Macromodelling and Systems Estimation

- There has been considerable debate over the years on how the macroeconomy should be modelled....
- With structural models when you estimate systems of equations it is important to determine whether they are identified, whether you can move from the reduced form model to the underlying structural model.
- This approach requires that you establish prior to any econometric investigation which variables are endogeous and which are exogenous. Taking a system

$$y_t'\Gamma + x_t'B = u_t'$$

  where the dimensions are

$$\begin{array}{ccccc} y_t' & \Gamma & x_t' & B & u_t' \\ 1xM & MxM & 1xH & HxM & 1xM \end{array}$$

- The reduced form of the system is

$$y_t' = -x_t'\Pi + u_t'$$

  with

$$\Pi = B\Gamma^{-1}$$

$$v_t = u'\Gamma^{-1}$$

- The Cowles Commission researchers argued that the structural equation system was the appropriate means of modelling the economy.
- In this the role of economic theory was explicit, it provided the endogenous variables to be determined and their determinants.
- Which variables were endogenous and which exogenous thus played the role of auxiliary hypotheses as did the restrictions on each equation specified.
- The spirit was one of verification of theoretical models.
- Koopmans identified exogeneity
  - variables which influence the endogenous variables but are not influenced thereby. This 'causal principle' is restrictive.
  - generalised predetermined variables
- Key features of the Cowles Commission approach
  - restrictions from economic theory
  - presumtions about the direction of causation
  - exogeneity
- Data assumed generated by a simultaneous equation system with endogenous/exogenous dichotomy the causality running from the exogenous variables to the endogenous variables. These are given a priori and are untestable.
- This methodology underlies the well known Klein-Goldberger model, the first to be used for exante forecasting
- This led to a reallocation of effort away from economic theory and the development of statistically efficient methods of estimation to the regular updating of models and more practical considerations (see discussion in Berndt)
- Systems estimation methods: FIML versus reality of large macroeconometric models

Objections to this approach

- Rational expectations: understanding processes/model including government policy do 'structural' parameter might change when government policy changes

- Lucas critique: no reason structure of the economic relations should be invariant under policy intervention

Responses

- Make models with expectations and forecasts generated that are consistent. But still identification problems
- Sims: reject whole approach
  - theoretical models 'incredible'
  - should include all variable in all structural equations as a priori theory can lead to identification of which is which
  - virtually all variables are endogenous so underidentifiaction is rife
  - anyway, structural identification is not necessary for forecasting and policy analysis
- So proposes
  - no apriori endogenous/exogenous divisions
  - no imposition of zero restrictions
  - no strict (prior to modelling) economic theory within which the model is grounded
- Criticised as 'atheoretical'

VARs

- This approach implies setting up a Vector Autoregression Model (VAR):

$$Z_t = \sum_{i=1}^{k} A_i Z_{t-i} + \varepsilon_t$$

$$Z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}$$

  with y_t the current endogenous and x_t the current exogenous variables and ε_t a column vector of random errors, usually assumed to be contemporaneously correlated, but not autocorrelated. This implies a non-diagonal covariance matrix
- Often the model is completed by the addition of deterministic components, intercept, trend seasonal dummies
- Can clearly see the advantage of this approach for forecasting.
  - Taking a single structural equation and attempting to forecast the endogenous variable (defined from theory):

$$y_t = \alpha x_t + \varepsilon_t$$

$$y_t^F = \alpha x_t$$

  - so you need to forecast x_t.
- In a VAR you don't need to worry about theory and you don't need forecasts of the exogenous variables.
  - In the first period

$$Z_t^F = \sum_{i=1}^{k} A_i Z_{T-i+1}$$

- Then

$$Z^F_{T+j} = \sum_{i=1}^{k} A_i Z^F_{T+j-i}$$

- so it is recursive.

In Practice

- In practice have to impose some restrictions on VARs
  - number of variables included
  - lag lengths
- If there are 6 variables and 5 lags this means there are at least 30 regressors in each equation.
- An intuitive way to choose lag length is to try to get residuals without serial correlation.
- Note that ε_t can have non zero covariances, so can have a 'structural alternative' consistent with a particular economic theory.
- To make interpretations more straightforward it is common to transform the model into one with 'othogonal\innovations', so that the error terms are no longer contemporaneously uncorrelated and you end up with a scalar variance covariance matrix.
- So take

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\begin{bmatrix} x_{t-2} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

$$E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0$$

$$E(\varepsilon_{1t}^2) = \sigma_{11}; E(\varepsilon_{2t}^2) = \sigma_{22}; E(\varepsilon_{1t}\varepsilon_{1t}) = \sigma_{12}$$

- multiply the first row by

$$\delta = \frac{\sigma_{12}}{\sigma_{11}}$$

  and subtract the result form the second row.
  This will give contemporaneously uncorrelated errors.

$$\begin{bmatrix} x_t \\ y_t - \delta x_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 - \delta a_1 & d_1 - \delta b_1 \end{bmatrix}\begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix}$$

$$+ \begin{bmatrix} a_2 & b_2 \\ c_2 - \delta a_2 & d_2 - \delta b_2 \end{bmatrix}\begin{bmatrix} x_{t-2} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} - \delta\varepsilon_{1t} \end{bmatrix}$$

  so

$$\begin{aligned} &E(\varepsilon_{1t}(\varepsilon_{2t} - \delta\varepsilon_{1t})) \\ &= E((\varepsilon_{1t}\varepsilon_{2t}) - \frac{\sigma_{12}}{\sigma_{11}})E(\varepsilon_{1t}^2) \\ &= \sigma_{12} - \sigma_{12} = 0 \end{aligned}$$

  so uncorrelated.
- The idea of this is to allow the equations to be used separately for policy analysis
  - in the sense of looking at the impact of a known innovation (random shock) to the system
  - problem is that the results may be sensitive to the ordering of the VAR equations

- in practice may be able to decide on the equation ordering by applying causality tests

- Pagan describes the process of setting up the VAR:
  - Transform the data -need stationary series for the techniques used
  - Choose lag length and the number of variables compatible with the number of observations
  - Try to simplify by reducing the lag length, imposing arbitrary restrictions
  - Use orthogonal innovation representation to address questions.

1.

Criticism of VARS

Darnell and Evans provide a good critique of the approach with references

1. Criticism of the transformations undertaken to yield causal chains -orthogonal innovations
2. Requirement of stationarity means transform data at outset, but this is neither easy nor without risk of misleading e.g. differencing can be over utilized
3. Selecting the included variables -results can be sensitive to inclusion of another variable and may be sensitive to choice of lag length
4. Sims tended to prefer models which are triangularised but symmetric (all the same lagged values in each equation). Cooley an LeRoy argue should justify as these are substantial restrictions.
5. VAR has no role it the hypothetico-deductive method of economics and can give little to our understanding of economic phenomena. Blaug described as 'mindless instrumentalism'.

VARs, Causality and cointegration

VARs

- The generalisation of an AR2 to a vector is the VAR2:

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \varepsilon_t$$

where y_t is now a $m \times 1$ vector, A_0 a $m \times 1$ vector, A_1 and A_2 are $m \times m$ matrices.

- For $m = 2, ... y_t = (y_{1t}, y_{2t})'$ the VAR is:

$$y_{1t} = a_1^0 + a_{11}^1 y_{1t-1} + a_{12}^1 y_{2t-1} + a_{11}^2 y_{1t-2} + a_{12}^2 y_{2t-2} + \varepsilon_{1t},$$
$$y_{2t} = a_2^0 + a_{21}^1 y_{1t-1} + a_{22}^1 y_{2t-1} + a_{21}^2 y_{1t-2} + a_{22}^2 y_{2t-2} + \varepsilon_{2t}.$$

- Each equation of the VAR can be estimated consistently by OLS and the covariance matrix can be estimated from the OLS residuals.

Granger Causality

- A variable y_{2t} is said to Granger cause y_{1t} if knowing current values of y_2 helps you to predict future values of y_1 equivalently, current y_1 is explained by past y_2. In this case, y_2 is Granger causal with respect to y_1 if either a_{12}^1 or a_{12}^2 are non zero.
- Can test that they are both zero with a standard F test of linear restrictions. The restricted model just excludes y_{2t-1} and y_{2t-2} from the equation for y_{1t}.
- Granger causality is rarely the same as economic causality.
- More lags can be included and you can decide the appropriate lag length by Likelihood Ratio tests or model selection criteria like the AIC or SBC.
- Make sure that you use the same sample for the restricted and unrestricted model; i.e. do not use the extra observation that becomes available when you shorten the lag length.
- If the lag lenth is p, each equation of the VAR has $1 + mp$ parameters. This can get large, 4 lags in a 4 variable VAR gives 17 parameters in each equation.
- Be careful about degrees of freedom.

A pth order VAR

- Taking the general VAR

$$y_t = A_0 + \sum_{i=1}^{p} A_i y_{t-i} + \varepsilon_t$$

- This is stable if all the roots of the determinantal equation $| \ I - A_1 z - A_2 z^2 - \ldots - A_p z^p \ | = 0$ lie outside the unit circle.
- If we reparameterise the VAR2:

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \varepsilon_t$$

as:

$$y_t - y_{t-1} = A_0 - (I - A_1 - A_2) y_{t-1} - A_2 (y_{t-1} - y_{t-2}) + \varepsilon_t$$
$$\Delta y_t = A_0 - \Pi y_{t-1} + \Gamma \Delta y_{t-1} + \varepsilon_t$$

- and the VARp as:

$$\Delta y_t = A_0 - \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t.$$

- Which is called the ’cointegrating transformation’
- Notice that this is the vector equivalent of the ADF above for testing for unit roots. Express the Γ_i in terms of the A_i.

Cointegration

- Taking this form of the VAR

$$\Delta y_t = A_0 - \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t$$

- If all the variables, the m elements of y_t, are I(0), Π is a full rank matrix.
- If all the variables are I(1) and not cointegrated, $\Pi = 0$, and a VAR in first differences is appropriate.
- If the variables are I(1) and cointegrated, with r cointegrating vectors, then there are r cointegrating relations, combinations of y_t that are I(0),

$$z_t = \beta' y_t$$

where z_t is a $r \times 1$ vector and β' is a $r \times m$ matrix.
- Then we can write the model as:

$$\Delta y_t = A_0 - \alpha z_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t,$$

in which the I(0) dependent variable is only explained by I(0) variables and α is a $m \times r$ matrix of ‘adjustment coefficients’ which measure how the deviations from equilibrium (the r I(0) variables z_{t-1}) feed back on the changes.
- This is called the ’cointegrating transformation’

- It can also be written:

$$\Delta y_t = A_0 - \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t,$$

  so $\Pi = \alpha\beta'$ has rank $r < m$ if there are r cointegrating vectors.
- If there are $r < m$ cointegrating vectors, then y_t will also be determined by $m - r$ stochastic trends.
- If there is cointegration, some of the α must be non-zero, there must be some feedback on the y_t to keep them from diverging, i.e. there must be some Granger causality in the system.
- If there are r cointegrating vectors and Π has rank r, it will have r non-zero eigenvalues
- Johansen provided a way of estimating the eigenvalues and two tests for determining how many of the eigenvalues are different from zero.
- These allow us to determine r, though the two tests may give different answers.
- The Johansen estimates of the cointegrating vectors β are the associated eigenvectors but there is an ‘identification’ problem, since the α and β are not uniquely determined.
- We can always choose a non-singular $r \times r$ matrix P such that $(\alpha P)(P^{-1}\beta) = \Pi$ and the new estimates $\alpha^* = (\alpha P)$ and $\beta^* = (P^{-1}\beta)$ would be equivalent, though they might have very different economic interpretations.
- Put differently, if $z_{t-1} = \beta' y_{t-1}$ are I(0) so are $z^*_{t-1} = P^{-1}\beta' y_{t-1}$, since any linear combination of I(0) variables is I(0).
- We need to choose the appropriate P matrix to allow us to interpret the estimates.
- This requires r^2 restrictions, r on each cointegrating vector.
- One of these is provided by normalisation, we set the coefficient of the ‘dependent variable’ to unity, so if $r = 1$ this is straightforward; but if there is more than one cointegrating vector it requires prior economic assumptions.
- The Johansen identification assumption, that the β are eigenvectors with unit length and orthogonal, do not allow an economic interpretation.
- The EViews identifying assumptions (that the first *rxr* block of the β matrix is the identity matrix) are only rarely appropriate.
- Microfit allows you to specify the r^2 just identifying restrictions and test any extra ‘over-identifying’ restrictions.

- As we saw above with the Dickey Fuller regression, there is also a problem with the treatment of the deterministic elements.
- If we have a linear trend in the VAR, and do not restrict the trends, the variables will be determined by $m - r$ quadratic trends.
- To avoid this (economic variables tend to show linear not quadratic trends), we enter the trends in the cointegrating vectors.
- Most programs give you a choice of how you enter trends and intercepts;
- unrestricted intercepts and restricted trends is a good choice for trended economic data.
- Takingthe VAR

$$Z_t = \alpha_0 + \alpha_{1t} + \sum_{i=1}^{k} \Phi_i Z_{t-i} + \psi_y w_t + u_t$$

and corresponding cointegrating VAR

$$\Delta y_t = \alpha_{0y} + \alpha_{0x} t - \Pi_y z_{t-1} + \sum_{i=1}^{p-1} \Gamma_{iy} \Delta z_{t-i} + \psi_y w_t + \varepsilon_t$$

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

- y_t is an m_y vector of jointly determined (endogenous) $I(1)$ variables
- x_t is an m_x vector of exogenous $I(1)$ variables
- w_t is a $qx1$ vector of exogenous/deterministic $I(0)$ variables
- allow for feedbacks Δy to Δx but not for levels feedbacks
- assumes the xs are not themselves cointegrated

$$\Delta x_t = \alpha_{0x} + \sum_{i=1}^{p-1} \Gamma_{ix} \Delta z_{t-i} + \psi_x w_t + \epsilon_t$$

- So for:

$$\Delta y_t = \alpha_{0y} + \alpha_{1y} t - \Pi_y z_{t-1} + \sum_{i=1}^{p-1} \Gamma_{iy} \Delta z_{t-i} + \psi_y w_t + \varepsilon_t$$

- Options are:
  - $\alpha_{0y} = \alpha_{1y} = 0$ (no intercepts and no trends)
  - $\alpha_{1y} = 0$ and $\alpha_{0y} = \Pi_y \mu_y$ (restricted intercepts and no trends)
    - $\alpha_{0y} = \Pi_y \mu_y$ meaning the intercepts are part of the cointegrating vectors
  - $\alpha_{1y} = 0$ and $\alpha_{0y} \neq 0$(unrestricted intercepts and no trends)
  - $\alpha_{0y} \neq 0$ and $\alpha_{1y} = \Pi_y \gamma_y$ (unrestricted intercepts and restricted trends)
    - $\alpha_{1y} = \Pi_y \gamma_y$ meaning the trends are part of the cointegrating vectors
  - $\alpha_{0y} \neq 0$ and $\alpha_{1y} \neq 0$ (unrestricted intercepts and unrestricted trends)

Example

- Consider a VAR1 in the logarithms of real money and income, which are both I(1) with a linear trend:

$$y_t = a_{10} + a_{11}y_{t-1} + a_{12}m_{t-1} + \gamma_1 t + \varepsilon_{1t}$$

$$m_t = a_{20} + a_{21}y_{t-1} + a_{22}m_{t-1} + \gamma_2 t + \varepsilon_{2t}$$

  and $z_t = m_t - \beta y_t$ is I(0).
- The cointegrating vector is $(1, -\beta)$ and we have normalised the equation by setting the coefficient of m_t to unity. This just identifies the cointegrating vector for r=1.
- The VECM is:

$$\Delta y_t = a_{10} + (a_{11} - 1)y_{t-1} + a_{12}m_{t-1} + \gamma_1 t + \varepsilon_{1t}$$

$$\Delta m_t = a_{20} + a_{21}y_{t-1} + (a_{22} - 1)m_{t-1} + \gamma_2 t + \varepsilon_{2t}.$$

- Imposing the cointegration restriction, it becomes:

$$\Delta y_t = a_{10} - \alpha_1(m_{t-1} - \beta y_{t-1}) + \gamma_1 t + u_{1t}$$

$$\Delta m_t = a_{20} - \alpha_2(m_{t-1} - \beta y_{t-1}) + \gamma_2 t + u_{2t}$$

  thus

$$\Pi = \begin{bmatrix} \alpha_1 & -\alpha_1\beta \\ \alpha_2 & -\alpha_2\beta \end{bmatrix}$$

  which is clearly of rank 1, since a multiple of the first column equals the second column.
- A natural over-identifying restriction to test in this context would be that $\beta = 1$.
- To restrict the trend we could put it in the cointegrating vector, saving one parameter:

$$\Delta y_t = a_{10} - \alpha_1(m_{t-1} - \beta y_{t-1} + \gamma t) + u_{1t}$$

$$\Delta m_t = a_{10} - \alpha_2(m_{t-1} - \beta y_{t-1} + \gamma t) + u_{2t}$$