Average effect among compliers



Four principal strata

Compliers	$(Z \rightarrow A) = +1$	$(Z \rightarrow Y) = (A \rightarrow Y)$
Always takers	$(Z \rightarrow A) = 0$	$(Z \rightarrow Y) = 0$
Never takers	$(Z \rightarrow A) = 0$	$(Z \rightarrow Y) = 0$
Defiers	$(Z \rightarrow A) = -1$	$(Z \rightarrow Y) = -(A \rightarrow Y)$

Assume **no defiers** in the population

Hypothetical:

Population is 50% compliers, 25% always takers, 25% never takers

Average effect of $Z \rightarrow Y$ among compliers is 0.6

What is the average effect of $Z \rightarrow Y$ in the population? **0.3**

Instrumental variables: 2) Average effect among compliers

Deriving the general case:

$$
\begin{aligned}
&E(Y^{Z=1} - Y^{Z=0}) \\
&= \sum_s E(Y^{Z=1} - Y^{Z=0} \mid S = s) \underbrace{\mathbb{P}(S = s)}_{\substack{\text{Denote} \\ \pi_s}} \\
&= E(Y^{Z=1} - Y^{Z=0} \mid S = \text{Complier})\pi_{\text{Complier}} \\
&\quad + E(Y^{Z=1} - Y^{Z=0} \mid S = \text{Always-Taker})\pi_{\text{Always-Taker}} \qquad (= 0) \\
&\quad + E(Y^{Z=1} - Y^{Z=0} \mid S = \text{Never-Taker})\pi_{\text{Never-Taker}} \qquad (= 0) \\
&\quad + E(Y^{Z=1} - Y^{Z=0} \mid S = \text{Defier})\pi_{\text{Defier}} \qquad (= 0)
\end{aligned}
$$

Instrumental variables: 2) Average effect among compliers

Average effect of instrument	=	Average among compliers	×	Proportion who comply

$$E(Y^{Z=1} - Y^{Z=0}) = E(Y^{Z=1} - Y^{Z=0} \mid S = \text{Complier})\pi_{\text{Complier}}$$

Among compliers, $(Z \to Y) = (A \to Y)$.

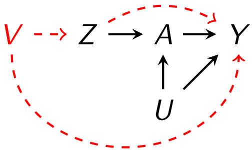
$$E(Y^{Z=1} - Y^{Z=0}) = E(Y^{A=1} - Y^{A=0} \mid S = \text{Complier})\pi_{\text{Complier}}$$

Rearrange to get the complier average treatment effect

$$\begin{aligned} E(Y^{A=1} - Y^{A=0} \mid S = \text{Complier}) &= \frac{E(Y^{Z=1} - Y^{Z=0})}{\pi_{\text{Complier}}} \\ &= \frac{E(Y^{Z=1} - Y^{Z=0})}{E(A^{Z=1} - A^{Z=0})} \end{aligned}$$

Three assumptions that make IV work

Exclusion Restriction:
Instrument must affect outcome only through treatment
(this edge cannot exist)



Exogeneity:
Instrument must be unconfounded
(this path cannot exist)

Monotonicity:
Instrument must affect treatment in the same direction for everyone

Given these extra assumptions,
the average effect among compliers
is identified by
$E(Y^{A=1} - Y^{A=0} \mid S = \text{Complier})$

$$= \frac{E(Y^{Z=1} - Y^{Z=0})}{E(A^{Z=1} - A^{Z=0})} \qquad \begin{array}{l} \leftarrow \text{Effect of } Z \text{ on } Y \\ \leftarrow \text{Effect of } Z \text{ on } A \end{array}$$

$$= \frac{E(Y|Z=1) - E(Y|Z=0)}{E(A|Z=1) - E(A|Z=0)}$$

Two equivalent estimators for the simple setting

We have been studying the **Wald estimator**

$$\frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{E(A \mid Z=1) - E(A \mid Z=0)} = \frac{\mathrm{Cov}(Z, Y)}{\mathrm{Cov}(Z, A)}$$

where the version at the right generalizes continuous variables in a linear system

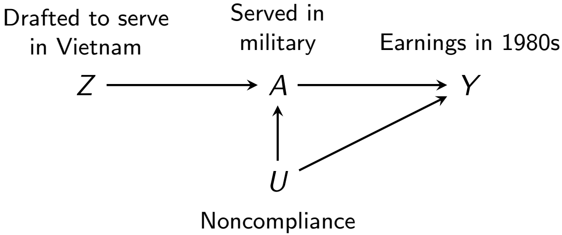
A numerically equivalent estimator is **two-stage-least-squares**:

1. Regress A on Z
2. Predict $\hat{A}$
   - Intuition: Isolates Z-generated variation in A
3. Regress Y on $\hat{A}$

Angrist (1990): Draft lottery as an instrument to study the relationship between military service and income

https://en.wikipedia.org/wiki/Draft_lottery_(1969)#/media/File:1969_draft_lottery_photo.jpg

Assumptions are less plausible in observational studies

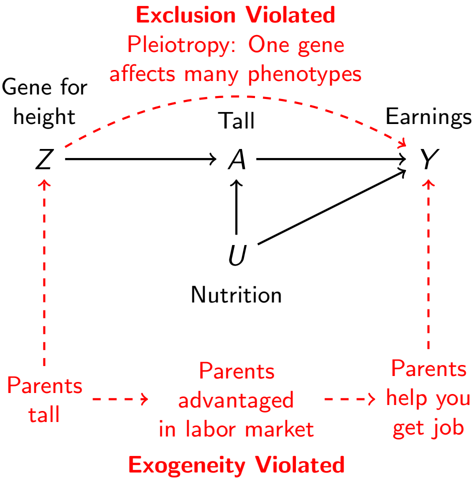


This is fairly credible

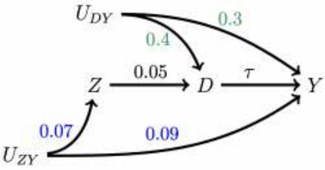
- Exogeneity: Randomly selected draft numbers
- Exclusion: Being drafted affects earnings only through service
- Monotonicity: No one joins the military in defiance of not being drafted

Angrist, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security administrative records. The American Economic Review, 313-336.

A difficult example: Genes as instruments

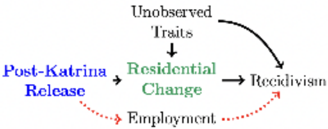


Warning: IV is very sensitive to assumptions



Asymptotic Bias of OLS
$= 0.3 \times 0.4 = \mathbf{0.12}$

Asymptotic Bias of IV
$= (0.07 \times 0.09)/0.05 = \mathbf{0.126}$

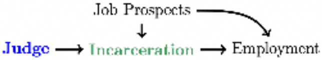
Study	Causal Structure	Identification Assumptions
Kirk (2009)		(1) Post-Katrina release induces moves away from a parolee's home neighborhood. (2) Post-Katrina release affects recidivism *only through* residential change and not, e.g., employment or police resources. (3) The timing of release shares no common causes with recidivism. (4) Post-Katrina release *never discourages* residential change.

Laidley and Conley (2018)

(1) More sunlight causes kids to exercise more. (2) Sunlight affects test scores *only through* exercise and not, e.g., mood. (3) When person fixed effects are included, within-person sunlight variation shares no common causes with test scores. (4) Sunnier weather *never discourages* exercise.

Sampson and Winter (2018)		(1) Proximity to a smelting plant increases lead exposure. (2) Proximity affects deliquency *only through* lead exposure. (3) Proximity shares no common causes—such as neighborhood disadvantage—with deliquency. (4) Moving closer to a smelting plant *never* causes someone to experience less lead exposure.

Harding et al. (2018)



(1) Judge assignment affects the probability of being incarcerated. (2) Judge assignment affects employment *only through* incarceration. (3) Judge assignment shares no common causes with future employment. (4) If assigned to a more lenient judge, a defendant will *never* receive a harsher sentence.

Effects for Compliers

- In all these cases we are estimating the effect for compliers, sometimes called the local average treatment effect (LATE) — or as we have here: the average effect among compliers.
- That is, the LATE theorem (Angrist, Imbens and Rubin 1996) states::

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[A_i|Z_i = 1] - E[A_i|Z_i = 0]} = E[Y_i(1) - Y_i(0)|A_i(1) > A_i(0)]$$

- Compliers are an unknown subset of the data.
  - Treated units are a mix of always takers and compliers.
  - Control units are a mix of never takers and compliers.
- Without further assumptions, Local Average Treatment Effect (LATE) and Average Treatment Effect (ATE) are not equal.
- Complier group depends on the instrument $\rightsquigarrow$ different IVs will lead to different estimands!

  How much we care about LATE turns on what the instrument is!

Who are the Compliers?

Study	Outcome	Treatment	Instrument
Angrist and Evans (1998)	Earnings	More than 2 Children	Multiple Second Birth (Twins)
Angrist and Evans (1998)	Earnings	More than 2 Children	First Two Children are Same Sex
Levitt (1997)	Crime Rates	Number of Policemen	Mayoral Elections
Angrist and Krueger (1991)	Earnings	Years of Schooling	Quarter of Birth
Angrist (1990)	Earnings	Veteran Status	Vietnam Draft Lottery
Miguel, Satyanath and Sergenti (2004)	Civil War Onset	GDP per capita	Lagged Rainfall
Acemoglu, Johnson and Robinson (2001)	Economic performance	Current Institutions	Settler Mortality in Colonial Times
Cleary and Barro (2006)	Religiosity	GDP per capita	Distance from Equator

Randomized Trials with One-Sided Noncompliance

- Will the LATE ever be equal to a usual causal quantity?
- When non-compliance is one-sided, then the LATE is equal to the ATT.
- Think of a randomized experiment:
  - Randomized treatment assignment = instrument (Z_i)
  - Non-randomized actual treatment taken = treatment (A_i)
- One-sided noncompliance: only those assigned to treatment (control) can actually take the treatment (control). Or

$$A_i(0) = 0 \forall i \quad \rightsquigarrow \quad P(A_i = 1 | Z_i = 0) = 0$$

- Maybe this is because only those treated actually get pills or only they are invited to the job training location. But this can be very difficult in many settings.
- See also, Imbens 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)" `http://dx.doi.org/10.1257/jel.48.2.399`

Appendix: Proof

One-sided noncompliance $\rightsquigarrow$ no "always-takers" and since there are no defiers,

- Treated units must be compliers.
- ATT is the same as the LATE.

Proof.

$$
\begin{aligned}
E[Y_i|Z_i=1]-E[Y_i|Z_i=0] =& E[Y_i(0)+(Y_i(1)-Y_i(0))T_i|Z_i=1]-E[Y_i(0)|Z_i=0] \\
& \text{(exclusion restriction + one-sided noncompliance)} \\
=& E[Y_i(0)|Z_i=1]+E[(Y_i(1)-Y_i(0))T_i|Z_i=1]-E[Y_i(0)|Z_i=0] \\
=& E[Y_i(0)]+E[(Y_i(1)-Y_i(0))T_i|Z_i=1]-E[Y_i(0)] \\
& \text{(randomization)} \\
=& E[Y_i(1)-Y_i(0)|T_i=1,Z_i=1]P(T_i=1|Z_i=1) \\
& \text{(law of iterated expectations + binary treatment)} \\
=& E[Y_i(1)-Y_i(0)|T_i=1]P(T_i=1|Z_i=1) \\
& \text{(one-sided noncompliance)}
\end{aligned}
$$

Noting that $P(T_i=1|Z_i=0)=0$, then the Wald estimator is just the ATT: $\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{P(T_i=1|Z_i=1)} = E[Y_i(1)-Y_i(0)|T_i=1]$ Thus, under the additional assumption of one-sided compliance, we can estimate the ATT using the usual IV approach □

Size of Complier Group

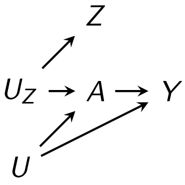
TABLE 4.4.2
Probabilities of compliance in instrumental variables studies

Source (1)	Endogenous Variable (D) (2)	Instrument (Z) (3)	Sample (4)	$P[D=1]$ (5)	First Stage, $P[D_1 > D_0]$ (6)	$P[Z=1]$ (7)	Compliance Probabilities: $P[D_1 > D_0 \mid D=1]$ (8)	Compliance Probabilities: $P[D_1 > D_0 \mid D=0]$ (9)
Angrist (1990)	Veteran status	Draft eligibility	White men born in 1950	.267	.159	.534	.318	.101
			Non-white men born in 1950	.163	.060	.534	.197	.033
Angrist and Evans (1998)	More than two children	Twins at second birth	Married women aged 21–35 with two or more children in 1980	.381	.603	.008	.013	.966
		First two children are same sex		.381	.060	.506	.080	.048
Angrist and Krueger (1991)	High school graduate	Third- or fourth-quarter birth	Men born between 1930 and 1939	.770	.016	.509	.011	.034
Acemoglu and Angrist (2000)	High school graduate	State requires 11 or more years of school attendance	White men aged 40–49	.617	.037	.300	.018	.068

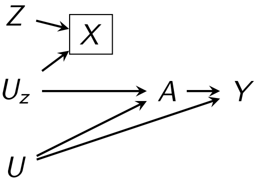
Notes: The table computes the absolute and relative size of the complier population for a number of instrumental variables. The first stage, reported in column 6, gives the absolute size of the complier group. Columns 8 and 9 show the size of the complier population relative to the treated and untreated populations.

Wait, What Were The Assumptions?

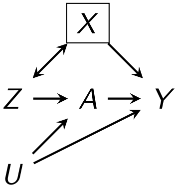
- Assumptions given have typically been sufficient but not necessary.
- Easy to state assumptions for linear structural equation models (where we can use covariance of error and variable), harder in general.
- Technically can use the following in place of assumptions 1-3: (see technical point 16.1 of Hernán and Robins)
  1) Exogeneity: $Y_i(a, z) \perp\!\!\!\perp Z_i$ for all a, z.
  2) Exclusion: $Y_i(a, z) = Y_i(a, z') = Y_i(t)$ for all z, z', a and i
  3) Relevance: $Z \not\perp\!\!\!\perp A$
  - all possibly conditional on X
- This allows some graphs that don't meet our original conditions, but satisfy new assumptions given particular edge configurations.



(proxy instrument)



(instrument via collider)



(exclusion via control)

Concluding Thoughts on Instrumental Variables

- Strong assumptions but powerful results
- Enormous care is required in the interpretation
- Questions to Always Ask
  1. is the instrument weak?
     (does Z predict A)
  2. is the instrument exogenous?
     (was Z randomly assigned?)
  3. does the exclusion restriction hold?
     (is there a path from Z to Y not through A)
  4. do we believe monotonicity?
     (are there units where the instrument discourages treatment)
  5. do the assumptions identify an effect for the subpopulation of interest?
     (is the instrument such that we care about compliers?)
- Be sure to evaluate all conditions and remember randomization of Z does not guarantee the exclusion restriction.

1 Natural Experiments

2 Instrumental Variables

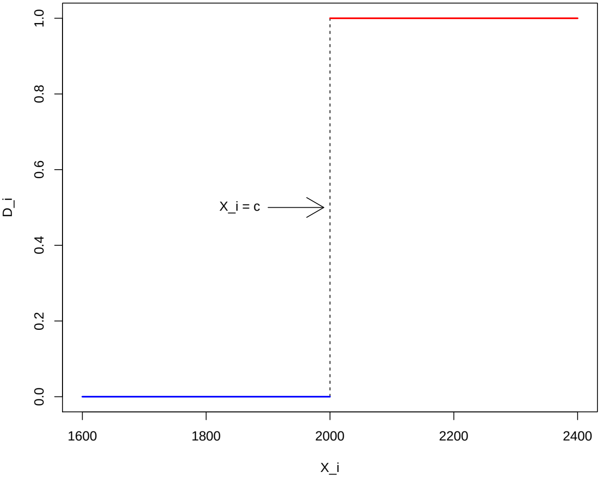
3 Regression Discontinuity

4 Interrupted Time-Series

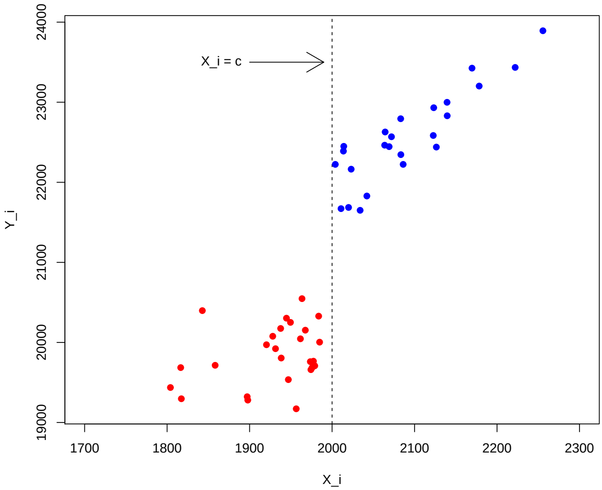
Regression Discontinuity

- A different strategy where the core intuition is that identification comes in a discontinuity in treatment assignment
- A widely applicable strategy in rule-based systems or allocations of limited resources (e.g. administrative programs, elections, admission systems)
- It is a fairly old idea, generally credited to education research by Thistlethwaite and Campbell 1960 but with a dynamic and interesting recent history (Hahn et al 2001 and Lee 2008 were big jumps forward).
- The goal here is to get you up to speed with the core idea: if you want to know how to do this in practice read *A Practical Introduction to Regression Discontinuity Designs* Volumes I and II by Matias Cattaneo, Nicolás Idrodo and Rocío Titiunik

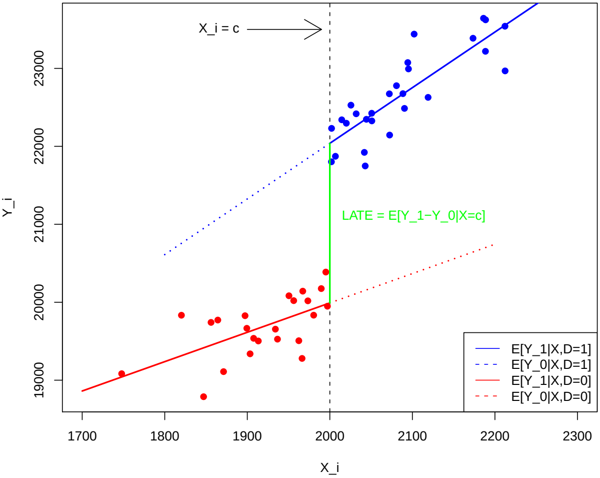
Graphical Illustration



Graphical Illustration



Graphical Illustration



Setup

- The basic idea behind RDDs:
  - X_i is a forcing variable.
  - Treatment assignment is determined by a cutoff in X_i.

$$T_i = 1\{X_i > c\} \text{ so } T_i = \begin{cases} T_i = 1 & \text{if } X_i > c \\ T_i = 0 & \text{if } X_i < c \end{cases}$$

- X_i can be related to the potential outcomes and so comparing treated and untreated units does not provide causal estimates
- assume relationship between X and the potential outcomes Y_1 and Y_0 is smooth around the threshold $\rightsquigarrow$ discontinuity created by the treatment to estimate the effect of T on Y at the threshold

Design

- Sharp RD: treatment assignment is a deterministic function of the forcing variable and the threshold.
- Key assumption: no compliance problems (deterministic)
- At the threshold, c, we only see treated units and below the threshold, we only see control values:

$$P(T_i = 1|X_i = c) = 1$$
$$P(T_i = 1|X_i < c) = 0$$

- Intuitively, we are interested in the discontinuity in the outcome at the discontinuity in the treatment assignment.
- We want to investigate the behavior of the outcome around the threshold: $\lim_{x \downarrow c} E[Y_i|X_i = x] - \lim_{x \uparrow c} E[Y_i|X_i = x]$
- Under certain assumptions, this quantity identifies the ATE at the threshold: $\tau_{SRD} = E[Y_i(1) - Y_i(0)|X_i = c]$

Identification

Identification Assumption

1. $Y(1), Y(0) \perp\!\!\!\perp T|X$ *(trivially met by construction)*
2. $0 < P(T = 1|X = x) < 1$ *(always violated in Sharp RDD)*
3. $E[Y(1)|X, T]$ *and* $E[Y(0)|X, T]$ *are continuous in* X *around the threshold* $X = c$ *(individuals have imprecise control over* X *around the threshold)*

Identification Result

The treatment effect is identified at the threshold as:

$$\begin{aligned}
\alpha_{SRDD} &= E[Y(1) - Y(0)|X = c] \\
&= E[Y(1)|X = c] - E[Y(0)|X = c] \\
&= \lim_{x \downarrow c} E[Y(1)|X = x] - \lim_{x \uparrow c} E[Y(0)|X = x]
\end{aligned}$$

Without further assumptions α_{SRDD} *is only identified at the threshold.*

Extrapolation and smoothness

- Remember the quantity of interest here is the effect at the threshold:

$$\begin{aligned}\tau_{SRD} &= E[Y_i(1) - Y_i(0)|X_i = c] \\ &= E[Y_i(1)|X_i = c] - E[Y_i(0)|X_i = c]\end{aligned}$$

- But we don't observe $E[Y_i(0)|X_i = c]$ ever due to the design, so we're going to extrapolate from $E[Y_i(0)|X_i = c - |\varepsilon|]$ for $\varepsilon \approx 0$.
- Extrapolation, even at short distances, requires smoothness in the functions we are extrapolating.

What can go wrong?

- If the potential outcomes change at the discontinuity for reasons other than the treatment, then smoothness will be violated.
- For instance, if people sort around threshold, then you might get jumps other than the one you care about.
- If things other than the treatment change at the threshold, then that might cause discontinuities in the potential outcomes.

General estimation strategy

- The main goal in RD is to estimate the limits of various CEFs such as:

$$\lim_{x \uparrow c} E[Y_i | X_i = x]$$

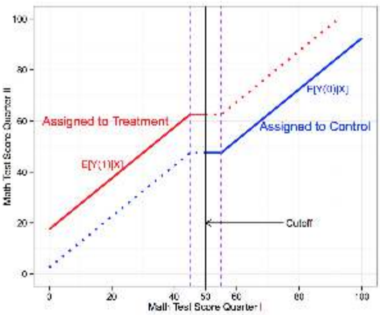
- It turns out that this is a hard problem because we want to estimate the regression at a single point and that point is a boundary point.
- As a result, the usual kinds of nonparametric estimators perform poorly (polynomials and kernels are particularly bad)
- In general, we are going to have to choose some way of estimating the regression functions around the cutpoint.
- Using the entire sample on either side will obviously lead to bias because those values that are far from the cutpoint are clearly different than those nearer to the cutpoint.
- $\rightarrow$ restrict our estimation to units close to the threshold.
- Local linear regression is a good way to go: see `rdrobust` package in `R` (Calonico et al 2015)

Misconceptions

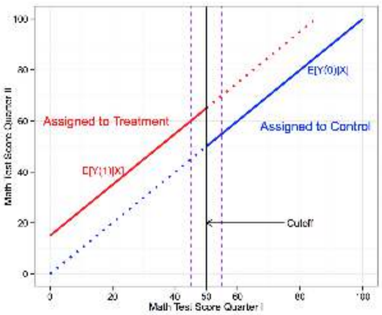
- Continuity of the potential outcomes does not imply local randomization
- This has caused a lot of confusion in the literature particularly in testing with background covariates
- Local statistical independence does not imply exclusion restriction (i.e. forcing variable not directly affecting the outcome)
- If you are doing an RDD: be sure to do balance checks and sensitivity checks (read-up on best practices first!)

Local Randomization vs. Continuity (Sekhon and Titiunik 2017)

Figure 1: Two Scenarios with Randomly Assigned Score



(a) Test scores locally unrelated to potential outcomes



(b) Test scores locally related to potential outcomes

Fuzzy RD

- With fuzzy RD, the treatment assignment is no longer a deterministic function of the forcing variable, but there is still a discontinuity in the probability of treatment at the threshold:

Assumption 1 of Fuzzy RD

$$\lim_{x \downarrow c} P[T_i = 1 | X_i = x] \neq \lim_{x \uparrow c} P[T_i = 1 | X_i = x]$$

- In the sharp RD, this is also true, but it further required the jump in probability to be from 0 to 1.
- Fuzzy RD is often useful when the threshold encourages participation in program, but does not actually force units to participate.
- Sound familiar? Fuzzy RD is just IV!

Fuzzy RD is IV

- Forcing variable is an instrument: affects Y_i, but only through T_i (at the threshold)
- Let $T_i(x)$ be the potential value of treatment when we set the forcing variable to x, for some small neighborhood around c.
- $T_i(x) = 1$ if unit i would take treatment when X_i was x
- $T_i(x) = 0$ if unit i would take control when X_i was x

Fuzzy RD assumptions

Assumption 2: Monotonicity

There exists ε such that $T_i(c+e) \geq T_i(c-e)$ for all $0 < e < \varepsilon$

No one is discouraged from taking the treatment by crossing the threshold.

Assumption 3: Local Exogeneity of Forcing Variable

In a neighborhood of c,

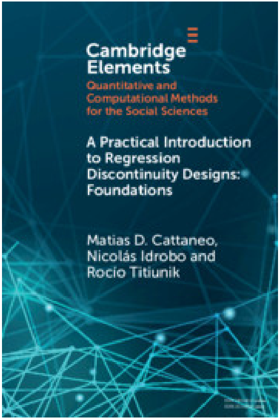
$$\{\tau_i, T_i(x)\} \perp\!\!\!\perp X_i$$

Basically, in an ε-ball around c, the forcing variable is randomly assigned.

Regression Discontinuity Conclusions

- Key idea is to exploit an arbitrary assignment rule to identify a causal quantity.
- Remember that we are only identifying an effect at the boundary.
- There are many other nuances to estimation and choosing an appropriate bandwidth for the comparison- be sure to read more before trying this at home.
- There is an interesting literature on geographic regression discontinuity designs as well. These are harder but can be useful!

What to read next?



- A Practical Introduction to Regression Discontinuity Designs, Volumes I and II by Matias Cattaneo, Nicolas Idrodo and Rocio Titunik
- Angrist and Pishke Chapter 6 Regression Discontinuity Designs

Interrupted Time Series

- A simple construction often used with natural experiment is the Interrupted Time Series (ITS)
- ITS designs convey the basic intuition that when an event abruptly occurs, we can compare results immediately before and immediately afterwards.
- We can write this as a model indexing the date with d:

$$Y_d = f(d) + T_d \beta + \epsilon_d$$

- The key identifying assumption is that the observed values of y_d before the treatment status switches at d^* can be used to specify $f(d)$ for the rest of the series used.

Interrupted Time Series Example

American Political Science Review (2018) 112, 3, 621–636
doi:10.1017/S0003055418000084

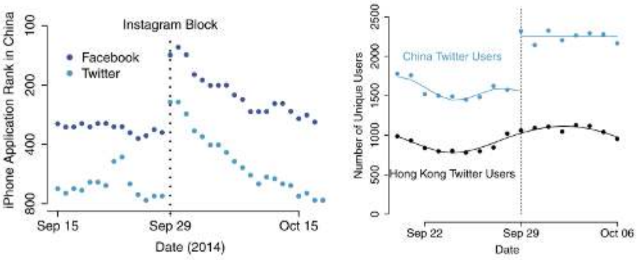
How Sudden Censorship Can Increase Access to Information

WILLIAM R. HOBBS *Northeastern University*
MARGARET E. ROBERTS *University of California, San Diego*

Conventional wisdom assumes that increased censorship will strictly decrease access to information. We delineate circumstances when increases in censorship expand access to information for a substantial subset of the population. When governments suddenly impose censorship on previously uncensored information, citizens accustomed to acquiring this information will be incentivized to learn methods of censorship evasion. These evasion tools provide continued access to the newly blocked information—and also extend users' ability to access information that has long been censored. We illustrate this phenomenon using millions of individual-level actions of social media users in China before and after the block of Instagram. We show that the block inspired millions of Chinese users to acquire virtual private networks, and that these users subsequently joined censored websites like Twitter and Facebook. Despite initially being apolitical, these new users began browsing blocked political pages on Wikipedia, following Chinese political activists on Twitter, and discussing highly politicized topics such as opposition protests in Hong Kong.

Interrupted Time Series Example

FIGURE 3. Left: The Instagram block's effect on the rank of Facebook and Twitter on iPhones from mainland China, from AppAnnie.com. Right: Comparison of tweets per day from Mainland China and Hong Kong before and after the Instagram block.



The left panel of this figure shows the change in download ranks for Facebook and Twitter before and after Instagram was blocked. The right panel of this figure shows that the Chinese Twitter users in our sample increased 30% the same day that we observe a spike in Instagram mentions and several days after the beginning of the Hong Kong protests. This increase only occurred in China and not in Hong Kong. The lines in this panel were fit using a smoothing spline.

This Week in Review

- This week we covered approaches to unmeasured confounding
- The trick is to exploit some other feature but there is No Free Lunch.
- Now that you have seen a few examples, hopefully you can be on the lookout for your own research.
- We talked about natural experiments, instrumental variables, regression discontinuity and interrupted time series.

Next Week: Difference in Differences, Fixed Effects, and Review