Instrumental Variables
2-Stage Least Squares

Research Snippets
American Taxation Association
Mid-Year Meeting

Jennifer Blouin
The Wharton School

Wharton
University of Pennsylvania

Economic Issues

- Omitted Correlated Variables
- Endogeneity
  - Which way is the relation?
  - Debt levels = f(Marginal tax rates)

    OR

    Marginal tax rates = f(Debt levels)
- Errors in Variables
  - Firm Value = f(Economic Earnings)

    But we don’t observe economic earnings but rather accounting earnings, so

    accounting earnings = economic earnings + error

Wharton
UNIVERSITY of PENNSYLVANIA

Endogeneity
What is the basic issue?

- Which way is the relation?

$$y = \beta x + \varepsilon$$

OR

$$x = \phi y + v$$

Endogeneity

- Instrumental Variables (IV) estimation is used when your model has endogenous *x*'s
- That is, whenever Cov(x, ε) ≠ 0
- To tackle this issue, we attempt to identify an instrument

$$y = \beta_0 + \beta_1 x + \varepsilon$$

$$x = \pi_0 + \pi_1 z + v$$

Choosing an instrument

- In order for a variable, z, to serve as a valid instrument for x, the following must be true
  - The instrument must be exogenous
    - That is, $\text{Cov}(z,\varepsilon) = 0$
  - The instrument must be correlated with the endogenous variable x
    - That is, $\text{Cov}(z,x) \neq 0$

Wharton
University of Pennsylvania

Choosing an instrument

- We have to use common sense and economic theory to decide if it makes sense to assume $\text{Cov}(z,\varepsilon) = 0$
- We can test if $\text{Cov}(z,x) \neq 0$
  - Just testing H_0: $\pi_1 = 0$ in $x = \pi_0 + \pi_1 z + v$
  - Often refer to this regression as the first-stage regression

Wharton
UNIVERSITY *of* PENNSYLVANIA

IV estimation

- For $y = \beta_0 + \beta_1 x + \varepsilon$, and given our assumptions
- Cov(z,y) = β_1Cov(z,x) + Cov(z,ε), so
- β_1 = Cov(z,y) / Cov(z,x)
- Then the IV estimator for β_1 is

$$\hat{\beta}_1 = \frac{\sum (z_i - \bar{z})(y_i - \bar{y})}{\sum (z_i - \bar{z})(x_i - \bar{x})}$$

Wharton
University of Pennsylvania

IV versus OLS

- IV is consistent, while OLS is inconsistent, when Cov(x, ε) $\neq$ 0
- The stronger the correlation between z and x, the smaller the IV standard errors

Wharton
UNIVERSITY of PENNSYLVANIA

The potential pitfalls

- What if Cov(z, ε) = 0 is true but that z has weak explanatory power for x?
  - Standard error in IV case differs from OLS only in the R^2 from regressing x on z
  - Since $R^2 < 1$, IV standard errors are larger
    - Example: if x and z have a correlation of only 25%, then the IV std error will be larger than the OLS std error by a factor of 2 (= $\sqrt{1/0.25}$)

Wharton
University of Pennsylvania

The potential pitfalls con't

- What if our assumption that Cov(z, ε) = 0 is false?
  - The IV estimator will be inconsistent, too
  - We Can compare asymptotic bias in OLS and IV
  - Prefer IV if Corr(z, ε)/Corr(z, x) < Corr(x, ε)

$$\text{IV: plim}\hat{\beta}_1 = \beta_1 + \frac{Corr(z,\varepsilon)}{Corr(z,x)} \bullet \frac{\sigma_\varepsilon}{\sigma_x}$$

$$\text{OLS: plim } \tilde{\beta}_1 = \beta_1 + Corr(x,\varepsilon) \bullet \frac{\sigma_\varepsilon}{\sigma_x}$$

Wharton
UNIVERSITY of PENNSYLVANIA

Testing for endogeneity

- Since OLS is preferred to IV if we do not have an endogeneity problem, then we'd like to be able to test for endogeneity
- If we do not have endogeneity, both OLS and IV are consistent
- Idea of Hausman test is to see if the estimates from OLS and IV are different

Testing for endogeneity

- If x is endogenous, then estimate $\hat{x}$ (from the reduced-form equation – 1st stage).
- If β_2 is significant, then you can reject the null of no endogeneity

$$y = \beta_0 + \beta_1 x + \beta_2 \hat{x} + \varepsilon$$

Wharton
University of Pennsylvania

Errors in variables

- Remember the classical errors-in-variables problem is where we observe x (accounting earnings) instead of x^* (economic earnings)
- It leads to coefficient estimates that are biased towards zero
- Where $x = x^* + e$, and e is uncorrelated with x^*
  - If there is a z, such that Corr(z,e) = 0 and Corr(z,x^*) $\neq$ 0, then the IV will eliminate the attenuation bias

Wharton
UNIVERSITY of PENNSYLVANIA

Summary

- The IV approach can be useful in mitigating econometric problems only if the instruments are carefully identified.
- The incorrect instruments create bias that will lead to inappropriate inferences.
- See Larcker and Rusticus (2005)

Wharton
University of Pennsylvania