**Assignment**

**Semester II, 2018**

This assignment is worth 10% of the assessment for this unit. Submit a hard copy of your assignment on or before the due date. Do NOT send the soft copies through email. However, please keep the soft copies of all of your files (e.g., MS Word, Excel, R script etc.) with your workings. They may be requested, if necessary, during marking process. **The due date for the submission is 17th of September 2018, by 16:00.** Assignment box number TBA later which is located on level 5, building 6B.

This assignment focuses on financial asset return forecasting. A financial return, in its simplest terms, is the money made or lost on an investment. This concept is one of the important topics in Finance or Financial Econometrics because the investors and portfolio managers are interested in accurate return forecasting to make important financial decisions. In addition, financial return forecast is used to calculate a value-at-risk which is the most common measure of risk in practice.

The following learning objectives will be addressed in this assignment.

1. To understand the properties and distributional characteristics of financial returns.

2. To understand and model the first moment (mean) and second moment (volatility) processes. Note that this assignment focuses only on first moment (mean) process.

5. To understand and conduct diagnostic tests of financial time series models and generate forecasts for returns and volatility. Note that this assignment focuses only on return forecasts.

The data for this assignment consists of observations on the daily closing futures price returns on West Texas Intermediate crude oil, heating oil #2, and natural gas (see the .txt files). These futures contracts trade on the New York Mercantile Exchange (NYMEX) and contract specifications and trading details are available from their website (www.nymex.com). The data set for crude oil and heating oil #2 covers the period January 1, 2001 to December 29, 2017 for a total of 4435 observations. The data have been downloaded from Datastream. Futures continuously compounded price returns are calculated as, where is the closing price on day . As the full sample comprises 4435 observations, we get 4434 return observations.

1. Perform each of the following tasks and provide the answers for the questions.
2. Compute the sample mean, standard deviation, skewness, excess kurtosis, minimum, and a maximum of each log return series. Interpret the mean, standard deviation, skewness and kurtosis from the point of view of financial analysis.

Provide R codes and/or Excel calculations and outputs in your hard-copy.

1. Obtain the empirical density plot of the daily log returns of all series. Add the theoretical normal density plot to the existing empirical plot for the comparison purposes. Are the log returns normally distributed? Why? Perform a normality test to justify your answer.

Provide R codes and/or Excel calculations and outputs in your hard-copy.

**[30%]**

1. A rolling window method of obtaining out-of-sample forecasts has been commonly used as an approach for competing models in practice. A rolling window is one where the length of the in-sample period used to estimate the model is fixed, so that the start date and end date successively increase by one observation.

The beginning and ending dates of the in-sample estimation period are then moved forward one day. The model coefficients are re-estimated, and finally, these new estimates are utilized to forecast daily one-step ahead log returns over the out-of sample period. This procedure is repeated until the available out-of sample period is exhausted. The out-of-sample period includes the last 1250 observations of the full-sample period of 1 January 2001 to 29 December 2018 and covers the period 18 March 2013 to 29 December 2018 for each commodity.

At time period *t*, a 1-day forecast is made. Models are estimated with 3184 return observations. The estimation period is then rolled forward by adding one new day and dropping the most distant day. In this way the sample size used in estimating the models stays at a fixed length and the forecasts do not overlap. Thus there are 1250 one-day return forecasts for each of petroleum futures price.

The models to be considered:

**AR(1) model:**

|  |  |
| --- | --- |
|  | (1) |

where is assumed to be a white noise series with mean zero and variance .

Note: Arima() or arima() commands in R can be used to estimate the AR(1) model.

**MA(1) model:**

|  |  |
| --- | --- |
|  | (2) |

where is assumed to be a white noise series with mean zero and variance .

Note: Arima() or arima() commands in R can be used to estimate the MA(1) model.

**ARIMA(1,0,1) model:**

|  |  |
| --- | --- |
|  | (3) |

Note: Arima() or arima() commands in R can be used to estimate the ARIMA(1,0,1) model.

**Naïve model:**

From a naïve model, the best forecast of next period's return is this period's actual return.

|  |  |
| --- | --- |
|  | (4) |

Note: Excel spreadsheet will be sufficient to obtain the forecasts.

**Historical mean approach:**

From the historical approach, the best forecast of next period’s return is the average of the previous returns.

|  |  |
| --- | --- |
|  | (5) |

Note: Excel spreadsheet will be sufficient to obtain the forecasts.

**Simple moving average technique:**

A simple moving average (MA) technique is widely used in time series forecasting. In this exercise a moving average of length m where m=20, 60, 180 days is used to generate return forecasts. These values of m correspond to one month, three months and six months of trading days respectively. The expression for the m day moving average is shown below

|  |  |
| --- | --- |
|  | (6) |

Note: Excel spreadsheet will be sufficient to obtain the forecasts.

1. You are trying to determine the accurate forecasting model for the energy commodity futures returns among the models and techniques specified earlier. Use the mean squared error (MSE) as well as mean absolute deviation (MAD) loss functions for the model comparison purposes. Report the MSE and MAD values for each model and rank all the competing models. What is the ‘best’ and ‘worst’ model for each commodity futures return?

If is a vector of one-step ahead forecasts generated from in-sample rolling windows, and is the vector of observed returns, then the out-of-sample MSE and MAD of the predictor is computed as:

Provide R codes and/or Excel calculations and outputs in your hard-copy.

1. Are the conclusions obtained from using MSE and MAD forecasting accuracy measures identical? Explain discrepancies, if there is any.

Provide R codes and/or Excel calculations and outputs in your hard-copy.

1. Do you think the complex models like AR(1), MA(1) and ARIMA(1,0,1) always produce best return forecasts compared to naïve, historical, and simple moving average models? Justify your answers based on your results.

Provide R codes and/or Excel calculations and outputs in your hard-copy.

**[70%]**