# UGANDA MARTYRS UNIVERSITY

# NKOZI

**UNIVERSITY EXAMINATIONS**

**END OF SEMESTER TWO FINAL ASESSMENT 2022/23**

**BSC III GEN & BSC III ECON**

ECO 3202

**ECONOMETRICS II**

**DATE:** **Wednesday 24 May 2023**

**TIME:** **09:30AM – 12:30PM**

***Instructions:***

1. *Question **ONE** is **COMPULSORY***
2. *Answer **FOUR (04)** questions in TOTAL*
3. *All questions carry equal marks*
4. *Do not write anything on the questions paper.*
5. *Carefully read through ALL the questions before attempting.*
6. *No **names** should be written anywhere on the examination booklet.*
7. *Ensure your work is **clear** and **readable**. Untidy work shall be penalized.*
8. *Any type of examination Malpractice will lead to automatic disqualification.*
9. *Ensure that your **ID number** is indicated on all pages of the examination answer booklet.*

## QUESTION ONE (COMPULSORY)

Suppose that a researcher collects data on houses that have sold in a particular neighborhood over the past year and obtains the regression results in the table shown below.

| Dependent Variable: *ln(Price)* | | | | | |
|---|---|---|---|---|---|
| **Regressor** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| *Size* | 0.00042<br>(0.000038) | | | | |
| *ln(Size)* | | 0.69<br>(0.054) | 0.68<br>(0.087) | 0.57<br>(2.03) | 0.69<br>(0.055) |
| *ln(Size)*$^2$ | | | | 0.0078<br>(0.14) | |
| *Bedrooms* | | | 0.0036<br>(0.037) | | |
| *Pool* | 0.082<br>(0.032) | 0.071<br>(0.034) | 0.071<br>(0.034) | 0.071<br>(0.036) | 0.071<br>(0.035) |
| *View* | 0.037<br>(0.029) | 0.027<br>(0.028) | 0.026<br>(0.026) | 0.027<br>(0.029) | 0.027<br>(0.030) |
| *Pool* × *View* | | | | | 0.0022<br>(0.10) |
| *Condition* | 0.13<br>(0.045) | 0.12<br>(0.035) | 0.12<br>(0.035) | 0.12<br>(0.036) | 0.12<br>(0.035) |
| *Intercept* | 10.97<br>(0.069) | 6.60<br>(0.39) | 6.63<br>(0.53) | 7.02<br>(7.50) | 6.60<br>(0.40) |
| **Summary Statistics** | | | | | |
| *SER* | 0.102 | 0.098 | 0.099 | 0.099 | 0.099 |
| $\bar{R}^2$ | 0.72 | 0.74 | 0.73 | 0.73 | 0.73 |
| **Variable definitions** | | | | | |
| Price=sale price (Ush); Size=house size (in square feet); Bedrooms=number of bedrooms; Pool=binary variable (1 if house has a swimming pool, 0 otherwise); View=binary variable (1 if house has a nice view, 0 otherwise); Condition=binary variable (1 if house is in excellent condition, 0 otherwise)<br>Note: ln = Natural Logarithm; SER = Standard Error of the Regression | | | | | |

(a) Using the results in column (1), what is the expected change in price of building a 500-square-foot addition to a house? Construct a 95% confidence interval for the percentage change in price. **(05marks)**

(b) Comparing columns (1) and (2), is it better to use *Size* or ln(*Size*) to explain house prices? **(05marks)**

(c) Using column (2), what is the estimated effect of pool on price? (Make sure you get the units right.) Construct a 95% confidence interval for this effect. **(03marks)**

(d) The regression in column (3) adds the number of bedrooms to the regression. How large is the estimated effect of an additional bedroom? Is the effect statistically significant? Why do you think the estimated effect is so small? (*Hint:* Which other variables are being held constant?) **(05marks)**

(e) Is the quadratic term ln(*Size*)$^2$ important? **(02marks)**

(f) Use the regression in column (5) to compute the expected change in price when a pool is added to a house that doesn't have a view. Repeat the exercise for a house that has a view. Is there a large difference? Is the difference statistically significant? **(05marks)** od over the past

QUESTION TWO

Write brief notes on the following as applied in Econometric modelling

(a) Adjusted $R^2$ (05marks)

(b) Multicollinearity (05marks)

(c) Generalized Least Squares model (05marks)

(d) Endogeneity Problem (05marks)

(e) Gaussian white noise process (05marks)

QUESTION THREE

Given the following two-variable model:

$Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i$ *Where* $X_{0i} = 1$ *for each* $i$.

*And, the variance of the disturbance term* $u_i$ *is heteroscedastic.*

(a) Give six (06) reasons why the variances of $u_i$ may be heteroscedastic (variable). (12marks)

(b) Assuming that the heteroscedastic variances $\delta_i^2$ are known. Show how the method of Generalized Least Squares (GLS) can be used to obtain variables that satisfy the standard least-squares assumptions. Hence producing estimators that are BLUE, when the variance of the disturbance term in the GLS equation is constant.

(13marks)

QUESTION FOUR

This following problem compares total compensation among top executives in a large set of Uganda's public corporations in the 2000. (Each year these publicly traded corporations must report total compensation levels for their top five executives.)

(a) Let *Female* be an indicator variable that is equal to 1 for females and 0 for males. A regression of the logarithm of earnings onto *Female* yields

$$\widehat{\ln(Earnings)} = \underset{(0.01)}{6.48} - \underset{(0.05)}{0.44}\, Female,\ SER = 2.65.$$

i) The estimated coefficient on *Female* is -0.44. Explain what this value means. (02marks)

ii) The Standard Error of the Regression (*SER*) is 2.65. Explain what this value means. (03marks)

iii) Does this regression suggest that female top executives earn less than top male executives? Explain. (03marks)

iv) Does this regression suggest that there is gender discrimination? Explain. (06marks)

(b) Two new variables, the market value of the firm (a measure of firm size, in millions of Ushs) and stock return (a measure of firm performance, in percentage points), are added to the regression:

$$\widehat{\ln(Earnings)} = \underset{(0.03)}{3.86} - \underset{(0.04)}{0.28}\, Female + \underset{(0.004)}{0.37}\ln(MarketValue) + \underset{(0.003)}{0.004}\, Return,$$

$$n = 46{,}670,\ \bar{R}^2 = 0.345.$$

i) The coefficient on ln(*MarketValue*) is 0.37. Explain what this value means. (03marks)

ii) The coefficient on *Female* is now -0.28. Explain why it has changed from the regression in (a) (03marks)

(c) Are large firms more likely than small firms to have female top executives? Explain. (05marks)

## QUESTION FIVE

(a) The most celebrated test for detecting serial correlation is that developed by statisticians Durbin and Watson popularly known as the Durbin Watson ***d*** statistics. Define Durbin Watson ***d*** statistics and give the assumptions underlying the ***d*** statistic. (12marks)

(b) Using relevant illustrations, explain the procedure followed in carrying out the Durbin Watson ***d*** test and the corresponding decision rules. (13marks)

## QUESTION SIX

(a) Assume the following 'single set-dummy variable' models with several categories:

Model A: $Y_t = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \ldots\ldots\ldots + \alpha_m D_{mt} + u_t$

Model B: $Y_t = \gamma_1 + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \ldots\ldots\ldots + \gamma_m D_{mt} + \beta_i X_{it} + u_t$

$$t = 1,2,\ldots\ldots,n, \qquad i = 1,2,\ldots\ldots,k$$

(i) Explain the difference between the two models in terms of type and interpretation (06marks)

(ii) Why is it not possible to estimate Model A with a general constant term? What should be done to make the estimation possible? (03marks)

(iii) What statistics and null hypotheses are used in Model B to test for, one, *differential intercepts from the base category* and two, *differential intercepts.* (04marks)

(b) The following model has two sets of dummy variables (i.e. gender and colour) each with two categories (male, female) and (white, black), that is,

$$Y_t = \alpha_1 + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \beta_i X_{it} + u_t, \qquad t = 1,2,\ldots\ldots,n$$

Where;

$Y_t$ = annual salary of a UMU lecturer;

$X_t$ = the years of teaching experience

$$D_2 = \begin{cases} 1 & \text{if male} \\ 0 & \text{if female} \end{cases} \quad \text{and} \quad D_3 = \begin{cases} 1 & \text{if white} \\ 0 & \text{if black} \end{cases}$$

(i) Specify the base category lecturer and determine his/her mean salary? (02marks)

(ii) Determine the mean salary of a black male lecturer, a white female lecturer and a white male lecturer? (06marks)

(iii) Assume the model is estimated by OLS, briefly interpret the tests on coefficients $\alpha_2$ and $\alpha_3$. (04marks)

**END**