

**ASSESSMENT 1**  
**INDIVIDUAL COURSEWORK 2023/ 24**  
**2198 words**

**MSc FinTech and Business Analytics**  
**Predictive Analysis for Decision Making**  
**Module Code: FNCE044W**  
**Module Leader: Dr Issam Malki**  
**Deadline: 11th April 2024**

## Table of Contents

Question 1.....	3
1.a .....	3
1.b .....	4
1.c .....	5
1.d .....	7
1.e .....	9
1.f.i .....	10
1.f.ii .....	11
1.f.iii .....	12
1.g .....	13
Question 2 .....	14
2.a.i .....	14
2.a.ii .....	15
2.b .....	16
2.c.i .....	17
2.c.ii .....	19
Question 3 .....	20
3.a.i .....	20
3.a.ii .....	21
3.b .....	22
3.c.i .....	23
3.c.ii .....	24
3.c.iii .....	25

1 a) From the graphs provided, the main properties of the data and the stationarity can be concluded as follows :

1)  $\ln_t$  (logarithm of money)

When we access this graph we can easily say that this logarithm graph follows a trend and there is no mean reversion  
 $\therefore$  We can say that the logarithm of money is non stationary

2)  $\log_t$  (logarithm of output)

Upon assessment of this graph it is fair to say that this graph also has a trend and no mean reversion to the equilibrium.

$\therefore$  We can say that logarithm of output is non stationary as well.

3)  $\ln_p$  (logarithm of price)

Accessing logarithms of price graph tells us that it follows a trend as well leading to the conclusion that it displays non stationarity as well.

4)  $r_t$  (interest rate)

The interest rate graph does not show us a clear trend so when we take the logarithm of the interest rate we might identify a trend and the ACF for the graph will fall leading us to conclude that the interest rate is also non stationary.

1b) The mathematical model that can be expressed from Table 1 is as follows:

$$\ln_i = \beta_0 + \beta_1 r_t + \beta_2 \ln p_t + \beta_3 \ln o_t + \varepsilon_t$$

where

$\ln_i$  is the logarithm of money

$r_t$  is the interest rate

$\ln p_t$  is the logarithm of price

$\ln o_t$  is the logarithm of order

$\varepsilon_t$  is the residual error term

The main assumptions in the mathematical expression for explanatory variables are:

- (i) The set of variables that are chosen are linearly independent. Violating this assumption may lead to rise in multicollinearity which further leads to our OLS estimator to be inefficient. Formally this assumption is also called as  $X$  being a full column rank.
- (ii) The set of variables are assumed to be fixed and non random. This implies that explanatory variables are uncontrolled with the errors. Violating this assumption may lead to or result in endogeneity. This implies the OLS estimator is no longer unbiased and consistent.  
This can be expressed in the form of equation as;  
$$E[x|u] = E[x] = x$$

1c) Since we have not been provided with the t-value of coefficient of  $r_t$ . Let's calculate the coefficient of  $r_t$  first before interpretation of the coefficients.

We can calculate the coefficient of  $r_t$  using the formula

$$t\text{-value} = \frac{\hat{\beta}_{r_t}}{Se(r_t)}$$

where

$\hat{\beta}_{r_t}$  is the coefficient of  $r_t$

$Se(r_t)$  is the standard error of  $r_t$

t value is the t value of  $r_t$

$$\begin{aligned}\therefore \hat{\beta}_{r_t} &= t\text{-value} \times Se(r_t) \\ &= -26.01 \times 0.001 \\ &= -0.026\end{aligned}$$

$\therefore$  We can interpret that :

- (i)  $-0.026$  is the estimated effect of the rise in interest rate on the demand of money. As the interest rate increases by 1% the demand for money decreases by a factor of  $-0.026$ .
- (ii)  $0.961$  is the estimated effect of increase in price on the demand for money. It is estimated to be positively related. Therefore we can say the higher the prices the more is the demand for money.
- (iii)  $0.439$  is the estimated average of US goods and services output impact on demand for money. As the output increases by a factor of 1, the demand for money increases by a factor of  $0.439$ .
- (iv)  $0.416$  is the demand for money supply even when all the other factors are kept constant.

For each coefficient we have the following hypothesis to test

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

where  $i = 0, 1, 2, 3$

Since we are given  $\alpha = 5\%$ .

For all the values of  $Z_i$  (critical values),

$$|Z_i| = 1.96 \quad (z_{\alpha/2} = 1.96)$$

where  $|Z_i|$  can be calculated by the given formula:

$$|\hat{Z}_i| = \left| \frac{\hat{\beta}_i - \beta_i}{Se(\hat{\beta}_i)} \right| \sim |Z_{\alpha/2}|$$

Rejection rule:

For rejecting  $H_0$   $|\hat{Z}_i| > |Z_{\alpha/2}|$

After calculating the critical values  $|Z_i|$  for all the explanatory variables (ie.  $Z_{R_t} = 26.01$ ,  $Z_{Lp_t} = 68.6$ ,  $Z_{Lw_t} = 17.56$ ).

We can certainly say that we can reject  $H_0$  for all the explanatory variables as all of them have a p-value = 0.000 and their  $|\hat{Z}_i| > 1.96$ .

1d) The joint significance hypothesis can be expressed as follows:

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_1: \text{at least } 1 \beta_i \neq 0$$

where  $i = 1, 2, 3$

Failing to reject  $H_0$  implies model is jointly insignificant.

Rejecting the  $H_0$  implies the model is jointly significant (ie  $H_1$ )

The decision rule is as follows:

Reject the null hypothesis  $H_0$ , if

$$F > F^* \\ (\alpha, df)$$

where  $\alpha$  is significance level

$\gamma$  is number of restrictions

$df$  is degrees of freedom

$k$  is number of slopes  
 $df = N - (k + 1)$

where  $N$  is the sample size

$K$  is the number of slopes

and  $F$  can be calculated as :

$$F = \frac{(RRSS - URSS)/\gamma}{URSS/df}$$

where RRSS is Restricted sum of Squared residuals

URSS is Unrestricted sum of squared residuals

$r$  is the number of restrictions

$df$  is the degrees of freedom

o: for our given problem

$$F = \frac{(249.548 - 1.754)/3}{1.754/476} \\ = 22445.1687$$

$$(df = N - (k + 1)) \\ = 480 - (3 + 1) \\ = 476$$

$$\therefore F > F_{(3,476)}^{5\%} \approx 2.623$$

We reject the null hypothesis and conclude the model is jointly significant.

The coefficient of determination ( $R^2$ ) suggests that about 99.2% of the variations are explained by the model. Still with such high value of  $R^2$  we should also keep in mind that there could be a possibility of ~~overfitting~~ overfitting.

e) Based on the economists suggestions we formulate the hypothesis as follows:

$$H_0: \beta_2 = 2\beta_3 \text{ or } \beta_2 - 2\beta_3 = 0$$

$H_1: H_0$  is false

Our Unrestricted Model is as follows (URM):

$$lm_t = \beta_0 + \beta_1 r_t + \beta_2 l_p_t + \beta_3 l_o_t + \varepsilon_t$$

Whereas the restricted model provided by the economists is (RM):

$$lm_t = \beta_0 + \beta_1 r_t + \beta_2 l_p_t + 0.5\beta_2 l_o_t + \varepsilon_t$$

We are provided with the values of URSS and RRSS  
(ie  $URSS = 1.754$  and  $RRSS = 1.760$ )

$$\begin{aligned} \therefore F &= \frac{(RRSS - URSS)/df}{URSS/df} & \because df = N - (k+1) \\ &= \frac{1.760 - 1.754/1}{1.754/477} \\ &= 1.628 \end{aligned}$$

$\therefore$  The decision rule states

Reject the  $H_0$  if  $F > F_{(1,474)}^{5\%} \approx 8.524$

$\therefore$  We cannot reject  $H_0$  since  $F$  is not greater than the critical value.

1 f) i) Unit roots are referred to a characteristic of time series where the series is non stationary due to the presence of a root which is equal to 1 in the autoregressive model of the series

Consider the Data Generating Process (DGP) for a process  $y_t$ :

$$y_t = \rho y_{t-1} + u_t \text{ with } -1 < \rho \leq 1$$

where  $u_t$  is the white noise and

$\rho$  is the persistence term. If  $\rho = 1$  the process  $y_t$  becomes a Random Walk Process and  $y_t$  contains a unit root.

The main implication of the process  $y_t$  or the process which contains a unit root is that the process will never converge to a long run level or equilibrium.

In context of the estimated model in Table 1. If the variables exhibit unit roots, the models estimations and inferences could be unreliable and will become necessary to establish that the model is cointegrated before proceeding with the analysis.

1 f) ii) The Augmented Dickey Fuller (ADF) test is a statistical test that checks whether a series has a unit root. The steps involved to perform ADF Test is as follows:

Step 1: Identify the Form

ADF proposes three (DGPs)

a)  $y_t$  is a random walk (no constant or trend)

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^k \theta_i \Delta y_{t-i} + u_t$$

b)  $y_t$  is a random walk with drift (with constant)

$$\Delta y_t = \beta_1 + \delta y_{t-1} + \sum_{i=1}^k \theta_i \Delta y_{t-i} + u_t$$

c)  $y_t$  is a random walk with drift around a stochastic trend (with drift and trend)

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^k \theta_i \Delta y_{t-i} + u_t$$

where  $k$  is the lag length in each equation

Step 2: Identify for unit root with ADF

$$H_0: \delta = 0 \Rightarrow \text{unit root}$$

$$H_1: \delta \neq 0 \Rightarrow \text{No unit root}$$

Step 3: Identify stationarity using KPSS

$$H_0: \delta < 0 \Rightarrow \text{stationarity}$$

$$H_1: \delta = 0 \Rightarrow \text{unit root}$$

Step 4: Conclude the results

(i) Reject  $H_0$  for ADF and fail to reject  $H_0$  for KPSS for both tests to conclude stationarity

(ii) For both tests Reject  $H_0$  if  $P < \alpha$ , where  $\alpha$  is the level of significance (ie 5%)

(iii) and fail to reject  $H_0$  when  $P \geq \alpha$

1 f) iii) The ADF test statistic is compared to the critical value at the 1%, 5%, 10% significance levels.

If the test statistic is greater than the critical value at a given significance level, we reject the null hypothesis of a unit root, indicating that the series is stationary.

If the test statistic is less than the critical value, we fail to reject the null hypothesis, suggesting the presence of a unit root.

Now let's analyze each variable:

i)  $r_t$ : Test statistic: -2.612

Critical Value: -2.334 (1%), -1.648 (5%), -1.203 (10%)

Conclusion: Reject the null hypothesis  $H_0$ , the series is stationary.

ii)  $lm_t$ : Test statistic: 0.048

Critical Value: -2.334 (1%), -1.648 (5%), -1.203 (10%)

Conclusion: Fail to reject the null hypothesis, indicating the presence of a unit root.

iii)  $lp_t$ : Test statistic: -0.258

Critical Value: -2.334 (1%), -1.648 (5%), -1.203 (10%)

Conclusion: Fail to reject the null hypothesis, indicating the presence of a unit root.

iv)  $lo_t$ : Test statistic: -0.909

Critical Value: -2.334 (1%), -1.648 (5%), -1.203 (10%)

Conclusion: Fail to reject the null hypothesis, indicating the presence of a unit root.

1g)

The Durbin Watson value given in the table: - 0.122

Interpretation of the Durbin Watson Statistic

The hyopesis:

$H_0$ : No positive first-order autocorrelation

$H_1$ : Positive first order autocorrelation

The decision can be taken as:

- Value approximately 2 indicates  $\rightarrow$  no autocorrelation

- Value  $< 2$  indicates - positive correlation

- Value  $> 2$  indicates - negative correlation

$\therefore$  The DW Statistic  $< 2$ , We can clearly say that there is a positive correlation in the residuals, violating the assumption of the independent errors which could impact the validity of statistical inferences from the model.

Therefore we can conclude that from the Durbin Watson test results, that there is proof of positive first order autocorrelation in the errors and the problem needs to be fixed before further analysis.

2 a) i) Given that  $\epsilon_t \sim \text{iid}(0, \sigma^2)$

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

Mean

$$E[y_t] = E[\epsilon_t] + \theta_1 E[\epsilon_{t-1}] + \theta_2 E[\epsilon_{t-2}]$$

$\because \epsilon_t$  is iid, for all values of  $\epsilon_t$  the mean is same

$$\therefore E[y_t] = 0 + \theta_1(0) + \theta_2(0)$$

$$= 0$$

Variance

$$\text{Var}[y_t] = \text{Var}[\epsilon_t] + \theta_1^2 \text{Var}[\epsilon_{t-1}] + \theta_2^2 \text{Var}[\epsilon_{t-2}]$$

similarly since  $\epsilon_t$  is iid all the  $\epsilon_t$  variance values are same

$$\begin{aligned}\therefore \text{Var}[y_t] &= \sigma^2 + \theta_1^2 \sigma^2 + \theta_2^2 \sigma^2 \\ &= \sigma^2(1 + \theta_1^2 + \theta_2^2)\end{aligned}$$

Covariance

$$\begin{aligned}\text{Cov}[y_t, y_{t-1}] &= E[(y_t - E[y_t])(y_{t-1} - E[y_{t-1}])] \\ &= E[\epsilon_t \epsilon_{t-1} + \theta_1 \epsilon_{t-1}^2 + \theta_2 \epsilon_{t-1} \epsilon_{t-2}] \\ &= E[\epsilon_t \epsilon_{t-1}] + \theta_1 E[\epsilon_{t-1}^2] + \theta_2 E[\epsilon_{t-1} \epsilon_{t-2}] \\ &= 0 + \theta_1 \sigma^2 + 0\end{aligned}$$

$$\begin{aligned}\text{Cov}[y_{t-1}, y_{t-2}] &= E[(y_t - E[y_t])(y_{t-2} - E[y_{t-2}])] \\ &= E[\epsilon_t \epsilon_{t-2} + \theta_1 \epsilon_{t-1} \epsilon_{t-2} + \theta_2 \epsilon_{t-2}^2] \\ &= E[\epsilon_t \epsilon_{t-2}] + \theta_1 E[\epsilon_{t-1} \epsilon_{t-2}] + \theta_2 E[\epsilon_{t-2}^2] \\ &= 0 + \theta_1 0 + \theta_2 \sigma^2 \\ &= \theta_2 \sigma^2\end{aligned}$$

2a) ii)  $\epsilon_t \sim \text{iid}(0, \sigma^2)$

$$w_t = w_{t-1} + \epsilon_t$$

where  $w_0 = 0$

using recursive substitution we can say that

$$w_1 = w_0 + \epsilon_1$$

$$w_2 = w_1 + \epsilon_2 = w_0 + \epsilon_1 + \epsilon_2$$

$$w_3 = w_2 + \epsilon_3 = w_0 + \epsilon_1 + \epsilon_2 + \epsilon_3$$

$$\therefore w_t = w_0 + \epsilon_1 + \epsilon_2 + \dots + \epsilon_{t-1} + \epsilon_t$$

since  $w_0 = 0$

$$w_t = \epsilon_1 + \epsilon_2 + \dots + \epsilon_{t-1} + \epsilon_t$$

$\therefore$  Mean

$$\begin{aligned} E[w_t] &= E[\epsilon_1 + \epsilon_2 + \epsilon_3 + \dots + \epsilon_{t-1} + \epsilon_t] \\ &= E[\epsilon_1] + E[\epsilon_2] + \dots + E[\epsilon_{t-1}] + E[\epsilon_t] \end{aligned}$$

since  $\epsilon_t$  follows  $\text{iid}(0, \sigma^2)$  therefore  $E[\epsilon_t] = 0$

$$\therefore E[w_t] = 0$$

Variance

$$\begin{aligned} \text{Var}[w_t] &= \text{Var}[\epsilon_1 + \epsilon_2 + \dots + \epsilon_{t-1} + \epsilon_t] \\ &= \text{Var}[\epsilon_1] + \text{Var}[\epsilon_2] + \dots + \text{Var}[\epsilon_{t-1}] + \text{Var}[\epsilon_t] \end{aligned}$$

since  $\epsilon_t$  follows  $\text{iid}(0, \sigma^2)$  therefore  $\text{Var}[\epsilon_t] = \sigma^2$

$$\begin{aligned} \therefore \text{Var}[w_t] &= \sigma^2 + \sigma^2 + \dots + \sigma^2 + \sigma^2 \\ &= t\sigma^2 \end{aligned}$$

Covariance

$$\text{Cov}[w_t, w_{t-j}] = E[(w_t - E[w_t])(w_{t-j} - E[w_{t-j}])]$$

if  $j = 0$

$$\begin{aligned} \text{Cov}[w_t, w_{t-j}] &= E[(w_t - E[w_t])(w_t - E[w_t])] \\ &= E[w_t^2] = \text{Var}(w_t) = t\sigma^2 \end{aligned}$$

if  $j \neq 0$

$$\text{Cov}[w_t, w_{t-j}] = E[(w_t - E[w_t])(w_{t-j} - E[w_{t-j}])]$$

$$= E[w_t] E[w_{t-j}]$$
$$= 0$$

$\therefore$  Covariance

$$\text{Cov}[w_t, w_{t-j}] = \begin{cases} t\sigma^2 & \text{if } j=0 \\ 0 & \text{if } j \neq 0 \end{cases}$$

$$2 b) y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

The process  $y_t$  is weakly stationary. This can be concluded from the derived expression of mean, variance and covariance of  $y_t$ .

The mean is constant and equal to 0. independent of time  $t$ .

The variance is constant and equal to  $\sigma^2(1+\theta_1^2+\theta_2^2)$ . independent of time  $t$ .

The covariances  $\text{cov}(y_t, y_{t-j})$  depends only on the lag  $j$ , not on the time  $t$ .

Therefore, these properties satisfy the definition of weak stationarity where all mean, variance and covariances do not depend on time.

$$w_t = w_{t-1} + \epsilon_t$$

The process  $w_t$  is not weakly stationary. This can be concluded from the derived values of mean, variance and covariances of  $w_t$ .

The mean is constant and equal to 0. independent of  $t$ .

The variance  $\text{Var}(w_t)$  is however dependent on time  $t$ , since  $\text{Var}(w_t) = t\sigma^2$  which violates the requirement for weak stationarity.

The covariance  $\text{cov}(w_t, w_{t-j})$  depends on time  $t$ , not just the lag  $j$ , again violating the requirement of weak stationarity.

Therefore we can conclude that the process  $w_t$  is not weak stationary.

2 c) i) Detect first order autocorrelation using Durbin Watson (DW) test (given DW = 1.20)

### Hypothesis

Null hypothesis

$H_0 : \rho = 0$  first order autocorrelation does not exist in the model

Alternate Hypothesis

$H_1 : \rho \neq 0$  there is evidence of first order autocorrelation in the model.

The critical values derived from the DW statistics table is as follows:

$$\alpha = 5\%$$

$$K = 2$$

$$N = 90$$

$$\therefore d_L = 1.61, d_U = 1.70$$

The following interpretation can be used to test the hypothesis.

$0 - 1.61(d_L) \rightarrow \text{Reject } H_0$  (Positive autocorrelation)

$1.61(d_L) - 1.70(d_U) \rightarrow \text{cannot conclude}$

$1.70(d_U) - 2.3(4 - d_U) \rightarrow \text{Do not reject } H_0$

$2.3 - 2.39(4 - d_L) \rightarrow \text{Cannot conclude}$

$2.39 - 4 \rightarrow \text{Reject } H_0$  (Negative correlation)

The DW value 1.20 lies between  $0 - 1.61(d_L)$  denoting positive correlation, therefore we reject  $H_0$  showing evidence of first order autocorrelating in the model.

2c) ii) We are given that

$$\hat{b}_t = 0.49 - 0.27 p_t + 0.22 y_t$$
$$(0.84) \quad (0.27) \quad (0.71)$$
$$R^2 = 0.28, DW = 1.20, LM(2) = 7.42$$

where  $b_t$  is demand for Test books

$p_t$  is price of a book

$y_t$  is total level of income.

LM test can be done using the following steps:

1) Hypothesis

$H_0$ : Autocorrelation of 2<sup>nd</sup> order does not exist in the model

$H_1$ : There is evidence of 2<sup>nd</sup> order autocorrelation in the model

2) Residuals Calculation

for every observation in the model the residuals  $\hat{u}_t$

$$\hat{u}_t = b_t - \hat{b}_t$$

3) Auxiliary regression

We perform the regression using the residuals  $\hat{u}_t$  as the dependent variable against the original explanatory variables ( $p_t$  &  $y_t$ ) and the lagged residuals ( $\hat{u}_{t-1}, \hat{u}_{t-2}$ )

$$\hat{u}_t = \rho_0 + \rho_1 p_t + \rho_2 y_t + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \epsilon_t$$

From this regression we derive the coefficient of determination  $R^2$

4) Deriving the LM statistics value

The LM statistics can be calculated as

$$LM = R^2 \times n$$

where LM is given 7.42

### 5) Deriving the critical values using chi square distribution

The LM statistics follow Chi-Square distribution, with degrees of freedom equal to the number of lags which in our case is 2, with a 5% significance level, the critical value from the chi-square distribution with two degrees of freedom is 5.99.

### 6) Decision rule and Conclusion

As the LM is higher as compared to the critical value (i.e  $LM: 7.42 > \text{critical value: } 5.99$ ), therefore we reject the null hypothesis ( $H_0$ ). It indicates that the model exhibits 2<sup>nd</sup> order autocorrelation.

3a) i) Given

URM

$$\hat{y}_t = 2.3 + 0.2x_{1i} - 6.14x_{2i} - 0.01x_{3i} + 1.5x_{4i}$$

RM

$$\hat{y}_t = 2.6 + 0.25x_{1i} - 6.14x_{2i}$$

The joint significance hypothesis are :

$$H_0: \beta_3 = \beta_4 = 0$$

$$H_1: \text{at least 1 of } \beta_3 \text{ or } \beta_4 \neq 0$$

Failing to reject the null hypothesis, implies model is jointly insignificant.

Rejecting the null hypothesis, implies the model is jointly significant.

The decision rule is as follows:

Reject the null hypothesis  $H_0$ , if

$$F > F^*_{(r, df)}$$

where  $F$  can be calculated as

$$\begin{aligned} F &= \frac{(RRSS - URSS)/r}{URSS/df} \\ &= \frac{(0.70 - 0.60)/2}{0.60/47} \\ &= 3.937 \end{aligned}$$

$$\therefore F < F^*_{(2, 47)} \approx 19.48$$

Therefore we fail to reject the null hypothesis and conclude that the model is jointly insignificant

3a)ii) To access whether the financial firms have average ROE, we use a dummy variable in the equation, that can be explained by the following steps:

- Define the dummy variable

$$D_{\text{financial}} = 0 \text{ for non financial firms}$$

$$D_{\text{financial}} = 1 \text{ for financial firms}$$

- Define the model with dummy variable

$$\hat{Y}_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k D_{\text{financial}} + \epsilon_t$$

- Define hypothesis

$$H_0: \beta_k = 0 \rightarrow \text{financial firms do not have higher ROE compared to non financial firms}$$

$$H_1: \beta_k > 0 \rightarrow \text{financial firms have higher average ROE than non financial firms.}$$

- Model Estimation & Mean consideration

Model's coefficient are calculated using regression. The coefficient  $\beta_k$  of the dummy variable measures average difference between the average ROE between financial and non financial firms, while the other parameters remain constant. The p value associated with the coefficient  $\beta_k$  will show pre-determined significance level ( $< 0.05$ ), based on which we can determine whether to reject or accept  $H_0$ .

Mean is higher: This will tell us that the financial firms have higher average ROE compared to non financial firms. We reject the null hypothesis  $H_0$ .

Mean is lower : This will tell us that the model is insignificant and financial firms have lower or similar average ROE compared to non financial firms. Accept the H<sub>0</sub>.

3 b) In order to test the presence of 4<sup>th</sup> order ARCH effects in stock market index, the ARCH LM test proposed by Engle shall can be applied, which we can do by following the below steps.

i) Conditional mean specification

A regression or ARIMA model is run to fit the return ( $R_t$ ), thus obtaining the residuals  $\hat{E}_t$ .

ii) Auxiliary regression

Residual obtained from the regression above are squared to estimate the variance followed by regressing the squared residuals on their own past values. it can be denoted by the expression below:

$$\hat{\varepsilon}_t^2 = a_1 \hat{\varepsilon}_{t-1}^2 + a_2 \hat{\varepsilon}_{t-2}^2 + a_3 \hat{\varepsilon}_{t-3}^2 + a_4 \hat{\varepsilon}_{t-4}^2 + u_t$$

iii) hypothesis

$$H_0: a_1 = a_2 = a_3 = a_4 = 0$$

$$H_1: a_i \neq 0$$

$\rightarrow$  No ARCH effects are present

$\rightarrow$  ARCH effect exists in atleast one

$i (1, 2, 3, 4)$

iv) Statistic & Distribution

The LM statistic can be calculated where  $LM = nR^2$

It follows a chi-square ( $\chi^2$ ) distribution with 4 degrees of freedom referring to the number of lagged squared residual in the regression.

v) Decision Rule and Conclusion

In this we will compare  $\chi^2$  value with LM statistics at 5% significance levels.

If  $LM >$  critical value  $\rightarrow$  Reject  $H_0$ , 4<sup>th</sup> arch effect is present

If  $LM \leq$  critical value  $\rightarrow$  fail to reject  $H_0$ , no evidence of 4<sup>th</sup> arch effect in stock market index returns.

3 c i)  $\hat{y}_t = 0.75 \rightarrow$  conditional mean

$$\hat{\sigma}_t^2 = 1.76 + 0.07 u_{t-1}^2 + 0.24 \hat{\sigma}_{t-1}^2 + 0.87 u_{t-1}^2 I_{t-1}$$

The last term in the equation captures the asymmetric effect in the market, if positive, the impact of bad news and good news is different.

$$\text{Good news} \rightarrow 0.07 u_{t-1}^2 + 0.87 u_{t-1}^2 \times 0$$

$\therefore$  The magnitude of the shock is 0.07

$$\text{Bad news} \rightarrow 0.07 u_{t-1}^2 + 0.87 u_{t-1}^2 \times 1$$

$\therefore$  The magnitude of the shock is 0.94

### Interpretation

The asymmetric effect is significant statistically.  
Also positive shock is 0.07 and negative shock is 0.94.  
Therefore when we have a positive news we have a  
restricted volatility and during negative news it has  
a greater susceptibility to unfavourable occurrences.

3c ii) The estimated coefficient of 0.24 for the lagged conditional variance ( $\sigma_{t-1}^2$ ) in a GARCH model can be interpreted as shown in the GARCH effect, or the persistence of shocks to volatility. This view is consistent with the measurement of the impact of volatility from earlier eras on the volatility of the FTSE 100 index in the current time.

0.24 is the persistence of the ~~shock~~ shock or GARCH effect. The closer to 1 the higher is the persistence, as the shock lasts for a longer time.

Hence in this case the shock is moderately persistent.

3 ciii)

Given

$$\sigma_{t-1}^2 = 0.74$$

$$\hat{u}_{t-1} = \pm 0.5$$

For positive shock  $\hat{u}_{t-1} = 0.5$

For negative shock  $\hat{u}_{t-1} = -0.5$

$\sigma_t^2$  for good news

$$I_{t-1} = 0, \hat{u}_{t-1} = 0.5$$

$$\begin{aligned}\sigma^2 &= 1.76 + 0.07(0.5)^2 + 0.24(0.74) \\ &= 1.955\end{aligned}$$

Interpretation

There is a volatility of 1.955 in the market due to good news

$\sigma_t^2$  for bad news

$$I_{t-1} = 1, \hat{u}_{t-1} = -0.5$$

$$\begin{aligned}\sigma^2 &= 1.76 + 0.07(-0.5)^2 + 0.24(0.74) + 0.87(-0.5)(1) \\ &= 2.176\end{aligned}$$

Interpretation

There is a volatility of 2.176 in the market due to a bad news.