ECON 4160: Econometrics–Modelling and Systems Estimation:

Computer Class # 4

André Anundsen

Department of Economics, University of Oslo

October 7, 2013

Outline

Data sets for today:

- *forlengetNORKPI_agg0* in *pcmbynls.zip*
- *KonsData2Nor.in7* in *KonsData2Nor_old*

Batch files for today:

- GNC.fl
- Super_exo.fl

The last exercise from Seminar # 2 I

At the previous seminar, we looked at two econometric Phillips Curve models:

Static:

$$(\pi_t - \pi^*) = \beta_0 + \beta_1 u_t + \varepsilon_t, \;\; \varepsilon_t \sim N(\mu, \sigma^2) \qquad (1)$$

where we considered the cases with ($\pi^* = 0$ and $\pi^* = 2.5$). We also showed that the unemployment rate consistent with $\pi = \pi^*$ was given by $u^* = -\frac{\beta_0}{\beta_1}$. This model had problems with autocorrelation, so we also considered an alternative dynamic model.

ARDL(1,1):

$$(\pi_t - \pi^*) = \beta_0 + \rho(\pi_{t-1} - \pi^*) + \beta_1 u_t + \beta_2 u_{t-1} + \epsilon_t \;\; \epsilon_t \sim N(\mu, \sigma^2) \qquad (2)$$

The unemployment rate consistent with inflation being on the target in this model is $u^* = -\frac{\beta_0}{\beta_1 + \beta_2}$. This model solved the problems with autocorrelation

The last exercise from Seminar # 2 II

The relevant model might be somewhere in between. Consider the case where:

$$(\pi_t - \pi^*) = \beta_0 + \beta_1 u_t + \varepsilon_t \quad (3)$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t, \quad v_t \sim N(\mu, \sigma^2) \quad (4)$$

so that the relevant model is a static model with disturbances following an AR(1) process! Note that (3) implies:

$$\varepsilon_t = (\pi_t - \pi^*) - \beta_0 - \beta_1 u_t$$

But this also means that:

$$\varepsilon_{t-1} = (\pi_{t-1} - \pi^*) - \beta_0 - \beta_1 u_{t-1}$$

Substitute this into (4), and we get:

$$\varepsilon_t = \rho\left((\pi_{t-1} - \pi^*) - \beta_0 - \beta_1 u_{t-1}\right) + v_t$$

The last exercise from Seminar # 2 III

Then, substitute the last expression into (3):

$$(\pi_t - \pi^*) = \beta_0 + \beta_1 u_t + \rho\left((\pi_{t-1} - \pi^*) - \beta_0 - \beta_1 u_{t-1}\right) + v_t$$
$$\Rightarrow$$
$$(\pi_t - \pi^*) = \beta_0(1-\rho) + \beta_1\left(u_t - \rho u_{t-1}\right) + \rho(\pi_{t-1} - \pi^*) + v_t$$

This model may also be written as (using lag operators):

$$(1-\rho L)(\pi_t - \pi^*) = \beta_0(1-\rho) + \beta_1(1-\rho L)u_t + v_t$$

This is called a common factor (COMFAC) model! So, a static model with AR(1) disturbances is in a sense an ARDL(1,1) with an implied COMFAC restriction!

The last exercise from Seminar # 2 IV

Recap the unrestricted ARDL(1,1) model:

$$(\pi_t - \pi^*) = \beta_0 + \rho(\pi_{t-1} - \pi^*) + \beta_1 u_t + \beta_2 u_{t-1} + \epsilon_t$$

This model nests both the static model ($\rho = \beta_2 = 0$):

$$(\pi_t - \pi^*) = \beta_0 + \beta_1 u_t + \epsilon_t$$

and the COMFAC model ($\beta_2 = -\rho\beta_1$):

$$(\pi_t - \pi^*) = \beta_0 + \rho(\pi_{t-1} - \pi^*) + \beta_1 u_t - \beta_1 \rho u_{t-1} + \epsilon_t$$
$$\Rightarrow$$
$$(1 - \rho L)(\pi_t - \pi^*) = \beta_0 + \beta_1(1 - \rho L)u_t + \epsilon_t$$

Our aim is then to test the validity of the COMFAC restriction! The test is asymptotically χ^2. How many degrees of freedom?

- Open the data set *forlengetNORKPI_agg0.xls*, which is contained in *pcmbynls.zip*, and let's test the validity of imposing the COMFAC restriction on our ARDL(1,1) model

Four concepts of exogeneity

1. Weak exogeneity
2. Strong exogeneity
3. Super exogeneity
4. Strict exogeneity and pre-determinedness

VAR

At CC #2, we considered a bi-variate VAR model of first order:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \tag{5}$$

A more compact notation gives:

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{6}$$

Assume that $\boldsymbol{\varepsilon}_t \sim MVN(\mathbf{0}_{2\times 1}, \boldsymbol{\Sigma})$, i.e.:

$$\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim MVN\left(\mathbf{0}_{2\times 1}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}\right) \tag{7}$$

We have also shown that the bi-variate VAR(1) represented by (5) can expressed in **model form**:

$$y_{1,t} = \phi_0 + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \phi_1 y_{1,t-1} + \epsilon_t \quad (8)$$

$$y_{2,t} = \mu_2 + a_{21} y_{1,t-1} + a_{22} y_{2,t-1} + \varepsilon_{2,t} \quad (9)$$

$$\begin{aligned} E(\epsilon_t) &= 0 \;\forall\; t \\ Var(\epsilon_t) &= \sigma_1^2 (1 - \rho_{12}) \;\forall\; t \\ E(y_{2,t}\epsilon_t) &= 0 \;\forall\; t \\ E(\varepsilon_{2,t}\epsilon_t) &= 0 \;\forall\; t \end{aligned}$$

Weak exogeneity

If there are no efficiency gains from considering the system ((8) and (9)) rather than only considering (8), we say that $y_{2,t}$ is weakly exogenous (WE) with respect to the parameters of interest, *i.e.* we don't "loose" anything by abstracting from the marginal model for $y_{2,t}$!

Granger non-causality (GNC)

- We see that in one sense (8) and (9) define a recursive model, since given the history (represented by $y_{1,t-1}$ and $y_{2,t-1}$), $y_{2,t}$ is determined first, and given this, $y_{1,t}$ is determined. Moreover, $Cov(\varepsilon_{2,t}, \epsilon_t) = 0$
- However, the two variables are clearly jointly determined over time, since in general $y_{2,t-1} \longrightarrow y_{1,t}$ and $y_{1,t-1} \longrightarrow y_{2,t}$. In econometrics we call this joint Granger causality. Only if $a_{21} = 0$ can we say that we have a recursive causal chain
- With $a_{21} = 0$ imposed, we say that $y_{1,t-1}$ is not Granger-causing $y_{2,t}$ while $y_{2,t-1}$ is Granger-causing $y_{1,t}$.

Testing for GNC between consumption and income

- Open the data set *KonsData2Nor.in7* contained in *KonsData2Nor_old*

We will use these data to test whether income is Granger non-causal for consumption, and whether consumption is Granger non-causal for income using the sample 1968q2–2004q4! We will also play with the batch file *GNC.fl*

Strong exogeneity and super exogeneity

Strong exogeneity:

If $y_{2,t}$ is WE wrt. the parameters of interest and $y_{1,t}$ is not Granger causing $y_{2,t}$, we say that the variable $y_{2,t}$ is strongly exogenous (SE)! (How to test?)

Super exogeneity:

If the coefficient on $y_{2,t}$ in the conditional model for $y_{1,t}$ (β_0 in equation (8)) is invariant to structural breaks in the marginal equation for $y_{2,t}$, then we say that $y_{2,t}$ is super exogenous (SpE) in the conditional model for $y_{1,t}$!

Invariance, super exogeneity and autonomy

- Super-exogeneity is the property that we have constant parameters in the conditional model for $y_{1,t}$ even in periods where there is a structural break in the marginal equation for $y_{2,t}$
- Super exogeneity is defined for conditional models, but the concept is related to the more general idea of autonomy
- Econometric models with parameters that are invariant in the face of wide range of structural breaks have a high degree of autonomy

Testing invariance and super exogeneity and autonomy

To test a hypothesis of lack of invariance, we need to investigate two issues:

1. Test the null hypothesis of no-structural breaks in the marginal model
2. Test the null hypothesis of stability in the conditional model

We have several tools available:

- Recursive estimation and recursive graphs
- Recursive Chow tests
- Find significant dummy variables that represent structural breaks in a marginal model, and test the significance of those dummies in the conditional model (here we can e.g. use Autometrics in combination with our institutional knowledge)

Testing for super exogeneity of wealth in consmption equation

- Again, open the data set *KonsData2Nor.in7* contained in *KonsData2Nor_old*
- Also open the batch file called *Super_exo.fl*

We shall now test whether the wealth variable is super exogenous in the conditional consumption function using the sample 1968q2–2004q4

Strict exogeneity and pre-determinedness

Strict exogeneity:

With reference to (8), if $E(\epsilon_t y_{2,t}) = 0$, $y_{2,t}$ is a strictly exogenous variable

Pre-determinedness:

Again, if in equation (8), we have that $E(\epsilon_{t+j} y_{1,t-1}) = 0 \; \forall \; j \geq 0$, we say that $y_{1,t-1}$ is a pre-determined variable

The Order Condition for identification

Suppose you have a system of N endogenous variables and K exogenous variables. Equation i in that system is exactly identified if the number of exogenous variables omitted from that the equation $K - K_i$ (K_i is the number of exogenous variables in equation i) is equal to the number of endogenous variables in that equation, N_i, less one, i.e.

$$K - K_i = N_i - 1 \text{ ,or equivalently } (K - K_i) + (N - N_i) = N - 1$$

In a similar way we get that equation is overidentified if

$$K - K_i > N_i - 1 \text{ ,or equivalently } (K - K_i) + (N - N_i) > N - 1$$

And it is not identified if:

$$K - K_i < N_i - 1 \text{ ,or equivalently } (K - K_i) + (N - N_i) < N - 1$$

The rank condition for identification

- The order condition is necessary, but not sufficient for identification
- If the coefficient of the omitted exogenous variable in an equation i of the system *is zero also in the equation(s)* where it is not omitted, its omission does not help for identification of equation i
- The omitted variable is in that case not a part of the statistical system (the reduced form), and it is not a valid instrument

Consider the following exactly identified simultaneous (supply and demand) equation system:

$$x_t = \alpha_1 + \beta_{11} p_t + \gamma_{11} z_{1t} + \varepsilon_{1t} \tag{10}$$

$$p_t = \alpha_2 + \beta_{22} x_t + \gamma_{22} z_{2t} + \varepsilon_{2t} \tag{11}$$

where the z's are exogenous and the disturbances have the following joint distribution

$$\boldsymbol{\varepsilon}_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim N\left(\mathbf{0}_{2\times 1}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right)$$

Let us generate artificial data with the following structural parameters

$$\alpha_1 = 1.2, \alpha_2 = 0.8$$
$$\beta_{11} = 0.6, \beta_{22} = -1.4$$
$$\gamma_{11} = 2, \gamma_{22} = 1.5$$
$$\sigma_1^2 = 1, \sigma_2^2 = 1, \sigma_{12} = 0.7$$

Thus, our system looks like this:

$$
\begin{aligned}
x_t &= 1.2 + 0.6 p_t + 2 z_{1t} + \varepsilon_{1t} \\
p_t &= 0.8 - 1.4 x_t + 1.5 z_{2t} + \varepsilon_{2t}
\end{aligned}
$$

$$\boldsymbol{\varepsilon}_t = \sim N\left(\mathbf{0}_{2\times 1}, \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}\right)$$

How well is OLS eq.-by-eq. doing in pinning down the coefficients as specified in the DGP?

We have a full simultaneous structure, and would therefore expect OLS on either the demand or the supply equation to lead to a bias (confer the notes to Lecture # 6)

Let us concentrate on the supply equation (eq. (10)), and do a **Monte Carlo** to get an impression of the "size" of the bias!

A consistent estimator in the exactly identified case: ILS and IV

- When a structural equation is exactly identified, we can retrieve the true structural parameters of that equation if we know the coefficients of the reduced form (i.e, the system, or the VAR)
- If we do not know the reduced form coefficients, but have a consistent estimator of the reduced form coefficients, we also have a consistent estimator for the structural form coefficients
- The OLS estimator is consistent for the reduced form, which we call the "solved out" estimator or the *Indirect Least Squares Estimator*, ILS
- In the exactly identified case, ILS is numerically identical to the IV estimator

As we know theoretically and as we shown by simulations, we get a bias when using OLS to pin down the coefficients in the supply equation! But we also know that there is still hope if we can find an instrument for p_t! What is the most natural instrument for p_t in the supply equation? Remember that an instrument for p_t, call it w_t, should have the following properties:

1. It must be correlated with the variable it is acting as an instrument for, i.e. $cov(p_t, w_t) \neq 0$
2. It must be uncorrelated with the disturbance in the equation where it is acting as an instrument, i.e. $cov(\varepsilon_{1t}, w_t) = 0$

Given this, there is one natural candidate: z_{2t}!

Let us do a new **Monte Carlo**, where we use z_{2t} as an instrument for p_t!

Identification in a standard (old) Keynesian model: The order condition

Consider the following dynamic (stochastic) Keynes-model

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

With reference to the OC, we have $N = 3$ and $K = 2$
And we find:

$$
\begin{aligned}
C_t &: K - K_C = 2 - 0 > N_C - 1 = 2 - 1 = 1 \Rightarrow \text{Overidentified} \\
I_t &: K - K_I = 2 - 1 = 1 = N_I - 1 = 2 - 1 = 1 \Rightarrow \text{Exactly identified} \\
Y_t &: \text{This is an identity}
\end{aligned}
$$

Identification table

	C_t	I_t	Y_t	Y_{t-1}	X_t
C-equation	1	0	0.75	0	0
I- equation	0	1	-1.5	-0.5	0
Identitiy	-1	-1	1	0	1

The Rank condition I

- We know that the order condition is only necessary
- The rank condition says that an equation in a system of N equations is identified *if and only if we can construct at least one non-zero* $(N-1)\times(N-1)$ *determinant from the coefficients excluded from that equation, but that are still contained in other equations in the system*
  Another way to think about this is that the matrix of the coefficients on the variables excluded from that equation should have rank $= N-1$

The Rank condition II

An intuitive way of understanding the rank condition: *All variables excluded from an equation in a system must appear in at least one other equation in that system (non-zero columns), i.e. it must be part of the reduced form. In addition, at least one of the variables excluded from that equation must appear in all other equations in the system*

The Rank condition II

Recall:

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

For the rank condition to be satisfied,we must have in both the consumption equation and in the investment equation that the coefficient matrix of the variables excluded from the equation under consideration has a rank equal to $(N-1) = 2$

The Rank condition III

Now, express this system in matrix notation in the following way;

$$\begin{pmatrix} 1 & 0 & 0.75 & 0 & 0 \\ 0 & 1 & -1.5 & 0.5 & 0 \\ -1 & -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} C_t \\ I_t \\ Y_t \\ Y_{t-1} \\ X_t \end{pmatrix} = \begin{pmatrix} \varepsilon_{Ct} \\ \varepsilon_{It} \\ 0 \end{pmatrix}$$

Rank condition: C-equation I

The coefficient matrix is:

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

Form a matrix of the coefficients of the variables excluded from the consumption equation that are contained in other equations:

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

Rank condition: C-equation II

Look at the under determinants of order 2:

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

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

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

Since at least one of these is $\neq 0$, the rank i 2 $\Rightarrow$ the rank condition is fulfilled!

Rank condition: Inv-equation

Remember the coefficient matrix:

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

Form a matrix of the coefficients of the variables excluded from the investment equation that are contained in the other two equations:

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

The determinant is:

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

and the rank condition is fulfilled

What do we take away form this?

- The *Order condition for identification* is a necessary, but not sufficient, 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 the main distinction is between non-identified and identified.

Recursive graphs of parameters

Consider the model

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \;\; t = 1, 2, \ldots, T, \tag{12}$$

A simple and intuitive method for assessing the hypothesis that the parameters β_1, β_2, $\sigma = \sqrt{Var(\varepsilon_t)}$ are stable, is to graph the output from a recursive estimation:

1. Estimate on a short sample: $t = 1, 2, ..., T_1, \; T_1 < T$
2. Extend the sample with one single observation, and estimate the model on the sample $t = 1, 2, ..., T_1 + 1, \; T_1 + 1 < T$
3. Continue until all observations are part of the sample, i.e. $t = 1, 2, ..., T$
4. The sequence of estimates, $\hat{\beta}_1(j)$, $\hat{\beta}_2(j)$, and $\hat{\sigma}(j)$ $(j = T_1, T_1 + 1, \ldots, T)$, can be plotted against time

Pc-Give plots e.g., $\hat{\beta}_2(j)$ together with $\pm 2\sqrt{Var(\hat{\beta}_1(j))}$, and $\hat{\sigma}(j)$ together with the *1-step residuals*, which are

$$y_j - \hat{\beta}_1(j) - \hat{\beta}_2(j)x_j, \quad j = T_1, T_1 + 1, ..., T$$

Chow-tests of parameter stability

- If we have a hypothesis about when a structural break occurs we can test that hypothesis
- Let T_1 denote the last period with the "old" regime and let $T_1 + 1$ denote the first period of the "new";

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \; t = 1, 2, \ldots, T_1 \text{ and}$$
$$y_t = \gamma_1 + \gamma_2 x_t + \epsilon_s, \; s = T_1 + 1, 2, \ldots, T.$$

then

$$H_0\colon \beta_1 = \gamma_1, \, \beta_2 = \gamma_2 \text{ vs } H_1\colon \beta_1 \neq \gamma_1, \, \beta_2 \neq \gamma_2.$$

- In the multivariate case:

$$H_0\colon \beta_1 = \gamma_1, \, \beta_2 = \gamma_2, \beta_3 = \gamma_3, \ldots, \beta_K = \gamma_K$$

- There are two well known statistics for these cases, both due to Chow (1960) and referred to as *Chow tests*.

2-sample Chow-test

Let SSE be the Sum of Squared Errors (which of course corresponds to RSS or SSR):
SSE_1 is for the first sample ($t = 1, 2, .., T_1$) and SSE_2 is for the second.
$SSE_U = SSE_1 + SSE_2$. SSE_R is the SSE when the whole sample is used, i.e. under H_0

$$F_{Chow2} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{(T-4)}{2} \sim F(2, T-4).$$

In general

$$F_{Chow2} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T-2K}{K} \sim F(K, T-2K).$$

Insufficient observations Chow-test

- Consider $T - T_1 < K$. Same SSE_R (full sample) but SSE_U is only on the basis of the first T_1 observations. This "predictive" Chow-test is given as

$$F_{ChowP} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T_1 - K}{T - T_1} \sim F(T - T_1, T_1 - K)$$

- If we have no clear idea about the dating of a regime shift, a graph with the whole sequence of predictive Chow tests is useful.
- Chow tests rely on constant and equal variances of the disturbances. Hence, it is good practice to plot the recursively estimated $\hat{\sigma}$.

Recursive Chow-test

The Test-Menu Graphics in PcGive contains three version of the F_{ChowP} test!

- *1-step test*: the sample size T_1 is increased by one observation sequentially, and $N = T - T_1$ is always 1.
- *Break-point Chow test* (called *N-dn CHOWS* in the graphs): Here T_1 is increased by one observation, then by 2 and so on until $T_1 = T$. Hence, $N = T - T_1$ is decreasing through the sequence of tests.
- Forecast Chow test (called *N-up CHOWS* in the graphs): Here T_1 is kept fixed and $T - T_1$ is first 1, then 2, and so on. Hence, $N = T - T_1$ is increasing through the sequence of tests.
- All the Recursive Chows are scaled by (1-off) critical values. (1% level is the default) so that the critical values becomes a straight line at unity.