2021

ECONOMICS — HONOURS — PRACTICAL

Paper : DSE-A-1

(Applied Econometrics)

Full Marks : 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Create a log file (.smcl format), put your registration number and roll number, without any blank space, as file name. If your Registration Number is XXX-XXXX-XXX-XX and Roll Number is XXXXXX-XXXX your file name will be REG XXX-XXXX-XXXX-XX ROLL XXXXXX-XXX. If there is no log file your answer-script will not be evaluated.

Answer any three questions.

(Using STATA or R)

Group - A

- 1. Dataset Question 1 represents sex (1 for male and 2 for female), age, wage and size of household (hhd_size) of 16853 respondents. Use the dataset Question 1 to
 - (a) Calculate the number of male and female respondents.
 - (b) Calculate the mean, variance, range, minimum value and maximum value of wage and age.
 - (c) Draw a histogram for wage by plotting percentage of wage on the vertical axis.
 - (d) Create a variable to know the number of respondents who have the household size (hhd_size) of at least 7 persons.
 - (e) Get the non statistical description of the dataset.
- 2. Dataset Question 2 represents wage, age and general education (gen_edu) of 16853 respondents. Use the dataset to answer the following questions.
 - (a) Create a variable (lnwage) to store natural log values of wage.
 - (b) Estimate the linear regression model where 'lnwage' is explained by age and general education of the respondents. Interpret the results.
 - (c) Are the coefficients of age and gen edu individually significant?
 - (d) What do you say about the overall significance of the estimated model?
 - (e) Conduct a formal test for heteroscedasticity in the residuals and interpret the result.

1+(1+2)+2+2+(1+1)

2+2+2+2+2

Please Turn Over

V(5th Sm.)-Economics-H/Pr./DSE-A-1/CBCS (2)

- **3.** Dataset Question 3 represents natural log of wage (lwage), age, general education (gen_edu), technical education (tech-edu) of 16853 respondents. Use the dataset to answer the following questions.
 - (a) Estimate a linear regression model where lwage is explained by age, general education and technical education of the respondents.
 - (b) Get the 95% confidence interval for general education.
 - (c) Compute the t statistic for the coefficient of wage from the estimated result. Compare it with the displayed value of t statistic.
 - (d) Compute the R^2 (*R* squared) from the displayed values of sum of squares. Compare this value with the displayed value of *R*-squared.
 - (e) Create a series of predicted values of 'lwage'. Compute the predicted value of 'lwage' for a person aged 25 years and with general education 10 and technical education 2.

1+1+2+3+(1+2)

Group - B

- 4. Dataset Question 4 represents annual data of private final consumption expenditure (pfce) and gross fixed capital formation (gfcf) of a country for the time period (1950 to 2012). Use the dataset to
 - (a) Set the data as yearly time series data.
 - (b) Get the lineplot of 'pfce' for the period 1950 to 2012.
 - (c) Get a series two period lagged values of 'pfce'.
 - (d) Create a series of 'year', 'pfce', and two period lagged values of pfce for first 10 years.
 - (e) Create a series of natural log of 'pfce'.
 - (f) Create a variable to contain the time period, starting 1950 as 1 and ending 2012 as 63.
 - (g) Estimate the time trend of 'pfce' by estimating a regression of 'pfce' taking the variable created above (4f) as regressor. 1+2+1+2+1+1+2

Group - C

- 5. Dataset Question 5 represents annual data of GDP growth rate (gdp_growth) and employment population ratio (employment_population_ratio) of 10 countries for the time period (1991 to 2018). Use this dataset to answer the following questions.
 - (a) Encode country name.
 - (b) Set the data as panel data taking time variable as year.
 - (c) Run the Pooled Regression model assuming GDP growth is explained by employment-populationratio. Interpret the result.
 - (d) Estimate the above model by applying fixed effect error component model assuming significant differences among countries but no significant temporal effect. Interpret the result.
 - (e) State whether pooled regression model is better than the fixed effect error component model on the basis of your results. 1+1+2+(2+1)+(1+2)