Introduction to VAR Models

Nicola Viegi
University of Pretoria

July 2010

Introduction

Origins of VAR models

Sims "Macroeconomics and Reality" Econometrica 1980

It should be feasible to estimate large macromodels as unrestricted reduced forms, treating all variables as endogenous

- Natural extension of the univariate autoregressive model to multivariate time series
- Especially useful for describing the dynamic behaviour of economic and financial time series
- Benchmark in forecasting
- Used for structural inference

Consider a bivariate $Y_t=(y_t, z_t)$, first-order VAR model:

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

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

The two variables y and z are endogenous.

The error terms (structural shocks) ε_{yt} and ε_{zt} are white noise innovations with standard deviations σ_y and σ_z and a **zero** covariance.

Shock ε_{yt} affects y directly and z indirectly.

There are 10 (8 coefficients and two standard deviations of the errors) parameters to estimate.

- The structural is not estimable directly
- VAR in reduced form is estimable.
- In a reduced form representation y and z are just functions of lagged y and z.
- To solve for a reduced form write the structural VAR in matrix form as:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

or, in short

$$BY_t = \Gamma_0 + \Gamma_1 Y_{t-1} + \varepsilon_t$$

- Premultipication by B^{-1} allow us to obtain a standard VAR(1):

$$BY_t = \Gamma_0 + \Gamma_1 Y_{t-1} + \varepsilon_t$$
$$Y_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 Y_{t-1} + B^{-1}\varepsilon_t$$
$$Y_t = A_0 + A_1 Y_{t-1} + a_t$$

- This reduced form can be estimated (by OLS equation by equation)
- Before estimating
  - Determine the optimal lag length of the VAR
  - Determine stability conditions (roots of the system inside the unit circle)
- After estimating the reduced form
  - Hypothesis Testing – Granger Casuality
  - Impulse Response Function
  - Variance Decomposition
  - Identification of Structural VAR

Information Criteria (IC) can be used to choose the "right" number of lags in a VAR(p): that minimizes IC(p) for p=1, ..., P .

$$AIC = \ln\left|\Sigma(p)\right| + \frac{2}{T}(n^2 p + n)$$

$$SBC = \ln\left|\Sigma(p)\right| + (n^2 p + n)\frac{\ln(T)}{T}$$

AIC criterion asymptotically overestimates the order with positive probability

SBC criterion estimates the order consistently if the true p is less than the p(max)

Same concept of stability than in univariate time series

The system is stable if the root of the matrix A_1 of the system are all less than 1 in absolute value

$$Y_t = A_0 + A_1 Y_{t-1} + a_t$$

If not Vector Error Correction Models and Cointegration Models more appropriate.

Granger (1969) “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods”, *Econometrica, 37*

Consider two random variables X_t, Y_t

Forecast of X_t, s periods ahead

$$\hat{X}_t(s)^{(1)} = E(X_{t+s} \mid X_t, X_{t-1}, \ldots.) \quad \hat{X}_t(s)^{(2)} = E(X_{t+s} \mid X_t, X_{t-1}, \ldots Y_t, Y_{t-1}, \ldots.)$$

Define Minimum square error $MSE(\hat{X}_t(s)) = E(X_{t+s} - \hat{X}_t(s))^2$

If $MSE(\hat{X}_t(s)^{(1)}) = MSE(\hat{X}_t(s)^{(2)})$

then Y_t does not Granger-cause X_t $\forall s > 0$

$\Leftrightarrow X_t$ is exogenous with respect to Y_t

$\Leftrightarrow Y_t$ is not linearly informative to forecast X_t

Granger Causality Test

Assume a lag length of p

$$X_t = c_1 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ...\alpha_p X_{t-p} + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ...\beta_p X_{t-p} + a_t$$

Estimate by OLS and test for the following hypothesis

$$H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0 \;\; (Y_t \text{ does not Granger - cause } X_t)$$

$$H_1 : \text{any } \beta_i \neq 0$$

Unrestricted sum of squared residuals $RSS_1 = \sum_t \hat{a}_t^{\,2}$

Restricted sum of squared residuals $RSS_2 = \sum_t \hat{\hat{a}}_t^{\,2}$

$$F = \frac{(RSS_2 - RSS_1)/p}{RSS_1/(T-2p-1)}$$

reject if $F > F_{\alpha,(p,T-2p-1)}$

Important – Stationarity of the data

Objective: Show the reaction of the system to a shock

$$Y_t = c + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + a_t$$

If the system is covariance - stationary,

$$Y_t = \mu + \Psi(L) a_t = \mu + a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + ...$$

$$\Psi(L) = [\Phi(L)]^{-1}$$

Redating at time $t+s$:

$$Y_{t+s} = \mu + a_{t+s} + \Psi_1 a_{t+s-1} + \Psi_2 a_{t+s-2} + ... + \Psi_s a_t + \Psi_{s+1} a_{t-1} + ...$$

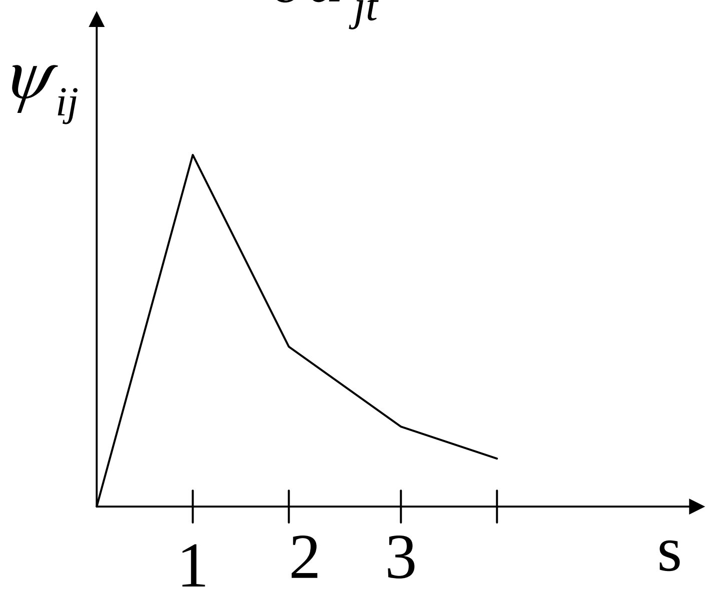
$$\frac{\partial Y_{t+s}}{\partial a'_t} = \Psi_s = \left[\psi_{ij}^{(s)}\right]_{n \times n} \qquad \text{(multipliers)}$$

$$\frac{\partial y_{i,t+s}}{\partial a_{jt}} = \psi_{ij}^{(s)} \longrightarrow$$

Reaction of the i-variable to a unit change in innovation j

Impulse-response function: response of $y_{i,t+s}$ to one-time impulse in y_{jt} with all other variables dated t or earlier held constant.

$$\frac{\partial y_{i,t+s}}{\partial a_{jt}} = \psi_{ij}$$



Example: IRF for a VAR(1)

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2_{t-1}} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}; \quad \Sigma_a = \begin{bmatrix} \sigma_1{}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2{}^2 \end{bmatrix}$$

$$t < 0 \quad y_{1t} = y_{2t} = 0$$

$$t = 0 \quad a_{20} = 1 \quad (y_{2t} \text{ increases by 1 unit})$$

(no more shocks occur)

Reaction of the system

$$\begin{bmatrix} y_{10} \\ y_{20} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad (\textit{impulse})$$

$$\begin{bmatrix} y_{11} \\ y_{21} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \phi_{11} \\ \phi_{22} \end{bmatrix}$$

$$\begin{bmatrix} y_{12} \\ y_{22} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{21} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} y_{1s} \\ y_{2s} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}^s \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \Phi_1{}^s \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Contribution of the j-th orthogonalized innovation to the MSE of the s-period ahead forecast

If shocks does not explain none of the forecast error variance of at all forecast horizon we can say that the sequence is exogenous

- Is it possible to recover the parameters in the structural VAR from the estimated parameters in the standard VAR? No!!
- There are 10 parameters in the bivariate structural VAR(1) and only 9 estimated parameters in the standard VAR(1).
- The VAR is underidentified.
- If one parameter in the structural VAR is restricted the standard VAR is exactly identified.
- Sims (1980) suggests a recursive system to identify the model letting $b_{21}=0$. **Choleski decomposition**.

$$\begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

- **b_{21}=0** implies

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

The parameters of the structural VAR can now be identified from the following 9 equations

$$a_{10} = b_{10} - b_{12} b_{20} \quad a_{20} = b_{20} \quad \mathrm{var}(e_1) = \sigma_y^2 + b_{12}^2 \sigma_z^2$$

$$a_{11} = \gamma_{11} - b_{12} \gamma_{21} \quad a_{21} = \gamma_{21} \quad \mathrm{var}(e_2) = \sigma_z^2$$

$$a_{12} = \gamma_{12} - b_{12} \gamma_{22} \quad a_{22} = \gamma_{22} \quad \mathrm{cov}(e_1, e_2) = -b_{12} \sigma_z^2$$

- Both structural shocks can now be identified
- b_{21}=0 implies y does not have a contemporaneous effect on z.
- Both ε_{yt} and ε_{zt} affect y contemporaneously but only ε_{zt} affects z contemporaneously.
- The residuals of e_{2t} are due to pure shocks to z.
- There are other methods used to identify models – Restrictions coming from theory – (Sims Bernake, Blanchard and Quah etc)

Critics on VAR

- A VAR model can be a good forecasting model, but it is an atheoretical model (as all the reduced form models are).
- To calculate the IRF, the order matters: Identification not unique.
- Sensitive to the lag selection
- Dimensionality problem.

Standard Tool for Macroeconomic Analysis

An example of VAR analysis

Leeper Sims and Zha (1996)
"What does monetary policy do"

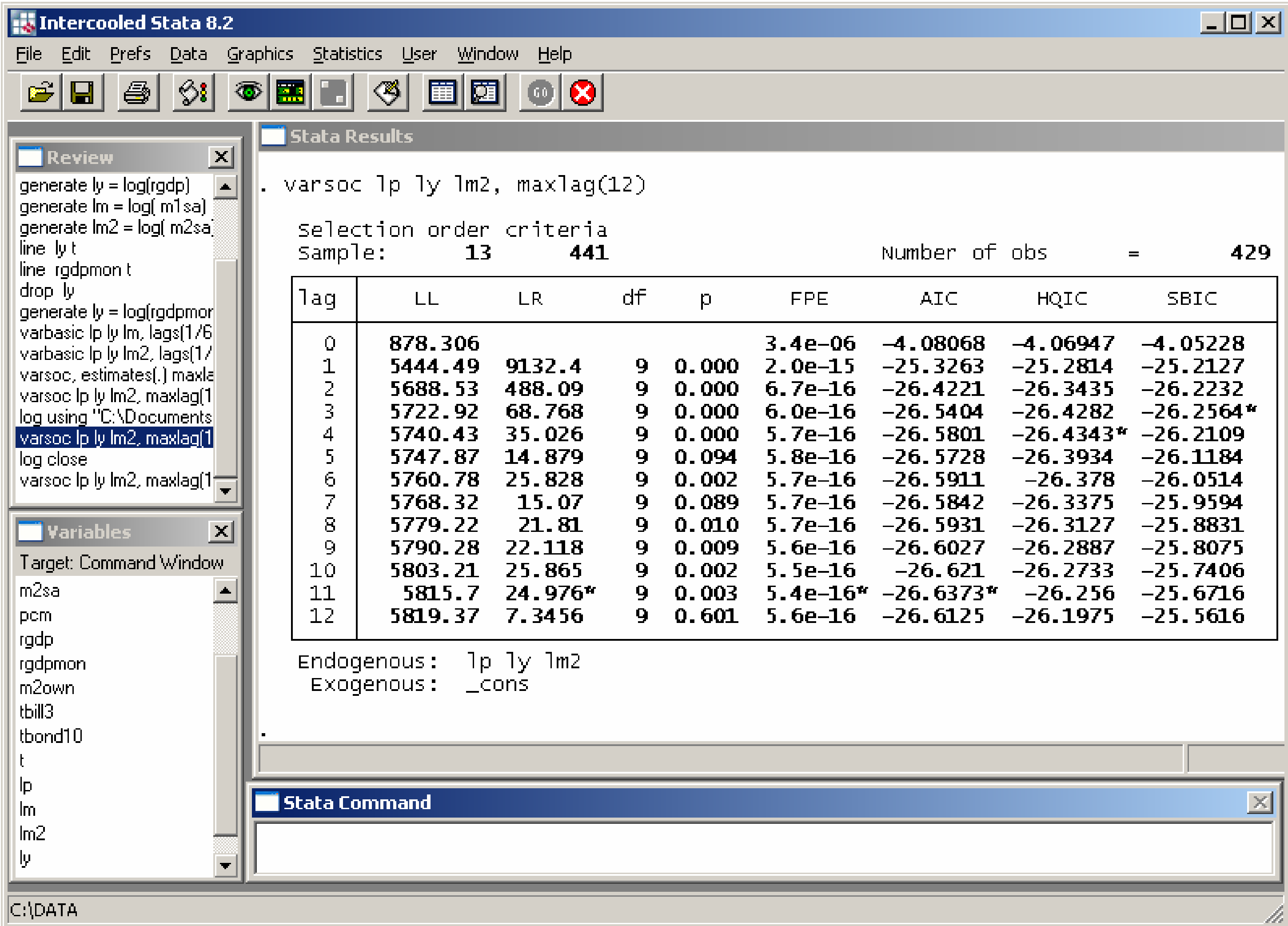
The model

$$P_t = \text{Prices}, X_t = \text{Income}, M_t = \text{Money}$$

$$Y_t = \begin{bmatrix} P_t \\ X_t \\ M_t \end{bmatrix} \quad Y_t = c + \Phi_1 \begin{bmatrix} P_{t-1} \\ X_{t-1} \\ M_{t-1} \end{bmatrix} + \Phi_2 \begin{bmatrix} P_{t-2} \\ X_{t-2} \\ M_{t-2} \end{bmatrix} + ..\Phi_p \begin{bmatrix} P_{t-p} \\ X_{t-p} \\ M_{t-p} \end{bmatrix} + a_t$$

Where $\Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}^{(1)}$ And $E(a_t) = 0 \quad E(a_t a_\tau') = \begin{cases} \Omega & t = \tau \\ 0 & t \neq \tau \end{cases}$

Steps:
- Optimal Lag Length
- Stability
- Residual Analysis
- Identification (if interest in structural form or in structural shocks)
- Impulse Response Function
- Variance Decomposition



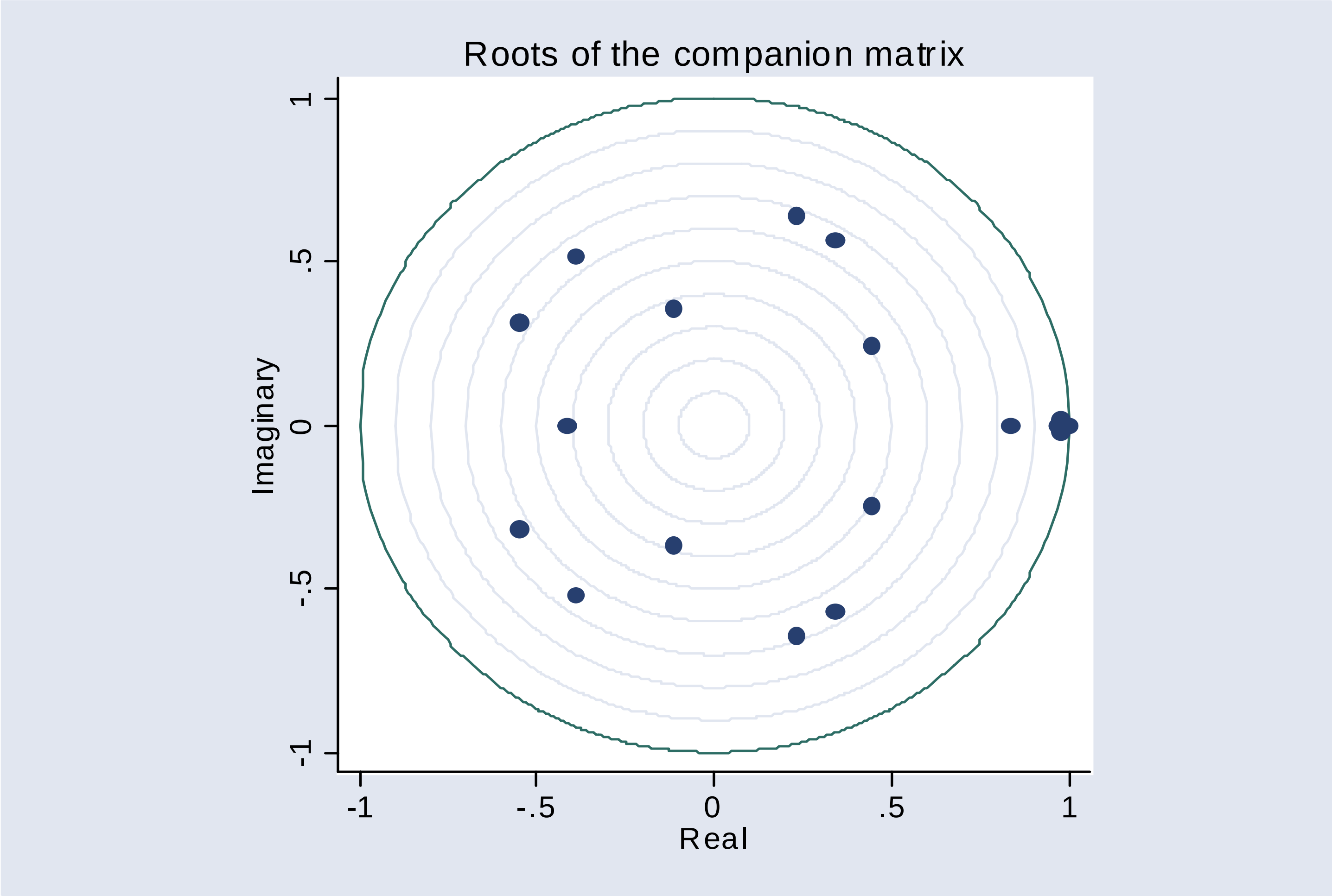

```
. varsoc lp ly lm2, maxlag(12)

  Selection order criteria
  Sample:       13       441                     Number of obs      =       429
```

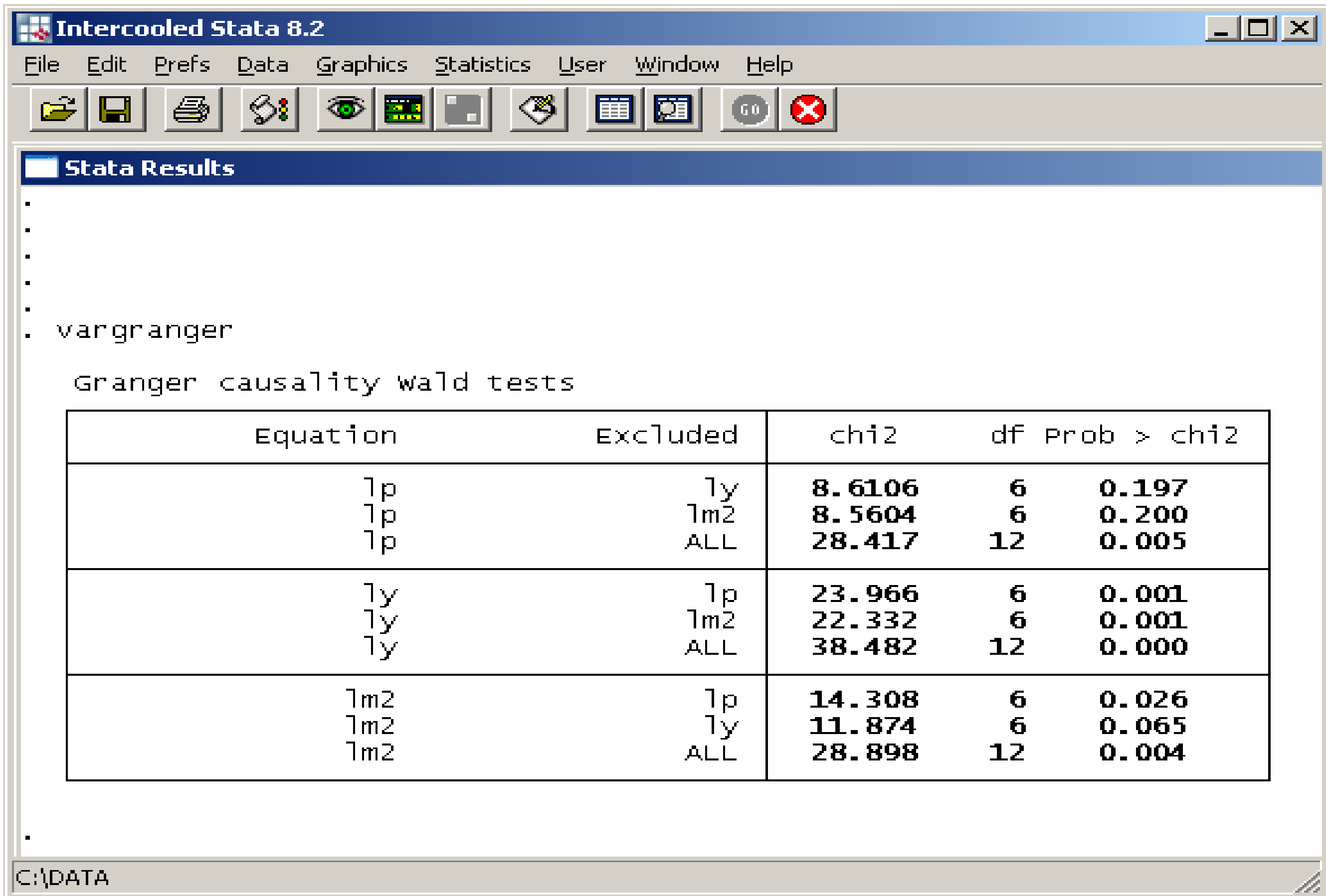
lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	878.306				3.4e-06	-4.08068	-4.06947	-4.05228
1	5444.49	9132.4	9	0.000	2.0e-15	-25.3263	-25.2814	-25.2127
2	5688.53	488.09	9	0.000	6.7e-16	-26.4221	-26.3435	-26.2232
3	5722.92	68.768	9	0.000	6.0e-16	-26.5404	-26.4282	-26.2564*
4	5740.43	35.026	9	0.000	5.7e-16	-26.5801	-26.4343*	-26.2109
5	5747.87	14.879	9	0.094	5.8e-16	-26.5728	-26.3934	-26.1184
6	5760.78	25.828	9	0.002	5.7e-16	-26.5911	-26.378	-26.0514
7	5768.32	15.07	9	0.089	5.7e-16	-26.5842	-26.3375	-25.9594
8	5779.22	21.81	9	0.010	5.7e-16	-26.5931	-26.3127	-25.8831
9	5790.28	22.118	9	0.009	5.6e-16	-26.6027	-26.2887	-25.8075
10	5803.21	25.865	9	0.002	5.5e-16	-26.621	-26.2733	-25.7406
11	5815.7	24.976*	9	0.003	5.4e-16*	-26.6373*	-26.256	-25.6716
12	5819.37	7.3456	9	0.601	5.6e-16	-26.6125	-26.1975	-25.5616

```
  Endogenous:  lp ly lm2
   Exogenous:  _cons
.
```

Stability of the VAR – Roots of the companion matrix



Command **varstable, graph** after **var** or **svar**



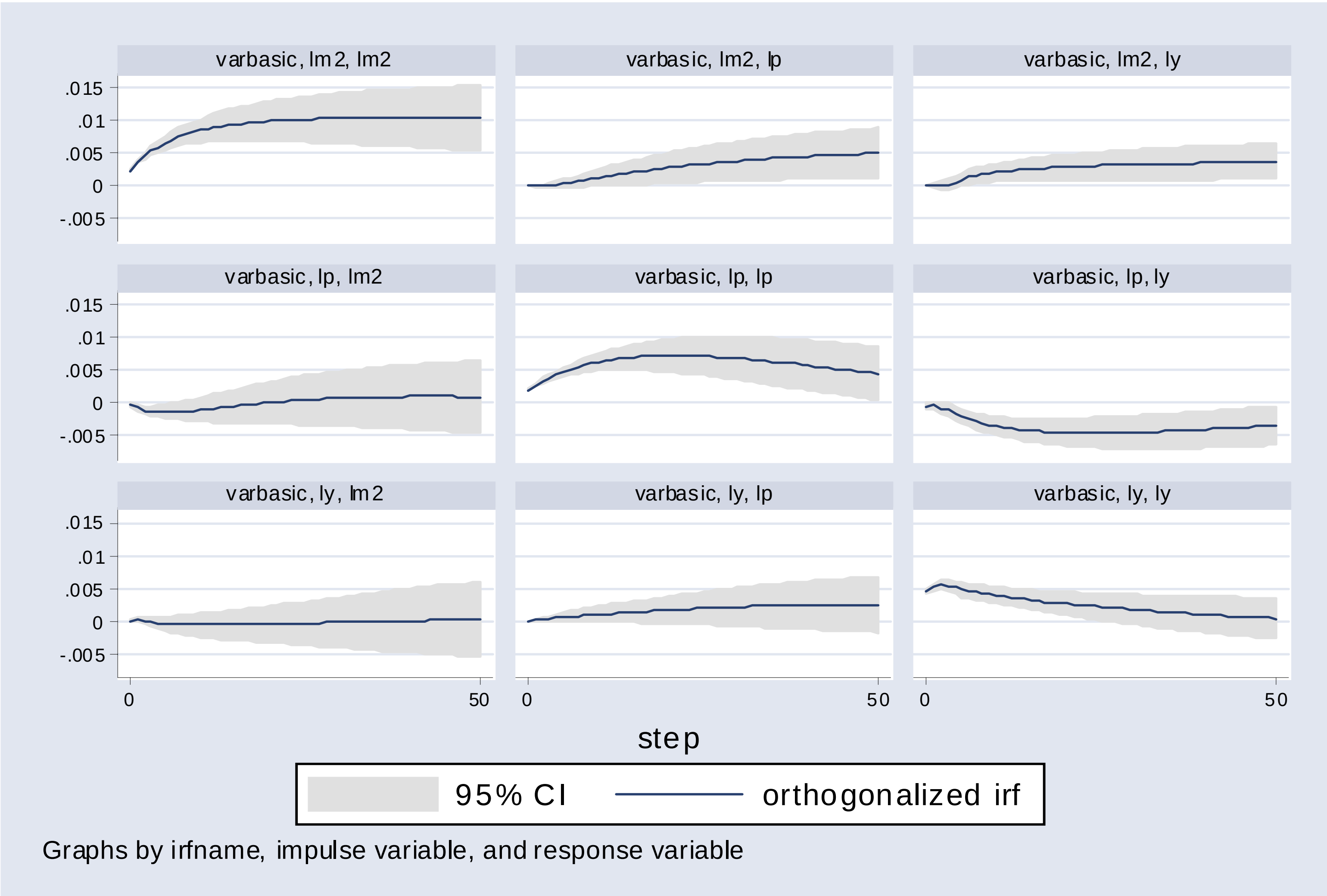
. vargranger

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
lp	ly	8.6106	6	0.197
lp	lm2	8.5604	6	0.200
lp	ALL	28.417	12	0.005
ly	lp	23.966	6	0.001
ly	lm2	22.332	6	0.001
ly	ALL	38.482	12	0.000
lm2	lp	14.308	6	0.026
lm2	ly	11.874	6	0.065
lm2	ALL	28.898	12	0.004

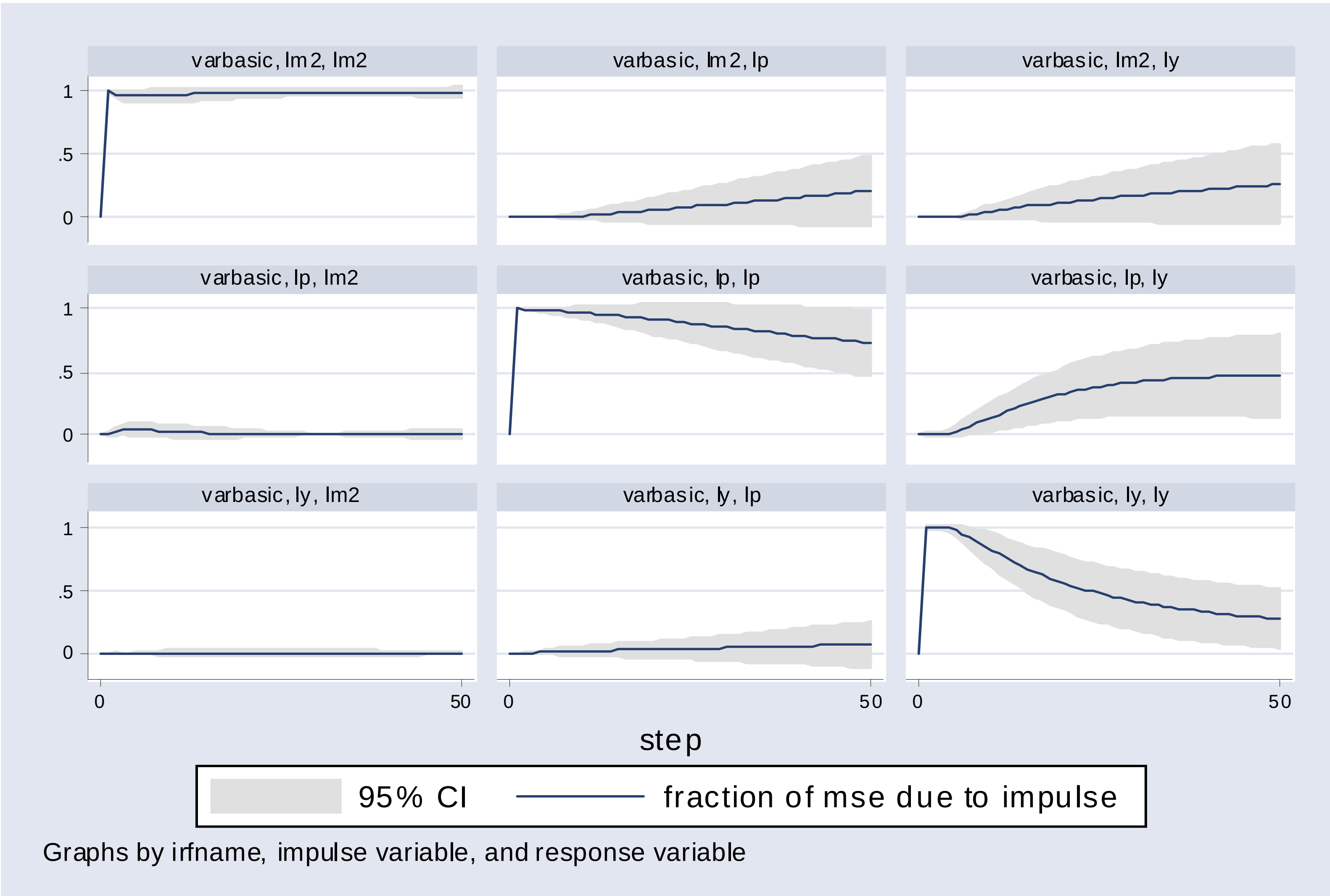
Command – vargranger after varbasic, var or svar

Cholesky Decomposition – Order Prices/Income/Money



varbasic lp ly lm2, lags(1/6) step(50) oirf

Forecast Error Variance Decomposition - FEVD



varbasic lp ly lm2, lags(1/6) step(50) fevd