MACHAKOS UNIVERSITY

University Examinations for 2022/2023 Academic Year

SCHOOL OF BUSINESS, ECONOMICS AND HOSPITALITY AND TOURISM MANAGEMENT

DEPARTMENT OF ECONOMICS

……… YEAR SECOND SEMESTER EXAMINATION FOR

MASTER OF SCIENCE (AGRICULTURAL ECONOMICS)

AED 806: APPLIED ECONOMETRICS

DATE: **TIME:**

INSTRUCTIONS

    (i)   **Answer question ONE and any other TWO questions**

    (ii)   **Show your workings clearly**

QUESTION ONE (COMPULSORY) (30 MARKS)

a)   An agricultural economist sought to investigate whether there was a relationship between consumption and income. The following specified model was adopted

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

Where Y represents consumption and X represents income

To achieve this, the following data was collected from five households.

Household	A	B	C	D	E
Income (X) in Kshs.'000s	2	5	8	10	15
Consumption(Y) in Kshs.'000s	1	2	3	4	5

    (i) Fit the sample consumption function for the 5 households and interpret your results. From the marginal propensity to consume, what can you say about the nature of households in the sample?   (5 marks)

(ii) Compute the ESS and RSS for this data and explain how the two relate   (4 marks)

(iii)Test whether the slope is significantly different from zero at the 5% level of significance.
       (4 marks)

b)   A researcher estimated the performance of master degree students in the department of agriculture in Machakos University and obtained the following equation. Standard errors are in parenthesis.

$$Y_i = 1.00 + 0.005X_1 - 0.20X_2 + 0.10X_3$$

$$(0.01) \qquad (0.20) \qquad (0.10)$$

Where Y is performance of master degree student measured by average marks obtained at the end of the course, X_1 is average marks obtained at the end first degree course, X_2 is student's work experience in years, X_3 student's age in years.

(i)   Does multicollinearity appear to exist in this equation? Explain your answer
       (2 marks)

(ii)   What suggestions (if any) would you have for another run of this equation? (2 marks)

(iv)In running panel data, fixed effects model is preferred over random effects model. Is this statement true or false? Explain your answer.   (3 marks)

QUESTION TWO (20 MARKS)

a) Consider the model
$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 Y_{t-1} + V_t$$
Suppose Y_{t-1} and V_t are correlated. To remove the correlation, suppose we use the following instrumental variable approach. First, we regress Y_{t-1} on X_{1t} and X_{2t} and obtain $\hat{Y}_{t-1}$ the estimated from this regression. Then regress
$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 \hat{Y}_{t-1} + V_t$$
where $\hat{Y}_{t-1}$ are estimated from the first stage regression.

(i)   How does this procedure remove the correlation between Y_{t-1} and V_t in the original model?   (5 marks)

(ii)   What good attributes do the coefficients of equation 2 depict that were lacking in equation 1?   (5 marks)

b) A modeler estimated tractor ownership as a function logarithm of income. Tractor ownership was a binary variable, Y=1 if a household owns a tractor, zero otherwise.

The following estimated equation was obtained.

$$\hat{L}_i = -2.72 + 0.35 lnincome$$
$$t \qquad (-3.35) \quad (4.05)$$
$$\chi^2(1\ df) = 16.81 \quad (p\ value) = 0.00)$$

Where $\hat{L}_i$ estimated logit and *lnincome* is log of income. The χ^2 measures the goodness of fit of the model.

(i) Interpret the estimated logit model (3 marks)

(ii) From the estimated logit model, how would you obtain the expression for the probability of tractor ownership? (2 marks)

(iii) What is the probability that a household with income of KES 20000 will own a tractor? (3 marks)

(iv) Comment on statistical significance of the estimated logit model. (2 marks)

QUESTION THREE (20 MARKS)

a) Using data from three agribusiness firms in Kiambu County, a researcher estimated the total cost function shown below.

$$TC_{it} = \beta_{1i} + \beta_2 Q_{1t} + \beta_3 PF_{it} + u_{it}$$

Where TC is total cost, Q is output, PF is fuel price and u is the error term.

The researcher obtained the results shown below.

Variable	Fixed Effects	Random Effects
Constant	- -	0.8499
		(0.0433)
Q	0.0557	0.0393
	(0.0042)	(0.0033)
PF	0.0351	0.0092
	(0.0051)	(0.0036)
p value for Hausman test	0.0000	

(i) Do the obtained slope coefficients make economic sense? (4 marks)

(ii) Is there a vast difference in the results produced by the two models? If so, what might account for these differences? (3 marks)

(iii) Which results should the researcher report? Explain. (3 marks)

b) Joshua, an agricultural researcher studied the effect of marital status and experience on wages in the agricultural sector. The results are shown below.

lnwage= 1.14 + 0.16*married* + 0.119*exper* - 0.0067*expersq*

se (0.0197) (0.0164) (0.01065) (0.0007305)

Using Breusch-Pagan test, heteroscedasticity was revealed to be present. To remedy this heteroscedasticity consistent (robust) standard errors were adopted and the following results were reported.

lnwage = 1.14 + 0.16*married* + 0. 12*exper* - 0.0067*expersq*

se (0.0197) (0. .0159)   (0.0101)   (0. 000646)

   i.   Why is heteroscedasticity of concern while conducting this research?   (4 marks)

  ii.   Has the problem of heteroscedasticity been corrected in the second model? How do you know?   (3 marks)

 iii.   Apart from robust standard error regression, what other solutions can be used to handle the problem heteroscedasticity?   (3 marks)

QUESTION FOUR (20 MARKS)

a)   Consider the following regression results

$$UN = 2.7491 + 1.1507D - 1.5294V - 0.8511DV$$
$$t = \quad (26.89) \quad (3.63) \qquad (\text{-}12.56) \qquad (\text{-}2.34)$$
$$R^2 = 0.9128$$

Where UN is unemployment rate, V is job vacancy rate in agricultural sector, $D = 1$ for period beginning 2000, $D = 0$ for period before 2000.

(i) What is your *a priori* expectation about the relationship between unemployment rate and vacancy rates in the agricultural rate?   (2 marks)

(ii) Holding job vacancy rate constant, what is the average unemployment rate in the period beginning 2000? Is it statistically different from the period before 2000? How do you know?   (3 marks)

(iii) What could be the implication of not including DV in the regression model?

          (2 marks)

(iv) Are the slopes in the previous pre- and-post 2000 statistically different? How do you know? Show your working.   (3 marks)

b) Despite presence of endogeneity, an econometrics student estimated a model using Ordinary Least Square. Critique this student's results. What other option(s) to estimate the model could you have suggested? Justify each.   (10 marks)

QUESTION FIVE (20 MARKS)

(a)   Explain whether each of the following statements is true or false.   (10 marks)

      (i) A high R^2 implies no omitted variables bias.

      (ii) A high p-value indicates that we cannot reject the null hypothesis.

      (iii) In a regression model with no explanatory variables the R^2 is equal to 0.

      (iv) Ceteris paribus, the higher the VIF is, the larger the variances of OLS estimators.

      (v) Autocorrelation is a common problem in time series data.

(b)   Covid-19 pandemic has affected small and medium enterprises firms in the Kenyan agricultural sector negatively. Explain the methodology of econometrics in this kind of study.

          (10 marks)