"Investing in Africa's Future"

COLLEGE OF HEALTH, AGRICULTURE & NATURAL SCIENCES

NAEC 503 APPLIED ECONOMETRICS

END OF FIRST SEMESTER EXAMINATIONS

NOVEMBER 2021

LECTURER: PROF L. DUBE

DURATION: 5 HOURS

INSTRUCTIONS

1. Do not write your name on the answer sheet
2. Use Answer Sheets Provided
3. Begin your answer for Each Question on a New Page
4. Credit is Given for Neat Presentation

INSTRUCTIONS TO CANDIDATES

Answer any one question. All questions carry equal marks. (Show all workings, steps and formulas to obtain full scores)

Question 1

a) From a sample of 10 observations, the following statistics descriptions were obtained from the EVIEWS as:

Group: GROUP01 Workfile: TABLE3-5

View | Procs | Objects | Print | Name | Freeze | Sample | Sheet | Stats | Spec

	X	X_SQUARE	XY	Y	Y_SQUARE
Mean	5.500000	38.50000	37.20000	5.500000	38.50000
Median	5.500000	30.50000	27.00000	5.500000	30.50000
Maximum	10.00000	100.0000	81.00000	10.00000	100.0000
Minimum	1.000000	1.000000	2.000000	1.000000	1.000000
Std. Dev.	3.027650	34.17358	31.76931	3.027650	34.17358
Skewness	0.000000	0.568676	0.301763	0.000000	0.568676
Kurtosis	1.775758	2.031676	1.489464	1.775758	2.031676
Jarque-Bera	0.624487	0.929676	1.102484	0.624487	0.929676
Probability	0.731803	0.628237	0.576234	0.731803	0.628237
Sum	55.00000	385.0000	372.0000	55.00000	385.0000
Sum Sq. Dev.	82.50000	10510.50	9083.600	82.50000	10510.50
Observations	10	10	10	10	10

Consider the regression model: $Y_i = \beta_1 + \beta_2 X_i + u_i$

i) What are the values of $\hat{\beta}_1$ and $\hat{\beta}_2$? How to interpret the $\hat{\beta}_2$? (10%)

ii) Are the estimated coefficients statistically different from zero at 5% level of significance? (10%)

iii) What is the value of R^2? (10%)

iv) Suppose you re-check the sample data and find that the last two pairs of observations were wrongly typed. The correct data should be

Obs.	Y	X		Y	X
9	5	7	Instead of	2	6
10	2	4		1	4

What will be the corrected $\hat{\beta}_2$ and R^2? (10%)

v) Construct the ANOVA table. (10%)

b) Suppose you used the dummy variable to run the saving-income regression for the years from 1980 to 1995 and obtained the estimated results as

$\widehat{\textbf{Saving}}_t$	= 1.0161	+152.478 $\mathbf{D}_t$	+0.0803 $\textbf{Income}_t$	-0.0051($\mathbf{D}_t \bullet \textbf{Income}_t$)
se	= (0.0503)	(160.6090)	(0.0401)	(0.0021)

N=30 R^2=0.936 $\bar{R}^2$=0.9258 SEE=0.1217 DW=0.9549

Where D_t = 1 for 1982-1995
= 0 for 1970-1981

i) Write the estimated saving-income relationships for the two different periods? (6%)
ii) Test whether there is any structural change in the regression? (7%)
iii) If your professor said that you should use the "Chow" test to carry out the test of stability among the data, clearly state are the procedures of "Chow" test and the F-statistic? (7%)

c) A researcher run the OLS regression and got his computer printout, but somehow the printer had some problem and smudged some information in the ANOVA table as following:

Source of Variations	Sum of Squares	Degree of Freedom	MSS
Due to Regression (ESS)	800	$%^&	&%$#
Due to Residual (RSS)	%#*&	45	#$%&
Total variation (TSS)	1200	49	@%&@

i) How many observations are there in the sample? (2%)
ii) How many independent variables are used in the regression equation? (2%)
iii) What is the $\sum \hat{\varepsilon}^2$? (4%)
iv) What is the $\hat{\sigma}^2$? (4%)
v) What is the standard error of regression (SEE)? (4%)
vi) What is the R^2? (4%)
vii) Calculate the F-statistic to test the null hypothesis that all $\hat{\beta}_i$'s are equal to zero at the 5 % level of significance. (5%)
viii) Briefly state the assumptions underlying the classical linear regression model. (10%)

Question 2

a) A researcher used OLS to estimate the following demand for money function for Mutare town over the 39 year period 1969 to 2007.

$$\log (M/P) = b_0 + b_1 \log (i) + b_2 \log (Y/P) + b_3 \log (L/P) + u$$

$$\log (M/P) = 2.310 - 0.761 \log(i) + 0.008 \log(Y/P) + 0.012 \log(L/P)$$
s.e. (0.11) (0.44) (0.001) (0.006)

where: M = quantity of nominal money demanded,
i = interest rate,
L = nominal amount of liquid assets, and
P =price index (to deflate nominal values, so as to eliminate changes in purchasing power).

The residuals (e_t) showed that $\Sigma e_t^2 = 1.0567$ and $\Sigma (e_t - e_{t-1})^2 = 0.2240$.

On the basis of this information, the researcher decided to re-estimate the demand for money function, using the first differences of the original variables.

i) Calculate the Durbin-Watson d statistic. (10%)
ii) Does the value of d justify the 'first differences' solution, adopted by the researcher? (10%)

b) The following questions are based on the study of demand for chicken in the United States, 1960-1982, and the EVIEWS results of the estimated model are as

Equation: PART02 Workfile: TABLE7-9

View | Procs | Objects | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LOG(Y)
Method: Least Squares
Date: 01/28/03 Time: 17:30
Sample: 1960 1982
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.125498	0.137882	15.41533	0.0000
LOG(X2)	0.405924	0.044791	[illegible]535	0.0000
LOG(X3)	-0.438825	0.083332	[illegible]956	0.0000
LOG(X4)	0.106656	0.087838	[illegible]14228	0.2395

R-squared	0.981509	Mean dependent var	3.663887
Adjusted R-squared	0.978590	S.D. dependent var	0.187659
S.E. of regression	[illegible]	Akaike info criterion	-4.195488
Sum squared resid	0.014326	Schwarz criterion	-3.998011
Log likelihood	52.24812	F-statistic	336.1808
Durbin-Watson stat	1.778678	Prob(F-statistic)	0.000000

Where Y = per capita consumption of chickens, lb.
X2 = real disposable income per capita, $

X3 = real retail price of chicken per lb, ¢
X4 = real retail price of pork per lb, ¢

Unfortunately, the printer had malfunctioned and cannot print out the exact number of t-statistics and the SEE.

i) How would you interpret the coefficients of LOG(X2), LOG(X3) and LOG(X4) in this model with economic meaning? (10%)
ii) Comment on the results in view of your prior expectation and judge on whether it is good or not. (10%) (Show the step to get the t-vales.)

iii) Establish a 95% confidence interval for $\hat{\beta}_2$. (10%)
iv) Set up the ANOVA table for the above three-variable model. (10%)
v) Would you reject the hypothesis that the true coefficient of $\hat{\beta}_4$ is *greater* than zero at 10% significant level? (Remember to state the null hypothesis and alternative hypothesis and show the steps of test.) (10%)
vi) In the regression model, suppose you divide each log(Y) value by a constant, say 100. Will it change the results of the estimated coefficients and statistics? Explain. (10%)

c) The following "demand-for-money (M1)" equation was estimated in logarithms from annual data for the periods 1960-91, 1960-75 and 1976-91:

$$M1 = \beta_1 + \beta_2 Rate + \beta_3 Disc + \varepsilon$$

Where R = interest rate, Disc = discount rate. The estimated results were reported in the following table. (t-values are in parentheses)

Dependent Variable	Constant	Rate	Disc	R^2	SEE	RSS
1960-91:						
M1	4.985	0.405	0.380	0.8931	0.2639	2.0202
	(22.039)	(2.603)	(10.375)			
1960-75:						
M1	4.609	0.941	0.139	0.9064	0.1203	0.1883
	(28.113)	(8.343)	(4.241)			
1976-91:						
M1	5.164	-0.120	0.527	0.7543	0.1941	0.4898
	(8.148)	(-0.067)	(6.025)			

i) Test the hypothesis that the demand for money function has shifted between the two sub-periods 1919-39 and 1940-57. (10%)
ii) Instead of running three regressions to test structural stability, explain which simple method you would suggest to carry out the similar test. (10%)

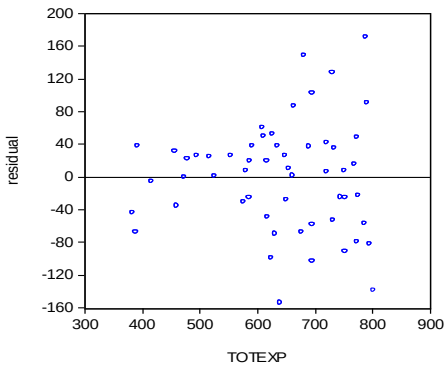
Question 3

a) A researcher, Mr. Wasu, collected data on 55 households in a developing country and got the following regression results, where *foodexp* and *totexp* are food expenditure and total expenditure, respectively.

$$\underset{se}{\widehat{foodexp}} = \underset{(50.8564)}{94.2088} + \underset{(0.0783)}{0.4368}\ totexp$$

$$R^2 = 0.3698,\ N = 55.$$

i) The above results are derived by applying the method of least squares. Mr. AAA then plots the residuals against *totexp* in a diagram as shown below. He thinks that a classical assumption may not hold in the above regression. What is the assumption violated? Explain. (6%)



ii) In view of the problem in (a), what would be the consequences of the least square estimators? (6%)

iii) Mr. Wasu finally comes up with the following regression results and gets rid of the problem in (a). (8%)

$$\underset{se}{\widehat{ln(foodexp)}} = \underset{(0.7780)}{1.1543} + \underset{(0.1207)}{0.7363}\ ln(totexp)$$

$$R^2 = 0.4125,\ N = 55.$$

When totexp = 640, find a prediction for the value of foodexp (10%). (Write your answer in 2 decimal places.)

b) The quantity supplied of a commodity X is assumed to be a linear function of the price of X and the wage rate of labour used in the production of X. The population supply equation is given as:

$$Q = b_0 + b_1P_x + b_2W + \mu$$

Where Q = quantity supplied of X

Px = price of X

W = wage rate

Using the sample data of the following table,

Y= Q	20	35	30	47	60	68	76	90	100	105	130	140	125	120	135
$X_1=P_x$	10	15	21	26	40	37	42	33	30	38	60	65	50	35	42
X_2=W	12	10	9	8	5	7	4	5	7	5	3	4	3	1	2

Intermediate results:

$\sum Y = 1\ 281$ $\sum X_1 = 544$ $\sum X_2 = 85$

$\sum X_1 Y = 53\ 665$ $\sum X_1^2 = 22\ 922$ $\sum X_1 X_2 = 2\ 568$

$\sum X2Y = 5\ 706$ $\sum Y^2 = 132\ 609$ $\sum X_2^2 = 617$

i) Estimate the parameters by OLS. (10%)
ii) Test the statistical significance of the individual coefficients at the 5 percent level of significance. (10%)
iii) What percentage of the total variation in the quantity supplied is explained by both P_x and W? (10%)
iv) Compute the price elasticity of supply at the mean price and mean quantity traded. (10%)

c) Suppose you regress Y on X_2, X_3 X_4, and X_5 as following:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + u_1$$

Y = the number of wildcats drilled (Thousands)
X_2 = price at the wellhead in the previous period (in constant dollars, 1972=100)
X_3 = domestic output ($millions of barrels per day)
X_4 = GNP constant dollars ($billions)
X_5 = trend variable
The regression result is obtained from EVIEWS as follow:

Equation: EQ01 Workfile: TABLE7-7

View | Procs | Objects | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: Y
Method: Least Squares
Date: 02/25/03 Time: 10:18
Sample: 1 31
Included observations: 31

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9.854596	8.895196	-1.107856	0.2781
X2	2.701012	0.695769		0.0006
X3		0.937314	3.264226	0.0031
X4	-0.016060	0.008179		0.0604
X5	-0.022701	0.272306	-0.083368	0.9342

R-squared	0.580377	Mean dependent var	10.64613
Adjusted R-squared	0.515819	S.D. dependent var	2.351515
S.E. of regression	1.[illegible]57	Akaike info criterion	3.969390
Sum squared resid	69.61077	Schwarz criterion	4.200678
Log likelihood	-56.52554	F-statistic	
Durbin-Watson stat	0.933888	Prob(F-statistic)	0.000107

i) Fill in the missing numbers due to the malfunction of printer. (5%)

ii) How would you interpret this result is good or not? How would you interpret the coefficients $\hat{\beta}_2$ and $\hat{\beta}_3$? (7%)

iii) Would you reject the hypothesis that the domestic output (X_3) has the effect of 3.00 on wildcat drilled (Y)? (5%)

iv) Why are you using the t-test in (iii) above but not using the normal distribution test? (3%)

v) Set up the ANOVA table for this example. (10%)