**HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY**

**BIG DATA TECHNOLOGY**

**MSBD5006: Quantitative Analysis of Financial Time Series**

**Stock Market Index Trading Strategies with ARIMA+GARCH Model**

**Group 2**

**AU YEUNG Ka Ho 20736576**

**HO Kwan Yu 07300837**

**TSE Man Chim 20547664**

**WONG Chin Hang 20123808**

**WONG Tsz Ho 20725187**

**Content**

[**1.** **INTRODUCTION** 3](#_Toc57762264)

[**1.1** **Introduction to Index Investing** 3](#_Toc57762265)

[**1.2** **Motivation** 4](#_Toc57762266)

[**2.** **EXPLORATORY DATA ANALYSIS** 5](#_Toc57762267)

[**2.1** **Data-at-a-glance** 5](#_Toc57762268)

[**2.2** **Data Cleaning** 6](#_Toc57762269)

[**2.3** **Unit Root Test: Dickey-Fuller Test** 6](#_Toc57762270)

[**2.4** **Serial Correlation Checking: Ljung–Box test** 7](#_Toc57762271)

[**2.5** **ARCH Effect Detection: Ljung–Box test** 8](#_Toc57762272)

[**2.6** **Cointegration Test: Johansen Test** 8](#_Toc57762273)

[**3.** **MODEL FITTING AND VALIDATION** 9](#_Toc57762274)

[**3.1** **Univariate ARIMA Model** 9](#_Toc57762275)

[**3.1.1 Hang Seng Index** 9](#_Toc57762276)

[**3.1.2 S&P 500** 9](#_Toc57762277)

[**3.1.3 FTSE100** 10](#_Toc57762278)

[**3.1.5 Conclusion** 11](#_Toc57762279)

[**3.2** **ARIMA-GARCH Model** 12](#_Toc57762280)

[**3.2.1 Hang Seng Index** 12](#_Toc57762281)

[**3.2.2 S&P 500 Index** 13](#_Toc57762282)

[**3.2.3 FTSE 100 Index** 13](#_Toc57762283)

[**3.2.4 SSE Index** 14](#_Toc57762284)

[**3.3** **Multivariate Autoregressive Moving Average (ARMA) Models** 15](#_Toc57762285)

[**3.3.1 Vector Autoregressive (VAR) Model** 15](#_Toc57762286)

[**3.3.2 Vector Moving Average (VMA) model** 18](#_Toc57762287)

[**3.3.3 Vector-Autoregressive Moving Average (VARMA) Model** 21](#_Toc57762288)

[**3.3.4 Model Summary** 23](#_Toc57762289)

[**3.4** **Multivariate GARCH Model – BEKK** 23](#_Toc57762290)

[**4.** **TRADING STRATEGIES AND BACKTESTING RESULT** 24](#_Toc57762291)

[**4.1** **Implementation of the strategy** 24](#_Toc57762292)

[**4.2** **Back-testing Result –S&P 500 Index** 25](#_Toc57762293)

[**5.** **DISCUSSION** 26](#_Toc57762294)

[**5.1** **Data Limitation** 26](#_Toc57762295)

[**5.2** **Model Limitation** 26](#_Toc57762296)

[**6.** **CONCLUSION** 27](#_Toc57762297)

[**Reference** 28](#_Toc57762298)

[**Appendix** 30](#_Toc57762299)

# **INTRODUCTION**

## **Introduction to Index Investing**

An index is a way to track the performance of certain group of assets, which typically use a basket of securities to replicate the market for performance measurement. There are few index weight methodologies, including Market-Capitalization Weighting which is to weight constituent stocks by its market capitalization (e.g., Hang Seng Index), Price Weighting which is to weight constituent stocks by its price per share (e.g., Dow Jones Industrial Average) and last but not the least Equal Weighting which is to give each constituent stocks equal weights (e.g., S&P 500 equally weighted index).

Stock Market Index investing is a passive investment strategy which aims to generate similar returns as the market index. The buy-and-hold strategy can be used to replicate the performance by investing in exchange-traded fund (ETF) that closely tracks the underlying index or using the Index futures that the investors can buy or sell the index today to be settled at a future date.

There are several advantages of index investing. Firstly, the index investing can offer greater diversification than the stock-picking strategies. The index trading can help to reduce the uncertainty due to the unsystematic risk which is the opposite of the systematic risk. Unsystematic risk is the company or industry specific hazards which affect the company due to certain factors, no matter external or internal. For instance, the internal frauds can negatively impact the valuation of the company but not for the other companies. Secondly, the index investing can provide lower management fees and lower expense ratio than the actively managed funds. This is because of the simplicity of this strategy which is to track the market index. From the investment research firm Morningstar Research report[[1]](#footnote-2), the average expense ratio of passive funds was 0.15% in 2018, compared to 0.67% for active funds. Thirdly, indexing intends to match the risk and return of the overall market. From the empirical studies, this strategy tends to perform better over the long term than the actively managed funds. This is because it is hard for a person to pick stocks that will beat the market and even harder to do so consistently over decades.

## **Motivation**

As there are many advantages of index investing over the active funds, the index investing is gaining popularity in the recent days. From the Financial Times article[[2]](#footnote-3), the assets managed by global index funds have smashed through the $10tn level amid the investors shying away from the active funds.

This project aims to identify the way to improve the return from the index investing strategy with the help of technical analysis, concretely using the time series model including univariate ARIMA-GARCH model and multivariate GARCH model. In our report, Hang Seng Index would be our focus which is the free float-adjusted market capitalization-weighted stock market index in Hong Kong. To assist our analysis, the other major stock market indexes will also be included in our studies. Due to the possible tracking error of the ETFs and the transaction costs incurred when rolling over the future contracts, the closing price of the indexes will directly be used in our analysis.

# **EXPLORATORY DATA ANALYSIS**

## **Data-at-a-glance**

**Chart, scatter chart

Description automatically generated**Chart, line chart, histogram

Description automatically generatedChart, histogram

Description automatically generatedChart, line chart, histogram

Description automatically generatedIn our studies, the data from 2007 to 2020 will be used. It can be easily observed from the graphs below that the HSI index closing price is not stationary. As we are going to employ multivariate time series model in our analysis, the closing price and log return are also plotted for the other major market indexes including Shanghai Composite Index, S&P 500 Index and FTSE 100 Index

Chart, line chart

Description automatically generated**Chart, line chart, histogram

Description automatically generatedChart, line chart, histogram

Description automatically generatedChart, line chart

Description automatically generated**Figure 1: Time Plot for Market Indexes

Figure 2: Time Plot for Log Return of Market Indexes

## **Data Cleaning**

Since holidays in different countries/cities are not the same, some markets are open while a market may be closed. The closed market does not provide any index data on that day. To align the all the index data on the same day, we used Hang Seng Index as a primary index, if other market index data do not exist, we will substitute the missing data with the previous valid market index.

|  |  |
| --- | --- |
| Before | After |
| |  |  |  | | --- | --- | --- | | **Date** | **HSI** | **SSE** | | 3-May | 20681.58 | - | | 2-May | 20388.49 | - | | 30-Apr | 20318.98 | 3841.27 | | |  |  |  | | --- | --- | --- | | **Date** | **HSI** | **SSE** | | 3-May | 20681.58 | 3841.27 | | 2-May | 20388.49 | 3841.27 | | 30-Apr | 20318.98 | 3841.27 | |
| Table 1a | Table 1b |

## **Unit Root Test: Dickey-Fuller Test**

The Dickey-Fuller test is used to determine if the unit root is present in the time series data. The null hypothesis is that the data contains a unit root while the alternative is that the data is a stationary process. From the R output below, it can be concluded that the HSI Index closing price is not stationary while the HSI Index Log Return is stationary as the p-value is below 5%. Similarly, the index closing for SSE Index, S&P500 and FTSE100 are also not stationary.

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Result** |
| HSI Index | 0.1254 (>0.05) | Not Stationary |
| Log Return of HSI Index | 0.01 (<0.05) | Stationary |
| SSE Index | 0.4351 (>0.05) | Not Stationary |
| Log Return of SSE Index | 0.01 (<0.05) | Stationary |
| S&P500 Index | 0.289 (>0.05) | Not Stationary |
| Log Return of S&P500 Index | 0.01 (<0.05) | Stationary |
| FTSE100 Index | 0.2323 (>0.05) | Not Stationary |
| Log Return of FTSE100 Index | 0.01 (<0.05) | Stationary |

Table 2: R Output for Dickey-Fuller Test

## **Serial Correlation Checking: Ljung–Box test**

The Ljung-box test is used to test if any of autocorrelations of the variables are different from zero. The null hypothesis is that whole group of autocorrelations is zero while the alternative is that at least one of them is non-zero. From the R output below, it can be concluded that all the market indexes exhibit the serial correlations as the p-value are all below 0.05.

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Result** |
| Log Return of HSI Index | 0.03457 (<0.05) | Not all ACF are zero |
| Log Return of SSE Index | 0.000363 (<0.05) | Not all ACF are zero |
| Log Return of S&P500 Index | < 2.2e-16 (<0.05) | Not all ACF are zero |
| Log Return of FTSE Index | < 2.2e-16 (<0.05) | Not all ACF are zero |

Chart

Description automatically generatedChart, histogram

Description automatically generatedChart, histogram

Description automatically generatedChart

Description automatically generatedTable 3: R Output for Serial Correlation Check

Figure 3: ACF Plot for Log-return of Market Indexes

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generatedChart, histogram

Description automatically generatedChart, histogram

Description automatically generated

Figure 4: Partial ACF Plot for Log-return of Market Indexes

## **ARCH Effect Detection: Ljung–Box test**

Autoregressive conditional heteroscedastic (ARCH) effect is referring to the time series exhibiting autocorrelation in the squared residual series. Ljung-Box test is conducted on first 10 legs of the squared residual series to detect if any of autocorrelations of the squared residuals are different from zero. The null hypothesis is that whole group of autocorrelations is zero . From the R output below, it can be concluded that all the market indexes exhibit the ARCH effects as the p-value are all below 0.05.

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Result** |
| Log Return of HSI Index | < 2.2e-16 (<0.05) | Not all ACF are zero |
| Log Return of SSE Index | < 2.2e-16 (<0.05) | Not all ACF are zero |
| Log Return of S&P500 Index | < 2.2e-16 (<0.05) | Not all ACF are zero |
| Log Return of FTSE Index | < 2.2e-16 (<0.05) | Not all ACF are zero |

Table 4: R Output for ARCH Effect Detection

## **Cointegration Test: Johansen Test**

In this section, we are going to analyse whether two or more of the market indexes can form a cointegration relationship. This would allow us to form a mean reversion trading strategy if any. There are two types of Johansen Test, namely trace test and maximum eigenvalue test. In this section, trace test will be used. From the R Output below, it shows the trace test statistics for four hypotheses of and . For the tests below, the test statistics are all below the critical levels at different confidence level (1%, 5% and 10%). As a result, we are not able to reject the null hypothesis of no integration. In another words, there is concrete evidence to support the existence of a stationary linear combination of the major market indexes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Eigenvalues** | | | |
| 0.0061077034 | 0.0028737321 | 0.0027190337 | 0.0007067396 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test** | **Test Statistics** | **10pct** | **5pct** | **1pct** |
| r <= 3 | 2.23 | 6.5 | 8.18 | 11.65 |
| r <= 2 | 10.80 | 15.66 | 17.95 | 23.52 |
| r <= 1 | 19.86 | 28.71 | 31.52 | 37.22 |
| r = 0 | 39.14 | 45.23 | 48.28 | 55.43 |

Table 5: R Output for Johansen Test

# **MODEL FITTING AND VALIDATION**

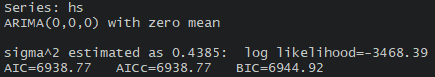
## **Univariate ARIMA Model**

In this part, different combination of p and q in ARMA(p,q) models will be fitted on the market indexes and subsequently model checking will be done to validate the fitting result.

### **3.1.1 Hang Seng Index**

All AR and MA models, including the AR(1), AR(2), MA(1), MA(2), ARMA(1,1), ARMA(2,2), failed in Ljung-Box test in 10% significant level. It means we cannot reject the null hypothesis that all AR and MA models' residuals are white noise at 10% significant level.

Second, we employ auto.arima from R package. The result is as follow:



Output 1a: R output on ARIMA model of HSI

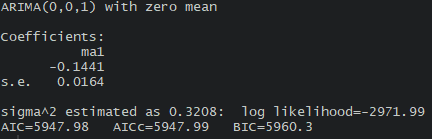
The program calculated the best fitted model, based on AIC, is ARIMA(0,0,0).

We can say the Hang Seng Indexes is not well fitted with the simple ARMA model. We will try different models in the follow sections.

### **3.1.2 S&P 500**

Many AR and MA models fit in S&P 500. We decide to pick the model with highest AIC.

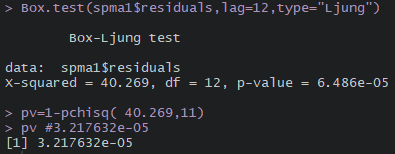
The model is MA(1):



Output 1b: R output on ARIMA model of SP500

The fitted model is:

We employed Ljung-Box test to examine the residuals of the fitted model:



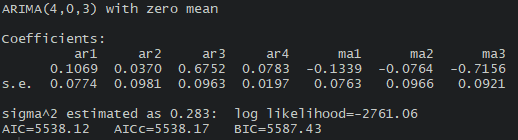
Output 1c: R output on Ljung Box Test of SP500

Since p- value is 3.217632e-05, we can reject the null hypothesis that the model’s residuals are white noise at 10% significant level. We will try different models in the follow sections.

### **3.1.3 FTSE100**

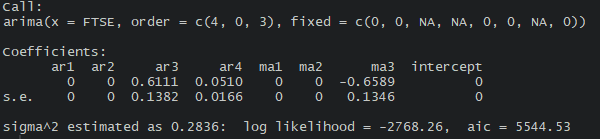
Many AR and MA models fit in FTSE100. We decide to pick the model with highest AIC.

The model is ARIMA(4,0,3):



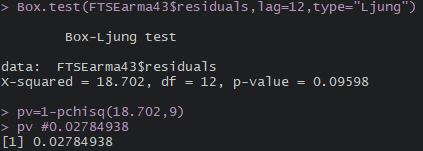
Output 2a: R output on ARIMA model of FTSE100

Since some of the parameters are not significant, we set them to zero and recalculate the function:



Output 2b: R output on pruned ARIMA model of FTSE100

The fitted model is:



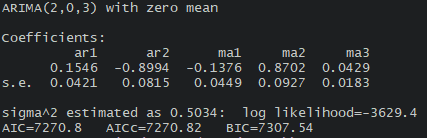
Output 2c: R output on Ljung Box Test of FTSE100

Since p- value is 0.02784938, we can reject the null hypothesis that the model’s residuals are white noise at 10% significant level. We will try different models in the follow sections.

**3.1.4 Shanghai Composite Index**

Many AR and MA models fit in Shanghai Composite Index. We decide to pick the model with highest AIC.

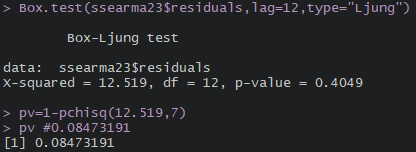
The model is ARIMA(2,0,3):



Output 3a: R output on ARIMA model of SSE

The fitted model is:





Output 3b: R output on Ljung Box Test of SSE

Since p- value is 0.08473191, we can reject the null hypothesis that the model’s residuals are white noise at 10% significant level. We will try different models in the follow sections.

### **3.1.5 Conclusion**

We cannot find any ARMA model that is significant for any indexes. In the following section, we will examine GARCH models to see whether we can get better models.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Best fitted model** | **Residuals' P-Value** | **Result** |
| Log Return of HSI Index | None | NIL | No suitable ARMA model |
| Log Return of SSE Index | ARIMA(2,0,3) | 0.08473191 | Residuals are not white noise |
| Log Return of S&P500 Index | MA(1) | 3.217632e-05 | Residuals are not white noise |
| Log Return of FTSE Index | ARIMA(4,0,3) | 0.02784938 | Residuals are not white noise |

Table 6: Summary of the fitted univariate models

## **ARIMA-GARCH Model**

As the ARIMA model fits on the stationary and linear time series data, the non-linear part of the data will remain in the residuals of the model. Meanwhile, the GARCH model can represent the non-linear residual patterns well. These two models need to be combined to capture both linear and non-linear part of the data.

### **3.2.1 Hang Seng Index**

As shown in the earlier section, there exists ARCH effect on the log return of HSI time series. To capture this effect, GARCH model must be combined with the existing model. Having tested with different GARCH Model, it is found that iGARCH [[3]](#footnote-4)(1,1) fit the data series the best. The fitted model is shown as below:

,

,

,

From the R output below, as the p-value of Ljung–Box test for the standardized residuals is 0.2314 and the squared standardized residuals is 0.1398 which are all above 0.05, these show that the residuals are not autocorrelated and do not exhibit ARCH effect so the model is adequate.

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Estimate** | **P-value** |
| Alpha1 | 0.069228 | 0.000016 |
| Beta1 | 0.930772 | N/A |

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Intrepretation** |
| Residual | 0.2314 | Residuals are not autocorrelated |
| Squared Residual | 0.1398 | Residuals have no ARCH pattern |

Table 7: ARIMA(0,0,0)-iGARCH(1,1) Model for HSI Index:

### **3.2.2 S&P 500 Index**

Similarly, there exists ARCH effect on the log return of S&P 500 time series. To capture this effect, GARCH model is also combined with the existing model. Having tested with different GARCH Model, it is found that eGARCH [[4]](#footnote-5)(1,1) fit the data series the best. The fitted model is shown as below:

,

,

,

From the R output below, as the p-value of Ljung–Box test for the standardized residuals is 0.3959 and the squared standardized residuals is 0.6925 which are all above 0.05, these show that the residuals are not autocorrelated and do not exhibit ARCH effect so the model is adequate.

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Estimate** | **P-value** |
| MA1 | -0.062438 | 0.000212 |
| Omega | -0.316975 | 0 |
| Alpha1 | -0.149711 | 0 |
| Beta1 | 0.964378 | 0 |
| Gamma1 | 0.199993 | 0 |

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Intrepretation** |
| Residual | 0.395868 | Residuals are not autocorrelated |
| Squared Residual | 0.692483 | Residuals have no ARCH pattern |

Table 8: ARIMA(0,0,1)-eGARCH(1,1) Model for S&P 500 Index:

### **3.2.3 FTSE 100 Index**

To capture the ARCH effect for log return of FTSE 100 time series, GARCH model is also combined with the existing model. Having tested with different GARCH Model, it is found that eGARCH (1,1) fit the data series the best. As the p-value for AR4 term is greater than 0.05, the term has been dropped. The fitted model is shown as below:

,

,

,

From the R output below, as the p-value of Ljung–Box test for the standardized residuals is 0.606153 and the squared standardized residuals is 0.151819 which are all above 0.05, these show that the residuals are not autocorrelated and do not exhibit ARCH effect so the model is adequate.

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Estimate** | **P-value** |
| AR1 | -0.35307 | 0 |
| AR2 | -0.44693 | 0 |
| AR3 | 0.57066 | 0 |
| MA1 | 0.34789 | 0 |
| MA2 | 0.43625 | 0 |
| MA3 | -0.58221 | 0 |
| Omega | -0.20495 | 0 |
| Alpha1 | -0.12936 | 0 |
| Beta1 | 0.97756 | 0 |
| Gamma1 | 0.14935 | 0 |

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Intrepretation** |
| Residual | 0.606153 | Residuals are not autocorrelated |
| Squared Residual | 0.151819 | Residuals have no ARCH pattern |

Table 9: ARIMA(3,0,3)-eGARCH(1,1) Model for FTSE100 Index:

### **3.2.4 SSE Index**

To capture the ARCH effect for log return of SSE Index time series, GARCH model is also combined with the existing model. Having tested with different GARCH Model, it is found that sGARCH (1,1) fit the data series the best. The fitted model is shown as below:

,

,

,

From the R output below, as the p-value of Ljung–Box test for the standardized residuals is 0.0338 which is below 0.05 while the squared standardized residuals is 0.172990 which is above 0.05, these show that the residuals do not exhibit ARCH effect but are autocorrelated so the model is not adequate.

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Estimate** | **P-value** |
| AR1 | 0.587176 | 0.006751 |
| AR2 | -0.703278 | 0.001649 |
| MA1 | -0.559731 | 0.009952 |
| MA2 | 0.684215 | 0.002292 |
| MA3 | 0.050076 | 0.010260 |
| Alpha1 | 0.066595 | 0.050493 |
| Beta1 | 0.932293 | 0 |

|  |  |  |
| --- | --- | --- |
| **Variable** | **P-Value** | **Intrepretation** |
| Residual | 0.033796 | Residuals are not autocorrelated |
| Squared Residual | 0.172990 | Residuals have no ARCH pattern |

Table 10: ARIMA(2,0,3)-sGARCH(1,1) Model for SSE Index:

## **Multivariate Autoregressive Moving Average (ARMA) Models**

In this section, we are trying to develop a multivariate vector-ARMA model for predicting the market indexes close. As our focus is on Hang Seng Index, three indexes pairs (HSI vs SSE; HSI vs S&P 500; HSI vs FTSE100) will be considered in our analysis.

### **3.3.1 Vector Autoregressive (VAR) Model**

Firstly, we are trying to explore multivariate vector-AR model to examine the relationship between different variables over time. VAR Model is useful as it can predict multiple time series variable using a single model. The finite VAR(q) model has the general representation as shown below

AIC is calculated below for each of the pairs with different lag order.

|  |  |  |  |
| --- | --- | --- | --- |
| **AIC** | | | |
| **Model** | **HSI & SSE** | **HSI & SP500** | **HSI & FTSE100** |
| VAR(1) | -16.89061 | -17.34055 | -17.49471 |
| VAR(2) | -16.89008 | -17.36437 | -17.50082 |
| VAR(3) | -16.8945 | -17.36618\* | -17.50413 |
| VAR(4) | -16.89977\* | -17.36527 | -17.51064 |
| VAR(5) | -16.89855 | -17.36388 | -17.51205\* |

Table 9. AIC of estimated VAR models

The residuals of the model were checked by the Ljung-Box test to see if they are white noise. If so, the VMA model should be adequate.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: Residuals of HSI & SSE** | | | | |
|  | **Residuals of HSI** | | **Residuals of SSE** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VAR(1) | 18.103 | 0.079221 | 32.558 | 0.00062 |
| VAR(2) | 17.497 | 0.064065 | 31.892 | 0.000417 |
| VAR(3) | 15.994 | 0.067007 | 22.762 | 0.006754 |
| VAR(4) | 12.737 | 0.121225 | 11.376 | 0.181291 |
| VAR(5) | 12.519 | 0.084732 | 6.3541 | 0.499064 |

Table 10. Results of Ljung-Box Test for residuals of VAR models of HSI and SSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: Residuals of HSI & SP500** | | | | |
|  | **Residuals of HSI** | | **Residuals of SP500** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VAR(1) | 43.6 | 8.536e-06 | 32.82 | 5.627e-04 |
| VAR(2) | 41.781 | 8.202e-06 | 32.935 | 2.793e-04 |
| VAR(3) | 37.427 | 2.208e-05 | 30.981 | 2.983e-04 |
| VAR(4) | 36.509 | 1.417e-05 | 29.088 | 3.060e-04 |
| VAR(5) | 34.066 | 1.674e-05 | 26.846 | 3.552e-04 |

Table 11. Results of Ljung-Box Test for residuals of VAR models of HSI and SP500

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: Residuals of HSI & FTSE100** | | | | |
|  | **Residuals of HSI** | | **Residuals of FTSE100** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VAR(1) | 35.062 | 2.420e-04 | 72.497 | 4.085e-11 |
| VAR(2) | 37.046 | 5.557e-05 | 58.019 | 8.574e-09 |
| VAR(3) | 34.847 | 6.340e-05 | 54.62 | 1.439e-08 |
| VAR(4) | 30.537 | 1.698e-04 | 28.621 | 3.695e-04 |
| VAR(5) | 32.527 | 3.242e-05 | 27.866 | 2.325e-04 |

Table 12. Results of Ljung-Box Test for residuals of VAR models of HSI and FTSE

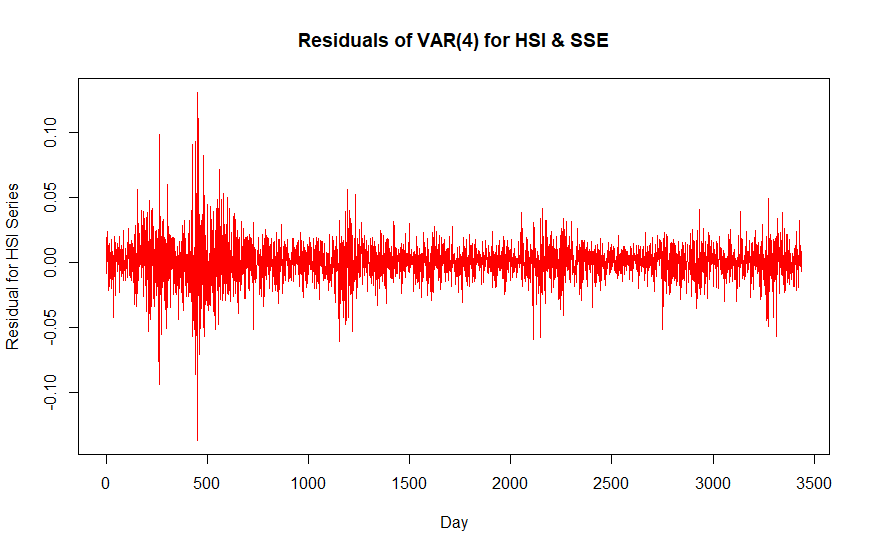
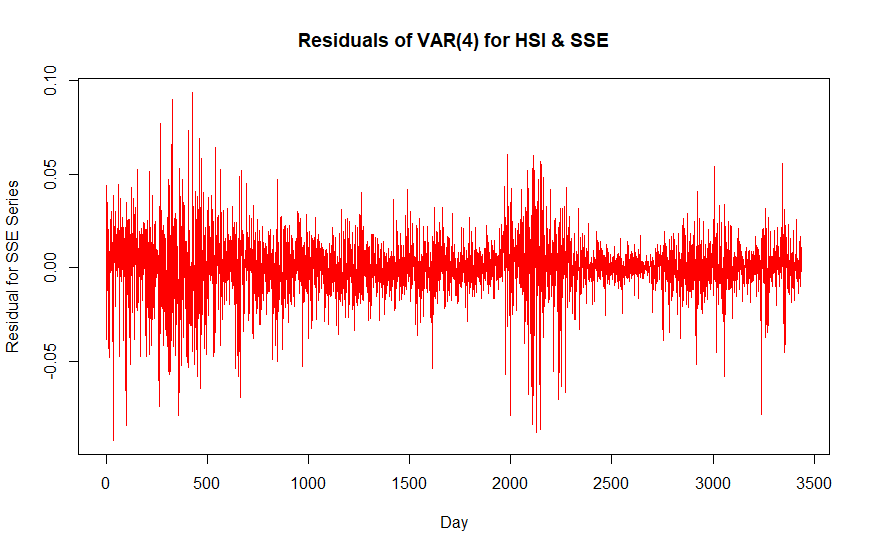
 

Figure 5. Residuals of the selected VAR(4) model for HSI & SSE

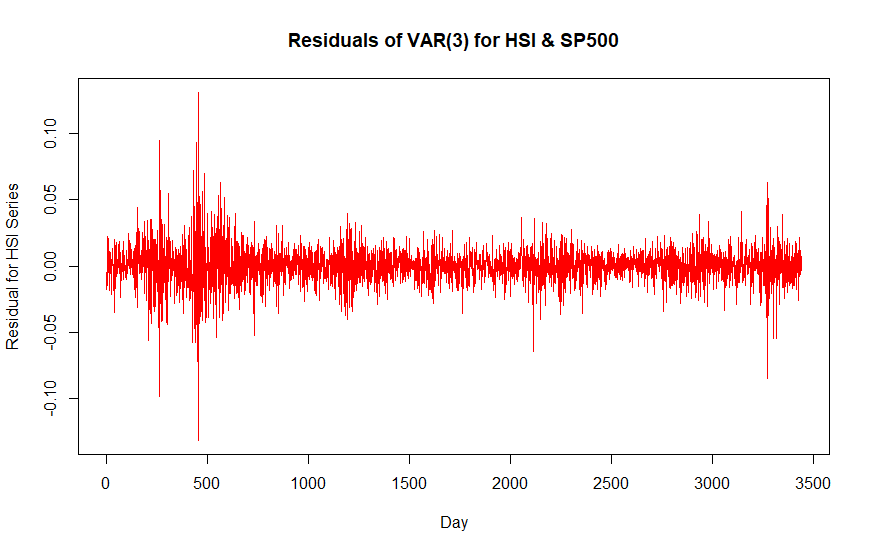
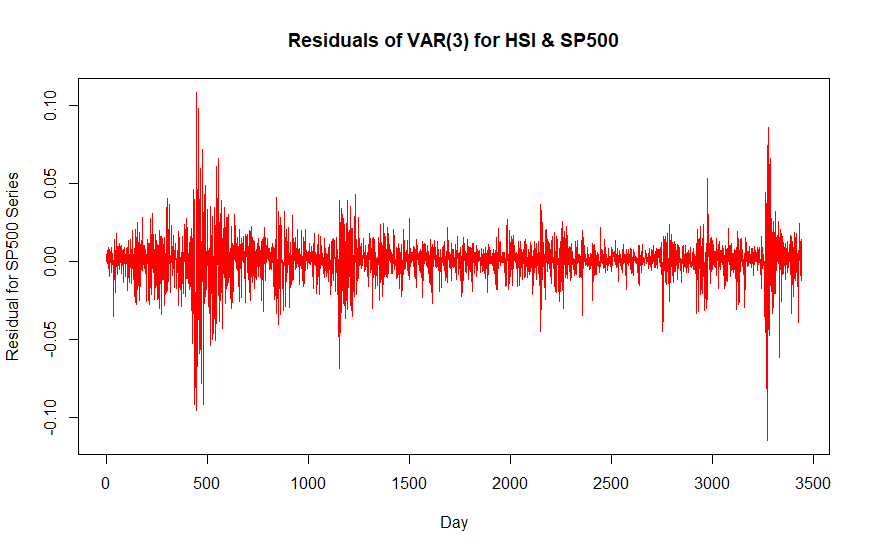
 

Figure 6. Residuals of the selected VAR(3) model for HSI & SP500 – Rejected

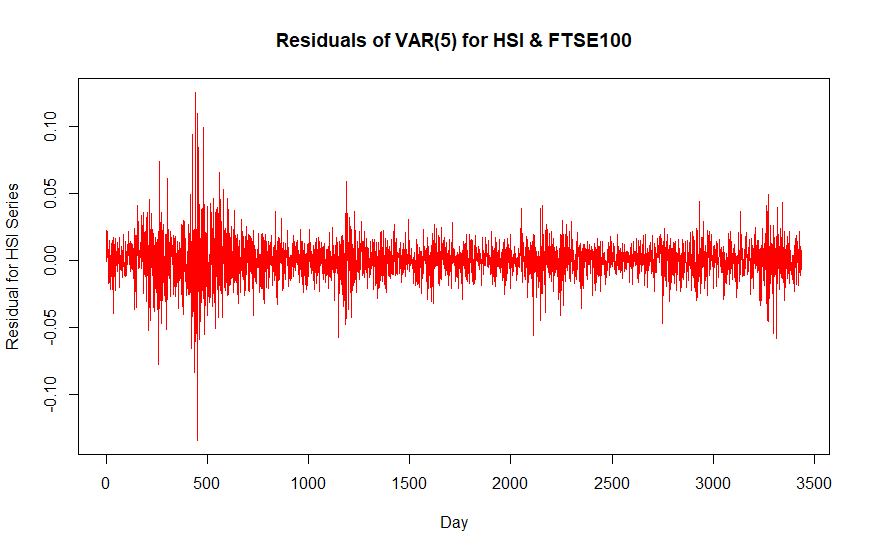
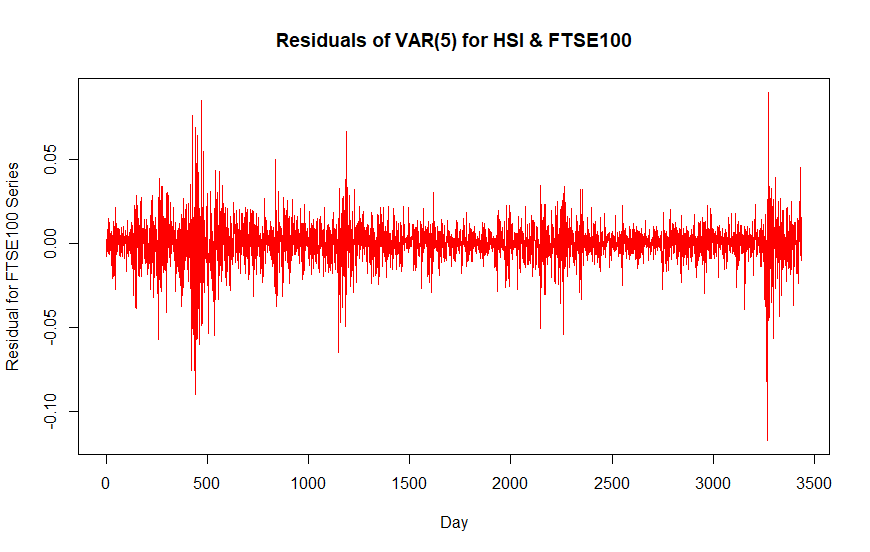
 

Figure 7. Residuals of the selected VAR(5) model for HSI & FTSE100 - Rejected

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **SSE** | **SP500** | **FTSE100** |
| HSI | VAR(4) | VAR (3) - rejected | VAR(5) - rejected |

Table 13. The summary of selected VAR models after checked by Ljung-Box test

The selected VAR(4) model for HSI & SSE is

The selected VAR(3) model for HSI & SP500 is

The selected VAR(5) model for HSI & FTSE100 is

From the Ljung-Box test result of the HSI’s residuals from fitted model, there is no significant evidence showing that VAR model is adequate to predict HSI from SP500 and FSET100 indexes.

On the other hand, for HSI and SSE pair, p-value of HSI’s and SSE’s residuals are 0.1212 and 0.1813 respectively, both are larger than 0.05, the VAR model is adequate to predict HSI from SSE.

### **3.3.2 Vector Moving Average (VMA) model**

The VAR model discussed above is useful in examining the associations between different variables. However, it cannot describe the relationship between a group of variables and the corresponding shocks at different time periods. Therefore, another finite multivariate time series model, namely vector moving average (VMA) model, can be adopted to determine the relationship between the data series and its various shocks at time lags. The finite VMA(q) model has the general representation as shown below

In this section, we performed the conditional maximum likelihood estimation to estimate three VMA models of the bivariate system composed of HSI and other indexes, which are HSI & SSE, HSI & SP500 and HSI & FTSE100.

|  |  |  |  |
| --- | --- | --- | --- |
| **AIC** | | | |
| **Model** | **HSI & SSE** | **HSI & SP500** | **HSI & FTSE100** |
| VMA(1) | -16.89109 | -16.88395\* | -17.50169 |
| VMA(2) | -16.89069 | -16.87641 | -17.50264 |
| VMA(3) | -16.89497 | -16.87356 | -17.5051 |
| VMA(4) | -16.89972\* | -16.87117 | -17.51029 |
| VMA(5) | -16.89932 | -16.86363 | -17.51233\* |

Table 14. AIC of estimated VMA models

The process of testing the adequacy of the VMA(q) model is also like testing the adequacy of VAR(p) mode. The residuals of the model were checked by the Ljung-Box test to see if they are white noise. If so, the VMA(q) model should be adequate.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: Residuals of HSI & SSE** | | | | |
|  | **Residuals of HSI** | | **Residuals of SSE** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VMA(1) | 17.941 | 0.1175 | 33.178 | 0.0009081 |
| VMA(2) | 16.99 | 0.1082 | 30.628 | 0.001262 |
| VMA(3) | 15.328 | 0.1206 | 21.052 | 0.02073 |
| VMA(4) | 5.9029 | 0.7496 | 13.261 | 0.1512 |
| VMA(5) | 4.112 | 0.8469 | 5.5326 | 0.6994 |

Table 15. Results of Ljung-Box Test for residuals of VMA models of HSI and SSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: Residuals of HSI & SP500** | | | | |
|  | **Residuals of HSI** | | **Residuals of SP500** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VMA(1) | 39.03 | 0.0001041 | 35.666 | 0.0003666 |
| VMA(2) | 37.883 | 8.188e-05 | 32.667 | 0.0005957 |
| VMA(3) | 31.82 | 0.000429 | 27.771 | 0.001964 |
| VMA(4) | 12.17 | 0.2039 | 21.558 | 0.01039 |
| VMA(5) | 10.333 | 0.2425 | 21.086 | 0.006923 |

Table 16. Results of Ljung-Box Test for residuals of VMA models of HSI and SP500

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: Residuals of HSI & FTSE100** | | | | |
|  | **Residuals of HSI** | | **Residuals of FTSE100** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VMA(1) | 35.321 | 0.0004162 | 69.93 | 3.302e-10 |
| VMA(2) | 37.4 | 9.873e-05 | 59.437 | 1.18e-08 |
| VMA(3) | 32.968 | 0.0002758 | 46.545 | 1.142e-06 |
| VMA(4) | 18.018 | 0.03496 | 24.088 | 0.004164 |
| VMA(5) | 15.882 | 0.0441 | 17.898 | 0.02201 |

Table 17. Results of Ljung-Box Test for residuals of VMA models of HSI and FTSE100

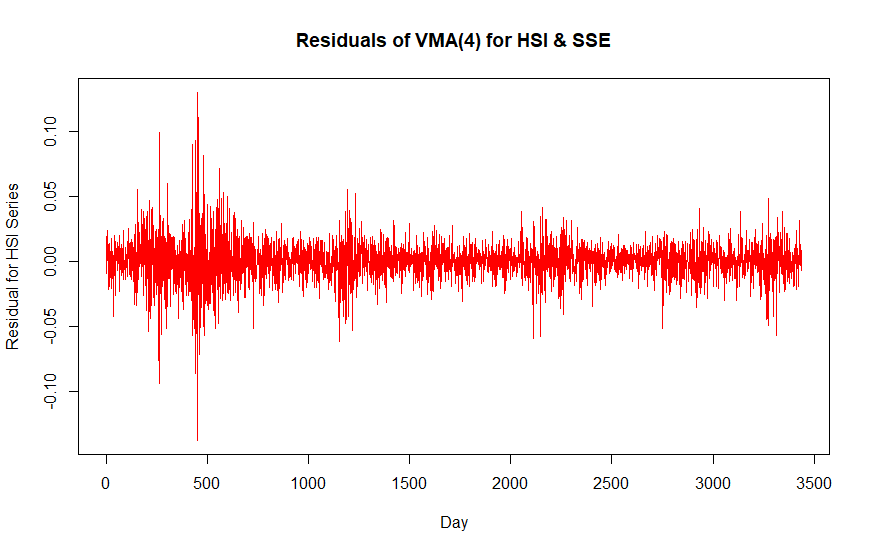
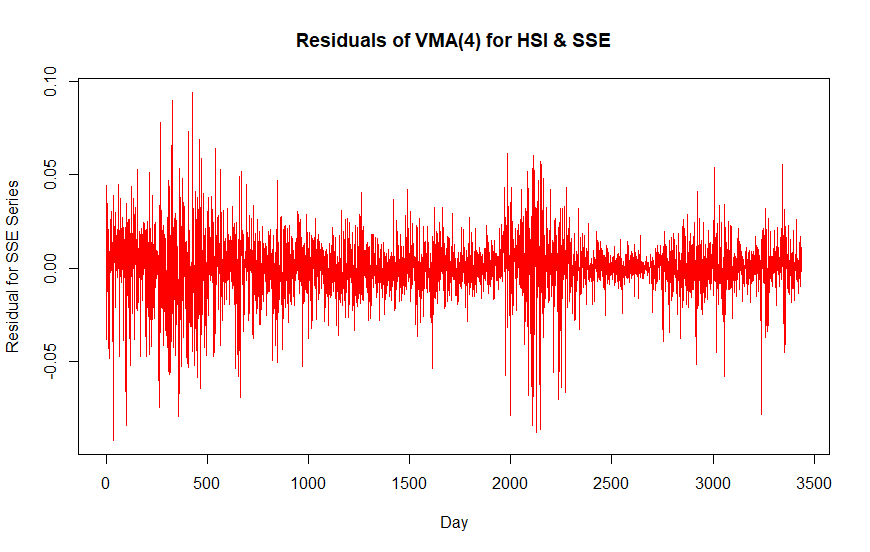
 

Figure 8. Residuals of the selected VMA(4) model for HSI & SSE

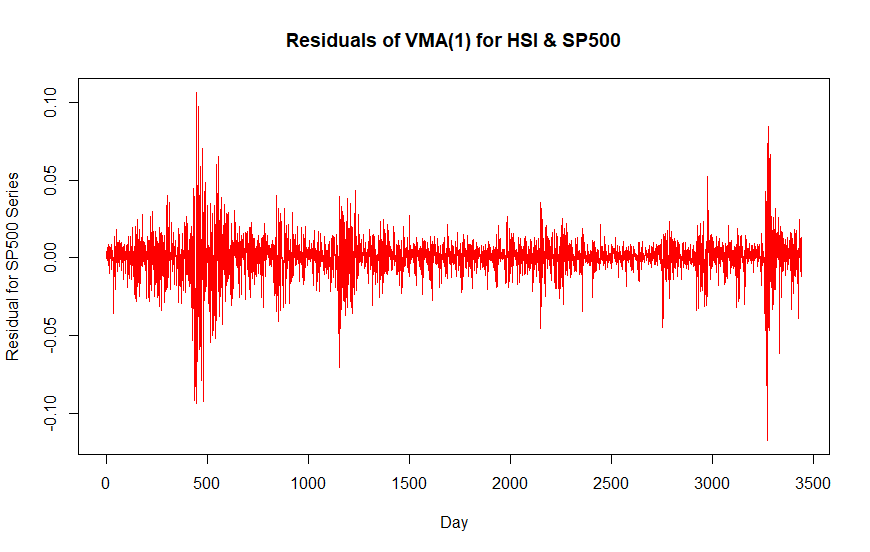
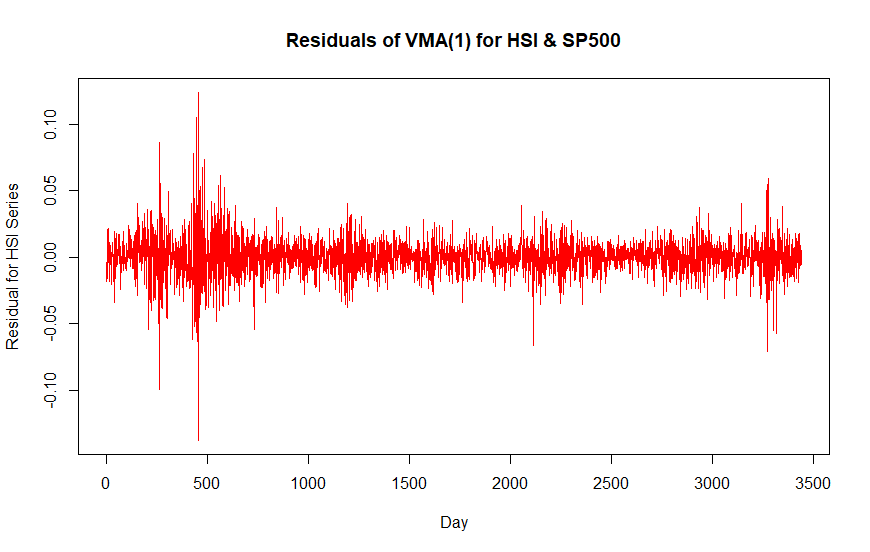


Figure 9. Residuals of the selected VMA(1) model for HSI & SP500 - Rejected

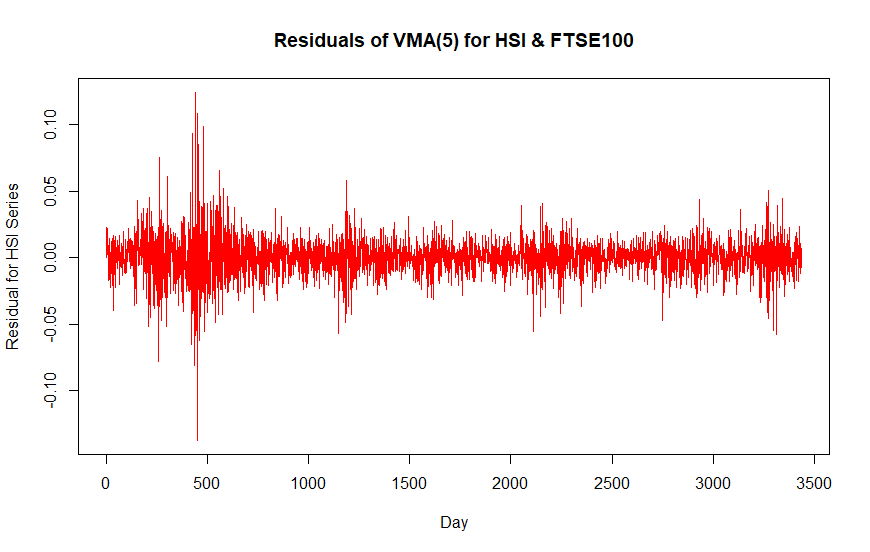
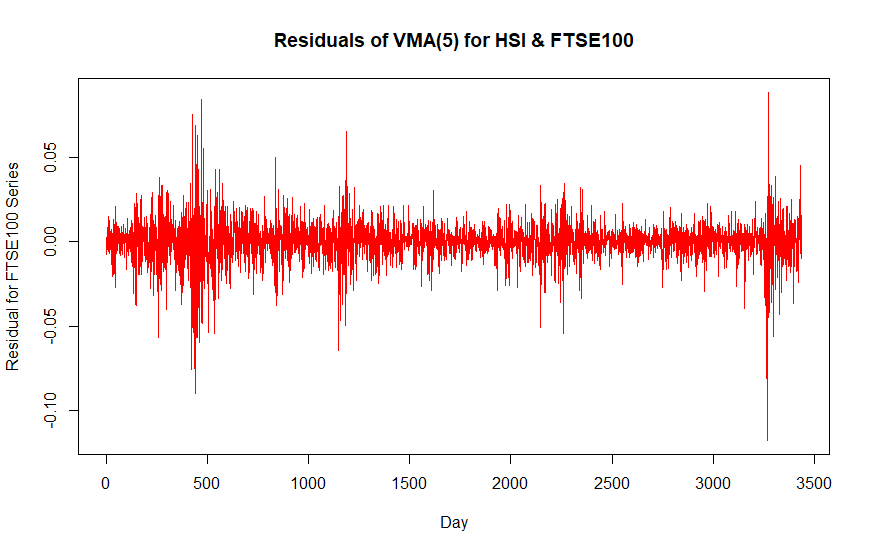
 

Figure 10. Residuals of the selected VMA(5) model for HSI & FTSE100 - Rejected

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **SSE** | **SP500** | **FTSE100** |
| HSI | VMA(4) | VMA(1) - rejected | VMA(5) - rejected |

Table 18. The summary of selected VMA models after checked by Ljung-Box test

The selected VMA(4) model for the bivariate series of HSI & SSE is:

The selected VMA(1) model for the bivariate series of HSI & SP500 is:

The selected VMA(5) model for the bivariate series of HSI & FTSE100 is:

The result of Ljung-Box test for the residuals of the VMA models is shown in Table 10, 11 and 12. In table 11 and 12, it clearly indicates that there is no significant evidence showing the VMA models are adequate to represent the bivariate series of HSI & SP500 pairs and HSI & FTSE100 pairs. Meanwhile, the results shown in Table 10 indicates that both VMA(4) and VMA(5) are adequate to model the bivariate series of HSI & SSE. In overall consideration, the VMA(4) model become our best selected VMA model for forecasting HSI with SSE as its AIC is -16.89972 and the corresponding p-values of HSI’s residuals and SSE’s residuals are 0.7496 and 0.1512 in the test.

### **3.3.3 Vector-Autoregressive Moving Average (VARMA) Model**

In this section, we will apply VARMA model to see if the model can benefit from the VAR and VMA models to examine the relationship between different variables and the shocks over time. The finite VARMA(p,q) model has the general representation as shown below

Where and

AIC is calculated below for each of the pairs with different lag order.

|  |  |  |  |
| --- | --- | --- | --- |
| **AIC** | | | |
| **Model** | **HSI & SSE** | **HSI & SP500** | **HSI & FTSE100** |
| VARMA (1,1) | -16.89105 | -17.36786 | -17.50045 |
| VARMA (1,2)\* | -16.89861 | -17.36581 | -17.50428 |
| VARMA (2,1) | -16.89393 | -17.36581 | -17.50366 |
| VARMA (2,2) | -16.89839 | -17.36654 | -17.51026 |

Table 19. AIC of estimated VARMA models

The residuals of the model were checked by the Ljung-Box test to see if they are white noise. If so, the VARMA model should be adequate.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: HSI & SSE** | | | | |
|  | **Residuals of HSI** | | **Residuals of SSE** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VARMA (1,1) | 17.535 | 0.09302 | 32.876 | 0.000551 |
| VARMA (1,2) | 8.0163 | 0.6272 | 20.41 | 0.0256 |
| VARMA (2,1) | 16.738 | 0.08037 | 26.469 | 0.003159 |
| VARMA (2,2)\* | 9.5458 | 0.3885 | 9.2789 | 0.4119 |

Table 20. Results of Ljung-Box Test for residuals of VARMA models of HSI and SSE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: HSI & SP500** | | | | |
|  | **Residuals of HSI** | | **Residuals of SP500** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VARMA (1,1) | 36.543 | 0.0001374 | 32.76 | 0.0005755 |
| VARMA (1,2) | 34.558 | 0.0001485 | 26.483 | 0.003142 |
| VARMA (2,1) | 32.748 | 0.0003002 | 27.157 | 0.002459 |
| VARMA (2,2) | 16.177 | 0.06328 | 22.25 | 0.00812 |

Table 21. Results of Ljung-Box Test for residuals of VARMA models of HSI and SP500

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ljung-Box test: HSI & FTSE100** | | | | |
|  | **Residuals of HSI** | | **Residuals of FTSE100** | |
| **Model** | **X-squared** | **p-value** | **X-squared** | **p-value** |
| VARMA (1,1) | 33.701 | 0.0004047 | 63.301 | 2.246e-09 |
| VARMA (1,2) | 38.796 | 2.756e-05 | 39.978 | 1.71e-05 |
| VARMA (2,1) | 41.119 | 1.075e-05 | 46.935 | 9.705e-07 |
| VARMA (2,2) | 27.313 | 0.001242 | 26.064 | 0.001994 |

Table 22. Results of Ljung-Box Test for residuals of VARMA models of HSI and FTSE100

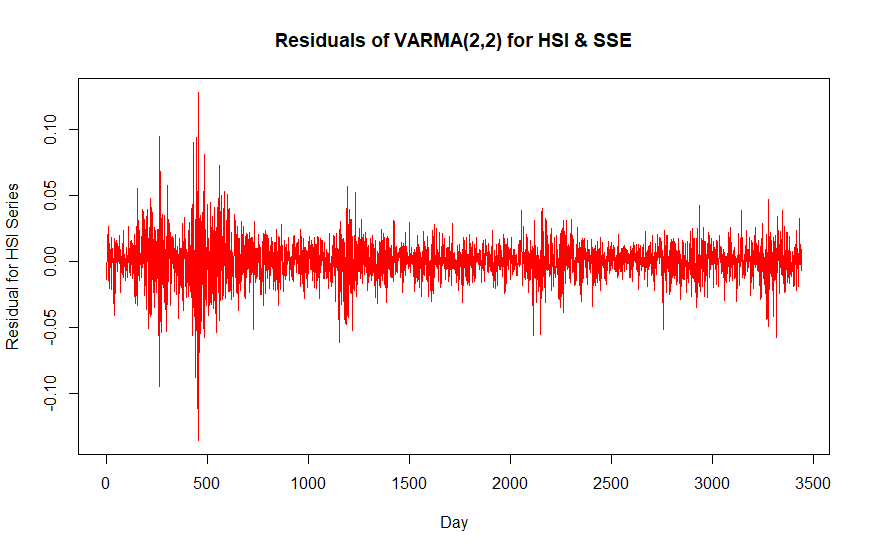
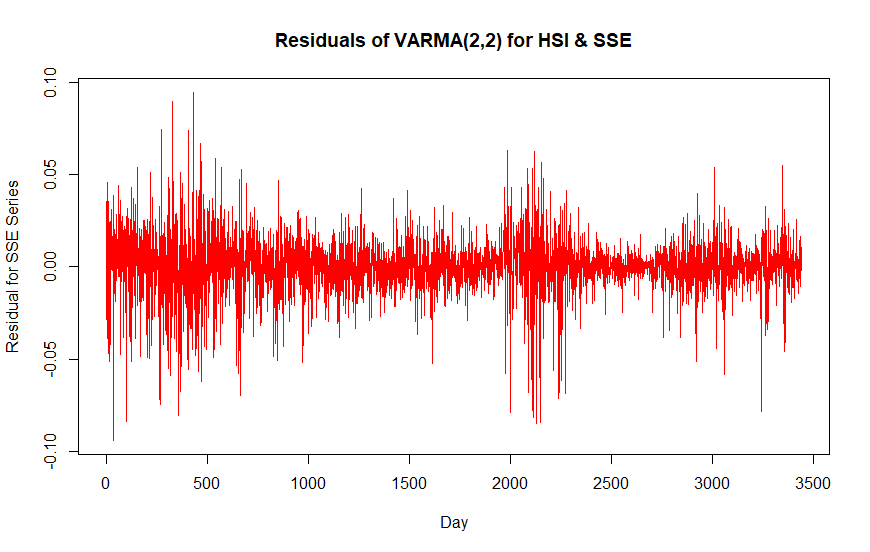
 

Figure 11. Residuals of the selected VARMA(2,2) model for HSI & SSE

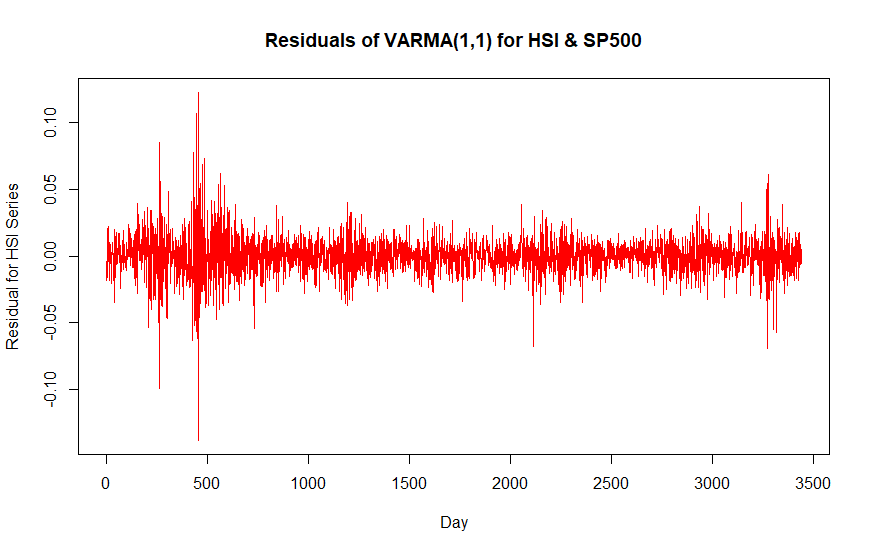
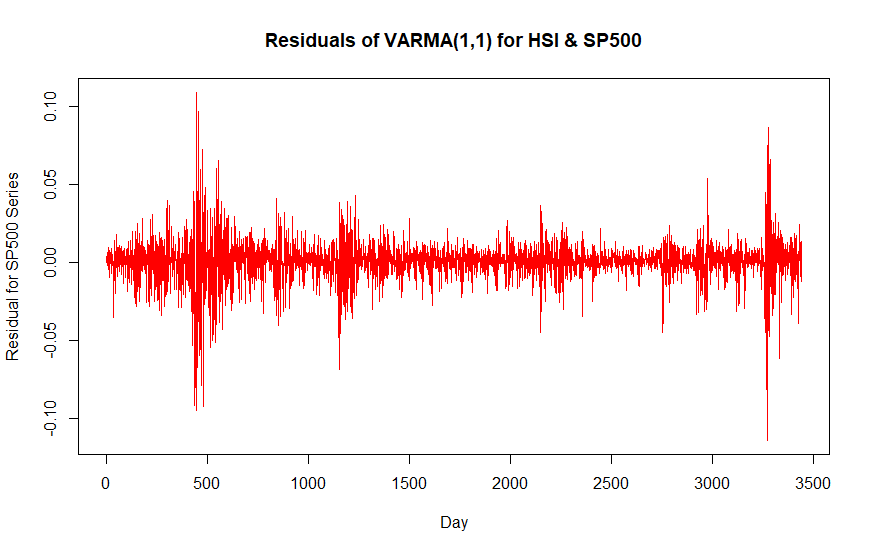
 

Figure 12. Residuals of the selected VARMA(1,1) model for HSI & SP500 - Rejected

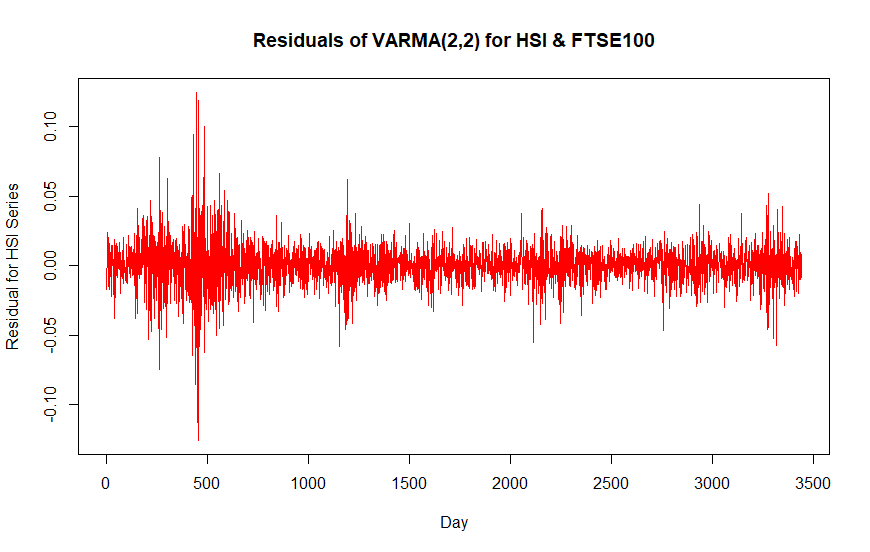
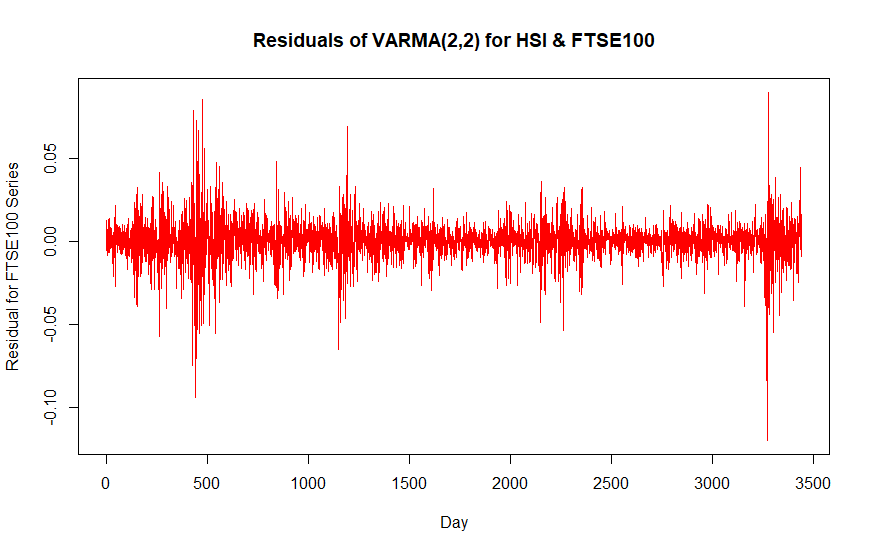
 

Figure 13. Residuals of the selected VARMA(2,2) model for HSI & FTSE100 - Rejected

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **SSE** | **SP500** | **FTSE100** |
| HSI | VARMA(2,2) | VARMA(1,1)- rejected | VARMA(2,2)- rejected |

Table 23. The summary of selected VARMA models after checked by Ljung-Box test

The selected VARMA(2,2) mode for HSI & SSE is

The selected VARMA(1,1) mode for HSI & SP500 is

The selected VARMA(2,2) mode for HSI & FTSE100 is

From the Ljung-Box test result of the HSI’s residuals from fitted model, there is no significant evidence showing that VARMA model is adequate to predict HSI from SP500 and FSET100 indexes.

On the other hand, for HSI and SSE pair, p-value of HSI’s and SSE’s residuals are 0.3885 and 0.4119 respectively, both are larger than 0.05, the VARMA model is adequate to predict HSI from SSE.

### **3.3.4 Model Summary**

The section summarized the model result from VAR, VMA, and VARMA for 3 pairs of indexes.

|  |  |  |  |
| --- | --- | --- | --- |
| Index | VAR | VMA | VARMA |
| HSI & SSE | VAR(4) | VMA(4) | VARMA(2,2) |
| HSI & SP500 | No adequate model found | | |
| HSI & FTSE100 | No adequate model found | | |

Table 24: Table summary of VAR, VMA, VARMA model

## **Multivariate GARCH Model – BEKK**

There are two most widely used multivariate GARCH models namely Dynamic Conditional Correlation (DCC) and BEKK. As there are more parameters in the BEKK Model, we believe it can explain the information hidden in the market indexes in a better manner. As a result, BEKK (1,1) is chosen in our analysis.

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Estimate** | **P-value** |
| Mu1\_HSI | 0.000362 | 0.272229 |
| Mu2\_SSE | 0.000101 | 0.721910 |
| A011 | 0.002305 | 0.024950 |
| A021 | 0.001359 | 0.165655 |
| A022 | 0.001740 | 2.49E-09 |
| A11 | 0.180948 | 0.034314 |
| A21 | -0.158708 | 0.002545 |
| A12 | 0.074496 | 0.343443 |
| A22 | 0.453980 | 2.66E-13 |
| B11 | 0.959223 | <2.22E-16 |
| B21 | 0.032510 | 0.054280 |
| B12 | -0.011267 | 0.760017 |
| B22 | 0.894866 | <2.22E-16 |

Table 25: BEKK-GARCH Model for HSI-SSE Index

# **TRADING STRATEGIES AND BACKTESTING RESULT**

In this section, we will compare the portfolio performance between the buy-and-hold strategy and trading strategies using the fitted time series model. The strategy is carried out as follows:

1. The time series model is used to predict for the next day returns
2. If the predicted return is less than the threshold, we will sell the index at previous close level.
3. If the predicted return is higher than the threshold, we will buy index at previous close level.
4. If the predicted return is within the threshold, no action is done.

## **Implementation of the strategy**

To implement this strategy, we need to calibrate the parameters of the ARIMA-GARCH model every day with the rolling window of 2-years as the new information comes in. The model can then predict tomorrow’s return with the updated parameters. From the earlier section, it is found that the ARIMA-GARCH model fit the S&P 500 the best. As a result, we will illustrate this strategy using S&P 500 index.

## **Back-testing Result –S&P 500 Index**

From the graph below, it can be observed that the returns for the trading strategy using ARIMA-GARCH model as the signal beat the performance from the buy-and-hold strategy. The reason why there is no big difference before 2020 for S&P 500 index mainly because of the effect from Quantitative Easing which started from 2009. Quantitative Easing pushes down the interest and thus lowers the returns investors can get from the low-risk investments including Government bonds. Investors are then forced into relatively riskier investments to find stronger returns and thus pushing up stock market prices. As ARIMA-GARCH model is trying to predict the shocks at different time periods, it will not perform very well under the Quantitative Easing environment.

However, it is noted that the ARIMA-GARCH model performs very well during COVID-19 which the ARIMA-GARCH strategy has significantly outperformed buy-and-hold strategy. This makes sense since it is likely to have significant serial correlation during this period which can be well captured by the ARIMA-GARCH models. It is expected that the model performance begins to suffer again when the market gets recovered from COVID 19.

Figure 14: ARIMA-GARCH strategy outperform buy-and-hold strategy

# **DISCUSSION**

## **Data Limitation**

The index is using a basket of securities to replicate the market performance. When the underlying securities have corporate action (e.g. dividend payment, right issues), the effect may not be incorporated in the research. As a result, the analysis may not reflect the total returns of the investment. Apart from this, there are time-zone difference between Asian, European, and American index closing times. The Asian market index may not be adjusted for the news which arrive at U.S. hours. This will possibly impact the analysis result including the cointegration test.

The analysis is based on the closing price of the market indexes without taking into the consideration of the transaction cost which plays as one of the major roles in the investment decisions. The transaction cost includes the brokerage fee and the stamp duty.

## **Model Limitation**

The biggest drawback of the standard ARCH and GARCH model is that the asymmetries of the volatility cannot be modelled with regards with the sign of past shocks. As a result, positive and negative shocks have the same effects on volatility. From the empirical studies in the financial market, it is found that the bad news has a bigger impact on the volatility compared with the good news. As a result, the model will likely be underperformed upon the big news.

Another disadvantage of the ARCH and GARCH model is that the model does not consider of what the market might do in the future. As the time series model is trying to forecast the future values from the past values, the model may not perform well under the unexpected events.

Furthermore, the computational costs of the VARMA Model and BEKK Model are very expensive which is difficult for us to develop the trading strategy using these models which involve continuously calibrating the parameters.

# **CONCLUSION**

In this paper, we had explored four main markets, HSI, SSE, SP500, FTSE100. Applied basic statistical analysis on the closing price of those four markets, we can observe log return of all market pass the Dickey-Fuller Test, meaning that they are stationary. Next, we applied Ljung-Box test on the log return of indexes, discovering that all of them have some component of serial correlation. We noticed that HSI had a greater p-value, meaning that almost all component is close to zero. Ljung-Box test was also conducted on the residuals are discovered that all the market indexes exhibit ARCH effect. Finally, we performed a cointegration test between market indexes and we are not able to reject the null hypothesis of no integration, meaning that there is some relation between different markets.

We had fitted both univariate and multivariate model into these four markets. The model we fitted for SP500 is MA(1), for FTSE100 is ARMA(4,3), for SSE is ARMA(2,3). There are no model fitted for HSI. The residuals of the fitted model are not white noise, meaning that it had GARCH effect. We had fit a iGARCH(1,1) to HSI index, eGARCH(1,1) to SP500, sGARCH (1,1) to SSE and eGARCH (1,1) FTSE index. Next, we explored multivariate vector ARMA model on indexes. We focus on HSI, and we consider three pairs, HSI vs SSE, HSI vs S&P500, HSI vs FTSE100. The most suitable model for HSI & SSE is VAR(4), VMA(4) and for HSI & SP500 is VAR(3), VMA(1), and for HSI & FTSE100 is VAR(5), VMA(5). After that we take advantage of both multivariate models and fit VARMA to three pair of indexes. Only the HSI & SSE pair is not rejected and the corresponding model is VARMA(2,2) and we then fit BEKK(1,1) multivariate GARCH model to HSI & SSE pair to obtain a BEKK-GARCH model result. We developed a strategy base on the model we obtained, and we did some back-testing on S&P500 using ARIMA-GARCH model and the strategy is outperforming the buy-and-hold strategy.

# **Reference**

1. US Fund Fee Study (Morningstar Research, Apr 2019). Available: https://www.morningstar.com/content/dam/marketing/shared/pdfs/Research/USFundFeeStudyApr2019.pdf
2. Index funds break through $10tn-in-assets mark amid active exodus (Financial Times, Jan 2020). Available: https://www.ft.com/content/a7e20d96-318c-11ea-9703-eea0cae3f0de
3. Caporale, Guglielmo Maria & Pittis, Nikitas & Spagnolo, Nicola. (2003). IGARCH models and structural breaks. Applied Economics Letters. 10. 765-768. 10.1080/1350485032000138403.
4. Changli He, Timo Teräsvirta and Hans Malmsten (2002), Moment Structure of a Family of First-Order Exponential GARCH Models https://www.jstor.org/stable/3533416?seq=1
5. Lütkepohl, H. (2006). Forecasting with VARMA models. *Handbook of economic forecasting, 1*, 287-325.
6. Ghani, I. M., & Rahim, H. A. (2019, June). Modeling and Forecasting of Volatility using ARMA-GARCH: Case Study on Malaysia Natural Rubber Prices. *In IOP Conference Series: Materials Science and Engineering (Vol. 548, No. 1, p. 012023).* IOP Publishing.
7. Dufour, J. M., & Pelletier, D. (2008). Practical methods for modelling weak VARMA processes: Identification, estimation and specification with a macroeconomic application. *Manuscript*, McGill University.
8. Simionescu, M. (2013). The use of VARMA models in forecasting macroeconomic indicators. *Economics & Sociology, 6(2),* 94.
9. Cromwell, J. B., & Terraza, M. (1994). *Multivariate tests for time series models* (No. 100). Sage.
10. Francq, C., & Zakoian, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, *10*(4), 605-637.
11. Zhou, J. (2009). *Modeling S&P 500 stock index using ARMA-asymmetric power ARCH models* (Doctoral dissertation, Master Thesis in Statistics. School of Economics and Social Science Hogskolan Dalarna, Sweden).
12. Salisu, A. A., & Oloko, T. F. (2015). Modeling oil price–US stock nexus: A VARMA–BEKK–AGARCH approach. *Energy Economics*, *50*, 1-12.
13. Quantstart (2015), ARIMA+GARCH Trading Strategy on the S&P500 Stock Market Index Using R
14. Harris, R. I. (1992). Testing for unit roots using the augmented Dickey-Fuller test: Some issues relating to the size, power and the lag structure of the test. *Economics letters*, *38*(4), 381-386.
15. Pham, D. T., Roy, R., & Cédras, L. (2003). Tests for non‐correlation of two cointegrated ARMA time series. *Journal of Time Series Analysis*, *24*(5), 553-577.
16. Solibakke, P. B. (2001). Efficiently ARMA–GARCH estimated trading volume characteristics in thinly traded markets. *Applied Financial Economics*, *11*(5), 539-556.
17. Chang, C. L., McAleer, M., & Tansuchat, R. (2011). Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics*, *33*(5), 912-923.
18. So, M. K., & Philip, L. H. (2006). Empirical analysis of GARCH models in value at risk estimation. *Journal of International Financial Markets, Institutions and Money*, *16*(2), 180-197.

# **Appendix**

|  |
| --- |
| R Script - EDA |
| install.packages('forecast') #for ARIMA forecasting  library(forecast)  HSI<-read.csv("./data/Hang Seng Historical Data.csv",header=T)  SSE<-read.csv("./data/Shanghai Historical Data.csv",header=T)  SP500<-read.csv("./data/SP500 Historical Data.csv",header=T)  FTSE<-read.csv("./data/FTSE100 Historical Data.csv",header=T)  HSI\_df=HSI[dim(HSI)[1L]:1,]  SSE\_df=SSE[dim(SSE)[1L]:1,]  SP500\_df=SP500[dim(SP500)[1L]:1,]  FTSE\_df=FTSE[dim(FTSE)[1L]:1,]  #  Plot Index Close  plot(as.numeric(sub(",","",HSI\_df[,2]))~as.Date(HSI\_df[,1], "%d/%m/%Y"),xlab="Date",ylab=" ",main="Time plot for HSI Index",type="l",col="red",xlim=as.Date(c("2007-01-01", "2020-11-19")))  plot(as.numeric(sub(",","",SSE\_df[,2]))~as.Date(SSE\_df[,1], "%d/%m/%Y"),xlab="Date",ylab=" ",main="Time plot for SSE Index",type="l",col="red",xlim=as.Date(c("2007-01-01", "2020-11-19")))  plot(as.numeric(sub(",","",SP500\_df[,2]))~as.Date(SP500\_df[,1], "%d/%m/%Y"),xlab="Date",ylab=" ",main="Time plot for S&P 500 Index",type="l",col="red",xlim=as.Date(c("2007-01-01", "2020-11-19")))  plot(as.numeric(sub(",","",FTSE\_df[,2]))~as.Date(FTSE\_df[,1], "%d/%m/%Y"),xlab="Date",ylab=" ",main="Time plot for FTSE100 Index",type="l",col="red",xlim=as.Date(c("2007-01-01", "2020-11-19")))  #  Calculate Log Return  HSI\_log=diff(log(as.numeric(sub(",","",HSI\_df[,2]))))  SSE\_log=diff(log(as.numeric(sub(",","",SSE\_df[,2]))))  SP500\_log=diff(log(as.numeric(sub(",","",SP500\_df[,2]))))  FTSE\_log=diff(log(as.numeric(sub(",","",FTSE\_df[,2]))))  #  Plot Log Return  plot(HSI\_log~as.Date(HSI\_df[-1,1], "%d/%m/%Y"),xlab=" ",ylab=" ",main="Time plot for daily log returns of HSI Index",type="l",col="red",ylim=c(-0.2,0.2))  plot(SSE\_log~as.Date(SSE\_df[-1,1], "%d/%m/%Y"),xlab=" ",ylab=" ",main="Time plot for daily log returns of SSE Index",type="l",col="red",ylim=c(-0.2,0.2))  plot(SP500\_log~as.Date(SP500\_df[-1,1], "%d/%m/%Y"),xlab=" ",ylab=" ",main="Time plot for daily log returns of S&P 500 Index",type="l",col="red",ylim=c(-0.2,0.2))  plot(FTSE\_log~as.Date(FTSE\_df[-1,1], "%d/%m/%Y"),xlab=" ",ylab=" ",main="Time plot for daily log returns of FTSE100 Index",type="l",col="red",ylim=c(-0.2,0.2))  #  Plot ACF & PACF  acf(HSI\_log,20,main="ACF Plot for log return of HSI Index",col="red",ylim=c(-0.5,1))  acf(SSE\_log,20,main="ACF Plot for log return of SSE Index",col="red",ylim=c(-0.5,1))  acf(SP500\_log,20,main="ACF Plot for log return of S&P500 Index",col="red",ylim=c(-0.5,1))  acf(FTSE\_log,20,main="ACF Plot for log return of FTSE100 Index",col="red",ylim=c(-0.5,1))  pacf(HSI\_log,20,main="Partial ACF Plot for log return of HSI Index",col="red",ylim=c(-0.5,1))  pacf(SSE\_log,20,main="Partial ACF Plot for log return of SSE Index",col="red",ylim=c(-0.5,1))  pacf(SP500\_log,20,main="Partial ACF Plot for log return of S&P500 Index",col="red",ylim=c(-0.5,1))  pacf(FTSE\_log,20,main="Partial ACF Plot for log return of FTSE100 Index",col="red",ylim=c(-0.5,1))  #Test the Serial Correlation for HSI, SSE, FTSE and SP500  Box.test(HSI\_log,lag=10,type="Ljung")  Box.test(SSE\_log,lag=10,type="Ljung")  Box.test(SP500\_log,lag=10,type="Ljung")  Box.test(FTSE\_log,lag=10,type="Ljung")  #Test the ARCH Effect for HSI, SSE, FTSE and SP500  at1=HSI\_log-mean(HSI\_log)  Box.test(at1^2,lag=10,type="Ljung")  at2=SSE\_log-mean(SSE\_log)  Box.test(at2^2,lag=10,type="Ljung")  At3=SP500\_log-mean(SP500\_log)  Box.test(at3^2,lag=10,type="Ljung")  At4=FTSE\_log-mean(FTSE\_log)  Box.test(at4^2,lag=10,type="Ljung")  #Fit ARIMA Model for HSI, SSE, FTSE and SP500  auto.arima(HSI\_log)  auto.arima(SP500\_log)  auto.arima(SSE\_log)  auto.arima(FTSE\_log) |

|  |
| --- |
| R Script (VAR & VMA & VARMA.ipynb) |
| install.packages('rugarch') #for GARCH  install.packages('forecast') #for ARIMA forecasting  install.packages('tseries') #for Dickey-Fuller Test  install.packages('fGarch')  install.packages('MTS') #for Multivariate Time Series  install.packages('vars')  library(rugarch)  library(forecast)  library(tseries)  library(fGarch)  library(MTS)  library(vars)  library(ggplot2)  library(xts)  #read data  data = read.csv("./data/cleaned\_data.csv",header=T)  mts = xts(x = data[,2:7], order.by = as.Date(data[,1], "%d/%m/%Y"))  plot(mts)  #log return  mts\_return = diff(log(mts))[-1,]  plot(mts\_return)  #index pairs  hsi\_sse\_return = mts\_return[,c("HSI","SSE")]  hsi\_sp500\_return = mts\_return[,c("HSI","SP500")]  hsi\_ftse100\_return = mts\_return[,c("HSI","FTSE100")]  hsi\_dowjones\_return = mts\_return[,c("HSI","DowJones")]  hsi\_dax\_return = mts\_return[,c("HSI","DAX")]  plot(hsi\_sse\_return)  plot(hsi\_sp500\_return)  plot(hsi\_ftse100\_return)  plot(hsi\_dowjones\_return)  plot(hsi\_dax\_return)  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  #  fit VMA model  #HSI & SSE  fit\_vma\_hsi\_sse\_q1=VARMA(as.data.frame(hsi\_sse\_return),p=0,q=1,include.mean=FALSE)  fit\_vma\_hsi\_sse\_q2=VARMA(as.data.frame(hsi\_sse\_return),p=0,q=2,include.mean=FALSE)  fit\_vma\_hsi\_sse\_q3=VARMA(as.data.frame(hsi\_sse\_return),p=0,q=3,include.mean=FALSE)  fit\_vma\_hsi\_sse\_q4=VARMA(as.data.frame(hsi\_sse\_return),p=0,q=4,include.mean=FALSE)  fit\_vma\_hsi\_sse\_q5=VARMA(as.data.frame(hsi\_sse\_return),p=0,q=5,include.mean=FALSE)  #HSI & SP500  fit\_vma\_hsi\_sp500\_q1=VARMA(as.data.frame(hsi\_sp500\_return),p=0,q=1,include.mean=FALSE)  fit\_vma\_hsi\_sp500\_q2=VARMA(as.data.frame(hsi\_sp500\_return),p=0,q=2,include.mean=FALSE)  fit\_vma\_hsi\_sp500\_q3=VARMA(as.data.frame(hsi\_sp500\_return),p=0,q=3,include.mean=FALSE)  fit\_vma\_hsi\_sp500\_q4=VARMA(as.data.frame(hsi\_sp500\_return),p=0,q=4,include.mean=FALSE)  fit\_vma\_hsi\_sp500\_q5=VARMA(as.data.frame(hsi\_sp500\_return),p=0,q=5,include.mean=FALSE)  #HSI & FTSE100  fit\_vma\_hsi\_ftse100\_q1=VARMA(as.data.frame(hsi\_ftse100\_return),p=0,q=1,include.mean=FALSE)  fit\_vma\_hsi\_ftse100\_q2=VARMA(as.data.frame(hsi\_ftse100\_return),p=0,q=2,include.mean=FALSE)  fit\_vma\_hsi\_ftse100\_q3=VARMA(as.data.frame(hsi\_ftse100\_return),p=0,q=3,include.mean=FALSE)  fit\_vma\_hsi\_ftse100\_q4=VARMA(as.data.frame(hsi\_ftse100\_return),p=0,q=4,include.mean=FALSE)  fit\_vma\_hsi\_ftse100\_q5=VARMA(as.data.frame(hsi\_ftse100\_return),p=0,q=5,include.mean=FALSE)  #HSI & DowJones  fit\_vma\_hsi\_dowjones\_q1=VARMA(as.data.frame(hsi\_dowjones\_return),p=0,q=1,include.mean=FALSE)  fit\_vma\_hsi\_dowjones\_q2=VARMA(as.data.frame(hsi\_dowjones\_return),p=0,q=2,include.mean=FALSE)  fit\_vma\_hsi\_dowjones\_q3=VARMA(as.data.frame(hsi\_dowjones\_return),p=0,q=3,include.mean=FALSE)  fit\_vma\_hsi\_dowjones\_q4=VARMA(as.data.frame(hsi\_dowjones\_return),p=0,q=4,include.mean=FALSE)  fit\_vma\_hsi\_dowjones\_q5=VARMA(as.data.frame(hsi\_dowjones\_return),p=0,q=5,include.mean=FALSE)  #HSI & DAX  fit\_vma\_hsi\_dax\_q1=VARMA(as.data.frame(hsi\_dax\_return),p=0,q=1,include.mean=FALSE)  fit\_vma\_hsi\_dax\_q2=VARMA(as.data.frame(hsi\_dax\_return),p=0,q=2,include.mean=FALSE)  fit\_vma\_hsi\_dax\_q3=VARMA(as.data.frame(hsi\_dax\_return),p=0,q=3,include.mean=FALSE)  fit\_vma\_hsi\_dax\_q4=VARMA(as.data.frame(hsi\_dax\_return),p=0,q=4,include.mean=FALSE)  fit\_vma\_hsi\_dax\_q5=VARMA(as.data.frame(hsi\_dax\_return),p=0,q=5,include.mean=FALSE)  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  # Diagnostic Checking of a VMA Model  #HSI & SSE  apply(fit\_vma\_hsi\_sse\_q1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q2$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q3$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q4$residuals,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q5$residuals,2,Box.test,lag=8 ,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q1$residuals^2,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q2$residuals^2,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q3$residuals^2,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q4$residuals^2,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_sse\_q5$residuals^2,2,Box.test,lag=8 ,type="Ljung")  #HSI & SP500  apply(fit\_vma\_hsi\_sp500\_q1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q2$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q3$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q4$residuals,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q5$residuals,2,Box.test,lag=8 ,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q1$residuals^2,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q2$residuals^2,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q3$residuals^2,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q4$residuals^2,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_q5$residuals^2,2,Box.test,lag=8 ,type="Ljung")  #HSI & FTSE100  apply(fit\_vma\_hsi\_ftse100\_q1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q2$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q3$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q4$residuals,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q5$residuals,2,Box.test,lag=8 ,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q1$residuals^2,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q2$residuals^2,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q3$residuals^2,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q4$residuals^2,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_q5$residuals^2,2,Box.test,lag=8 ,type="Ljung")  #HSI & DowJones  apply(fit\_vma\_hsi\_dowjones\_q1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q2$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q3$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q4$residuals,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q5$residuals,2,Box.test,lag=8 ,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q1$residuals^2,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q2$residuals^2,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q3$residuals^2,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q4$residuals^2,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_q5$residuals^2,2,Box.test,lag=8 ,type="Ljung")  #HSI & DAX  apply(fit\_vma\_hsi\_dax\_q1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q2$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q3$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q4$residuals,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q5$residuals,2,Box.test,lag=8 ,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q1$residuals^2,2,Box.test,lag=12,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q2$residuals^2,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q3$residuals^2,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q4$residuals^2,2,Box.test,lag=9 ,type="Ljung")  apply(fit\_vma\_hsi\_dax\_q5$residuals^2,2,Box.test,lag=8 ,type="Ljung")  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  # Plot residuals of selected VMA models  #HSI & SSE  plot(fit\_vma\_hsi\_sse\_q4$residuals[,1],type="l", col="red", main = "Residuals of VMA(4) for HSI & SSE", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_vma\_hsi\_sse\_q4$residuals[,2],type="l", col="red", main = "Residuals of VMA(4) for HSI & SSE", xlab="Day", ylab="Residual for SSE Series")  #HSI & SP500  plot(fit\_vma\_hsi\_sp500\_q1$residuals[,1],type="l", col="red", main = "Residuals of VMA(1) for HSI & SP500", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_vma\_hsi\_sp500\_q1$residuals[,2],type="l", col="red", main = "Residuals of VMA(1) for HSI & SP500", xlab="Day", ylab="Residual for SP500 Series")  #HSI & FTSE100  plot(fit\_vma\_hsi\_ftse100\_q5$residuals[,1],type="l", col="red", main = "Residuals of VMA(5) for HSI & FTSE100", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_vma\_hsi\_ftse100\_q5$residuals[,2],type="l", col="red", main = "Residuals of VMA(5) for HSI & FTSE100", xlab="Day", ylab="Residual for FTSE100 Series")  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  ####VAR####  #HSI & SSE  fit\_var\_hsi\_sse\_p1=VARMA(as.data.frame(hsi\_sse\_return),p=1,q=0,include.mean=FALSE)  fit\_var\_hsi\_sse\_p2=VARMA(as.data.frame(hsi\_sse\_return),p=2,q=0,include.mean=FALSE)  fit\_var\_hsi\_sse\_p3=VARMA(as.data.frame(hsi\_sse\_return),p=3,q=0,include.mean=FALSE)  fit\_var\_hsi\_sse\_p4=VARMA(as.data.frame(hsi\_sse\_return),p=4,q=0,include.mean=FALSE)  fit\_var\_hsi\_sse\_p5=VARMA(as.data.frame(hsi\_sse\_return),p=5,q=0,include.mean=FALSE)  #HSI & SP500  fit\_var\_hsi\_sp500\_p1=VARMA(as.data.frame(hsi\_sp500\_return),p=1,q=0,include.mean=FALSE)  fit\_var\_hsi\_sp500\_p2=VARMA(as.data.frame(hsi\_sp500\_return),p=2,q=0,include.mean=FALSE)  fit\_var\_hsi\_sp500\_p3=VARMA(as.data.frame(hsi\_sp500\_return),p=3,q=0,include.mean=FALSE)  fit\_var\_hsi\_sp500\_p4=VARMA(as.data.frame(hsi\_sp500\_return),p=4,q=0,include.mean=FALSE)  fit\_var\_hsi\_sp500\_p5=VARMA(as.data.frame(hsi\_sp500\_return),p=5,q=0,include.mean=FALSE)  #HSI & FTSE100  fit\_var\_hsi\_ftse100\_p1=VARMA(as.data.frame(hsi\_ftse100\_return),p=1,q=0,include.mean=FALSE)  fit\_var\_hsi\_ftse100\_p2=VARMA(as.data.frame(hsi\_ftse100\_return),p=2,q=0,include.mean=FALSE)  fit\_var\_hsi\_ftse100\_p3=VARMA(as.data.frame(hsi\_ftse100\_return),p=3,q=0,include.mean=FALSE)  fit\_var\_hsi\_ftse100\_p4=VARMA(as.data.frame(hsi\_ftse100\_return),p=4,q=0,include.mean=FALSE)  fit\_var\_hsi\_ftse100\_p5=VARMA(as.data.frame(hsi\_ftse100\_return),p=5,q=0,include.mean=FALSE)  # Diagnostic Checking of a VAR Model  #HSI & SSE  apply(fit\_var\_hsi\_sse\_p1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sse\_p2$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sse\_p3$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sse\_p4$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_var\_hsi\_sse\_p5$residuals,2,Box.test,lag=10,type="Ljung")  #HSI & SP500  apply(fit\_var\_hsi\_sp500\_p1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sp500\_p2$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sp500\_p3$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sp500\_p4$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_sp500\_p5$residuals,2,Box.test,lag=12,type="Ljung")  #HSI & FTSE100  apply(fit\_var\_hsi\_ftse100\_p1$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_ftse100\_p2$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_ftse100\_p3$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_ftse100\_p4$residuals,2,Box.test,lag=12,type="Ljung")  apply(fit\_var\_hsi\_ftse100\_p5$residuals,2,Box.test,lag=12,type="Ljung")  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  # Plot residuals of selected VAR models  #HSI & SSE  plot(fit\_var\_hsi\_sse\_p4$residuals[,1],type="l", col="red", main = "Residuals of VAR(4) for HSI & SSE", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_var\_hsi\_sse\_p4$residuals[,2],type="l", col="red", main = "Residuals of VAR(4) for HSI & SSE", xlab="Day", ylab="Residual for SSE Series")  #HSI & SP500  plot(fit\_var\_hsi\_sp500\_p3$residuals[,1],type="l", col="red", main = "Residuals of VAR(3) for HSI & SP500", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_var\_hsi\_sp500\_p3$residuals[,2],type="l", col="red", main = "Residuals of VAR(3) for HSI & SP500", xlab="Day", ylab="Residual for SP500 Series")  #HSI & FTSE100  plot(fit\_var\_hsi\_ftse100\_p5$residuals[,1],type="l", col="red", main = "Residuals of VAR(5) for HSI & FTSE100", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_var\_hsi\_ftse100\_p5$residuals[,2],type="l", col="red", main = "Residuals of VAR(5) for HSI & FTSE100", xlab="Day", ylab="Residual for FTSE100 Series")  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  #  fit VARMA model  #HSI & SSE  fit\_varma\_hsi\_sse\_11=VARMA(as.data.frame(hsi\_sse\_return),p=1,q=1,include.mean=FALSE)  fit\_varma\_hsi\_sse\_12=VARMA(as.data.frame(hsi\_sse\_return),p=1,q=2,include.mean=FALSE)  fit\_varma\_hsi\_sse\_21=VARMA(as.data.frame(hsi\_sse\_return),p=2,q=1,include.mean=FALSE)  fit\_varma\_hsi\_sse\_22=VARMA(as.data.frame(hsi\_sse\_return),p=2,q=2,include.mean=FALSE)  #HSI & SP500  fit\_varma\_hsi\_sp500\_11=VARMA(as.data.frame(hsi\_sp500\_return),p=1,q=1,include.mean=FALSE)  fit\_varma\_hsi\_sp500\_12=VARMA(as.data.frame(hsi\_sp500\_return),p=1,q=2,include.mean=FALSE)  fit\_varma\_hsi\_sp500\_21=VARMA(as.data.frame(hsi\_sp500\_return),p=2,q=1,include.mean=FALSE)  fit\_varma\_hsi\_sp500\_22=VARMA(as.data.frame(hsi\_sp500\_return),p=2,q=2,include.mean=FALSE)  #HSI & FTSE100  fit\_varma\_hsi\_ftse100\_11=VARMA(as.data.frame(hsi\_ftse100\_return),p=1,q=1,include.mean=FALSE)  fit\_varma\_hsi\_ftse100\_12=VARMA(as.data.frame(hsi\_ftse100\_return),p=1,q=2,include.mean=FALSE)  fit\_varma\_hsi\_ftse100\_21=VARMA(as.data.frame(hsi\_ftse100\_return),p=2,q=1,include.mean=FALSE)  fit\_varma\_hsi\_ftse100\_22=VARMA(as.data.frame(hsi\_ftse100\_return),p=2,q=2,include.mean=FALSE)  #HSI & DowJones  fit\_varma\_hsi\_dowjones\_11=VARMA(as.data.frame(hsi\_dowjones\_return),p=1,q=1,include.mean=FALSE)  fit\_varma\_hsi\_dowjones\_12=VARMA(as.data.frame(hsi\_dowjones\_return),p=1,q=2,include.mean=FALSE)  fit\_varma\_hsi\_dowjones\_21=VARMA(as.data.frame(hsi\_dowjones\_return),p=2,q=1,include.mean=FALSE)  fit\_varma\_hsi\_dowjones\_22=VARMA(as.data.frame(hsi\_dowjones\_return),p=2,q=2,include.mean=FALSE)  #HSI & DAX  fit\_varma\_hsi\_dax\_11=VARMA(as.data.frame(hsi\_dax\_return),p=1,q=1,include.mean=FALSE)  fit\_varma\_hsi\_dax\_12=VARMA(as.data.frame(hsi\_dax\_return),p=1,q=2,include.mean=FALSE)  fit\_varma\_hsi\_dax\_21=VARMA(as.data.frame(hsi\_dax\_return),p=2,q=1,include.mean=FALSE)  fit\_varma\_hsi\_dax\_22=VARMA(as.data.frame(hsi\_dax\_return),p=2,q=2,include.mean=FALSE)  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  # Diagnostic Checking of a VARMA Model  #HSI & SSE  apply(fit\_varma\_hsi\_sse\_11$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_varma\_hsi\_sse\_12$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_varma\_hsi\_sse\_21$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_varma\_hsi\_sse\_22$residuals,2,Box.test,lag=9 ,type="Ljung")  #HSI & SP500  apply(fit\_vma\_hsi\_sp500\_11$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_12$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_21$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_sp500\_22$residuals,2,Box.test,lag=9 ,type="Ljung")  #HSI & FTSE100  apply(fit\_vma\_hsi\_ftse100\_11$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_12$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_21$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_ftse100\_22$residuals,2,Box.test,lag=9 ,type="Ljung")  #HSI & DowJones  apply(fit\_vma\_hsi\_dowjones\_11$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_12$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_21$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dowjones\_22$residuals,2,Box.test,lag=9 ,type="Ljung")  #HSI & DAX  apply(fit\_vma\_hsi\_dax\_11$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dax\_12$residuals,2,Box.test,lag=11,type="Ljung")  apply(fit\_vma\_hsi\_dax\_21$residuals,2,Box.test,lag=10,type="Ljung")  apply(fit\_vma\_hsi\_dax\_22$residuals,2,Box.test,lag=9 ,type="Ljung")  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  # Plot residuals of selected VARMA models  #HSI & SSE  plot(fit\_varma\_hsi\_sse\_22$residuals[,1],type="l", col="red", main = "Residuals of VARMA(2,2) for HSI & SSE", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_varma\_hsi\_sse\_22$residuals[,2],type="l", col="red", main = "Residuals of VARMA(2,2) for HSI & SSE", xlab="Day", ylab="Residual for SSE Series")  #HSI & SP500  plot(fit\_vma\_hsi\_sp500\_11$residuals[,1],type="l", col="red", main = "Residuals of VARMA(1,1) for HSI & SP500", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_vma\_hsi\_sp500\_11$residuals[,2],type="l", col="red", main = "Residuals of VARMA(1,1) for HSI & SP500", xlab="Day", ylab="Residual for SP500 Series")  #HSI & FTSE100  plot(fit\_vma\_hsi\_ftse100\_22$residuals[,1],type="l", col="red", main = "Residuals of VARMA(2,2) for HSI & FTSE100", xlab="Day", ylab="Residual for HSI Series")  plot(fit\_vma\_hsi\_ftse100\_22$residuals[,2],type="l", col="red", main = "Residuals of VARMA(2,2) for HSI & FTSE100", xlab="Day", ylab="Residual for FTSE100 Series")  #\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* |

1. US Fund Fee Study (Morningstar Research, Apr 2019). Available: https://www.morningstar.com/content/dam/marketing/shared/pdfs/Research/USFundFeeStudyApr2019.pdf [↑](#footnote-ref-2)
2. Index funds break through $10tn-in-assets mark amid active exodus (Financial Times, Jan 2020). Available:

   https://www.ft.com/content/a7e20d96-318c-11ea-9703-eea0cae3f0de [↑](#footnote-ref-3)
3. Caporale, Guglielmo Maria & Pittis, Nikitas & Spagnolo, Nicola. (2003). IGARCH models and structural breaks. Applied Economics Letters. 10. 765-768. 10.1080/1350485032000138403. [↑](#footnote-ref-4)
4. Changli He, Timo Teräsvirta and Hans Malmsten (2002), Moment Structure of a Family of First-Order Exponential GARCH Models https://www.jstor.org/stable/3533416?seq=1 [↑](#footnote-ref-5)