Quantitative Economic Analysis Ec504: Econometric Computer Exercise 0
Revising 2 variable OLS Estimation

The purpose of this exercise is to familiarise you with the basics of estimation and interpretation of the 2 variable regression model using a regression package. The economic issue tested in this exercise is to try to establish a relationship between the rate of inflation and the level of unemployment, (the Phillip's Curve).

Read in the Stata data set, cex0.dta, from the course website:

http://personal.rhul.ac.uk/uhte/006/ec5040/indexa.html

Find the data and then right click using the "save as" option to save the data set onto your Y drive space
And then use the "use" or "file open" commands to read the data into Stata (see Stata worksheet)

The data contain annual observations on U.S. inflation and unemployment from 1957 - 1998

I	inf	annual inflation rate (% change)
	unr	annual unemployment rate (%)
	year	year indicator

1. As a means of familiarising yourself with the data, graph the inflation and unemployment series over time. Use the Stata command

   gra inf year,c(m) xlab ylab

   What do you conclude about the likely process underlying these series? What is the average value of each series over the sample period? How do the average series values change over time?
   [Use the Stata command
   su inf unr
   to obtain mean values]

2. Now begin to test the relationship between these series more formally.

   Regress inflation on unemployment over the entire sample period. What is the estimated coefficient on unemployment? How do you interpret this estimate?
   Is it statistically significant? What is the 95% confidence interval around this estimate?
   How well does your model fit the data?

   Graph the fitted line from your regression along with the original data. (You can do this by
   1st saving the estimated values of the dependent variable from your regression using the commands
   reg inf unr

predict *variable name*

and then graph this variable alongside the true values of inflation and unemployment

gra *variable name* inf unr,c(m)

(The graph shows the scatter diagram of your data set with the estimated regression line super-imposed)

Does the model fit the data at all levels of unemployment?

Now look at the residual values from your regression over time. You can do this by 1st saving the residuals from your regression

reg inf unr
predict *residual variable name*, resid

and then graphing these against the year variable.

gra *residual variable name* year, yline(0) xlab ylab

Is the relation stable over time?

3. Suppose you rescale the independent variable so that it unemployment is measured as a fraction of the labour force ie as a *rate* (so it takes a value between 0 and 1). You can do this by the command

replace unr=unr/100

How do your regression estimates change?
What happens if both variables are measured as rates?

4. Now test the functional form of your model by regressing the inflation rate on the **log** of unemployment.

gen *variable name* = log(unr)

Repeat the steps in section 2. Does this alternative specification give a better fit? How do you interpret your estimated coefficients?
What happens if you restore the unemployment variable