Microeconometrics Module

Lecture 8: Instrumental Variables

Swapnil Singh

Lietuvos Bankas and KTU

Course Link

Selection on Observables and Unobservables

Understanding the Assumptions:

- Observables:

  $$(Y_{0i}, Y_{1i}) \perp D_i \mid X_i$$

  Assumes that all relevant variables influencing treatment assignment, X_i, are observed.

- Unobservables:
  1. $(Y_{0i}, Y_{1i}) \perp Z_i$
  2. $\text{Cov}(Z_i, D_i) \neq 0$

  Moves beyond observable characteristics, relying on instruments (Z_i) that affect treatment but are independent from the outcomes except through treatment.

The Goals of Causal Inference

Isolating Variation:

- We aim to separate the exogenous (good) variation in D_i from the endogenous (bad) variation that correlates with outcomes Y_{0i} and Y_{1i}.
- **Strategy:**
  - Use observables X_i to control for all observable confounding.
  - Utilize unobservables Z_i to capture the pure effects of D_i on Y_i, filtering out the confounding.

Choosing the Right Approach

Which approach better handles confounding?

1. **Observables:** Assumes perfect knowledge of confounders—challenging and often unrealistic in non-experimental settings.
2. **Unobservables (Instrumental Variables):** More plausible as it uses instruments to isolate exogenous variation, though it may sometimes be underpowered.

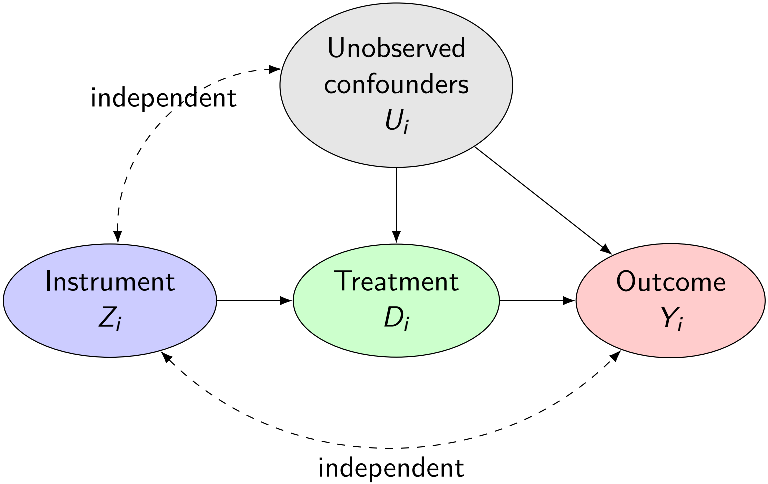
Instrumental Variables: An Introduction

Isolating Good Variation:

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \tag{1}$$

- Instrumental Variables (IV) use tools or instruments (Z_i) to ensure that $\text{Cov}(D_i, \varepsilon_i) = 0$.
- This approach allows us to estimate β_1 consistently, avoiding the bias that occurs in ordinary least squares (OLS) when D_i is correlated with ε_i.

Directed Acyclic Graph: Instrumental Variable Approach



Defining a Valid Instrument

Criteria for Instrument Validity:

1. $\text{Cov}(Z_i, D_i) \neq 0$ — The instrument must influence the treatment.
2. $\text{Cov}(Z_i, \varepsilon_i) = 0$ — The instrument must not be related to the outcome other than through the treatment.

These conditions ensure that Z_i is a suitable instrument for isolating the causal effect of D_i on Y_i.

First condition is easy to check. Second condition – known as **exclusion restriction** – can only be reasoned through. You cannot check it. Why?

Two-Stage Least Squares (2SLS)

Operationalizing IV:

- **First Stage:** Estimate the effect of Z_i on D_i (possibly with other covariates X_i), obtaining $\widehat{D_i}$.
- **Second Stage:** Use $\widehat{D_i}$ to estimate the causal effect on Y_i.

Consistency:

$$\hat{\beta}_{\text{2SLS}} = \left(D' P_Z D\right)^{-1} \left(D' P_Z Y\right)$$

where $P_Z = Z(Z'Z)^{-1}Z'$ is the projection matrix of Z. This estimator is consistent for β_1, assuming valid instruments.

IV Estimation Essentials

Key Considerations in Two-Stage Least Squares (2SLS) and IV Estimation:

- The controls (X_i) must be consistent across both stages to maintain coherence in the model.
- With one instrument and one endogenous variable, 2SLS and IV yield identical results.
- Standard errors from the second stage are typically biased unless corrected for the procedures used in the first stage.

Understanding the Reduced Form

The Reduced Form Equation:

$$Y_i = \pi_1 Z_i + \pi_2 X_i + u_i$$

- This regression links the outcome directly to the instrument and control variables.
- It provides a consistent estimate of the instrument's effect on the outcome, bypassing endogeneity issues.

The Value of the Reduced Form

Advantages of the Reduced Form:

- Clarifies the source of identifying variation in the model.
- Avoids complications associated with weak instruments.
- Simplifies interpretation by directly estimating the effect of the instrument on the outcome.

$$\widehat{\beta}_1^{2SLS} = \frac{\widehat{\pi}_1}{\widehat{\gamma}_1}$$

This ratio illustrates the causal effect of treatment, scaled by the influence of the instrument on treatment.

Intuitive Understanding of Reduced Form

Estimating Impact via 2SLS:

$$\widehat{\beta}_1^{2SLS} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}} = \frac{\widehat{\pi}_1}{\widehat{\gamma}_1}$$

- $\widehat{\pi}_1$: Effect of the instrument on the outcome.
- $\widehat{\gamma}_1$: Effect of the instrument on the treatment.
- This ratio adjusts the direct instrument effect on the outcome to reflect its indirect effect via treatment.

Reduced Form Example: Scholarship Impact on Income

Scenario Analysis:

- Suppose 50% of lottery winners graduate due to the scholarship ($\widehat{\gamma}_1 = 0.50$).
- Reduced form estimate shows a $5,000 income increase for winners ($\widehat{\pi}_1 = \$5,000$).
- Actual effect of graduation on income, when adjusted for the scholarship effect on graduation, is:

$$\frac{\$5,000}{0.50} = \$10,000$$

This calculation adjusts for the proportion of scholarship recipients who graduate, doubling the perceived income benefit.

Deep Dive into IV Mechanics

Further Insights: Explore the nuances of IV estimation, focusing on the subtleties and implications of instrument strength and exogeneity.