1.

Consider the following equations

i) $u_i = Y_i - ß_1 - ß_2 X_i$

ii) $Y_i = b_1 + \hat{b}_2 X_i + \hat{u}_i$

iii) $\hat{Y}_i = \hat{b}_1 + \hat{b}_2 X_i + u_i$

where ^ indicates an estimate, otherwise the variable (and/or coefficient) is the true value

a) Say, with reason, whether each equation is true or false.

(6 marks)

b) What are the 5 main assumptions that underlie the OLS estimation technique?

(10 marks)

c) What do you understand by the terms

i) Bias of an estimator
ii) the efficiency of an estimator
iii) the Gauss-Markov property of an OLS estimate.

(9 marks)

2. Given the following regression output

$$\widehat{Investment} = 20000 - 1.0 * Interest\_rates$$

R^2=0.90   TSS = 1000 ESS = 885 Var(Interest_rate)=0.8

estimated over the period 1980-2004

i) Using the information above, find the standard error on the estimate of interest_rates in the model

(6 marks)

ii) Test the null hypothesis that the true effect of interest rates on investment is zero

(4 marks)

iv) what is the 95% confidence interval around this estimate?

(3 marks)

v) What do you understand by the terms Size of a Test, Type I and Type II error?

(9 marks)

vi) What is the approximate p value for the null hypothesis that the true coefficient on interest rates is zero?

(3 marks)

3. Interpret the meaning of the OLS estimates of the slope in each of the following time series regressions of consumption on income, (where both consumption and income are measured in £million). Some of the regression output has been hidden.

```
i) reg cons income

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------
       Model |  12000.0
    Residual |  3000.0                                 R-squared     =  0.80
-------------+------------------------------           Adj R-squared =  0.82
       Total |  15000.0                                Root MSE      =
------------------------------------------------------------------------------
        cons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |   .9172948   .0118722    77.26   0.000     .8933523    .9412372
       _cons |   13.49616   4.025455     3.35   0.002     5.378055    21.61426
```

(2 marks)

```
. reg LnCons LnIncome

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(   ,      ) = 8575.43
       Model |  4.95869491        4.95869491           Prob > F      =  0.0000
    Residual |  .024864517        .000578245           R-squared     =  0.9950
-------------+------------------------------           Adj R-squared =  0.9949
       Total |  4.98355943        .113262714           Root MSE      =  .02405
------------------------------------------------------------------------------
      LnCons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    LnIncome |   .9448621   .0102033    92.60   0.000     .9242852     .965439
       _cons |   .2790579   .0583503     4.78   0.000     .1613834    .3967325
```

(2 marks)

```
iii) reg LnCons income

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(   ,      ) = 1761.65
       Model |  4.86481426        4.86481426           Prob > F      =  0.00
    Residual |  .118745167        .002761516           R-squared     =  0.97
-------------+------------------------------           Adj R-squared =  0.97
       Total |  4.98355943        .113262714           Root MSE      =  .052
------------------------------------------------------------------------------
      LnCons |      Coef.   Std.8Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |   .0029489   .0000703    41.97   0.000     .0028072    .0030906
       _cons |   4.728025   .0238226   198.47   0.000     4.679983    4.776068
```

(2 marks)

b) Test the null hypothesis that the model as a whole in specification (i) has no explanatory power.

(4 marks)

c) Suppose specification (i) is estimated over 35 time periods (rather than the original 45) with the following results

```
. reg cons income if year<90

      Source |       SS       df       MS              Number of obs =      35
-------------+------------------------------           F(    ,      ) =
       Model |  157400.0                               Prob > F      =  0.0000
    Residual |  1000.0                                 R-squared     =
-------------+------------------------------           Adj R-squared =
       Total |  158400.0                               Root MSE      =
------------------------------------------------------------------------------
        cons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |   .9467358   .0167604    56.49   0.000     .9126366    .9808351
       _cons |   6.366222   4.704141     1.35   0.185    -3.204424    15.93687
```

Do the Chow forecast test of the hypothesis that the OLS coefficients estimated over the shorter period predict well out of sample.

(6 marks)

d) Show that the R^2 is equal to the square of the correlation coefficient between the actual and predicted values of y , ie R^2 = corr(y, $\hat{y}$)^2 where

$$corr(y, \hat{y}) = \frac{Cov(y, \hat{y})}{\sqrt{Var(y) * Var(\hat{y})}}$$

(9 marks)

4. The following output is taken from regressions of a) the annual employment growth rate (measured in % points) b) the log of annual employment growth rate on annual gdp growth rate (measured in % points) using a sample of 20 countries

```
a) reg empl gdp

      Source |       SS       df       MS              Number of obs =      16
-------------+------------------------------           F(  1,    14) =   25.56
       Model |  8.31647759     1  8.31647759           Prob > F      =  0.0001
    Residual |  5.85581708    14  .325323171           R-squared     =  0.5868
-------------+------------------------------           Adj R-squared =  0.5639
       Total |  14.1722947    15  .745910246           Root MSE      =  .57037
------------------------------------------------------------------------------
        empl |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         gdp |   .3970308   .0785257     5.06   0.000     .2320544    .5620073
       _cons |  -.1091787    .272871    -0.40   0.694    -.6824595     .464102
```

```
b) reg Lnempl gdp

      Source |       SS       df       MS              Number of obs =      16
-------------+------------------------------           F(  1,    14) =    5.71
       Model |  6.92550263     1  6.92550263           Prob > F      =  0.0280
    Residual |  21.8251244    14  1.21250691           R-squared     =  0.2409
-------------+------------------------------           Adj R-squared =  0.1987
       Total |   28.750627    15   1.5131909           Root MSE      =  1.1011
------------------------------------------------------------------------------
      Lnempl |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         gdp |   .3623099   .1515991     2.39   0.028      .043812    .6808078
       _cons |  -1.467879   .5267954    -2.79   0.012    -2.574635   -.3611229
```

i) Interpret the estimated effect of gdp in both regressions
(4 marks)

ii) Write down the equations needed to calculate a) the arithmetic mean b) the geometric mean
(4 marks)

iii) The following is output from regressions of employment growth and the log of employment growth divided by the geometric mean, (empladj and lempladj respectively). Do the Box-Cox test of the null hypothesis that the RSS from the two regressions are the same. Which specification do you prefer?

(4 marks)

```
. reg empladj gdp

      Source |       SS       df       MS              Number of obs =      16
-------------+------------------------------           F(  1,    14) =   23.25
       Model |  16.2996276     1  16.2996276           Prob > F      =  0.0003
    Residual |  9.81551482    14  .701108202           R-squared     =  0.6241
-------------+------------------------------           Adj R-squared =  0.5973
       Total |  26.1151424    15   1.7410095           Root MSE      =  .83732
------------------------------------------------------------------------------
     empladj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         gdp |   .5962491   .1236606     4.82   0.000     .3310234    .8614747
       _cons |  -.3921257    .450953    -0.87   0.399    -1.359324    .5750723
```

```
. reg lempladj gdp

      Source |       SS       df       MS              Number of obs =      16
-------------+------------------------------           F(  1,    14) =    5.73
       Model |  7.77622121     1  7.77622121           Prob > F      =  0.0313
    Residual |  19.0038339    14  1.35741671           R-squared     =  0.2904
-------------+------------------------------           Adj R-squared =  0.2397
       Total |  28.7500552    15  1.78533701           Root MSE      =  1.1651
------------------------------------------------------------------------------
    lempladj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         gdp |   .4118349   .1720662     2.39   0.031     .0427897    .7808801
       _cons |  -1.448807   .6274735    -2.31   0.037    -2.794603   -.1030099
```

iv) The following output is taken from the distribution of OLS residuals from the two regressions, resa and resb respectively

```
. su resa, det
                          Residuals
      Percentiles      Smallest
 1%    -.8867749      -.8867749
 5%     -.800374      -.7139732
10%    -.6933984      -.6728235       Obs                  16
25%    -.4225267      -.6685373       Sum of Wgt.          16
50%    -.0230367                      Mean          -2.00e-09
                        Largest       Std. Dev.      .5551584
75%      .471398       .5822049
90%     .7580767       .6624828       Variance       .3082009
95%     .8717977       .8536705       Skewness        .600000
99%     .8899248       .8899248       Kurtosis       1.200000

. su resb, det
                          Residuals
      Percentiles      Smallest
 1%     -3.17601       -3.17601
 5%     -2.60786      -2.039709
10%    -1.345669      -.6516281       Obs                  16
25%    -.4761055      -.6334555       Sum of Wgt.          16
50%     .3739764                      Mean           1.12e-08
                        Largest       Std. Dev.       1.07177
75%     .7360109       .8852508
90%     .9456291       .8893937       Variance       1.148691
95%     1.032405       1.001864       Skewness      -3.000000
99%     1.062945       1.062945       Kurtosis       6.000000
```

What is skewness?
(2 marks)

What is kurtosis?
(2 marks)

Do the Jarque-Bera test for normality in the OLS residuals for both equations. Which model do you prefer?
(6 marks)

v) If a model fails this test what does this mean for OLS estimation?
(3 marks)

END OF TEST