Econometrics A – Seminar 8
Silmultaneous equations models

3SLS
Over-identification test
Hausman‘s specification test

Lukáš Rečka
(with thanks to Ladislav Krištoufek)

Institute of Economic Studies, Charles University, Prague
lukasrecka@gmail.com

3.12.2012

2 Good news

- Last seminar will be in the 1st week of January
- You will have change to choose the day

What have we learnt so far?

- Order & Rank condion
- OLS vs. 2SLS
- OLS biased and inconsistent for simultaneous eq.
- 2SLS used for single equation estimation

- Whole system? → Three stage least squares (3SLS)
  - Baltagi (2008, pp. 267-273/286)

- Bonus question – what is "$\otimes$" and it properties?

Properties of Kronecker product

- Inverse: $(\Sigma \bigotimes I_n)^{-1} = \Sigma^{-1} \bigotimes I_n$
- Determinant: $|\Sigma \bigotimes I_n| = |\Sigma|^m |I_n|^n = |\Sigma|^m$
- Trace: $tr(\Sigma \bigotimes I_n) = tr(\Sigma) tr(I_n) = n * tr(\Sigma)$

3SLS – derivation (1)

- Consider the following system:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{pmatrix} = \begin{pmatrix} Z_1 & 0 & \dots & 0 \\ 0 & Z_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & Z_G \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_G \end{pmatrix}$$

- Or simply $y = Z\delta + u$
- u has a zero mean and variance-covariance matrix $\Sigma \bigotimes I_T \rightarrow$ possible correlations between disturbances
- We use GLS on the whole system and get

$$\hat{\delta}_{GLS} = (Z^T(I_G \bigotimes X)[\Sigma^{-1} \bigotimes (X^T X)^{-1}](I_G \bigotimes X^T)Z)^{-1}$$

$$(Z^T(I_G \bigotimes X)[\Sigma^{-1} \bigotimes (X^T X)^{-1}](I_G \bigotimes X^T)y)$$

3SLS – derivation (2)

- If simplified, we get

$$\hat{\delta}_{GLS} = (Z^T[\Sigma^{-1} \bigotimes P_X]Z)^{-1}(Z^T[\Sigma^{-1} \bigotimes P_X]y)$$

- However, we need an estimate of Σ
- For i-th equation, we get residuals by running 2SLS so that $\hat{u}_i = y_i - Z_i\hat{\delta}_{i,2SLS}$ and estimate Σ by $\hat{\Sigma} = |\hat{\sigma}_{ij}|$ where

$$\hat{\sigma}_{ij} = \left| \frac{\hat{u}_i^T \hat{u}_j}{\sqrt{T - g_i - k_i}\sqrt{T - g_j - k_j}} \right|$$

  for $i, j = 1, 2, \ldots, G$
- 3SLS estimator then turns to

$$\hat{\delta}_{3SLS} = (Z^T[\hat{\Sigma}^{-1} \bigotimes P_X]Z)^{-1}(Z^T[\hat{\Sigma}^{-1} \bigotimes P_X]y)$$

2SLS vs 3SLS

- If the system is properly identified, 3SLS is more efficient than 2SLS
- 2SLS and 3SLS are equivalent if
  - Σ is diagonal
  - Each equation is just identified
  - If second stage estimators of i-th equation are perfect linear combination of the ones of j-th eqaution

Test for over-identification

- We test the following hypotheses:
  - H_0: $y_1 = Z_1\delta_1 + u_1$
  - H_1: $y_1 = Z_1\delta_1 + W^*\gamma + u_1$
- W is a matrix of instruments of full rank l, W^* is a subset of W
- If W^* is correlated with u_1 or the first structural equation is misspecified (by not containing some variables of W^*), then $\gamma \neq 0$
- Thus we test $\gamma = 0$ if the system is properly identified

Hausman's specification test

- We test two hypotheses:
  - H_0: $E(u|X) = 0$
  - H_1: $E(u|X) \neq 0$
- We need two estimators for the test:
  - One must be consistent and efficient estimator of β under H_0 which becomes inconsistent under H_1, say $\hat{\beta}_0$
  - One must be consistent estimator of β under both H_0 and H_1 but inefficient under H_0, say $\hat{\beta}_1$
- We define difference $\hat{q} = \hat{\beta}_1 - \hat{\beta}_0$ and $m = \hat{q}^T var(\hat{q})^{-1} \hat{q}$ which converges to χ^2_k where k is a dimension of β

Empirical example

- Let's consider the following model (Laffer, 1970):

$$\log(TM/P) = \alpha_0 + \alpha_1 \log(RM/P) + \alpha_2 \log i + u_1$$

$$log(TM/P) = \beta_0 + \beta_1 \log(Y/P) + \beta_2 \log i + \beta_3 \log S_1 + \beta_4 \log S_2 + u_2$$

  - TM – nominal total trade money
  - RM – nominal effective reserve money
  - Y – GNP in current USD
  - S_1 – mean real size of the representative economic unit ($1939 = 100$)
  - S_2 – degree of market utilization
  - i – short-term rate of interest
  - P – GNP price deflator ($1958 = 100$)
- Is the system identified?
- Let's estimate OLS, 2SLS and 3SLS

Results

- OLS:

$$\log(TM/P) = 3.66 + 0.55\log(RM/P) + 0.17\log i + u_1$$

$$log(TM/P) = 3.06 + 0.62\log(Y/P) - 0.02\log i - 0.31\log S_1 + 0.15\log S_2 + u_2$$

- 2SLS:

$$\log(TM/P) = 3.74 + 0.52\log(RM/P) + 0.18\log i + u_1$$

$$log(TM/P) = 1.56 + 0.76\log(Y/P) - 0.08\log i - 0.19\log S_1 + 0.23\log S_2 + u_2$$

- 3SLS:

$$\log(TM/P) = 3.74 + 0.52\log(RM/P) + 0.18\log i + u_1$$

$$log(TM/P) = 1.89 + 0.76\log(Y/P) - 0.08\log i - 0.23\log S_1 + 0.13\log S_2 + u_2$$

- Hausman's test:
  - OLS vs 2SLS: reject the null for the first equation, not for the second $\rightarrow$ 2SLS estimate might be significantly different
  - 2SLS vs 3SLS: cannot reject the null but 3SLS still more efficient

Thank you for your attention!

Next seminar: Panel data