Advanced time-series analysis (University of Lund, Economic History Department)

30 Jan-3 February and 26-30 March 2012

Lecture 9 Vector Autoregression (VAR) techniques: motivation and applications. Estimation procedure. IRF and motivations for SVAR.

9.a Defining a VAR, notation:

Vector autoregressive models can generally be defined as follows:

$$\mathbf{Y_t} = \delta + \sum_{i=1}^{p} \mathbf{\Theta_i} \mathbf{Y_{t-i}} + \boldsymbol{\beta} \mathbf{X_t} + \boldsymbol{\varepsilon}_\mathbf{t}$$

where

$\mathbf{Y_t} = (y_{1t}, y_{2t}, \ldots, y_{kt})$ is a vector k endogenous variables (k-variate VAR). These variables are thought to be determined within the system. (Remember what you learnt during the lecture on simultaneous equation models.) What we here define is that we have k equations, one for each endogenous variable, δ is a vector of equation specific constants (so it is again a vector with k elements). What makes the system autoregressive is that each endogenous variable is regressed on its own lagged values plus the lagged values of all other endogenous variables ($\mathbf{Y_{t-i}}$). The lag length (p) determines the order of the VAR.

$\mathbf{\Theta_i}$ is obviously a kxk matrix of coefficients. We can augment the equation by including exogenous variables denoted by x. These variables now included in matrix $\mathbf{X_t}$. $\boldsymbol{\beta}$ is the coefficient vector of the exogenous variables. $\boldsymbol{\varepsilon}_\mathbf{t}$ is a vector of k elements, denoting the residuals.

This is now a bit too difficult to see, so let us simplify this system to a two-variate case:

Then we have y_{1t} and y_{2t} as dependent variables.

The system, without exogenous variables, can be written as:

$$y_{1t} = \delta_1 + \theta_{111} y_{1t-1} + \theta_{112} y_{2t-1} + \theta_{211} y_{1t-2} + \theta_{212} y_{2t-2} + \ldots + \theta_{p11} y_{1t-p} + \theta_{p12} y_{2t-p} + \varepsilon_{1t}$$
$$y_{2t} = \delta_2 + \theta_{121} y_{1t-1} + \theta_{122} y_{2t-1} + \theta_{221} y_{1t-2} + \theta_{222} y_{2t-2} + \ldots + \theta_{p21} y_{1t-p} + \theta_{p22} y_{2t-p} + \varepsilon_{2t}$$

so:

$$\mathbf{Y_t} = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \ \boldsymbol{\delta} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}, \ \mathbf{\Theta}_p = \begin{pmatrix} \theta_{p11} & \theta_{p12} \\ \theta_{p21} & \theta_{p22} \end{pmatrix}, \boldsymbol{\varepsilon}_\mathbf{t} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

You can write this in a polynomial form:

$\mathbf{\Theta}(L) = \mathbf{I_k} - \mathbf{\Theta}_1 - \mathbf{\Theta}_2 - \ldots - \mathbf{\Theta}_p$ then:

$\mathbf{\Theta}(L)\mathbf{Y_t} = \delta + \boldsymbol{\beta}\mathbf{X_t} + \boldsymbol{\varepsilon}_\mathbf{t}$, this is now an object that you can already treat similarly to the univariate time series models:

$\mathbf{Y_t} = \left(\mathbf{\Theta}(L)\right)^{-1} \delta + \left(\mathbf{\Theta}(L)\right)^{-1} \boldsymbol{\beta}\mathbf{X_t} + \left(\mathbf{\Theta}(L)\right)^{-1} \boldsymbol{\varepsilon}_\mathbf{t}$ this is going to be an infinite order VMA (vector moving average) model.

The expected value of a VAR(p) is:

$E(\mathbf{Y_t}) = \left(\mathbf{\Theta}(L)\right)^{-1} \delta + \left(\mathbf{\Theta}(L)\right)^{-1} \boldsymbol{\beta} E(\mathbf{X_t})$ in the presence of exogenous variables and

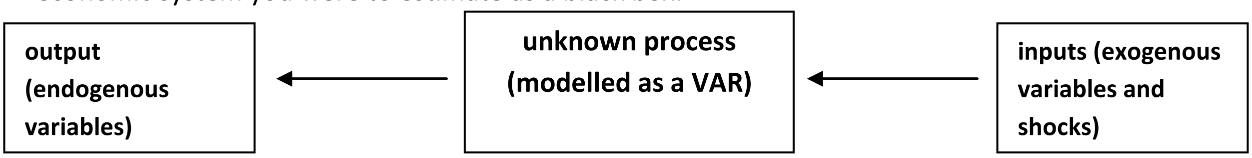
$E(\mathbf{Y_t}) = \left(\mathbf{\Theta}(L)\right)^{-1} \delta$ in the presence of constants only

Obviously, a finite expected value will require that $\mathbf{\Theta}(L)$ is invertible, so there are stationarity conditions here as well. The equations of the VAR system can be estimated with an OLS, since they are ARDLX(p,p) type models, and the residual will be stationary. You do not have to estimate the coefficients simultaneously, you can go by equations. Of course, since the residuals may be correlated with each other, OLS is not guaranteed to be the best estimator (you can use a SUR instead).

Obviously you have to make some choices when estimating a VAR. You need to decide which variables are endogenous and which are exogenous. Since the right hand side of the equations will contain only predetermined and exogenous variables, you do not need to worry about simultaneity bias. You need to choose the order of the VAR system (p). This is done with information criteria, just like you would do in the single equation case.

9.b. The idea behind a VAR

Until the 1970s macroeconometrics had used simultaneous equation systems that we already covered earlier during this course. This required that one precisely determined the relationship among variables, and also required that you had at least the key equations identified so that you can estimate the coefficients. This approach has been criticized following the Lucas critique that raised the problem that even the coefficients of such equation (behavioral parameters) may change over time, so policy recommendations and forecasting from such systems may be misleading and invalid.
During the 1980s Sims suggested the VAR as an alternative solution. This theoretically treated the economic system you were to estimate as a black box:



The main idea was that you do not precisely specify a system of equations but rather you use a very general VAR system and you draw conclusions based on the reaction of the endogenous variables on changes in exogenous factors. These are the impulse response functions.
Sims argued that using a VAR has obvious advantages, like: you need not to be very certain which of your variables are exogenous and endogenous, it is “theory free” (now, it is disputable what this means), and you need not to worry much about identification. (My personal view is that the last two statements are simply false.)

But what is a VAR? So far VAR has been defined as a completely new type of model but this is misleading: VAR is actually a set of reduced form equations form a system of simultaneous equations. (This is the reason why it is better if you learn simultaneous equations methods first…)
Let us assume that we have two macroeconomic variables, say, real aggregate income (y_t) and rate of inflation (π_t). We believe that these are in simultaneous relationship but have a degree of hysteresis (so they can partly be explained by their past values).

$$y_t = \beta_{10} + \beta_{11}\pi_t + \beta_{12}y_{t-1} + u_t \, , u_t \sim IID(0, \sigma_u^2)$$
$$\pi_t = \beta_{20} + \beta_{21}y_t + \beta_{22}\pi_{t-1} + v_t \, , v_t \sim IID(0, \sigma_v^2)$$

Of course, if you have some pre-knowledge, you see that both equations are exactly identified (if you do not see this, please read the notes for lecture 8 again).
Let us derive the reduced form equations:

$$y_t = \delta_1 + \theta_{11} y_{t-1} + \theta_{12}\pi_{t-1} + \eta_{1t}$$

$$\pi_t = \delta_2 + \theta_{21} y_{t-1} + \theta_{22}\pi_{t-1} + \eta_{2t}$$

What you have obtained is a VAR(1) system with:

$$\eta_{1t} = \frac{u_t + \beta_{11} v_t}{1-\beta_{11}\beta_{21}}, \delta_1 = \frac{\beta_{10} + \beta_{11}\beta_{20}}{1-\beta_{11}\beta_{21}}, \theta_{11} = \frac{\beta_{12}}{1-\beta_{11}\beta_{21}}, \theta_{12} = \frac{\beta_{11}\beta_{22}}{1-\beta_{11}\beta_{21}}, \eta_{2t} = \frac{v_t + \beta_{21} u_t}{1-\beta_{21}\beta_{11}},$$

$$\delta_2 = \frac{\beta_{20} + \beta_{21}\beta_{10}}{1-\beta_{11}\beta_{21}}, \theta_{21} = \frac{\beta_{21}\beta_{12}}{1-\beta_{11}\beta_{21}}, \theta_{22} = \frac{\beta_{22}}{1-\beta_{11}\beta_{21}}$$

So all VAR systems are already based on an assumed system of equations, which necessarily involves some theoretical considerations and a set of assumptions. At most you are not aware of them, but that will not make the method free of theory. This is one principal reason why VAR should not be labeled "theory free". You should be aware that η_{1t} and η_{2t} are correlated, since they both depend on u_t and v_t.
The practical application of VAR is like follows: first you choose a proper VAR representation and you draw conclusions on the relationship among the endogenous variables based on IRFs. The IRFs for y are simply: $\frac{\partial y_t}{\partial \eta_{1t-i}}, \frac{\partial y_t}{\partial \eta_{2t-i}}$ and for π: $\frac{\partial \pi_t}{\partial \eta_{1t-i}}, \frac{\partial \pi_t}{\partial \eta_{2t-i}}$ for i=0,…,T. In words, the IRF will show you how an endogenous variable reacts on innovations in its own value (own error term) or on an innovation in another endogenous variables (error term of another equation).
The problem of simultaneity and IRFs becomes apparent when you look at the values of the η's. IRF is only useful if the two etas can be subject to a shock independently. But how is it possible that only η_1 experiences a shock if it is correlated with η_2? Well, the answer is that it is not possible. So, unless your endogenous variables are uncorrelated (so they are not simultaneously correlated…) you need to tell how they are related. When you take account with the contemporaneous correlation structure of your shocks, what you are doing is a Structural Vector Autoregression or SVAR. For a SVAR you need to have some ideas (a theory) in order to explain the contemporaneous correlation structure of your residuals. This is the second reason why you cannot call a VAR "theory free".

9.c Basics of a SVAR

For thinking in terms of an SVAR we use the above example of real GDP and inflation.
you have the following primitive (or structural) form of the VAR system:

$$y_t = \beta_{10} + \beta_{11}\pi_t + \beta_{12} y_{t-1} + u_t$$

$$\pi_t = \beta_{20} + \beta_{21} y_t + \beta_{22}\pi_{t-1} + v_t$$

This can be rewritten as:

$$\begin{pmatrix} y_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} 0 & \beta_{11} \\ \beta_{21} & 0 \end{pmatrix}\begin{pmatrix} y_t \\ \pi_t \end{pmatrix} + \begin{pmatrix} 0 & \beta_{11} \\ \beta_{21} & 0 \end{pmatrix}\begin{pmatrix} y_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}$$

We can rearrange this as follows:

$$\begin{pmatrix} 1 & -\beta_{11} \\ -\beta_{21} & 1 \end{pmatrix}\begin{pmatrix} y_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} 0 & \beta_{11} \\ \beta_{21} & 0 \end{pmatrix}\begin{pmatrix} y_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}$$

The square matrix at the left-hand side contains the coefficients of the contemporary relationships (this is often referred to as matrix A. You can arrive at the VAR (or reduced) form by taking the inverse of this matrix A and premultiply both sides by it.

$$\begin{pmatrix} y_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} 1 & -\beta_{11} \\ -\beta_{21} & 1 \end{pmatrix}^{-1}\begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} 1 & -\beta_{11} \\ -\beta_{21} & 1 \end{pmatrix}^{-1}\begin{pmatrix} 0 & \beta_{11} \\ \beta_{21} & 0 \end{pmatrix}\begin{pmatrix} y_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & -\beta_{11} \\ -\beta_{21} & 1 \end{pmatrix}^{-1}\begin{pmatrix} u_t \\ v_t \end{pmatrix}$$

We know that $\begin{pmatrix} 1 & -\beta_{11} \\ -\beta_{21} & 1 \end{pmatrix}^{-1} = \dfrac{1}{1+\beta_{11}\beta_{21}}\begin{pmatrix} 1 & \beta_{11} \\ \beta_{21} & 1 \end{pmatrix} = \begin{pmatrix} \dfrac{1}{1+\beta_{11}\beta_{21}} & \dfrac{\beta_{11}}{1+\beta_{11}\beta_{21}} \\ \dfrac{\beta_{21}}{1+\beta_{11}\beta_{21}} & \dfrac{1}{1+\beta_{11}\beta_{21}} \end{pmatrix}$

Which will give you actually the same coefficients that we already had earlier (check it!).
So in general, all VAR can be written in a primitive and in a reduced form:
Primitive form:

$$\mathbf{AY}_t = \boldsymbol{\alpha} + \sum_{i=1}^{p} \mathbf{B}_i \mathbf{Y}_{t-i} + \mathbf{u}_t$$

reduced form (which is the VAR):

$$\mathbf{Y}_t = \mathbf{A}^{-1}\boldsymbol{\alpha} + \sum_{i=1}^{p} \mathbf{A}^{-1}\mathbf{B}_i \mathbf{Y}_{t-i} + \mathbf{A}^{-1}\mathbf{u}_t = \boldsymbol{\delta} + \sum_{i=1}^{p} \boldsymbol{\Theta}_\mathbf{i} \mathbf{Y}_\mathbf{t-i} + \mathbf{e}_\mathbf{t}$$

Obviously, unless matrix A is an identity matrix (having ones in the main diagonal and nulls elsewhere), your residuals in the reduced form (your VAR) will be correlated. In the simplest case you do not care about this, this is called the basic case of VAR. This is very useful if you wish to make forecasts, but once you wish to know something the relation of your variables (and this is what you in most cases want) the IRF will be misleading. So actually you need to know the elements of matrix A, and when you do this, you are doing a Structural VAR. And SVAR involves a lot of theory. Using VAR will not save you from having some ideas about the underlying structure of your data.

9.d A basic VAR

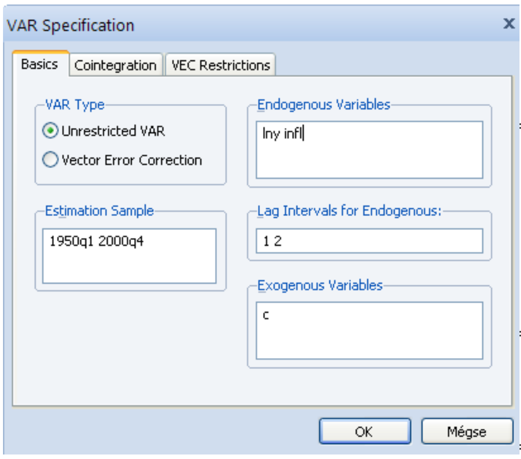
Let us get down to business! We have data for inflation (infl) and the log of real GDP (lny). We will now estimate a two-variate VAR system.

First we need to choose the order of the system. For this purpose we estimate a simple VAR(2), which is the default setting in Eviews:

Vector Autoregression Estimates
Date: 02/29/12 Time: 19:01
Sample (adjusted): 1950Q3 2000Q4
Included observations: 202 after adjustments
Standard errors in ( ) & t-statistics in [ ]

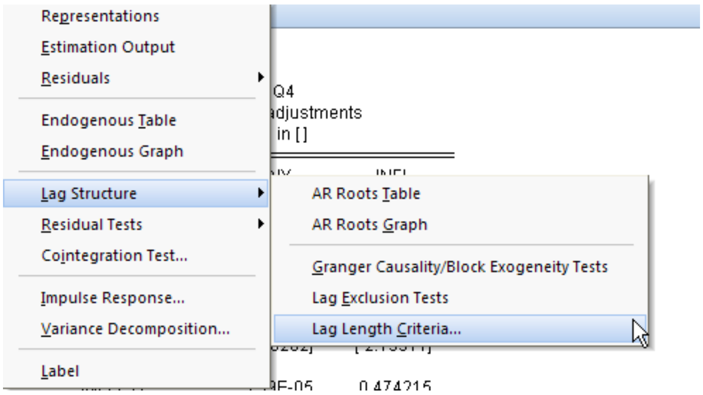
	LNY	INFL
LNY(-1)	1.302451	38.67961
	(0.06586)	(17.7907)
	[19.7749]	[2.17414]
LNY(-2)	-0.302839	-38.37315
	(0.06580)	(17.7727)
	[-4.60262]	[-2.15911]
INFL(-1)	-5.99E-05	0.474215
	(0.00025)	(0.06761)
	[-0.23915]	[7.01439]
INFL(-2)	-0.000580	0.296985
	(0.00025)	(0.06817)
	[-2.29857]	[4.35684]
C	0.011663	-1.989840
	(0.01123)	(3.03287)
	[1.03869]	[-0.65609]
R-squared	0.999650	0.490536
Adj. R-squared	0.999643	0.480192
Sum sq. resids	0.016292	1188.684
S.E. equation	0.009094	2.456406
F-statistic	140855.3	47.42028
Log likelihood	665.3345	-465.6313
Akaike AIC	-6.537965	4.659716
Schwarz SC	-6.456077	4.741604
Mean dependent	8.321353	3.935925
S.D. dependent	0.481558	3.407050

Determinant resid covariance (dof adj.)	0.000499
Determinant resid covariance	0.000474
Log likelihood	199.7817
Akaike information criterion	-1.879027
Schwarz criterion	-1.715251



You could add exogenous variables as well, like seasonal dummies, additional dummies for policy changes, trend or any other variables as your model requires. We do not do so for this exercise, but feel free to experiment!

We have significant coefficients, high R-squared (do not be surprised, the log of real GDP is likely to be non-stationary). Obviously, a good strategy would be to add lags as long as the information criteria are falling. Fortunately, Eviews has a built-in process for this:

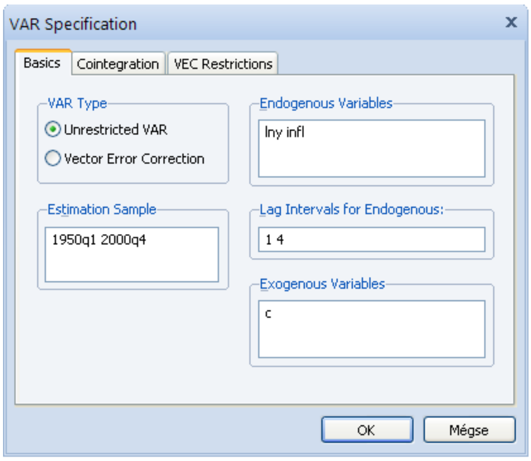


VAR Lag Order Selection Criteria
Endogenous variables: LNY INFL
Exogenous variables: C
Date: 02/29/12 Time: 19:05
Sample: 1950Q1 2000Q4
Included observations: 196

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-636.5025	NA	2.315389	6.515332	6.548782	6.528874
1	182.9504	1613.820	0.000563	-1.805616	-1.705265	-1.764989
2	206.1090	45.13576	0.000463	-2.001112	-1.833862	-1.933401
3	220.5352	27.82187	0.000417	-2.107502	-1.873351*	-2.012706
4	228.8892	15.94082*	0.000399*	-2.151931*	-1.850879	-2.030051*
5	232.7728	7.331222	0.000399	-2.150742	-1.782791	-2.001778
6	236.3812	6.738229	0.000401	-2.146747	-1.711895	-1.970698
7	238.2269	3.408902	0.000410	-2.124764	-1.623012	-1.921631
8	238.9096	1.247006	0.000424	-2.090915	-1.522262	-1.860697

* 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

You can see that the software puts a star to the lag-length that has been found to yield the lowest information criterion (or prediction error) value. Most criteria suggest a VAR(4) model be the best, except for the Schwarz, but this is not surprising, this criterion often suggest a more parsimonious model than the rest. Since we have quarterly data, common sense also suggest that we should move for a VAR(4).



It is worthwhile to test if the residuals fulfill the assumption of the classical linear model:

VAR Residual Serial Correlation LM T...
Null Hypothesis: no serial correlation ...
Date: 02/29/12 Time: 19:18
Sample: 1950Q1 2000Q4
Included observations: 200

Lags	LM-Stat	Prob
1	8.910945	0.0634
2	5.644488	0.2273
3	11.13398	0.0251
4	9.141276	0.0577
5	5.068440	0.2803

Probs from chi-square with 4 df.

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)
Date: 02/29/12 Time: 19:18
Sample: 1950Q1 2000Q4
Included observations: 200

Joint test:

Chi-sq	df	Prob.
119.9939	48	0.0000

Individual components:

Dependent	R-squared	F(16,183)	Prob.	Chi-sq(16)	Prob.
res1*res1	0.152703	2.061301	0.0118	30.54050	0.0154
res2*res2	0.319911	5.380152	0.0000	63.98220	0.0000
res2*res1	0.143590	1.917675	0.0212	28.71807	0.0259

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: residuals are multivariate normal
Date: 02/29/12 Time: 19:19
Sample: 1950Q1 2000Q4
Included observations: 200

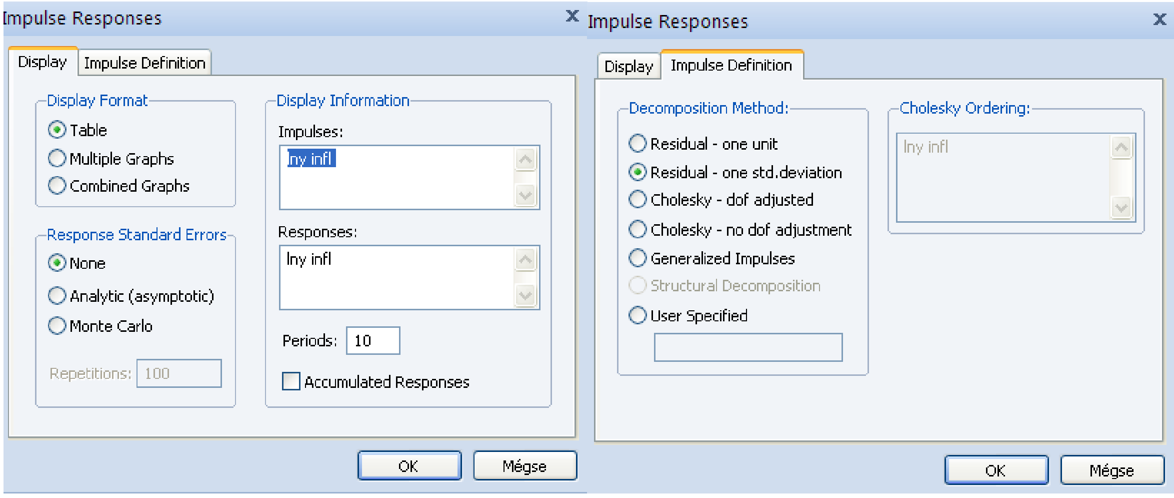
Component	Skewness	Chi-sq	df	Prob.
1	0.055072	0.101096	1	0.7505
2	-0.132448	0.584748	1	0.4445
Joint		0.685844	2	0.7097

Component	Kurtosis	Chi-sq	df	Prob.
1	4.409710	16.56069	1	0.0000
2	4.734731	25.07743	1	0.0000
Joint		41.63812	2	0.0000

Component	Jarque-Bera	df	Prob.
1	16.66178	2	0.0002
2	25.66218	2	0.0000
Joint	42.32396	4	0.0000

It seems that the residuals have no serial correlation at 1% level of significance, but are heteroscedastic and in case of the second equation (inflation) they are also not normally distributed. We can live with these, especially since we cannot do much about these (you could identify possible breakpoints). If you find the residual exhibiting serial correlation you can experiment with adding further lags.

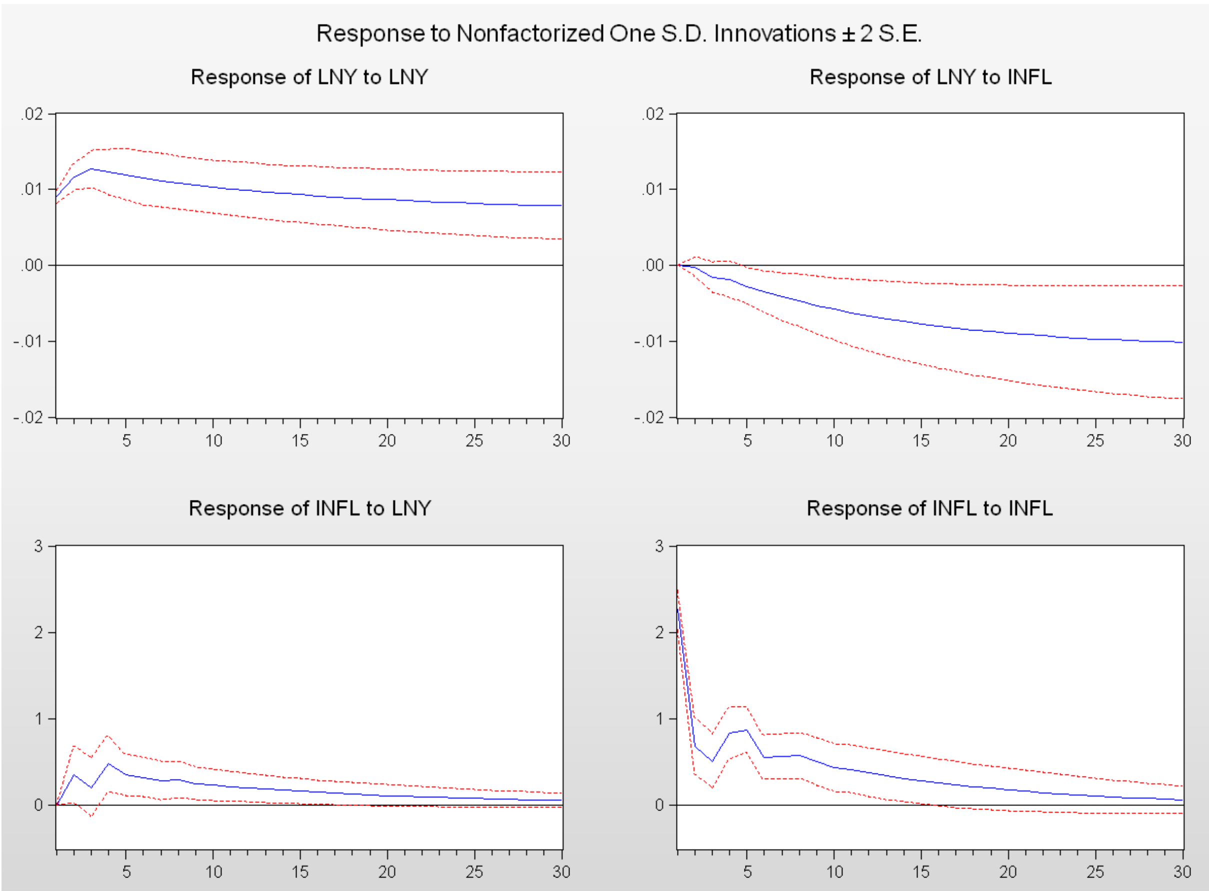
Now, if you are interested in the relationship between the rate of inflation and the log level of real GDP, you normally estimate an impulse-response function (IRF).



You can determine the time horizon of the IRF, you can ask for CIRF. In the Impulse Definition table you can tell Eviews what type of shocks you have in mind.

If you have good reason to believe that the residuals are uncorrelated (no contemporary relation between the two endogenous variables or the matrix A is an identity matrix), you can go for a simple one standard deviation innovation in the residuals. Of course, you can have a one unit shock as well, but do not forget that increasing lny by one unit means a growth by a factor of 2.71, while in the case of the rate of inflation it means one percentage point. So it seems a better idea to set the magnitude of the shock to one standard deviation of the variable, which is about 0.49 for the log GDP and 3.4 percentage points for inflation.

If we ask for a horizon of 30 quarters (7.5 years) we obtain the following IRFs:



The results tells us that in this case of a one-time shock of one std. deviation in log GDP the effect on the log of GDP will not revert to zero even after 30 periods (it is non-stationary…), while it seems to have a transitory positive effect on the rate of inflation. In case of a one-time shock in inflation, we find that it has a negative, seemingly permanent effect on log GDP, and a transitory effect on itself.

9.e. Structural factorization

But are these results satisfactory? It depends on our beliefs about the underlying economic process. Do you believe that real GDP does not affect inflation and vice versa in the same quarter? If the answer is no, you need to have a structural factorization (that is, you need to tell how the matrix A looks). From this point on, you are doing SVAR.

What are the possible solutions?

First, you can assume, that both inflation and GDP affect each other. Then your matrix A is:

$\mathbf{A} = \begin{pmatrix} 1 & a_{12} \\ a_{21} & 1 \end{pmatrix}$ this will cause a problem, however, since this means that you cannot separate the two residuals.

We already found out that the residuals ($\mathbf{e}$) from the VAR system are going to be:

$\mathbf{A}^{-1}\mathbf{u} = \mathbf{e}$ from this you arrive at the following: $\mathbf{u} = \mathbf{Ae}$. You already know that in the primitive form the residuals $\mathbf{u}$ were independent of each other, so $E(\mathbf{uu}') = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}$.(this is the covariance matrix of $\mathbf{u}$). Let us not forget that you do not know $\mathbf{u}$ so here you have two unknowns. We also know that the VAR residuals ($\mathbf{e}$) can be correlated if **A** is not an identity matrix: $\mathbf{A}^{-1}\mathbf{u} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} u_t \\ v_t \end{pmatrix} = \begin{pmatrix} c_{11} \cdot u_t + c_{12} \cdot v_t \\ c_{21} \cdot u_t + c_{22} \cdot v_t \end{pmatrix}$. Which yields k equations: $\begin{matrix} e_1 = c_{11} \cdot u_t + c_{12} \cdot v_t \\ e_2 = c_{21} \cdot u_t + c_{22} \cdot v_t \end{matrix}$, obviously this will not be identifiable, since e_1 and e_2 are both unknown.

Now you need to take the covariance matrixes to identify the SVAR:

$\mathbf{A}^{-1}\mathbf{uu}'\mathbf{A}'^{-1} = \mathbf{ee}' \rightarrow \mathbf{A}^{-1}E(\mathbf{uu}')(\mathbf{A}^{-1})' = \mathbf{\Sigma}$ where:

$\mathbf{\Sigma} = E(\mathbf{ee}') = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_1e_2} \\ \sigma_{e_2e_1} & \sigma_{e_2}^2 \end{bmatrix}$ and $E(\mathbf{uu}') = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}$

this yields:

$$\begin{pmatrix} \sigma_{e_1}^2 & \sigma_{e_1e_2} \\ \sigma_{e_2e_1} & \sigma_{e_2}^2 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} = \begin{pmatrix} c_{11}\sigma_u^2 & c_{12}\sigma_v^2 \\ c_{21}\sigma_u^2 & c_{22}\sigma_v^2 \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} c_{11}^2\sigma_u^2 + c_{12}^2\sigma_v^2 & c_{11}c_{21}\sigma_u^2 + c_{12}c_{22}\sigma_v^2 \\ c_{11}c_{21}\sigma_u^2 + c_{12}c_{22}\sigma_v^2 & c_{21}^2\sigma_u^2 + c_{22}^2\sigma_v^2 \end{pmatrix}$$

Which leads to four equations:

$\sigma_{e_1}^2 = c_{11}^2\sigma_u^2 + c_{12}^2\sigma_v^2$

$\sigma_{e_1e_2} = c_{11}c_{21}\sigma_u^2 + c_{12}c_{22}\sigma_v^2$

$\sigma_{e_2e_1} = c_{11}c_{21}\sigma_u^2 + c_{12}c_{22}\sigma_v^2$

$\sigma_{e_2}^2 = c_{21}^2\sigma_u^2 + c_{22}^2\sigma_v^2$

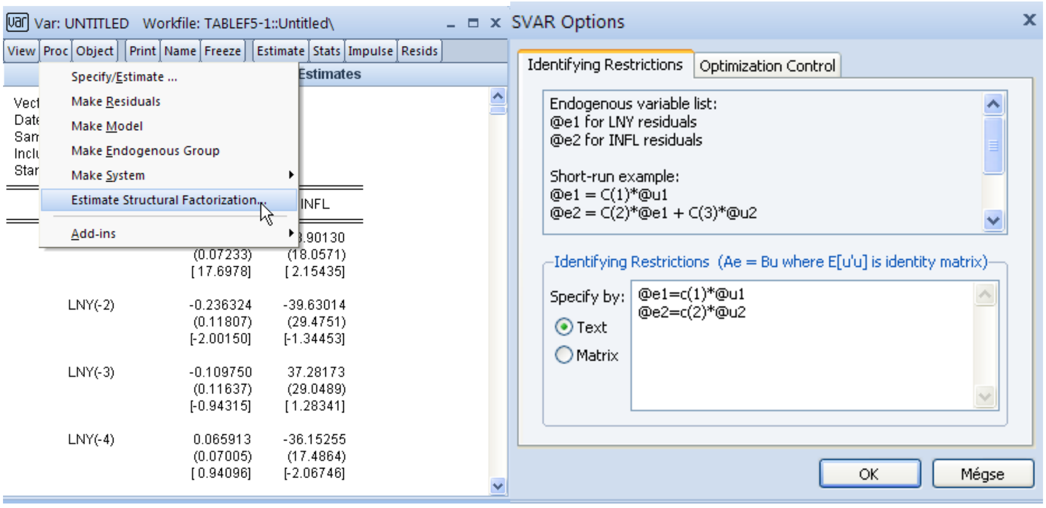
Due to symmetry the 2[nd] and 3[rd] equation are the same. You have now six unknowns (c_{11},c_{12},c_{21},c_{22}, σ_u^2 , σ_v^2) and three equations so you cannot solve this. Of course, some restrictions are straightforward: the main diagonal has only ones (this introduces two restrictions: c_{11}= c_{22}=1), so now you have three equations for four unknowns. This is still not enough you need at least one restriction yet.

Possible solution strategies:

1. Arguing that the two variables (log of GDP and rate of inflation) are nor contemporaneously correlated. This is tantamount with setting **A**=**I** (c_{12}= c_{21}=0, c_{11}= c_{22}=1) then you have: $\sigma_{e_1}^2 = \sigma_u^2, \sigma_{e_1e_2} = 0, \sigma_{e_2e_1} = 0, \sigma_{e_2}^2 = \sigma_v^2$. Now you have two additional restrictions, yo you overidentify your SVAR. This can be done, so this would lead to valid (but not necessarily correct) results.
2. Arguing that changes in one variable can affect the other variable immediately, but not vice versa. This is the lower triangular matrix approach, a.k.a., Cholesky factorization. If we assume that the first variable affects the second but not vice versa. Then **A** is: $\begin{pmatrix} 1 & 0 \\ a_{21} & 1 \end{pmatrix}$ so $A^{-1} = \begin{pmatrix} 1 & 0 \\ -a_{21} & 1 \end{pmatrix}$, so the equations are: $\sigma_{e_1}^2 = \sigma_u^2, \sigma_{e_1e_2} = \sigma_{e_2e_1} = -a_{21}\sigma_u^2, \sigma_{e_2}^2 = a_{21}^2\sigma_u^2 + \sigma_v^2$
   This also means that you argue that the a shock in the value of the first variable is completely responsible for its measured variance, while observed variation in the second variable is a linear combination of the shocks to both variables. That is, e_2 depends on e_1, while e_1 does not depend on e_2. Now you have identified your SVAR exactly, as you introduced only a single additional restriction.
3. Finally, you can have any theoretically based restriction as long as the SVAR is identified. In this case you should add one restriction. If you had 3 endogenous variables, you needed to have 3 restrictions on A besides setting the elements in the main diagonal to zero. The general rule is to have k(k+1)/2 restrictions.

> The Eviews has a bit different notation:
> $\mathbf{Ae}_t = \mathbf{Bu}_t$ where $\mathbf{e}_t$ denotes you residuals from the VAR and $\mathbf{u}_t$ are the shocks. We assume that where $E(\mathbf{uu}') = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$. So matrix **B** is set in that was that: $E(\mathbf{BB}') = \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}$. Actually what we called $\mathbf{u}_t$ earlier is now denoted as $\mathbf{Bu}_t$. For the rest it is the same as before.

Let us apply this knowledge to our data! Structural factorization is done as follows. You can define restrictions on A in two ways: by equations or by matrices.



Now I show you the equation version (for the matrix version you need to define a matrix A and another Matrix B, with NAs for the elements that you want to estimate).

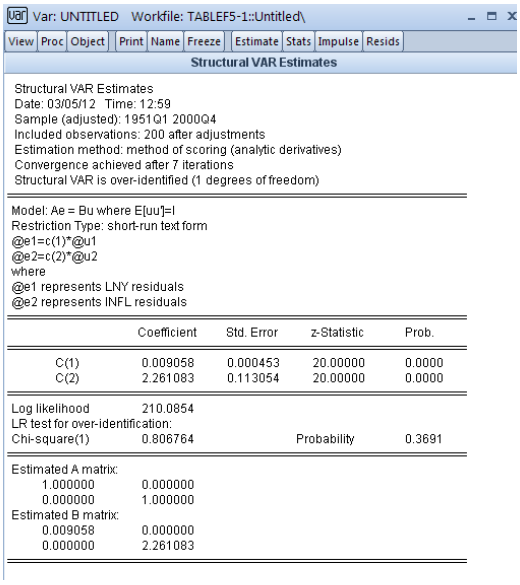
@e denotes elements of the VAR residual vector, @u are the shocks (basically, the error terms from the primitive form equations).

When you assume no contemporary relationship you would write:

$$@e1 = c(1)*@u1$$

$$@e2 = c(2)*@u2$$

That is, the observed VAR residuals for each equation depend only on the shocks of their respective errors.


Var: UNTITLED Workfile: TABLEF5-1::Untitled\
View Proc Object Print Name Freeze Estimate Stats Impulse Resids

Structural VAR Estimates

Structural VAR Estimates
Date: 03/05/12 Time: 12:59
Sample (adjusted): 1951Q1 2000Q4
Included observations: 200 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 7 iterations
Structural VAR is over-identified (1 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@e1=c(1)*@u1
@e2=c(2)*@u2
where
@e1 represents LNY residuals
@e2 represents INFL residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.009058	0.000453	20.00000	0.0000
C(2)	2.261083	0.113054	20.00000	0.0000

Log likelihood	210.0854		
LR test for over-identification:			
Chi-square(1)	0.806764	Probability	0.3691

Estimated A matrix:

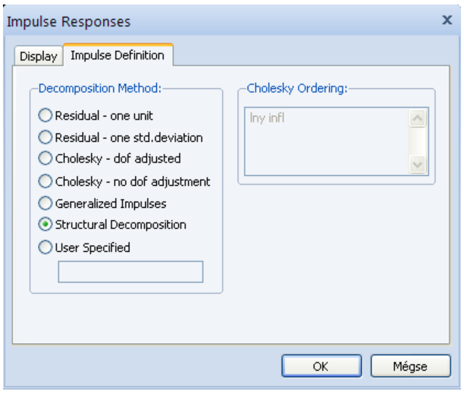
1.000000	0.000000
0.000000	1.000000

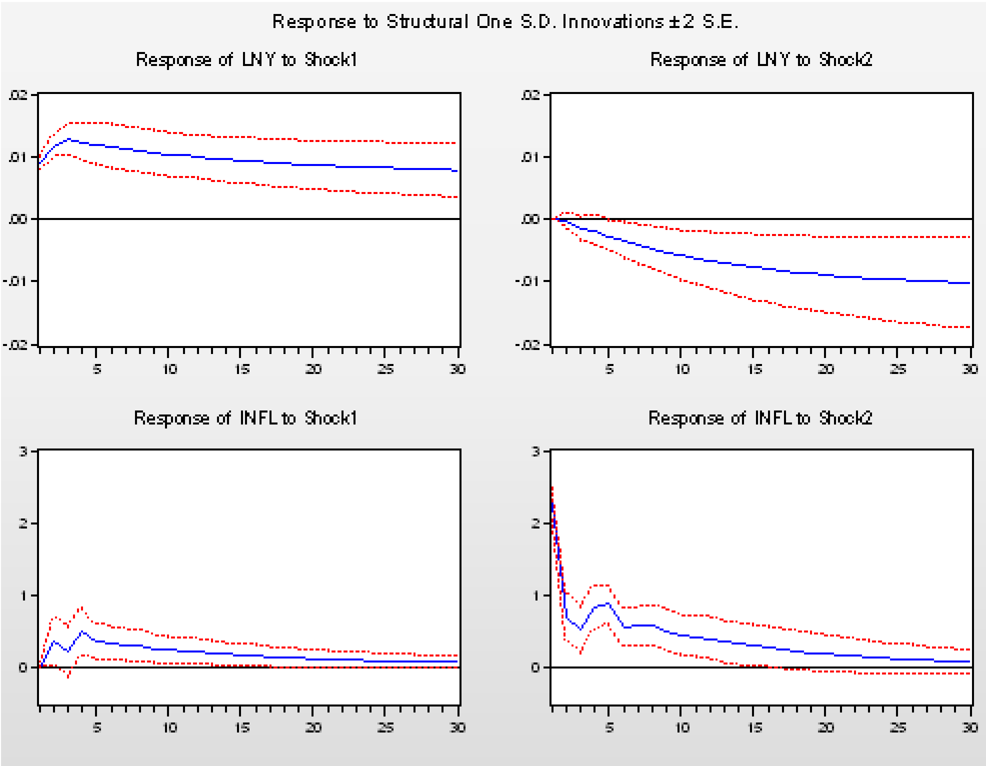
Estimated B matrix:

0.009058	0.000000
0.000000	2.261083

The output that you receive shows you that the SVAR is now overidentified (you needed at least 1 restriction but we have two). Plus, you obtain the restrictions in equivalent matrix form (matrix A and matrix B) as well.

Now you obtain the IRFs as follows:





Not surprisingly this is exactly the same as before, without structural factorization, since this is the baseline case.

You can assume that innovations in the log of GDP affect the innovation sin inflation but not vice versa. This is written as:

$$@e1 = c(1) * @u1$$

$$@e2 = c(2) * @u2 + c(3) * @e1$$

Now we get:

Structural VAR Estimates
Date: 03/05/12 Time: 12:59
Sample (adjusted): 1951Q1 2000Q4
Included observations: 200 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 7 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@e1=c(1)*@u1
@e2=c(2)*@u2+c(3)*@e1
where
@e1 represents LNY residuals
@e2 represents INFL residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(3)	-15.83881	17.61615	-0.899107	0.3686
C(1)	0.009058	0.000453	20.00000	0.0000
C(2)	2.256527	0.112826	20.00000	0.0000
Log likelihood	210.4888			

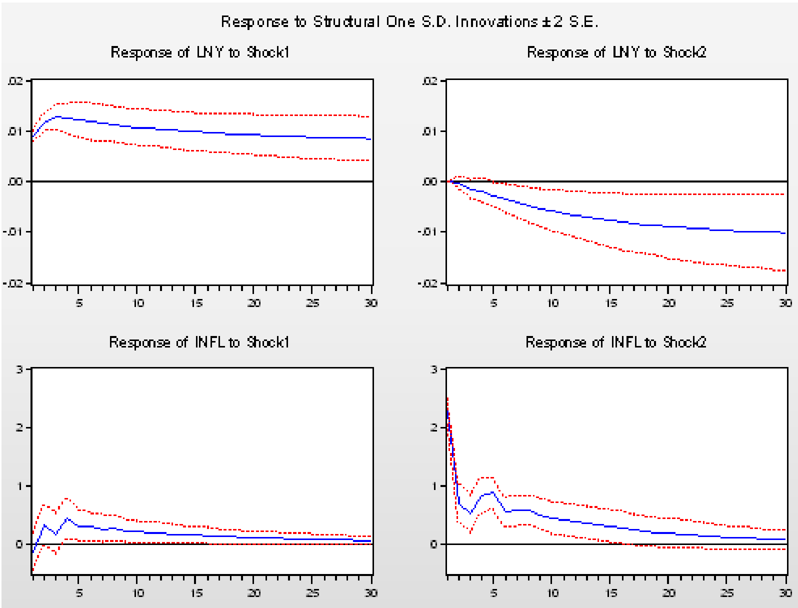
Estimated A matrix:

1.000000	0.000000
15.83881	1.000000

Estimated B matrix:

0.009058	0.000000
0.000000	2.256527

The SVAR is now exactly identified, but we find the estimated element of the A matrix statistically insignificant which suggests that our assumption regarding the contemporaneous correltion structure may be wrong.



The IRFs are just a little bit different. You can have Cholesky-type factorization directly from your response option, without the need to estimate a factorization like this. Do not forget, however, that with this type of factorization the order of your variables counts. So you should have the "most exogenous" variable first and so on.

9.f. Another example

Let us generalize what we learnt here to a more complex model.

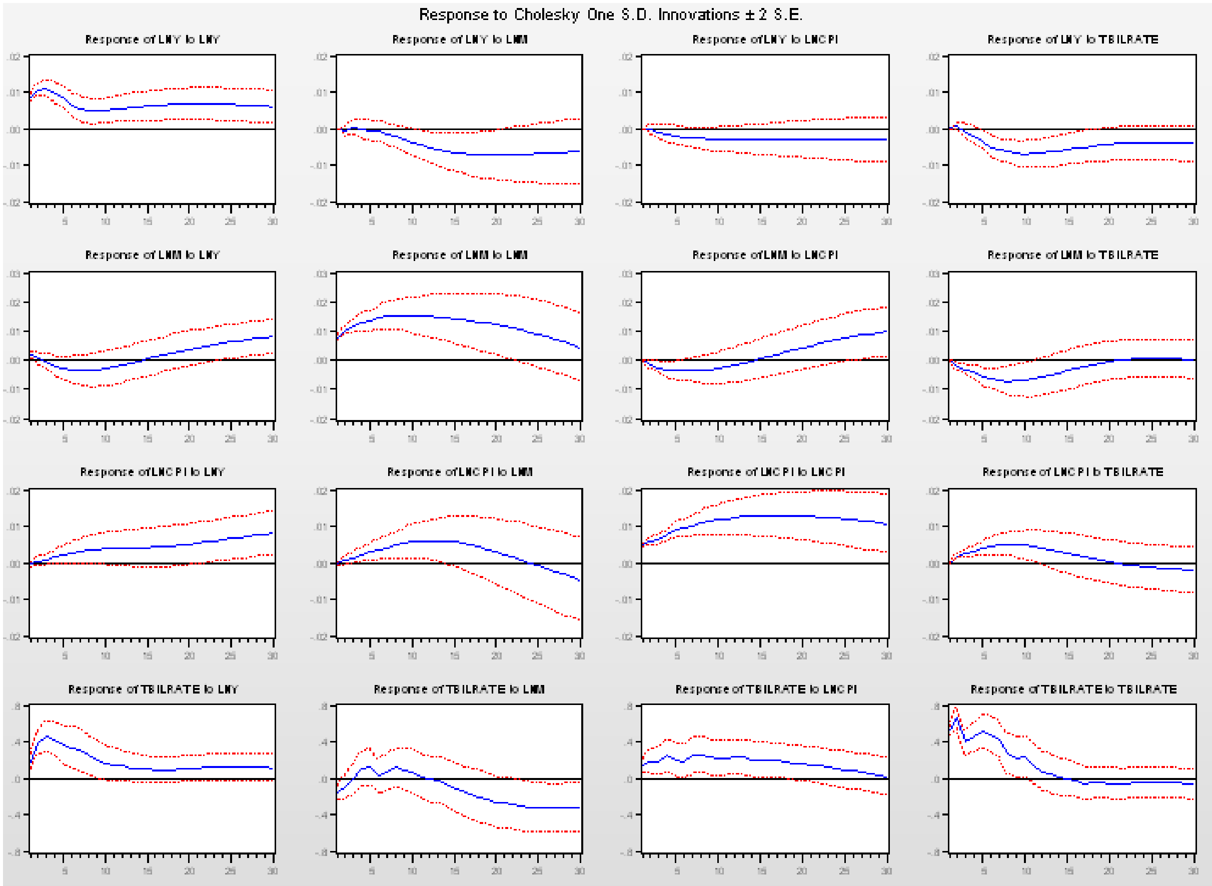
We will estimate a system with log real gdp, log of M1, log of CPI and tbilrate. (You can guess that this is now a standard money demand equation.) For lag length we obtain:

VAR Lag Order Selection Criteria
Endogenous variables: LNY LNM LNCPI TBILRATE
Exogenous variables: C
Date: 03/05/12 Time: 16:22
Sample: 1950Q1 2000Q4
Included observations: 196

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-205.7798	NA	9.99e-05	2.140611	2.207511	2.167695
1	1789.014	3887.813	1.70e-13	-18.05117	-17.71667	-17.91575
2	1881.796	177.0425	7.77e-14	-18.83465	-18.23255	-18.59089
3	1931.692	93.17209	5.50e-14	-19.18053	-18.31082*	-18.82843*
4	1956.707	45.69119	5.02e-14	-19.27252	-18.13521	-18.81208
5	1981.326	43.96234	4.61e-14	-19.36047	-17.95556	-18.79169
6	1997.935	28.98240*	4.59e-14*	-19.36669*	-17.69418	-18.68958
7	2008.031	17.20411	4.89e-14	-19.30644	-17.36633	-18.52099
8	2021.065	21.67861	5.06e-14	-19.27617	-17.06846	-18.38239

* 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

Now we prefer a VAR system of order 6. Do not forget to look at the residual statistics! For the IRF we obtain:

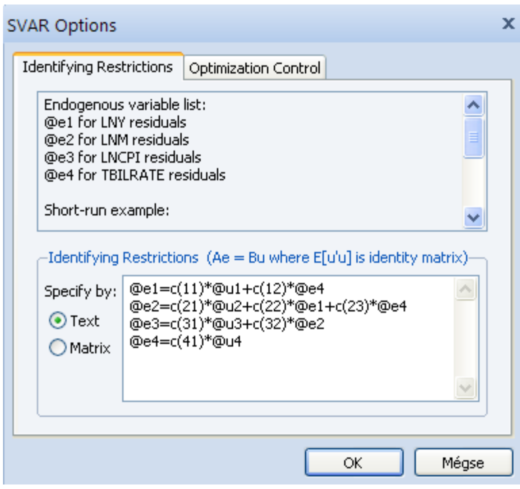


This is now under the assumption that the endogenous variables of our system are contemporarily uncorrelated, so it is possible to have a shock in one of them without necessarily having a shock in any other.

Let us think about the feasibility of this! We can indeed assume that the treasury bill rate is indeed contemporary exogenous, since its value is set for 3 months. You can assume that lnCPI depends on lnM1, the log of real GDP and lnCPI depends on interest rate.

Of course a lot of other strategies are possible, but let us take this now. What we had just said about the variables yields the following matrix **A**:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & a_{14} \\ a_{21} & 1 & 0 & a_{24} \\ 0 & a_{32} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$ or written as equations:



We obtain:

Structural VAR Estimates
Date: 03/05/12 Time: 16:24
Sample (adjusted): 1951Q3 2000Q4
Included observations: 198 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 14 iterations
Structural VAR is over-identified (2 degrees of freedom)

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@e1=c(11)*@u1+c(12)*@e4
@e2=c(21)*@u2+c(22)*@e1+c(23)*@e4
@e3=c(31)*@u3+c(32)*@e2
@e4=c(41)*@u4
where
@e1 represents LNY residuals
@e2 represents LNM residuals
@e3 represents LNCPI residuals
@e4 represents TBILRATE residuals

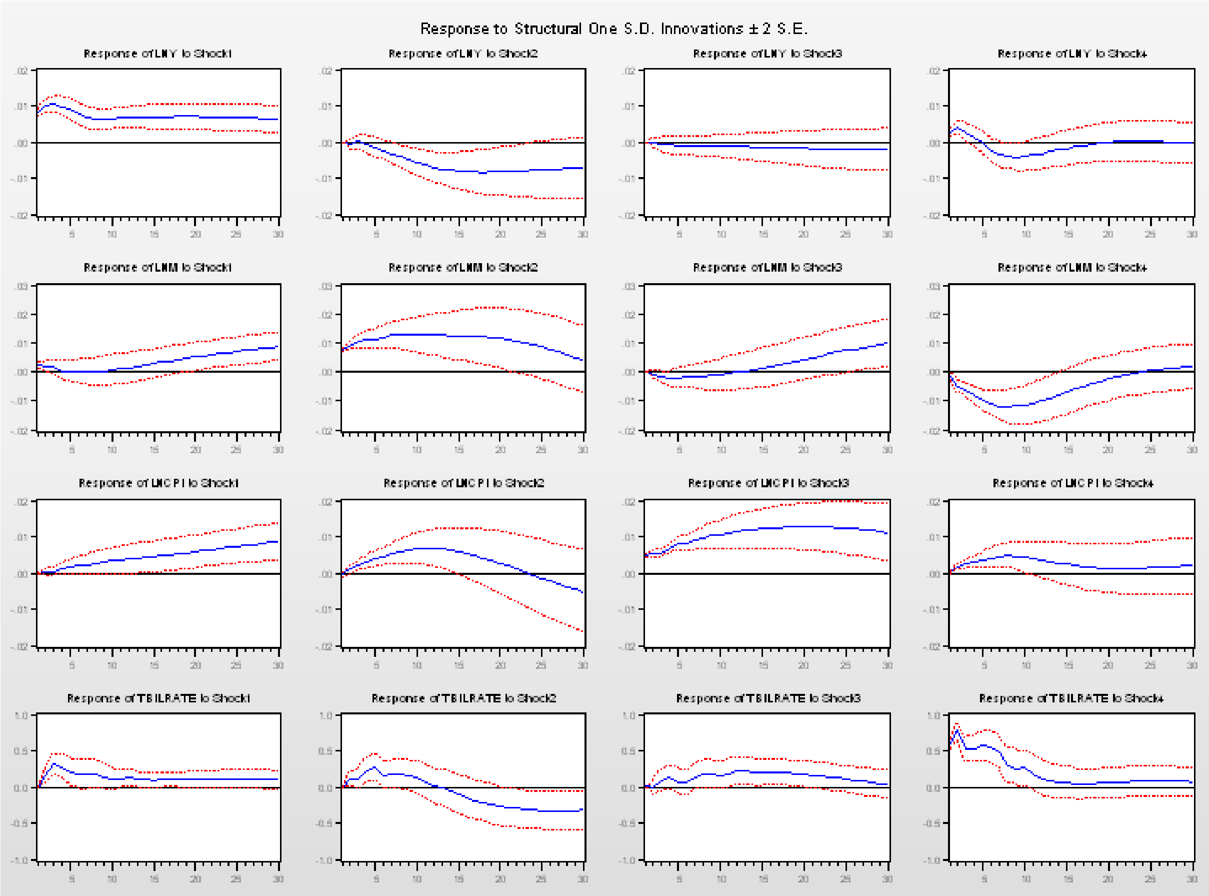
	Coefficient	Std. Error	z-Statistic	Prob.
C(12)	0.004567	0.000987	4.627977	0.0000
C(22)	0.268693	0.064992	4.134228	0.0000
C(23)	-0.003986	0.000950	-4.195732	0.0000
C(32)	-0.038869	0.042325	-0.918346	0.3584
C(11)	0.008253	0.000415	19.89975	0.0000
C(21)	0.007548	0.000379	19.89975	0.0000
C(31)	0.004786	0.000241	19.89975	0.0000
C(41)	0.594358	0.029868	19.89975	0.0000

Log likelihood	1954.312			
LR test for over-identification:				
Chi-square(2)	13.77818		Probability	0.0010

Estimated A matrix:			
1.000000	0.000000	0.000000	-0.004567
-0.268693	1.000000	0.000000	0.003986
0.000000	0.038869	1.000000	0.000000
0.000000	0.000000	0.000000	1.000000
Estimated B matrix:			
0.008253	0.000000	0.000000	0.000000
0.000000	0.007548	0.000000	0.000000
0.000000	0.000000	0.004786	0.000000
0.000000	0.000000	0.000000	0.594358

Since we used 8 restrictions (and we need at least 4*3/2=6), we overidentifed the system. We find that only one of the expected correlations (between lnCPI and lnM1) is not significant, we could even set it to zero. Feel free to experience further with reasonable restrictions!

The IRFs are:



You find just small differences compared to the baseline case, but now, for example, the reaction of real GDP on a shock in treasury bill rate is different than in the first case.

9.g. Long-run restrictions

Restricting matrix **A** was about defining immediate of contemporary relationships. For this reason this is called a short-run identification approach of SVARs. You can alternatively go for defining long-run relationships as well. Using the matrix A we could write a VAR(1) system in the following form:

$$\mathbf{Y_t} = \mathbf{A}^{-1}\boldsymbol{\delta} + \mathbf{A}^{-1}\boldsymbol{\Theta}_1\mathbf{Y_{t\text{-}1}} + \mathbf{A}^{-1}\mathbf{u}_t$$

In order to be able to say something about the IRFs, you need to convert this to an VMA(∞) form:

$$(I - \mathbf{A}^{-1}\boldsymbol{\Theta}_1 L)\mathbf{Y_t} = \mathbf{A}^{-1}\boldsymbol{\delta} + \mathbf{A}^{-1}\mathbf{u}_t \rightarrow \mathbf{Y_t} = (I - \mathbf{A}^{-1}\boldsymbol{\Theta}_1 L)^{-1}\mathbf{A}^{-1}\boldsymbol{\delta} + (I - \mathbf{A}^{-1}\boldsymbol{\Theta}_1 L)^{-1}\mathbf{A}^{-1}\mathbf{u}_t$$

So:

$$\mathbf{e_t} = (I - \mathbf{A}^{-1}\boldsymbol{\Theta}_1 L)^{-1}\mathbf{A}^{-1}\mathbf{u_t} = \mathbf{A}^{-1}\mathbf{u_t} + \psi_1\mathbf{A}^{-1}\mathbf{u}_{t-1} + \psi_2\mathbf{A}^{-1}\mathbf{u}_{t-2} + \ldots$$

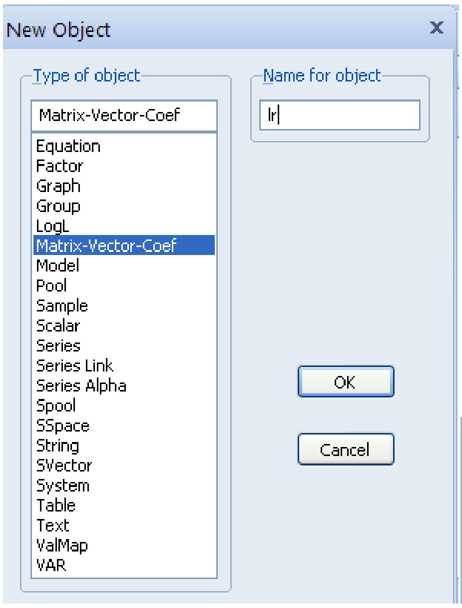
Let $\boldsymbol{\Psi}_\infty = \frac{\partial \mathbf{e_t}}{\partial \mathbf{u_t}} + \frac{\partial \mathbf{e_t}}{\partial \mathbf{u_{t\text{-}1}}} + \frac{\partial \mathbf{e_t}}{\partial \mathbf{u_{t\text{-}2}}} + \ldots = \left(I - \mathbf{A}^{-1}\boldsymbol{\Theta}_1\right)^{-1}$ which is the accumulated IRF or CIRF. This says us to how much the effect of a shock in the residual is going to sum up in the long-run. Or alternatively, this is the long-run effect of a permanent increase in one of the endogenous variables. We call this matrix that has the long-run effect **C**.

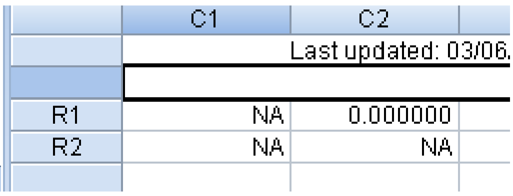
$\mathbf{C} = \boldsymbol{\Psi}_{\infty}\mathbf{A}^{-1}$, matrix **C** will be a kxk matrix with each element having the long-run impact of shock in variable j on variable i.

Let us returns to our log GDP inflation example, where k=2.

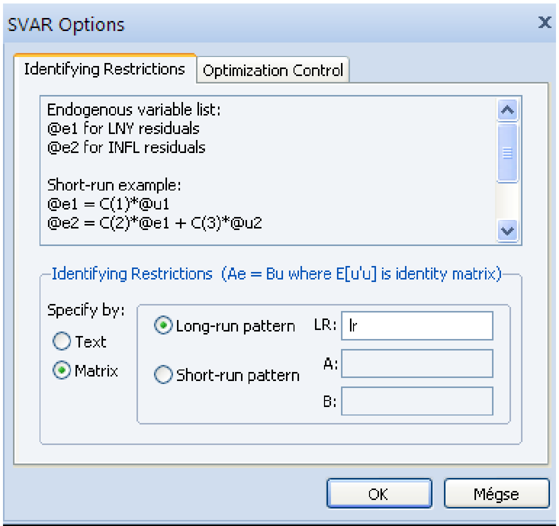
If we believe that inflation should have no effect on the log of real GDP in the long-run (classical dichotomy) than you should use a long-run restriction matrix like this:

$\mathbf{C} = \begin{pmatrix} NA & NA \\ 0 & NA \end{pmatrix}$ remember, by setting an element of a matrix to NA in Eviews, you ask the software to estimate the value of that element. So here you make a single restriction. First, define a new matrix called lr (you cannot use C as a name of a new object since it is preserved for constant):





Now you can use it for identifying the VAR:



We obtain the following solution:

Structural VAR Estimates
Date: 03/06/12 Time: 12:59
Sample (adjusted): 1951Q1 2000Q4
Included observations: 200 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 8 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I
Restriction Type: long-run pattern matrix
Long-run response pattern:

C(1)	0
C(2)	C(3)

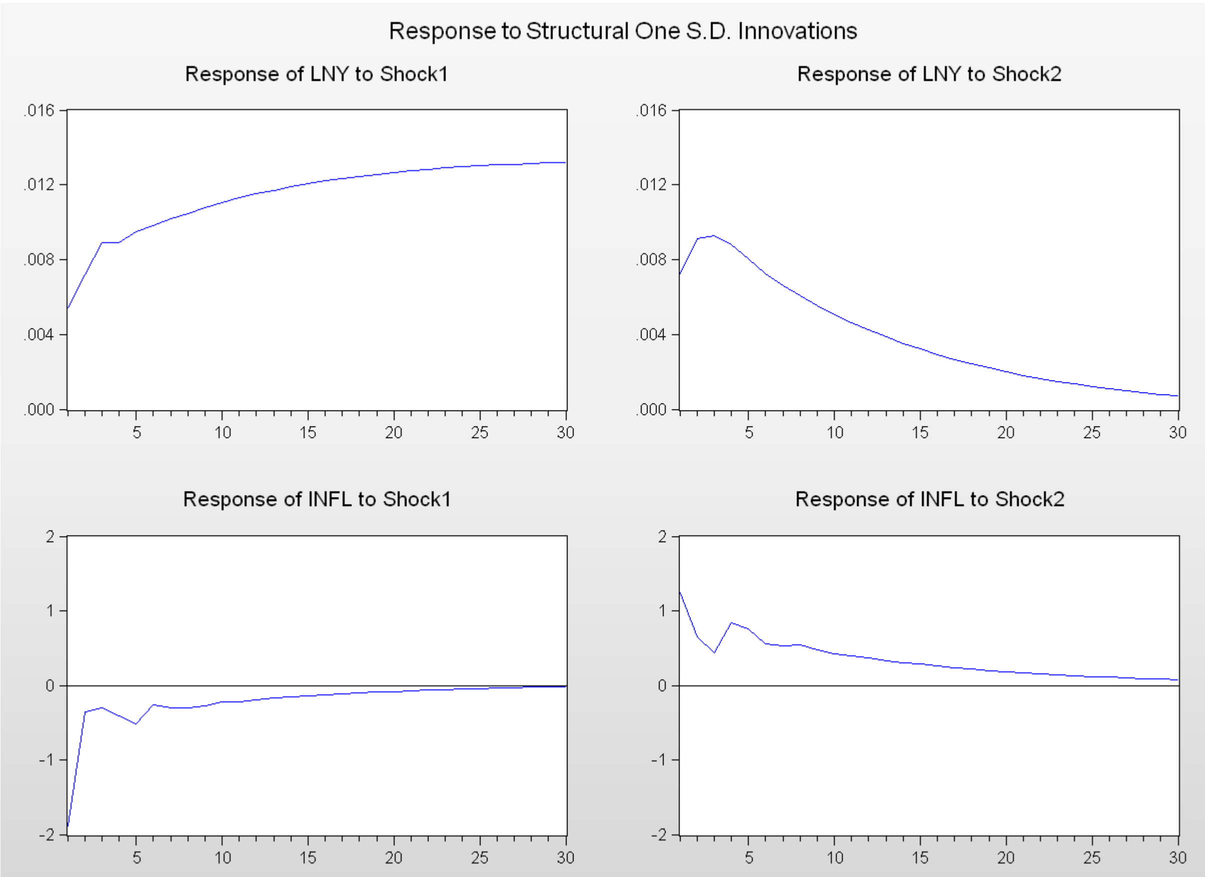
	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	7.045819	0.352291	20.00000	0.0000
C(2)	8.244436	0.879757	9.371261	0.0000
C(3)	10.99133	0.549566	20.00000	0.0000
Log likelihood	210.4888			

Estimated A matrix:

1.000000	0.000000
0.000000	1.000000

Estimated B matrix:

0.005433	0.007247
-1.891553	1.238758



Comparing these results with the baseline case you will find a major difference! Higher rate of inflation will in the short-run contribute to a higher real GDP but this effect wears out (this is what we forced on our data by the long run restrictions). Shock in real GDP also leads to a positive surge in inflation rate and this is also transitory.

Feel free to play with different restrictions. Using long-run restrictions can be very useful, because sometimes we can be unsure about the contemporaneous correlation structure of the data, while we can be quite sure about long-run responses (classical dichotomy was a quite certain point of departure in this example).