Problem Set 5 Suggested Answers

Ch 11. Exercises, Problems and Complements:

11. (p. 254) [Hint: The data for housing starts, completions is in the Diebold data sets.]

Univariate autoregessions are just single-equation AR models. If we model starts and completions as separate equations with each a function of just four lags of itself and forecast the 1992:1-1996:6 period, we get the following results:

Dependent Variable: STARTS
Method: Least Squares
Date: 11/15/07 Time: 15:23
Sample (adjusted): 1968M05 1991M12
Included observations: 284 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.063134	0.033156	1.904175	0.0579
STARTS(-1)	0.693428	0.059678	11.61950	0.0000
STARTS(-2)	0.220292	0.072355	3.044590	0.0026
STARTS(-3)	0.116594	0.072257	1.613611	0.1077
STARTS(-4)	-0.071625	0.059787	-1.198012	0.2319

R-squared	0.892165	Mean dependent var	1.574771
Adjusted R-squared	0.890619	S.D. dependent var	0.382362
S.E. of regression	0.126458	Akaike info criterion	-1.280365
Sum squared resid	4.461659	Schwarz criterion	-1.216123
Log likelihood	186.8119	Hannan-Quinn criter.	-1.254609
F-statistic	577.0717	Durbin-Watson stat	2.004557
Prob(F-statistic)	0.000000		

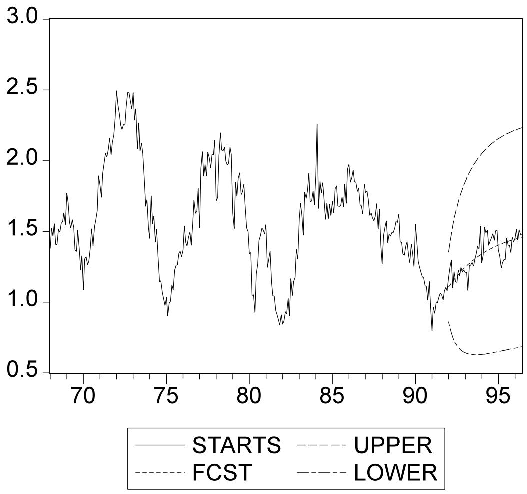
Dependent Variable: COMPS
Method: Least Squares
Date: 11/15/07 Time: 15:28
Sample (adjusted): 1968M05 1991M12
Included observations: 284 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.033288	0.028839	1.154272	0.2494
COMPS(-1)	0.550397	0.059865	9.194037	0.0000
COMPS(-2)	0.353461	0.068151	5.186420	0.0000
COMPS(-3)	0.090228	0.067797	1.330860	0.1843
COMPS(-4)	-0.016631	0.060175	-0.276380	0.7825

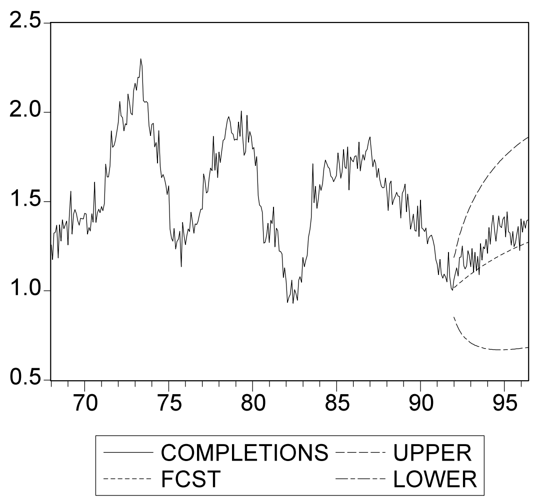
R-squared	0.912703	Mean dependent var	1.547958
Adjusted R-squared	0.911452	S.D. dependent var	0.286689
S.E. of regression	0.085310	Akaike info criterion	-2.067597
Sum squared resid	2.030513	Schwarz criterion	-2.003355
Log likelihood	298.5988	Hannan-Quinn criter.	-2.041841
F-statistic	729.2499	Durbin-Watson stat	2.006328
Prob(F-statistic)	0.000000		

Note that I just used the same four period lag length as in the textbook's VAR. If you use AIC and SIC to chose the lag length it selects a lag length of 2 for both starts and comps. In both cases, these appear to be acceptable models, with a relatively good fit and little remaining residual serial correlation.

Standalone Model of Starts w/o Completions Lags



Standalone Model of Completions w/o Lagged Starts



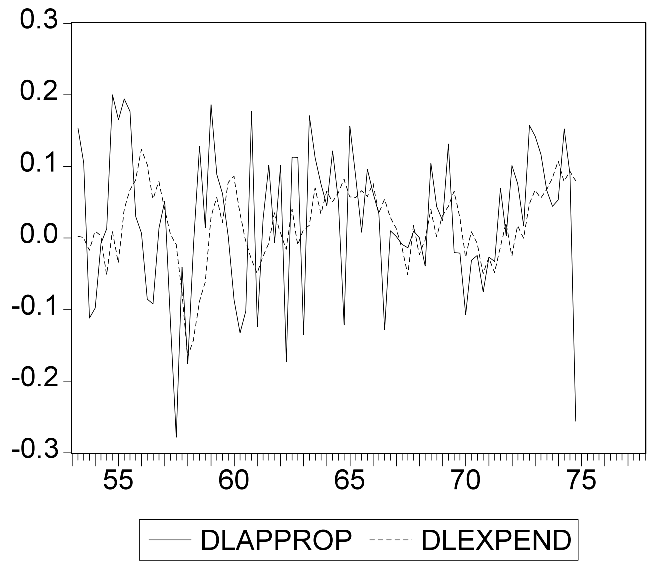
We might expect the comps VAR to outperform the univariate autoregression because the VAR analysis and Granger causality tests (see the book) indicate that starts help to predict completions but not vice versa.

But notice that the VAR forecasts (see the book diagrams) are clearly NOT better than the forecasts from either univariate model, at least in terms of forecasting the short-run 1992:1-1996:6 period. (You could calculate the MSE of each forecast to show this formally.) But the VAR does have the advantage of incorporating economically-meaningful information about the relationship between the two series, which may help us to construct <u>consistent</u> forecasts of the two series. This also may help us to communicate our results to forecast consumers.

Extra application problem:

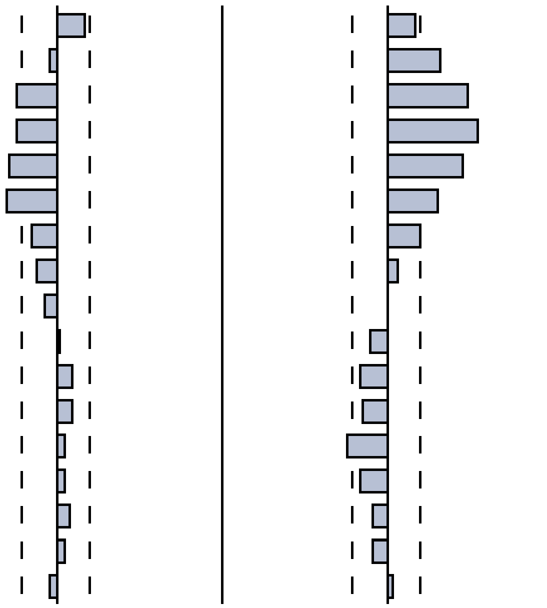
Under Supplements, you will find a data file, capital.dat, with data for quarterly seasonally-adjusted capital expenditures and appropriations for U.S. manufacturing firms between 1953 and 1974.

a. *Because the raw data are non-stationary in levels, we will work with then in differenced logarithms (remember these are then approximately growth rates). Use series commands and the dlog() function to create the appropriate transformed series.* The commands are "series dlapprop = dlog(approp)" and "series dlexpend = dlog(expend)" to create the differenced log series. Here is a graph:



b. *Create a cross correlogram for the two differenced log series [Hint: create a "group" in EViews consisting of the two differenced log series. Open the group and from the View menu choose "cross correlation"] What do the graphs tell us?*

Date: 04/15/10 Time: 20:51
Sample: 1953Q1 1977Q4
Included observations: 87
Correlations are asymptotically consistent approximations



DLAPPROP,DLEXPEND(-i)	DLAPPROP,DLEXPEND(+i)	i	lag	lead
		0	0.1866	0.1866
		1	-0.0421	0.3363
		2	-0.2527	0.5097
		3	-0.2377	0.5742
		4	-0.2948	0.4738
		5	-0.3047	0.3225
		6	-0.1523	0.2057
		7	-0.1149	0.0663
		8	-0.0701	0.0055
		9	0.0262	-0.0990
		10	0.0995	-0.1688
		11	0.1009	-0.1476
		12	0.0564	-0.2470
		13	0.0604	-0.1631
		14	0.0894	-0.0963
		15	0.0605	-0.0969
		16	-0.0395	0.0460

The first cross correlogram related current growth of capital appropriations to past values of expenditures. The negative correlations suggest that past growth in actual expenditures reduces appropriations (firms have built up enough capital so they can cut back on new appropriations.) The second cross correlogram shows that recent expenditure growth tends to be positively related to subsequent expenditure, which also makes good intuitive sense.

c. *Conduct Granger causality tests for the two differenced log series [Hint: open your group and under View choose "Granger causality"] What do the tests indicate? Are they sensitive to the number of lags included?*

Here are Granger causality test using different lag lengths:

Pairwise Granger Causality Tests
Date: 05/04/03 Time: 20:43
Sample: 1953:1 1977:4
Lags: 8

Null Hypothesis:	Obs	F-Statistic	Probability
DLEXPEND does not Granger Cause DLAPPROP	79	1.35881	0.23243
DLAPPROP does not Granger Cause DLEXPEND		7.27730	9.1E-07

Pairwise Granger Causality Tests
Date: 05/04/03 Time: 20:44
Sample: 1953:1 1977:4
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Probability
DLEXPEND does not Granger Cause DLAPPROP	83	2.10616	0.08852
DLAPPROP does not Granger Cause DLEXPEND		9.91598	1.8E-06

Pairwise Granger Causality Tests
Date: 05/04/03 Time: 20:44
Sample: 1953:1 1977:4
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
DLEXPEND does not Granger Cause DLAPPROP	85	4.90054	0.00983
DLAPPROP does not Granger Cause DLEXPEND		13.4123	9.5E-06

All three models can reject the null that dlapprop does not Granger cause dlexpend, in other words providing support for our casual reading of the cross correlograms that changes in appropriations tend to “cause” increase expenditure in subsequent quarters. The evidence going the other direction is mixed and depends on the number of lags of each variable included.

d. *Set up a VAR model for the two differenced log series. [Hint: Under Quick, choose "Estimate VAR"] Select an appropriate lag order for your VAR by estimating a range of VARS with 1 to 8 lags and tabulating AIC and SIC values. [You set the lag order in the "lag intervals for endog" box. Entering "1 4" would indicate a 4th order VAR. You want the summary AIC and SIC values at the very bottom of the regression output. Set your estimation sample for 1955:2 1974:4 to allow for lags.] Comment on the economic meaning of the coefficient estimates for your selected VAR.*

Number of lags

1	
Akaike Information Criteria	-5.795837
Schwarz Criteria	-5.615879
2	
Akaike Information Criteria	-6.073104
Schwarz Criteria	-5.773173
3	
Akaike Information Criteria	-6.199632
Schwarz Criteria	**-5.779729**
4	
Akaike Information Criteria	-6.183406
Schwarz Criteria	-5.643532
5	
Akaike Information Criteria	-6.115524
Schwarz Criteria	-5.455678
6	
Akaike Information Criteria	-6.104796
Schwarz Criteria	-5.324978
7	
Akaike Information Criteria	-6.243359
Schwarz Criteria	-5.343569
8	
Akaike Information Criteria	**-6.291160**
Schwarz Criteria	-5.271398

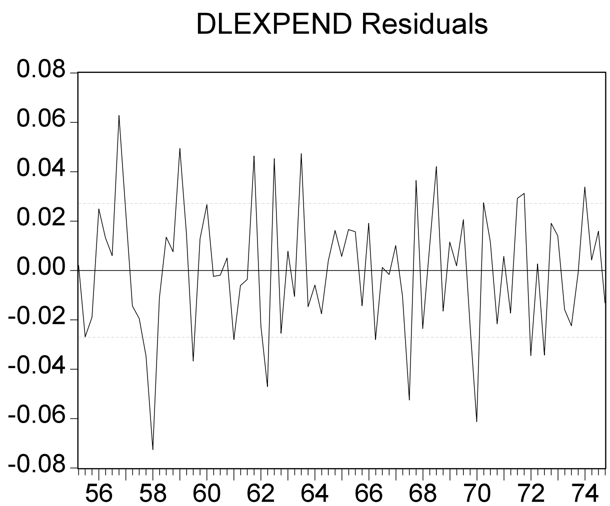
The AIC criteria would select the largest model considered—one with eight lags of both variables. SIC picks a much more parsimonious three lag VAR. (Note that I set the same beginning point for estimation for each model, 1955:2 for comparability with the largest models.) We will opt for the smaller model as long as other diagnostics look OK, in particular serial correlation properties of residuals.

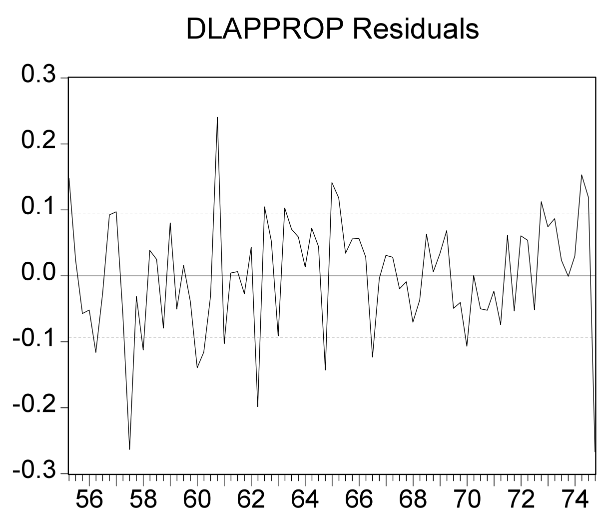
Here is the selected model:

Date: 05/04/03 Time: 21:22
Sample: 1955:2 1974:4
Included observations: 79
Standard errors & t-statistics in parentheses

	DLEXPEND	DLAPPROP
DLEXPEND(-1)	0.351602	0.272144
	(0.11001)	(0.38063)
	(3.19597)	(0.71498)
DLEXPEND(-2)	0.053857	-0.925762
	(0.11220)	(0.38821)
	(0.47999)	(-2.38470)
DLEXPEND(-3)	0.112617	-0.030969
	(0.10199)	(0.35285)
	(1.10424)	(-0.08777)
DLAPPROP(-1)	0.118393	0.070463
	(0.03738)	(0.12935)
	(3.16686)	(0.54477)
DLAPPROP(-2)	0.160383	0.129358
	(0.03496)	(0.12097)
	(4.58706)	(1.06934)
DLAPPROP(-3)	0.151581	0.275877
	(0.03736)	(0.12924)
	(4.05779)	(2.13454)
C	0.001190	0.021223
	(0.00347)	(0.01200)
	(0.34319)	(1.76888)
R-squared	0.764795	0.215315
Adj. R-squared	0.745194	0.149925
Sum sq. resids	0.052908	0.633332
S.E. equation	0.027108	0.093788
F-statistic	39.01926	3.292765
Log likelihood	176.5954	78.53912
Akaike AIC	-4.293555	-1.811117
Schwarz SC	-4.083603	-1.601166
Mean dependent	0.022804	0.019636
S.D. dependent	0.053702	0.101723
Determinant Residual Covariance		4.88E-06
Log Likelihood		258.8854
Akaike Information Criteria		-6.199632
Schwarz Criteria		-5.779729

The residuals have some large outliers, but otherwise “look like” white noise:





We choose “Proc/make residuals” to save the residuals as time series for a closer look. Here are the corellograms and partial autocorrelation functions, with Q-Statistics for both residuals:

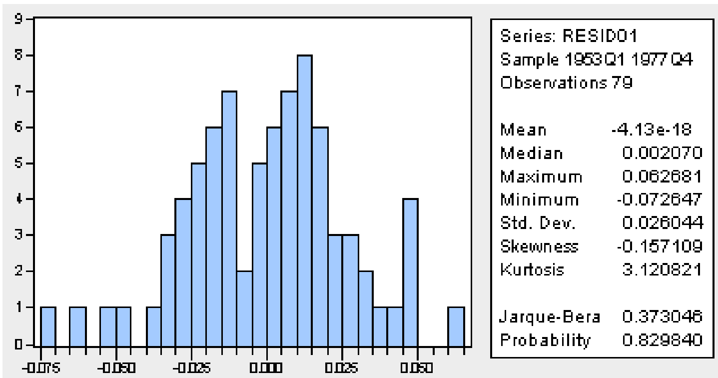
Date: 04/15/10 Time: 20:42
Sample: 1953Q1 1977Q4
Included observations: 79

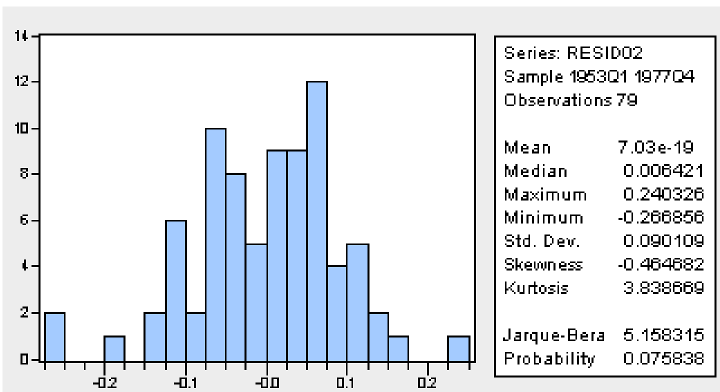
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.072	-0.072	0.4247	0.515
		2	-0.117	-0.123	1.5707	0.456
		3	-0.014	-0.033	1.5874	0.662
		4	-0.190	-0.213	4.6648	0.323
		5	-0.076	-0.125	5.1654	0.396
		6	-0.171	-0.268	7.7385	0.258
		7	0.042	-0.070	7.8943	0.342
		8	-0.073	-0.242	8.3741	0.398
		9	-0.013	-0.173	8.3898	0.495
		10	0.214	-0.004	12.618	0.246
		11	-0.029	-0.143	12.695	0.314
		12	-0.014	-0.168	12.714	0.390
		13	0.056	-0.091	13.015	0.447
		14	0.070	-0.003	13.495	0.488
		15	0.013	-0.017	13.511	0.563
		16	-0.123	-0.123	15.058	0.520

Date: 04/15/10 Time: 20:43
Sample: 1953Q1 1977Q4
Included observations: 79

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.005	-0.005	0.0023	0.962
		2	-0.093	-0.093	0.7220	0.697
		3	-0.002	-0.003	0.7223	0.868
		4	-0.081	-0.090	1.2770	0.865
		5	0.166	0.167	3.6476	0.601
		6	-0.116	-0.139	4.8182	0.567
		7	0.253	0.311	10.500	0.162
		8	-0.084	-0.178	11.137	0.194
		9	-0.051	0.104	11.375	0.251
		10	0.125	0.003	12.831	0.233
		11	-0.132	-0.028	14.481	0.208
		12	0.054	-0.064	14.761	0.255
		13	-0.274	-0.227	22.027	0.055
		14	0.078	0.062	22.632	0.067
		15	0.069	-0.028	23.114	0.082
		16	-0.114	-0.028	24.427	0.081

While there are a few significant partial autocorrelations, the Q-statistics do not reject white noise errors. Looking at descriptive statistics, we nearly reject the null of normality for the residual from the dlapprop equations, but there is no evidence of non-normality for the dlexpend residuals:

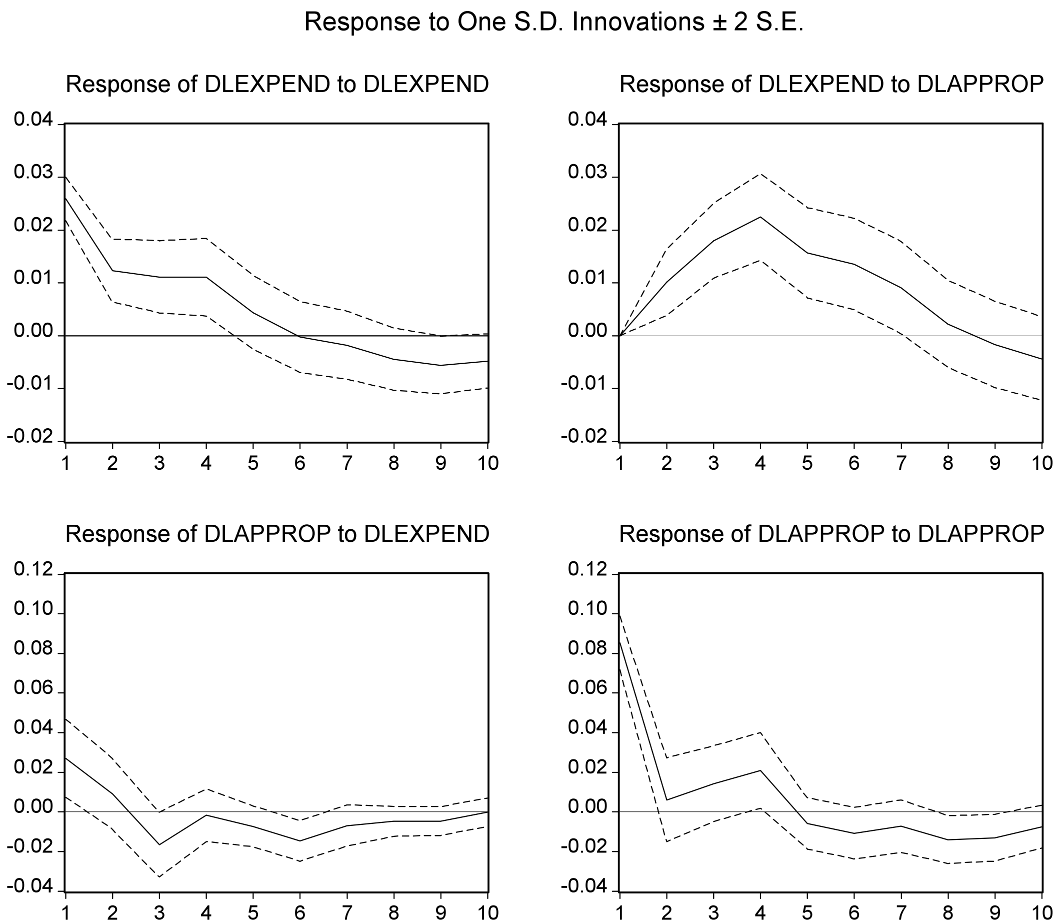




In sum, this model looks satisfactory.

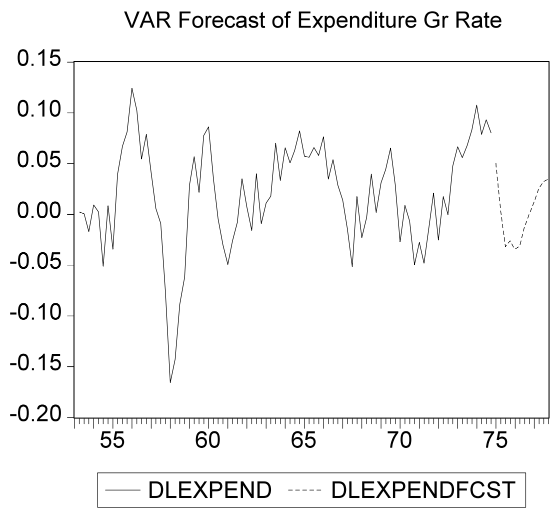
Looking at the estimation results, note that the model explains much more of the variation in expenditures than appropriations, and the standard error of the latter is correspondingly larger. Looking at the coefficients, we see that lagged growth in appropriations tends to raise the growth rate of expenditure, as expected, and that lagged expenditure tends to reduce growth in appropriations, particularly after two quarters. There is more “inertia” in expenditure, evidenced by the large and significant lagged dependent variable.

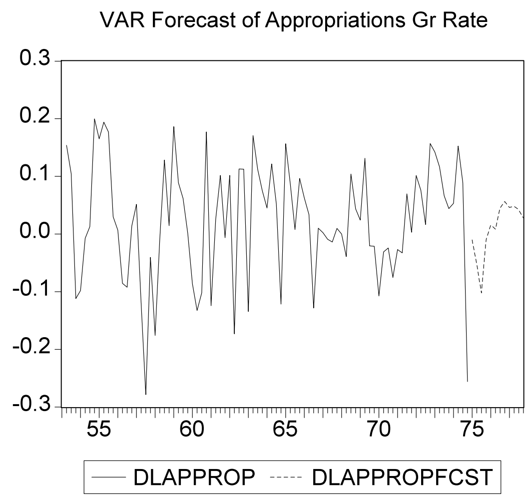
e. *Generate impulse responses [Hint: click "Impulse" button] and interpret.*



Again, the impulse response analysis—which shows the implied response of each variable to a one-standard-deviation positive shock to itself and the other variable—confirm the economic relationships discussed above. Focusing on the important upper-right and lower-left quadrants, we see that growth in appropriations causes a build up of expenditure that peaks after four quarters, and an increase in expenditure leads to a small decline in appropriations after three periods (although this is not significantly less than zero, except at lag 6, judging from the 95% confidence interval). Depending on the number of lags and the sample you used, your graphs may look a bit different.

f. *Produce a forecast for the period 1975:1 - 1977:4 for the two differenced log series. [Hint: you will need to set up a model from the two equations of the VAR. Select "Make Model" from the "Proc" menu in the VAR window.]*





Notice that the model predicts negative growth rates for both expenditures and appropriations in the near term before both turn positive again. This is presumably because the recent surge in expenditure reduces appropriations and then this lead to less expenditure in subsequent quarters.

If we want to see forecasts of the log levels, we can add identities to the model specification to calculate them (e.g. lexpend = lexpend(-1)+dlexpend). Here are what they look like: