Causality, Introduction

Econ 2560, Fall 2023

Prof. Josh Abel

(Chapters 9.2, 13.1-2, Appendix 13.3)

Introduction

- We have discussed how regression coefficients often capture non-causal effects
  - E.g. The "effect" of education on income from a univariate regression likely captures things like family background or AFQT score
- Multivariate regression can help strip some of those "confounding factors" out
  - E.g. Including AFQT in the regression:
    1. "Effect of education, holding AFQT constant"
    2. "Effect of the part of education that cannot be explained by AFQT"

Control Variables Approach vs. Research Design Approach

- "Control Variable Approach": Rather than using all variation in Education, we try to strip out the tainted variation and just use the cleaner residual variation
  - That's multivariate regression
  - But we do not know where the residual variation is coming from. It could still be tainted.
    - Even if we focus on biological twins, they must differ in some ways if one gets more education...
- "Research Design Approach": rather than starting with all variation and stripping out the bad, let's just find some good variation!

Experiments

- In the hard sciences, causal relationships are established with lab experiments
- To understand how applied economists think about causality, it will be useful to think formally about *why* an experiment is so compelling

Abadie et al (2001)

	Entire Sample	Assignment: Treatment	Assignment: Control	Assignment: Diff. (*t*-stat.)	Treatment: Trainees	Treatment: Non-Trainees	Treatment: Diff. (*t*-stat.)
B. Women							
Number of observations	6,102	4,088	2,014		2,722	3,380	
Treatment							
Training	.45	.66	.02	.64			
	[.50]	[.47]	[.13]	(80.24)			
Outcome Variable							
30 month earnings	13,029	13,439	12,197	1,242	14,211	12,078	2,133
	[13,415]	[13,614]	[12,964]	(3.46)	[13,550]	[13,230]	(6.18)
Baseline Characteristics							
Age	33.33	33.33	33.35	−.02	33.11	33.52	−.41
	[9.78]	[9.77]	[9.81]	(−.09)	[9.71]	[9.84]	(−1.62)
High school or GED	.72	.73	.70	.03	.75	.70	.05
	[.43]	[.43]	[.44]	(2.01)	[.42]	[.45]	(5.07)
Married	.22	.22	.21	.01	.22	.21	.01
	[.40]	[.40]	[.39]	(1.55)	[.40]	[.39]	(1.35)
Black	.26	.27	.26	.01	.26	.27	−.01
	[.44]	[.44]	[.44]	(.95)	[.44]	[.44]	(−.97)
Hispanic	.12	.12	.12	−.00	.12	.11	.01
	[.32]	[.32]	[.33]	(−.89)	[.33]	[.32]	(1.29)
Worked less than 13 weeks in past year	.52	.52	.52	−.00	.51	.53	−.02
	[.47]	[.47]	[.47]	(−.08)	[.47]	[.47]	(−1.52)
AFDC	.31	.30	.31	−.01	.32	.30	.02
	[.46]	[.46]	[.46]	(−1.03)	[.47]	[.46]	(1.92)

Potential Outcomes

- Consider workers $i = 1, ..., n$ who are being considered for a job training program
- Each worker has 2 potential outcomes:
  - $Y_i^p(1)$: Earnings if received training
  - $Y_i^p(0)$: Earnings if no training
- **Treatment Effect** is $\beta = Y_i^p(1) - Y_i^p(0)$
  - This is how we'll define a "causal effect"

Assignment of treatment

- We can define an indicator variable for treatment

$$T_i = \begin{cases} 1 \text{ if receives training} \\ 0 \text{ if no training} \end{cases}$$

- We can relate the observed outcome (Y_i) to the potential outcomes and treatment:

$$Y_i = \begin{cases} Y_i^p(1) \text{ if } T_i = 1 \\ Y_i^p(0) \text{ if } T_i = 0 \end{cases} = Y_i^p(T_i)$$

Randomization

- Consider the difference in means between treatment and control:

$$\begin{aligned}\hat{\beta} &= E[Y_i|T_i = 1] - E[Y_i|T_i = 0] \\ &= E[Y_i^p(1)|T_i = 1] - E[Y_i^p(0)|T_i = 0]\end{aligned}$$

- Now suppose that treatment was randomly assigned, as in an experiment
- Then $E[Y_i^p(1)|T_i = 1] = E[Y_i^p(1)]$
- Then $E[Y_i^p(0)|T_i = 0] = E[Y_i^p(0)]$
    - Random assignment means that the average potential outcomes in the treatment/control groups are the same as in the population at large

Randomization

- Consider the difference in means between treatment and control:

$$\begin{aligned}\hat{\beta} &= E[Y_i | T_i = 1] - E[Y_i | T_i = 0] \\ &= E[Y_i^p(1) | T_i = 1] - E[Y_i^p(0) | T_i = 0]\end{aligned}$$

- Now suppose that treatment was randomly assigned, as in an experiment
- Then $E[Y_i^p(1) | T_i = 1] = E[Y_i^p(1)]$
- Then $E[Y_i^p(0) | T_i = 0] = E[Y_i^p(0)]$
  - Random assignment means that the average potential outcomes in the treatment/control groups are the same as in the population at large

$$\hat{\beta} = E[Y_i^p(1)] - E[Y_i^p(0)] = \beta$$

The experimental ideal

- A randomized experiment identifies causality because we know the only systematic difference between the groups is the treatment
- If there were no random component of the treatment, we could try to control for confounding factors (age, education, etc.)
    - But we could never be sure that the drivers of residual variation in job training were not also driving differences in Y_i
- Experiments exploit only variation in treatment that we know to be random!

Partial compliance

- Experiments might not always go as planned, if subjects don't comply
  - Some assigned to attend the training might not show up
  - Some not assigned the training might find a way in anyhow
- Working through this more complicated case not only shows us how to handle experiments with imperfect compliance, but also – much more importantly – how to work with non-experimental data

Partial compliance (2)

- Let Z_i be the randomized assignment:

$$Z_i = \begin{cases} 1 \text{ if assigned treatment} \\ 0 \text{ if not assigned treatment} \end{cases}$$

- T_i still shows whether subject *actually* received treatment
- Y_i is still observed outcome; $Y_i^p(0)$ and $Y_i^p(1)$ are still potential outcomes

The Identification Problem

- Consider the difference in means from before:

$$\begin{aligned}\hat{\beta} &= E[Y_i|T_i = 1] - E[Y_i|T_i = 0] \\ &= E[Y_i^p(1)|T_i = 1] - E[Y_i^p(0)|T_i = 0]\end{aligned}$$

- Can we still consider it a causal effect?

The Identification Problem

- Consider the difference in means from before:

$$\begin{aligned}\hat{\beta} &= E[Y_i | T_i = 1] - E[Y_i | T_i = 0] \\ &= E[Y_i^p(1) | T_i = 1] - E[Y_i^p(0) | T_i = 0]\end{aligned}$$

- Can we still consider it a causal effect?
- No.
- People who receive treatment might have different potential outcomes than those who do not ("selection bias")
  - Maybe really motivated people make sure to get training, and they also have strong earnings potential
  - Or maybe people with low earnings potential make sure to get training because they really need it

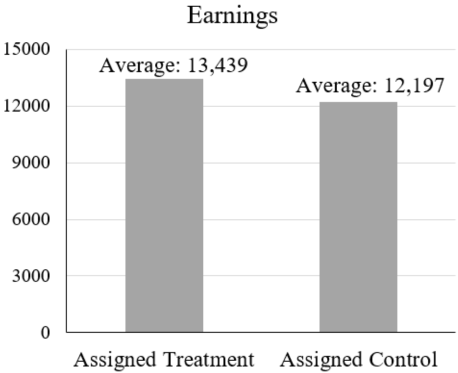
An instrumental variable

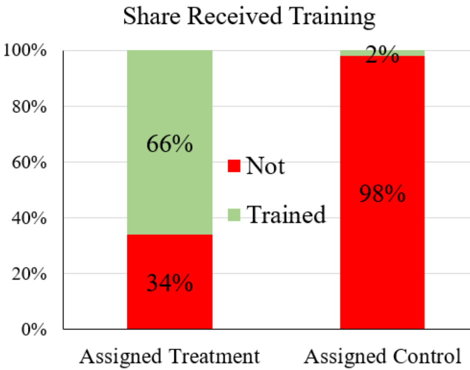
- The non-randomness of T_i will prevent us from simply comparing treatment and control groups
- Instead, consider: $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$
- Effect of *assigned* treatment (random) rather than *actual* treatment (non-random)
    - Randomization of Z_i will give this a causal interpretation!
    - But it's subtle...

Formalizing partial compliance

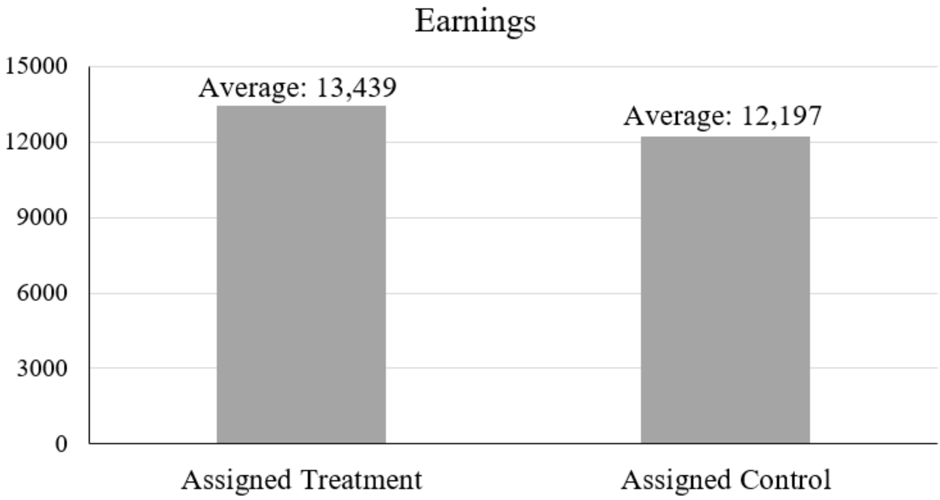
- Can think of 4 types of people
  1. "Always takers": $T_i = 1$ regardless of Z_i
     - They get treatment regardless of whether they are assigned to it
     - Denote their share as s_{AT}
  2. "Never takers": $T_i = 0$ regardless of Z_i
     - They never get treatment regardless of whether they are assigned to it
     - Denote their share as s_{NT}
  3. "Compliers": $T_i = Z_i$
     - They get treatment if and only if they are assigned to it
     - Denote their share as s_C
  4. "Defiers": $T_i = 1 - Z_i$
     - They always do the opposite of their assignment
     - We will assume they don't exist ($s_D = 0$)

Visual Analysis of Abadie et al (2001) Data

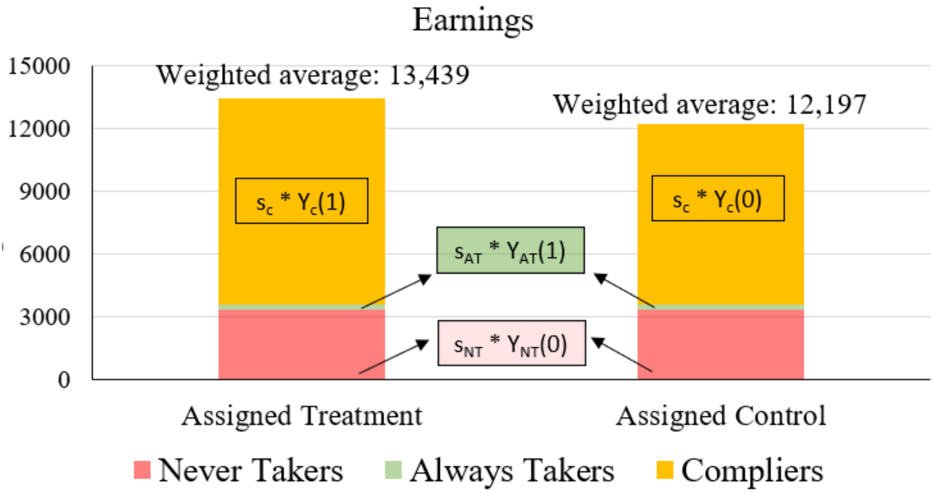




Visual Analysis of Abadie et al (2001) Data

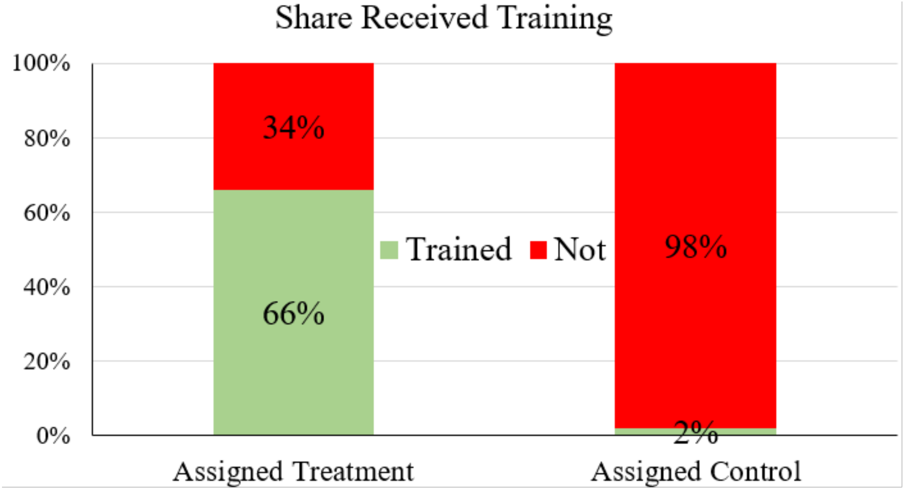


Visual Analysis of Abadie et al (2001) Data



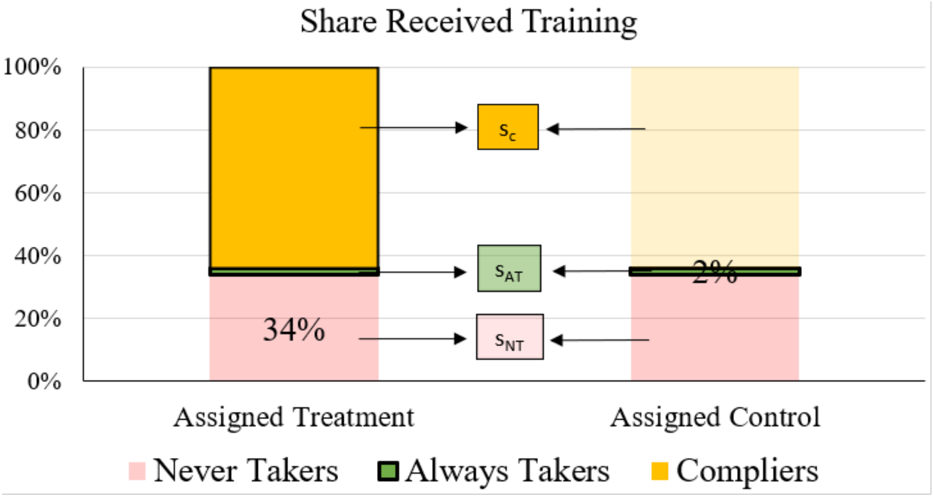
$$13,439 - 12,197 = s_c \cdot \left(Y_c(1) - Y_c(0)\right)$$

Visual Analysis of Abadie et al (2001) Data



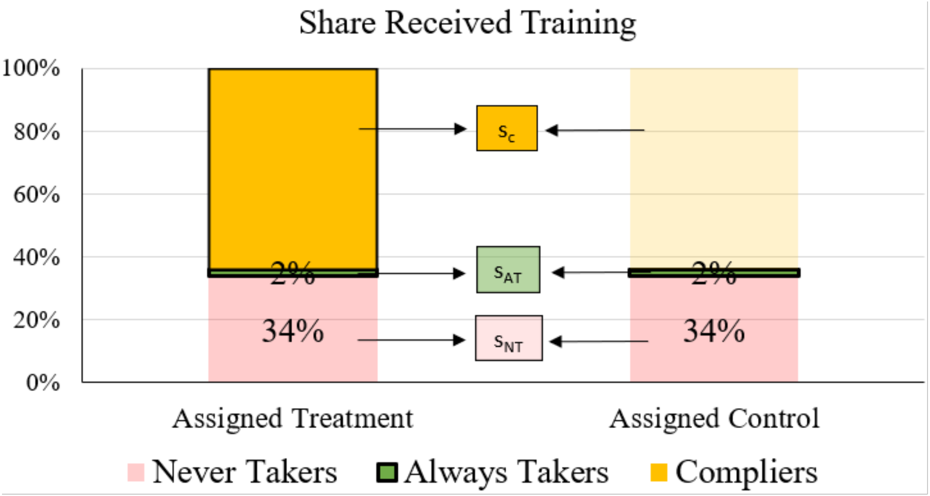
$$13,439 - 12,197 = s_c \cdot \left(Y_c(1) - Y_c(0)\right)$$

Visual Analysis of Abadie et al (2001) Data



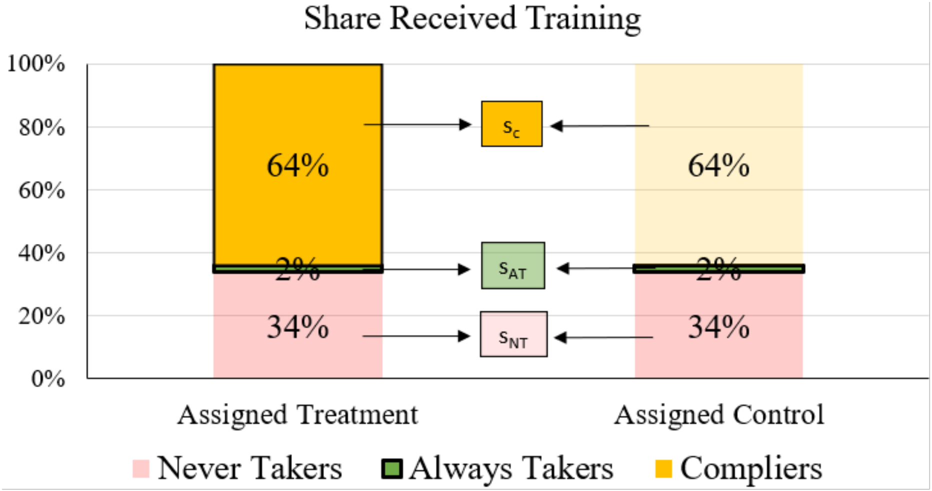
$$13,439 - 12,197 = s_c \cdot \big(Y_c(1) - Y_c(0)\big)$$

Visual Analysis of Abadie et al (2001) Data



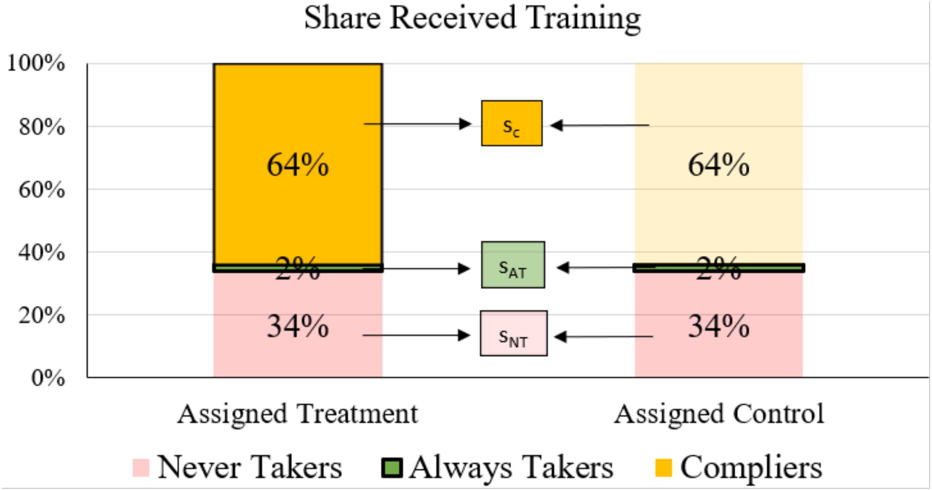
$$13,439 - 12,197 = s_c \cdot \big(Y_c(1) - Y_c(0)\big)$$

Visual Analysis of Abadie et al (2001) Data



$$13,439 - 12,197 = (0.66 - 0.02) \cdot \big(Y_c(1) - Y_c(0)\big)$$

Visual Analysis of Abadie et al (2001) Data



$$Y_c(1) - Y_c(0) = \frac{13{,}439 - 12{,}197}{0.66 - 0.02} = \frac{1{,}242}{0.64} = 1,941$$

Some important algebra

So what is $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$?

Some important algebra

So what is $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$?

$$E[Y_i|Z_i = 1] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(1)|\text{C}] \cdot s_C$$

Some important algebra

So what is $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$?

$$E[Y_i|Z_i = 1] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(1)|\text{C}] \cdot s_C$$
$$E[Y_i|Z_i = 0] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(0)|\text{C}] \cdot s_C$$

Some important algebra

So what is $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$?

$$E[Y_i|Z_i = 1] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(1)|\text{C}] \cdot s_C$$
$$E[Y_i|Z_i = 0] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(0)|\text{C}] \cdot s_C$$

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = \Big(E[Y_i^p(1)|C] - E[Y_i^p(0)|C]\Big) \cdot s_C$$

Some important algebra

So what is $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$?

$$E[Y_i|Z_i = 1] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(1)|\text{C}] \cdot s_C$$
$$E[Y_i|Z_i = 0] = E[Y_i^p(1)|\text{AT}] \cdot s_{AT} + E[Y_i^p(0)|\text{NT}] \cdot s_{NT} + E[Y_i^p(0)|\text{C}] \cdot s_C$$

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = \Big(E[Y_i^p(1)|C] - E[Y_i^p(0)|C]\Big) \cdot s_C$$

Will show on problem set that: $s_C = E[T_i|Z_i = 1] - E[T_i|Z_i = 0]$

$$E[Y_i^p(1)|C] - E[Y_i^p(0)|C] = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[T_i|Z_i = 1] - E[T_i|Z_i = 0]}$$

Interpreting the algebra

$$E[Y_i^p(1)|C] - E[Y_i^p(0)|C] = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[T_i|Z_i = 1] - E[T_i|Z_i = 0]}$$

- We have identified the causal effect *for compliers*
- Numerator: Difference in outcome from being *assigned treatment*
  - Often called "reduced form" or "intent to treat (ITT)"
  - Because Z_i is random, ITT is a causal effect.
  - But it understates the true effect of the treatment, because the true effect is diluted by ATs and NTs who are not responsive to the assignment (Z)
- Denominator: Difference in treatment likelihood based on assignment
  - Often called "first stage"
  - Scaling factor to correct for the dilution mentioned above
- Intuitively, only using the variation in T_i that is randomly assigned

Local Average Treatment Effect (LATE)

- If the treatment is non-random but there is a random component, we can identify the effect of the treatment for the group that is responsive to the random part
- Imagine if we randomly assigned college scholarships to high school seniors

Local Average Treatment Effect (LATE)

- If the treatment is non-random but there is a random component, we can identify the effect of the treatment for the group that is responsive to the random part
- Imagine if we randomly assigned college scholarships to high school seniors
  - Some would have attended college regardless (Always Takers)
  - Some will not attend even with the scholarship (Never Takers)
  - But some will be induced to attend because of the scholarship (Compliers)

Local Average Treatment Effect (LATE)

- If the treatment is non-random but there is a random component, we can identify the effect of the treatment for the group that is responsive to the random part
- Imagine if we randomly assigned college scholarships to high school seniors
  - Some would have attended college regardless (Always Takers)
  - Some will not attend even with the scholarship (Never Takers)
  - But some will be induced to attend because of the scholarship (Compliers)
- We can compute the causal effect of college on Compliers!
  - Compare earnings of scholarship recipients to those who did not get the scholarship (reduced form)
  - Divide by difference in college attendance rates between recipients and non-recipients (first stage)
- Note that we are able to do this despite not knowing which specific people are Compliers

Abadie et al (2001) – table for men

	Entire Sample	Assignment: Treatment	Assignment: Control	Assignment: Diff. (t-stat.)	Treatment: Trainees	Treatment: Non-trainees	Treatment: Diff. (t-stat.)
A. Men							
Number of observations	5,102	3,399	1,703		2,136	2,966	
Treatment							
Training	.42	.62	.01	.61			
	[.49]	[.48]	[.11]	(70.34)			
Outcome variable							
30 month earnings	19,147	19,520	18,404	1,116	21,455	17,485	3,970
	[19,540]	[19,912]	[18,760]	(1.96)	[19,864]	[19,135]	(7.15)
Baseline Characteristics							
Age	32.91	32.85	33.04	−.19	32.76	33.02	−.26
	[9.46]	[9.46]	[9.45]	(−.67)	[9.64]	[9.32]	(−.95)
High school or GED	.69	.69	.69	−.00	.71	.68	.03
	[.45]	[.45]	[.45]	(−.12)	[.44]	[.45]	(2.46)
Married	.35	.36	.34	.02	.37	.34	.03
	[.47]	[.47]	[.46]	(1.64)	[.47]	[.46]	(2.82)
Black	.25	.25	.25	.00	.26	.25	.01
	[.44]	[.44]	[.44]	(.04)	[.44]	[.43]	(.48)
Hispanic	.10	.10	.09	.01	.10	.09	.01
	[.30]	[.30]	[.29]	(.70)	[.31]	[.29]	(1.60)
Worked less than 13 weeks in past year	.40	.40	.40	.00	.40	.40	−.00
	[.47]	[.47]	[.47]	(.56)	[.47]	[.47]	(−.32)

3 Steps for an Instrumental Variable (IV) Analysis

1. Argue that your "instrument" is effectively random
    - In an experiment like Abadie et al (2001), this is basically trivial
    - In most IV analyses, this is by far the hardest part
2. Calculate the Local Average Treatment Effect (LATE)
    - In Abadie et al (2001), this was $1,941 for women and $1,830 for men
3. Consider who the Compliers are and how this affects your analysis
    - In Abadie et al (2001), these are people who will attend a job training program if they are admitted
        - Do you think their LATE is higher, lower, or the same as the overall ATE?
        - Which is more policy-relevant: ATE or Compliers' LATE?

Abel and Fuster (2021)

TABLE 2—COMPARISON OF HARP-ELIGIBLE VERSUS HARP-INELIGIBLE BORROWERS IN OUR SAMPLE

	Mean: Eligible ($N = 115{,}130$)	Mean: Ineligible ($N = 104{,}277$)	Ineligible − Eligible	SE on Diff.
CLTV (%)	87.3	88.5	1.2	0.1
Orig. CLTV (%)	86.6	87.4	0.8	0.1
Credit score	763.6	759.6	−4.0	0.2
Interest rate (%)	4.98	4.93	−0.05	0.00
Credit utilization (%)	24.5	24.7	0.2	0.1
First mortgage balance (1,000\$)	225.9	225.1	−0.8	0.4
All other debt balances (1,000\$)	24.3	25.0	0.8	0.1
Purchase mortgage (%)	25.8	29.6	3.8	0.2
Refinanced (%)	53.17	34.96	−18.21	0.21
Defaulted (%)	3.53	3.64	0.11	0.08
Servicing transferred from McDash (%)	6.15	8.22	2.07	0.11
Refinanced out of McDash (%)	11.37	11.82	0.45	0.14
Non-refi prepay (%)	21.10	19.27	−1.83	0.17
Active (%)	57.85	57.05	−0.79	0.21

Notes: The top half of the table checks for balance in our sample between borrowers with mortgages purchased by a GSE before the cutoff date (eligible) and mortgages purchased after (ineligible). Variables in the top half are measured in March 2010. The bottom half of the table shows the fraction that ever refinanced in the sample and the termination status of the mortgages as of February 2016.

3 Steps for Abel and Fuster (2021)

1. Argue that your "instrument" is effectively random
   - Buyers just after cutoff date should be similar to those just before
   - Plus, "observables" look similar
2. Calculate the Local Average Treatment Effect (LATE)
   - $\frac{0.0353-0.0364}{0.5317-0.3496} = -0.0060 \approx 0.6\%$
3. Consider who the Compliers are and how this affects your analysis
   - People who refinance but only because of the program
     - This is the most policy-relevant group!

Viewed as regressions

- Because Z_i is binary, we can think of the reduced form as the following regression:

$$E[Y_i|Z_i] = \beta_0^{RF} + \beta_1^{RF} \cdot Z_i$$

  - $\beta_1^{RF} = E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$

- Similarly, the first stage can be done as a regression:

$$E[T_i|Z_i] = \beta_0^{FS} + \beta_1^{FS} \cdot Z_i$$

  - $\beta_1^{FS} = E[T_i|Z_i = 1] - E[T_i|Z_i = 0]$

- The Local Average Treatment Effect is then:

$$\beta = \frac{\beta_1^{RF}}{\beta_1^{FS}}$$

Viewed as regressions with controls

- The regression framework is useful because then we can add controls:

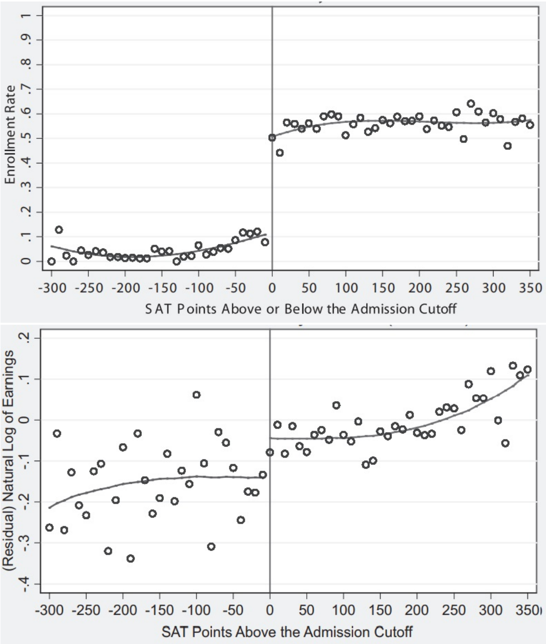
$$E[Y_i|Z_i, X_{1i}, ...] = \beta_0^{RF} + \beta_1^{RF} \cdot Z_i + \beta_2^{RF} \cdot X_{1i} + ...$$

$$E[T_i|Z_i, X_{1i}, ...] = \beta_0^{FS} + \beta_1^{FS} \cdot Z_i + \beta_2^{FS} \cdot X_{1i} + ...$$

$$\beta = \frac{\beta_1^{RF}}{\beta_1^{FS}}$$

- Why include controls if we have randomized?
  - Maybe in our finite sample, Z ended up correlating with some X by random chance
  - Maybe we were only able to randomize within certain X categories, but Z is not randomly assigned across X
  - If X is good at explaining Y, then including X could help lower the standard errors on β_1^{RF} and β_1^{FS}!

Hoekstra (2009)



3 Steps for Hoekstra (2009)

1. Argue that your “instrument” is effectively random

3 Steps for Hoekstra (2009)

1. Argue that your "instrument" is effectively random
    - Someone who scores 1 point above the threshold is effectively the same as someone who scores 1 point below

3 Steps for Hoekstra (2009)

1. Argue that your “instrument” is effectively random
   - Someone who scores 1 point above the threshold is effectively the same as someone who scores 1 point below
2. Calculate the Local Average Treatment Effect (LATE)
   - $\frac{0.095}{0.426} = 0.223$

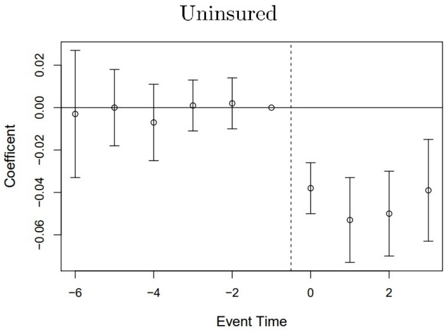
3 Steps for Hoekstra (2009)

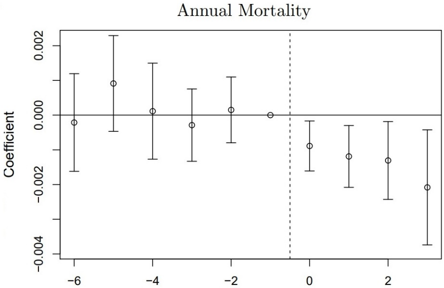
1. Argue that your "instrument" is effectively random
   - Someone who scores 1 point above the threshold is effectively the same as someone who scores 1 point below
2. Calculate the Local Average Treatment Effect (LATE)
   - $\frac{0.095}{0.426} = 0.223$
3. Consider who the Compliers are and how this affects your analysis

3 Steps for Hoekstra (2009)

1. Argue that your "instrument" is effectively random
    - Someone who scores 1 point above the threshold is effectively the same as someone who scores 1 point below
2. Calculate the Local Average Treatment Effect (LATE)
    - $\frac{0.095}{0.426} = 0.223$
3. Consider who the Compliers are and how this affects your analysis
    - People who would attend flagship state school if they get admitted due to SAT score, not otherwise
        - Non-Compliers are mostly Never Takers.
        - Could be people who would go to a more prestigious school, so their treatment effect might be larger than for Compliers...
        - Could be people who would not attend college at all, so their treatment effect might be lower than for Compliers...

Miller et al (2019)





3 Steps for Miller et al (2019)

1. Argue that your "instrument" is effectively random

3 Steps for Miller et al (2019)

1. Argue that your "instrument" is effectively random
    - Non-Medicaid expansion states are the same, at least once you control for stuff...

3 Steps for Miller et al (2019)

1. Argue that your "instrument" is effectively random
    - Non-Medicaid expansion states are the same, at least once you control for stuff...
2. Calculate the Local Average Treatment Effect (LATE)
    - $\approx \frac{-0.002}{-0.040} = 0.050$

3 Steps for Miller et al (2019)

1. Argue that your “instrument” is effectively random
   - Non-Medicaid expansion states are the same, at least once you control for stuff...
2. Calculate the Local Average Treatment Effect (LATE)
   - $\approx \frac{-0.002}{-0.040} = 0.050$
3. Consider who the Compliers are and how this affects your analysis

3 Steps for Miller et al (2019)

1. Argue that your "instrument" is effectively random
   - Non-Medicaid expansion states are the same, at least once you control for stuff...
2. Calculate the Local Average Treatment Effect (LATE)
   - $\approx \frac{-0.002}{-0.040} = 0.050$
3. Consider who the Compliers are and how this affects your analysis
   - People who are insured only because of the Medicaid expansion
     - Likely lower-income, but not the lowest of the income distribution. So maybe in the middle of the pack in terms of treatment effect
   - But they are the policy-relevant group for Medicaid expansion!