Computer Handout 12: Vector Autoregressions

Diego Escobari

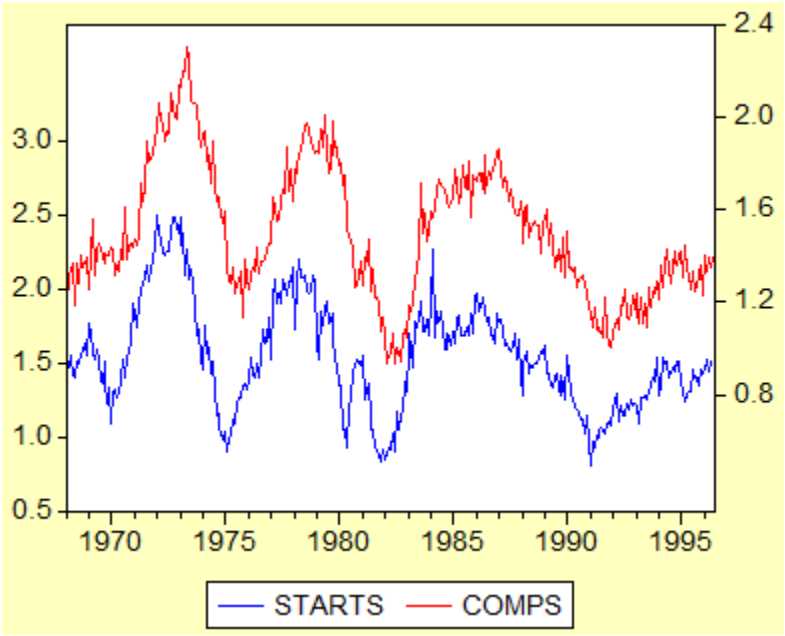
Econ 3342

This Computer Handout 12 will replicate the results in Chapter 11 showing the computer commands for the estimation of Vector Autoregressions (VAR), forecasting with regression models and Impulse-response functions (IRF).

The data we will use includes two variables: (1) the seasonally adjusted housing starts and (2) housing completions. These are monthly observations from January 1968 through June 1996.

A graph of both variables can be obtained by:

```
group both starts comps
both.line(d)
```

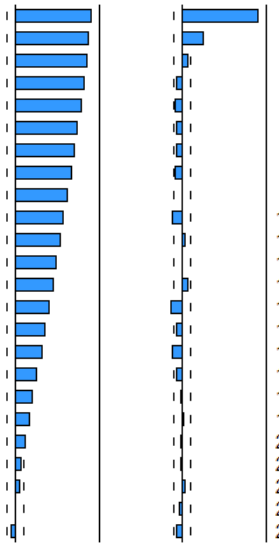


We will use the data from January 1968 through December 1991 for model estimation and the forecast will be done for the period from January 1992 through June 1996.

The correlograms for both variables are:

Correlogram of STARTS

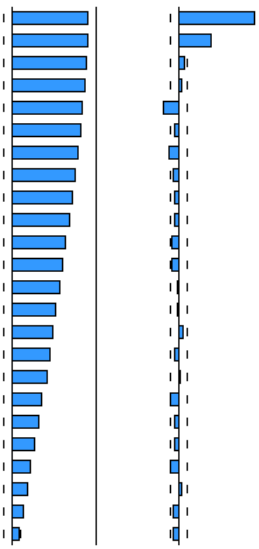
Date: 11/22/10 Time: 17:43
Sample: 1968:01 1996:06
Included observations: 342



		AC	PAC	Q-Stat	Prob
	1	0.940	0.940	304.66	0.000
	2	0.914	0.263	593.59	0.000
	3	0.888	0.070	866.99	0.000
	4	0.853	-0.062	1120.5	0.000
	5	0.816	-0.086	1352.9	0.000
	6	0.778	-0.061	1564.7	0.000
	7	0.737	-0.062	1755.3	0.000
	8	0.689	-0.097	1922.6	0.000
	9	0.648	-0.004	2071.0	0.000
	10	0.594	-0.120	2196.2	0.000
	11	0.553	0.038	2304.9	0.000
	12	0.510	0.008	2397.6	0.000
	13	0.475	0.076	2478.3	0.000
	14	0.422	-0.146	2542.0	0.000
	15	0.375	-0.064	2592.5	0.000
	16	0.321	-0.116	2629.8	0.000
	17	0.268	-0.063	2655.9	0.000
	18	0.221	-0.013	2673.5	0.000
	19	0.174	0.017	2684.5	0.000
	20	0.127	-0.022	2690.4	0.000
	21	0.079	-0.017	2692.7	0.000
	22	0.041	0.038	2693.3	0.000
	23	-0.005	-0.025	2693.3	0.000
	24	-0.047	-0.060	2694.1	0.000

Correlogram of COMPS

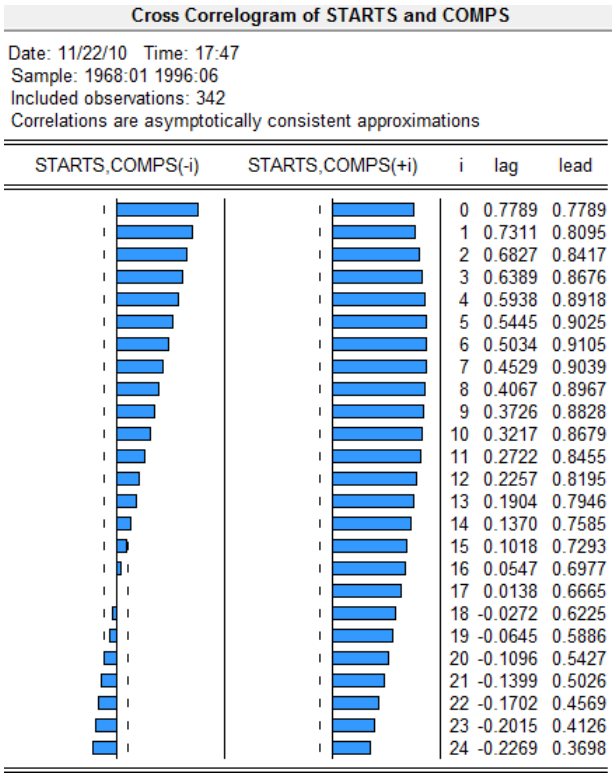
Date: 11/22/10 Time: 17:45
Sample: 1968:01 1996:06
Included observations: 342



		AC	PAC	Q-Stat	Prob
	1	0.947	0.947	309.40	0.000
	2	0.939	0.406	614.29	0.000
	3	0.919	0.073	907.27	0.000
	4	0.904	0.027	1191.5	0.000
	5	0.870	-0.196	1455.6	0.000
	6	0.848	-0.057	1707.4	0.000
	7	0.813	-0.121	1939.8	0.000
	8	0.784	-0.061	2156.3	0.000
	9	0.748	-0.052	2354.1	0.000
	10	0.714	-0.056	2534.6	0.000
	11	0.672	-0.090	2694.9	0.000
	12	0.630	-0.092	2836.6	0.000
	13	0.591	-0.015	2961.3	0.000
	14	0.549	-0.022	3069.2	0.000
	15	0.511	0.050	3163.1	0.000
	16	0.465	-0.056	3241.1	0.000
	17	0.427	0.010	3307.1	0.000
	18	0.377	-0.112	3358.6	0.000
	19	0.335	-0.055	3399.4	0.000
	20	0.287	-0.050	3429.4	0.000
	21	0.237	-0.109	3450.1	0.000
	22	0.194	0.037	3463.9	0.000
	23	0.146	-0.060	3471.7	0.000
	24	0.097	-0.061	3475.2	0.000

Both show a strong cyclical component.

The cross-correlation function shows the correlation between a variable and the lags of another variable. To obtain it open both variables as a group, then go to “view” and then “cross-correlation” to obtain:



This cross-correlation shows a strong correlation between both variables.

The VAR(4) as presented in the textbook (chapter 11) can be estimated using EViews in two different ways: (1) Equation by equation and (2) jointly. Equation by equation we simply

smpl 1968m01 1991m12

ls starts c starts(-1) starts(-2) starts(-3) starts(-4) comps(-1) comps(-2) comps(-3) comps(-4)

ls comps c starts(-1) starts(-2) starts(-3) starts(-4) comps(-1) comps(-2) comps(-3) comps(-4)

that yield the following output:

Dependent Variable: STARTS
Method: Least Squares
Date: 11/22/10 Time: 21:21
Sample (adjusted): 1968M05 1991M12
Included observations: 284 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.146871	0.044235	3.320264	0.0010
STARTS(-1)	0.659939	0.061242	10.77587	0.0000
STARTS(-2)	0.229632	0.072724	3.157587	0.0018
STARTS(-3)	0.142859	0.072655	1.966281	0.0503
STARTS(-4)	0.007806	0.066032	0.118217	0.9060
COMPS(-1)	0.031611	0.102712	0.307759	0.7585
COMPS(-2)	-0.120781	0.103847	-1.163069	0.2458
COMPS(-3)	-0.020601	0.100946	-0.204078	0.8384
COMPS(-4)	-0.027404	0.094569	-0.289779	0.7722

R-squared	0.895566	Mean dependent var	1.574771
Adjusted R-squared	0.892528	S.D. dependent var	0.382362
S.E. of regression	0.125350	Akaike info criterion	-1.284241
Sum squared resid	4.320952	Schwarz criterion	-1.168605
Log likelihood	191.3622	F-statistic	294.7796
Durbin-Watson stat	1.991908	Prob(F-statistic)	0.000000

Dependent Variable: COMPS
Method: Least Squares
Date: 11/22/10 Time: 21:22
Sample (adjusted): 1968M05 1991M12
Included observations: 284 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.045347	0.025794	1.758045	0.0799
STARTS(-1)	0.074724	0.035711	2.092461	0.0373
STARTS(-2)	0.040047	0.042406	0.944377	0.3458
STARTS(-3)	0.047145	0.042366	1.112805	0.2668
STARTS(-4)	0.082331	0.038504	2.138238	0.0334
COMPS(-1)	0.236774	0.059893	3.953313	0.0001
COMPS(-2)	0.206172	0.060554	3.404742	0.0008
COMPS(-3)	0.120998	0.058863	2.055593	0.0408
COMPS(-4)	0.156729	0.055144	2.842160	0.0048

R-squared	0.936835	Mean dependent var	1.547958
Adjusted R-squared	0.934998	S.D. dependent var	0.286689
S.E. of regression	0.073093	Akaike info criterion	-2.362995
Sum squared resid	1.469205	Schwarz criterion	-2.247359
Log likelihood	344.5453	F-statistic	509.8375
Durbin-Watson stat	2.013370	Prob(F-statistic)	0.000000

These are tables 11.4 and 11.6 in your textbook.

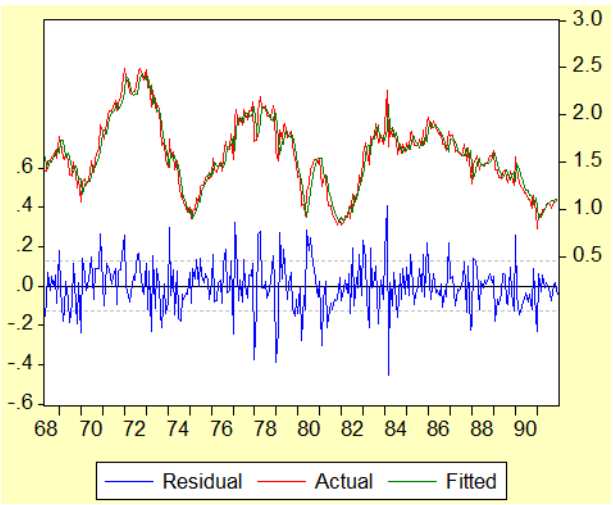
For the correlogram of the residuals (see previous handouts) we have:

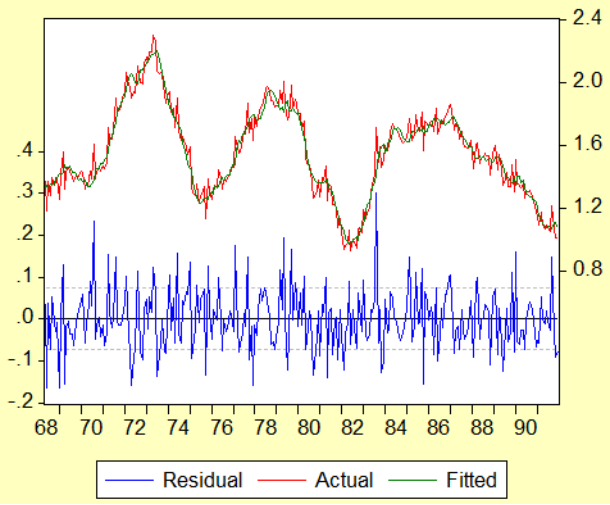
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.009	-0.009	0.0238	0.877
		2	-0.035	-0.035	0.3744	0.829
		3	-0.037	-0.037	0.7640	0.858
		4	-0.088	-0.090	3.0059	0.557
		5	-0.105	-0.111	6.1873	0.288
		6	0.012	-0.000	6.2291	0.398
		7	-0.024	-0.041	6.4047	0.493
		8	0.041	0.024	6.9026	0.547
		9	0.048	0.029	7.5927	0.576
		10	0.045	0.037	8.1918	0.610
		11	-0.009	-0.005	8.2160	0.694
		12	-0.050	-0.046	8.9767	0.705
		13	-0.038	-0.024	9.4057	0.742
		14	-0.055	-0.049	10.318	0.739
		15	0.027	0.028	10.545	0.784
		16	-0.005	-0.020	10.553	0.836
		17	0.096	0.082	13.369	0.711
		18	0.011	-0.002	13.405	0.767
		19	0.041	0.040	13.929	0.788
		20	0.046	0.061	14.569	0.801
		21	-0.096	-0.079	17.402	0.686
		22	0.039	0.077	17.875	0.713
		23	-0.113	-0.114	21.824	0.531
		24	-0.136	-0.125	27.622	0.276

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.001	0.001	0.0004	0.985
		2	0.003	0.003	0.0029	0.999
		3	0.006	0.006	0.0119	1.000
		4	0.023	0.023	0.1650	0.997
		5	-0.013	-0.013	0.2108	0.999
		6	0.022	0.021	0.3463	0.999
		7	0.038	0.038	0.7646	0.998
		8	-0.048	-0.048	1.4362	0.994
		9	0.056	0.056	2.3528	0.985
		10	-0.114	-0.116	6.1868	0.799
		11	-0.038	-0.038	6.6096	0.830
		12	-0.030	-0.028	6.8763	0.866
		13	0.192	0.193	17.947	0.160
		14	0.014	0.021	18.010	0.206
		15	0.063	0.067	19.199	0.205
		16	-0.006	-0.015	19.208	0.258
		17	-0.039	-0.035	19.664	0.292
		18	-0.029	-0.043	19.927	0.337
		19	-0.010	-0.009	19.959	0.397
		20	0.010	-0.014	19.993	0.458
		21	-0.057	-0.047	21.003	0.459
		22	0.045	0.018	21.644	0.481
		23	-0.038	0.011	22.088	0.515
		24	-0.149	-0.141	29.064	0.218

These last ones correspond to Figures 11.7 and 11.9 and to Tables 11.5 and 11.7.

Figures 11.6 and 11.8 are:





To estimate both equations of the VAR(4) at the same time we need to type the following command:

```
var bookfigure.ls 1 4 starts comps
```

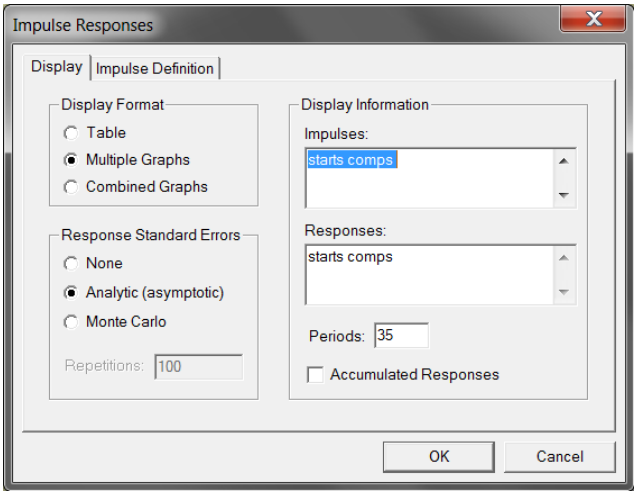
This estimates both equations and stores the VAR(4) in the workfile under the name "bookfigure." The output is the following:

Vector Autoregression Estimates
Date: 11/22/10 Time: 21:54
Sample (adjusted): 1968M05 1991M12
Included observations: 284 after adjustments
Standard errors in ( ) & t-statistics in [ ]

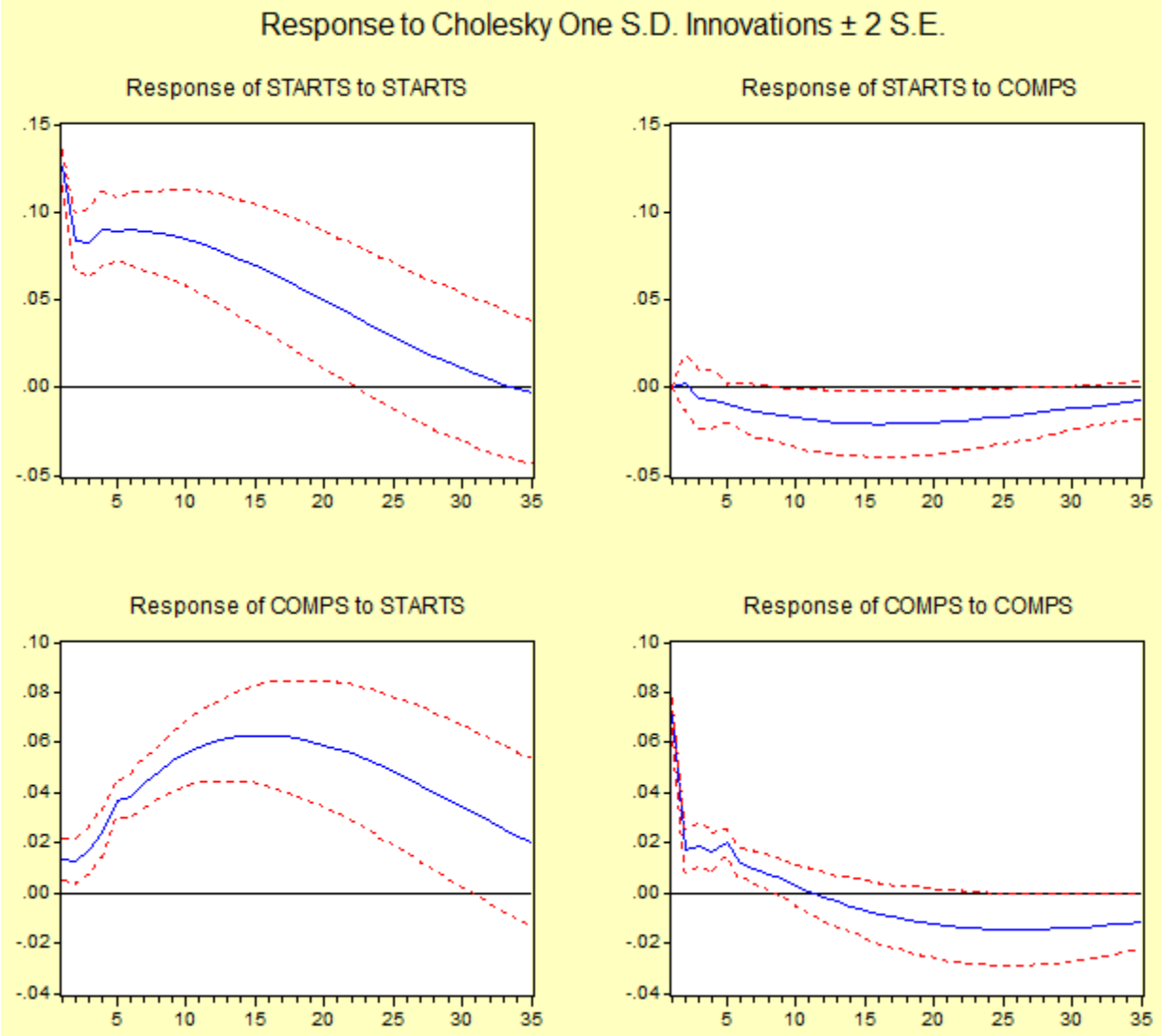
	STARTS	COMPS
STARTS(-1)	0.659939	0.074724
	(0.06124)	(0.03571)
	[10.7759]	[2.09246]
STARTS(-2)	0.229632	0.040047
	(0.07272)	(0.04241)
	[3.15759]	[0.94438]
STARTS(-3)	0.142859	0.047145
	(0.07265)	(0.04237)
	[1.96628]	[1.11280]
STARTS(-4)	0.007806	0.082331
	(0.06603)	(0.03850)
	[0.11822]	[2.13824]
COMPS(-1)	0.031611	0.236774
	(0.10271)	(0.05989)
	[0.30776]	[3.95331]
COMPS(-2)	-0.120781	0.206172
	(0.10385)	(0.06055)
	[-1.16307]	[3.40474]
COMPS(-3)	-0.020601	0.120998
	(0.10095)	(0.05886)
	[-0.20408]	[2.05559]
COMPS(-4)	-0.027404	0.156729
	(0.09457)	(0.05514)
	[-0.28978]	[2.84216]
C	0.146871	0.045347
	(0.04423)	(0.02579)
	[3.32026]	[1.75804]
R-squared	0.895566	0.936835
Adj. R-squared	0.892528	0.934998
Sum sq. resids	4.320952	1.469205
S.E. equation	0.125350	0.073093
F-statistic	294.7796	509.8375
Log likelihood	191.3622	344.5453
Akaike AIC	-1.284241	-2.362995
Schwarz SC	-1.168605	-2.247359
Mean dependent	1.574771	1.547958
S.D. dependent	0.382362	0.286689

Determinant resid covariance (dof adj.)	8.11E-05
Determinant resid covariance	7.61E-05
Log likelihood	540.7183
Akaike information criterion	-3.681115
Schwarz criterion	-3.449842

Notice that this is exactly the same result we obtained before. The benefit from this second approach is that the impulse-response functions can then be easily estimated by going to "View" and then "Impulse Response" to obtain:



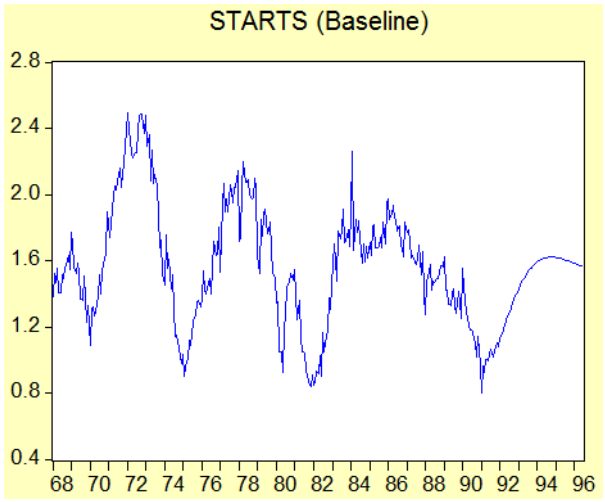
The corresponding impulse-response are:



For forecasting using the estimated VAR(4) and replicate the results in Figure 11.12 we use the following:

```
bookfigure.makemodel(varmod) @prefix s_

smpl 1992m01 1996m06

varmod.solveopt(s=d, d=d)

solve varmod

smpl 1968m01 1996m06

varmod.makegraph(g=v) finalfigure starts
```

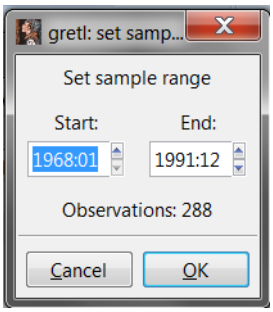
To obtain:



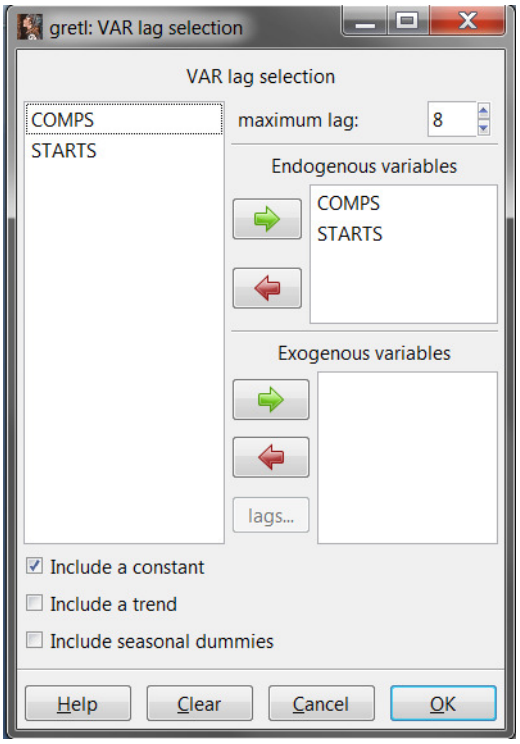
Alternatively, one can use the VAR(4) and obtain forecasts equation by equation using the same tools described in previous handouts. Just go to "Forecast" right after the estimation of each of the VAR equations.

gretl

To estimate the VAR(4) using gretl, first open the file from the class website. Then, make sure you select the correct sample for the estimation of the VAR. That is, go to 'Sample' and select 'Set range' to type:



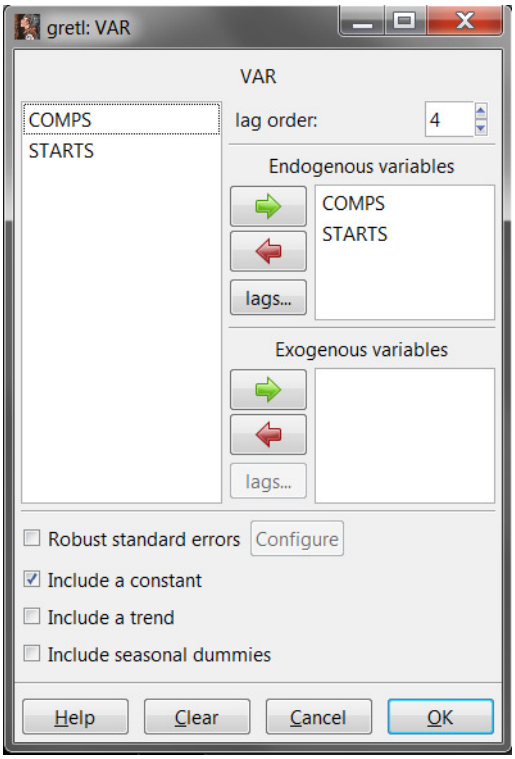
For the lag selection gretl has an automatic procedure that estimates the VAR(p) for various values of p. To do this just go to 'Model' then 'Time series' and then select 'VAR lag selection.' You will then have to choose the following options:



To obtain:

```
VAR system, maximum lag order 8

The asterisks below indicate the best (that is, minimized) values
of the respective information criteria, AIC = Akaike criterion,
BIC = Schwarz Bayesian criterion and HQC = Hannan-Quinn criterion.

lags        loglik    p(LR)       AIC          BIC          HQC

   1     469.76173             -3.312584    -3.234695    -3.281343
   2     514.50003  0.00000    -3.603572    -3.473758    -3.551503
   3     526.47574  0.00008    -3.660541    -3.478802*   -3.587645
   4     533.68942  0.00605    -3.683496*   -3.449831    -3.589772*
   5     534.51241  0.80051    -3.660803    -3.375212    -3.546252
   6     537.26899  0.23857    -3.651921    -3.314405    -3.516543
   7     538.60850  0.61289    -3.632918    -3.243476    -3.476712
   8     540.62332  0.40201    -3.618738    -3.177371    -3.441705
```

Notice that gretl already shows with a * the minimum AIC and BIC. To be consistent with the results obtained with EViews and the textbook we follow the AIC and estimate the VAR(4) by going to 'Model' then 'Time series' and finally selecting 'Vector Autoregression' to select the following options:

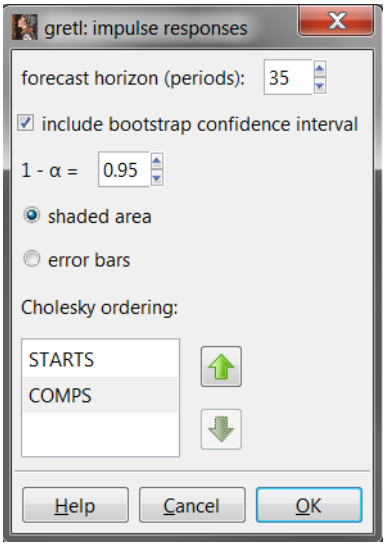


and obtain the following VAR:

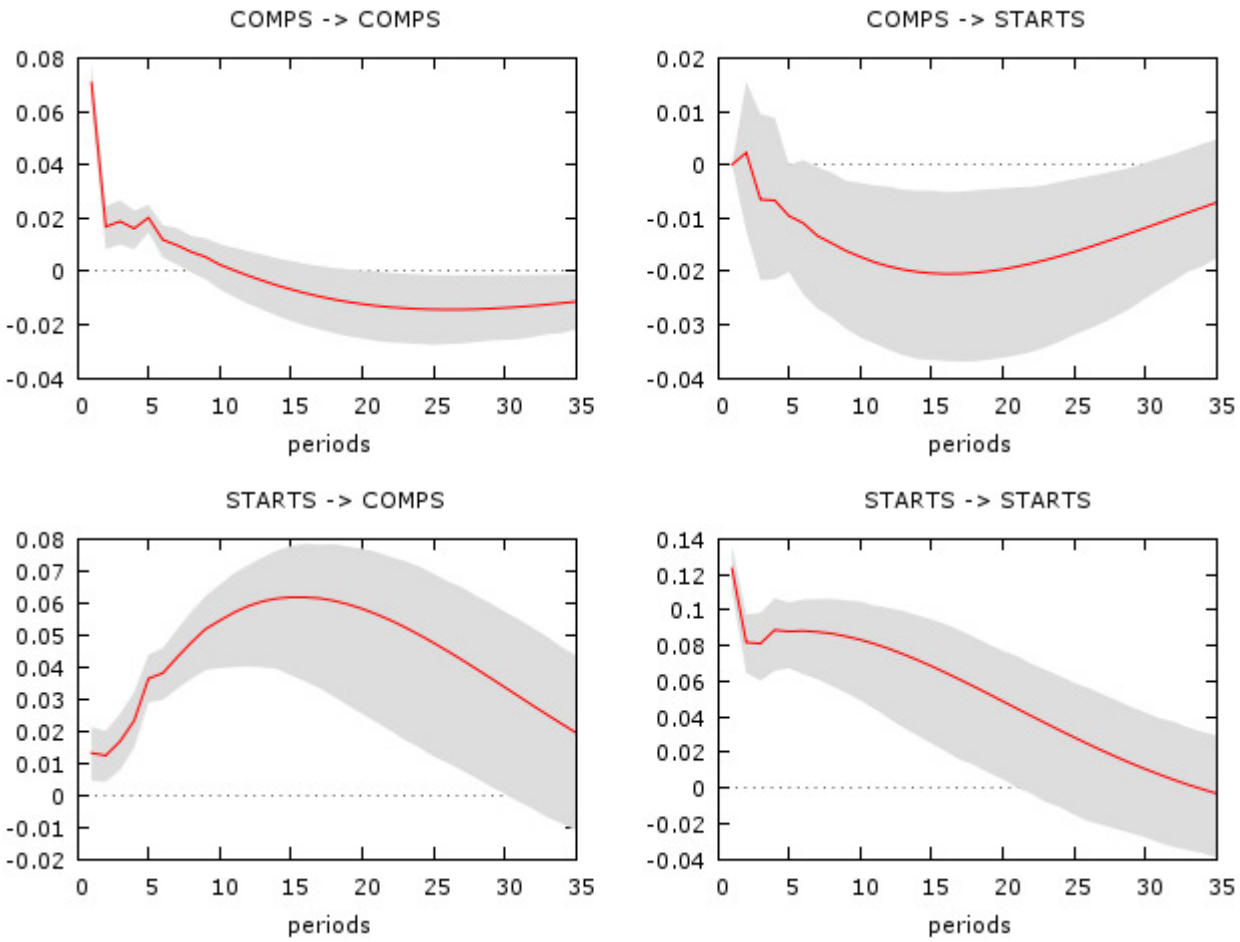
```
VAR system, lag order 4
OLS estimates, observations 1968:05-1991:12 (T = 284)
Log-likelihood = 540.71827
Determinant of covariance matrix = 7.6087294e-005
AIC = -3.6811
BIC = -3.4498
HQC = -3.5884
Portmanteau test: LB(48) = 204.246, df = 176 [0.0712]

Equation 1: COMPS

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       0.0453465    0.0257937    1.758     0.0799   *
  COMPS_1     0.236774     0.0598926    3.953     9.81e-05 ***
  COMPS_2     0.206172     0.0605545    3.405     0.0008   ***
  COMPS_3     0.120998     0.0588627    2.056     0.0408   **
  COMPS_4     0.156729     0.0551442    2.842     0.0048   ***
  STARTS_1    0.0747240    0.0357111    2.092     0.0373   **
  STARTS_2    0.0400474    0.0424062    0.9444    0.3458
  STARTS_3    0.0471448    0.0423657    1.113     0.2668
  STARTS_4    0.0823313    0.0385043    2.138     0.0334   **

Mean dependent var   1.547958   S.D. dependent var   0.286689
Sum squared resid    1.469205   S.E. of regression   0.073093
R-squared            0.936835   Adjusted R-squared   0.934998
F(8, 275)            509.8375   P-value(F)           4.3e-160
rho                 -0.009141   Durbin-Watson        2.013370

F-tests of zero restrictions:

All lags of COMPS           F(4, 275) =   149.10 [0.0000]
All lags of STARTS          F(4, 275) =   26.266 [0.0000]
All vars, lag 4             F(2, 275) =   6.9352 [0.0012]

Equation 2: STARTS

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       0.146871     0.0442346     3.320    0.0010    ***
  COMPS_1     0.0316106    0.102712      0.3078   0.7585
  COMPS_2    -0.120781     0.103847     -1.163    0.2458
  COMPS_3    -0.0206008    0.100946     -0.2041   0.8384
  COMPS_4    -0.0274041    0.0945689    -0.2898   0.7722
  STARTS_1    0.659939     0.0612423    10.78     8.08e-023 ***
  STARTS_2    0.229632     0.0727240     3.158    0.0018    ***
  STARTS_3    0.142859     0.0726546     1.966    0.0503    *
  STARTS_4    0.00780619   0.0660325     0.1182   0.9060

Mean dependent var   1.574771   S.D. dependent var   0.382362
Sum squared resid    4.320952   S.E. of regression   0.125350
R-squared            0.895566   Adjusted R-squared   0.892528
F(8, 275)            294.7796   P-value(F)           4.0e-130
rho                  0.001120   Durbin-Watson        1.991908

F-tests of zero restrictions:

All lags of COMPS           F(4, 275) =   2.2388 [0.0651]
All lags of STARTS          F(4, 275) =   203.90 [0.0000]
All vars, lag 4             F(2, 275) = 0.046240 [0.9548]

For the system as a whole:

  Null hypothesis: the longest lag is 3
  Alternative hypothesis: the longest lag is 4
  Likelihood ratio test: Chi-square(4) = 14.7643 [0.0052]

  Comparison of information criteria:
  Lag order 4: AIC = -3.68111, BIC = -3.44984, HQC = -3.58839
  Lag order 3: AIC = -3.65730, BIC = -3.47742, HQC = -3.58518
```

which yields the same results as EViews.

To obtain the Impulse-response functions, in the Vector Autoregression window select 'Graphs' and then 'Impulse responses (combined)' to get to the following menu:

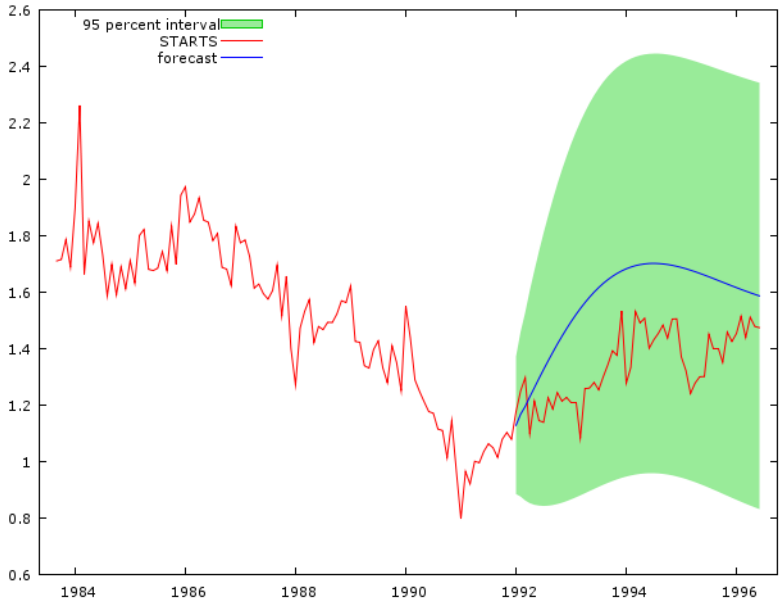


Just make sure you select the correct ordering: STARTS causes COMPS, so STARTS appears first. Once you click OK you will obtain:



which is the same as the one obtained with EViews.

For the forecasts with the VAR just o to 'Analysis' then 'Forecasts' then select 'STARTS' and the default options in the following menu to obtain:


95 percent interval
STARTS
forecast
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
1984
1986
1988
1990
1992
1994
1996