_____

FACULTY/COLLEGE	College of Business and Economics
SCHOOL	School of Economics
CAMPUS	APK
MODULE NAME	Econometrics 3A
MODULE CODE	ECM03A3
SEMESTER	First
ASSESSMENT OPPORTUNITY, MONTH AND YEAR	Summative Assessment Opportunity A

ASSESSMENT DATE		SESSION	08:30 – 11:30
ASSESSOR(S)	Dr M Pretorius and Dr C Saba		
MODERATOR(S)	Ms B Thobejane (UJ) and Ms N Cattaneo (Rhodes University)		
DURATION	3 hours (180 min)	TOTAL MARKS	150

NUMBER OF PAGES OF QUESTION PAPER (Including cover page)	5

INSTRUCTIONS:

_____
- This is a closed-book assessment.
- There are three questions (A to C). **Answer all questions.**
- Read the questions carefully and only answer what is required.
- Number your answers clearly and correctly as per the question paper.
- Write neatly and legibly on both sides of the paper in the answer book, starting on the first page.
- Calculators may be used.
- Please round your final answers off to 2 decimal places.

_____

_____

_____

QUESTION A [50 MARKS]

1. State whether the following statements are true or false. (20)
   a. The standard errors of OLS estimators measure their precision. **True**
   b. If the correlation between two variables is zero, there is no relationship between the two variables whatsoever. **False**
   c. The t-test of significance requires that the sampling distributions of estimators follow the normal distribution. **True**
   d. Even though the disturbance term in the Classical Linear Regression Model (CLRM) is not normally distributed, the OLS estimators are still unbiased. **True**
   e. If there is no intercept in the regression model, the estimated u_i will not sum to zero. **True**
   f. The conditional and unconditional means of a random variable are the same things. **False**
   g. In the case of high multicollinearity, it is not possible to assess the individual significance of one or more partial regression coefficients. **False**
   h. In the presence of heteroscedasticity, OLS estimators are biased and inefficient. **False**
   i. If residuals estimated from an OLS regression exhibit a systematic pattern, heteroscedasticity is present in the data. **False**
   j. The consequences of including irrelevant variables in the model are less serious than omitting relevant variables. **True**

2. Complete the missing parts in the following sentences. (6)
   a. In a regression model that contains an intercept, the sum of the residuals is always zero.
   b. The multiple coefficient of determination is computed by dividing SSR by SST.
   c. A type I error occurs when rejecting a true hypothesis.
   d. A/an indepedent variable is also known as an explanatory variable.
   e. According to the Gauss-Markov theorem, given the assumptions of the CLRM, the least-squares estimators, in the class of unbiased linear estimators, have minimum variance. That is, they are BLUE.
   f. A regression through the origin does not contain an explicit intercept term.

3. Explain what a multiple linear regression examines. (2)
   The relationship between one dependent and more than one independent variables

4. Give a brief outline of the nine assumptions that underpin the least-squares method (9).
   a. Linear regression model, or linear in the parameters.
   b. Fixed X values or X values independent of the error term. Here, this means we require zero covariance between Ui and each X variables. cov (Ui, X2i) = cov (Ui, X3i) = 0
   c. Zero mean value of disturbance Ui. E(Ui |X2i, X3i)=0 for each i
   d. Homoscedasticity or constant variance of Ui. var (Ui)= σ2
   e. No autocorrelation, or serial correlation, between the disturbances. cov (Ui, Uj) =0   i = j

_____

_____

f. The number of observations n must be greater than the number of parameters to be estimated, which is 3 in our current case.
g. There must be variation in the values of the X variables.
h. No exact collinearity between the X variables. No exact linear relationship between X2 and X3.
I. There is no specification bias. The model is correctly specified.

5. Explain the difference between a dummy variable and a dummy variable trap. (4)

   Dummy variable are variables that assume the values of 0 or 1 where 1 indicates the presence of a certain attribute and 0 indicates the absence of that attribute while dummy variable trap is the situation of perfect collinearity or perfect multicollinearity, that is, if there is more than one exact relationship among the variables.

6. Why is the number of dummy variables entered into a regression model always equal to the number of groups (m) minus 1, i.e. (m-1)? (2)

   To avoid the situation of perfect multicollinearity

7. Explain what is meant with ANCOVA? (3)

   Regression models containing an admixture of quantitative and qualitative variables are called analysis of covariance (ANCOVA) models. ANCOVA models are an extension of the ANOVA models in that they provide a method of statistically controlling the effects of quantitative regressors, called covariates or control variables, in a model that includes both quantitative and qualitative, or dummy, regressors.

8. According to Montgomery and Peck, multicollinearity may be due to four factors, briefly highlight these factors. (4)

   a. The data collection method employed. For example, sampling over a limited range of the values taken by the regressors in the population.
   b. Constraints on the model or in the population being sampled. For example, in the regression of electricity consumption on income (X2) and house size (X3) there is a physical constraint in the population in that families with higher incomes generally have larger homes than families with lower incomes.
   c. Model specification. For example, adding polynomial terms to a regression model, especially when the range of the X variable is small.
   d. An overdetermined model. This happens when the model has more explanatory variables than the number of observations. This could happen in medical research where there may be a small number of patients about whom information is collected on a large number of variables.

QUESTION B   **[50 MARKS]**

Many countries use the revenue from their exports to finance their imports (i.e. imports depend on exports). Make use of the data provided in the Excel file "Question B" to answer the following questions.

_____

_____

1. Consider the following table. Estimate the necessary regression models in Eviews and complete the missing information in the table (use letters a. to l. on your answer sheet). (24)

Name of Model	Specification of Model	Estimated equation	Interpretation of partial slope coefficient
a. Regression through the origin	$Imports = \beta_2 Exports + u_t$	b. Imports = 1.090143Exports	c. If Exports increase with one unit, Imports increase with 1.090143 units.
Regression on standardised variables	d. $Imports * = \beta_2 Exports * + u_t$ where $Imports * = \dfrac{Y_i - \bar{Y}}{S_Y}$	e. Imports_st = 0.951331Exports_st	f. If Exports increase with one standard deviation unit, Imports will increase with 0.951331 standard deviation units.
g. Double log / Log linear model	h. $\ln Imports = \ln \beta_1 + \beta_2 \ln Exports + u_i$	$\log(Imports) = -0.558177 + 1.059032 \log(Exports)$	i. If Exports increase with one percent, imports will increase with 1.059032 percent.
j. Reciprocal model	k. $Imports = \beta_1 + \beta_2 \left(\dfrac{1}{X_i}\right) + u_i$	l. Imports = 88644.2 − 1.57E+09 (1/Exports)	As exports increase indefinitely, imports will tend towards the intercept value of xxxxx (Please note that this value is not given here, values should be given in answers c, f and i)

2. Estimate the following regression: $Imports = \beta_1 + \beta_2 Exports$. Write down your regression results. (2)
Imports = -2052.820 + 1.132629Exports

_____

_____

3. Interpret the slope coefficient of the model in 2. Does this relationship make economic sense? Explain. (5)
If exports increases with one unit, imports will increase with 1.13 units, ceteris paribus. This positive relationship makes economic sense. As exports increase, more forex will flow into the country which can be used to fund imports from other countries.

4. Construct a 95% confidence interval for the slope coefficient. Show all your steps. (5)

$$\hat{\beta}_2 \pm t_{\frac{\alpha}{2}} \, se(\hat{\beta}_2)$$

$1.132629 \pm 1.97897(0.032686)$

1.132629 ± 0.064685

$[1.067944; 1.197314]$

5. Is the exports variable statistically significant? Explain by making use of the confidence interval you constructed in question 4. (3)
Yes it is. Zero is not included in the confidence interval.

6. Comment on the goodness of fit of this model. (3)
$R^2 = 0.905031$, this is a good model. 90.5% of the variation in imports is explained by exports.

7. Make use of an appropriate test and determine if autocorrelation is present in the model. Show all your steps. (3)
Durbin-Watson = 1.575860, which lies relatively close to two – no autocorrelation is expected.
OR
H0: No autocorrelation (Breusch-Godfrey)
Reject if p-value of Obs*R-squared < 0.05
p = 0.0181
Reject, the model suffers from autocorrelation

8. Make use of an appropriate test and determine if heteroscedasticity is present in the model. Show all your steps. (3)
White test / Breusch-Pagan
H0: No heteroscedasticity / Homoscedasticity
Reject if p-value of Obs*R-squared < 0.05
p = 0.008 (0.0023 for Breusch-Pagan)
Reject, the model suffers from heteroscedasticity.

9. Consider your answers in 7. and 8. How would you remedy the problems (if any) identified in these two questions? (2)
Autocorrelation and heteroscedasticity is present and therefore we can make use of the Newey-West method to adjust the standard errors.

QUESTION C   **[50 MARKS]**

Make use of the Excel file "Question C" to answer the following questions. The data ranges from 1950 to 2019.

| q_x is quantity of exports for South Africa (dependent variable) |
| pl_x is price of exports for South Africa |
| xr is the exchange rate for South Africa |
| dum is a dummy variable (dum = 0 before the 2008 financial crisis and dum = 1 after the 2008 financial crisis |

_____

_____

1. Obtain the descriptive statistics for the quantity of exports for South Africa. Give and interpret two measures of central tendency. (2)
   The average quantity of exports for the sample period was 687494.3.
   The middle value (when the data is ordered) of the quantity of exports for the sample period was 518546.0.

2. Is the quantity of exports variable normally distributed? Use the Jarque-Bera statistic to support your claim. Show all your steps. (4)
   H0: quantity of exports is normally distributed. H0: quantity of exports is not normally distributed.
   Jacque-Bera value is 6.89 with a probability of 0.03.
   Since the p-value is less than 0.05, we reject the null hypothesis and conclude that q_x does not have a normal distribution.

3. Plot the quantity of exports against the price of exports. Describe the relationship. (2)
   A positive to quadratic relationship is observed
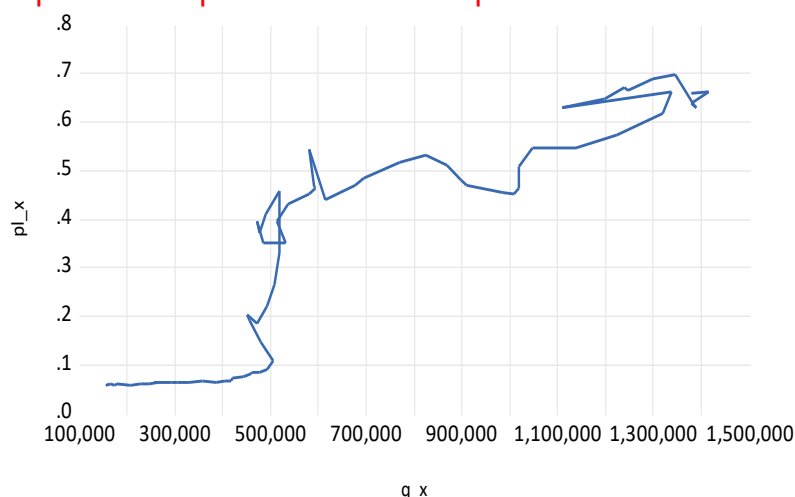
4. Specify the regression model to show the impact of the price of exports and the exchange rate on the quantity of exports. Estimate your regression and give your results. (4)
   q_x = b$_1$ + b$_2$*PL_X + b$_3$*XR +U$_1$
   q_x = 235652.1+ 664421.6*PL_X + 58339.21*XR

5. Give the interpretation of the partial slope coefficient of the price of exports. Does it make economic sense? (3)
   When the price of exports increases by one unit, the quantity of exports increases by 664,421 units, holding the exchange rate (or holding other factors constant).

6. Give the interpretation of the partial slope coefficient of the exchange rate. Does it make economic sense? (3)
   When the exchange rate increases (or rand depreciates) by one unit, the quantity of exports increases by 58,339 units, holding the exchange rate (or holding other factors constant).

7. Make use of the t-statistic to test for the individual statistical significance of the two independent variables. Show all your steps. (8)
   For price of exports
   H0: B2 is equal to zero. H0: B2 is not equal to zero.
   Computed t-statistic is 7.24. Critical value is 2 (rule of thumb).

_____

Conclusion: Since the computed t-value is greater than the critical t-value, reject the null hypothesis and conclude that b2 is statistically significant at 5 per cent level.

For exchange rate
H0: B3 is equal to zero. H0: B3 is not equal to zero.
Computed t-statistic is 11.26. Critical value is 2 (rule of thumb).
Conclusion: Since the computed t-value is greater than the critical t-value, reject the null hypothesis and conclude that b3 is statistically significant at 5 per cent level.

8. Make use of the F-statistic to test for the statistical significance of the overall model. Show all your steps. (4)
   H0: B2 = B3 is equal to zero. H0: Not all Bs are equal to zero.
   Computed F-statistic 448.09 with p-value of zero.
   Conclusion: Since the p-value of the F-stat is zero, we strongly reject the null hypothesis and conclude that the overall model is statistically significant at 1 per cent level.

9. Test whether there was a structural change in the regression in 2008 due to the financial crisis by making use of the Chow Test. Show all your steps. (4)
   H0: there was no structural change. H1: there was a structural change
   The F-statistic from Chow Breakpoint is 4.27 with a probability of 0.0082.
   Conclusion: We reject the null hypothesis since the prob-value of F-stat is zero, and less than 0.05. We conclude that there was a structural change during the 2008 financial crisis.

10. Estimate a new regression with the quantity of exports as the dependent variable, price of exports, the exchange rate, a differential intercept dummy (dum) and a multiplicative dummy (dum interacted with exchange rate only). Specify your regression first before reporting the estimated regression equation. (5)
    Q_X = b1 + b2*PL_X + b3*XR + b4 dum+ b5 dum*XR+ U1
    Q_X = 244597.6 + 514850.5*PL_X + 71046.59*XR + 433809.3*DUM - 43623.76*DUM*XR

11. Give the interpretation of the differential intercept dummy (dum) coefficient. (2)
    During the period after the 2008 financial crisis, the quantity of exports increased (alternatively, the intercept increased by 433,809.3 units) holding other variables constant.

12. Give the interpretation of the differential slope dummy (dum*xr) coefficient. (2)
    During the period after the 2008 financial crisis, the differential slope with exchange rate decreased by 43,623 units, holding other variables constant.

13. Use the correlation matrix to test for multicollinearity between the exchange rate, the price of exports and the dummy variable. What is your conclusion? Make sure to report all your steps and results from the matrix. (4)
    Correlation matrix results

	PL_X	XR	DUM
PL_X	1	0.81	0.62
XR	0.81	1	0.78
DUM	0.62	0.78	1

    Conclusion: The price of exports and the exchange rate seem to be highly related, with a correlation of 0.80. Hence, there is possible collinearity problem.

14. How would you remedy the problem if multicollinearity is present in a model? (3)

I would keep the model as is to avoid specification error.

-----o0o-----