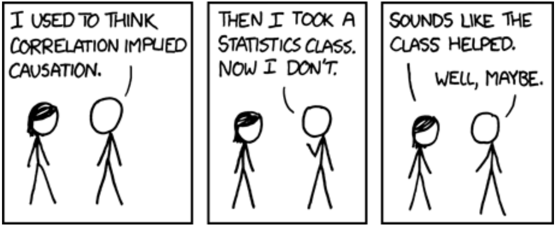
Advanced Applied Econometrics

Teacher: Felix Weinhardt



Slides heavily borrow from Mixed Tape
and Mostly Harmless Econometrics textbooks

Get some intution for subgroup populations: binary encouragement designs

- Four types of people by potential outcomes
  - **Compliers** - treatment if encouraged, control if not
  - **Always-takers** - treatment whether encouraged or not
  - **Never-takers** - control whether encouraged or not
  - **Defiers** - control if encouraged, treatment if not encouraged
- Not all of these may exist in a particular study
- In a trial of a new drug or offering (removing) a new (existing) feature, there are neither always-takers nor defiers

Heterogenous treatment effects

- *Monotonicity*: With probability 1, $D_i^z \geq D_i^{z'}$ for all $z \geq z'$ and all i
- Then local average treatment effect (LATE) is identified
- In binary Z, D case, LATE is the average treatment effect for the population of compliers
- We always have to ask ourselves: are we interested in the LATE?

Reduced form: Intent-to-treat (ITT)

- One option is forego the IV analysis and instead estimate the ITT
- In other words, analyze ego outcome Y as a function of the number of peers assigned Z
- What else?
  - Compute probabilities of assignment, $Pr(Z)$
  - Conduct complier analysis

Complier analysis: Intuition

- We cannot identify the correct group at the individual level
- But we can compare compliers again other populations at the group level
- This will give us a better sense for the types of individuals who are driving the LATE effect sizes

Complier analysis: size of complier population

- In case of a binary IV and treatment the size of the complier population is given by the WALD- first stage
- How many individuals got moved into treatment from having the different instrument assigned (taking the pill)?
- This will give us a better sense for the types of individuals who are driving the LATE effect sizes

Size of the complier group

$$P[D_i^1 > D_i^0] = E[D_i^1 - D_i^0]$$
$$= E[D_i^1] - E[D_i^0]$$
$$= E[D^i|Z_i = 1] - E[D^i|Z_i = 0]$$

Complier analysis: proportion of complier population among treated

- Note for compliers treatment status is completely determined by Z_i

Proportion of the complier group (among treated)

$$P[D_i^1 > D_i^0 | D_i = 1]$$
$$= \frac{P[D_i^1 = 1 | D_i^1 > D_i^0] P[D_i^1 > D_i^0]}{P[D_i = 1]}$$
$$= E[D_i^1] - E[D_i^0]$$
$$= E[D^i | Z_i = 1] - E[D^i | Z_i = 0]$$

- The proportion of the treated who are compliers is given by the first stage, times the probability of the instrument switched on, divided by the proportion treated.
- Tough to understand intuitively but this can be calculated!

Examples of quantifying compliers

Source	Endogenous Variable	Instrument	Sample	Size of complier group	Proportion compliers among treated
Angrist (1990)	Veteran Status	Draft Eligibility	White men born 1950	0.534	0.101
			Non-white men born 1950	0.534	0.033
Angrist and Evans (1998)	More than 2 children	Twins as second birth	Married women with 2 or more children	0.603	0.966
Angrist and Krueger (1991)	High school graduate	3rd/4th quarter birth	Men born 1930-39	0.016	0.034

Charaterising compliers

- Despite the fact that we cannot identify individual compliers, we can describe the distribution of complier characteristics
- This can be compared to the general population/sample
- As a result we can get a sense of how the complier population is selected with respect to certain X vars.

Charaterising compliers

Example: Are sex composition compliers more or less likely to be college graduates than other comen with two children?

Charaterising compliers (relative first stages)

$$\frac{P[X_i^1=1|D_i^1>D_i^0]}{P[X_i^1=1]} = \frac{P[D_i^1>D_i^0|X_i^1=1]}{P[D_i^1>D_i^0]}$$
$$= \frac{E[D_i|Z_i=1,X_i^1=1]-E[D_i|Z_i=0,X_i^1=1]}{E[D_i|Z_i=1]-E[D_i|Z_i=1]}$$

- This is the ratio of the first stage for college graduates to the overall first stage.
- This gives the relative likelihood that a complier is a college graduate.

Examples: charaterising compliers for twin instruments

Variable	$P[X_i^1 = 1]$	$P[X_i^1 = 1 \vert D_i^1 > D_i^0]$	$\frac{P[X_i^1=1\vert D_i^1>D_i^0]}{P[X_i^1=1]}$
Age 30 or older at first birth	0.0029	0.004	1.39
Black or hispanic	0.125	0.103	0.822
College graduate	0.132	0.151	1.14

- Twin compliers are more likely to be over 30 at first birth compared to sample.
- They are also less likely to be black or hispanic and more likely to have a college degree.

Beyond complier analysis: the value of having multipe instruments

- In the end complier analysis gives us a sense for the population the LATE is estimated for. We still do not know if LATE generalises.
- Imagine you have multiple valid instruments for the same variable but with different complier populations
  - If you find different effects - these are driven by effect heterogeneity
  - If you find similar effects - effect homogeneity assumptions probably holds
  - Remember the Hausman test that we did to check for exogeneity of instruments when you have multiple instruments there? Conditional on all instruments being valid, this equivalent test can now be used for testing against H0: homogeneity.