University of York

Department of Economics

PhD Course 2006

VAR ANALYSIS IN MACROECONOMICS

Lecturer: Professor Mike Wickens

Lecture 4

Macroeconomic Modeling with a VAR

Contents

1. The uses of a VAR

The VAR has been put to a large number of uses in macroeconomics, and has been used elsewhere in economics, for example, in finance.

Broadly there are four uses:

1. To represent an economic theory through a VAR

2. To represent data through a VAR without attempting to relate the VAR to a structural economic model

3. To estimate the response of economic variables to identifiable shocks estimated from a VAR

4. To estimate the reponse of key variables to changes of economic policy.

We will focus mainly on 2, 3 and 4. Also, to start we will consider the issues in the abstract, and then add the economics after.

When Sims (1980) first advocated the use of a VAR in economics it was in response to the prevailing orthodoxy that all econometric models should be structural models, i.e. should include identifying restrictions - these were mainly exclusion type of restrictions.

We can include the classification of variables into endogenous and exogenous as another set of restrictions.

In Sims's view most of these restrictions were "incredible" and should not be used.

Instead, he argued for the use of an unrestricted VAR with no distinction being made in the model between endogenous and exogenous variables.

The aim was to free-up econometric modeling from the constraints applied by economic theory and, in effect, "let the data speak" for themselves.

Unfortunately, this is not possible.

We have already shown that in fact some restrictions are needed.

1. We need to choose the variables to be included

2. We have to choose a restricted lag length for the VAR due to their being more coefficients to estimate than data to estimate them on.

There are also the issues of whether we are better off using

(i) a CVAR than a VAR in levels

(ii) a VMA than a VAR or a CVAR.

2. Traditional VAR analysis

Sims's starting point is the VAR for the vector of n variables z_t

$$\begin{aligned} A(L)z_t &= e_t \\ \text{where } A_0 &= I \text{ and } e_t \sim i.i.d(0, \Sigma) \end{aligned}$$

If $A(L)$ has all its roots outside the unit circle then the z_t are all stationary variables

The VAR can then be inverted as a VMA

$$z_t = A(L)^{-1}e_t = C(L)e_t, \qquad C_0 = I$$

If $A(L)$ roots has $p \leq n$ roots on the unit circle then p of the z_t variables are non-stationary.

If $p = n$ then all of the variables are I(1).

In this case we can consider re-writing the VAR as a CVAR

In either event, the shocks e_t are in general correlated with each other.

This means that in general every shock affects every other shock, and hence every variable z_t, instantly.

The impulse response functions for the VAR would give the response of $z_{i,t+s}$ to e_{jt}. These would be the coefficients of $C(L)$. Thus $\frac{\partial z_{i,t+s}}{\partial e_{jt}} = C_s(i,j)$. Note that $\frac{\partial z_{i,t}}{\partial e_{jt}} = 0$ for $i \neq j$, i.e. the j^{th} disturbance does not affect the i^{th} variable contemporaneously.

If we treat the VAR residuals as the shocks then we cannot determine any causal structure in the model. We cannot say which shock is affecting which variable, or when a variable changes what it is due to. This is not very useful for economists who tend to think of shocks as being independent innovations and would like to know their transmission mechanism.

It should, however, be noted that in an SEM the structural disturbances are typically contemporaneously correlated and not independent innovations.

And in a DSGEM the shocks affecting an economic agent such as a household will in general affect every one of their decision variables (consumption, savings, money demand, labour supply).

This all raises the issue of whether we should be thinking in terms of shocks as innovations

Choleski decomposition

Sims's response to this problem was implicitly to retain the notion of shocks as innovations but to concede the need for just one identification condition, to solve the problem of causality.

He proposed replacing e_t by

$$e_t = K\varepsilon_t$$

where

(i) ε_{it} $\{i = 1, ..n\}$, the elements of $\varepsilon_t = (\varepsilon_{1t},...,\varepsilon_{nt})'$ are distributed independently of each other and $\varepsilon_{it} \sim i.i.d.(0, 1)$.

(ii) K is a triangular matrix

$$K = \begin{bmatrix} k_{11} & 0 & ... & 0 \\ k_{21} & k_{22} & ... & 0 \\ . & . & ... & . \\ k_{n1} & k_{n2} & ... & k_{nn} \end{bmatrix}$$

This imparts a recursive structure to the errors as we can write

$$\begin{aligned} e_{1t} &= k_{11}\varepsilon_{1t} \\ e_{2t} &= k_{21}\varepsilon_{1t} + k_{22}\varepsilon_{2t} \\ &... \\ e_{nt} &= k_{n1}\varepsilon_{1t} + k_{n2}\varepsilon_{2t} + ... + k_{nn}\varepsilon_{nt} \end{aligned}$$

Hence,

$$
\begin{aligned}
E(e_t e_t') &= \Sigma \\
E(e_t e_t') &= E(K\varepsilon_t \varepsilon_t' K') \\
&= KE(\varepsilon_t \varepsilon_t')K' = KK'
\end{aligned}
$$

K places no restrictions on Σ, so the VAR can still be estimated by OLS.

Hence having estimated the model, and obtained an estimate of Σ, we can then derive K as a Choleski decomposition of Σ.

Scaling the Choleski matrix

Although K as defined above is unique, it can still be scaled to alter the size of the shock. Three common choices are:

(i) a unit shock

(ii) a one standard error shock

(iii) normalising the leading diagonal of K.

A unit shock implies that $E(\varepsilon_t \varepsilon_t') = I$. This is what we have assumed above.

A one standard error shock implies that $E(\varepsilon_t \varepsilon_t')$ is a diagonal matrix with the variances of the VAR shocks on the diagonal, i.e. $E(\varepsilon_t \varepsilon_t') = diag\{\sigma_{11}, \sigma_{22}, ..., \sigma_{nn}\}$. This is equivalent to replacing K with $D^{-1}K$, where D is a diagonal matrix with elements $\{\sqrt{\sigma_{11}}, \sqrt{\sigma_{22}}, ..., \sqrt{\sigma_{nn}}\}$.

Sims's original specification was to normalise the leading diagonal of K. As a result, from $\Sigma = KE(\varepsilon_t \varepsilon_t')K'$, we can show that $E(\varepsilon_{1t}^2) = \sigma_{11}$, $E(\varepsilon_{2t}^2) = \sigma_{22} - k_{21}^2 \sigma_{11}$, ... In other words, apart from the first shock, the variances of the shocks are difficult to interpret.

Causality and the Choleski decomposition

Another way of writing the relation between e_t and ε_t is as

$$K^{-1}e_t = \varepsilon_t$$

where K^{-1} is also lower triangular. This implies that

$$\begin{aligned}
e_{1t} &= k_{11}\varepsilon_{1t} \\
e_{2t} &= \frac{k_{21}}{k_{11}}e_{1t} + k_{22}\varepsilon_{2t} \\
e_{3t} &= (\frac{k_{31}}{k_{11}} - \frac{k_{21}k_{32}}{k_{11}k_{22}})e_{1t} + \frac{k_{32}}{k_{22}}e_{2t} + ... + k_{33}\varepsilon_{3t} \\
& \quad etc
\end{aligned}$$

This shows that the i^{th} VAR error depends on the $(i-1)^{th}$, $(i-2)^{th}$,... VAR errors plus an innovation ε_{it}.

Equivalently, we can re-write the VAR as

$$\begin{aligned}
K^{-1}A(L)z_t &= \varepsilon_t \\
B(L)z_t &= \varepsilon_t
\end{aligned}$$

where, as K^{-1} is lower triangular, so is B_0.

This implies that z_{it} depends on $z_{i-1,t}$, $z_{i-2,t}$,..., z_{1t}, on lags of all z_t and on ε_{it}.

Thus the shock ε_{1t} affects z_{1t}, but the other shocks have no effect on z_{1t} in period t.

As z_{1t} affects z_{2t} in period t, the shock ε_{1t} also affects z_{2t} in period t, and in turn all of the other z variables.

The order of the variables

The practical problem revealed by these alternative representations of the Choleski decomposition is knowing in what order to place the variables in the VAR as the order determines the causal structure.

We could either use economic theory as a guide, or we could try every possible ordering and see what difference it makes to the results, for example, to the impulse response functions.

If Σ were diagonal, or if the off-diagonal elements were small compared with those on the diagonal, then the ordering would make little difference to the results.

We return to the issue of the order later.

Economic example

A typical economic example is a VAR in three variables: the logs of M1, the CPI and real GDP with the variables in this order.

If K is lower triangular, the implication would be

(i) that shocks to M1 in period t affect the CPI and GDP in period t, but shocks to the CPI and GDP in period t do not affect M1 in period t.

(ii) that shocks to CPI in period t affect GDP in period t, but shocks to GDP in period t do not affect the CPI in period t.

This ordering raises a number of questions:

1. Is it reasonable from an economic point of view?

2. Is the money shock a money demand shock or a money supply shock?

3. If the monetary authority is using the interest rate as the monetary instrument money becomes endogenous. The CPI and GDP shocks will then affect the money stock. So does this identification scheme only work if the money supply is the policy instrument?

4. Is it really true that there is no time lag in the effect of money on prices and output?

This would, of course, depend on the data. If it were annual data then the answer is almost certainly YES, but if the data were monthly data then the answer would probably be NO. In this case we might want to restrict Σ to be diagonal.

5. Is there any cointegration - or long-run relation - between these variables that should be taken into account?

But now we are starting to use economic theory to find identifying restrictions, which Sims was trying to avoid.

Even this very simple example illustrates the practical problems that must be dealt with.

Alternative identification schemes

Once we see that all our identification scheme is doing is to find a way of defining independent shocks using a linear relation $e_t = K\varepsilon_t$ and choosing K such that $KK' = \Sigma$, then we may note that there are many other ways of doing this.

We have seen that making K lower triangular is one alternative.

Another identification scheme is to use an orthonormal transformation- for a recent application see Leeper, Sims and Zha (1997). This has the additional property that $K' = K^{-1}$.

It allows every shock ε_{it} to affect each z_{jt} variable in the current period.

In some ways this scheme is preferable. At least it avoids the problem we faced with the economic example as it allows CPI and GDP shocks to affect the money stock.

The drawback with this identification scheme is that giving an economic interpretation to the shocks becomes more problematical.

Summary of impulse response functions (IRFs)

In general the IRF is $\frac{\partial z_{i,t+s}}{\partial (shock)_{jt}}$

Type of IRF	Transformation	Restrictions
Unrestricted - uses OLS residuals	e_t	$E(e_t e_t') = \Sigma$
Choleski - uses lower triangular transformation	$e_t = K\varepsilon_t$	$E(\varepsilon_t \varepsilon_t') = I, \quad KK' = \Sigma$
Orthonormal - orthogonal transformation	$e_t = K\varepsilon_t$	$E(\varepsilon_t \varepsilon_t') = I, \quad K^{-1}\Sigma K = I$

Note: In EViews4

(i) unrestricted shocks are called RESIDUALS. There is the option to have one unit or one standard error residual shocks. The ordering of the variables does not affect the resulting impulse response functions.

(ii) the Choleski shocks are constructed so that $E(\varepsilon_t \varepsilon_t') = I$. A degrees of freedom correction is available, but will only make a difference for small values of T.

3. The Structural VAR (SVAR)

An alternative identification scheme is to use

$$Pe_t = Q\varepsilon_t$$

Thus the SVAR is just an intermediate case between Sims's position and full structural modeling.

If P and Q are both triangular then we can write

$$\begin{aligned} e_t &= K\varepsilon_t \\ K &= P^{-1}Q \end{aligned}$$

where K is still lower triangular.

In the SVAR of Blanchard (1989) P and Q were lower triangular except for a single element in each above the diagonal, which was non-zero.

More generally we can impose any restrictions on P and Q that we wish. The VAR would take the form

$$\begin{aligned} A(L)z_t &= Q\varepsilon_t \\ A_0 &= P \text{ and } \varepsilon_t \sim i.i.d.(0, I) \end{aligned}$$

where the coefficient matrices of lags of z_t would be unrestricted still.

This is known as a structural VAR because $A_0 \neq I$.

In choosing P and Q we would have to be careful to produce a unique identification of the model. In general this is not an easy thing to do

Example

Blanchard (AER 1989) proposed a Keynesian model of the macroeconomy and used a structural VAR.

Model

$$
\begin{aligned}
AD &: \quad y = c_{12}e_s + e_d \\
OL &: \quad u = a_{21}y + e_s \\
PS &: \quad p = a_{34}w + a_{31}y + c_{32}e_s + e_p \\
WS &: \quad w = a_{43}p + a_{42}u + c_{42}e_s + e_w \\
MR &: \quad m = a_{51}y + a_{52}u + a_{53}p + a_{54}w + e_m
\end{aligned}
$$

where y =log GDP, u =unemployment rate, p =log price level, w =log wage rate, m =log M1.

The AD equation is the aggregate demand function, OL is Okun's law, the remaining equations are the price, the wage and the money rule

(i) Structural VAR

$$
\begin{aligned}
Az_t &= A(L)z_{t-1} + Bx_t + Ce_t \\
E(e_t e_t') &= D, \text{ diagonal matrix}
\end{aligned}
$$

where $z_t' = (\Delta y_t, u_t, \Delta p_t, \Delta w_t, \Delta m_t)$ and $e_t' = (e_d, e_s, e_p, e_w, e_m)$.

Thus, where necessary, non-stationary variables are first differenced to make them stationarity (i.e. all except unemployment)

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 & 0 \\ -a_{31} & 0 & 1 & -a_{34} & 0 \\ 0 & -a_{42} & -a_{43} & 1 & 0 \\ -a_{51} & -a_{52} & -a_{53} & -a_{54} & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & c_{12} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & c_{32} & 1 & 0 & 0 \\ 0 & c_{42} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(ii) Reduced form VAR

$$\begin{aligned} z_t &= A^{-1}A(L)z_{t-1} + A^{-1}Bx_t + A^{-1}Ce_t \\ &= F(L)z_{t-1} + Gx_t + v_t \\ E(v_t v_t') &= \Sigma, \text{ unrestricted matrix} \end{aligned}$$

Thus, in this SVAR there are two elements that break the recursive structure of the model: a_{34} and c_{12}.

Given c_{12}, and either a_{34} or a_{43}, the SVAR is just identified. In the absence of this knowledge it is necessary to use structural estimation and not OLS.

EViews4 has an SVAR option which allows the user to specify the elements of P and Q.

4. VAR modeling with non-stationary data but without cointegration

If the variables are I(1) but are not cointegrated then the shocks will all have a permanent effect.

The model can be written (including a possible constant)

$$\begin{aligned} A(L)\Delta z_t &= \mu + e_t \\ \text{where } A_0 &= I,\ e_t \sim i.i.d(0, \Sigma) \end{aligned}$$

and the roots of $|A(L)| = 0$ all lie outside the unit circle.

Thus

$$\begin{aligned} \Delta z_t &= A(1)^{-1}\mu + A(L)^{-1}e_t = \nu + C(L)e_t, \quad C_0 = I \\ &= \upsilon + C(1)e_t + C^*(L)\Delta e_t \end{aligned}$$

where, in the absence of cointegration, $C(1)$ is of full rank.

This implies that there are n common stochastic trends (permanent shocks) which we denote by τ_t and define by $\Delta\tau_t = e_t$.

Integrating gives the solution for z_t as

$$z_t = z_0 + \upsilon t + C(1)\tau_t + C^*(L)e_t$$

As noted earlier the trend in z_t can be written as z_t^T which is comprised of a linear trend and a stochastic trend (the last term):

$$z_t^T = z_0 + \upsilon t + C(1)\tau_t$$

The impulse response functions will show that in general each shock will have a permanent effect on each variable, z_t.

Estimation

The model can be estimated in two ways: using

(i) a VAR in levels, despite the non-stationarity of the data

(ii) first differencing the variables then using a standard stationary VAR

It is common to use (ii), but Christiano, Eichenbaum and Evans (1998) use (i).

Identification

(i) Using Σ

The problem of identifying the shocks remains. Writing the VAR in first-differenced form shows that the same choices as for the stationary VAR exist. These focus on the covariance matrix of e_t.

(ii) Long-run restrictions

A new type of restriction is now possible, based on the long-run impact of a shock.

It is possible to impose the restriction that some shocks have no long-run effect on particular variables, z_t.

In this case, we would need to restrict certain elements of $C(1)$ to be zero.

If the j^{th} shock e_{jt} has no effect on the long-run on z_{it} then $C(1)_{ij} = 0$. At the same time we may wish to impose the restriction that the shocks are uncorrelated.

For example, we could assert that money shocks have no long-run effect on output - the super-neutrality condition of money.

Robertson and Wickens (1997) use this approach. They consider $z_t = (y_t, p_t)'$ where $y =$ log GDP and $p =$ log of the price level. It can be shown that y_t and p_t are I(1) but not cointegrated.

Consider a VAR in first differences.

This has two shocks: a real shock and a nominal shock.

Suppose that

(i) they are uncorrelated

(ii) the nominal shock has no long-run effect on output.

How do we construct such shocks from VAR residuals? We must construct orthogonal shocks and impose the additional restriction that one of these orthogonal shocks has no long-run effect on one of the variables.

1. Define K such that

$$\begin{aligned} e_t &= K\varepsilon_t \\ K^{-1'}\Sigma K &= I \text{ or } K'K = \Sigma \end{aligned}$$

2. Define

$$D(L) = C(L)K$$

Hence

$$\begin{aligned} \Delta z_t &= \nu + C(L)e_t \\ &= \upsilon + C(L)KK^{-1}e_t \\ &= \nu + D(L)\varepsilon_t \end{aligned}$$

We now have orthogonal shocks

K has $n^2 = 4$ elements and the restriction $K^{-1'}\Sigma K = I$ provides $\frac{n(n+1)}{2} = 3$ of these. We therefore have one degree of freedom left which we use to impose the long-run restriction.

3. Define

$$\Delta\tau_t = \varepsilon$$

Then

$$z_t = z_0 + \nu t + D(1)\tau_t + D^*(L)\Delta\tau_t$$

4. Now impose the long-run restriction

$$D(1)_{12} = 0$$

This implies that $C(1)$ and K must be defined to satisfy the restriction

$$D(1)_{12} = C(1)_{11}K_{21} + C(1)_{12}K_{22} = 0$$

This is the 4^{th} restriction required to identify K.

5. Thus we estimate the VAR in first differences

Then we solve for K using these restrictions.

Now we can use K and e_t to calculate ε_t and τ_t

Finally, we compute the impulse response functions.

5. VAR modeling using a cointegrated VAR (CVAR)

We recall that the CVAR can be written

$$\begin{aligned} A^*(L)\Delta z_t &= \alpha\beta' z_{t-1} + e_t, \quad A_0^* = I \\ &= \alpha w_{t-1} + e_t \\ w_t &= \beta' z_t \end{aligned}$$

and corresponding to this there is a VMA

$$\begin{aligned} \Delta z_t &= \gamma\theta'\Delta\tau_t + R^*(L)\Delta e_t \\ \Delta\tau_t &= e_t \end{aligned}$$

A number of different impulse response functions can be calculated

The response of z_t (or Δz_t) to

(i) the unrestricted shocks e_t

(ii) transformations of e_t, such as the Choleski decomposition or the orthonormal transformation

(iii) w_t, the cointegrating residuals. This measures the response of z_t to deviations from equilibrium

(iv) τ_t, the stochastic trends - these will be permanent shocks

We can also compute the response of w_t to shocks.

If we define the non-singular matrix

$$G = \begin{bmatrix} \beta' \\ 0 \quad I \end{bmatrix}$$

then

$$Gz_t = \begin{bmatrix} w_t \\ z_{2t} \end{bmatrix}$$

Hence

$$\begin{aligned}
A^*(L)G^{-1}G\Delta z_t &= \alpha\beta' z_{t-1} + e_t \\
A^*(L)G^{-1}\begin{bmatrix} \Delta w_t \\ \Delta z_{2t} \end{bmatrix} &= \alpha w_{t-1} + e_t \\
H(L)\begin{bmatrix} w_t \\ \Delta z_{2t} \end{bmatrix} &= H^*\begin{bmatrix} w_{t-1} \\ \Delta z_{2,t-1} \end{bmatrix} + e_t
\end{aligned}$$

Premultiplying by H_0^{-1} would give a VAR in the stationary variables w_t and Δz_{2t}.

Impulse response functions can then be computed in the usual way.

All of these IRFs suffer from the same identification as discussed earlier.

We now propose a new way to identify the shocks based on Wickens and Motto (2001).

Estimating shocks and impulse response functions using a CVAR

Explains a new way to identify VAR models

- Uses two pieces of information:

1. Can classify the variables into endogenous and exogenous
2. Have knowledge of the long-run structure

- Puts no restrictions on the covariance structure of the errors

- Can deal with stationary and non-stationary variables

Econometric model

Need to consider the equations for the endogenous and exogenous variables separately.

1. Endogenous variables: the simultaneous equation model (SEM)

$$B(L)y_t + C(L)x_t + Rd_t = e_t \tag{1}$$

y_t is a p x 1 vector of I(1) endogenous variables

x_t is a q x 1 vector of I(1) exogenous variables ($q \geq 1$)

d_t represents a vector deterministic variables (intercepts, trends and seasonals etc.)

e_t is distributed i.i.d (0,Σ)

roots of $|B(L)| = 0$ lie outside the unit circle

The long-run solution defines p cointegrating vectors β (an n x p matrix)

$\beta' = [B(1) \;\; C(1)]$.

The normalization is made on the long-run model

$\beta_{ii} = 1$ *for* $\{i = 1, .., p\}$

2. Exogenous variables: the incomplete VAR

$$D(L)\Delta x_t + E(L)\Delta y_{t-1} = Sd_t + \varepsilon_t \tag{2}$$

ε_t is distributed i.i.d(0,Ψ)

e_t and ε_t are uncorrelated

$D(0) = I$ and the roots of $|D(L)| = 0$ lie outside the unit circle

NB. if $E(L) = 0$ then x_t will be strongly exogenous

3. SEM written in generalised form with CVs

$$
\begin{aligned}
[B(0)\ C(0)]\Delta z_t &= -\beta' z_{t-1} + [\widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1} \\
&\quad + Rd_t + e_t
\end{aligned} \tag{3}
$$

$z_t = (y_t', x_t')'$.

4. Identification

Use only long-run restrictions and leave the short-run dynamics unrestricted

i.e. $B(0)$, $C(0)$ and the higher order lag coefficient matrices are unrestricted

5. The long-run structural model

This is derived by eliminating all time subscripts by setting $L = 1$ in the SEM to obtain

$$
B(1)y_t + C(1)x_t = \beta' z_t = Rd_t + e_t^L \tag{4}
$$

- diagonal elements of $B(1)$ are normalized to unity

- e_t^L is the cointegrating residual, or long-run disturbance

- elements of β are assumed to be restricted so that equation (4) is identified

- identification of β is achieved by applying the usual rank and order conditions

for simultaneous equation systems to $\beta' z_t$

6. SEM written as a VECM

The aim here is to retain the dynamics but to separate the long-run from the short run.

$$
\begin{aligned}
[I \quad B(0)^{-1}C(0)]\Delta z_t &= -B(0)^{-1}\beta' z_{t-1} + B(0)^{-1}Rd_t \\
&\quad + B(0)^{-1}[\widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1} + B(0)^{-1}e_t
\end{aligned} \tag{5}
$$

or

$$
\begin{aligned}
\Delta y_t &= -B(0)^{-1}C(0)\Delta x_t - B(0)^{-1}w_{t-1} + B(0)^{-1}Rd_t \\
&\quad +B(0)^{-1}[\widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1} + u_t
\end{aligned}
$$

7. Complete model written as a VECM

The complete model contains both endogenous and exogenous variables and can also be written as VECM

$$
\begin{aligned}
\begin{bmatrix} I & B(0)^{-1}C(0) \\ 0 & I \end{bmatrix} \Delta z_t &= \begin{bmatrix} -B(0)^{-1} \\ 0 \end{bmatrix} \beta' z_{t-1} + \begin{bmatrix} B(0)^{-1}R \\ S \end{bmatrix} d_t \\
&\quad + \begin{bmatrix} B(0)^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \widetilde{B}(L) & \widetilde{C}(L) \\ \widetilde{E}(L) & \widetilde{D}(L) \end{bmatrix} \Delta z_{t-1} \\
&\quad + \begin{bmatrix} B(0)^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} e_t \\ \varepsilon_t \end{bmatrix}
\end{aligned}
\tag{6}
$$

where $\widetilde{E}(L) = -E(L)$

8. Complete model written as a cointegrated VAR (CVAR)

This is in, effect, the reduced form of the VECM as each equation has only one contemporaneous endogenous variable

$$\Delta z_t = \begin{bmatrix} -B(0)^{-1} \\ 0 \end{bmatrix} \beta' z_{t-1} + \begin{bmatrix} B(0)^{-1}[R - C(0)S] \\ S \end{bmatrix} d_t$$
$$+ \begin{bmatrix} B(0)^{-1} & -B(0)^{-1}C(0) \\ 0 & I \end{bmatrix} \begin{bmatrix} \widetilde{B}(L) & \widetilde{C}(L) \\ \widetilde{E}(L) & \widetilde{D}(L) \end{bmatrix} \Delta z_{t-1}$$
$$+ \begin{bmatrix} B(0)^{-1} & -B(0)^{-1}C(0) \\ 0 & I \end{bmatrix} \begin{bmatrix} e_t \\ \varepsilon_t \end{bmatrix} \qquad (7)$$

CVAR in familiar notation

More compactly, we can write

$$\Delta z_t = -\alpha\beta' z_{t-1} + A(L)\Delta z_{t-1} + Qd_t + v_t \qquad (8)$$

$$\alpha = \begin{bmatrix} -B(0)^{-1} \\ 0 \end{bmatrix},$$

$$A(L) = \begin{bmatrix} B(0)^{-1} & -B(0)^{-1}C(0) \\ 0 & I \end{bmatrix} \begin{bmatrix} \widetilde{B}(L) & \widetilde{C}(L) \\ \widetilde{E}(L) & D(L) \end{bmatrix}$$

$$Q = \begin{bmatrix} B(0)^{-1}[R - C(0)S] \\ S \end{bmatrix}$$

$$v_t = \begin{bmatrix} B(0)^{-1} & -B(0)^{-1}C(0) \\ 0 & I \end{bmatrix} \begin{bmatrix} e_t \\ \varepsilon_t \end{bmatrix}$$

$$
\begin{aligned}
E(v_t v_t') &= \Omega \\
&= \begin{bmatrix} B(0)^{-1}(\Sigma + C(0)\Psi C(0)')B(0)^{-1\prime} & -B(0)^{-1}C(0)\Psi \\ -\Psi C(0)' B(0)^{-1\prime} & \Psi \end{bmatrix}
\end{aligned}
$$

α and β are restricted

$A(L)$ and Ω are unrestricted.

Estimation of the CVAR

- usually performed on an unrestricted CVAR using the Johansen estimator
- no distinction is made between endogenous and exogenous variables
- the cointegrating vectors are estimated unrestrictedly
- CVs are not identified when there is more than one cointegrating vector and α is unrestricted
- Alternative estimator proposed by Pesaran and Shin (1999)

- impose identifying restrictions on the cointegrating vectors

- use maximum likelihood estimation subject to this constraint

- To use here must take account of the restrictions associated with α as well as those on β.
- NB: Christiano, Eichenbaum and Evans (1999) in effect ignore the presence of cointegrating vectors in equation (8) by not factorising the coefficient matrix of z_{t-1}.since they are primarily interested in estimating the impulse response functions, and not the individual coefficients. They re-write and estimate the equation in levels.

Estimating the system shocks

- Ideally, we would like to estimate the shocks to the system, e_t and ε_t
- e_t are the stationary structural disturbances, and are temporary shocks
- ε_t give the non-stationary stochastic trends, and are permanent
- Problem with the Johansen and the Pesaran-Shin approach is that they provide an estimate of the cointegrated VAR disturbances v_t and NOT an estimate of the structural disturbances e_t, or of the common stochastic trends ε_t.

- To identify and estimate e_t and ε_t need estimates of $B(0)$ and $C(0)$

Can then transform v_t and to e_t and ε_t.

BUT this is not possible from the unrestricted cointegrated VAR.

- The usual approach followed instead is

- assume a particular ordering of the variables in z_t

- use a Choleski decomposition of the covariance matrix of v_t

- this gives a recursive structure to the model

- Problem

- multiplying the estimate of v_t by the inverse of the Choleski matrix

is assumed to produce estimates of e_t and ε_t

- BUT in general this gives a *linear transformation* of e_t and ε_t

- need a recursive structure to estimate e_t and ε_t

- ## 6. Proposed new approach to estimating shocks

This is based on Wickens and Motto (2001).

1. Assume that the variables z_t can be classified as endogenous or exogenous

2. Consider the equations for the endogenous and exogenous variables separately.

- Endogenous variables: re-write as the VECM

$$\begin{aligned}\Delta y_t &= -B(0)^{-1}C(0)\Delta x_t - B(0)^{-1}w_{t-1} + B(0)^{-1}Rd_t \\ &\quad +B(0)^{-1}[\widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1} + u_t \end{aligned} \tag{9}$$

$u_t = B(0)^{-1}e_t$

Δx_t is treated as an explanatory variable

$w_t = \beta' z_t$ is treated as observable

- NB: Either β is assumed known

  Or, it is assumed that there is an *a prior* (super-consistent) estimate of β

  Or, β can be estimated at the same time as the other coefficients

- Can estimate equation (9) by OLS as all of the variables in it are stationary, Δx_t is exogenous (and hence uncorrelated with the equation disturbances) and the coefficient matrices and error covariance matrix are unrestricted

- Assuming that any estimate of β is super-consistent, the asymptotic properties of this OLS estimator will be unaffected by treating w_{t-1} as observed

- Now then have a consistent estimator of an unrestricted $B(0)$ from the term in w_{t-1}

- We can now estimate e_t from u_t using $e_t = B(0)u_t$

- Estimates of all of the remaining coefficient matrices of equation (1) can also be derived

- Exogenous variables: estimate equation (2) by OLS

The residuals will be a direct estimate of ε_t.

Impulse response functions

- Often the main object of a VAR analysis is to calculate the impulse response functions of the variables to the system shocks

- This means estimating the responses of y_t, x_t and w_t to the structural shocks e_t and the stochastic trends ε_t.

- Problem is how to identify and estimate the shocks

- Usually

- no distinction is made between endogenous and exogenous variables

- arbitrary identifying schemes are introduced in order to try to give economic meaning to the shocks

- Even bigger problem

- even if the shocks are identified, the impulse response functions are not

- this is because, in general, the structural shocks e_t will be correlated with each other, as will the stochastic trends ε_t

- hence, part of the transmission mechanism of each shock is via their correlation with the other shocks

- In order to estimate the full effect of a shock (even when it is identified) it is therefore necessary to know how to transform the shocks to orthogonality

-There is no unique way to do this without additional *a priori* information, which is usually not available

- For example

1. A canonical factorization makes the eigenvectors the transformation matrix, but these have no economic meaning as they are ordered according to the size of the eigenvalues
2. A Choleski decomposition requires an ordering of the structural shocks, and implies that each shock is a linear function of the previous shock and a component orthogonal to this.

- Ignoring any correlation between the shocks has the virtue of making it clear what the impulse response functions mean. But the fact that this would then give the same impulse response functions as assuming that the shocks are uncorrelated shows what the disadvantage is

- It may therefore be better to accept that the state of nature is that the shocks are contemporaneously correlated, and instead calculate standard dynamic multipliers which ignore the correlation. This is the approach proposed.

- NB: Pesaran and Shin (1999) and Garratt, Lee, Pesaran and Shin (1999) have proposed computing the persistence profile

- This is the response of w_t - a measure of the deviation of each structural relation from long-run equilibrium - to the shocks v_t.
- Unfortunately, it is not possible in general to give an economic interpretation to the persistence profile. To do so one would need to impose identifying restrictions on the v_t.

Dynamic multiplier analysis proposed

- We can compute dynamic multipliers (impulse response functions) for y_t and w_t

  - these would also be the persistent profiles for each type of shock e_t and ε_t

- In general the temporary shocks e_t will be contemporaneously correlated, and so it makes no sense to try to orthogonalise them. Recall they are the structural disturbances.

- It is not clear whether ε_t, the shocks to the exogenous variables, are correlated or uncorrelated. If correlated then it might be considered necessary to derive uncorrelated shocks from ε_t in the usual way. It probably makes more sense, however, to impose the restriction that the ε_t are uncorrelated. This would have to done at the initial estimation stage by the use of restricted maximum likelihood.

- *A deeper consideration is that perhaps there is little point in finding the response to the ε_t shocks as the real interest lies in the dynamic response of y_t to x_t. This can be obtained directly from the VECM using a conventional dynamic multiplier analysis. Thus there is no need to use the VAR at all.*

- If the responses of z_t to e_t and ε_t is required then they can be computed using the companion form

$$
\begin{bmatrix} z_t \\ z_{t-1} \\ z_{t-2} \\ . \\ z_{t-k} \end{bmatrix} = \begin{bmatrix} I - \alpha\beta' + A(0) & -A(1) + A(2) & ... & -A(k) \\ I & 0 & ... & 0 \\ 0 & I & ... & 0 \\ . & . & ... & 0 \\ 0 & 0 & .. & 0 \end{bmatrix} \begin{bmatrix} z_{t-1} \\ z_{t-2} \\ z_{t-3} \\ . \\ z_{t-k-1} \end{bmatrix} + \begin{bmatrix} B(0)^{-1} & -B(0)^{-1}C(0) \\ 0 & I \\ \multicolumn{2}{c}{0} \\ \multicolumn{2}{c}{0} \\ \multicolumn{2}{c}{.} \\ \multicolumn{2}{c}{0} \end{bmatrix} \begin{bmatrix} e_t \\ \varepsilon_t \end{bmatrix} \tag{10}
$$

where k is assumed to be the maximum lag and the deterministic variables d_t are omitted.

- Can write more compactly as the $(p+q)k$ system

$$
\bar{z}_t = \Gamma \bar{z}_{t-1} + \bar{v}_t \tag{11}
$$

- The impulse response functions for y_{t+s} are

$$\frac{dy_{t+s}}{de_t} = \begin{bmatrix} I_p & 0 \end{bmatrix} \begin{bmatrix} I_{p+q} & 0 & ... & 0 \end{bmatrix} \Gamma^s \begin{bmatrix} B(0)^{-1} \\ 0 \\ 0 \\ . \\ 0 \end{bmatrix} \tag{12}$$

$$\frac{dy_{t+s}}{d\varepsilon_t} = \begin{bmatrix} 0 & I_q \end{bmatrix} \begin{bmatrix} I_{p+q} & 0 & ... & 0 \end{bmatrix} \Gamma^s \begin{bmatrix} -B(0)^{-1}C(0) \\ I \\ 0 \\ . \\ 0 \end{bmatrix} \tag{13}$$

- As as $s \to \infty$

  (i) $\frac{dy_{t+s}}{de_t}$ will converge to zero because the e_t are temporary shocks

  (ii) $\frac{dy_{t+s}}{d\varepsilon_t}$ will not converge to zero as the ε_t are permanent.

- For $s = 0$

(i) $\frac{dy_t}{de_t} = B(0)^{-1}$

(ii) $\frac{dy_t}{d\varepsilon_t} = -B(0)^{-1}C(0)$

  - can also obtain from $\frac{dy_t}{d\varepsilon_t} = \frac{dy_t}{dx_t}\frac{dx_t}{d\varepsilon_t}$ using $\frac{dy_t}{dx_t} = -B(0)^{-1}C(0)$ and $\frac{dx_t}{d\varepsilon_t} = I$.

  - thus the stochastic trends of y_t, and hence the source of the non-stationarity of y_t are the reduced form shocks to the exogenous variables x_t.

- The impulse response functions of w_{t+s} give the persistence profiles, or paths back to equilibrium, of each of the long-run relations

- $\frac{dw_{t+s}}{de_t} = \beta' \frac{dz_{t+s}}{de_t}$ and $\frac{dw_{t+s}}{d\varepsilon_t} = \beta' \frac{dz_{t+s}}{de_t}$ and both converge to zero.

Conclusions so far:

1 By

(i) identifying which variables are endogenous and which exogenous

(ii) imposing identifying restrictions on the long-run structural coefficients

it is possible to

(a) identify and obtain estimates of all of the coefficients of the model

(b) identify and estimate the shocks e_t and ε_t using OLS

2.The CVAR model implied by this approach will have unrestricted short-run dynamics and an unrestricted covariance matrix of the errors.

3. Since the structural shocks and the common stochastic trends are in general intra-correlated (though not inter-correlated), to calculate the full impulse response of the variables to the shocks it is necessary to decompose them into their orthogonal components.

4. Since the absence of *a priori* restrictions makes the decomposition arbitrary from an economic point of view, we suggest simply estimating the dynamic multipliers for each shock (i.e. the impulse response functions ignoring any correlation among the shocks) on the grounds that these do have economic meaning.

Including stationary variables in a CVAR

Usually in CVAR analysis it is assumed that all of the variables are I(1). In practice in economics some variables may be I(0). How do we deal combine I(1) and I(0) variables when there is also cointegration?

If there is no cointegration then we can simply form a VAR from first differences of the I(1) variables and levels of the I(0) variables.

When there are I(0) variables we must include them in the structural equations and the equations for the exogenous variables. We must also specify equations for the stationary variables. We look at each set of equations in turn.

1. Simultaneous equation model (SEM) determining the I(1) variables

$$F(L)s_t + B(L)y_t + C(L)x_t = e_t \tag{14}$$

where s_t is an r x 1 vector of stationary endogenous variables.

Can re-write as

$$\begin{aligned}[F(0)\ B(0)\ C(0)]\Delta z_t^* &= -F(1)s_{t-1} - \beta' z_{t-1} \\ &\quad +[\widetilde{F}(L)\ \widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1}^* + e_t\end{aligned}$$

$$\begin{aligned}[F(0)\ B(0)\ C(0)]\Delta z_t^* &= -[F(1)\ I]\beta^{*\prime} z_{t-1}^* \\ &\quad +[\widetilde{F}(L)\widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1}^* + e_t\end{aligned} \tag{15}$$

where

$$\beta^{*\prime} = \begin{bmatrix} I & 0 \\ 0 & \beta' \end{bmatrix}, \; y_t^* = (s_t' \;\; y_t')' \text{ and } z_t^* = (s_t' \;\; y_t' \;\; x_t')'$$

We can distinguish between

- the long-run structural equations (which may include stationary variables)

- the cointegrating vectors (which will not include stationary variables)

Long-run structure

$$\begin{aligned} \tilde{\beta}' z_t^* &= [F(1)\ I] \begin{bmatrix} I & 0 \\ 0 & \beta' \end{bmatrix} z_t^* \\ &= [F(1)\ I] \beta^{*\prime} z_t^* \\ &= [F(1)\ I]\, w_t^* \end{aligned}$$

$w_t^{*\prime} = [s_t'\ w_t']$

$\tilde{\beta}' = [F(1)\ B(1)\ C(1)] = [F(1)\ \beta']$

- can be interpreted as the long-run coefficient matrix

- as before, it will be restricted *a priori.*

2. Model for the stationary endogenous variables

$$J(L)\Delta s_t + G(L)\Delta y_t + H(L)\Delta x_t + M s_{t-1} + K\beta' z_{t-1} = \xi_t \tag{16}$$

- ξ_t is distributed i.i.d(0,Φ) and is independent of ε_t, but not in general of e_t

- the roots of $|J(L)(1-L)+ML| = 0$ lie outside the unit circle

Can re-write as

$$\begin{aligned} [J(0)\ G(0)\ H(0)]\Delta z_t^* &= -[M\ K]\beta^{*\prime} z_{t-1}^* \\ &\quad +[\widetilde{J}(L)\ \widetilde{G}(L)\ \widetilde{H}(L)]\Delta z_{t-1}^* + e_t \end{aligned} \tag{17}$$

3. Complete system

$$\begin{aligned} \begin{bmatrix} J(0) & G(0) & H(0) \\ F(0) & B(0) & C(0) \\ 0 & 0 & I \end{bmatrix} \Delta z_t^* &= -\begin{bmatrix} M & K \\ F(1) & I \\ 0 & 0 \end{bmatrix} \beta^{*\prime} z_{t-1}^* \\ &\quad - \begin{bmatrix} \widetilde{J}(L) & \widetilde{G}(L) & \widetilde{H}(L) \\ \widetilde{F}(L) & \widetilde{B}(L) & \widetilde{C}(L) \\ \widetilde{E}_s(L) & \widetilde{E}_y(L) & \widetilde{D}(L) \end{bmatrix} \Delta z_{t-1}^* \\ &\quad + \begin{bmatrix} \xi_t \\ e_t \\ \varepsilon_t \end{bmatrix} \end{aligned} \tag{18}$$

Can now re-write as the CVAR

$$\Delta z_t^* = -\alpha^* \beta^{*\prime} z_{t-1}^* + A^*(L) \Delta z_{t-1}^* + v_t^* \tag{19}$$

where

$$\alpha^* = \begin{bmatrix} B^*(0)^{-1} \begin{bmatrix} M & K \\ F(1) & I \end{bmatrix} \\ 0 \end{bmatrix}$$

$$B^*(0) = \begin{bmatrix} J(0) & G(0) \\ F(0) & B(0) \end{bmatrix}$$

and

$$v_t^* = \begin{bmatrix} B^*(0)^{-1} & -B^*(0)^{-1} \begin{bmatrix} H(0) \\ C(0) \end{bmatrix} \\ 0 & I \end{bmatrix} \begin{bmatrix} \xi_t \\ e_t \\ \varepsilon_t \end{bmatrix}$$

Note:

- Equation (19) will NOT be a standard cointegrated VAR as it contains equations for the stationary as well as the non-stationary variables

.

- β^* is NOT the matrix of cointegrating vectors.

Estimation

- Can't use the standard Johansen maximum likelihood estimator as there are I(0) *endogenous* variables

- The Johansen estimator could be used on the sub-system of non-stationary variables if the stationary variables were also included in these equations.

- But unless ξ_t is uncorrelated with e_t the resulting estimators of the coefficients of all the stationary variables, the first difference terms and the cointegrating residuals will not be efficient. The sub-system for the stationary variables would then need to be estimated separately by LIML.

- Constrained maximum likelihood on the whole system would be efficient.

- The problem with both of these approaches is that, as before, it would not be possible to recover estimates of the structural disturbances ξ_t and e_t and the stochastic trends ε_t from estimates of v_t^*.

1. Estimating equations for y_t

Write as the VECM

$$\begin{aligned}\Delta y_t &= -B(0)^{-1}F(0)\Delta s_t - B(0)^{-1}C(0)\Delta x_t - B(0)^{-1}F(1)s_{t-1} - B(0)^{-1}w_{t-1} \\ &\quad +B(0)^{-1}[\widetilde{F(L)}\ \widetilde{B}(L)\ \widetilde{C}(L)]\Delta z_{t-1}^* + u_t \end{aligned} \tag{20}$$

- assume as before that a prior estimate of β is available

- if a prior estimate is also available for $F(1)$

- then have an estimate of $\tilde{\beta}$ and can treat w^*_{t-1} as observed in

$$\begin{aligned}\Delta y_t &= -B(0)^{-1}F(0)\Delta s_t - B(0)^{-1}C(0)\Delta x_t - B(0)^{-1}w^*_{t-1} \\ &\quad + B(0)^{-1}[\widetilde{F(L)}\ \widetilde{B}(L)\ \widetilde{C}(L)]\Delta z^*_{t-1} + u_t \qquad (21)\end{aligned}$$

- The choice of estimator will depend on whether e_t and ξ_t are correlated

- if they are uncorrelated, Δs_t will be uncorrelated with u_t and OLS can be used

- if they are correlated, Δs_t will need to be instrumented and a sub-system estimator used

2. Estimation of the equations for the stationary variables

- Need identifying restrictions.

  - In this paper the aim is to achieve identification through the use of long-run restrictions, not restrictions on either the covariance matrix of the disturbances, or on the short-run dynamics

  - implies need restrictions on K and M

  - the leading diagonal of M is normalized

  - remaining restrictions on K and M will need to satisfy the order condition for identification that each equation requires as many (long-run) linear restrictions as the number of stationary variables less one

  -NB if there is only one stationary variable, then no restrictions are required on the long-run solution

- the short-run coefficient matrices are all unrestricted, including that of Δs_t

- Equation (16) is re-written as

$$\begin{aligned} Ms_t &= (M - J(0))\Delta s_t - G(L)\Delta y_t - H(L)\Delta x_t \\ &\quad -Kw_{t-1} + \widetilde{J}(L)\Delta s_{t-1} + \xi_t \end{aligned} \tag{22}$$

with $M_{ii} = 1$, and M and K restricted to achieve identification

- Will need to use a sub-system estimator as instrumental variables are needed for each included s_t, for Δs_t and for Δy_t

- if e_t and ξ_t are uncorrelated, Δy_t can be treated as exogenous

- the explanatory variables of equation (21) are the

appropriate instruments together with s_{t-1}

3. Exogenous variables

The equations for the exogenous variables can still be estimated by OLS.

Additional notes

- If e_t and ξ_t are correlated it would be more efficient to estimate equations (20) and (22) jointly by a simultaneous equations estimator

- If the simultaneous estimation of β is preferred to using a prior estimate, the systems estimator will need to be further constrained.

- If $F(1)$, M and K are known then can use OLS

- Can also use prior estimates

- Basic idea is to write the equation in terms of $y_t^* = (s_t' \;\; y_t')'$ instead of y_t.

- Now estimate

$$\begin{aligned} \Delta y_t^* &= -B^*(0)^{-1}C^*(0)\Delta x_t - B^*(0)^{-1}w_{t-1}^{\#} \\ &\quad +B^*(0)^{-1}[\widetilde{B}^*(L)\,\widetilde{C}^*(L)]\Delta z_{t-1}^* + u_t^* \end{aligned} \tag{23}$$

where $w_t^{\#} = Q\beta^{*\prime} z_t^*$ is treated as observed,

$$u_t^* = B^*(0)^{-1}\begin{bmatrix} \xi_t \\ e_t \end{bmatrix}$$

$$C^*(0) = \begin{bmatrix} H(0) \\ C(0) \end{bmatrix}$$

$$Q = \begin{bmatrix} M & K \\ F(1) & I \end{bmatrix}$$

and the coefficient matrices of Δz^*_{t-1} are unrestricted.

4. Impulse response functions

Ignoring any correlation among the shocks, these can be obtained from the companion form of the complete model

$$\begin{bmatrix} z_t^* \\ z_{t-1}^* \\ z_{t-2}^* \\ . \\ z_{t-k}^* \end{bmatrix} = \begin{bmatrix} I - \alpha^* \beta^{*\prime} + A^*(0) & -A^*(1) + A^*(2) & ... & -A^*(k) \\ I & 0 & ... & 0 \\ 0 & I & ... & 0 \\ . & . & ... & 0 \\ 0 & 0 & .. & 0 \end{bmatrix} \begin{bmatrix} z_{t-1}^* \\ z_{t-2}^* \\ z_{t-3}^* \\ . \\ z_{t-k-1}^* \end{bmatrix}$$

$$+ \begin{bmatrix} B^*(0)^{-1} & -B^*(0)^{-1} \begin{bmatrix} H(0) \\ C(0) \end{bmatrix} \\ 0 & I \\ \multicolumn{2}{c}{0} \\ \multicolumn{2}{c}{0} \\ \multicolumn{2}{c}{.} \\ \multicolumn{2}{c}{0} \end{bmatrix} \begin{bmatrix} \xi_t \\ e_t \\ \varepsilon_t \end{bmatrix} \tag{24}$$

The remaining calculations are similar to those described earlier.

Example (Wickens and Motto, 2001)

Small stylised model of the US economy estimated from quarterly data over the period 1960.1-1998.4

$$i_t = \rho + \Delta p_t + \xi_t$$

$$m_t - p_t = \upsilon + y_t - \lambda i_t + e_t$$

$$\Delta y_t = \gamma + \alpha \Delta y_{t-1} + \varepsilon_{yt}$$

$$\Delta m_t = \mu + \theta \Delta m_{t-1} + \varepsilon_{mt}$$

where

i=federal funds rate - I(0)

p=log GDP price deflator - I(1)

y=log GDP - I(1)

m=log M1 - I(1)

Long-run structure: $m_t - p_t - y_t$ is I(0)

$$
\begin{bmatrix} 1 & -1 & 0 & 0 \\ -\lambda & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \Delta z_t^* = \begin{bmatrix} \rho \\ -\upsilon \\ \gamma \\ \mu \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\lambda & 1 & 1 & -1 \end{bmatrix} z_{t-1}^*
$$

$$
+ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \mu \end{bmatrix} \Delta z_{t-1}^* + v_t
$$

$z_t^{*\prime} = \{i_t \; p_t \; y_t \; m_t\}$ and $v_t' = \{\xi_t \; -e_t \; \varepsilon_{yt} \; \varepsilon_{mt}\}$

- Only one coefficient, λ, requires prior estimation
- A standard Johansen test does not reject this long-run structure.

- Equations for the endogenous variables

$$\begin{bmatrix} \Delta i_t \\ \Delta p_t \end{bmatrix} = B^*(0)^{-1} \begin{bmatrix} \rho \\ \upsilon \end{bmatrix} - B^*(0)^{-1} \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \Delta m_t \end{bmatrix}$$
$$+B^*(0)^{-1} w^{\#}_{t-1} + B^*(0)^{-1} \begin{bmatrix} \xi_t \\ -e_t \end{bmatrix}$$

where

$$B^*(0) = \begin{bmatrix} 1 & -1 \\ -\lambda & 1 \end{bmatrix}$$

$$w^{\#}_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\lambda & 1 & 1 & -1 \end{bmatrix} z^*_t$$

- In the model actually estimated

  - its short-run dynamics were generalized

  - $B^*(0)$ was not restricted

  - two lags in Δz^*_t were added

- An estimate of $\lambda = 8.9$ was obtained from equation (20)

7. Rational Expectations Models

- It is increasingly common in modern monetary policy analysis to use small structural models instead of VAR models. These are often based on DSGE models and have New Keynesian price equations. A feature of these models is the presence of variables that are the rational expectation of future variables using current information. This is unlike the previous use of RE where the rational expectations were usually of *current* variables and based on past information.

- The assumption of rational expectations has proved highly controversial due to the extremely strong informational requirements they entail. Arguably, however, this criticism has been overdone. The assumption of RE is better thought of as analagous to that of perfect competition: it is a limiting but useful benchmark.

- In practice, of course, it is highly improbable that expectations can be fully rational. But the aim of these models is more to capture forward-lookingness in decision making than strong rationality. Expectations can be based on the organizing structure of the model and the optimal use (in a statistical sense) of current and past data. They are therefore better thought of as consistent than rational.

- At this stage we consider the general problem of analysing RE models. In later lectures we shall look at particular models.

- We shall examine

(i) the solution of forward-looking RE models

(ii) the implications of such structural models for VAR analysis

(iii) the estimation of these models. The problem of estimating RE models where the expectation is of a current variable and the information set is dated in the past is well-known (see Wallis (1982) and Wickens (1982) and Pesaran (1987)).

First we consider single equation models to establish the issues, then we turn to simultaneous equation models.

Single equation forward-looking RE models.

Consider the model

$$
\begin{aligned}
y_t &= \alpha E_t[y_{t+1}] + \delta y_{t-1} + \beta x_t + e_t \\
x_t &= \theta_1 x_{t-1} + \theta_2 x_{t-2} + \varepsilon_t
\end{aligned}
$$

where e_t and ε_t are assumed to be serially and mutually uncorrelated.

Using the lag operator L which enables us to write $x_{t-n} = L^n x_t$ and $E_t[x_{t+n}] = L^{-n} x_t$, we can express the difference equation as

$$(-\alpha L^{-1} + 1 - \delta L)x_t = \beta x_t + e_t$$

This has the characteristic equation

$$-\alpha L^{-1} + 1 - \delta L = 0$$

implying that

$$-\delta L^{-1}[L^2 - \frac{1}{\delta}L + \frac{\alpha}{\delta}] = 0$$

The solution of

$$L^2 - \frac{1}{\delta}L + \frac{\alpha}{\delta} = 0$$

can be written

$$(L-\gamma_1)(L-\gamma_2)=0$$

If the solution has saddlepath dynamics then one root is stable and the other is unstable. Let γ_1 be the stable root and γ_2 the unstable root, then $|\gamma_1|<1$ and $|\gamma_2|>1$. In this case

$$(L-\gamma_1)(L-\gamma_2)|_{L=1}<0$$

Thus we can quickly check whether the difference equation has saddlepath dynamics from

$$\begin{aligned}(L-\gamma_1)(L-\gamma_2)|_{L=1} &= (L^2-\frac{1}{\delta}L+\frac{\alpha}{\delta})|_{L=1}\\ &= 1-\frac{1}{\delta}+\frac{\alpha}{\delta}<0\end{aligned}$$

We can now re-write the difference equation as

$$-\delta L^{-1}(L-\gamma_1)(L-\gamma_2)\,y_t=\beta x_t+e_t$$

or as

$$(1-\gamma_1 L^{-1})\left(1-\gamma_2^{-1}L\right)y_t=\frac{\beta}{\delta\gamma_2}x_t+\frac{1}{\delta}e_t$$

Hence

$$\begin{aligned}y_t &= \frac{\beta}{\delta\gamma_2}\sum_{s=0}^{\infty}\gamma_1^s L^{-s}x_t+\frac{1}{\gamma_2}Ly_t+\frac{1}{\delta}e_t\\ &= \frac{\beta}{\delta\gamma_2}\sum_{s=0}^{\infty}\gamma_1^s E_t[x_{t+s}]+\frac{1}{\gamma_2}y_{t-1}+\frac{1}{\delta}e_t\end{aligned}$$

y_t can also be written as the partial adjustment model

$$\Delta y_t=(1-\frac{1}{\gamma_2})[y_t^*-y_{t-1}]+\frac{1}{\delta}e_t$$

where y_t^*, the long-run solution for y_t, is

$$y_t^*=\frac{\beta}{\delta(\gamma_2-1)}\sum_{s=0}^{\infty}\gamma_1^s E_t[x_{t+s}]$$

We now need to replace $E_t[x_{t+s}]$ using the equation for x_t. If x_t is an AR(2) then we can write

$$E_t[x_{t+s}] = \phi_{1s}x_t + \phi_{2s}x_{t-1}$$

The solution for y_t then takes the form

$$y_t = \lambda_1 x_t + \lambda_2 x_{t-1} + \lambda_3 y_{t-1} + u_t$$

where $u_t = \frac{1}{\delta}e_t$.

We have therefore transformed the RE model to a standard dynamic structural model without RE variables of the sort that we considered previously.

We can use the solution to solve for $E_t[y_{t+1}]$. This is

$$\begin{aligned} E_t[y_{t+1}] &= \lambda_1 E_t[x_{t+1}] + \lambda_2 x_t + \lambda_3 y_t \\ &= (\lambda_1\theta_1 + \lambda_2)x_t + \lambda_1\theta_2 x_{t-1} + \lambda_3 y_t \end{aligned}$$

Substituting this into the original model gives

$$\begin{aligned} y_t &= \alpha[(\lambda_1\theta_1 + \lambda_2)x_t + \lambda_1\theta_2 x_{t-1} + \lambda_3 y_t] + \delta y_{t-1} + \beta x_t + e_t \\ &= \frac{\alpha(\lambda_1\theta_1 + \lambda_2) + \beta}{1 - \alpha\lambda_3}x_t + \frac{a\lambda_1\theta_2}{1 - \alpha\lambda_3}x_{t-1} + \frac{\delta}{1 - \alpha\lambda_3}y_{t-1} + \frac{1}{1 - \alpha\lambda_3}e_t \end{aligned}$$

If this is solved for y_t then we simply return to the previous solution.

We note two further things about the solution:

1. Identification:

   The presence of x_{t-2} in the equation for x_t is crucial for identification. In its absence $E_t[y_{t+1}]$ would be determined just by x_t and y_t, implying that $E_t[y_{t+1}]$ would be perfectly correlated with the other variables in the model for y_t. Estimation of α would then be impossible.

2. Estimation:

   The presence of y_t in $E_t[y_{t+1}]$ implies that it will correlated with e_t. Thus one can't simply substitute a forecast for $E_t[y_{t+1}]$ based on regressing y_{t+1} on x_t, x_{t-1} and y_t. It will be necessary to instrument the forecast too. Valid instruments are x_t, x_{t-1} and y_{t-1}.

Thus, if the forecast for y_{t+1} is $\hat{y}_{t+1}$we substitute this into the original equation to obtain

$$\begin{aligned} y_t &= \alpha\hat{y}_{t+1} + \delta y_{t-1} + \beta x_t + v_t \\ v_t &= e_t + \alpha(E_t[y_{t+1}] - \hat{y}_{t+1}) \end{aligned}$$

which can be consistently estimated by IV (GMM) using (x_t, x_{t-1}, y_{t-1}) as instruments.

Simultaneous systems with future expectations

It is possible to generalise the methodology to the case of a system of equations

$$\begin{aligned} FE_t[y_{t+1}] + B(L)y_t + C(L)x_t &= e_t \\ D(L)x_t &= \varepsilon_t \end{aligned}$$

Using the lag operator the model can be written

$$A(L)y_t = -C(L)x_t + e_t$$

where

$$A(L) = B(L) + C(L) + FL^{-1}$$

We denote the roots of

$$|A(L)| = 0$$

as λ_i $(i = 1, .., p)$ and $\gamma_j = (j = 1, ..., q)$ where $p + q = n$, λ_i are the unstable roots and γ_j are the stable roots. Thus $|\lambda_i| \leq 1$ and $|\gamma_j| > 1$ and

$$\begin{aligned} |A(L)| &= aL^{-p} \textstyle\prod_{i=1}^{p}(L - \lambda_i) \prod_{i=1}^{q}(L - \gamma_j) \\ &= b \textstyle\prod_{i=1}^{p}(1 - \lambda_i L^{-1}) \prod_{i=1}^{q}(1 - \gamma_j^{-1} L) \end{aligned}$$

where

$$b = a \textstyle\prod_{i=1}^{q}(-\gamma_j)$$

Hence we can write

$$\begin{aligned} A(L) &= \frac{adj A(L)}{b \prod_{i=1}^{p}(1 - \lambda_i L^{-1}) \prod_{i=1}^{q}(1 - \gamma_j^{-1} L)} \\ &= \frac{adj A(L)}{b \prod_{i=1}^{q}(1 - \gamma_j^{-1} L)} \textstyle\sum_{i=1}^{p} \frac{b_i}{1 - \lambda_i L^{-1}} \end{aligned}$$

Thus the model can be written

$$\textstyle\prod_{i=1}^{q}(1 - \gamma_j^{-1} L) y_t = \frac{adj A(L)}{b} \sum_{i=1}^{p} \frac{b_i}{1 - \lambda_i L^{-1}} [-C(L)x_t + e_t]$$

or

$$y_t = G(L)y_{t-1} + \frac{1}{b} \textstyle\sum_{i=1}^{p} \sum_{s=0}^{\infty} b_i \lambda_i^s L^s adj A(L)[-C(L)x_t + e_t]$$

If $D(L)$ is of order n then

$$E_t[x_{t+s}] = D_s(L)x_t$$

and so

$$y_t = G(L)y_{t-1} + H(L)x_t + K(L)e_t$$

Hence, $E_t[y_{t+1}]$ is obtained from

$$E_t[y_{t+1}] = G(L)y_t + H^*(L)x_t + K^*(L)e_t$$

We note that due to the presence of $K^*(L)e_t$, in general this is not a VAR for y_{t+1}, but a VARMA.

Substituting $E_t[y_{t+1}]$ into the original structural model gives

$$B(L)y_t + C(L)x_t + F[G(L)y_t + H^*(L)x_t + K^*(L)e_t] = e_t$$

implying that the solution can be written

$$B^*(L)y_t + C^*(L)x_t = M(L)e_t$$

To estimate the model we use

$$B(L)y_t + C(L)x_t + FE_t[y_{t+1}] = e_t$$

replacing $E_t[y_{t+1}]$ with the predicted value from the VARMA and instrumenting it due to the presence of y_t and e_t.

We have shown, therefore, that the solution is a SEM *but* with a VMA error structure.

Pre-multiplying by $M(L)^{-1}$ and then normalising would produce a VAR of infinite order.

If all of the variables are I(0), pre-multiplying by $B^*(L)^{-1}$ would yield an infinite order VMA.

If some of the variables are I(1) and there is cointegration then the Granger Representation theorem would need to be used to obtain the VMA.

It follows that it is possible to use the VAR models discussed previously even when the underlying structural model has forward-looking RE variables.

Impluse response function analysis

Consider the two equation model again. This is

$$\begin{aligned} y_t &= \alpha E_t[y_{t+1}] + \delta y_{t-1} + \beta x_t + e_t \\ x_t &= \theta_1 x_{t-1} + \theta_2 x_{t-2} + \varepsilon_t \end{aligned}$$

and has the solution for y_t:

$$y_t = \lambda_1 x_t + \lambda_2 x_{t-1} + \lambda_3 y_{t-1} + \lambda_4 e_t$$

Thus the VAR in (y_t, x_t) is

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} 1 & -\lambda_1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_3 & \lambda_2 \\ \theta_1 & \theta_2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & -\lambda_1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_4 e_t \\ \varepsilon_t \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_3 + \lambda_1\theta_1 & \lambda_2 + \lambda_1\theta_2 \\ \theta_1 & \theta_2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \lambda_4 & \lambda_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e_t \\ \varepsilon_t \end{bmatrix}$$

The error term of the VAR therefore has a recursive structure, but is upper triangular and not lower triangular, implying that causality goes from x_t to y_t, and not vice-versa. If a Cholseki decomposition is employed, the ordering of the variables in the VAR therefore needs reversing.

According to the structural solution, y_t responds instantaneously to changes in x_t; part of this is transmitted directly, and part through $E_t[y_{t+1}]$. The latter is picking up the fact that current changes in x_t have a predictable effect on future values of x_t, and these affect $E_t[y_{t+1}]$, and hence y_t. As a result, the VAR accurately captures impact effects on y_t: the impact effect from the solution is λ_1, as is the impact of ε_t on y_t in the VAR. Lagged effects will also be captured correctly by the VAR.

VAR analysis and the Lucas critique

If the Lucas critique dealt a severe blow to the SEM of the Cowles Commission, the blow to the VAR was even more severe.

To see why, suppose that one of the equations is a policy rule. For example, in the single equation model, x_t might be a policy instrument and the equation for x_t a policy rule. Suppose that at some point in the sample the rule was changed.

Consider the two equation model again. A change in the x_t equation (that was known) would affect y_t directly, as before, and through $E_t[y_{t+1}]$. The direct effect could not be distinguished from a shock to ε_t, but the change in the x_t equation would cause the relation between $E_t[y_{t+1}]$ and x_t forecasts of future values of x_t to change. This would cause the forecast for y_{t+1} to be revised. All of this would occur the moment the change is anticipated and, due the forward-looking expectations, this could be prior to the change occurring. As result, the backward-looking solution for y_t would change. This implies that the VAR would become structurally unstable.

In contrast, the original structural equation for y_t would remain unchanged. As a consequence, when it is thought that the underlying structural equations involve expectations variables, it might be better to estimate the structural model than a VAR and to carry out impulse response function analysis on a model formed from the structural equation and whatever equation for x_t is expected to be in force.