Instrumental Variables

Dr. Katrina Kosec
Senior Research Fellow
IFPRI, Washington, DC, USA

Regional Training Course on Applied Econometric Analysis

June 11, 2018

WIUT, Tashkent, Uzbekistan

Readings

- Angrist, J.D. and J.-S. Pischke (2009). Chapter 4 "Instrumental variables in action: Sometimes you get what you need" *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press
- Gertler, P. J.; Martinez, S., Premand, P., Rawlings, L. B. and Christel M. J. Vermeersch, 2010, *Impact Evaluation in Practice*. Washington, DC: The World Bank.
- Wooldridge, J. *Introductory Econometrics* (4th ed). "Chapter 15: Instrumental variables and two stage least squares"
- Hansen, C. "Introduction to Instrumental Variables Methods." Booth School of Business: University of Chicago.

Outline

- What is endogeneity
- Introduction to IV as a way to address endogeneity
- What is an instrument and where do we find it?
- IV in an experimental setting: A job training program
- Mechanics of two-stage least squares
- Test for Endogeneity: Durbin-Wu-Hausman Test (DWH)
- Testing Overidentifying Restrictions
- IV in an observational setting:
  - Example 1: Returns to schooling
  - Example 2: Teenage work and deviant behavior

Recap: What is Endogeneity?

- Broadly refers to situations in which an explanatory variable is correlated with the error term
- Three main sources of endogeneity:
  1) Reverse causality, or “simultaneity”
  2) Omitted variables
  3) Measurement error
- All endogeneity sources will bias the coefficient on the affected RHS variable, and potentially any other variables that are correlated with the endogenous variable
- It is thus crucial to determine whether your model may suffer from one of these endogeneity issues

Reverse Causality, or “Simultaneity”

- In general, if you are unsure whether X causes Y or Y causes X, your regression suffers from simultaneity bias (i.e., because X and Y are determined together)
- If this happens, it is impossible to use OLS to determine whether and by how much a change in X will affect Y

Omitted Variables

- There are many reasons to be parsimonious about what control variables you use
- But there are good reasons to include controls, also:
  - It improves the predictive power of your model and thus the precision of your estimates
  - Including them amounts to moving unobservable variables from the error term to the specification itself, reducing the likelihood that any explanatory variable is correlated with the error term
- Excluding relevant variables can bias the coefficients on the included variables—known as "omitted variable bias" (OVB)
  - OVB appears when an explanatory variable must do "double duty"; its coefficient captures both its direct effect on Y as well as its indirect effect through some omitted variable that is correlated with it
- Fortunately, omitting explanatory variables correlated with Y results in biased coefficients *only* if the omitted variables are correlated with included variables

Measurement Error

- Data are often measured with error (e.g. reporting errors, coding errors)
- When the measure error is in the independent variable, endogeneity arises (our estimate of the effect of X is biased toward zero – also known as "attenuation bias")
- Most of the time, you should not be overly concerned about attenuation bias
  - It is inevitable that you will measure some predictor variables with error
  - If the measurement errors are relatively small, the bias is small as well
  - Moreover, this biases you against finding an effect (i.e. not much risk of it leading you to think a variable X has an effect on Y when it does not)

Which Variables are Endogenous and Which are Exogenous?

- In reality: In a non-experimental setting, almost all variable are likely to be endogenous
- Further, even if you have many exogenous variable in a regression, including a single endogenous variable can bias the coefficients on *all* independent variables
- For any variable for which we are not *explicitly addressing* its endogeneity (e.g., by instrumenting for it) we are *explicitly assuming* it is exogenous
  - This assumption may be objectionable to some! Hence why people often like to show that their results are robust to inclusion or exclusion of a set of control variables!
- Context matters: assess whether endogeneity exists, you need to know what is the outcome, what are the other independent variables in the equation
  - E.g., a variable may be less likely to be endogenous if village fixed effects are included

Which Variables are Endogenous and Which are Exogenous? (continued…)

- You will usually have one independent (explanatory) variables you care about most; addressing endogeneity of it is of prime interest
  - (E.g., through experimental methods, difference-in-differences, fixed effects, instrumental variables, RDD, etc)
- After doing so, if you wish to include control variables (e.g., to increase precision of your estimates), show that your results do not disappear if you omit them
  - Nothing makes a control variable more objectionable than if results vanish without it!
  - Note: year and geographic fixed effects are (sometimes) a valid exception; they can be critical to addressing endogeneity of your main independent variable

Instrumental Variables (IV): One Way to Address Endogeneity

- The method of instrumental variables (IV) can help to address endogeneity and establish causality, even when using observational data
- An “instrument” (also known as an “instrumental variable,” or an “excluded instrument”) is used to provide consistent estimates of the effect some endogenous independent variable X has on outcome Y
- Intuitively, when you have an endogeneity problem, you want to separate out that part of an endogenous variable X that is correlated with the error term
    - Once that part is separated out, you can get an unbiased causal estimate of the effect of the “uncorrelated portion” of X on Y

What is An Instrument?

- Consider the following model:

$$y = \alpha + \beta_1 x + \varepsilon$$

- An instrument is simply some variable z that affects x, but which only influences y through its effect on x
- Another way of saying this is that z is highly correlated with x, but uncorrelated with the error term, ε
- Very important:
  - You can PROVE that z is highly correlated with x, and you *must* do so (you will present a first stage regression as evidence)
  - You cannot PROVE that z is uncorrelated with ε; you instead argue this, and can provide suggestive empirical evidence (e.g., test of overidentifying restrictions)
- Note: There's no unique instrument; many variables may satisfy these properties!

The Inclusion and the Exclusion Restriction

- Common terminology for an IV being 'valid' is that it satisfies the inclusion and the exclusion restrictions:
- Inclusion restriction: instrument is "sufficiently" correlated with the endogenous variable
  - If the correlation is low, you have problems of "weak instruments"; fortunately this is something you can test for (e.g., Stata gives critical values following use of ivreg2)
  - See Stock and Yogo (2005), "Testing for Weak Instruments in Linear IV Regression" (http://mayoral.iae-csic.org/IV_2015/stock_yogo_2005.pdf)
- Exclusion restriction: instrument is uncorrelated with the error term
  - Largely argued, but you can do test of overidentifying restrictions (more later)

Where Do We Get an Instrument?

- IV is useful not only in observational studies, but also in experiments
- An instrument can be generated ex-ante:
  - Encouragement design (randomize provision of encouragement to get treatment)
  - "Randomized offering" of a program (offer treatment to some and not others)
- An instrument can be used ex-post, either to correct for non-compliance in a randomized experiment or deal with endogeneity in observational data
  - When you randomize Treatment (T) and have some non-compliance, instrument for receiving treatment with having been assigned to treatment (thus estimate "the effect of treatment on the treated" (TOT) rather than "intent to treat" (ITT) estimates
  - Fuzzy Regression Discontinuity Designs (more on this tomorrow)
  - Identify some other exogenous source of variation in an endogenous variable

Number of Endogenous Variables and Number of Instruments

- We first focus on the simple case where there is one endogenous explanatory variable and one instrument
- We then expand to the case of multiple endogenous variables and/or multiple instruments
- Note: you always need at least as many instruments as you have endogenous variables
    - # instruments = # endogenous variables: “exactly identified”
    - # instruments > # endogenous variables: “over-identified” (some benefits for efficiency and ability to test for overidentifying restrictions to help you argue that an instrument is exogenous)

Example: A Voluntary Job Training Program

- Say we wish to evaluate a voluntary job training program
    - Any unemployed person is eligible (Universal eligibility)
    - Some people choose to register (Participants)
    - Other people choose not to register (Non-participants)

- Some simple (but not-so-good) ways to evaluate the program:
    - Compare before and after situation in the participant group
    - Compare situation of participants and non-participants after the intervention
    - Compare situation of participants and non-participants before and after (difference in differences)

Example: A Voluntary Job Training Program

Say we decide to compare outcomes, y (e.g., income later in life) for those who participate and those who do not:

- A simple model to do this:

$$y = \alpha + \beta_1 T + \beta_2 x + \varepsilon$$

$$T = \begin{cases} 1 & \text{If person receives training} \\ 0 & \text{If person does not receive training} \end{cases}$$

x = Control variables (exogenous & observed)

- What's wrong with this? Two main problems:
  - o The decision to participate in training is endogenous – that is, it is correlated with factors that affect our outcome y
  - o Even if we add controls, we may still be omitting some variables that affect y and are correlated with T

Problem #1: Omitted Variables

- Even if we try to control for "everything", we may miss:
  (1) Characteristics that we did not know mattered
  (2) Characteristics that are too complicated to measure (not observed):
    - Talent, motivation
    - Level of information and access to services
    - Income level
    - Level of support from family
    - Opportunity cost of participation in training
- Full model would be:

$$y = \gamma_0 + \gamma_1 x + \gamma_2 T + \gamma_3 M_1 + \eta$$

But we cannot observe M_1, the "missing" and unobserved variables

Omitted variable bias

- True model is: $y = \gamma_0 + \gamma_1 x + \gamma_2 T + \gamma_3 M_1 + \eta$
- But we estimate: $y = \beta_0 + \beta_1 x + \beta_2 T + \varepsilon$
- If there is a correlation between M_1 and T, then the *OLS* estimator of β_2 will not be a consistent estimator of γ_2, the true impact of T
- Why?

  When M_1 is missing from the regression, the coefficient of P will "pick up" some of the effect of M_1

Problem #2: Endogenous Decision to Participate in Training

- True model is: $$y = \gamma_0 + \gamma_1 x + \gamma_2 T + \eta$$

  with $$T = \pi_0 + \pi_1 x + \pi_2 M_2 + \xi$$

  M_2 = Vector of unobserved / missing characteristics

  (i.e. we don't fully know why people decide to participate or not)

- Since we don't observe M_2 , we can only estimate a simplified model:

$$y = \beta_0 + \beta_1 x + \beta_2 T + \varepsilon$$

- Is $\beta_{2,\, OLS}$ an unbiased estimator of γ_2? *NO!*

Problem #2: Endogenous Decision to Participate in Training

- To see why $\beta_{2,\ OLS}$ is a biased estimator of γ_2 *note:*

$$
\begin{aligned}
Corr\ (\varepsilon, T) \quad &= corr\ (\varepsilon, \pi_0 + \pi_1 x + \pi_2 M_2 + \xi) \\
&= \pi_1\ corr\ (\varepsilon, x) + \pi_2\ corr\ (\varepsilon, M_2) \\
&= \pi_2\ corr\ (\varepsilon, M_2)
\end{aligned}
$$

- If there is a correlation between the missing variables that determine participation in training (e.g. talent) and outcomes not explained by observed characteristics (i.e. in the error term), then the OLS estimator will be biased

What can we do to solve this problem?

- We estimate: $$y = \beta_0 + \beta_1 x + \beta_2 \boldsymbol{T} + \boldsymbol{\varepsilon}$$

- But the problem is the correlation between ***T*** and ***ε***

- So replace ***T*** with "something else" that is actually exogenous (uncorrelated with ***ε***); we can find "something else" if we look for some ***Z*** *that is:*
    - Similar to ***T***
    - But is not correlated with ***ε***

Back to the job training program

- ***T*** = participation in training
- ***ε*** = that part of outcomes that is not explained by program participation or by observed characteristics we are controlling for, X
- I'm looking for a variable ***Z*** that is:
  1. Closely related to participation in training, *T*
  2. But does not directly affect people's outcomes *Y, other than through its effect on participation in training*
- So this variable must be coming from outside (i.e. from some "exogenous" source that is not correlated with the error term)
  - We call this "an exogenous source of variation in T"

Generating an Exogenous Source of Variation in Access to Job Training

- Say that a social worker visits unemployed persons to encourage them to participate.
  - She only visits 50% of persons on her roster, and
  - Which 50% is visited is chosen randomly
- If she is effective, many people she was assigned to visit will enroll; there will be a correlation between receiving a visit and enrolling
- But being visited does not directly affect outcomes (e.g. income) apart from its effect on the likelihood of enrolling in training; why?
- Randomized “encouragement” or “promotion” visits are thus a valid “instrumental variable,” or “instrument”

Characteristics of a Valid Instrumental Variable

- Define a new variable *Z*

$Z =$ 1 — If person was randomly chosen to receive the encouragement visit from the social worker

0 — If person was randomly chosen *not* to receive the encouragement visit from the social worker

- $Corr\ (Z, T) > 0$

People who receive the encouragement visit are more likely to participate than those who do not

- $Corr\ (Z, \varepsilon) = 0$

No correlation between receiving a visit and benefit to the program apart from the effect of the visit on participation.

- *Z* is called an instrumental variable

Popular Form of IV: Two-stage Least Squares (2SLS)

- A popular form of the instrumental variables estimator, often employed in the context of endogeneity, is known as two-stage least squares (2SLS)
- Just as in ordinary least squares (OLS), we take a linear regression model and try to minimize the sum of the squares of the residuals

Two-stage Least Squares (2SLS)

Remember the original model with endogenous *T*:

$$y = \beta_0 + \beta_1 x + \beta_2 T + \varepsilon$$

Step 1

Regress the endogenous variable *T* on the instrumental variable(s) *Z* and other exogenous variables x

$$T = \delta_0 + \delta_1 x + \delta_2 Z + u$$

- Calculate the predicted value of T for each observation: $\hat{T}$
- Since *Z* and *x* are not correlated with *ε*, $\hat{T}$ will not be correlated with *ε* either. So we are getting rid of our endogenous regressor (T) in favor of an exogenous one ($\hat{T}$).

Two-stage Least Squares (2SLS)

Step 2

Regress *y* on the predicted variable $\hat{T}$ and the other exogenous variables

$$y = \beta_0 + \beta_1 x + \beta_2 \hat{T} + \varepsilon$$

- Note: The standard errors of the second stage *OLS* need to be corrected because $\hat{T}$ is a predicted regressor.
- In Practice: Use *STATA* ivreg2 command, which does the two steps at once and reports correct standard errors (in addition to computing useful post-estimation statistics!)
- Intuition: By using *Z* for *P*, we cleaned *P* of its correlation with η
- It can be shown that (under certain conditions) $\beta_{2,IV}$ yields a consistent estimator of γ_2 (large sample theory)

Is the IV Estimator Unbiased?

- No! This is a common misconception, since many are aware that IV is used to address problems like omitted variable bias
- The IV estimator is consistent, as long as the two key assumptions about the instrument's properties are satisfied
- But the IV estimator is not an unbiased estimator, and in small samples, its bias may be substantial
- The hope is that IV will be less biased than OLS – which should hold as long as an instrument is sufficiently strong

Link Back to the Estimation Formula

Stage 1

- Regress participation in training on a dummy for whether person received a visit from a social work encouraging enrolment (linear model)
- Compute predicted value of participation

Stage 2

Regress wages on the predicted value of participation

Reminder: The Exclusion and Inclusion Restrictions

- Exclusion Restriction: *corr* $(Z,\varepsilon) = 0$
  - If *corr* $(Z, \varepsilon) \neq 0$, "Bad instrument"

- Inclusion Restriction: *corr* $(Z,T) \neq 0$
  - "Weak instruments": the correlation between Z and T needs to be sufficiently strong. If not, the bias stays large, even for large sample sizes
  - Stock and Watson's rule of thumb: the first-stage F-statistic testing the hypothesis that the coefficients on the instruments are jointly zero should be at least 10
  - As mentioned previously Stock and Yogo (2005) provide critical values (smaller or larger than 10, depending on your regression); Stata gives you these critical values following use of ivreg2

Finding Instrumental Variables

- Searching for an IV ex post is hard and risky (and readers are skeptical!)
  - This does not mean you should not do it, but be prepared to defend your instrument!
- Using an IV you designed ex-ante (e.g., an information campaign) may be less objectionable
  - If everyone is eligible to participate in treatment
  - But having information about the program increases your likelihood of participating
  - And you provide "additional information" (think of this as lowering the costs – in terms of time – to an individual of finding out about the program) on a random basis

Using IV with an Experiment: Recovering TOT from ATE in Case of Non-compliance

- In advance, you decide that one randomly-selected group will be offered treatment and a second will not. But sometimes eligible units refuse treatment, or there are spillovers and a unit in the control group accidentally gets treated.
- Computing the Average Treatment Effect (ATE) (i.e. "intent to treat" estimate): Straight difference in average outcomes between the group offered treatment and the group not offered treatment
  → Tells us the likely impact of a program (messy or not in its execution!)
- Computing the Effect of Treatment on the Treated (TOT)
  Use the randomized offering as an instrumental variable *(Z)* for whether people receive treatment *(T)*
  → *Tells us the effect of being treated*
- Question: What if compliance is perfect; what does IV (computing TOT) give us?

Some Examples of the Use of IV in Randomized Experiments

OutcomeVariable	Endogenous Variable	Source of Instrumental Variable(s)	Reference
Earnings	Participation in job training program	Random assignment of admission to training program	Bloom et al. (1997)
Earnings	Participation in job training program	Random assignment of admission to training program	Burghardt et al. (2001)
Achieventment test scores	Enrollment in private school	Randomly selected offer of school voucher	Howell et al. (2000)
Achieventment test scores	Class size	Random assignment to a small or normal-size class	Krueger (1999)
Achieventment test scores	Hours of study	Random mailing of test preparation materials	Powers and Swinton (1984)
Birth weight	Maternal smoking	Random assignment of free smoker's counseling	Permutt and Hebel (1989)

Note: IV is a 'Local' Effect

- IV methods identify the average gains to persons induced to change their choice by a change of the instrument (referred to as compliers)
  - We estimate a "local average treatment effect," or LATE
  - We cannot identify who these people are (e.g., would someone who was visited by a social worker have taken up the job training even if they had not been visited? We cannot observe that and so do not know)
  - Also: different instruments will identify different parameters and answer different questions
- So: Be cautious when extrapolating to the whole population

Multiple Endogenous Variables

- Suppose you have k >1 endogenous variables
- Requires you have at least k > 1 instrumental variables (or more is fine, too)
- Despite the apparent 1:1 correspondence, you are not picking separate instrument for each endogenous variable; you will end up using *all* instrumental variables to predict each endogenous variable
    - You have one first stage equation for each endogenous variable; in it, you regress the endogenous variable on all instruments and on all control variables

Test for Endogeneity: Durbin-Wu-Hausman Test (DWH)

- Why might we want to test for endogeneity?
  1. Finding an instrument can be hard
  2. Use of IV inflates the variances of estimators, and thus weakens our ability to make inferences
- Durbin-Wu-Hausman Test (DWH) is a popular test for endogeneity

Test for Endogeneity: Durbin-Wu-Hausman Test (DWH)

- Imagine that we have the equation:

$$y = \beta_0 + \beta_1 x + \beta_2 T + \varepsilon \qquad (1)$$

where T is an endogenous explanatory variable and x is exogenous

- In the first stage, 2SLS would have us estimating (where Z is our instrument):

$$T = \delta_0 + \delta_1 x + \delta_2 Z + u \qquad (2)$$

- This allows us to compute the OLS residuals, $\hat{u}$ (a consistent estimator of u)
- If T is endogenous in (1), it will occur because cov(u,ε)≠0
- So add $\hat{u}$ to (1) and see if its coefficient is statistically significant:

$$y = \alpha_0 + \alpha_1 x + \alpha_2 T + \alpha_3 \hat{u} + v \qquad (3)$$

- If cov(u,ε)=0, $\widehat{\alpha_3}$ should not be significantly different from zero
- In that case, there is no evidence that T is endogenous in (1), and we can use OLS

Test for Endogeneity: Durbin-Wu-Hausman Test (DWH)

- This test may also be generalized for the presence of multiple included endogenous variables in (1); the relevant test is then an F-test, jointly testing that a set of coefficients are all zero
- The test is available within Stata as the estat endog command following ivreg

Testing Overidentifying Restrictions

- This is a test you can perform only if you have more instruments than you have endogenous variables, and thus more instruments than you need to estimate the causal effect of some endogenous variable
- The number of “overidentifying restrictions” is just the number of excess instruments (beyond the minimum number needed)
- If all instruments satisfy the exclusion restriction, all subsets should (asymptotically) return the same estimate of the treatment effect
- Idea: Obtain multiple estimates of the treatment effect and test that they are the same
- Rejection implies some subset of exclusion restrictions may be invalid

Testing Overidentifying Restrictions

- Regress the residuals from the second stage regression on the instruments (and any exogenous control variables) and test whether the coefficients on the instruments are all zero
  - Helps us see: Do the instruments truly satisfy the condition that they need to be uncorrelated with the error term?
- If we *fail to reject* that they are jointly equal to zero (i.e. if they are likely 0!), then the instruments are more likely to be reliably ‘exogenous’
- This test is available within Stata as the estat overid command following ivreg

Testing Overidentifying Restrictions

- But please note: You can never tell you that the exclusion restriction ($E[z\varepsilon] = 0$) is satisfied
  - Failure to reject does not imply true
  - Even if it did, we only learn that probability limits of various IV estimators are the same; maybe all are the same and wrong!
- Rejection indicates that some subset of instruments may be invalid
  - Does not indicate which subset
  - Does not mean *all* exclusion restrictions are invalid
- So, you still need to argue (and strongly) for exogeneity (i.e. that your instruments satisfy the exclusion restriction)

Example 1: Returns to Schooling

- Goal: Estimate the value added of additional years of schooling in terms of wages

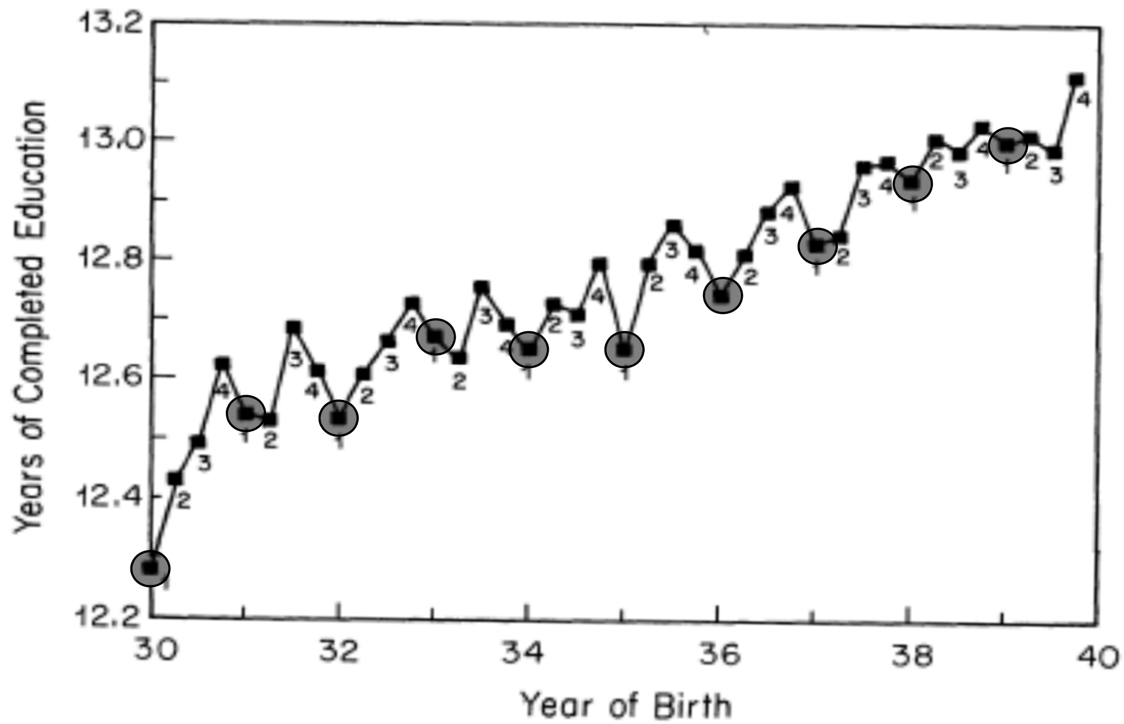
- Problem: Years of completed schooling is not randomly assigned. May be endogenous.
  - E.g. maybe academic ability is related to qualities that relate to job performance/salary (motivation, intelligence, task orientation, etc.)
- Instrument: Quarter of birth (Angrist and Krueger, 1991)

Example 1: Returns to Schooling

- Plausibility of instrument:
  - Compulsory schooling laws in the U.S. are typically based on age, not number of years of school. People born at different times of the year can drop out after receiving different amounts of school.
  - When a person is born is unrelated to inherent traits (e.g. motivation, intelligence, …) and so should not have a direct effect on wages but only affect wages through the relationship to completed schooling induced by compulsory education laws.
    - Untestable, but we do have overidentifying restrictions coming from different birth quarters.
    - Validity has been questioned. E.g. winter birth may be correlated to increased exposure to early health problems; more conscientious parents may respond by timing birth; …

Example 1: Returns to Schooling

- But we get a compelling picture nonetheless; people born in Q1 do obtain less schooling
  - But pay close attention to the scale of the y-axis
  - Mean difference between Q1 and Q4 is only 0.124, or 1.5 months
- So...need large N
  - Angrist and Krueger (1991) had a sample of over 329,000



Source: Angrist and Krueger (1991), Figure I

Example 1: Returns to Schooling

- Structural Equation:

$$\log(wage_i) = \alpha School_i + \beta\, x_i + \varepsilon_i$$

- First-Stage Equation:

$$School_i = \pi_1 Q1_i + \pi_2 Q2_i + \pi_3 Q3_i + \propto x_i + u_i$$

Example 1: Returns to Schooling

- Data from 1980 Census for men aged 40-49 in 1980
- Variables:
  - *Wage* – hourly wage
  - *School* – reported years of completed schooling
  - *Q1-Q3* – dummies for quarter of birth
  - *x* – 59 control variables. Dummies for state of birth and year of birth

Example 1: Returns to Schooling

- OLS Results (from Stata):

```
xi: reg lwage educ i.yob i.sob , robust
i.yob             _Iyob_30-39        (naturally coded; _Iyob_30 omitted)
i.sob             _Isob_1-56         (naturally coded; _Isob_1 omitted)

Linear regression                                      Number of obs =  329509
                                                       F( 60,329448) =  649.29
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.1288
                                                       Root MSE      =  .63366

------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .067339   .0003883   173.40   0.000     .0665778    .0681001
.
.
.
```

Example 1: Returns to Schooling

- First-Stage Results (from Stata):

```
 xi: regress educ i.qob i.sob i.yob , robust
Linear regression                                      Number of obs =  329509
                                                       F( 62,329446) =  292.87
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0572
                                                       Root MSE      =  3.1863

------------------------------------------------------------------------------
             |               Robust
        educ |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     _Iqob_2 |   .0455652    .015977     2.85   0.004     .0142508    .0768797
     _Iqob_3 |   .1060082   .0155308     6.83   0.000     .0755683     .136448
     _Iqob_4 |   .1525798   .0157993     9.66   0.000     .1216137    .1835459

.
.
.
 testparm _Iqob*

 ( 1)  _Iqob_2 = 0
 ( 2)  _Iqob_3 = 0
 ( 3)  _Iqob_4 = 0

       F(  3,329446) =   36.06
            Prob > F =    0.0000
```

First-stage F-statistic.

Example 1: Returns to Schooling

- 2SLS Results (from Stata):

```
xi: ivregress 2sls lwage (educ = i.qob) i.yob i.sob , robust

Instrumental variables (2SLS) regression               Number of obs =  329509
                                                       Wald chi2(60) = 9996.12
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.0929
                                                       Root MSE      =  .64652

------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1076937   .0195571     5.51   0.000     .0693624     .146025
.
.
.
```

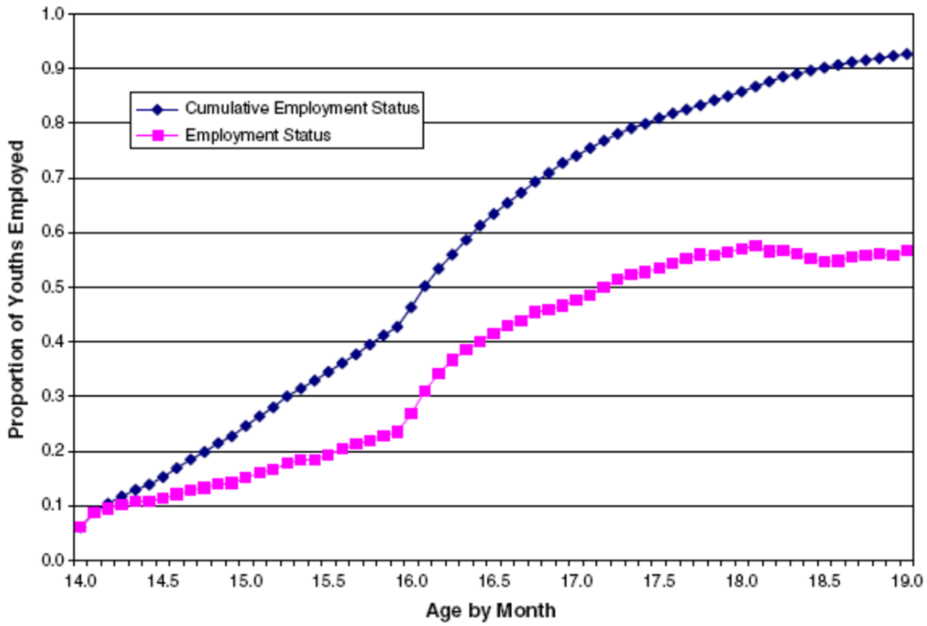
Bigger than OLS

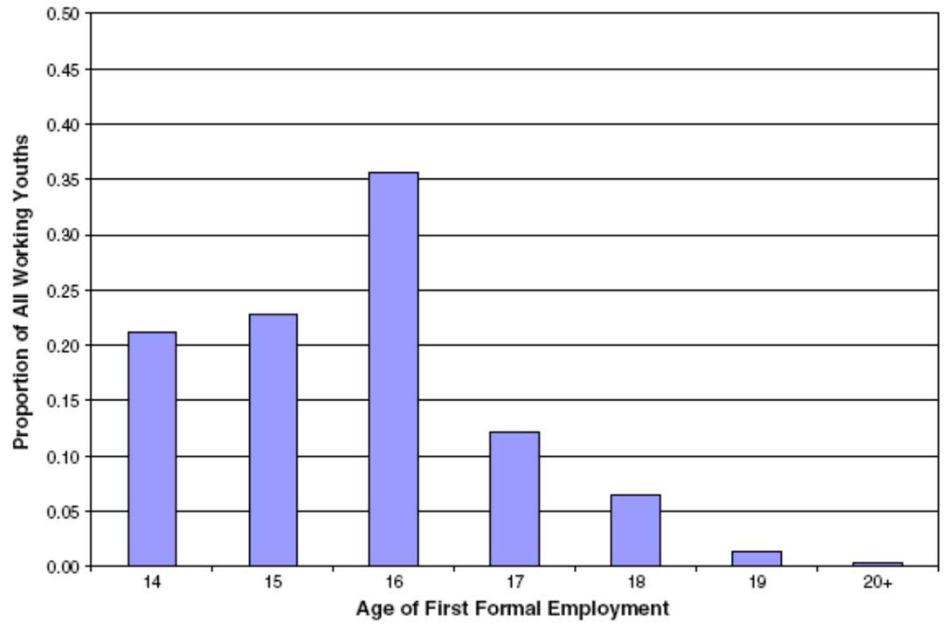
Example 2: Teenage work and deviant behavior

- What effect does working have on teenagers' behavior in the United States?
  - Prior research suggests the consequences of work are uniformly negative
  - But examining a dummy variable for working masks work intensity…
  - Thus focus on "work intensity" rather than work per se

- But there is an endogeneity problem here: Non-random selection into the youth labor market
  - Especially pronounced for high-intensity workers (think of children who have to work a lot because their families are too poor to pay for basic necessities)

Example 2: Teenage work and deviant behavior





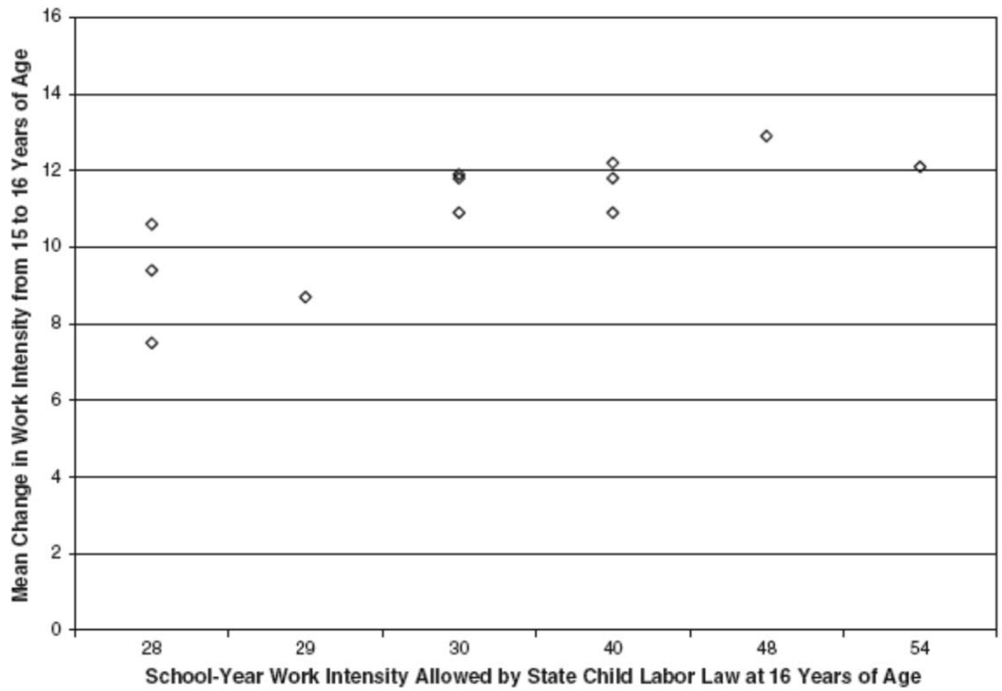
- Something interesting happens at age 16
  - Youth work is no longer governed by the federal Fair Labor Standards Act (F.L.S.A.)

Example 2: Teenage work and deviant behavior

- F.L.S.A. governs employment of all 15 year olds during the school year
  - No work past 7:00 pm
  - Maximum 3 hours/day and 18 hours/week
- But, F.L.S.A. expires for 16 year olds
  - And...every state has its own law governing 16-year-old employment
  - Thus, youth age into less restrictive regimes that vary across jurisdictions

Example 2: Teenage work and deviant behavior

- Change in work intensity at 15-16 transition among 15-year-old non-workers



Magnitude of change is an increasing function of the number of hours individuals are allowed to work at age 16

Example 2: Teenage work and deviant behavior

	Model 1
State child labor laws	
Hours per week	.317 (.049)***
Hours per weekday	
Work curfew	
No hours per week limit	11.426 (1.58)***
No hours per weekday limit	
No work curfew limit	
Individual-level controls	
Residential Location:	
Central City (ref.)	
Suburbs	−.207 (1.80)
Rural	−2.947 (3.97)
Dwelling type	
House, condo, or farm	−1.176 (1.28)
Apartment or flat	−1.573 (1.52)
Other dwelling (ref.)	
Residential mobility	7.877 (3.39)*
Household size	.144 (.287)

- Instrumental variables used in primary model:
  1. Hours per week state labor laws allow 16 year olds to work
  2. A dummy for no limit on hours worked per week
- Both instruments positively correlated with # hours of labor supplied by 16 year olds

Example 2: Teenage work and deviant behavior

Table 3 Comparative models of deviant behavior and academic achievement at the 15-to-16 transition

Dependent variable	N	Panel A: structural coefficients for work intensity		
		(1) Random effects	(2) Fixed effects	(3) Fixed effects IV
Delinquent behavior	2,207	.0034 (.0014)*	.0000 (.0017)	−.0233 (.0089)**
Arrest	2,199	.0007 (.0004)*	−.0006 (.0005)	−.0000 (.0023)
Substance use	2,203	.0070 (.0013)***	.0034 (.0014)*	.0067 (.0073)
School suspension	1,905	.0185 (.0051)***	.0005 (.0064)	−.0765 (.0372)*
Transcript grades	1,264	−.0040 (.0012)***	−.0020 (.0012)†	−.0073 (.0069)
School dropout	2,217	.0021 (.0003)***	.0017 (.0004)***	.0109 (.0025)***

- A 10-hour increase in the number of hours worked per week:
  - Reduces the "variety" of delinquent behavior by 0.23 (–.0233×10) (on a 0-7 scale)
  - Reduces the incidence of school dropout by 0.109 (0.0109x10) (on a 0-1 scale)

Recommendations Regarding IVs

1. Before you talk about the instrument you use, talk about your endogeneity problem and show that you understand it
   - Do not just acknowledge endogeneity—describe exactly where it comes from (e.g., unobserved variable called w may influence both the outcome and your main independent variable of interest)
2. Present both OLS and IV results – at least for your main table(s)
3. Especially if OLS and IV results are different, discuss why – linking back to your discussion of the sources of endogeneity
4. Explain why your instrument is a good one – both why it meets the inclusion restriction and why it meets the exclusion restriction

Recommendations Regarding IVs

5. Try to describe the randomized experiment your IV is mimicking
6. Show first stage regression results (a table in your paper)
7. State clearly what is your identifying assumption and argue for its plausibility
   - Usually something along the lines of “we assume that Z only influences Y through its effect on X”
   - Check examples from papers using IV successfully (i.e. in good journals/ which you personally find convincing)
8. Report test(s) of overidentifying restrictions if you have more instruments than endogenous variables
9. You may wish to report results from a test for endogeneity (Durbin-Wu-Hausman)

Recommendations Regarding IVs

10. As you would with any non-experimental paper, think of other creative ways to validate your modeling choice
    - E.g., If examining the impact of income on children's height and instrumenting for income, let's say OLS results are negative and insignificant, but IV results are positive and highly significant. Is IV doing something weird that is biasing the coefficients upward (and biasing standard errors downward)? Showing us that income has no effect on adult height in IV results would be reassuring and make any identification strategy more convincing.
11. Be prepared: Whether your OLS or IV results form the central results of the paper may end up being up to the editor/ referees
12. A lagged value of the endogenous regressor is not a good instrument; all sorts of omitted variables can affect both lags of X and current values of Y
13. If you use a particular instrument, it always helps to cite other papers that have used the same or a similar instrument (provided that they are published in reputable journals and/or by reputable people)

Recommendations Regarding IVs

14. It is okay to search for an instrument that gives you a strong first stage; this is not 'data mining' in a bad sense; you should always want the instrument that performs well in the first stage and thus most closely mimics the process generating the data!
15. If you have an endogenous variable x that is interacted with some other variable r, x*r is also endogenous, but if you have a first instrument z, you can use z*r as a second instrument
16. Even if your endogenous variable does not involve an interaction, you can still use an instrumental variable that is an interaction
    - But still need to include the levels of these two interacted variables in the regression – either as instruments or as controls
17. We cannot say that an instrument is 'good' for a given endogenous variable (e.g., 'rainfall is a good instrument for farm income'); it matters what the outcome variable is, since that determines what is in the error term

Recommendations Regarding IVs

18. Attacking someone's instrumental variable is the what seminar participants and referees do; you should defend well your modeling assumptions and where possibly show that your results go through even without a certain assumption