Introductory Econometrics

Tutorial 10

PART A: To be done before you attend the tutorial. The solutions will be made available at the end of the week.

1. (*Reminding ourselves of interpretation of parameters*): In the following equation, *gdp* refers to gross domestic product, *bankcredit* refers to domestic bank credit issued and FDI refers to foreign direct investment.

$$\log(gdp) = \underset{(0.13)}{2.65} + \underset{(0.022)}{0.527}\log(bankcredit) + \underset{(0.017)}{0.222}\log(FDI).$$

   Which of the following statements is then true?

   (a) If FDI increases by 1%, *gdp* increases by approximately 22.2%, the amount of bank credit remaining constant.

   (b) If FDI increases by 1%, *gdp* increases by approximately 0.22%, the amount of bank credit remaining constant.

   (c) If FDI increases by 1 dollar, *gdp* increases by approximately 22 cents, the amount of bank credit remaining constant.

   (d) If *gdp* increases by 1%, FDI increases by approximately 0.22%, the amount of bank credit remaining constant.

2. *Correlograms are useful for modelling time series*: The 4 plots on the next page show the correlograms of 4 time series. These correlograms are estimated based on 196 observations. Which of these series is likely to be white noise? For each of the series that are not white noise, determine the order of the AR model which would be suitable for that series.

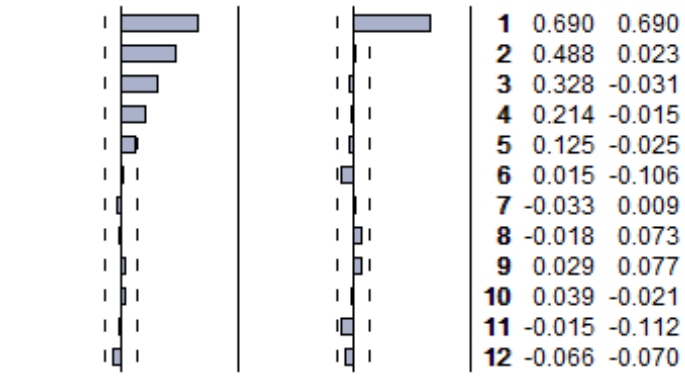
3. (*Serially correlated errors*) The "Phillips curve" in economics is a theory due to Bill Phillips, a New Zealand economist, about the trade-off between inflation and unemployment rate. According to this, lower than normal unemployment rate goes hand in hand with higher than normal inflation, and vice versa. The file Phillips.wf1 include data on the US annual inflation (*inf*) and unemployment (*unem*) rates from 1948 to 2003. We want to see if this theory is supported by the data or not.

   (a) Look at the scatter plot of inflation against the unemployment rate. Is there any visually recognisable trade-off?

   (b) Run a regression on inflation on a constant and unemployment rate. Since these are time series, we worry about serial correlation in errors. Use visual tools and the BG test to test that errors are white noise against the alternative that they are generated by an AR(1) process. If you find evidence of serial correlation, amend your model and re-estimate it.

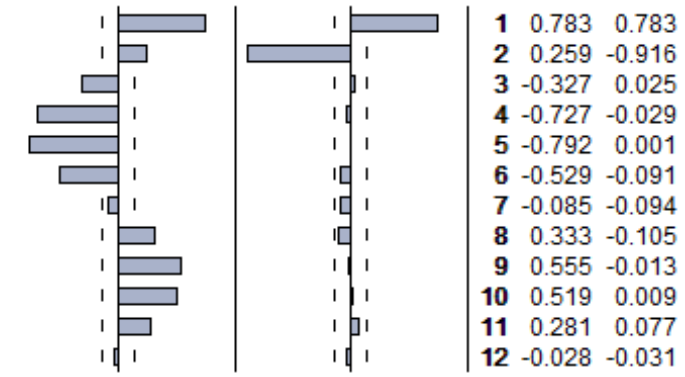
   (c) What conclusion do you obtain from your analysis about the Phillips curve hypothesis?

Correlogram of Y1
Sample: 5 200
Included observations: 196



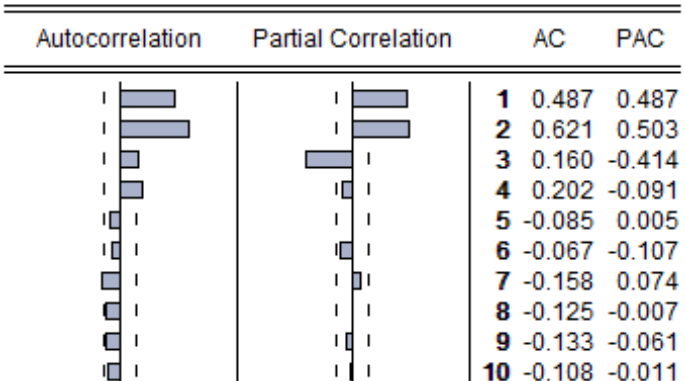
	AC	PAC
1	0.690	0.690
2	0.488	0.023
3	0.328	-0.031
4	0.214	-0.015
5	0.125	-0.025
6	0.015	-0.106
7	-0.033	0.009
8	-0.018	0.073
9	0.029	0.077
10	0.039	-0.021
11	-0.015	-0.112
12	-0.066	-0.070

Correlogram of Y2
Sample: 5 200
Included observations: 196



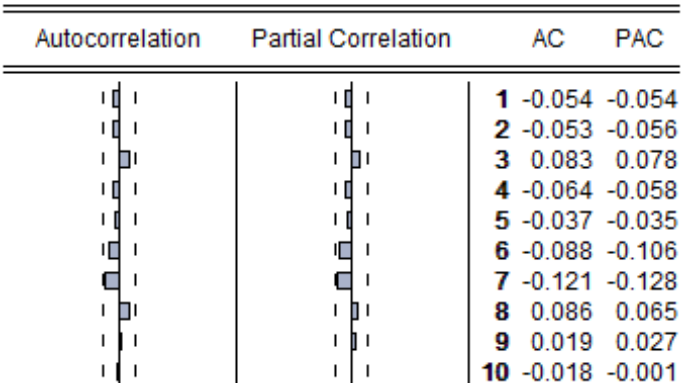
	AC	PAC
1	0.783	0.783
2	0.259	-0.916
3	-0.327	0.025
4	-0.727	-0.029
5	-0.792	0.001
6	-0.529	-0.091
7	-0.085	-0.094
8	0.333	-0.105
9	0.555	-0.013
10	0.519	0.009
11	0.281	0.077
12	-0.028	-0.031

Correlogram of Y3
Sample: 5 200
Included observations: 196



	AC	PAC
1	0.487	0.487
2	0.621	0.503
3	0.160	-0.414
4	0.202	-0.091
5	-0.085	0.005
6	-0.067	-0.107
7	-0.158	0.074
8	-0.125	-0.007
9	-0.133	-0.061
10	-0.108	-0.011
11	-0.098	-0.038
12	-0.080	-0.015

Correlogram of Y4
Sample: 5 200
Included observations: 196



	AC	PAC
1	-0.054	-0.054
2	-0.053	-0.056
3	0.083	0.078
4	-0.064	-0.058
5	-0.037	-0.035
6	-0.088	-0.106
7	-0.121	-0.128
8	0.086	0.065
9	0.019	0.027
10	-0.018	-0.001
11	-0.076	-0.115
12	-0.013	-0.044

Do not forget to bring your answers to PART A and a copy of the tutorial questions to your tutorial.

<u>Part B:</u> This part will be covered in the tutorial. It is still a good idea to attempt these questions before the tutorial.

The purpose of this tutorial is to use dummy variables when using time series data.

1. The data set SeattleElectric2005-6.xlsx contains data on average hourly electricity usage in cold months in Seattle, USA, from 1 Oct 2005 to 31 March 2006. The variables in this data set are:

date	:	date
dow	:	day of the week (1 for Sunday to 7 for Saturday)
pubhol	:	=1 if that day was a public holiday, =0 otherwise
avetemp	:	average daily temperature in degrees Fahrenheit
aveload	:	average hourly electricity load on that day in megawatts

(a) Use *dow* to create dummy variables *sat* for Saturday (day 7) and *sun* for Sunday (day 1). In addition, create a dummy variable called *wknd* which is 1 if the day is a Saturday or a Sunday, and 0 for any other day.

(b) The goal is to predict *aveload* in winter months (Seattle, USA, is in the Northern hemisphere) based on temperature and other available information (our assumption is that the weather bureau can produce temperature forecasts that are pretty accurate). Given this goal, have an appropriate look at the data. In particular, investigate differences in average load in different days of the week and look at the scatter plot of *aveload* against *avetemp* to familiarise yourself with the data. Remember that a good analyst does not confine himself or herself to any particular software. If a pivot chart would be better for visualising differences in average load in different days of the week, then use it! If you have learnt another software (for example R, SPSS or STATA) in another unit that you think will be more helpful, use it!

(c) Estimate a regression of *aveload* on a constant, *avetemp*, *wknd* and the public holiday dummy. We know that electricity usage, like anything else in life, is likely to be related to what the usage was in previous days. As a result, we suspect that part of the electricity load that cannot be explained by temperature is likely to be correlated over time. If that is the case, what consequences would that have for the OLS estimator? In such a case, can we use the model that you estimated to test that there is no difference in the intercept in weekdays and weekends?

(d) Use visual aids and a formal test to test the hypothesis that errors of this model are white noise against the alternative that they are generated by an AR(7). You should reject the null. Then investigate what kind of AR model would be sufficient for capturing the dynamics of the errors (the t-statistics in your BG auxiliary regression may give you a hint, and the partial autocorrelations of the residuals in the correlogram are informative as well). Re-estimate the model by adding an AR equation for u_t.

(e) Using the regression model with AR errors, investigate if the sensitivity of electricity load to temperature (i.e. the coefficient of temperature) is different in weekends relative to the rest of the week.

(f) Using the regression model with AR errors, investigate if having dummies for Saturday and Sunday separately instead of *wknd* would improve the model. In other words, test the hypothesis $H_0 : \beta_{SAT} = \beta_{SUN}$.