Vector Autoregression (VAR) Models
Quick review

Ugo Panizza

The origins of VAR

- The VAR methodology was originally developed by Sims in a 1980 *Econometrica* paper titled "Macroeconomics and **Reality**"
- In that paper Sims said that standard multi-equation macroeconomic models are full of "**incredible**" restrictions
- His approach is the opposite of the Box and Jenkins.
  - Rather than being parsimonious, let's be profligate and insert more variables

Structural VAR

Can we estimate this?

- Consider a bivariate, first-order VAR

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

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

  - y and z are stationary and the ε are uncorrelated white noise disturbances: $\text{var}(\varepsilon_{yt}) = \sigma_y^2 \quad \text{var}(\varepsilon_{zt}) = \sigma_z^2 \quad \text{cov}(\varepsilon_{yt}\varepsilon_{zt}) = 0$
- In this system y and z affect each other; -b_{12} is the contemporaneous effect of z on y, and -b_{21} is the contemporaneous effect of y on z
  - Note that ε_y (ε_z) is an innovation to y (z), but if b_{21} (b_{12}) is different from zero ε_y (ε_z) has an indirect contemporaneous effect on z (y).

- Can we estimate this?

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

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

or primitive form

- **NO!**
  - Why? Because it's not a reduced form equation (in fact, it's called "structural" VAR). Let's rewrite it 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 $Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

VAR in **standard** form

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_\text{t}$$

- Premultiply by $\mathbf{B}^{-1}$ and get the VAR in standard (or reduced) form

$$x_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 x_{t-1} + B^{-1}\varepsilon_\text{t}$$

$$x_t = A_0 + A_1 x_{t-1} + e_\text{t}$$

$$A_0 = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}; A_1 = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix};$$

$$e_t = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$A_0 = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix}; A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

VAR in **standard** form

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$

- Premultiply by **B⁻¹** and get the VAR in <u>standard (or reduced) form</u>

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

$$x_t = A_0 + A_1 x_{t-1} + e_t \qquad A_0 = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix}; A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

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

We can estimate this with OLS equation by equation
(but with SUR could be more efficient)

The error term is composed of two shocks

$$e_t = B^{-1}\varepsilon_t, \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

VAR in **standard** form

– The error term is composed of **two shocks**

$$e_t = B^{-1}\varepsilon_t, \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$\text{Remember if } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = (ad - bc)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

VAR in **standard** form

– The error term is composed of two shocks

$$e_t = B^{-1}\varepsilon_t, \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$B^{-1} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} = (1 - b_{12}b_{21})^{-1} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{1-b_{12}b_{21}} & \dfrac{-b_{12}}{1-b_{12}b_{21}} \\ \dfrac{-b_{21}}{1-b_{12}b_{21}} & \dfrac{1}{1-b_{12}b_{21}} \end{bmatrix}$$

$$e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{1-b_{12}b_{21}} & \dfrac{-b_{12}}{1-b_{12}b_{21}} \\ \dfrac{-b_{21}}{1-b_{12}b_{21}} & \dfrac{1}{1-b_{12}b_{21}} \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} = \begin{bmatrix} \dfrac{\varepsilon_{yt} - b_{12}\varepsilon_{zt}}{1-b_{12}b_{21}} \\ \dfrac{-b_{21}\varepsilon_{yt} + \varepsilon_{zt}}{1-b_{12}b_{21}} \end{bmatrix}$$

Variances

$$e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{1-b_{12}b_{21}} & \dfrac{-b_{12}}{1-b_{12}b_{21}} \\ \dfrac{-b_{21}}{1-b_{12}b_{21}} & \dfrac{1}{1-b_{12}b_{21}} \end{bmatrix} \quad \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} = \begin{bmatrix} \dfrac{\varepsilon_{yt} - b_{12}\varepsilon_{zt}}{1-b_{12}b_{21}} \\ \dfrac{-b_{21}\varepsilon_{yt} + \varepsilon_{zt}}{1-b_{12}b_{21}} \end{bmatrix}$$

$$\operatorname{var}(e_{1t}) = E\left(\frac{\varepsilon_{yt} - b_{12}\varepsilon_{zt}}{1-b_{12}b_{21}}\right)^2 = \frac{\sigma_y^2 + b_{12}^2\sigma_z^2}{(1-b_{12}b_{21})^2}$$

$$\operatorname{var}(e_{2t}) = E\left(\frac{-b_{21}\varepsilon_{yt} + \varepsilon_{zt}}{1-b_{12}b_{21}}\right)^2 = \frac{b_{21}^2\sigma_y^2 + \sigma_z^2}{(1-b_{12}b_{21})^2}$$

$$\operatorname{cov}(e_{1t}e_{2t}) = E\left[\left(\frac{\varepsilon_{yt} - b_{12}\varepsilon_{zt}}{1-b_{12}b_{21}}\right)\left(\frac{-b_{21}\varepsilon_{yt} + \varepsilon_{zt}}{1-b_{12}b_{21}}\right)\right] = -\frac{b_{21}\sigma_y^2 + b_{12}\sigma_z^2}{(1-b_{12}b_{21})^2}$$

$$\text{Variance covariance matrix } \Sigma = \begin{vmatrix} \operatorname{var}(e_{1t}) & \operatorname{cov}(e_{1t}e_{2t}) \\ \operatorname{cov}(e_{1t}e_{2t}) & \operatorname{var}(e_{2t}) \end{vmatrix} = \begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{vmatrix}$$

Stationarity

- An AR(1) process is stationary if ϕ<1
  - Here, we have a similar concept
- Start with $x_t = A_0 + A_1 x_{t-1} + e_\mathrm{t}$
- and iterate backward

$$x_t = A_0 + A_1\left(A_0 + A_1 x_{t-2} + e_{\mathrm{t\text{-}1}}\right) + e_\mathrm{t}$$

$$x_t = (I + A_1)A_0 + A_1^2 x_{t-2} + A_1 e_{\mathrm{t\text{-}1}} + e_\mathrm{t}$$ …after n iterations:

$$x_t = (I + A_1 + A_1^2 + \ldots + A_1^n)A_0 + A_1^{n+1} x_{\mathrm{t\text{-}n\text{-}1}} + \sum_{i=0}^{n} A_1^i e_{t-i}$$

This will converge if A_1^n goes to zero when n goes to infinity.

In a VAR with two variables it can be shown that this requires that the roots of $(1\text{-}a_{11}L)(1\text{-}a_{22}L)\text{-}(a_{12}a_{21}L^2)$ lie outside the unit circle

Counting parameters

- Let's step back
- What we care about is the structural VAR, but we can only estimate the reduced form VAR
- But here we have a **problem**
- In the structural VAR we have 10 parameters (4 bs, 4 γs and the variances of the error terms)

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

Counting parameters

- Let's step back
- What we care about is the structural VAR, but we can only estimate the reduced form VAR
- But here we have a problem
- In the reduced form VAR we can estimate 9 parameters (6 a parameters, and the variances and covariance of the error terms)

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

Counting parameters

- Let’s step back
- What we care about is the structural VAR, but we can only estimate the reduced form VAR
- But here we have a problem
  - In the structural VAR we have 10 parameters
  - In the reduced form VAR we can estimate 9 parameters
  - There is no way we can recover the 10 parameters of the structural form with the 9 parameters of the reduced form
- BIG PROBLEM!!!!

Counting parameters

- How do we deal with this issue?
- We impose a **restriction**
- The usual (but probably wrong) way to do this is to impose an **ordering**
    - This is a bit funny, is this restriction "credible"?
  - Assume that you are ready to claim that y has no direct impact on z (or $b_{21}=0$).
- The VAR becomes recursive

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

Counting parameters

- **B** is no longer $\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}$ but it becomes $\begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}$
- and $B^{-1} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix}$

We made up a parameter!

Thus:

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

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

The z shock has a contemporaneous effect on y, but the y shock has no contemporaneous effect on z

Counting parameters

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

We estimate

$$y_t = \quad a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{yt}$$

$$z_t = \quad a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{zt}$$

Counting parameters

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

We estimate

$$y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{yt}$$

$$z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{zt}$$

$$\mathrm{var}(e_1) = \sigma_y^2 + b_{12}^2 \sigma_z^2$$

$$\mathrm{var}(e_2) = \sigma_z^2$$

$$\mathrm{cov}(e_1 e_2) = -b_{12}\sigma_z^2$$

Now we are fine, because we have 9 unknown and we are estimating 9 things

Identifying a VAR

- Note that we **always** have this problem.
- If we have a VAR with **N** variables, **B** is always an **NxN** matrix
  - Since the elements of the diagonal are always ones, **B** has **NxN-N** unknowns
  - We also have **N** unknown variances of ε
- So we have **NxN-N+N** unknowns ie $\mathbf{N^2}$ unknowns

- There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know.
  - **Donald Rumsfeld**

- It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so**.**
  - **Mark Twain**

Identifying a VAR

- We have to recover these unknowns with the estimates of the variance-covariance matrix (the other estimates are already used to recover the gammas and the constant)
  - Since the variance-covariance is symmetric it only has **$(N^2+N)/2$** distinct elements
  - **N** elements along the principal diagonal, **N-1** elements along the first off diagonal, **N-2** elements in the next diagonal, … one element in the corner. The total is **$(N^2+N)/2$**
- So we need **N^2** and we only have **$(N^2+N)/2$**.
  - The difference is **$N^2 - (N^2+N)/2 = (N^2-N)/2$**
- We need to impose **$(N^2-N)/2$** restrictions in the system
  - If we have two variables 1 restriction, if we have 3 variables we need 3 restrictions, if we have 4 variables we need 6 restrictions, if we have 10 variables we need 45 restrictions!

Identifying a VAR

- 5 variables

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ b_{21} & 1 & 0 & 0 & 0 \\ b_{31} & b_{32} & 1 & 0 & 0 \\ b_{41} & b_{42} & b_{43} & 1 & 0 \\ b_{51} & b_{52} & b_{53} & b_{54} & 1 \end{bmatrix}$$

(25-5)/2=10 restrictions

Count the zeros:

$b_{12}=b_{13}=b_{14}=b_{15}=b_{23}=b_{24}=b_{25}=b_{34}=$

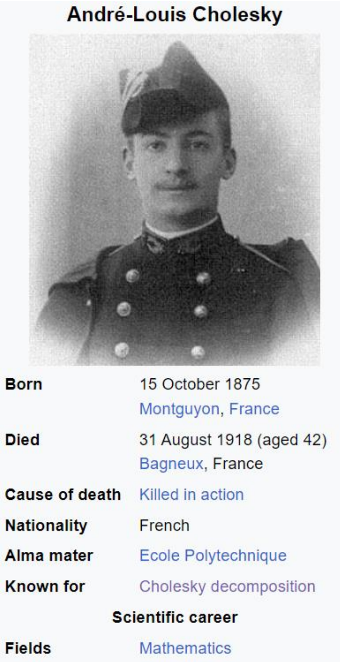
$=b_{35}=b_{45}=0$

Identifying a VAR

- If we use the recursive system adopted before and have **N** variables, we need to write the B matrix as

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{21} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{31} & b_{32} & 1 & 0 & 0 & 0 & 0 & 0 \\ b_{41} & b_{42} & b_{43} & 1 & 0 & 0 & 0 & 0 \\ b_{51} & \dots & \dots & b_{54} & 1 & 0 & 0 & 0 \\ b_{61} & \dots & \dots & \dots & b_{65} & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & 1 & 0 \\ b_{n1} & b_{n2} & b_{n3} & \dots & \dots & \dots & b_{nn-1} & 1 \end{bmatrix}$$

This is called a Choleski decomposition



Estimation

- We can estimate the system one equation at a time with OLS
  - SUR does not make the estimation more efficient because we have the same variables in the RHS
    - However, if some variables (or lags) are not included in all equations (i.e., you have a Near VAR) SUR yields more efficient estimates
- Key issues
- What variables should you include?
  - Economic theory should guide you
- How do you decide lag length?
  - If **p** is too small, the model is misspecified
  - If **p** is too big you lose *dof*
    - Note that if you have **N** variables and **p** lags, you need to estimate **n*p+1** parameters for each equation (in total we estimate $\mathbf{p*n^2+n}$ parameters) and you may soon run out of degrees of freedom
- Ordering
  - This is irrelevant for estimation purposes
    - (IRRELEVANT FOR ESTIMATIONS PURPOSES, BUT KEY FOR INTERPRETATION)

How do you decide lag length?

- The lagged variables are highly collinear so test on individual lags do not make much sense
- You may be tempted to use an F test equation by equation,
- But this is **not** appropriate because the lag length affects all the equations simultaneously
- So you want a **joint** test for all equations

How do you decide lag length?

- A proper test for cross-equation restrictions is a **likelihood ratio test**. Do as follows:
  - Start with the longest plausible lag length
    - Quarterly data, multiples of 4, monthly data, multiples of 12
  - Estimate the VAR and compute the VAR-COV matrix Σ_u (u for unrestricted). Compute the determinant of Σ_u
  - Estimate the VAR with a shorter lag length (OVER THE SAME SAMPLE PERIOD) and compute the VAR-COV matrix. Call it Σ_r (r for restricted), Compute the determinant of Σ_r
  - The likelihood ratio test is

$$LR = (T)\left(\ln|\Sigma_r| - \ln|\Sigma_u|\right) \sim \chi^2$$

  - However Sims (1980) suggests to use

$$LR = (T - c)\left(\ln|\Sigma_r| - \ln|\Sigma_u|\right) \sim \chi^2$$

    - **T** is the number of usable observations, **c** is the number of parameters to be estimated in the unrestricted system and $|\Sigma|$ is the determinant of Σ.
    - **LR** has a chi squared distribution with DOF equal to the number of restrictions in the system ***(u-r)n^2***.

How do you decide lag length?

- Example assume you have 20 years of quarterly data, a system with 3 variables (**n=3**)
- You start with **p=12**. You compute $|\Sigma_u|=15$
- Then, you estimate the model with **p=8** and find that $|\Sigma_r|=20$
  - **T = 20*4-12 =68; c = 1+12n=1+12*3=37**;
  - **LR = (68-37)(ln(20)-ln(15))=8.92**
  - DOF:
    - **(12-8)** restrictions in each variable, so **4*n^2 =36** restrictions
  - The 5% critical value (under the null that the right model is the restricted model) for a chi squared with 36 degrees of freedom is approximately 50
  - So, we do not reject the null that the appropriate lag length is 8, so I can try with 4 lags

$$LR = (T-c)\left(\ln\left|\Sigma_r\right| - \ln\left|\Sigma_u\right|\right) \sim \chi^2_{(u-r)n^2}$$

How do you decide lag length?

- There are some problems with the LR test
  - You can only make pairwise comparisons
  - It requires normally distributed errors in each equation
  - It may say that you are OK when you compare 12 with 8 lags and when you compare 8 with 4 but that you are not OK when you compare 12 with 4
    - More in general, in econometric the fact that x is SS and y is not SS does not mean that x-y is SS

How do you decide lag length?

- An alternative is to use one of the following information criteria
  - $MAIC = \ln|\Sigma| + 2k/T$
  - $MSBIC = \ln|\Sigma| + (k/T)\ln(T)$
  - $MHQIC = \ln|\Sigma| + (2k/T)\ln(\ln(T))$
    - **k** total number of regressors in the system: **k=n*p+n**
  - Note that these are not tests, they just give you a ranking

Once we estimate the model what do we do with it?

- Maybe you have a system with three variables and 2 lags
- We can stare really hard at the parameters

$$y_t = \gamma_{11} y_{t-1} + \gamma_{12} y_{t-2} + \gamma_{13} z_{t-1} + \gamma_{14} z_{t-2} + \gamma_{15} m_{t-1} + \gamma_{16} m_{t-2} + \varepsilon_{yt}$$

$$z_t = \gamma_{21} y_{t-1} + \gamma_{22} y_{t-2} + \gamma_{23} z_{t-1} + \gamma_{24} z_{t-2} + \gamma_{25} m_{t-1} + \gamma_{26} m_{t-2} + \varepsilon_{zt}$$

$$m_t = \gamma_{31} y_{t-1} + \gamma_{32} y_{t-2} + \gamma_{33} z_{t-1} + \gamma_{34} z_{t-2} + \gamma_{35} m_{t-1} + \gamma_{36} m_{t-2} + \varepsilon_{mt}$$

Once we estimate the model what do we do with it?

- Maybe you have a system with three variables and 2 lags
- We can stare really hard at the parameters

$$y_t = \gamma_{11} y_{t-1} + \gamma_{12} y_{t-2} + \gamma_{13} z_{t-1} + \gamma_{14} z_{t-2} + \gamma_{15} m_{t-1} + \gamma_{16} m_{t-2} + \varepsilon_{yt}$$

$$z_t = \gamma_{21} y_{t-1} + \gamma_{22} y_{t-2} + \gamma_{23} z_{t-1} + \gamma_{24} z_{t-2} + \gamma_{25} m_{t-1} + \gamma_{26} m_{t-2} + \varepsilon_{zt}$$

$$m_t = \gamma_{31} y_{t-1} + \gamma_{32} y_{t-2} + \gamma_{33} z_{t-1} + \gamma_{34} z_{t-2} + \gamma_{35} m_{t-1} + \gamma_{36} m_{t-2} + \varepsilon_{mt}$$

- Keep staring!
  - You will find that parameters at some lags are positive, at other lags they are negative. You will have no clue about the total effect
  - Some of them are SS other are not, but huge amount of multicollinerity
- What you want to know is how a **shock to one** variable affects the **other variables** in the system

EXAMPLE

We use a dataset that contains quarterly data on national income, private consumption, and nonresidential investment for the United States from 1947 through the first quarter of 2006.

The variables pincome, pinvestment, and pconsumption contain percent changes at an annualized rate.

The three series are correlated for several reasons. For example, if residents' incomes increase, then they will likely go out and purchase more goods. Increased investment today may lead to higher productivity and income later and hence higher future consumption.

```
. var pincome pinvestment pconsumption, lags(1/2)

Vector autoregression

Sample:  1947q4 - 2006q1                         No. of obs      =       234
Log likelihood = -1101.056                       AIC             =  9.590226
FPE            =  2.934825                       HQIC            =  9.715255
Det(Sigma_ml)  =  2.452498                       SBIC            =  9.900318

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
pincome              7     1.15742   0.2207   66.26631   0.0000
pinvestment          7     2.29811   0.2904   95.77788   0.0000
pconsumption         7     .805211   0.1142   30.1806    0.0000
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
pincome      |
     pincome |
         L1. |   .4389707    .075006     5.85   0.000     .2919617    .5859797
         L2. |    .029891   .0791689     0.38   0.706    -.1252772    .1850592
             |
 pinvestment |
         L1. |  -.0309568   .0379123    -0.82   0.414    -.1052635    .0433499
         L2. |  -.0054707   .0349477    -0.16   0.876     -.073967    .0630256
             |
pconsumption |
         L1. |    .125813   .1067887     1.18   0.239    -.0834891    .3351151
         L2. |   .0838814   .1018411     0.82   0.410    -.1157236    .2834863
             |
       _cons |   .7741577    .170533     4.54   0.000     .4399192    1.108396
-------------+----------------------------------------------------------------
pinvestment  |
     pincome |
         L1. |   .6284843    .148928     4.22   0.000     .3365908    .9203779
         L2. |  -.1449329   .1571937    -0.92   0.357    -.4530269    .1631612
             |
 pinvestment |
         L1. |   .1753363   .0752767     2.33   0.020     .0277966    .3228759
         L2. |  -.0040007   .0693905    -0.06   0.954    -.1400035    .1320021
             |
pconsumption |
         L1. |   .3245674   .2120343     1.53   0.126    -.0910122    .7401469
         L2. |   .6925128   .2022106     3.42   0.001     .2961873    1.088838
             |
       _cons |  -.7613954   .3386016    -2.25   0.025    -1.425042   -.0977484
-------------+----------------------------------------------------------------
pconsumption |
     pincome |
         L1. |   .0123896   .0521814     0.24   0.812    -.0898841    .1146633
         L2. |  -.0769745   .0550776    -1.40   0.162    -.1849245    .0309756
             |
 pinvestment |
         L1. |   .0029507   .0263755     0.11   0.911    -.0487443    .0546456
         L2. |  -.0438997    .024313    -1.81   0.071    -.0915524     .003753
             |
pconsumption |
         L1. |    .008587   .0742926     0.12   0.908    -.1370238    .1541979
         L2. |   .3824941   .0708506     5.40   0.000     .2436295    .5213587
             |
       _cons |   .6917505   .1186393     5.83   0.000     .4592218    .9242793
------------------------------------------------------------------------------
```

We use two lags

Can you **interpret** the regression coefficients? (assuming that you can see them)

Impulse response function

- We want to express the **variables** in the system as a **function of the shocks**
- This can be done because, like an AR has an MA representation, a VAR has a VMA representation
- The **VMA** is an essential feature of Sims's methodology because it allows to **trace the effect of the shocks** on the variable contained in the VAR system

Impulse response function

- We start with

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

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

- We rewrite it as:

$$x_t = (I + A_1 + A_1^2 + \ldots + A_1^n) A_0 + A_1^{n+1} x_{\text{t-n-1}} + \sum_{i=0}^{n} A_1^i e_{t-i}$$

- If the stability conditions are met, we can rewrite it as

$$x_t = \mu + \sum_{i=0}^{n} A_1^i e_{t-i}$$

with $\mu = [\bar{y}\ \bar{z}]'$, and

$$\bar{y} = [a_{10}(1 - a_{22}) + a_{12} a_{20}] / \Delta \quad \text{unconditio nal mean of y}$$

$$\bar{z} = [a_{20}(1 - a_{11}) + a_{21} a_{10}] / \Delta \quad \text{unconditio nal mean of z}$$

$$\Delta = (1 - a_{11})(1 - a_{22}) - a_{12} a_{21}$$

Impulse response function

- In other words, we go from

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

- to

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix}$$

Impulse response function

- Rewrite the errors as

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

- And plug it into

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix}$$

- We get

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{1-b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix}$$

Impulse response function

- Now, this is becoming messy. Write

$$\phi_i = \frac{A_1^i}{1-b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

- And plug it into

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{1-b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix}$$

- We get

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix}$$

Impulse response function

- Now, this is becoming messy. Write

$$\phi_i = \frac{A_1^i}{1-b_{12}b_{21}}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

More compact notation

- And plug it into

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{1-b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix}$$

- We get

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix} \qquad x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$

Impulse response function

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$

- This is cool because you can use the ϕ to look at the effect of ε on the entire time paths of y and z

- $\phi_{jk}(0)$ are the impact multipliers.
  - For instance, $\boldsymbol{\phi_{12}(0)}$ is the impact of a change in ε_{zt} on y_t.
  - Similarly, $\boldsymbol{\phi_{12}(1)}$ is the impact of a change in ε_{zt} on y_{t+1}, and $\boldsymbol{\phi_{12}(2)}$ is the impact of a change in ε_{zt} on y_{t+2}
  - You can find the **cumulative** impact by summing the various ϕ. After n periods, the cumulate effect of ε_{zt} on the $\{y_t\}$ sequence is

$$\sum_{i=0}^{n} \phi_{12}(i)$$

When *n* goes to infinite, we have the long-run multiplier

Impulse response function

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$

- Again:
- The $\phi_{jk}(0)$ are the impact multipliers
- The $\sum_{i=0}^{n} \phi_{12}(i)$ are the cumulative multipliers
- The $\sum_{i=0}^{n} \phi_{12}(i)$ for n➔oo are the long-run multipliers

Impulse response function

- The four sets of coefficients $\phi_{11}(i)$; $\phi_{22}(i)$; $\phi_{21}(i)$; $\phi_{12}(i)$ are called impulse response functions (IRF)
- Plotting the impulse response function is a practical way to look at the impact of ε on y and z.
- Note that in order to build the IRF and trace out the effect of the pure shocks we need to know all the parameters of the **primitive** system
- But we **don't know** them!
  - The estimated VAR is underidentified
    - We are missing **$(N^2 - N)/2$** parameters

Identifying IRF

- One possibility is to use the same Cholesky decomposition discussed before
- Remember that with our n=2 VAR the errors were $$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$
- If we set $b_{21}=0$

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \qquad \begin{aligned} e_{1t} &= \varepsilon_{yt} - b_{12}\varepsilon_{zt} \\ e_{2t} &= \varepsilon_{zt} \end{aligned}$$

Identifying IRF

- By doing this we **equate** e_{2t} (which we can estimate) to ε_{zt}.
- Then, we can recover b_{12} by computing the variance of e_2 and the covariance of $e_1 e_2$ and recalling that

$$\operatorname{cov}(e_1 e_2) = -b_{12}\sigma_z^2$$

- Notice that the Cholesky decomposition does not allow for any direct effect of ε_y on z. But, we still have an indirect lagged effect because y_{t-1} has an effect on z_t.

Identifying IRF

- So, the decomposition generates an **asymmetry** in the system.
  - One shock has a contemporaneous effect on all variables but the other shock has only a lagged effect on one of the two variables
- This is why this is called an ordering

$$e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$$

$$e_{2t} = \varepsilon_{zt}$$

A unit shock to z causes z to jump by one unit and y to jump by $-b_{12}$

A unit shock to y causes y to jump by one unit but has no contemporaneous effect on z

- In our case, z is said to be **causally prior** to y

Example

$$e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$$
$$\text{cov}(e_1 e_2) = -b_{12}\sigma_z^2$$

- Assume we estimate a VAR and we find $a_{10}=a_{20}=0$; $a_{11}=a_{22}=0.7$, and $a_{12}= a_{21}=0.2$. $var(e_1)=var(e_2)=1$, $cov(e_1,e_2)=0.8$

$$\begin{aligned} y_t &= 0.7y_{t-1} + 0.2z_{t-1} + e_{1t} & e_{1t} &= \varepsilon_{yt} + 0.8\varepsilon_{zt} \\ z_t &= 0.2y_{t-1} + 0.7z_{t-1} + e_{2t} & e_{2t} &= \varepsilon_{zt} \end{aligned}$$

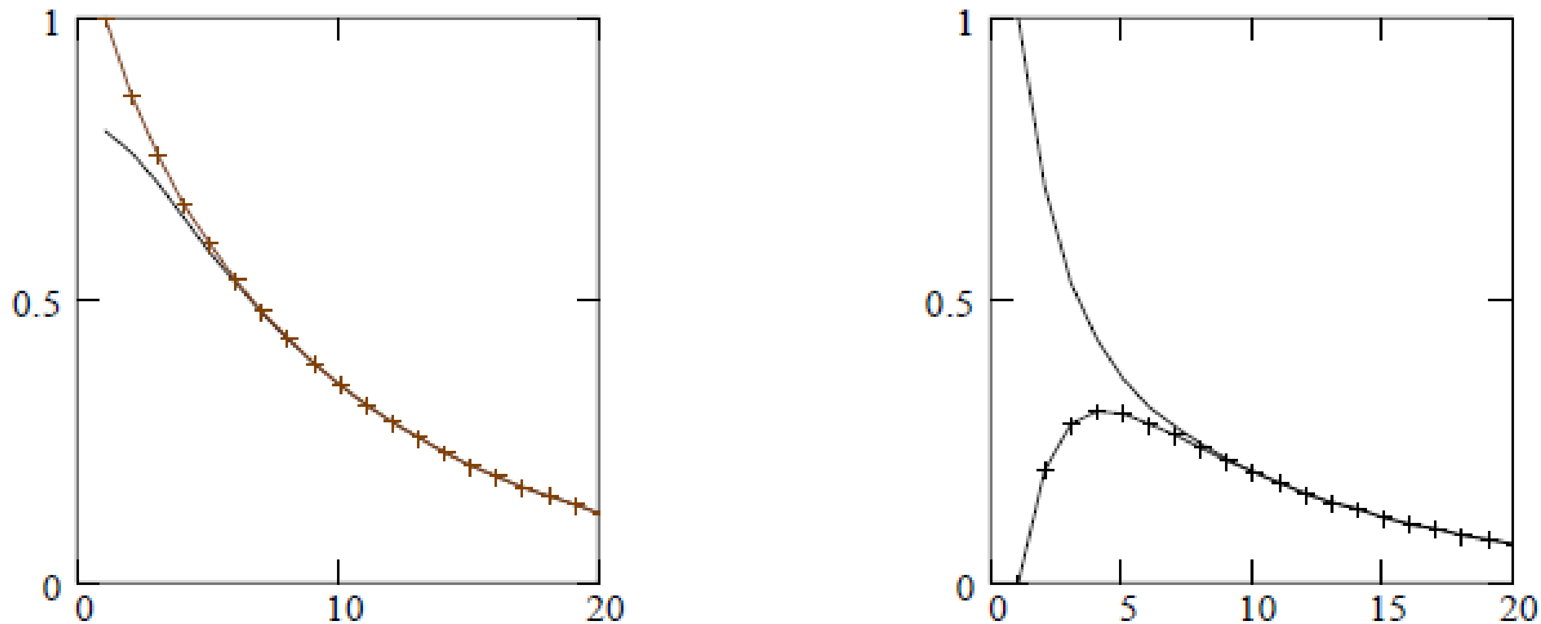
  – We have a unit shock to ε_{zt}.
    - In period 0 z goes up by one and y goes up by 0.8. In period 1 z goes up by 0.7+0.2*0.8=0.86 and y goes up by 0.7*0.8+0.2*1=0.76. In period 2 z goes up by 0.2*0.76+0.7*0.86=0.754 and y goes up by 0.7*0.76+0.2*0.86=0.704. They converge

  – We have a unit shock to ε_{yt}.
    - In period 0 y goes up by one and z doesn't change. In period 1 z goes up by 0.2 and y goes up by 0.7. In period 2 z goes up by 0.2*0.7+0.7*0.2=0.28 and y goes up by 0.7*0.7+0.2*0.2=0.53. They converge

Model 1: $$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}$$

Response to ε_{zt} shock

Response to ε_{yt} shock



(a) *(b)*

Legend: Solid line = $\{y_t\}$ sequence Cross-hatch = $\{z_t\}$ sequence

$u_t = 0.8v_t + \varepsilon_{yt}$ and $v_t = \varepsilon_{zt}$

- What happens if we **reverse** the Cholesky decomposition (i.e.. we set b_{21}=0).
  - In this case, since **we assumed symmetry**, we would just reverse the IRF of the two shocks
  - But, in **general**, this is **not** the case
- What happens if we set $a_{12} = a_{21} = -0.2$

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}$$

- The importance of the ordering depends on the magnitude of $corr(e_1,e_2)$
- If this correlation is **zero**, the ordering is **irrelevant** (because $\varepsilon_{yt}=e_{1t}$ and $\varepsilon_{zt}=e_{2t}$)
  - If there is no correlation $b_{12}=b_{21}=0$
- The higher the correlation the more important the ordering.
  - If corr(e_1,e_2)=1 and we set b_{21}=0 then $\varepsilon_{zt}=e_{2t}=e_{1t}$. If we set b_{12}=0 then $\varepsilon_{yt}=e_{1t}=e_{2t}$.
  - What if corr(e_1,e_2)=-1?

Test the hypothesis that $\rho_{12}=0$

- Assuming that the correlation coefficient is distributed normally with mean zero and standard deviation of $\mathbf{1/(T^{0.5})}$ you can use a *t* table to test if ρ is different from zero

How do we choose the ordering

- Economic theory
- Try with different orderings and if the results change dramatically investigate more

How do you build confidence intervals for the IRF?

- IRF are built using **estimated** coefficients and since each coefficient is estimated imprecisely, the IRF also contains errors
- How do we test the significance of IRF?
  - Delta method
    - The delta method expands a function of a random variable around its mean, usually with a one-step Taylor approximation, and then takes the variance.
    - If we want to approximate the variance of $G(X)$ where X is a random variable with mean μ and $G()$ is differentiable, we can use:
    - $G(X) = G(\mu) + (X\text{-}\mu)G'(\mu)$ +remainder
    - $Var(G(X)) = Var(X)*[G'(\mu)]^2$ (approximately)
    - This is a good approximation if X has a high probability of being close enough to its mean (μ)
  - Bootstrapping

Inference with montecarlo simulation

- Use the estimated variance covariance matrix $\widehat{\Sigma}$ to generate a sequence of forecast errors for each variable in your VAR
- For $j = 1, \ldots M$ generate the following sequence of forecast errors $e_t^{(j)} \tilde{} i.i.d \mathcal{N}(0, \widehat{\Sigma}), t = 1, \ldots T*$, with $T* >>> T$
- Use the generated forecast errors to generate artificial data using the parameters of the estimated model

$$y_t^{(j)} = \widehat{c} + \sum_{i=1}^{p} \widehat{A}_i y_{t-i}^{(j)} + e_t^{(j)}$$

Inference with montecarlo simulation

1. Use the generated forecast errors to generate artificial data using the parameters of the estimated model

$$y_t^{(j)} = \widehat{c} + \sum_{i=1}^{p} \widehat{A}_i y_{t-i}^{(j)} + e_t^{(j)}$$

2. Set the initial conditions $y_1^{(j)}, ..., y_p^{(j)}$ equal to the observed values and then discard the observations at the beining of the period and only keep the last T observatiosn.
3. Use the generated data to estimate the IRF functions
4. Build confidence intervalst using the quantiles of the empirical distrbution fo each element $\widehat{b}_{ij,s}^{(j)}$ of $\widehat{B}(L)^{(j)}$ for $j = 1...m$

Variance decomposition

- Unrestricted VARs are overparametrized and they are not very good for forecasting
- But, we can learn about the relationship among the variables in the system by looking at the properties of the forecast error
- The forecast error variance decomposition shows the **proportion** of the movements in a sequence which is due to its own shock versus the shocks to other variables
- If shocks to y explain none of the forecast errors variance of x, we say that x is exogenous with respect to y

Variance decomposition

- Suppose we estimate our model
- $x_t = A_0 + A_1 x_{t-1} + e_t$
- We use our estimates of A_0 and A_1 to take the conditional expectation
- $Ex_{t+1}=A_0+A_1x_t$
- The one-step-ahead forecast error is
- $x_{t+1}-Ex_{t+1}=e_{t+1}$
- Updating two periods
- $x_{t+2}=A_0+A_1x_{t+1}+e_{t+2}=A_0+A_1(A_0+A_1x_t+e_{t+1})+e_{t+2}$
- $Ex_{t+2}=(I+A_1)\ A_0+A_1^2x_t$
- The two-step-ahead forecast error is
- $x_{t+2}-Ex_{t+2}=A_1e_{t+1}+e_{t+2}$

NB:

ALL expectations are taken at time *t*

Variance decomposition

- The *n*-step-ahead forecast is

$$Ex_{t+n}=(I+A_1+A^2{}_1+A^3{}_1+A^4{}_1+\ldots A^{n-1}{}_1)\,A_0+A_1{}^n x_t$$

and the forecast error is

$$x_{t+n}\text{-}Ex_{t+n}=A_1{}^{n-1}e_{t+1}+A_1{}^{n-2}e_{t+2}+..+A_1{}^2e_{t+n-2}+A_1{}^1e_{t+n-1}+e_{t+n}$$

NB:

ALL expectations are taken at time *t*

Variance decomposition

- An alternative way is to start from the VMA form of the structural model

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix}$$

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$ **OR** $$x_{t+n} = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t+n-i}$$

- and write the n-step-ahead forecast error as

$$x_{t+n} - Ex_{t+n} = \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i}$$

NB:

ALL expectations are taken at time *t*

Variance decomposition

- Focus on y_t, the **n**-step-ahead forecast error is

$$y_{t+n}-Ey_{t+n}=\phi_{11}(0)\varepsilon_{yt+n}+\phi_{11}(1)\varepsilon_{yt+n-1}+..+\phi_{11}(n-1)\varepsilon_{yt+1}+$$
$$+\ \phi_{12}(0)\varepsilon_{zt+n}+\phi_{12}(1)\varepsilon_{zt+n-1}+..+\phi_{12}(n-1)\varepsilon_{zt+1}$$

- And its variance:

$$\sigma_y(n)^2=\sigma_y^2(\phi_{11}(0)^2+\phi_{11}(1)^2+..+\phi_{11}(n-1)^2)+$$
$$+\ \sigma_z^2(\phi_{12}(0)^2+\phi_{12}(1)^2+..+\phi_{12}(n-1)^2)+$$

  - Since $\phi_{ij}(i)^2$ are non-negative the variance of the forecast error increases with **n**

Variance decomposition

- And we can decompose the variance of the forecast error into proportions due to shocks in the ε_{yt} and ε_{zt} sequences

$$1=[\sigma_y^2(\phi_{11}(0)^2+\phi_{11}(1)^2+..+\phi_{11}(n-1)^2)]/\sigma_y(n)^2+$$
$$+[\sigma_z^2(\phi_{12}(0)^2+\phi_{12}(1)^2+..+\phi_{12}(n-1)^2)]/\sigma_y(n)^2+$$

- The <u>forecast error variance decomposition</u> shows the proportion of the movements in a sequence which is due to its own shock versus the shocks to other variables

Variance decomposition

- If shocks to y explain **none** of the forecast errors variance of x, we say that x is **exogenous** with respect to y
- If shocks to y explain 100% the forecast errors variance of x at all **n**, then we can say that x is fully endogenous with respect to y
- Normally, a variable explains most of its variance at short horizons and a decreasing share at longer horizons

Variance decomposition

- Note that in order to do a variance decomposition we also need to assume an ordering or some other set of restrictions
  – If not, we cannot recover the ϕ
- If we set $b_{21}=0$ we are forcing ε_{zt} to explain 100% of one-step ahead forecast error variance of z_t. (the opposite happens if we set b_{12}=0)
- This dramatic effect of ordering should disappear at longer forecast horizons

Variance decomposition

- It is useful to check what happens with different orderings
- If things remain different even at long forecast horizons we really need to think hard about the ordering
- Again, the magnitude of ρ_{12} is very important
  - If the correlation among innovations is small, identification is not a big issue and alternative orderings should yield similar results

Innovation accounting

- Impulse response functions and forecast error variance decomposition are jointly called **innovation accounting**
- They are useful to examine the **interrelation** among variables

Now, let's do it (example 1)

- We will follow Leeper, Sims and Zha (LSZ) (1996) "What Does Monetary Policy Do?"
- We will use monthly data (July 1959-March 1996 period) and consider a VAR with 6 variables
  1) Log of US output (ly)
  2) Log of US prices (lp)
  3) Log of commodity prices ($lpcm$)
  4) US Fed funds rate ($usff$)
  5) Seasonally adjusted total bank reserves ($smtr$)
  6) Seasonally adjusted non-borrowed reserves ($smnbr$)

How many lags?

- We have a list of variables, but how many lags should we include?
- We can use information criteria
- STATA has a command called `VARSOC` which produces various information criteria and suggests the best lag length
  - IC produced by VARSOC:
    - Final prediction error (FPE); Akaike's information criterion (AIC); Schwarz's Bayesian information criterion (SBIC); Hannan and Quinn information criterion (HQIC)
  - AIC chooses longer lag length; SBIC and HQIC are more parsimonious
  - With monthly data people usually look at SBIC

```
. varsoc lpcm lp ly usff smtr smnbr, maxlag(24)
```

Up to two years

```
   Selection-order criteria
   Sample:  1961m7 - 1996m3                    Number of obs      =       417
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	589.669				2.5e-09	-2.79937	-2.77643	-2.74134
1	6776.36	12373	36	0.000	3.8e-22	-32.2991	-32.1385	-31.8929
2	6914.74	276.75	36	0.000	2.3e-22	-32.7901	-32.4919*	-32.0357*
3	6961.01	92.537	36	0.000	2.2e-22*	-32.8394*	-32.4035	-31.7368
4	6985.18	48.34	36	0.082	2.3e-22	-32.7826	-32.2091	-31.3319
5	7003.4	36.444	36	0.448	2.5e-22	-32.6974	-31.9861	-30.8984
6	7033.88	60.97	36	0.006	2.6e-22	-32.6709	-31.822	-30.5238
7	7061.34	54.907	36	0.023	2.7e-22	-32.6299	-31.6434	-30.1346
8	7095.17	67.671	36	0.001	2.8e-22	-32.6195	-31.4953	-29.7761
9	7122.55	54.759	36	0.023	2.9e-22	-32.5782	-31.3163	-29.3865
10	7147.99	50.881	36	0.051	3.0e-22	-32.5275	-31.128	-28.9877
11	7175.42	54.85	36	0.023	3.2e-22	-32.4864	-30.9493	-28.5984
12	7215.76	80.688	36	0.000	3.1e-22	-32.5073	-30.8324	-28.2711
13	7234.48	37.44	36	0.403	3.4e-22	-32.4244	-30.6119	-27.84
14	7269.85	70.745	36	0.000	3.5e-22	-32.4214	-30.4712	-27.4888
15	7292.52	45.33	36	0.137	3.7e-22	-32.3574	-30.2696	-27.0767
16	7329.93	74.82	36	0.000	3.7e-22	-32.3642	-30.1387	-26.7353
17	7363.39	66.912	36	0.001	3.8e-22	-32.352	-29.9889	-26.3749
18	7395.47	64.166	36	0.003	3.9e-22	-32.3332	-29.8324	-26.0079
19	7423.11	55.283	36	0.021	4.2e-22	-32.2931	-29.6547	-25.6196
20	7456.08	65.945	36	0.002	4.3e-22	-32.2786	-29.5025	-25.2569
21	7488.65	65.129	36	0.002	4.4e-22	-32.2621	-29.3484	-24.8923
22	7520.79	64.29	36	0.003	4.6e-22	-32.2436	-29.1922	-24.5256
23	7535.95	30.306	36	0.736	5.2e-22	-32.1436	-28.9546	-24.0775
24	7568.27	64.64*	36	0.002	5.4e-22	-32.126	-28.7993	-23.7116

```
   Endogenous:  lpcm lp ly usff smtr smnbr
    Exogenous:  _cons
```

Let’s try with a VAR2

```
var lpcm lp ly usff smtr smnbr, lags(2)
```

Does the ordering matter?

Not, now, but it will matter later

```
Vector autoregression

Sample:  1959m9 - 1996m3                       No. of obs      =       439
Log likelihood =  5997.578                     AIC             = -27.13248
FPE            =  6.63e-20                     HQIC            =  -26.9783
Det(Sigma_ml)  =  5.48e-20                     SBIC            =  -26.7417

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
lpcm                  7     .036029   0.9942   75721.43   0.0000
lp                    7     .003304   1.0000   1.39e+07   0.0000
ly                    7     .007451   0.9995   841072.9   0.0000
usff                  7     1.02139   0.9101   4445.641   0.0000
smtr                  7     .012848   0.9994   774436.8   0.0000
smnbr                 7     .029101   0.9972   155974.2   0.0000
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lpcm         |
        lpcm |
         L2. |   1.008313   .0121597    82.92   0.000     .9844806    1.032146
          lp |
         L2. |  -.0583564   .0192418    -3.03   0.002    -.0960696   -.0206433
          ly |
         L2. |   .0942555   .0256698     3.67   0.000     .0439436    .1445674
        usff |
         L2. |  -.0021731   .0010309    -2.11   0.035    -.0041937   -.0001526
        smtr |
         L2. |   .0058835   .0512282     0.11   0.909    -.0945219    .1062889
       smnbr |
         L2. |  -.0023575    .050335    -0.05   0.963    -.1010123    .0962974
       _cons |   -.567391   .1608495    -3.53   0.000    -.8826501   -.2521318
-------------+----------------------------------------------------------------
lp           |
        lpcm |
         L2. |   .0156976   .0011152    14.08   0.000     .0135119    .0178833
          lp |
```

Is the model stable?

- We can use the command `VARSTABLE` to check whether all the roots are inside the unit circle

```
. varstable

  Eigenvalue stability condition
```

Eigenvalue	Modulus
-.9985384	.998538
.9985384	.998538
-.9882019	.988202
.9882019	.988202
.9865945 + .02854773*i*	.987007
.9865945 - .02854773*i*	.987007
-.9865945 + .02854773*i*	.987007
-.9865945 - .02854773*i*	.987007
.9322356	.932236
-.9322356	.932236
-.8574726	.857473
.8574726	.857473

Complex roots

```
All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.
```

Are the residuals normal?

```
. varnorm
```

Jarque-Bera test

Equation	chi2	df	Prob > chi2
lpcm	135.315	2	0.00000
lp	30.781	2	0.00000
ly	21.959	2	0.00002
usff	3243.001	2	0.00000
smtr	22.239	2	0.00001
smnbr	2.5e+04	2	0.00000
ALL	2.9e+04	12	0.00000

Skewness test

Equation	Skewness	chi2	df	Prob > chi2
lpcm	.27377	5.484	1	0.01919
lp	-.20688	3.131	1	0.07680
ly	.06776	0.336	1	0.56218
usff	-.00208	0.000	1	0.98577
smtr	.41482	12.590	1	0.00039
smnbr	-3.5873	941.563	1	0.00000
ALL		963.105	6	0.00000

Kurtosis test

Equation	Kurtosis	chi2	df	Prob > chi2
lpcm	5.6642	129.831	1	0.00000
lp	4.2295	27.650	1	0.00000
ly	4.0873	21.623	1	0.00000
usff	16.315	3243.000	1	0.00000
smtr	3.7263	9.649	1	0.00189
smnbr	39.332	2.4e+04	1	0.00000
ALL		2.8e+04	6	0.00000

Skewness

Kurtosis

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right),$$

Normal variables have skewness=0 and kurtosis=3

We always reject the null of normality

Are the residuals autocorrelated?

```
. varlmar, mlag(12)
```

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	1.4e+03	36	0.00000
2	218.8432	36	0.00000
3	115.2873	36	0.00000
4	103.1177	36	0.00000
5	133.4064	36	0.00000
6	127.3183	36	0.00000
7	122.3013	36	0.00000
8	120.5973	36	0.00000
9	129.8093	36	0.00000
10	103.7364	36	0.00000
11	68.5218	36	0.00087
12	62.9145	36	0.00362

H0: no autocorrelation at lag order

OK, diagnostics really suck
The residuals are not normal and
they are autocorrelated

Since we have annual data,
lets' try with a model with 12 lags
(VAR(12))

We always reject H0, we have big time
Autocorrelation!

```
. var lpcm lp ly usff smtr smnbr, lags(1/12)

Vector autoregression

Sample:  1960m7 - 1996m3                        No. of obs      =       429
Log likelihood =  7416.107                      AIC             = -32.53197
FPE            =  3.06e-22                      HQIC            = -30.89442
Det(Sigma_ml)  =  3.89e-23                      SBIC            = -28.38531

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
lpcm                 73     .020137   0.9984   274934.4   0.0000
lp                   73     .001929   1.0000   4.65e+07   0.0000
ly                   73     .004581   0.9998    2415767   0.0000
usff                 73     .528117   0.9799   20876.42   0.0000
smtr                 73     .008533   0.9998    2004728   0.0000
smnbr                73     .017071   0.9992   517232.9   0.0000
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lpcm         |
        lpcm |
         L1. |   1.405009   .0485778    28.92   0.000     1.309798    1.500219
         L2. |   -.410722   .0839434    -4.89   0.000     -.575248    -.246196
         L3. |  -.0582564   .0863925    -0.67   0.500    -.2275827    .1110698
         L4. |   .0060963   .0860305     0.07   0.944    -.1625204     .174713
         L5. |   .1898023   .0846223     2.24   0.025     .0239456     .355659
         L6. |  -.1352143   .0843991    -1.60   0.109    -.3006335    .0302049
         L7. |  -.0221461    .084104    -0.26   0.792    -.1869869    .1426946
         L8. |   .1088265   .0834062     1.30   0.192    -.0546467    .2722997
         L9. |  -.0908905   .0830278    -1.09   0.274    -.2536219    .0718409
        L10. |  -.0087207   .0820738    -0.11   0.915    -.1695823    .1521409
        L11. |   .0530312   .0801067     0.66   0.508    -.1039751    .2100376
        L12. |  -.0604962   .0491042    -1.23   0.218    -.1567387    .0357463
          lp |
         L1. |  -.8525089   .5058086    -1.69   0.092    -1.843875    .1388577
         L2. |   2.357894   .7653501     3.08   0.002     .8578352    3.857952
         L3. |  -1.547534   .7682057    -2.01   0.044    -3.053189   -.0418782
         L4. |   .1369851   .7722047     0.18   0.859    -1.376508    1.650479
         L5. |  -.7643444   .7735541    -0.99   0.323    -2.280483    .7517939
         L6. |   .8417005   .7827753     1.08   0.282    -.6925108    2.375912
         L7. |    .322795   .7820853     0.41   0.680    -1.210064    1.855654
         L8. |  -.8197013   .7794201    -1.05   0.293    -2.347337     .707934
         L9. |   1.207514   .7791782     1.55   0.121     -.319647    2.734675
        L10. |  -.7059256   .7798005    -0.91   0.365    -2.234307    .8224553
        L11. |  -.1697698   .7722171    -0.22   0.826    -1.683288    1.343748
        L12. |   .0117781   .4894804     0.02   0.981    -.9475858    .9711419
          ly |
```

```
. var lpcm lp ly usff smtr smnbr, lags(1/12)

Vector autoregression

Sample:  1960m7 - 1996m3                        No. of obs      =       429
Log likelihood =  7416.107                      AIC             = -32.53197
FPE            =  3.06e-22                      HQIC            = -30.89442
Det(Sigma_ml)  =  3.89e-23                      SBIC            = -28.38531
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
lpcm	73	.020137	0.9984	274934.4	0.0000
lp	73	.001929	1.0000	4.65e+07	0.0000
ly	73	.004581	0.9998	2415767	0.0000
usff	73	.528117	0.9799	20876.42	0.0000
smtr	73	.008533	0.9998	2004728	0.0000
smnbr	73	.017071	0.9992	517232.9	0.0000

		Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lpcm							
lpcm	L1.	1.405009	.0485778	28.92	0.000	1.309798	1.500219
	L2.	-.4107222	.0839434	-4.89	0.000	-.5752448	-.2461996
	L3.	-.0582564	.0863925	-0.67	0.500	-.2275827	.1110698
	L4.	.0060963	.0860305	0.07	0.944	-.1625204	.1747133
	L5.	.1898023	.0846221	2.24	0.025	.0239456	.3556589
	L6.	-.1352143	.0843991	-1.60	0.109	-.3006335	.0302049
	L7.	-.0221461	.0841041	-0.26	0.792	-.1869869	.1426946
	L8.	.1088265	.0834066	1.30	0.192	-.0546467	.2722997
	L9.	-.0908905	.0830278	-1.09	0.274	-.2536219	.0718409
	L10.	-.0087207	.0820738	-0.11	0.915	-.1695823	.1521409
	L11.	.0530312	.0801067	0.66	0.508	-.1039751	.2100376
	L12.	-.0604962	.0491042	-1.23	0.218	-.1567387	.0357463
lp	L1.	-.8525089	.5058086	-1.69	0.092	-1.843875	.1388577
	L2.	-2.357894	.7653501	-3.08	0.002	[illegible]	[illegible]
	L3.	-1.547534	.7682057	-2.01	0.044	[illegible]	[illegible]
	L4.	.1369851	.7722047	0.18	0.859	[illegible]	[illegible]
	L5.	-.7643444	.7735541	-0.99	0.323	[illegible]	[illegible]
	L6.	-.8417005	.7782753	-1.08	0.280	[illegible]	[illegible]
	L7.	.3222795	.7820853	0.41	0.680	[illegible]	[illegible]
	L8.	-.8197013	.7794201	-1.05	0.293	[illegible]	[illegible]
	L9.	1.207514	.7791782	1.55	0.121	[illegible]	[illegible]
	L10.	-.7059256	.7798005	-0.91	0.365	[illegible]	[illegible]
	L11.	-.1697698	.7722171	-0.22	0.826	[illegible]	[illegible]
	L12.	.0117781	.4894804	0.02	0.981	[illegible]	[illegible]
ly	L1.	-.1193677	.212606	-0.56	0.574	-.5360678	.2973324
	L2.	-.4105978	.3172038	-1.29	0.196	[illegible]	[illegible]
	L3.–L12.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
usff	L1.	.000868	.0019825	0.44	0.661	-.0030175	.0047536
	L2.–L12.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
smtr	L1.–L12.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
smnbr	L1.–L12.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	_cons	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
lp							
lpcm, lp, ly, usff, smtr, smnbr	L1.–L12.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	_cons	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ly							
lpcm	L1.–L5.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

--more--

```
. varstable
```

Eigenvalue stability condition

Eigenvalue	Modulus
.9994007	.999401
.9878918 + .0215195*i*	.988126
.9878918 - .0215195*i*	.988126
.9649073	.964907
.9527284 + .04779166*i*	.953926
.9527284 - .04779166*i*	.953926
.943027 + .1011819*i*	.94844
.943027 - .1011819*i*	.94844
.16277 + .9020018*i*	.91657
.16277 - .9020018*i*	.91657
-.6048426 + .6876819*i*	.915828
-.6048426 - .6876819*i*	.915828
-.3877779 + .8181855*i*	.905428
-.3877779 - .8181855*i*	.905428
-.04668451 + .8843734*i*	.885605
-.04668451 - .8843734*i*	.885605
.2762283 + .8410597*i*	.885259
.2762283 - .8410597*i*	.885259
-.5109755 + .72245*i*	.88489
-.5109755 - .72245*i*	.88489
.7299002 + .4990881*i*	.884219
.686893 + .3817634*i*	.785853
.686893 - .3817634*i*	.785853
.1730042 + .763779*i*	.783128
.1730042 - .763779*i*	.783128
-.7180819	.718082
-.6166154 + .2458618*i*	.663824
-.6166154 - .2458618*i*	.663824
-.1336691 + .3698518*i*	.393266
-.1336691 - .3698518*i*	.393266
-.1020844	.102084

All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.

OK

```
. varnorm
```

Jarque-Bera test

Equation	chi2	df	Prob > chi2
lpcm	174.968	2	0.00000
lp	107.445	2	0.00000
ly	17.504	2	0.00016
usff	5382.831	2	0.00000
smtr	7.667	2	0.02164
smnbr	5062.135	2	0.00000
ALL	1.1e+04	12	0.00000

NO GOOD!

Skewness test

Equation	Skewness	chi2	df	Prob > chi2
lpcm	-.26049	4.852	1	0.02762
lp	-.02579	0.048	1	0.82737
ly	-.08678	0.538	1	0.46307
usff	-1.1191	89.539	1	0.00000
smtr	.25366	4.601	1	0.03196
smnbr	-2.2566	364.085	1	0.00000
ALL		463.663	6	0.00000

Kurtosis test

Equation	Kurtosis	chi2	df	Prob > chi2
lpcm	6.085	170.117	1	0.00000
lp	5.4512	107.397	1	0.00000
ly	3.9742	16.965	1	0.00004
usff	20.208	5293.292	1	0.00000
smtr	3.4141	3.066	1	0.07995
smnbr	19.212	4698.050	1	0.00000
ALL		1.0e+04	6	0.00000

```
. varlmar, mlag(24)
```

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	27.6464	36	0.83956
2	44.5814	36	0.15439
3	29.7153	36	0.76080
4	25.4750	36	0.90431
5	41.0927	36	0.25727
6	45.3367	36	0.13686
7	56.1594	36	0.01730
8	41.4924	36	0.24361
9	41.6707	36	0.23767
10	38.7024	36	0.34863
11	35.1426	36	0.50919
12	42.4407	36	0.21315
13	33.7081	36	0.57807
14	32.3312	36	0.64381
15	35.6894	36	0.48324
16	45.6283	36	0.13052
17	41.4994	36	0.24337
18	38.4841	36	0.35774
19	45.3969	36	0.13553
20	48.4022	36	0.08110
21	49.4428	36	0.06709
22	42.8626	36	0.20048
23	45.3814	36	0.13587
24	48.3907	36	0.08127

H0: no autocorrelation at lag order

GOOD!

So maybe we should use a VAR(12), In fact several people suggest that with monthly data one should use a VAR(12) or even a VAR(13) (one year plus one month)

But let's forget about this and follow LSZ and use a VAR 6

(note that a VAR 6 still fails normality tests and sometimes fails Autocorrelation tests)

Why these test failures? Probably because we had structural changes in the US economy

```
. var lpcm lp ly usff smtr smnbr, lags(1/6)

Vector autoregression

Sample:  1960m1 - 1996m3                      No. of obs       =       435
Log likelihood =  7339.575                     AIC              = -32.72448
FPE            =  2.48e-22                     HQIC             =  -31.9036
Det(Sigma_ml)  =  8.91e-23                     SBIC             = -30.64465

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
lpcm                 37     .020122   0.9983   257208.5   0.0000
lp                   37     .001971   1.0000   4.15e+07   0.0000
ly                   37     .004653   0.9998    2238974   0.0000
usff                 37     .560679   0.9749   16891.04   0.0000
smtr                 37     .008787   0.9998    1755766   0.0000
smnbr                37     .017063   0.9991   481192.2   0.0000
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lpcm         |
        lpcm |
         L1. |   1.392091   .0475966    29.25   0.000     1.298804    1.485379
         L2. |  -.3841456   .0820621    -4.68   0.000    -.5449844   -.2233069
         L3. |  -.0420422   .0838941    -0.50   0.616    -.2064717    .1223872
         L4. |  -.0101772   .0826745    -0.12   0.902    -.1722162    .1518617
         L5. |   .1765003   .0806509     2.19   0.029     .0184274    .3345732
         L6. |  -.1360474   .0488753    -2.78   0.005    -.2318412   -.0402536
          lp |
         L1. |  -.8195481   .4904325    -1.67   0.095    -1.780778     .141682
         L2. |   1.926339   .7693656     2.50   0.012     .4184105    3.434268
         L3. |  -1.129228   .7770207    -1.45   0.146     -2.65216    .3937051
         L4. |   .1052967   .7783661     0.14   0.892    -1.420273    1.630866
         L5. |  -.5971992   .7640633    -0.78   0.434    -2.094736    .9003373
         L6. |   .5020343    .476431     1.05   0.292    -.4317533    1.435822
          ly |
         L1. |  -.1149547   .2055704    -0.56   0.576    -.5178653    .2879559
         L2. |   .3639539   .3094284     1.18   0.240    -.2425147    .9704224
         L3. |  -.1731068   .3158735    -0.55   0.584    -.7922076     .445994
```

IRF

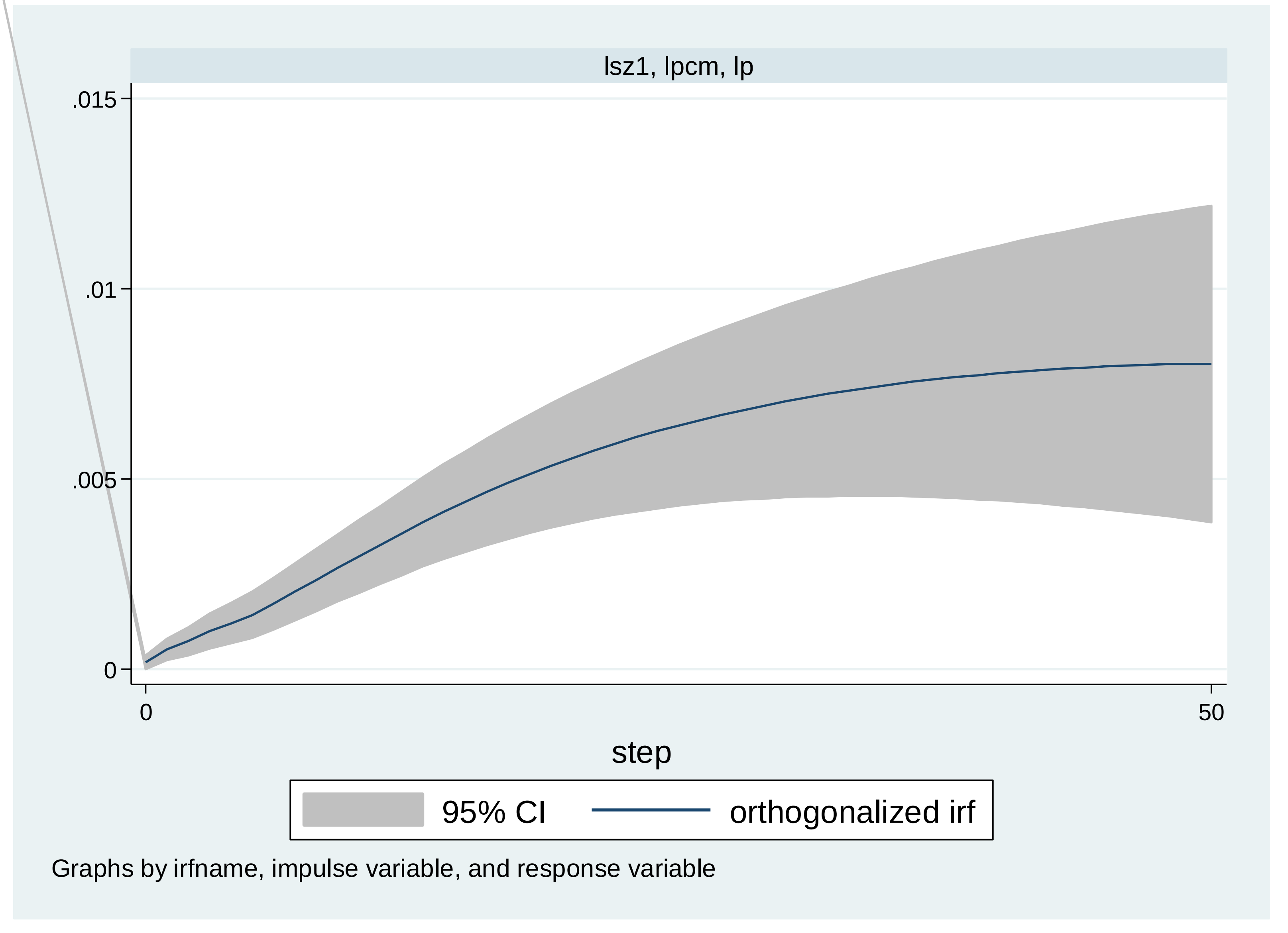
```
. irf create lsz1, set(lsz1) step(50)
(file lsz1.irf created)
(file lsz1.irf now active)
```

- This command creates the IRF file with 50 step forward forecasts (it **assumes** an ordering)
- Now we can graph the IRF. If want to graph them one by one we do the following

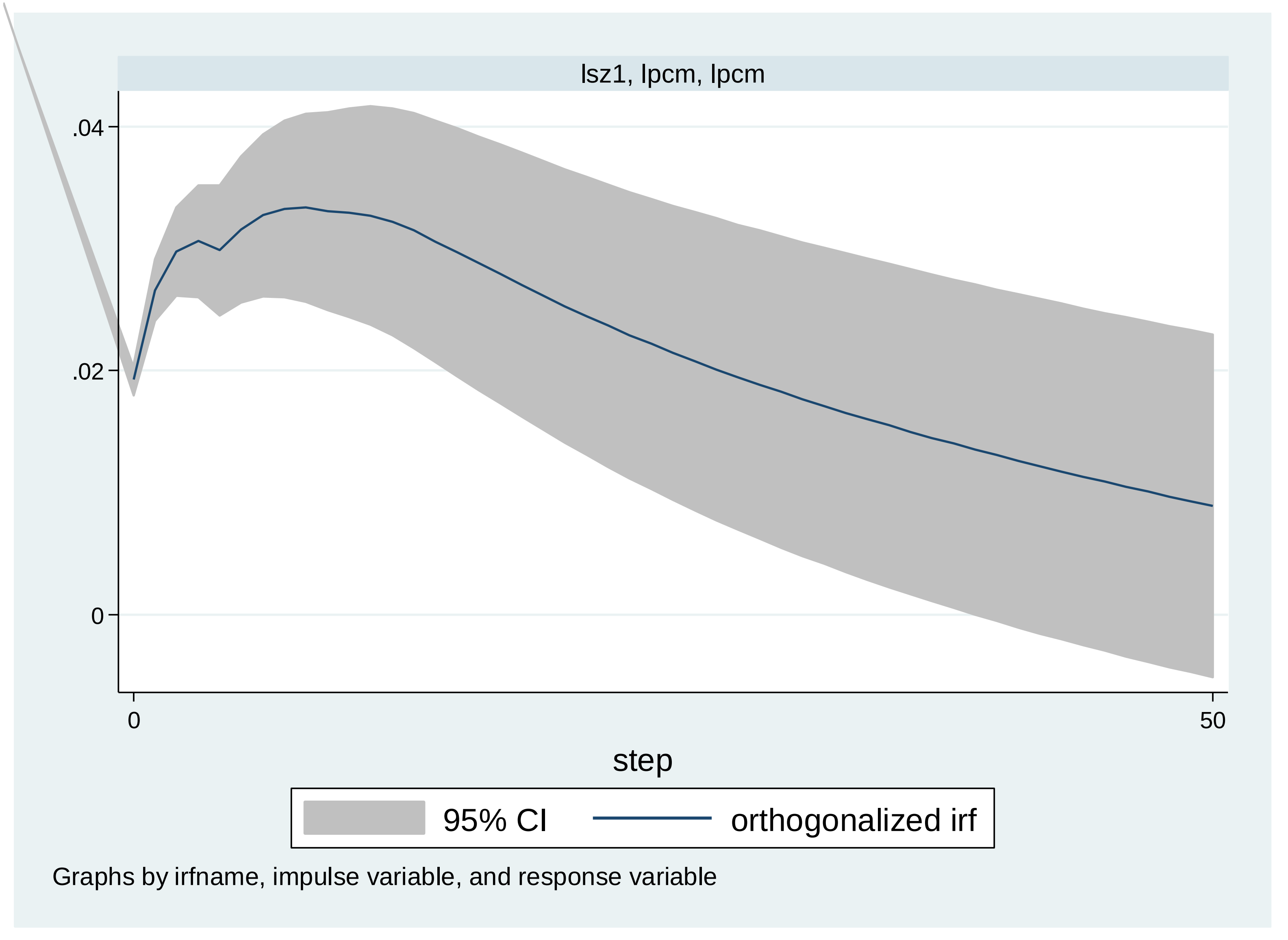
```
irf graph oirf, impulse(var) response(var)
```

This is telling STATA to orthogonalize the shocks using a Cholesky decomposition with the ordering that we used to estimated the VAR

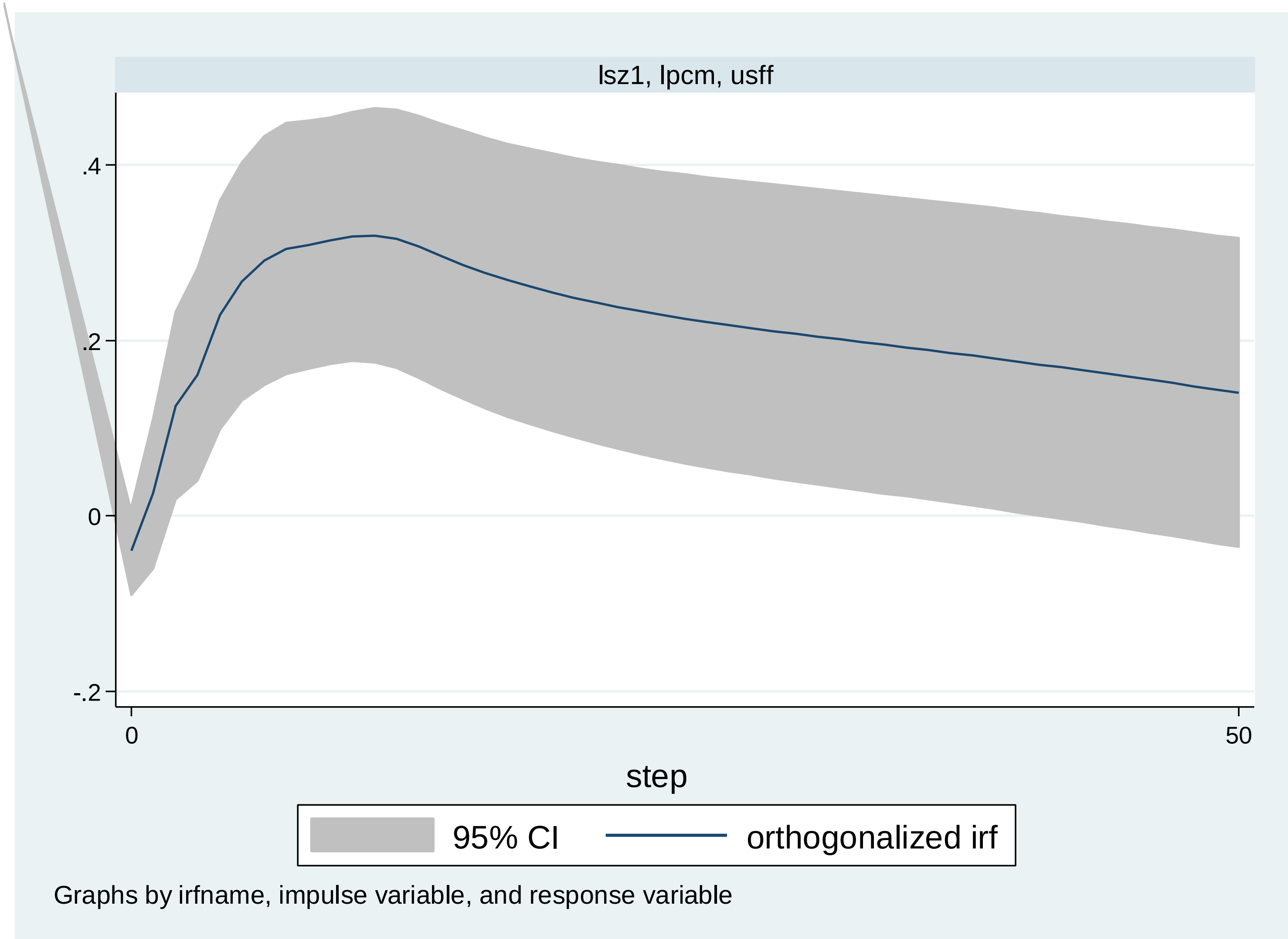
```
irf graph oirf, impulse(lpcm) response(lp)
```



```
irf graph oirf, impulse(lpcm) response(lpcm)
```



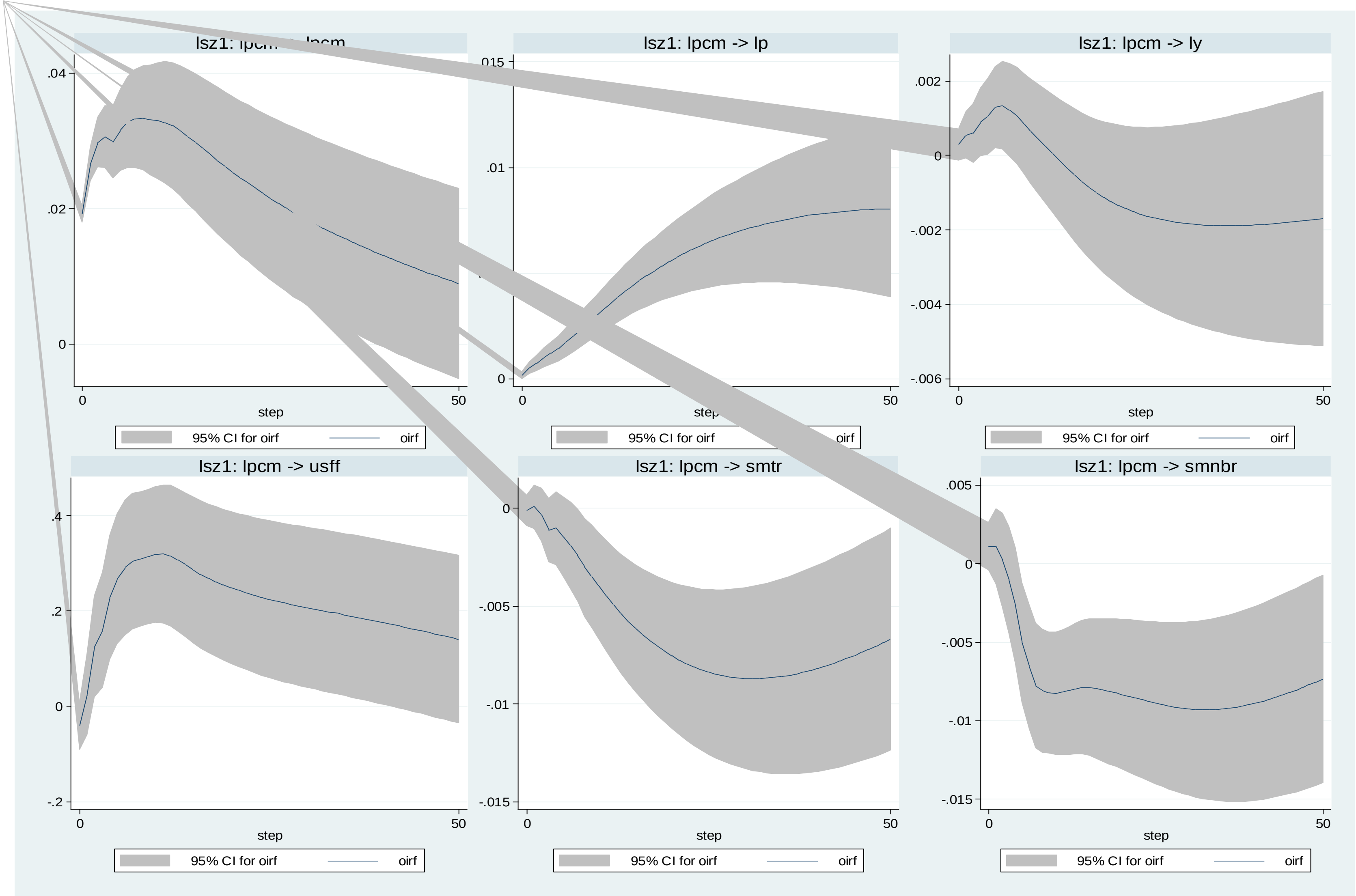
```
irf graph oirf, impulse(lpcm) response(usff)
```



OK, this is getting boring. Luckily we have a command to combine graphs

```
irf cgraph (lsz1 lpcm lpcm oirf) (lsz1 lpcm lp oirf)
(lsz1 lpcm ly oirf) (lsz1 lpcm usff oirf)
(lsz1 lpcm smtr oirf) (lsz1 lpcm smnbr oirf)
```

COMBINE



```
. irf table oirf, response (lpcm) noci
```

Results from lsz1

step	(1) oirf	(2) oirf	(3) oirf	(4) oirf	(5) oirf	(6) oirf
0	.019248	0	0	0	0	0
1	.026586	-.001505	-.000399	.000371	-.001738	-.000027
2	.029761	-.000446	.000046	-.001096	-.00071	-.000144
3	.030603	-.000167	.001198	-.00169	-.001295	.000133
4	.029859	.000645	.000979	-.002778	-.000329	.000734
5	.031568	-.000037	.002153	-.002691	.001191	.000096
6	.032715	-.000777	.00254	-.002413	.000766	.000038
7	.033248	-.001137	.002613	-.002827	.001148	.000355
8	.033342	-.001353	.002913	-.003708	.00129	.000905
9	.033048	-.001321	.002718	-.004781	.0012	.00197
10	.032938	-.001418	.002719	-.005639	.001219	.002785
11	.032702	-.001565	.002727	-.006322	.000927	.003378
12	.032195	-.001611	.002701	-.006882	.000596	.003883
13	.031467	-.001571	.002714	-.00739	.000276	.004233
14	.030578	-.001474	.002693	-.007932	-.000069	.004489
46	.010475	-.00155	.002468	-.012276	-.004576	.001325
47	.010075	-.001542	.002465	-.012219	-.004587	.001253
48	.009684	-.001533	.002462	-.012156	-.004595	.001184
49	.009303	-.001522	.002459	-.012089	-.0046	.001118
50	.00893	-.001511	.002455	-.012016	-.004602	.001053

(1) irfname = lsz1, impulse = lpcm, and response = lpcm
(2) irfname = lsz1, impulse = lp, and response = lpcm
(3) irfname = lsz1, impulse = ly, and response = lpcm
(4) irfname = lsz1, impulse = usff, and response = lpcm
(5) irfname = lsz1, impulse = smtr, and response = lpcm
(6) irfname = lsz1, impulse = smnbr, and response = lpcm

response of lpcm

```
. irf table oirf, impulse (lpcm) noci
```

Results from lsz1

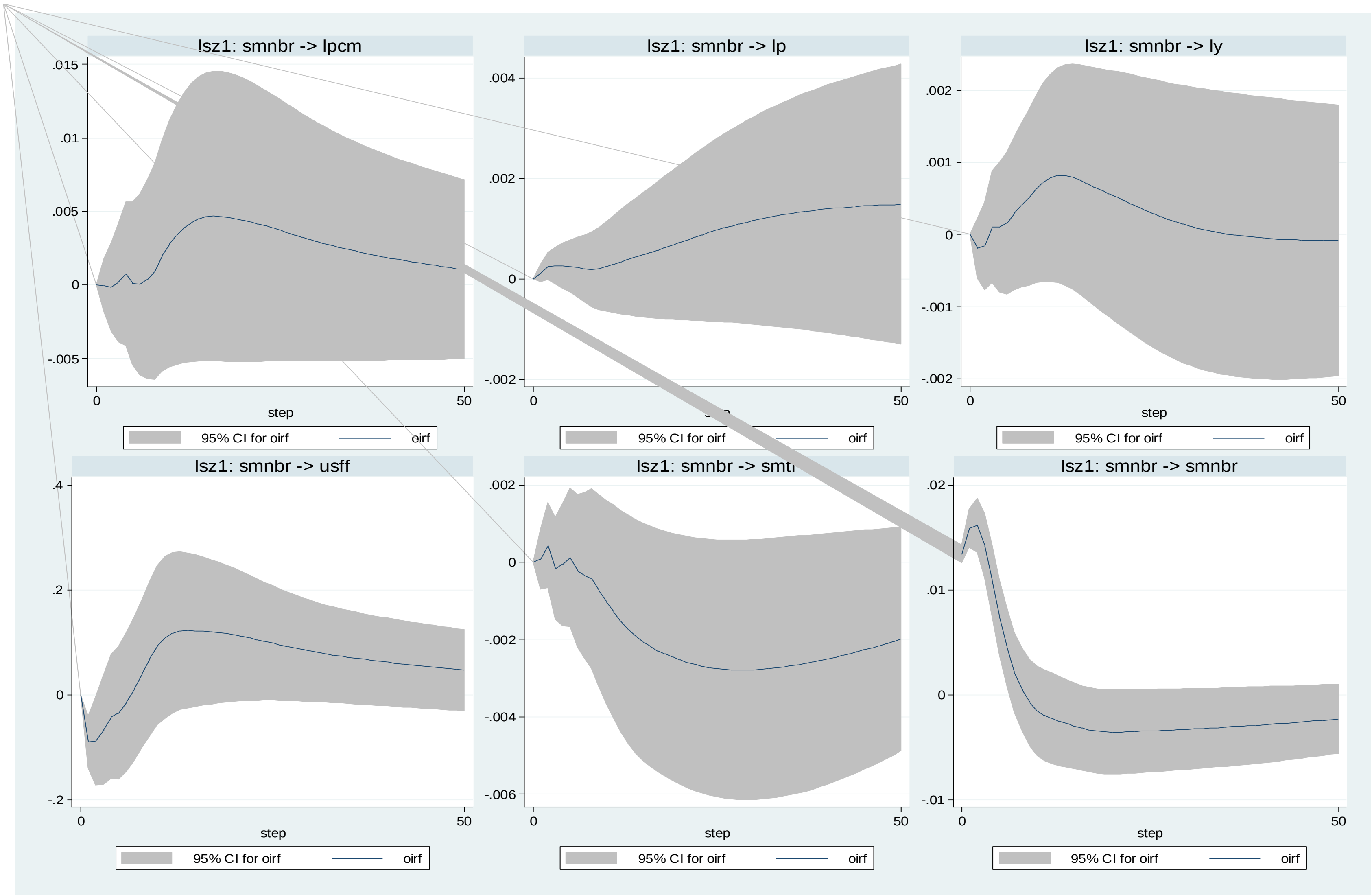
step	(1) oirf	(2) oirf	(3) oirf	(4) oirf	(5) oirf	(6) oirf
0	.019248	.000176	.000296	-.040024	-.000094	.001094
1	.026586	.000519	.000551	.025581	.000075	.001075
2	.029761	.000732	.000599	.125392	-.00031	.000199
3	.030603	.001	.000902	.160628	-.001112	-.00106
4	.029859	.001207	.001055	.228547	-.001021	-.002714
5	.031568	.001425	.001294	.267032	-.00146	-.005042
6	.032715	.001727	.001342	.291159	-.001899	-.006589
7	.033248	.002033	.001219	.304512	-.002415	-.007781
47	.010075	.007999	-.001766	.151405	-.007201	-.007894
48	.009684	.008013	-.001744	.147757	-.007035	-.007719
49	.009303	.008023	-.00172	.144086	-.006864	-.007537
50	.00893	.00803	-.001696	.140399	-.006687	-.00735

(1) irfname = lsz1, impulse = lpcm, and response = lpcm
(2) irfname = lsz1, impulse = lpcm, and response = lp
(3) irfname = lsz1, impulse = lpcm, and response = ly
(4) irfname = lsz1, impulse = lpcm, and response = usff
(5) irfname = lsz1, impulse = lpcm, and response = smtr
(6) irfname = lsz1, impulse = lpcm, and response = smnbr

Impulse of lpcm

```
. irf cgraph (lsz1 smnbr lpcm oirf) (lsz1 smnbr lp oirf)
 (lsz1 smnbr ly oirf) (lsz1 smnbr usff oirf)
(lsz1 smnbr smtr oirf) (lsz1 smnbr smnbr oirf)
```

Do you remember the original ordering



```
. irf table oirf, impulse (smnbr) noci
```

Results from lsz1

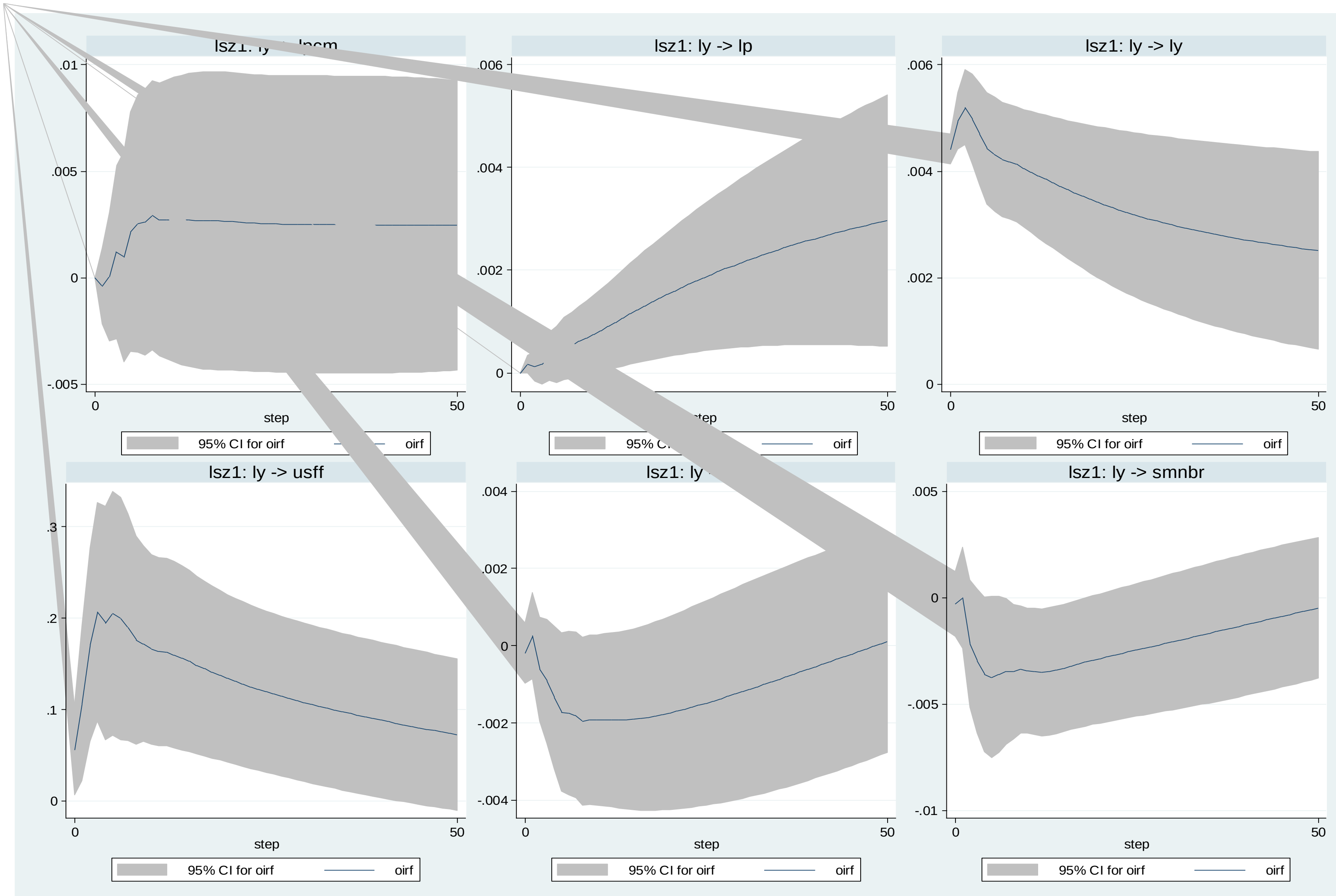
step	(1) oirf	(2) oirf	(3) oirf	(4) oirf	(5) oirf	(6) oirf
0	0	0	0	0	0	.013418
1	-.000027	.000114	-.000194	-.090441	.000076	.015854
2	-.000144	.000248	-.000158	-.088047	.000434	.016124
3	.000133	.000258	.000104	-.066569	-.000163	.014165
4	.000734	.000253	.0001	-.041628	-.000061	.010892

```
. irf table oirf, response (smnbr) noci
```

Results from lsz1

step	(1) oirf	(2) oirf	(3) oirf	(4) oirf	(5) oirf	(6) oirf
0	.001094	-.001249	-.000292	-.004515	.007944	.013418
1	.001075	-.000487	-.00002	-.007281	.009181	.015854
2	.000199	.001417	-.002163	-.007334	.008182	.016124
3	-.00106	.001724	-.002981	-.006563	.00805	.014165
4	-.002714	.001873	-.003609	-.004787	.008571	.010892
5	-.005042	.003164	-.003734	-.004621	.007735	.007373
6	-.006589	.003701	-.003626	-.004929	.007771	.004558
7	-.007781	.003809	-.003465	-.005005	.008038	.002109
8	-.008088	.003542	-.003484	-.004942	.008155	.000429

```
. irf cgraph (lsz1 ly lpcm oirf) (lsz1 ly lp oirf)
 (lsz1 ly ly oirf) (lsz1 ly usff oirf) (lsz1 ly smtr oirf)
 (lsz1 ly smnbr oirf)
```



```
. irf table oirf, impulse (ly) noci
```

Results from lsz1

step	(1) oirf	(2) oirf	(3) oirf	(4) oirf	(5) oirf	(6) oirf
0	0	0	.004419	.056465	-.000197	-.000292
1	-.000399	.000167	.004943	.106525	.000252	-.00002
2	.000046	.000126	.005196	.170771	-.000617	-.002163
3	.001198	.000168	.004975	.20663	-.000926	-.002981
4	.000979	.000318	.004695	.194093	-.001341	-.003609
5	.002153	.000361	.004429	.205233	-.00172	-.003734
6	.00254	.000474	.004309	.199397	-.001748	-.003626
7	.002613	.000536	.00422	.189895	-.001799	-.003465
8	.002913	.000613	.004176	.175287	-.001958	-.003484
9	.002718	.000681	.004129	.171957	-.00193	-.003384
10	.002719	.000746	.004054	.165588	-.001929	-.003447
11	.002727	.000826	.003984	.163205	-.001925	-.003473

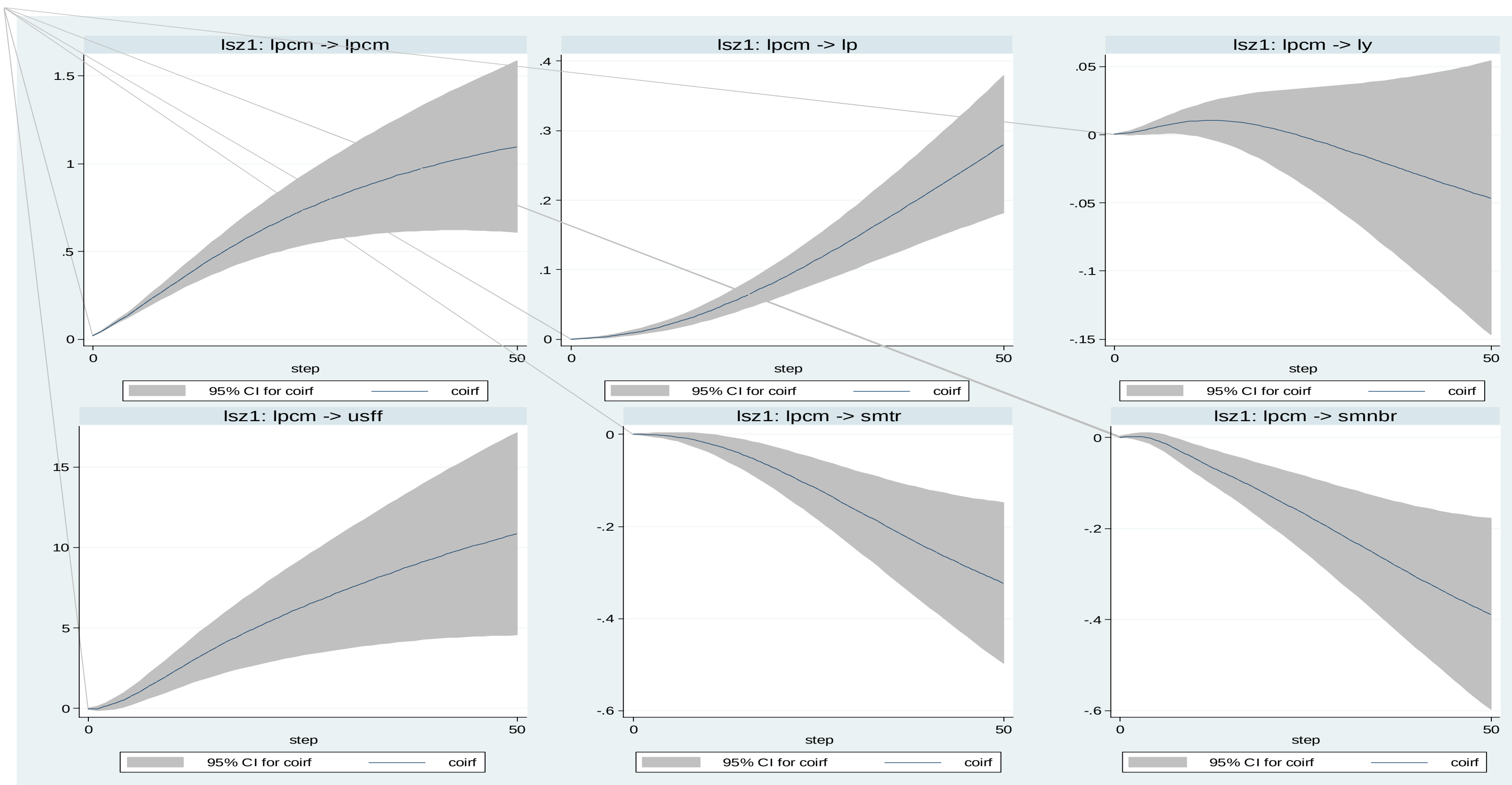
```
. irf table oirf, response (ly) noci
```

Results from lsz1

step	(1) oirf	(2) oirf	(3) oirf	(4) oirf	(5) oirf	(6) oirf
0	.000296	-.000439	.004419	0	0	0
1	.000551	-.000424	.004943	-5.3e-06	-.00031	-.000194
2	.000599	-.000684	.005196	.000058	-.000181	-.000158
3	.000902	-.000631	.004975	.000083	-.000271	.000104
4	.001055	-.000924	.004695	-.000215	-.000319	.0001
5	.001294	-.001022	.004429	-.001018	-.000644	.000155
6	.001342	-.001192	.004309	-.001691	-.000863	.000293

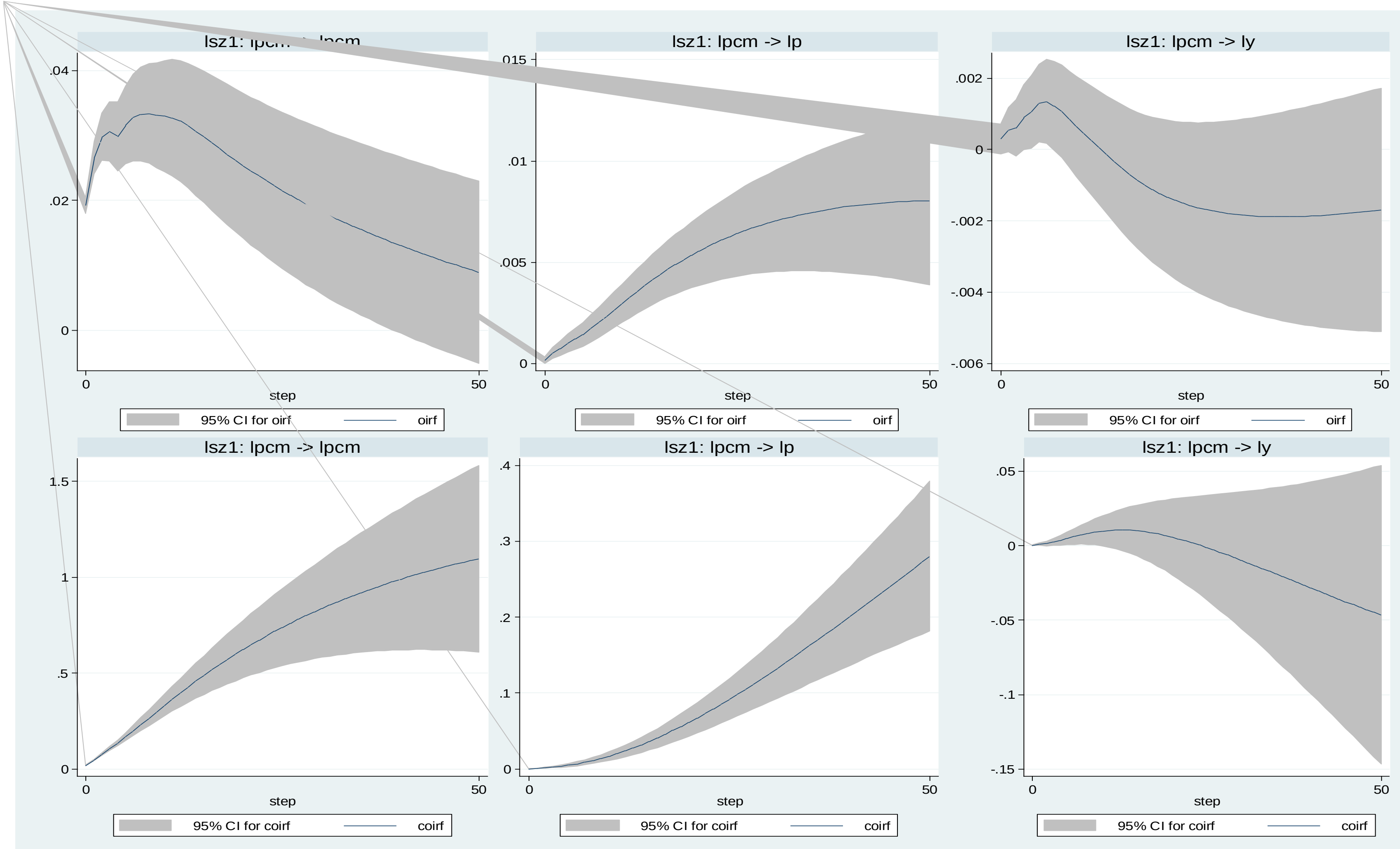
We can also look at the cumulative effect

```
irf cgraph (lsz1 lpcm lpcm coirf) (lsz1 lpcm lp coirf)
(lsz1 lpcm ly coirf) (lsz1 lpcm usff coirf)
 (lsz1 lpcm smtr coirf) (lsz1 lpcm smnbr coirf)
```



Or both

```
irf cgraph (lsz1 lpcm lpcm oirf) (lsz1 lpcm lp oirf)
(lsz1 lpcm ly oirf) (lsz1 lpcm lpcm coirf)
(lsz1 lpcm lp coirf)  (lsz1 lpcm ly coirf)
```

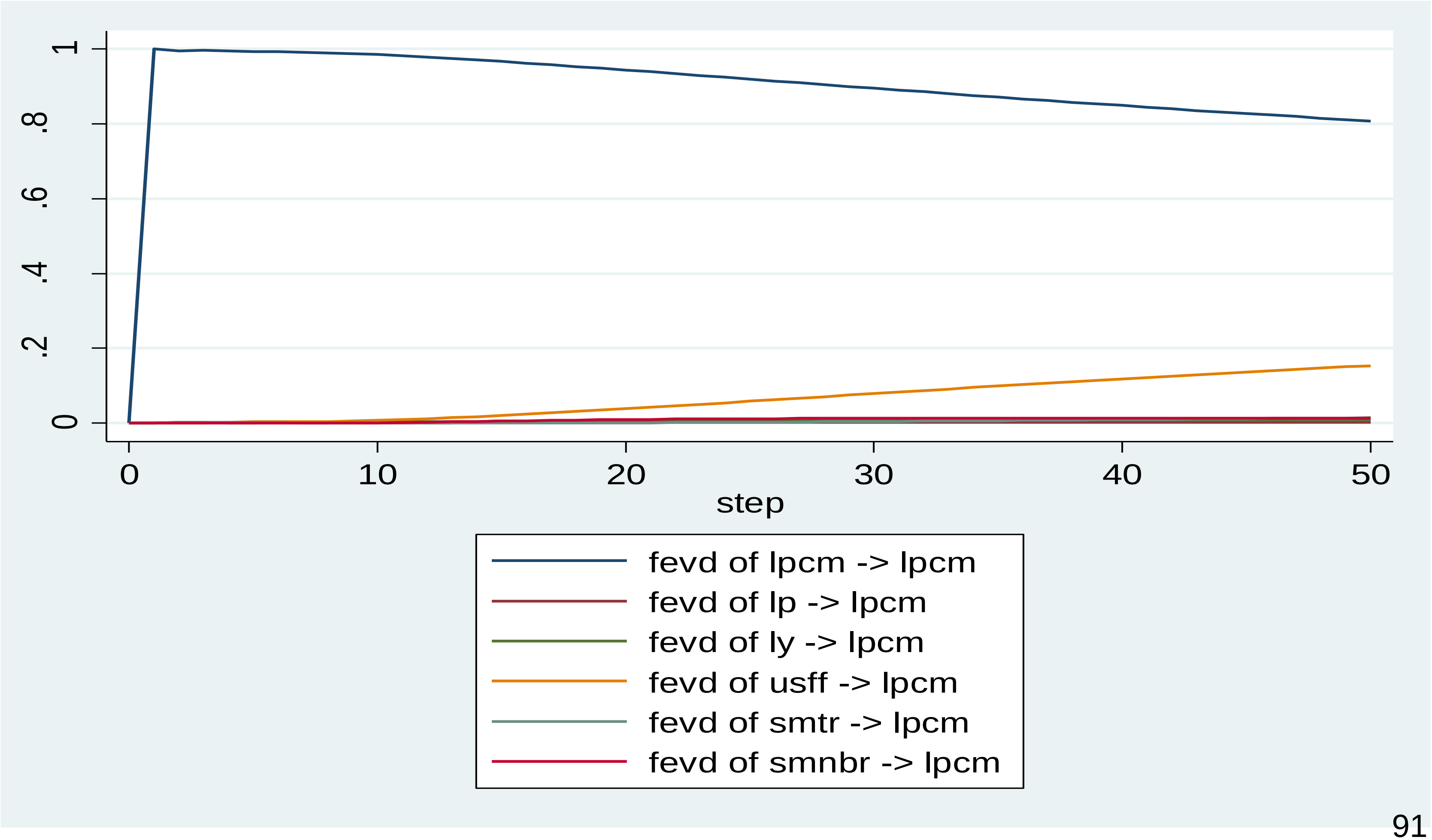


Forecast error variance decomposition

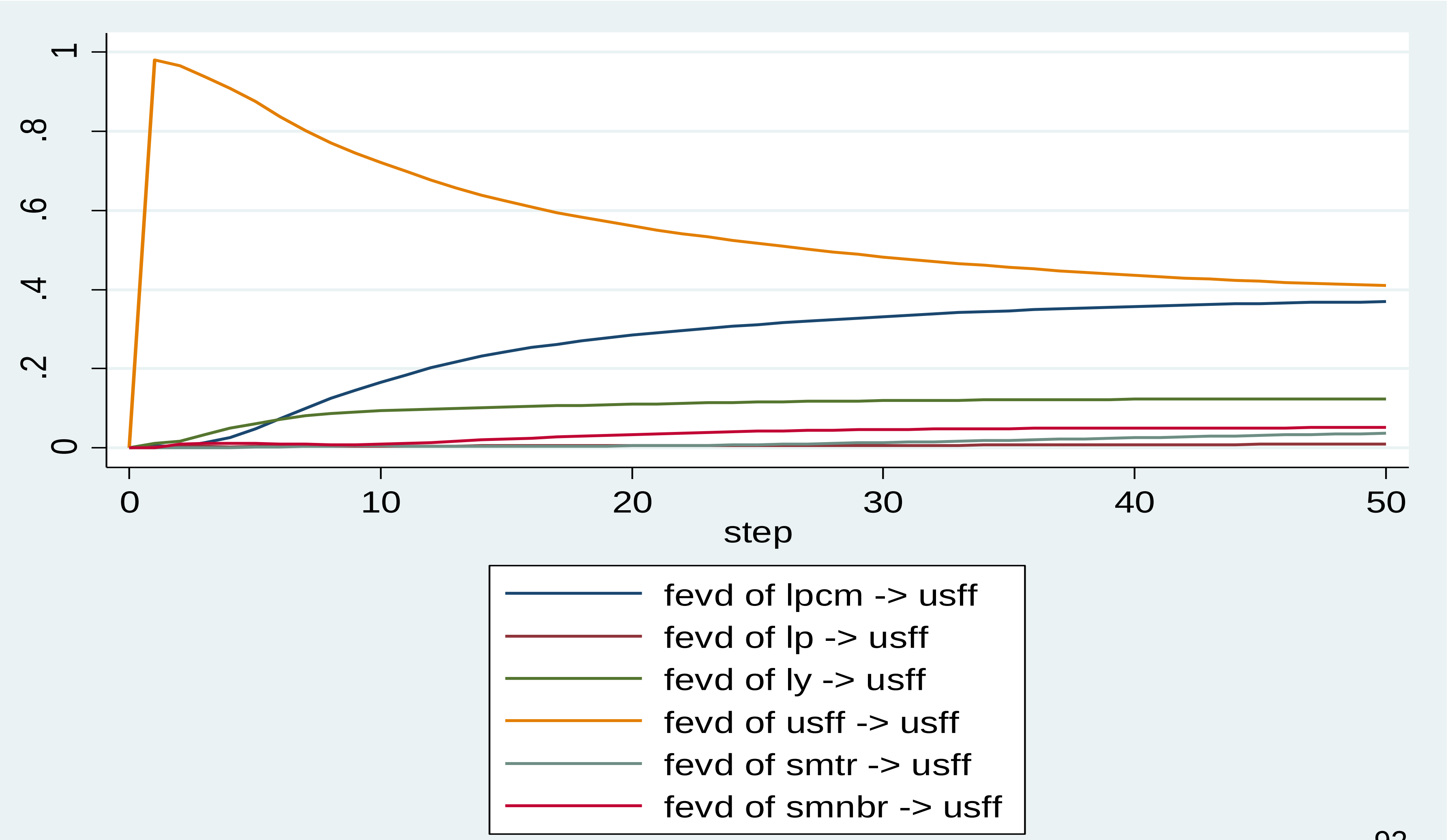
- The forecast error variance decomposition shows the proportion of the movements in a sequence which is due to its own shock versus the shocks to other variables
- If shocks to *y* explain none of the forecast errors variance of *x*, we say that *x* is exogenous with respect to *y*
- If shocks to *y* explain 100% the forecast errors variance of *x* at all **n**, then we can say that *x* is fully endogenous with respect to *y*
- Normally a variable explains most of its variance at short horizons and a decreasing share at longer horizons

Forecast error variance decomposition

```
irf ograph (lsz1 lpcm lpcm fevd) (lsz1 lp lpcm  fevd)
 (lsz1 ly lpcm fevd) (lsz1 usff lpcm fevd)
(lsz1 smtr lpcm  fevd) (lsz1 smnbr lpcm  fevd)
```



```
. irf ograph (lsz1 lpcm usff fevd) (lsz1 lp usff fevd)
 (lsz1 ly usff fevd) (lsz1 usff usff fevd)
 (lsz1 smtr usff  fevd) (lsz1 smnbr usff fevd)
```



```
. irf table fevd, impulse (usff) noci
```

Results from lsz1

step	(1) fevd	(2) fevd	(3) fevd	(4) fevd	(5) fevd	(6) fevd
0	0	0	0	0	0	0
1	0	0	0	.981187	.024948	.076539
2	.000127	.009338	6.4e-07	.964931	.013008	.11182
3	.00068	.04193	.000046	.937376	.010691	.121839
4	.00144	.078688	.000104	.90912	.015494	.124697
5	.003123	.103676	.00046	.875648	.023183	.120412
6	.003971	.116206	.007441	.837181	.026896	.119901
7	.004228	.119967	.022947	.801337	.031147	.121542

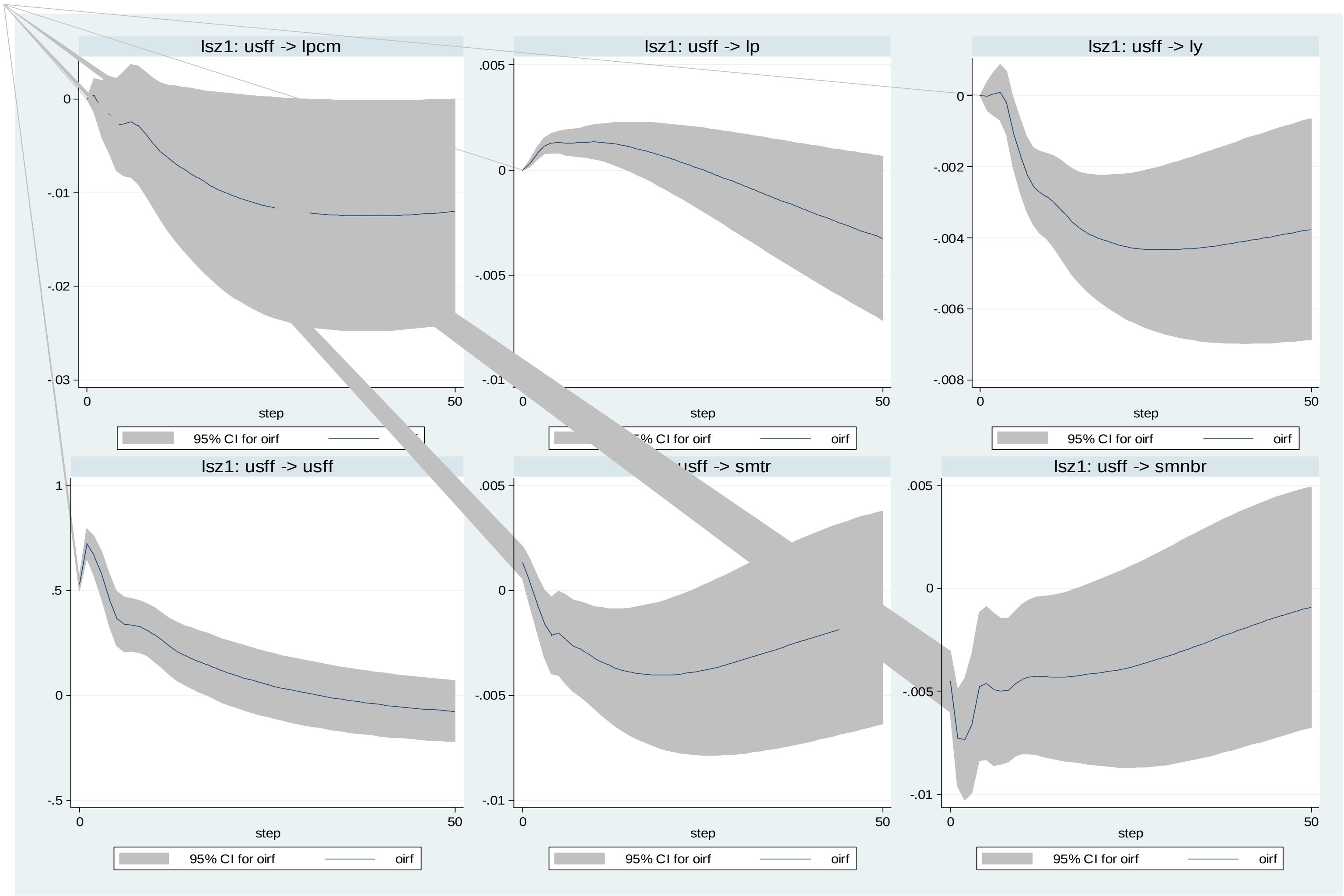
```
. irf table fevd, response (usff) noci
```

Results from lsz1

step	(1) fevd	(2) fevd	(3) fevd	(4) fevd	(5) fevd	(6) fevd
0	0	0	0	0	0	0
1	.00557	.002158	.011085	.981187	0	0
2	.002701	.004998	.017399	.964931	.00018	.009791
3	.01353	.003523	.032884	.937376	.000698	.011989
4	.025389	.002716	.050101	.90912	.000865	.011809
5	.047361	.003241	.061199	.875648	.001651	.0109
6	.073358	.003721	.072861	.837181	.002657	.010222

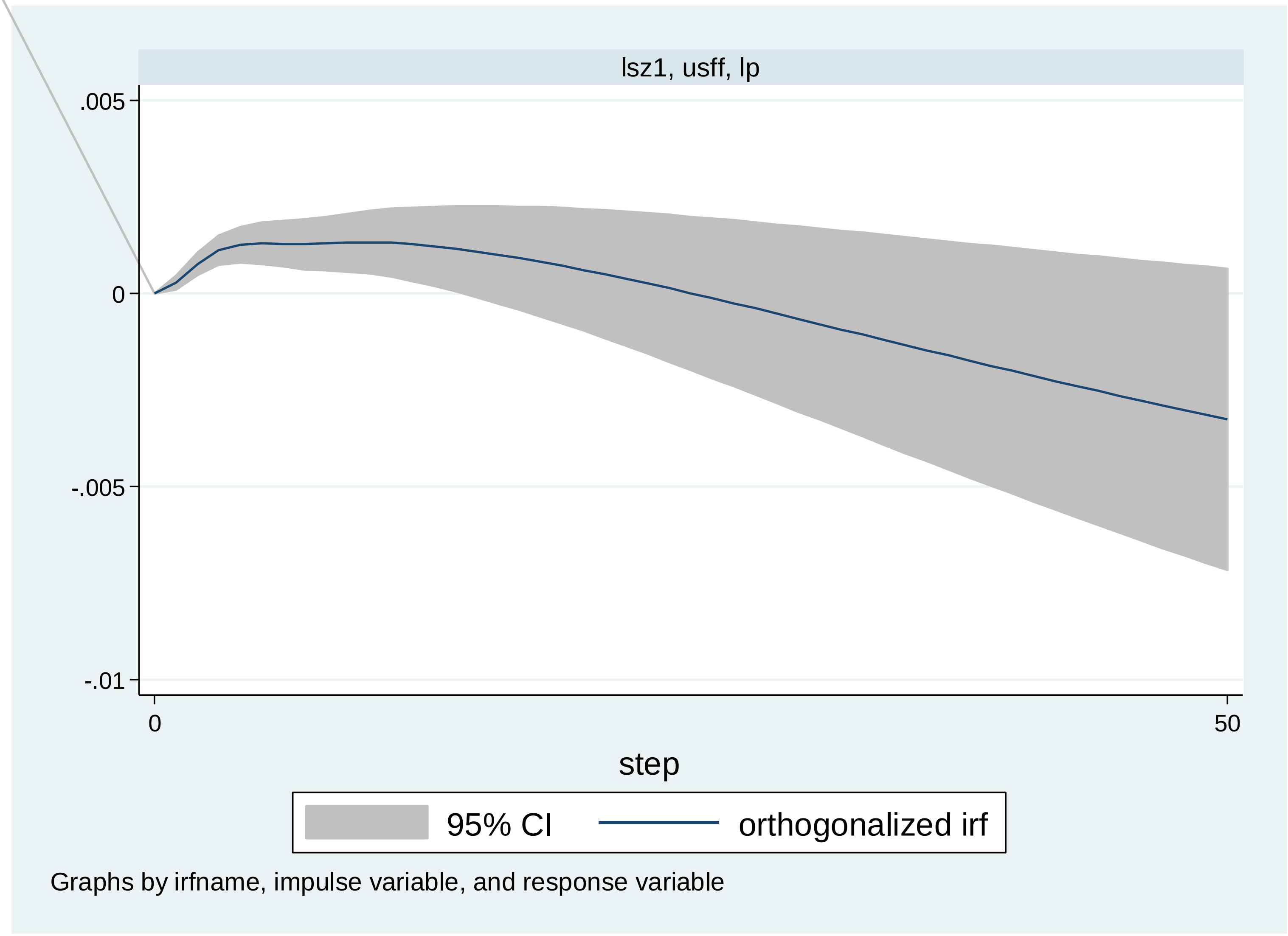
Always zero at zero because there is no forecast error

```
irf cgraph (lsz1 usff lpcm oirf) (lsz1 usff lp oirf)
 (lsz1 usff ly oirf) (lsz1 usff usff oirf)
(lsz1 usff smtr oirf) (lsz1 usff smnbr oirf)
```



Anything strange?

```
irf graph oirf, impulse(usff) response(lp)
```



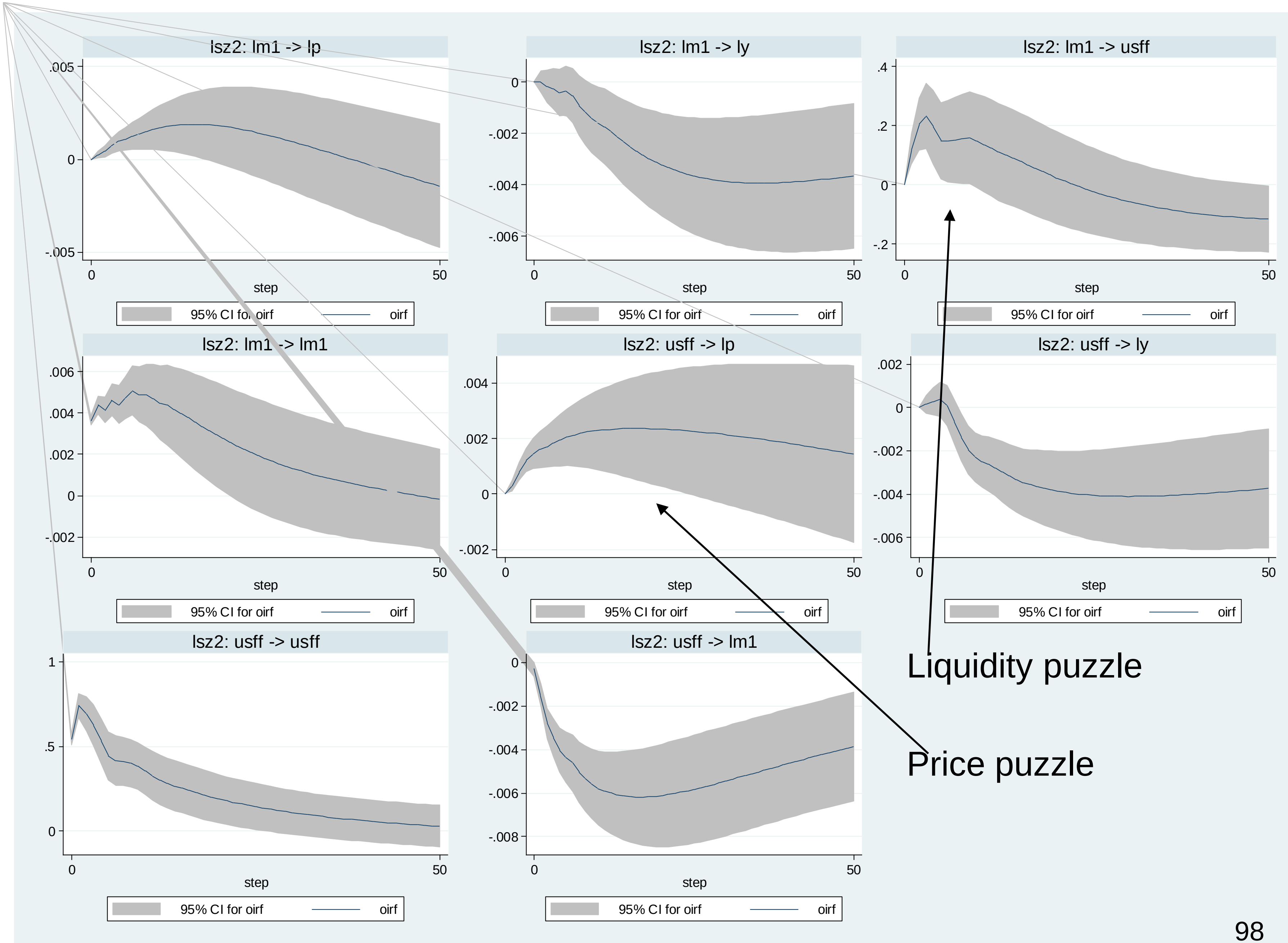
Price puzzle

Let's try a VAR4

- We substitute the two measures of reserves with the log of M1 and we also drop commodity prices

```
. var lp ly usff lm1, lag(1/6)

Vector autoregression

Sample:  1960m1 - 1996m3                        No. of obs      =       435
Log likelihood =  5294.272                      AIC             = -23.88171
FPE            =  5.00e-16                      HQIC            = -23.51194
Det(Sigma_ml)  =  3.15e-16                      SBIC            = -22.94485

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
lp                   25     .002056   1.0000   3.70e+07   0.0000
ly                   25     .004652   0.9998    2174782   0.0000
usff                 25     .564034   0.9738    16184.4   0.0000
lm1                  25     .003769   1.0000   1.50e+07   0.0000
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lp           |
          lp |
         L1. |   1.271131   .0475633    26.73   0.000     1.177909    1.364353
         L2. |  -.1278407   .0775241    -1.65   0.099     -.279785    .0241037
         L3. |  -.1442782   .0779459    -1.85   0.064    -.2970493    .0084929
```

```
. irf create lsz2, set(lsz2) step(50)
(file lsz2.irf created)
(file lsz2.irf now active)
(file lsz2.irf updated)

. irf cgraph (lsz2 lm1 lp oirf) (lsz2 lm1 ly oirf)
 (lsz2 lm1 usff oirf) (lsz2 lm1 lm1 oirf)
(lsz2 usff lp oirf) (lsz2 usff ly oirf)
(lsz2 usff usff oirf) (lsz2 usff lm1 oirf)
```


lsz2: lm1 -> lp
lsz2: lm1 -> ly
lsz2: lm1 -> usff
lsz2: lm1 -> lm1
lsz2: usff -> lp
lsz2: usff -> ly
lsz2: usff -> usff
lsz2: usff -> lm1
step
95% CI for oirf
oirf
Liquidity puzzle
Price puzzle

Why do we get these puzzles?

- Maybe our Cholesky ordering is not appropriate
- Can we improve things if we move to a structural VAR?
  - Language is confusing.
  - Cholesky is also structural, but for some reason people sometime use structural when they identify the VAR with something different from Cholesky

Structural VARs

- VARs are often criticized for having little economics
- Economics dictates the choice of variables and, maybe, the ordering. Everything else is mechanical
  - If the ordering is not justified by sound economic theory and the residuals are correlated we may get bad results
  - Of course you can try different orderings, but when you have more than two variables you need to look at a lot of stuff
    - (with **n** variables you have **n!** possible orderings)
  - And what if different orderings yield dramatically different results?
- Can we use economic theory instead of Cholesky to recover the structural shocks (ε) from the residuals (e)?
- Yes, as long as economic theory allows us to impose $\mathbf{(N^{2}-N)/2}$ restrictions

Structural VARs

- Types of restrictions
  - Coefficient restrictions
    - You may know that one coefficient is one (2, 7, -10, whatever)
    - You may know that the sum of a set of coefficients is zero
  - Variance restrictions
    - You may know that the variance of a given structural shock is 1 (3, 7, 19, whatever)
  - Sign restrictions
    - You may know that the price effect of a demand shock is positive
  - Symmetry restrictions
    - You may know that a given parameter is equal to another parameter (half, twice, the negative, whatever)
    - You may know that the sum of some parameters are equal to another parameter
    - You may know that the variance of a given structural shock is equal to the variance of another structural shock (half, twice, whatever)

The Structural VAR (SVAR)

- The aim of a structural VAR is to use economic theory (rather than Cholesky) to recover the structural innovation