Econometrics V Lecture 11
Vector Auto Regressions VARs

- In the two variable case we can let the time path of (y_t) be affected by the current and past realisations of the (z_t) sequence and let the time path of the (z_t) sequence be affected by current and past realisations of the (y_t) sequence.
- In effect treat both variables as being endogenous.
- Consider the following bivariate system

Econometrics V Lecture 11
Vector Auto Regressions VARs

$$y_t = b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt} \qquad (1)$$

$$z_t = b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \qquad (2)$$

Where it is assumed

(1) that both y_t and z_t are stationary

(2) ε_{yt} and εt_z are white noise disturbances with standard deviations of σ_y σ_z respectively, and

(3) (ε_{yt})and (ε_{zt}) are uncorrelated white noise disturbances

Econometrics V Lecture 11
Vector Auto Regressions VARs

- Eqns (1) and (2) constitute a first order vector autoregression (VAR) since the longest lag length is unity.
- The structure of the system incorporates feedback since y_t and z_t are allowed to affect each other.
- Eg. $-b_{12}$ is the contemporaneous effect of a unit change of z_t on y_t, and γ_{21} the effect of a unit change in y_{t-1} on z_t.
- ε_{yt} and ε_{zt} are pure innovations or shocks in y_t and z_t respectively.

Econometrics V Lecture 11
Vector Auto Regressions VARs

- Equations (1) and (2) are not reduced form equations since y_t has a contemporaneous effect on z_t and vice- versa
- Using matrix algebra we can write the system in a compact form.

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

or

$$\mathrm{B}x_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$

Econometrics V Lecture 11
Vector Auto Regressions VARs

- Where

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}. \quad x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

- Premutiplication by B^{-1} allows us to obtain the vector autoregression model (VAR) in its standard form

Econometrics V Lecture 11
Vector Auto Regressions VARs

- $X_t = \boldsymbol{A}_0 + \boldsymbol{A}_1 x_{t-1} + e_t$ (3)
- Where

$\boldsymbol{A}_0 = \boldsymbol{B}^{-1}\Gamma_0$

$\boldsymbol{A}_1 = \boldsymbol{B}^{-1}\Gamma_1$

$e_t = \boldsymbol{B}^{-1}\varepsilon_t$

For notational purposes we can define a_{i0} as element i of the vector $\boldsymbol{A}_0$, and e_{it} as the element i of the vector e_t

- We can use this notation to write (3) in the equivalent form

$$y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{1t} \qquad (4a)$$

$$z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{2t} \qquad (4b)$$

- To distinguish between the systems represented by (1) and (2) versus (4a) and (4b), the first is called a structural VAR or the primitive system, the second is called a VAR in standard form.

Econometrics V Lecture 11
Vector Auto Regressions VARs

- It is important to note that the error terms eit and e_{2t} are composites of the two shocks ε_{yt} and ε_{zt} since et = $\boldsymbol{B}^{-1}\varepsilon_t$
- We can compute e_{1t} and e_{2t} as

$$e_{1t} = (\varepsilon_{yt} - b_{12}\varepsilon_{zt})/(1 - b_{12}b_{21}) \qquad (5)$$

$$e_{2t} = (\varepsilon_{zt} - b_{21}\varepsilon_{yt})/(1 - b_{12}b_{21}) \qquad (6)$$

Econometrics V Lecture 11
Vector Auto Regressions VARs

- Since ε_{yt} and ε_{zt} are white noise processes it follows that both e_{1t} and e_{2t} have zero means, constant variances, and are individually serially uncorrelated.
- Take the expected value of (5)
- $Ee_{it} = E(\varepsilon_{yt} - b_{12}\varepsilon_{zt})/(1-b_{12}b_{21})=0$
- The variance of e_{it} is given by

$$Ee_{1t}^2 = E\left[(\varepsilon_{yt} - b_{12}\varepsilon_{zt})/(1-b_{12}b_{21})\right]^2$$

$$=(\sigma_y^2 + b_{12}^2\sigma_z^2)/(1-b_{12}b_{21})^2$$

Econometrics V Lecture 11
Vector Auto Regressions VARs

- Thus the variance of e_{it} is time independent. The autocovariances of of e_{1t} and e_{1t-i} are

$$Ee_1 e_{1,t-i} = E\left[(\varepsilon_{yt} - b_{12}\varepsilon_{zt})(\varepsilon_{y,t-i} - b_{12}\varepsilon_{z,t-i})\right]/(1-b_{12}b_{21})^2 = 0 \quad for\, i \neq 0$$

- Similarly (6) can be used to demonstrate that e2t is a stationary process with a zero mean, constant variance, and having all autocovariances equal to zero.

Econometrics V Lecture 11
Vector Auto Regressions VARs: More formally

Structural VARs

Structural Representation

Consider the structural VAR (SVAR) model

$$\begin{aligned} y_{1t} &= \gamma_{10} - b_{12}y_{2t} + \gamma_{11}y_{1t-1} + \gamma_{12}y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= \gamma_{20} - b_{21}y_{1t} + \gamma_{21}y_{1t-1} + \gamma_{22}y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

where

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \text{iid}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right).$$

Remarks:

- ε_{1t} and ε_{2t} are called structural errors
- In general, $\text{cov}(y_{2t}, \varepsilon_{1t}) \neq 0$ and $\text{cov}(y_{1t}, \varepsilon_{2t}) \neq 0$
- All variables are endogenous - OLS is not appropriate!

In matrix form, the model becomes

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

or

$$\begin{aligned} \mathbf{B}\mathbf{y}_t &= \gamma_0 + \Gamma_1 \mathbf{y}_{t-1} + \varepsilon_t \\ E[\varepsilon_t \varepsilon_t'] &= \mathbf{D} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \end{aligned}$$

In lag operator notation, the SVAR is

$$\begin{aligned} \mathbf{B}(L)\mathbf{y}_t &= \gamma_0 + \varepsilon_t, \\ \mathbf{B}(L) &= \mathbf{B} - \Gamma_1 L. \end{aligned}$$

Reduced Form Representation

Solve for $\mathbf{y}_t$ in terms of $\mathbf{y}_{t-1}$ and ε_t :

$$\begin{aligned} \mathbf{y}_t &= \mathbf{B}^{-1}\gamma_0 + \mathbf{B}^{-1}\Gamma_1\mathbf{y}_{t-1} + \mathbf{B}^{-1}\varepsilon_t \\ &= \mathbf{a}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{u}_t \\ \mathbf{a}_0 &= \mathbf{B}^{-1}\gamma_0, \mathbf{A}_1 = \mathbf{B}^{-1}\Gamma_1, \mathbf{u}_t = \mathbf{B}^{-1}\varepsilon_t \end{aligned}$$

or

$$\begin{aligned} \mathbf{A}(L)\mathbf{y}_t &= \mathbf{a}_0 + \mathbf{u}_t \\ \mathbf{A}(L) &= \mathbf{I}_2 - \mathbf{A}_1 L \end{aligned}$$

Note that

$$\mathbf{B}^{-1} = \frac{1}{\Delta}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}, \ \Delta = \det(\mathbf{B}) = 1 - b_{12}b_{21}$$

The reduced form errors $\mathbf{u}_t$ are linear combinations of the structural errors ε_t and have covariance matrix

$$\begin{aligned} E[\mathbf{u}_t\mathbf{u}_t'] &= \mathbf{B}^{-1}E[\varepsilon_t\varepsilon_t']\mathbf{B}^{-1\prime} \\ &= \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime} \\ &= \Omega. \end{aligned}$$

Remark: Parameters of RF may be estimated by OLS equation by equation

Identification Issues

Without some restrictions, the parameters in the SVAR are not identified. That is, given values of the reduced form parameters $\mathbf{a}_0, \mathbf{A}_1$ and $\mathbf{\Omega}$, it is not possible to uniquely solve for the structural parameters $\mathbf{B}, \gamma_0, \mathbf{\Gamma}_1$ and $\mathbf{D}$.

- 10 structural parameters and 9 reduced form parameters

- Order condition requires at least 1 restriction on the SVAR parameters

Typical identifying restrictions include

- Zero (exclusion) restrictions on the elements of $\mathbf{B}$; e.g., $b_{12} = 0$.

- Linear restrictions on the elements of $\mathbf{B}$; e.g., $b_{12} + b_{21} = 1$.

MA Representations

Wold representation

Multiplying both sides of reduced form by $\mathrm{A}(L)^{-1} = (\mathrm{I}_2 - \mathrm{A}_1 L)^{-1}$ to give

$$
\begin{aligned}
\mathrm{y}_t &= \mu + \Psi(L)\mathrm{u}_t \\
\Psi(L) &= (\mathrm{I}_2 - \mathrm{A}_1 L)^{-1} \\
&= \sum_{k=0}^{\infty} \Psi_k L^k, \;\; \Psi_0 = \mathrm{I}_2, \Psi_k = \mathrm{A}_1^k \\
\mu &= \mathrm{A}(1)^{-1}\mathrm{a}_0 \\
E[\mathrm{u}_t \mathrm{u}_t'] &= \Omega
\end{aligned}
$$

Remark: Wold representation may be estimated using RF VAR estimates

Structural moving average (SMA) *representation*

SMA of $\mathbf{y}_t$ is based on an infinite moving average of the structural innovations ε_t. Using $\mathbf{u}_t = \mathbf{B}^{-1}\varepsilon_t$ in the Wold form gives

$$
\begin{aligned}
\mathbf{y}_t &= \mu + \Psi(L)\mathbf{B}^{-1}\varepsilon_t \\
&= \mu + \Theta(L)\varepsilon_t \\
\Theta(L) &= \sum_{k=0}^{\infty} \Theta_k L^k \\
&= \Psi(L)\mathbf{B}^{-1} \\
&= \mathbf{B}^{-1} + \Psi_1\mathbf{B}^{-1}L + \cdots
\end{aligned}
$$

That is,

$$
\begin{aligned}
\Theta_k &= \Psi_k\mathbf{B}^{-1} = \mathbf{A}_1^k\mathbf{B}^{-1}, \; k = 0, 1, \ldots . \\
\Theta_0 &= \mathbf{B}^{-1} \neq \mathbf{I}_2
\end{aligned}
$$

Example: SMA for bivariate system

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} \\ \theta_{21}^{(1)} & \theta_{22}^{(1)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix} + \cdots$$

Notes

- $\mathbf{\Theta}_0 = \mathbf{B}^{-1} \neq \mathbf{I}_2$. $\mathbf{\Theta}_0$ captures initial impacts of structural shocks, and determines the contemporaneous correlation between y_{1t} and y_{2t}.

- Elements of the $\mathbf{\Theta}_k$ matrices, $\theta_{ij}^{(k)}$, give the dynamic multipliers or impulse responses of y_{1t} and y_{2t} to changes in the structural errors ε_{1t} and ε_{2t}.

Econometrics V Lecture 11

Impulse Response Functions

Consider the SMA representation at time $t+s$

$$\begin{bmatrix} y_{1t+s} \\ y_{2t+s} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \cdots + \begin{bmatrix} \theta_{11}^{(s)} & \theta_{12}^{(s)} \\ \theta_{21}^{(s)} & \theta_{22}^{(s)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \cdots .$$

The *structural dynamic multipliers* are

$$\frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}} = \theta_{11}^{(s)}, \frac{\partial y_{1t+s}}{\partial \varepsilon_{2t}} = \theta_{12}^{(s)}$$

$$\frac{\partial y_{2t+s}}{\partial \varepsilon_{1t}} = \theta_{21}^{(s)}, \frac{\partial y_{2t+s}}{\partial \varepsilon_{2t}} = \theta_{22}^{(s)}$$

The *structural impulse response functions* (IRFs) are the plots of $\theta_{ij}^{(s)}$ vs. s for $i, j = 1, 2$. These plots summarize how unit impulses of the structural shocks at time t impact the level of y at time $t+s$ for different values of s.

Stationarity of y_t implies

$$\lim_{s \to \infty} \theta_{ij}^{(s)} = 0, \quad i, j = 1, 2$$

The *long-run cumulative impact* of the structural shocks is captured by

$$\Theta(1) = \begin{bmatrix} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{\infty} \theta_{11}^{(s)} & \sum_{s=0}^{\infty} \theta_{12}^{(s)} \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} & \sum_{s=0}^{\infty} \theta_{22}^{(s)} \end{bmatrix}$$

$$\Theta(L) = \begin{bmatrix} \theta_{11}(L) & \theta_{12}(L) \\ \theta_{21}(L) & \theta_{22}(L) \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{\infty} \theta_{11}^{(s)} L^s & \sum_{s=0}^{\infty} \theta_{12}^{(s)} L^s \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} L^s & \sum_{s=0}^{\infty} \theta_{22}^{(s)} L^s \end{bmatrix}$$

Digression: Dynamic Regression Models

In the SVAR every variable is engodenous. Suppose, for example, y_{2t} is strictly exogenous which implies $b_{21} = 0$ and $\gamma_{21} = 0$. Then, the first equation is an ADL(1,1)

$$\begin{aligned} y_{1t} &= \alpha + \phi y_{1t-1} + \beta_0 y_{2t} + \beta_1 y_{2t-1} + \varepsilon_{1t} \\ \mathrm{cov}(y_{2t}, \varepsilon_{1t}) &= 0 \end{aligned}$$

In lag operator notation the equation becomes

$$\begin{aligned} \phi(L) y_{1t} &= \alpha + \beta(L) y_{2t} + \varepsilon_{1t} \\ \phi(L) &= 1 - \phi L, \ \beta(L) = \beta_0 + \beta_1 L \end{aligned}$$

The second equation is an AR(1) model for y_{2t}

$$y_{2t} = c + \rho y_{2t-1} + \varepsilon_{2t}$$

Stationarity now only requires $|\phi| < 1$ and $|\rho| < 1$.

The first equation may then be solved for y_{1t} as a function of y_{2t} and ε_{1t}

$$
\begin{aligned}
y_{1t} &= \frac{\alpha}{\phi(1)} + \phi(L)^{-1}\beta(L)y_{2t} + \phi(L)^{-1}\varepsilon_{1t} \\
&= \mu + \psi_\beta(L)y_{2t} + \psi(L)\varepsilon_t \\
\mu &= \frac{\alpha}{\phi(1)} \\
\psi_\beta(L) &= \phi(L)^{-1}\beta(L),\ \psi(L) = \phi(L)^{-1}
\end{aligned}
$$

Since y_{2t} is exogenous, we have two sources of shocks.
Note: there can be four types of dynamic multipliers :

$$\frac{\partial y_{1t+s}}{\partial y_{2t}}, \frac{\partial y_{1t+s}}{\partial \varepsilon_{2t}}, \frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}}, \frac{\partial y_{2t+s}}{\partial \varepsilon_{2t}}$$

The short-run dyamic multipliers with respect to y_{2t} and ε_{1t} are

$$\begin{aligned}\frac{\partial y_{1t+s}}{\partial y_{2t}} &= \frac{\partial y_{1t}}{\partial y_{2t-s}} = \psi_{\beta,s} \\ \frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}} &= \frac{\partial y_{1t}}{\partial \varepsilon_{1t-s}} = \psi_s\end{aligned}$$

In the *steady state* or *long-run equilibrium* all variables are constant

$$\begin{aligned}y_1^* &= \mu + \psi_\beta(L)y_2^* = \mu + \psi_\beta(1)y_2^* \\ y_2^* &= \frac{c}{1-\rho} \\ \psi_\beta(1) &= \phi(1)^{-1}\beta(1) = \frac{\beta_0+\beta_1}{1-\phi}\end{aligned}$$

The long-run impact of a change in y_2 on y_1 is then

$$\frac{\partial y_1^*}{\partial y_2^*} = \psi_\beta(1) = \frac{\beta_0+\beta_1}{1-\phi} = \sum_{s=0}^{\infty} \frac{\partial y_{1t+s}}{\partial y_{2t}}$$

Identification issues

In some applications, identification of the parameters of the SVAR is achieved through restrictions on the parameters of the SMA representation.

Identification through contemporaneous restrictions

Suppose that ε_{2t} has no contemporaneous impact on y_{1t}. Then $\theta_{12}^{(0)} = 0$ and

$$\Theta_0 = \begin{bmatrix} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix}.$$

Since $\Theta_0 = \mathbf{B}^{-1}$ then

$$\begin{bmatrix} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

$$\Rightarrow \quad b_{12} = 0$$

Hence, assuming $\theta_{12}^{(0)} = 0$ in the SMA representation is equivalent to assuming $b_{12} = 0$ in the SVAR representation.

Identification through long-run restrictions

Suppose ε_{2t} has no long-run cumulative impact on y_{1t}. Then

$$\begin{aligned}\theta_{12}(1) &= \sum_{s=0}^{\infty}\theta_{12}^{(s)} = 0 \\ \mathbf{\Theta}(1) &= \begin{bmatrix}\theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1)\end{bmatrix}.\end{aligned}$$

This type of long-run restriction places nonlinear restrictions on the coefficients of the SVAR since

$$\begin{aligned}\mathbf{\Theta}(1) &= \mathbf{\Psi}(1)\mathbf{B}^{-1} = \mathbf{A}(1)^{-1}\mathbf{B}^{-1} \\ &= (\mathbf{I}_2 - \mathbf{B}^{-1}\mathbf{\Gamma}_1)^{-1}\mathbf{B}^{-1}\end{aligned}$$

Estimation Issues

In order to compute the structural IRFs, the parameters of the SMA representation need to be estimated. Since

$$\begin{aligned}\Theta(L) &= \Psi(L)\mathbf{B}^{-1} \\ \Psi(L) &= \mathbf{A}(L)^{-1} = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}\end{aligned}$$

the estimation of the elements in $\Theta(L)$ can often be broken down into steps:

- $\mathbf{A}_1$ is estimated from the reduced form VAR.
- Given $\widehat{\mathbf{A}}_1$, the matrices in $\Psi(L)$ can be estimated using $\widehat{\Psi}_k = \widehat{\mathbf{A}}_1^k$.
- $\mathbf{B}$ is estimated from the identified SVAR.
- Given $\hat{\mathbf{B}}$ and $\hat{\Psi}_k$, the estimates of Θ_k, $k = 0, 1, \ldots,$ are given by $\hat{\Theta}_k = \widehat{\Psi}_k \hat{\mathbf{B}}^{-1}$.

Forecast Error Variance Decompositions

Idea: determine the proportion of the variability of the errors in forecasting y_1 and y_2 at time $t+s$ based on information available at time t that is due to variability in the structural shocks ε_1 and ε_2 between times t and $t+s$.

To derive the FEVD, start with the Wold representation for $\mathbf{y}_{t+s}$

$$\begin{aligned}\mathbf{y}_{t+s} &= \mu + \mathbf{u}_{t+s} + \Psi_1 \mathbf{u}_{t+s-1} + \cdots \\ &\quad + \Psi_{s-1}\mathbf{u}_{t+1} + \Psi_s \mathbf{u}_t + \Psi_{s+1}\mathbf{u}_{t-1} + \cdots.\end{aligned}$$

The best linear forecast of $\mathbf{y}_{t+s}$ based on information available at time t is

$$\mathbf{y}_{t+s|t} = \mu + \Psi_s \mathbf{u}_t + \Psi_{s+1}\mathbf{u}_{t-1} + \cdots$$

and the forecast error is

$$\mathbf{y}_{t+s} - \mathbf{y}_{t+s|t} = \mathbf{u}_{t+s} + \Psi_1 \mathbf{u}_{t+s-1} + \cdots + \Psi_{s-1}\mathbf{u}_{t+1}.$$

Econometrics V Lecture 11

Using

$$\varepsilon_t = \mathbf{B}^{-1}\mathbf{u}_t, \ \Theta_k = \Psi_k \mathbf{B}^{-1}$$

The forecast error in terms of the structural shocks is

$$\begin{aligned} \mathbf{y}_{t+s} - \mathbf{y}_{t+s|t} &= \mathbf{B}^{-1}\varepsilon_{t+s} + \Psi_1 \mathbf{B}^{-1}\varepsilon_{t+s-1} + \\ &\quad \cdots + \Psi_{s-1}\mathbf{B}^{-1}\varepsilon_{t+1} \\ &= \Theta_0 \varepsilon_{t+s} + \Theta_1 \varepsilon_{t+s-1} + \cdots + \Theta_{s-1}\varepsilon_{t+1} \end{aligned}$$

The forecast errors equation by equation are

$$\begin{bmatrix} y_{1t+s} - y_{1t+s|t} \\ y_{2t+s} - y_{2t+s|t} \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \cdots + \begin{bmatrix} \theta_{11}^{(s-1)} & \theta_{12}^{(s-1)} \\ \theta_{21}^{(s-1)} & \theta_{22}^{(s-1)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+1} \\ \varepsilon_{2t+1} \end{bmatrix}$$

For the first equation

$$y_{1t+s} - y_{1t+s|t} = \theta_{11}^{(0)}\varepsilon_{1t+s} + \cdots + \theta_{11}^{(s-1)}\varepsilon_{1t+1} + \theta_{12}^{(0)}\varepsilon_{2t+s} + \cdots + \theta_{12}^{(s-1)}\varepsilon_{2t+1}$$

Since it is assumed that $\varepsilon_t \sim i.i.d.\ (0, \mathbf{D})$ where $\mathbf{D}$ is diagonal, the variance of the forecast error in may be decomposed as

$$var(y_{1t+s} - y_{1t+s|t}) = \sigma_1^2(s)$$
$$= \sigma_1^2\left(\left(\theta_{11}^{(0)}\right)^2 + \cdots + \left(\theta_{11}^{(s-1)}\right)^2\right)$$
$$+\sigma_2^2\left(\left(\theta_{12}^{(0)}\right)^2 + \cdots + \left(\theta_{12}^{(s-1)}\right)^2\right).$$

The proportion of $\sigma_1^2(s)$ due to shocks in ε_1 is then

$$\rho_{1,1}(s) = \frac{\sigma_1^2\left(\left(\theta_{11}^{(0)}\right)^2 + \cdots + \left(\theta_{11}^{(s-1)}\right)^2\right)}{\sigma_1^2(s)}$$

the proportion of $\sigma_1^2(s)$ due to shocks in ε_2 is

$$\rho_{1,2}(s) = \frac{\sigma_2^2\left(\left(\theta_{12}^{(0)}\right)^2 + \cdots + \left(\theta_{12}^{(s-1)}\right)^2\right)}{\sigma_1^2(s)}.$$

The forecast error variance decompositions (FEVDs) for y_{2t+s} are

$$\rho_{2,1}(s) = \frac{\sigma_1^2\left(\left(\theta_{21}^{(0)}\right)^2 + \cdots + \left(\theta_{21}^{(s-1)}\right)^2\right)}{\sigma_2^2(s)},$$

$$\rho_{2,2}(s) = \frac{\sigma_2^2\left(\left(\theta_{22}^{(0)}\right)^2 + \cdots + \left(\theta_{22}^{(s-1)}\right)^2\right)}{\sigma_2^2(s)},$$

where

$$var(y_{2t+s} - y_{2t+s|t}) = \sigma_2^2(s)$$
$$= \sigma_1^2\left(\left(\theta_{21}^{(0)}\right)^2 + \cdots + \left(\theta_{21}^{(s-1)}\right)^2\right)$$
$$+\sigma_2^2\left(\left(\theta_{22}^{(0)}\right)^2 + \cdots + \left(\theta_{22}^{(s-1)}\right)^2\right).$$

SIMS (1972, 1980, 1986, 1992)

- Large-scale statistical macroeconomic models

Inappropriate / incredible identification (a priori restrictions)

- Simultaneous equations

Two or more endogenous variables (simultaneous equation bias)

VAR modelling

Alternative style of macroeconometrics proposed

Estimate dynamic systems 'without using theoretical perspectives'

Estimate dynamic systems 'using lagged (exogenous) variables'

SIMPLE 2-VARIABLE (= BIVARIATE) MODEL [y1, y2]

$$y_{1t} = b_{11}\, y_{1t-1} + \dots + b_{1q}\, y_{1t-q} + b_{21}\, y_{2t-1} + \dots + b_{2q}\, y_{2t-q} + e_{y1t}$$

$$y_{2t} = c_{11}\, y_{1t-1} + \dots + c_{1q}\, y_{1t-q} + c_{21}\, y_{2t-1} + \dots + c_{2q}\, y_{2t-q} + e_{y2t}$$

(for simplicity, excl. constant term / intercepts)

vector/matrix representation: $y_t = B(L)\, y_{t-1} + e_t$

where $B(L) = B_1 + B_2 L + B_3 L^2 + \dots + B_q L^{q-1}$

y_t is a vector of (endogenous) variables

y_t depends on own lags, autoregression

- FORECASTING

VAR forecasts extrapolate expected values of current and future values of each of the variables using observed lagged values of all variables, assuming no further shocks

- IMPULSE RESPONSE FUNCTIONS (IRFs)

IRFs trace out the expected responses of current and future values of each of the variables to a shock in one of the VAR equations (note: shocks can be defined/measured in different ways)

- FORECAST ERROR DECOMPOSITION OF VARIANCE (FEDVs)

FEDVs provide the percentage of the variance of the error made in forecasting a variable at a given horizon due to a specific shock. Thus, the FEDV is like a (partial) R^2 for the forecast error

- GRANGER-CAUSALITY TESTS

Granger-causality requires that lagged values of variable A are related to subsequent values in variable B, keeping constant the lagged values of variable B and any other explanatory variables

Assume

$y_{1t} = b_{11}\ y_{1t-1} + b_{21}\ y_{2t-1} + e_{y1t}$

$y_{2t} = c_{11}\ y_{1t-1} + c_{21}\ y_{2t-1} + e_{y2t}$

Or $y_t = B\ y_{t-1} + e_t$

Then $E_{t-1}(y_{1t}) = b_{11}\ y_{1t-1} + b_{21}\ y_{2t-1}$ where $E_{t-1}(e_{y1t}) = E_{t-1}(e_{y1t+j}) = 0$

$E_{t-1}(y_{2t}) = c_{11}\ y_{1t-1} + c_{21}\ y_{2t-1}$ where $E_{t-1}(e_{y2t}) = E_{t-1}(e_{y2t+j}) = 0$

Or $E_{t-1}(y_t) = B\ y_{t-1}$

Then $E_{t-1}(y_{1t+1}) = b_{11}\ E_{t-1}(y_{1t}) + b_{21}\ E_{t-1}(y_{2t}) =$

$= b_{11}^2\ y_{1t-1} + b_{11}\ b_{21}\ y_{2t-1} + b_{21}\ c_{11}\ y_{1t-1} + b_{21}\ c_{21}\ y_{2t-1}$

$= [b_{11}^2 + b_{21}\ c_{11}]\ y_{1t-1} + [b_{11}\ b_{21} + b_{21}\ c_{21}]\ y_{2t-1}$

Or $E_{t-1}(y_{t+1}) = B^2\ y_{t-1}$

In general: $E_{t-1}(y_{t+k}) = B^{k+1}\ y_{t-1}$
and more lagged variables $E_{t-1}(y_{t+k}) = B(L)^{k+1}\ y_{t-1}$

VAR IMPULSE RESPONSE FUNCTIONS (IRFs)

IRFs: expected k-period ahead prediction error of a variable produced by an innovation in another variable.

$y_{1t} = b_{11}\, y_{1t-1} + b_{21}\, y_{2t-1} + e_{y1t}$

$y_{2t} = c_{11}\, y_{1t-1} + c_{21}\, y_{2t-1} + e_{y2t}$

Or $y_t = B\, y_{t-1} + e_t$

Then $E(y_{1t}-E_{t-1}(y_{1t})) = e_{y1t}$

$E(y_{2t}-E_{t-1}(y_{2t})) = e_{y2t}$

Or $E(y_t-E_{t-1}(y_t)) = e_t$

Problem: e_{y1t} and e_{y2t} may be correlated
Structural model: identification problem!

$E(y_{1t+1}-E_{t-1}(y_{1t+1})) = b_{11}\, e_{y1t} + b_{21}\, e_{y2t}$

$E(y_{2t+1}-E_{t-1}(y_{2t+1})) = c_{11}\, e_{y1t} + c_{21}\, e_{y2t}$ where $E(e_{y1t+j}) = E(e_{y2t+j}) = 0$

Or $E(y_{t+1}-E_{t-1}(y_{t+1})) = B\, e_t$

In general: $E(y_{t+k}-E_{t-1}(y_{t+k})) = B^k\, e_t$

FEDVs: percentage of the expected k-period ahead squared prediction error of a variable produced by an innovation in another variable.

Ignore formulas, related to IRFs

Granger causality: If the history (i.e. lagged observations) of variable x does not help to predict the future values of variable y (given lagged values of y and lagged values of other variables), we say that x does not Granger-cause y. (Granger, 1969; Sims, 1972)

Alternative test statistics

- Bivariate tests

$$y_{1t} = b_{11}\, y_{1t-1} + b_{21}\, y_{2t-1} + e_{y1t}$$
$$y_{2t} = c_{11}\, y_{1t-1} + c_{21}\, y_{2t-1} + e_{y2t}$$

H_0: y_2 does not Granger-cause y_1, test $b_{21} = 0$
H_0: y_1 does not Granger-cause y_2, test $c_{11} = 0$

using t-statistics (single coefficient), or F-test,
Log-likelihood ratio-test, Wald-test (multiple coefficients)

Block exogeneity tests:
Does a variable belong in the VAR somewhere because it affects variables through other variables

e.g. test coefficient for y_{1t-1} in equations for y_{2t}, <u>and</u> y_{3t}, <u>and</u> … etc.

Log-likelihood ratio-test

VARX is a VAR model with endogenous and exogenous variables

Vector/matrix representation: $y_t = B(L)\, y_{t-1} + C(L)\, x_t + e_t$

No feedback from variable y_t to $x_t, x_{t+1}, \ldots x_{t+k}$

For example,

$x_t =$ - trends, seasonal dummies, shock dummies
- oil price
- foreign variables (for small open country models)
etc.

Vector Error Correction Model (VECM) is a VAR model with error-correction mechanism based on cointegration relationships between variables

Vector/matrix representation: $\Delta y_t = B(L)\, \Delta y_{t-1} + \Pi\, y_{t-1} + e_t$

Coefficient matrix Π describes long-run equilibrium (= cointegration) relationships between variables y_t

For example, $\Pi\, y_t$

money demand	$m - p = c + \alpha\, y - \gamma\, i$
share prices and dividends	$sp = div - \ln(r - g)$
exchange rate, domestic and foreign price levels	$s = q + p^f - p$
etc.	

ISSUES IN VAR MODELLING

- Selection of VAR variables (2, 3, …; which ones?)
- Selection of VAR variables levels or differences
- Selection of VAR lag lengths
- Identification scheme

  * Variables ordering (= Choleski decomposition)

  * Structural VARs
    - contemporaneous restrictions (short-run)
    - long-run restrictions
    - mix

VAR CHOICE OF LEVELS, 1ST DIFFERENCES

- in levels, if all variables are stationary (I(0))
- in first differences, if some variables have a unit root (I(1)) and the series are not cointegrated

Note: with I(1) variables a VAR in levels or 1st differences makes no difference asymptotically (e.g. Sims, Stock, Watson, 1990), but 1st differences is better in small samples (Hamilton, 1994: 553, 652)

But, if 2 or more variables I(1) and cointegrated

- 1st difference estimates are biased if there is cointegration because ECM is omitted
- levels estimates implicitly incorporate cointegration relationship but standard errors are unreliable, inefficient

→ VECM is preferred

Alternative criteria for finding “best” model

- LR: Likelihood ratio test criterion (sensitive to errors being normal distributed)
- FPE: Final prediction error criterion
- AIC: Akaike information criterion
- SIC: Schwarz information criterion
- HQ: Hannan-Quin information criterion

The “best” fitting model is the one that maximizes the LR, or minimizes the FPE criterion function (in essence, the overall sum of squared residuals) or AIC, SIC or HQ.

Alternative criteria imply different tradeoffs between better fit (smaller residuals) and loss of degrees of freedom (due to number of estimated parameters). Difficult to say which one is “best”.

Additional requirement that VAR residuals are not autocorrelated (and normal distributed).

INTERPRETATIONS OF IRFs and FEDVs

IRF: $E(y_{t+k}-E_{t-1}(y_{t+k})) = B(L)^k e_t$

Effects of shock to e_{1t}

- But e_{1t} and e_{2t} may be correlated. How to separate?
- Structural interpretation
  E.g. Is a shock to e_{jt} in the interest rate equation actually measuring monetary policy? Interest rates also respond to news about inflation, economic growth, exchange rates, etc.

VARs ARE NOT ATHEORETICAL MODELLING!

Structural models	(1a) $A_0\, y_t = A(L)\, y_{t-1} + D\, \varepsilon_t$.
Reduced form	(1b) $y_t = A_0^{-1} A(L)\, y_{t-1} + A_0^{-1} D\, \varepsilon_t$.
VAR model	(2) $y_t = B(L)\, y_{t-1} + e_t$.

- VAR equals reduced-form of structural model with $B(L) = A_0^{-1} A(L)$ and $e_t = A_0^{-1} D\, \varepsilon_t$
- VAR reduced-form residuals e_t ARE NOT structural shocks ε_t

To derive structural shocks ε_t (and structural coefficients generally) we must identify $A_0^{-1} D$: cannot be estimated directly, identifying assumptions required

From estimated parameters to structural parameters

Assume number of variables n, number of lags q

Unknown structural parameters	Estimated VAR parameters
$A_0 y_t = A(L) y_{t-1} + D \varepsilon_t$	$y_t = B(L) y_{t-1} + e_t$
A_0: n*n	*
A(L): n*n*q	B(L): n*n*q
D: n*n	*
Σ_ε: n(n+1)/2 unique elements	Σ_e: n(n+1)/2 unique elements

Too few estimated parameters to solve unknown structural parameters

Still 2*(n*n) further 'identification restrictions' required

VAR MODEL IDENTIFICATION (Cont)

A_0: unit diagonal elements → n restrictions (normalizations, explain y_t)

D: identity matrix (0, 1) → n*n restrictions (structural shocks do not affect other variables directly)

Σ_ε: diagonal matrix → n(n+1)/2 – n restrictions (structural shocks uncorrelated)

$2n^2 – (n + n^2 + n(n+1)/2 – n) = \mathbf{n(n-1)/2}$

We need additional n(n-1)/2 restrictions to identify

- Standard VAR models — 'Wold-ordering' and 'Choleski-decomposition'
  Recursive identification
- Structural VAR models — alternative restrictions
  Nonrecursive identification

VAR model [y1, y2, y3]

Standard or non-structural Wold-ordering en Choleski decomposition referred to as recursive identification

$e_{1t} = \varepsilon_{1t}$

$e_{2t} = \alpha_{21}\, \varepsilon_{1t} + \varepsilon_{2t}$

$e_{3t} = \alpha_{31}\, \varepsilon_{1t} + \alpha_{32}\, \varepsilon_{2t} + \varepsilon_{3t}$

with n(n-1)/2 [i.e. 3*2/2=3] additional zero restrictions [i.e. α_{12}, α_{13}, α_{23}=0]

'Ordering' of the variables y1, y2, y3 determines how they affect each other

Note: Stock/Watson (2001) recursive VAR model

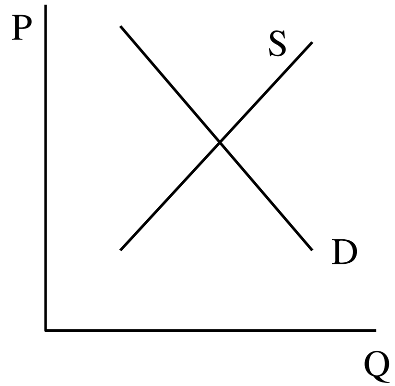
$y_t = B(L)\, y_{t-1} + C\, y_t + \varepsilon_t$ with $C = \begin{bmatrix} 0 & 0 & 0 & \dots \\ c_{21} & 0 & 0 & \dots \\ c_{31} & c_{32} & 0 & \dots \end{bmatrix}$

equations estimated OLS

exogeneity assumed

PROBLEM RECURSIVE VAR IDENTIFICATION

Example, normal market where price and quantity are determined by demand and supply. Recursive VAR uses unacceptable restrictions.



Quantity Q and price P

Demand function $Q^d = -\delta P + \varepsilon^d$

Supply function $Q^s = P + \varepsilon^s$

Equilibrium $Q^d = Q^s$

SOLUTIONS: price $P = [1+\delta]^{-1}\varepsilon^d - [1+\delta]^{-1}\varepsilon^s$

quantity $Q = [1+\delta]^{-1}\varepsilon^d + \delta\,[1+\delta]^{-1}\varepsilon^s$

P-Q VAR model has no Wold-ordering; no triangular matrix for shock effects, but simultaneous relationships

Only special market conditions allow triangular matrix (i.e. vertical or horizontal demand or supply functions)

VAR model [y1, y2, y3]

Identification using contemporaneous (short-run) restrictions

$$e_{1t} = \varepsilon_{1t} + \alpha_{12}\,\varepsilon_{2t} + \alpha_{13}\,\varepsilon_{3t}$$
$$e_{2t} = \alpha_{21}\,\varepsilon_{1t} + \varepsilon_{2t} + \alpha_{23}\,\varepsilon_{3t}$$
$$e_{3t} = \alpha_{31}\,\varepsilon_{1t} + \alpha_{32}\,\varepsilon_{2t} + \varepsilon_{3t}$$

Requiring n(n-1)/2 [i.e. 3*2/2=3] additional zero restrictions [e.g. $\alpha_{ij}=0$]

Note: Stock/Watson (2001) structural VAR model

$y_t = B(L)\, y_{t-1} + C\, y_t + \varepsilon_t$ with $C = [0\ c_{12}\ c_{13} \ldots]$

simultaneous equations estimated IV $[c_{21}\ 0\ c_{23} \ldots]$

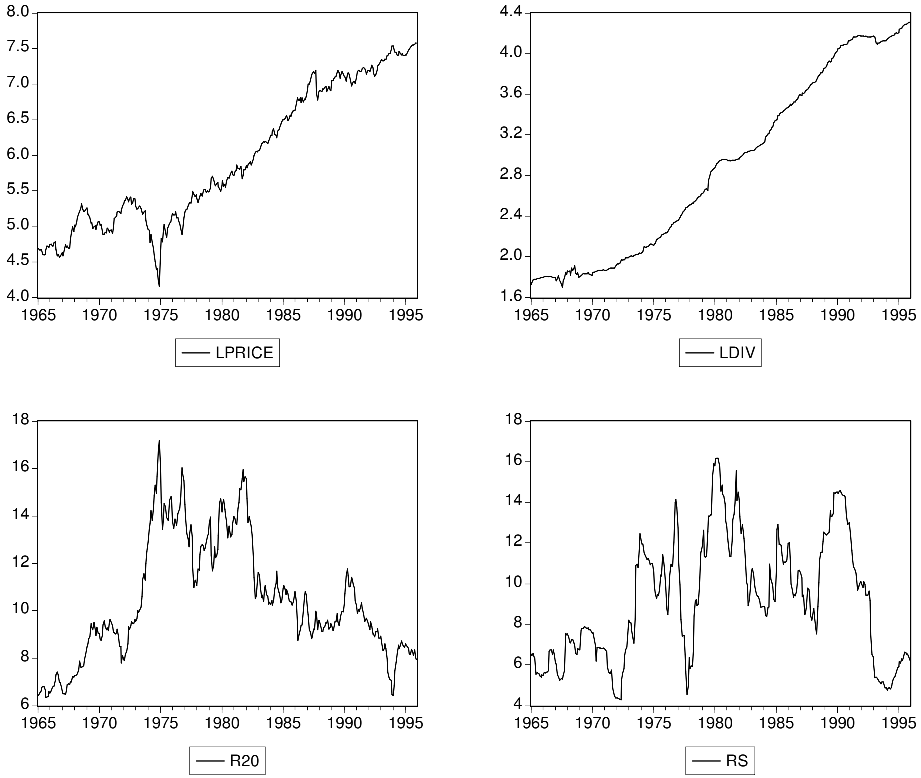
$[c_{31}\ c_{32}\ 0 \ldots]$ with n(n-1)/2 extra restrictions

Alternative structural identification
* long-run restrictions (= sum coefficients) required to be zero, one, …
* mixed short-run and long-run restrictions

EXAMPLE VAR

Monthly observations on
* FTA All Share index (FTAprice),
* FTA Dividend index (FTAdiv),
* yield on 20 year UK Gilts (R20),
* 91 day Treasury bills (RS)

January 1965 to December 1995
(372 months)



Basic dividend discount model for stock market

$$P_t = E_t\left[\sum_{i=1}^{\infty}\left(\frac{1}{(1+R)}\right)^i D_{t+i}\right]$$

can be written (in logs) as

$$p_t = \frac{k}{1-\rho} + E_t\left[\sum_{j=0}^{\infty} \rho^j [(1-\rho)\, d_{t+j+1} - r_{t+j+1}]\right]$$

p = log share price
d = log dividend
r = one-period return

VAR forecasting model for future d and r

Rational expectations, efficient markets present-value model implies certain restrictions on VAR model [p, d, r]

VAR VARIABLES UNIT ROOT TESTS

	Levels		First-differences	
Test	Constant, no trend	Constant, trend	Constant, no trend	Constant, trend
ADF				
log FTAprice	-0.278 [0.925] (3)	-2.536 [0.311] (3)	-10.217 [0.000] (2)	-10.216 [0.000] (2)
log FTAdiv	0.460 [0.985] (8)	-2.665 [0.252] (8)	-4.537 [0.000] (7)	-4.624 [0.001] (7)
R20	-1.989 [0.292] (2)	-1.899 [0.653] (2)	-13.119 [0.000] (1)	-13.189 [0.000] (1)
RS	-2.612 [0.091] (1)	-2.536 [0.311] (1)	-13.575 [0.000] (0)	-13.578 [0.000] (0)
PP				
log FTAprice	-0.156 [0.941]	-2.422 [0.368]	-16.905 [0.000]	-16.874 [0.000]
log FTAdiv	0.887 [0.995]	-2.268 [0.450]	-18.900 [0.000]	-18.811 [0.000]
R20	-1.926 [0.320]	-1.843 [0.682]	-13.654 [0.000]	-13.623 [0.000]
RS	-2.437 [0.132]	-2.341 [0.410]	-13.575 [0.000]	-13.578 [0.000]
KPSS				
log FTAprice	2.316	0.355	0.083	0.046
log FTAdiv	2.270	0.346	0.442	0.272
R20	0.482	0.483	0.232	0.036
RS	0.514	0.300	0.081	0.030
crit val 1%	0.739000	0.216000	0.739000	0.216000
5%	0.463000	0.146000	0.463000	0.146000

VAR LAG LENGTH SELECTION

VAR Lag Order Selection Criteria
Endogenous variables: DLDIV DLPRICE DR20 DRS
Exogenous variables: C

31*12 = 372 – 1 for 1st diff – 12 for max lag

Sample: 1965:01 1995:12
Included observations: 359

Lag	LogL	LR	FPE	AIC	SC	HQ
0	1102.694	NA	2.58E-08	-6.120858	-6.077589	-6.103651
1	1157.779	108.6361	2.08E-08	-6.338603	-6.122262*	-6.252572*
2	1174.263	32.14157	2.07E-08*	-6.341299*	-5.951885	-6.186445
3	1184.995	20.68635	2.13E-08	-6.311950	-5.749463	-6.088271
4	1188.681	7.022523	2.28E-08	-6.243347	-5.507787	-5.950844
5	1196.543	14.80367	2.39E-08	-6.198008	-5.289376	-5.836681
6	1210.511	25.99237	2.42E-08	-6.186693	-5.104988	-5.756542
7	1217.982	13.73406	2.54E-08	-6.139175	-4.884397	-5.640200
8	1233.620	28.40201*	2.54E-08	-6.137161	-4.709310	-5.569362
9	1240.182	11.77085	2.68E-08	-6.084580	-4.483656	-5.447956
10	1248.197	14.19882	2.81E-08	-6.040094	-4.266097	-5.334646
11	1254.971	11.84952	2.96E-08	-5.988695	-4.041625	-5.214423
12	1264.552	16.54766	3.07E-08	-5.952938	-3.832796	-5.109842

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

LAG LENGTH AND RESIDUAL CORRELATION
1 LAG MODEL

VAR Residual Portmanteau Tests for Autocorrelations
H0: no residual autocorrelations up to lag h

Sample: 1965:01 1995:12
Included observations: 370

31*12 = 372 – 1 for 1st diff - 1 lag model

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	2.322186	NA*	2.328479	NA*	NA*
2	32.02764	0.0099	32.19538	0.0094	16
3	52.35899	0.0130	52.69292	0.0121	32
4	65.63168	0.0462	66.11067	0.0425	48
5	85.36840	0.0384	86.11776	0.0341	64
6	109.9050	0.0149	111.0588	0.0124	80
7	120.7004	0.0449	122.0624	0.0375	96
8	150.5627	0.0088	152.5847	0.0065	112
9	164.6290	0.0161	167.0016	0.0117	128
10	184.7106	0.0125	187.6410	0.0085	144
11	194.8653	0.0315	198.1069	0.0218	160
12	207.7349	0.0511	211.4079	0.0352	176

*The test is valid only for lags larger than the VAR lag order.
df is degrees of freedom for (approximate) chi-square distribution

Q-Stat = B-P
Adj Q-Stat = L-B

Residual correlation remains with 1 lag

LAG LENGTH AND RESIDUAL CORRELATION
2 LAG MODEL

VAR Residual Portmanteau Tests for Autocorrelations
H0: no residual autocorrelations up to lag h

Sample: 1965:01 1995:12
Included observations: 369

31*12 = 372 – 1 for 1st diff - 2 lags model

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.582023	NA*	0.583605	NA*	NA*
2	1.794387	NA*	1.802575	NA*	NA*
3	18.77958	0.2803	18.92699	0.2725	16
4	32.45501	0.4443	32.75229	0.4299	32
5	49.08332	0.4295	49.60901	0.4089	48
6	70.06338	0.2815	70.93585	0.2576	64
7	82.40645	0.4048	83.51760	0.3720	80
8	106.4813	0.2182	108.1260	0.1872	96
9	122.4393	0.2354	124.4829	0.1979	112
10	141.1596	0.2012	143.7246	0.1619	128
11	150.9869	0.3284	153.8539	0.2719	144
12	163.9779	0.3983	167.2816	0.3306	160

*The test is valid only for lags larger than the VAR lag order.
df is degrees of freedom for (approximate) chi-square distribution

Residual correlation removed with 2 lags

VAR ESTIMATES

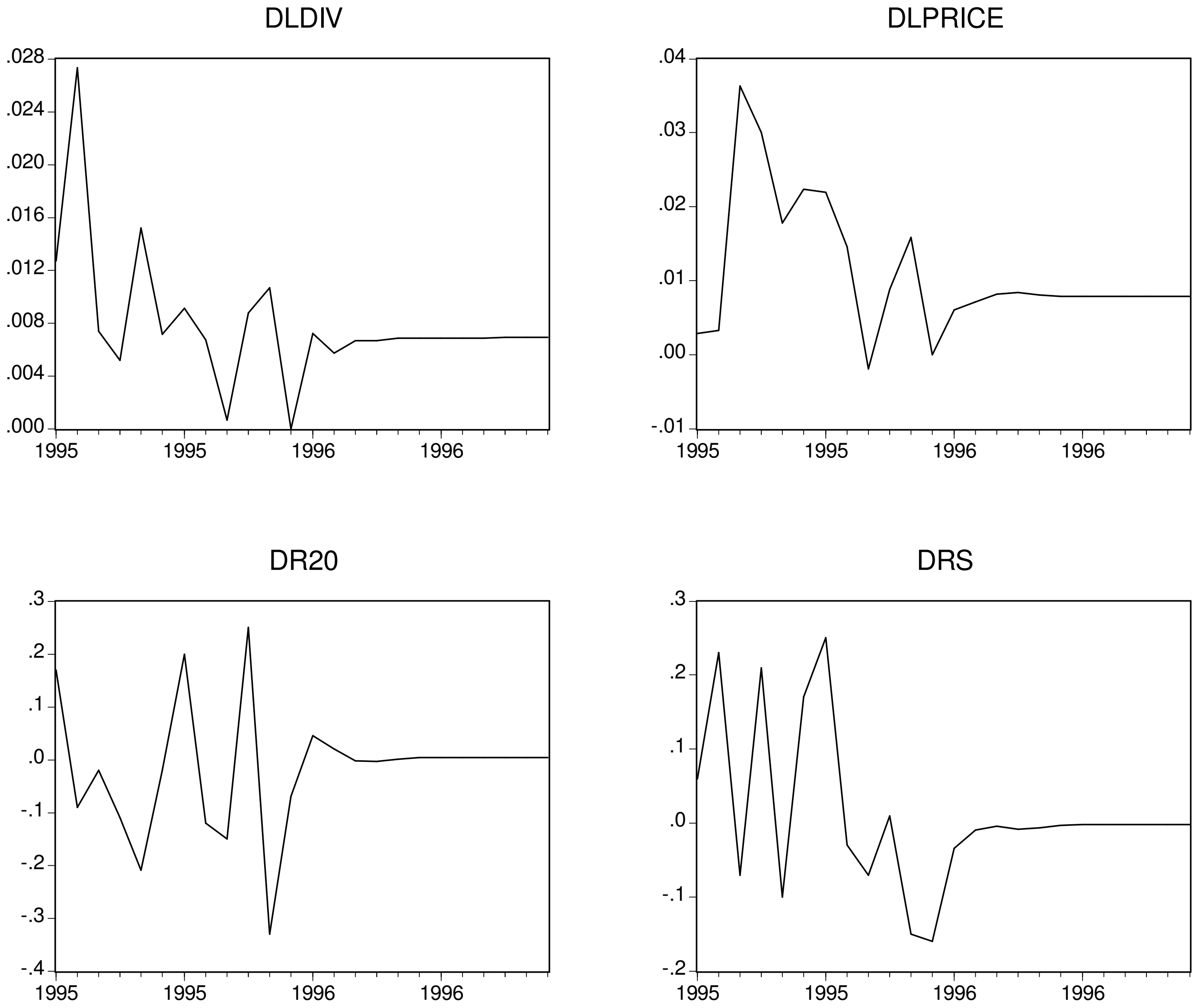
Vector Autoregression Estimates
Sample(adjusted): 1965:04 1995:12
Included observations: 369 after adjusting endpoints
Standard errors in ( ) & t-statistics in [ ]

	DLDIV	DLPRICE	DR20	DRS
DLDIV(-1)	0.031048	-0.123668	2.322616	2.567405
	(0.05250)	(0.22135)	(1.30558)	(1.97302)
	[0.59139]	[-0.55869]	[1.77899]	[1.30126]
DLDIV(-2)	0.104663	0.067811	1.004643	1.989944
	(0.05263)	(0.22189)	(1.30873)	(1.97778)
	[1.98876]	[0.30561]	[0.76765]	[1.00615]
DLPRICE(-1)	-0.006030	0.096377	-1.741237	-0.996496
	(0.01295)	(0.05458)	(0.32195)	(0.48654)
	[-0.46577]	[1.76564]	[-5.40841]	[-2.04814]
… [cut]				
R-squared	0.031060	0.063462	0.187060	0.162649
Adj. R-squared	0.009528	0.042650	0.168995	0.144042
Sum sq. resids	0.072843	1.294899	45.04720	102.8782
S.E. equation	0.014225	0.059975	0.353739	0.534577
F-statistic	1.442497	3.049281	10.35465	8.740932
Log likelihood	1050.241	519.2728	-135.5690	-287.9355
Akaike AIC	-5.643583	-2.765706	0.783572	1.609406
Schwarz SC	-5.548197	-2.670321	0.878957	1.704792
Mean dependent	0.006884	0.007887	0.003821	-0.000892
S.D. dependent	0.014293	0.061296	0.388044	0.577809
Determinant Residual Covariance			1.72E-08	
Log Likelihood (d.f. adjusted)			1204.059	
Akaike Information Criteria			-6.330943	
Schwarz Criteria			-5.949402	

VAR FORECASTS

1996:01-1996:12

Baseline



Actual data 1995:01-1995:12, forecasts starting 1996:01

VAR RESIDUAL CROSS-CORRELATIONS

VAR Residual Cross-Correlations
Ordered by: variables

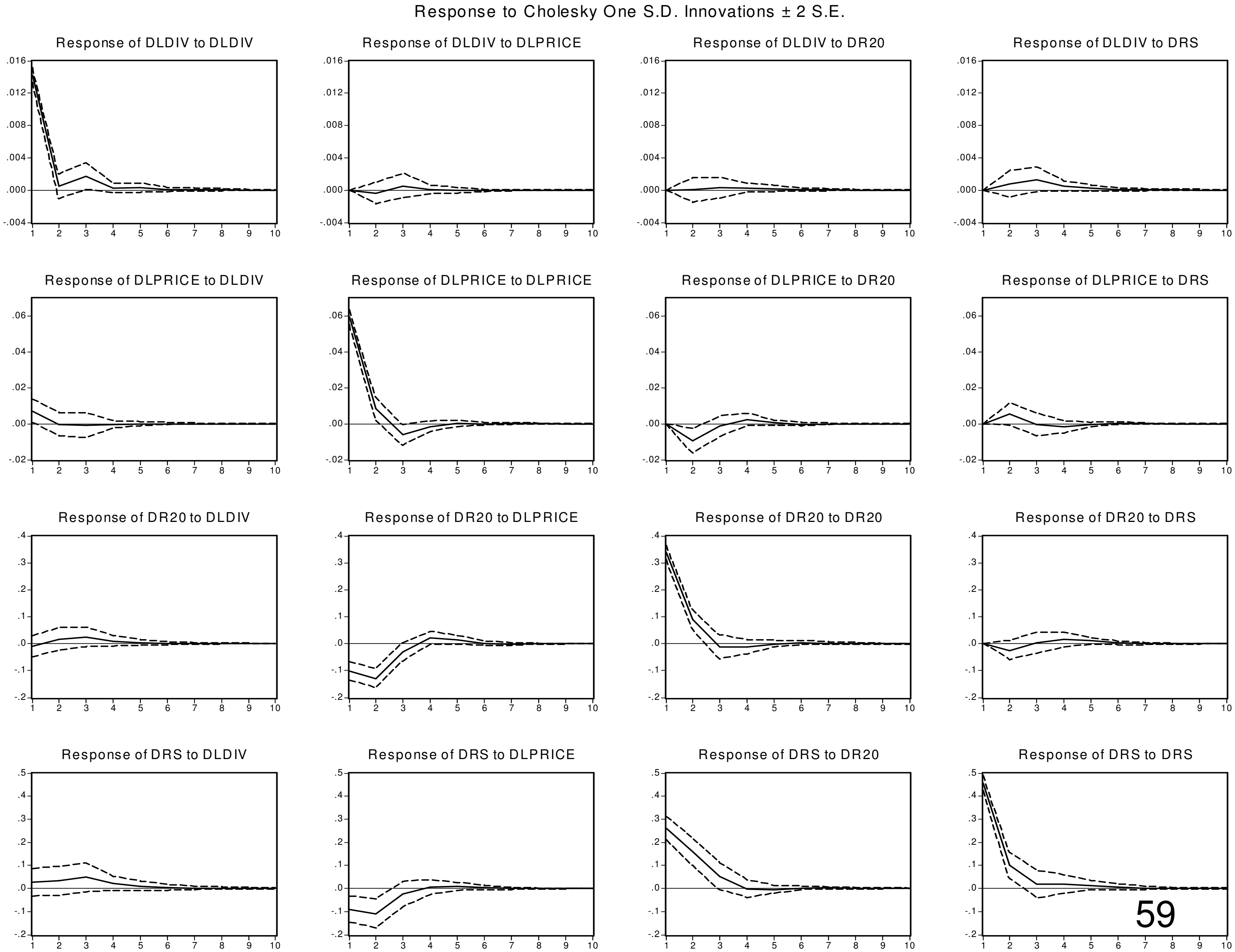
Sample: 1965:01 1995:12
Included observations: 369

	DLDIV	DLPRICE	DR20	DRS
DLDIV	1.000000			
DLPRICE	0.118511	1.000000		
DR20	-0.027883	-0.286931	1.000000	
DRS	0.050364	-0.163683	0.513983	1.000000

Orthogonalization of impulses necessary

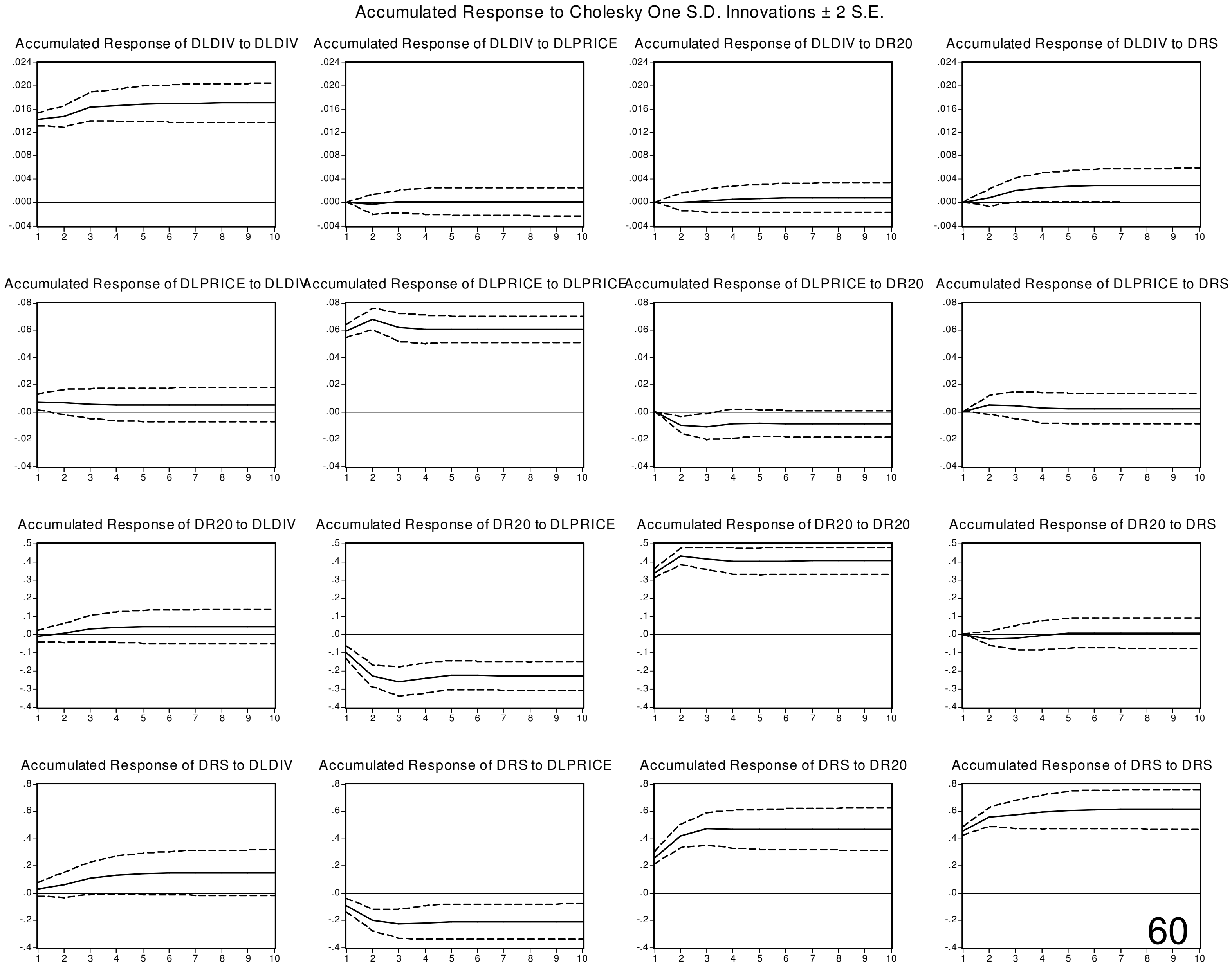
Selected identification scheme: In this example recursive VAR: Cholesky decomposition

IMPULSE RESPONSE FUNCTIONS (10 months)

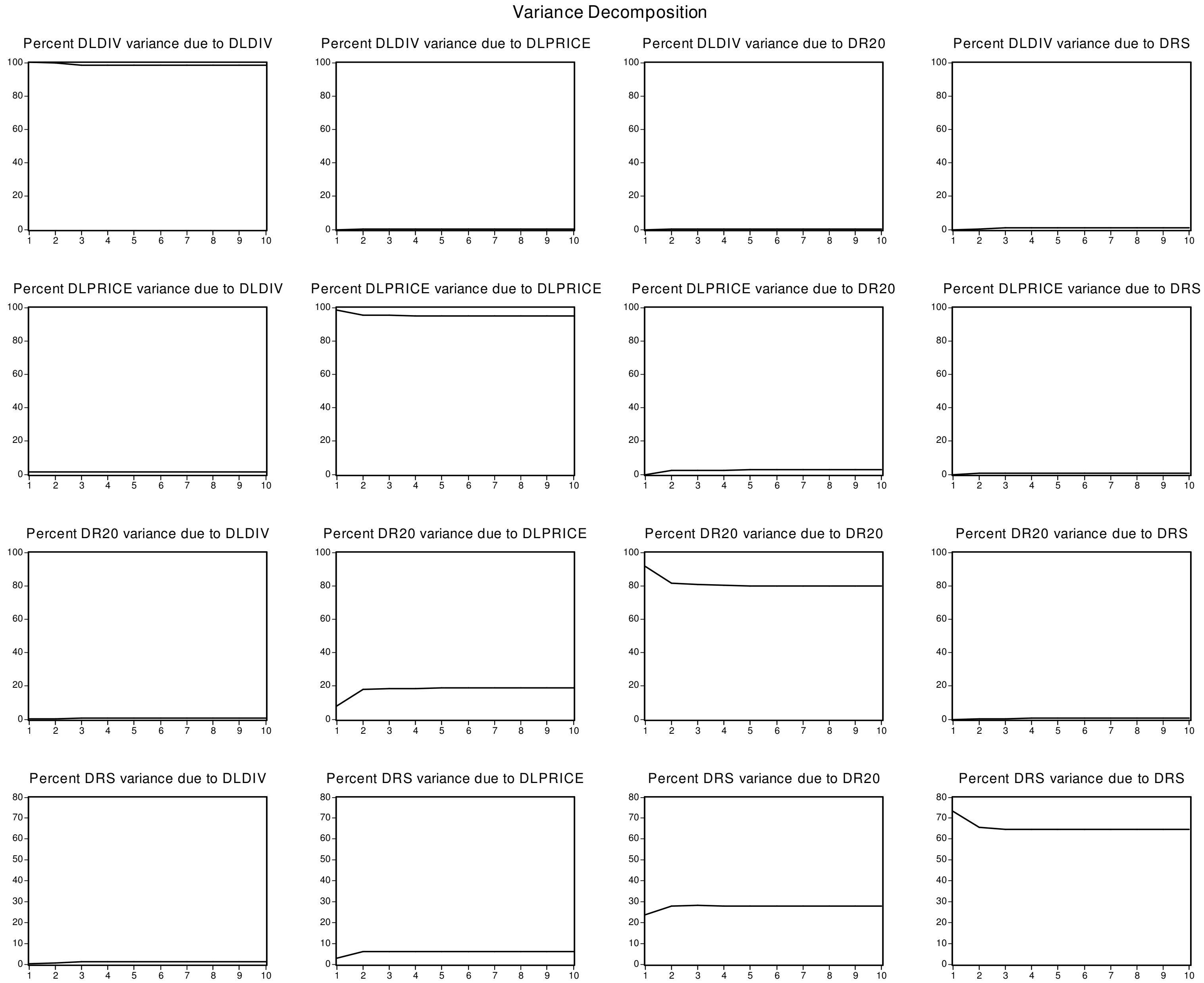


Edith Cowan University
Facu

IMPULSE RESPONSE FUNCTIONS (10 months, accumulated)



FORECAST ERROR DECOMPOSITION



VAR CAUSALITY TESTS

Test restriction for each equation in the VAR that coefficients for selected variable(s) all equal zero

Alternative to standard Wald test in EViews: use F-test or Log-likelihood ratio test for individual equations estimated separately

VAR Pairwise Granger Causality/Block Exogeneity Wald Tests

Sample: 1965:01 1995:12
Included observations: 369

Dependent variable: DLDIV

Exclude	Chi-sq	df	Prob.
DLPRICE	0.908277	2	0.6350
DR20	0.841612	2	0.6565
DRS	4.516248	2	0.1045
All	5.362081	6	0.4983

Dependent variable: DLPRICE

Exclude	Chi-sq	df	Prob.
DLDIV	0.391903	2	0.8221
DR20	12.43508	2	0.0020
DRS	3.126158	2	0.2095
All	12.78468	6	0.0466

Tests joint significance of all other lagged endogenous variables in the equation

Dependent variable: DR20

Exclude	Chi-sq	df	Prob.
DLDIV	3.875408	2	0.1440
DLPRICE	29.26330	2	0.0000
RS	2.817665	2	0.2444
All	33.28872	6	0.0000

Dependent variable: DRS

Exclude	Chi-sq	df	Prob.
DLDIV	2.820340	2	0.2441
DLPRICE	6.125972	2	0.0467
DR20	9.990261	2	0.0068
All	21.16888	6	0.0017

How well do VARs perform 4 tasks? (Stock/Watson, 2001)

- Data description

Useful description of comovements in variables. E.g. theoretical models that imply Granger causality could be tested. But, limitations on statistical inference. E.g. calculating standard errors for IRFs

- Forecasting

Useful as a simple benchmark. But, limited numbers of variables and only lagged information

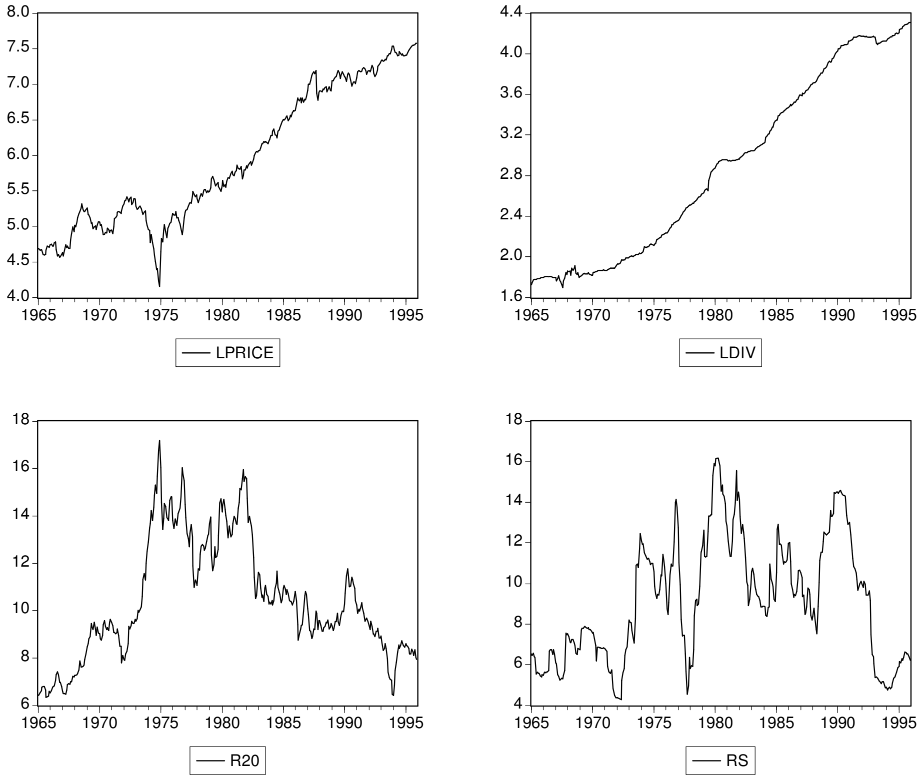
- Structural inference

Sensitive to identification of structural shocks. E.g. monetary policy shocks, demand-supply shocks. Omitted variables bias. E.g. price puzzle for monetary policy.

- Policy analysis

One-off shocks – represented in IRFs
Changes in policy rules – VAR models also subject to Lucas critique

EXAMPLE VAR:
LEVELS



Monthly observations on
* FTA All Share index (FTAprice),
* FTA Dividend index (FTAdiv),
* yield on 20 year UK Gilts (R20),
* 91 day Treasury bills (RS)

January 1965 to December 1995
(372 months)

Some results are crucially different

EXAMPLE VAR IN LEVELS:
VAR LAG LENGTH SELECTION

VAR Lag Order Selection Criteria
Endogenous variables: LDIV LPRICE R20 RS
Exogenous variables: C

Sample: 1965:01 1995:12
Included observations: 360

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-1892.193	NA	0.441710	10.53441	10.57759	10.55158
1	1128.801	5958.073	2.48E-08	-6.160008	-5.944113	-6.074164
2	1187.997	115.4306*	1.95E-08*	-6.399981*	-6.011370*	-6.245462*
3	1201.042	25.14930	1.99E-08	-6.383568	-5.822242	-6.160374
4	1212.196	21.25389	2.04E-08	-6.356644	-5.622602	-6.064775
5	1217.481	9.952685	2.17E-08	-6.297114	-5.390357	-5.936570
6	1225.126	14.22837	2.27E-08	-6.250698	-5.171225	-5.821479
7	1237.898	23.48779	2.31E-08	-6.232769	-4.980580	-5.734875
8	1246.976	16.49175	2.40E-08	-6.194314	-4.769409	-5.627745
9	1259.053	21.67070	2.46E-08	-6.172517	-4.574896	-5.537273
10	1266.942	13.98144	2.57E-08	-6.127457	-4.357121	-5.423538
11	1273.491	11.46007	2.72E-08	-6.074949	-4.131897	-5.302355
12	1281.702	14.18673	2.84E-08	-6.031677	-3.915909	-5.190408

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

EXAMPLE VAR IN LEVELS:
LAG LENGTH AND RESIDUAL CORRELATION
2 LAG MODEL

VAR Residual Portmanteau Tests for Autocorrelations
H0: no residual autocorrelations up to lag h

Sample: 1965:01 1995:12
Included observations: 370

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	1.744362	NA*	1.749090	NA*	NA*
2	26.25784	NA*	26.39579	NA*	NA*
3	45.69866	0.0001	45.99553	0.0001	16
4	59.65049	0.0021	60.09984	0.0019	32
5	77.26963	0.0047	77.96034	0.0040	48
6	97.09965	0.0048	98.11722	0.0039	64
7	113.9212	0.0076	115.2632	0.0060	80
8	134.9972	0.0054	136.8049	0.0040	96
9	146.0473	0.0169	148.1305	0.0126	112
10	164.1074	0.0172	166.6923	0.0122	128
11	174.5405	0.0423	177.4450	0.0304	144
12	188.4672	0.0615	191.8386	0.0436	160

*The test is valid only for lags larger than the VAR lag order.
df is degrees of freedom for (approximate) chi-square distribution

Residual correlation remains with 2 lags, but is usually ignored in levels models

EXAMPLE VAR IN LEVELS:
VAR RESIDUAL CROSS-CORRELATIONS

VAR Residual Cross-Correlations
Ordered by: variables

Sample: 1965:01 1995:12
Included observations: 370

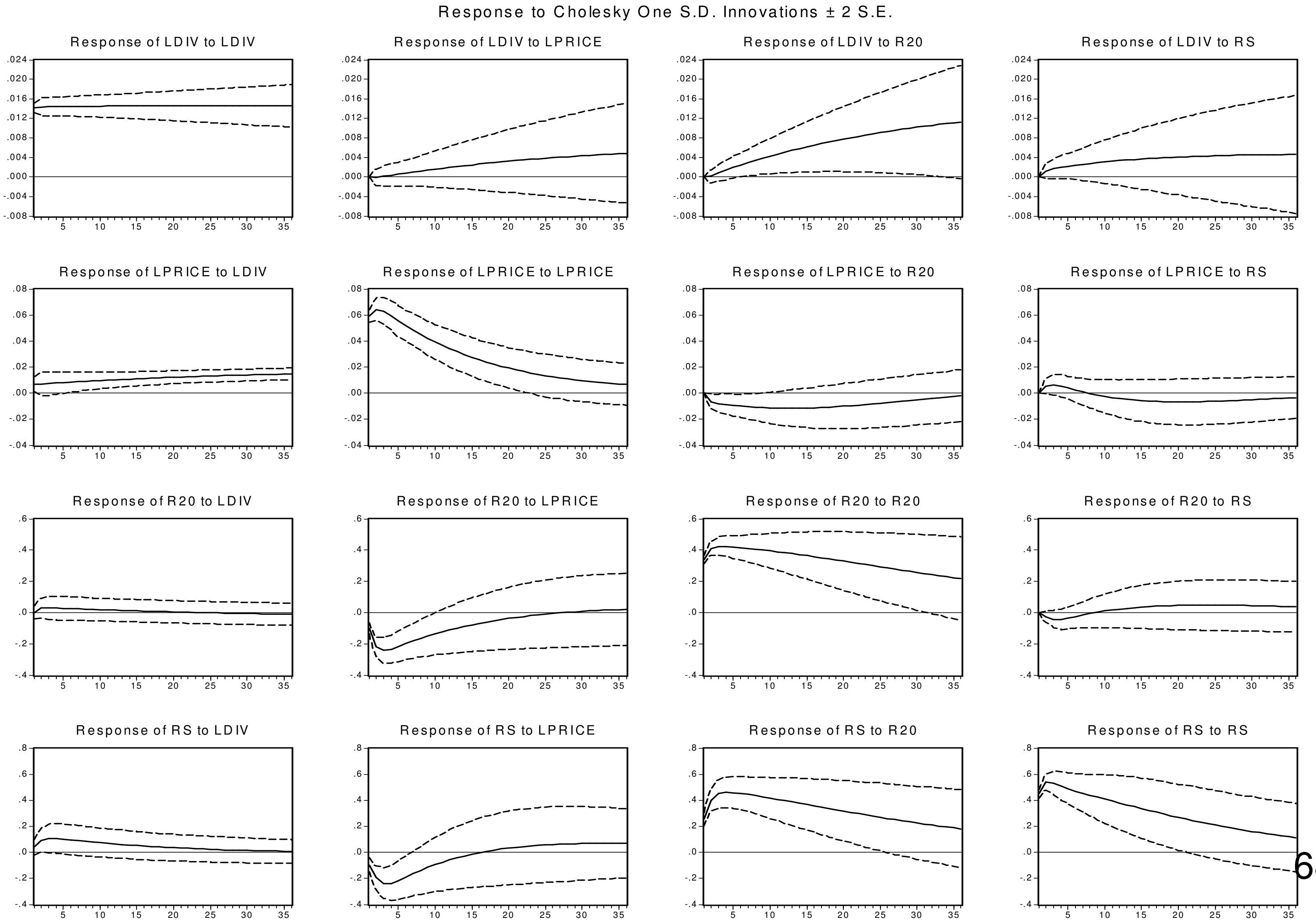
	LDIV	LPRICE	R20	RS
LDIV	1.000000			
LPRICE	0.117045	1.000000		
R20	0.002553	-0.283717	1.000000	
RS	0.078986	-0.170897	0.517853	1.000000

Orthogonalization of impulses necessary

Selected identification scheme: In this example recursive VAR, Cholesky decomposition

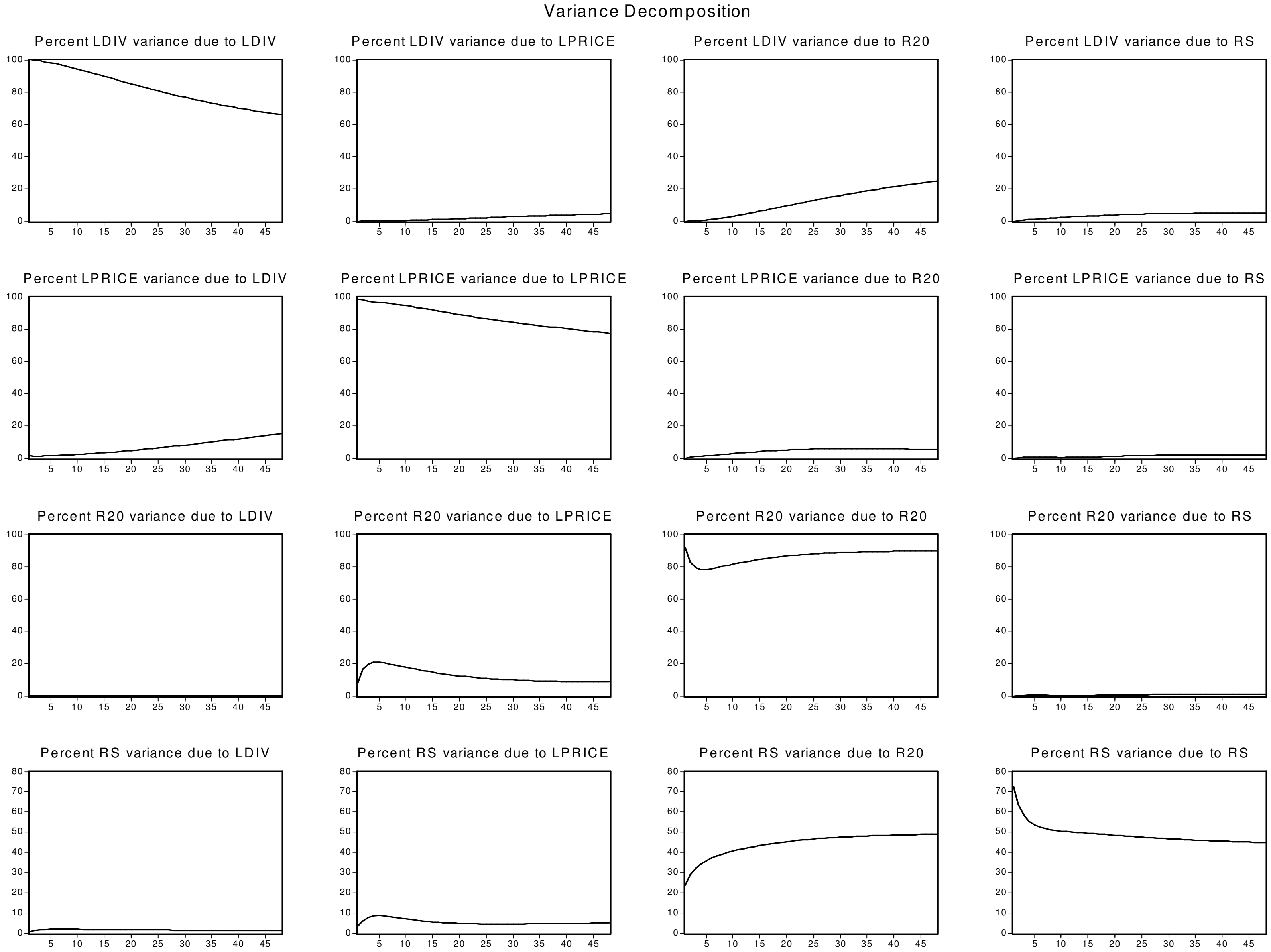
EXAMPLE VAR IN LEVELS:

IMPULSE RESPONSE FUNCTIONS (36 months)



EXAMPLE VAR IN LEVELS:

FORECAST ERROR DECOMPOSITION (48 months)



EXAMPLE VAR IN LEVELS:
VAR CAUSALITY TESTS

Test restriction for each equation in the VAR that coefficients for selected variable(s) all equal zero

VAR Pairwise Granger Causality/Block Exogeneity Wald Tests

Sample: 1965:01 1995:12
Included observations: 370

Dependent variable: LDIV

Exclude	Chi-sq	df	Prob.
LPRICE	3.528107	2	0.1713
R20	3.406864	2	0.1821
RS	4.131428	2	0.1267
All	14.13748	6	0.0281

Dependent variable: LPRICE

Exclude	Chi-sq	df	Prob.
LDIV	13.05868	2	0.0015
R20	7.651568	2	0.0218
RS	4.766564	2	0.0922
All	21.23616	6	0.0017

Tests joint significance of all other lagged endogenous variables in the equation

Dependent variable: R20

Exclude	Chi-sq	df	Prob.
LDIV	6.424003	2	0.0403
LPRICE	27.63004	2	0.0000
RS	2.976666	2	0.2257
All	34.30814	6	0.0000

Dependent variable: RS

Exclude	Chi-sq	df	Prob.
LDIV	4.238411	2	0.1201
LPRICE	5.592535	2	0.0610
R20	9.606175	2	0.0082
All	20.38451	6	0.0024

Balke/Emery (1994) The Federal Funds Rate as an Indicator of Monetary Policy: Evidence from the 1980s

Table 1
Marginal Significance Levels for Exclusion of Lags, 1959–79

Three-variable system (federal funds rate, unemployment, inflation)

	Marginal Significance Levels		
		Lags of	
Equation	Federal funds	Unemployment	Inflation
Federal funds	.0000	.0000	.0002
Unemployment	.0092	.0000	.2300
Inflation	.0000	.0698	.1304

Four-variable system (federal funds rate, M2, unemployment, inflation)

	Marginal Significance Levels			
			Lags of	
Equation	Federal funds	M2	Unemployment	Inflation
Federal funds	.0000	.0357	.0003	.0035
M2	.0000	.0000	.0897	.4875
Unemployment	.1140	.1032	.0000	.1032
Inflation	.0000	.0063	.2523	.2376

Tests for dropping variable from four-variable system

M2	$\chi^2(18) = 40.88$	(.0016)
Federal funds	$\chi^2(18) = 85.50$	(.0000)

Balke/Emery (1994)

Figure 2
Responses of Federal Funds Rate to Inflation and Unemployment Shocks

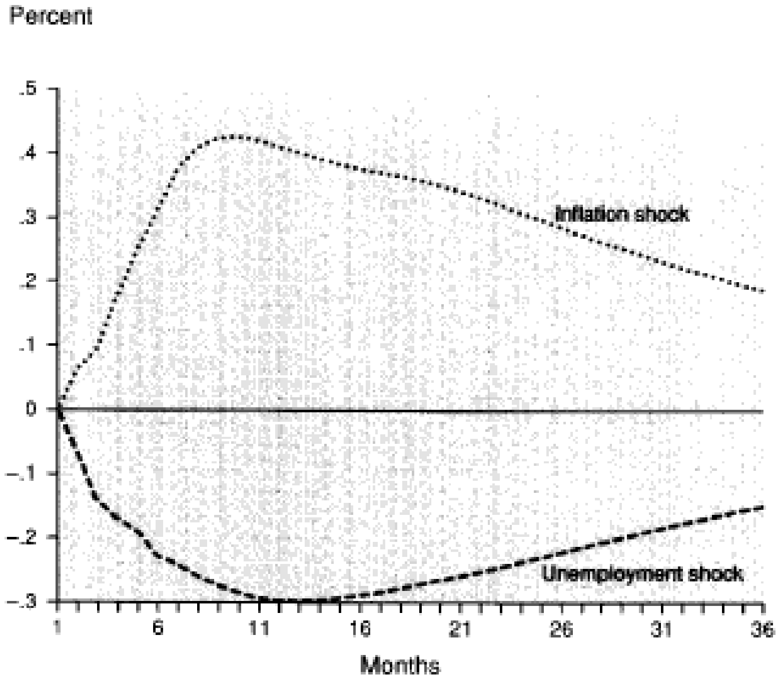
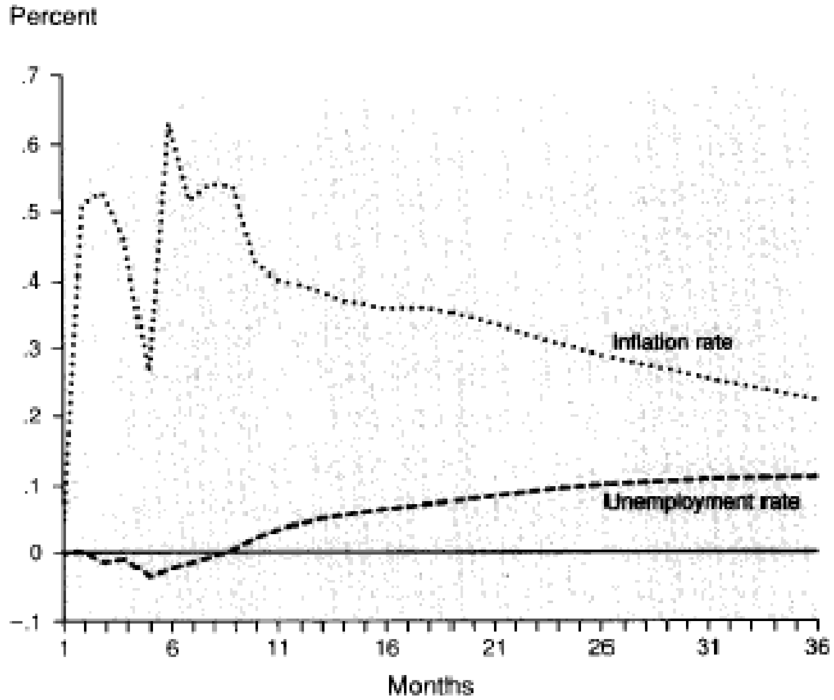


Figure 3
Responses to Federal Funds Rate Shocks



Balke/Emery (1994)

Table 2

Forecast Error Variance Decompositions, 1959–79

Federal funds rate

Forecast horizon	Percentage of forecast error variance explained by: Federal funds	M2	Unemployment	Inflation
6	77.7	2.9	8.9	10.5
12	51.7	9.3	18.7	20.3
24	30.5	31.9	22.2	15.5
36	23.9	41.6	21.8	12.7

M2

Forecast horizon	Percentage of forecast error variance explained by: Federal funds	M2	Unemployment	Inflation
6	22.5	70.3	3.3	4.0
12	21.1	64.0	8.1	6.8
24	23.1	60.0	9.8	7.1
36	24.3	58.0	9.4	8.3

Unemployment rate

Forecast horizon	Percentage of forecast error variance explained by: Federal funds	M2	Unemployment	Inflation
6	1.1	2.8	96.2	.3
12	1.7	6.0	90.4	.6
24	13.2	10.2	68.1	3.6
36	23.8	8.7	57.5	8.0

Inflation rate

Forecast horizon	Percentage of forecast error variance explained by: Federal funds	M2	Unemployment	Inflation
6	12.9	7.7	2.9	76.5
12	17.2	11.8	4.7	66.2
24	15.5	24.5	6.7	52.3
36	14.2	33.6	8.0	44.2

Balke/Emery (1994)

Figure 4
Responses of Federal Funds Rate to Inflation and Unemployment Shocks