UNIVERSITY OF WESTMINSTER
WESTMINSTER BUSINESS SCHOOL
DATA ANALYSIS 7ECON001W

SET Fall 2022/23

Submission Instructions

1. Only an electronic copy should be handed-in via the blackboard site of the module by **13:00** (UK time) on **10 JANUARY 2023**. This copy will automatically be scanned through a text matching system (designed to check for possible plagiarism).

2. To avoid mark penalisation for late submission ensure you submit your coursework in time according to the deadline above. See module handbook for details on mark penalisation for late submission and step by step on-line submission instructions.

3. The name and the registration number of the student should be clearly shown on the first page of the assignment.

General Instructions

1. This assignment is **individual.**

2. Provide and explain all calculations and relevant EViews outputs.

3. Use a significance level of 5% for all tests.

4. Presentation is worth 10 marks of the total grade —note that a good presentation requires clear and concise answers avoiding redundant information.

5. Word limit: 4,000. The rule of 10% extra words apply. Footnotes, appendices and references do not count.

6. This assignment is worth 60% of the total module mark. A qualifying mark of a minimum of 40% is required in this piece of assessment to pass the module.

Exercise 1

The Excel file "Data SET.xlsx" contains closing prices (in \$) $\{P_t\}_{t=1}^{T}$ of stocks constituents of the Standard and Poor's 500 (S&P500) index. The time period comprises from December 10, 1999 to December 10, 2019. All data series were downloaded from Datastream.
For the series randomly allocated to you (see surnames in first row of the Excel sheet) answer the following questions:

1. Apply the Box and Jenkins methodology to select an appropriate specification for the conditional mean of the series. Include plots of the level of the series in each step as well as the correlograms to show and illustrate your answers.

2. Test for ARCH effects in the residuals of your selected conditional mean model. Include and interpret the corresponding EViews output.

3. What is the difference between heteroscedasticity and ARCH? Use an example to explain your answer.

4. Answer both parts:

   (a) Define ARIMA-GARCH-type specifications you know that may be appropriate to explain the conditional variance of your chosen series.

   (b) Fit the ARIMA-GARCH-type models defined above. Interpret the results pointing to the main differences between the estimated models. Provide the corresponding EViews outputs.

5. Using the estimation results in question 4, answer both parts:

   (a) Compare the goodness-of-fit of the models in the previous question. Select the most appropriate specification. Explain your answer using appropriate tests and statistics.

   (b) Does your series returns volatility present the "leverage effect"? Explain your answer using your estimation results in the previous question.

6. Using your selected model in question 5, calculate a one-day-ahead forecast of the conditional variance of the returns using an appropriate GARCH-type model under

the normal distribution, as well as a 95% confidence interval for the conditional mean forecast.

7. Define VaR and calculate a one-day-head 1% Value-at-Risk forecast under Normal errors.

8. Explain the implications for regulatory capital of underestimating or overestimating VaR.

9. Do the residuals of the model estimated in question 4 come for a normal distribution? Perform an appropriate test to answer the question.

Notes: (1) Describe all tests in detail, step by step. (2) For each question provide the relevant EViews outputs or/and plots.

Exercise 2

Consider the following model:

$$vote_t = \beta_0 + \beta_1 g\,\mathrm{row}\, th_t + \beta_2 goodnews_t + \beta_3 war_t + u_t$$

where

vote	= percentage share of the popular vote won by the incumbent[1] party.
growth	= annual growth rate in real per capita GDP in the first 15 quarters of an administration.
goodnews	= number of quarters in the first 15 quarters of the administration in which the growth rate of real per capita GDP is greater than 3.2% at an annual rate.
war	= dummy variable equal to 1 for the elections of 1920, 1944 and 1948, and zero otherwise.

1. Perform a Goldfeld-Quandt (G-Q) test for heteroscedasticity in the residuals of the model above.

2. Why is the White test preferred to *the* Goldfeld-Quandt test, for heteroscedasticity. Explain your answer.

3. Assuming that there is autocorrelation in the residuals of the model:

   (a) Use EViews to perform the Cochrane-Orcutt (C-O) procedure to resolve autocorrelation of order 1 and interpret the results.

   (b) Explain analytically step by step the C-O procedure to resolve first-order autocorrelation assuming that the coefficient of autocorrelation in the residuals is unknown.

4. Test whether the growth rate has a different effect on *vote* for election years 1920, 1944 and 1948 and the rest of the years in the study?

Notes: (1) The dataset for this exercise, called "*data_ elections.wf1*", contains data for the United States presidential elections years from 1880 to 2008 (31 observations). (2) Describe the implementation of all tests in detail step by step. (3) For each question provide the relevant EViews outputs or/and plots.

[1] Incumbent means the party in power at the time of the election.

Exercise 3

Consider the analysis of quarterly data, from 1980 to 2018, of the variables GDP_t (income), CAP_t (stock of capital) and LAB_t (stock of labour).

1. Employ the Engle and Granger (EG) procedure to find out whether the model below constitutes a cointegrating relationship.

$$GDP_t = \gamma_0 + \gamma_1 CAP_t + \gamma_2 LAB_t + \varepsilon_t$$

2. Explain the consequences of the result of the EG test for the reliability of the regression above.

3. Specify the equation of the error correction model (ECM) using the model above as an example. How can one interpret the sign and magnitude of this ECM error correction coefficient? You should use equations to illustrate your answer.

Notes (1) The dataset for this exercise is "output.wf1". (2) Describe the implementation of all tests in detail step by step. (3) For each test provide the relevant EViews outputs or/and plots.

Exercise 4

Consider the following simultaneous equations model (SEM)

$$\begin{aligned} P_t &= \beta_0 + \beta_1 W_t + \beta_2 I_t + \beta_3 im_t + u_t \\ W_t &= \alpha_0 + \alpha_1 P_t + \alpha_2 U_t + \alpha_3 X_t + u_t \end{aligned}$$

where P is the rate of growth of prices, W is the rate of growth of wages, I is the rate of growth of investment, U is the rate of unemployment, im is the rate of growth of imports and X is the rate of growth of productivity.

Answer all questions below:

1. Classify the variables in the model above according to whether they are endogenous or exogenous.

2. Derive the reduced-form equations and determine whether the equations are identified.

3. Explain how the classical linear regression model assumption that "*the error term is uncorrelated with the explanatory variables*" is not met in the SEM above, and discuss the consequences of this violation on the ordinary least squares (OLS) estimator.

4. Can indirect least squares (ILS) be used to estimate the parameters of this SEM? Why?

5. Explain the two-stages least squares (TSLS) method of estimation for this SEM.

[The end]