THE CATHOLIC UNIVERSITY OF EASTERN AFRICA

A. M. E. C. E. A

P.O. Box 62157
00200 Nairobi - KENYA
Telephone: 891601-6
Fax: 254-20-891084
E-mail:academics@cuea.edu

MAIN EXAMINATION

AUGUST - DECEMBER 2016 TRIMESTER

FACULTY OF ARTS AND SOCIAL SCIENCES

DEPARTMENT OF SOCIAL SCIENCES

REGULAR PROGRAMME

SEC 409: ADVANCED ECONOMETRICS

Date: DECEMBER 2016	**Duration: 2 Hours**
INSTRUCTIONS: Answer Question ONE and ANY other TWO Questions	

Q1. a) Consider the error term $\varepsilon = Y - X\beta$. Show that the matrix derivation of the β is $\beta = (X'X)^{-1}X'Y$ and prove that the OLS estimators are unbiased. **(8 marks)**

b) Consider the linear regression model

$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$

Where $\varepsilon_t \sim N(0, \sigma^2)$

You are given the following observations

Y_t	X_t
5	3
2	2
3	1
2	-1
-2	0

i) Find the Least Squares Estimates using the matrix approach **(5 marks)**
ii) Find the estimated variance of the least squares estimator of β_2 **(3 marks)**
iii) Test the significance of β_2 **(2 marks)**

c) In a two variable regression model, what assumptions do make about the stochastic variable in order to use the ordinary least squares (OLS) econometric method to estimate its parameters **(6 marks)**

d) With use of examples, distinguish between an economic model and an econometric model. **(6 marks)**

Q2. One school of thought alleges that the number of packets of cigarettes bought by smokers is a direct function of income. The table below provides time series data over a ten year period on annual averages of the number of packets bought and income levels.

Packets of Cigarettes	160	170	190	200	210	220	240	250	260	300
Income ($)	220	260	270	290	260	230	280	250	240	300

a) Prove the allegation wrong or right by estimating the following linear regression model

$$Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t$$

Where Y_t is number of packets of cigarettes, and X_t is income level **(10arks)**

b) Suppose the standard errors of the estimated coefficients are found to be

$S(\alpha_0) = 150$

$S(\alpha_1) = 0.08$

Test the hypothesis that α_i; i = 0, 1 are not different from zero at the 5% level of significance. Interpret your results graphically. **(5 marks)**

c) Test the overall significance of the regression model. **(3 marks)**

d) Find and interpret the coefficient of determination **(2 marks)**

Q3. a) Briefly explain how you can estimate and interpret a model with a set of two dummy variables, say gender and race as explanatory variables. **(6 marks)**

b) Explain briefly the following:

i) Gauss Markov Theorem **(2 marks)**

ii) Lagged variables **(2 marks)**

iii) Simultaneous Equations Bias **(2 marks)**

iv) Non-stationary series **(3 marks)**

c) Given the following set of simultaneous equations

$$P = \alpha_1 R + \varepsilon_1$$

$$Q = \alpha_2 R + \beta_1 X + \varepsilon_2$$

$$R = \alpha_3 P + \alpha_4 Q + \beta_2 Y + \varepsilon_3$$

Show whether the equations are identified, over identified or not identified **(5 marks)**

Q4. a) An investigator estimated the parameters in the equation

$$\ln Y_t = \alpha + \beta \ln X_t + u_t$$

by Ordinary Least Squares using quarterly observations for 1972 to 1984 inclusive. This resulted in a residual sum of squares (RSS) of 0.78. When 3 dummy variables representing the first 3 quarters of the year were added to the equation the RSS fell to 0.56. Test for the presence of seasonality. **(10 marks)**

b) Briefly demonstrate your understanding of Two-stage least squares method. **(5 marks)**

c) Differentiate between autoregressive and distributed lagged models and indicate the problems associated with lagging. **(5 marks)**

Q5. a) In econometric modeling, the omission of relevant independent variables has relatively more serious consequences than the inclusion of irrelevant independent variables. Explain (also providing formal proof) the theoretical basis for this statement. **(8 marks)**

b) Explain how you would carry out a cointegration test using the Engle-Granger procedure. What are the main defects of the procedure? **(8 marks)**

c) Briefly explain the criteria you would use to carry out identification of equations to be estimated in simultaneous equations system. **(4 marks)**

END