Simultaneous Systems

Prof. Mike Aguilar

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

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings
- Treatment effects are difficult to estimate due to endogeneity of regressors and errors

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings
- Treatment effects are difficult to estimate due to endogeneity of regressors and errors
- IV provides a solution

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings
- Treatment effects are difficult to estimate due to endogeneity of regressors and errors
- IV provides a solution
- As always with IV, our instruments must be carefully chosen

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Housing & Saving



Case Study Details

Background:

- Your client is a multifamily residential real estate developer/operator.

Case Study Details

Background:

- Your client is a multifamily residential real estate developer/operator.
- They are concerned that if savings drop (e.g. impending recession), their tenants may have less money for rent.

Case Study Details

Background:

- Your client is a multifamily residential real estate developer/operator.
- They are concerned that if savings drop (e.g. impending recession), their tenants may have less money for rent.

Data:

- Unit of measurement is a tenant (i) in a residential apartment building

Case Study Details

Background:

- Your client is a multifamily residential real estate developer/operator.
- They are concerned that if savings drop (e.g. impending recession), their tenants may have less money for rent.

Data:

- Unit of measurement is a tenant (i) in a residential apartment building
- Data structure is a panel across several hundred tenants, spanning 12 months, and covering various apartment buildings in a single city

Case Study Details

Background:

- Your client is a multifamily residential real estate developer/operator.
- They are concerned that if savings drop (e.g. impending recession), their tenants may have less money for rent.

Data:

- Unit of measurement is a tenant (i) in a residential apartment building
- Data structure is a panel across several hundred tenants, spanning 12 months, and covering various apartment buildings in a single city
- We have data on housing expenditures (i.e. the dollar amount of rent paid), savings (i.e. stock of personal savings measured in dollars), as well as several controls

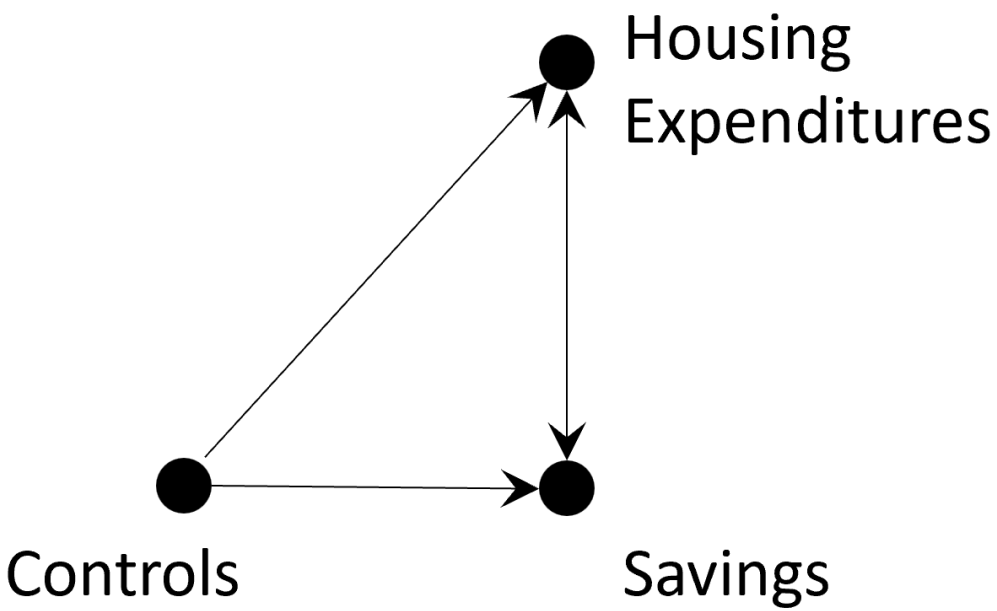
Drawing the Causal System

Q: Can you draw a DAG? What's wrong?

Drawing the Causal System

Q: Can you draw a DAG? What's wrong?

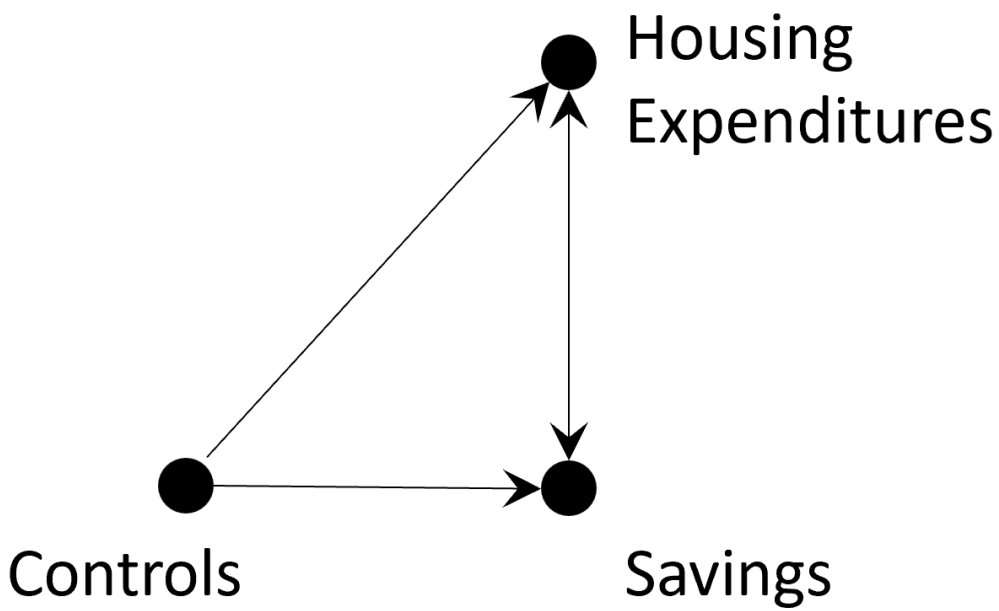
A:



Drawing the Causal System

Q: Can you draw a DAG? What's wrong?

A:

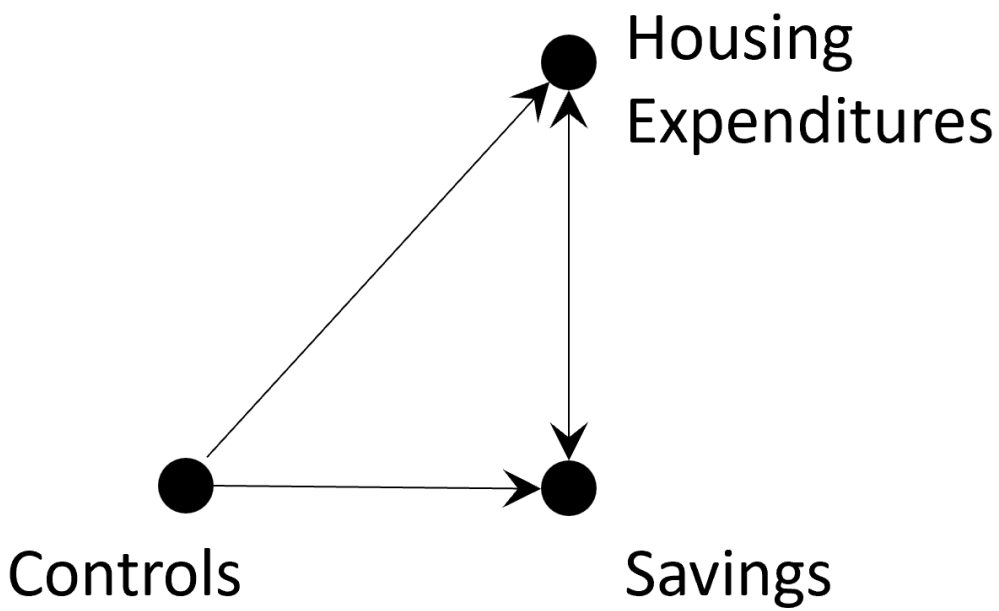


Notice that the causal graph isn't "directed". There are arrows that go both ways.
Savings drives how much is available for Housing Expenditures, and

Drawing the Causal System

Q: Can you draw a DAG? What's wrong?

A:



Notice that the causal graph isn't "directed". There are arrows that go both ways.
Savings drives how much is available for Housing Expenditures, and Housing Expenditures can drive how much is available to be Saved.

An Expenditure Model

You structure the following model

$$hsngexp = \alpha + \beta saving + \gamma Z + u$$

where

- *hsngexp*: housing expenditures ($)
- *savings*: savings ($)
- Z: any control variables of interest

Interpreting the Model

Q: How interpret β?

$$hsngexp = \alpha + \beta saving + \gamma Z + u$$

Interpreting the Model

Q: How interpret β?

$$hsngexp = \alpha + \beta saving + \gamma Z + u$$

A:

Interpreting the Model

Q: How interpret β?

$$hsngexp = \alpha + \beta saving + \gamma Z + u$$

A:

- $\frac{\partial HsngExp}{\partial Savings} = \beta$

Interpreting the Model

Q: How interpret β?

$$hsngexp = \alpha + \beta saving + \gamma Z + u$$

A:

- $\frac{\partial HsngExp}{\partial Savings} = \beta$
- (Average) marginal effect of savings upon housing expenditure, c.p.

Trouble with the model

Q: Why might the interpretation of β not be reasonable?

Trouble with the model

Q: Why might the interpretation of β not be reasonable?

A: β doesn't provide a (c.p.) interpretation.

Trouble with the model

Q: Why might the interpretation of β not be reasonable?

A: β doesn't provide a (c.p.) interpretation.

$$hsngexp = \alpha_1 + \beta_1 saving + \gamma Z_1 + u_1$$

Trouble with the model

Q: Why might the interpretation of β not be reasonable?

A: β doesn't provide a (c.p.) interpretation.

$$\begin{aligned} \textit{hsngexp} &= \alpha_1 + \beta_1 \textit{saving} + \gamma Z_1 + u_1 \\ \textit{saving} &= \alpha_2 + \beta_2 \textit{hsngexp} + \delta Z_2 + u_2 \end{aligned}$$

Trouble with the model

Q: Why might the interpretation of β not be reasonable?

A: β doesn't provide a (c.p.) interpretation.

$$\begin{aligned} \textit{hsngexp} &= \alpha_1 + \beta_1 \textit{saving} + \gamma Z_1 + u_1 \\ \textit{saving} &= \alpha_2 + \beta_2 \textit{hsngexp} + \delta Z_2 + u_2 \end{aligned}$$

The feedback between housing and savings makes this a "Simultaneous System" of equations, plagued by a "Simultaneity Bias". This bias is similar to OVB.

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$\begin{aligned} \textit{hsngexp} &= 1 + .5 * \textit{saving} + 0 * Z_1 + u_1 \\ \textit{saving} &= .25 + 1 * \textit{hsngexp} + 0 * Z_2 + u_2 \end{aligned}$$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp =$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$
- ...and so $\Delta savings =$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$
- ...and so $\Delta savings = 1 * \$0.5 = \$0.5$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$
- ...and so $\Delta savings = 1 * \$0.5 = \$0.5$
- ...and so $\Delta hsnexp =$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$
- ...and so $\Delta savings = 1 * \$0.5 = \$0.5$
- ...and so $\Delta hsnexp = .5 * \$0.5 = \$0.25$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$
- ...and so $\Delta savings = 1 * \$0.5 = \$0.5$
- ...and so $\Delta hsnexp = .5 * \$0.5 = \$0.25$
- ...and so $\Delta savings = \cdots$

Quick Example to Illustrate Intuition of the Problem

Suppose the "true" DGP was...

$$hsngexp = 1 + .5 * saving + 0 * Z_1 + u_1$$
$$saving = .25 + 1 * hsngexp + 0 * Z_2 + u_2$$

If savings rises by \$1,

- ..... $\Delta hsngexp = .5 * \$1 = \$0.5$
- ...and so $\Delta savings = 1 * \$0.5 = \$0.5$
- ...and so $\Delta hsnexp = .5 * \$0.5 = \$0.25$
- ...and so $\Delta savings = \cdots$

The .5 in front of *savings* is NOT the marginal effect of savings upon housing expenditure because all is not c.p.

Caution: In this simple example, we are not considering realistic frictions that may create lags in this process.

Causality in Simultaneous Systems

- Our typical notion of causality doesn't apply in a simultaneous system

Causality in Simultaneous Systems

- Our typical notion of causality doesn't apply in a simultaneous system
- Independent sources of variation are no longer the causal variables

Causality in Simultaneous Systems

- Our typical notion of causality doesn't apply in a simultaneous system
- Independent sources of variation are no longer the causal variables
- (ala Cameron and Trivedi) We require an independent and exogenous source of variation

Causality in Simultaneous Systems

- Our typical notion of causality doesn't apply in a simultaneous system
- Independent sources of variation are no longer the causal variables
- (ala Cameron and Trivedi) We require an independent and exogenous source of variation
- Tip: Talk to your SMEs. Draw your DAG. If an arrow goes both ways, you have a simultaneous system. You can't proceed as normal because you'll have a special type of bias.

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Outline

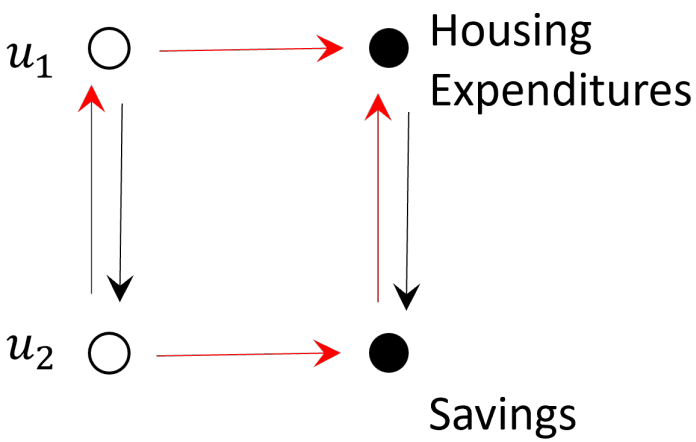
1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*

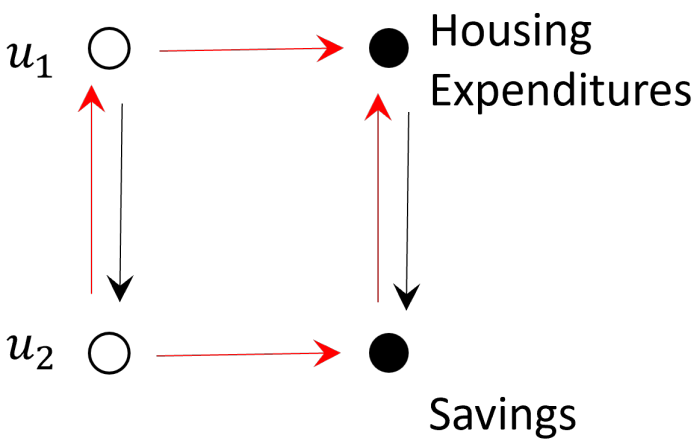
Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*



Biasing the Impact of Savings On Housing

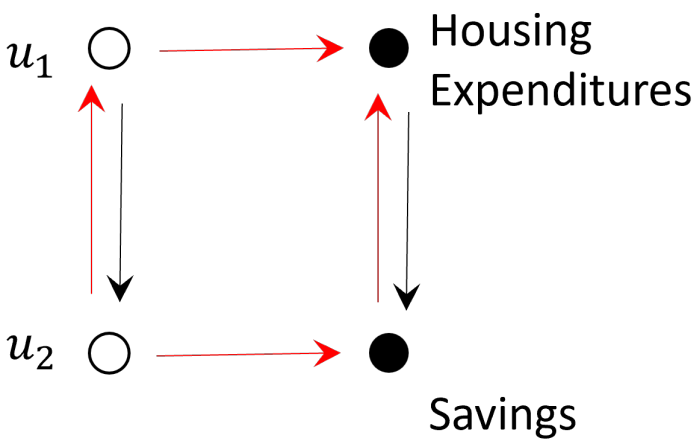
Define TE as impact of *savings* on *housing*



Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*

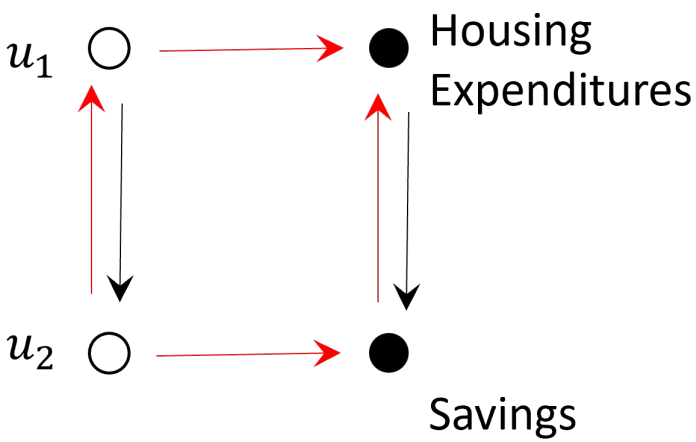


Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)

A: u_2 because 1)

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*

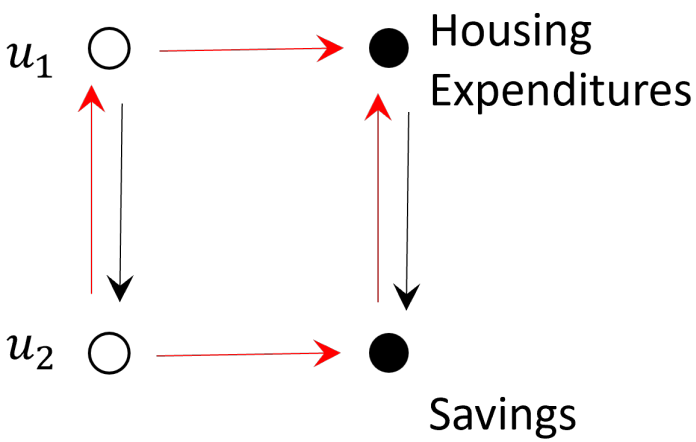


Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)

A: u_2 because 1) it is unobserved, and 2)

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*

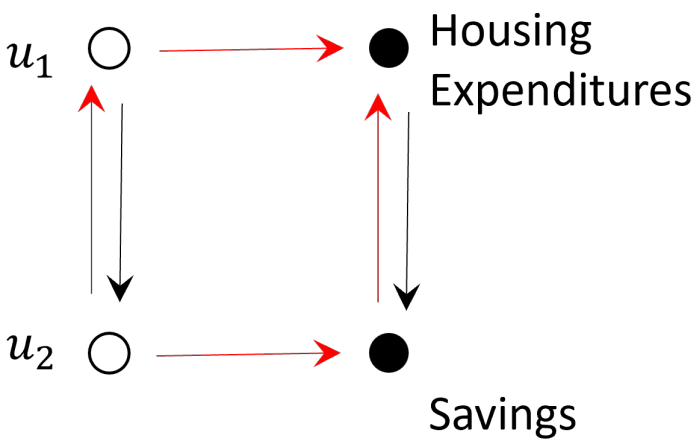


Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)

A: u_2 because 1) it is unobserved, and 2) it impacts Housing and Savings

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*



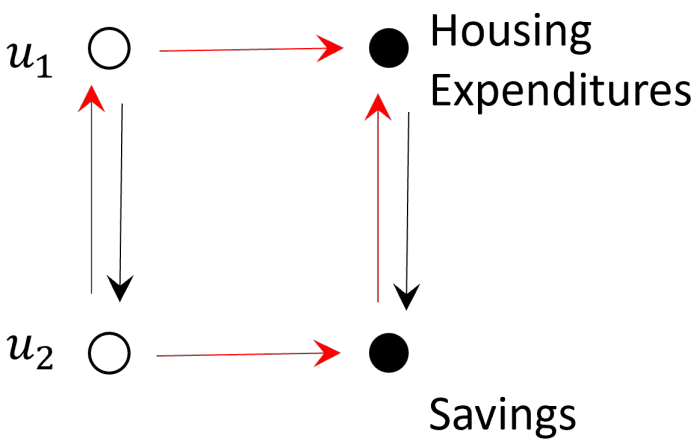
Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)

A: u_2 because 1) it is unobserved, and 2) it impacts Housing and Savings

Q: Meaning of u_2 and the OVB?

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*

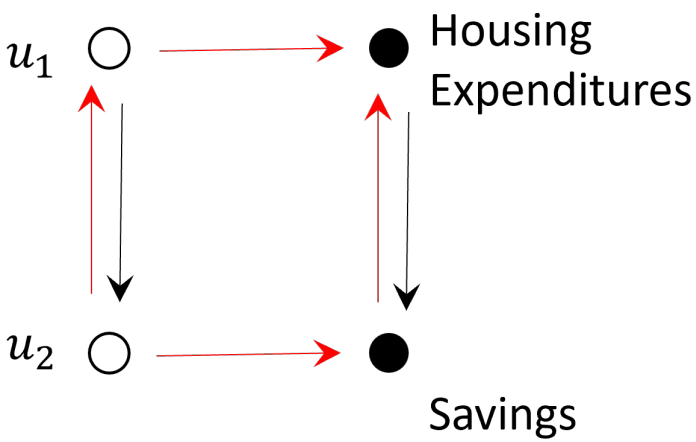


Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)
A: u_2 because 1) it is unobserved, and 2) it impacts Housing and Savings

Q: Meaning of u_2 and the OVB?
A: If the unobserved part of savings influences (in)directly housing expenditures, the treatment effect is biased.

Biasing the Impact of Savings On Housing

Define TE as impact of *savings* on *housing*



Q: Which unobserved variable causes an "OVB"? (hint: follow the red arrows)
A: u_2 because 1) it is unobserved, and 2) it impacts Housing and Savings

Q: Meaning of u_2 and the OVB?
A: If the unobserved part of savings influences (in)directly housing expenditures, the treatment effect is biased. *We will brainstorm on examples of u_2 shortly*

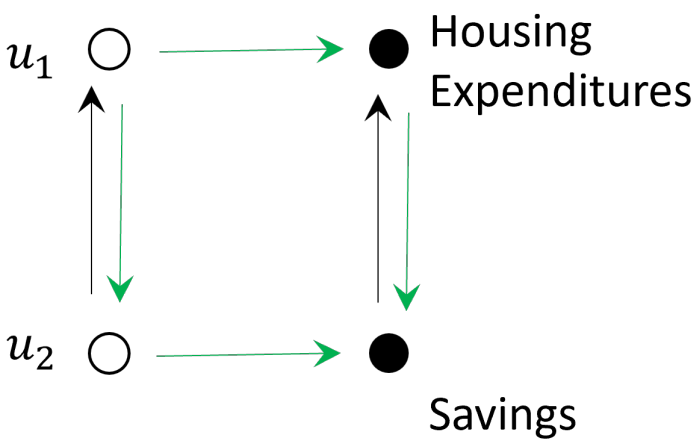
Arrows in the DAG are colored to highlight their importance in this case.

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*

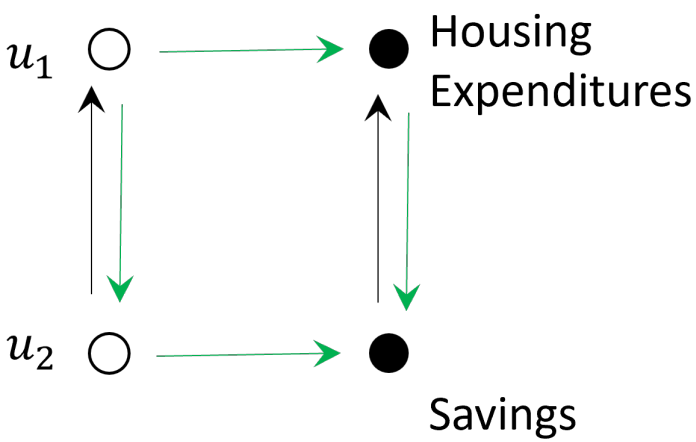
Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*



Bias the Impact of Housing on Savings

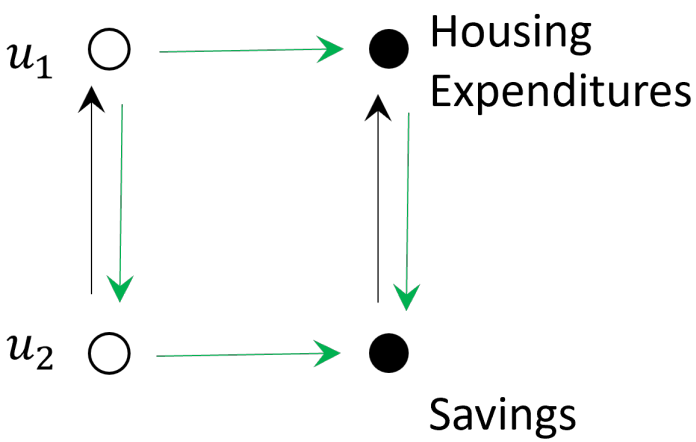
Define TE as impact of *housing* on *savings*



Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*

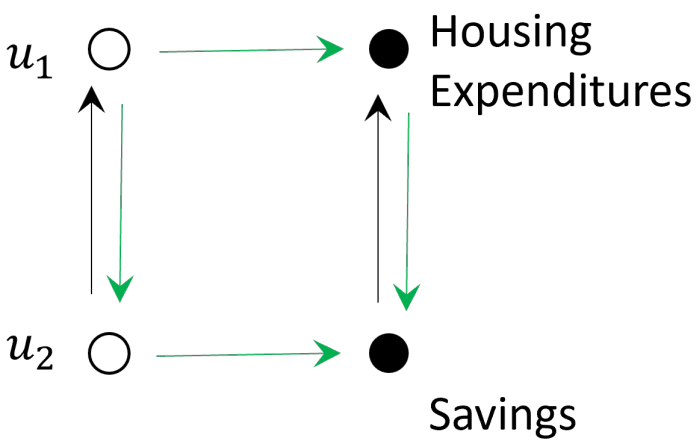


Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)

A: u_1 because 1)

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*

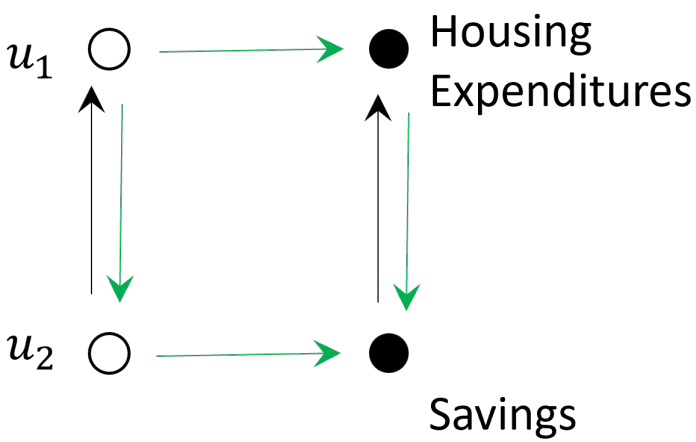


Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)

A: u_1 because 1) it is unobserved, and 2)

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*

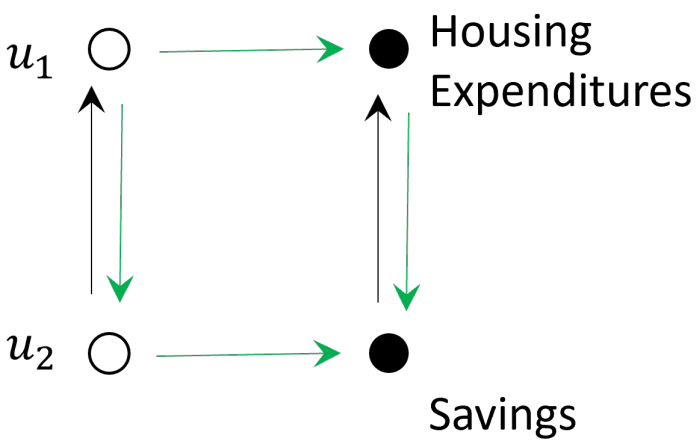


Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)

A: u_1 because 1) it is unobserved, and 2) it impacts Housing and Savings

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*

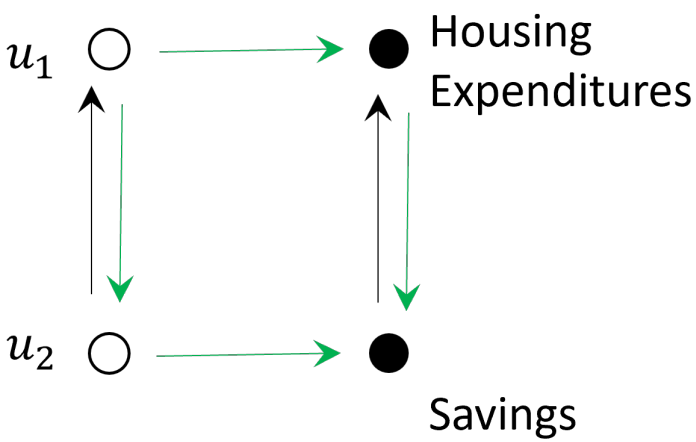


Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)
A: u_1 because 1) it is unobserved, and 2) it impacts Housing and Savings

Q: Meaning of u_1 and the OVB?

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*

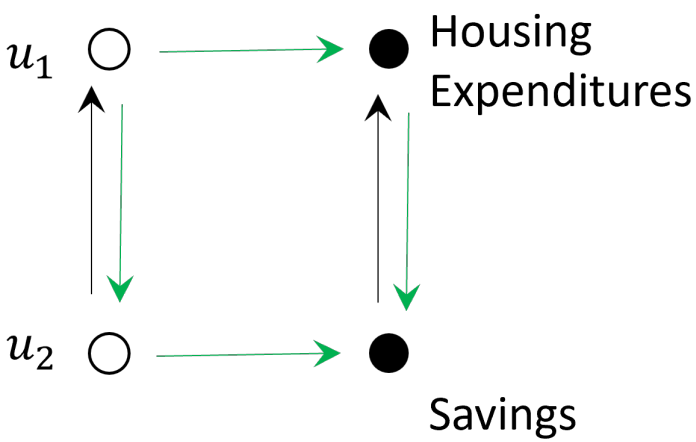


Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)
A: u_1 because 1) it is unobserved, and 2) it impacts Housing and Savings

Q: Meaning of u_1 and the OVB?
A: If the unobserved part of housing expenditures influences (in)directly savings, the treatment effect is biased.

Bias the Impact of Housing on Savings

Define TE as impact of *housing* on *savings*



Q: Which unobserved variable causes an "OVB"? (hint: follow the green arrows)
A: u_1 because 1) it is unobserved, and 2) it impacts Housing and Savings

Q: Meaning of u_1 and the OVB?
A: If the unobserved part of housing expenditures influences (in)directly savings, the treatment effect is biased. *We will brainstorm on examples of u_1 shortly*

Simultaneity Bias

- The visual depictions above are one type of bias path (indirect). We will explore a direct path as well.

Simultaneity Bias

- The visual depictions above are one type of bias path (indirect). We will explore a direct path as well.
- The simultaneity in our causal system biases our traditional methods of eliciting causal effects

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Impact of Housing on Savings

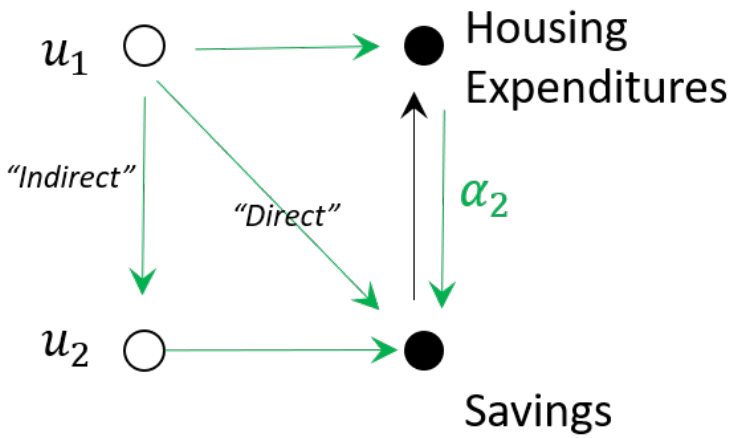
Q: Which parameter captures the impact of *housing* upon *savings*?

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

Impact of Housing on Savings

Q: Which parameter captures the impact of *housing* upon *savings*?

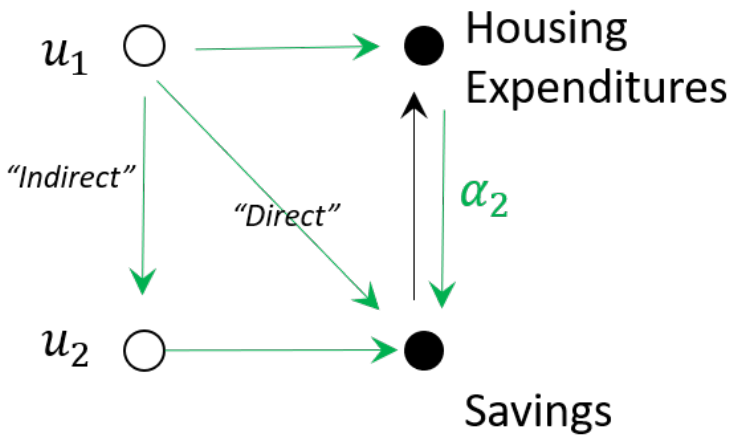


$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_1 is related to savings, α_2 is plagued by a bias.

Impact of Housing on Savings

Q: Which parameter captures the impact of *housing* upon *savings*?



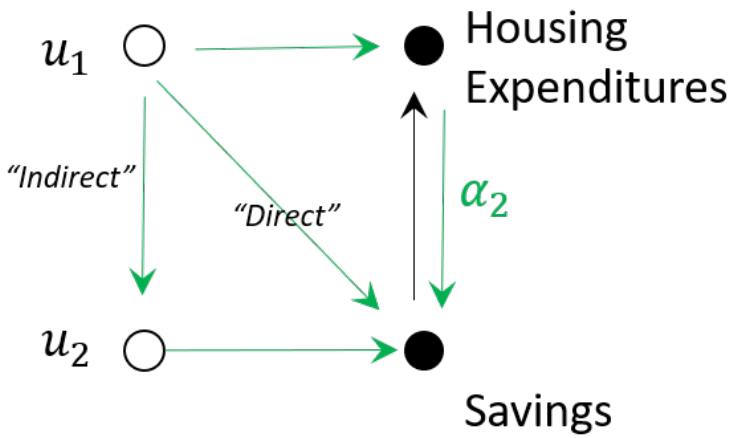
$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_1 is related to savings, α_2 is plagued by a bias.
- Q: Example of Direct bias in our case?

Impact of Housing on Savings

Q: Which parameter captures the impact of *housing* upon *savings*?



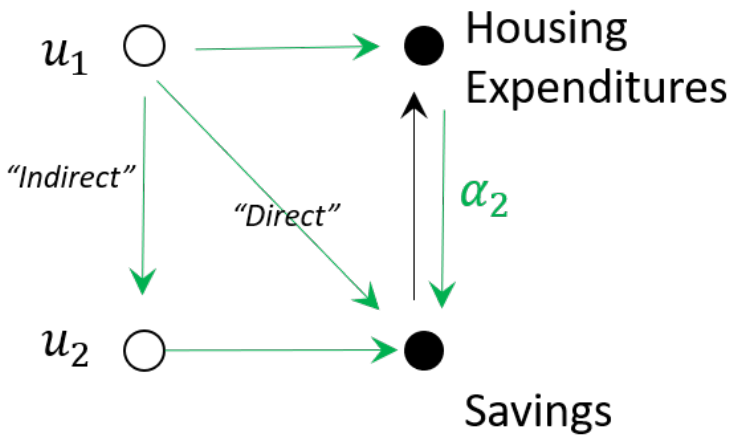
$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_1 is related to savings, α_2 is plagued by a bias.
- Q: Example of Direct bias in our case?
- A: Local GDP u_1 impacts housing and savings

Impact of Housing on Savings

Q: Which parameter captures the impact of *housing* upon *savings*?

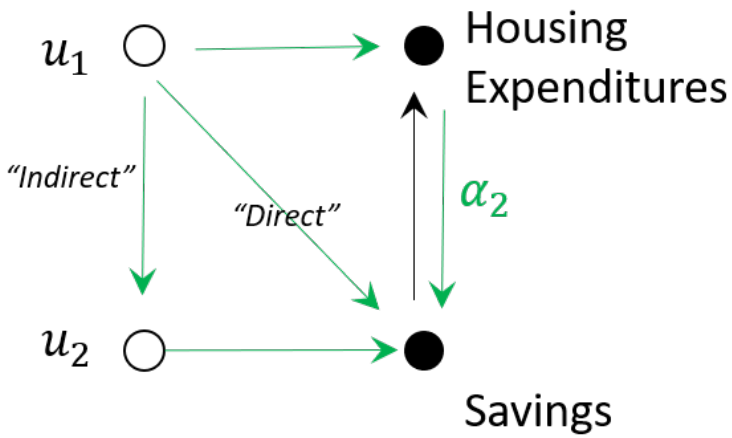


$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_1 is related to savings, α_2 is plagued by a bias.
- Q: Example of Direct bias in our case?
- A: Local GDP u_1 impacts housing and savings
- Q: Example of Indirect bias in our case?

Impact of Housing on Savings

Q: Which parameter captures the impact of *housing* upon *savings*?



$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_1 is related to savings, α_2 is plagued by a bias.
- Q: Example of Direct bias in our case?
- A: Local GDP u_1 impacts housing and savings
- Q: Example of Indirect bias in our case?
- A: Mortgage rates u_1 may impact housing expenditure directly AND impact savings interest rates u_2, which impacts savings indirectly.

Impact of Savings on Housing

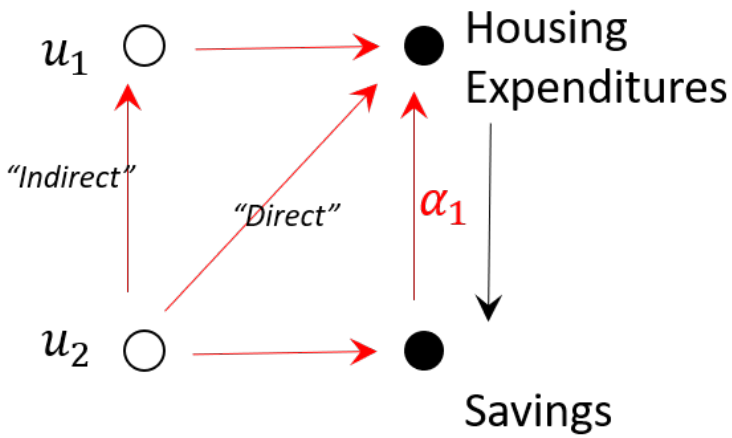
Q: Which parameter captures the impact of *savings* upon *housing*?

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

Impact of Savings on Housing

Q: Which parameter captures the impact of *savings* upon *housing*?

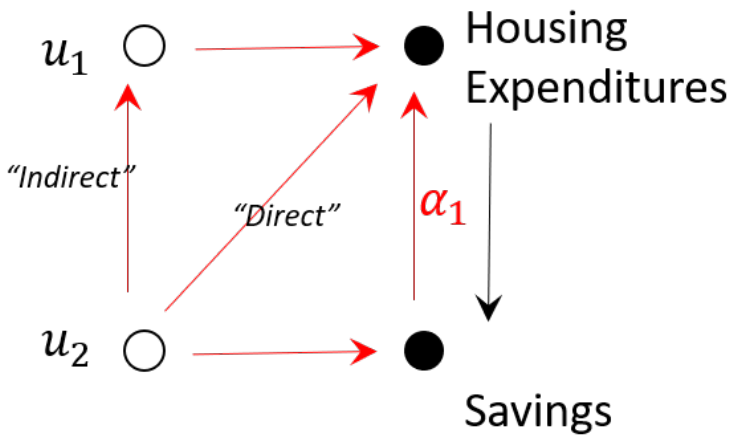


$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_2 is related to savings, α_1 is plagued by a bias.

Impact of Savings on Housing

Q: Which parameter captures the impact of *savings* upon *housing*?



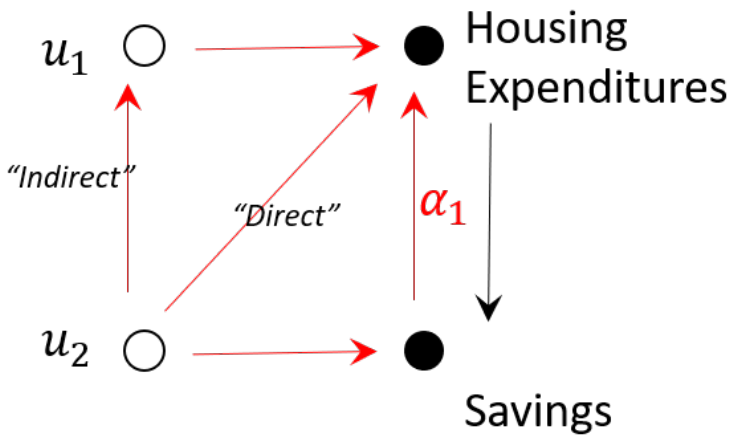
$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_2 is related to savings, α_1 is plagued by a bias.
- Q: Example of Direct bias in our case?

Impact of Savings on Housing

Q: Which parameter captures the impact of *savings* upon *housing*?



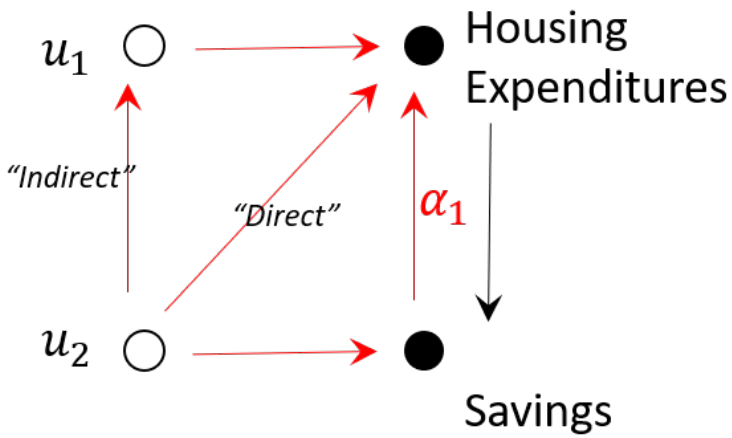
$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_2 is related to savings, α_1 is plagued by a bias.
- Q: Example of Direct bias in our case?
- A: Local GDP u_2 impacts housing and savings

Impact of Savings on Housing

Q: Which parameter captures the impact of *savings* upon *housing*?



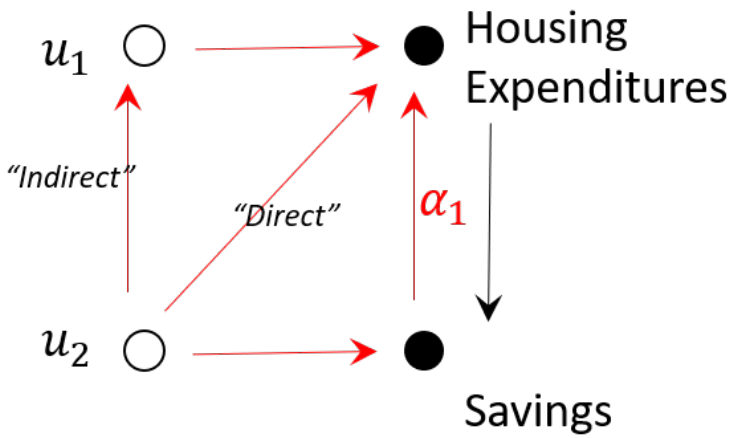
$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_2 is related to savings, α_1 is plagued by a bias.
- Q: Example of Direct bias in our case?
- A: Local GDP u_2 impacts housing and savings
- Q: Example of Indirect bias in our case?

Impact of Savings on Housing

Q: Which parameter captures the impact of *savings* upon *housing*?



$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

- If u_2 is related to savings, α_1 is plagued by a bias.
- Q: Example of Direct bias in our case?
- A: Local GDP u_2 impacts housing and savings
- Q: Example of Indirect bias in our case?
- A: Availability of high yield savings accounts u_2 impacts savings AND may impact disposable income available for consuming housing u_1, which impacts housing expenditures.

Bias Summary

- A Direct bias is generating by something omitted that drives both treatment and outcome

Bias Summary

- A Direct bias is generating by something omitted that drives both treatment and outcome
- Direct bias is the most common source of simultaneity bias

Bias Summary

- A Direct bias is generating by something omitted that drives both treatment and outcome
- Direct bias is the most common source of simultaneity bias
- An Indirect bias is generated by an omitted driver of the treatment that also impacts the outcome via an omitted mediator

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Understand the System. Understand the Bias.

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \tag{1}$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \tag{2}$$

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Q: Can we elicit a causal effect of hsngexp upon saving by estimating (2)?

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Q: Can we elicit a causal effect of hsngexp upon saving by estimating (2)?
A:

- No

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Q: Can we elicit a causal effect of hsngexp upon saving by estimating (2)?
A:

- No
- There is a feedback: hsngexp $\rightarrow$ saving $\rightarrow$ hsngexp $\rightarrow \cdots$

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Q: Can we elicit a causal effect of hsngexp upon saving by estimating (2)?
A:

- No
- There is a feedback: hsngexp $\rightarrow$ saving $\rightarrow$ hsngexp $\rightarrow \cdots$

Q: Which variables are endogenous? Exogenous?

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Q: Can we elicit a causal effect of hsngexp upon saving by estimating (2)?
A:

- No
- There is a feedback: hsngexp $\rightarrow$ saving $\rightarrow$ hsngexp $\rightarrow \cdots$

Q: Which variables are endogenous? Exogenous?
A:

- Endogenous: hsnexp and saving

Understand the System. Understand the Bias.

SME's help us define our system via "Structural Equations"

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Q: Can we elicit a causal effect of hsngexp upon saving by estimating (2)?
A:

- No
- There is a feedback: hsngexp $\rightarrow$ saving $\rightarrow$ hsngexp $\rightarrow \cdots$

Q: Which variables are endogenous? Exogenous?
A:

- Endogenous: hsnexp and saving
- Exogenous: Z_1, Z_2

Reducing the Simultaneous System

A "Reduced Form" system helps identify the bias

Structural Equations

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Reducing the Simultaneous System

A "Reduced Form" system helps identify the bias

Structural Equations

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Plug (1) into (2)

Reducing the Simultaneous System

A "Reduced Form" system helps identify the bias

Structural Equations

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Plug (1) into (2)

$$\text{saving} = \alpha_2(\alpha_1 \text{saving} + \beta_1 Z_1 + u_1) + \beta_2 Z_2 + u_2$$

Reducing the Simultaneous System

A "Reduced Form" system helps identify the bias

Structural Equations

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Plug (1) into (2)

$$\text{saving} = \alpha_2(\alpha_1 \text{saving} + \beta_1 Z_1 + u_1) + \beta_2 Z_2 + u_2$$
$$(1 - \alpha_2\alpha_1)\text{saving} = \alpha_2\beta_1 Z_1 + \beta_2 Z_2 + \alpha_2 u_1 + u_2$$

Reducing the Simultaneous System

A "Reduced Form" system helps identify the bias

Structural Equations

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Plug (1) into (2)

$$\text{saving} = \alpha_2(\alpha_1 \text{saving} + \beta_1 Z_1 + u_1) + \beta_2 Z_2 + u_2$$
$$(1 - \alpha_2\alpha_1)\text{saving} = \alpha_2\beta_1 Z_1 + \beta_2 Z_2 + \alpha_2 u_1 + u_2$$

If we assume $\alpha_1\alpha_2 \neq 1$ we can write the structural system in a reduced form.

Reduced Form for Savings

Reduced Form Equation of saving

$$\text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Reduced Form for Savings

Reduced Form Equation of saving

$$\text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

where

- $\pi_{21} = \frac{\alpha_2 \beta_1}{1 - \alpha_2 \alpha_1}$
- $\pi_{22} = \frac{\beta_2}{1 - \alpha_2 \alpha_1}$
- $v_2 = \frac{\alpha_2 u_1 + u_2}{1 - \alpha_2 \alpha_1}$

Reduced Form for Housing Expenditures

Similarly, the Reduced Form Equation of hsngexp

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

Reduced Form for Housing Expenditures

Similarly, the Reduced Form Equation of hsngexp

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

where

- $\pi_{11} = \frac{\beta_1}{1-\alpha_1\alpha_2}$
- $\pi_{12} = \frac{\alpha_1\beta_2}{1-\alpha_1\alpha_2}$
- $v_1 = \frac{\alpha_1 u_2 + u_1}{1-\alpha_1\alpha_2}$

Reduced Form for Housing Expenditures

Similarly, the Reduced Form Equation of hsngexp

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

where

- $\pi_{11} = \frac{\beta_1}{1-\alpha_1\alpha_2}$
- $\pi_{12} = \frac{\alpha_1\beta_2}{1-\alpha_1\alpha_2}$
- $v_1 = \frac{\alpha_1 u_2 + u_1}{1-\alpha_1\alpha_2}$

Q: What is a key feature of these reduced form specifications?

Reduced Form for Housing Expenditures

Similarly, the Reduced Form Equation of hsngexp

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

where

- $\pi_{11} = \frac{\beta_1}{1-\alpha_1\alpha_2}$
- $\pi_{12} = \frac{\alpha_1\beta_2}{1-\alpha_1\alpha_2}$
- $v_1 = \frac{\alpha_1 u_2 + u_1}{1-\alpha_1\alpha_2}$

Q: What is a key feature of these reduced form specifications?
A: The endogenous variables hsngexp and saving are expressed purely as functions of the exogenous variables.

A Reduced Form Solution?

Reduced Form:

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

A Reduced Form Solution?

Reduced Form:

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

Q: Can we apply OLS to the reduced form equations?

A Reduced Form Solution?

Reduced Form:

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

Q: Can we apply OLS to the reduced form equations?
A: Yes. The errors are unrelated to the regressors.

A Reduced Form Solution?

Reduced Form:

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

Q: Can we apply OLS to the reduced form equations?
A: Yes. The errors are unrelated to the regressors.

Q: Will this provide causal effects of interest?

A Reduced Form Solution?

Reduced Form:

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

Q: Can we apply OLS to the reduced form equations?
A: Yes. The errors are unrelated to the regressors.

Q: Will this provide causal effects of interest?
A:

- Unlikely
- The reduced form parameters are nonlinear functions of the structural parameters.

A Reduced Form Solution?

Reduced Form:

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

Q: Can we apply OLS to the reduced form equations?
A: Yes. The errors are unrelated to the regressors.

Q: Will this provide causal effects of interest?
A:

- Unlikely
- The reduced form parameters are nonlinear functions of the structural parameters.
- We want the structural parameters (e.g. α_1), but only have the reduced form parameters (e.g. π_{11})

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1$$
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2$$

$$\text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$
$$\text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \qquad \text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \qquad \text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Let's focus on α_1 as our parameter of interest.

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \qquad \text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \qquad \text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Let's focus on α_1 as our parameter of interest. By the exogeneity assumption, $Cov(\text{saving}, u_1) = 0$.

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \qquad \text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \qquad \text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Let's focus on α_1 as our parameter of interest. By the exogeneity assumption, $Cov(\text{saving}, u_1) = 0$. This may be violated if

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \qquad \text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \qquad \text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Let's focus on α_1 as our parameter of interest. By the exogeneity assumption, $Cov(\text{saving}, u_1) = 0$. This may be violated if

1. u_1 is directly related to v_2, or

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \qquad \text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \qquad \text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Let's focus on α_1 as our parameter of interest. By the exogeneity assumption, $Cov(\text{saving}, u_1) = 0$. This may be violated if

1. u_1 is directly related to v_2, or
2. u_1 is related to u_2, which is in turn related to v_2

Isolating Two Channels of Bias

Reduced Form can't eliminate the bias, but it can help us understand the bias

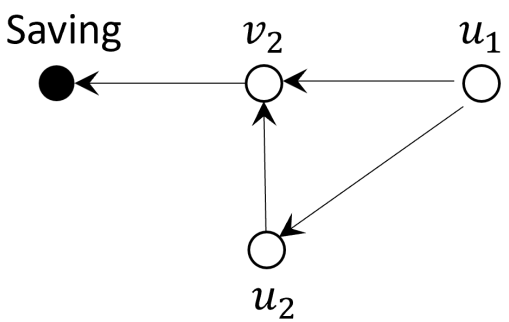
Reminder

$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \qquad \text{hsngexp} = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \qquad \text{saving} = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Let's focus on α_1 as our parameter of interest. By the exogeneity assumption, $Cov(\text{saving}, u_1) = 0$. This may be violated if

1. u_1 is directly related to v_2, or
2. u_1 is related to u_2, which is in turn related to v_2

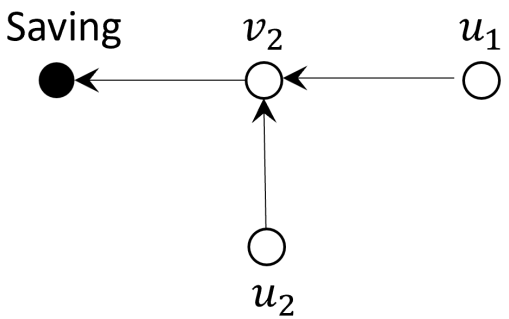


A direct channel of Bias

Assume for a moment that u_1 is not related to u_2.

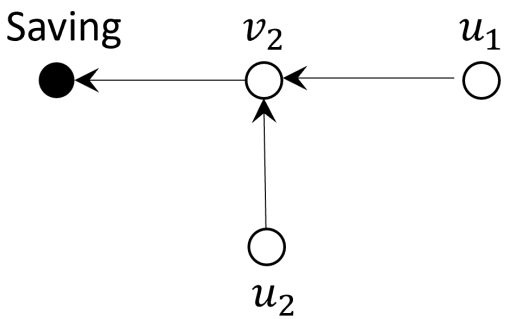
A direct channel of Bias

Assume for a moment that u_1 is not related to u_2.



A direct channel of Bias

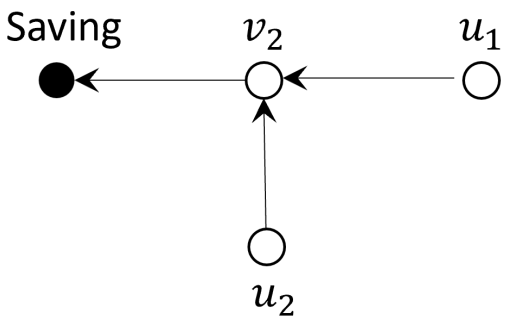
Assume for a moment that u_1 is not related to u_2.



Then bias exists (i.e. saving is related to u_1) only if

A direct channel of Bias

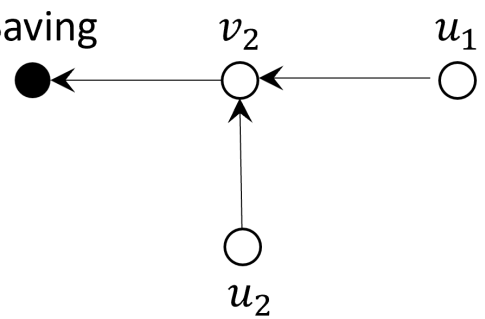
Assume for a moment that u_1 is not related to u_2.



Then bias exists (i.e. saving is related to u_1) only if u_1 is related to v_2.

A direct channel of Bias

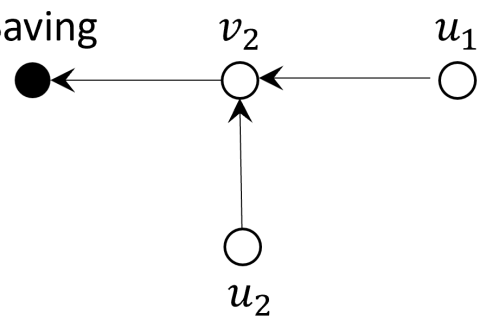
Assume for a moment that u_1 is not related to u_2.



Then bias exists (i.e. saving is related to u_1) only if u_1 is related to v_2.
Recall $v_2 = \frac{\alpha_2 u_1 + u_2}{1 - \alpha_2 \alpha_1}$.

A direct channel of Bias

Assume for a moment that u_1 is not related to u_2.



Then bias exists (i.e. saving is related to u_1) only if u_1 is related to v_2.
Recall $v_2 = \frac{\alpha_2 u_1 + u_2}{1 - \alpha_2 \alpha_1}$.
u_1 is related to v_2 if and only if $\alpha_2 \neq 0$.

A direct channel of Bias - Cont'd

Q: What does $\alpha_2 \neq 0$ mean?

A direct channel of Bias - Cont'd

Q: What does $\alpha_2 \neq 0$ mean?
A:

- Recall structural equation (2) saving $= \alpha_2$hsngexp $+ \beta_2 Z_2 + u_2$

A direct channel of Bias - Cont'd

Q: What does $\alpha_2 \neq 0$ mean?
A:

- Recall structural equation (2) saving $= \alpha_2$hsngexp $+ \beta_2 Z_2 + u_2$
- If $\alpha_2 = 0$, then

A direct channel of Bias - Cont'd

Q: What does $\alpha_2 \neq 0$ mean?
A:

- Recall structural equation (2) saving $= \alpha_2$hsngexp $+ \beta_2 Z_2 + u_2$
- If $\alpha_2 = 0$, then saving was not simultaneously determined by hsngexp.

A direct channel of Bias - Cont'd

Q: What does $\alpha_2 \neq 0$ mean?
A:

- Recall structural equation (2) saving $= \alpha_2$hsngexp $+ \beta_2 Z_2 + u_2$
- If $\alpha_2 = 0$, then saving was not simultaneously determined by hsngexp.
- We could then estimate structural equation (1) hsngexp $= \alpha_1$saving $+ \beta_1 Z_1 + u_1$ via OLS without trouble.

A direct channel of Bias - Cont'd

Q: What does $\alpha_2 \neq 0$ mean?
A:

- Recall structural equation (2) saving $= \alpha_2$hsngexp $+ \beta_2 Z_2 + u_2$
- If $\alpha_2 = 0$, then saving was not simultaneously determined by hsngexp.
- We could then estimate structural equation (1) hsngexp $= \alpha_1$saving $+ \beta_1 Z_1 + u_1$ via OLS without trouble.
- But, if our "DAG" is correct, and we truly have a simultaneous system, then $\alpha_2 \neq 0$, implying that we do have a bias in our estimate of α_1.

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.
Suppose for simplicity that we drop Z_1 from the structural equation (1) so that hsngexp $= \alpha_1$saving $+ u_1$ and that $corr(u_1, u_2) = 0$.

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.
Suppose for simplicity that we drop Z_1 from the structural equation (1) so that hsngexp $= \alpha_1$saving $+ u_1$ and that $corr(u_1, u_2) = 0$.Then we have

$$\begin{aligned} Cov(\text{saving}, u_1) &= Cov(v_2, u_1) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]E(u_1^2) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]\sigma_1^2 \end{aligned}$$

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.
Suppose for simplicity that we drop Z_1 from the structural equation (1) so that hsngexp $= \alpha_1$saving $+ u_1$ and that $corr(u_1, u_2) = 0$.Then we have

$$\begin{aligned} Cov(\text{saving}, u_1) &= Cov(v_2, u_1) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]E(u_1^2) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]\sigma_1^2 \end{aligned}$$

Again, this quantity should be 0 to avoid the simultaneity bias.

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.
Suppose for simplicity that we drop Z_1 from the structural equation (1) so that hsngexp $= \alpha_1$saving $+ u_1$ and that $corr(u_1, u_2) = 0$.Then we have

$$\begin{aligned} Cov(\text{saving}, u_1) &= Cov(v_2, u_1) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]E(u_1^2) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]\sigma_1^2 \end{aligned}$$

Again, this quantity should be 0 to avoid the simultaneity bias.
Suppose $\alpha_2 > 0$ (i.e. impact of hsngexp upon saving is positive) and $\alpha_2\alpha_1 < 1$, then the bias is

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.
Suppose for simplicity that we drop Z_1 from the structural equation (1) so that hsngexp $= \alpha_1$saving $+ u_1$ and that $corr(u_1, u_2) = 0$.Then we have

$$\begin{aligned} Cov(\text{saving}, u_1) &= Cov(v_2, u_1) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]E(u_1^2) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]\sigma_1^2 \end{aligned}$$

Again, this quantity should be 0 to avoid the simultaneity bias.
Suppose $\alpha_2 > 0$ (i.e. impact of hsngexp upon saving is positive) and $\alpha_2\alpha_1 < 1$, then the bias is positive (i.e.

Signing the Direct Bias Channel

Similar to our first encounter with OVB, we may not be able to determine the magnitude, but domain knowledge might help us determine the sign of the bias.
Suppose for simplicity that we drop Z_1 from the structural equation (1) so that hsngexp $= \alpha_1$saving $+ u_1$ and that $corr(u_1, u_2) = 0$.Then we have

$$\begin{aligned} Cov(\text{saving}, u_1) &= Cov(v_2, u_1) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]E(u_1^2) \\ &= [\alpha_2/(1 - \alpha_2\alpha_1)]\sigma_1^2 \end{aligned}$$

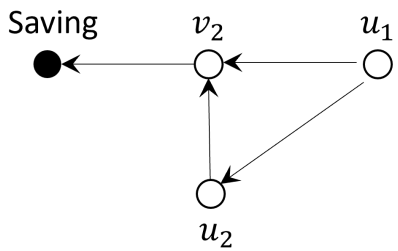
Again, this quantity should be 0 to avoid the simultaneity bias.
Suppose $\alpha_2 > 0$ (i.e. impact of hsngexp upon saving is positive) and $\alpha_2\alpha_1 < 1$, then the bias is positive (i.e. our estimate of α_1 is too large).

Indirect Channel of Bias

The other possible channel to induce bias is that saving is related to u_1 indirectly via u_2

Indirect Channel of Bias

The other possible channel to induce bias is that saving is related to u_1 indirectly via u_2

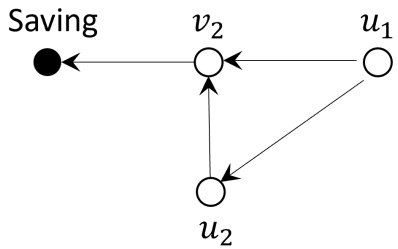


$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

Indirect Channel of Bias

The other possible channel to induce bias is that saving is related to u_1 indirectly via u_2



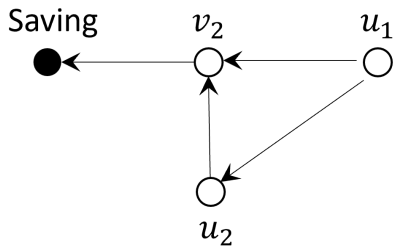
$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

This relationship would imply that the structural errors are related to one another.

Indirect Channel of Bias

The other possible channel to induce bias is that saving is related to u_1 indirectly via u_2



$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

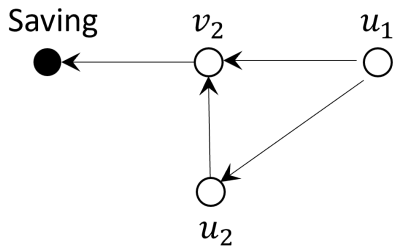
$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

This relationship would imply that the structural errors are related to one another.

A typical example might be an omitted variable that is common to both structural equations (e.g.

Indirect Channel of Bias

The other possible channel to induce bias is that saving is related to u_1 indirectly via u_2



$$\text{hsngexp} = \alpha_1 \text{saving} + \beta_1 Z_1 + u_1 \quad (1)$$

$$\text{saving} = \alpha_2 \text{hsngexp} + \beta_2 Z_2 + u_2 \quad (2)$$

This relationship would imply that the structural errors are related to one another.

A typical example might be an omitted variable that is common to both structural equations (e.g. GDP). Again, we likely can't determine the magnitude, but domain knowledge might tell us about the sign of the bias.

Bias Summary

Bias Summary

- A Direct Bias is almost ensured in a simultaneous system

Bias Summary

- A Direct Bias is almost ensured in a simultaneous system
- An Indirect Bias is akin to a typical OVB problem

Bias Summary

- A Direct Bias is almost ensured in a simultaneous system
- An Indirect Bias is akin to a typical OVB problem
- Direct and Indirect can plague a simultaneous system at the same time

Bias Summary

- A Direct Bias is almost ensured in a simultaneous system
- An Indirect Bias is akin to a typical OVB problem
- Direct and Indirect can plague a simultaneous system at the same time
- A well crafted IV solution should be able to deal with both problems

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

A Familiar Problem

- Instrumental Variables (IV) is a familiar tool to solve the problem of Omitted Variable Bias (OVB)
- We can employ the same IV solution here to simultaneity bias.

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Identification

Reminder on Identification with IV

- For our purposes, identification means we can estimate the model to elicit causality

Identification

Reminder on Identification with IV

- For our purposes, identification means we can estimate the model to elicit causality
- For identification, we need at least as many instruments as there are included endogenous variables.

Identification

Reminder on Identification with IV

- For our purposes, identification means we can estimate the model to elicit causality
- For identification, we need at least as many instruments as there are included endogenous variables.
- For instance, if we have two variables in the same equation that are omitted "good controls", then we would need at least two instruments.

Identification in a 2-Equation System

Consider the structural equations

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \tag{3}$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \tag{4}$$

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \tag{3}$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \tag{4}$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \tag{3}$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \tag{4}$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?
A:

- We can use Z_1 as an instrument in (4), and hence it is "identified".

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \tag{3}$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \tag{4}$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?
A:

- We can use Z_1 as an instrument in (4), and hence it is "identified".
- However, since Z_1 appears in equation (3), it cannot be used as an instrument there, and hence (3) is not identified.

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \tag{3}$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \tag{4}$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?
A:

- We can use Z_1 as an instrument in (4), and hence it is "identified".
- However, since Z_1 appears in equation (3), it cannot be used as an instrument there, and hence (3) is not identified.

Q: How does this impact our Client?

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \quad (3)$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \quad (4)$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?
A:

- We can use Z_1 as an instrument in (4), and hence it is "identified".
- However, since Z_1 appears in equation (3), it cannot be used as an instrument there, and hence (3) is not identified.

Q: How does this impact our Client?
A: Our Client seems most interested in α_1. But we can only estimate α_2 reliably.

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \quad (3)$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \quad (4)$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?
A:

- We can use Z_1 as an instrument in (4), and hence it is "identified".
- However, since Z_1 appears in equation (3), it cannot be used as an instrument there, and hence (3) is not identified.

Q: How does this impact our Client?
A: Our Client seems most interested in α_1. But we can only estimate α_2 reliably. We need to search for more instruments.

Identification in a 2-Equation System

Consider the structural equations

$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \quad (3)$$

$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \quad (4)$$

Q: Suppose hsngexp, savings, and Z_1 are the only variables available to us. Which equation is identified?
A:

- We can use Z_1 as an instrument in (4), and hence it is "identified".
- However, since Z_1 appears in equation (3), it cannot be used as an instrument there, and hence (3) is not identified.

Q: How does this impact our Client?
A: Our Client seems most interested in α_1. But we can only estimate α_2 reliably. We need to search for more instruments. Let's continue with equation (4) for now.

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4)

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4) (i.e. you have something leftover in your dataset that you can use as an instrument)

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4) (i.e. you have something leftover in your dataset that you can use as an instrument)
2. Rank Condition: at least one exogenous variable excluded from the second equation has a non zero population coefficient in the first equation.

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4) (i.e. you have something leftover in your dataset that you can use as an instrument)
2. Rank Condition: at least one exogenous variable excluded from the second equation has a non zero population coefficient in the first equation. (i.e. your instrument for hsngexp in the savings equation must be related to hsnexp)

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4) (i.e. you have something leftover in your dataset that you can use as an instrument)
2. Rank Condition: at least one exogenous variable excluded from the second equation has a non zero population coefficient in the first equation. (i.e. your instrument for hsngexp in the savings equation must be related to hsnexp) We can test via t-test or F-test by estimating the first equation.

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4) (i.e. you have something leftover in your dataset that you can use as an instrument)
2. Rank Condition: at least one exogenous variable excluded from the second equation has a non zero population coefficient in the first equation. (i.e. your instrument for hsngexp in the savings equation must be related to hsnexp) We can test via t-test or F-test by estimating the first equation.

   For instance, in our system

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

   we would test

Rank and Order Conditions

To identify Equation(4) in our two equation system requires

1. Order Condition: at least one exogenous variable is excluded from Equation(4) (i.e. you have something leftover in your dataset that you can use as an instrument)
2. Rank Condition: at least one exogenous variable excluded from the second equation has a non zero population coefficient in the first equation. (i.e. your instrument for hsngexp in the savings equation must be related to hsnexp) We can test via t-test or F-test by estimating the first equation.

   For instance, in our system

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

   we would test $\beta_{11} = 0$. If reject, then using Z_1 as an instrument in the second equation satisfies the Rank Condition.

- Order Condition $\approx$ Instrument Exogeneity

- Order Condition $\approx$ Instrument Exogeneity
  - Being omitted from Equation(4)(i.e. Order Condition) is akin to the instrument not being related to saving (i.e. Instrument Exogeneity)

- Order Condition $\approx$ Instrument Exogeneity
  - Being omitted from Equation(4)(i.e. Order Condition) is akin to the instrument not being related to saving (i.e. Instrument Exogeneity)
- Rank Condition $\approx$ Instrument Relevance

- Order Condition $\approx$ Instrument Exogeneity
  - Being omitted from Equation(4)(i.e. Order Condition) is akin to the instrument not being related to saving (i.e. Instrument Exogeneity)
- Rank Condition $\approx$ Instrument Relevance
  - Having a nonzero population coefficient in Equation(3) (i.e. Rank Condition) is akin to being related to the potentially endogenous hsngexp (i.e. Instrument Relevance).

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

IV Solution

Consider again

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

To implement IV on the second equation we need an instrument that is

IV Solution

Consider again

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

To implement IV on the second equation we need an instrument that is

- Correlated with hsngexp

IV Solution

Consider again

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

To implement IV on the second equation we need an instrument that is

- Correlated with hsngexp (i.e. Relevance)

IV Solution

Consider again

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

To implement IV on the second equation we need an instrument that is

- Correlated with hsngexp (i.e. Relevance)
- Not correlated with u_2

IV Solution

Consider again

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

To implement IV on the second equation we need an instrument that is

- Correlated with hsngexp (i.e. Relevance)
- Not correlated with u_2 (i.e. Independence)

IV Solution

Consider again

$$\begin{aligned}\text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2\end{aligned}$$

To implement IV on the second equation we need an instrument that is

- Correlated with hsngexp (i.e. Relevance)
- Not correlated with u_2 (i.e. Independence)
- Does not appear in the second equation (i.e. Exclusive)

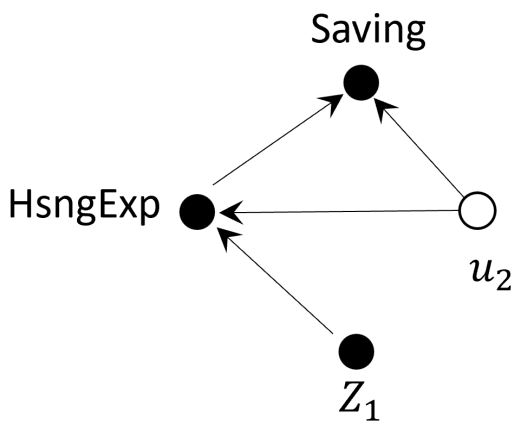
IV Solution

Consider again

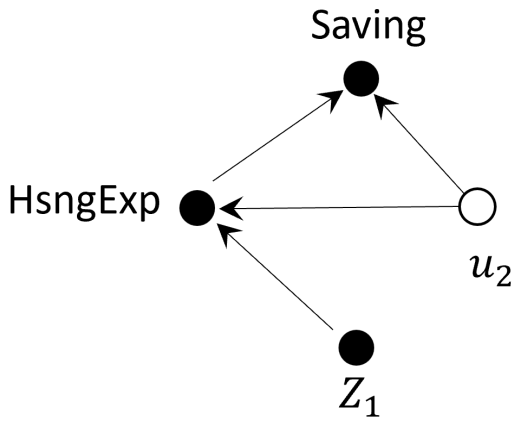
$$\text{hsngexp} = \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1$$
$$\text{saving} = \beta_{20} + \alpha_2 \text{hsngexp} + u_2$$

To implement IV on the second equation we need an instrument that is

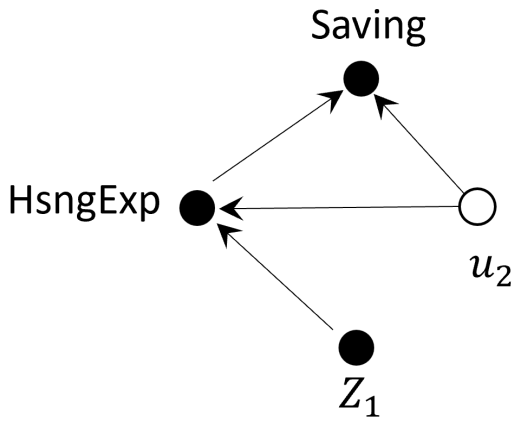
- Correlated with hsngexp (i.e. Relevance)
- Not correlated with u_2 (i.e. Independence)
- Does not appear in the second equation (i.e. Exclusive)



A potential instrument

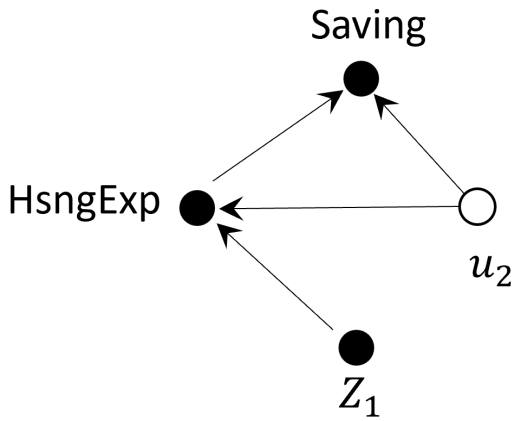


A potential instrument



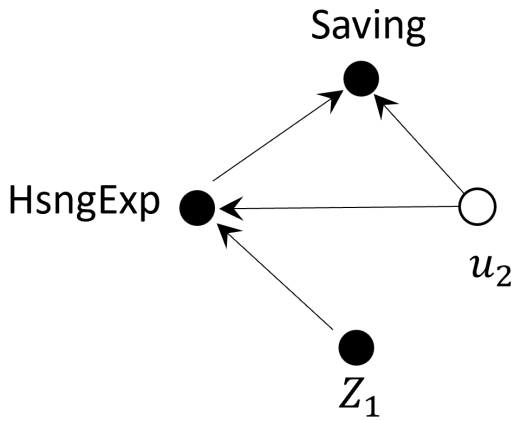
Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t.

A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

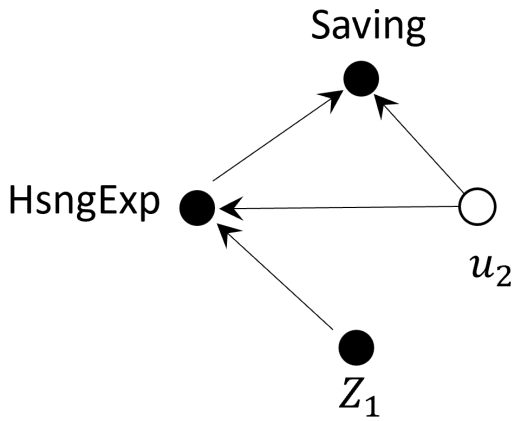
A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

- Relevance:

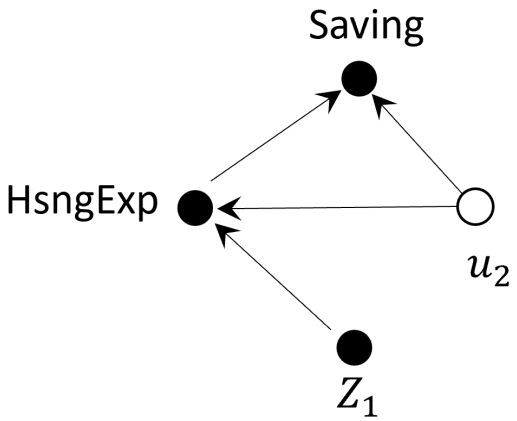
A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

- Relevance: $Z_{1,it}$ shifts hsngexp_{it} because renewed leases incorporate the higher amenity fee

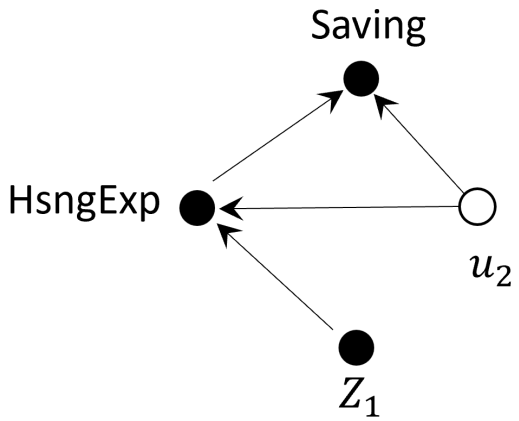
A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

- Relevance: $Z_{1,it}$ shifts hsngexp_{it} because renewed leases incorporate the higher amenity fee
- Independence:

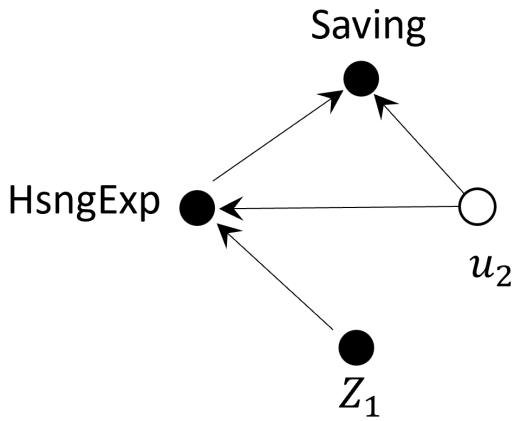
A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

- Relevance: $Z_{1,it}$ shifts hsngexp_{it} because renewed leases incorporate the higher amenity fee
- Independence: $Z_{1,it}$ is uncorrelated with the unobserved determinants of savings, i.e. $Cov(Z_{it}, u_{2,it}) = 0$, such as income, financial literacy, etc..

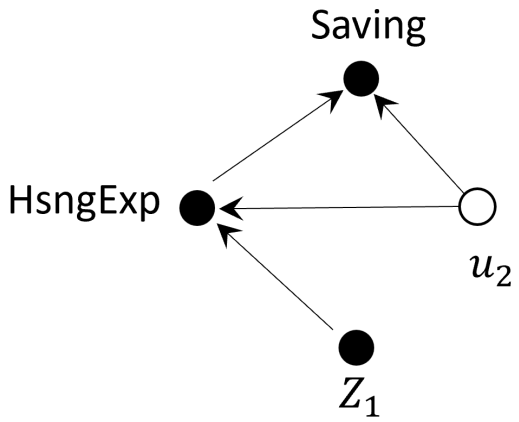
A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

- Relevance: $Z_{1,it}$ shifts hsngexp_{it} because renewed leases incorporate the higher amenity fee
- Independence: $Z_{1,it}$ is uncorrelated with the unobserved determinants of savings, i.e. $Cov(Z_{it}, u_{2,it}) = 0$, such as income, financial literacy, etc..
- Exclusion:

A potential instrument



Instrument: $Z_{1,it}$ = Amenity-fee increase *applied to tenant i at lease renewal* in month t. We can assume that the fee schedule is set administratively and renewal timing is predetermined by lease start dates.

- Relevance: $Z_{1,it}$ shifts hsngexp_{it} because renewed leases incorporate the higher amenity fee
- Independence: $Z_{1,it}$ is uncorrelated with the unobserved determinants of savings, i.e. $Cov(Z_{it}, u_{2,it}) = 0$, such as income, financial literacy, etc..
- Exclusion: $Z_{1,it}$ affects saving_{it} only through its effect on hsngexp_{it} (no other channel)

No Bias with IV

Recall the IV estimator

$$\hat{\alpha}_2^{IV} = \frac{Cov(Z_1, \text{saving})}{Cov(Z_1, \text{hsngexp})}$$

No Bias with IV

Recall the IV estimator

$$\begin{aligned}\hat{\alpha}_2^{IV} &= \frac{Cov(Z_1, \text{saving})}{Cov(Z_1, \text{hsngexp})} \\ &= \frac{Cov(Z_1, \beta_{20} + \alpha_2 \text{hsngexp} + u_2)}{Cov(Z_1, \text{hsngexp})}\end{aligned}$$

No Bias with IV

Recall the IV estimator

$$\begin{aligned}\hat{\alpha}_2^{IV} &= \frac{Cov(Z_1, \text{saving})}{Cov(Z_1, \text{hsngexp})} \\ &= \frac{Cov(Z_1, \beta_{20} + \alpha_2 \text{hsngexp} + u_2)}{Cov(Z_1, \text{hsngexp})} \\ &= \frac{\alpha_2 Cov(Z_1, \text{hsngexp}) + Cov(Z_1, u_2)}{Cov(Z_1, \text{hsngexp})}\end{aligned}$$

No Bias with IV

Recall the IV estimator

$$\begin{aligned}
\hat{\alpha}_2^{IV} &= \frac{Cov(Z_1, \text{saving})}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{Cov(Z_1, \beta_{20} + \alpha_2 \text{hsngexp} + u_2)}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{\alpha_2 Cov(Z_1, \text{hsngexp}) + Cov(Z_1, u_2)}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{\alpha_2 Cov(Z_1, \text{hsngexp})}{Cov(Z_1, \text{hsngexp})}
\end{aligned}$$

No Bias with IV

Recall the IV estimator

$$\begin{aligned}
\hat{\alpha}_2^{IV} &= \frac{Cov(Z_1, \text{saving})}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{Cov(Z_1, \beta_{20} + \alpha_2 \text{hsngexp} + u_2)}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{\alpha_2 Cov(Z_1, \text{hsngexp}) + Cov(Z_1, u_2)}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{\alpha_2 Cov(Z_1, \text{hsngexp})}{Cov(Z_1, \text{hsngexp})} \\
&= \alpha_2
\end{aligned}$$

No Bias with IV

Recall the IV estimator

$$\begin{aligned}
\hat{\alpha}_2^{IV} &= \frac{Cov(Z_1, \text{saving})}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{Cov(Z_1, \beta_{20} + \alpha_2 \text{hsngexp} + u_2)}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{\alpha_2 Cov(Z_1, \text{hsngexp}) + Cov(Z_1, u_2)}{Cov(Z_1, \text{hsngexp})} \\
&= \frac{\alpha_2 Cov(Z_1, \text{hsngexp})}{Cov(Z_1, \text{hsngexp})} \\
&= \alpha_2
\end{aligned}$$

Implying that the IV estimator of α_2 is unbiased

IV Estimation Summary

In summary, theory and SME's produce the structural equations

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

IV Estimation Summary

In summary, theory and SME's produce the structural equations

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

- Our objects of interest α_1, α_2 cannot be estimated by OLS due to a simultaneity bias

IV Estimation Summary

In summary, theory and SME's produce the structural equations

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

- Our objects of interest α_1, α_2 cannot be estimated by OLS due to a simultaneity bias
- The second equation is identified because Z_1 is omitted (i.e. Order Condition), but Z_1 is included in the first equation (i.e. Rank Condition).

IV Estimation Summary

In summary, theory and SME's produce the structural equations

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

- Our objects of interest α_1, α_2 cannot be estimated by OLS due to a simultaneity bias
- The second equation is identified because Z_1 is omitted (i.e. Order Condition), but Z_1 is included in the first equation (i.e. Rank Condition).
- We can overcome the bias in our α_2 estimate via IV by using Z_1 as an instrument for hsngexp in the second equation.

IV Estimation Summary

In summary, theory and SME's produce the structural equations

$$\begin{aligned} \text{hsngexp} &= \beta_{10} + \alpha_1 \text{saving} + \beta_{11} Z_1 + u_1 \\ \text{saving} &= \beta_{20} + \alpha_2 \text{hsngexp} + u_2 \end{aligned}$$

- Our objects of interest α_1, α_2 cannot be estimated by OLS due to a simultaneity bias
- The second equation is identified because Z_1 is omitted (i.e. Order Condition), but Z_1 is included in the first equation (i.e. Rank Condition).
- We can overcome the bias in our α_2 estimate via IV by using Z_1 as an instrument for hsngexp in the second equation.
- The parameter α_1 cannot be estimated without bias.

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Impact of Price Changes on Milk Demand/Supply



Image: Bing Image Creator

Supply

Consider the market for milk in the USA.

$$q_s = \tilde{\alpha}_0 + \tilde{\alpha}_1 p + \tilde{\alpha}_2 z_1 + u_s$$

where

- q_s is quantity of milk supplied by farmers (millions of gallons) each year
- p is price per gallon
- z_1 is price of cow feed

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Q: What are the meanings and signs of α_1 and α_2?

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Q: What are the meanings and signs of α_1 and α_2?

A:

- α_1 is

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Q: What are the meanings and signs of α_1 and α_2?

A:

- α_1 is elasticity of supply with respect to price (aka price elasticity of supply), which we expect to be

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Q: What are the meanings and signs of α_1 and α_2?
A:

- α_1 is elasticity of supply with respect to price (aka price elasticity of supply), which we expect to be positive

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Q: What are the meanings and signs of α_1 and α_2?
A:

- α_1 is elasticity of supply with respect to price (aka price elasticity of supply), which we expect to be positive
- α_2 is

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

Q: What are the meanings and signs of α_1 and α_2?
A:

- α_1 is elasticity of supply with respect to price (aka price elasticity of supply), which we expect to be positive
- α_2 is elasticity of supply with respect to the price of cow feed, which we expect to be

Recasting Supply for Elasticities

Q: How recast the supply equation for elasticity?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$

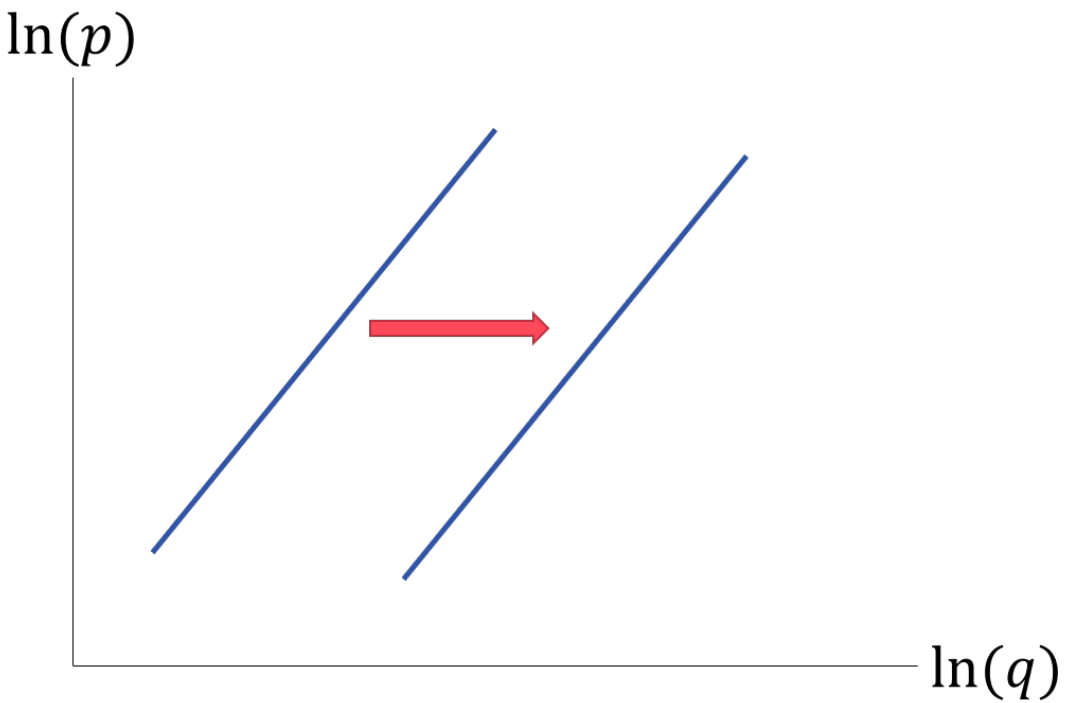
Q: What are the meanings and signs of α_1 and α_2?
A:

- α_1 is elasticity of supply with respect to price (aka price elasticity of supply), which we expect to be positive
- α_2 is elasticity of supply with respect to the price of cow feed, which we expect to be negative

Shifting the Supply Curve

Q: What might cause Supply to shift right?

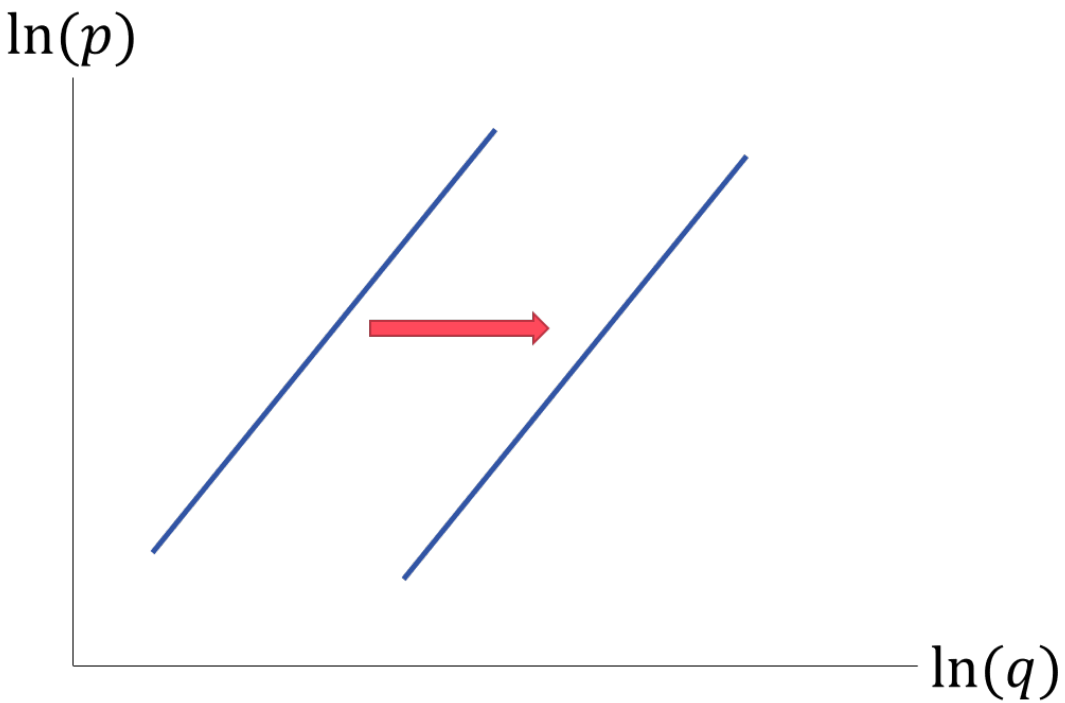
$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$



Shifting the Supply Curve

Q: What might cause Supply to shift right?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$



A: As the price of cow feed z_1 declines, the quantity supplied rises for all price levels of milk.

Demand

Consider the market for milk in the USA.

$$q_d = \tilde{\beta}_0 + \tilde{\beta}_1 p + u_d$$

where

- q_d is quantity of milked demanded (millions of gallons) each year
- p is price per gallon

Recasting Demand for Elasticities

By transforming all variables of the demand equation into logs, the coefficients are now elasticities

$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

Recasting Demand for Elasticities

By transforming all variables of the demand equation into logs, the coefficients are now elasticities

$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

Q: What is the meaning and sign of β_1?

Recasting Demand for Elasticities

By transforming all variables of the demand equation into logs, the coefficients are now elasticities

$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

Q: What is the meaning and sign of β_1?
A:

- β_1 is

Recasting Demand for Elasticities

By transforming all variables of the demand equation into logs, the coefficients are now elasticities

$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

Q: What is the meaning and sign of β_1?
A:

- β_1 is elasticity of demand with respect to price (aka price elasticity of demand), which we expect to be

Recasting Demand for Elasticities

By transforming all variables of the demand equation into logs, the coefficients are now elasticities

$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

Q: What is the meaning and sign of β_1?
A:

- β_1 is elasticity of demand with respect to price (aka price elasticity of demand), which we expect to be negative

Identification

Q: Which equation is identified?

$$
\begin{aligned}
ln(q_s) &= \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s \\
ln(q_d) &= \beta_0 + \beta_1 ln(p) + u_d
\end{aligned}
$$

Identification

Q: Which equation is identified?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$
$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

A:

- Demand curve is identified

Identification

Q: Which equation is identified?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$
$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

A:

- Demand curve is identified
- Order Condition: $ln(z_1)$ is omitted from demand

Identification

Q: Which equation is identified?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$
$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

A:

- Demand curve is identified
- Order Condition: $ln(z_1)$ is omitted from demand
- Rank Condition: $ln(z_1)$ is in supply (assuming non zero α_2)

Identification

Q: Which equation is identified?

$$ln(q_s) = \alpha_0 + \alpha_1 ln(p) + \alpha_2 ln(z_1) + u_s$$
$$ln(q_d) = \beta_0 + \beta_1 ln(p) + u_d$$

A:

- Demand curve is identified
- Order Condition: $ln(z_1)$ is omitted from demand
- Rank Condition: $ln(z_1)$ is in supply (assuming non zero α_2)
- Supply curve is not identified since we don't have a z_2 within demand but not supply.

Equilibrium

Equilibrium

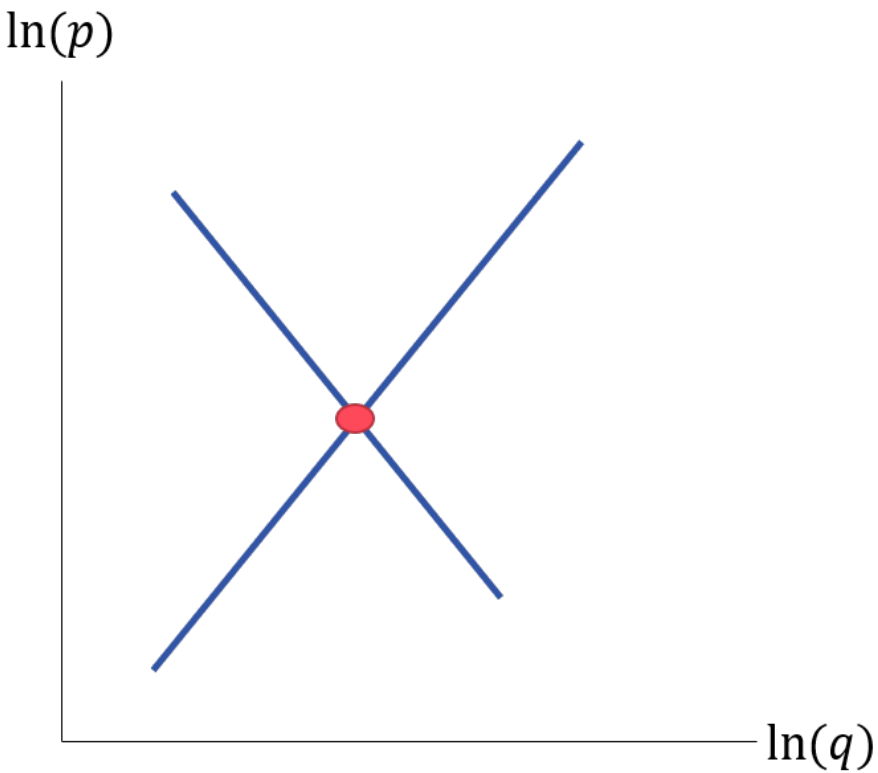
- We don't observe supply

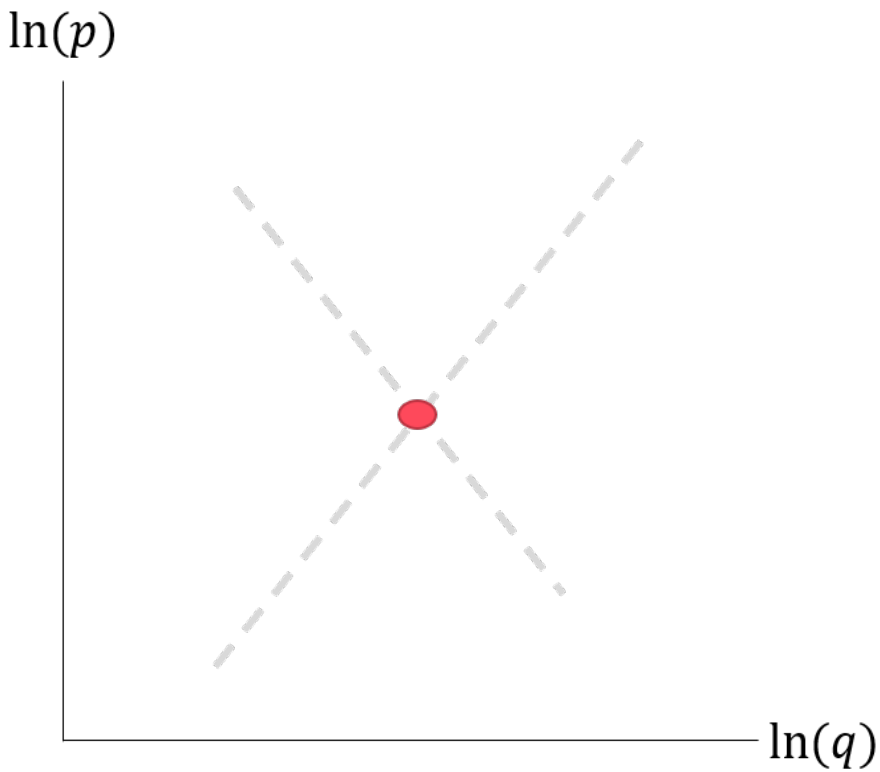
Equilibrium

- We don't observe supply
- We don't observe demand

Equilibrium

- We don't observe supply
- We don't observe demand
- We only observe equilibrium; e.g. $(q_s = q_d = q)$





Simultaneity Bias

Suppose we observe these (p, q) pairs for several countries (i). Then we can think of our equilibrium system as

$$\begin{aligned} ln(q_{si}) =& \alpha_0 + \alpha_1 ln(p_i) + \alpha_2 ln(z_{1i}) + u_{si} \\ ln(q_{di}) =& \beta_0 + \beta_1 ln(p_i) + u_{di} \end{aligned}$$

Simultaneity Bias

Suppose we observe these (p, q) pairs for several countries (i). Then we can think of our equilibrium system as

$$\begin{aligned} ln(q_{si}) =& \alpha_0 + \alpha_1 ln(p_i) + \alpha_2 ln(z_{1i}) + u_{si} \\ ln(q_{di}) =& \beta_0 + \beta_1 ln(p_i) + u_{di} \end{aligned}$$

Estimating each equation separately by OLS generates a simultaneity bias

IV Solution

Recall the Demand Equation was Identified

IV Solution

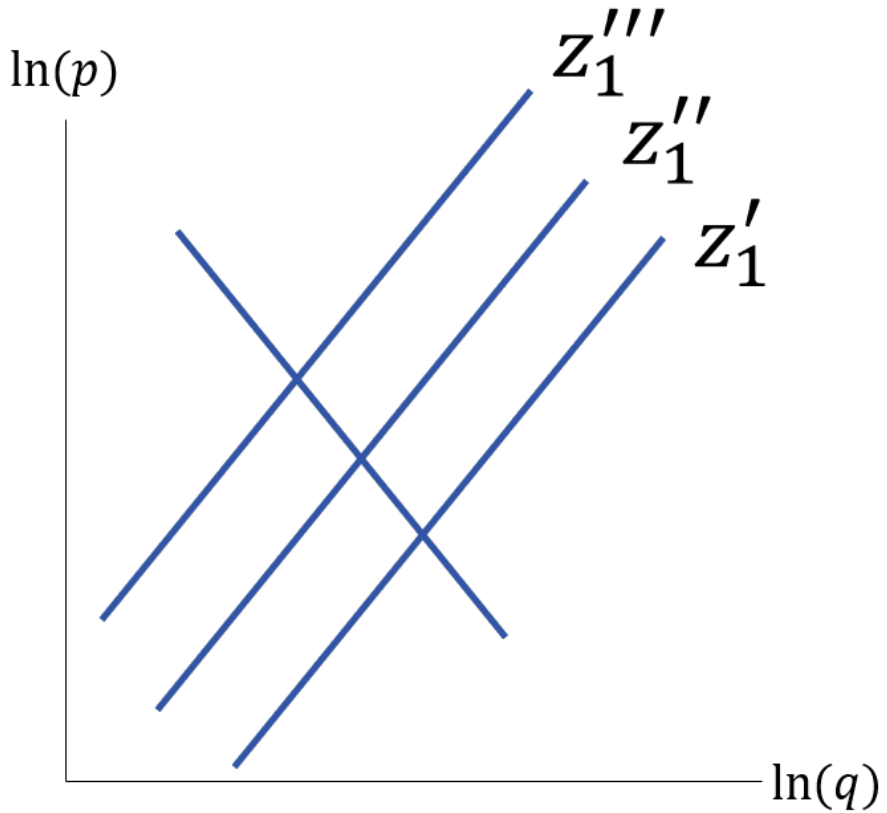
Recall the Demand Equation was Identified

By shifting *supply*, the instrument $ln(z_1)$ allows us to identify the demand curve and thereby estimate the price elasticity of demand β_1

IV Solution

Recall the Demand Equation was Identified

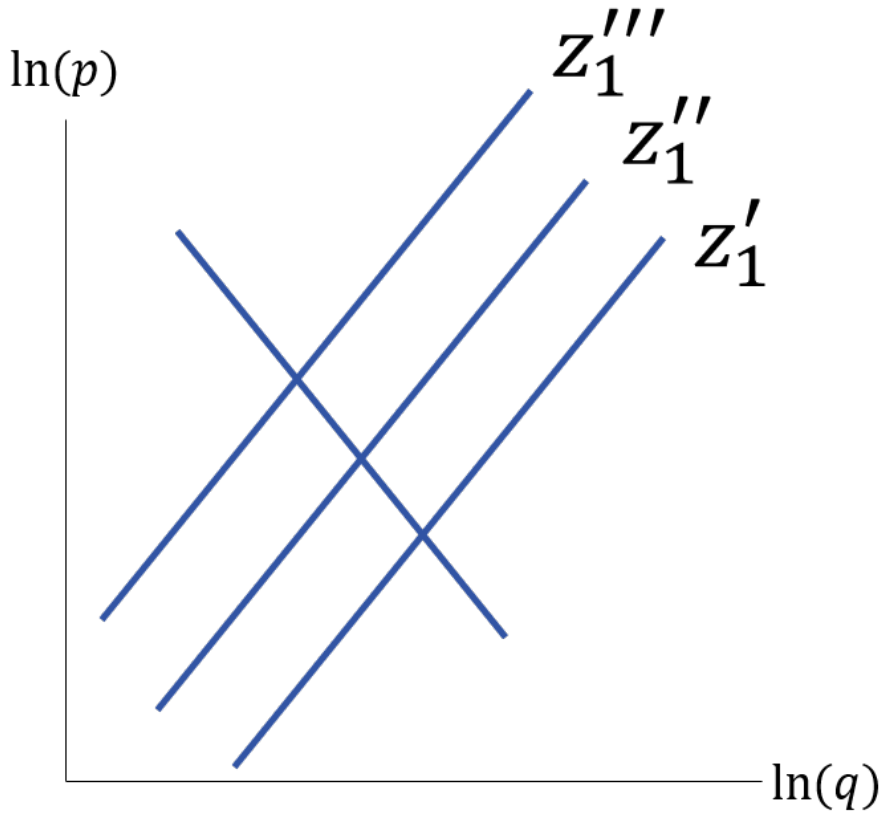
By shifting *supply*, the instrument $ln(z_1)$ allows us to identify the demand curve and thereby estimate the price elasticity of demand β_1



IV Solution

Recall the Demand Equation was Identified

By shifting *supply*, the instrument $ln(z_1)$ allows us to identify the demand curve and thereby estimate the price elasticity of demand β_1



Hence, IV helps to overcome simultaneity bias.

Outline

1. Case 1: Housing & Savings
2. Simultaneity Bias
   - Visualizing the Bias
   - Understanding the Bias
   - Bias Details
3. IV Solution
   - Identification
   - Estimation
4. Case 2: Supply & Demand
5. Case 3: Prices & Wages

Impact of Wage on Consumer Prices



Source: Bing Image Creator

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?
A:

- β_2

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?
A:

- $\beta_2 > 0$ via Wage Push Inflation

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?
A:

- $\beta_2 > 0$ via Wage Push Inflation
- α_2

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?
A:

- $\beta_2 > 0$ via Wage Push Inflation
- $\alpha_2 > 0$ via cost of living adjustments and union bargaining

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?
A:

- $\beta_2 > 0$ via Wage Push Inflation
- $\alpha_2 > 0$ via cost of living adjustments and union bargaining
- α_3

Prices and Wages

Define p as annual growth rate of consumer prices, w as the annual growth rate of wages, and UR as the unemployment rate.
Structural equations

$$p = \beta_1 + \beta_2 w + u_p \tag{5}$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \tag{6}$$

Q: Expected signs & meanings for β_2,α_2, α_3?
A:

- $\beta_2 > 0$ via Wage Push Inflation
- $\alpha_2 > 0$ via cost of living adjustments and union bargaining
- $\alpha_3 < 0$ via labor market slack

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.
Q: What does exogeneity assumption require?

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.
Q: What does exogeneity assumption require?
A: $Cov(w, u_p) = 0$.

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.
Q: What does exogeneity assumption require?
A: $Cov(w, u_p) = 0$.
Q: Why is this violated?

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.
Q: What does exogeneity assumption require?
A: $Cov(w, u_p) = 0$.
Q: Why is this violated?
A: Rewrite the wage question (6)

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.
Q: What does exogeneity assumption require?
A: $Cov(w, u_p) = 0$.
Q: Why is this violated?
A: Rewrite the wage question (6)

$$\begin{aligned} w =& \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \\ =& \alpha_1 + \alpha_2(\beta_1 + \beta_2 w + u_p) + \alpha_3 UR + u_w \end{aligned}$$

Isolating the Bias

Let's focus on the price equation (5) $p = \beta_1 + \beta_2 w + u_p$.
Q: What does exogeneity assumption require?
A: $Cov(w, u_p) = 0$.
Q: Why is this violated?
A: Rewrite the wage question (6)

$$\begin{aligned} w =& \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \\ =& \alpha_1 + \alpha_2(\beta_1 + \beta_2 w + u_p) + \alpha_3 UR + u_w \end{aligned}$$

We can see that w is a function of u_p, i.e. $Cov(w, u_p) \neq 0$, and hence w is endogenous in the price equation.

Checking Order and Rank Conditions

Q: Are the Order and Rank conditions met for (5)?

$$p = \beta_1 + \beta_2 w + u_p \quad (5)$$

$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \quad (6)$$

Checking Order and Rank Conditions

Q: Are the Order and Rank conditions met for (5)?

$$p = \beta_1 + \beta_2 w + u_p \quad (5)$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \quad (6)$$

A:

1. Order Condition for (5): at least one exogenous variable is excluded from the first equation.

Checking Order and Rank Conditions

Q: Are the Order and Rank conditions met for (5)?

$$p = \beta_1 + \beta_2 w + u_p \quad (5)$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \quad (6)$$

A:

1. Order Condition for (5): at least one exogenous variable is excluded from the first equation. *UR is not in price equation* ✓

Checking Order and Rank Conditions

Q: Are the Order and Rank conditions met for (5)?

$$p = \beta_1 + \beta_2 w + u_p \quad (5)$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \quad (6)$$

A:

1. Order Condition for (5): at least one exogenous variable is excluded from the first equation. *UR is not in price equation* ✓
2. Rank Condition for (5): at least one exogenous variable excluded from the first equation has a non zero population coefficient in the second equation.

Checking Order and Rank Conditions

Q: Are the Order and Rank conditions met for (5)?

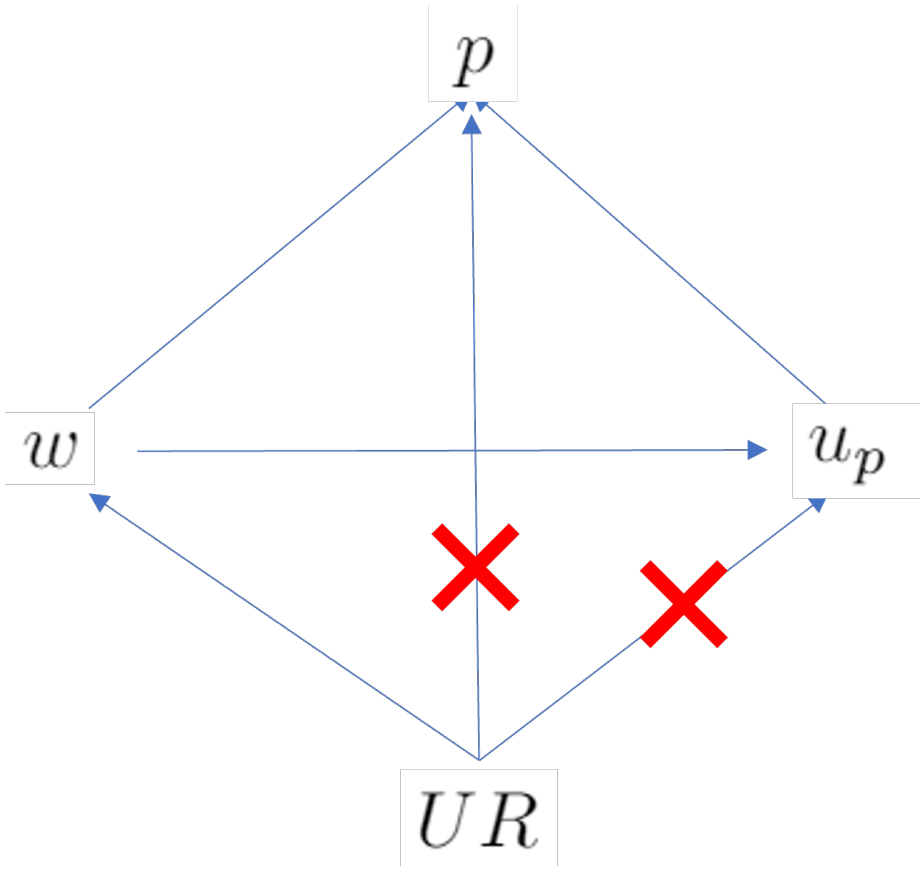
$$p = \beta_1 + \beta_2 w + u_p \quad (5)$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w \quad (6)$$

A:

1. Order Condition for (5): at least one exogenous variable is excluded from the first equation. *UR is not in price equation* ✓
2. Rank Condition for (5): at least one exogenous variable excluded from the first equation has a non zero population coefficient in the second equation. *SMEs tell us UR is important for wages* ✓

UR as an Instrument for Wages

Rank&Order conditions met. (5) is identified. *UR* potential instrument for *w*



Monte Carlo Evidence

Reduced Form

$$w = \frac{\alpha_1 + \alpha_2\beta_1}{1 - \alpha_2\beta_2} + \frac{\alpha_3 UR}{1 - \alpha_2\beta_2} + \frac{u_w + \alpha_2 u_p}{1 - \alpha_2\beta_2}$$

$$w = \frac{\alpha_1 + \alpha_2\beta_1}{1 - \alpha_2\beta_2} + \frac{\alpha_3 UR}{1 - \alpha_2\beta_2} + \frac{u_w + \alpha_2 u_p}{1 - \alpha_2\beta_2}$$
$$p = \frac{\beta_1 + \alpha_1\beta_2}{1 - \alpha_2\beta_2} + \frac{\alpha_3\beta_2 UR}{1 - \alpha_2\beta_2} + \frac{u_p + \beta_2 u_w}{1 - \alpha_2\beta_2}$$

$$w = \frac{\alpha_1 + \alpha_2\beta_1}{1 - \alpha_2\beta_2} + \frac{\alpha_3 UR}{1 - \alpha_2\beta_2} + \frac{u_w + \alpha_2 u_p}{1 - \alpha_2\beta_2}$$
$$p = \frac{\beta_1 + \alpha_1\beta_2}{1 - \alpha_2\beta_2} + \frac{\alpha_3\beta_2 UR}{1 - \alpha_2\beta_2} + \frac{u_p + \beta_2 u_w}{1 - \alpha_2\beta_2}$$

With these reduced form specifications, wages and prices are functions only of exogenous variables *UR* and errors. We can conduct a Monte Carlo experiment to illustrate the simultaneity bias.

Monte Carlo Evidence

Calibrated System

structural equations	$p = \beta_1 + \beta_2 w + u_p$	$w = \alpha_1 + \alpha_2 p + \alpha_3 U + u_w$

Monte Carlo Evidence

Calibrated System

structural equations	$p = \beta_1 + \beta_2 w + u_p$	$w = \alpha_1 + \alpha_2 p + \alpha_3 U + u_w$
	$p = 1.5 + 0.5w + u_p$	$w = 2.5 + 0.5p - 0.4U$

Monte Carlo Evidence

Calibrated System

structural equations

$$p = \beta_1 + \beta_2 w + u_p \qquad w = \alpha_1 + \alpha_2 p + \alpha_3 U + u_w$$

$$p = 1.5 + 0.5w + u_p \qquad w = 2.5 + 0.5p - 0.4U$$

reduced form equation

$$p = \frac{\beta_1 + \alpha_1\beta_2 + \alpha_3\beta_2 U + u_p + \beta_2 u_w}{1 - \alpha_2\beta_2} = 3.67 - 0.27U + 1.33u_p$$

reduced form equation

$$w = \frac{\alpha_1 + \alpha_2\beta_1 + \alpha_3 U + \alpha_2 u_p + u_w}{1 - \alpha_2\beta_2} = 4.33 - 0.53U + 0.67u_p$$

Reference: Dougherty "Econometrics" Ch9. U=UR from our price and wage example. u_w omitted for simplicity.

Monte Carlo Evidence

Errors Impact Both Equations

reduced form equation

$$p = \frac{\beta_1 + \alpha_1\beta_2 + \alpha_3\beta_2 U + u_p + \beta_2 u_w}{1 - \alpha_2\beta_2}$$

reduced form equation

$$w = \frac{\alpha_1 + \alpha_2\beta_1 + \alpha_3 U + \alpha_2 u_p + u_w}{1 - \alpha_2\beta_2}$$

Monte Carlo Evidence

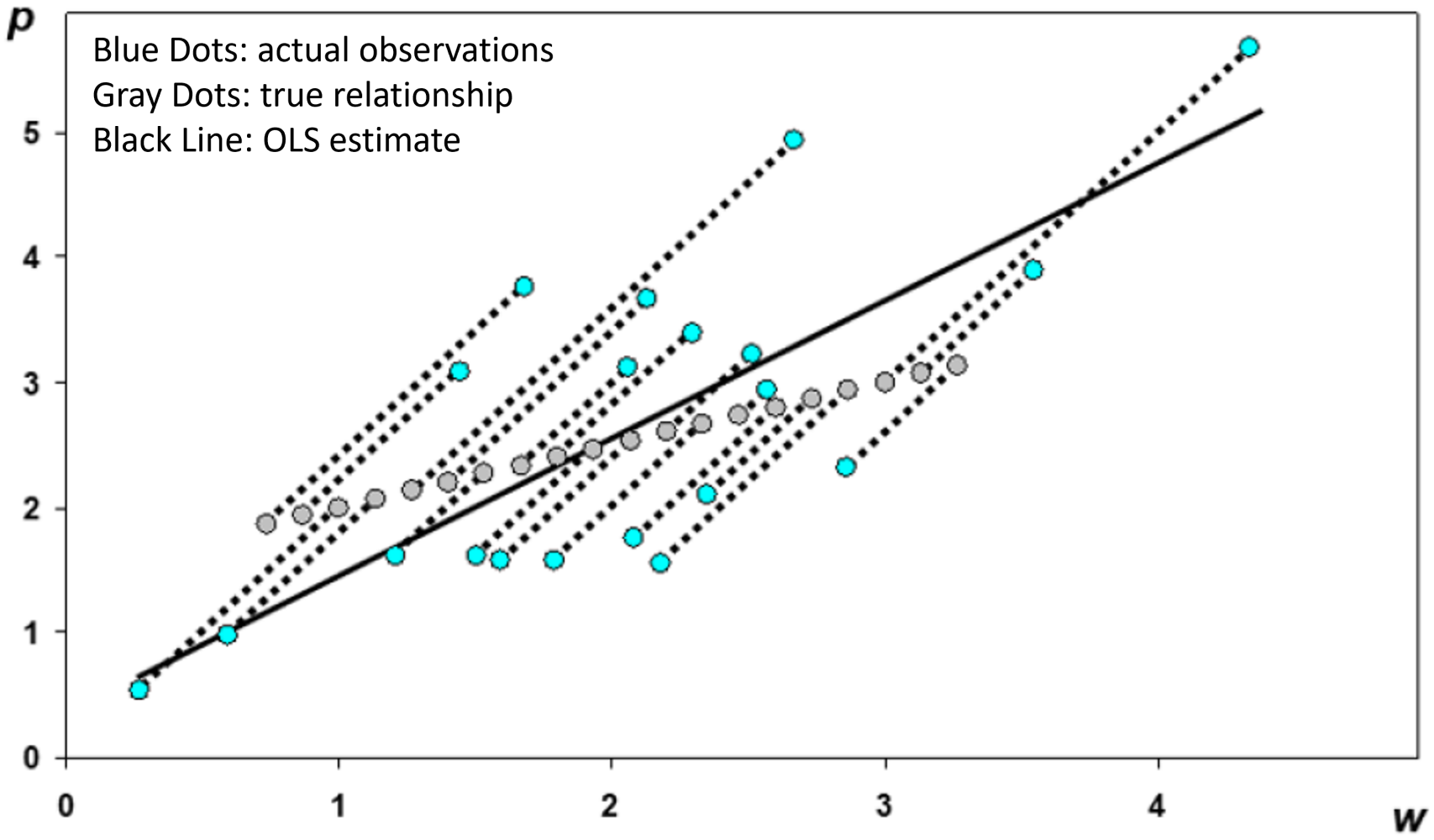
Errors Impact Both Equations

reduced form equation	$p = \frac{\beta_1 + \alpha_1\beta_2 + \alpha_3\beta_2 U + u_p + \beta_2 u_w}{1 - \alpha_2\beta_2}$
reduced form equation	$w = \frac{\alpha_1 + \alpha_2\beta_1 + \alpha_3 U + \alpha_2 u_p + u_w}{1 - \alpha_2\beta_2}$

- Both p and w are impacted by variations in u_p (i.e. the errors in the price equation impact the price AND wage equations)

Monte Carlo Evidence

Errors in Prices Impact Both Equations



Monte Carlo Evidence

Simulations

structural equations

$$p = \beta_1 + \beta_2 w + u_p \qquad w = \alpha_1 + \alpha_2 p + \alpha_3 U + u_w$$

$$p = 1.5 + 0.5w + u_p \qquad w = 2.5 + 0.5p - 0.4U$$

reduced form equation

$$p = \frac{\beta_1 + \alpha_1\beta_2 + \alpha_3\beta_2 U + u_p + \beta_2 u_w}{1 - \alpha_2\beta_2} = 3.67 - 0.27U + 1.33u_p$$

reduced form equation

$$w = \frac{\alpha_1 + \alpha_2\beta_1 + \alpha_3 U + \alpha_2 u_p + u_w}{1 - \alpha_2\beta_2} = 4.33 - 0.53U + 0.67u_p$$

n = 20.

U = 2 to 6.75 in steps of 0.25. Variance 2.08.

u_p ~ iid $N(0, 0.64)$.

u_w dropped.

Monte Carlo Evidence

Intercept is underestimated every time

structural equations

$$p = \beta_1 + \beta_2 w + u_p \qquad w = \alpha_1 + \alpha_2 p + \alpha_3 U + u_w$$

$$p = 1.5 + 0.5w + u_p \qquad w = 2.5 + 0.5p - 0.4U$$

	$\hat{\beta}_1$	s.e.($\hat{\beta}_1$)	$\hat{\beta}_2$	s.e.($\hat{\beta}_2$)
1	0.36	0.49	1.11	0.22
2	0.45	0.38	1.06	0.17
3	0.65	0.27	0.94	0.12
4	0.41	0.39	0.98	0.19
5	0.46	0.92	0.77	0.22
6	0.26	0.35	1.09	0.16
7	0.31	0.39	1.00	0.19
8	1.06	0.38	0.82	0.16
9	–0.08	0.36	1.16	0.18
10	1.12	0.43	0.69	0.20

Monte Carlo Evidence

Slope is overestimated every time

structural equations

$$p = \beta_1 + \beta_2 w + u_p \qquad w = \alpha_1 + \alpha_2 p + \alpha_3 U + u_w$$

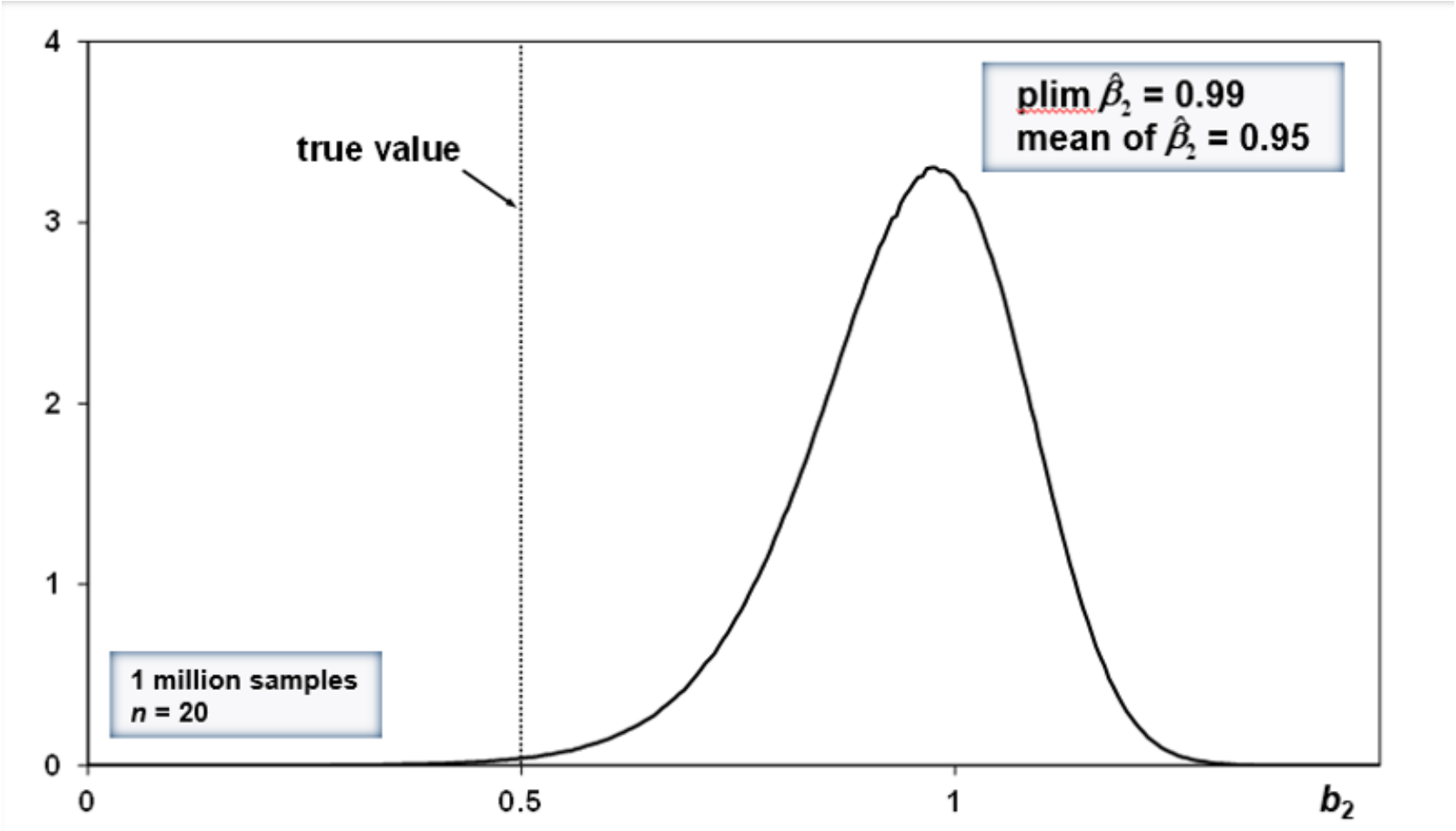
$$p = 1.5 + 0.5 w + u_p \qquad w = 2.5 + 0.5 p - 0.4 U$$

1st 10 samples

	$\hat{\beta}_1$	s.e.($\hat{\beta}_1$)	$\hat{\beta}_2$	s.e.($\hat{\beta}_2$)
1	0.36	0.49	1.11	0.22
2	0.45	0.38	1.06	0.17
3	0.65	0.27	0.94	0.12
4	0.41	0.39	0.98	0.19
5	0.46	0.92	0.77	0.22
6	0.26	0.35	1.09	0.16
7	0.31	0.39	1.00	0.19
8	1.06	0.38	0.82	0.16
9	–0.08	0.36	1.16	0.18
10	1.12	0.43	0.69	0.20

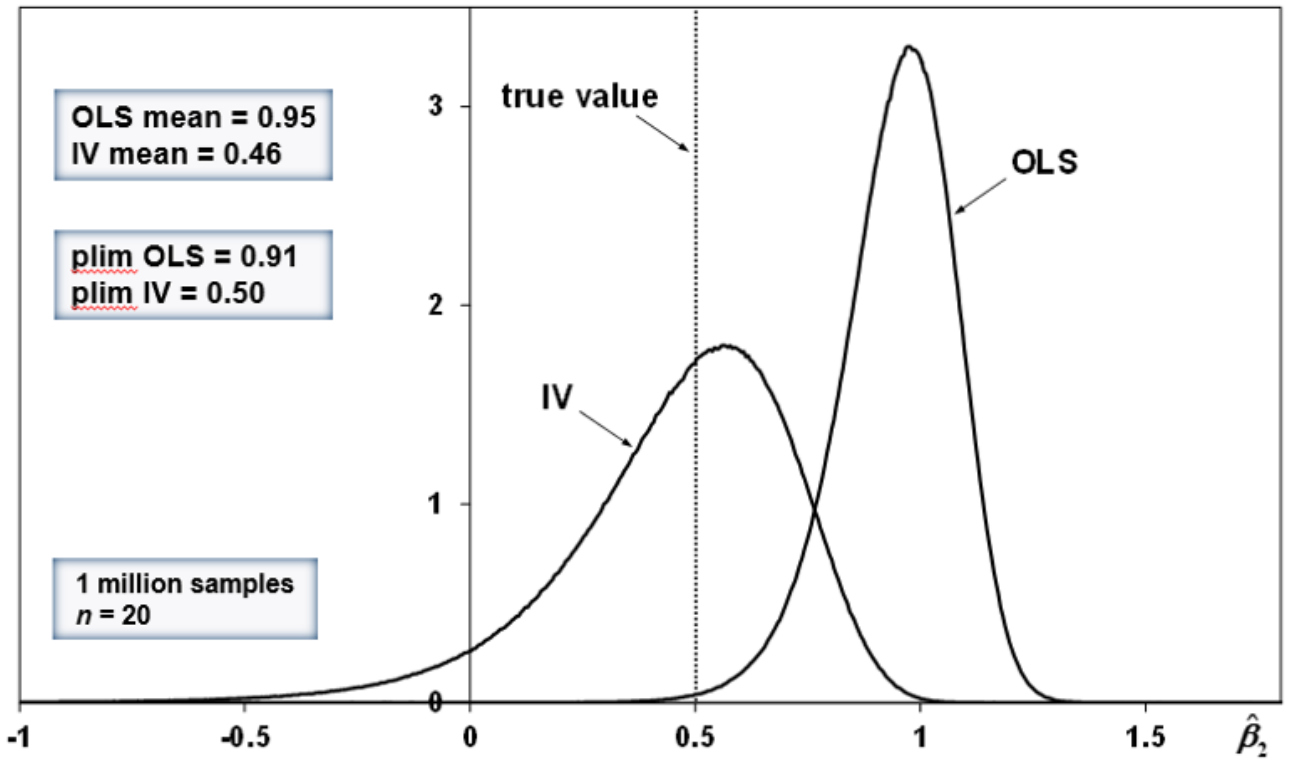
Monte Carlo Evidence

OLS slope (ATE) is overestimated



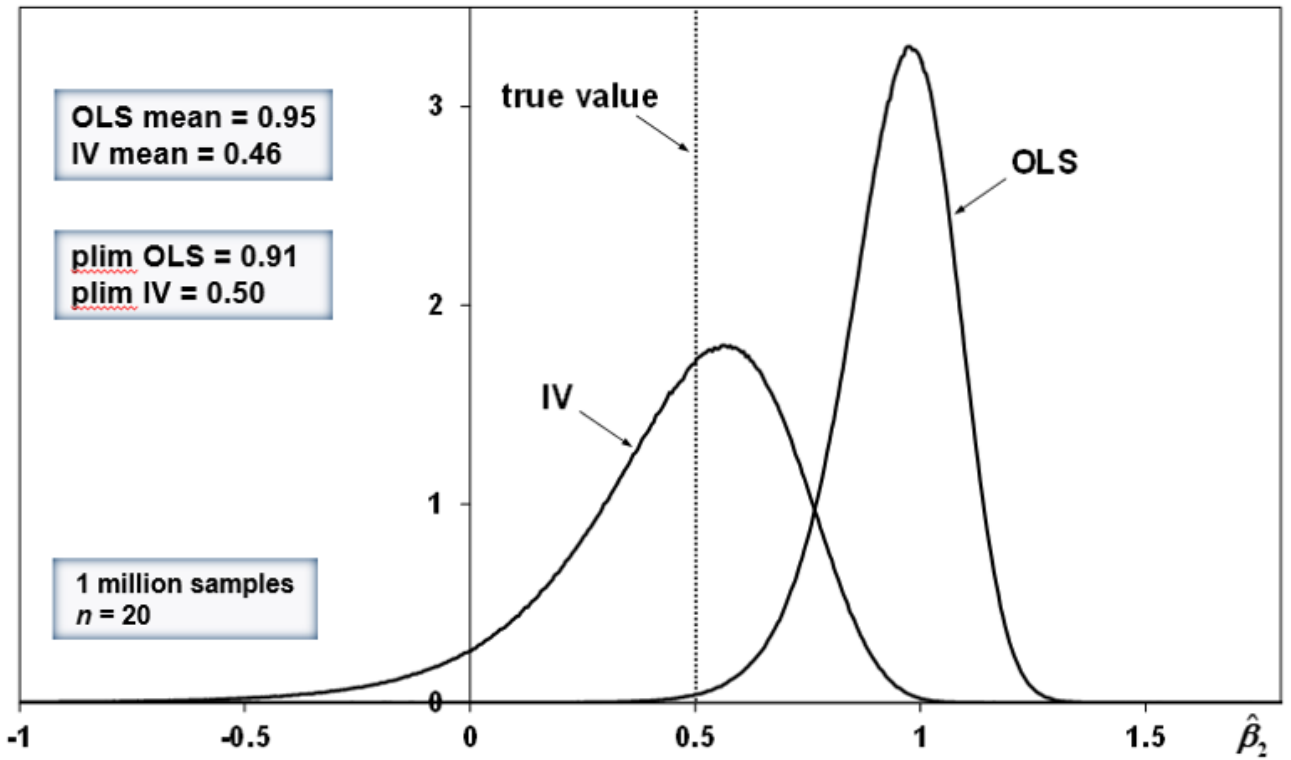
Monte Carlo Evidence

IV Solution



Monte Carlo Evidence

IV Solution



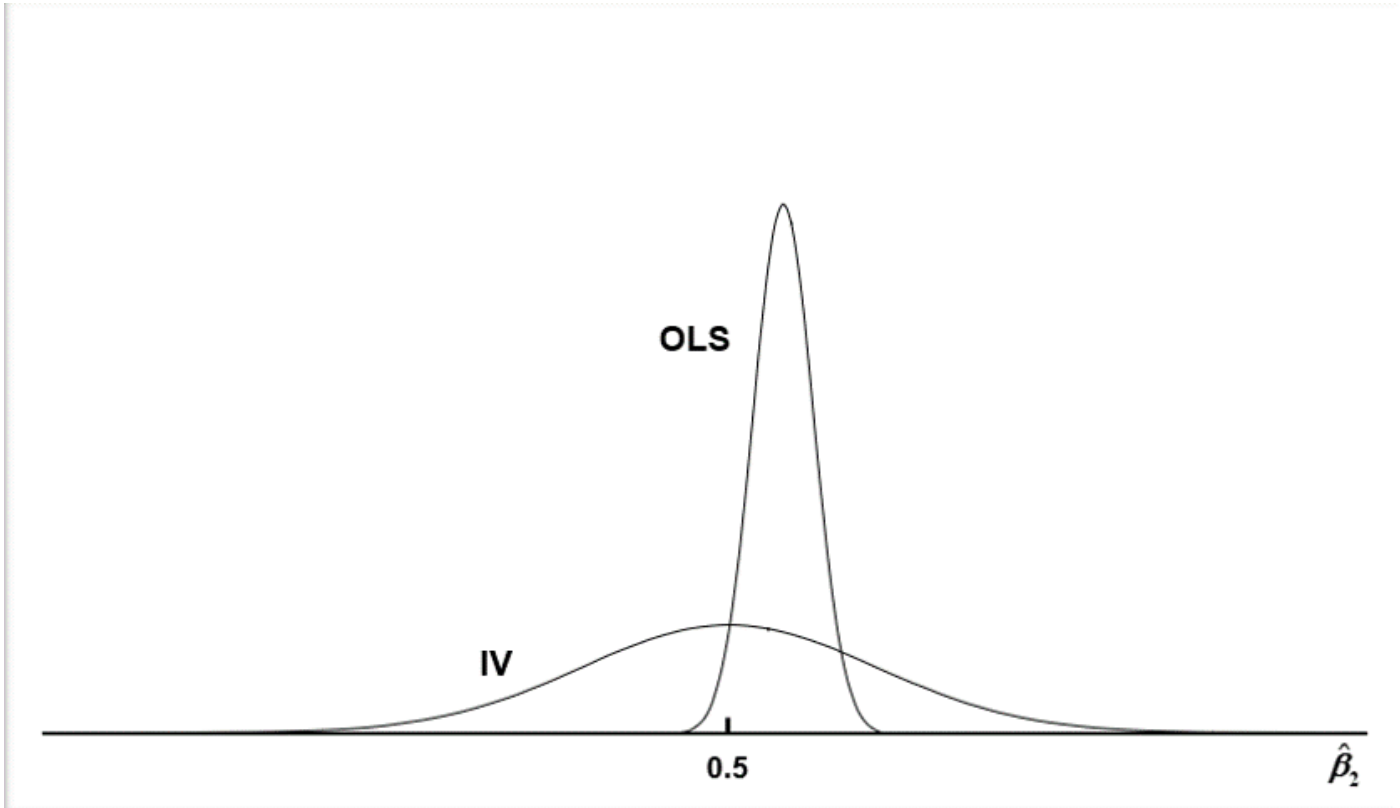
The IV estimator reduced the bias in the OLS estimator.

Monte Carlo Evidence

Q: What happens if the instruments are "weak"?

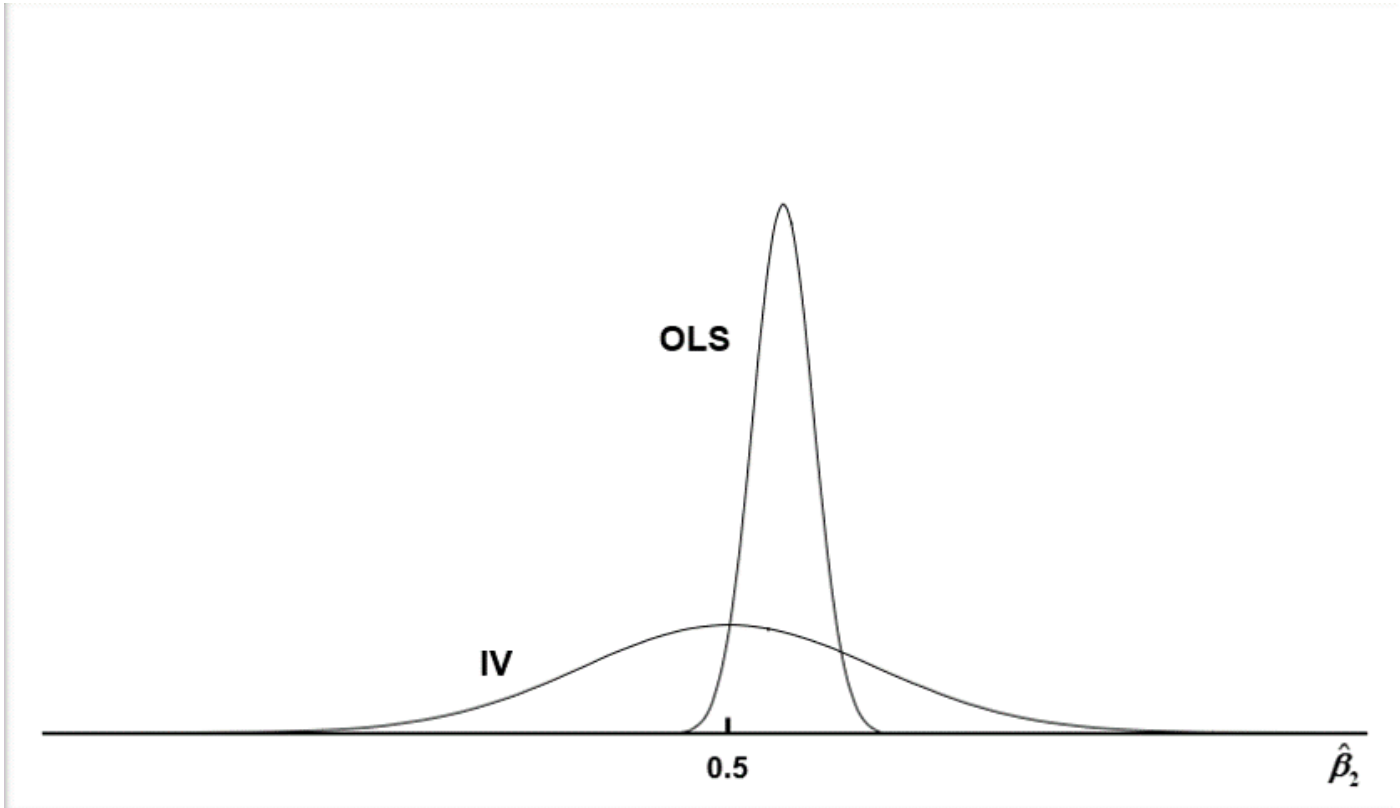
Monte Carlo Evidence

Q: What happens if the instruments are "weak"?



Monte Carlo Evidence

Q: What happens if the instruments are "weak"?



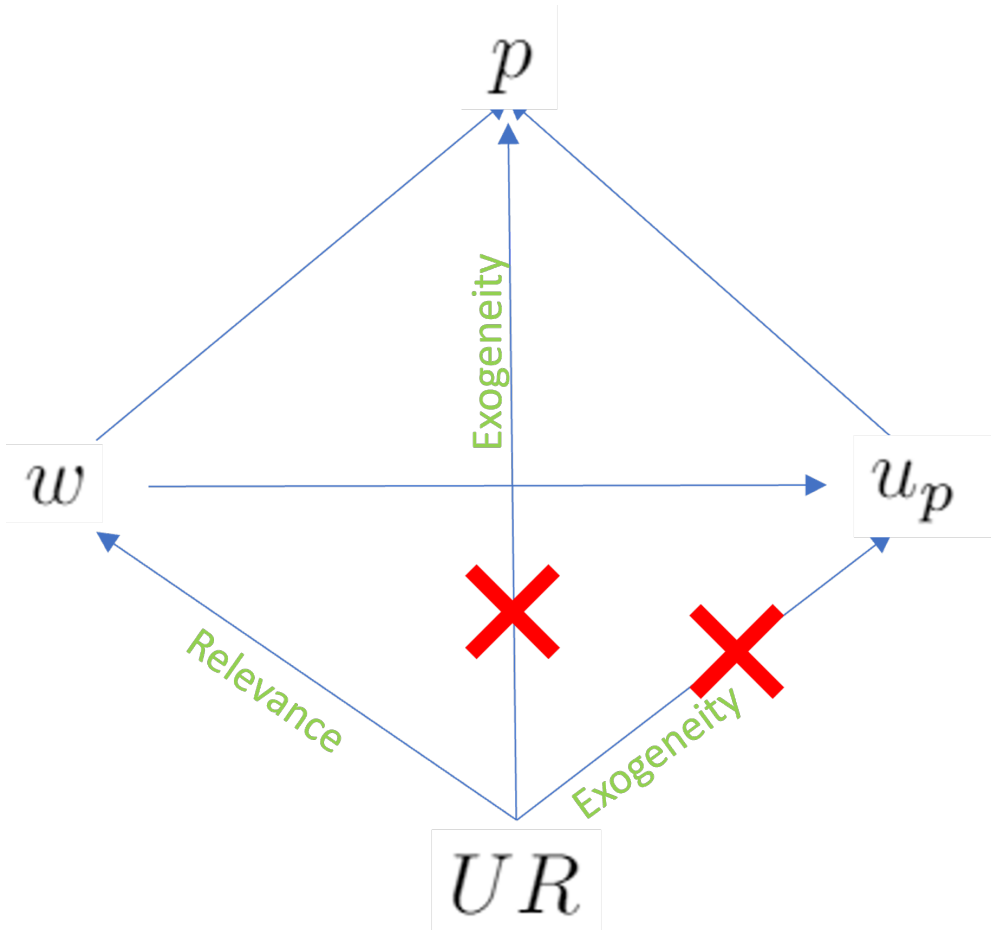
Despite its bias, OLS might be preferred in a Mean Squared Error sense.

Visualizing Weak Instruments

Q: How see weak instruments on the DAG?

Visualizing Weak Instruments

Q: How see weak instruments on the DAG?



A: Via the Relevance condition.

Testing for Weak Instruments

Q: How test for weak instruments when using UR to instrument for w?

$$p = \beta_1 + \beta_2 w + u_p$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w$$

Testing for Weak Instruments

Q: How test for weak instruments when using UR to instrument for w?

$$p = \beta_1 + \beta_2 w + u_p$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w$$

A:

- Auxiliary Regression: $w = \pi_0 + \pi_1 UR + e$

Testing for Weak Instruments

Q: How test for weak instruments when using UR to instrument for w?

$$p = \beta_1 + \beta_2 w + u_p$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w$$

A:

- Auxiliary Regression: $w = \pi_0 + \pi_1 UR + e$
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$

Testing for Weak Instruments

Q: How test for weak instruments when using UR to instrument for w?

$$p = \beta_1 + \beta_2 w + u_p$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w$$

A:

- Auxiliary Regression: $w = \pi_0 + \pi_1 UR + e$
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject then "weak"

Testing for Weak Instruments

Q: How test for weak instruments when using UR to instrument for w?

$$p = \beta_1 + \beta_2 w + u_p$$
$$w = \alpha_1 + \alpha_2 p + \alpha_3 UR + u_w$$

A:

- Auxiliary Regression: $w = \pi_0 + \pi_1 UR + e$
- $H_0 : \pi_1 = 0$ vs $H_1 : \pi_1 \neq 0$
- If fail to reject then "weak"

Note that if we have multiple instruments, then an F-test of overall goodness of fit is useful.
The test should be rerun for each potentially endogenous regressor.

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$
3. Test the coefficient on the residuals $H_0 : \beta_3 = 0$ vs $H_1 : \beta_3 \neq 0$

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$
3. Test the coefficient on the residuals $H_0 : \beta_3 = 0$ vs $H_1 : \beta_3 \neq 0$

Q: What's the consequence if we fail to reject?

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$
3. Test the coefficient on the residuals $H_0 : \beta_3 = 0$ vs $H_1 : \beta_3 \neq 0$

Q: What's the consequence if we fail to reject?

A:

- If $\beta_3 = 0$ then wages are exogenous in the price equation

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$
3. Test the coefficient on the residuals $H_0 : \beta_3 = 0$ vs $H_1 : \beta_3 \neq 0$

Q: What's the consequence if we fail to reject?

A:

- If $\beta_3 = 0$ then wages are exogenous in the price equation
- There is no simultaneity bias

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$
3. Test the coefficient on the residuals $H_0 : \beta_3 = 0$ vs $H_1 : \beta_3 \neq 0$

Q: What's the consequence if we fail to reject?

A:

- If $\beta_3 = 0$ then wages are exogenous in the price equation
- There is no simultaneity bias
- There is no need for IV

Probing Exogeneity

Q: How test if wage is endogenous in the price equation?

A: (Durbin) Wu-Hausman test

1. Run the reduced form model $w = \pi_0 + \pi_1 UR + e$
2. Include the reduced form residuals in the original model $p = \beta_1 + \beta_2 w + \beta_3 \hat{e} + v$
3. Test the coefficient on the residuals $H_0 : \beta_3 = 0$ vs $H_1 : \beta_3 \neq 0$

Q: What's the consequence if we fail to reject?

A:

- If $\beta_3 = 0$ then wages are exogenous in the price equation
- There is no simultaneity bias
- There is no need for IV
- CAUTION: This does NOT suggest that the wage equation can be estimated safely. We need to conduct a separate exploration.

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings
- Treatment effects are difficult to estimate due to endogeneity of regressors and errors

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings
- Treatment effects are difficult to estimate due to endogeneity of regressors and errors
- IV provides a solution

Simultaneous Systems - Summary

- DAG is bi directional: Treatment $\rightarrow$ Outcome $\rightarrow$ Treatment $\rightarrow \cdots$
- Typically seen in observational settings
- Treatment effects are difficult to estimate due to endogeneity of regressors and errors
- IV provides a solution
- As always with IV, our instruments must be carefully chosen