Problem Set 1. Least Squares Estimation of the 2 Variable Model

1. Using the data given in the Table in Question 1 of Problem Set 0, plot a scatter diagram of y against x. Next, calculate the least squares estimates of the constant, a, and the slope coefficient, b, from a regression of y on x, y = a + bx.

Compute and plot the residuals for each observation.
Find the R^2 from this regression.

Now see if you can replicate your findings using Stata.

2. Show that $Cov(\hat{Y}\,\hat{u}) = 0$

3. The table below shows the average annual percentage change in GDP and employment for 25 OECD countries for the period 1988-1997. Using the Stata handout as a guideline, enter the data for the first 23 countries into the *Stata* Data Editor and regress Employment on GDP growth. Interpret your coefficients.

Average Annual % Growth of Employment Growth and GDP Growth, 1988-1997

	employment	*GDP*		*employment*	*GDP*
Australia	1.68	3.04	Korea	2.57	7.73
Austria	0.65	2.55	Luxembourg	3.02	5.64
Belgium	0.34	2.16	Netherlands	1.88	2.86
Canada	1.17	2.03	New Zealand	0.91	2.01
Denmark	0.02	2.02	Norway	0.36	2.98
Finland	–1.06	1.78	Portugal	0.33	2.79
France	0.28	2.08	Spain	0.89	2.60
Germany	0.08	2.71	Sweden	–0.94	1.17
Greece	0.87	2.08	Switzerland	0.79	1.15
Iceland	–0.13	1.54	Turkey	2.02	4.18
Ireland	2.16	6.40	UK	0.66	1.97
Italy	–0.30	1.68	USA	1.53	2.46
Japan	1.06	2.81			

Now add data for the United States and Japan into your data area and do the same regression. What happens to your estimated coefficients? Is there anything else that changes in your regression output?

Now using all 25 observations, reverse the order of the variables so that you now regress GDP growth on employment. Use these estimates to derive another estimate of the effect of GDP on employment. What do you find and why?

Now show, with a proof, what your estimates of the regression coefficients would be if employment and GDP growth were both measured as rates, (i.e. dividing

the data by 100 to give numbers between 0 and 1) rather than in percentage point terms. What happens to the R^2? What happens to the estimated coefficients if only the employment growth is measured as a rate?

4. Interpret the meaning of the OLS estimates of the constant and the slope in the following prediction equations.

$$\hat{Wage} = 5 + 1.2 * Age$$

(where wage is measured in £ an hour and age is measured in years)

$$\hat{Consumption} = 3{,}000 + 0.82 * Income$$

(where annual consumption and income levels are measured in £000)

$$\hat{GDP} = -5{,}000 + 1000 * Population$$

(where GDP is measured in $ and population is measured in millions)

$$\hat{Weight} = -210 + 0.51 * Height$$

(where weight is measured in kilograms and height in centimetres)

5. Consider the following equations

a) $Y_t = \beta_1 + \beta_2 X_t + u_t$ b) $Y_t = \hat{\beta}_1 + \hat{\beta}_2 X_t + \hat{u}_t$ c) $\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t$

d) $Y_t = \hat{\beta}_1 + \hat{\beta}_2 X_t + u_t$ e) $Y_t = \beta_1 + \beta_2 X_t$ f) $Y_t = \beta_1 + \beta_2 X_t + \hat{u}_t$

Say whether each equation is true or false.

6. Suppose a regression of wages on height produces the following result

$$\hat{Wage} = 5 + 0.6 * Height, \quad R^2 = 0.4$$

Wages are measured in £ an hour and height in meters. Suppose the height data were measured in centimeters instead of meters and wages were measured in pence an hour, (ie multiplying each variable by 100). What would happen to the estimates of the slope? What would happen if wages were measured in pence an hour and height in meters?