E 4160: Econometrics–Modelling and Systems Estimation
Computer Class # 5

Ragnar Nymoen

Department of Economics, University of Oslo

31 October 2011

- KeynesID1.zip
  Unzip everything in a single directory

Model from Seminar 4, with numbers for coefficients for concreteness

$$C_t = 0.75 Y_t + \varepsilon_{ct}$$
$$I_t = 1.5 Y_t - 0.5 Y_{t-1} + \varepsilon_{It}$$
$$Y_t = C_t + I_t + X_t$$

Start with "identification table" based on all 3 structural equations:

	C	I	Y	Y_{-1}	X_1
C-equation	1	0	0.75	0	0
I- equation	0	1	-1.5	-0.5	0
Identitiy	-1	-1	1	0	1

- C-eq: # of excluded exogenous is 2 while # of included endogenous minus one is 1 $\Longrightarrow$ Overidentification.
- The alternative form of the order condition: exclude $3-1=2$ of the variable in the structural model, also gives overidentification, since # of excluded variables (endo and ex) is 3.
- But order condition is only necessary.
- The sufficient rank condition says that C-eq is identified *if and only if at least one non-zero* $(3-1) \times (3-1)$ *determinant is contained in the array of coefficients with which those variables excluded from the equation appear in the other equations.*

To check this, write the table more compactly as a matrix:

$$\begin{pmatrix} 1 & 0 & 0.75 & 0 & 0 \\ 0 & 1 & -1.5 & 0.5 & 0 \\ -1 & -1 & 1 & 0 & 1 \end{pmatrix}$$

then delete first row and the columns which have non-zeros in the first row:

$$\begin{pmatrix} 1 & 0.5 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

- The rank of this matrix is 2 meaning that the order of the largest non-zero determinant is 2.

- This means that the rank condition for identification is fulfilled. For example

$$\begin{vmatrix} 1 & 0.5 \\ -1 & 0 \end{vmatrix} = 0.5 \neq 0$$

and also

$$\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

$$\begin{pmatrix} 1 & 0 & 0.75 & 0 & 0 \\ 0 & 1 & -1.5 & 0.5 & 0 \\ -1 & -1 & 1 & 0 & 1 \end{pmatrix}$$

and delete second row and columns with non-zeros in the second row:

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

We can only construct one non-zero $(3-1)\times(3-1)$ determinant in this case, but since

$$\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

the rank condition is fulfilled

System where we have discriminate between exogenous expenditure components:

$$C_t = 15 + 0.75 Y_t + \varepsilon_{ct}$$
$$I_t = -0.2 + 1.5 Y_t - 0.5 Y_{t-1} + \varepsilon_{It}$$
$$Y_t = C_t + I_t + X_{1t} + X_{2t} - X_{3t}$$

	C	I	Y	Y_{-1}	X_1	X_2	X_3
C-equation	1	0	0.75	0	0	0	0
I-equation	0	1	−1.5	−0.5	0	0	0
	−1	−1	1	0	1	1	−1

C-equation

$$\begin{pmatrix} 1 & 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1.5 & -0.5 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 & 1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -0.5 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & -1 \end{pmatrix}$$

rank $= 2$. Can now form

$$\frac{5!}{2! \cdot 3!} = \frac{5 \cdot 4}{1 \cdot 2} = 10$$

2×2 determinants. Have the same two (0.5 and 1) as before, but also -1 and so on.

I-equation

$$\begin{pmatrix} 1 & 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1.5 & -0.5 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 & 1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{pmatrix}$$

Rank= 2. Can form 6 2×2 determinants. One of these are 1 but now also -1 for example

What do we take away form this? I

- The *Order condition for identification* is a necessary condition for identification.
- Any conclusion about identification on the basis of the Order condition is therefore provisional, until the Rank condition has been checked (in practice that is often a small step though).
- The Rank condition is sufficient for identification
- Note that main distinction is between non-identified and identified. And that Rank condition is sufficient for identification.

What do we take away form this? II

- The amount of (or degree of overidentification) is another issue: An over-identified equation is identified so there are no principal difference between identification and over-identification.
- What we can gain from overidentification is twofold:
  - We have more than one consistent estimator of the structural parameters—these can be weighted together in an efficient way: Achieved by 2SLS in the case of independent structural disturbances.
  - The validity of the overidentifying restrictions can be tested.

We will look at these two gains in PcGive. We will use the Seminar 4 macro model as our example.

- In the case of over-identification we can test the assumption of independence between the instruments and the disturbances of the structural equation we estimate.
- In the IVE output from PcGive this test is found as the

Specification test

It is calculated as

$$\text{number of observations} * R^2_{IVres}$$

where R^2_{IVres} is the R^2 from the regression between the residuals from the IV estimated structural equation and the variables that used as instruments (an auxiliary regression)

- Under the H_0 of correct specification, the test statistic is Chi-squared with degrees of freedom equal to the number of overidentifying instruments.
- This test is one of the most cited in single equation IV estimation, see for example Stata output where it is referred to as *Hansen J statstic [Overidentifying test of all instruments]*
- It is basically the same test as in Give (Bårdsen and Nymoen (2011), p 231 contains a clarification), which is attributed to an earlier contribution by Denis Sargan.
- We can use the consumption function of Seminar 4 model as an example.
- Since the test is based on residuals, we take to include intercept both in the structural equation and in the auxiliary regression).

We then get

$$100 * 0.00125149 = 0.125149$$

from the auxiliary regression and

$$\text{Specification test } \chi^2(1) = 0.12515[0.7235]$$

from the IV estimation

- When we estimated systems of regression equations by OLS we saw that the joint significance of the restrictions could be testes by comparing the log likelihood of the system (which often is an unrestricted VAR) with the likelihood of the restricted system (restricted reduced form).
- We can perform a similar type of test of the restrictions placed on the reduced form by the overidentifying restrictions of a simultaneous equation model.

Calculate manually the LR test of over-identifying restrictions

$$-2 * (L_{RRF} - L_{URF}) = -2 * (-267.422604 + 267.391052) = 0.063104$$

and PcGive gives

$$LR\ test\ of\ over-identifying\ restrictions : Chi\hat{}2(1) = 0.063104[0.8017]$$

when we estimate the simultaneous equations model with 2SLS.

- What happens for the system and model when X is split into $X1$, $X2$, and $X3$?