Lecture 5:
Instrumental Variables, 1 of 2

February 12, 2025

Course Administration

1. Hopefully you've turned in PS 2
2. PS 3 posted, due March 5
3. Lab after class next week
4. If you still need approval for your replication paper, or need to choose a new one, do it ASAP
5. March 19: quantitative progress report due
6. Let's divide next week's articles
7. Please come see me about your replication paper
8. Any other issues?

Plan for Today

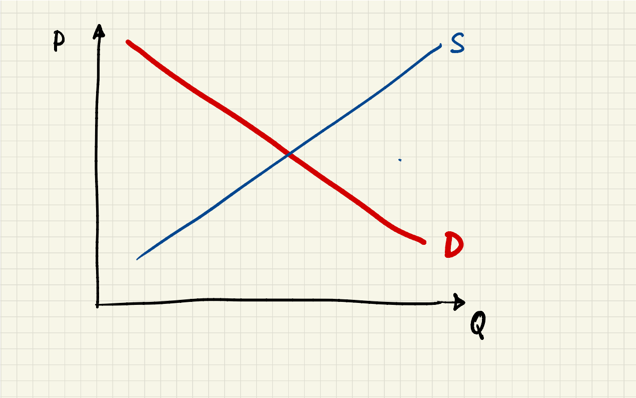
1. IV Overview
2. A&K: Oldie but Goodie

IV Background

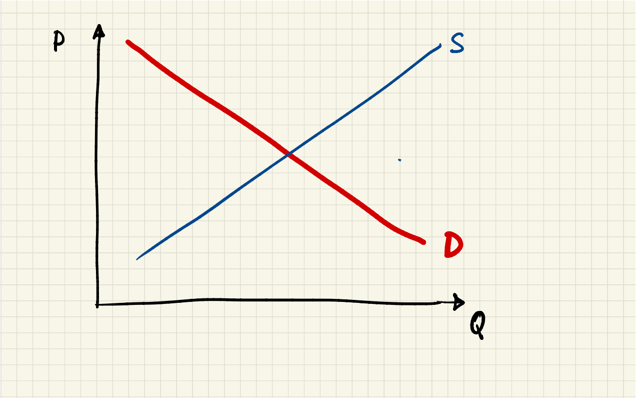
1. Origins and motivation of IV
2. More general formulation
3. Regression framework: Wald estimate and 2SLS
4. Testing assumptions underlying IV

IV Origins

IV Origin Story



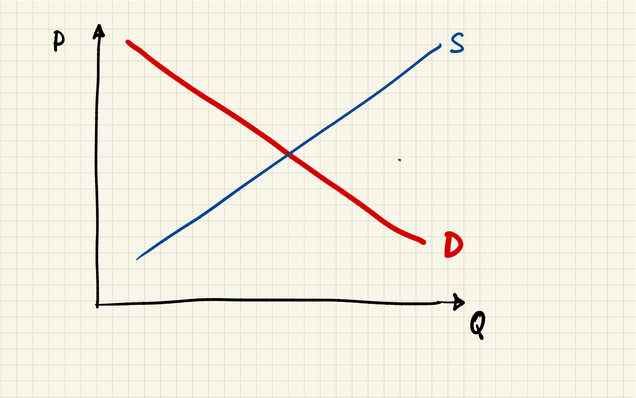
IV Origin Story



- We want to know the impact of price on quantity demanded, or

$$Q_t^D = \alpha P_t + \epsilon_t$$

IV Origin Story



- We want to know the impact of price on quantity demanded, or

$$Q_t^D = \alpha P_t + \epsilon_t$$

- But we only observe equilibrium where $S = D$

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)}$$

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)}$$

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)} = \alpha \frac{\text{cov}(P_t, P_t)}{\text{var}(P_t)} + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)} = \alpha + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)}$$

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)} = \alpha \frac{\text{cov}(P_t, P_t)}{\text{var}(P_t)} + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)} = \alpha + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)}$$

- Estimating the true α depends on the assumption that $\frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P)} = 0$

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)} = \alpha \frac{\text{cov}(P_t, P_t)}{\text{var}(P_t)} + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)} = \alpha + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)}$$

- Estimating the true α depends on the assumption that $\frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P)} = 0$
- When might this not be the case?
    - suppose a study shows that chocolate is good for your health

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)} = \alpha \frac{\text{cov}(P_t, P_t)}{\text{var}(P_t)} + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)} = \alpha + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)}$$

- Estimating the true α depends on the assumption that $\frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P)} = 0$
- When might this not be the case?
  - suppose a study shows that chocolate is good for your health
  - P increases and Q increases – does increase in P cause increase in Q?

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)} = \alpha \frac{\text{cov}(P_t, P_t)}{\text{var}(P_t)} + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)} = \alpha + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)}$$

- Estimating the true α depends on the assumption that $\frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P)} = 0$
- When might this not be the case?
    - suppose a study shows that chocolate is good for your health
    - P increases and Q increases – does increase in P cause increase in Q? no!

What's the Trouble with this Estimation?

$$Q_t^D = \alpha P_t + \epsilon_t$$

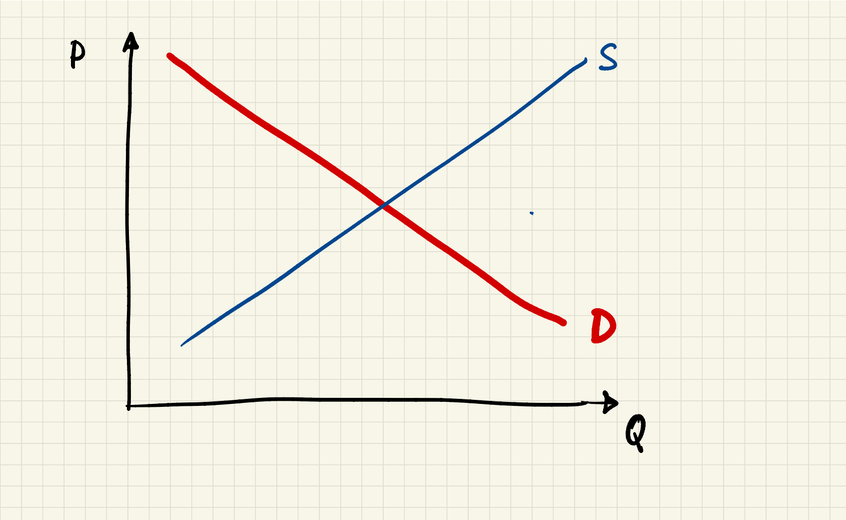
- We estimate $\hat{\alpha}$:

$$\hat{\alpha} = \frac{\text{cov}(Q_t, P_t)}{\text{var}(P_t)} = \frac{\text{cov}(\alpha P_t + \epsilon_t, P_t)}{\text{var}(P_t)} = \alpha \frac{\text{cov}(P_t, P_t)}{\text{var}(P_t)} + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)} = \alpha + \frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P_t)}$$

- Estimating the true α depends on the assumption that $\frac{\text{cov}(\epsilon_t, P_t)}{\text{var}(P)} = 0$
- When might this not be the case?
  - suppose a study shows that chocolate is good for your health
  - P increases and Q increases – does increase in P cause increase in Q? no!
  - an omitted variable – new info – causes both P and Q^D to increase

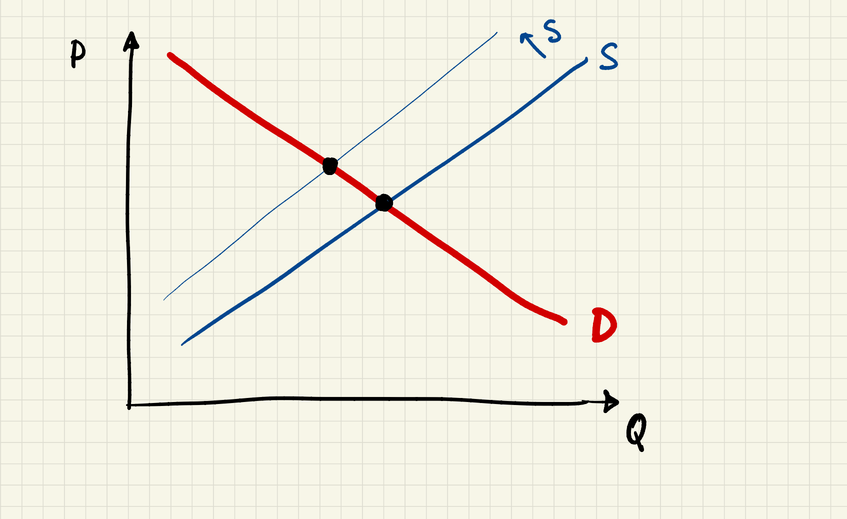
Potential Solutions to this Problem

We need something that shifts P AND does not affect Q_D



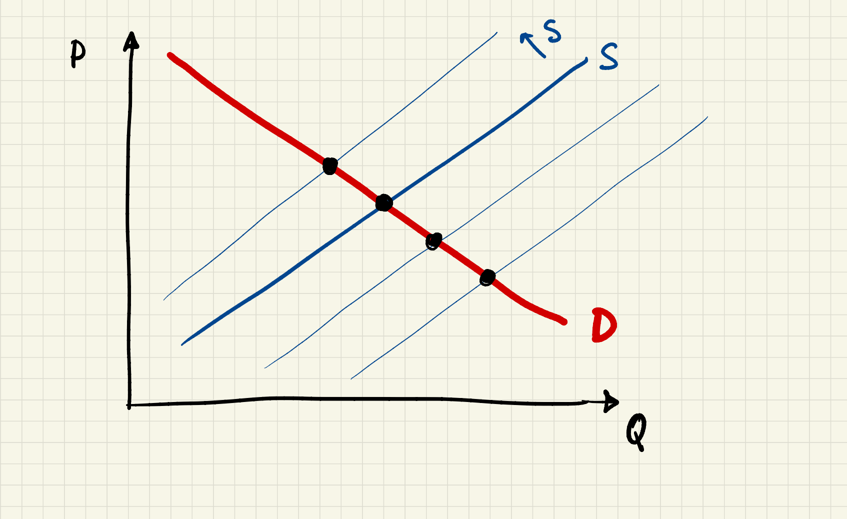
Potential Solutions to this Problem

We need something that shifts P AND does not affect Q_D

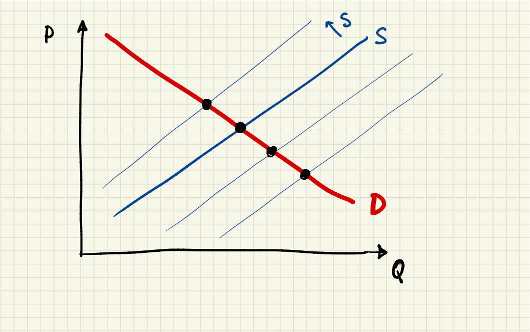


Potential Solutions to this Problem

We need something that shifts P AND does not affect Q_D

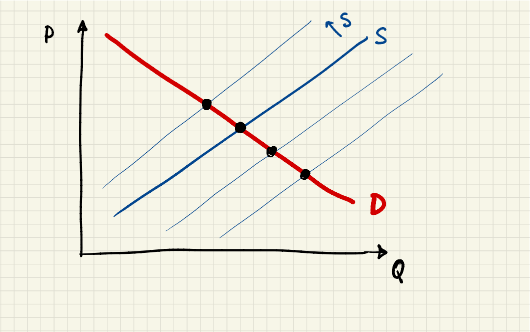


Potential Solutions to this Problem



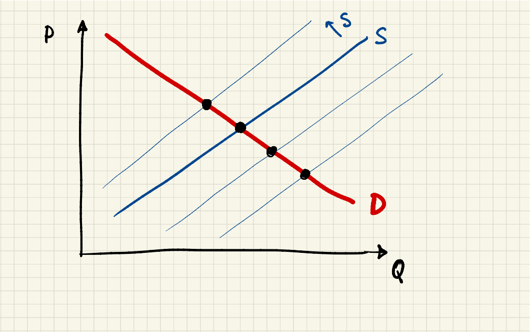
- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve

Potential Solutions to this Problem



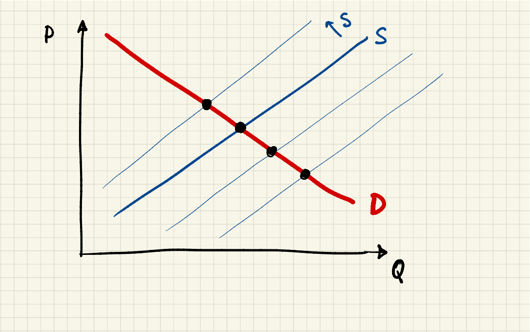
- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve
- What could such a thing be?

Potential Solutions to this Problem



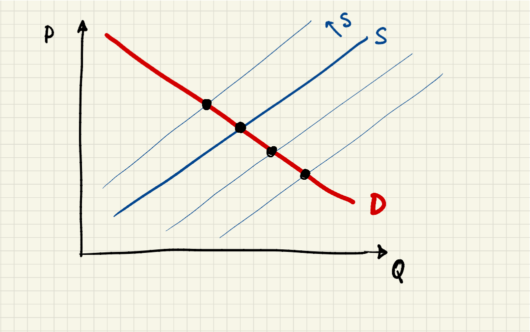
- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve
- What could such a thing be? input prices

Potential Solutions to this Problem



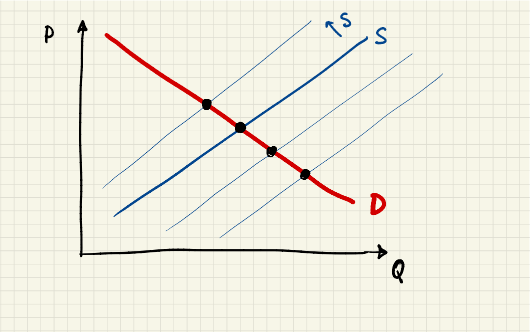
- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve
- What could such a thing be? input prices
- What if this is chocolate?

Potential Solutions to this Problem



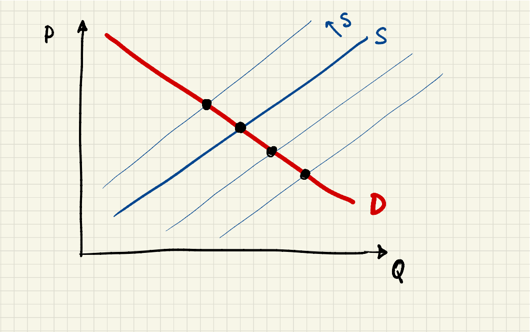
- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve
- What could such a thing be? input prices
- What if this is chocolate?
    - price of cocoa beans
    - price of labor in Ivory Coast

Potential Solutions to this Problem



- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve
- What could such a thing be? input prices
- What if this is chocolate?
  - price of cocoa beans
  - price of labor in Ivory Coast
  - price of labor in US is no good – why?

Potential Solutions to this Problem



- Something that shifts P that doesn't affect Q_D
- If supply shifts, we trace out the demand curve
- What could such a thing be? input prices
- What if this is chocolate?
  - price of cocoa beans
  - price of labor in Ivory Coast
  - price of labor in US is no good – why?

We need something that generates variation in P unrelated to outcome Q_D

IV: More General Problem

General Formulation of Basic Problem

- We want to know β:

$$Y = \beta X + \nu$$

General Formulation of Basic Problem

- We want to know β:

$$Y = \beta X + \nu$$

- But we know that

$$\nu = \alpha A + \phi$$

General Formulation of Basic Problem

- We want to know β:

$$Y = \beta X + \nu$$

- But we know that

$$\nu = \alpha A + \phi$$

- This implies that we should really estimate

$$Y = \beta X + \alpha A + \phi$$

General Formulation of Basic Problem

- We want to know β:

$$Y = \beta X + \nu$$

- But we know that

$$\nu = \alpha A + \phi$$

- This implies that we should really estimate

$$Y = \beta X + \alpha A + \phi$$

- Problem: We don't observe A
- Are we stuck?

Instrument as an Imperfect Experiment

Imperfect Experiment

- imperfect
  because it's not entirely
  randomized
- experiment
  because it contains some
  element of randomness

Instrument as an Imperfect Experiment

Imperfect Experiment

- imperfect
  because it's not entirely randomized
- experiment
  because it contains some element of randomness

Two motivating examples

- A and K: X is schooling, Y is wages – what's A?

Instrument as an Imperfect Experiment

Imperfect Experiment

- imperfect
  because it's not entirely randomized
- experiment
  because it contains some element of randomness

Two motivating examples

- A and K: X is schooling, Y is wages – what's A?
- Angrist et al: Y is test scores, X is charter school attendance – what's A?

Instrument as an Imperfect Experiment

Imperfect Experiment

- imperfect
  because it's not entirely randomized
- experiment
  because it contains some element of randomness

Two motivating examples

- A and K: X is schooling, Y is wages – what's A?
- Angrist et al: Y is test scores, X is charter school attendance – what's A?
- instrument for A and K is

Instrument as an Imperfect Experiment

Imperfect Experiment

- imperfect
  because it's not entirely randomized
- experiment
  because it contains some element of randomness

Two motivating examples

- A and K: X is schooling, Y is wages – what's A?
- Angrist et al: Y is test scores, X is charter school attendance – what's A?
- instrument for A and K is quarter of birth

Instrument as an Imperfect Experiment

Imperfect Experiment

- imperfect
  because it's not entirely randomized
- experiment
  because it contains some element of randomness

Two motivating examples

- A and K: X is schooling, Y is wages – what's A?
- Angrist et al: Y is test scores, X is charter school attendance – what's A?
- instrument for A and K is quarter of birth
- instrument for Angrist et al is lottery winning

IV: Regression Framework

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta$

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta \rightarrow \hat{X} =$

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta \rightarrow \hat{X} = \hat{\gamma} Z$

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta \rightarrow \hat{X} = \hat{\gamma} Z$
2. reduced form: $Y = \lambda Z + \epsilon$

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta \rightarrow \hat{X} = \hat{\gamma} Z$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu$

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta \rightarrow \hat{X} = \hat{\gamma} Z$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$

IV Regression Framework

Let Z be the instrument, and X be the endogenous variable.

Three key equations

1. first stage: $X = \gamma Z + \delta \rightarrow \hat{X} = \hat{\gamma} Z$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$

Instead of full variation in X, use variation in X that comes from Z

Conditions for an instrument, Z

1. Instrument must be correlated with endogenous variable:

$$\text{cov}(X, Z) \neq 0$$

2. Instrument must not be correlated with error in estimating equation:

$$\text{cov}(Z, \nu) = 0$$

Conditions for an instrument, Z

1. Instrument must be correlated with endogenous variable:

$$\text{cov}(X, Z) \neq 0$$

2. Instrument must not be correlated with error in estimating equation:

$$\text{cov}(Z, \nu) = 0$$

Conditions for an instrument, Z

1. Instrument must be correlated with endogenous variable:

$$\text{cov}(X, Z) \neq 0$$

2. Instrument must not be correlated with error in estimating equation:

$$\text{cov}(Z, \nu) = 0$$

- Angrist and Pischke divide (2) in two
  - independence assumption: "instrument is as good as randomly assigned" conditional on covariates
  - exclusion restriction: instrument impacts dependent variable Y **only** through its relationship with endogenous variable X

Conditions for an instrument, Z

1. Instrument must be correlated with endogenous variable:

$$\text{cov}(X, Z) \neq 0$$

2. Instrument must not be correlated with error in estimating equation:

$$\text{cov}(Z, \nu) = 0$$

- Angrist and Pischke divide (2) in two
  - independence assumption: "instrument is as good as randomly assigned" conditional on covariates
  - exclusion restriction: instrument impacts dependent variable Y **only** through its relationship with endogenous variable X
- Between (1) and (2), which one is hard?

Conditions for an instrument, Z

1. Instrument must be correlated with endogenous variable:

$$\text{cov}(X, Z) \neq 0$$

2. Instrument must not be correlated with error in estimating equation:

$$\text{cov}(Z, \nu) = 0$$

- Angrist and Pischke divide (2) in two
  - independence assumption: "instrument is as good as randomly assigned" conditional on covariates
  - exclusion restriction: instrument impacts dependent variable Y **only** through its relationship with endogenous variable X
- Between (1) and (2), which one is hard? the second!

Think About Limiting Cases

1. first stage: $X = \gamma Z + \delta$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$

- what if $\hat{\gamma} = 1$?

Think About Limiting Cases

1. first stage: $X = \gamma Z + \delta$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$

- what if $\hat{\gamma} = 1$? Just use Z directly – and your instrument may be crap

Think About Limiting Cases

1. first stage: $X = \gamma Z + \delta$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$

- what if $\hat{\gamma} = 1$? Just use Z directly – and your instrument may be crap
- Note that $\hat{\lambda} = \hat{\beta}\hat{\gamma}$

Think About Limiting Cases

1. first stage: $X = \gamma Z + \delta$
2. reduced form: $Y = \lambda Z + \epsilon$
3. second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$

- what if $\hat{\gamma} = 1$? Just use Z directly – and your instrument may be crap
- Note that $\hat{\lambda} = \hat{\beta}\hat{\gamma}$
- this means that

$$\hat{\beta} = \frac{\hat{\lambda}}{\hat{\gamma}} = \frac{\text{cov}(Y, Z)/\text{var}(Z)}{\text{cov}(X, Z)/\text{var}(Z)} = \frac{\text{cov}(Y, Z)}{\text{cov}(X, Z)}$$

- in words, how much of the change in Y due to Z is explained by the change in X due to Z

Simplifying with a Binary Instrument

- Let's think of Z as being binary: won lottery or not
- Write the Wald Estimate as

$$\hat{\beta}_{IV} = \frac{\text{cov}(Y, Z)}{\text{cov}(X, Z)} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(X|Z=1) - E(X|Z=0)}$$

- In words?

Simplifying with a Binary Instrument

- Let's think of Z as being binary: won lottery or not
- Write the Wald Estimate as

$$\hat{\beta}_{IV} = \frac{\text{cov}(Y, Z)}{\text{cov}(X, Z)} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(X|Z=1) - E(X|Z=0)}$$

- In words? mean difference in test scores for students who won KIPP lottery and those who didn't divided by mean difference in KIPP attendance between those who won KIPP lottery and those who didn't

Simplifying with a Binary Instrument

- Let's think of Z as being binary: won lottery or not
- Write the Wald Estimate as

$$\hat{\beta}_{IV} = \frac{cov(Y, Z)}{cov(X, Z)} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(X|Z=1) - E(X|Z=0)}$$

- In words? mean difference in test scores for students who won KIPP lottery and those who didn't divided by mean difference in KIPP attendance between those who won KIPP lottery and those who didn't
- if the instrument has no effect the denominator goes to zero (!!)
- you can do this with means or a regression

Implementing More General IV

- first stage: $X = \gamma Z + \delta, \rightarrow \hat{X}$
- second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$
- note that you can do this in two steps – that's why it's called two-stage

Implementing More General IV

- first stage: $X = \gamma Z + \delta, \rightarrow \hat{X}$
- second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$
- note that you can do this in two steps – that's why it's called two-stage
- estimate first stage, get $\hat{X}$, then put $\hat{X}$ as regressor
- this will give you the right coefficient and the **wrong** standard error

Implementing More General IV

- first stage: $X = \gamma Z + \delta, \rightarrow \hat{X}$
- second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$
- note that you can do this in two steps – that's why it's called two-stage
- estimate first stage, get $\hat{X}$, then put $\hat{X}$ as regressor
- this will give you the right coefficient and the **wrong** standard error
- Stata can fix the standard error for you – we won't go into more detail

Implementing More General IV

- first stage: $X = \gamma Z + \delta, \rightarrow \hat{X}$
- second stage: $Y = \beta \hat{X} + \nu = \beta(\hat{\gamma} Z) + \nu$
- note that you can do this in two steps – that's why it's called two-stage
- estimate first stage, get $\hat{X}$, then put $\hat{X}$ as regressor
- this will give you the right coefficient and the **wrong** standard error
- Stata can fix the standard error for you – we won't go into more detail

Note that β_{IV} is biased in finite samples, but consistent – β_{OLS} is both unbiased and consistent

IV: Testing Underlying Assumptions

Testing Assumptions Underlying IV

- Correlation between Z and error is untestable

Testing Assumptions Underlying IV

- Correlation between Z and error is untestable
- However, we can test
  1. correlation between instrument and things that should not be affected by treatment – recall that instrument works *only* through treatment
  2. relationship betwen Z and Y where Z should not impact treatment

Testing Assumptions Underlying IV

- Correlation between Z and error is untestable
- However, we can test
  1. correlation between instrument and things that should not be affected by treatment
     – recall that instrument works *only* through treatment
  2. relationship betwen Z and Y where Z should not impact treatment
- Section III in A and K is an attempt to do these things
- Any credible IV paper should some argument along these lines

Angrist and Kreuger

Second Half of Lecture 5

1. Why do we read this old paper?
2. Causal question and endogeneity concerns
3. Data
4. Instruments and validity
5. Results
6. Reflection, many years on

1. Why do we read this old paper?

Why Do We Read This Old Paper?

- Enormously influential in use of instruments and random variation more broadly
- Well-written
- Clear exposition of how the instrument works

Why Do We Read This Old Paper?

- Enormously influential in use of instruments and random variation more broadly
- Well-written
- Clear exposition of how the instrument works
- Angrist won 2021 Nobel Prize

Why Do We Read This Old Paper?

- Enormously influential in use of instruments and random variation more broadly
- Well-written
- Clear exposition of how the instrument works
- Angrist won 2021 Nobel Prize
- Not quite right, so there's room to think

2. Causal question and endogeneity

What is this paper about?

Causal question?

What is this paper about?

Causal question?

- General:
  What is the impact of education on earnings?

What is this paper about?

Causal question?

- General:
  What is the impact of education on earnings?
- Specific:
  How does the impact of additional year of schooling, driven by birth quarter, impact wages?

What is this paper about?

Causal question?

- General:
  What is the impact of education on earnings?
- Specific:
  How does the impact of additional year of schooling, driven by birth quarter, impact wages?

Identification issues

- why don't we just estimate

$$\text{income}_i = \beta_0 + \beta_1 \text{schooling}_i + \beta_2 X_i + \epsilon_i$$

?

3. Data

What are the data?

What are the data?

- 1960, 1970 and 1980 5 percent samples from Census
- Cross-section or panel?

What are the data?

- 1960, 1970 and 1980 5 percent samples from Census
- Cross-section or panel? cross-section

What are the data?

- 1960, 1970 and 1980 5 percent samples from Census
- Cross-section or panel? cross-section
- What's the unit of observation?

What are the data?

- 1960, 1970 and 1980 5 percent samples from Census
- Cross-section or panel? cross-section
- What's the unit of observation? person (not in a year)

3. Instruments

What are the instruments?

What are the instruments?

- generally, quarter of birth
- what's the story about how this works?

What are the instruments?

- generally, quarter of birth
- what's the story about how this works?
- the story is that this works through compulsory schooling laws + different age of starting school, aka quarter of birth

What are the instruments?

- generally, quarter of birth
- what's the story about how this works?
- the story is that this works through compulsory schooling laws + different age of starting school, aka quarter of birth
- See Eq. 1, p. 997: quarter*year (10 years), so 40 instruments

What are the instruments?

- generally, quarter of birth
- what's the story about how this works?
- the story is that this works through compulsory schooling laws + different age of starting school, aka quarter of birth
- See Eq. 1, p. 997: quarter*year (10 years), so 40 instruments
- or 50 states * 4 quarters + 10 years * 4 quarters = 240 instruments

Showing the Easy IV Condition: $\text{corr}(X, Z) \neq 0$

IV condition No. 1: $\text{corr}(X, Z) \neq 0$. What evidence do they offer?

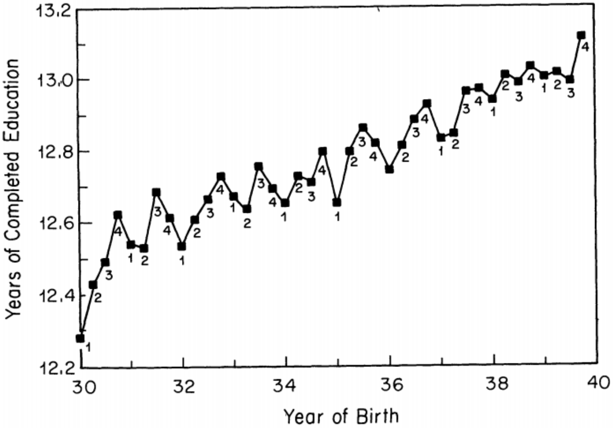
- First, the story linking quarter of birth to compulsory schooling laws

Showing the Easy IV Condition: $\text{corr}(X, Z) \neq 0$

IV condition No. 1: $\text{corr}(X, Z) \neq 0$. What evidence do they offer?

- First, the story linking quarter of birth to compulsory schooling laws
- Figures 1-3, first part of Table 1, and Figure 4 (which is Figures 1 to 3 in differences)
- and we have to believe that the diff-in-diff with compulsory schooling laws is not endogenous
- note that Section 1 is mostly dedicated to this

Figure 1: Quarter of Birth and Education



Data from 1980 Census, ages 50 (left) to 40 (right)

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$. What evidence do they offer?

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$. What evidence do they offer?

- Table 1: effect doesn't work through higher ed, suggesting compulsory schooling
- Table 2: laws are important, and they used to be more compelling
  - what's the hypothesis they are trying to test here?
  - QOB matters for attendance at age 16 when school leaving age is 16, but not when school leaving age is 17 or 18 – suggesting that the law is the mechanism

TABLE I
THE EFFECT OF QUARTER OF BIRTH ON VARIOUS EDUCATIONAL OUTCOME VARIABLES

Outcome variable	Birth cohort	Mean	Quarter-of-birth effect[a] I	II	III	F-test[b] [P-value]
Total years of education	1930–1939	12.79	−0.124	−0.086	−0.015	24.9
			(0.017)	(0.017)	(0.016)	[0.0001]
	1940–1949	13.56	−0.085	−0.035	−0.017	18.6
			(0.012)	(0.012)	(0.011)	[0.0001]
High school graduate	1930–1939	0.77	−0.019	−0.020	−0.004	46.4
			(0.002)	(0.002)	(0.002)	[0.0001]
	1940–1949	0.86	−0.015	−0.012	−0.002	54.4
			(0.001)	(0.001)	(0.001)	[0.0001]
Years of educ. for high school graduates	1930–1939	13.99	−0.004	0.051	0.012	5.9
			(0.014)	(0.014)	(0.014)	[0.0006]
	1940–1949	14.28	0.005	0.043	−0.003	7.8
			(0.011)	(0.011)	(0.010)	[0.0017]
College graduate	1930–1939	0.24	−0.005	0.003	0.002	5.0
			(0.002)	(0.002)	(0.002)	[0.0021]
	1940–1949	0.30	−0.003	0.004	0.000	5.0
			(0.002)	(0.002)	(0.002)	[0.0018]
Completed master's degree	1930–1939	0.09	−0.001	0.002	−0.001	1.7
			(0.001)	(0.001)	(0.001)	[0.1599]
	1940–1949	0.11	0.000	0.004	0.001	3.9
			(0.001)	(0.001)	(0.001)	[0.0091]
Completed doctoral degree	1930–1939	0.03	0.002	0.003	0.000	2.9
			(0.001)	(0.001)	(0.001)	[0.0332]
	1940–1949	0.04	−0.002	0.001	−0.001	4.3
			(0.001)	(0.001)	(0.001)	[0.0050]

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$. What evidence do they offer?

- Table 1: effect doesn't work through higher ed, suggesting compulsory schooling
- Table 2: laws are important, and they used to be more compelling
  - what's the hypothesis they are trying to test here? and method?

TABLE II
PERCENTAGE OF AGE GROUP ENROLLED IN SCHOOL BY BIRTHDAY AND LEGAL DROPOUT AGE[a]

	Type of state law[b]		
Date of birth	School-leaving age: 16 (1)	School-leaving age: 17 or 18 (2)	Column (1) − (2)
	Percent enrolled April 1, 1960		
1. Jan 1–Mar 31, 1944	87.6	91.0	−3.4
(age 16)	(0.6)	(0.9)	(1.1)
2. Apr 1–Dec 31, 1944	92.1	91.6	0.5
(age 15)	(0.3)	(0.5)	(0.6)
3. Within-state diff.	−4.5	−0.6	−4.0
(row 1 − row 2)	(0.7)	(1.0)	(1.2)

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$. What evidence do they offer?

- Table 1: effect doesn't work through higher ed, suggesting compulsory schooling
- Table 2: laws are important, and they used to be more compelling
  - what's the hypothesis they are trying to test here? and method?
  - QOB matters for attendance at age 16 when school leaving age is 16, but not when school leaving age is 17 or 18 – suggesting that the law is the mechanism

TABLE II
PERCENTAGE OF AGE GROUP ENROLLED IN SCHOOL BY BIRTHDAY AND LEGAL DROPOUT AGE[a]

Date of birth	Type of state law[b]: School-leaving age: 16 (1)	School-leaving age: 17 or 18 (2)	Column (1) − (2)
	Percent enrolled April 1, 1960		
1. Jan 1–Mar 31, 1944	87.6	91.0	−3.4
(age 16)	(0.6)	(0.9)	(1.1)
2. Apr 1–Dec 31, 1944	92.1	91.6	0.5
(age 15)	(0.3)	(0.5)	(0.6)
3. Within-state diff.	−4.5	−0.6	−4.0
(row 1 − row 2)	(0.7)	(1.0)	(1.2)

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$ What evidence do they offer?

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$ What evidence do they offer?

- Sec 3, p. 1005 is dedicated to this.
  - suggest that argument that students that start older do better only works against their finding
  - argue that there is no relationship between socio-economic status and quarter of birth
  - say that they look for season of birth effects on earnings for college graduates, and find nothing

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$ What evidence do they offer?

- Sec 3, p. 1005 is dedicated to this.
  - suggest that argument that students that start older do better only works against their finding
  - argue that there is no relationship between socio-economic status and quarter of birth
  - say that they look for season of birth effects on earnings for college graduates, and find nothing
- What if people who are born in the first quarter are more likely to be mentally ill? should this challenge their estimation strategy?

Showing the Hard IV Condition: $\text{corr}(Z, \epsilon) = 0$

IV condition No. 2: $\text{corr}(Z, \epsilon) = 0$ What evidence do they offer?

- Sec 3, p. 1005 is dedicated to this.
  - suggest that argument that students that start older do better only works against their finding
  - argue that there is no relationship between socio-economic status and quarter of birth
  - say that they look for season of birth effects on earnings for college graduates, and find nothing
- What if people who are born in the first quarter are more likely to be mentally ill? should this challenge their estimation strategy?
  - no, if the effect works through education
  - yes, if it works through pathways other than education

4. Results

Interpreting the Results: Wald Estimate

OLS estimate:

$$\hat{\beta}_{OLS} = \frac{\Delta Y}{\Delta X}$$

Interpreting the Results: Wald Estimate

OLS estimate:

$$\hat{\beta}_{OLS} = \frac{\Delta Y}{\Delta X}$$

Wald estimate:

$$\hat{\beta}_{IV} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(X|Z=1) - E(X|Z=0)}$$

Interpreting the Results: Wald Estimate

OLS estimate:

$$\hat{\beta}_{OLS} = \frac{\Delta Y}{\Delta X}$$

Wald estimate:

$$\hat{\beta}_{IV} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(X|Z=1) - E(X|Z=0)}$$

TABLE III

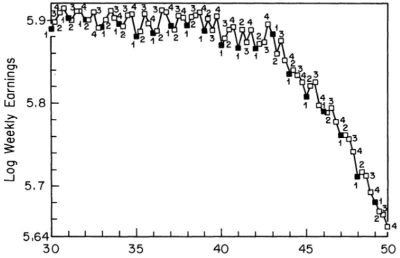
PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 1920–1929[a]

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year	(3) Difference (std. error) (1) − (2)
ln (wkly. wage)	5.1484	5.1574	−0.00898 (0.00301)
Education	11.3996	11.5252	−0.1256 (0.0155)
Wald est. of return to education			0.0715 (0.0219)
OLS return to education[b]			0.0801 (0.0004)

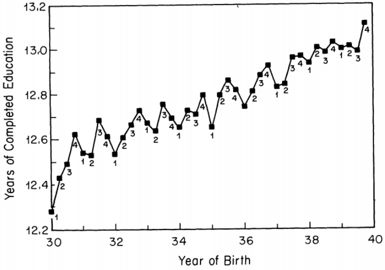
Visual Wald Estimate

Divide changes due to QOB in Figure 5 by changes due to QOB in Figure 1

Numerator: ΔY due to ΔZ



Denominator: ΔX due to ΔZ



Interpreting the Results: OLS and IV

- Interpret OLS coefficient
- Interpret IV coefficient

TABLE IV
OLS AND TSLS ESTIMATES OF THE RETURN TO EDUCATION FOR MEN BORN 1920–1929: 1970 CENSUS[a]

Independent variable	(1) OLS	(2) TSLS	(3) OLS	(4) TSLS	(5) OLS	(6) TSLS	(7) OLS	(8) TSLS
Years of education	0.0802	0.0769	0.0802	0.1310	0.0701	0.0669	0.0701	0.1007
	(0.0004)	(0.0150)	(0.0004)	(0.0334)	(0.0004)	(0.0151)	(0.0004)	(0.0334)
Race (1 = black)	—	—	—	—	0.2980	−0.3055	−0.2980	−0.2271
					(0.0043)	(0.0353)	(0.0043)	(0.0776)
SMSA (1 = center city)	—	—	—	—	0.1343	0.1362	0.1343	0.1163
					(0.0026)	(0.0092)	(0.0026)	(0.0198)
Married (1 = married)	—	—	—	—	0.2928	0.2941	0.2928	0.2804
					(0.0037)	(0.0072)	(0.0037)	(0.0141)
9 Year-of-birth dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8 Region of residence dummies	No	No	No	No	Yes	Yes	Yes	Yes
Age	—	—	0.1446	0.1409	—	—	0.1162	0.1170
			(0.0676)	(0.0704)			(0.0652)	(0.0662)
Age-squared	—	—	−0.0015	−0.0014	—	—	−0.0013	−0.0012
			(0.0007)	(0.0008)			(0.0007)	(0.0007)
χ^2 [dof]	—	36.0 [29]	—	25.6 [27]	—	34.2 [29]	—	28.8 [27]

5. A and K, 30 Years On

A and K, 30 Years On

- Compulsory schooling part still has traction
- Quarter of birth part not as much
- Subject to a scathing critique (BBJ, 1995), but method much copied
  - quarter of birth impacts school performance
  - health differences by quarter of birth
  - regional patterns in quarter of birth
- Next class: we'll discuss problems of using many and weak instruments

Next Lecture

- Read
  - the instrument paper to which you're assigned
  - trade if you'd like
  - skim the intro of the other
- Summary due next week if you're on the list