Lecture 6: Causality

Mitch Downey[1]

February 13, 2024

[1]IIES, Stockholm University.

Today

- "Traditional" teaching of causality
  - Very common in undergraduate courses
  - Many of you have seen
  - Very closely connected to the material we've already covered
  - Existing examples and problem sets show some limitations
- Problems with traditional "definition"
- An alternative approach
  - Rubin potential outcomes framework
  - Conceptually straightforward, builds upon conditional expectation functions we've already discussed
  - Simple to nest all the endogeneity concerns we're used to
  - Starting point of Econometrics II
- Informal formalization of mainstream identification strategies
- Link between conditional expectation functions and causal inference

Mainstream teaching of econometrics

- The true model is:
$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$
where everything is a scalar
- β_1 is true causal effect of x_1: Effect of increasing x_1 while holding all else equal
- Definition: x_1 is exogenous if and only if $E(x_1 \varepsilon) = 0$
  - Otherwise it is endogenous
  - Note: Error term (ε) is a never observable object that includes all determinants of y other than the columns of X
  - $\hat{\varepsilon}$ from OLS is still mechanically uncorrelated with X
- If x_1 is exogenous, then $\hat{\beta}_1^{OLS}$ is a consistent estimator for β_1
- Difference from our previous lecture: What is β?
  - Lecture 5: Population projection coefficient
    - Linear projection of Y onto space spanned by columns of X
    - Linear approximation of conditional mean function
    - Minimizes length of error vector
    - $E(X'\varepsilon) = 0$ is a result
  - Traditional model: Causal effect of X
    - $E(X'\varepsilon) = 0$ is an assumption necessary for $\hat{\beta}^{OLS}$ to be consistent

Endogeneity

- The true model is:
$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$
- β_1 is true causal effect of x_1: Effect of increasing x_1 while holding all else equal
- Definition: x_1 is exogenous if and only if $E(x_1 \varepsilon) = 0$
- If x_1 is exogenous, then $\hat{\beta}_1^{OLS}$ is consistent estimator for β_1
- Main two types of endogeneity:
- OVB:
  - If $\beta_j \neq 0$ so that x_j has a "true" effect on y, then excluding x_j from the regression means it's in the error term ε
  - If x_j is correlated with x_1, then x_1 is correlated with the error term
- Reverse causality: Exogenously increasing y causes x_1 to increase, so documented conditional correlation is causal effect of y, not of x_1

My two problems with the original model: Truth

- The **true model** is:

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$

- β_1 is true causal effect of x_1: Effect of increasing x_1 while holding all else equal

1. The true model surely isn't linear
   - Standard response: "Well sure, but this is a linear approximation"
   - So then what's the true model?
   - What are the properties of the approximation?
   - Is it possible that those properties don't depend on the true model?
2. β surely depends on why x is changing
   - Some things are misspecification
     - Non-linearities, heterogeneous effects, etc.
   - But what if it depends on *why* x is changing?
     - Example from problem set: Capital gains and labor market earnings have different effects on outcomes, but we don't think about that and pool them into "income"
     - Maybe using an erroneously aggregated variable is a type of misspecification
     - Then is "causal effect of income" not well defined?
     - There doesn't exist a true β we're trying to estimate?

My two problems with the original model: Causality

- The true model is:

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$

- β_1 is true causal effect of x_1: Effect of increasing x_1 while **holding all else equal**
- Surely you cannot mean "*all else*" equal?
  - Defining the causal effect of income requires holding investments, spending, and housing choices equal?
  - Standard response: No, those are mechanisms
  - Does that mean don't control for them at all? Surely it depends on what kind of causal effect I want to estimate.
- If spending affects outcomes, then either it's an x or it's in ε (in which case it's correlated with income)
- Intuitively: It's a bad control, don't control for it
- But within this framework, what does that actually mean? How can you formalize that? How can you represent that?
- Core problem: β is the causal effect, the causal effect is $\partial y / \partial x = \beta$
  - It's all very circular
  - We're never taking a stand and proposing a principled definition of causal effect

Rubin potential outcomes framework: Main idea

- (Sometimes called Rubin causal model)
- Define two states of the world
    - Example: You took this course or you didn't
- For each state of the world, there is a **potential outcome**, which is what *would have* happened under that state of the world
- Only one state of the world actually happened, so you (exactly) observe one and only one of the two potential outcomes
- As a definition, the **causal effect** of an event is the difference between the potential outcome in which that event happened and the potential outcome in which it didn't
    - Very intuitive, broadly accessible, normal-person definition of causal effect
- In practice, under what conditions can we estimate the average causal effect?

Rubin potential outcomes framework: Formally

- Let Y be the outcome of interest
- Let S be the state of the world
  - $S = 1$: You took this course, $S = 0$: You didn't
- The potential outcomes of individual i are given by the function $Y_i(\cdot)$:
  - $Y_i(S = s_1)$ is outcome of individual i if the state of the world is s_1
  - $Y_i(S = s_2)$ is outcome of individual i if the state of the world is s_2
  - Note: Generally, you cannot observe both of these
- $\Delta_i \equiv Y_i(s_1) - Y_i(s_2)$: The causal effect for individual
  - Note 1: If you cannot observe both $Y_i(s_1)$ and $Y_i(s_2)$ then you cannot observe this
  - Note 2: The definition of a treatment effect depends on both s_1 and s_2 (as is realistic: The effect of taking this course *instead of what*?)
- Main goal of econometrics: Estimate $E(\Delta_i \mid i \in C)$
  - Average treatment effect/average causal effect
  - Generally, we estimate this only for some subset C of the full population
  - Examples of C when S is participation in a program respondents are assigned to:
    - People who participated in the program: Treatment on the Treated (TOT)
    - People assigned to participate in the program: Intent to Treat (ITT)
    - People who would participate when assigned to do so, but would not have participated when not assigned to: Compliers

Conceptual advantages of the potential outcomes framework: Non-linearity

- Traditional framework:
  - The true model is:
    $$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$
  - β_1 is true causal effect of x_1: Effect of increasing x_1 while holding all else equal
- Potential outcomes framework (POF)
  - Causal effect of $x_1 = a$ instead of b: $Y_i(x_1 = a) - Y_i(x_1 = b)$
- POF makes no assumptions about the conditional expectation function's form
- They aren't necessary to *define* causal effects
  - Need to be more specific about relative to what
- They might be necessary to *estimate* it

Conceptual advantages of the potential outcomes framework: Why is x changing?

- Traditional framework:
  - The true model is:
    $$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$
  - β_1 is true causal effect of x_1: Effect of increasing x_1 while holding all else equal
- Potential outcomes framework (POF)
  - Causal effect of wage increase:
    $Y_i(income = income_0 + wageincrease) - Y_i(income = income_0)$
  - Causal effect of capital gains windfall:
    $Y_i(income = income_0 + highcapgains) - Y_i(income = income_0)$
  - Causal effect of winning the lottery:
    $Y_i(income = income_0 + lottery) - Y_i(income = income_0)$
- If you cannot express your thought experiment in the potential outcomes framework, you probably won't be able to estimate it
  - The effect of higher income for no reason

Conceptual advantages of the potential outcomes framework: What else equal?

- Traditional framework:
  - The true model is:
    $$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon$$
  - β_1 is true causal effect of x_1: Effect of increasing x_1 while holding all else equal
- Potential outcomes framework (POF)
  - Causal effect of unanticipated capital gains that agents didn't respond to: $Y_i\big(inc = inc_0 + windfall, spend = spend(inc_0)\big) - Y_i\big(inc = inc_0, spend = spend(inc_0)\big)$
  - Causal effect of capital gains windfall that agents could fully respond to: $Y_i\big(inc = inc_0 + windfall\big) - Y_i\big(inc = inc_0\big)$
  - Causal effect of *lower* than expected inc. because anticipated windfall didn't happen:

$$Y_i\big(inc = inc_0 + windfall, spend = spend(inc_0 + windfall)\big)$$
$$- Y_i\big(inc = inc_0, spend = spend(inc_0 + windfall)\big)$$

- Note: Derivatives are symmetric (increase: β, decrease: $-\beta$), so it's hard for the traditional framework to express something like this

Simplest case: Experimental variation in treatment

- **Estimand**: The object that we are trying to estimate
- The average treatment effect:

$$E\Big[Y_i(T=1) - Y_i(T=0)\Big]$$

- Conceptual comments:
  1. Allows for arbitrary heterogeneity because treatment effects are subscripted by i
  2. Easy to express restrictions on what actions are allowed.
     - Let B be some behavior, and suppose your experiment doesn't allow agents to change B
     - We can still estimate $E\big(Y_i(T=1, B=b_0) - Y_i(T=0, B=b_0)\big)$, which will be different than what would happen if agents could change B
  3. Causal inference is sometimes described as being only about "internal validity," but the POF creates an obvious bridge to external validity
     - If the POF expression doesn't look anything like the program you want to create, then the causal effect it estimates isn't relevant for you

Simplest case: Experimental variation in treatment

- The average treatment effect:

$$E\big[Y_i(T=1) - Y_i(T=0)\big]$$

- Practical comments: Expectations is a linear operator

$$E\big[Y_i(T=1) - Y_i(T=0)\big] = E\big[Y_i(T=1)\big] - E\big[Y_i(T=0)\big]$$

- If the program is randomly assigned, then treatment status is independent of potential outcomes:[2]

$$E\big[Y_i(T=1)\big] = E\big[Y_i(T=1)|T_i=1\big]$$
$$E\big[Y_i(T=0)\big] = E\big[Y_i(T=0)|T_i=0\big]$$

- We can estimate this with a basic plug-in estimator:

$$\left[\frac{1}{n_T}\sum_{i\in\mathcal{T}} y_i\right] - \left[\frac{1}{n_C}\sum_{i\in\mathcal{C}} y_i\right] \equiv \bar{y}_T - \bar{y}_C$$

$\mathcal{T}, \mathcal{C}$ are sets of treatment and control participants, resp., with $|\mathcal{T}| = n_T$, $|\mathcal{C}| = n_C$

[2]In general, we only need that treatment status is uncorrelated with potential outcomes, but random assignment implies independence, which is stronger

Omitted variable bias

- Suppose treatment status is correlated with some variable W:

$$E(W_i|T_i = 1) = w_1$$
$$E(W_i|T_i = 0) = w_0 \neq w_1$$

- Then what is $\bar{y}_T - \bar{y}_C$ estimating?

$$\begin{aligned}
&E\left[Y_i(T=1)|T_i=1\right] - E\left[Y_i(T=0)|T_i=0\right] \\
\equiv &E\left[Y_i(1)|T_i=1\right] - E\left[Y_i(0)|T_i=0\right] \\
= &E\left[Y_i(1)|W_i=w_1\right] - E\left[Y_i(0)|W_i=w_0\right] \\
= &E\left[Y_i(1)|W_i=w_1\right] - E\left[Y_i(1)|W_i=w_0\right] \\
&+ E\left[Y_i(1)|W_i=w_0\right] - E\left[Y_i(0)|W_i=w_0\right]
\end{aligned}$$

- First line: Causal effect of shifting W from w_0 to w_1 on expected outcomes conditional on participating in treatment
- Second line: Causal effect of treatment given people with $W = w_0$ (i.e., among the control group)
- Sum: treatment effect + effect of omitted variable (just like traditional framewrk)
  - Note that if treatment has low correlation with W then $|w_1 - w_0|$ is very small

Selection bias: Selection on gains

- Selection on gains:[3] People who would benefit more from the program are more likely to participate in it.
- Suppose the causal effect of the program on individual i is either Δ_H or $\Delta_L < \Delta_H$
- Suppose there is selection into treatment ($T_i = 1$) on gains:

$$Pr(T_i = 1|\Delta_i = \Delta_H) \equiv p_H$$
$$Pr(T_i = 1|\Delta_i = \Delta_L) \equiv p_L$$
$$p_H > p_L$$

- Simplify math: Assume Δ_H, Δ_L are equally common, half the sample participates
  - $Pr(\Delta_i = \Delta_H) = Pr(T_i = 1) = 1/2 \Rightarrow (1/2)p_H + (1/2)p_L = 1/2 \Rightarrow p_L = 1 - p_H$
- Then what does $\bar{y}_T - \bar{y}_C$ estimate?

[3] This has long been a focal case in the structural labor literature on returns to education. See Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. "Estimating marginal returns to education." *American Economic Review* 101.6 (2011): 2754-2781.

Selection bias: Selection on gains

$$
\begin{aligned}
E\big[Y_i(1)|T_i=1\big] - E\big[Y_i(0)|T_i=0\big] =& E\big[Y_i(1)|\Delta_H\big]p_H + E\big[Y_i(1)|\Delta_L\big]p_L \\
& - E\big[Y_i(0)|\Delta_H\big](1-p_H) - E\big[Y_i(0)|\Delta_L\big](1-p_L) \\
=& \big(E\big[Y_i(0)|\Delta_H\big] + \Delta_H\big)p_H + \big(E\big[Y_i(0)|\Delta_L\big] + \Delta_L\big)p_L \\
& - E\big[Y_i(0)|\Delta_H\big](1-p_H) - E\big[Y_i(0)|\Delta_L\big](1-p_L) \\
=& p_H\Delta_H + p_L\Delta_L \\
& + \big(p_H - (1-p_H)\big)E\big[Y_i(0)|\Delta_H\big] \\
& + \big(p_L - (1-p_L)\big)E\big[Y_i(0)|\Delta_L\big] \\
=& p_H\Delta_H + p_L\Delta_L \\
& + (p_H - p_L)\big(E\big[Y_i(0)|\Delta_H\big] - E\big[Y_i(0)|\Delta_L\big]\big)
\end{aligned}
$$

Selection bias: Selection on gains

$$E\big[Y_i(1)|T_i = 1\big] - E\big[Y_i(0)|T_i = 0\big] = p_H\Delta_H + p_L\Delta_L + (p_H - p_L)\big(E\big[Y_i(0)|\Delta_H\big] - E\big[Y_i(0)|\Delta_L\big]\big)$$

- $p_H\Delta_H + p_L\Delta_L$: Average treatment effect
  - **Not** population average
  - Average treatment on the treated
- $p_H - p_L$: Selection

  If high and low types are equally likely to participate (i.e., no selection), this term is zero and the difference between group outcomes is just an average treatment effect
- $E\big[Y_i(0)|\Delta_H\big] - E\big[Y_i(0)|\Delta_L\big]$: Bias
  - Note that if this is zero, then there is no bias
  - In that case, the outcomes of high types and low types would have been the same in the absence of treatment, so it doesn't matter that there's selection ($p_H - p_L \neq 0$)
  - Otherwise, the difference in what would have happened *anyway* causes a bias
  - Might under- or over-estimate treatment effects, depending on the sign of this term
  - Example: We say we have *positive selection* if $E\big[Y_i(0)|T_i = 1\big] > E\big[Y_i(0)|T_i = 0\big]$

Summary so far

- Because regressions are conditional means, all of the above insights apply to regressions instead of two-sample comparisons of means
  - They generalize to continuous "treatments" too, but the notation and the math is somewhat cumbersome
- Defining what you *want* to estimate is straightforward with the potential outcomes framework
- Formalizing various biases is generally also pretty easy
  - This formalization can be done without specifying functional forms
  - Decomposing terms is often informative and intuitive, and clarifies which assumptions would be sufficient
- Clearly does away with the notion of "***the***" causal effect
  - Always keeping track of heterogeneity
  - Also clear mapping to talking about policy questions
- To me, these are a big deal and the potential outcomes framework helps me think and communicate
- It is ok to be underwhelmed and continue to use the traditional framework if that makes sense to you
- But it is important to know what this is so you can read others' work (applied stuff and econometrics papers)

Causal inference in economics: My classification

1. True experiments
   - Researcher designed and implemented experiments are popular in development economics, but increasingly common in labor and public and political
   - Experiments also happen in the real world
     - People winning the lottery: Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace. "Wealth, health, and child development: Evidence from administrative data on Swedish lottery players." *The Quarterly Journal of Economics* 131, no. 2 (2016): 687-738.
     - Firms winning a procurement lottery: Carrillo, Paul, Dave Donaldson, Dina Pomeranz, and Monica Singhal. Misallocation in Firm Production: A Nonparametric Analysis Using Procurement Lotteries. NBER Working Paper 31311, 2023.
     - Kids winning access to an oversubscribed charter school: Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt. "The effect of school choice on participants: Evidence from randomized lotteries." *Econometrica* 74, no. 5 (2006): 1191-1230.
   - In general, real-world experiments and researcher-designed experiments are both only *conditionally* exogenenous
     - Lottery winnings are only exogenous conditional on lottery ticket purchase behavior
     - Implication 1: All our lessons about aggregation still apply (e.g., charter school lotteries)
     - Implication 2: Despite x being conditionally exogenous, an OLS regression will not, in general, recover the average treatment effect because OLS is variance-weighted

Causal inference in economics: My classification

1. True experiments
2. Assumed exogeneity (note: this stuff won't be on the exam)
    - In OLS, you can assume X is exogeneous or conditionally exogenous
        - Very unpopular for a long time
        - Somewhat more common: I only care about how something changes over time, and I assume that the bias is constant over time
    - In IV, the *exclusion restriction* states that the instrument only enters the potential outcomes expressions via the endogenous variable you want to study
        - Very unpopular now days
        - People often still believe the instrument is uncorrelated with potential outcomes (exogenous)
        - This allows you to estimate the causal effects of the instrument (i.e., the reduced form), but not the causal effect of the endogenous variable (people reject the IV interpretation)
    - Famous example: Acemoglu, Johnson, Robinson (2001)[4]
        - Look at African countries that were colonized
        - Instrument for the type of legal frameworks established during colonization using the mortality rate among settlers at the beginning of colonization
        - **Goal**: Estimate causal effects of one type of legal framework (rather than another) on outcomes today

[4] Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The colonial origins of comparative development: An empirical investigation." *American Economic Review* 91.5 (2001): 1369-1401.

Understanding IV assumptions: Acemoglu, Johnson, Robinson (2001)

- Three types of criticisms of the settler mortality instrument:
  1. Data issues
  2. Settler mortality is endogenous
     - Correlated with potential outcomes: High mortality places would be poorer regardless
     - Jeffrey Sachs: Mortality 200 years ago was driven by diseases which differ across geography and still affect outcomes today (e.g., malaria)
  3. The exclusion restriction is violated: Settler mortality has a causal effect on modern outcomes through multiple channels, not only the legal framework
     - Examples: trade links with Europe, investments in infrastructure, conflict with colonizers
- We spend our careers on criticism 2: It is often solvable, though that's difficult
- Criticism 3 is harder to solve because important things affect multiple channels
- It is also completely separate from criticism 2 and identification
  - You can't instrument for spending using lottery winnings, even though it's exogenous
  - Lottery winnings also affect savings and wealth, hours worked and labor market earnings, residential choice, financial stress, etc.
  - Each would violate the exclusion restriction required by one of the other variables
- Very rare now days to actually defend an exclusion restriction
- Instead, people define the estimand of interest to be the effects of the exogenous variation (i.e., the reduced form)
  - This blends "internal" and "external" validity: They are not distinct

Understanding IV assumptions: Bartik shocks

- Special case of more general "shift-share" identification strategy
- All of this will be covered in Econometrics II
- My goal: Relate this common IV to our argument from last time: The advantage of IV is it lets you restrict to one specific source of variation
  - Thus, this will be a very brief introduction

Understanding IV assumptions: Bartik shocks

- Suppose you are interested in the effects of city-level wage growth on some city-level outcome $Y_{c,t}$ (t: time)
- You could estimate: $\Delta Y_{c,t} = \beta_0 + \beta_1 \Delta w_{c,t} + \varepsilon_{c,t}$ where $\Delta Z_{c,t} \equiv Z_{c,t} - Z_{c,t-1}$ and $w_{c,t}$ is average wages, averaged across all workers in the city
- But city-level wage growth comes from many sources and is endogenous
  - Note that those are two different statements
- A Bartik IV strategy's first stage is given by:

$$\Delta w_{c,t} = \gamma_0 + \gamma_1 Z_{c,t} + \nu_{c,t}$$
$$\text{where } Z_{c,t} \equiv \sum_i s_{i,c,t-1} \Delta w_{i,-c,t}$$

  where $s_{i,c,t}$ is the share of all workers in city c at time t who work in industry i, and $\Delta w_{i,-c,t}$ is the change in average wages seen in industry i from $t-1$ to t across all cities *except* city c (a leave-one-out estimate)
- Where does variation come from? Cities that had, at $t-1$, types of industries that saw the most aggregate (nationwide) wage growth from $t-1$ to t
- Is this causal? Maybe. Under some assumptions, yes. Which ones? See Econometrics II or references in Lecture 1.
- More important (my opinion): Do you understand what this variation is? Yes.

Causal inference in economics: My classification

1. True experiments
2. Assumed exogeneity
3. Imputing the counterfactual
   - Main example: Difference-in-difference
     - Example: New Jersey (NJ) increased minimum wage in 1994, Pennsylvania (PA) did not
     - Define $T_s = 1$ if state s increases its minimum wage in 1994
     - $Y_{NJ,94}(T_s = 1)$ is observable
     - Impute $Y_{NJ,94}(T_s = 0)$ as $\left[Y_{NJ,93}(1) - Y_{PA,93}(0)\right] + \left[Y_{PA,94}(0) - Y_{PA,93}(0)\right]$
     - Now generalized to multiple treatment events, multiple comparison "states," longer time periods and dynamic treatment effects, anticipation effects $\left(Y_{NJ,93}(1) \neq Y_{NJ,93}(0)\right)$, continuous treatments, etc.
   - Plenty of criticism, but still widely used and accepted within economics
   - Major focus of Econometrics II
   - More straightforward example: Matching (less common in economics, but econometric theorists typically really like it)

Causal inference in economics: My classification

1. True experiments
2. Assumed exogeneity
3. Imputing the counterfactual
4. Control functions
    - Generally considered most convincing non-experimental approach
    - Key idea:
        - Assume that potential outcomes are smooth functions
        - Find instances where treatment assignment is not smooth

Causal inference in economics: My classification

1. True experiments
2. Assumed exogeneity
3. Imputing the counterfactual
4. Control functions
    - Simple example: Regression Discontinuity
        - If you get 2250+ points on the Missouri high school equivalence exam, you get a degree
        - Assume that the potential outcomes are continuous across that threshold ($S = 2250$)
        - With that, we can use any consistent estimator for the conditional expectation function, combined with our assumption of continuity (which ensures that we can approximate the either potential outcome at the discontinuity by approaching it only from one side):

$$\text{estimand: } E(\Delta_i|S_i = 2250) \equiv E(Y_i(\text{Pass=1})|S_i = 2250) - E(Y_i(\text{Pass=0})|S_i = 2250)$$

$$E(Y_i(0)|S_i = 2250) \overset{bc}{=} \lim_{x \to^- 2250} \hat{E}(Y_i|S_i = x)$$

$$E(Y_i(1)|S_i = 2250) \overset{bc}{=} \lim_{x \to^+ 2250} \hat{E}(Y_i|S_i = x)$$

where $\overset{bc}{=}$ denotes "by the continuity assumption: both potential outcomes are continuous across the threshold", $\hat{E}(\cdot)$ is a consistent estimate of a CEF, and $\to^-$ and $\to^+$ mean "the limit approaching from below" and "above," respectively

    - Is this *a* causal effect? Yes. Is it *the* causal effect? No.
    - Local average treatment effect, conditional on being asymptotically close to $S = 2250$

Regression discontinuity: Jepsen, Christopher, Peter Mueser, and Kenneth Troske. "Labor market returns to the GED using regression discontinuity analysis." *Journal of Political Economy* 124, no. 3 (2016): 621-649.

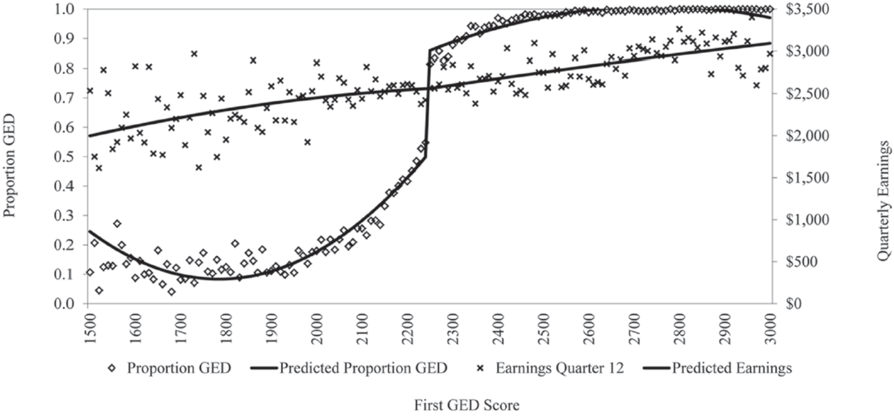
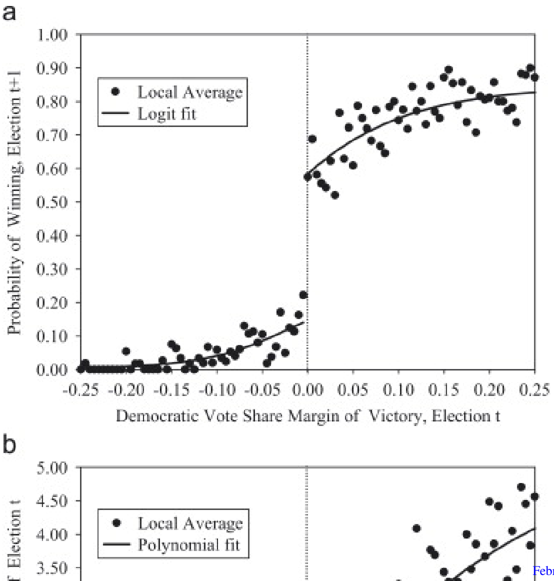


FIG. 4.—Regression discontinuity models predicting GED and quarterly earnings, men

Regression discontinuity: Lee, David S. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics* 142, no. 2 (2008): 675-697.



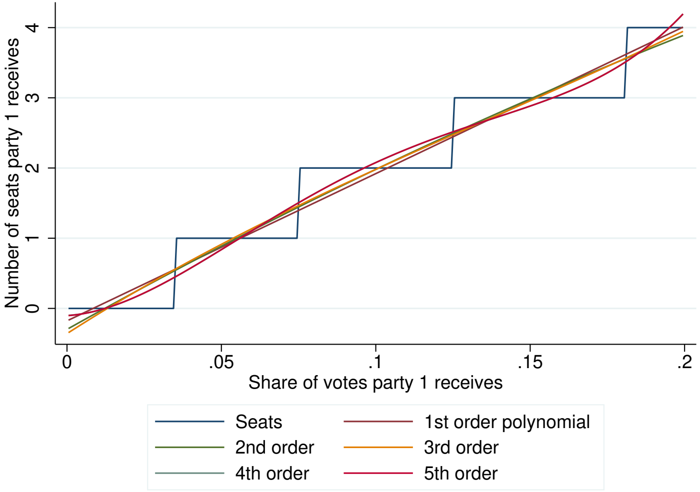
Regression discontinuity and control functions

- Regression discontinuity is the most popular control function
  - Transparent and easy to understand
  - Straightforward figures facilitate non-parameteric analysis and visual tests of identification assumptions
- In practice, most people collapse complete control functions into sets of discontinuities and use RD (example in 2 slides)
- But the *idea* is more general, and it's worth understanding it because...
  - It helps us understand what we're doing and looking for
  - It is a natural implication of the OLS material we've covered

Control functions: Basic idea

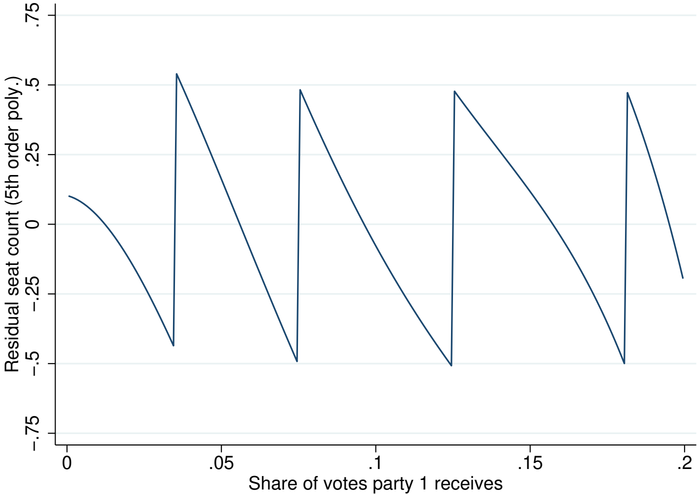
- (Example on the next slide)
- Assume that all potential outcomes are smooth across all non-treatment variables
- But there are “non-smoothies” in the treatment assignment rule
    - discontinuities, non-linearities in excess of what a plausible conditional expectation function would have, etc.
- We can use points of non-smoothness to generate *exogenous* variation in treatment assignment
- How? Use the variation that persists in excess of the best smooth approximation of the treatment assignment rule
- Why does this work to control for other determinants of potential outcomes, even if we don’t observe them? Because of FWL!

Seat assignments and vote share: Swedish Municipal Councils (20 seats)

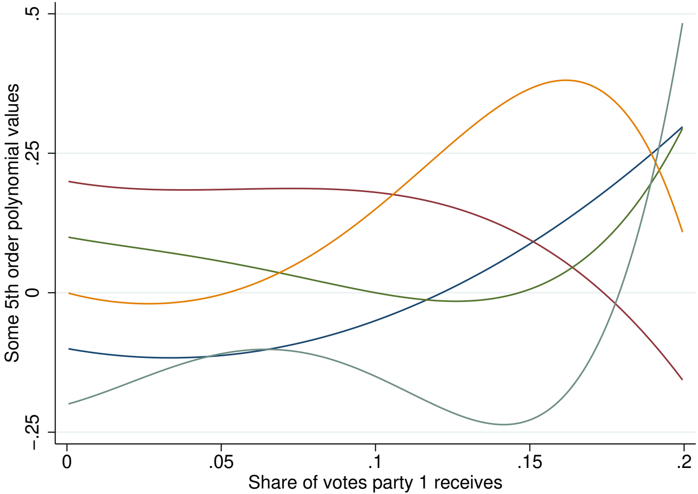


$v_2 = .6(1 - v_1),\ v_3 = .3(1 - v_1),\ v_4 = .1(1 - v_1)$

Residual seat assignments: Swedish Municipal Councils



5th order polynomials



Control functions

- Suppose we have k variables, and that x_j is our "treatment" variable
    - i.e., the main variable of interest that we want to estimate the causal effect of
- FWL: The variation that identifies the coefficient on x_j is the variation orthorgonal to all other variables x_m, $m \neq j$
- If you control for x_m, you're just projecting out of your treatment variable a *linear* function of x_m
- Because of OV'B', this also projects out any *unobserved* variable linear in x_m
    - Estimate $x_j = x_m\gamma_{obs} + \nu$ and γ will be $\gamma_{obs} + \frac{cov(x_m, x_{k+1})}{var(x_m)}\gamma_{unobs}$ where x_{k+1} is some unobserved variable, and γ_{unobs} is the coefficient that it *would* have if you included x_{k+1} in this regression
    - This is a classic result: If you cannot control for something because you don't observe it, you can control for a bunch of stuff correlated with it
- That's ok if you believe $E\big(Y(1)|x_m\big)$ and $E\big(Y(0)|x_m\big)$ are linear in x_m and x_{k+1}
- You could control for some crazy super flexible polynomial, but we generally assume that eliminates all of our identifying variation

Control functions

- Suppose we have k variables, and that x_j is our "treatment" variable
  - i.e., the main variable of interest that we want to estimate the causal effect of
- FWL: The variation that identifies the coefficient on x_j is the variation orthorgonal to all other variables x_m, $m \neq j$
- If you control for x_m, you're just projecting out of your treatment variable a *linear* function of x_m
- Because of OV'B', this also projects out any *unobserved* variable linear in x_m
  - Estimate $x_j = x_m \gamma_{obs} + \nu$ and γ will be $\gamma_{obs} + \frac{cov(x_m, x_{k+1})}{var(x_m)} \gamma_{unobs}$ where x_{k+1} is some unobserved variable, and γ_{unobs} is the coefficient that it *would* have if you included x_{k+1} in this regression
  - This is a classic result: If you cannot control for something because you don't observe it, you can control for a bunch of stuff correlated with it
- That's ok if you believe $E\left(Y(1)|x_m\right)$ and $E\left(Y(0)|x_m\right)$ are linear in x_m and x_{k+1}
- You could control for some crazy super flexible polynomial, but we generally assume that eliminates all of our identifying variation
- **Control functions are about finding treatment assignment rules that are so non-smooth that flexible polynomials don't eliminate identifying variation in our treatment of interest**

Control functions

- Basic idea: Find treatment assignment rules that include non-smoothies[5]
- Observation 1: A gram of institutional knowledge is worth a kilogram of econometric sophistication
  - Most of the most convincing causal research in modern economics uses OLS
  - But it finds institutional contexts where non-smoothies create exogenous variation
- Observation 2: This has huge effects on the questions we ask
  - Non-smoothies come up a lot (education policy, tax policy, public assistance, etc.)
  - But they almost always come from policy
  - Non-smoothies don't happen with equilibrium market objects because they move slowly due to no arbitrage conditions and general equilibrium equalization
    - Sometimes example: Wage setting
  - This means doing causal research in fields focused on market equilibrium objects (like macro and traditional labor) is necessarily harder

[5]For a formal discussion of control functions, see Wooldridge, Jeffrey M. "Control function methods in applied econometrics." *Journal of Human Resources* 50, no. 2 (2015): 420-445. For an applied discussion, see Dahl, Gordon B., and Lance Lochner. "The impact of family income on child achievement: Evidence from the earned income tax credit." *American Economic Review* 102, no. 5 (2012): 1927-1956.

Summary of approaches to modern causal inference (and lessons from our course)

1. True experiments
   - Experiments are only *conditionally* exogenous
   - You still have to think about OLS as variance-weighted averaging over potentially heterogeneous treatment effects
2. Assumed exogeneity
   - IV's don't have to be perfect to be valuable
   - One view: We cannot learn anything from an IV that isn't causal
   - Alternative view 1 (lottery example):
     - Even completely exogenous random variation rarely makes a good instrument
     - We should define our causal estimands of interest to be things we can actually capture
     - Focus on the reduced form, and eliminate your mental internal/external validity distinction
   - Alternative view 2 (Bartik example):
     - OLS is pooling over tons of variation and its confusing
     - It's useful to know the projection onto *one* source of variation
3. Imputing the counterfactual
4. Control functions

Summary of approaches to modern causal inference (and lessons from our course)

1. True experiments
2. Assumed exogeneity
3. Imputing the counterfactual
   - Diff-in-diff is about imputing coutnerfactuals using additive separability assumptions
   - In their pure form, these assumptions are not testable
   - But there are lots of realistic, reasonable "hints" we can test
   - Potential outcomes makes it easy to formalize all of this stuff, including the rapid econometric advances
4. Control functions
   - OLS projects outcome variance onto treatment variables' residual variance
   - Strategy: Find instances where decision-rules generate treatment variation that is more non-smooth than any reasonable conditional expectation function could be
   - This is more about understanding institutions than econometrics

Conclusion

- Thanks for taking this course seriously and being attentive and active
- It means a lot, and I had a lot of fun
- Get excited about Markus' lectures and Econometrics II
- I teach in the labor and political economy sequences
    - I also participate in those research groups
- I hope to talk to many of you again as you continue your research careers