Instrumental Variables

Prof. Mike Aguilar

https:/www.linkedin.com/in/mike-aguilar-econ

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Instrumental Variables - Summary

- IV is an approach to mitigate OVB
- Within our hierarchy of methods, IV is an Observational Study
- It can be very powerful, but should not be used unless necessary
- Valid instruments (relevant, independent, exogenous) can be difficult to find

Housekeeping

- Throughout the slide deck you will see "Q", which indicates a question to you, the reader
- You will also see "A", which indicates the associated answer
- It is generally most efficient to learn this material through active participation. Whenever you encounter a "Q", be sure to try and develop your answer before turning the page to the provided "A" answer
- I recommend reading through these slides before engaging in associated coding exercises
- This slide deck is part of a broader curriculum in causal estimation and inference
- Suggested pre-requisite: foundations of regression, omitted variables bias

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Private School vs Public School

Your team is hired by the Development Officer of the national council of Private Schools, who wants to be able to say that students who attend private schools have standardized test scores x percentage points higher than students who attend public school. Your task is to find the x.



Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome:

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:Attending Private School
- Treated:

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:Attending Private School
- Treated: Students who attended a Private School
- Control:

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:Attending Private School
- Treated: Students who attended a Private School
- Control: Students who attended public school
- Treatment Effect:

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:Attending Private School
- Treated: Students who attended a Private School
- Control: Students who attended public school
- Treatment Effect: The impact of private school on test scores
- Confounder:

Case Study Set Up

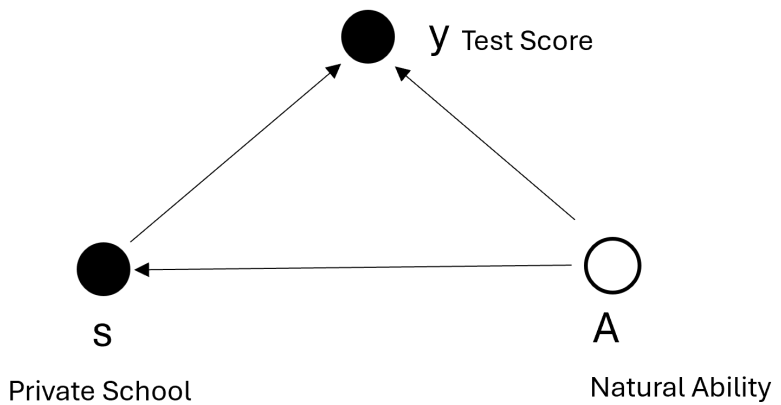
Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:Attending Private School
- Treated: Students who attended a Private School
- Control: Students who attended public school
- Treatment Effect: The impact of private school on test scores
- Confounder: Natural ability (among others)

Case Study Set Up

Outcome? Treatment? Treated? Control? Treatment Effect? Confounders? DAG?

- Outcome: Test Scores
- Treatment:Attending Private School
- Treated: Students who attended a Private School
- Control: Students who attended public school
- Treatment Effect: The impact of private school on test scores
- Confounder: Natural ability (among others)



The IV Solution

Natural Ability is an omitted variable.

Suppose there is no viable proxy available for Ability.

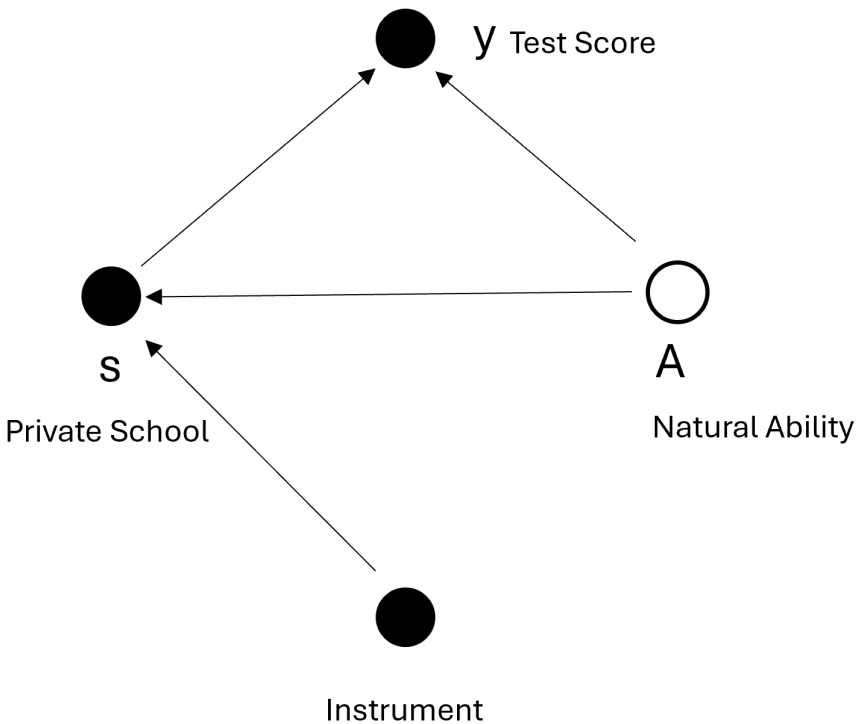
Instrumental Variables (IV) may be a solution.

The IV Solution

Natural Ability is an omitted variable.

Suppose there is no viable proxy available for Ability.

Instrumental Variables (IV) may be a solution.



Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Specifying the Schooling Problem

$$y = \beta_0 + \beta_1 s + \gamma A + e$$

where

- $y =$ Test Score
- $s =$ Dummy=1 if Private; 0 if Public
- $A =$ Ability (unobserved)
- $e =$ Error

Specifying the Schooling Problem

$$y = \beta_0 + \beta_1 s + \gamma A + e$$

where

- $y =$ Test Score
- $s =$ Dummy=1 if Private; 0 if Public
- $A =$ Ability (unobserved)
- $e =$ Error

Q: Which parameter is the ATE?

Specifying the Schooling Problem

$$y = \beta_0 + \beta_1 s + \gamma A + e$$

where

- $y =$ Test Score
- $s =$ Dummy=1 if Private; 0 if Public
- $A =$ Ability (unobserved)
- $e =$ Error

Q: Which parameter is the ATE?
A: β_1

Misspecifying the Schooling Problem

Since Ability is unobserved, we can only estimate

$$\begin{aligned} y &= \beta_0 + \beta_1 s + \nu \\ \nu &= \beta_2 A + e \end{aligned}$$

Misspecifying the Schooling Problem

Since Ability is unobserved, we can only estimate

$$\begin{aligned} y &= \beta_0 + \beta_1 s + \nu \\ \nu &= \beta_2 A + e \end{aligned}$$

Ability may impact test scores directly. However, Ability may also be related to schooling

Misspecifying the Schooling Problem

Since Ability is unobserved, we can only estimate

$$\begin{aligned} y &= \beta_0 + \beta_1 s + \nu \\ \nu &= \beta_2 A + e \end{aligned}$$

Ability may impact test scores directly. However, Ability may also be related to schooling (perhaps more able students seek more/better education via private schooling).

Misspecifying the Schooling Problem

Since Ability is unobserved, we can only estimate

$$\begin{aligned} y &= \beta_0 + \beta_1 s + \nu \\ \nu &= \beta_2 A + e \end{aligned}$$

Ability may impact test scores directly. However, Ability may also be related to schooling (perhaps more able students seek more/better education via private schooling).

The regression $y = \beta_0 + \beta_1 s + \nu$ is misspecified. We can't use OLS to determine β_1. The regressor (s) is correlated with the error (ν), inducing an Omitted Variable Bias.

Using the Mastering Metrics Notation

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

When studying OVB found that $\hat{\beta}_1^s = \hat{\beta}_1^l + \hat{\pi}_1 \hat{\gamma}^l$ where $\hat{\pi}_1$ comes from the auxiliary regression

Using the Mastering Metrics Notation

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

When studying OVB found that $\hat{\beta}_1^s = \hat{\beta}_1^l + \hat{\pi}_1 \hat{\gamma}^l$ where $\hat{\pi}_1$ comes from the auxiliary regression $A = \pi_0 + \pi_1 s + u$.

Using the Mastering Metrics Notation

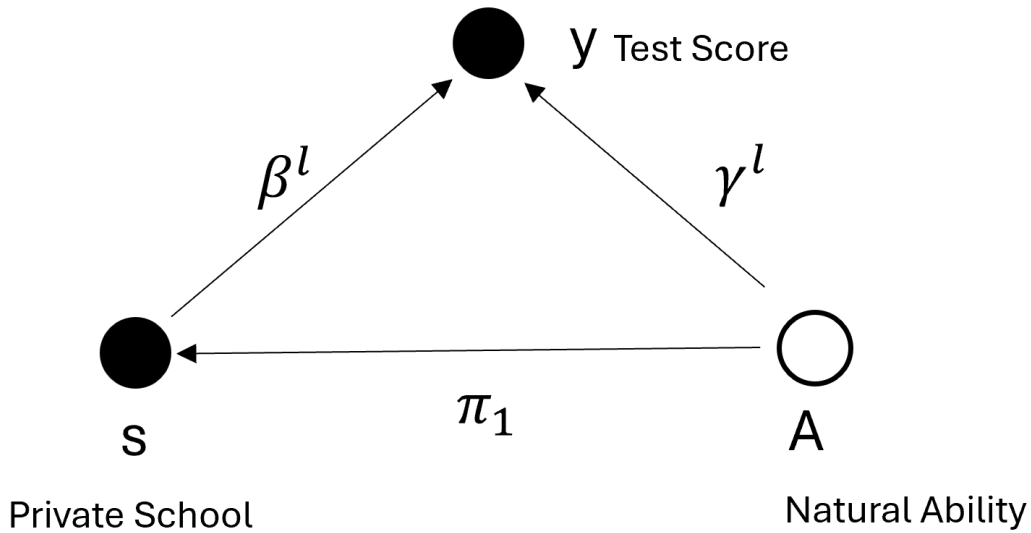
$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

When studying OVB found that $\hat{\beta}_1^s = \hat{\beta}_1^l + \hat{\pi}_1 \hat{\gamma}^l$ where $\hat{\pi}_1$ comes from the auxiliary regression $A = \pi_0 + \pi_1 s + u$.

$OVB =$ $\hat{\pi}_1 \hat{\gamma}^l$

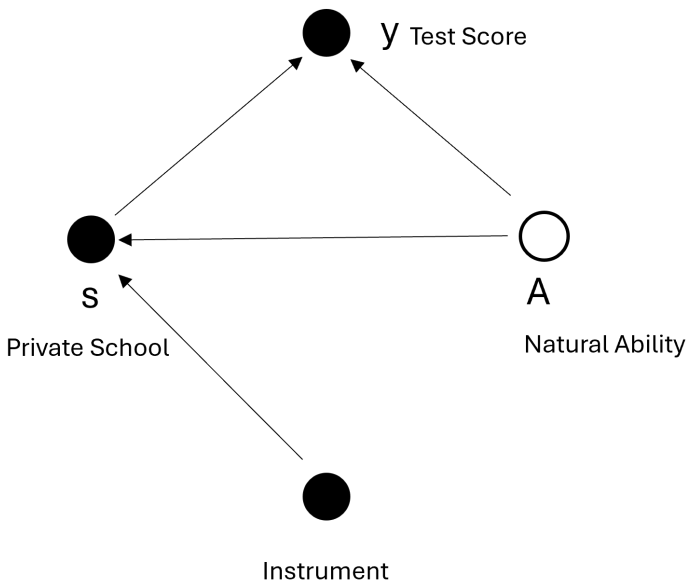
$OVB =$ Regression of Omitted on Included $\times$ Effect of Omitted in Long

Visualizing the Bias within a DAG

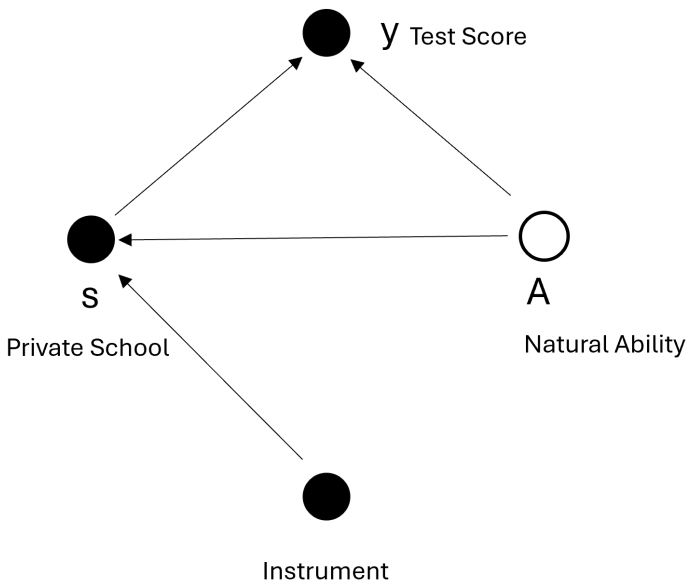


CAUTION: The auxiliary regression that produced π_1 was not causal. The existence of that arrow connecting ability and private school is more important than its direction.

The IV Solution

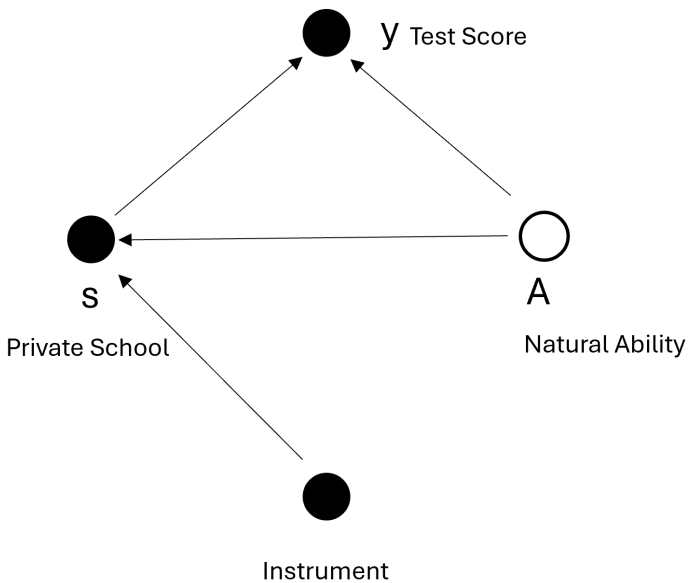


The IV Solution



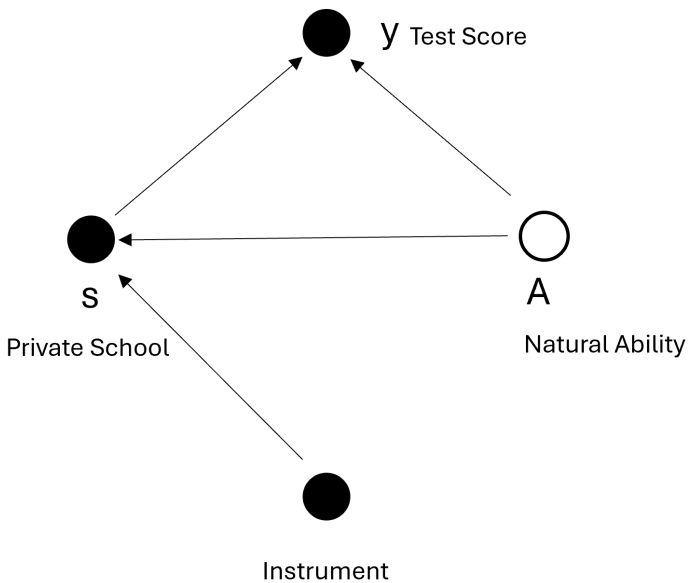
- Given the omitted Ability, Private School is Endogenous.

The IV Solution



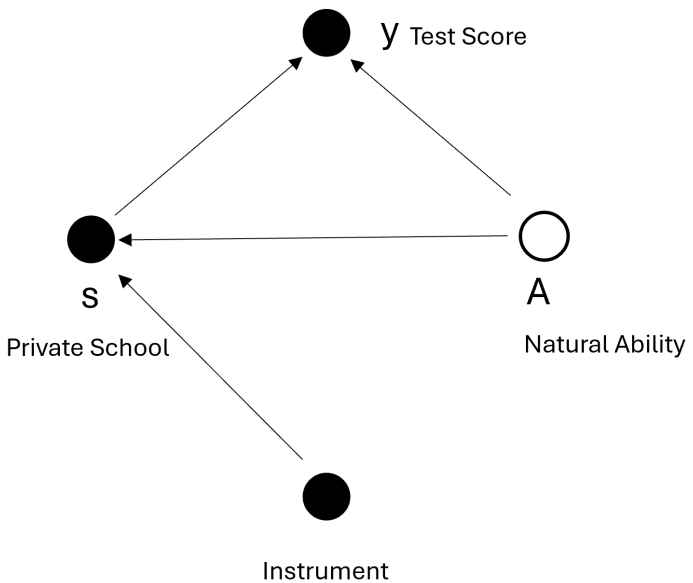
- Given the omitted Ability, Private School is Endogenous.
- We seek an "instrument" for Private School

The IV Solution



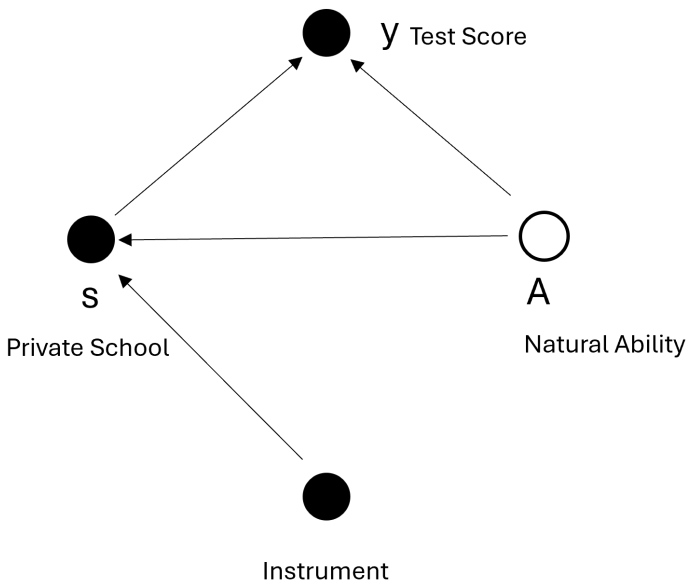
- Given the omitted Ability, Private School is Endogenous.
- We seek an "instrument" for Private School
- The instrument needs very special properties.

The IV Solution



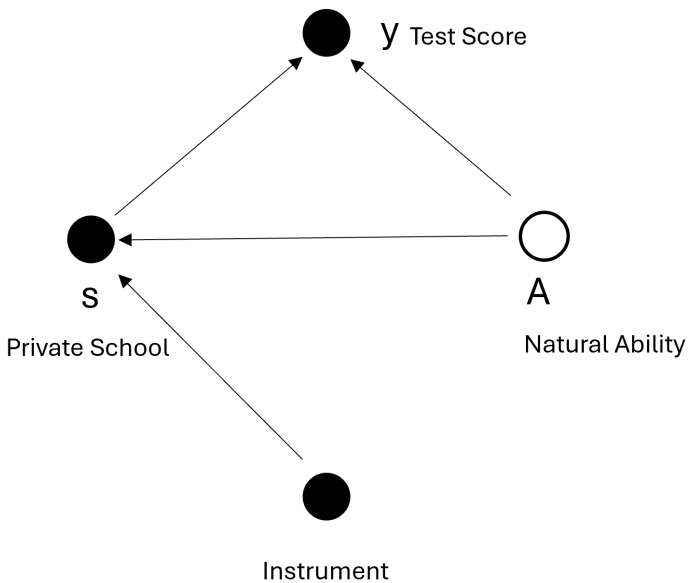
- Given the omitted Ability, Private School is Endogenous.
- We seek an "instrument" for Private School
- The instrument needs very special properties.
- Valid instrument are difficult to find.

The IV Solution



- Given the omitted Ability, Private School is Endogenous.
- We seek an "instrument" for Private School
- The instrument needs very special properties.
- Valid instrument are difficult to find.
- If you can find an instrument, you can bypass the OVB problem.

The IV Solution



- Given the omitted Ability, Private School is Endogenous.
- We seek an "instrument" for Private School
- The instrument needs very special properties.
- Valid instrument are difficult to find.
- If you can find an instrument, you can bypass the OVB problem.
- But if your instrument is not valid, IV can make your problems worse.

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Instrument Properties

We seek an instrument z for the treatment variable s

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \quad \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \quad \nu \end{aligned}$$

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \quad \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \quad \nu \end{aligned}$$

Q: What are the conditions for an instrument (z) to be valid?

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \quad \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \quad \nu \end{aligned}$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction:

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction: instrument only related to outcome via treatment

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction: instrument only related to outcome via treatment
   - Independence restriction:

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction: instrument only related to outcome via treatment
   - Independence restriction: instrument not related to omitted variable

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction: instrument only related to outcome via treatment
   - Independence restriction: instrument not related to omitted variable
2. $Cov(z, s) \neq 0$

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction: instrument only related to outcome via treatment
   - Independence restriction: instrument not related to omitted variable
2. $Cov(z, s) \neq 0$ (Instrument Relevance)

Instrument Properties

We seek an instrument z for the treatment variable s

Recall

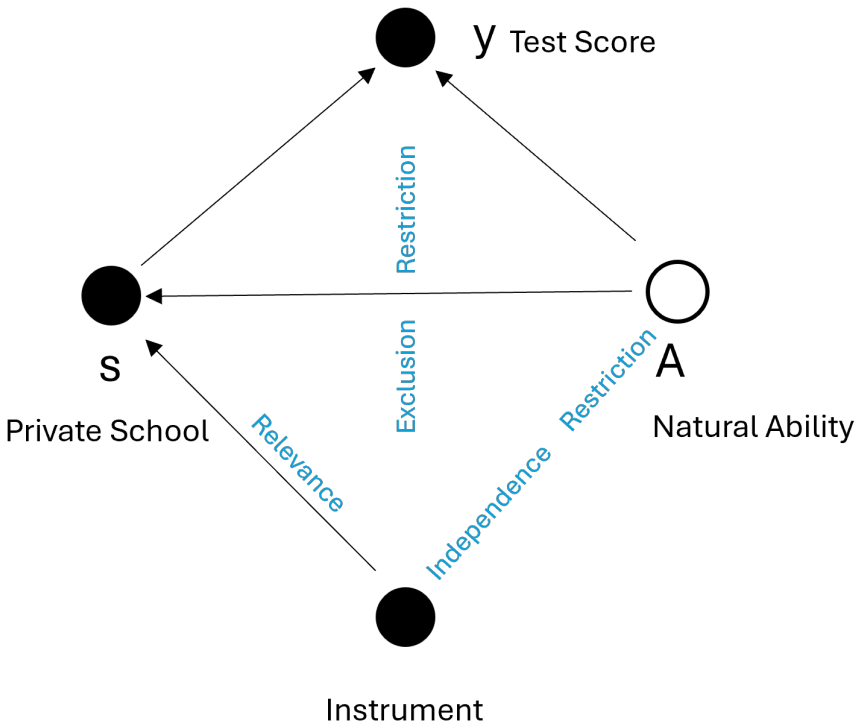
$$Long: y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \quad \nu$$

Q: What are the conditions for an instrument (z) to be valid?
A:

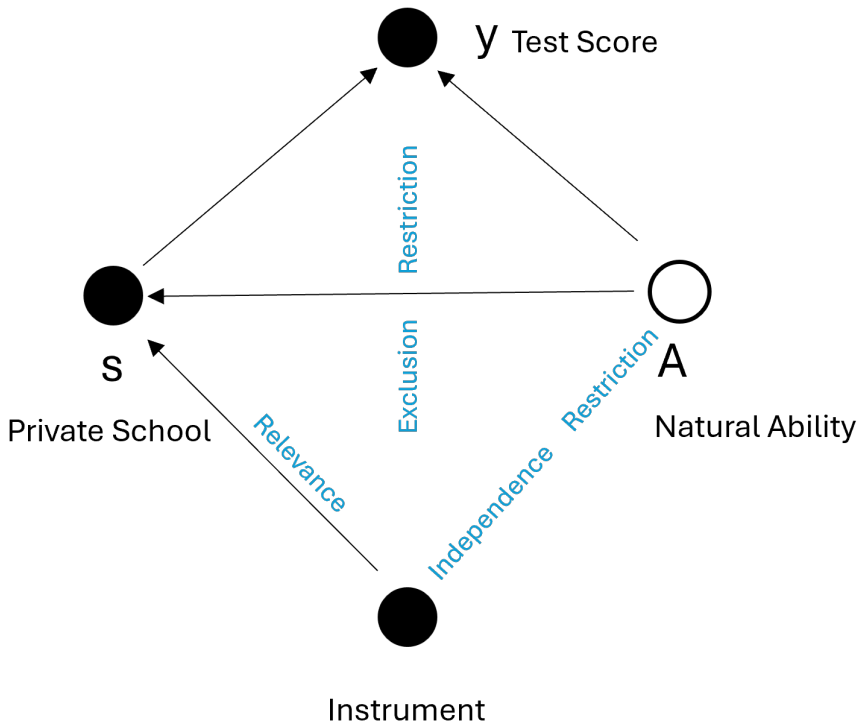
1. $Cov(z, \nu) = 0$ (Instrument Exogeneity)
   - Exclusion restriction: instrument only related to outcome via treatment
   - Independence restriction: instrument not related to omitted variable
2. $Cov(z, s) \neq 0$ (Instrument Relevance)
   - Instrument is related to the treatment

Q: Can you relate these instrument properties to our case?

Q: Can you relate these instrument properties to our case?

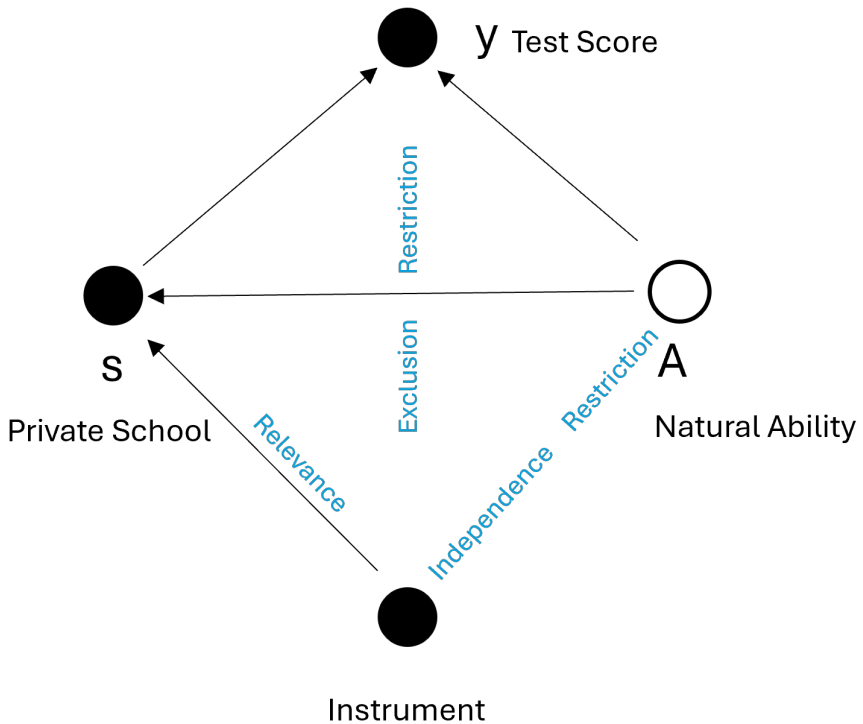


Q: Can you relate these instrument properties to our case?



Q: Can you think of a suitable instrument for our case study?

Q: Can you relate these instrument properties to our case?



Q: Can you think of a suitable instrument for our case study?

A: Possible instrument: Suppose a town held a Random Lottery. Winners were admitted into the private school with all tuition paid.

Validating the lottery as an instrument

Is the lottery a valid instrument?

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance:

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance: the lottery impacts private school attendance

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance: the lottery impacts private school attendance
- Exclusion:

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance: the lottery impacts private school attendance
- Exclusion: being chosen for the lottery doesn't impact test scores after we have controlled for the impact that attending the private school has on test scores

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance: the lottery impacts private school attendance
- Exclusion: being chosen for the lottery doesn't impact test scores after we have controlled for the impact that attending the private school has on test scores
- Independence:

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance: the lottery impacts private school attendance
- Exclusion: being chosen for the lottery doesn't impact test scores after we have controlled for the impact that attending the private school has on test scores
- Independence: since the lottery is random, it doesn't impact omitted student ability

Validating the lottery as an instrument

Is the lottery a valid instrument?

A:

- Relevance: the lottery impacts private school attendance
- Exclusion: being chosen for the lottery doesn't impact test scores after we have controlled for the impact that attending the private school has on test scores
- Independence: since the lottery is random, it doesn't impact omitted student ability

Let's explore each of the three conditions more carefully

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \quad \nu$$

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \quad \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \qquad \nu \end{aligned}$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?
A:

- $Cov(z, A) = 0 \Rightarrow$

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?
A:

- $Cov(z, A) = 0 \Rightarrow$ instrument uncorrelated with omitted variable which in our case study means

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?
A:

- $Cov(z, A) = 0 \Rightarrow$ instrument uncorrelated with omitted variable which in our case study means the lottery is not related to student ability

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?
A:

- $Cov(z, A) = 0 \Rightarrow$ instrument uncorrelated with omitted variable which in our case study means the lottery is not related to student ability
- $Cov(z, e) = 0 \Rightarrow$

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?
A:

- $Cov(z, A) = 0 \Rightarrow$ instrument uncorrelated with omitted variable which in our case study means the lottery is not related to student ability
- $Cov(z, e) = 0 \Rightarrow$ instrument uncorrelated with y once control for s, which in our case study means

Instrument Exogeneity

The Instrument Exogeneity condition states that $Cov(z, \nu) = 0$
Recall

$$Long : y = \beta_0^l + \beta_1^l s + \quad \gamma^l A + e$$
$$Short : y = \beta_0^s + \beta_1^s s + \qquad \nu$$

Notice: $Cov(z, \nu) = \gamma^l Cov(z, A) + Cov(z, e) = 0$
Q: What does $Cov(z, A) = 0$ and $Cov(z, e) = 0$ mean?
A:

- $Cov(z, A) = 0 \Rightarrow$ instrument uncorrelated with omitted variable which in our case study means the lottery is not related to student ability
- $Cov(z, e) = 0 \Rightarrow$ instrument uncorrelated with y once control for s, which in our case study means the lottery is not related to test scores once we've controlled for private school attendance

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?
A:

- If the $Cov(z, s)$ is too high,

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?
A:

- If the $Cov(z, s)$ is too high, the instrument may be plagued by the same issues as the endogenous variable s

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?
A:

- If the $Cov(z, s)$ is too high, the instrument may be plagued by the same issues as the endogenous variable s
- If the $Cov(z, s)$ is too low,

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?
A:

- If the $Cov(z, s)$ is too high, the instrument may be plagued by the same issues as the endogenous variable s
- If the $Cov(z, s)$ is too low, the instrument may not be relevant (i.e. Weak Instruments)

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?
A:

- If the $Cov(z, s)$ is too high, the instrument may be plagued by the same issues as the endogenous variable s
- If the $Cov(z, s)$ is too low, the instrument may not be relevant (i.e. Weak Instruments). Under weak instruments two problems plague the parameter estimates
  - high variance

Instrument Relevance

The Instrument Relevance condition states that $Cov(z, s) \neq 0$

The lottery instrument (z) must be related, either positively or negatively, to the endogenous treatment variable of private schooling (s).

Q: What is the tradeoff we face with $Cov(z, s) \neq 0$?
A:

- If the $Cov(z, s)$ is too high, the instrument may be plagued by the same issues as the endogenous variable s
- If the $Cov(z, s)$ is too low, the instrument may not be relevant (i.e. Weak Instruments). Under weak instruments two problems plague the parameter estimates
  - high variance
  - large bias

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \qquad \frac{\sigma_\nu^2}{SST_s}$$

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$

$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?
A: Low

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?
A: Low
Q: What does a lower $R_{s,z}^2$ imply about $Var(\beta_1^{IV})$?

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?
A: Low
Q: What does a lower $R_{s,z}^2$ imply about $Var(\beta_1^{IV})$?
A: Higher

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?
A: Low
Q: What does a lower $R_{s,z}^2$ imply about $Var(\beta_1^{IV})$?
A: Higher

- Weak Instruments generates high variance of the estimator, which impacts inference

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?
A: Low
Q: What does a lower $R_{s,z}^2$ imply about $Var(\beta_1^{IV})$?
A: Higher

- Weak Instruments generates high variance of the estimator, which impacts inference
- The variance of the IV estimator is (almost) always larger than that of OLS, suggesting that we should run OLS when possible.

Weak Instruments - High Variance

$$Var(\beta_1^{OLS}) = \frac{\sigma_\nu^2}{SST_s}$$
$$Var(\beta_1^{IV}) = \frac{\sigma_\nu^2}{SST_s \times R_{s,z}^2}$$

$R_{s,z}^2$ comes from an auxiliary regression of s upon z.
Q: Does a weak instrument have a high/low $R_{s,z}^2$?
A: Low
Q: What does a lower $R_{s,z}^2$ imply about $Var(\beta_1^{IV})$?
A: Higher

- Weak Instruments generates high variance of the estimator, which impacts inference
- The variance of the IV estimator is (almost) always larger than that of OLS, suggesting that we should run OLS when possible.
- We will develop a test of Weak Instruments later.

Weak Instruments - Bias

Weak Instruments also generates a bias of the estimator.

Weak Instruments - Bias

Weak Instruments also generates a bias of the estimator.
Suppose the Exogeneity assumption is violated (i.e. $Cov(z, \nu) \neq 0$). Then we can show that (NOTE: you don't need to know the derivation)

$$\beta_1^{s,OLS} \rightarrow \beta_1 + corr(s, \nu) \times \frac{\sigma_\nu}{\sigma_s}$$

$$\beta_1^{s,IV} \rightarrow \beta_1 + \frac{corr(z, \nu)}{corr(z, s)} \times \frac{\sigma_\nu}{\sigma_s}$$

Q: Can you fill in the sentence? If the instrument is "weak" $corr(z, s)$ is large/small, which makes the bias in the IV estimator large/small?

Weak Instruments - Bias

Weak Instruments also generates a bias of the estimator.
Suppose the Exogeneity assumption is violated (i.e. $Cov(z, \nu) \neq 0$). Then we can show that (NOTE: you don't need to know the derivation)

$$\beta_1^{s,OLS} \rightarrow \beta_1 + corr(s, \nu) \times \frac{\sigma_\nu}{\sigma_s}$$

$$\beta_1^{s,IV} \rightarrow \beta_1 + \frac{corr(z, \nu)}{corr(z, s)} \times \frac{\sigma_\nu}{\sigma_s}$$

Q: Can you fill in the sentence? If the instrument is "weak" $corr(z, s)$ is large/small, which makes the bias in the IV estimator large/small?
A: If the instrument is weak $corr(z, s)$ is small, which makes the bias in the IV estimator large.

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

- $Cov(\beta_0^l, z) = 0$ because

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

- $Cov(\beta_0^l, z) = 0$ because β_0 is a constant

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

- $Cov(\beta_0^l, z) = 0$ because β_0 is a constant
- $\gamma^l Cov(A, z) = 0$ because

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

- $Cov(\beta_0^l, z) = 0$ because β_0 is a constant
- $\gamma^l Cov(A, z) = 0$ because of instrument exogeneity (independence)

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

- $Cov(\beta_0^l, z) = 0$ because β_0 is a constant
- $\gamma^l Cov(A, z) = 0$ because of instrument exogeneity (independence)
- $Cov(e, z) = 0$ because

Deriving the IV Estimator

Begin with the "long" equation

$$y = \beta_0^l + \beta_1^l s + \beta_2^l A + e$$

Take the covariance of both sides with respect to the instrument

$$Cov(y, z) = Cov(\beta_0^l, z) + \beta_1^l Cov(s, z) + \gamma^l Cov(A, z) + Cov(e, z)$$
$$Cov(y, z) = 0 + \beta_1^l Cov(s, z) + 0 + 0$$

Q: Can you explain the prior line of math? A:

- $Cov(\beta_0^l, z) = 0$ because β_0 is a constant
- $\gamma^l Cov(A, z) = 0$ because of instrument exogeneity (independence)
- $Cov(e, z) = 0$ because of instrument exogeneity (exclusion)

Deriving the IV Estimator

Continuing from where we left off

$$Cov(y, z) = \beta_1^I Cov(s, z)$$

Deriving the IV Estimator

Continuing from where we left off

$$Cov(y, z) = \beta_1^I Cov(s, z)$$

$$\boxed{\beta_1^{I,IV} = \frac{Cov(y, z)}{Cov(s, z)}}$$

Deriving the IV Estimator

Continuing from where we left off

$$Cov(y, z) = \beta_1^I Cov(s, z)$$

$$\boxed{\beta_1^{I,IV} = \frac{Cov(y, z)}{Cov(s, z)}}$$

Note: The IV estimator for the intercept is nearly identical to that of OLS: $\beta_0^{I,IV} = \bar{y} - \beta_1^{I,IV} \bar{s}$

Deriving the IV Estimator

Continuing from where we left off

$$Cov(y, z) = \beta_1^I Cov(s, z)$$

$$\boxed{\beta_1^{I,IV} = \frac{Cov(y, z)}{Cov(s, z)}}$$

Note: The IV estimator for the intercept is nearly identical to that of OLS: $\beta_0^{I,IV} = \bar{y} - \beta_1^{I,IV} \bar{s}$

IV does NOT use the instrument as a proxy for the endogenous treatment.

Deriving the IV Estimator

Continuing from where we left off

$$Cov(y, z) = \beta_1^I Cov(s, z)$$

$$\boxed{\beta_1^{I,IV} = \frac{Cov(y, z)}{Cov(s, z)}}$$

Note: The IV estimator for the intercept is nearly identical to that of OLS: $\beta_0^{I,IV} = \bar{y} - \beta_1^{I,IV} \bar{s}$

IV does NOT use the instrument as a proxy for the endogenous treatment. IV uses the instrument to isolate the parts of the treatment that are unrelated to the omitted variable.

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Implementing IV

- There are several ways to implement IV
- We will cover a few, but not all (e.g. we won't cover method of moments approaches)
- They each produce the same/similar coefficient estimates, but standard errors may differ

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

IV#1: Direct

We could compute the IV estimator directly from

$$\beta_1^{I,IV} = \frac{Cov(y,z)}{Cov(s,z)}$$

IV#1: Direct

We could compute the IV estimator directly from

$$\beta_1^{I,IV} = \frac{Cov(y,z)}{Cov(s,z)}$$

Q: What's a weakness of this approach?

IV#1: Direct

We could compute the IV estimator directly from

$$\beta_1^{I,IV} = \frac{Cov(y,z)}{Cov(s,z)}$$

Q: What's a weakness of this approach?
A: Inference isn't easy.

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

A: They are coefficients from auxiliary regressions.

- Reduced Form:

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

A: They are coefficients from auxiliary regressions.

- Reduced Form: $y = \alpha_0 + \alpha_1 z + e \rightarrow \alpha_1 = Cov(y,z)/Var(z)$

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

A: They are coefficients from auxiliary regressions.

- Reduced Form: $y = \alpha_0 + \alpha_1 z + e \rightarrow \alpha_1 = Cov(y,z)/Var(z)$
- First Stage:

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

A: They are coefficients from auxiliary regressions.

- Reduced Form: $y = \alpha_0 + \alpha_1 z + e \rightarrow \alpha_1 = Cov(y,z)/Var(z)$
- First Stage:$s = \gamma_0 + \gamma_1 z + e \rightarrow \gamma_1 = Cov(s,z)/Var(z)$

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

A: They are coefficients from auxiliary regressions.

- Reduced Form: $y = \alpha_0 + \alpha_1 z + e \rightarrow \alpha_1 = Cov(y,z)/Var(z)$
- First Stage:$s = \gamma_0 + \gamma_1 z + e \rightarrow \gamma_1 = Cov(s,z)/Var(z)$

Q: What is a potential weakness of this approach?

First Stage & Reduced Form

Divide numerator and denominator by variance of instrument

$$\beta_1^{I,IV} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

Q: Do the numerator and denominator look familiar?

A: They are coefficients from auxiliary regressions.

- Reduced Form: $y = \alpha_0 + \alpha_1 z + e \rightarrow \alpha_1 = Cov(y,z)/Var(z)$
- First Stage:$s = \gamma_0 + \gamma_1 z + e \rightarrow \gamma_1 = Cov(s,z)/Var(z)$

Q: What is a potential weakness of this approach?

A:

- If include observed confounders, the estimator doesn't take this nice form
- If we are lucky enough to have more than one instrument, the estimator doesn't take this nice form
- This estimator isn't as (statistically) efficient as other methods

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Estimating IV in 2 Stages

First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + error$$

Estimating IV in 2 Stages

First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + error$$

Gather the estimated endogenous variable s

$$\hat{s} = \hat{\pi}_0 + \hat{\pi}_1 z_1 + \hat{\pi}_2 z_2$$

Estimating IV in 2 Stages

First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + error$$

Gather the estimated endogenous variable s

$$\hat{s} = \hat{\pi}_0 + \hat{\pi}_1 z_1 + \hat{\pi}_2 z_2$$

Second Stage

Plug $\hat{s}$ into the Short regression

$$y = \beta_0^s + \beta_1^s \hat{s} + \nu$$

Estimating IV in 2 Stages

First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + error$$

Gather the estimated endogenous variable s

$$\hat{s} = \hat{\pi}_0 + \hat{\pi}_1 z_1 + \hat{\pi}_2 z_2$$

Second Stage

Plug $\hat{s}$ into the Short regression

$$y = \beta_0^s + \beta_1^s \hat{s} + \nu$$

The result is the Two Stage Least Squares (2SLS) estimator

$$\boxed{\hat{\beta}_1^s = \hat{\beta}_1^{2SLS}}$$

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem:

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility
- Instrument Exogeneity:

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility
- Instrument Exogeneity: $Cov(z, A) = 0$

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility
- Instrument Exogeneity: $Cov(z, A) = 0$
- First stage generates the part of s only related to instrument z

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility
- Instrument Exogeneity: $Cov(z, A) = 0$
- First stage generates the part of s only related to instrument z
- So, $\hat{s}$ is the part of schooling

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility
- Instrument Exogeneity: $Cov(z, A) = 0$
- First stage generates the part of s only related to instrument z
- So, $\hat{s}$ is the part of schooling NOT related to (A)bility

Intuition of 2SLS

Q: Why isn't β_1^s in $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ plagued by OVB?

A:

- Problem: private (s)chooling related to omitted (A)bility
- Instrument Exogeneity: $Cov(z, A) = 0$
- First stage generates the part of s only related to instrument z
- So, $\hat{s}$ is the part of schooling NOT related to (A)bility
- So, second stage $y = \beta_0^s + \beta_1^s \hat{s} + \nu$ has no OVB

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Instrument Relevance with Multiple Regressors

Suppose we have

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \beta_2^l X + \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \beta_2^s X + \nu \end{aligned}$$

where X is an additional exogenous regressor.

Instrument Relevance with Multiple Regressors

Suppose we have

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \beta_2^l X + \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \beta_2^s X + \nu \end{aligned}$$

where X is an additional exogenous regressor.
Then the notion of instrument relevance $Cov(z, s) \neq 0$ is altered slightly to a concept of partial correlation.

Instrument Relevance with Multiple Regressors

Suppose we have

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \beta_2^l X + \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \beta_2^s X + \nu \end{aligned}$$

where X is an additional exogenous regressor.
Then the notion of instrument relevance $Cov(z, s) \neq 0$ is altered slightly to a concept of partial correlation.

$$s = \pi_0 + \pi_1 z + \pi_2 X + \nu$$

Instrument Relevance with Multiple Regressors

Suppose we have

$$\begin{aligned} Long: y &= \beta_0^l + \beta_1^l s + \beta_2^l X + \gamma^l A + e \\ Short: y &= \beta_0^s + \beta_1^s s + \beta_2^s X + \nu \end{aligned}$$

where X is an additional exogenous regressor.
Then the notion of instrument relevance $Cov(z, s) \neq 0$ is altered slightly to a concept of partial correlation.

$$s = \pi_0 + \pi_1 z + \pi_2 X + \nu$$

Q: How do we interpret the requirement that $\pi_1 \neq 0$?

Instrument Relevance with Multiple Regressors

Suppose we have

$$\begin{aligned} Long: y =& \beta_0^l + \beta_1^l s + \beta_2^l X + \gamma^l A + e \\ Short: y =& \beta_0^s + \beta_1^s s + \beta_2^s X + \nu \end{aligned}$$

where X is an additional exogenous regressor.
Then the notion of instrument relevance $Cov(z, s) \neq 0$ is altered slightly to a concept of partial correlation.

$$s = \pi_0 + \pi_1 z + \pi_2 X + \nu$$

Q: How do we interpret the requirement that $\pi_1 \neq 0$?
A: After controlling for X, z and s remain correlated.

2SLS with Controls

If our short equation has other exogenous control variables (X)

2SLS with Controls

If our short equation has other exogenous control variables (X)

$$y = \beta_0^s + \beta_1^s s + \beta_2^s X + \nu$$

The 2SLS procedure follows analogously. Estimate the First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 X + error$$

2SLS with Controls

If our short equation has other exogenous control variables (X)

$$y = \beta_0^s + \beta_1^s s + \beta_2^s X + \nu$$

The 2SLS procedure follows analogously. Estimate the First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 X + \textit{error}$$

Plug the estimated s into the short equation for the Second Stage

$$y = \beta_0^s + \beta_1^s \hat{s} + \beta_2^s X + \nu$$

2SLS with Controls

If our short equation has other exogenous control variables (X)

$$y = \beta_0^s + \beta_1^s s + \beta_2^s X + \nu$$

The 2SLS procedure follows analogously. Estimate the First Stage

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 X + error$$

Plug the estimated s into the short equation for the Second Stage

$$y = \beta_0^s + \beta_1^s \hat{s} + \beta_2^s X + \nu$$

The result is the Two Stage Least Squares (2SLS) estimator

$$\boxed{\hat{\beta}_1^s = \hat{\beta}_1^{2SLS}}$$

More than one Endogenous Variable

Suppose we have a short equation with endogenous variables s_1 and s_2 (e.g. private school and GPA, both of which impact test scores and are related to the omitted ability)

$$y = \beta_0 + \beta_1 s_1 + \beta_2 s_2 + \nu$$

More than one Endogenous Variable

Suppose we have a short equation with endogenous variables s_1 and s_2 (e.g. private school and GPA, both of which impact test scores and are related to the omitted ability)

$$y = \beta_0 + \beta_1 s_1 + \beta_2 s_2 + \nu$$

To estimate the marginal effects of interest β_1 and β_2 we need to satisfy the Order Condition:

More than one Endogenous Variable

Suppose we have a short equation with endogenous variables s_1 and s_2 (e.g. private school and GPA, both of which impact test scores and are related to the omitted ability)

$$y = \beta_0 + \beta_1 s_1 + \beta_2 s_2 + \nu$$

To estimate the marginal effects of interest β_1 and β_2 we need to satisfy the Order Condition:

- Within the short equation, we need at least as many excluded exogenous variables as there are included endogenous variables

More than one Endogenous Variable

Suppose we have a short equation with endogenous variables s_1 and s_2 (e.g. private school and GPA, both of which impact test scores and are related to the omitted ability)

$$y = \beta_0 + \beta_1 s_1 + \beta_2 s_2 + \nu$$

To estimate the marginal effects of interest β_1 and β_2 we need to satisfy the Order Condition:

- Within the short equation, we need at least as many excluded exogenous variables as there are included endogenous variables

In our case, we need at least two instruments in order to estimate β_1 and β_2.

Caution over R^2

Caution over R^2

Recall,

$$R^2 = 1 - \frac{SSR}{SST}$$

where SSR is the sum of the squared IV residuals, and SST is the total sum of squares of y. In the face of endogeneity, we cannot decompose the variance of y in the usual manner. So, R^2 doesn't have the usual meaning. Unlike in OLS, in IV the SSR can be larger than the SST, meaning the R^2 can be negative.

Caution over R^2

Recall,

$$R^2 = 1 - \frac{SSR}{SST}$$

where SSR is the sum of the squared IV residuals, and SST is the total sum of squares of y. In the face of endogeneity, we cannot decompose the variance of y in the usual manner. So, R^2 doesn't have the usual meaning. Unlike in OLS, in IV the SSR can be larger than the SST, meaning the R^2 can be negative.

For these reasons, the R^2 typically isn't used when running an IV regression. Remember, the purpose of IV is assessing marginal effects, not maximizing goodness of fit.

Using IV for Errors in Variables

Suppose we are interested in the marginal effect of college major GPA (s^*) upon test scores (y)

$$y = \beta_0 + \beta_1 s^* + \nu$$

Using IV for Errors in Variables

Suppose we are interested in the marginal effect of college major GPA (s^*) upon test scores (y)

$$y = \beta_0 + \beta_1 s^* + \nu$$

where s^* meets all of the standard exogeneity assumptions. Unfortunately, we don't observe s^*, but instead we observe the student's overall GPA (s) (i.e. college major GPA is measured with error).

Using IV for Errors in Variables

Suppose we are interested in the marginal effect of college major GPA (s^*) upon test scores (y)

$$y = \beta_0 + \beta_1 s^* + \nu$$

where s^* meets all of the standard exogeneity assumptions. Unfortunately, we don't observe s^*, but instead we observe the student's overall GPA (s) (i.e. college major GPA is measured with error). Further,

$$s = s^* + u_s$$

where u_s is the measurement error.

Using IV for Errors in Variables

Suppose we are interested in the marginal effect of college major GPA (s^*) upon test scores (y)

$$y = \beta_0 + \beta_1 s^* + \nu$$

where s^* meets all of the standard exogeneity assumptions. Unfortunately, we don't observe s^*, but instead we observe the student's overall GPA (s) (i.e. college major GPA is measured with error). Further,

$$s = s^* + u_s$$

where u_s is the measurement error. Notice that

$$y = \beta_0 + \beta_1 s + (\nu - \beta_1 u_s)$$

Using IV for Errors in Variables

Suppose we are interested in the marginal effect of college major GPA (s^*) upon test scores (y)

$$y = \beta_0 + \beta_1 s^* + \nu$$

where s^* meets all of the standard exogeneity assumptions. Unfortunately, we don't observe s^*, but instead we observe the student's overall GPA (s) (i.e. college major GPA is measured with error). Further,

$$s = s^* + u_s$$

where u_s is the measurement error. Notice that

$$y = \beta_0 + \beta_1 s + (\nu - \beta_1 u_s)$$

We can see that s is related to the composite error term $(\nu - \beta_1 u_s)$, and thus OLS will produce biased results for β_1.

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.
We can use this second measurement as an instrument if

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.
We can use this second measurement as an instrument if

- $Cov(z, \nu) = 0$ (Instrument Exogeneity)

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.
We can use this second measurement as an instrument if

- $Cov(z, \nu) = 0$ (Instrument Exogeneity)

$$Cov(z, \nu) = Cov(s^* + u_z, \nu) = Cov(s^*, \nu) + Cov(u_z, \nu) = 0$$

- $Cov(z, s) \neq 0$ (Instrument Relevance)

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.
We can use this second measurement as an instrument if

- $Cov(z, \nu) = 0$ (Instrument Exogeneity)

  $$Cov(z, \nu) = Cov(s^* + u_z, \nu) = Cov(s^*, \nu) + Cov(u_z, \nu) = 0$$

- $Cov(z, s) \neq 0$ (Instrument Relevance)

  $$Cov(z, s) = Cov(s^* + u_z, s^* + u_s)$$
  $$= Cov(s^*, s^*) + Cov(s^*, u_s) + Cov(u_z, s^*) + Cov(u_z, u_s)$$

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.
We can use this second measurement as an instrument if

- $Cov(z, \nu) = 0$ (Instrument Exogeneity)
$$Cov(z, \nu) = Cov(s^* + u_z, \nu) = Cov(s^*, \nu) + Cov(u_z, \nu) = 0$$
- $Cov(z, s) \neq 0$ (Instrument Relevance)
$$Cov(z, s) = Cov(s^* + u_z, s^* + u_s)$$
$$= Cov(s^*, s^*) + Cov(s^*, u_s) + Cov(u_z, s^*) + Cov(u_z, u_s) = Var(s^*) \neq 0$$

Using IV for Errors in Variables

Consider another measure of college major GPA (like GPA from Master's program z), but suppose this too is measured with error.
Suppose

$$z = s^* + u_z$$

where u_z is an error term such that $Cov(u_z, u_s) = 0$ and $Cov(u_z, \nu) = 0$.
We can use this second measurement as an instrument if

- $Cov(z, \nu) = 0$ (Instrument Exogeneity)
$$Cov(z, \nu) = Cov(s^* + u_z, \nu) = Cov(s^*, \nu) + Cov(u_z, \nu) = 0$$
- $Cov(z, s) \neq 0$ (Instrument Relevance)
$$Cov(z, s) = Cov(s^* + u_z, s^* + u_s)$$
$$= Cov(s^*, s^*) + Cov(s^*, u_s) + Cov(u_z, s^*) + Cov(u_z, u_s) = Var(s^*) \neq 0$$

Caution: Difficult to find good instruments to solve EIV. e.g. Wouldn't we want to include Master's GPA in the "long" regression?

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

IV Testing

In this subsection we will discuss

- Do I suspect OVB?
- Do we need IV in the first place?
- Are my instruments exogenous?
- Are the instruments valid?

IV Testing

In this subsection we will discuss

- Do I suspect OVB?
- Do we need IV in the first place?
- Are my instruments exogenous?
- Are the instruments valid?

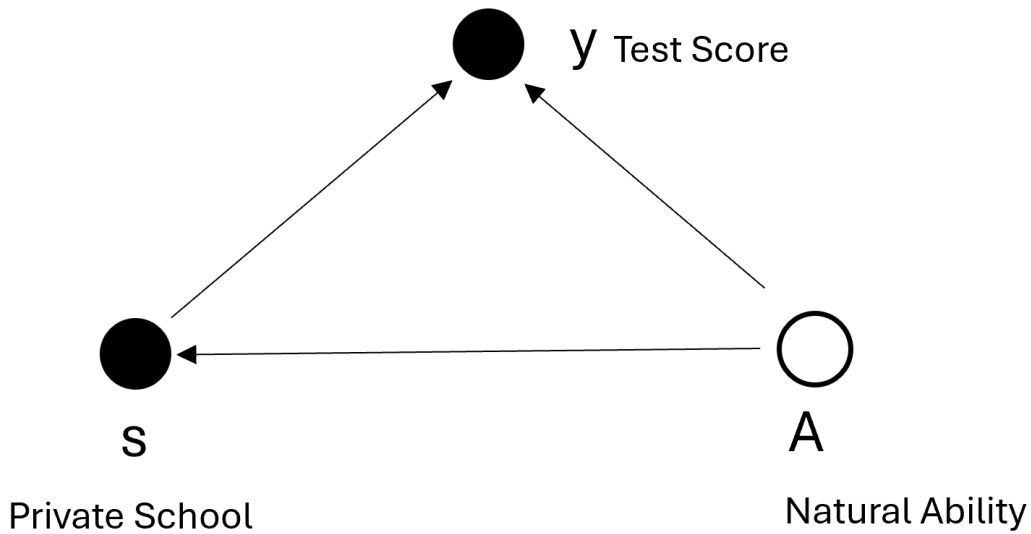
Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Q: How determine if OVB exists in my case?

Q: How determine if OVB exists in my case?

A: You can't say for sure. Need to rely upon DAG and subject matter expertise.



Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)

We will focus on the OVB case, but the approach is broadly applicable

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed?

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous?

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous? i.e. is my treatment related to my error?

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous? i.e. is my treatment related to my error? i.e. is private school attendance related to ability?

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous? i.e. is my treatment related to my error? i.e. is private school attendance related to ability?
Consider the long and short equations

$$Long: y = \beta_0^l + \beta_1^l s + \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \nu$$

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous? i.e. is my treatment related to my error? i.e. is private school attendance related to ability?
Consider the long and short equations

$$Long: y = \beta_0^l + \beta_1^l s + \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \nu$$

We saw earlier that IV might generate estimators with large standard errors, so we shouldn't apply IV unless necessary.

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous? i.e. is my treatment related to my error? i.e. is private school attendance related to ability?
Consider the long and short equations

$$Long: y = \beta_0^l + \beta_1^l s + \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \nu$$

We saw earlier that IV might generate estimators with large standard errors, so we shouldn't apply IV unless necessary.

If $Cov(s, \nu) = 0$ then we should estimate the short via OLS (in fact, then there is no OVB).

Detecting Endogeneity

OVB is one way Endogeneity can occur (e.g. measurement error, simultaneity)
We will focus on the OVB case, but the approach is broadly applicable

Is IV needed? i.e. is my treatment endogenous? i.e. is my treatment related to my error? i.e. is private school attendance related to ability?
Consider the long and short equations

$$Long: y = \beta_0^l + \beta_1^l s + \gamma^l A + e$$
$$Short: y = \beta_0^s + \beta_1^s s + \nu$$

We saw earlier that IV might generate estimators with large standard errors, so we shouldn't apply IV unless necessary.

If $Cov(s, \nu) = 0$ then we should estimate the short via OLS (in fact, then there is no OVB).
2 approaches to determine this: 1) Heuristic, 2) Test

Heuristic to Detect Endogeneity

Hausman (1978) suggests comparing the IV and OLS estimates under the notion that if all variables were exogenous then both estimators approach the true marginal effect.

Heuristic to Detect Endogeneity

Hausman (1978) suggests comparing the IV and OLS estimates under the notion that if all variables were exogenous then both estimators approach the true marginal effect.
Recall

$$\beta_1^{s,OLS} \rightarrow \beta_1 + corr(s, \nu) \times \frac{\sigma_\nu}{\sigma_s}$$

$$\beta_1^{s,IV} \rightarrow \beta_1 + \frac{corr(z, \nu)}{corr(z, s)} \times \frac{\sigma_\nu}{\sigma_s}$$

Heuristic to Detect Endogeneity

Hausman (1978) suggests comparing the IV and OLS estimates under the notion that if all variables were exogenous then both estimators approach the true marginal effect.
Recall

$$\beta_1^{s,OLS} \rightarrow \beta_1 + corr(s, \nu) \times \frac{\sigma_\nu}{\sigma_s}$$

$$\beta_1^{s,IV} \rightarrow \beta_1 + \frac{corr(z, \nu)}{corr(z, s)} \times \frac{\sigma_\nu}{\sigma_s}$$

If s were indeed exogenous (i.e. $Cov(s, \nu) = 0$), and our instruments were strong and exogenous (i.e. $Cov(z, \nu) = 0$), then IV and OLS approach the true marginal effect.

Heuristic to Detect Endogeneity

Hausman (1978) suggests comparing the IV and OLS estimates under the notion that if all variables were exogenous then both estimators approach the true marginal effect.
Recall

$$\beta_1^{s,OLS} \rightarrow \beta_1 + corr(s,\nu) \times \frac{\sigma_\nu}{\sigma_s}$$

$$\beta_1^{s,IV} \rightarrow \beta_1 + \frac{corr(z,\nu)}{corr(z,s)} \times \frac{\sigma_\nu}{\sigma_s}$$

If s were indeed exogenous (i.e. $Cov(s,\nu) = 0$), and our instruments were strong and exogenous (i.e. $Cov(z,\nu) = 0$), then IV and OLS approach the true marginal effect.

Quick Diagnostic: Estimate via OLS AND via IV. If they are similar, then you may not have an endogeneity problem.

Test to Detect Endogeneity

To test more formally, consider

$$y = \beta_0^s + \beta_1^s s + \nu$$

Test to Detect Endogeneity

To test more formally, consider

$$y = \beta_0^s + \beta_1^s s + \nu$$

We are concerned that $Cov(s, \nu) \neq 0$.

Test to Detect Endogeneity

To test more formally, consider

$$y = \beta_0^s + \beta_1^s s + \nu$$

We are concerned that $Cov(s, \nu) \neq 0$. We can consider the auxiliary regression

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$$

Test to Detect Endogeneity

To test more formally, consider

$$y = \beta_0^s + \beta_1^s s + \nu$$

We are concerned that $Cov(s, \nu) \neq 0$. We can consider the auxiliary regression

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$$

We can write the condition above as

$$Cov(s, \nu) \neq 0$$
$$Cov(\pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1, \nu) \neq 0$$

Test to Detect Endogeneity

To test more formally, consider

$$y = \beta_0^s + \beta_1^s s + \nu$$

We are concerned that $Cov(s, \nu) \neq 0$. We can consider the auxiliary regression

$$s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$$

We can write the condition above as

$$Cov(s, \nu) \neq 0$$
$$Cov(\pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1, \nu) \neq 0$$

Given the instrument exogeneity conditions we know that

$$Cov(\pi_0 + \pi_1 z_1 + \pi_2 z_2, \nu) = 0$$

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?
- Q: Why is $Cov(\textit{instument}, \textit{ability}) = 0$?

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?
- Q: Why is $Cov(\textit{instument}, \textit{ability}) = 0$? A: independence

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?
- Q: Why is $Cov(\textit{instument}, \textit{ability}) = 0$? A: independence
- So the question is whether $Cov(\textit{noise}, \textit{ability}) = 0$?

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?
- Q: Why is $Cov(\textit{instument}, \textit{ability}) = 0$? A: independence
- So the question is whether $Cov(\textit{noise}, \textit{ability}) = 0$?
- We don't observe ability, but we do observe y. Since ability is related to y we can proxy the last condition as $Cov(\textit{noise}, y) = 0$?

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?
- Q: Why is $Cov(\textit{instument}, \textit{ability}) = 0$? A: independence
- So the question is whether $Cov(\textit{noise}, \textit{ability}) = 0$?
- We don't observe ability, but we do observe y. Since ability is related to y we can proxy the last condition as $Cov(\textit{noise}, y) = 0$?
- If $Cov(\textit{noise}, y) \neq 0$, then we have an endogeneity problem.

Leaves us with the question: $Cov(v_1, \nu) = 0$?

Casting within our case study

- We are worried that $Cov(\text{private school}, \textit{ability}) \neq 0$
- Notice private school s is comprised of instrument and noise
- $Cov(\textit{instrument} + \textit{noise}, \textit{ability}) = 0$?
- $Cov(\textit{instrument}, \textit{ability}) + Cov(\textit{noise}, \textit{ability}) = 0$?
- Q: Why is $Cov(\textit{instument}, \textit{ability}) = 0$? A: independence
- So the question is whether $Cov(\textit{noise}, \textit{ability}) = 0$?
- We don't observe ability, but we do observe y. Since ability is related to y we can proxy the last condition as $Cov(\textit{noise}, y) = 0$?
- If $Cov(\textit{noise}, y) \neq 0$, then we have an endogeneity problem.
- i.e. if the part of private schooling s that's not related to the lottery instrument z aka *noise* is related to ability and test scores, then we still have an endogeneity problem.

Hausman Test for Endogeneity of Treatment Variable

1. $s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$

Hausman Test for Endogeneity of Treatment Variable

1. $s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$
2. $y = \beta_0^s + \beta_1^s s + \delta \hat{v}_1 + \text{error}$

Hausman Test for Endogeneity of Treatment Variable

1. $s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$
2. $y = \beta_0^s + \beta_1^s s + \delta \hat{v}_1 + \text{error}$
3. $H_0 : \delta = 0$ vs $H_1 : \delta \neq 0$
4. If reject, then s is endogenous (i.e. we might need IV to avoid OVB)

Hausman Test for Endogeneity of Treatment Variable

1. $s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$
2. $y = \beta_0^s + \beta_1^s s + \delta \hat{v}_1 + \text{error}$
3. $H_0 : \delta = 0$ vs $H_1 : \delta \neq 0$
4. If reject, then s is endogenous (i.e. we might need IV to avoid OVB)

We can extend to multiple potential endogenous variables via $y = \beta_0^s + \beta_1^s s_1 + \beta_2^s s_2 + \cdots + \delta_1 \hat{v}_1 + \delta_2 \hat{v}_2 + \cdots + \text{error}$ and we can test the joint significance of $\delta_1 = \delta_2 = \cdots = 0$ via an F test.

Hausman Test for Endogeneity of Treatment Variable

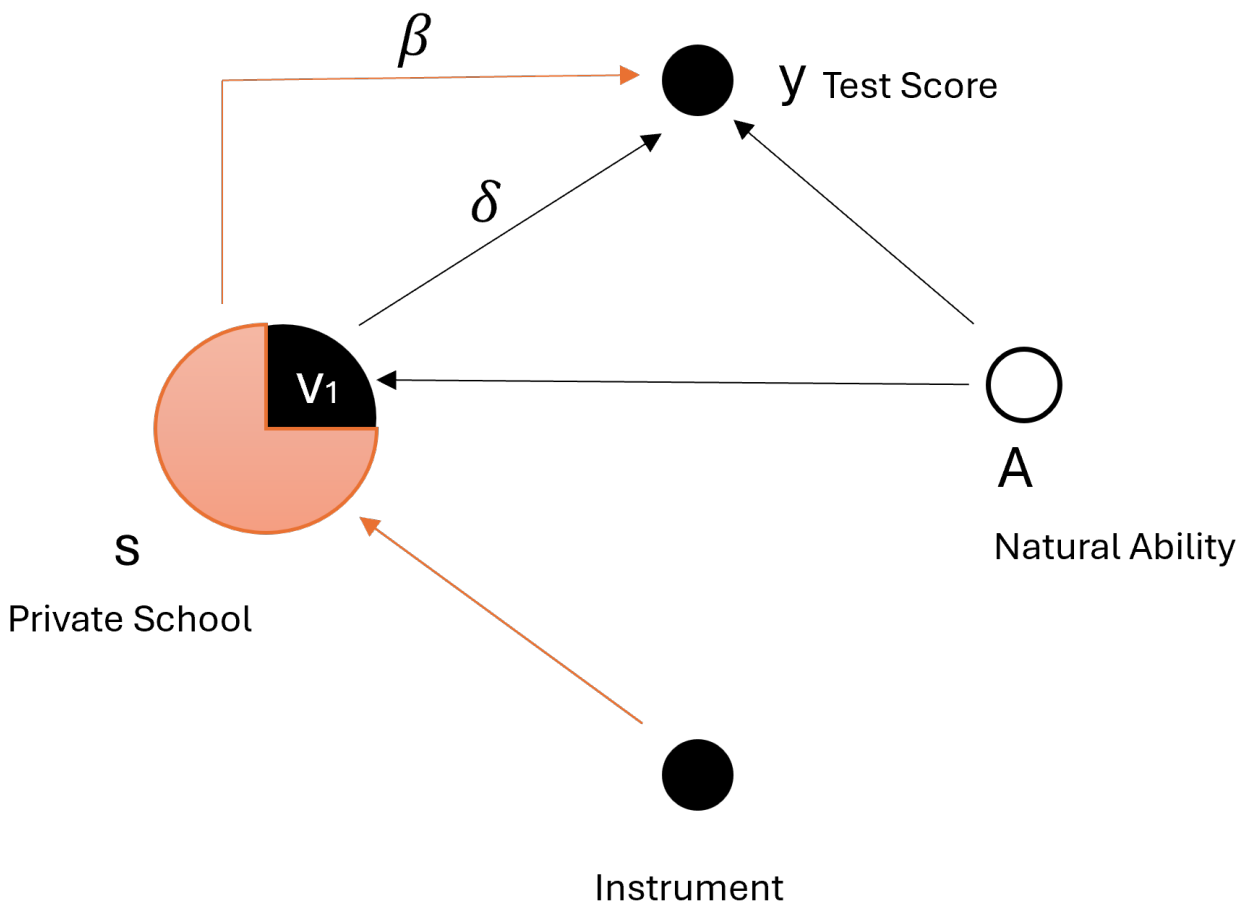
1. $s = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_1$
2. $y = \beta_0^s + \beta_1^s s + \delta \hat{v}_1 + \text{error}$
3. $H_0 : \delta = 0$ vs $H_1 : \delta \neq 0$
4. If reject, then s is endogenous (i.e. we might need IV to avoid OVB)

We can extend to multiple potential endogenous variables via $y = \beta_0^s + \beta_1^s s_1 + \beta_2^s s_2 + \cdots + \delta_1 \hat{v}_1 + \delta_2 \hat{v}_2 + \cdots + \text{error}$ and we can test the joint significance of $\delta_1 = \delta_2 = \cdots = 0$ via an F test.

Why keep the suspected endogenous variable in the second step regression? — The estimated error is the part of the endogenous variable (s) that is NOT related to the instruments. The regression y=f(v) is misspecified under the alternative. The only way that would be correct is if the instruments are exogenous, so we'd actually be conducting a joint test of exogeneity of s and z.

Visualizing the Endogeneity Test

If $\delta \neq 0$ we might have an Endogeneity problem



Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Verifying Exogeneity of Instrument

Verifying Exogeneity of Instrument

Recall Exogeneity $Cov(z, \nu) = 0$ had two conditions:

Verifying Exogeneity of Instrument

Recall Exogeneity $Cov(z, \nu) = 0$ had two conditions:

1. Q: Can you test the Independence Restriction $Cov(z, ability) = 0$?

Verifying Exogeneity of Instrument

Recall Exogeneity $Cov(z, \nu) = 0$ had two conditions:

1. Q: Can you test the Independence Restriction $Cov(z, ability) = 0$?
   A: No. Ability is unobserved.

Verifying Exogeneity of Instrument

Recall Exogeneity $Cov(z, \nu) = 0$ had two conditions:

1. Q: Can you test the Independence Restriction $Cov(z, ability) = 0$?
   A: No. Ability is unobserved.
2. Q: Can you test the Exclusion Restriction $Cov(z, y)|s = 0$?

Verifying Exogeneity of Instrument

Recall Exogeneity $Cov(z, \nu) = 0$ had two conditions:

1. Q: Can you test the Independence Restriction $Cov(z, ability) = 0$?
   A: No. Ability is unobserved.
2. Q: Can you test the Exclusion Restriction $Cov(z, y)|s = 0$?
   A: Not reliably. Our OVB problem plagues our estimates and inference.

Verifying Exogeneity of Instrument

Recall Exogeneity $Cov(z, \nu) = 0$ had two conditions:

1. Q: Can you test the Independence Restriction $Cov(z, ability) = 0$?
   A: No. Ability is unobserved.
2. Q: Can you test the Exclusion Restriction $Cov(z, y)|s = 0$?
   A: Not reliably. Our OVB problem plagues our estimates and inference.

The best we can do is rely upon subject matter expertise.

Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject, then instrument is

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject, then instrument is NOT relevant (i.e. it is

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject, then instrument is NOT relevant (i.e. it is Weak)

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject, then instrument is NOT relevant (i.e. it is Weak)
- Caution: Stock & Yogo (2001) show that the critical value for a t-test is 3.2 and for an F-test is 10.

Verifying Instrument Relevance

Weak Instruments Test

- Run auxiliary regression: $s = \pi_0 + \pi_1 z + \pi_2 X + error$
- $\pi_1 \approx Cov(s, z)/Var(z)$, and so z is Relevant if and only if $\pi_1 \neq 0$.
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject, then instrument is NOT relevant (i.e. it is Weak)
- Caution: Stock & Yogo (2001) show that the critical value for a t-test is 3.2 and for an F-test is 10.
- Easily extended to multiple instruments, each of which would have their own test of relevance. Can also extend to many exogenous features X.

Testing Summary

If you have an omitted variable, then your treatment effect is biased. IV may be a solution.

Testing Summary

If you have an omitted variable, then your treatment effect is biased. IV may be a solution.

- Do I suspect OVB? — Use subject matter expertise.

Testing Summary

If you have an omitted variable, then your treatment effect is biased. IV may be a solution.

- Do I suspect OVB? — Use subject matter expertise.
- Do I need IV? — i) Hausman Test for endogeneity. ii) If OLS and IV "close" then no need for IV.

Testing Summary

If you have an omitted variable, then your treatment effect is biased. IV may be a solution.

- Do I suspect OVB? — Use subject matter expertise.
- Do I need IV? — i) Hausman Test for endogeneity. ii) If OLS and IV "close" then no need for IV.
- Is my instrument exogenous (i.e. unrelated to y and ν?) — Cannot check $Cov(z, \nu) = 0$ easily, so rely upon subject matter expertise.

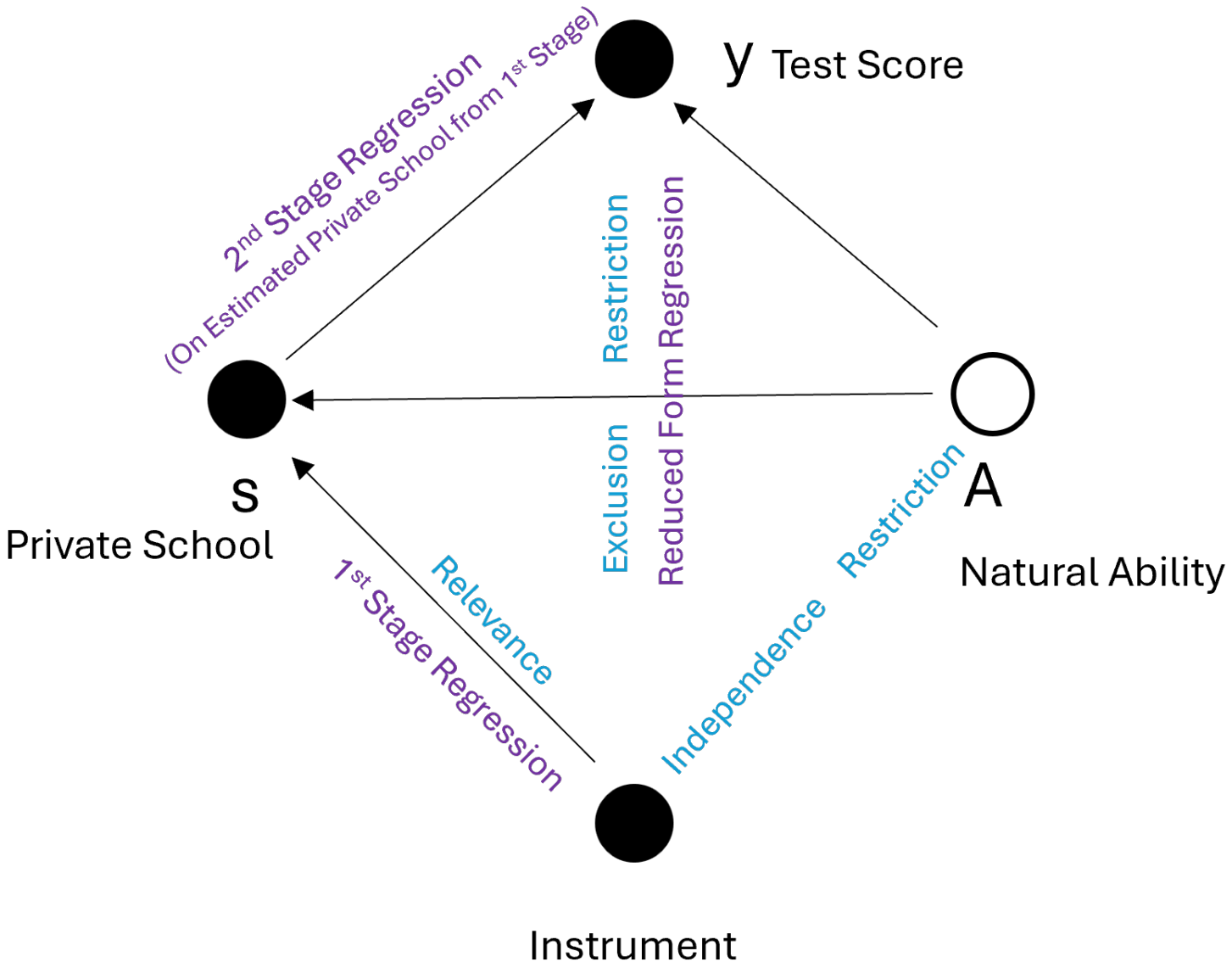
Testing Summary

If you have an omitted variable, then your treatment effect is biased. IV may be a solution.

- Do I suspect OVB? — Use subject matter expertise.
- Do I need IV? — i) Hausman Test for endogeneity. ii) If OLS and IV "close" then no need for IV.
- Is my instrument exogenous (i.e. unrelated to y and ν?) — Cannot check $Cov(z, \nu) = 0$ easily, so rely upon subject matter expertise.
- Is my instrument relevant? — If $Cov(z, s)$ small, then the instrument is "weak" which causes bias and large standard errors.

Implementation

Summary



Outline

1. Case Study Set Up
2. Understanding IV
   - How does IV "solve" OVB?
   - What makes for a good instrument?
   - What is the IV Estimator?
3. Implementing IV
   - *IV*#1: Direct
   - *IV*#2: First Stage & Reduced Form
   - *IV*#3: Two Stage Least Squares (2SLS)
   - Implementation Appendix
4. Testing IV
   - Do I suspect OVB?
   - Do I have an endogeneity problem?
   - Is my instrument exogenous?
   - Is my instrument relevant?
5. Case Study Results

KIPP Charter School

- Within Mastering Metrics we see our case in action. Let's use their notation.
- What is the impact of attending a special charter school (KIPP) on educational outcomes (e.g. test scores)?
- Outcome y: test score
- Treatment D: attending KIPP school

The Problem

$$y = \alpha + \beta D + e$$

Q: What's the problem with this specification?

The Problem

$$y = \alpha + \beta D + e$$

Q: What's the problem with this specification?
A:

- There are likely many omitted features that are relevant for explaining test scores

The Problem

$$y = \alpha + \beta D + e$$

Q: What's the problem with this specification?
A:

- There are likely many omitted features that are relevant for explaining test scores
- If these are related to D (i.e. attending the KIPP school), then we have OVB

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school.

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance $Cov(D, z) \neq 0$

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance $Cov(D, z) \neq 0$: Attendance depends strongly upon winning lottery

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance $Cov(D, z) \neq 0$: Attendance depends strongly upon winning lottery
  - Independence

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance $Cov(D, z) \neq 0$: Attendance depends strongly upon winning lottery
  - Independence $Cov(z, e) = 0$ The lottery should be unrelated to other drivers of test scores due to random nature

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance $Cov(D, z) \neq 0$: Attendance depends strongly upon winning lottery
  - Independence $Cov(z, e) = 0$ The lottery should be unrelated to other drivers of test scores due to random nature
  - Exclusion

An IV Solution

- IV might be able to help
- In order to get into KIPP students needed to win a lottery
- Consider the instrument $z = 1$ if student was randomly selected as a "winner" in the KIPP lottery and offered a spot in the charter school. (CAUTION: attendance D is not the same as random offer z)
- Q: Is this a valid instrument?
  A:
  - Relevance $Cov(D, z) \neq 0$: Attendance depends strongly upon winning lottery
  - Independence $Cov(z, e) = 0$ The lottery should be unrelated to other drivers of test scores due to random nature
  - Exclusion $Cov(z, y)|D = 0$ The lottery should be unrelated to test scores after controlling for impact of KIPP schooling

The IV Estimator

Recall from our previous notation of Reduced Form / First Stage

$$\beta^{s}_{1,IV} = \frac{Cov(y,z)}{Cov(s,z)} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

The IV Estimator

Recall from our previous notation of Reduced Form / First Stage

$$\beta^s_{1,IV} = \frac{Cov(y,z)}{Cov(s,z)} = \frac{Cov(y,z)/Var(z)}{Cov(s,z)/Var(z)}$$

In our current setting this translates to (impact of attendance on scores):

$$\begin{aligned}\beta^s_{1,IV} &= \frac{Cov(Scores, Offers)}{Cov(Attendance, Offers)} \\ &= \frac{Cov(Scores, Offers)/Var(Offers)}{Cov(Attendance, Offers)/Var(Offers)}\end{aligned}$$

The IV Estimate

TABLE 3.1
Analysis of KIPP lotteries

		KIPP applicants			
	Lynn public fifth graders (1)	KIPP Lynn lottery winners (2)	Winners vs. losers (3)	Attended KIPP (4)	Attended KIPP vs. others (5)
Panel A. Baseline characteristics					
Hispanic	.418	.510	−.058 (.058)	.539	.012 (.054)
Black	.173	.257	.026 (.047)	.240	−.001 (.043)
Female	.480	.494	−.008 (.059)	.495	−.009 (.055)
Free/Reduced price lunch	.770	.814	−.032 (.046)	.828	.011 (.042)
Baseline (4th grade) math score	−.307	−.290	.102 (.120)	−.289	.069 (.109)
Baseline (4th grade) verbal score	−.356	−.386	.063 (.125)	−.368	.088 (.114)
Panel B. Outcomes					
Attended KIPP	.000	.787	.741 (.037)	1.000	1.000 —
Math score	−.363	−.003	.355 (.115)	.095	.467 (.103)
Verbal score	−.417	−.262	.113 (.122)	−.211	.211 (.109)
Sample size	3,964	253	371	204	371

Notes: This table describes baseline characteristics of Lynn fifth graders and reports estimated offer effects for Knowledge Is Power Program (KIPP) Lynn applicants. Means appear in columns (1), (2), and (4). Column (3) shows differences between lottery winners and losers. These are coefficients from regressions that control for risk sets, namely, dummies for year and grade of application and the presence of a sibling applicant. Column (5) shows differences between KIPP students and applicants who did not attend KIPP. Standard errors are reported in parentheses.

From Mastering 'Metrics: The Path from Cause to Effect. © 2015 Princeton University Press. Used by permission. All rights reserved.

- From (3):

  $Cov(Scores, Offers)/Var(Offers) \rightarrow .355$

- From (3):

  $Cov(Attendance, Offers)/Var(Offers) \rightarrow .741$

The IV Estimate

TABLE 3.1
Analysis of KIPP lotteries

		KIPP applicants			
	Lynn public fifth graders (1)	KIPP Lynn lottery winners (2)	Winners vs. losers (3)	Attended KIPP (4)	Attended KIPP vs. others (5)
Panel A. Baseline characteristics					
Hispanic	.418	.510	−.058 (.058)	.539	.012 (.054)
Black	.173	.257	.026 (.047)	.240	−.001 (.043)
Female	.480	.494	−.008 (.059)	.495	−.009 (.055)
Free/Reduced price lunch	.770	.814	−.032 (.046)	.828	.011 (.042)
Baseline (4th grade) math score	−.307	−.290	.102 (.120)	−.289	.069 (.109)
Baseline (4th grade) verbal score	−.356	−.386	.063 (.125)	−.368	.088 (.114)
Panel B. Outcomes					
Attended KIPP	.000	.787	.741 (.037)	1.000	1.000 —
Math score	−.363	−.003	.355 (.115)	.095	.467 (.103)
Verbal score	−.417	−.262	.113 (.122)	−.211	.211 (.109)
Sample size	3,964	253	371	204	371

Notes: This table describes baseline characteristics of Lynn fifth graders and reports estimated offer effects for Knowledge Is Power Program (KIPP) Lynn applicants. Means appear in columns (1), (2), and (4). Column (3) shows differences between lottery winners and losers. These are coefficients from regressions that control for risk sets, namely, dummies for year and grade of application and the presence of a sibling applicant. Column (5) shows differences between KIPP students and applicants who did not attend KIPP. Standard errors are reported in parentheses.

From Mastering 'Metrics: The Path from Cause to Effect. © 2015 Princeton University Press. Used by permission. All rights reserved.

- From (3):

  $Cov(Scores, Offers)/Var(Offers) \rightarrow .355$

- From (3):

  $Cov(Attendance, Offers)/Var(Offers) \rightarrow .741$

- $\beta^s_{1,IV} = \frac{.355}{.741} \approx .48$

The IV Estimate

TABLE 3.1
Analysis of KIPP lotteries

		KIPP applicants			
	Lynn public fifth graders (1)	KIPP Lynn lottery winners (2)	Winners vs. losers (3)	Attended KIPP (4)	Attended KIPP vs. others (5)
Panel A. Baseline characteristics					
Hispanic	.418	.510	−.058 (.058)	.539	.012 (.054)
Black	.173	.257	.026 (.047)	.240	−.001 (.043)
Female	.480	.494	−.008 (.059)	.495	−.009 (.055)
Free/Reduced price lunch	.770	.814	−.032 (.046)	.828	.011 (.042)
Baseline (4th grade) math score	−.307	−.290	.102 (.120)	−.289	.069 (.109)
Baseline (4th grade) verbal score	−.356	−.386	.063 (.125)	−.368	.088 (.114)
Panel B. Outcomes					
Attended KIPP	.000	.787	.741 (.037)	1.000	1.000 —
Math score	−.363	−.003	.355 (.115)	.095	.467 (.103)
Verbal score	−.417	−.262	.113 (.122)	−.211	.211 (.109)
Sample size	3,964	253	371	204	371

Notes: This table describes baseline characteristics of Lynn fifth graders and reports estimated offer effects for Knowledge Is Power Program (KIPP) Lynn applicants. Means appear in columns (1), (2), and (4). Column (3) shows differences between lottery winners and losers. These are coefficients from regressions that control for risk sets, namely, dummies for year and grade of application and the presence of a sibling applicant. Column (5) shows differences between KIPP students and applicants who did not attend KIPP. Standard errors are reported in parentheses.

From Mastering 'Metrics: The Path from Cause to Effect. © 2015 Princeton University Press. Used by permission. All rights reserved.

- From (3):

  $Cov(Scores, Offers)/Var(Offers) \rightarrow .355$

- From (3):

  $Cov(Attendance, Offers)/Var(Offers) \rightarrow .741$

- $\beta^s_{1,IV} = \frac{.355}{.741} \approx .48$
- Test scores were .48 higher for students who attended KIPP (Note: data was standardized, so this is really $.48\sigma$).

The IV Estimate

TABLE 3.1
Analysis of KIPP lotteries

		KIPP applicants			
	Lynn public fifth graders (1)	KIPP Lynn lottery winners (2)	Winners vs. losers (3)	Attended KIPP (4)	Attended KIPP vs. others (5)
Panel A. Baseline characteristics					
Hispanic	.418	.510	−.058 (.058)	.539	.012 (.054)
Black	.173	.257	.026 (.047)	.240	−.001 (.043)
Female	.480	.494	−.008 (.059)	.495	−.009 (.055)
Free/Reduced price lunch	.770	.814	−.032 (.046)	.828	.011 (.042)
Baseline (4th grade) math score	−.307	−.290	.102 (.120)	−.289	.069 (.109)
Baseline (4th grade) verbal score	−.356	−.386	.063 (.125)	−.368	.088 (.114)
Panel B. Outcomes					
Attended KIPP	.000	.787	.741 (.037)	1.000	1.000 —
Math score	−.363	−.003	.355 (.115)	.095	.467 (.103)
Verbal score	−.417	−.262	.113 (.122)	−.211	.211 (.109)
Sample size	3,964	253	371	204	371

Notes: This table describes baseline characteristics of Lynn fifth graders and reports estimated offer effects for Knowledge Is Power Program (KIPP) Lynn applicants. Means appear in columns (1), (2), and (4). Column (3) shows differences between lottery winners and losers. These are coefficients from regressions that control for risk sets, namely, dummies for year and grade of application and the presence of a sibling applicant. Column (5) shows differences between KIPP students and applicants who did not attend KIPP. Standard errors are reported in parentheses.

From Mastering 'Metrics: The Path from Cause to Effect. © 2015 Princeton University Press. Used by permission. All rights reserved.

- From (3):

  $Cov(Scores, Offers)/Var(Offers) \rightarrow .355$

- From (3):

  $Cov(Attendance, Offers)/Var(Offers) \rightarrow .741$

- $\beta^s_{1,IV} = \frac{.355}{.741} \approx .48$
- Test scores were .48 higher for students who attended KIPP (Note: data was standardized, so this is really $.48\sigma$).
- From (5): We can run the OLS of math scores on KIPP and get .467, which is close to our IV estimate, suggesting

The IV Estimate

TABLE 3.1
Analysis of KIPP lotteries

		KIPP applicants			
	Lynn public fifth graders (1)	KIPP Lynn lottery winners (2)	Winners vs. losers (3)	Attended KIPP (4)	Attended KIPP vs. others (5)
Panel A. Baseline characteristics					
Hispanic	.418	.510	−.058 (.058)	.539	.012 (.054)
Black	.173	.257	.026 (.047)	.240	−.001 (.043)
Female	.480	.494	−.008 (.059)	.495	−.009 (.055)
Free/Reduced price lunch	.770	.814	−.032 (.046)	.828	.011 (.042)
Baseline (4th grade) math score	−.307	−.290	.102 (.120)	−.289	.069 (.109)
Baseline (4th grade) verbal score	−.356	−.386	.063 (.125)	−.368	.088 (.114)
Panel B. Outcomes					
Attended KIPP	.000	.787	.741 (.037)	1.000	1.000 —
Math score	−.363	−.003	.355 (.115)	.095	.467 (.103)
Verbal score	−.417	−.262	.113 (.122)	−.211	.211 (.109)
Sample size	3,964	253	371	204	371

Notes: This table describes baseline characteristics of Lynn fifth graders and reports estimated offer effects for Knowledge Is Power Program (KIPP) Lynn applicants. Means appear in columns (1), (2), and (4). Column (3) shows differences between lottery winners and losers. These are coefficients from regressions that control for risk sets, namely, dummies for year and grade of application and the presence of a sibling applicant. Column (5) shows differences between KIPP students and applicants who did not attend KIPP. Standard errors are reported in parentheses.

From Mastering 'Metrics: The Path from Cause to Effect. © 2015 Princeton University Press. Used by permission. All rights reserved.

- From (3):

  $Cov(Scores, Offers)/Var(Offers) \rightarrow .355$

- From (3):

  $Cov(Attendance, Offers)/Var(Offers) \rightarrow .741$

- $\beta^s_{1,IV} = \frac{.355}{.741} \approx .48$
- Test scores were .48 higher for students who attended KIPP (Note: data was standardized, so this is really $.48\sigma$).
- From (5): We can run the OLS of math scores on KIPP and get .467, which is close to our IV estimate, suggesting selection bias was not severe.

Local Avg Treatment Effect (LATE)

Our IV estimates are in general NOT recovering an "overall" average treatment effect" $E[Y_1|D = 1] - E[Y_0|D = 0]$

Instead, we are capturing a "local average treatment effect" (LATE)

Local implies we have a causal effect for the subset of the population known as compliers

LATE

TABLE 3.2
The four types of children

		Lottery losers $Z_i = 0$	
		Doesn't attend KIPP $D_i = 0$	Attends KIPP $D_i = 1$
Lottery winners $Z_i = 1$	Doesn't attend KIPP $D_i = 0$	Never-takers (*Normando*)	Defiers
	Attends KIPP $D_i = 1$	Compliers (*Camila*)	Always-takers (*Alvaro*)

Note: KIPP = Knowledge Is Power Program.

LATE

TABLE 3.2
The four types of children

		Lottery losers $Z_i = 0$	
		Doesn't attend KIPP $D_i = 0$	Attends KIPP $D_i = 1$
Lottery winners $Z_i = 1$	Doesn't attend KIPP $D_i = 0$	Never-takers (*Normando*)	Defiers
	Attends KIPP $D_i = 1$	Compliers (*Camila*)	Always-takers (*Alvaro*)

Note: KIPP = Knowledge Is Power Program.

Hint: read this table with an "if" statement. If Camila won the lottery should would attend. If Camila lost the lottery she would not attend. She always complies to the "treatment" of winning the lottery.

LATE

The instrument does not affect the treatment status for "never-takers" nor "always-takers".

The instrument changes treatment status for Compliers

Assuming monotonicity (i.e. winning the lottery only makes you more likely to attend) eliminates "defiers"

TABLE 3.2
The four types of children

	Lottery losers $Z_i = 0$	
	Doesn't attend KIPP $D_i = 0$	Attends KIPP $D_i = 1$
Lottery winners $Z_i = 1$		
Doesn't attend KIPP $D_i = 0$	Never-takers (*Normando*)	Defiers
Attends KIPP $D_i = 1$	Compliers (*Camila*)	Always-takers (*Alvaro*)

Note: KIPP = Knowledge Is Power Program.

LATE

- The prior matrix illustrates hypothetical scenarios. You do not need to alter your data.

LATE

- The prior matrix illustrates hypothetical scenarios. You do not need to alter your data.
- The main idea is that IV is not estimating a causal effect that applies to everyone. Rather, it applies only to compliers.

LATE

- The prior matrix illustrates hypothetical scenarios. You do not need to alter your data.
- The main idea is that IV is not estimating a causal effect that applies to everyone. Rather, it applies only to compliers.
- Q: Why is it fortunate for decision makers that LATE applies to compliers?

LATE

- The prior matrix illustrates hypothetical scenarios. You do not need to alter your data.
- The main idea is that IV is not estimating a causal effect that applies to everyone. Rather, it applies only to compliers.
- Q: Why is it fortunate for decision makers that LATE applies to compliers?
- A: These are the individuals that the policy can influence; i.e. the lottery would impact Camila, but not the never takers, always takers, nor defiers.

LATE

- The prior matrix illustrates hypothetical scenarios. You do not need to alter your data.
- The main idea is that IV is not estimating a causal effect that applies to everyone. Rather, it applies only to compliers.
- Q: Why is it fortunate for decision makers that LATE applies to compliers?
- A: These are the individuals that the policy can influence; i.e. the lottery would impact Camila, but not the never takers, always takers, nor defiers.

Note: we can still think of 2SLS as capturing the average of the LATEs even when endogenous variables (s) and instrument (z) are continuous, rather than binary

Instrumental Variables - Summary

- IV is a approach to mitigate OVB
- It can be very powerful, but should not be used unless necessary
- Valid instruments (relevant, independent, exogenous) can be difficult to find