Endogeneity, LIML, testing

Jerzy Mycielski

Microeconometrics,November 17, 2025

Limited Information Maximum Likelihood (LIML) estimator

- Alternative estimator which can be used to estimate regression equation if endegeneity problem is present is LIML
- This estimator is a special case of Maximum Likelihood (ML) estimator which is based on maximisation of the (log)likelihood function
- Model which is estimated is the

$$Y_1 = \mathbf{Z}_1\beta_1 + \mathbf{Y}_2\beta_2 + e$$

$$\mathbf{Y}_2 = \mathbf{Z}_1\mathbf{\Gamma}_{12} + \mathbf{Z}_2\grave{}_{22} + \mathbf{u}_2$$

  and we assume that $(e, \mathbf{u}_2)$ has multivariate normal distribution
- Substituting the second equation to the first we obtain reduced form of the first equation:

$$Y_1 = \mathbf{Z}_1\beta_1 + \left(\mathbf{Z}_1\mathbf{\Gamma}_{12} + \mathbf{Z}_2\grave{}_{22} + \mathbf{u}_2\right)\beta_2 + e$$

$$= \mathbf{Z}_1\left(\beta_1 + \mathbf{Z}_1\mathbf{\Gamma}_{12}\right) + \mathbf{Z}_2\grave{}_{22}\beta_2 + \left(\mathbf{u}_2\beta_2 + e\right)$$

LIML

- Denoting $\boldsymbol{\lambda}_1 = \boldsymbol{\beta}_1 + \mathbf{Z}_1\boldsymbol{\Gamma}_{12}$, $\boldsymbol{\lambda}_2 = \grave{}_{22}\boldsymbol{\beta}_2$ and by $u_1 = \mathbf{u}_2\boldsymbol{\beta}_2 + e$ we obtain reduced form of the model

$$Y_1 = \mathbf{Z}_1\boldsymbol{\lambda}_1 + \mathbf{Z}_2\boldsymbol{\lambda}_2 + u_1$$

$$\mathbf{Y}_2 = \mathbf{Z}_1\boldsymbol{\Gamma}_{12} + \mathbf{Z}_2\grave{}_{22} + \mathbf{u}_2$$

- Notice that $(Y_1, \mathbf{Y}_2) \sim N\left((\mathbf{Z}_1\boldsymbol{\lambda}_1 + \mathbf{Z}_2\boldsymbol{\lambda}_2, \mathbf{Z}_1\boldsymbol{\Gamma}_{12} + \mathbf{Z}_2\grave{}_{22}), \boldsymbol{\Sigma}\right)$ with parameters $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2$ restricted satisfy the equations:

$$\boldsymbol{\lambda}_1 = \boldsymbol{\beta}_1 + \mathbf{Z}_1\boldsymbol{\Gamma}_{12}$$

$$\boldsymbol{\lambda}_2 = \grave{}_{22}\boldsymbol{\beta}_2$$

- Notice that for overidentified model $dim(\boldsymbol{\beta}_2) < dim(\boldsymbol{\lambda}_2)$ so that this system of equations indeed imposes restrictions on $\boldsymbol{\lambda}_2$
- The above system of equations can estimated with restricted ML estimator

Endogeneity testing - Hausman-Wu test

- Tested hypothesis:
  - H_0 all explanatory variables are exogenous $E(\mathbf{x}_i u_i) = 0$
  - H_1 same explanatory variables are endogenous $E(\mathbf{x}_i u_i) \neq 0$
- In order to test this hypothesis we use Hausman-Wu test
- It is based on observation that IV and OLS estimators are both consistent if H_0 is valid, but only OLS estimator is efficient.
- However, if H_0 is invalid than under H_1 only IV estimator is consistent:

	$\widehat{\beta}_{OLS}$	$\widehat{\beta}_{IV}$	$\widehat{\beta}_{OLS} - \widehat{\beta}_{IV}$
H_0 : *true*	$\widehat{\beta}_{OLS} \xrightarrow{p} \beta$	$\widehat{\beta}_{IV} \xrightarrow{p} \beta$	$\widehat{\beta}_{OLS} - \widehat{\beta}_{IV} \xrightarrow{p} 0$
H_0 : *false*	$\widehat{\beta}_{OLS} \xrightarrow{p} \beta^* \neq \beta$	$\widehat{\beta}_{IV} \xrightarrow{p} \beta$	$\widehat{\beta}_{OLS} - \widehat{\beta}_{IV} \xrightarrow{p} \beta^* - \beta \neq 0$

- Therefore, to verify exogeneity we test for significance of the difference $\widehat{\beta}_{OLS} - \widehat{\beta}_{IV}$
- Test statistic

$$\left(\widehat{\beta}_{OLS} - \widehat{\beta}_{IV}\right) \widehat{\Sigma}^{-1}_{\widehat{\beta}_{OLS} - \widehat{\beta}_{IV}} \left(\widehat{\beta}_{OLS} - \widehat{\beta}_{IV}\right) \xrightarrow{d} \chi^2_K$$

Overidentification test/validity of instruments test

- Crucial assumption of the IV estimation

$$E\left(\mathbf{z}_i u_i\right) = 0$$

- This assumption can be tested but only if the model is overidentified
- Model
  - is exactly identified if the number of instruments is equal to the number of endogenous explanatory variables $L_Z = G$
  - is overidentified if the number of instruments is larger the number of endogenous explanatory variables $L_Z > G$
- Extra restrictions implied by overidentification can be tested using Sargan test:

$$\frac{\widehat{\mathbf{e}}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\widehat{\mathbf{e}}}{\sigma^2} \xrightarrow{d} \chi^2_{L_z - G}$$

- if $H_0 : E\left(\mathbf{z}_i u_i\right) = 0$ is rejected this suggests that some instruments are invalid.

Testing for weak instruments (Stock and Yogo 2005)

- If instruments are weak (small correlation with endogenous variable) than small sample distributions of test statistics can be heavily distorted
- Consider simple model with one endogenous variable and one instrument

$$\begin{aligned} Y &= X\beta + e \\ X &= Z\Gamma + u_2 \end{aligned}, \; (e, u_2) \sim N\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- It was shown that the t-statistic have the asymptotic distribution for hypothesis $H_0 : \beta = 0$ is

$$T \xrightarrow{d} \frac{\xi_1}{\sqrt{1 - 2\rho\frac{\xi_1}{\mu+\xi_2} + \left(\frac{\xi_1}{\mu+\xi_2}\right)^2}} \stackrel{def}{=} S$$

  where $\Gamma = n^{-1/2}\mu$ and (ξ_1, ξ_2) is bivariate normal. This distribution is obviously not gaussian except for $\mu \to \infty$.
- Notice that here small μ is small in proportion to $n^{-1/2}$.

Testing for weak instruments

- We choose the value of ρ for which the distortion is the most serious ($\rho = 1$)
- For this case formula simplifies to

$$S = \xi \left| 1 + \frac{\xi}{\mu} \right|, \ \xi \sim N(0, 1)$$

- We choose such $\mu^2 = \tau^2$ that

$$Pr(|S| \geq t_\alpha) \leq \alpha^*$$

where α is a nominal significance level and α^* is the actual significance level when testing the $H_0 : \beta = 0$.

- If indeed the H_0 is valid then even in the worst case the actual size of the test does not exceed α^*.

Testing for weak instruments

- The idea of Stock and Yogo (2005) is to test $H_0 : \mu^2 = \tau^2$ against $H_1 : \mu^2 > \tau^2$ using F-test statistic

$$Pr\left(F > c|\,\mu^2 = \tau^2\right) \xrightarrow{d} \chi^2\left(\tau^2\right) = \alpha$$

  where $\chi^2\left(\tau^2\right)$ is non-central χ^2 distribution with non-centrality parameter τ^2
- If we are not rejecting $H_0 : \mu^2 = \tau^2$ then maximum distortion of the test for significance made on nominal α is equal to $\alpha^* - \alpha$

Critical values of Stock and Yogo (1943) test

Table 12.4: 5% Critical Value for Weak Instruments, $k = 1$

	Maximal Size r							
	2SLS				LIML			
$\grave{}_2$	0.10	0.15	0.20	0.25	0.10	0.15	0.20	0.25
1	16.4	9.0	6.7	5.5	16.4	9.0	6.7	5.5
2	19.9	11.6	8.7	7.2	8.7	5.3	4.4	3.9
3	22.3	12.8	9.5	7.8	6.5	4.4	3.7	3.3
4	24.6	14.0	10.3	8.3	5.4	3.9	3.3	3.0
5	26.9	15.1	11.0	8.8	4.8	3.6	3.0	2.8
6	29.2	16.2	11.7	9.4	4.4	3.3	2.9	2.6
7	31.5	17.4	12.5	9.9	4.2	3.2	2.7	2.5
8	33.8	18.5	13.2	10.5	4.0	3.0	2.6	2.4
9	36.2	19.7	14.0	11.1	3.8	2.9	2.5	2.3
10	38.5	20.9	14.8	11.6	3.7	2.8	2.5	2.2
15	50.4	26.8	18.7	12.2	3.3	2.5	2.2	2.0
20	62.3	32.8	22.7	17.6	3.2	2.3	2.1	1.9
25	74.2	38.8	26.7	20.6	3.8	2.2	2.0	1.8
30	86.2	44.8	30.7	23.6	3.9	2.2	1.9	1.7

Source: Hansen (2022)