PHS 2000 Econometrics
Instrumental Variables

Maggie McConnell

Department of Global Health and Population

Harvard T. H. Chan School of Public Health

Plan of presentation

1. Correlation between errors and explanatory variables
2. Example
3. Why instruments work
4. Instrumental Variable estimation
5. Required assumptions for Instrumental Variable estimation
6. Testing the required assumptions
7. Local Average Treatment Effect

My office hours will be today 1:10-145pm in my office (Building 1, room 1217) or we can chat directly after class

Unbiasedness of Ordinary Least Squares

$$Y = X\beta + \varepsilon$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

$$\hat{\beta} = (X^T X)^{-1} X^T (X\beta + \varepsilon)$$

$$\hat{\beta} = (X^T X)^{-1} X^T X\beta + (X^T X)^{-1} X^T \varepsilon$$

$E(\hat{\beta}_{OLS}) = \beta$ *since* $(X^T X)^{-1}(X^T X)\beta = \beta$

$E(X^T \varepsilon) = 0$

Reasons OLS Fails

- OLS estimation will yield biased and inconsistent results if one of the regressors is not orthogonal with the error term (endogeneity)

- For example if we have:
  - **Omitted variable bias:** We omit a relevant variable which is correlated with one or several of the included variables

- This will induce correlation between the X variable of interest and the error term
  - Economists might say – X is endogenous

Instrumental Variables: Intuition

- Let's assume we care about the relationship between health care utilization and housing insecurity
- Suppose we observed lower rates of primary care utilization and higher rates of ED use among individuals reporting housing insecurity
- What kind of confounding would you worry about?
    - One possibility is that there is an omitted variable that is strongly associated with both health care utilization and housing insecurity
        - Income, neighborhood / community access issues
- Is this problem solved with longitudinal data? Suppose we saw reductions in primary care utilization and increases in ED use following eviction?
    - There could still be omitted variables that predate both changes: mental health crises, experience of intimate partner violence, challenges with substances

Instrumental variables: Intuition

- Many reasons for a relationship between health care utilization and housing insecurity
  - Valuable to describe this relationship to better understand challenges, drive priority setting, target resources, etc
- For *causal inference* – we want to focus on a narrow question
  - Does addressing housing insecurity *cause* changes in health care utilization?
  - State Medicaid program wants to fund programs that provide supportive housing but they want to be confident this is a *causal* relationship
    - They also want to know how large this effect might be to figure out how much to spend
- We want to find something that creates variation in housing insecurity that is totally unrelated to income (or other determinants of health care utilization)

Instrumental variables: Intuition

- It's helpful to think of instruments as things that are "as good as randomly assigned", or quasi-random
- IVs often have a plausibly random component, like whether or not it rains on a certain day, whether you have twins or not, which case officer is assigned your case, what month you were born in, etc.
- One example of an instrumental variable is a treatment that is randomly assigned
  - Sometimes RCTs get analyzed using IV techniques (more later)

Instrumental variables

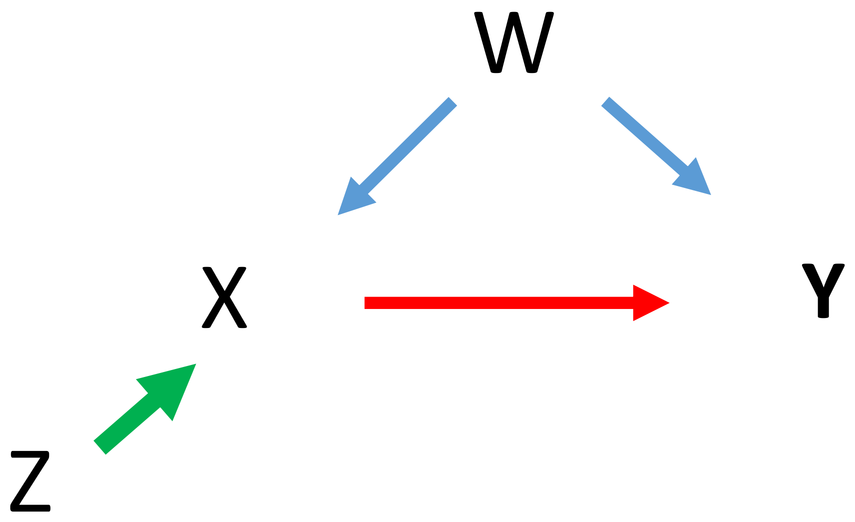
- IV can be used to estimate causal effects when we worry about selection, in the presence of measurement error and as a tool for analyzing experimental data

- IV can be useful but it can be hard to find a good instrument
- IV with an invalid instrument doesn’t help you very much
- Today we will cover
  - Understanding what a good instrument is
  - Knowing what to do with a good instrument
  - Knowing how to interpret IV results

Example: Adherence Adjustment

- Suppose we have a randomized trial that allocates people randomly to treatment and control groups.
- However "adherence" to random assignment not perfect, why?
    - Sometimes people are missed, drop out, refuse treatment
    - Some people in the control arm may get the treatment from other sources
- Intention to treat analysis: comparing those *assigned* to treatment and those *assigned* to control
    - Sometimes this *is* the policy relevant coefficient to estimate, why?
- We may also want to know about the average effect of the treatment on those who actually *received* the treatment
    - i.e. in settings where we want to decide if it is cost effective to scale up

Adherence adjustment - Pathways



- Random intervention Z (offer of supportive housing)
- Housing insecurity X
- Health care utilization Y
- Omitted variable W (income)

- OLS estimate of X on Y biased due to correlation with W

Instrumental Variables: Why it Works

- Assume we are in a setting where

$$E(X'\epsilon) \neq 0$$

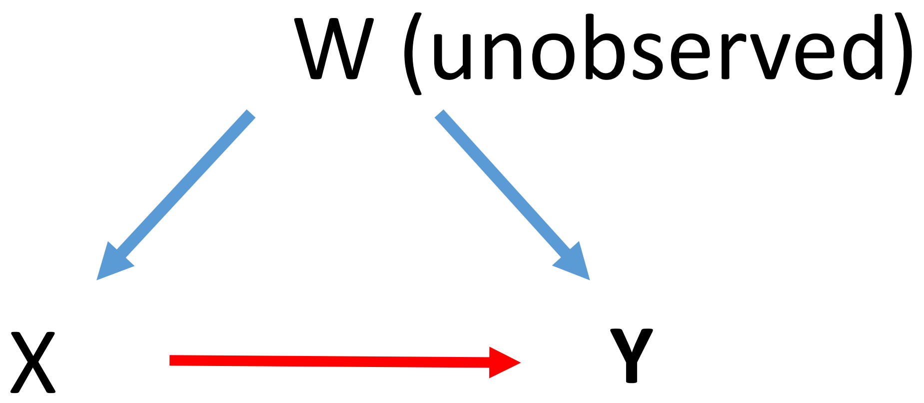
- Further assume that there is some variable Z (which we call the "instrument") that satisfies the following two conditions

1. Z is correlated with X: $cov(x, z) \neq 0$
2. Z is orthogonal to the error: $cov(z, \epsilon) = 0$

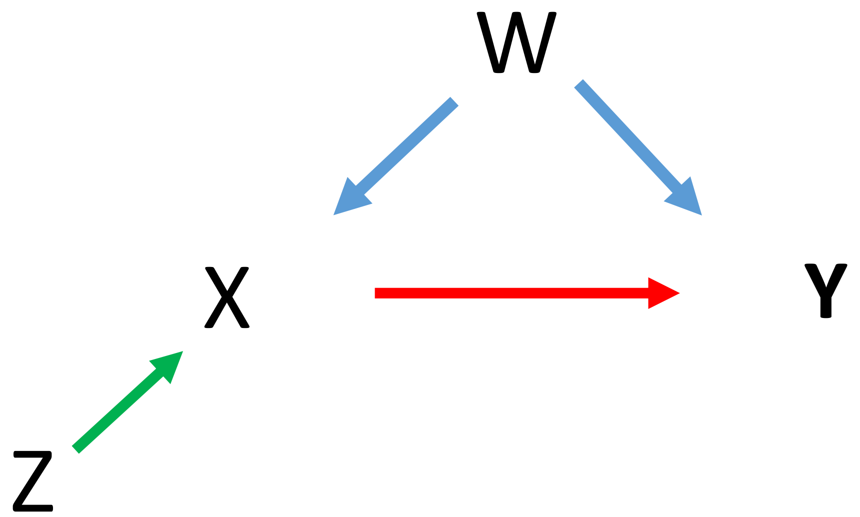
OLS setup

$$X \longrightarrow \mathbf{Y}$$

Setup with a violation of OLS assumptions



IV setup



IV estimation in the single instrument case

- Suppose the model we want to estimate is:

$$y_i = \alpha + \beta x_i + \epsilon_i \quad [1]$$

- However, suppose that in this model, $cov(x, \epsilon) \neq 0$

- But, assume that we have identified an instrument (Z) such that

$$x_i = \gamma z_i + u_i \quad [2]$$

with $cov(z, u) = 0, and\ cov(z, \epsilon) = 0$

IV estimation using Indirect Least Squares

- Since $cov(z, u) = 0$, we can estimate $x_i = \gamma z_i + u_i$ using OLS to get an unbiased and consistent estimator $\hat{\gamma}$

- Now, substituting $x_i = \gamma z_i + u_i$ in $y_i = \alpha + \beta x_i + \epsilon_i$, we get

$$y_i = \alpha + \beta(\gamma z_i + u_i) + \epsilon_i$$

$$y_i = \alpha + (\beta\gamma) z_i + (\beta u_i + \epsilon_i)$$

with $cov(z, \beta u + \epsilon) = 0$

- We can therefore use OLS to get an unbiased and consistent estimator $\widehat{\beta\gamma}$ by regressing the outcome y on the instrument z
- We can then consistently estimate $\hat{\beta} = \frac{\widehat{\beta\gamma}}{\hat{\gamma}}$ as the ratio of the coefficients estimated in the two regressions
- Pause for questions

Housing-Health example

- Estimate $\hat{\gamma} \; = 0.26$ effect of intervention on housing insecurity
- Estimate $\widehat{\beta\gamma} = 0.04$ effect of intervention on adherence with follow-up care after surgery

- Both these estimates are consistent since the assumption of OLS are satisfied by randomization of the intervention

- Hence estimate $\hat{\beta} = \frac{\widehat{\beta\gamma}}{\hat{\gamma}}$ = $\frac{0.04}{0.26} = 0.15$ effect of housing insecurity on adherence to follow-up care after surgery

IV estimation using Two Stage Least Squares

- First stage: Estimate

$$x_i = \gamma z_i + u_i$$

with $cov(z, u) = 0, cov(z, \epsilon) = 0$

- Second stage: Get predictions from the first stage and estimate

$$y_i = \alpha + \beta \widehat{x_i} + \epsilon_i'; \quad \widehat{x_i} = \hat{\gamma} z_i$$

with $cov(\hat{x}, \epsilon') = cov(\hat{\gamma} z, \epsilon') = 0$

- Same numerical estimate as Indirect Least Squares with one instrument
- More efficient with multiple instruments in first stage if we can find several instruments.
- Can apply to the case where we want to instrument several variables at the same time (we must have at least as many instruments as endogenous variables)

Pause for questions

When does this work?

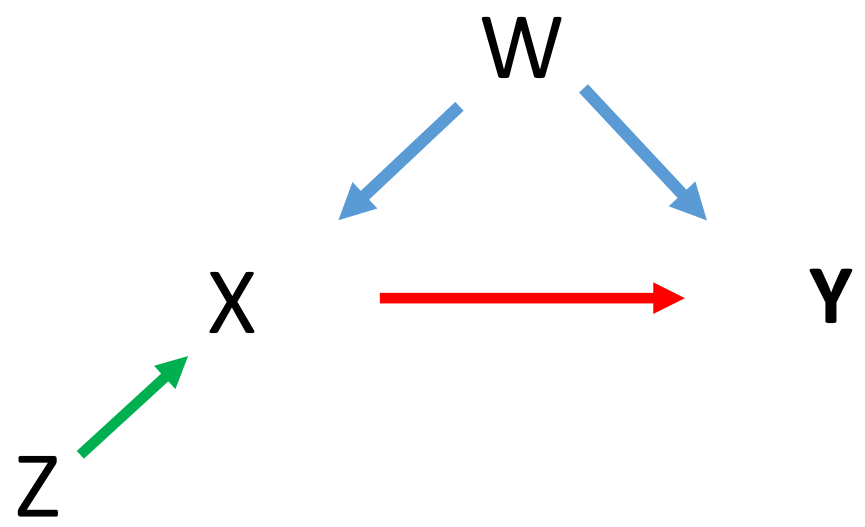
- There are FOUR key assumptions for IV inference to be valid:

1. **Relevance:** The instrument Z must be correlated with the variable of interest X
2. **Excludability:** The instrument Z must NOT have a direct causal effect on the outcome Y
3. **Independence:** The instrument cannot be correlated with any other unobserved determinant W of the outcome Y
4. **Homogeneity:** The treatment effect is homogenous across all study subjects

In economics usually stated differently

- We are estimating a linear regression (homogeneity)
- The explanatory variable is correlated with the instrument (relevance)
- The instrument is uncorrelated with both the error terms in the first second stage regression (excludability and independence)
- Also stated as the exclusion restriction: the instrument affects the outcome *only* through the instrument

Condition 1: Relevance   Z -> X



Note: we only need correlation not causality

Condition 1: Relevance

- The relevance assumption is relatively easy to both satisfy and verify
- In most contexts, other factors are correlated with the variable of interest X
- The “strength” of the instrument is determined by its predictive power in a multivariate setting
- A good instrument explains a significance part of the variation in the endogenous variable *once we control for all other factors in our empirical model*
- *This is testable*

Two Stage Least Squares: validating the relevance condition

- To validate the **relevance condition**:
  - Look at the first stage predictive power
- If predictive power is low, then we have a “weak instrument”, i.e., an instrument which is weakly correlated with the endogenous variable of interest

- Recall that:

$$\hat{\beta} = \frac{\widehat{\beta\gamma}}{\hat{\gamma}}$$

- If the first stage is weak, then $\hat{\gamma} \approx 0$ which means that the 2SLS estimate explodes! Violates boundedness assumptions for valid standard errors

IV estimation in practice using Two Stage Least Squares

- In multivariate models, we may have several endogenous variables and instrumental variables and some covariates

- IV estimation is done in two steps:
  - **Step 1:** regress the endogenous variables and covariates on the instrument and predict the “exogenous” part of the variation for these variables
  - **Step 2:** regress the dependent variable on the predicted values from the first stage to get the (ideally) unbiased IV estimator
  - *Note:* manually fitting the two stages will result in estimating incorrect standard errors since fitted values in second stage depend on first stage parameters that have estimation error
- Since the two stage approach is fairly standard, IV estimation is also referred to as Two Stage Least Squares estimation (2SLS, or TSLS)

Two Stage Least Squares: validating the relevance condition

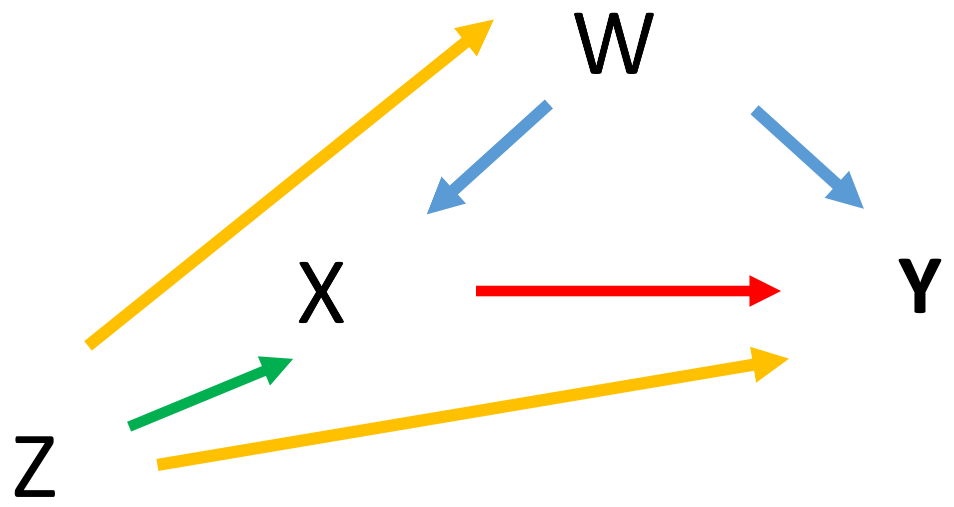
- Check check the predictive power of the first stage:
  - If the first stage estimate is possibly close to zero the instrument is weak
  - Even if we reject zero at 95% confidence, there is a 5% chance first stage is actually zero (or different sign) – too much!
  - This means that statistical significance of instruments ($t>1.96$) in the first stage is not enough!
  - Rule of thumb: If the F-statistic is greater than 10, then the instrument is not weak. If you have only one instrument, the F-statistic is the square of the t-statistic ($t> 3.16$)
  - Formally - use the Cragg-Donald statistic for critical values F-test of the joint significance of the instruments in predicting the endogenous variable
  - STATA IVREG2 reports Cragg –Donald statistic and critical values for first stage

Pause for questions

Conditions 2/3: Excludability/Independence

- The excludability/independence conditions are more complex since orthogonality to (i.e., being uncorrelated with) the error requires that:
  - The instrument does not have a direct effect on the dependent variable
  - The instrument is not correlated with other omitted variables that have a direct effect
  - The instrument is not correlated with measurement error in X
  - The instrument is not “caused” by the dependent variable

Excludability/Independence



- If the yellow lines are pathways the instrument is invalid
- Did all the effect of the supportive housing intervention go through housing insecurity? Where their other pathways from the intervention to the health care utilization outcome? Example?
- Maybe you got into supportive housing and found out about a free medical clinic that offered more sensitive care from a new contact living next door

Randomization and IV

- Randomization of Z helps rule out correlation with W and reverse causality from Y in adherence adjustment
- Still issue of other pathways Z to Y except through X
- IV analysis essentially assumes that all of the effect of Z on Y is via its effect on X – but if this is not the case our adjusted estimate will be invalid

Conducting an IV analysis in observational studies

- Suppose we want to estimate the effect of education on health in a cross-section of adults:
  - Research question: How much will adult health (eg probability of surviving to age x) improve by if we increase a person's schooling by one year?

- We are worried that simple regression of survival on years of schooling may not capture the true effect
  - Omitted variable bias – variables that affect both schooling and health
  - Examples?

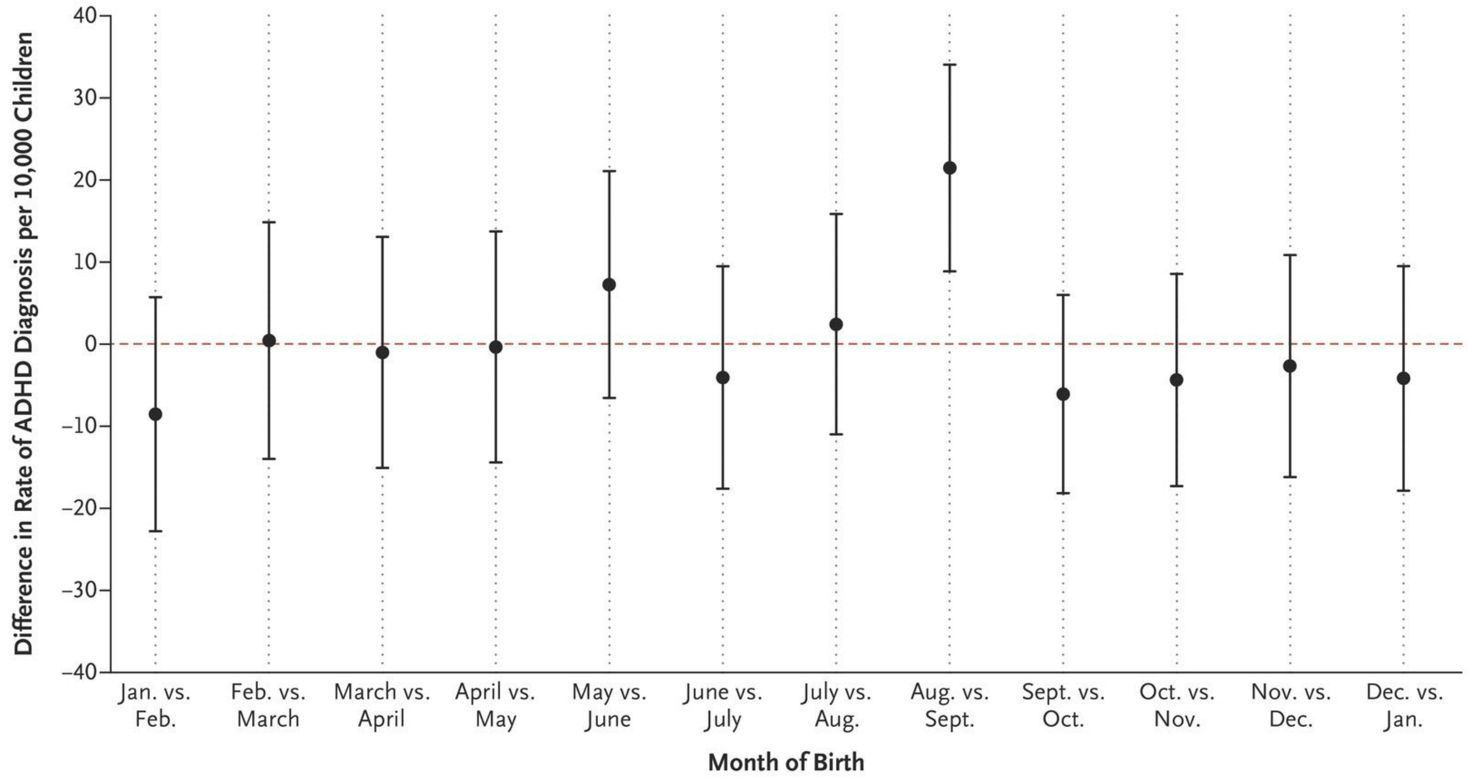
Education and Health

- Discuss the validity of the following potential instruments:
- Parents' education
- State mandatory schooling laws when the person was of school age
- Distance to a four-year college at age 18
- Randomization to a head start program (free pre-school program below age 5)
- Child's month of birth

Instrument example – month of birth

- For many research questions, month of birth is used as an instrument
- Examples
  - Month of birth as instrument for how much schooling you will get
  - Month of birth as instrument for being youngest vs oldest in class
  - Month of birth as instrument for being born before or after development of new vaccine

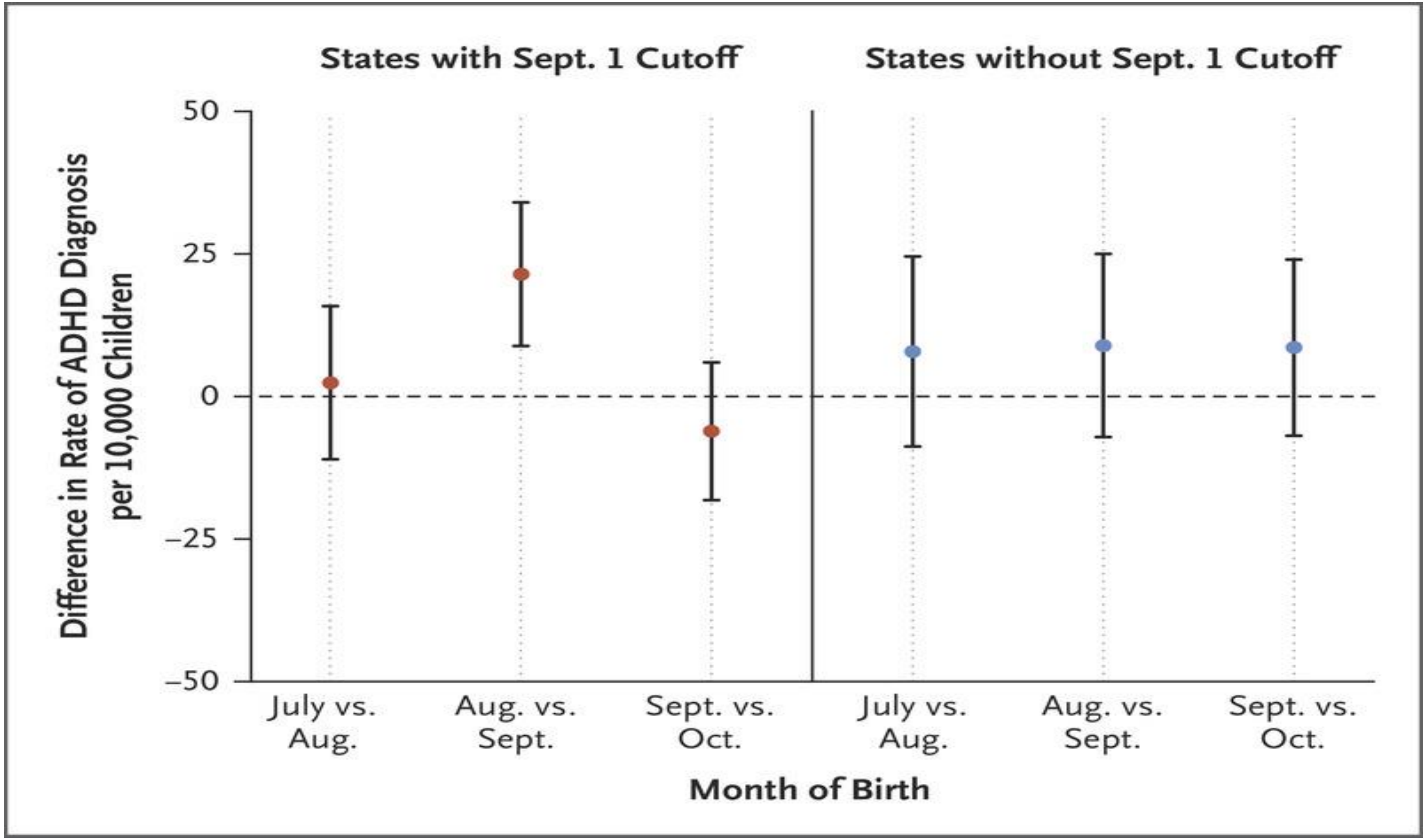
Differences in Diagnosis Rates of ADHD



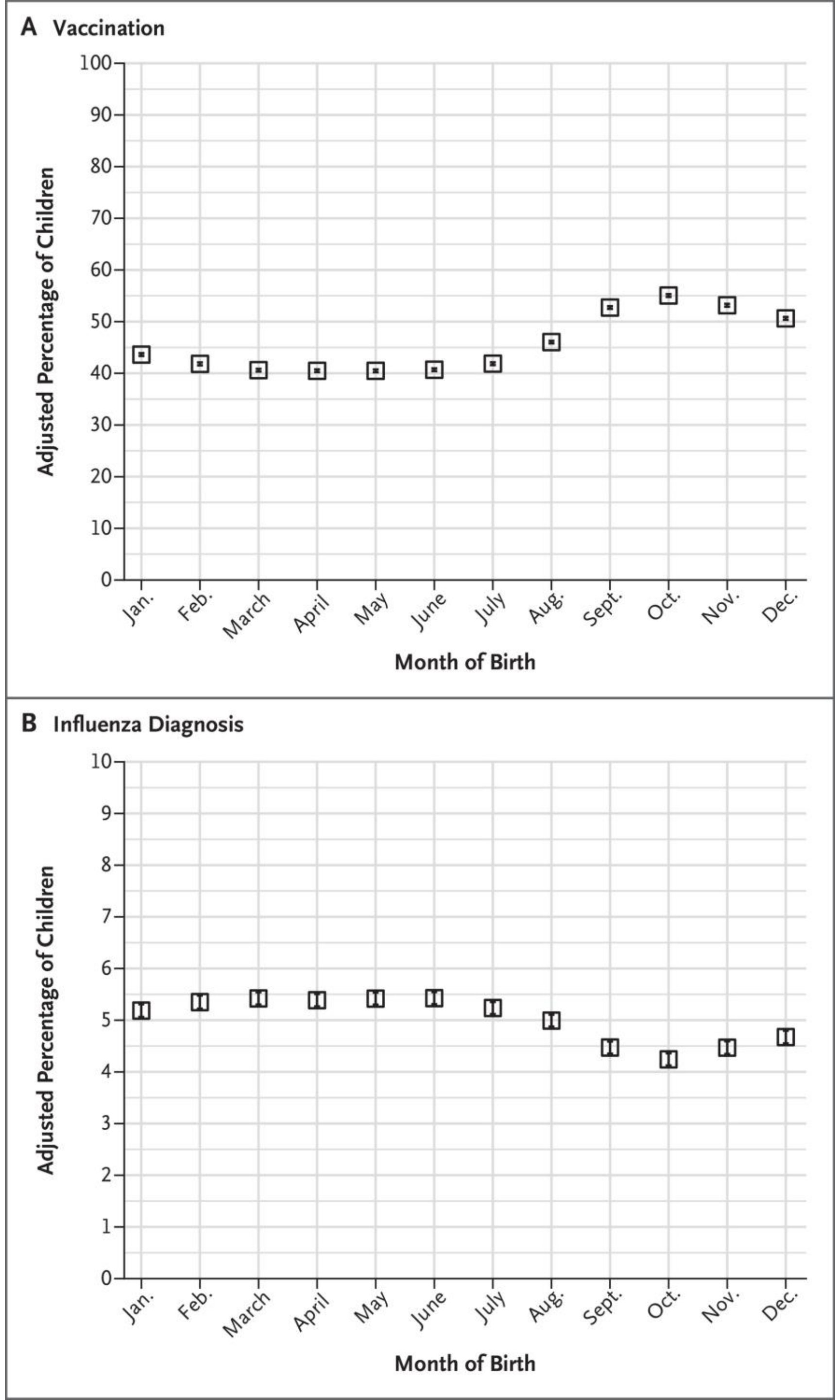
	January	February	March	April	May	June	July	August	September	October	November	December
Total no. of children	32,690	31,238	34,405	34,565	34,977	34,415	36,577	36,319	35,353	34,405	31,285	31,617
No. of children with ADHD	265	280	307	312	287	317	320	309	225	240	232	243
Rate per 10,000 children	81.1	89.6	89.2	90.3	90.6	83.4	87.5	85.1	63.6	69.8	74.2	76.9

[Attention Deficit–Hyperactivity Disorder and Month of School Enrollment]

Differences in Diagnosis Rates of ADHD



Attention Deficit–Hyperactivity Disorder and Month of School Enrollment



[Birth Month and Influenza Vaccination in Children]

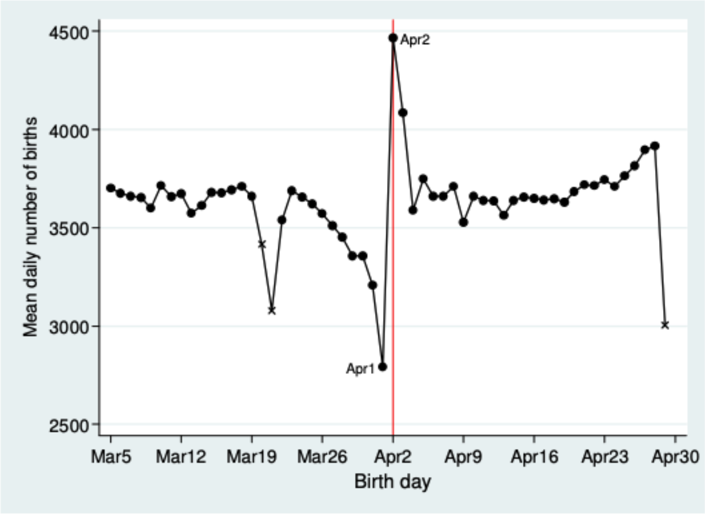
Instrument example – month of birth

- For many research questions, month of birth is used as an instrument

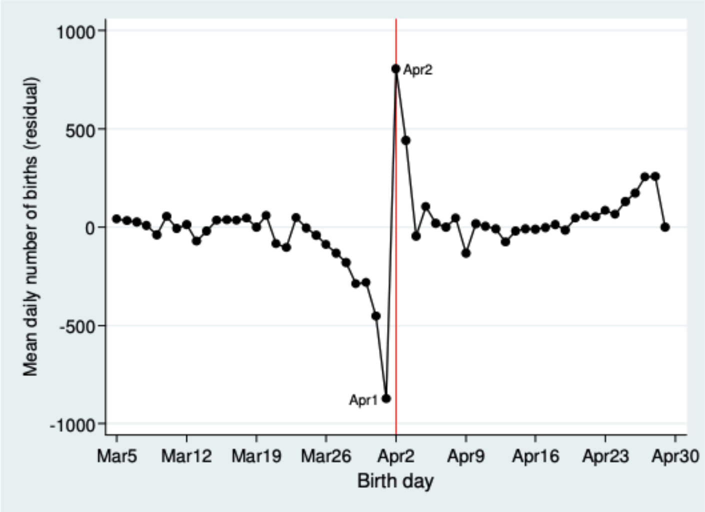
- Is birth month a good instrumental variable?
  - Some evidence that births in US timed around tax advantages
    - New Evidence on Taxes and the Timing of Birth
  - Some evidence that the characteristics of families with births at different times of year may differ
    - Season of Births and Later Outcomes, Old Questions, New Answers
  - Evidence from some settings that parents may be aware that older children have advantages and time births based on school cutoff

Figure 2: Mean daily number of births around April 2

A. Raw data



B. Adjusted

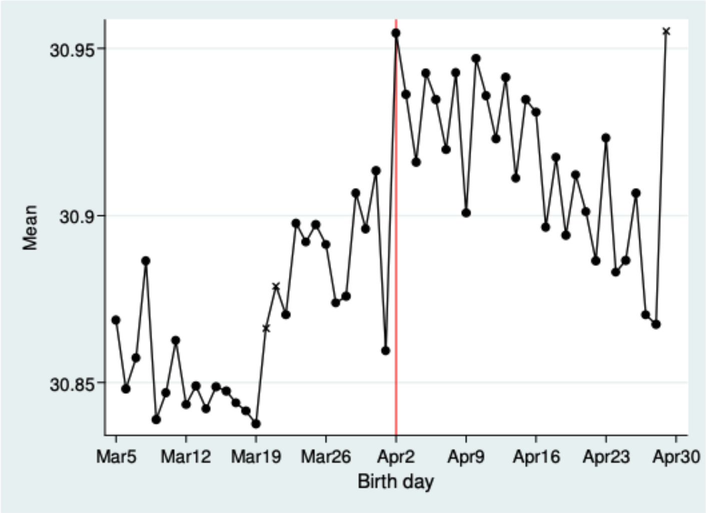


Note: Each plot is the mean daily number of births. The vertical line corresponds to April 2, which is the school entry cutoff date in Japan. The markers with cross signs in Panel A are holidays (March 20 or March 21 depending on the year, and April 29). Panel B adjusts for holidays, day of week, and year fixed effects. The data come from pooled 1974–2010 birth data.

30

Figure 4: Child Outcomes

A. Birth weight (100 grams)



[School Entry Cutoff Date and the Timing of Births]

Heterogeneity of effect and the Local Average Treatment Effect (LATE)

- So far, we have assumed homogenous treatment effect β; but, what if it is heterogeneous?
- In most settings, the instrument may only affect a specific sub-population
- Compulsory schooling to age 16 – affects mainly people who would otherwise leave school before age 16
  - Sometimes called “compliers”
- In such a situation, IV estimation identifies the causal effect of X on Y ONLY in this sub-population
  - The treatment effect on this specific sub-population is known as the Local Average Treatment Effect (LATE). LATE may be different from the Average Treatment Effect (ATE)
  - LATE will converge to ATE with very good adherence

Heterogeneity of effect and the Local Average Treatment Effect (LATE)

- Relevant heterogeneity might be as follows:
  - **Compliers:** only take the treatment if assigned to treatment group
  - **Always-takers:** always take the treatment even if assigned to the control group
  - **Never-takers:** always refuse treatment, even if assigned to the treatment group
  - **Defiers:** do the opposite of their treatment assignment

If any of the last 3 are meaningfully different from the compliers and they are populations of interest for policy-makers → LATE estimates not very helpful

Assumptions for IV estimation of the LATE

- In addition to the first three IV assumptions, we need an extra assumption to for LATE to be well-identified for the local population of compliers:

1. **Relevance:** The instrument must be correlated with the variable of interest
2. **Excludability:** The instrument must NOT have a direct causal effect on the outcome Y
3. **Independence:** The instrument cannot be correlated with any other unobserved determinant of the outcome Y
4. **Monotonicity:** Assignment to a treatment unit does not make anyone *less* a member of their treatment unit than they are supposed to be
   - When treatment is binary this means no defiers

- Pause for questions

Special case of LATE: Treatment on the Treated (TOT)

- Treatment on the Treated (TOT): Effect of treatment on those who receive the program
- Can only be estimated when no one in the control group is treated
  - Such as for example in a randomized-controlled trial
  - No “always takers”

IV coefficients versus OLS coefficients

- IV estimation can be used to overcome omitted variable bias, measurement error, and reverse causality
- Depending on which concern is larger, IV estimates can be larger or smaller than (the biased) OLS coefficients
- In many published papers, you will find IV estimates that are <u>much</u> larger than OLS estimates
  - This could be due to measurement error and attenuation bias
  - This could also reflect a violation of the excludability condition or the use of weak instruments

Examples of instruments

- Research question: [What is the effect of placement in foster care on subsequent use of emergency department care?]
  - Instrument: assignment of placement officer

- Research question: [What is the relationship between the number of children a family has and the amount of education children receive]?
  - Instrument: indicator of whether first child is male in a context with strong son preference

- Research question: [What is the impact of requiring workers to be physically present on COVID-19 death rates?]
  - Instrument: index of how many local workers pre-pandemic can work from home, based on differences in county occupational mix

Credibility of IV estimation in the economics literature

- Many older papers that used IV estimation now seem invalid
- A lot of concern about weak instruments
- Most challenging issue is exclusion/independence assumptions
- Difficult to publish an IV analysis today unless the instrument is very convincing

Pause for questions

Additional Slides

Readings

Wooldridge, *Introductory Econometrics*, Chapter 15

Hernán, Miguel A., and James M. Robins. "Instruments for causal inference: an epidemiologist's dream?." *Epidemiology* (2006): 360-372.

Murray, Michael P. "Avoiding invalid instruments and coping with weak instruments." *Journal of Economic Perspectives* 20.4 (2006): 111-132.

Two Stage Least Squares: validating the excludability condition

- If there is only one instrument, there is no test of the excludability condition
  - Validity rests on a priori assumptions

- With multiple instruments, one could use the Sargan-Hansen “over-identifying restriction” test
  - Works under the assumption that you have at least one valid instrument, which is generally not testable!

- Sargan-Hansen idea - Instruments should have no predictive power on the outcome Y once we control for predicted X in the main regression – they only work through X
  - So, take residuals from the second stage and regress on instruments. You should see no predictive power of the instruments
  - note that for one instrument the instrument has no predictive power on 2nd stage residuals by construction

Pause for questions

Example of an IV analysis: Mendelian randomization

- One of the most exciting data developments in recent years has been the increasing availability of genome data

- Genetic traits of genotypes are in principle excellent candidates for IV estimation as:
    - They are, by construction, randomly assigned (“mendelian randomized” within families)
    - They often show robust associations with traits or behaviors of interest

- Despite this, the mendelian randomization IV estimates may not yield unbiased estimates in practice because…

Mendelian randomization concerns

- The presence of certain genotypes may be correlated with other genotypes unobserved family traits
- Eg genotype for education identified by GWAS– use as IV?
- Chromosomes split means genes randomized in groups not individually – genes for education correlated with other genes
- Genes for education may also affect other phenotypes eg patience
- Both of these are violations of the second validity criterion of IV estimation and lead to bias

- Some helpful references: Thomas & Conti (2004), Glymour et. al. (2012)

Example: Effect of intensive care on mortality

- Estimating the effect of health care on health outcomes is difficult
- Treatment is given to the most at risk but these may also have worse outcomes
- We can control for initial observed health but unmeasured confounding can still exist
- Intensive care patients have much higher mortality than those not in intensive care
  - Correct for age, health condition, health status, referral to intensive care, but still higher mortality – intensive care is associated with high mortality but is intensive care really bad for patients?
- Harris, Steve, et al. "Impact on mortality of prompt admission to critical care for deteriorating ward patients: an instrumental variable analysis using critical care bed strain." *Intensive care medicine* 44.5 (2018): 606-615.
  - Outcome: 7 / 90 day mortality – binary variable
  - Endogenous variable: admission to ICU within 4 hours of referral
  - Instrumental variable: number of free ICU beds in hospital (0, 1, 2+) at time of referral

Effects of strain on the admission pathway

	Critical care beds			Test for trend
	≤ 0	1	≥ 2	*p* value
Patients referred	988 (8.0%)	1353 (10.9%)	10,039 (81.1%)	
Critical care				
Recommended	354 (35.8%)	471 (34.8%)	3735 (37.2%)	0.1407
Admitted	247 (25.0%)	425 (31.4%)	3775 (37.6%)	< 0.0001
Prompt admission	87 (8.8%)	196 (14.5%)	2128 (21.2%)	< 0.0001
Death without critical care	92 (9.3%)	79 (5.8%)	614 (6.1%)	0.0002
Time to critical care, hours	5.0 (2.2–15.8)	4.0 (1.0–12.0)	3.0 (1.0–8.0)	0.0009
ICNARC physiology score				
At referral	15.2 (7.1)	15.1 (7.2)	15.2 (7.2)	0.8266
Change between referral and admission	4.5 (9.2)	3.3 (9.1)	3.1 (9.2)	0.0301
Mortality				
7-day	147 (14.9%)	179 (13.2%)	1391 (13.9%)	0.6226
90-day	312 (31.6%)	417 (30.8%)	3007 (30.0%)	0.2326

Instrumental variable model for the effect of prompt admission on 90-day mortality among those recommended for critical care

	Odds of Mortality	
	Odds ratio	*p* value
Critical care recommended level 0-2	Reference	
Critical Care level Recommended level 3	1.53 (1.17–2.02)	0.002
Prompt admission (within 4 h)	0.46 (0.22–0.96)	0.036

Prompt admission instrumented with number of critical care beds available