Business and Economics Forecasting
Econon 3342

Spring, 2019
Diego Escobari

Assignment 4 - Suggested Solutions

- Due Monday December 9 (at 3:30 pm).
- You can work in groups of up to three students.
- Send your PDF responses by email and make sure you copy all members when submitting your PDF file.
- Make sure your PDF file shows your work on EViews.

In this assignment you will be working with two variables. There are many very interesting data sets online and I encourage you to use something you find unique and interesting. You are free to use any two variables you want. In case you need an example, we have the following two data sources that we used before:

1) Data from yahoo finance. (http://finance.yahoo.com/) You can use the stocks of any public company you may be interested in (there are also bonds, mutual funds, indices, cryptocurrencies). Try typing, for example, Microsoft. Once you obtain the data for that stock just go to 'Historical Data' and follow the instructions to download the period you want as a MS Excel file.

2) Data from Google trends. (http://www.google.com/trends) Just type any keyword you think is interesting.

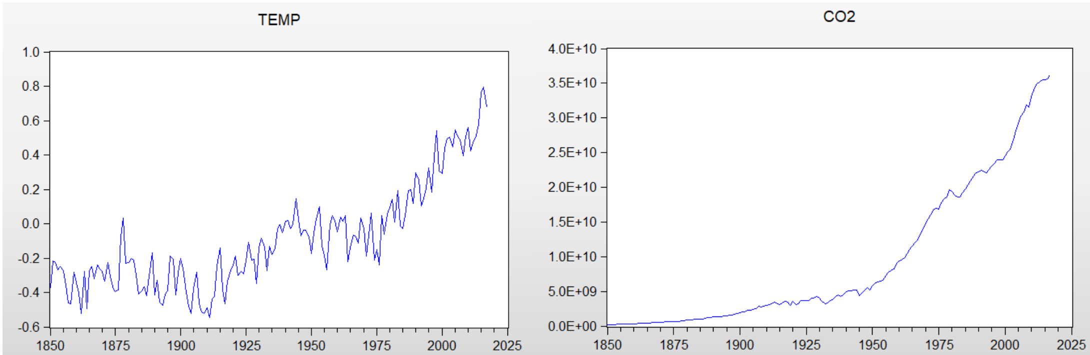
For example, you can match weekly data from Yahoo Finance and Google or use both series from the same data source. You need to obtain at least two series.

1. Describe both of your variables. Why are they interesting?

For this project I decided to assess whether CO2 emissions have an effect on global temperatures. Hence, I obtained the data from https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions This is yearly time series data from 1850 until 2017. We have two variables: annual CO2 emissions (measured in tons) and annual average temperatures measured in Celsius and relative to 1961-90 average.

These two variables are interesting due to the dynamics and the potential effect that CO2 might have on global temperatures.

2. Obtain the time series graph of both of the series. Any insights?



From the times series graphs of both of these variables are trending upwards. This means that we would need to control for a trend in our model. Alternatively, we can just work with the growth rates of each of the variables.

3. Obtain the cross-correlations and interpret.

Sample: 1850 2025
Included observations: 168
Correlations are asymptotically consistent approximations

TEMP,CO2(-i)	TEMP,CO2(+i)	i	lag	lead
		0	0.8903	0.8903
		1	0.8708	0.8652
		2	0.8529	0.8339
		3	0.8340	0.8037
		4	0.8134	0.7817
		5	0.7932	0.7634
		6	0.7730	0.7481
		7	0.7523	0.7345
		8	0.7306	0.7122
		9	0.7106	0.6922
		10	0.6923	0.6776
		11	0.6724	0.6559
		12	0.6507	0.6331
		13	0.6290	0.6087
		14	0.6130	0.5882
		15	0.5940	0.5646
		16	0.5804	0.5417
		17	0.5682	0.5199
		18	0.5549	0.5047
		19	0.5420	0.4884
		20	0.5279	0.4603

Sample: 1850 2025
Included observations: 167
Correlations are asymptotically consistent approximations

RGCO2,RGTEMP(-i)	RGCO2,RGTEMP(+i)	i	lag	lead
		0	0.0376	0.0376
		1	0.2345	-0.0498
		2	-0.0888	0.0100
		3	-0.1100	-0.0009
		4	-0.0574	-0.0673
		5	0.0799	-0.0410
		6	-0.1425	0.0962
		7	-0.0262	-0.0942
		8	0.0274	-0.1873
		9	-0.0482	-0.0599
		10	-0.0554	-0.0729
		11	-0.0905	-0.0456
		12	-0.0813	0.1911
		13	-0.0018	0.2333
		14	0.0016	0.1185
		15	-0.0928	-0.0233
		16	-0.0240	0.1182
		17	0.0590	-0.0577
		18	-0.0220	-0.0425
		19	-0.0378	0.0698
		20	0.0023	0.0377

We present two cross-correlations. The one on the left-hand side is the cross-correlations for the variables in levels (as presented in the figures on question 3). The cross-correlogram on the right-hand side is for the growth rates of the variables. Both cross-correlograms appear to show some ling between CO2 and temperatures, and of course, the link is much stronger for the variables in levels as both trend upwards.

4. Estimate the most appropriate VAR (selected AIC or BIC).

	VAR(1)	VAR(2)	VAR(3)	VAR(4)	VAR(5)	VAR(6)	VAR(7)
AIC	2.1555	2.1826	2.2228	2.2068	2.2261	2.2718	2.2592
BIC	2.2679	2.3709	2.4874	2.5485	2.6454	2.7694	2.8358

The table above shows the summary of the estimation of various VAR models, from VAR(1) to VAR(7). Both, the AIC and the BIC select the same model, the VAR(1). The regression output for the estimation of these VAR(1) is as follows:

Vector Autoregression Estimates
Sample (adjusted): 1852 2017
Included observations: 166 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	RGCO2	RGTEMP
RGCO2(-1)	-0.020718	-3.563701
	(0.07616)	(5.58064)
	[-0.27203]	[-0.63858]
RGTEMP(-1)	0.003302	0.003912
	(0.00107)	(0.07828)
	[3.09038]	[0.04997]
C	0.034721	-0.181370
	(0.00452)	(0.33087)
	[7.68942]	[-0.54817]
R-squared	0.055481	0.002500
Adj. R-squared	0.043892	-0.009740
Sum sq. resids	0.376132	2019.558
S.E. equation	0.048037	3.519933
F-statistic	4.787330	0.204229
Log likelihood	269.9098	-442.9314
Akaike AIC	-3.215781	5.372668
Schwarz SC	-3.159540	5.428909
Mean dependent	0.033040	-0.300210
S.D. dependent	0.049127	3.502916
Determinant resid covariance (dof adj.)		0.028551
Determinant resid covariance		0.027528
Log likelihood		-172.9072
Akaike information criterion		2.155508
Schwarz criterion		2.267990
Number of coefficients		6

For example, the temperature equation is:

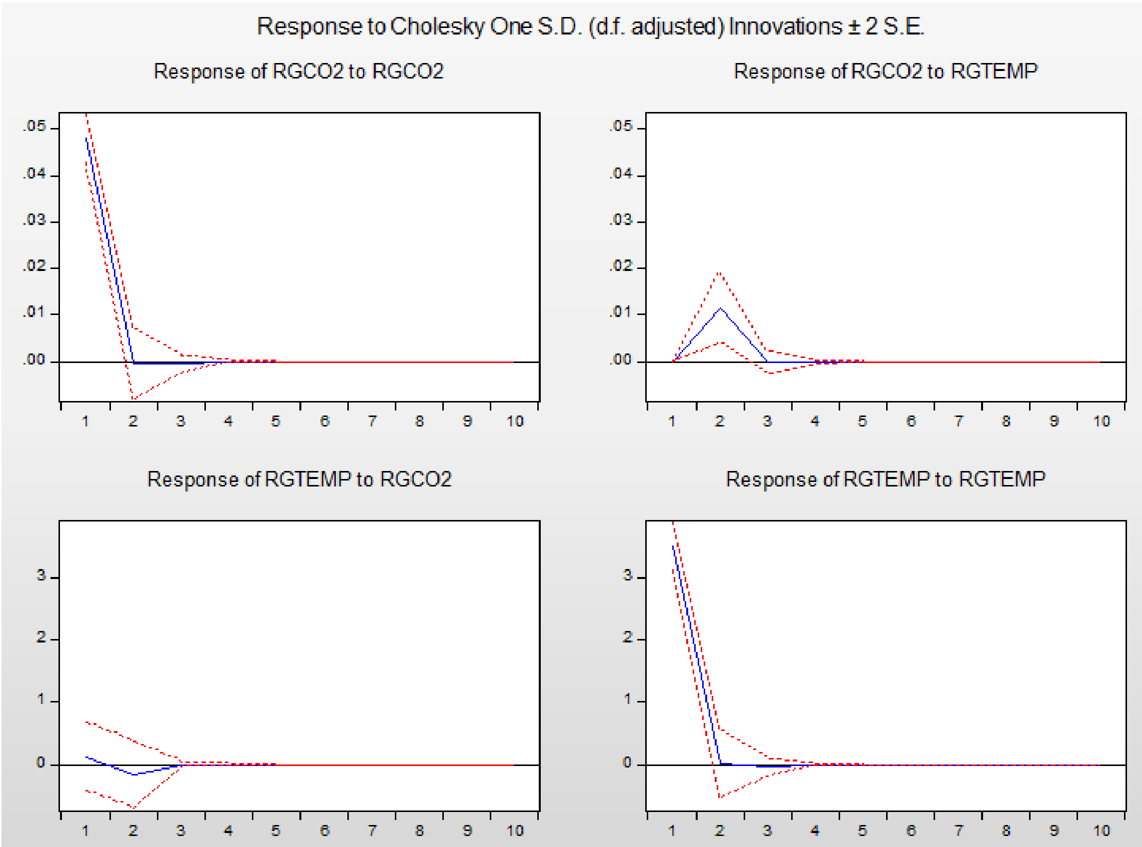
$$RGTEMP_t = -0.181 - 3.564RGCO2_{t-1} + 0.00391RGTEMP_{t-1}$$

Where RGTEMP is the rate of growth of TEMP and is given by:

$$RGTEMP_t = (TEMP_t - TEMP_{t-1})/TEMP_t$$

5. Obtain the impulse response functions. Interpret your findings for at least two of the quadrants.

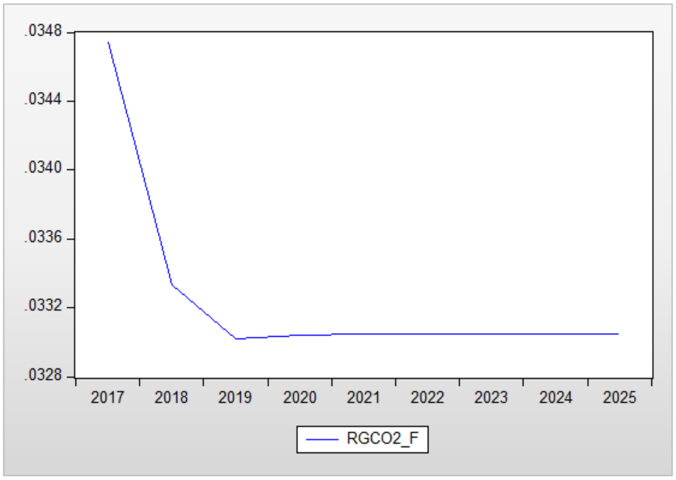
The impulse response functions are presented below:



From the lower left quadrant, we see that the response of the RGTEMP to a shock in RGCO2 is no statistically significant effect. This is because zero is always between the lower and upper bands (the +/- 2 standard error red bands).

From the upped left quadrant we can see that the effect of RGCO2 on RGCO2 is positive and statistically significant only for the first year. The effect quickly drops to zero.

6. Obtain the out-of-sample forecast for one of your variables. Interpret.



The out-of-sample forecast for the yearly growth rate of CO2 emissions show that it is expected to be on average about 0.033 between 2019 and 2025. This means that CO2 emissions are expected to growth at a rate of 3.3% per year between 2019 and 2025.

7. Estimate only one of your VAR equations via OLS. Are the error terms White Noise?

When estimating only one RGTEMP equation we obtain:

Dependent Variable: RGTEMP
Method: Least Squares
Sample (adjusted): 1852 2017
Included observations: 166 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.181370	0.330867	-0.548167	0.5843
RGCO2(-1)	-3.563701	5.580638	-0.638583	0.5240
RGTEMP(-1)	0.003912	0.078284	0.049966	0.9602

R-squared	0.002500	Mean dependent var	-0.300210
Adjusted R-squared	-0.009740	S.D. dependent var	3.502916
S.E. of regression	3.519933	Akaike info criterion	5.372668
Sum squared resid	2019.558	Schwarz criterion	5.428909
Log likelihood	-442.9314	Hannan-Quinn criter.	5.395497
F-statistic	0.204229	Durbin-Watson stat	2.000087
Prob(F-statistic)	0.815484		

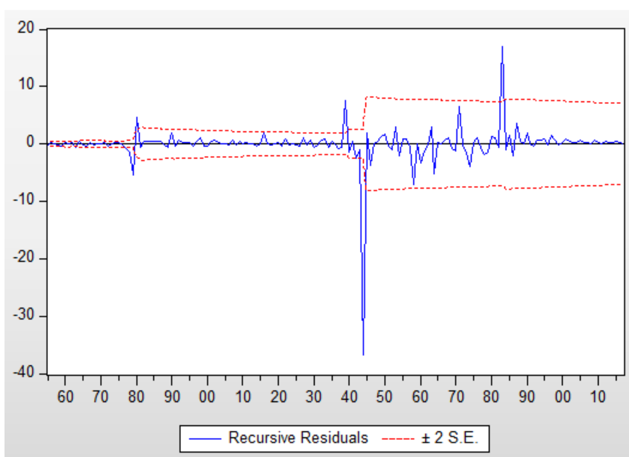
The correlogram of the regression residuals is presented below. The relatively high p-values associated with the Q-statistics fail to reject the null hypothesis of White Noise. This means that the residuals resemble a White Noise process, which validates our model.

Sample: 1850 2025
Included observations: 166

	AC	PAC	Q-Stat	Prob
1	-0.000	-0.000	6.E-07	0.999
2	0.141	0.141	3.4004	0.183
3	-0.037	-0.037	3.6294	0.304
4	0.048	0.029	4.0324	0.402
5	-0.186	-0.180	9.9951	0.075
6	0.003	-0.006	9.9972	0.125
7	-0.013	0.041	10.025	0.187
8	0.023	0.013	10.115	0.257
9	-0.099	-0.094	11.868	0.221
10	0.052	0.018	12.351	0.262
11	-0.055	-0.033	12.898	0.300
12	0.069	0.066	13.751	0.317
13	-0.024	-0.001	13.858	0.384
14	0.172	0.126	19.262	0.155
15	0.004	0.018	19.264	0.202
16	0.089	0.038	20.733	0.189
17	-0.000	0.023	20.733	0.238
18	0.047	0.016	21.156	0.272
19	-0.118	-0.074	23.805	0.204
20	0.122	0.124	26.661	0.145
21	-0.022	0.024	26.754	0.179
22	0.001	-0.041	26.754	0.221

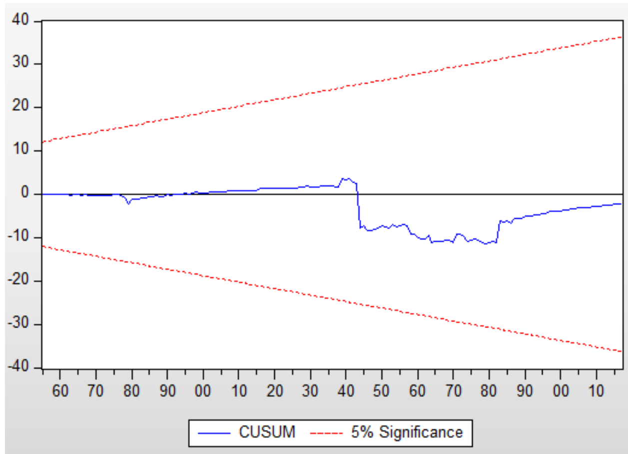
8. Assess the stability of your model:

    a. Recursive residuals, interpret.



The recursive residuals graph presented above shows evidence of a structural break around 1944. From the temperature time series graph (question 2) we see that 1944 is about the year where temperatures start to climb.

b. CUSUM, interpret.



The CUSUM test shows evidence that the model is stable. We conclude this with the observation that the CUSUM (blue line) never crosses the 5% confidence bands (red lines). This also means that there appears to be no structural breaks in the sample (1850-2019) that we are studying.

c. Recursive coefficients, interpret.

For the recursive coefficients, note that our model estimates three coefficients. That is, the constant and the coefficients on the lags of RGCO2 and RGTEMP.

From all three recursive coefficient estimates we see that our three coefficients are fairly stable. Note that around 1944 there appears to be a sudden change in the estimates (blue line), which is consistent with the recursive residuals results.