Introductory Econometrics

Introduction to Generalized Method of Moments

Instructor: Ma, Jun

Renmin University of China

December 10, 2021

2SLS for overidentified IV models

- We consider the simple model with endogeneity with the intercept known to be zero:

$$\begin{aligned} Y_i &= \alpha D_i + U_i \\ \mathrm{E}\left[U_i\right] &= 0 \\ \mathrm{Cov}\left[D_i, U_i\right] &\neq 0. \end{aligned}$$

- Suppose that we have l IVs $(Z_{i1}, ..., Z_{il})$ which satisfies $\mathrm{E}\left[U_i Z_{ij}\right] = 0$, for $j = 1, 2, ..., l$.
- The first-stage (the reduced-form equation) of 2SLS uses the linear projection of D_i on 1 and Z_i:

$$\begin{aligned} D_i &= \pi_0 + \pi_1 Z_{i1} + \cdots + \pi_l Z_{il} + V_i \\ \mathrm{E}\left[V_i\right] &= 0 \\ \mathrm{E}\left[Z_{ij} V_i\right] &= 0, j = 1, 2, ..., l. \end{aligned}$$

- Then,

$$\begin{aligned} Y_i &= \alpha D_i + U_i \\ D_i &= \pi_0 + \pi_1 Z_{i1} + \cdots + \pi_l Z_{il} + V_i \end{aligned} \implies$$

$$Y_i = \alpha \left(\pi_0 + \pi_1 Z_{i1} + \cdots + \pi_l Z_{il}\right) + \alpha V_i + U_i.$$

  Regression of Y_i on $\pi_0 + \pi_1 Z_{i1} + \cdots + \pi_l Z_{il}$ consistently estimates α.
- 2SLS replaces $(\pi_0, \pi_1, ..., \pi_l)$ with the first-stage OLS estimates.
- We know that the true parameter α satisfies the following equations simultaneously:

$$\begin{aligned} \mathrm{E}\left[U_i\right] = 0 &\implies \mathrm{E}\left[Y_i - \alpha D_i\right] = 0 \\ \mathrm{E}\left[U_i Z_{i1}\right] = 0 &\implies \mathrm{E}\left[\left(Y_i - \alpha D_i\right) Z_{i1}\right] = 0 \\ \vdots \quad & \vdots \quad \vdots \\ \mathrm{E}\left[U_i Z_{il}\right] = 0 &\implies \mathrm{E}\left[\left(Y_i - \alpha D_i\right) Z_{il}\right] = 0. \end{aligned}$$

Generalized method of moments

- By the method of moments principle, an estimator $\widehat{\alpha}$ can be constructed as the solution to the sample moment equations:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} g_i^0\left(\widehat{\alpha}\right) &= 0 \\
\frac{1}{n}\sum_{i=1}^{n} g_i^1\left(\widehat{\alpha}\right) &= 0 \\
\vdots \quad &\vdots \quad \vdots \\
\frac{1}{n}\sum_{i=1}^{n} g_i^l\left(\widehat{\alpha}\right) &= 0
\end{aligned}$$

where

$$\begin{aligned}
g_i^0(a) &= Y_i - aD_i \\
g_i^1(a) &= (Y_i - aD_i)\, Z_{i1} \\
\vdots \quad &\vdots \quad \vdots \\
g_i^l(a) &= (Y_i - aD_i)\, Z_{il}.
\end{aligned}$$

- There may not exist a solution to the sample moment equations.
- Alternatively we solve

$$\min_{a\in\mathbb{R}} \left\| \begin{pmatrix} n^{-1}\sum_{i=1}^{n} g_i^0(a) \\ n^{-1}\sum_{i=1}^{n} g_i^1(a) \\ \vdots \\ n^{-1}\sum_{i=1}^{n} g_i^l(a) \end{pmatrix} \right\|^2,$$

where $\left\|(x_1, ..., x_l)^\top\right\|^2 = \sum_{i=1}^{l} x_i^2$
- Let W_n be a $(l+1)\times(l+1)$ positive definite matrix, which means that $t^\top W_n t > 0$ for all $(l+1)$-dimensional vectors $t \neq 0$. W_n is called a weighting matrix.
- The generalized method of moment (GMM) estimator $\widehat{\alpha}(W_n)$ is the solution to

$$\min_{a\in\mathbb{R}} \begin{pmatrix} n^{-1}\sum_{i=1}^{n} g_i^0(a) \\ n^{-1}\sum_{i=1}^{n} g_i^1(a) \\ \vdots \\ n^{-1}\sum_{i=1}^{n} g_i^l(a) \end{pmatrix}^\top W_n \begin{pmatrix} n^{-1}\sum_{i=1}^{n} g_i^0(a) \\ n^{-1}\sum_{i=1}^{n} g_i^1(a) \\ \vdots \\ n^{-1}\sum_{i=1}^{n} g_i^l(a) \end{pmatrix}.$$

- Here, $\begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix}^\top = \begin{pmatrix} x_1 & \cdots & x_l \end{pmatrix}$ denotes the vector transpose.
- Suppose that $W_n \to_p W$ for some nonrandom positive definite matrix W.
- We can show that

$$\sqrt{n}\left(\widehat{\alpha}\left(W_n\right)-\alpha\right) \to_d \mathrm{N}\left(0, \sigma^2\left(W\right)\right)$$

and $\sigma^2(W) \geq \sigma^2(W^*)$, where

$$W^* = \left(\mathrm{E}\left[\begin{pmatrix} g_i^0(\alpha) \\ g_i^1(\alpha) \\ \vdots \\ g_i^l(\alpha) \end{pmatrix} \begin{pmatrix} g_i^0(\alpha) \\ g_i^1(\alpha) \\ \vdots \\ g_i^l(\alpha) \end{pmatrix}^\top \right] \right)^{-1}.$$

Efficient GMM

- The efficient GMM uses the weighting matrix $\widehat{W}^* \to_p W^*$, where

$$\widehat{W}^* = \left(\frac{1}{n} \sum_{i=1}^{n} \begin{pmatrix} g_i^0(\widetilde{\alpha}) \\ g_i^1(\widetilde{\alpha}) \\ \vdots \\ g_i^l(\widetilde{\alpha}) \end{pmatrix} \begin{pmatrix} g_i^0(\widetilde{\alpha}) \\ g_i^1(\widetilde{\alpha}) \\ \vdots \\ g_i^l(\widetilde{\alpha}) \end{pmatrix}^{\top} \right)^{-1}$$

  and $\widetilde{\alpha}$ is a preliminary consistent estimator, which can be the 2SLS.
- The efficient GMM estimator $\widehat{\alpha}\left(\widehat{W}^*\right)$ has the lowest asymptotic variance.
- The 2SLS is an GMM estimator which uses

$$W_n^{2SLS} = \left(\frac{1}{n} \sum_{i=1}^{n} \begin{pmatrix} 1 \\ Z_{i1} \\ \vdots \\ Z_{il} \end{pmatrix} \begin{pmatrix} 1 \\ Z_{i1} \\ \vdots \\ Z_{il} \end{pmatrix}^{\top} \right)^{-1}.$$