Instrumental Variables

Long Hong

Nov. 19, 2021

IV summary

Standard problem

▶ Interested: what is the effect of X on Y?

▶ The OLS estimate is is biased due to various reasons (e.g., OVB, reverse causality).

▶ IV method is one of the tools that you could correct the bias.

Example: Return to schooling

$$Y_i = \alpha + \beta S_i + \varepsilon_i$$

▶ Y is log(earning), S is years of schooling, β is the return to schooling

▶ $\hat{\beta}_{OLS}$ is biased - why?
  - primarily it's an omitted variable bias problem
  - ability is unobserved in ε, but it affects **both** Y and S

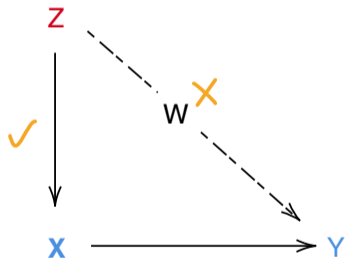We need an IV to get an **unbiased** estimate for β

IV Approach

Assumptions:

1. Relevance
   - ▶ Z must be correlated with X

2. Exclusion Restriction (Exogeneity)
   - ▶ Z can only affect Y through X
     (no other ways)

Example: Card (1993)

Instrumental variable Z for schooling S_i

- Z_i: whether i live **near** a 4-year college, e.g., UW-Madison

Discussion:

A1. Relevance: i is more likely to attend college if college is nearby (e.g., living cost, etc.)
[we can test it in data]

A2. Exogeneity: whether i live near a college can only affect my earnings through education
[we cannot test]

Can you think of a reason why it might not be a good IV?

IV in practtice

A1. Relevance - can be tested
- ▶ Show your first stage regression
- ▶ Show your F-stat > 10 to avoid **weak instrument** issue (Stock and Yogo, 2005)

A2. Exclusion Restriction - **cannot** be tested
- ▶ Argue: Why Z can affect Y only through X?
- ▶ Think: Is there any reason Z **cannot** be a **perfect** IV?
- ▶ Defend: It is still a decent IV (hardest part)

IV Implementation

First stage:

$$S_i = \alpha + \beta Z_i + \gamma \mathbf{x}_{it} + \varepsilon_{it}$$

$\Rightarrow$ get $\hat{S}_i$ (predicted S) and plug it into the second stage.

Second stage:

$$Y_{it} = \alpha + \beta \hat{S}_i + \gamma \mathbf{x}_{it} + \varepsilon_{it}$$

Caveats:
- ▶ Run the two stages together for correct standard errors (detail in Stata)
- ▶ Report first stage, including F-stat (check weak IV)