NATIONAL UNIVERSITY OF LESOTHO

MSc EXAMINATIONS

EC6507 – Advanced Econometrics Theory and Practice

January 2024	**100 Marks**	**3 Hours**

INSTRUCTIONS:

1) Answer Question one and any other three questions

2) Explain your answers and show workings (*useful information is provided in the appendix*).

3) Students are allowed to use calculators and statistical tables. However, no additional material should be brought into and used in the examination.

Question 1

a. The Russian Ukraine war led to several external shocks in many economies across the globe. This is not unrelated to the fact that Russia and Ukraine export nearly a third of the world's wheat and barley and more than 70 percent of its sunflower oil. The war disrupted global supply chains across many countries in Africa that are largely self in-sufficient in food such as Lesotho. In relation to this,

i. State a valid research question and objective that applies to the Lesotho economy **[2]**

ii. Specify the economic or conceptual model **[2]**

iii. Specify the econometric model **[2]**

iv. Specify the null and alternative hypothesis **[2]**

v. Specify the nature and source of data required for such a study **[2]**

vi. Explain whether panel data and spatial data can be employed for this study **[2]**

vii. What kind of model will you be estimating? **[2]**

viii. How will you justify the variables that you will be selecting? **[2]**

ix. What are the dangers with omitting a variable on one hand and over-specifying a model on the other hand. In your view, which is worse? **[4]**

b. Likotsi, Libuseng and Hlalefang investigated the expenditure pattern for Lesotho consumers using a sample of 6334 households in the 2021 Lesotho Consumer Expenditure Survey. They performed the following three regressions. The first regression shows the result of regressing $\boldsymbol{LGFDHO}$, the logarithm of annual household expenditure on food eaten at home, on $\boldsymbol{LGEXP}$, the logarithm of total annual household expenditure, and $\boldsymbol{LGSIZE}$, the logarithm of the number of persons in the household. The second regression shows results of regressing $\boldsymbol{LGFDHOPC}$, the logarithm of food expenditure per capita ($\boldsymbol{FDHO/SIZE}$) on $\boldsymbol{LGEXPPC}$, the logarithm of total expenditure per capita ($\boldsymbol{EXP/SIZE}$). The third regression shows results of $\boldsymbol{LGFDHOPC}$ regressed on $\boldsymbol{LGEXPPC}$ and $\boldsymbol{LGSIZE}$. The STATA output results of the three regressions they performed are shown in Appendix I below.

i. Explain why the second model is a restricted version of the first, stating the restriction. [**4**]

ii. Perform an $F\ test$ of the restriction. Use 1 percent level of confidence. [**5**]

iii. Perform a $t\ test$ of the restriction. Use 1 percent level of confidence. [**4**]

iv. Summarize your conclusions from the analysis of the regression results. [**7**]

Question 2

a. Assume that you want to estimate the relationship between the age of the CEO and stock market return using the regression model based on a sample containing cross sectional data for 100 companies in 2023. In the sample you find that your data for age of the CEOs all happen to be in the range 45-47. Will this be a concern? Explain why or why not? **[4]**

b. Increasing the sample size is said to alleviate multicollinearity problem by reducing the degree of micro-numerosity.

i. But what about increasing samples by identical observations? **[3]**

ii. What happens to the variance of the coefficients? **[3]**

iii. Comment on the claim that more data is no remedy for the multi-collinearity problem if the data are simply "more of the same". **[3]**

iv. What is the biggest danger with multicolinearity? **[3]**

c. The variance of the estimator is derived as

$$Var[\hat{\beta}|X] = E\left[(\hat{\beta}-\beta)(\hat{\beta}-\beta)^{\mathrm{T}}\right] = \sigma^2(X^TX)^{-1}$$

Where, the assumptions of homoscedasticity and zero autocorrelation is used in the derivation. State and explain the assumptions for homoscedasticity and autocorrelation that validate this expression. **[4]**

Question 3

a. Mahapa, Matseliso and Mpho investigating whether government expenditure tends to crowd out investment fits the regression (standard errors in parentheses):

$$\hat{\boldsymbol{I}} = \underset{(7.79)}{\mathbf{18.10}} - \underset{(0.14)}{\mathbf{1.97}}\boldsymbol{G} + \underset{(0.02)}{\mathbf{0.36}}\boldsymbol{Y} \qquad \boldsymbol{R^2} = \mathbf{0.99}$$

where $\boldsymbol{G}$ is government recurrent expenditure, $\boldsymbol{Y}$ is gross domestic product and $\boldsymbol{I}$ is investment. Instead Nkobolo, Tlhalefo and Ntsoaki run the following regressions as alternative specifications of the model ran above (standard errors in parentheses):

$$\frac{\hat{I}}{P} = -\underset{(0.28)}{0.03}\frac{1}{P} - \underset{(0.16)}{0.69}\frac{G}{P} + \underset{(0.03)}{0.34}\frac{Y}{P} \qquad R^2 = 0.97$$

$$\frac{\hat{I}}{P} = \underset{(0.04)}{0.39} + \underset{(0.42)}{0.03}\frac{1}{Y} - \underset{(0.22)}{0.93}\frac{Y}{P} \qquad R^2 = 0.78$$

$$\widehat{\log I} = -\underset{(0.26)}{2.44} - \underset{(0.12)}{0.63}\log G + \underset{(0.12)}{1.60}\log Y \qquad R^2 = 0.98$$

where G, Y, and I are as before and P is population.

In each case the regression is run again for the subsamples of observations with the 11 smallest and 11 greatest values of the sorting variable, after sorting by G/Y, Y/P and $\log Y$, respectively. The residual sums of squares are as shown in table 1.

Table 1

	11 *smallest*	11 *largest*
(1)	1.43	12.63
(2)	0.0223	0.0155
(3)	0.573	0.155

Perform a Goldfeld–Quandt test at 1 percent level of confidence for each model specification and discuss the merits of each specification. **[14]**

Is there evidence that investment is an inverse function of government expenditure? **[1]**

b. Rabeah, Mamanyena and Mamazondo save the residuals from the full-sample regression in **(a)** and regress their squares on G, Y, their squares, and their product. The STATA output results are shown in Appendix II below. Perform a White test for heteroskedasticity at 1 percent level of confidence. **[5]**

Question 4

a. Describe all the steps that are carried out when running a two stage least square (2SLS) regression with one endogenous variable, one exogenous variable and one instrument. **[5]**

b. What are the identification classifications of parameters in regression and briefly explain them? **[6]**

c. What are the two conditions that a valid instrument must satisfy? **[2]**

d. Mohapi, Mabela and Nthabeleng perform an OLS regression of the logarithm of hourly earnings on $\boldsymbol{S}$, $\boldsymbol{EXP}$, $\boldsymbol{ASVABC}$, $\boldsymbol{MALE}$, $\boldsymbol{ETHBLACK}$S and $\boldsymbol{ETHHISP}$ using EAWE data set and an IV regression using $\boldsymbol{SM}$, $\boldsymbol{SF}$, and $\boldsymbol{SIBLINGS}$ as instruments for $\boldsymbol{ASVABC}$. The STATA outputs of the two regressions (IV and OLS) are shown in Appendix III below. Perform a Durbin–Wu–Hausman test at 5 percent confidence to evaluate whether $\boldsymbol{ASVABC}$ appears to be subject to measurement error. **[7]**

Question 5

a. What is autocorrelation and why is it a problem in time series data? **[2]**

b. What is an autoregressive (AR) model, distributed lag (DL) model and autoregressive distributed lag (ADL) model. Give an example of each in terms of $\boldsymbol{X}$ and $\boldsymbol{Y}$, where $\boldsymbol{Y}$ is the dependent variable and $\boldsymbol{X}$ is the independent variable. **[3]**

c. The tables in Appendix IV show the Eviews output results of a regression model explaining the relationship between food consumption and disposable income. Using the Eviews output results results in Appendix IV, perform the Breusch-Godfrey test for autocorrelation at 1 percent level of confidence. **[7]**

d. Using the same results in Appendix IV, perform the Durbin-Watson test for autocorrelation at 1 percent level of confidence. **[8]**

APPENDICES

Appendix I

```
. reg LGFDHO LGEXP LGSIZE

  Source |       SS       df       MS              Number of obs =     868
---------+------------------------------           F(  2,   865) =  460.92
   Model |  138.776549     2  69.3882747           Prob > F      =  0.0000
Residual |  130.219231   865  .150542464           R-squared     =  0.5159
---------+------------------------------           Adj R-squared =  0.5148
   Total |  268.995781   867  .310260416           Root MSE      =    .388

------------------------------------------------------------------------------
  LGFDHO |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   LGEXP |   .2866813   .0226824     12.639   0.000       .2421622    .3312003
  LGSIZE |   .4854698   .0255476     19.003   0.000       .4353272    .5356124
   _cons |   4.720269   .2209996     21.359   0.000       4.286511    5.154027
------------------------------------------------------------------------------

. reg LGFDHOPC LGEXPPC

  Source |       SS       df       MS              Number of obs =     868
---------+------------------------------           F(  1,   866) =  313.04
   Model |  51.4364364     1  51.4364364           Prob > F      =  0.0000
Residual |  142.293973   866  .164311747           R-squared     =  0.2655
---------+------------------------------           Adj R-squared =  0.2647
   Total |   193.73041   867  .223449146           Root MSE      =  .40535

------------------------------------------------------------------------------
LGFDHOPC |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
 LGEXPPC |    .376283   .0212674     17.693   0.000       .3345414    .4180246
   _cons |   3.700667   .1978925     18.700   0.000       3.312262    4.089072
------------------------------------------------------------------------------

. reg LGFDHOPC LGEXPPC LGSIZE

  Source |       SS       df       MS              Number of obs =     868
---------+------------------------------           F(  2,   865) =  210.94
   Model |  63.5111811     2  31.7555905           Prob > F      =  0.0000
Residual |  130.219229   865  .150542461           R-squared     =  0.3278
---------+------------------------------           Adj R-squared =  0.3263
   Total |   193.73041   867  .223449146           Root MSE      =    .388

------------------------------------------------------------------------------
LGFDHOPC |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
 LGEXPPC |   .2866813   .0226824     12.639   0.000       .2421622    .3312004
  LGSIZE |  -.2278489   .0254412     -8.956   0.000      -.2777826   -.1779152
   _cons |   4.720269   .2209996     21.359   0.000       4.286511    5.154027
------------------------------------------------------------------------------
```

Appendix II

```
reg EI2 G Y G2 Y2 GYPROD
------------------------------------------------------------------------------
    Source |       SS       df       MS              Number of obs =      30
-----------+------------------------------           F(  5,    24) =  389.39
     Model |   229715064     5  45943012.8           Prob > F      =  0.0000
  Residual |  2831715.29    24  117988.137           R-squared     =  0.9878
-----------+------------------------------           Adj R-squared =  0.9853
     Total |   232546779    29  8018854.45           Root MSE      =  343.49
------------------------------------------------------------------------------
       EI2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------+------------------------------------------------------------------
         G |  -36.11647   7.400742    -4.88   0.000    -51.39085   -20.84209
         Y |   5.522015   1.188795     4.65   0.000     3.068463    7.975567
        G2 |   .5463273    .027054    20.19   0.000     .4904907     .602164
        Y2 |    .007948   .0003913    20.31   0.000     .0071403    .0087556
    GYPROD |  -.1350344   .0060013   -22.50   0.000    -.1474204   -.1226485
     _cons |   196.4158   92.97307     2.11   0.045     4.528833    388.3028
------------------------------------------------------------------------------
```

Appendix III

```
. ivreg LGEARN S EXP MALE ETHBLACK ETHHISP (ASVABC=SM SF SIBLINGS)

Instrumental variables (2SLS) regression
------------------------------------------------------------------------------
    Source |       SS       df       MS              Number of obs =     500
-----------+------------------------------           F(  6,   493) =   22.29
     Model |   27.631679     6  4.60527983           Prob > F      =  0.0000
  Residual |  121.501359   493  .246453061           R-squared     =  0.1853
-----------+------------------------------           Adj R-squared =  0.1754
     Total |  149.133038   499  .298863804           Root MSE      =  .49644
------------------------------------------------------------------------------
    LGEARN |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------+------------------------------------------------------------------
    ASVABC |  -.0938253   .1319694    -0.71   0.477    -.3531172    .1654666
         S |   .1203265   .0251596     4.78   0.000     .0708931    .1697599
       EXP |   .0444094   .0092246     4.81   0.000      .026285    .0625338
      MALE |   .1909863   .0456252     4.19   0.000     .1013424    .2806302
  ETHBLACK |  -.1678914   .1355897    -1.24   0.216    -.4342963    .0985136
   ETHHISP |    .075698   .0828383     0.91   0.361    -.0870617    .2384576
     _cons |   .6503199   .3570741     1.82   0.069    -.0512548    1.351895
------------------------------------------------------------------------------
Instrumented:  ASVABC
Instruments:   S EXP MALE ETHBLACK ETHHISP SM SF SIBLINGS
------------------------------------------------------------------------------
. estimates store IV1
```

```
reg LGEARN S EXP ASVABC MALE ETHBLACK ETHHISP

------------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =     500
-------------+------------------------------           F(  6,   493) =   23.81
       Model |  33.5095496     6  5.58492493           Prob > F      =  0.0000
    Residual |  115.623489   493  .234530403           R-squared     =  0.2247
-------------+------------------------------           Adj R-squared =  0.2153
       Total |  149.133038   499  .298863804           Root MSE      =  .48428
------------------------------------------------------------------------------
      LGEARN |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           S |   .0953713   .0106101     8.99   0.000     .0745246    .1162179
         EXP |    .043139   .0089279     4.83   0.000     .0255976    .0606805
      ASVABC |   .0477892   .0282877     1.69   0.092      -.00779    .1033685
        MALE |   .1954406   .0443323     4.41   0.000     .1083371    .2825441
    ETHBLACK |  -.0448382    .074738    -0.60   0.549    -.1916824     .102006
     ETHHISP |   .1226463   .0692577     1.77   0.077    -.0134303     .258723
       _cons |   .9766376   .1938648     5.04   0.000     .5957345    1.357541
------------------------------------------------------------------------------
. estimates store OLS1


hausman IV1 OLS1, constant

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      IV1          OLS1        Difference          S.E.
-------------+----------------------------------------------------------------
      ASVABC |   -.0938253     .0477892       -.1416145        .1289021
           S |    .1203265     .0953713        .0249552         .022813
         EXP |    .0444094      .043139        .0012704        .0023208
        MALE |    .1909863     .1954406       -.0044543        .0107847
    ETHBLACK |   -.1678914    -.0448382       -.1230532        .1131318
     ETHHISP |     .075698     .1226463       -.0469484        .0454484
       _cons |    .6503199     .9766376       -.3263177        .2998639
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from ivreg
          B = inconsistent under Ha, efficient under Ho; obtained from regress
    Test:  Ho:  difference in coefficients not systematic
                  chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        1.21
                Prob>chi2 =      0.9908
```

Appendix IV

```
=============================================================
Dependent Variable: RLGFOOD
Method: Least Squares
Sample(adjusted): 1961 2003
Included observations: 43 after adjusting endpoints
=============================================================
     Variable      Coefficient Std. Error t-Statistic  Prob.
=============================================================
        C            0.071220   0.277253   0.256879   0.7987
      LGDPI          0.000251   0.006491   0.038704   0.9693
     LGPRFOOD       -0.015572   0.051617  -0.301695   0.7645
    RLGFOOD(-1)      1.009693   0.163240   6.185318   0.0000
    RLGFOOD(-2)     -0.289159   0.171960  -1.681548   0.1009
=============================================================
R-squared            0.602010    Mean dependent var 0.000149
Adjusted R-squared   0.560117    S.D. dependent var 0.020368
S.E. of regression   0.013509    Akaike info criter-5.661981
Sum squared resid    0.006935    Schwarz criterion -5.457191
Log likelihood       126.7326    F-statistic         14.36996
Durbin-Watson stat   1.892212    Prob(F-statistic)  0.000000
=============================================================
```

```
=============================================================
Dependent Variable: LGFOOD
Method: Least Squares
Sample: 1959 2003
Included observations: 45
=============================================================
     Variable      Coefficient Std. Error t-Statistic  Prob.
=============================================================
        C            2.236158   0.388193   5.760428   0.0000
      LGDPI          0.500184   0.008793   56.88557   0.0000
     LGPRFOOD       -0.074681   0.072864  -1.024941   0.3113
=============================================================
R-squared            0.992009    Mean dependent var 6.021331
Adjusted R-squared   0.991628    S.D. dependent var 0.222787
S.E. of regression   0.020384    Akaike info criter-4.883747
Sum squared resid    0.017452    Schwarz criterion -4.763303
Log likelihood       112.8843    Hannan-Quinn crite-4.838846
F-statistic          2606.860    Durbin-Watson stat 0.478540
Prob(F-statistic)    0.000000
=============================================================
```