Lecture 8: Instrumental Variables (Part III)

Jaakko Meriläinen

5304 Econometrics @ Stockholm School of Economics

Introduction: What We Learned in the Previous Lecture

- IV is a powerful method to deal with bias in OLS
- IV relies on two conditions:
  - Relevance (which can be directly tested)
  - Validity (which is not directly testable)
  - Under maintained assumptions, IV is consistent
- The IV estimator is easily extended to multiple endogenous variables, multiple instruments
- When IVs are weak, even small violations of validity can cause large inconsistency
- This should be tested for and first-stage should always be presented

Plan for Today

- We will extend the discussion in two directions:
  - The interpretation of IV results when potential outcomes are heterogeneous
  - The use of IV methods in randomized trials

- The aim is to get a deeper understanding of what IV methods identify and the specific relationship to different causal quantities ("treatment effects") and external validity

Plan for Today

1. Introduction
2. Heterogenous potential outcomes
   - Compliance groups
   - The LATE (= local average treatment) theorem
   - LATE with multiple instruments
   - LATE with multivalued treatments
3. Examples of LATE and ACR (= average causal response)
   - Oreopoulos (2006)
   - Muralidharan, Singh, and Ganimian (2019)
4. Making sense of LATE
5. Summary

Heterogenous Potential Outcomes

- So far, we have worked under the assumption of constant causal effects
- Any treatment had the same value for all individuals; our parameter (β) did not vary across individuals
- But in practice, this may not be reasonable to assume
- In fact, in general, we expect that treatment effects will vary across individuals!
- E.g., the effect of giving a specialized textbook to a high-performing student could be very different from giving one to a low-performing student!
- Interpreting IV results becomes trickier then

Heterogeneity by Known Characteristics

- The most straightforward way to think of heterogeneity is where treatment effects differ across known (and exogenously determined) sub-groups
- For instance, the returns to education could different for men and women
- At that point, think of running the IV estimates in separate regressions for women and men
- You need first stage variation and validity for both groups
- This could also be estimated with an interaction term
  - Two endogenous variables: you can generate two IVs
  - Particularly important when considering heterogeneity by continuous covariates
- Our major concern, however, relates to heterogeneity of a different kind...

Compliance Groups

- Let us stay with the case of a binary instrument (Z_i) and a binary endogenous variable (X_i)
- Imagine X is finishing college, Z is getting a (randomized) scholarship
- We can think of four types of compliance units:
  - **Always-takers:** $D = 1$, whether $Z = 1$ or $Z = 0$
    - I always go to college, whether or not I get a scholarship
  - **Never-takers:** $D = 0$, whether $Z = 1$ or $Z = 0$
    - I never go to college, whether or not I get a scholarship
  - **Compliers:** $D = 1$ when $Z = 1$; $D = 0$ when $Z = 0$
    - Getting a scholarship makes me go to college
  - **Defiers:** $D = 0$ when $Z = 1$; $D = 1$ when $Z = 0$
    - I would have gone to college otherwise; but now that I have a scholarship, I will not go!

Compliance Groups

Compliance Type by Treatment and Instrument

		Instrument: Z_i 0	1
Treatment: D_i	0	complier/never-taker	never-taker/defier
	1	always-taker/defier	complier/always-taker

- Compliance groups cannot be inferred from data (without adding more assumptions)!

The LATE Assumptions

1. **Independence:** $\{Y_i(D_{1i}, 1), Y_i(D_{0i}, 0), D_{1i}, D_{0i}\} \perp Z_i$
   - This means that Z_i is as good as randomly assigned
2. **Exclusion:** $Y_i(d, 0) = Y_i(d, 1) \equiv Y_{di} \; for \; d = 0, 1$
   - The instrument Z_i only affects Y_i through D_i
3. **First stage:** $E[D_{1i} - D_{0i}] \neq 0$
   - This is the equivalent of what we have called the relevance condition
   - The instrument has some explanatory power for D_i
4. **Monotonicity:** $D_{1i} - D_{0i} \geq 0 \; for \; all \; i$ or vice versa
   - This means that the instrument affects the probability of $D_i = 1$ in the same direction for all individuals
   - This is also called the "no-defiers" assumption

The LATE Theorem

- Under the four assumptions on the previous slide, it can be shown that the IV estimator

$$\beta = \frac{E(Y_i|Z_i = 1) - E(Y_i|Z_i = 0)}{E(D_i|Z_i = 1) - E(D_i|Z_i = 0)}$$

  identifies the effect of treatment on those individuals whose treatment status has been changed by the instrument (i.e. the compliers)

- This parameter is called the **Local Average Treatment Effect**

The LATE Theorem

- Formally, it can be shown that

$$\begin{aligned} \beta_{LATE} &= \frac{E(Y_i|Z_i=1)-E(Y_i|Z_i=0)}{E(D_i|Z_i=1)-E(D_i|Z_i=0)} \\ &= E(Y_{1i}-Y_{0i}|D_{1i}>D_{0i}) \end{aligned}$$

- $D_{1i} > D_{0i}$ is the group of individuals for whom the instrument changes the schooling decision
- For the proof of the LATE theorem, see Angrist and Pischke (2009, p. 155)

- Note that, since different instruments will move different individuals into treatment, **the LATE is specific to the instrument chosen**!

What Does the LATE Tell You?

- Given heterogeneous potential outcomes, the LATE is the best we get without additional assumptions
- It is a well-defined causal estimate but what does it tell us and is it useful?
- This depends entirely on what you want to use the causal estimate for!
- Understanding what LATE does/does not tell you is **the key aspect to learn** about treatment effects when potential outcomes vary across individuals

What the LATE Is Not: LATE$\neq$ATT

- In the first lecture, we had defined the average treatment effect on the treated (abbreviated as ATET or ATT or ToT)—**LATE $\neq$ ATT!**

- To see why, just note that:

$$
\begin{aligned}
ATT &= \underbrace{E[Y_{1i} - Y_{0i}|D_i = 1]}_{\textit{effect on treated}} \\
&= \underbrace{E[Y_{1i} - Y_{0i}|D_{0i} = 1]}_{\textit{effect on always-takers}} P[D_{0i} = 1|D_i = 1] \\
&\quad + \underbrace{E[Y_{1i} - Y_{0i}|D_i > D_{0i}]}_{\textit{effect on compliers}} P[D_i > D_{0i},\ Z_i = 1|D_i = 1]
\end{aligned}
$$

- The effect on the treated is a weighted average of the effect on always-takers (which the LATE does not tell us anything about) and the effect on compliers (which is what the LATE is)

What the LATE Is Not: The Effect on the Un-Treated

- By a similar logic, the effect on the untreated is a weighted average of the treatment effect on never-takers and the treatment effect on compliers

$$
\begin{aligned}
ATU &= \underbrace{E[Y_{1i} - Y_{0i} | D_i = 0]}_{\textit{effect on nontreated}} \\
&= \underbrace{E[Y_{1i} - Y_{0i} | D_{1i} = 0]}_{\textit{effect on never-takers}} P[D_{1i} = 0 | D_i = 0] \\
&\quad + \underbrace{E[Y_{1i} - Y_{0i} | D_i > D_{0i}]}_{\textit{effect on compliers}} P[D_i > D_{0i}, Z_i = 0 | D_i = 0]
\end{aligned}
$$

What the LATE Is Not: The Average Treatment Effect

- This follows naturally from the previous two slides

$$\begin{aligned} ATE &= \underbrace{E[Y_{1i} - Y_{0i}]}_{\text{average treatment effect}} \\ &= \underbrace{E[Y_{1i} - Y_{0i} | D_i = 1]}_{ATT} P[D_i = 1] \\ &+ \underbrace{E[Y_{1i} - Y_{0i} | D_i = 0]}_{ATU} P[D_i = 0] \end{aligned}$$

- The ATE is the weighted average of the ATT and the ATU

- LATE does not pin this down

Generalizing LATE

- We stated the LATE theorem in terms of a binary treatment with a single, binary instrument with no covariates
  - This is the standard illustration of LATE (and communicates most relevant ideas)
  - Note: **all** of this is about interpreting the results from IV estimates, not how they are estimated

- But, as we saw, the 2SLS estimator is more general: it allows for multiple instruments, multiple endogenous variables, covariates etc.

- We will now discuss how 2SLS can be interpreted in these general cases with heterogeneous potential outcomes

LATE with Multiple Instruments

- Imagine if we had two instruments Z_1 and Z_2 which satisfied the LATE assumptions
- We could use either of them as an IV and recover a LATE
- These LATEs are **instrument-specific**: they are defined over the subpopulation moved into treatment by Z_1 or Z_2
- Using only one instrument leaves potentially useful information on the table
- But how should the resulting 2SLS estimate be interpreted?

LATE with Multiple Instruments: Two Binary Instruments

- It can be shown that the 2SLS estimate obtained using both instruments is a weighted-average of the two instrument-specific LATEs
- This is simplest to see in the case of two binary instruments (Z_1 and Z_2) which are uncorrelated with each other
- Define the instrument-specific LATEs as:

$$\rho_j = \frac{Cov(Y_i, Z_{ji})}{Cov(D_i, Z_{ji})};\ j = 1, 2$$

LATE with Multiple Instruments: Two Binary Instruments

- The population first-stage fitted values are:

$$\widehat{D_i} = \pi_{11} z_{1i} + \pi_{12} z_{2i}$$

- Then it can be shown that $\rho_{2SLS} = \psi \rho_1 + (1 - \psi)\rho_2$ where

$$\psi = \frac{\pi_{11} Cov(D_i, Z_{1i})}{\pi_{11} Cov(D_i, Z_{1i}) + \pi_{12} Cov(D_i, Z_{2i})}$$

2SLS Estimates with Multi-Valued Treatments

- A final generalization is to the case of an endogenous variable that takes on many values
- One example of such a variable is years of schooling
  - Clearly the returns to schooling could be heterogeneous across individuals
  - Also, there is no guarantee, except by assumption, that these returns are linear
  - How should we then interpret the 2SLS estimate of the returns to an additional year of schooling?

2SLS Estimates with Multi-Valued Treatments: Average Causal Response

- Angrist and Imbens (1995) discuss this case in depth and show that the 2SLS estimate can be interpreted as an average causal response (ACR)

  > *This parameter captures a weighted average of causal responses to a unit change in treatment, for those whose treatment status is affected by the instrument.*

- The interpretation is thus similar to that of a LATE in the binary treatment case

ATE and LATE of the Returns to Schooling (Oreopoulos 2006)

- Oreopoulos focuses on the distinction between LATE and ATE and how much this might matter in practice
- The paper builds on a large literature (such as Angrist and Krueger, 1991) which use IV methods to recover the returns to schooling in wage regressions
- What is the concern?
  - Treatment effects affect $<10\%$ of the population exposed to the IV
  - IV models are frequently greater than OLS (which is perhaps counter-intuitive)
- One possibility is that OLS, in the absence of bias, approximates the effect on everyone but LATE is on a "small and peculiar group" $\Rightarrow$ The difference between IV and OLS is not about bias but about different populations!

ATE and LATE of the Returns to Schooling (Oreopoulos 2006)

- Oreopoulos (2006) studies this explicitly by looking at returns from schooling using changes in compulsory schooling laws in UK, US and Canada
  - The compulsory schooling age changed, e.g., in Britain from 14 to 15 years in 1947
  - This forces some people to stay in school longer

- Oreopoulos points out that degree of compliance varies a great deal across the three contexts!
  - Because dropout rates in US and Canada were lower, a much greater proportion of the sample can be known as "always-takers"
  - Dropout rates in the UK were much higher, so ~40% of the sample was in the compliers!

- Oreopoulos's argument is that comparing LATEs across samples with very different compliance rates is informative of how much ATEs vary

- This implicitly assumes effects for always-takers/compliers are stable across settings

The Effect of Compulsory Schooling Laws

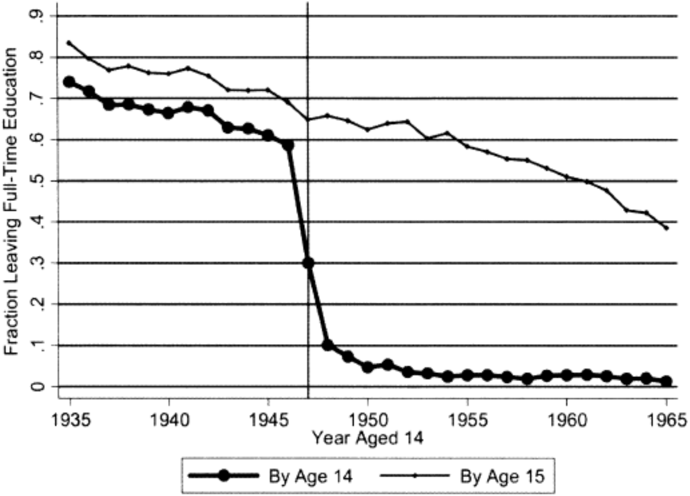


FIGURE 1. FRACTION LEFT FULL-TIME EDUCATION BY YEAR AGED 14 AND 15
(Great Britain)

LATE Results Similar(-ish) Across Countries and Subgroups, IV > OLS Everywhere

	(1) OLS full sample	(2) IV with regional controls	(3) IV with regional trends	(4) IV with regional trends and regional controls
Dependent variable	United States [1901–1961 birth cohorts aged 25–64 in the 1950–2000 censuses]			
Log weekly earnings (all workers)	0.078	0.142	0.175	0.405
	[0.0005]***	[0.0119]***	[0.0426]***	[0.7380]
Log weekly earnings (males)	0.070	0.127	0.074	0.235
	[0.0004]***	[0.0145]***	[0.0384]*	[0.1730]
Log weekly earnings (black males)	0.074	0.172	0.119	0.264
	[0.0004]***	[0.0137]***	[0.0306]***	[0.1295]**
	Canada [1911–1961 birth cohorts aged 25–64 in the 1971–2001 censuses]			
Log annual earnings (all workers)	0.099	0.096	0.095	0.142
	[0.0007]***	[0.0254]***	[0.1201]	[0.0652]**
Log annual earnings (males)	0.087	0.124	−0.383	0.115
	[0.0008]***	[0.0284]***	[1.1679]	[0.0602]*
	United Kingdom [1921–1951 birth cohorts aged 32–64 in the 1983–1998 GHHS]			
Log annual earnings (all workers)	0.079	0.158	0.195	NA
	[0.0024]***	[0.0491]***	[0.0446]***	
Log annual earnings (males)	0.055	0.094	0.066	NA
	[0.0017]***	[0.0568]	[0.0561]	
	Britain [1921–1951 birth cohorts aged 32–64 in the 1983–1998 GHHS]			
	OLS	RD		
Log annual earnings (all workers)	0.078	0.147	NA	NA
	[0.002]***	[0.061]**		
Log annual earnings (males)	0.055	0.150	NA	NA
	[0.0017]***	[0.130]		

Summarizing Oreopoulos (2006)

- The pattern, that IV estimates of returns to schooling exceed OLS, is perhaps unintuitive
  - We typically worry about OLS schooling coefficient being biased upwards
  - If omitted ability is positively correlated with schooling and with independent returns on the labor market

- Oreopoulos (2006) does not tell us why IV $>$ OLS

- But it does suggest that the reason is not that the LATE is being estimated on some peculiar sample
  - This pattern is similar even when the degree of compliance is a lot more!

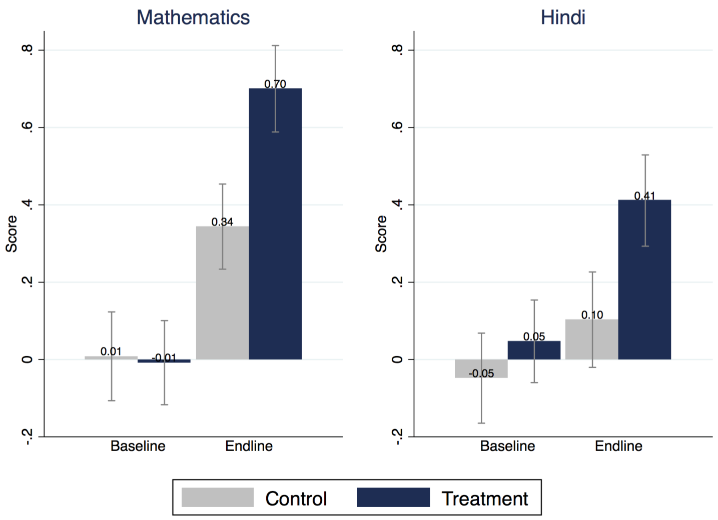
IV in RCTs with a Multivalued treatment

- Muralidharan, Singh and Ganiman (2019) report results from a randomized trial of a personalized learning program in Delhi
  - This is motivated by very low learning levels, very high within-grade heterogeneity
  - The program targets middle school students in Grades 6-9, outside of school hours and premises
  - Participation is individually randomized

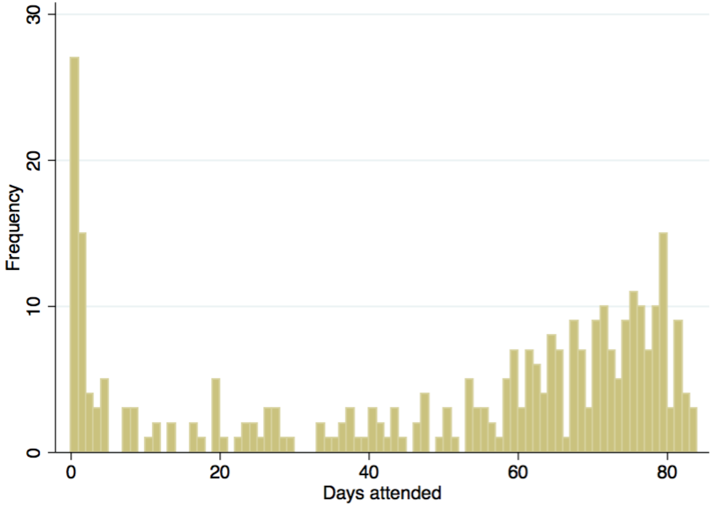
- Lottery-winners were offered free access to the program for ˜1/2 a school year
  - High take-up (among lottery participants) if defined as showing up ever
  - Variable take-up when looking at number of days actually attended

- The **intent-to-treat effect** is the reduced form effect of the lottery on the outcome

Intent-to-Treat Effects



Number of Days Attended (Max. Possible 86 Days)



What Do We want to Get at?

- For most purposes, the ITT itself is very informative here (and very big)

- Treating take-up as a binary decision could lead us to think of “scaling up” ITT
  - That is a minor adjustment since attending >1 day is almost 90%!
  - Note there are no “always-takers”, so $LATE = ATT$

- 58% attendance among lottery-winners over a total program duration of ˜86 days
  - Merely a function of how things panned out
  - Could have started earlier, or had higher/lower attendance

- What we want, if extrapolating to different intensities, is a dose-response function:
  - How does attending x days affect achievement?
  - E.g. imagine if we wanted to think how such program would work if delivered as summer vacation remedial instruction?

IV Estimates of Dose-Response: Specification

- Muralidharan et al. (2019) estimate:

$$Y_{is2} = \alpha + \gamma Y_{is1} + \mu_1 Attendance_i + \eta_{is2}$$

where Y_{ist} is student test score in subject s at time t and *Attendance* is the number of days a student logged in to the Mindspark system (which is zero for all lottery-losers)

- Since program attendance may be endogenous to expected gains from the program, they instrument for *Attendance* with the randomized offer of a voucher

Estimates of Dose-Response

	(1)	(2)	(3)	(4)	(5)	(6)
		Dep var: Standardized IRT scores (endline)				
	IV estimates		OLS VA (full sample)		OLS VA (Treatment group)	
VARIABLES	Math	Hindi	Math	Hindi	Math	Hindi
Attendance (days)	0.0065***	0.0040***	0.0068***	0.0037***	0.0075***	0.0033*
	(0.0011)	(0.0011)	(0.00087)	(0.00090)	(0.0018)	(0.0020)
Baseline score	0.53***	0.67***	0.54***	0.69***	0.57***	0.68***
	(0.036)	(0.037)	(0.039)	(0.039)	(0.062)	(0.056)
Constant			0.35***	0.16***	0.31***	0.18
			(0.040)	(0.042)	(0.12)	(0.13)
Observations	529	533	529	533	261	263
R-squared	0.422	0.460	0.413	0.468	0.413	0.429
Angrist-Pischke F-statistic for weak instrument	1238	1256				
Diff-in-Sargan statistic for exogeneity (p-value)	0.26	0.65				
Extrapolated estimates of 90 days' treatment (SD)	0.585	0.36	0.612	0.333	0.675	0.297

Note: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Estimates of Dose-Response: Interpretation

- IV estimates identify an "Average Causal Response" (Angrist and Imbens, 1995)
- "Captures a weighted average of causal responses to a unit change in treatment (in this case, an extra day of attendance), for those whose treatment status is affected by the instrument"
- Using IV estimates to predict the effect of varying treatment intensity requires further assumptions about
  - Heterogeneity in treatment effects—ACR is identified off compliers
  - Functional form of relationship between attendance and the treatment effect—ACR averages effects over different intensities of treatment
- Muralidharan et al. (2019) show evidence in our data in support of constant treatment effects across compliance groups and a linear dose-response relationship

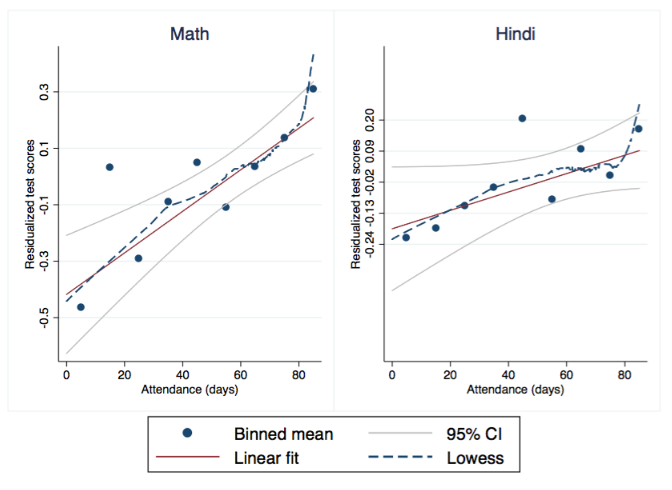
Evidence in Favor of Constant Treatment Effects

- ITT effects were constant over the full range of the baseline distribution of test scores
  - Typically a good summary measure of individual-specific heterogeneity
  - No significant evidence of heterogeneity on wealth and gender either

- Muralidharan et al. cannot reject equality of estimates between OLS and IV; suggests ATE and LATE are similar in this setting

- Constant term in OLS specifications (corresponding to zero attendance) is identical when estimated using full sample and when only using lottery-winners
  - This suggests equality of potential outcomes between "never-takers" and "compliers"
  - No "always-takers" in this setting

Dose-Response Relationship in Value-Added: Math and Hindi



Is LATE a Useful Parameter?

I find it hard to make any sense of the LATE. We are unlikely to learn much about the processes at work if we refuse to say anything about what determines j; heterogeneity is not a technical problem calling for an econometric solution, but is a reflection of the fact that we have not started on our proper business, which is trying to understand what is going on.

Angus Deaton (2011)

Is Heterogeneity of Outcomes First-Order or Not?

- The merit of LATE is that it makes explicit what is the most you can hope to identify in general without making any further restrictions on heterogeneity
- This is very important since these distinctions were not made in much of the traditional IV literature
- With constant causal effects, these concerns of interpretation are assumed away
- But in the general case, you should worry about how first-order the issues of heterogeneity are
- Are the compliers a particularly non-representative subsample that we should be worried about generalizing from them?

Is Heterogeneity of Outcomes First-Order or Not?

- There is no general formulation that allows you to see this
- One alternative is to take a more explicit theory-driven approach
  - E.g., the marginal treatment effects work done by Heckman and co-authors
  - This is more advanced than the level of this course and we will not be studying these
- But sometimes you can evaluate this indirectly even in reduced-form approaches
  - How heterogeneous are treatment effects in settings with different levels of compliance?
  - How heterogeneous are first-stage effects and treatment effects for different (identifiable) subgroups?
  - How heterogeneous are treatment effects by observables?

Summary: What We Covered Today

- IV estimators offer solutions for tackling noncompliance in randomized trials
  - The assigned treatment $\neq$ the delivered treatment
  - The effect of assigning the treatment is the **intention-to-treat effect**
  - The (randomly) assigned treatment may be used as an IV for actually delivered treatment

- Interpreting IV estimates with heterogeneous potential outcomes is tricky
  - Assuming independence, exclusion, relevance and **monotonicity,** the IV estimates can be interpreted as estimating a **Local Average Treatment Effect**
  - This parameter is estimated only over the **compliers**
  - It is not (necessarily) the ATT, the ATE, or the ATU!

Readings

- Angrist and Pischke (2009) Chapter 4, from Section 4.4 up till Section 4.5.1 (p. 150-175)
- Angrist, J. D. (2006). Instrumental variables methods in experimental criminological research: what, why and how. Journal of Experimental Criminology, 2(1), 23-44.
- Cunningham (2009) Chapter on IV, section on Heterogeneous Potential Outcomes, p. 232-243
- Watch **this 4 minute video**!

Readings: Is LATE a Useful Parameter?

Recommended readings:

- Deaton, A. (2010). Instruments, randomization, and learning about development. Journal of Economic Literature, 48(2), 424-455.
- Imbens, G. W. (2010). Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). Journal of Economic Literature, 48(2), 399-423.

Additional readings:

- Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. The American Economic Review, 96(1), 152-175.
- Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India, American Economic Review.