Introductory Econometrics

Lecture 21: Robustness, Efficiency and Test of Endogeneity

Instructor: Ma, Jun

Renmin University of China

November 25, 2021

Robustness versus efficiency

- The IV estimator is robust to the presence of endogeneity, since the usual OLS estimator is inconsistent under endogeneity.
- Suppose there is no endogeneity. Let us compare in this situation the quality of the OLS estimator ($\hat{\beta}_{1,n}^{OLS}$) and the IV estimator ($\hat{\beta}_{1,n}^{IV}$) of the slope.
- Under exogeneity, both estimators are consistent.
- For simplicity, let us assume that the errors are homoskedasticity:

$$\mathrm{E}\left[U_i^2 | X_i, Z_i\right] = \mathrm{E}\left[U_i^2 | X_i\right] = \mathrm{E}\left[U_i^2 | Z_i\right] = \sigma^2.$$

Comparison of asymptotic variances

- We have seen that under exogeneity and homoskedasticity,

$$\sqrt{n}\left(\hat{\beta}_{1,n}^{OLS}-\beta_1\right)\to_d \mathrm{N}\left(0,\frac{\sigma^2}{\mathrm{Var}\left[X_i\right]}\right)$$

  and

$$\sqrt{n}\left(\hat{\beta}_{1,n}^{IV}-\beta_1\right)\to_d \mathrm{N}\left(0,\frac{\mathrm{Var}\left[Z_i\right]\sigma^2}{\mathrm{Cov}\left(Z_i,X_i\right)^2}\right),$$

  since

$$\frac{\mathrm{E}\left[\left(Z_i-\mathrm{E}Z_i\right)^2U_i^2\right]}{\mathrm{Cov}\left(Z_i,X_i\right)^2}=\frac{\mathrm{Var}\left[Z_i\right]\sigma^2}{\mathrm{Cov}\left(Z_i,X_i\right)^2}$$

  if $\mathrm{E}\left[U_i^2|Z_i\right]=\sigma^2$.
- We show that $V^{IV}=\frac{\mathrm{Var}[Z_i]\sigma^2}{\mathrm{Cov}(Z_i,X_i)^2}$ is greater than or equal to $V^{OLS}=\frac{\sigma^2}{\mathrm{Var}[X_i]}$.

- This follows easily from the Cauchy-Schwarz inequality:

$$\operatorname{Cov}(Z_i, X_i)^2 \leq \operatorname{Var}[X_i] \operatorname{Var}[Z_i] .$$

- When there is no endogeneity, it is better to use the OLS estimator than to use the IV. When there is no endogeneity, both estimators are consistent and the standard errors are correct. However, the standard error of $\hat{\beta}_{1,n}^{OLS}$ tends to be smaller than that of $\hat{\beta}_{1,n}^{IV}$. Hence the quality of $\hat{\beta}_{1,n}^{OLS}$ is better than $\hat{\beta}_{1,n}^{IV}$.
- When there is endogeneity, we should use the IV estimator because the OLS estimator is inconsistent.
- The IV estimator is robust to endogeneity, but loses efficiency (i.e. the asymptotic variance is larger than the OLS estimator) when there is no endogeneity. The OLS estimator is not robust to endogeneity, but more efficient than the IV estimator, when there is no endogeneity.

Hausman-Wu test of endogeneity

- One can make use of the tradeoff between robustness and efficiency to construct a test that tests the exogeneity assumption.
- The null hypothesis and the alternative hypothesis are set to be:

$$H_0 : \mathrm{Cov}\,(U_i, X_i) = 0$$

and

$$H_1 : \mathrm{Cov}\,(U_i, X_i) \neq 0.$$

- Under H_0 exogeneity, both estimators are consistent and should be close to each other when the sample size is large.
- We need to construct a test statistic and define a rejection region. It appears reasonable to base our test statistic upon the difference between $\hat{\beta}_{1,n}^{OLS}$ and $\hat{\beta}_{1,n}^{IV}$, since under H_0 the difference should be small but it should be large under H_1, when the sample size is large.

- One can show that under H_0,

$$\sqrt{n}\left(\hat{\beta}_{1,n}^{IV} - \hat{\beta}_{1,n}^{OLS}\right) \rightarrow_d \mathrm{N}\left(0, V^{IV} - V^{OLS}\right).$$

- Define the following testing statistic:

$$T_n = \frac{\left|\sqrt{n}\left(\hat{\beta}_{1,n}^{IV} - \hat{\beta}_{1,n}^{OLS}\right)\right|}{\sqrt{\hat{V}^{IV} - \hat{V}^{OLS}}},$$

  where $\hat{V}^{IV}$ and $\hat{V}^{OLS}$ are consistent estimators of V^{IV} and V^{OLS} under both H_0 and H_1.

- If the significance level is 5%, we reject H_0 when T_n is greater than 1.96. This test is called the Hausman-Wu test.