(This question paper contains printed pages)

Roll Number:

Serial Number of question paper:

Unique Paper Code: **12277502_NC**

Name of the Paper: **Applied Econometrics**

Name of the Course: **B.A. (Honours) Economics CBCS**

Semester **Semester V**

Duration: **3 hours**

Maximum Marks: **75**

<u>Instructions for Candidates</u>

1. **Write your Roll No. on the top immediately on receipt of this question paper.**
2. **Answers may be written in English or Hindi but the same medium should be used throughout the paper.**
3. **The question paper consists of six questions. Answer any *four* questions.**
4. **All questions carry equal marks.**
5. **Use of simple non programmable calculator is allowed.**
6. **Statistical tables are attached for your reference.**

<u>परीक्षार्थियों हेतु अनुदेश</u>

1. इस प्रश्न-पत्र के प्राप्त होते ही तुरन्त सबसे ऊपर अपना रोल नम्बर लिखिए।
2. उत्तर अंग्रेजी या हिन्दी में दिए जा सकते हैं परन्तु पूरे पेपर में एक ही माध्यम का उपयोग किया जाना चाहिए।
3. इस प्रश्न-पत्र में **छः** प्रश्न हैं। किन्हीं **चार** प्रश्नों के उत्तर दीजिए।
4. सभी प्रश्नों के बराबर अंक हैं।
5. साधारण अप्रोग्रामनीय कैलकुलेटर का प्रयोग मान्य है।
6. आपके सन्दर्भ हेतु सांख्यिकीय सारिणियाँ संलग्न हैं।

Q1. (a) Consider the following regression model for wages:

$$lnW_i = \alpha_1 + \alpha_2 EDU_i + \alpha_3 EXP_i + \alpha_4 D_i + u_i \qquad (1)$$

where W_i=wages, EDU_i= number of years of education, EXP_i= number of years of experience and D_i=1 if female and 0 otherwise.

(i) The regression model was estimated using data on 125 observations. To test for heteroscedasticity, White's Test for heteroscedasticity was conducted. Describe the test procedure and write down the most general form of the auxiliary equation (which includes squares and cross products) for the variance of u_i. If $R^2 = 0.2251$ for this auxiliary regression, conduct the test at 1% level of significance. Do you find evidence of heteroscedasticity?

(ii) Using the estimated residuals from equation (1), the LM test for omitted variables was conducted to see if EXP_i^2 should be added to the regression model. Describe the test procedure. If $R^2 = 0.0389$ for the auxiliary regression of the LM test, conduct the test at 1% level of significance and state your conclusion.

Q1. (b) Consider the structural equation in the context of omitted variables:

$$Y_1 = \beta_1 + \beta_2 Y_2 + \beta_3 Z_3 + \beta_4 Z_4 + u$$

where Y_1 and Y_2 are endogenous variables and Z_3 and Z_4 are exogenous variables and u is the disturbance term and $E(u) = 0$.

(i) 'If the structural equation is estimated using OLS then the estimators will be biased and inconsistent'. Explain.

(ii) In the context of omitted variables explain the statement-' A proxy variable for the omitted variable will not be a good instrumental variable (IV).'

(iii) Let Z_2 be an IV for Y_2. What conditions must Z_2 satisfy to be used as an IV for Y_2? Explain how the IV estimators can be obtained.

(9.5+9.25)

Q2. (a) Consider the infinite lag distributed lag model:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots\dots + u_t$$

(i) Show that in the Koyck model, the assumption regarding the β coefficients results in an exponentially weighted average of current and past values of the variable X. Show that the sum of the coefficients i.e. the long run multiplier is finite. Write down the estimable form of the model as an autoregressive model using the Koyck transformation.

(ii) Using annual data on per capita consumption expenditure in rupees (C) and per capita disposable income in rupees (DI) for 1968 to 2018, the following results were obtained after estimation of the autoregressive model:

(iii)

$$\hat{C}_t = 1239.35 + 0.4501 DI_t + 0.5502 C_{t-1}$$

se= (1998.53) (0.0821) (0.1244)

d=1.1145 n= 50

Using the Koyck model write down the estimated short run and long run MPC and interpret them. What is the estimated long run equation?

(iv) If the results are interpreted using the adaptive expectations hypothesis, what is the estimated coefficient of expectation? How will you interpret it?

(v) Conduct an appropriate test for serial correlation at 5% level of significance.

Q2. (b) Consider the following regression model:

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

where $Y_i = 1$ if the loan application is approved and 0 if it is not approved and X_i is monthly income of the applicant in thousand rupees.

(i) Explain the linear probability model (LPM) approach to estimation of the model using OLS. What are the problems with the LPM? Show that the disturbances are heteroscedastic and explain how the method of weighted least squares can be applied.

(ii) The LPM model was estimated using data on 1050 loan applicants. After dropping observations where the fitted values lie outside the unit interval, the results are given below:

$$\widehat{Y_i} = -2.3322 + 0.1210\ X_i$$
$$\text{se} = \quad (0.1112) \quad (0.0222)$$

Figures in brackets are heteroscedasticity-robust standard errors. Interpret the estimated slope coefficient from the above regression. For an income level of 20,000 rupees, interpret the estimated value of the dependent variable.

(9.5+9.25)

Q3. (a) A researcher formulates the following simultaneous equation model:

$$P = \beta_1 + \beta_2 W + \epsilon_1$$

$$W = \alpha_1 + \alpha_2 P + \alpha_3 N + \epsilon_2$$

where P is the annual rate of growth of prices, W is the annual rate of growth of wages, N is the rate of unemployment. P and W are endogenous variables while N is exogenous. The disturbance terms ϵ_1 and ϵ_2 are assumed to be uncorrelated and

$$E(\epsilon_1) = 0\ ;\ V(\epsilon_1) = \sigma_{\epsilon 1}^2\ ;\ E(\epsilon_2) = 0\ ;\ V(\epsilon_2) = \sigma_{\epsilon 2}^2\ .$$

(i) Derive the reduced form equations for P and W.

(ii) Show that $cov(W, \epsilon_1) \neq 0$. Hence show that the OLS estimator for β_2 is inconsistent. What is the restriction for the simultaneous equation bias to be positive if it is assumed that $\alpha_2 > 0$?

Q3. (b) Using quarterly data on GDP and M (money supply) from 2002-Q1 to 2010-Q4, the Granger test is conducted to examine the nature of causality between the two variables.

(i) Write down the relevant pair of regression equations necessary to conduct the F test for causality. Describe the test procedure for testing $M \rightarrow GDP.$

(ii) The table below presents the results of the F test for different quarterly lags.

Direction of causality	Number of lags	F value
$GDP \rightarrow M$	2	10.53
$M \rightarrow GDP$	2	3.93
$GDP \rightarrow M$	4	4.96
$M \rightarrow GDP$	4	2.36
$GDP \rightarrow M$	6	3.76
$M \rightarrow GDP$	6	1.50

For each of the lags, conduct the F test for Granger causality at 5% level of significance stating the null hypothesis clearly. Comment on the results for the different lags and in this context explain the use of Akaike or Schwarz information criteria.

(9.5+9.25)

Q4. (a) Consider the following panel data regression model for 5 mobile phone manufacturing companies over 20 years.

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it}$$

where Y represents cost, X_2 represents index of raw material costs and X_3 represents wages.

(i) What are the consequences of using Pooled OLS regression model if the cost structure of firms depends on the nature of ownership/management which is time invariant and is not directly observable?

(ii) Rewrite the model to show how heterogeneity can be taken into account in the one-way Fixed Effects Least Squares Dummy Variable (LSDV) model.

(iii) Outline the test procedure to choose between Pooled and fixed effects LSDV models. If R^2 for the two models are 0.9765 and 0.9852 respectively, conduct the test at 1% level of significance and state your conclusion.

Q4. (b) Consider the 3-variable regression model $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$. Assume u_i satisfy the classical OLS assumptions. Given that X_{2i} and X_{3i} are uncorrelated and

$$\overline{X_2} = \overline{X_3} = 0 \ ; \ \bar{Y} = 5 \ ; \sum X_{2i}^2 = 100 \ ; \sum X_{3i}^2 = 50; n = 10;$$

$$\sum X_{2i} Y_i = 200 \ ; \sum X_{3i} Y_i = 50 \ ; \boldsymbol{y}'\boldsymbol{y} = 800$$

(i) Using matrix algebra find the estimated $\hat{\boldsymbol{\beta}}$ vector .

(ii) What is the calculated value of R^2? Write down the estimated standard error of the regression.

(iii) Conduct the F test at 5% level of significance.

(9.5+9.25)

Q5. (a) Using grouped data on number of families having medical insurance at different income (X_i) levels, the following logit model was estimated:

$$L_i = ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_i + u_i$$

(i) The model was estimated using WLS and after appropriate transformation the estimated logits for two different income levels are given below:

Income(in Rs 1000)	Estimated Logit
20	-0.3478
32	0.2352

Calculate the estimated probabilities for the two income classes and interpret them. Can the estimated probabilities be less than 0 or greater than 1? Explain.

(ii) Show that $\frac{dP_i}{dX_i} = \beta_2 P_i(1 - P_i)$. What is the estimated change in probabilities for the two income levels given above if estimated $\beta_2 = 0.1891$?

Q5. (b) The following savings-income relationship is considered by a researcher:

$$Savings_t = \beta_1 + \beta_2 Income_t + u_t$$

Using data from1991-2018, the researcher wants to test if there was a structural break in the savings-income relationship in 2008. The results from the regression for the pooled data and the two sub-periods is given below.

	OLS using pooled data(N=28)	OLS for first sub-period(N=18)	OLS for second sub-period(N=10)
ESS	26040		
RSS		2345	
TSS			5150
R^2	0.8		0.7

where ESS is explained sum of squares, RSS is the residual sum of squares and TSS is the total sum of squares.

(i) Conduct the Chow Test at 1% level of significance and state your conclusion.

(ii) If the date of the structural break is uncertain, explain how Recursive Least Squares can be used to check for the point of structural break.

(9.5+9.25)

Q6. (a) A researcher collects data on hourly compensation rate in manufacturing sector (Y_{it}, in rupees) and unemployment rate (X_{it}, in percentage) for 4 countries for the period 1990-2016:

$$Y_{it} = \beta_1 + \beta_2 X_{it} + u_{it}$$

(i) Describe how you can apply the Random Effects Model (REM) to the panel data. Clearly state the assumptions made by the REM.

(ii) Show that the error term in the REM is homoscedastic.

(iii) Show that the correlation between error terms at two different times remains the same, no matter how far apart the two time periods are and that the correlation structure remains the same for all cross-sectional units.

(iv) To choose between the Fixed Effects and Random Effects model, the Hausman Test was conducted. If the value of this test statistic is 7.9821 specify the null and alternate hypothesis clearly and state the conclusion. Choose 1% level of significance.

Q6. (b) Consider the regression model relating savings (S) and income (I) of households.

$$S_i^* = \alpha + \beta\, I_i + u_i$$

Suppose data on actual savings S_i^* is not available but data on reported savings of households i.e. S_i is available.

$$S_i = S_i^* + \epsilon_i$$

where ϵ_i is the measurement error. Given that $E(u_i) = 0$; $E(\epsilon_i) = 0$; and that $cov(I_i, u_i) = 0$; $cov(\ \epsilon_i, u_i) = 0$.

(i) If the measurement errors ϵ_i are uncorrelated with the explanatory variable I_i, show that the OLS estimator of the slope coefficient is unbiased.

(ii) Suppose the measurement errors ϵ_i vary systematically with I_i, so that ϵ_i and I_i, are correlated, is the OLS estimator of β unbiased?

(iii) If $cov(\ \epsilon_i, I_i) = 0$ but $S_i = S_i^* + c$ where c is some constant, then is the OLS estimator $\hat{\beta}$ unbiased? Explain.

(9.5+9.25)