STUDENT NAME……………………………...…

Hemis No………………………

ECONOMICS DEPT

ASSESSED TEST

For Internal Students of
Royal Holloway

COURSE UNIT : **EC5040**

TITLE: **Econometrics**

Date of Test **17th November 2004**

Time Allowed: 55 minutes

Instructions to candidates:

WRITE ALL YOUR ANSWERS (INCLUDING ROUGH WORKING) ON THE PAPER PROVIDED

WHEN YOU HAVE FINISHED ATTACH THIS QUESTION PAPER TO YOUR ANSWERS

STATISTICAL TABLES ARE PROVIDED

SILENT NON-PROGRAMMABLE CALCULATORS MAY BE USED

DO NOT TURN OVER UNTIL TOLD TO BEGIN

1. Given the general linear model y = XB + u where y is n_x1, B is n_x1 , X is n_xk and u ~ N(0, s^2),

show that

a) the OLS estimate of the coefficient vector B, $\hat{b}$ = $(X'X)^{-1}X'y$,

(7 marks)

b) $X'\hat{u} = 0$

(8 marks)

c) $(y'y - n\bar{y}^2) = (\hat{b}'X'X\hat{b} - n\bar{y}^2) + \hat{u}'\hat{u}$ **(10 marks)**

Give a short verbal description of what these results mean in your answers

2. The following is taken from a regression of the log of hourly wages on the number of years of work experience, (exper) the square of experience (exper2) and a dummy variable (female) that takes the value 1 if the individual is female and 0 otherwise log of labour input and log of capital input. Some of the information has been concealed.

$$\widehat{Ln(hourwage)} = 2.00 + 0.036 * \exp er - 0.001 * \exp er2 - 0.30\, female$$

N=1004 **R^2 =0.090** $\bar{R}^2 = 0.080$

The variance/co-variance matrix of the parameter estimates, Var ($\hat{b}$), is given by

	exper	exper2	female
exper	0.029		
exper2	-0.005	0.001	
female	-0.090	0.002	.015625

a) Interpret the coefficient on the female dummy variable (3 marks)

b) Find the standard error of the estimate on the female variable

the estimated t value under the null that $\beta_{female} = 0$

the 95% confidence interval around this estimate

(10 marks)

c) Test the hypothesis that the coefficients on exper and exper2 add to zero

(5 marks)

d) Test the significance of the goodness of fit of the model as a whole

(4 marks)

e) Outline the intuition that lies behind the tests used in parts b) and c)

(3 marks)

3. Given the model

$$y = b_1 + b_2X_2 + b_3X_3 + u$$

and the following information

$$X'X = \begin{bmatrix} 103 & 0 & 0 \\ 0 & 10 & 5 \\ 0 & 5 & 3 \end{bmatrix} \qquad X'y = \begin{bmatrix} 10 \\ 5 \\ 1 \end{bmatrix} \qquad \sum_i (Y_i - \bar{Y})^2 = 17$$

Find the OLS estimates of

a) **the slope coefficients** **(6 marks)**

b) **the residual sum of squares (RSS)** **(4 marks)**

c) **the standard error of the slope coefficients, b_2 and b_3** **(5 marks)**

d) **What do you understand by the term "encompassing principle"** **(4 marks)**

e) **Outline the form of an encompassing test that you might use to choose between y and log(y) as a dependent variable** **(6 marks)**

4. The following regression output is based on estimates from a data set containing 30 yearly observations on total consumption and income over the period 1970:1999

```
. reg cons income if year<97

      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  1,    25) = 1280.91
       Model |  1.3386e+11     1  1.3386e+11           Prob > F      =  0.0000
    Residual |  2.6125e+09    25   104501857           R-squared     =  0.9809
-------------+------------------------------           Adj R-squared =  0.9801
       Total |  1.3647e+11    26  5.2488e+09           Root MSE      =   10223
------------------------------------------------------------------------------
        cons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |   .9016052   .0251917    35.79   0.000     .8497219    .9534886
       _cons |   17898.81   9314.261     1.92   0.066    -1284.271    37081.89


. reg cons income if year>=97
      Source |       SS       df       MS              Number of obs =       3
-------------+------------------------------           F(  1,     1) =    3.66
       Model |   706995633     1   706995633           Prob > F      =  0.3065
    Residual |   192948516     1   192948516           R-squared     =  0.7856
-------------+------------------------------           Adj R-squared =  0.5712
       Total |   899944149     2   449972074           Root MSE      =   13891
------------------------------------------------------------------------------
        cons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |    1.75263   .9155935     1.91   0.306    -9.881088    13.38635
       _cons |  -422003.4   486938.4    -0.87   0.545      -6609142     5765135

reg cons income

      Source |       SS       df       MS              Number of obs =      30
-------------+------------------------------           F(  1,    28) = 1454.04
       Model |  2.0574e+11     1  2.0574e+11           Prob > F      =  0.0000
    Residual |  3.9618e+09    28   141494469           R-squared     =  0.9811
-------------+------------------------------           Adj R-squared =  0.9804
       Total |  2.0970e+11    29  7.2311e+09           Root MSE      =   11895
------------------------------------------------------------------------------
        cons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |   .9196488   .0241176    38.13   0.000     .8702462    .9690514
       _cons |   11536.83   9381.629     1.23   0.229    -7680.561    30754.23
```

a) **Write down the restricted and unrestricted models in stacked matrix form** **(5 marks)**

b) **Do the Chow forecast test of parameter constancy** **(6 marks)**

turn over

c) A dummy variable that takes the value 1 if the observation belongs to the year 1991 and 0 otherwise is now added to the regression.

```
reg cons income d91

      Source |       SS       df       MS              Number of obs =      30
-------------+------------------------------           F(  2,    27) =  859.61
       Model |  2.0646e+11     2  1.0323e+11           Prob > F      =  0.0000
    Residual |  3.2424e+09    27   120088034           R-squared     =  0.9845
-------------+------------------------------           Adj R-squared =  0.9834
       Total |  2.0970e+11    29  7.2311e+09           Root MSE      =   10958
------------------------------------------------------------------------------
        cons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |   .9273132    .022438    41.33   0.000     .8812742    .9733522
         d91 |  -27551.04   11255.95    -2.45   0.021    -50646.34   -4455.744
       _cons |   9554.774   8680.729     1.10   0.281    -8256.611    27366.16
```

Explain, (without giving formal proof), what the addition of the dummy variable does to the residual sum of squares in the model and to the estimated coefficient on income (6 marks)

d) Now suppose the regression equation in part c were projected beyond the year 1999 on the following data

year	consumption	income
2000	55000	56000
2001	57000	60000

find the forecast values of consumption for these two periods (to the nearest whole number)

(3 marks)

write down the statistics that you would need to enable you to calculate the 95% confidence interval for the true value of consumption in each year based around these forecasts

(5 marks)

END OF TEST