(a) Write the regression relating Log (W) to EDU for males and females separately.

(b) The returns to education are measured by the percentage increase in wages due to an extra year of education. Find the returns to education, for males and females.

(c) Interpret the coefficient of interactive dummy.

(d) Is the difference between returns to education for males and females statistically significant at 5% level of significance?

6. Explain the followings :

(a) Explain the problem of Multicollinearity? How can it be solved?

(b) Explain the problem of Autocorrelation? How can it be solved?

(c) Akaike information criterion (AIC).

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 5105 **H**

Unique Paper Code : 2922072401

Name of the Paper : Basics of Econometrics

Name of the Course : **Bachelor of Business administration (Financial Investment Analysis) [BBA (FIA)]**

Semester : IV

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question **1** is compulsory.
3. Attempt any **four** questions from question **2** - question **6**.
4. **All** questions carry equal weightage.
5. Use of simple calculator is allowed.

1. Explain the following with appropriate examples :

   (a) Explain the different functional form of regression models.

   (b) Dummy Variable Regression

   (c) Interaction Variable

   (d) Logistic regression

   (e) P-value

2. (a) Explain the problem of endogeneity? How can it be solved?

   (b) Consider the following model:

$$Yi = \alpha 0 + \alpha 1 D1 + \alpha 2\, D2 + \alpha 3\, D3 + \alpha 4\, D1\, *X + \beta 1\, Xi + Ui$$

Where Y = annual salary of college professor depending upon region of residence of professor

D1 = dummy variable for region, 1 for North; 0 otherwise

D2 = dummy variable for region, 1 for South; 0 otherwise

D3 = dummy variable for region, 1 for East; 0 otherwise

X*D1 = Interactive Dummy

Xi = years of teaching experience

An interaction term; AGE * Female is included in the regression. Give an interpretation of the coefficient on this variable. Does this result add to our understanding of the gender differences in the income-age relationship? How would you now test the null hypothesis of no gender differences?

5. Using data for 500 individuals the following model of wage determination was estimated :

$$\text{Log } (W) = \beta_0 + \beta_1\, D_i + \beta_2\, EDU_i + \beta_3\, (D_i * EDU_i) + u_i$$

Where,

W = Daily wages in rupees

D = Dummy variable for gender, D=1 for females and 0 for males

EDU = Years of education

D*EDU = Interactive dummy

The table below gives the estimated regression coefficients and their standard errors :

Variables	Estimated coefficients	Standard errors
C	0.40000	0.1290
D	-0.23000	0.1780
EDU	0.0800	0.0090
D*EDU	-0.006	0.0151

The estimation sample is: 1 to 12237

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-125862.	9347.	-13.5	0.000	0.0146
age	10373.8	316.3	32.8	0.000	0.0808
sigma	242135	RSS		7.17328365e+014	
R^2	0.0808074	F(1,12235) =	1076	[0.000]**	
log-likelihood	-169068	DW		1.1	
no. of observations	12237	no. of parameters		2	
mean(inc)	172161	var(inc)		6.3773e+010	

Review the estimated coefficients. What is the precise interpretation of the coefficient X? Test the hypothesis that the relation between age and income is the same for men and women. Suppose the result looks like the one given below:

The estimation sample is: 1 to 12237

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-190025.	1.273e+004	-14.9	0.000	0.0179
age	13676.2	434.4	31.5	0.000	0.0749
fem	124028.	1.849e+004	6.71	0.000	0.0037
fem*age	-6504.33	625.3	-10.4	0.000	0.0088
sigma	239051	RSS		6.99060616e+014	
R^2	0.104216	F(3,12233) =	474.4	[0.000]**	
log-likelihood	-168910	DW		1.13	
no. of observations	12237	no. of parameters		4	
mean(inc)	172161	var(inc)		6.3773e+010	

Salary	Coeff	p>|t|
D_1	15400	0.000
D_2	12300	0.000
D_3	10240	0.091
X_i	3000	0.000
X_i*D_1	1000	0.000
α_0	-5000	0.043

Write the regression equation corresponding to each of the following: annual Salary of a professor who is residing in North, Salary of a professor who is residing in South. Salary of a professor who is residing in East, Salary of a professor who is residing in West. How would you infer the significance of coefficient of D3 using the p-value? What impact does it have on the model? How would the model look like if we add the dummy variable for 'WEST? Write the equation for it and explain the 'dummy variable trap' in this model's context. Interpret the Coeff of interactive dummy.

3. A researcher, using a sample of 2000 individuals, is investigating how the probability of a respondent obtaining a bachelor's degree from a three-year college is related to the respondent's score on ASVXYZ. 30 percent of the respondents earned bachelor's degrees. ASVXYZ ranged from 20 to 70, with mean value 52, and most scores were in the

range 45 to 65. Defining a variable DEGR to be equal to 1 if the respondent has a bachelor's degree (or higher degree) and 0 otherwise, the researcher fitted the OLS regression (standard errors in parentheses):

$$\underset{(0.000)}{DEGR = -0.863 + 0.022\ ASV\ XYZ}$$

She also fitted the following logit regression:

$$\underset{(0.004)}{Z = -12.103 + 0.179\ ASV\ XYZ}$$

Where Z is the variable in the logit function.

(a) Give an interpretation of the OLS regression, and explain why OLS is not a satisfactory Estimation method for this kind of model.

(b) What does the slope coefficient in Logit model measures? Interpret the estimated logit model. From the estimated logit model, how would you obtain the expression for the probability of earning a bachelor's degree?

(c) What is the probability that an individual with a score of 40 on ASV XYZ will have a bachelor's degree? What is the rate of change of probability at the score of 55?

4. The relation between two variables age and income is investigated if it is same for men and women. You have the following relations:

$$Income_i = \alpha_1 + \alpha_2\ Age_i + u_i \qquad \text{----- (1), i=1,2...n}$$

$$Income_i = \beta_1 + \beta_2\ Age_i + \lambda\ Fem_i + v_i \qquad \text{---- (2), i=1,2,...,n}$$

Where u_i and v_i are stochastic error terms. Fem_i is female and takes the value 1 if the observation unit i is of female and zero for a male. Below is the regression output from OLS applied to equation (1) and (2) :

The estimation sample is: 1 to 12237

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-100578.	9431.	-10.7	0.000	0.0092
age	10536.8	313.8	33.6	0.000	0.0844
fem	-62966.0	4349.	-14.5	0.000	0.0168
sigma	240096	RSS	7.05242974e+014		
R^2	0.0962938	F(2,12234) =	651.8	[0.000]**	
log-likelihood	-168964	DW	1.12		
no. of observations	12237	no. of parameters	3		
mean(inc)	172161	var(inc)	6.3773e+010		

P.T.O.