2.3 *Exogeneity and Causality*

Consider the following extension of the VAR-model (*multivariate dynamic regression model*)

$$\mathbf{y}_t = \mathbf{C} + \sum_{i=1}^{p} \mathbf{A}_i' \mathbf{y}_{t-i} + \sum_{i=0}^{p} \mathbf{B}_i' \mathbf{x}_{t-i} + \epsilon_t,$$

where $p + 1 \leq t \leq T$, $\mathbf{y}_t' = (y_{1t}, \ldots, y_{mt})$, $\mathbf{C}$ is an $m \times 1$ vector of constants, $\mathbf{A}_1, \ldots, \mathbf{A}_p$ are $m \times m$ matrices of lag coefficients, $\mathbf{x}_t' = (x_{1t}, \ldots, x_{kt})$ is a $k \times 1$ vector of (possibly stochastic) regressors, $\mathbf{B}_0, \mathbf{B}_1, \ldots, \mathbf{B}_p$ are $k \times m$ coefficient matrices, and ϵ_t is an $m \times 1$ vector of errors having the properties

$$E(\epsilon_t) = E\{E(\epsilon_t | \mathbf{Y}_{t-1}, \mathbf{x}_t)\} = 0$$

and

$$E(\epsilon_t \epsilon_s') = E\{E(\epsilon_t \epsilon_s' | \mathbf{Y}_{t-1}, \mathbf{x}_t)\} = \begin{cases} \Sigma_\epsilon & t = s \\ \mathbf{0} & t \neq s, \end{cases}$$

where

$$\mathbf{Y}_{t-1} = (\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_1).$$

We can compile this in matrix form (exercise)

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{U},$$

where

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}'_{p+1} \\ \vdots \\ \mathbf{y}'_T \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}'_{p+1} \\ \vdots \\ \mathbf{X}'_T \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} \epsilon'_{p+1} \\ \vdots \\ \epsilon'_T \end{pmatrix}$$

with

$$\mathbf{X}_t = (1, \mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p}, \mathbf{x}'_t, \ldots, \mathbf{x}'_{t-p})$$

and

$$\mathbf{B} = (\mathbf{C}, \mathbf{A}'_1, \ldots, \mathbf{A}'_p, \mathbf{B}'_0, \ldots, \mathbf{B}'_p)'.$$

The estimation theory for this model is basically the same as for the univariate linear regression. For example the LS and (approximate) ML estimator of $\mathbf{B}$ is

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y},$$

and the ML estimator of Σ_ϵ is

$$\hat{\Sigma}_\epsilon = \frac{1}{T}\hat{\mathbf{U}}'\hat{\mathbf{U}}, \qquad \hat{\mathbf{U}} = \mathbf{Y} - \mathbf{X}'\hat{\mathbf{B}}.$$

Above we have made the crucial assumption that $\mathbf{x}_t$ is weakly exogenous for the parameters of interest, $\mathbf{B}$, and Σ_ϵ*. That is, the likelihood function of $\mathbf{x}$ contains no information that is relevant for the *estimation of the parameters of interest*. Otherwise we could not estimate them without taking the stochastic structure of $\mathbf{x}_t$ into account.

If furthermore $\mathbf{x}_t|\mathbf{Y}_{t-1}, \mathbf{x}_{t-1}, \ldots, \mathbf{x}_1$ is statistically independent of $\mathbf{Y}_{t-1}$, such that

$$f(\mathbf{x}_t|\mathbf{Y}_{t-1}, \mathbf{x}_{t-1}, \ldots, \mathbf{x}_1) = f(\mathbf{x}_t|\mathbf{x}_{t-1}, \ldots, \mathbf{x}_1)$$

(there is no feedback from $\mathbf{Y}_{t-1}$ to $\mathbf{x}_t$), then we may forecast $\mathbf{x}_t$ seperately from $\mathbf{y}_t$ and treat it as given when attempting to *forecast future observations* of $\mathbf{y}_t$. In that case $\mathbf{x}_t$ is said to be strongly exogenous to $\mathbf{y}_t$. We shall turn to that issue below.

*For a thorough discussion of *Exogeneity* see Engle, R.F., D.F. Hendry and J.F. Richard (1985). Exogeneity. *Econometrica*, **51**, 277–304.

Granger-causality and measures of feedback

One of the key questions that can be addressed with VAR-models is how useful some variables are in forecasting others.

If the history of y does not help to predict the future values of x, we say that y does not Granger-cause x. * Usually the prediction ability is measured in terms of the MSE (Mean Square Error). Hence, y fails to Granger-cause x, if for all $s > 0$

$$\begin{aligned}&\mathsf{MSE}(\widehat{x}_{t+s}|x_t, x_{t-1}, \ldots) \\ &\qquad = \mathsf{MSE}(\widehat{x}_{t+s}|x_t, x_{t-1}, \ldots, y_t, y_{t-1}, \ldots),\end{aligned}$$

where (e.g.)

$$\begin{aligned}&\mathsf{MSE}(\widehat{x}_{t+s}|x_t, x_{t-1}, \ldots) \\ &\qquad = E\left((x_{t+s} - \widehat{x}_{t+s})^2 | x_t, x_{t-1}, \ldots\right).\end{aligned}$$

It is then also said that $\mathbf{x}$ is block-exogenous with respect to $\mathbf{y}$.

*Granger, C.W. (1969). *Econometrica* **37**, 424–438. Sims, C.A. (1972). *American Economic Review*, **62**, 540–552.

Note: This is equivalent to strong exogeneity of x, provided that x is weakly exogenous for the parameters of interests (weak exogeneity + block exogeneity = strong exogeneity).

In terms of VAR models this can be expressed as follows:

Consider the $g = m + k$ dimensional vector $\mathbf{z}'_t = (\mathbf{y}'_t, \mathbf{x}'_t)$, which is assumed to follow a VAR(p) model

$$\mathbf{z}_t = \sum_{i=1}^{p} \Pi_i \mathbf{z}_{t-i} + \nu_t$$

where

$$\begin{aligned} E(\nu_t) &= \mathbf{0} \\ E(\nu_t \nu'_s) &= \begin{cases} \Sigma_\nu, & t = s \\ \mathbf{0}, & t \neq s. \end{cases} \end{aligned}$$

Partition the VAR of $\mathbf{z}$ as

$$\begin{aligned}\mathbf{y}_t &= \sum_{i=1}^{p} \mathbf{C}_{2i}\mathbf{x}_{t-i} + \sum_{i=1}^{p} \mathbf{D}_{2i}\mathbf{y}_{t-i} + \nu_{1t} \\ \mathbf{x}_t &= \sum_{i=1}^{p} \mathbf{E}_{2i}\mathbf{x}_{t-i} + \sum_{i=1}^{p} \mathbf{F}_{2i}\mathbf{y}_{t-i} + \nu_{2t}\end{aligned}$$

where $\nu_t' = (\nu_{1t}', \nu_{2t}')$ and $\mathbf{\Sigma}_\nu$ are correspondingly partitioned as

$$\Sigma_\nu = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

with $E(\nu_{it}\nu_{jt}') = \Sigma_{ij}$, $i, j = 1, 2$.

Now $\mathbf{x}$ does not Granger-cause $\mathbf{y}$ if and only if $\mathbf{C}_{2i} \equiv \mathbf{0}$, or equivalently, if and only if $|\Sigma_{11}| = |\Sigma_1|$, where $\mathbf{\Sigma}_1 = E(\eta_{1t}\eta_{1t}')$ with η_{1t} from the regression

$$\mathbf{y}_t = \sum_{i=1}^{p} \mathbf{C}_{1i}\mathbf{y}_{t-i} + \eta_{1t}.$$

Changing the roles of the variables we get the necessary and sufficient condition of $\mathbf{y}$ *not* Granger-causing $\mathbf{x}$.

Testing for the Granger-causality of $\mathbf{x}$ on $\mathbf{y}$ reduces to testing for the hypothesis

$$H_0 \; : \; \mathbf{C}_{2i} = \mathbf{0}.$$

This can be done with the likelihood ratio test by estimating with OLS the restricted * and non-restricted † regressions, and calculating the respective residual covariance matrices:

Unrestricted:

$$\hat{\Sigma}_{11} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{\nu}_{1t} \hat{\nu}'_{1t}$$

Restricted:

$$\hat{\Sigma}_{1} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{\eta}_{1t} \hat{\eta}'_{1t}.$$

*Perform OLS regressions of each of the elements in $\mathbf{y}$ on a constant, p lags of the elements of $\mathbf{x}$ and p lags of the elements of $\mathbf{y}$.

†Perform OLS regressions of each of the elements in $\mathbf{y}$ on a constant and p lags of the elements of $\mathbf{y}$.

The LR test is then

$$\text{LR} = (T-p)\left(\ln|\hat{\Sigma}_1| - \ln|\hat{\Sigma}_{11}|\right) \sim \chi^2_{mkp},$$

if H_0 is true.

<u>Example</u>. Granger causality between pairwise equity-bond market series

```
Pairwise Granger Causality Tests
Sample: 1965:01 1995:12
Lags: 12
```

Null Hypothesis:	Obs	F-Statistic	Probability
DFTA does not Granger Cause DDIV	365	0.71820	0.63517
DDIV does not Granger Cause DFTA		1.43909	0.19870
DR20 does not Granger Cause DDIV	365	0.60655	0.72511
DDIV does not Granger Cause DR20		0.55961	0.76240
DTBILL does not Granger Cause DDIV	365	0.83829	0.54094
DDIV does not Granger Cause DTBILL		0.74939	0.61025
DR20 does not Granger Cause DFTA	365	1.79163	0.09986
DFTA does not Granger Cause DR20		3.85932	0.00096
DTBILL does not Granger Cause DFTA	365	0.20955	0.97370
DFTA does not Granger Cause DTBILL		1.25578	0.27728
DTBILL does not Granger Cause DR20	365	0.33469	0.91843
DR20 does not Granger Cause DTBILL		2.46704	0.02377

The p-values indicate that FTA index returns Granger cause 20 year Gilts, and Gilts lead Treasury bill.

Let us next examine the block exogeneity between the bond and equity markets (two lags). Test results are in the table below.

Direction	LoglU	LoglR	2(LU-LR)	df	p-value
(Tbill, R20) --> (FTA, Div)	-1837.01	-1840.22	6.412	8	0.601
(FTA,Div) --> (Tbill, R20)	-2085.96	-2096.01	20.108	8	0.010

The test results indicate that the equity markets are Granger-causing bond markets. That is to some extend previous changes in stock markets can be used to predict bond markets.

2.4 *Geweke's* measures of Linear Dependence*

Above we tested Granger-causality, but there are several other interesting relations that are worth investigating.

Geweke has suggested a measure for linear feedback from $\mathbf{x}$ to $\mathbf{y}$ based on the matrices Σ_1 and Σ_{11} as

$$F_{\mathbf{x}\to\mathbf{y}} = \ln(|\Sigma_1|/|\Sigma_{11}|),$$

so that the statement that "$\mathbf{x}$ does not (Granger) cause $\mathbf{y}$" is equivalent to $F_{\mathbf{x}\to\mathbf{y}} = 0$. Similarly the *measure of linear feedback from* $\mathbf{y}$ *to* $\mathbf{x}$ is defined by

$$F_{\mathbf{y}\to\mathbf{x}} = \ln(|\Sigma_2|/|\Sigma_{22}|).$$

*Geweke (1982) *Journal of the American Statistical Association*, **79**, 304–324.

It may also be interesting to investigate the *instantaneous causality* between the variables. For the purpose, premultiplying the earlier VAR system of $\mathbf{y}$ and $\mathbf{x}$ by

$$\begin{pmatrix} \mathbf{I}_m & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_k \end{pmatrix}$$

gives a new system of equations, where the first m equations become (exercise)

$$\mathbf{y}_t = \sum_{i=0}^{p} \mathbf{C}_{3i}\mathbf{x}_{t-i} + \sum_{i=1}^{p} \mathbf{D}_{3i}\mathbf{y}_{t-i} + \omega_{1t},$$

with the error $\omega_{1t} = \nu_{1t} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\nu_{2t}$ that is uncorrelated with ν_{2t} * and consequently with $\mathbf{x}_t$ (important!). That is, we may describe the same structural relationship between $\mathbf{y}_t$ and $\mathbf{x}_t$ with contemporenously uncorrelated error terms for the price of including the current value x_t as additional explanatory variable.

$^{*}\mathrm{Cov}(\omega_{1t}, \nu_{2t}) = \mathrm{Cov}(\nu_{1t} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\nu_{2t}, \nu_{2t}) = \mathrm{Cov}(\nu_{1t}, \nu_{2t}) - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathrm{Cov}(\nu_{2t}, \nu_{2t}) = \boldsymbol{\Sigma}_{12} - \boldsymbol{\Sigma}_{12} = \mathbf{0}$

Similarly, the last k equations can be written as

$$\mathbf{x}_t = \sum_{i=1}^{p} \mathbf{E}_{3i}\mathbf{x}_{t-i} + \sum_{i=0}^{p} \mathbf{F}_{3i}\mathbf{y}_{t-i} + \omega_{2t}.$$

Denoting $\Sigma_{\omega i} = E(\omega_{it}\omega'_{it})$, $i = 1, 2$, there is instantaneous causality between $\mathbf{y}$ and $\mathbf{x}$ if and only if $\mathbf{C}_{30} \neq \mathbf{0}$ and $\mathbf{F}_{30} \neq \mathbf{0}$ or, equivalently, $|\Sigma_{11}| > |\Sigma_{\omega 1}|$ and $|\Sigma_{22}| > |\Sigma_{\omega 2}|$. Analogously to the linear feedback we can define instantaneous linear feedback

$$F_{\mathbf{x}\cdot\mathbf{y}} = \ln(|\Sigma_{11}|/|\Sigma_{\omega 1}|) = \ln(|\Sigma_{22}|/|\Sigma_{\omega 2}|).$$

A concept closely related to the idea of linear feedback is that of linear dependence, a measure of which is given by

$$F_{\mathbf{x},\mathbf{y}} = F_{\mathbf{x}\to\mathbf{y}} + F_{\mathbf{y}\to\mathbf{x}} + F_{\mathbf{x}\cdot\mathbf{y}}.$$

Consequently the linear dependence can be decomposed additively into three forms of feedback. Absence of a particular causal ordering is then equivalent to one of these feedback measures being zero.

Using the method of least squares we get estimates for the various matrices above as

$$\widehat{\Sigma}_i = (T-p)^{-1} \sum_{t=p+1}^{T} \widehat{\eta}_{it}\widehat{\eta}'_{it},$$

$$\widehat{\Sigma}_{ii} = (T-p)^{-1} \sum_{t=p+1}^{T} \widehat{\nu}_{it}\widehat{\nu}'_{it},$$

$$\widehat{\Sigma}_{\omega i} = (T-p)^{-1} \sum_{t=p+1}^{T} \widehat{\omega}_{it}\widehat{\omega}'_{it},$$

for $i = 1, 2$. For example

$$\widehat{F}_{\mathbf{x}\to\mathbf{y}} = \ln(|\widehat{\Sigma}_1|/|\widehat{\Sigma}_{11}|).$$

With these estimates one can test the particular dependencies,

No Granger-causality: $\mathbf{x} \to \mathbf{y}$ $H_{01} : F_{\mathbf{x}\to\mathbf{y}} = 0$

$$(T-p)\widehat{F}_{\mathbf{x}\to\mathbf{y}} \sim \chi^2_{mkp}.$$

No Granger-causality: $\mathbf{y} \to \mathbf{x}$ $H_{02} : F_{\mathbf{y}\to\mathbf{x}} = 0$

$$(T-p)\widehat{F}_{\mathbf{y}\to\mathbf{x}} \sim \chi^2_{mkp}.$$

No instantaneous feedback: $H_{03} : F_{\mathbf{x}\cdot\mathbf{y}} = 0$

$$(T-p)\widehat{F}_{\mathbf{x}\cdot\mathbf{y}} \sim \chi^2_{mk}.$$

No linear dependence: $H_{04} : F_{\mathbf{x},\mathbf{y}} = 0$

$$(T-p)\widehat{F}_{\mathbf{x},\mathbf{y}} \sim \chi^2_{mk(2p+1)}.$$

This last is due to the asymptotic independence of the measures $F_{\mathbf{x}\to\mathbf{y}}$, $F_{\mathbf{y}\to\mathbf{x}}$ and $F_{\mathbf{x}\cdot\mathbf{y}}$.

There are also so called Wald and Lagrange Multiplier (LM) tests for these hypotheses that are asymptotically equivalent to the LR test.

Note that in each case $(T-p)\hat{F}$ is the LR-statistic.

<u>Example</u>. The LR-statistics of the above measures and the associated χ^2 values for the equity-bond data are reported in the following table with $p = 2$.

[$\mathbf{y} = (\Delta \log \mathrm{FTA}_t, \Delta \log \mathrm{DIV}_t)$ and $\mathbf{x}' = (\Delta \log \mathrm{Tbill}_t, \Delta \log \mathrm{r20}_t)$]

	LR	DF	P-VALUE
x-->y	6.41	8	0.60118
y-->x	20.11	8	0.00994
x.y	23.31	4	0.00011
x,y	49.83	20	0.00023

Note. The results lead to the same inference as in Mills (1999), p. 251, although numerical values are different [in Mills VAR(6) is analyzed and here VAR(2)].