Department of Economics,
Management and Quantitative Methods

B-74-3-B Time Series Econometrics
Academic year 2019-2020

Computer Session 3

Exercise 8: Modelling Vector Autoregressions.

8.1 Start

Download the file regression.wf1 from the website, and open it.

8.2 Estimate a VAR model for series y and x.

From the Menu *Quick* select *Estimate VAR*. You are prompted by the window *VAR Specification*. We can choose between VAR and Vector Error Correction (and Bayesian VAR, but we do not cover this topic). At this stage we keep the option as *VAR*.

In the *endogenous variables* cell, you must specify the variables (y and x: notice that we do not specify the presence of the constant, that goes in another cell); in *Lag Intervals for endogenous* the intervals of lags we want, for example if we write *1 1* we mean that we want to start at lag 1 (so, with y_{t-1} and x_{t-1}) and finish at lag 1 (so, with y_{t-1} and x_{t-1}), meaning that y_{t-1} and x_{t-1} are all the lagged variables that we want; if we write *1 3* we mean that we want to start at lag 1 (so, with y_{t-1} and x_{t-1}) and finish at lag 3 (so, with y_{t-3} and x_{t-3}), meaning that y_{t-1}, x_{t-1} , y_{t-2}, x_{t-2} , y_{t-3} and x_{t-3} are all the lagged variables that we want. For example suppose that we want a VAR with two lags, and write *1 2* (this was the default, by the way). Finally, we specify if we want the constant or trends in the equations: these are not determined with the VAR, so they are

exogenous, hence they are mentioned in the cell *Exogenous variables*. For example, for the constant we write *c.*
Click *OK*. We are presented with the estimated VAR.
The estimated equations are listed as columns in the output. For example, the estimated equation for x is

$$\hat{x}_t = 1.34x_{t-1} - 0.33x_{t-2} + 0.32y_{t-1} - 0.01y_{t-2}$$

8.3 Select lag length

We estimated the VAR with 2 lags, but this was arbitrary. To select the lags we could use information criteria. Moreover, as the VAR models are nested (we can write a VAR with p lags as a restriction of a VAR with $p+1$ lags), we can test for example if the coefficients of lag $p+1$ are significant. We test this in several ways, the LR test being the natural procedure, as both the restricted and unrestricted models are estimated. To select the lags, in the Estimation Output do *View*, *lag structure*, *lag length criteria*: you are prompted with a mask asking how many lags to take as maximum: 8 seems reasonable so go for that one.

8.4 Estimate VAR with lags as selected.

From our example, the LR test / Information Criteria recommend 2 lags, so that is the model we estimate. Estimate the VAR with 2 lags and the constant.

8.5 Granger Causality test.

We sometimes are interested in testing if one variable anticipates the movement of another one. For example, if x anticipates y. This could be done by testing if the coefficients of past x are significantly different in the equation for y.
This may be interesting when forecasting; interpreting this as causality however requires additional assumptions.
To run the test, in the Estimation Output we do *View, lag structure, Granger causality/block exogeneity tests*.

In the output we can see for example that we can exclude past values of y in the equation of x (in our example, this is a test on two coefficients: notice that the P-value is approximately 0.13, so above 5%); we cannot exclude x from the equation of y. Thus, x Granger causes y, but y does not granger cause x.

8.6 Impulse Response Function.

In the Estimation Output window, choose *Impulse*.
We are prompted with a mask *Impulse Response.* In *Display Format* we can choose the type of output (Table, Multiple Graph, Combined Graph): Table is useful when we want to export the point values of the IRF for example to present them with another format, or when we want to comment on individual points, for example to say when is the maximum response and for which value; otherwise, Multiple Graph is set as default: this is OK as it shows all the graphs in the same sheet and we can have a quick view of everything. In Display information we may select if we want impulses of only a subset of variables, or responses of only a subset of variables: the default is that all the variables are considered for impulses and responses: this is usually OK but sometimes you may prefer to focus the output on only some variables (typically, when we have many variables, to avoid clutter in the output). Keep standard errors as asymptotics. The important tab is Impulse definition: as the correlation matrix may be identified in different ways, this is the stage where this decision is taken. Select *Cholesky* as Decomposition method: now the ordering: if we think that x causes y, we should order x first, otherwise, if we think that y causes x, we should order y first. Sometimes, economic theory may help us, for example there are good economic reasons to think that a short term interest rate determines the long term rate at simultaneous stage. Keep x causing y in this example.

8.7 Impulse response function, again.

In the Impulse definition menu, set *y x* instead, so y is assumed to cause x. Comparing the outcome we can see that they are fairly different, showing that the identification assumption plays a very important role.

8.8 Close file without saving.

Exercise 9: Modelling Vector Error Corrections.

9.1 Start

Download the file regression.wf1 from the website, and open it.

9.2 Select model for X1, Y1: lag selection

Estimate a VAR model for series *y1* and *x1*. Select lags using information criteria and LR. Notice that SIC selects 3 lags, LR selects 9 but if we ignore that one the choice would be for 4. On balance, 3 or 4 lags could be selected, here.

9.3 Johansen (full system) Cointegration test

Return to the *workfile* (in the window *window*, select *workfile*). From the workfile, select *Quick* and then *Group Statistics*, *Johansen cointegration test*. You are prompted by a mask asking for series: write *X1* and *Y1*. The next mask is called *Cointegration Test Specification*: you have to choose the model: take option 6, which considers all models. Specify 1 4 as lags. The *Schwarz information criterion* recommends model 1 (no intercept, no trend), and 1 cointegrating relation. This is not convincing, the Johansen statistics for the same model would recommend two cointegrating relations (thus, as if the series were I(0)).

Repeat the exercise with three lags: this time, SIC selects the same model, but the cointegration test has one cointegration relation only. Overall, this seems more convincing.

9.4 VAR benchmark

A VAR in level is still consistently estimated, even in presence of unit root and / or cointegration (notice, however, that the confidence intervals in the IRF are not correct). This could be a valuable guideline for what happens after cointegration is imposed. Estimate a VAR with 3 lags, and look at the IRF.

9.5 Estimating a cointegrated model

Select button Estimate, and Vector Error Correction as VAR type. Notice that the lags are now lags for first differences (notice the D in the window): three lags in levels correspond to two lags for the differences. Select panel *Cointegration*: we have to specify the cointegration model. Set the number of cointegrating relations to 1; for the deterministic component, the SIC selected the case with no intercept and no trend neither in the cointegrating equation nor in the VAR, however, to demonstrate what to do with the intercept select the case with intercept in CE but not in VAR (for interest rates, perhaps the case with intercept in the CE but not in VAR is more appropriate).

The estimated Cointegrating Equation is

$$ECM_t = x_t - 1.003y_t + 0.151$$

Corresponding to the estimated model

$$y_t = \frac{0.151}{1.003} + \frac{1}{1.003}x_t + \widehat{w}_t$$

where w_t is a I(0) process (not necessarily iid).

Notice that standard errors are also present in the estimation output. For example, this suggests that the coefficient of x could be restricted to be 1.

9.6 Testing cointegrating restrictions

The estimated system is

$$ECM_t = B_1 x_t + B_2 y_t + B_c$$

$$\Delta \mathrm{x}_t = A_1 ECM_{t-1} + \phi_{1,1}\Delta \mathrm{x}_{t-1} + \phi_{2,1}\Delta \mathrm{y}_{t-1} + \cdots + u_{1,t}$$

$$\Delta \mathrm{y}_t = A_2 ECM_{t-1} + \phi_{1,2}\Delta \mathrm{x}_{t-1} + \phi_{2,2}\Delta \mathrm{y}_{t-1} + \cdots + u_{2,t}$$

From the previous estimation output we see that $B_1 = 1$ (this is set by default as normalization); estimate $\hat{B}_2 = -1.003$: we may have some theory-based reasons to test $B_2 = -1$, corresponding to $\beta_1 = 1$ in model

$$y_t = \beta_0 + \beta_1 x_t + w_t$$

Finally, notice $\hat{A}_1$=-0.028, and not significantly different from 0 (notice the standard error in the estimation output). Restriction $A_1 = 0$ has a very important consequence: it means that variable X1 does not adjust toward the equilibrium in ECM, and all the adjustment is made by Y1. We thus estimate the three restrictions, $B_1 = 1$, $B_2 = -1$, $A_1 = 0$. For this purpose, press again button Estimate and select tab VEC restrictions and input restrictions

B(1,1)=1, B(1,2)=-1, A(1,1)=0

And press OK.

Check P-Value in the output, this is 0.74, so null hypothesis is not rejected (notice that two, not three restrictions were imposed, as B(1,1)=1 is a normalization condition).

9.7 IRF from cointegrated model

Press again Impulse from the estimation output. Check the residuals are orthogonolized ordering the variables as X1 first, Y1 afterwards (this reflects the fact that Y1 seems to follow X1 in this model, as we can see from the fact that the assumption that ECM does not lead X1 is not rejected).

Inspect the plot: shocks to X1 remain permanently in the system, with long run effects both to X1 and Y1. Shocks to Y1 do not remain permanently in the system. Most of the long term dynamics for both X1 and Y1 are in response to shocks of X1.

Notice that the responses of X1 and Y1 are the same after the 5^{th} period: this is because we imposed the cointegrating restriction that X1 and Y1 have the same ratio in the long run (see the restricted ECM equation).