2013

Improving volatility forecasting of GARCH models: applications to daily returns in emerging stock markets

Chaiwat Kosapattarapim

University of Wollongong

Recommended Citation


Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Improving Volatility Forecasting of GARCH Models: Applications to Daily Returns in Emerging Stock Markets

A thesis submitted in fulfilment of the requirements for the award of the degree

Doctor of Philosophy

from

University of Wollongong

by

Chaiwat Kosapattarapim

B.Sc., M.Sc. Statistics

School of Mathematics and Applied Statistics

June, 2013
I, Chaiwat Kosapattarapim, declare that this thesis, submitted in fulfillment of the requirement for the award of Doctor of Philosophy, in the School of Mathematics and Applied Statistics, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institutions.

Chaiwat Kosapattarapim

November 13, 2013
Abstract

The volatility modeling and forecasting of returns are essential for many areas of econometric and financial analysis. Volatility forecasting dramatically affects financial decisions, such as portfolio selection, option pricing, risk management and monetary policy making. Improving the modeling and forecasting of financial volatility remains an important issue. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is the most successful model to use for volatility modeling and forecasting of financial returns (Zakaria and Abdalla, 2012).

However, it is well known that financial return series generally exhibit non-normal characteristics while the typical GARCH model assumes a normal error distribution (Gokcan, 2000). Consequently, the typical GARCH model cannot well capture the stylized facts of return series such as heavy tails, excess kurtosis and skewness. This thesis will develop better GARCH models and use these models to improve the volatility forecasting of returns.

In this thesis, there are two main approaches for improving the volatility forecasting performance. The first approach combines GARCH model with various types of non-normal error distributions. There are a large number of non-normal distributions that can be applied to the error term in GARCH model. In this thesis, six different types of error distributions are considered. These are the normal, skewed normal, student-t, skewed student-t, generalized error and skewed generalized error distributions.
GARCH model with a normal error distribution is used as a benchmark to compare the volatility forecasting performance when competing GARCH models are fitted with the other five non-normal distributions. The impact of those error distributions on the best fitting model and the best performance model is studied in this thesis. The simulated results show that the best fitting GARCH(p,q) model is not necessarily the best volatility forecasting performance model. But the results from the paired t-test reveal that there are not greatly significant difference between the best fitting model and the best performance model in terms of Mean Square Error(MSE) and Mean Absolute Error(MAE). Therefore, it is still reliable in practice to use the best fitted model for volatility forecasting. The empirical results indicate that GARCH(p,q) models with non-normal distributions outperform GARCH(p,q) models with a normal distribution based on the three emerging indices from Thailand, Malaysia and Singapore.

The second approach considered in this thesis incorporates the GARCH error terms with the six types of error distributions into the cointegrating error terms in Error Correction Model (ECM). If the underlying financial time series are found to be cointegrated and each series can be well fitted by a univariate GARCH model with non-normal distributions, our study shows the knowledge of cointegration information among these series might result in further improvement in volatility forecasting based on univariate GARCH model.

There are several methods for detecting the cointegration relationships among financial time series. This thesis investigates which cointegration method is the most powerful to use for developing volatility forecasting models. The Johansen approach appears to provide superior results when the cointegrating errors are normally distributed. This thesis investigates whether the Johansen tests continue to be more powerful than another three tests when the cointegrating errors are non-normally distributed. The performance of the Johansen method is com-
pared with another three tests, the Dickey-Fuller test, the Cointegrating Regression Durbin-Watson test and the Wild Bootstrap test in terms of the size and power of the tests.

The simulation results reveal that the power of the Johansen tests is higher than that of other cointegration tests. Furthermore, the power of the Johansen tests slightly increases when the errors of the GARCH(1,1) model is given by the skewed student-t error distribution.

To investigate whether the knowledge of cointegration information can be beneficial to volatility forecasting performance, simulation studies are conducted to compare the performance in terms of the volatility forecasting between an individual univariate GARCH(p,q) model and cointegration-based ECM by taking into account alternative non-normal distribution assumptions. The results indicate that the model which contains the knowledge of cointegration information can further improve the volatility forecasting performance and provide better forecasts than the best fitting univariate GARCH model. A model with the non-normal error distributions tends to outperform a model with the normal error distribution. Therefore, the knowledge of cointegration relationship information among the underlying financial time series appears to provide certain benefits in volatility forecasting. Furthermore, using the non-normal error distributions such as skewed student-t and generalized error distributions in a GARCH model can improve accuracy of volatility forecasting.

This thesis also examines the comparisons of VaR estimations between the univariate GARCH model and the cointegration-based ECM by using the cointegrated indices of daily closing prices from Thailand and Malaysia. Two types of Backtesting used in this thesis for VaR evaluations are the unconditional coverage ($LR_{uc})$ and conditional coverage ($LR_{cc)$). VaR estimates calculated based on the knowledge of cointegration information (Model B) can produce adequate
VaR forecasts for 1-step ahead for both SET and KLCI. The results of VaR forecasting reveal that, if time series are cointegrated, the knowledge of cointegration information will help to improve the volatility forecasting and VaR forecasting for 1-step ahead.
Acknowledgements

This thesis would not have been possible without the help and support of many people. Firstly, I would like to thank my supervisors, Associate Professor Yan-Xia Lin, Dr. Chandra Gulati and Associate Professor Michael McCrae during my research period. I am extremely grateful for their wisdom, knowledge, guidance, great patience and kindness.

To Yan-Xia Lin, thank you very much for your precious time to teach me and support me, even when I had so many troubles. I appreciate all your teaching which I will adapt for my career. To Chandra, thank you so much for your support not only with regards to my thesis but also for the great encouragement you gave me. To Michael, many thanks for your contributions ideas for my thesis. Thank you for all of you for trying to understand and give me encouragement to finish my thesis.

I would like to thank to the Thai Government Science and Technology Scholarship including the Maejo University that I have received the financial supports for my PhD study.

Special thanks for my best friend Dane Lyddiard for devoting your time to correct my English grammar.

I would also like to thank all staff in the School of Mathematics and Applied Statistics for their support over years of my study.

Finally, I would like to thank to my family for your support and always encourage me in everything I have encountered during my stay in Australia.
## Contents

1 Introduction .................................................. 1
   1.1 The Modification of GARCH Models ....................... 1
   1.2 Cointegration and GARCH Models ....................... 4
   1.3 Forecasting of Value at Risk and GARCH Models .......... 5
   1.4 The Problems ........................................... 5

2 Preliminaries and Literature Review ................. 11
   2.1 Introduction ........................................... 11
   2.2 Returns and Volatility in Emerging Market ............. 12
   2.3 Some Stylized Facts on Returns ....................... 13
   2.4 Volatility Forecasting Models ....................... 17
      2.4.1 Development of Volatility Forecasting Models .......... 17
      2.4.2 Various GARCH Models with Alternative Distributional
            Assumptions ........................................... 21
   2.5 Cointegration of Financial Time Series ................ 24
   2.6 Value at Risk and Volatility Forecasting ............. 25

3 GARCH(p,q) Model with Alternative Error Distributions 29
   3.1 GARCH(p,q) Model with Normal Error Distribution ....... 29
   3.2 Non-normal Error Distributions Applied to GARCH Model .... 31
   3.3 Model Selection Using Information Criteria ............ 36
      3.3.1 Determining the Order of GARCH Models ............ 37
      3.3.2 Simulation Study on Order Determination ............ 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Volatility Forecasts via GARCH(p,q) Models</td>
<td>42</td>
</tr>
<tr>
<td>3.5</td>
<td>Evaluation of Volatility Forecasts</td>
<td>46</td>
</tr>
<tr>
<td>3.6</td>
<td>Simulation Study on the Performances of Volatility Forecasting: Comparing between the</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Best Fitting Model and the Best Performance Model</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td>Conclusion</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>Evaluating the Volatility Forecasting Performance of the Best Fitting GARCH Models in</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Emerging Asian Stock Markets</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>55</td>
</tr>
<tr>
<td>4.2</td>
<td>Data and Methodology</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>4.2.1 Descriptive Statistics of Three Emerging Indices</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>4.2.2 In-Sample Parameter Estimation and Model Diagnostics</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>Empirical Results for the Performance of Volatility Forecasting</td>
<td>65</td>
</tr>
<tr>
<td>4.4</td>
<td>Conclusion</td>
<td>72</td>
</tr>
<tr>
<td>5</td>
<td>Cointegration Tests with Non-Normal GARCH Error Distributions</td>
<td>75</td>
</tr>
<tr>
<td>5.1</td>
<td>Unit Root and Cointegration</td>
<td>76</td>
</tr>
<tr>
<td>5.2</td>
<td>The Residual-Based Tests for Cointegration</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>5.2.1 Dickey-Fuller Tests for the Test of Unit Root</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>5.2.2 Cointegrating Regression Durbin-Watson Test for the Test of Unit Root</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>5.2.3 The Wild Bootstrap Cointegration Test</td>
<td>84</td>
</tr>
<tr>
<td>5.3</td>
<td>The Johansen Cointegration Tests</td>
<td>85</td>
</tr>
<tr>
<td>5.4</td>
<td>An Empirical Example of Cointegrating Errors Fitted by a GARCH Model with Non-Normal</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Error Distributions</td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>Comparison of the Size and Power of Cointegration Tests with Various Distributional</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Assumptions of a GARCH(1,1) Model</td>
<td></td>
</tr>
</tbody>
</table>
5.6 The Results for the Size and Power of Cointegration Tests Based on Simulated Data ................................. 99
5.7 Conclusion .......................................................... 102

6 Improving Volatility Forecasting Based on Cointegration Information 111

6.1 Alternative Approaches Employed to Improve the Performance of Volatility Forecasting .................................. 112
6.2 Simulation Study for Evaluation of Volatility Forecasting: A Comparison between a GARCH Model and Cointegration based on ECM116
6.3 Simulation Results .................................................. 120
6.4 Empirical Study: Volatility Forecasting Performance using Emerging Stock Indices ............................... 138
6.5 Conclusion .......................................................... 143

7 Estimation of Value-at-Risk for Emerging Stock Markets 145

7.1 Definition of Value at Risk ........................................ 146
7.2 Estimation of VaR .................................................. 148
7.3 Evaluation of VaR Estimates by Backtesting Methods ........ 152
  7.3.1 The Statistical Framework of VaR Backtesting .......... 153
  7.3.2 The Kupiec Test ............................................. 155
  7.3.3 Christoffersen’s Interval Forecast Test ..................... 156
7.4 Empirical Results and Evaluation of VaR Estimates .......... 159
7.5 Conclusion .......................................................... 163

8 Conclusion and Further Research 165

8.1 Conclusion .......................................................... 165
8.2 Further Research .................................................. 169

A Program Files 171

A.1 Programs used in Chapter 3 ..................................... 171
A.1.1 Simulations for the True GARCH(1,3) and GARCH(2,1) Models .................................................. 172
A.2 Programs used in Chapter 4 ................................. 179
   A.2.1 Modeling and Forecasting for SET .................... 179
   A.2.2 Modeling and Forecasting for KLCI .................. 181
   A.2.3 Modeling and Forecasting for STI .................. 182
A.3 Programs used in Chapter 5 ................................. 184
   A.3.1 Unit Root Tests and Cointegration Analysis among SET,
        KLCI and STI .............................................. 184
   A.3.2 Simulation Studies on the Size of the Test for GARCH(1,1)
        Model ......................................................... 186
A.4 Programs used in Chapter 6 ................................. 213
   A.4.1 Unit Root Tests and Cointegration Analysis among SET,
        KLCI and STI .............................................. 213
A.5 Programs used in Chapter 7 ................................. 220
   A.5.1 Value at Risk Estimations ............................... 220

Bibliography ................................................. 229
List of Figures

2.1 Normal distribution and distribution with heavy tail ........... 14

2.2 Returns of SET from 24/11/1983 to 1/02/1996, 3,000 in total ob-
servations (excluding public holidays and weekends) ............. 16

4.1 Historical daily closing price movement and daily returns for the
SET; 6,536 in total observations (excluding public holidays and
weekends) ................................................................. 57

4.2 Historical daily closing price movement and daily returns for the
KLCI; 3,880 in total observations (excluding public holidays and
weekends) ................................................................. 58

4.3 Historical daily closing price movement and daily returns for the
STI; 5,407 in total observations (excluding public holidays and
weekends) ................................................................. 59

4.4 Volatility forecasting of the GARCH(1,3) model for the SET eval-
uated by MSE .......................................................... 68

4.5 Volatility forecasting of the GARCH(1,1) model for the KLCI eval-
uated by MSE .......................................................... 69

4.6 Volatility forecasting of the GARCH(2,1) model for the STI eval-
uated by MSE .......................................................... 69

4.7 Volatility forecasting of the GARCH(1,3) model for the SET eval-
uated by MAE .......................................................... 70
4.8 Volatility forecasting of the GARCH(1,1) model for the KLCI evaluated by MAE

4.9 Volatility forecasting of the GARCH(2,1) model for the STI evaluated by MAE

5.1 Plot log of closing price of SET and KLCI spanning from 1 July 1998 to 31 December 2002, 1,133 in total observations (excluding public holidays and weekends)

5.2 Plots of the first differences of LSET and LKLCI spanning from 1 July 1998 to 31 December 2002, 1,133 in total observations (excluding public holidays and weekends)
List of Tables

3.1 AIC values when true GARCH(p,q) models are from STD, SSTD and GED error distributions ........................................ 41
3.2 The values of AIC for simulated data from the GARCH(1,3) model ................................................................. 49
3.3 The values of AIC for simulated data from the GARCH(2,1) model ................................................................. 49
3.4 The values of MSE and MAE for simulated data from the GARCH(1,3) model ......................................................... 50
3.5 The values of MSE and MAE for simulated data from the GARCH(2,1) model ......................................................... 51
3.6 The P-values of paired t-test results between the best fitting model and the best performance model given by samples from the GARCH(1,3) model ................................................................. 52

4.1 Descriptive statistics and Jarque-Bera test statistic for normality of daily returns for the SET, KLCI and STI ................. 60
4.2 The AIC values given by models with different error distributions ................................................................. 60
4.3 Estimated parameters and diagnostics of the GARCH(1,3) model with different error distributions for the SET ...................... 62
4.4 Estimated parameters and diagnostics of the GARCH(1,1) model with different error distributions for the KLCI ............... 63
4.5 Estimated parameters and diagnostics of the GARCH(2,1) model with different error distributions for the STI ...................... 64
4.6 Out-of-sample volatility forecasting evaluated by MSE .................. 67
4.7 Out-of-sample volatility forecasting evaluated by MAE ........ 68
4.8 The percent error of MSE and MAE given by the best fitted model and the best performance model ......................... 71

5.1 Summary statistics .................................................. 89
5.2 Unit root test results for two stock indices ........................ 90
5.3 Autocorrelation tests of the residuals on LSET and LKLCI .... 91
5.4 Results and Critical Values for the $\lambda_{\text{trace}}$ and $\lambda_{\text{max}}$ test .................. 92
5.5 Comparative AIC values given by GARCH(1,1) with different type of error distributions ..................... 93
5.6 Estimated parameters and diagnostic of GARCH(1,1) model with SSTD .................................................. 94
5.7 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 100$ at 5% level ......................... 103
5.8 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.65, \beta = 0.05$ and $T = 100$ at 5% level ......................... 104
5.9 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 1,000$ at 5% level ......................... 105
5.10 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.65, \beta = 0.05$ and $T = 1,000$ at 5% level ......................... 106
5.11 The power of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 100$ at 5% level ......................... 107
5.12 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.65, \beta = 0.05$ and $T = 100$ at 5% level ......................... 108
5.13 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 1,000$ at 5% level ......................... 109
5.14 The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.65, \beta = 0.05$ and $T = 1,000$ at 5% level ......................... 110
6.1 AIC values for specifying the order of the GARCH(p,q) models used to fit $x_{1t}$, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.2, \beta = 0.3$ .................................................. 121

6.2 AIC values for specifying the order of the GARCH(p,q) models used to fit $x_{1t}$, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.5, \beta = 0.2$ .................................................. 122

6.3 AIC values for specifying the order of the GARCH(p,q) models used to fit $x_{1t}$, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ .................................................. 123

6.4 AIC values for specifying the order of the GARCH(p,q) models used to fit $x_{1t}$, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.1, \beta = 0.8$ .................................................. 124

6.5 AIC values for identifying appropriate error distributions in Model A, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.2, \beta = 0.3$ .................................................. 125

6.6 AIC values for identifying appropriate error distributions in Model A, where $x_{1t}$ was simulated from Equation (6.6) with $(\omega = 0.1, \alpha = 0.5, \beta = 0.2)$ .................................................. 126

6.7 AIC values for identifying appropriate error distributions in Model A, where $x_{1t}$ was simulated from Equation (6.6) with $(\omega = 0.1, \alpha = 0.3, \beta = 0.6)$ .................................................. 127

6.8 AIC values for identifying appropriate error distributions in Model A, where $x_{1t}$ was simulated from Equation (6.6) with $(\omega = 0.1, \alpha = 0.1, \beta = 0.8)$ .................................................. 128

6.9 The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 1-step ahead ................. 130
6.10 The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 2-step ahead .......... 131
6.11 The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 10-step ahead .......... 132
6.14 The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 2-step ahead .......... 135
6.15 The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 10-step ahead .......... 136
6.16 The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 15-step ahead .......... 137
6.17 The AIC values given by models with different error distributions 138
6.18 Estimated parameters by individual model fitting and the diagnostics from the GARCH(1,1)-SSTD model for SET for data spanning from 1/07/1998 to 31/12/2002 ............................................ 139
6.19 Estimated parameters by individual model fitting and diagnostics from the GARCH(1,1)-SSTD model for KLCI for data spanning from 1/07/1998 to 31/12/2002 ............................................ 139
6.20 The AIC values given by VECM(8) with different error distributions for SET and KLCI ............................................ 141
6.21 Estimated parameters and diagnostics of GARCH(1,1)-STD model in VECM(8) for SET, the data spanning from 1/07/1998 to 31/12/2002142
6.22 Estimated parameters and diagnostics of GARCH(1,1)-GED model in VECM(8) for KLCI, the data spanning from 1/07/1998 to 31/12/2002142
6.23 Comparisons of volatility forecasting performance of Model A and
B for emerging stock indices SET and KLCI covering the period
1/07/1998 to 31/12/2002 . . . . . . . . . . . . . . . . . . . . . . . . . . 143

7.1 Contingency table of possible outcome for independence test . . . 157
7.2 Out-of-sample Value-at-Risk for 1- and 2- steps ahead given by the
Kupiec test at 95% confidence level . . . . . . . . . . . . . . . . . . . . 160
7.3 Input data for calculating the Backtesting independence test . . . 161
7.4 Conditional coverage Backtesting results for 1- and 2- steps ahead
at 95% confidence level . . . . . . . . . . . . . . . . . . . . . . . . . . . 162
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF</td>
<td>Augmented Dickey-Fuller</td>
</tr>
<tr>
<td>AGARCH</td>
<td>Asymmetric GARCH</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike’s Information Criteria</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Networks</td>
</tr>
<tr>
<td>APARCH</td>
<td>Asymmetric Power ARCH</td>
</tr>
<tr>
<td>AR</td>
<td>Auto Regressive</td>
</tr>
<tr>
<td>ARCH</td>
<td>Auto Regressive Conditional Heteroskedasticity</td>
</tr>
<tr>
<td>ARMA</td>
<td>Auto Regressive Moving Average</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian Information Criteria</td>
</tr>
<tr>
<td>CRDW</td>
<td>Cointegrating Regression Durbin-Watson</td>
</tr>
<tr>
<td>DF</td>
<td>Dickey-Fuller</td>
</tr>
<tr>
<td>DIC</td>
<td>Draper’s Information Criteria</td>
</tr>
<tr>
<td>ECM</td>
<td>Error Correction Model</td>
</tr>
<tr>
<td>EGARCH</td>
<td>Exponential GARCH</td>
</tr>
<tr>
<td>EMD-NN</td>
<td>Empirical Mode Decomposition and Neural Network</td>
</tr>
<tr>
<td>EWMA</td>
<td>Exponentially Weighted Moving Average</td>
</tr>
<tr>
<td>FIGARCH</td>
<td>Fractionally Integrated GARCH</td>
</tr>
<tr>
<td>GARCH</td>
<td>Generalized Autoregressive Conditional Heteroskedasticity</td>
</tr>
<tr>
<td>GED</td>
<td>Generalized Error Distribution</td>
</tr>
<tr>
<td>HQC</td>
<td>Quinn Information Criteria</td>
</tr>
<tr>
<td>IGARCH</td>
<td>Integrated GARCH</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>KLKI</td>
<td>Kuala Lumpur Composite Index</td>
</tr>
<tr>
<td>LKLCI</td>
<td>log of daily closing price of KLCI</td>
</tr>
<tr>
<td>LM</td>
<td>Lagrange Multiplier</td>
</tr>
<tr>
<td>LSET</td>
<td>log of daily closing price of SET</td>
</tr>
<tr>
<td>MA</td>
<td>Moving Average</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean Absolute Error</td>
</tr>
<tr>
<td>MAPE</td>
<td>Mean Absolute Percentage Error</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MLR</td>
<td>Multiple Linear Regressions</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>N</td>
<td>Normal Distribution</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>PE</td>
<td>Percent Error</td>
</tr>
<tr>
<td>QGARCH</td>
<td>Quadratic GARCH</td>
</tr>
<tr>
<td>REIT</td>
<td>Real Estate Investment Trust</td>
</tr>
<tr>
<td>SET</td>
<td>Stock Exchange of Thailand</td>
</tr>
<tr>
<td>SGED</td>
<td>Skewed Generalized Error Distribution</td>
</tr>
<tr>
<td>SN</td>
<td>Skewed Normal Distribution</td>
</tr>
<tr>
<td>SSTD</td>
<td>Skewed Student-t Distribution</td>
</tr>
<tr>
<td>STD</td>
<td>Student-t Distribution</td>
</tr>
<tr>
<td>STI</td>
<td>Straits Time Index</td>
</tr>
<tr>
<td>TGARCH</td>
<td>Threshold GARCH</td>
</tr>
<tr>
<td>VaR</td>
<td>Value at Risk</td>
</tr>
<tr>
<td>VAR</td>
<td>Vector Auto Regressive</td>
</tr>
<tr>
<td>VECM</td>
<td>Vector Error-Correction Model</td>
</tr>
<tr>
<td>WB</td>
<td>Wild Bootstrap</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

This chapter will introduce important issues to improve the volatility forecasting performance of financial time series using Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models. This chapter also describes the problems that will be tackled in this thesis.

1.1 The Modification of GARCH Models

Statistical volatility plays a crucial role in modeling and forecasting of financial time series. Volatility forecasting is used as a measurement for financial decisions, such as portfolio selection, option pricing, risk management and monetary policy making. Modeling and forecasting of financial volatility remains an important issue and so it would be beneficial to identify a model which is able to improve the accuracy of volatility forecasting. A large number of financial volatility models have been developed since the Auto Regressive Conditional Heteroskedasticity (ARCH) model was proposed by Engle (1982). However, an ARCH model has some weaknesses. It becomes difficult to estimate parameters in ARCH models when higher orders are considered. Consequently, Bollerslev (1986) extended the ARCH model to the GARCH model, which is more parsimonious than the ARCH model.

The GARCH model is the most popular model for successfully capturing
volatility in financial time series. Gokcan (2000) stated that GARCH models can effectively remove the excess kurtosis in financial return series. However, the error terms in traditional symmetric GARCH models are based on the assumption of normal distribution while the property of financial returns clearly exhibit non-normal distribution with high kurtosis, a heavy tail and sometime skewness. The weakness of the standard GARCH model with normal error distribution is that it fails to capture the stylized properties of underlying financial time series.

The error terms of traditional GARCH models are typically assumed to be normally distributed. It is inadequate to use such traditional GARCH models to forecast volatility when the distribution of returns are characterized by stylized facts such as heavy tails, excess kurtosis, skewness, volatility clustering and the leverage effect. Therefore, various modifications of GARCH models will be discussed for improving volatility forecasting performance by taking into account these kinds of financial stylized characteristics of return series. GARCH models have been developed by allowing alternative non-normal distributions in their error terms. Numerous studies on the development of GARCH models aimed at improving the performance of volatility forecasting have been conducted by a large number of researchers. Three main modifications on GARCH models are considered.

The first modification on GARCH model is to allow the error terms in symmetric GARCH models to have non-normal distribution. The main advantage of adopting non-normal distribution is that it is able to model thicker tail, higher kurtosis and skewness. When a non-normal distribution is incorporated into the error terms in a symmetric GARCH model, the GARCH model is more flexible and is able to capture the stylized properties of financial return series. Some examples of the non-normal distributions that can be applied to the error terms of GARCH models include the skewed normal, normal inverse Gaussian, Student-t,
skewed Student-t, generalized error and skewed generalized error distributions.

The second modification on the GARCH models is to develop asymmetric GARCH models with various non-normal error distributions. The advantage of asymmetric GARCH model is that it allows flexible specification for the financial volatility modeling and forecasting. The most popular asymmetric GARCH model is called the Exponential GARCH (EGARCH) model (Nelson, 1991). The EGARCH model can cope with the stylized properties of returns, called leverage effects. It allows both positive and negative shocks to have a different impact on volatility forecasting. The other classes of asymmetric GARCH models include the Asymmetric GARCH (AGARCH) (Engle and Ng, 1993), the Threshold GARCH (TGARCH) (Zakoian, 1994), the GJR-GARCH model (Glosten et al., 1993), and the Quadratic GARCH (QGARCH) (Sentana, 1995).

The last class of modified GARCH models is one which incorporates the different efficient approaches into the symmetric and asymmetric GARCH models. For instance, Taylor (2004) adopted new smooth transition exponential smoothing method with different types of GARCH(1,1) models for volatility forecasting and considered 1-step ahead volatility prediction.

This thesis raises a problem of how to develop a GARCH model to improve the performance of volatility forecasting. Symmetric GARCH models with higher order degree are examined by taking into account the non-normal error distributions. GARCH models with alternative non-normal error distributions will also be applied to the cointegrating error terms in Error Correction Model (ECM) to examine whether the knowledge of cointegration information can be beneficial to improve the performance of volatility forecasting if time series are cointegrated. In addition, Value at Risk (VaR) estimates are calculated using the best fitting univariate GARCH model and ECM based on GARCH model. The VaR estimates determined by univariate GARCH and ECM model, respectively will be
evaluated by the Backtesting to examine which VaR estimate is more accurate.

1.2 Cointegration and GARCH Models

The cointegration method introduced by Engle and Granger (1987) is now widely employed in analysis of econometrics and financial time series. The method is used to investigate the linear relationship between non-stationary time series. A variety of methods have been developed for testing cointegration among financial time series. The most popular cointegration test is the Johansen approach (Johansen, 1988, 1991). The unit root test is used to test the stationarity of financial time series. This technique can be employed for cointegration test. The unit root tests applied in this thesis for cointegration test are the Dickey-Fuller test (Fuller, 1976) and (Dickey and Fuller, 1979), the Cointegrating Regression Durbin-Watson (CRDW) test (Sargan and Bhargawa, 1983) and the Wild Bootstrap method (Gerolimetto and Procidano, 2003).

The Johansen tests and other tests are based on the assumption that cointegrating error is normally distributed. This thesis incorporates GARCH models with different non-normal distributions to compare the size and power of these tests. The GARCH error terms consist of normal, skewed normal, student-t, skewed student-t, generalized error distribution and skewed generalized error distributions. Comparison of the size and power of these tests under GARCH model with these six different types of error distributions are carried out and examine which test is the most powerful in detecting cointegration relationships among the underlying time series. Our study shows that the Johansen approach is the most powerful for testing cointegration when the cointegrating errors follow a GARCH model with non-normal error distributions.

The symmetric GARCH models with different distributional assumptions can improve the performance of volatility forecasting when compared with the stan-
standard GARCH model with a normal distribution. It is of interest to examine whether the cointegration information can benefit the performance of volatility forecasting of underlying financial time series. This thesis compares the volatility prediction performance of the cointegration based on ECM and GARCH model with the individual univariate GARCH model by taking into account the six types of error distributions mentioned above.

1.3 Forecasting of Value at Risk and GARCH Models

Value at Risk (VaR) is a downside risk measurement used widely in financial risk management (Füss et al., 2007). The traditional assumption of standard VaR estimation is based on a normal distribution and might be inadequate for financial returns. In practice, this risk measurement is related to the volatility forecasting of underlying financial data. Making accurate forecasts of financial volatility are very important in controlling the downside risk in investment. The more accurate a volatility forecast is, the more it can improve the quality of the risk measures and lead to a successful implementation of risk management. By using better models for volatility forecasting, the forecasting of VaR can be more accurate. There are several other studies which are related to the improvement of VaR estimations associated with GARCH models. This thesis will study and compare the VaR estimates between the best fitting GARCH models and the ECM with GARCH errors by taking into account non-normal error distributions. The Backtesting approaches are used to evaluate the adequacy of the VaR estimations.

1.4 The Problems

This thesis will raise some issues related to the development of volatility forecasting model using a symmetric GARCH\((p,q)\) model with the alternative non-normal
error distributions. The cointegration method is also considered in developing the performance of volatility prediction in symmetric GARCH(p,q) models.

In Chapter 2, the introduction of financial return series and some stylized properties of returns will be presented. Likewise, the development of volatility forecasting models including the variety of GARCH models with alternative distributional assumptions, cointegration of financial time series and the estimation of Value at Risk with better volatility forecasting model will be introduced.

In Chapter 3, a theoretical background of alternative non-normal error density functions applied to the GARCH model will be introduced. To model and forecast the volatility of underlying financial time series, symmetric GARCH(p,q) models with higher order are considered by taking into account five different types of non-normal error distributions in addition to the normal distribution.

Most studies on volatility forecasting have used various GARCH(1,1) models with different error distributions to predict volatility. However, it has not mentioned how to determine the order of a GARCH(p,q) model under the non-normality assumption in literature. In this chapter, simulation studies on how to determine the order of GARCH(p,q) models when the GARCH error terms are non-normally distributed will be conducted and discussed. The simulation studies on the order determination of GARCH model will be carried out.

In Chapter 4, the performance of volatility forecasting using the GARCH(p,q) models is examined. The characteristic properties of return series in the emerging markets are different from the returns of capital markets in developing countries. Bekaert and Harvey (1995) mentioned that the volatility of emerging markets appeared much higher than that of developed markets. The returns appeared to have a low correlation and greater forecast predictability. Gokcan (2000) confirmed the existence of higher return volatilities among emerging markets. Most studies on volatility forecasting have focused on the capital markets in devel-
CHAPTER 1. INTRODUCTION

Developing countries such as European and US markets. The real observations employed in this chapter are the daily closing price indices of three emerging stock markets in South East Asia: the Stock Exchange of Thailand Index (SET), the Kuala Lumpur Composite Index (KLCI) from Malaysia and the Straits Times Index (STI) from Singapore. To investigate a better GARCH(p,q) model for volatility forecasting of emerging returns, this chapter will not only focus on the GARCH(1,1) model but also investigate whether it is more appropriate to use higher orders of GARCH model to fit some returns of emerging stock markets. The GARCH error terms will be allowed to have the different types of competing error distributions. Therefore, six competing GARCH error distributions, including the normal, skewed normal, student-t, skewed student-t, generalized error distribution and skewed generalized error distributions are used to compare the performance of volatility forecasting.

Chapters 3 - 4 focuses on the development of a symmetric GARCH(p,q) models by allowing the six types of error distributions in the GARCH error terms. This thesis will adopt the cointegration method by incorporating a GARCH(p,q) model into the cointegrating error terms in ECM. The error terms in the GARCH(p,q) model used in the cointegrating errors continue to employ the six different types of error distributions mentioned above. There are several other cointegration tests for investigating the relationship among financial time series and the tests for cointegration are related to the tests of unit root. Thus, it is important to examine which cointegration test is the most powerful to detect the cointegration relationship among time series.

Chapter 5 focuses on the investigation of the four cointegration tests which consist of the Dickey-Fuller test (Fuller, 1976) and (Dickey and Fuller, 1979), the Cointegrating Regression Durbin-Watson (CRDW) test (Sargan and Bhargawa, 1983), the Wild Bootstrap method for the unit root test followed by the work
CHAPTER 1. INTRODUCTION

of Gerolimetto and Procidano (2003) and the Johansen tests by Johansen (1988, 1991). The first three cointegration tests are called the residual-based tests which are related to the tests of unit root, see (Maddala and Kim, 1998). The last cointegration test is based on the ECM.

The simulation studies will be carried out to examine which cointegration test is the most powerful to detect the cointegration relationship among time series. To evaluate the performance of the cointegration tests, the size and power of these cointegration tests are considered. In this study, GARCH models with the six different types of error distributions continue to be used in the cointegrating error terms. Using the most powerful cointegration test in Chapter 5, the Johansen approach will be used to compare the volatility forecasting performance of ECM with GARCH error and the univariate best fitting GARCH(p,q) model in Chapter 6.

In Chapter 6, comparison of the performance of the volatility forecasting between individual univariate GARCH(p,q) model and cointegration-based ECM by taking into account alternative non-normally distributed assumptions is examined. The two main objectives in this chapter are as follows:

1. Does the knowledge of cointegration relationships among underlying financial time series make any contributions to volatility forecasting?

2. According to the empirical results from Chapters 4 - 5, return series exhibit non-normal innovations. A large amount of research has shown that GARCH model with non-normal distribution can further improve the volatility forecasting performance. In this study, it is of interest to improve the volatility forecasting performance when the knowledge of cointegration is considered in the presence of symmetric GARCH(1,1) model by taking into account the six different types of error distribution used in Chapter 4.

Chapter 7 investigates the comparisons of VaR estimations between the uni-
variate best fitting GARCH(p,q) model and the model that contains cointegration relationship by using the daily closing prices for SET and KLCI indices. The Backtesting methods are employed to evaluate which VaR estimations are more accurate and reliable to use for the underlying financial time series.

Chapter 8 provides a summary, conclusions and suggestions for further research on the GARCH models for improving the volatility forecasting performance of underlying financial time series. The study carried out in this thesis will add to the literature for the development of GARCH models and also can be beneficial for practitioners and financial investors.

Some results from this thesis have been published in refereed journals:


In this thesis, [1] is part of Chapters 3 and 4; [2] is parts of Chapters 5.

The following manuscripts are under preparation for journals:

[3] Improving Volatility Forecasting Based on Cointegration Information and GARCH Model with Non-Normal Distributions will be submitted to a journal.

[4] Estimation of Value-at-Risk for Emerging Stock Markets Based on Cointegration Information and GARCH Model with Non-Normal Distributions will be submitted to a journal.

Chapter 2

Preliminaries and Literature Review

This thesis focuses on the issues of volatility modeling, volatility forecasting and their applications to emerging stock markets in South East Asia. This chapter presents a literature review on the research which is relevant to this thesis.

2.1 Introduction

A return series is defined as the difference of the logarithms of a financial time series such as stock price indices, exchange rates and interest rate changes. In this thesis, $p_t$ is defined as the daily closing price of a stock index at time $t$ for $t = 1, 2, 3, \ldots T$ where $T$ is the total number of observations. The daily return of index is denoted as the following:

$$ r_t = \ln\left\{ \frac{p_t}{p_{t-1}} \right\} $$

and variance of return is referred as the volatility of $r_t$.

Different from the way defined in literature, the terminology “volatility” in this thesis is used to define the variance of return on the underlying financial daily closing price of the stock index. In addition, volatility is sometimes referred to as the conditional variance of return. The modeling and forecasting for the volatility of returns are essential for many areas of finance since volatility is widely...
used as the most important indicator in financial investment. The forecasted volatility of financial returns is routinely used as a measure of risk. Furthermore, these forecasts are used in risk management in areas such as Value at Risk (VaR), derivative pricing and hedging, portfolio selection and many activities of financial applications. A large volume of research have been devoted to establish better models for improving the forecast of the volatility of underlying returns. Currently, volatility modeling and forecasting of financial returns remain the attractive issues motivating researchers to develop models and improve volatility forecasting.

The GARCH model has become the most successful model for volatility forecasting since it was first introduced by Bollerslev (1986). Various modifications of GARCH models have been developed for improving the volatility forecasting in stock markets.

Most research discussing volatility analysis focuses on the developed capital markets. Recently, emerging markets have become more attractive since the potential growth rate of these markets has increased dramatically. The financial markets in emerging economies also have become larger and more sophisticated. Consequently, the studies in emerging markets volatility have become more important for researchers. However, previous research on volatility modeling and forecasting in emerging markets has remained inconclusive and more needs to conduct in these markets.

2.2 Returns and Volatility in Emerging Market

Emerging markets are the economic capital markets in developing countries that have tremendous growth rate expectations to share on the stage of a country’s economic growth (Mody, 2004). The potential role of these markets have increased in the international economy and the markets have become larger players in the
global economic world. The growth of emerging capital markets has attracted investors in the past few years. Previous studies have reported that emerging markets have experienced a high level of return, high level of volatility and have provided diversification benefits for investors who invest in developed markets.

Most previous research into the development of share market return volatility has occurred within well developed, mature markets such as European and US markets. However, the increasing maturity, size and sophistication of many emerging markets is now attracting greater attention. While the short history of many of these markets has curtailed research, some unique factors have emerged. Bekaert and Harvey (1995) maintained that while the levels and volatility (risk) of returns appeared much higher than in developed markets, they appeared to have a low correlation and greater forecast predictability. Gokcan (2000) confirmed the existence of higher return volatilities among emerging markets.

Furthermore, Bekaert et al. (1998) confirmed the characteristics of emerging market returns displayed a non-normal distribution with positive skewness. This thesis attempts to focus on the volatility modeling and forecasting using data from emerging stock markets in South East Asia.

The following section contains reviews of some characteristics of returns which can be generally found in both developed and emerging markets.

2.3 Some Stylized Facts on Returns

Stylized facts of returns are the general properties in financial returns that are accepted as truth. Empirical findings show that these properties are very consistent across a wide range of methods, markets and time periods. Some stylized facts of returns are described in this section.

1. Absence of autocorrelations
Asset returns typically do not exhibit autocorrelation. The linear autocorrelation of returns are often insignificant, except for returns with a small time scale ($\approx 20$ minutes) (Cont, 2001). The autocorrelations for the absolute returns and squared returns are always positive, significant and decay slowly.

2. Non-normal Distribution

Mandelbrot (1963) pointed out that the normal distribution is inadequate for modeling returns. A normal distribution has excess kurtosis and skewness of zero but the probability distributions of many returns have their kurtosis greater than three and taller narrower peaks than a normal distribution. A random variable that has this property is said to be leptokurtic (sharp-peaked and heavy tailed). In addition, the distributions of returns sometimes exhibit skewness.

3. Heavy tails

A probability distribution is said to have a heavy tail if it exhibits extremely
large kurtosis or skewness. Due to the non-normally distributed character, the probability density functions of returns tend to be leptokurtic. Figure 2.1 shows a heavy tail distribution compares with a normal distribution (adapted from http://www.amex.com/dictionary/charts/chart77.gif). Several heavy tail distributions are commonly used in financial applications such as the student-t distribution, generalized error distribution (GED), log gamma distribution and mixtures of normal distributions.

4. Aggregational Gaussianity

Aggregational Gaussianity is used to described the fact that when the time scale ($\Delta t$) for return calculation is increased, the distribution of returns tends to be more like a normal distribution. But the shapes of the distributions are not exactly the same at different time scales (Cont, 2001).

5. Volatility clustering

Volatility clustering in returns is one of the well-known stylized facts in financial markets. Mandelbrot (1963) noted that volatility clustering explains the movement of returns where large positive changes in returns tend to be followed by large negative changes and small positive changes tend to be followed by small negative changes. The amplitude of returns is sometimes large and sometimes small. Figure 2.2 shows a time series plot of daily closing price index for the Stock Exchange of Thailand (SET). From this plot, it is apparent that the amplitude of the returns is changing over time and the volatility clustering phenomena can be clearly observed.
6. Gain/loss asymmetry

It is more probable to find that the scale of move down in stock prices is not equal to the scale of move up (Cont, 2001). Gain/loss asymmetry refers to the probability of stock prices or indices values moving up and down unequally.

7. Leverage effect

Another important stylized fact in asset returns is leverage effect. The common explanations for the leverage effect of returns refer to the negative correlation between the past returns and future volatility. When bad news occurs in the market, it might lead to the decrease of the stock price. This tends to cause increased future volatility and makes decrease a higher risk of the stock price in investment.
In other words, volatility tends to increase rapidly in response to bad news but decrease when good news appears.

For the most part, the stylized facts mentioned above are generally found in financial returns. This thesis attempts to establish the models of returns by taking into account some of the stylized facts, particularly the non-normal distribution with high kurtosis, skewness and a heavy tail.

2.4 Volatility Forecasting Models

Modeling and forecasting the volatility of financial returns have attracted a great deal of attention in the field of financial research. Firstly, volatility of returns plays crucial role in the global economy because it is often used as a measurement of market risk and quantifies the risk of instrument over that time period. Secondly, greater changes of volatility of financial returns raise public policy issues about the stability of financial markets. Policy makers usually rely on the estimation of volatility. Finally, the theoretical perspective of volatility of returns also plays an important role as in investor’s sentiment. It is used as a key for many investor decisions, portfolio allocation and risk management. Appropriate modeling for volatility of financial returns is able to lead to accurate forecasts of volatility. Therefore, it is important to develop an adequate model for modeling volatility of financial returns.

2.4.1 Development of Volatility Forecasting Models

The first simple model for forecasting volatility was the Random Walk model where the standard deviations at time $t$ ($\sigma_t$) are forecasted by the standard deviation at time $t-1$ ($\sigma_{t-1}$). This idea was extended to the Historical Average method, the simple Moving Average method, the Exponential Smoothing method
and the Exponentially Weighted Moving Average (EWMA) method (Poon and Granger, 2003).

A subsequent group of models includes the Autoregressive (AR) model, Moving Average (MA) model and a combination of AR and MA models (ARMA). The ARMA model is sometimes called the Box-Jenkins model (Box and Jenkins, 1970). The ARMA model is used to predict the conditional mean of stationary time series. An ARMA(p,q) process of the order \( p \) and \( q \) is defined as the following equation:

\[
r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},
\]

where \( \phi_0 \) is constant, \( \phi_i \) are the parameters of the autoregressive component of order \( p \), and \( \theta_i \) are the parameters of the moving average component of order \( q \). The \( \varepsilon_t \) is called a noise process or errors at time \( t \); \( p \) and \( q \) are non-negative integers. Mean and variance of a noise process \( \varepsilon_t \) are 0 and \( \sigma^2 \), respectively. In this model, volatility of returns (a noise process) is assumed to be constant over time. This model states that the current value of \( r_t \) depends linearly on its previous values \( r_s, s \leq t - 1 \), as well as current and previous errors \( \varepsilon_t \). In other words, the statistical properties of the past behavior of a time series \( r_t \) can be used to predict its behavior in the future.

In practice, the volatility of returns tends to change over time. Consequently, the conventional time series and econometric models seem unattractive for the financial time series. Researchers were very much aware of the changes in variance and developed new methods to model conditional variance.

Engle (1982) was the first person who developed the Autoregressive Conditional heteroskedasticity (ARCH) model to predict the conditional variance of return series. An ARCH(p) model, where the conditional variance depends on \( p \)
lagged square errors, is given by:

\[ r_t = \mu + \varepsilon_t \]  
(2.3)

\[ \varepsilon_t = \eta_t \sqrt{h_t} \]  
(2.4)

\[ h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2. \]  
(2.5)

where \( r_t \) is a return series, \( \{\eta_t\} \) are i.i.d random variables with \( E(\eta_t) = 0 \) and \( Var(\eta_t) = 1 \) and \( h_t \) is a conditional variance of returns at time \( t \) which must be non-negative. Consequently the ARCH model is able to overcome the phenomenon of conditional heteroskedasticity of financial returns.

Numerous applications of the ARCH model have been reported. Engle (1983) discussed the conditional variance of inflation rate using the ARCH model and found that the inflation tended to change over time. In Weiss (1984), ARMA models and the ARCH model were found to be incorporated successfully in modeling sixteen U.S. macroeconomic time series. Lastrapes (1989) confirmed that the ARCH model can account for many of the empirical regularities of weekly exchange rate data. The daily trading volume of stock markets were applied into the ARCH model by Lamoureux and Lastrapes (1990). Their results demonstrated the motivation for using ARCH models to study the behavior of asset prices. Bera and Higgins (1993) provided a discussion of major contributions to the ARCH model. They noted that the ARCH model was useful to capture various stylized properties of time series data, such as the leverage effect in volatility, volatility clustering, excess kurtosis and heavy tails.

However, the ARCH model has some weaknesses for higher order models. It turns out to be difficult to estimate parameters because the process requires long-lag length and a large number of parameters to be estimated (Zakaria and Abdalla, 2012). Another weakness of the ARCH model is that the process often produces negative coefficient estimators in parameters (\( \alpha_i \)). The more parameters
to be estimated in the ARCH model, the more likely it can obtain a negative estimated value. In order to avoid the long lag structure of the ARCH model and solve the negative coefficient problem, a Generalized ARCH (GARCH) model was developed by Bollerslev (1986). The GARCH model has been modified to accommodate the possibility of serial correlation in volatility. It contains a linear combination of lags of the squared residuals from the conditional return equation and lags from the conditional variance (Goudarzi, 2010). In other words, the GARCH model turns the AR process of the ARCH model into an ARMA model by adding in MA process. Empirical studies have found that the GARCH is a more parsimonious model than the ARCH model (Poon and Granger, 2003) and becomes a valuable model for volatility forecasting in financial time series. In particular, the GARCH(1,1) is the most popular model for estimating and forecasting the volatility. A GARCH(p,q) model is defined as below, assuming a log return series \( r_t = \mu + \varepsilon_t \) where \( \varepsilon_t \) is the error term at time \( t \). The \( \varepsilon_t \) follows a GARCH(p,q) model if

\[
\varepsilon_t = \eta_t \sqrt{h_t}, \quad (2.6)
\]

\[
h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}, \quad (2.7)
\]

where \( \{\eta_t\} \) are i.i.d random variables with mean equal to 0 and variance equal to 1; and \( \omega, \alpha_i \) and \( \beta_j \) are non-negative constants with \( \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1 \) to ensure conditional variance to be positive as well as stationary. When \( q = 0 \) the process reduces to the ARCH model. The ARCH model incorporated the autoregressive term in return series where as the GARCH model is superior to ARCH because it adds the general feature of conditional heteroskedasticity terms. The conventional GARCH model, which is called symmetric GARCH model, is not always a perfect model and could be improved because the error terms \( (\varepsilon_t) \) are typically assumed to be normally distributed. Consequently, the symmetric
GARCH model is less than adequate to fully account for some stylized fact of returns. Therefore, various modifications of GARCH models are proposed for improving the volatility forecasting performance by taking into account such kinds of financial stylized characteristics of return series. The GARCH model has been developed by allowing alternative non-normal error distributions.

2.4.2 Various GARCH Models with Alternative Distributional Assumptions

As previously mentioned, the distribution of underlying returns displays excess kurtosis, a heavy tail, a leverage effect and at times skewness. Whilst GARCH error terms are typically assumed to be normally distributed, their distributions are symmetrical. Consequently, traditional GARCH models with a normal error distribution may fail to capture some characteristics of financial returns and may not reflect asymmetric behavior of returns. To solve these problems, alternative distributional assumptions of GARCH error terms are taken into account. In addition, different types of GARCH models are adopted to capture some properties of returns for improving volatility forecasting performance.

The alternative distributions, including non-normal error distributions such as the Student-t, Generalized Error Distribution (GED) and Skewed GED (SGED) have been applied to the GARCH model. Lee and Pai (2010) used GARCH(1,1) models with normal, Student-t and skewed generalized error distributions (GARCH-N, GARCH-ST and GARCH-SGED respectively) to investigate the performance of volatility forecasting of the Real Estate Investment Trust (REIT). Their paper used a symmetric GARCH (1,1) model with those three types of error distributions. To evaluate the performance of forecasting on 1, 2, 5, 10 and 20 horizons, Mean Square Error (MSE) and Mean Absolute Error (MAE) were employed. Their empirical results showed that the GARCH-SGED model outperformed the GARCH-N and GARCH-ST models in volatility forecasting. Liu et al. (2009) ex-
examined the performance of volatility forecasting on daily prices of Shanghai and Shenzhen composite stock indices using two GARCH(1,1) models (GARCH(1,1) models with N and SGED). Their results confirm that the GARCH model with SGED is superior to the GARCH model with normal distribution.

Another limitation of the conventional GARCH model is its failure to explain the leverage effects in financial returns. Consequently, symmetric GARCH models widely incorporate alternative non-normal distributions in order to improve the volatility forecasting. These extensions to symmetric GARCH models led to the development of asymmetric GARCH models. The most popular asymmetric GARCH model is the Exponential GARCH (EGARCH) model. Nelson (1991) first developed the EGARCH model which specified conditional variance in logarithmic form. The strength of the EGARCH model is the ability to capture the leverage effect of financial time series, in contrast to the original GARCH model which cannot cope with this problem. Other asymmetric GARCH models are the GJR-GARCH model (Glosten et al., 1993), the Threshold GARCH (TGARCH) model (Zakoian, 1994) and the Quadratic GARCH (QGARCH) model (Sentana, 1995). For further details on these and other asymmetric GARCH models refer to Bollerslev (2007).

Alberg et al. (2006) investigated the forecasting performance of GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1) and Asymmetric Power ARCH (APARCH(1,1)) models with different error distributions: normal, Student-t, and skewed Student-t. Their results showed that the EGARCH model using a skewed Student-t distribution was the most successful in volatility forecasting. Shamiri and Isa (2009) examined the relative efficiency of three different types of GARCH models in terms of their volatility forecasting performance. They compared the performance of symmetric GARCH(1,1), asymmetric EGARCH(1,1) and non-linear asymmetric NAGARCH(1,1) models with six error distributions, namely normal,
skewed normal, Student-t, skewed Student-t, normal inverse Gaussian and generalized error distributions. They claimed that the EGARCH model provided better performance on volatility forecasting than the GARCH(1,1) model. The comparison between models with different error distributions showed that non-normal error distributions outperformed the normal distribution. Awartani and Corradi (2005) compared the relative predictive ability of the GARCH(1,1) model with the alternative asymmetric GARCH models, such as EGARCH, GJR-GARCH, QGARCH, TGARCH and Asymmetric GARCH (AGARCH) models. They used the daily S&P-500 Composite Price Index to model and forecast the volatility at 1-, 5-, 10-, 15-, 20- and 30- steps ahead. All models were considered under the normal error distribution. They found that asymmetric GARCH models were superior to GARCH(1,1) models for all prediction horizons. In particular, EGARCH provided the smallest MSE values, followed by the other asymmetric models. Bali (2007) investigated performance on the volatility forecasting of the GARCH model against eight asymmetric GARCH models. Some parts of his work was similar to that of Awartani and Corradi (2005) but they used four different types of error distributions: GED, SGED, Student-t and skewed Student-t distribution. The weekly observations of 3-month U.S. Treasury bills were employed in this study. Results showed that the volatility forecasts from the GED and student-t performed slightly better than the normal distribution. Among alternative GARCH models, the EGARCH and TGARCH models were superior to the GARCH model and provided the most accurate forecasts of future realized volatility. Liu and Hung (2010) also investigated the performance of volatility forecasting for the Standard & Poor’s 100 stock index series. They compared the symmetric GARCH model with three different types of distributions (normal, Student-t and skewed generalized error distribution) against asymmetric GARCH (GJR-GARCH and EGARCH) models. Their empirical results indicated that the
GJR-GARCH model achieved the most accurate volatility forecasts.

2.5 Cointegration of Financial Time Series

To improve the performance of volatility forecasting, researchers have attempted to relax the restrictions on symmetric GARCH models and asymmetric GARCH models by allowing various non-normal error distributions. If each time series can be individually fitted by using the GARCH model and these time series are cointegrated, the question is addressed in our studies whether detected cointegration relationships between these time series can further improve the volatility forecasting.

The cointegration method introduced by Engle and Granger (1987) is now widely employed in econometrics and financial time series. The concept of cointegration analysis is used to describe the relationship among non stationary time series. If two or more time series are cointegrated, it implies that there are a long-run equilibrium relationships among the time series and the cointegrated time series will move closely together overtime. Many cointegration tests can be use to investigate cointegration relationships among time series. One of the most popular methods is called the Johansen’s tests (see Johansen, 1991). Following brief literature review aims to present some of empirical studies related to the Johansen conintegration test.

Gan et al. (2006) examined the relationships between the New Zealand Stock Index and seven macroeconomic variables: inflation rate, exchange rate, gross domestic product, money supply, long term interest rate, short term interest rate and domestic retail oil price. These researchers used the Johansen’s tests to determine whether the seven macroeconomic variable are cointegrated with share prices in the New Zealand stock exchange. Their results showed that a long run relationship exists between the New Zealand Stock Index and the macroeconomic
variables tested. Click and Plummer (2005) examined the five stock markets in the Southeast Asian Nations (ASEAN-5) for cointegration relationships. Their results showed that the stock markets of Indonesia, Malaysia, Philippines, Singapore and Thailand were cointegrated in the period after the Asian financial crisis (July 1, 1998 through December 31, 2002). They concluded that ASEAN-5 stock markets are integrated in the economic sense. Lamba (2005) investigated the long-run relationships between the three South Asian equity markets of India, Pakistan and Sri Lanka and the developed equity markets of the US, UK and Japan. The results indicated that the Indian market was influenced and receptive to changes in the developed equity markets of the US, UK and Japan. Pakistan and Sri Lanka, however showed no such trends.

Most research on cointegration analysis is concerned with the identification of the relationships among stock market series. None of the research uses the knowledge of detected cointegration relationships among time series to further improve the accuracy of volatility forecasting. Therefore, it is of interest to find out whether cointegration information among underlying time series can be beneficial to the improvement of performance of volatility forecasting.

2.6 Value at Risk and Volatility Forecasting

Value at Risk (VaR) is one of a number of important concepts in risk management. In the beginning, VaR was simply calculated by assuming a normal distribution of returns, but the normal distribution is not always true in practice. The assumption of normality leads to bias in the VaR estimation and the results can be an underestimation or overestimation. Therefore, it is very important to find out what is the appropriate distribution to calculate VaR.

GARCH models have also become popular tools in modern risk management because they make accurate estimations of downside risk which is a key step in risk
management. The specific features in financial time series such as a heavy tail and volatility clustering lead to difficulties in downside risk evaluation. GARCH models can capture these features by taking into account non-normal distributional assumptions in the innovation processes. In risk management, the knowledge of future volatility is crucial in controlling the risk in investment. Due to the excess kurtosis and skewness in financial time series, the normal VaR has its drawbacks particularly when it is applied to financial risk management. Therefore, variants of GARCH models have been important to apply to VaR applications. The literature on VaR applications is reviewed in this section.

Substantial empirical studies have examined VaR estimation related to volatility modeling and forecasting. An improvement in VaR estimations was developed in conjunction with GARCH models with non-normal error distributions. So and Yu (2006) investigated different GARCH models (the Integrated GARCH; IGARCH and the Fractionally Integrated GARCH; FIGARCH) in VaR estimations and found that the GARCH model with a Student-t distribution was superior to that with a normal error distribution in determining an appropriate value of VaR for a long position with a 99% confidence level. Hung et al. (2008) studied the influence of heavy-tailed innovation processes on the performance of one-day-ahead VaR estimations using three GARCH models (GARCH-N, GARCH-t and GARCH-HT) for energy commodities when asset returns exhibit leptokurtic and heavy-tailed features. They found evidence that the GARCH-HT model based on the VaR approach achieved good accuracy at both high and low confidence levels. Angelides et al. (2004) examined the performance of an extensive family of GARCH models with three different error distributions (normal, Student-t and GED) to estimate the daily VaR of five stock indices. They found that EGARCH models with a Student-t distribution produced the most adequate VaR forecasts for the majority of stock market data.
Despite an extensive literature review on VaR forecasting, no previous work was found which discusses the VaR performance with GARCH models by taking into account the knowledge of cointegration. This thesis will investigate whether the knowledge of cointegration can achieve a more accurate VaR estimate by allowing the GARCH(1,1) model with various error distributions to be incorporated into cointegrating errors.
Chapter 3

GARCH(p,q) Model with Alternative Error Distributions

The GARCH model by Bollerslev (1986) is originally based on a normal conditional distribution. It is widely accepted that financial returns often tend to exhibit stylized statistical properties such as heavy tails, leptokurtic distribution, skewness, volatility clustering and the leverage effect. Consequently, the early generation of GARCH models with conditional normal distributed errors failed to sufficiently capture those main stylized characteristics in financial time series. This chapter attempts to present modeling and forecasting using the symmetric GARCH(p,q) model with alternative error distributions. To begin, the traditional GARCH model with normal error distribution is briefly introduced then the other alternative types of error distributional assumptions are detailed in the chapter.

3.1 GARCH(p,q) Model with Normal Error Distribution

The GARCH model is generally and widely used to model the volatility of financial time series. The time series $\varepsilon_t$ following a GARCH(p,q) model is defined
as:

$$\varepsilon_t = \eta_t \sqrt{h_t},$$  \hspace{1cm} (3.1)

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j},$$  \hspace{1cm} (3.2)

where $\omega$, $\alpha_i$ and $\beta_j$ are non-negative constants with $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ in order to ensure that the conditional variance ($h_t$) is positive as well as stationary.

The constraint of $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ implies that the unconditional variance of $\varepsilon_t$ is finite, whereas its conditional variance $h_t$ evolves over time. The $\{\eta_t\}$ in the GARCH($p,q$) model are i.i.d random variables with mean equal to 0 and variance equal to 1, and are assumed to be standard normally distributed. For $q$ equal to zero the process reduces to the ARCH model.

A GARCH model allows for an infinite number of squared errors to influence the current conditional variance. The conditional variance determined through GARCH is a weighted average of past squared residuals. The weights on past squared residuals are assumed to decline geometrically at a rate to be estimated from the data. In other words, the conditional variance is modeled as a linear function of both the past squared errors and past conditional variances.

The GARCH(1,1) model is the most popular and simplest model for volatility forecasting. The GARCH(1,1) model can be simply written as follows:

$$\varepsilon_t = \eta_t \sqrt{h_t},$$  \hspace{1cm} (3.3)

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$  \hspace{1cm} (3.4)

where $\varepsilon_t$ is the underlying process and the GARCH model with higher orders can be derived from Equation (3.2).

In general, there are three types of parameters in the probability density function: location, scale and shape. The location parameter indicates the position of distribution on the X-axis: the scale parameter controls the spread of the
variability of distribution; and the shape parameter controls the distributed vari-
ation of density function around the location parameter. The location and scale
parameters correspond to the mean and standard deviation respectively for the
normal distribution but this is not always true for non-normal distributions. The
next section will describe the alternative distributional error assumptions in the
GARCH(p,q) model which consist of five non-normal density functions (including
a normal distribution). The non-normal density functions of the five distributions
are reviewed. They will be considered by the empirical and simulated studies in
this thesis.

3.2 Non-normal Error Distributions Applied to
GARCH Model

The inability of the traditional GARCH model to capture the volatility for some
stylized fact of returns is well known. The main drawback to the traditional
GARCH model with normal error distribution is that it often fails to capture
stylized properties of underlying time series. Various non-normal error distribu-
tions have been suggested to solve this problem.

Previous studies have clearly confirmed that the conventional GARCH model
with a normal distribution error has failed to capture the volatility of return se-
ries. The modifications to GARCH models with the alternative non-normal error
distributions have been used to model and forecast volatility of returns and to
improve the performance of volatility forecasting. To find a better model for ex-
cessive third and fourth moments, GARCH error term ($\eta_t$) in (3.1) is allowed to
have different types of error distributions. Distributions such as the stable Pare-
tian distribution, Logistic distribution, mixture of normal distributions, Student-t
distribution, generalized error distribution and skewed distributions have a heavy
tail in their density function and their tails are thicker than in a normal distribu-
CHAPTER 3. GARCH(P,Q) MODEL WITH ALTERNATIVE ERROR DISTRIBUTIONS

error terms in the GARCH model having these types of distributions are more flexible to capture the stylized properties of financial returns. Some previous studies are presented below.

Curto et al. (2009) developed GARCH(1,1) models with stable Paretian distribution and compared the models with normal and Student-t error distributions using the daily returns of US, German and Portuguese stock indices. A stable Paretian distribution can be defined by four parameters: zero location, unit scale, skewness $\beta \in [-1:1]$ and tail thickness $\alpha \in (0,2]$. They demonstrated that the volatility forecasting performance of GARCH model with stable Paretian distribution clearly performed better than the normal distribution and slightly better than the Student-t error distribution.

Lopez (2001) showed the performance of forecasting accuracy for volatility in GARCH(1,1) models with different error distributions using four exchange rates: the daily British pound, the Canadian dollar, the Deutschemark and the Japanese Yen. The first three volatility models were GARCH(1,1) models with the normal error distribution, the Student-t error distribution and the generalized error distribution. The fourth and fifth GARCH models were incorporated into the exponential smoothing model and the stochastic volatility model respectively. Five measurements were used to evaluate the performance of volatility forecasting which consisted of the mean square error, mean absolute error, logarithmic loss function, heteroscedasticity-adjusted mean square error and Gaussian quasi-maximum likelihood function. The results showed that the GARCH model with exponential smoothing model was better than the other competing GARCH models when comparing the performance of volatility forecasting. Lin and Yeh (2000) applied the mixture of two normal distributions combined with the GARCH(1,1) model to Taiwan stock market returns. They found that the GARCH-mixed-normal model outperformed the GARCH model with normal dis-
Researchers attempted to incorporate heavy tail distributions into GARCH models by adopting a variety of non-normal error distributions. The comparisons of competing GARCH models with complicated error distributions on the performance of volatility forecasting have been examined but this issue remains interesting. In this thesis, five different types of distributions in GARCH error terms ($\eta_t$) are considered in addition to the normal distribution: the skewed normal, Student-t, skewed Student-t, generalized error and skewed generalized error distributions. These are well known and allow the tail of the distributions to be thicker than the normal distribution and their tails sometimes exhibit skewness.

In general, a GARCH(1,1) model is the simplest and most successful model for volatility forecasting in financial returns. Investigations into the volatility forecasting performance of GARCH models were reported mainly on GARCH(1,1) models with different distributional assumptions. Shamiri and Isa (2009) examined three different types of GARCH(1,1) models for modeling and forecasting volatility using the Kuala Lumpur Composite Index. Gokcan (2000) compared the performance of volatility forecasting of the GARCH(1,1) model against the EGARCH(1,1) model using the monthly stock market returns of seven emerging countries. Chuang et al. (2007) investigated the volatility forecasting performance of the GARCH(1,1) model with various distributional assumptions on stock market indices and exchanges markets. Komain (2007) used the ARMA-GARCH(1,1) model to examine the behavior of the stock index of Thailand.

Interestingly, the GARCH(1,1) model may not necessarily be the best fitted model in practice. This chapter is not only focused on the GARCH(1,1) model but also investigates whether it is more appropriate to fit the returns series of emerging stock markets into a higher order GARCH model in order to forecast
the volatility of financial time series.

The six different types of probability density functions used in this thesis are presented below:

1. Normal Distribution (N)

\[ f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty, \quad (3.5) \]

2. Skewed Normal Distribution (SN) \(^1\)

\[ f(z) = \frac{1}{\omega \pi} e^{-\frac{(z-\xi)^2}{2\omega^2}} \int_{-\infty}^{\vartheta} e^{-\frac{t^2}{2\omega^2}} dt, \quad -\infty < z < \infty, \quad (3.6) \]

where \(\xi\) denotes the location, \(\omega\) denotes the scale and \(\vartheta\) denotes the shape of density.

3. Student-t Distribution (STD) \(^2\)

\[ f(z) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu \pi} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{z^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}, \quad -\infty < z < \infty, \quad (3.7) \]

where \(\nu\) denotes the number of degrees of freedom and \(\Gamma\) denotes the Gamma function.

4. Skewed Student-t Distribution (SSTD) \(^3\)

\[ f(z; \mu, \sigma, \nu, \lambda) = \begin{cases} 
bc(1 + \frac{1}{\nu-2}\left(\frac{b(z-\mu) + a}{1-\lambda}\right)^{\frac{\nu-1}{2}}), & \text{if } z < -\frac{a}{b}, \\
bc(1 + \frac{1}{\nu-2}\left(\frac{b(z-\mu) + a}{1+\lambda}\right)^{\frac{\nu-1}{2}}), & \text{if } z \geq -\frac{a}{b}, 
\end{cases} \quad (3.8) \]

where \(\nu\) is the shape parameter with \(2 < \nu < \infty\) and \(\lambda\) is the skewness parameter.

---

\(^1\) See Shamiri and Isa (2009)
\(^2\) See Shamiri and Isa (2009)
\(^3\) See Bali (2007)
with $-1 < \lambda < 1$. The constants $a$, $b$ and $c$ are given below

$$a = 4\lambda c\left(\frac{\nu - 2}{\nu - 1}\right), \quad b = 1 + 3\lambda^2 - a^2, \quad c = \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{\sqrt{\pi(\nu - 2)}\Gamma\left(\frac{\nu}{2}\right)}$$

where $\mu$ and $\sigma^2$ are the mean and variance of the skewed student-t distribution, respectively.

5. Generalized Error Distribution (GED) \(^4\)

$$f(z; \mu, \sigma, \nu) = \frac{\sigma^{-1}\nu e^{-0.5\left(\frac{z - \mu}{\sigma}\right)^{\nu}}}{\lambda^2(1 + (1/\nu))\Gamma(1/\nu)}, \quad 1 < z < \infty, \quad (3.9)$$

where $\nu > 0$ is the degrees of freedom or tail-thickness parameter and

$$\lambda = \sqrt{2^{-2/\nu} \Gamma(1/\nu)/\Gamma(3/\nu)}.$$ 

If $\nu = 2$, the GED yields the normal distribution. If $\nu < 1$, the density function has thicker tails than the normal density function, whereas for $\nu > 2$ it has thinner tails.

6. Skewed Generalized Error Distribution (SGED) \(^5\)

$$f(z; \nu, \xi) = \nu[2\theta\Gamma(1/\nu)]^{-1} \exp\left(-\frac{|z - \delta|^{\nu}}{1 - \text{sign}(z - \delta)\xi^{\nu}\theta^{\nu}}\right) \quad (3.10)$$

where

$$\theta = \Gamma(1/\nu)^{0.5}\Gamma(3/\nu)^{-0.5}S(\xi)^{-1},$$

$$\delta = 2\xi AS(\xi)^{-1},$$

$$S(\xi) = \sqrt{1 + 3\xi^2 - 4A^2\xi^2},$$

$$A = \Gamma(2/\nu)\Gamma(1/\nu)^{-0.5}\Gamma(3/\nu)^{-0.5},$$

where $\nu > 0$ is the shape parameter controlling the height and heavy-tail of the

\(^4\)See Bali (2007)

\(^5\)See Liu et al. (2009)
density function while $\xi$ is a skewness parameter of the density with $-1 < \xi < 1$.

### 3.3 Model Selection Using Information Criteria

Model selection is an important part of statistical forecasting. An appropriate model leads to accurate forecasts by regarding information criteria. Model selection involves the use of information criteria to identify the best fitting model from a set of competing models. The best fitting model is measured by the smallest value of the information criteria. Commonly used information criteria for model selection will be briefly presented in this section.

1. Akaike’s Information Criteria (AIC)

AIC is one of the most commonly used criteria for model selection. It is a statistical measure of likelihood of the set of competing models, penalized by the number of parameters in the models. AIC is defined as:

$$AIC = -2\log (L) + 2p$$  \hspace{1cm} (3.11)

where $L$ is the likelihood under the fitted model and $p$ is the number of parameters in the model. The smaller the AIC values, the better fitting the model will be.

2. Bayesian Information Criteria (BIC)

Another information criteria is called BIC. It is derived based on Bayes factors. BIC is defined as:

$$BIC = -2\log (L) + p \log(T)$$  \hspace{1cm} (3.12)

The BIC criterion is different from AIC only in the last term in (3.11) and (3.12).
BIC depends on sample size $T$ and $p$ while AIC depends on $p$ only.

There are other model selection criteria such as Draper’s Information Criteria or DIC (Draper, 1995) and Hannan Quinn Information Criteria or HQC (Hannan and Quinn, 1979) can be used.

It is important to decide which criterion will be appropriate to use for model selection. Markon and Krueger (2004) claimed that the ability for model selection of AIC performs well for a small sample sizes, but does not perform well for larger sample sizes. BIC performs well with larger sample sizes while AIC seems more popular to use as a measurement for model selection. For selecting the best model, AIC criteria is adopted in this thesis.

### 3.3.1 Determining the Order of GARCH Models

Most studies on volatility forecasting for financial time series use only lower order GARCH(1,1) models. To further improve the volatility forecasting performance of underlying financial time series, researchers have suggested different types of non-normal error distributions. However, the GARCH(1,1) model is still commonly used to identify the best fitting model. The GARCH(p,q) model with a higher order might be more appropriate for some financial time series than a GARCH(1,1) model. Therefore, it is important to know how to determine the appropriate order of a GARCH(p,q) model. There are two different situations that need consideration when determining the order of a GARCH(p,q) model. One case is when the error terms in the GARCH(p,q) model are normally distributed and the other is when a GARCH model has error terms that are non-normally distributed.

Under the normality assumption, the order determination for the GARCH(p,q) model is straightforward. To identify the best fitting GARCH model, the competing GARCH(p,q) models are examined by using the statistical measurements as following:
1. Using information criteria such as AIC, BIC or log-likelihood values to choose the appropriate order of a GARCH(p,q) model, determined by considering which GARCH(p,q) models provide the smallest information criteria values.

2. Considering Ljung-Box $Q^2$ statistics in order to test for serial correlation of the squared standardized residuals.

3. Using the Lagrange multiplier test (LM) to evaluate whether ARCH effects are removed from the standardized residuals.

However, the determination of the order of GARCH(p,q) models remains inconclusive under the non-normality assumption. In the literature, there is no detailed information on the process of choosing the best fitting GARCH(p,q) model from among the competing models.

This thesis will suggest a possible way for determining an appropriate order among the competing GARCH(p,q) models when the error terms are non-normally distributed. The suggestion is presented below:

Step 1. Use the AIC criterion to select an appropriate order among competing GARCH(p,q) models (and also check coefficients are significant) by assuming the error terms are normally distributed.

Step 2. Change the error distribution in the selected GARCH(p,q) model from step 1 and identify an appropriate error distribution for the GARCH(p,q) model in terms of AIC value.
Step 3. Use the Ljung-Box $Q^2$ statistics and LM test as diagnostics tests.

To better understanding this suggestion, simulation studies are demonstrated in the next subsection.

### 3.3.2 Simulation Study on Order Determination

In order to illustrate how to choose the order of GARCH(p,q) model under the non-normality assumption, observations from GARCH(p,q) with the Student-t, skewed Student-t and GED error distributions are generated respectively. Each set of observation has size 3,000. The simulations using the GARCH model start from the order (1,1) until order (2,2). There are five GARCH models: GARCH(1,1), GARCH(1,2), GARCH(1,3), GARCH(2,1) and GARCH(2,2) are considered in this simulation study. These five GARCH models are considered as the true GARCH models with three different types of error distributions mentioned above. The five sets of parameters in the true GARCH(p,q) models from Equation (3.2) are initially set as follows:

**GARCH(1,1) model,**

$$h_t = 0.01 + 0.3\varepsilon^2_{t-1} + 0.5h_{t-1}.$$  

**GARCH(1,2) model,**

$$h_t = 0.01 + 0.3\varepsilon^2_{t-1} + 0.5h_{t-1} + 0.02h_{t-2}.$$  

**GARCH(1,3) model**

$$h_t = 0.01 + 0.3\varepsilon^2_{t-1} + 0.5h_{t-1} + 0.02h_{t-2} + 0.01h_{t-3}.$$
GARCH(2,1) model

\[ h_t = 0.01 + 0.3\varepsilon_{t-1}^2 + 0.01\varepsilon_{t-2}^2 + 0.5h_{t-1}. \]

GARCH(2,2) model

\[ h_t = 0.01 + 0.3\varepsilon_{t-1}^2 + 0.01\varepsilon_{t-2}^2 + 0.5h_{t-1} + 0.02h_{t-2}. \]

To examine the order of GARCH models, the 3,000 observations drawn from each true GARCH(p,q) model are divided into two parts. The first part is called an in-sample data set and is used to build up a model for the data set. The second part is called an out-sample data set and is used for investigating the performance of forecasting, but this part is omitted because the volatility forecasting performance does not be considered in this chapter. Therefore, the first 1,500 observations of each true GARCH(p,q) model are used to build up a model that assumed the normal error distribution. Coefficients in all models must be significant. Then, the AIC criterion is used to identify whether the order of GARCH model continue to correspond to the true models. If the order of each fitting GARCH model corresponds to the true GARCH model, it indicates that the order determination in each GARCH model with non-normal error distribution is valid. This suggestion for order determination of the GARCH(p,q) model under non-normal error distribution is officially acceptable as discussed in Subsection 3.3.1. The results of these simulations are reported in Table 3.1.

Table 3.1 shows AIC values of GARCH(p,q) models with a normal error distribution when the distribution of the error terms of true GARCH(p,q) models are STD, SSTD and GED respectively. By regarding significant coefficients of parameters in each model, it can be seen that the orders of the true GARCH(p,q) models with STD, SSTD and GED distributions are identical to the order of GARCH(p,q) models by assuming a normal error distribution. Therefore, the re-
### Table 3.1: AIC values when true GARCH(p,q) models are from STD, SSTD and GED error distributions

<table>
<thead>
<tr>
<th>Normal-GARCH(p,q)</th>
<th>The order of true GARCH(p,q) model with STD</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>-0.2959*</td>
<td>-0.1981*</td>
<td>-0.1454*</td>
<td>-0.2555*</td>
<td>-0.1538*</td>
</tr>
<tr>
<td>(1,1)</td>
<td><strong>-0.3475</strong></td>
<td>-0.2599*</td>
<td>-0.2117*</td>
<td>-0.3164*</td>
<td>-0.2260*</td>
</tr>
<tr>
<td>(1,2)</td>
<td>-0.3466</td>
<td><strong>-0.2618</strong></td>
<td>-0.2110*</td>
<td>-0.3152</td>
<td>-0.2249</td>
</tr>
<tr>
<td>(1,3)</td>
<td>-0.3446</td>
<td>-0.2574</td>
<td><strong>-0.2119</strong></td>
<td>-0.3134</td>
<td>-0.2231</td>
</tr>
<tr>
<td>(2,1)</td>
<td>-0.3457</td>
<td>-0.2581</td>
<td>-0.2099</td>
<td><strong>-0.3166</strong></td>
<td>-0.2242*</td>
</tr>
<tr>
<td>(2,2)</td>
<td>-0.3453</td>
<td>-0.2578</td>
<td>-0.2096</td>
<td>-0.3139</td>
<td><strong>-0.2276</strong>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal-GARCH(p,q)</td>
<td>The order of true GARCH(p,q) model with SSTD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>-0.3864*</td>
<td>-0.2967*</td>
<td>-0.2477*</td>
<td>-0.3515*</td>
<td>-0.2967*</td>
</tr>
<tr>
<td>(1,1)</td>
<td><strong>-0.4260</strong></td>
<td>-0.3429*</td>
<td>-0.2973*</td>
<td>-0.3973*</td>
<td>-0.3429*</td>
</tr>
<tr>
<td>(1,2)</td>
<td>-0.4245</td>
<td><strong>-0.3434</strong></td>
<td>-0.2959*</td>
<td>-0.3957</td>
<td>-0.3414</td>
</tr>
<tr>
<td>(1,3)</td>
<td>-0.4232</td>
<td>-0.2947</td>
<td><strong>-0.2977</strong></td>
<td>-0.3943</td>
<td>-0.3402</td>
</tr>
<tr>
<td>(2,1)</td>
<td>-0.4244</td>
<td>-0.2957</td>
<td>-0.2945</td>
<td><strong>-0.3977</strong></td>
<td>-0.3413*</td>
</tr>
<tr>
<td>(2,2)</td>
<td>-0.4234</td>
<td>-0.2946</td>
<td>-0.2901</td>
<td>-0.3946</td>
<td><strong>-0.3435</strong>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal-GARCH(p,q)</td>
<td>The order of true GARCH(p,q) model with GED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>-0.2560*</td>
<td>-0.1592*</td>
<td>-0.1069*</td>
<td>-0.2152*</td>
<td>-0.1143*</td>
</tr>
<tr>
<td>(1,1)</td>
<td><strong>-0.2757</strong></td>
<td>-0.1822*</td>
<td>-0.1312*</td>
<td>-0.2391*</td>
<td>-0.1421*</td>
</tr>
<tr>
<td>(1,2)</td>
<td>-0.2743</td>
<td>-0.1813</td>
<td><strong>-0.1827</strong></td>
<td>-0.2387*</td>
<td>-0.1422*</td>
</tr>
<tr>
<td>(1,3)</td>
<td>-0.2743</td>
<td>-0.1813</td>
<td><strong>-0.1327</strong></td>
<td>-0.2373</td>
<td>-0.1408</td>
</tr>
<tr>
<td>(2,1)</td>
<td>-0.2740</td>
<td>-0.1805</td>
<td>-0.1295</td>
<td><strong>-0.2393</strong></td>
<td>-0.1404*</td>
</tr>
<tr>
<td>(2,2)</td>
<td>-0.2748</td>
<td>-0.1817</td>
<td>-0.1309</td>
<td>-0.2379</td>
<td><strong>-0.1423</strong>*</td>
</tr>
</tbody>
</table>

Notes: STD denotes Student-t distribution; SSTD denotes Skewed Student-t distribution; GED denotes Generalized error distribution; (*) indicates that coefficients are significant.
results from the simulations in Table 3.1 confirm that the order of a GARCH(p,q) model can be determined by considering the smallest AIC values and assuming normal error distribution although the true error distribution is the non-normally distributed assumptions.

### 3.4 Volatility Forecasts via GARCH(p,q) Models

The volatility forecasting for multiple steps ahead in the GARCH(1,1) model has been mentioned by Tsay (2010). However, to derive a general formula term of higher order in GARCH(p,q) models for multi step-ahead forecasts, the formulas for multiple steps ahead in GARCH(1,1) are used. In this section, the current time of volatility forecasting is defined as \( n \); 1-step ahead volatility in GARCH(1,1) model is calculated by,

\[
h_{n+1} = \omega + \alpha_1 \varepsilon_n^2 + \beta_1 h_n, \tag{3.13}
\]

where \( \varepsilon_n \) and \( h_n \) are known at the time \( n \). Denote the 1-step ahead forecast given current information as

\[
h_n(1) = \omega + \alpha_1 \varepsilon_n^2 + \beta_1 h_n. \tag{3.14}
\]

where \( h_n(1) = E(h_{n+1}|I_n) \) and \( I_n \) is denoted the information \( \varepsilon \) available updated to time \( n \).

For multi-step ahead forecasts, by using \( \varepsilon_t^2 = \eta_t^2 h_t \), the volatility equation from (3.13) is rewritten as

\[
h_{t+1} = \omega + (\alpha_1 + \beta_1)h_t + \alpha_1 h_t(\eta_t^2 - 1). \tag{3.15}
\]

When \( t = n + 1 \), Equation (3.15) becomes

\[
h_{n+2} = \omega + (\alpha_1 + \beta_1)h_{n+1} + \alpha_1 h_{n+1}(\eta_{n+1}^2 - 1). \tag{3.16}
\]
CHAPTER 3. GARCH(P,Q) MODEL WITH ALTERNATIVE ERROR DISTRIBUTIONS

Since \( E(\eta_{n+1}^2 - 1|I_n) = 0 \) where \( I_n \) is denoted the information \( \varepsilon \) available updated to time \( n \), the 2-step ahead volatility forecast at the origin \( n \) satisfies the equation

\[
h_n(2) = \omega + (\alpha_1 + \beta_1)h_n(1). \tag{3.17}\]

In general, we have

\[
h_n(k) = \omega + (\alpha_1 + \beta_1)h_n(k-1), \quad k > 1, \tag{3.18}\]

which can be written as

\[
h_n(k) = \frac{\omega[1 - (\alpha_1 + \beta_1)^{k-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{k-1}h_n(1). \tag{3.19}\]

Therefore,

\[
h_n(k) \to \frac{\omega}{1 - \alpha_1 - \beta_1}, \quad \text{as } k \to \infty, \tag{3.20}\]

where \( 0 \leq \alpha_1, \beta_1 \leq 1 \) and \( \alpha_1 + \beta_1 < 1 \).

Now we derive the formula for volatility forecasting for GARCH(1,2) model.

In GARCH(1,2), \( \varepsilon_t = \eta_t \sqrt{h_t} \) and

\[
h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2}. \tag{3.21}\]

The 1-step ahead forecast at the origin \( n \) is

\[
h_{n+1} = \omega + \alpha_1 \varepsilon_n^2 + \beta_1 h_n + \beta_2 h_{n-1}, \tag{3.22}\]

\[
= \omega + \alpha_1 (\varepsilon_n^2 - h_n) + \alpha_1 h_n + \beta_1 h_n + \beta_2 h_{n-1}, \tag{3.23}\]

\[
= \omega + \alpha_1 (\eta_n^2 h_n - h_n) + \alpha_1 h_n + \beta_1 h_n + \beta_2 h_{n-1}, \tag{3.24}\]

\[
= \omega + \alpha_1 h_n [\eta_n^2 - 1] + (\alpha_1 + \beta_1) h_n + \beta_2 h_{n-1}. \tag{3.25}\]

But we know \( E(\eta_n^2 - 1|I_n) = 0 \), then 1-step ahead forecast for GARCH(1,2)
CHAPTER 3. GARCH(P,Q) MODEL WITH ALTERNATIVE ERROR DISTRIBUTIONS

model can be written as

\[ h_n(1) = \omega + (\alpha_1 + \beta_1)h_n + \beta_2h_{n-1}. \] (3.26)

From (3.26), the multi-step ahead forecasts for GARCH(1,2) model can be obtained recursively as follows:

\[ h_n(2) = \omega + (\alpha_1 + \beta_1)h_n(1) + \beta_2h_n, \] (3.27)

\[ h_n(3) = \omega + (\alpha_1 + \beta_1)h_n(2) + \beta_2h_n(1), \] (3.28)

\[ h_n(4) = \omega + (\alpha_1 + \beta_1)h_n(3) + \beta_2h_n(2). \] (3.29)

If we denote \( h_{n-i} \) by \( h_n(-i) \), the general k-step ahead forecast is

\[ h_n(k) = \omega + (\alpha_1 + \beta_1)h_n(k-1) + \beta_2h_n(k-2). \] (3.30)

For GARCH(1,3) model, to derive k-step ahead forecasts for GARCH(1,3) model, a GARCH(1,3) model is considered and defined as follows:

\[ \epsilon_t = \eta_t \sqrt{h_t}, \] (3.31)

\[ h_t = \omega + \alpha_1\epsilon_{t-1}^2 + \beta_1h_{t-1} + \beta_2h_{t-2} + \beta_3h_{t-3}. \] (3.32)

where \( \eta_t \) is i.i.d with \( E(\eta_t)=0 \) and \( Var(\eta_t)=1 \).

The 1-step ahead forecast at the origin \( n \) is

\[ h_{n+1} = \omega + \alpha_1\epsilon_n^2 + \beta_1h_n + \beta_2h_{n-1} + \beta_3h_{n-2}, \] (3.33)

\[ = \omega + \alpha_1(\epsilon_n^2 - h_n) + \alpha_1h_n + \beta_1h_n + \beta_2h_{n-1} + \beta_3h_{n-2}, \] (3.34)

\[ = \omega + \alpha_1[\eta_n^2h_n - h_n] + \alpha_1h_n + \beta_1h_n + \beta_2h_{n-1} + \beta_3h_{n-2}, \] (3.35)

\[ = \omega + \alpha_1h_n[\eta_n^2 - 1] + (\alpha_1 + \beta_1)h_n + \beta_2h_{n-1} + \beta_3h_{n-2}. \] (3.36)

But we know \( E(\eta_n^2 - 1|I_n) = 0 \), then 1-step ahead forecast for GARCH(1,3)
model can be written as

\[ h_n(1) = \omega + (\alpha_1 + \beta_1)h_n + \beta_2h_{n-1} + \beta_3h_{n-2}. \] (3.37)

From (3.37), the multi-step ahead forecasts for GARCH(1,3) model can be obtained recursively as follows:

\[ h_n(2) = \omega + (\alpha_1 + \beta_1)h_n(1) + \beta_2h_n + \beta_3h_{n-1}, \] (3.38)
\[ h_n(3) = \omega + (\alpha_1 + \beta_1)h_n(2) + \beta_2h_n(1) + \beta_3h_n, \] (3.39)
\[ h_n(4) = \omega + (\alpha_1 + \beta_1)h_n(3) + \beta_2h_n(2) + \beta_3h_n(1). \] (3.40)

Therefore, the general k-step ahead forecast is

\[ h_n(k) = \omega + (\alpha_1 + \beta_1)h_n(k-1) + \beta_2h_n(k-2) + \beta_3h_n(k-3). \] (3.41)

For GARCH(2,1) model, to derive k-step ahead forecasts for GARCH(2,1) model, a GARCH(2,1) model is considered and defined as follows:

\[ \varepsilon_t = \eta_t \sqrt{h_t}, \] (3.42)
\[ h_t = \omega + \alpha_1 \varepsilon_t^2 + \alpha_2 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \] (3.43)

where \( \eta_t \) is i.i.d with \( E(\eta_t)=0 \) and \( Var(\eta_t)=1. \)

We assume that the forecast origin is \( n. \) The 1-step ahead forecast is

\[ h_{n+1} = \omega + \alpha_1 \varepsilon_n^2 + \alpha_2 \varepsilon_{n-1}^2 + \beta_1 h_n, \] (3.44)
\[ = \omega + \alpha_1(\varepsilon_n^2 - h_n) + \alpha_1 h_n + \alpha_2(\varepsilon_{n-1}^2 - h_{n-1}) + \alpha_2 h_{n-1} + \beta_1 h_n, \] (3.45)
\[ = \omega + \alpha_1[\eta_n^2 h_n - h_n] + \alpha_1 h_n + \alpha_2[\eta_{n-1}^2 h_{n-1} - h_{n-1}] + \alpha_2 h_{n-1} + \beta_1 h_n, \] (3.46)
\[ = \omega + \alpha_1 h_n[\eta_n^2 - 1] + (\alpha_1 + \beta_1)h_n + \alpha_2 h_{n-1}[\eta_{n-1}^2 - 1] + \alpha_2 h_{n-1}. \] (3.47)
We know $E(\eta_n^2 - 1|I_n) = 0$, then 1-step ahead forecast for GARCH(2,1) model can be written as

$$h_n(1) = \omega + (\alpha_1 + \beta_1)h_n + \alpha_2h_{n-1}. \hspace{1cm} (3.48)$$

From (3.48), the multi-step ahead forecasts for GARCH(2,1) model can be obtained recursively as follows:

$$h_n(2) = \omega + (\alpha_1 + \beta_1)h_n(1) + \alpha_2h_n, \hspace{1cm} (3.49)$$
$$h_n(3) = \omega + (\alpha_1 + \beta_1)h_n(2) + \alpha_2h_n(1), \hspace{1cm} (3.50)$$
$$h_n(4) = \omega + (\alpha_1 + \beta_1)h_n(3) + \alpha_2h_n(2). \hspace{1cm} (3.51)$$

Therefore, the general k-step ahead forecast is

$$h_n(k) = \omega + (\alpha_1 + \beta_1)h_n(k - 1) + \alpha_2h_n(k - 2) \hspace{1cm} (3.52)$$

Similarly, the k-step ahead forecasts for a general GARCH(p,q) model can be derived in the same way. In general, the volatility forecasting for GARCH(p,q) model can be written as

$$h_n(k) = \omega + (\alpha_1 + \beta_1)h_n(k - 1) + \alpha_2h_n(k - 2) + \ldots + \alpha_p h_n(k - p)$$
$$+ \beta_2 h_n(k - 2) + \ldots + \beta_q h_n(k - q) \hspace{1cm} (3.53)$$

### 3.5 Evaluation of Volatility Forecasts

To compare the volatility forecasting ability of competing GARCH models, the forecasting performance of the different models are evaluated using error measurements. Several criteria of error measurements can be used such as the Mean Squared Error (MSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). In this thesis, MSE and MAE criteria are used as measurements for evaluating volatility forecasting performances. The MSE and MAE for


$k$ steps ahead in a forecast are defined as follows:

\[
MSE(k) = \frac{1}{T} \sum_{t=1}^{T} \left[ (r_{t+k} - \bar{r})^2 - h_t(k) \right]^2, \quad (3.54)
\]

\[
MAE(k) = \frac{1}{T} \sum_{t=1}^{T} |(r_{t+k} - \bar{r})^2 - h_t(k)|. \quad (3.55)
\]

where

- \( r_{t+k} \) denotes the return over horizon \( k \) steps ahead at current time \( t \),
- \( \bar{r} \) denotes the mean of return,
- \( h_t(k) \) denotes the forecasted conditional variance over horizon \( k \) steps ahead at current time \( t \).

Since the true volatility cannot be observed, it is important to choose an unbiased estimator for the true underlying volatility. This thesis adopts the suggestion of Awartani and Corradi (2005) to use squared returns \((r_t - \bar{r})^2\) as a proxy of volatility for latent volatility in this scenario. To evaluate the performance of volatility forecasting of the best fitting model compared with the best performance model, simulation studies are demonstrated in the next section using R package.

### 3.6 Simulation Study on the Performances of Volatility Forecasting: Comparing between the Best Fitting Model and the Best Performance Model

Shamiri and Isa (2009) stated that the best fitting model based on AIC criterion is not necessarily a model that is able to provide the best forecast of volatility in terms of MSE and MAE. Their conclusion was based on the study on the KLCI fitted to GARCH(1,1), EGARCH(1,1) and NAGARCH(1,1) models. For some underlying series of financial data sets, a higher order GARCH might be more
appropriate than a GARCH(1,1) model. Therefore, it is of interest to investigate whether Shamiri and Isa’s statement is still acceptable when the underlying financial time series are used to fit by GARCH(p,q) models with a higher order.

Simulated data from the following two models are used to carry out this study. Let \( r_t = \mu + \varepsilon_t \) and \( \varepsilon_t \) follows these two GARCH models: GARCH(1,3) and GARCH(2,1) models which are defined as follows:

GARCH(1,3) model :

\[
h_t = 0.00007 + 0.02354\varepsilon_{t-1}^2 + 0.05387h_{t-1} + 0.00127h_{t-2} + 0.18574h_{t-3}\] (3.56)

GARCH(2,1) model :

\[
h_t = 0.00008 + 0.05334\varepsilon_{t-1}^2 + 0.06147\varepsilon_{t-2}^2 + 0.08599h_{t-1}.\] (3.57)

Both are higher order of GARCH models. The coefficients of two models are borrowed from the fitted models with normal error distribution for the real observations of the Stock Exchange of Thailand (SET) and Straits Time Index (STI) from Singapore, respectively. Simulated 6,536 and 5,407 observations from the GARCH(1,3) and GARCH(2,1) models are considered respectively. Six types of error distributions including Normal(N), Skewed Normal(SN), Student-t(STD), Skewed Student-t(SSTD), GED and Skewed GED(SGED) (see Chapter 3, Section 3.2) are employed for \( \varepsilon_t \) in these simulations. The main reason for choosing these six types of error distributions is to take into account some stylized facts of return such as excess kurtosis, skewness and heavy-tails. Each data set are divided into two parts. The first part is in-sample observations which are used to estimate the coefficients in the fitting model (3,535 and 2,500 observations for
CHAPTER 3. GARCH(P,Q) MODEL WITH ALTERNATIVE ERROR DISTRIBUTIONS

GARCH(1,3) and GARCH(2,1) model, respectively). The second part serve as out-sample observations and is used for investigating the performance of volatility forecasting.

Each data set is fitted by the same order of GARCH model where the data are simulated from, with six different types of error distribution respectively. The AIC values given by each fitting are used to determine which model is the best fitting model for the underlying data set. The out-of-sample on 1-step ahead volatility forecasting are evaluated and the performance of each GARCH model with different types error distributions measured by MSE and MAE are compared. The results are shown in Table 3.2 and Table 3.3.

Table 3.2: The values of AIC for simulated data from the GARCH(1,3) model

<table>
<thead>
<tr>
<th>Error in true distribution</th>
<th>The distribution used in the fitted model</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,3)</td>
<td>N</td>
</tr>
<tr>
<td>N</td>
<td>-6.3972</td>
</tr>
<tr>
<td>SN</td>
<td>-6.4646</td>
</tr>
<tr>
<td>STD</td>
<td>-6.3597</td>
</tr>
<tr>
<td>GED</td>
<td>-6.3884</td>
</tr>
<tr>
<td>SGED</td>
<td>-6.4419</td>
</tr>
</tbody>
</table>

Table 3.3: The values of AIC for simulated data from the GARCH(2,1) model

<table>
<thead>
<tr>
<th>Error in true distribution</th>
<th>The distribution used in the fitted model</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(2,1)</td>
<td>N</td>
</tr>
<tr>
<td>N</td>
<td>-6.3465</td>
</tr>
<tr>
<td>SN</td>
<td>-6.3469</td>
</tr>
<tr>
<td>GED</td>
<td>-6.3393</td>
</tr>
<tr>
<td>SGED</td>
<td>-6.3893</td>
</tr>
</tbody>
</table>

From Tables 3.2 and 3.3, it can be seen that the true model of GARCH(1,3)
and GARCH(2,1) are always the best fitting model with regards to the smallest AIC values. For example, if the true model is generated from GARCH(1,3) with normal error distribution, the comparisons among the AIC values of GARCH(1,3) models with other competing error distributions (SN, STD, SSTD, GED and SGED) are considered. It is found that the AIC value of the GARCH(1,3) model with normal error distribution is always the smallest value. Similarly, when GARCH(1,3) with skewed Student-t is the true model, it produces the smallest AIC value compared with the remaining error distributions.

Table 3.4: The values of MSE and MAE for simulated data from the GARCH(1,3) model

<table>
<thead>
<tr>
<th>Error in true GARCH(1,3)</th>
<th>The distribution used in the fitted model</th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td></td>
<td>5.7615</td>
<td>5.7552</td>
<td>6.2716</td>
<td>6.2712</td>
<td>5.7208</td>
<td>5.7147</td>
</tr>
<tr>
<td>SN: MSE</td>
<td></td>
<td>1.6924</td>
<td>1.9654</td>
<td>2.5887</td>
<td><strong>0.7488</strong></td>
<td>1.6929</td>
<td>1.9288</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td>1.2478</td>
<td>1.3656</td>
<td>1.5640</td>
<td><strong>0.8009</strong></td>
<td>1.2480</td>
<td>1.3523</td>
</tr>
<tr>
<td>STD: MSE</td>
<td></td>
<td><strong>0.0637</strong></td>
<td>0.0731</td>
<td>0.0745</td>
<td>4.5324</td>
<td>6.7831</td>
<td>6.7379</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td><strong>0.1358</strong></td>
<td>0.2120</td>
<td>0.2113</td>
<td>2.1119</td>
<td>2.5920</td>
<td>2.5832</td>
</tr>
<tr>
<td>SSTD: MSE</td>
<td></td>
<td>0.1841</td>
<td>22.0342</td>
<td>1.3260</td>
<td>2.9168</td>
<td><strong>0.1764</strong></td>
<td>0.1765</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td>0.28982</td>
<td>4.68293</td>
<td>1.09884</td>
<td>1.67549</td>
<td><strong>0.28451</strong></td>
<td>0.28541</td>
</tr>
<tr>
<td>GED: MSE</td>
<td></td>
<td>50.3319</td>
<td>50.4062</td>
<td><strong>6.8372</strong></td>
<td>47.5525</td>
<td>50.1022</td>
<td>50.1885</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td>7.0831</td>
<td>7.0883</td>
<td><strong>2.5859</strong></td>
<td>6.8843</td>
<td>7.0669</td>
<td>7.0730</td>
</tr>
<tr>
<td>SGED: MSE</td>
<td></td>
<td><strong>7.0684</strong></td>
<td>12.1011</td>
<td>7.6976</td>
<td>49.9598</td>
<td>7.1367</td>
<td>12.1642</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td><strong>2.6391</strong></td>
<td>3.4664</td>
<td>2.7568</td>
<td>7.0630</td>
<td>2.6521</td>
<td>3.4755</td>
</tr>
</tbody>
</table>

Notes: MSE figures are $\times 10^{-10}$ and MAE figures are $\times 10^{-5}$. Bold values in each row represent the minimum value of AIC.

Though Tables 3.2 and 3.3 indicate that the true model is always the best fitting model, Tables 3.4 and 3.5 show that the true model does not necessarily provide the minimum values of MSE and MAE and it might not have produced the best performance of forecasting volatility. For these particular samples, the simulation studies show that the statement “the best fitted model does not necessarily provide the best forecast on volatility” is true and also holds for the case
of higher order GARCH models.

Table 3.5: The values of MSE and MAE for simulated data from the GARCH(2,1) model

<table>
<thead>
<tr>
<th>Error in true GARCH(2,1)</th>
<th>The distribution used in the fitted model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>N: MSE</td>
<td>3.3119</td>
</tr>
<tr>
<td>MAE</td>
<td>1.7607</td>
</tr>
<tr>
<td>MAE</td>
<td>7.0373</td>
</tr>
<tr>
<td>MAE</td>
<td>0.9342</td>
</tr>
<tr>
<td>SSTD: MSE</td>
<td>4.3486</td>
</tr>
<tr>
<td>MAE</td>
<td>2.0523</td>
</tr>
<tr>
<td>GED: MSE</td>
<td>3.7552</td>
</tr>
<tr>
<td>MAE</td>
<td>1.8497</td>
</tr>
<tr>
<td>MAE</td>
<td>5.3006</td>
</tr>
</tbody>
</table>

Notes: MSE figures are \( \times 10^{-10} \) and MAE figures are \( \times 10^{-5} \). Bold values in each row represent the minimum value of AIC.

Shamiri and Isa (2009) argued that there are several plausible models that can be selected to use for forecasting and that one should not be fooled into thinking that the model with the best fit is the one that will forecast the best. However, it is of interest to examine how much difference exists between the best performance model and the forecast given by the best fitting model.

To investigate this question, 100 independently simulated samples from equation (3.56) are carried out, given by each of the six different type of error distribution of \( \varepsilon_t \). Each sample has size 6,536. The first 3,535 observations are considered as in-sample data and the latter observations are considered as out-sample data. For each set of simulated data, the data are fitted by the GARCH(1,3) model with the six different types of error distributions, then the performance of volatility forecasting is evaluated by calculating MSE and MAE values. The paired t-test
is carried out on the following hypothesis:

\[ H_0 : \mu_a - \mu_b = 0 \]

\[ H_1 : \mu_a - \mu_b > 0 \]

where \( \mu_a \) denotes the mean of MSE (MAE) given by the best fitted model and \( \mu_b \) denotes the mean of MSE (MAE) given by the best performance model.

The main objective of examining this paired t-test is to check whether the mean of MSE and MAE of the best fitting model is significantly much different from the best performance model for volatility forecasting. If the null hypothesis cannot be rejected, it means that the best performance model does not statistically provide better volatility forecasting than the best fitting model in terms of MSE and MAE measurements. The P-values of the positive one tail paired t-test for 1- and 10- steps ahead forecasts are shown in Table 3.6

Table 3.6: The P-values of paired t-test results between the best fitting model and the best performance model given by samples from the GARCH(1,3) model

<table>
<thead>
<tr>
<th>Error in the best fitting model</th>
<th>1-step ahead</th>
<th>10-step ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
<td>MAE</td>
</tr>
<tr>
<td>N</td>
<td>0.3254</td>
<td>0.2517</td>
</tr>
<tr>
<td>SN</td>
<td>0.5871</td>
<td>0.9945</td>
</tr>
<tr>
<td>STD</td>
<td>0.8124</td>
<td>0.2641</td>
</tr>
<tr>
<td>SSTD</td>
<td>0.1934</td>
<td>0.3454</td>
</tr>
<tr>
<td>GED</td>
<td>0.0842</td>
<td>0.1104</td>
</tr>
<tr>
<td>SGED</td>
<td>0.3891</td>
<td>0.4512</td>
</tr>
</tbody>
</table>

Based on the outcomes from Tables 3.2 to 3.3, the true model is always the best fitting model, but the best fitting GARCH model does not provide the best performance for volatility forecasting. Table 3.6 shows that all the paired t-tests are not significant. It indicates that there is not much difference between the best fitting model and the best performance model for volatility forecasting in terms of mean of MSE and MAE. Therefore, the best fitting model is still able to be
utilized for reasonable forecasting of volatility.

3.7 Conclusion

This chapter presented how to identify the order of the GARCH(p,q) model under non-normality error assumptions.

To study the volatility forecasting performance between the best fitting model and the best performance model, simulation studies are carried out. Simulation studies are demonstrated relevant to the comparison of the volatility forecasting performance between the best fitting models and the best performance model. The paired t-test is also used to examine how much difference there is between the means of MSE and MAE of volatility forecasting of the best fitting models compared with those of the best performance model.

The simulated results show that the values of MSE and MAE given by the best fitted model are not significantly different from those given by the best forecast performance model. The best fitting model is not necessary to provide the best forecast of volatility. However, the results from the paired t-test revealed that there is not much significant difference of the means of MSE and MAE between the best fitting model and the best performance model. Therefore, it is still reliable to use the best fitted model for volatility forecasting in practice.

In the next chapter, some empirical studies are carried out to investigate the performance of volatility forecasting using GARCH(p,q) models with different error distributions for the real life data from the daily closing price of emerging stock indices.
Chapter 4

Evaluating the Volatility Forecasting Performance of the Best Fitting GARCH Models in Emerging Asian Stock Markets

4.1 Introduction

As mentioned in Chapters 2 and 3, a number of researchers have investigated the performance of symmetric and asymmetric GARCH models with alternative non-normal error distributions in mature stock markets. Since emerging markets have grown rapidly and become more attractive to both individual and institutional investors, it is of interest to examine the volatility forecasting in such markets, especially in South East Asia. However, these previous studies mainly examined the performance of volatility forecasting using the GARCH(1,1) model for the mature markets. In addition, these investigations did not examine the performance of GARCH(p,q) models when \( p \) and/or \( q \) were greater than 1.

This chapter investigates whether using higher orders of symmetric GARCH models might be appropriate for some indices from emerging stock markets of South East Asia. By using simulation studies, this chapter aims to examine and compare the volatility forecasting performance among competing GARCH(p,q) models with six different types of error distributions as introduced in Chapter 3.
The six error distributions are the Normal (N), Skewed Normal (SN), Student-t Distribution (STD), Skewed Student-t Distribution (SSTD), Generalized Error Distribution (GED) and Skewed Generalized Error Distribution (SGED). Another important objective in this chapter is to investigate whether the best fitting model, in terms of the Akaike Information Criterion (AIC), also provides the best performance of volatility forecasting of the underlying financial time series measured by the Mean Squared Error (MSE) and the Mean Absolute Error (MAE) criteria. Empirical simulation studies will be carried out for these purposes.

The daily closing price indices of three emerging stock markets in South East Asia, the Stock Exchange of Thailand Index (SET), the Kuala Lumpur Composite Index (KLCI) from Malaysia and the Straits Time Index (STI) from Singapore will be employed for empirical studies. The daily closing data employed for SET is obtained from the Stock Exchange of Thailand and the daily closing data were retrieved from the database of the Yahoo! Finance website (http://finance.yahoo.com/) for both the KLCI and STI.

4.2 Data and Methodology

The data employed in this thesis comprise of 6,536 daily closing prices from the SET covering the period 4/01/1982 to 11/08/2008; 3,880 daily closing prices from the KLCI covering the period 3/12/1993 to 21/08/2009 and 5,407 daily closing prices from the STI covering the period 28/12/1987 to 21/08/2009.

The daily log returns from these three daily closing prices are calculated as difference of the logarithms from Equation (2.1). Figures 4.1 to 4.3 show the daily closing price movements and daily returns of the SET, KLCI and STI respectively across the sample periods. All daily returns of these three stock markets appear stationary and show evidence of volatility clustering phenomena where large or small price changes tend to be followed by other large or small price changes.
of either sign (positive or negative). The movements show that the volatility of returns changes over time and the GARCH models may be appropriate for explaining these three data series.

![Graph showing historical daily closing price movement and daily returns for the SET; 6,536 in total observations (excluding public holidays and weekends)](image)

**Figure 4.1:** Historical daily closing price movement and daily returns for the SET; 6,536 in total observations (excluding public holidays and weekends)

### 4.2.1 Descriptive Statistics of Three Emerging Indices

A preliminary analysis of daily return series \( \{r_t\} \) for the SET, KLCI and STI are displayed in Table 4.1. It shows that the mean of returns for the SET is slightly larger than the mean of returns for the KLCI and STI. Both of the return series for the SET and KLCI display negative skewness while the STI displays positive skewness. All returns display evidence of excess kurtosis. The Jarque-Bera test is the test for normality (see Cromwell et al., 1994). The null hypothesis of normality for this test at a 1% level is significantly rejected for all three indices. All return series have non-normal distributions with high kurtosis and skewness.
CHAPTER 4. EVALUATING THE VOLATILITY FORECASTING PERFORMANCE OF THE BEST FITTING GARCH MODELS IN EMERGING ASIAN STOCK MARKETS

4.2.2 In-Sample Parameter Estimation and Model Diagnostics

This section identifies whether the best fitting model is appropriate to be used for volatility forecasting based on the real observations of the three emerging indices. The best fitting GARCH(p,q) model of each stock index was identified from among the competing GARCH models with their six different types of error distributions based on in-sample data for parameter estimations and out-sample data for volatility forecasting.

Each data set of the three stock indices is divided into two parts. The first part is called the in-sample data set which is used to build up a model for underlying data. The second part is called the out-sample data set, used to investigate the performance of volatility forecasting. The in-sample period for the SET starts from 4/10/1982 to 14/05/1996 with 3,535 daily observations; for the KLCI starts
CHAPTER 4. EVALUATING THE VOLATILITY FORECASTING PERFORMANCE OF THE BEST FITTING GARCH MODELS IN EMERGING ASIAN STOCK MARKETS

Figure 4.3: Historical daily closing price movement and daily returns for the STI; 5,407 in total observations (excluding public holidays and weekends) from 3/12/1993 to 23/12/1999 with 1,500 daily observations and it starts from 28/12/1987 to 19/01/1998 with 2,500 daily observations for the STI.

The order determination process for the GARCH(p,q) model under the non-normality assumption suggested in Section 3.3 and is applied to the data sets of the three emerging stock indices. There are six possible orders of competing GARCH(p,q) models which are pre-set. The competing GARCH(p,q) models are fitted for each data set by assuming the error terms in GARCH model is normally distributed. The six possible orders of competing GARCH models consist of GARCH(1,1), GARCH(1,2), GARCH(1,3), GARCH(2,1), GARCH(2,2) and GARCH(3,1) models. To select an appropriate order among competing GARCH(p,q) models for each stock market, the AIC criterion is used by considering the smallest AIC value. It is found that GARCH(1,3) model is the most appropriate order for the SET, GARCH(1,1) and GARCH(2,1) models for the
CHAPTER 4. EVALUATING THE VOLATILITY FORECASTING PERFORMANCE OF THE BEST FITTING GARCH MODELS IN EMERGING ASIAN STOCK MARKETS

Table 4.1: Descriptive statistics and Jarque-Bera test statistic for normality of daily returns for the SET, KLCI and STI

<table>
<thead>
<tr>
<th></th>
<th>Sample Size (×10^{-3})</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Excess Kurtosis</th>
<th>Jarque-Bera Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>SET</td>
<td>6.536</td>
<td>0.289</td>
<td>0.01562</td>
<td>-0.06808</td>
<td>7.9577</td>
<td>17220**</td>
</tr>
<tr>
<td>KLCI</td>
<td>3.880</td>
<td>-0.030</td>
<td>0.01554</td>
<td>-0.43354</td>
<td>40.7329</td>
<td>267574**</td>
</tr>
<tr>
<td>STI</td>
<td>5.407</td>
<td>-0.200</td>
<td>0.01331</td>
<td>0.11628</td>
<td>8.3077</td>
<td>15526**</td>
</tr>
</tbody>
</table>

Note: (*),(**) and (***) denote significance at 5%, 1% and 0.1 % level, respectively.

KLCI and STI respectively.

To find out the best fitting GARCH(p,q) model with suitable error distribution for each stock index, the six types of error distributions are applied to the appropriate order in GARCH(p,q) models based on AIC criterion. The AIC values for three stock indices with six different error distributions are reported in Table 4.2.

Table 4.2: The AIC values given by models with different error distributions

<table>
<thead>
<tr>
<th>AIC</th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>SET:GARCH(1,3)</td>
<td>-6.4703</td>
<td>-6.4749</td>
<td>-6.5421</td>
<td><strong>-6.5432</strong></td>
<td>-6.5376</td>
<td>-6.5409</td>
</tr>
<tr>
<td>KLCI:GARCH(1,1)</td>
<td>-6.9142</td>
<td>-6.9137</td>
<td>-6.9583</td>
<td>-6.9575</td>
<td><strong>-6.9608</strong></td>
<td>-6.9596</td>
</tr>
<tr>
<td>STI:GARCH(2,1)</td>
<td>-6.0739</td>
<td>-6.0735</td>
<td><strong>-6.1104</strong></td>
<td>-6.1100</td>
<td>-6.1061</td>
<td>-6.1058</td>
</tr>
</tbody>
</table>

According to Table 4.2, the GARCH(1,3) model with skewed Student-t error distribution is the best fitting model for the SET, the GARCH(1,1) model with generalized error distribution for the KLCI and the GARCH(2,1) model with Student-t for the STI.

After the best fitting GARCH(p,q) model with an appropriate error distribution for the SET, KLCI and STI are determined, model estimates and diagnostic tests for each stock index during the in-sample period are carried out. Ljung-Box $Q^2$ statistics are applied to test for serial correlation of the squared standardized
residuals. The ARCH effect is also tested by using Lagrange Multiplier (LM) test.

Tables 4.3 - 4.5 show parameter estimations of GARCH models and diagnostics for the SET, KLCI and STI. $Q^2(15)$ is the Ljung-Box $Q^2$ statistics for serial correlation in the squared standardized residuals with 15 lags (see Ljung and Box, 1978). All results of the Ljung-Box test are not significant, indicating that there is no autocorrelation in the squared returns. The ARCH(1)-LM test is the Lagrange Multiplier test for a first order linear ARCH effect (see Engle, 1982). The LM test is used to examine whether the presence of ARCH effects are removed from models. There should be no ARCH effects left in the standardized residuals if the conditional variance model is correctly specified. In the case of the results of the LM tests, all tests indicate that the null hypotheses are accepted, which means that the ARCH effects are removed. Therefore, the best three fitting GARCH models for each stock market are adequate and appropriate to use for volatility forecasting.

The best fitting models for the three stock indices can be expressed by the following equations:

For the SET, the best fitting GARCH(1,3) model with skewed Student-t distribution is:

$$r_t = 0.0001193 + \varepsilon_t$$

$$\varepsilon_t = \eta_t \sqrt{h_t}$$

$$h_t = 0.0000007 + 0.2728\varepsilon_{t-1}^2 + 0.5839h_{t-1} + 0.0001h_{t-2} + 0.1897h_{t-3}.$$  

where $\{\varepsilon_t\}$ has a skewed Student-t distribution.

For the KLCI, the best fitting GARCH(1,1) model with generalized error
Table 4.3: Estimated parameters and diagnostics of the GARCH(1,3) model with different error distributions for the SET

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th>P-value</th>
<th>SN</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$1.637 \times 10^{-4}$</td>
<td>(0.1308)</td>
<td>$6.189 \times 10^{-5}$</td>
<td>(0.5750)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>$8.584 \times 10^{-7}$</td>
<td>(2.83$x\times 10^{-6}$)***</td>
<td>$8.584 \times 10^{-7}$</td>
<td>(3.34$x\times 10^{-6}$)***</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>$2.545 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
<td>$2.575 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$5.312 \times 10^{-1}$</td>
<td>(2.73$x\times 10^{-5}$)***</td>
<td>$5.278 \times 10^{-1}$</td>
<td>(1.01$x\times 10^{-5}$)***</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>$1.000 \times 10^{-8}$</td>
<td>(1.0000)</td>
<td>$1.000 \times 10^{-8}$</td>
<td>(1.0000)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>$2.434 \times 10^{-1}$</td>
<td>(0.0028)**</td>
<td>$1.000 \times 10^{-6}$</td>
<td>(1.0000)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>$9.210 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LM test</td>
<td>13.1934</td>
<td>(0.3551)</td>
<td>13.6131</td>
<td>(0.3260)</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>15.8785</td>
<td>(0.3901)</td>
<td>16.2687</td>
<td>(0.3644)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th>P-value</th>
<th>SSTD</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$2.372 \times 10^{-4}$</td>
<td>(0.0167)*</td>
<td>$1.193 \times 10^{-4}$</td>
<td>(0.2798)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>$7.247 \times 10^{-7}$</td>
<td>(0.0012)**</td>
<td>$7.128 \times 10^{-7}$</td>
<td>(0.0014)**</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>$2.652 \times 10^{-1}$</td>
<td>(8.88$x\times 10^{-16}$)***</td>
<td>$2.728 \times 10^{-1}$</td>
<td>(1.11$x\times 10^{-15}$)***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$5.874 \times 10^{-1}$</td>
<td>(0.0008)***</td>
<td>$5.839 \times 10^{-1}$</td>
<td>(0.0007)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>$1.000 \times 10^{-8}$</td>
<td>(1.0000)</td>
<td>$1.000 \times 10^{-3}$</td>
<td>(0.0752)**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>$1.000 \times 10^{-6}$</td>
<td>(1.0000)</td>
<td>$1.897 \times 10^{-1}$</td>
<td>(0.00812)***</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>$9.469 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>-</td>
<td>-</td>
<td>$5.0330$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>LM test</td>
<td>14.0056</td>
<td>(0.3003)</td>
<td>14.6032</td>
<td>(0.2638)</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>16.6525</td>
<td>(0.3400)</td>
<td>17.1584</td>
<td>(0.3094)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th>P-value</th>
<th>SGED</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$2.295 \times 10^{-4}$</td>
<td>(0.0251)*</td>
<td>$5.948 \times 10^{-5}$</td>
<td>(0.5700)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>$7.710 \times 10^{-7}$</td>
<td>(0.0004)***</td>
<td>$7.392 \times 10^{-7}$</td>
<td>(0.0006)***</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>$2.531 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
<td>$2.632 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$5.738 \times 10^{-1}$</td>
<td>(0.0011)</td>
<td>$5.687 \times 10^{-1}$</td>
<td>(0.0005)***</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>$1.000 \times 10^{-8}$</td>
<td>(1.0000)</td>
<td>$1.000 \times 10^{-8}$</td>
<td>(1.0000)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>$2.031 \times 10^{-1}$</td>
<td>(0.0731)*</td>
<td>$2.042 \times 10^{-1}$</td>
<td>(0.0509)*</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>$9.277 \times 10^{-1}$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$1.2600$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
<td>$1.2510$</td>
<td>(&lt; 2$x\times 10^{-16}$)***</td>
</tr>
<tr>
<td>LM test</td>
<td>13.4731</td>
<td>(0.0056)</td>
<td>14.3506</td>
<td>(0.2788)</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>16.1362</td>
<td>(0.3730)</td>
<td>16.9054</td>
<td>(0.3215)</td>
</tr>
</tbody>
</table>

Note: (*) ,(**) and (*** ) denote significance at 5%, 1% and 0.1% level, respectively. Highlighted cells in the table indicate the GARCH(1,3) model with the most appropriate error distribution.
Table 4.4: Estimated parameters and diagnostics of the GARCH(1,1) model with different error distributions for the KLCI

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th>( N )</th>
<th>P-value</th>
<th>( S )</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>-5.160 \times 10^{-4}</td>
<td>(0.0018)**</td>
<td>-5.317 \times 10^{-4}</td>
<td>(0.0013)**</td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>1.042 \times 10^{-6}</td>
<td>(0.0042)**</td>
<td>1.064 \times 10^{-6}</td>
<td>(0.0041)**</td>
<td></td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>1.353 \times 10^{-1}</td>
<td>(5.31 \times 10^{-14})***</td>
<td>1.378 \times 10^{-1}</td>
<td>(5.75 \times 10^{-14})***</td>
<td></td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>8.604 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td>8.579 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-</td>
<td>-</td>
<td>9.681 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>( \nu )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>LM test</td>
<td>13.2454</td>
<td>(0.3514)</td>
<td>13.1267</td>
<td>(0.3598)</td>
<td></td>
</tr>
<tr>
<td>( Q^2(15) )</td>
<td>13.4147</td>
<td>(0.5702)</td>
<td>13.3063</td>
<td>(0.5786)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th>STD</th>
<th>P-value</th>
<th>SSTD</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>-4.410 \times 10^{-4}</td>
<td>(0.0055)**</td>
<td>-4.967 \times 10^{-4}</td>
<td>(0.0035)**</td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>7.704 \times 10^{-7}</td>
<td>(0.0294)*</td>
<td>7.909 \times 10^{-7}</td>
<td>(0.0279)*</td>
<td></td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>9.995 \times 10^{-2}</td>
<td>(2.27 \times 10^{-6})***</td>
<td>1.014 \times 10^{-1}</td>
<td>(1.97 \times 10^{-6})***</td>
<td></td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>8.958 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td>8.941 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-</td>
<td>-</td>
<td>9.688 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>( \nu )</td>
<td>5.8050</td>
<td>(&lt; 2 \times 10^{-10})***</td>
<td>5.8690</td>
<td>(&lt; 3.55 \times 10^{-10})***</td>
<td></td>
</tr>
<tr>
<td>LM test</td>
<td>17.8703</td>
<td>(0.1196)</td>
<td>17.6690</td>
<td>(0.1261)</td>
<td></td>
</tr>
<tr>
<td>( Q^2(15) )</td>
<td>18.0208</td>
<td>(0.2615)</td>
<td>17.8106</td>
<td>(0.2727)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th>GED</th>
<th>P-value</th>
<th>SGED</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>-5.005 \times 10^{-4}</td>
<td>(0.0004)**</td>
<td>-5.343 \times 10^{-4}</td>
<td>(0.0014)**</td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>8.607 \times 10^{-7}</td>
<td>(0.0259)*</td>
<td>8.725 \times 10^{-7}</td>
<td>(0.0255)*</td>
<td></td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>1.136 \times 10^{-1}</td>
<td>(1.11 \times 10^{-7})***</td>
<td>1.146 \times 10^{-1}</td>
<td>(1.10 \times 10^{-7})***</td>
<td></td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>8.809 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td>8.798 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-</td>
<td>-</td>
<td>9.854 \times 10^{-1}</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>( \nu )</td>
<td>1.3110</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td>1.3150</td>
<td>(&lt; 2 \times 10^{-16})***</td>
<td></td>
</tr>
<tr>
<td>LM test</td>
<td>15.2401</td>
<td>(0.2286)</td>
<td>15.1528</td>
<td>(0.2331)</td>
<td></td>
</tr>
<tr>
<td>( Q^2(15) )</td>
<td>15.3440</td>
<td>(0.4269)</td>
<td>15.2608</td>
<td>(0.4327)</td>
<td></td>
</tr>
</tbody>
</table>

Note: (*),(**) and (*** denote significance at 5%, 1% and 0.1 % level respectively. Highlighted cells in the table indicate the GARCH(1,1) model with the most appropriate error distribution.
Table 4.5: Estimated parameters and diagnostics of the GARCH(2,1) model with different error distributions for the STI

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>P-value</td>
<td>SN</td>
<td>P-value</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-5.972 × 10^{-4}</td>
<td>(0.0016)**</td>
<td>-6.060 × 10^{-4}</td>
<td>(0.0013)**</td>
</tr>
<tr>
<td>$\omega$</td>
<td>2.166 × 10^{-6}</td>
<td>(0.0003)**</td>
<td>2.164 × 10^{-6}</td>
<td>(0.0003)**</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>6.291 × 10^{-2}</td>
<td>(0.0050)**</td>
<td>6.311 × 10^{-2}</td>
<td>(0.0051)**</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>6.230 × 10^{-2}</td>
<td>(0.0222)*</td>
<td>6.414 × 10^{-2}</td>
<td>(0.0192)*</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>8.695 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
<td>8.679 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>9.763 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>$\nu$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LM test</td>
<td>8.4034</td>
<td>(0.7528)</td>
<td>8.5674</td>
<td>(0.7393)</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>9.4684</td>
<td>(0.8517)</td>
<td>9.6154</td>
<td>(0.8432)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STD</td>
<td>P-value</td>
<td>SSTD</td>
<td>P-value</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-5.850 × 10^{-4}</td>
<td>(0.0015)**</td>
<td>-6.311 × 10^{-4}</td>
<td>(0.0008)**</td>
</tr>
<tr>
<td>$\omega$</td>
<td>1.962 × 10^{-6}</td>
<td>(0.0017)**</td>
<td>1.965 × 10^{-6}</td>
<td>(0.0018)**</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>5.339 × 10^{-2}</td>
<td>(0.0464)*</td>
<td>5.326 × 10^{-2}</td>
<td>(0.0473)*</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>6.893 × 10^{-2}</td>
<td>(0.0262)*</td>
<td>7.076 × 10^{-2}</td>
<td>(0.0231)*</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>8.733 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
<td>8.720 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>9.722 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>$\nu$</td>
<td>7.6680</td>
<td>(2.27 × 10^{-12})*</td>
<td>7.6430</td>
<td>(2.14 × 10^{-12})*</td>
</tr>
<tr>
<td>LM test</td>
<td>8.4077</td>
<td>(0.7525)</td>
<td>8.5157</td>
<td>(0.7436)</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>9.4368</td>
<td>(0.8535)</td>
<td>9.5584</td>
<td>(0.8465)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated parameter</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GED</td>
<td>P-value</td>
<td>SGED</td>
<td>P-value</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-6.013 × 10^{-4}</td>
<td>(0.0009)**</td>
<td>-6.512 × 10^{-4}</td>
<td>(0.0005)**</td>
</tr>
<tr>
<td>$\omega$</td>
<td>2.065 × 10^{-6}</td>
<td>(0.0018)**</td>
<td>2.053 × 10^{-6}</td>
<td>(0.0019)**</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>5.516 × 10^{-2}</td>
<td>(0.0359)*</td>
<td>5.527 × 10^{-2}</td>
<td>(0.0365)*</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>6.847 × 10^{-2}</td>
<td>(0.0286)*</td>
<td>7.048 × 10^{-2}</td>
<td>(0.0251)*</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>8.711 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
<td>8.696 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-</td>
<td>-</td>
<td>9.721 × 10^{-1}</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.4290</td>
<td>(2 × 10^{-16})*</td>
<td>1.4270</td>
<td>(2 × 10^{-16})*</td>
</tr>
<tr>
<td>LM test</td>
<td>8.4263</td>
<td>(0.7509)</td>
<td>8.5772</td>
<td>(0.7385)</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>8.1084</td>
<td>(0.8570)</td>
<td>9.6469</td>
<td>(0.8413)</td>
</tr>
</tbody>
</table>

Note: (*), (**) and (***) denote significance at 5%, 1% and 0.1% level respectively. Highlighted cells in the table indicate the GARCH(2,1) model with the most appropriate error distribution.
distribution is:

\[ r_t = 0.000505 + \varepsilon_t \]
\[ \varepsilon_t = \eta_t \sqrt{h_t} \]
\[ h_t = 0.0000008 + 0.1136\varepsilon_{t-1}^2 + 0.8809h_{t-1}. \]

where \( \{\varepsilon_t\} \) has a generalized error distribution.

For the STI, the best fitting GARCH(2,1) model with student-t distribution is:

\[ r_t = 0.000585 + \varepsilon_t \]
\[ \varepsilon_t = \eta_t \sqrt{h_t} \]
\[ h_t = 0.0000019 + 0.0534\varepsilon_{t-1}^2 + 0.0689\varepsilon_{t-2}^2 + 0.8733h_{t-1}. \]

where \( \{\varepsilon_t\} \) has a Student-t distribution.

In the next section, the GARCH(1,3) model with skewed Student-t distribution for the SET, the GARCH(1,1) model with generalized error distribution for the KLCI and the GARCH(2,1) model with Student-t distribution for the STI are adopted as the best fitting models for predicting the volatility of these three stock markets.

### 4.3 Empirical Results for the Performance of Volatility Forecasting

The out-of-sample predictive ability of the best fitting GARCH models for each stock market is employed to evaluate the performance of volatility forecasting. To compare this performance, the error distribution of each best fitting GARCH model is changed and replaced with the other five error distributions. For instance, the GARCH(1,3) model with skewed Student-t distribution (SSTD) is the best fitting model for the SET. The out-of-sample volatility forecasting for
the SET is carried out by using the same GARCH(1,3) model but replacing the SSTD with the other five error distributions (N, SN, STD, GED and SGED).

After obtaining the best fitting model by comparing the other five error distributions for each stock index, the best fitting GARCH model of each stock index is used to predict the performance of volatility. The four horizons step ahead (1, 2, 10 and 15) of forecasting are carried out. Mean square error (MSE) and mean absolute error (MAE) in Equations (3.54) and (3.55) respectively are used to evaluate the performance of volatility forecasting.

The empirical results for performance of volatility forecasting, in terms of MSE and MAE measurements are reported in Table 4.6 and Table 4.7. All results from these tables show that the best fitting models in each stock market do not necessarily provide the best performance of volatility forecasting in terms of the values of MSE and MAE.

For the SET, the best fitting model is GARCH(1,3) with skewed Student-t error distribution while the minimum values of MSE and MAE are produced by a skewed normal error distribution. Similarly, the best fitting model of the KLCI is GARCH(1,1) with a generalized error distribution and GARCH(2,1) with Student-t error distribution for the STI. But the results show that the best performance of volatility forecasting for the KLCI is shifted from the generalized error distribution to the skewed generalized error distribution and for the STI it is shifted from the Student-t error distribution to the skewed normal error distribution (See Figures 4.4 - 4.9).
Table 4.6: Out-of-sample volatility forecasting evaluated by MSE

<table>
<thead>
<tr>
<th>Step-Ahead</th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>SET:GARCH(1,3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.561</td>
<td>2.551</td>
<td>2.560</td>
<td>2.553</td>
<td>2.557</td>
<td>2.543</td>
</tr>
<tr>
<td>2</td>
<td>2.534</td>
<td>2.524</td>
<td>2.546</td>
<td>2.539</td>
<td>2.539</td>
<td>2.525</td>
</tr>
<tr>
<td>10</td>
<td>2.995</td>
<td>2.992</td>
<td>3.139</td>
<td>3.011</td>
<td>3.033</td>
<td>3.058</td>
</tr>
<tr>
<td>15</td>
<td>2.985</td>
<td>2.982</td>
<td>3.207</td>
<td>3.004</td>
<td>3.028</td>
<td>3.084</td>
</tr>
<tr>
<td>KLCI:GARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.050</td>
<td>1.051</td>
<td>1.032</td>
<td>1.032</td>
<td>1.035</td>
<td>1.027</td>
</tr>
<tr>
<td>10</td>
<td>3.614</td>
<td>3.771</td>
<td>2.126</td>
<td>1.942</td>
<td>1.561</td>
<td>1.544</td>
</tr>
<tr>
<td>15</td>
<td>5.586</td>
<td>5.001</td>
<td>5.639</td>
<td>5.153</td>
<td>4.122</td>
<td>4.096</td>
</tr>
<tr>
<td>STI:GARCH(2,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.929</td>
<td>1.922</td>
<td>1.926</td>
<td>1.934</td>
<td>1.933</td>
<td>1.929</td>
</tr>
<tr>
<td>2</td>
<td>1.927</td>
<td>1.918</td>
<td>1.933</td>
<td>1.929</td>
<td>1.928</td>
<td>1.924</td>
</tr>
<tr>
<td>10</td>
<td>2.694</td>
<td>2.688</td>
<td>2.710</td>
<td>2.708</td>
<td>2.703</td>
<td>2.701</td>
</tr>
<tr>
<td>15</td>
<td>2.629</td>
<td>2.624</td>
<td>2.644</td>
<td>2.643</td>
<td>2.637</td>
<td>2.636</td>
</tr>
</tbody>
</table>

Notes: The reported value is multiplied by \( \times 10^{-7} \). The minimum value of MSE in the same row is in bold type.

As can be seen, the empirical results for the evaluation of the volatility forecasting between the best fitting model and the best performance model appeared to be identical to the simulated results. These results clearly confirm that the best fitting model was not necessarily a model which could provide the best performance on volatility forecasting based on MSE and MAE measurements.

However, to investigate the difference of MSE and MAE values between the best fitted model and the best performance model, the evaluations of the Percent Error (PE) of MSE and MAE for each underlying case are examined. The formula of PE is defined as follows:

\[
PE = \frac{A - B}{A} \times 100\% ,
\]  

(4.1)

where

\( A \) denotes MSE (MAE) given by the best fitted model,
Table 4.7: Out-of-sample volatility forecasting evaluated by MAE

<table>
<thead>
<tr>
<th>Step-Ahead</th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>SET:GARCH(1,3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.244</td>
<td>2.241</td>
<td>2.325</td>
<td>2.453</td>
<td>2.281</td>
<td>2.287</td>
</tr>
<tr>
<td>2</td>
<td>2.312</td>
<td>2.311</td>
<td>2.419</td>
<td>2.334</td>
<td>2.361</td>
<td>2.372</td>
</tr>
<tr>
<td>10</td>
<td>2.614</td>
<td>2.601</td>
<td>2.857</td>
<td>2.622</td>
<td>2.688</td>
<td>2.745</td>
</tr>
<tr>
<td>15</td>
<td>2.771</td>
<td>2.763</td>
<td>3.018</td>
<td>2.792</td>
<td>2.831</td>
<td>2.903</td>
</tr>
<tr>
<td>KLCI:GARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.528</td>
<td>2.578</td>
<td>1.995</td>
<td>1.921</td>
<td>1.764</td>
<td><strong>1.747</strong></td>
</tr>
<tr>
<td>2</td>
<td>2.440</td>
<td>2.448</td>
<td>2.311</td>
<td>2.323</td>
<td>2.325</td>
<td><strong>2.305</strong></td>
</tr>
<tr>
<td>10</td>
<td>1.898</td>
<td>1.939</td>
<td>1.454</td>
<td>1.390</td>
<td>1.252</td>
<td><strong>1.238</strong></td>
</tr>
<tr>
<td>15</td>
<td>3.093</td>
<td>3.160</td>
<td>2.371</td>
<td>2.266</td>
<td>2.044</td>
<td><strong>2.020</strong></td>
</tr>
<tr>
<td>STI:GARCH(2,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.207</td>
<td>2.203</td>
<td>2.218</td>
<td>2.217</td>
<td>2.214</td>
<td>2.213</td>
</tr>
<tr>
<td>2</td>
<td>2.185</td>
<td>2.182</td>
<td>2.199</td>
<td>2.198</td>
<td>2.195</td>
<td>2.193</td>
</tr>
<tr>
<td>10</td>
<td>2.538</td>
<td>2.535</td>
<td>2.552</td>
<td>2.549</td>
<td>2.545</td>
<td>2.543</td>
</tr>
<tr>
<td>15</td>
<td>2.509</td>
<td>2.507</td>
<td>2.521</td>
<td>2.518</td>
<td>2.515</td>
<td>2.513</td>
</tr>
</tbody>
</table>

Notes: The reported value is multiplied by \((\times 10^{-7})\). The minimum value of MAE in the same row is in bold type.

Figure 4.4: Volatility forecasting of the GARCH(1,3) model for the SET evaluated by MSE
Note: A, B, C and D denote 1-, 2-, 10- and 15-steps ahead volatility forecasting
Figure 4.5: Volatility forecasting of the GARCH(1,1) model for the KLCI evaluated by MSE
Note: A, B, C and D denote 1-, 2-, 10- and 15-steps ahead volatility forecasting

Figure 4.6: Volatility forecasting of the GARCH(2,1) model for the STI evaluated by MSE
Note: A, B, C and D denote 1-, 2-, 10- and 15-steps ahead volatility forecasting
CHAPTER 4. EVALUATING THE VOLATILITY FORECASTING PERFORMANCE OF THE BEST FITTING GARCH MODELS IN EMERGING ASIAN STOCK MARKETS

Figure 4.7: Volatility forecasting of the GARCH(1,3) model for the SET evaluated by MAE
Note: A, B, C and D denote 1-, 2-, 10- and 15-steps ahead volatility forecasting

Figure 4.8: Volatility forecasting of the GARCH(1,1) model for the KLCI evaluated by MAE
Note: A, B, C and D denote 1-, 2-, 10- and 15-steps ahead volatility forecasting
CHAPTER 4. EVALUATING THE VOLATILITY FORECASTING PERFORMANCE OF THE BEST FITTING GARCH MODELS IN EMERGING ASIAN STOCK MARKETS

Figure 4.9: Volatility forecasting of the GARCH(2,1) model for the STI evaluated by MAE
Note: A, B, C and D denote 1-, 2-, 10- and 15-steps ahead volatility forecasting

B denotes MSE (MAE) given by best performance model.

Table 4.8: The percent error of MSE and MAE given by the best fitted model and the best performance model

<table>
<thead>
<tr>
<th>Step Ahead</th>
<th>SET</th>
<th>KLCI</th>
<th>STI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Difference</td>
<td>PE(%)</td>
<td>Difference</td>
</tr>
<tr>
<td>1</td>
<td>0.002</td>
<td>0.078</td>
<td>0.003</td>
</tr>
<tr>
<td>2</td>
<td>0.015</td>
<td>0.591</td>
<td>0.008</td>
</tr>
<tr>
<td>10</td>
<td>0.019</td>
<td>0.631</td>
<td>0.017</td>
</tr>
<tr>
<td>15</td>
<td>0.022</td>
<td>0.732</td>
<td>0.026</td>
</tr>
</tbody>
</table>

The percent error of MAE

<table>
<thead>
<tr>
<th>Step Ahead</th>
<th>SET</th>
<th>KLCI</th>
<th>STI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Difference</td>
<td>PE(%)</td>
<td>Difference</td>
</tr>
<tr>
<td>1</td>
<td>0.012</td>
<td>0.489</td>
<td>0.017</td>
</tr>
<tr>
<td>2</td>
<td>0.023</td>
<td>0.985</td>
<td>0.002</td>
</tr>
<tr>
<td>10</td>
<td>0.021</td>
<td>0.801</td>
<td>0.0014</td>
</tr>
<tr>
<td>15</td>
<td>0.029</td>
<td>1.039</td>
<td>0.024</td>
</tr>
</tbody>
</table>

The PE values are presented in Table 4.8. It shows that the majority of PE values are small and less than 1.2%. It also indicates that the MSE and MAE given by the best fitting model are not much different from those given by the
best performance model. Therefore, one can conclude that the best fitting model is still able to be used for volatility forecasting in practical situations. Even though the best fitting model does not provide the minimum values of MAE and MSE, but there is not much difference from MAE and MSE values of the best performance model.

4.4 Conclusion

In this chapter, the order determination of a GARCH(p,q) model as discussed in Chapter 3 is applied to the three South East Asian emerging stock indices. GARCH(p,q) models with six different types of error distributions are considered and the best fitting model used to forecast the future volatility. The performances of volatility forecasting are evaluated by MSE and MAE.

For the SET, the best fitting model is GARCH(1,3) with SSTD but the best performance of volatility forecasting shifts from the SSTD to SN distribution. The best fitting model for the KLCI is GARCH(1,1) with GED and GARCH(2,1) with STD for the STI but the best performance of volatility forecasting for the KLCI and STI shift to the SGED and SN distribution, respectively.

The empirical results have demonstrated that the best fitting model and the best performance model for volatility forecasting appeared to be identical to the simulated results. They show that the best fitting models in each stock index does not necessarily provide the best performance of volatility forecasting. However the differences of MSE and MAE values between the best fitting model and the best performance model are not large when evaluating the Percent Error(PE). Therefore, the best fitting model is still significantly acceptable to use for forecasting the future volatility in practical situations.

The results for the performance of volatility predictions show that a GARCH model with non-normal error distributions tend to provide better out-of-sample
forecast performance than a GARCH model with a normal error distribution as evaluated by MSE and MAE.
CHAPTER 4. EVALUATING THE VOLATILITY FORECASTING PERFORMANCE OF THE BEST FITTING GARCH MODELS IN EMERGING ASIAN STOCK MARKETS
Chapter 5

Cointegration Tests with Non-Normal GARCH Error Distributions

Time series models for estimating and forecasting volatility are very important in the field of econometrics. Volatility forecasting can be modeled by different models. It is of interest to find out better volatility models for predicting accurately future volatility of underlying time series. Numerous researchers have been attempting to develop models for volatility forecasting by investigating alternative volatility models with different error distributional assumptions.

In Chapter 4, an adequate GARCH\((p,q)\) models were investigated by considering higher order of GARCH model with different types of error distributions. To obtain the best fitting GARCH\((p,q)\) models, real life data from three emerging stock markets were applied and individually modeled for each stock market. The best fitting GARCH\((p,q)\) models were used to forecast the future volatility. It was concluded that a GARCH model with non-normal error distributions can improve the volatility forecasting performance of emerging stock markets.

In the second of this thesis, Chapter 5 only considers various tests for cointegration under non-normal GARCH error distributions. The effect of cointegration on forecasting stochastic volatility is actually treated in Chapter 6. We consider the following attempt, is it possible to improve the volatility forecasting perfor-
mance by incorporating the cointegration information among financial underlying time series into a GARCH model? As before, the six different types of error distributions in the error terms of GARCH model were considered. If there are cointegration relationships among underlying time series, this raises an interesting question whether the knowledge of cointegrated relationships of these underlying time series can benefit and improve the performance of volatility forecasting.

Firstly, the introduction of unit root and the cointegration concepts are briefly described in this chapter. The main objective of this chapter is to determine which cointegration method is the most powerful to detect the relationships of underlying time series by considering the size and power of cointegration tests. Also, the alternative six types of error distributions in the previous chapter are taken into account to examine whether the size and power of cointegration tests are still valid and useful for detecting cointegration relationships. Then the most powerful cointegration method is employed to develop the better model for volatility forecasting in Chapter 6.

5.1 Unit Root and Cointegration

Most statistical estimating and forecasting methods are based on the stationarity assumption. This implies that the mean and variance of the underlying time series are constant and independent of time. However, the mean and variance of these time series can change over time in practice. In other words, economic time series often occur to be non-stationary, or contain a unit root. Checking for unit root is the starting point of most empirical time series studies to examine whether underlying data series are non-stationary or there are no unit roots. A traditional approach for unit root test is to use differencing approaches. The degree of differencing is also importantly required in order to achieve stationarity. The concept of stationarity can be found in Maddala and Kim (1998).
Cointegration is now widely employed in econometrics and financial time series, introduced by Engle and Granger (1987). The concept of cointegration is used in order to analyze the equilibrium relationships among time series. If two or more time series are found to be cointegrated, there is a tendency for them to move very closely together in the long run movement, or they have a long-run equilibrium relationship among them. However, it is important to test whether these time series are stationary or not before using cointegration analysis. To examine the stationary property of a time series, the unit root test is employed.

There is some correspondence between the tests for unit root and cointegration. The tests for unit root are applied to univariate time series while the tests for cointegration are applied to more than two series, which have a unit root in their levels.

It is convenient to first give the definition of cointegration. Then, the different tests for cointegration which relevant to the test for unit root are presented in this chapter.

**An Introduction to Cointegration**

Consider the set of time series \( \{y_{i,t}\}, i = \{1,2,\ldots,k\} \) where the cointegration equation can be written as follows:

\[
\beta_1 y_{1,t} + \beta_2 y_{2,t} + \ldots + \beta_k y_{k,t} = \mu + u_t
\]  

(5.1)

where \( k \) is number of time series, \( \mu \) is long-run equilibrium and \( u_t \) is cointegrating error.

The cointegration equation can be written in matrix form as

\[
\beta' y_t = \mu + u_t
\]  

(5.2)
where $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ and $y_t = (y_{1,t}, y_{2,t}, \ldots, y_{k,t})'$.

The cointegrating error $u_t$ can be rewritten as the deviation from the long-run equilibrium

$$u_t = \beta'y_t - \mu$$

(5.3)

The components of vector $y_t$ are said to be cointegrated of order $(d, b)$, which is denoted as $y_t \sim CI(d, b)$, where $d \geq b > 0$ if:

1). All components $y_t$ are integrated of the same order $d$.

2). There exists a vector $\beta$ such that $\beta'y_t$ is integrated of order $(d-b)$, where $\beta$ is the cointegrating vector for $y_t$.

To illustrate the concept of cointegration, consider two time series $x_t \sim I(1)$ and $y_t \sim I(1)$. Regressing of $y_t$ on $x_t$ is given by:

$$y_t = \beta x_t + u_t$$

(5.4)

If the residual term $u_t$ is $I(0)$ then time series $x_t$ and time series $y_t$ are called cointegrated and $\beta$ is called cointegration coefficient of these two time series $x_t$ and $y_t$.

The cointegration tests used in this thesis can be classified into two groups. The first group of tests is residual-based tests. In this chapter, the residual-base test for the tests of cointegration consist of three tests: the Dickey-Fuller test (DF), Cointegrating Regression Durbin-Watson (CRDW) test and the Wild Bootstrap cointegration (WB) test. The second group of cointegration tests is based on the Error Correction Model in short ECM. The most popular cointegration test with ECM is the Johansen approach.

To examine which cointegration test is the most powerful for detecting cointegration of underlying time series, three tests based on residual-based tests (DF test, CRDW test and the WB test) are examined. The performance of three
residual-based tests for detecting cointegration are investigated in terms of the size and power of cointegration tests comparing with the Johansen tests.

The residual-based tests for cointegration is used to check whether the error in the cointegrating regression \( (u_t) \) is \( I(0) \). To illustrate the residual-based tests, consider the case of two time series \( x_t \) and \( y_t \) in (5.4). These cointegration tests were proposed in Engle and Granger (1987) which used Ordinary Least Squares (OLS) to obtain the cointegrating residual series \( (u_t) \). After obtaining the residuals \( u_t \), a hypothesis testing of \( u_t \) is carried out. The null hypothesis for the test of unit root is that \( u_t \) contain a unit root. The test of this hypothesis is related to the test for hypothesis that there is no cointegration between \( x_t \) and \( y_t \). While the alternative hypothesis is that \( u_t \) has no a unit root or there is a cointegration relationship between \( x_t \) and \( y_t \). If the null hypothesis is accepted, it implies that there is a unit root or there is no cointegration relationship between \( x_t \) and \( y_t \). The residual-based tests for cointegration can be extended to the case of more than two time series for hypothesis testing.

Another type of cointegration test is based on the ECM which is well-known as the “Granger representation theorem” by Engle and Granger (1987). Consider the case of two time series \( x_t \) and \( y_t \) in (5.4). If \( x_t \) and \( y_t \) are cointegrated, there must exist an error correction model (ECM) representation of the dynamic system of the joint behavior of \( x_t \) and \( y_t \) over time. The procedure for estimating the parameters in ECM is introduced by using the Johansen approach in Section 5.3.

### 5.2 The Residual-Based Tests for Cointegration

Three residual-based tests for cointegration are briefly introduced. They are DF test, CRDW test and WB test.

Under the assumption that the two time series \( x_t \) and \( y_t \) are individually \( I(1) \)
and they are cointegrated, the residual terms \((u_t)\) in Equation (5.4) is integrated of order zero and \(u_t\) is stationary. The OLS is applies to obtain the residual series \(u_t\) by regressing \(y_t\) on \(x_t\). The regression equation of \(y_t\) on \(x_t\) can be written as below:

\[
\hat{u}_t = y_t - \hat{\beta}x_t \tag{5.5}
\]

where \(\hat{\beta}\) is the parameter estimation and \(\hat{u}_t\) is the residual series from regressing \(y_t\) on \(x_t\) which is used for the test of unit root. In other words, \(\hat{u}_t\) can be used for cointegration relationship between \(y_t\) and \(x_t\) because the tests for unit root are strongly related to the test for cointegration.

Thus, \(\hat{u}_t\) is used to test whether \(u_t\) is \(I(0)\) by using three of residual-based tests mentioned above. The null hypothesis for the cointegration tests of \(x_t\) and \(y_t\) is defined as below:

\[
H_0 : x_t \text{ and } y_t \text{ are not cointegrated} \\
H_1 : x_t \text{ and } y_t \text{ are cointegrated.}
\]

If the null hypothesis can be accepted, it implies there is a unit root and there is no cointegration relationship between two time series \(x_t\) and \(y_t\). On the other hand, if the null hypothesis is rejected, this indicates that these two time series are cointegrated.

Maddala and Kim (1998) and Engle and Granger (1987) pointed out that the residual terms are obtained from estimating parameters. Different test statistics for unit root tests have different probability distribution. The asymptotic distribution of the t-statistic under the residual-based tests is not the standard t-distribution, using the conventional critical values can lead to incorrect conclusion of testing for unit root.

Fuller (1976) computed the critical values by simulations for the two test
statistics of Dicky-Fuller (DF) test (K test and $t_\rho$ test). These two test statistics for testing the unit root will be presented in the following subsection.

### 5.2.1 Dickey-Fuller Tests for the Test of Unit Root

The most commonly unit root test was developed by Fuller (1976) and Dickey and Fuller (1979). This test is based on underlying time series following a simple autoregression with or without a constant or time trend. It is generally called the Dickey-Fuller test. The autoregressive models of DF test used to test for cointegration are defined as below:

$$\hat{u}_t = \rho \hat{u}_{t-1} + w_t,$$

(5.6)

$$\hat{u}_t = \beta_0 + \rho \hat{u}_{t-1} + w_t,$$

(5.7)

$$\hat{u}_t = \beta_0 + \beta_1 t + \rho \hat{u}_{t-1} + w_t,$$

(5.8)

where $\hat{u}_t$ is the residual series from regressing the two underlying time series $x_t$ and $y_t$.

Equations (5.6), (5.7) and (5.8) are models without drift, with drift and the model both with drift and time trend respectively. While $\{w_t\}$ are i.i.d normally distributed random variables with mean zero and variance $\sigma^2$. In order to test whether $\hat{u}_t$ is stationary process ($\hat{u}_t \sim I(0)$), each equation can be formulated into the first differencing form ($\Delta \hat{u}_t$) for testing of $\rho = 1$. Equations (5.6), (5.7) and (5.8) can be rewritten as

$$\Delta \hat{u}_t = \theta \hat{u}_{t-1} + w_t,$$

(5.9)

$$\Delta \hat{u}_t = \beta_0 + \theta \hat{u}_{t-1} + w_t,$$

(5.10)

$$\Delta \hat{u}_t = \beta_0 + \beta_1 t + \theta \hat{u}_{t-1} + w_t.$$

(5.11)

where $\theta = \rho - 1$.

After using OLS to obtain the residual terms $\hat{u}_t$, the residual series are used to
construct a DF test to examine whether $\hat{u}_t$ has an unit root. The null hypothesis and alternative hypothesis of DF test are defined below:

\[ H_0 : \theta = 0, \]
\[ H_1 : \theta < 1. \]

The hypothesis test $\theta = 0$ and $\rho = 1$ are equivalent. In this chapter, the hypothesis test of $\rho = 1$ is used.

There are two types of statistics tests for DF test. The first one is called K-test statistic defined as follows:

\[ K = T(\hat{\rho} - 1), \quad (5.12) \]

where $T$ is the sample size and $\hat{\rho}$ is the maximum likelihood estimator of $\rho$ which is defined as follows:

\[ \hat{\rho} = \left( \sum_{t=1}^{T} \hat{u}_{t-1}^2 \right)^{-1} \sum_{t=1}^{T} \hat{u}_t \hat{u}_{t-1}. \quad (5.13) \]

Another test statistic for the DF test is the t-type test statistic. The t-statistic of DF test is normalized from K statistic in (5.12) with the standard error of the estimator of $\rho$ ($\hat{\rho}$). The $\hat{\rho}$ term is estimated by using OLS. The $t$-statistic of the test for $\rho = 1$ is

\[ t_{\hat{\rho}} = \frac{\hat{\rho} - 1}{\sqrt{T(1 - \hat{\rho}^2)}} \quad (5.14) \]

or

\[ t_{\hat{\rho}} = \frac{\hat{\rho} - 1}{SE(\hat{\rho})} \quad (5.15) \]

Both of the DF test statistics are used in this chapter. For further details, the asymptotic distributions of the Dickey-Fuller statistics are mentioned in Maddala and Kim (1998).
5.2.2 Cointegrating Regression Durbin-Watson Test for the Test of Unit Root

Another cointegration test related to the residual-based is Cointegrating Regression Durbin-Watson (CRDW) test (Sargan and Bhargawa, 1983). This test involves the series of residuals $\hat{u}_t$ from regressing time series $y_t$ on $x_t$. The estimated residuals are used to construct a DW statistic (know as the CRDW). The DW test is defined as follows:

$$DW = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{1=1}^{T} \hat{u}_t^2}$$

where $T$ is the sample size.

After obtaining the residual series by using OLS, the DW statistic is used to examine whether the residuals appear to be stationary. If the residuals are non-stationary, the DW statistic will be close to zero and thus the null hypothesis will be rejected. On the other hand, if the value of DW becomes larger, then the test tends to reject the null hypothesis of no cointegration (finds cointegration relationship). The one-sided alternative hypothesis of CRDW test is defined as below:

$$H_0 : DW = 0; \quad x_t \text{ and } y_t \text{ are not cointegrated}$$

$$H_1 : DW > 0; \quad x_t \text{ and } y_t \text{ are cointegrated}$$

The critical values of CRDW test for two time series were proposed by the work of Engle and Yoo (1987) used for sample sizes of 50, 100 and 200 of 1%, 5% and 10% significant level. The sample sizes considered in this chapter are 100 and 1,000. The critical value of CRDW test for sample size equal to 100 is available in the work of Engle and Yoo (1987) but for sample size equal to 1,000 is unavailable. Therefore, the critical value of CRDW test for sample size equal to 1,000 needs to be produced by Monte Carlo method.
5.2.3 The Wild Bootstrap Cointegration Test

Another cointegration test based on residual-based test is called the Wild Bootstrap test (WB). This test refers to the re-sampling method to obtain the acceptance region for testing unit roots. The WB requires appropriate estimations of the percentile of the left-tail from the distribution of the test statistic.

In this thesis, the WB test for unit root is that introduced by the work of Gerolimetto and Procidano (2003). The WB test is still carried out the null hypothesis of \( \rho = 1 \) by using the test statistics for the DF test. The DF statistics for the WB cointegration test is defined as follows:

\[
\tau = (\hat{\rho} - 1)S^{-1}\left(\sum_{t=1}^{T} \hat{u}_t^2\right)^{1/2} \quad t = 1, 2, \ldots, T \tag{5.17}
\]

where \( S^2 = (T - 2)^{-1}\sum_{t=1}^{T}(\hat{u}_t - \hat{\rho}\hat{u}_{t-1})^2 \), \( \hat{\rho} = (\sum_{t=1}^{T} \hat{u}_t\hat{u}_{t-1})(\sum_{t=1}^{T} \hat{u}^2_{t-1})^{-1} \) and \( T \) is the size of underlying sample.

The algorithm of the general WB procedure in Gerolimetto and Procidano (2003) is described as follows:

1. Consider the series of residuals \( \hat{u}_t \) from regressing time series \( y_t \) on \( x_t \) and \( \hat{u}_t \) follows AR(1) model \( \hat{u}_t = \rho\hat{u}_{t-1} + \varepsilon_t \), \( t = 1, 2, \ldots, T \) where \( \varepsilon_t \) is error term which is i.i.d.

2. Calculate the value of the DF statistic \( \tau \) which is defined below:

\[
\tau = (\hat{\rho} - 1)S^{-1}\left(\sum_{t=1}^{T} \hat{u}_t^2\right)^{1/2} \quad t = 1, 2, \ldots, T \tag{5.18}
\]

where \( S^2 = (T - 2)^{-1}\sum_{t=1}^{T}(\hat{u}_t - \hat{\rho}\hat{u}_{t-1})^2 \) and \( \hat{\rho} = (\sum_{t=1}^{T} \hat{u}_t\hat{u}_{t-1})(\sum_{t=1}^{T} \hat{u}^2_{t-1})^{-1} \).
3. Let $\varepsilon_t^B$ be a sample drawn from a standard normal distribution $(N(0, 1))$. Then generate sample of $\hat{u}_t^B$ with the wild bootstrap:

$$\hat{u}_t^B = \hat{u}_{t-1}^B + \frac{|\varepsilon_t|}{\sqrt{1 - h_{tt}}} \varepsilon_t^B \quad t = 1, 2, \ldots, T$$ (5.19)

where

$$h_{tt} = \hat{u}_{t-1}^2 \left( \sum_{t=1}^{T} \hat{u}_{t-1}^2 \right)^{-1}.$$  

and the initial value for $\hat{u}_1^B$ is set as $\hat{u}_1^B = \hat{u}_1$.

4. The DF test statistic $\tau^B$ was calculated from the re-sampling series of $\hat{u}_t^B$.

5. Repeating the steps 3 and 4, $J$ times to obtain the empirical distribution of $\tau^B$. Then calculating the percentile of the left tail of the distribution $\tau^B$ which is called the acceptance region in this thesis. After that, examining to decide whether statistics $\tau$ belong to the acceptance region.

6. Repeating the step 3 to 5, $N$ times to obtain the reject proportions from the wild bootstrap of the acceptance region.

5.3 The Johansen Cointegration Tests

Another type of cointegration test is based on ECM. The Johansen tests are the most popular cointegration methods with ECM which was developed by Johansen (1988, 1991). Typically, the Johansen method applies the Maximum Likelihood (ML) to the Vector Autoregressive (VAR) model in order to determine the number of cointegrating vectors $r$ for the vector time series system with the assumption that the errors in VAR model are normally distributed. Consider a VAR process
 CHAPTER 5. COINTEGRATION TESTS WITH NON-NORMAL GARCH ERROR DISTRIBUTIONS

\( y_t \) below:

\[
y_t = \sum_{j=1}^{k} A_j y_{t-j} + \epsilon_t
\]  
(5.20)

where \( y_t \) is the \( n \)-dimensional time series vector and all the components are integrated with the same order; \( A_j \) is a \((n \times n)\) coefficient matrix; \( k \) is an integer and \( \epsilon_t \) represents a vector of i.i.d. normal errors.

The Johansen method is a procedure for testing the number of cointegration relations among the underlying time series \( (y_t) \). If the components of \( y_t \) are cointegrated, the VAR in (5.20) can be transformed into a Vector Error-Correction Model (VECM) which can be used to analyze both of short-run and long-run relationship between tested time series. The VECM can be written as follows:

\[
\Delta y_t = \Pi y_{t-k} + \sum_{j=1}^{k-1} \Gamma_j \Delta y_{t-j} + \epsilon_t,
\]  
(5.21)

where \( \Gamma_j = \sum_{i=1}^{j} A_i - I \) and \( \Pi = \sum_{j=1}^{k} A_j - I \), \( I \) is an \((n \times n)\) identity matrix.

The matrix \( \Pi \) can be expressed as:

\[
\Pi = \alpha \beta'
\]  
(5.22)

where \( \alpha \) represents the \((n \times r)\) matrix of error-correction coefficients which refer to the speed of adjustment to the long-run equilibrium, \( \beta \) is the \((n \times r)\) matrix of cointegrating vectors with rank \( r \).

The Johansen method is called the reduced rank model by using ML to carried out a reduced rank regression. The first step is to calculate the residuals matrices of \( R_{0t} \) and \( R_{1t} \) from OLS regressions by regressing \( \Delta y_t \) and \( y_t \) on \( y_{t-1}, y_{t-2}, \ldots, y_{t-k+j} \) respectively. Next, compute the product moment matrices
given by
\[ S_{ij} = \frac{1}{T} \sum_{t=1}^{T} R_{it} R_{jt}, \quad i, j = 0, 1 \] (5.23)

Let \( \lambda_1 \geq \ldots \geq \lambda_p \) be the eigenvalue of \( S_{10}S_{00}^{-1}S_{01} - \lambda S_{11} \), and \( v_1, v_2, \ldots, v_p \) are the corresponding eigenvectors. The maximum likelihood estimator of \( \beta \) is defined as the eigenvectors corresponding that yields the first \( r \) largest eigenvalues. Two different likelihood ratio tests are suggested from the Johansen method. The first one is called the trace statistic (\( \lambda_{\text{trace}} \)) test. The \( \lambda_{\text{trace}} \) is given by
\[ \lambda_{\text{trace}} = -T \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i), \quad r = 0, 1, \ldots, (p-1) \] (5.24)

The trace statistic used to test that the null hypothesis of \( r \) cointegrated vectors against the alternative hypothesis of \( n \) cointegrating vectors (\( r < n \)). The second test statistic is called maximum eigenvalue test (\( \lambda_{\text{max}} \)). This statistic is used to test the null hypothesis of \( r \) cointegrating vectors against the alternative hypothesis of \( r + 1 \) cointegrating vectors. The \( \lambda_{\text{max}} \) is given by
\[ \lambda_{\text{max}} = -T \ln(1 - \hat{\lambda}_{r+1}), \quad r = 0, 1, \ldots, (p-1). \] (5.25)

The tests of trace statistic and maximum eigenvalue are analogous but their hypotheses are set differently. The null hypothesis of trace statistic test is that “There are at most \( r \) cointegration relations” against the alternative of “There are \( n \) cointegration relations” (\( r < n \)). For the maximum eigenvalue test, the null hypothesis is that “There are \( r \) cointegration relations” against the alternative of “There are \( r + 1 \) cointegration relations”.

Before examining the size and power of all cointegration tests by simulation studies, it is necessary to find an example that there exist cointegrated underlying financial time series for which residuals in VECM can be significantly fitted by a GARCH model with a non-normal error distribution. The next section demon-
strates an example of empirical cointegrating financial time series from the real observations when the errors are fitted by GARCH model with non-normal error distribution.

5.4 An Empirical Example of Cointegrating Errors Fitted by a GARCH Model with Non-Normal Error Distributions

The purpose of this section is to show if there exist financial time series which are cointegrated and the cointegrating errors can be fitted by a GARCH model with one of the six types of error distributions mentioned in Chapter 3.

To find an empirical example of cointegration, two daily closing price indices from Thailand (SET) and Malaysia (KLCI) are considered. The data set comprises daily closing prices of SET and KLCI spanning from 1 July 1998 to 31 December 2002. The Johansen cointegration tests are used in order to examine the cointegration relationship between these two emerging stock indices.

The log of daily closing prices of Thailand(SET) and Malaysia(KLCI) are denoted as $LSET$ and $LKLCI$ respectively while the first differencing for $LSET$ and $LKLCI$ are denoted as $\Delta LSET$ and $\Delta LKLCI$, respectively.

Figure 5.1 shows the plots of the log closing prices of SET and KLCI. The behavior and trend of the log closing prices of these two stock indices have similar co-movement over the period of time from 1 July 1998 to 31 December 2002. Figure 5.2 shows the plots of the first differencing of $LSET$ and $LKLCI$ respectively.

Table 5.1 reports the summary statistics for the log daily closing price of Thailand ($LSET$), Malaysia($LKLCI$) and their first differencing ($\Delta SET, \Delta KLCI$).

In preparation for the Johansen cointegration analysis, these two emerging stock indices need to be tested whether both series are non-stationary with the same integrated order. The Augmented Dicky-Fuller test (ADF) is applied to
CHAPTER 5. COINTEGRATION TESTS WITH NON-NORMAL GARCH ERROR DISTRIBUTIONS

Figure 5.1: Plot log of closing price of SET and KLCI spanning from 1 July 1998 to 31 December 2002, 1,133 in total observations (excluding public holidays and weekends)

Table 5.1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>LSET</th>
<th>LKLCI</th>
<th>ΔLSET</th>
<th>ΔLKLCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>5.3342</td>
<td>5.5710</td>
<td>-0.0777</td>
<td>-0.2415</td>
</tr>
<tr>
<td>Maximum</td>
<td>6.3024</td>
<td>6.9209</td>
<td>0.1022</td>
<td>0.2025</td>
</tr>
<tr>
<td>Mean</td>
<td>5.8432</td>
<td>6.5099</td>
<td>2.47E-04</td>
<td>3.08E-04</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.1843</td>
<td>0.2249</td>
<td>0.0186</td>
<td>0.0189</td>
</tr>
</tbody>
</table>

each log of closing price index and return series. These results are reported in Table 5.2.

“Lags” in Table 5.2 are the lags used in the Augmented Dickey-Fuller regression for the log of closing prices of two stock markets and their differencing. The ADF statistic of the log of closing prices of the two stock markets are clearly not significant but the ADF statistic of their first differencing are significant. This
CHAPTER 5. COINTEGRATION TESTS WITH NON-NORMAL GARCH ERROR DISTRIBUTIONS

90

CHAPTER 5. COINTEGRATION TESTS WITH NON-NORMAL GARCH ERROR DISTRIBUTIONS

Figure 5.2: Plots of the first differences of LSET and LKLCI spanning from 1 July 1998 to 31 December 2002, 1,133 in total observations (excluding public holidays and weekends)

Table 5.2: Unit root test results for two stock indices

<table>
<thead>
<tr>
<th>Stock market</th>
<th>Log Price</th>
<th>First differencing (Δ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-prob</td>
<td>t-ADF</td>
</tr>
<tr>
<td>SET</td>
<td>0.0416</td>
<td>0.2552</td>
</tr>
<tr>
<td>KLCI</td>
<td>0.0051</td>
<td>0.4673</td>
</tr>
</tbody>
</table>

Note: * and ** indicated the level of significance at 5% and 1% respectively.

implies that the log of closing prices of two stocks are non-stationary while their differencing are stationary. Therefore, both stock indices are integrated of the same order one (I(1)). Then cointegration test is now carried out.

In order to evaluate the possible cointegration relationship between these two stock markets, the Johansen tests for cointegration analysis between non-stationary time series (the log of closing prices of each stock index) are employed. Either the trace statistic ($\lambda_{trace}$) or the maximum eigenvalue statistic ($\lambda_{max}$) are
used.

The number of lags in the VECM (see Equation (5.21)) can affect the conclusion of cointegration analysis. Therefore, choosing an appropriate number of lags is the first priority step before testing for cointegration.

To choose an appropriate number of lags in the VECM, serial correlation tests have been carried out for lags up to 15. The tests are carried through the auxiliary regression of the residuals in VECM based on the two time series (LSET and LKLCI). The hypothesis testing is defined as follows:

\[ H_0 : \text{There is no autocorrelation at lag } k \text{ in VECM.} \]
\[ H_1 : \text{There is autocorrelation at lag } k \text{ in VECM.} \]

If the null hypothesis cannot be rejected, it means that statistically, there is no error autocorrelation in VECM. The residuals should not be serially correlated. The results of hypothesis testing for vector autocorrelation are reported in Table 5.3. The test statistics suggests that the appropriate lags used in VECM is 8 (VECM(8)), as it is the smallest k lag for which the null hypothesis cannot be rejected.

<table>
<thead>
<tr>
<th>Lag(k)</th>
<th>F-statistics</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.4089 [0.0000]**</td>
<td>139.42 [0.0000]**</td>
</tr>
<tr>
<td>2</td>
<td>2.1617 [0.0000]**</td>
<td>125.84 [0.0000]**</td>
</tr>
<tr>
<td>3</td>
<td>2.1691 [0.0000]**</td>
<td>126.54 [0.0000]**</td>
</tr>
<tr>
<td>4</td>
<td>2.0958 [0.0000]**</td>
<td>122.78 [0.0000]**</td>
</tr>
<tr>
<td>5</td>
<td>1.5475 [0.0048]**</td>
<td>92.29 [0.0047]**</td>
</tr>
<tr>
<td>6</td>
<td>1.4637 [0.0124]*</td>
<td>87.59 [0.0116]*</td>
</tr>
<tr>
<td>7</td>
<td>1.5350 [0.0415]*</td>
<td>91.88 [0.1245]*</td>
</tr>
<tr>
<td>8</td>
<td>1.5894 [0.3557]</td>
<td>95.19 [0.3741]</td>
</tr>
</tbody>
</table>

Note: (*) and (**) denote significantly at 5% and 1% level respectively.

<table>
<thead>
<tr>
<th>Lag(k)</th>
<th>F-statistics</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.4089 [0.0000]**</td>
<td>139.42 [0.0000]**</td>
</tr>
<tr>
<td>2</td>
<td>2.1617 [0.0000]**</td>
<td>125.84 [0.0000]**</td>
</tr>
<tr>
<td>3</td>
<td>2.1691 [0.0000]**</td>
<td>126.54 [0.0000]**</td>
</tr>
<tr>
<td>4</td>
<td>2.0958 [0.0000]**</td>
<td>122.78 [0.0000]**</td>
</tr>
<tr>
<td>5</td>
<td>1.5475 [0.0048]**</td>
<td>92.29 [0.0047]**</td>
</tr>
<tr>
<td>6</td>
<td>1.4637 [0.0124]*</td>
<td>87.59 [0.0116]*</td>
</tr>
<tr>
<td>7</td>
<td>1.5350 [0.0415]*</td>
<td>91.88 [0.1245]*</td>
</tr>
<tr>
<td>8</td>
<td>1.5894 [0.3557]</td>
<td>95.19 [0.3741]</td>
</tr>
</tbody>
</table>

Two tests, \( \lambda_{\text{trace}} \) and \( \lambda_{\text{max}} \) are employed to determine whether LSET and
LKLCI are cointegrated and the number of cointegrating vector. The values of the statistical tests are reported in Table 5.4

Table 5.4: Results and Critical Values for the $\lambda_{\text{trace}}$ and $\lambda_{\text{max}}$ test

<table>
<thead>
<tr>
<th>$r = 0$</th>
<th>$\lambda_{\text{trace}}$</th>
<th>$CV_{(\text{trace,5%})}$</th>
<th>$r = 0$</th>
<th>$\lambda_{\text{max}}$</th>
<th>$CV_{(\text{max,5%})}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r \leq 1$</td>
<td>6.82</td>
<td>8.18</td>
<td>$r = 1$</td>
<td>6.82</td>
<td>8.18</td>
</tr>
</tbody>
</table>

Notes: (**) denoted rejection of null hypothesis at 5% level.

Lütkepohl et al. (2001) showed the comparison results between the trace and maximum eigenvalue tests for the cointegrating rank of a VAR process. They reported that the power of the trace test was in some situations superior to that of the maximum eigenvalue test. Therefore, the two types of Johansen tests might not provide the same conclusion. The trace and maximum eigenvalue statistics might provide different cointegration test results.

From Table 5.4, the value of $\lambda_{\text{trace}}$ is 6.82 which is less than the critical value of 5% level (8.18). The null hypothesis of $r = 0$ cannot be accepted at 5% significant level. It indicates that there are not more than one cointegrating relationship between the stock market LSET and LKLCI at 5% level. However, the value of $\lambda_{\text{max}}$ is not significant at 5% level. By using $\lambda_{\text{trace}}$, it concludes that there is only one cointegrating vector for LSET and LKLCI.

The cointegrating vector is, $\beta' = (1.0000, -0.7532)$. The normalized cointegration equation with respect to LSET can be written below:

$$LSET = 0.7532LKLCI + \delta_t$$  \hspace{1cm} (5.26)

where $\delta_t$ is the cointegrating errors. The number of lag used in the VECM model is eight (i.e., VECM(8)). The residuals of the VECM(8) are saved and stored in order to check for model fitting.

After the cointegrating relationship between SET and KLCI is found, the
residuals of the VECM(8) are modeled to demonstrate that the residual series can be well fitted by a GARCH(p,q) model with non-normal error distribution.

According to the simulated results from Table 3.1 in Chapter 3, the order of a GARCH(p,q) model was not sensitive to the type of error distribution in the underlying GARCH(p,q) model.

To identify the best fitting GARCH(p,q) model, the process of determining the order of the best fitting GARCH model for the residuals is followed by the suggestion in Chapter 3. Competing GARCH(p,q) models are fitted by assuming the error terms are normally distributed. It is found that a GARCH(1,1) is the best fitting model for the residuals from the VECM(8). The determination of an appropriate error distribution from the six different types of error distributions for the GARCH(1,1) model based on AIC value is sequentially carried out. The results of AIC values are reported in Table 5.5.

Table 5.5: Comparative AIC values given by GARCH(1,1) with different type of error distributions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-5.2791</td>
<td>-5.2842</td>
<td>-5.3132</td>
<td><strong>-5.3166</strong></td>
<td>-5.3123</td>
<td>-5.3163</td>
</tr>
</tbody>
</table>

Table 5.5 shows that GARCH(1,1) model with skewed student-t distribution provides the smallest AIC values and it is the best fitting GARCH model for the residuals of the VECM(8). The estimation of parameters and diagnostics in the GARCH(1,1) model with SSTD error distribution are reported in Table 5.6.

All parameter estimations of GARCH(1,1) model with SSTD are significant at 5% level. Ljung-Box test is not significant, indicating that there is no autocorrelation in the residuals. The LM test indicates that the null hypothesis, which indicated that the ARCH effects are removed, can be accepted. It is clearly confirmed that the residuals in VECM can be well fitted by a GARCH model with a skewed Student-t error distribution.
Table 5.6: Estimated parameters and diagnostic of GARCH(1,1) model with SSTD

<table>
<thead>
<tr>
<th>GARCH(1,1) parameter</th>
<th>estimation of parameter</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
<td>$1.168 \times 10^{-5}$</td>
<td>0.06687*</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>$1.213 \times 10^{-1}$</td>
<td>0.00022***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$8.500 \times 10^{-1}$</td>
<td>&lt; $2 \times 10^{-16}$***</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.103</td>
<td>&lt; $2 \times 10^{-16}$***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>7.437</td>
<td>&lt; $2.01 \times 10^{-6}$***</td>
</tr>
<tr>
<td>ARCH(1)-LM test</td>
<td>6.396</td>
<td>0.9723</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>5.685</td>
<td>0.9311</td>
</tr>
</tbody>
</table>

Notes: The GARCH(1,1) model was defined by (3.1) and (3.4), the $\lambda$ and $\nu$ are skewness and shape parameter of SSTD respectively.

To investigate the performances of a variety of cointegration tests for detecting the relationship of time series, simulation approaches will be carried out in the next section.

5.5 Comparison of the Size and Power of Cointegration Tests with Various Distributional Assumptions of a GARCH(1,1) Model

In this chapter, three cointegration tests based on residual-based tests (DF test ($K$ and $t_\rho$ statistics), CRDW test and WB test) are compared with the Johansen likelihood ratio tests ($\lambda_{\text{trace}}$ and $\lambda_{\text{max}}$). In this regard, the performance of the size and power of these four cointegration tests are examined.

From literature, the Johansen tests appeared to provide superior results when the tests were originally applied to situations where the cointegrating errors were normally distributed. However, substantial empirical evidence shows that financial time series tend to be non-normal in their distribution which may, in turn, lead to non-normal GARCH type cointegrating error distributions. The question addresses in this chapter is whether the Johansen tests are still more powerful than the three alternative tests when the underlying cointegrating errors are
non-normally distributed.

The performance of size and power of cointegration tests have been examined with the variety of GARCH error assumptions. Kim and Schmidt (1993) investigated the size of cointegration tests when the cointegrating errors followed a GARCH model with normal random error distributions. They found that the DF tests tend to over-reject the null hypothesis of no cointegration in the presence of GARCH errors. Lee and Tse (1996) examined the performance of Johansen tests compared with DF tests and CRDW test when cointegrating errors were fitted by a GARCH(1,1) model when the GARCH error terms were normal and Student-t distributions. They compared the performance of the size and power of the Johansen tests with other cointegration tests, DF tests and CRDW test. Their conclusion was that although the Johansen cointegration tests tended to over-reject the null hypothesis of no cointegration, the power of the Johansen tests remain higher than the other two tests when the cointegrating errors were fitted by GARCH(1,1) model with normal error distributions.

Furthermore, Gerolimetto and Procidano (2003) examined the WB test compared with DF tests, CRDW test and Johansen tests when the cointegrating errors followed a GARCH(1,1) model with normal random error distributions. Their conclusions were the same as the works of Lee and Tse (1996) when comparing the Johansen tests with the other three cointegration tests.

As mentioned before, the empirical distributions of financial time series are significantly non-normally distributed which exhibit skewness, excess kurtosis and heavy tails. A large number of heavy tail distributions have been identified from financial time series and applied in modeling and estimating GARCH models. Cheung and Lai (1993) and Gonzalo (1994) investigated the effect of non-Gaussian error distribution on the performance of the Johansen cointegration tests. They found that the Johansen tests are robust to both skewness and
excess kurtosis in cointegrating errors. The study in this chapter is an extension of the previous works in order to examine whether the performance of power of Johansen tests remain higher compared with the other three cointegration tests when the cointegrating errors followed a GARCH(1,1) model with normal error distribution and the other five non-normal error distributions.

This chapter is the extensions of the work of Lee and Tse (1996) as well as the work of Gerolimetto and Procidano (2003) with a wide range of six types of error distributions mentioned in Chapter 3. These six types of error distributions are applied to the errors of GARCH(1,1) models. It is of interest whether the type of error distribution has any significant impact on the performance of the size and power of cointegration test analysis. Therefore, the objective of this chapter is to investigate which cointegration test is the most powerful in detecting cointegration relationship when the cointegrating errors follow a GARCH(1,1) model with normal error distribution and other five non-normal error distributions.

This section deals with the performance of size and power of the cointegration tests, the four tests were mentioned in 5.2 and 5.3 (Jonhansen, DF, CRDW and WB tests). All four cointegration approaches are examined when the cointegrating errors follow GARCH(1,1) model with the six different types of errors distributions respectively (N, SN, STD, SSTD, GED and SGED). The comparison of performances are carried out by simulation approach. The Monte Carlo simulations are conducted for cointegration tests when cointegrating errors follow a GARCH(1,1) model with six various types of errors distributions.

The Size of the Cointegration Tests

The size of the tests are evaluated by counting the frequency of rejections of the null hypothesis. The null hypothesis is stated that the true underlying series
are generated from non-cointegrated model or contain unit roots. If a test accepts the null hypothesis, it means that the true property of the underlying series is correctly identified. On the other hand, if a test rejects the null hypothesis, it means this test fails to identify the true property of underlying series.

To examine the size of the tests, simulated samples from a non-cointegrated system $X_t = (x_{1t}, x_{2t})$ with a GARCH(1,1) model when the GARCH errors follow the six types of error distributions (N, SN, STD, SSTD, GED and SGED). The non-cointegrated system is given by:

$$
\Delta x_{1t} = e_{1t} \tag{5.27}
$$

$$
\Delta x_{2t} = e_{2t} \tag{5.28}
$$

where the errors $e_{1t}$ and $e_{2t}$ follow GARCH(1,1) model with six types of error distributions mentioned above. A GARCH(1,1) model is defined as follows:

$$
e_{it} = \eta_{it}\sqrt{h_{it}}, \tag{5.29}
$$

$$
h_{it} = \omega_i + \alpha_i e_{it-1}^2 + \beta_i h_{it-1}, \tag{5.30}
$$

where $\eta_{it}$ are i.i.d. with $E(\eta_{it}) = 0$, $Var(\eta_{it}) = 1$ for $i=1,2$ and $t=1,2,\ldots,T+d$. The first $d = 500$ observations are discarded. Simulated 10,000 samples are generated independently with sample sizes $T$ equal to 100 and 1,000 respectively from that system and applied the four cointegration tests (Johansen, DF, CRDW and WB tests) to the simulated data. The comparisons of the frequency of rejection non cointegration crossing the four tests are carried out. In terms of size of the tests, for better performance of the test, it should have less frequency of rejections of the null hypothesis.

For the cointegration tests based on residual-based tests, the residuals from regressing these two variables ($x_{1t}$ and $x_{2t}$) are saved and residual series used to carried out the hypothesis testing of the DF, CRDW and WB tests.
The Power of the Cointegration Tests

For the power of the test, it is evaluated by counting the frequency of rejections of the null hypothesis. The null hypothesis is stated that the true underlying series are generated from non-cointegrated model. If a test rejects the null hypothesis, it means that the test can well detect the cointegration relationship. To have a better performance, a test should have high frequency of rejections of the null hypothesis.

To examine the power of a cointegration test, data from a bivariate cointegrated system are generated as follows:

\[ \Delta x_{1t} = -0.2(x_{1,t-1} - x_{2,t-1}) + e_{1t} \]  
\[ \Delta x_{2t} = e_{2t} \]

(5.31)  
(5.32)

where the errors \( e_{1t} \) and \( e_{2t} \) follow GARCH(1,1) model respectively. The frequency of rejections non cointegration are counted. If the test is robust and powerful then we expect to obtain the higher proportion of rejection the null hypothesis.

Similarly, an independently simulated 10,000 samples with sample sizes \( T \) equal to 100 and 1,000 observations from the power of the cointegration system are generated. Then the four cointegration tests are applied to the simulated data.

In these simulation studies, GARCH parameters are chosen when \( \omega_1 = \omega_2, \alpha_1 = \alpha_2 \) and \( \beta_1 = \beta_2 \) (see Equation 5.30). To simplify the notation, the index \( i \) from \( h_{it} \) and \( (\omega_i, \alpha_i, \beta_i) \) are omitted. The two sets of GARCH(1,1) parameters in this simulation studies are \((0.1, 0.3, 0.6)\) and \((0.1, 0.65, 0.05)\). The probability distributions of the errors in GARCH(1,1) models \((e_{1t} \text{ and } e_{2t})\) are normal, skewed
normal, Student-t, skewed Student-t, GED and skewed GED respectively.

The values of skewness of skewed normal distribution, skewed student-t distribution and skewed generalized error distribution are chosen as 0.1, 0.5 and 3 respectively. The values of degree of freedom $\nu$ are considered by the work of Lee and Tse (1996) which were equal to 5 and 8 respectively. The kurtosis of the student-t distribution is given by $3(\nu - 2)/(\nu - 4)$ for $\nu > 4$; it is 9 for $\nu = 5$ and 4.5 for $\nu = 8$. The shape parameter of skewed normal distribution, generalized error distribution and skewed generalized error distribution is equal to 3.

5.6 The Results for the Size and Power of Cointegration Tests Based on Simulated Data

This section reports the results from the size and power of four cointegration tests when cointegration errors follow a GARCH(1,1) model with one of six different types of error distribution. Both non-cointegrated and cointegrated systems are designed in the different situations of parameters to investigate how much different impact on the performance of size and power of these four cointegration tests.

The cointegrating error terms $e_{1t}$ and $e_{2t}$ in both cointegrated system and non-cointegrated system are generated from GARCH(1,1) model. To investigate the impact of parameters in GARCH(1,1) model, two different sample sizes and the six different error distributions in GARCH error terms on the performance of size and power of cointegration tests are used. The error terms of each two systems are changed by the different set of parameters of GARCH(1,1) model.

The different choice of parameters are set in these cointegrated simulations such as the two different set of parameter in GARCH(1,1) model ($\omega = 0.1, \alpha = 0.3, \beta = 0.6), (\omega = 0.1, \alpha = 0.65, \beta = 0.05)$, two different sample size ($T=100, 1,000$), the shape parameter is used as 3 for SN, GED and SGED, for STD and SSTD the shape parameters are 4.5 and 9, the skew parameter used in
these simulations for SN, SSTD and SGED are 0.1, 0.5 and 3. To compare the
performance of the size and power of four cointegration tests under the presence of
a GARCH(1,1) model with six types of error distributions, the simulated results
are reported by considering of these factors.

The size of cointegration tests are reported in Tables 5.7 to 5.10 by considering
the different situation of parameters in the non-cointegrated system from (5.27)
and (5.28).

For two different sample sizes, the size of the tests tend to decrease when the
sample size $T$ increases by comparing within the same set of GARCH(1,1) pa-
rameter while the error distributions are the same. For instance, the size of trace
statistic for 100 sample size is equal to 0.0930 when the cointegrating errors follow
GARCH(1,1) with normal distribution. For sample size 1,000, the size of trace
statistic decreases from 0.0931 to 0.0656 with the same error distribution. The
tendencies of the size of cointegration tests are similar when the other different
situation of parameters are considered. It indicates that the larger sample size
affected the change in the performance of the size of the tests.

For the two sets of GARCH(1,1) parameter, the size of the tests tend to
be larger when $(\alpha + \beta)$ is larger,. In terms of the six different types of error
distributions in GARCH(1,1) error, when the scale of skewness increases, the
size of the tests are not much different. However, the scale of shape parameter
increased in STD and SSTD, the size of the tests tend to decrease.

To compare the performance of the size of the tests by considering the four
different cointegration tests, the DF test with $T(\hat{\rho} - 1)$ statistic provides the
best performance with the lowest values (see the majority of the highlight values
appeared in DF $T(\hat{\rho} - 1)$ in Tables 5.7 - 5.10, but the size of the Johansen
tests with Trace($\lambda_{\text{trace}}$) statistic sometimes outperforms the DF test, when $(\omega =
0.1, \alpha = 0.65, \beta = 0.05)$ of GARCH(1,1) with SSTD are considered (see Table
5.8). The size of DF tests tend to be smaller when the cointegrating errors follow a GARCH(1,1) with SGED.

Tables 5.11 - 5.14 report the performance of the power of the tests for all cointegration tests under the presence of GARCH(1,1) model. Six different types of error distributions and different situations of parameters are considered. The results of power of the tests are very consistent in all different features.

The power of all cointegration tests are not much different in all situations of parameters when $T = 1,000$ (this results are exactly same as the work by Lee and Tse, 1996). Therefore, to compare the performance of power of the tests under different parameters, the results are only described for $T = 100$.

When the sample size is equal to 100 in Tables 5.11 - 5.12, the power of the Johansen tests are higher than all other cointegration tests. Particularly, the performance of the $\lambda_{max}$ statistic is slightly higher than the $\lambda_{trace}$ statistic, but not much different.

This result suggests that the Johansen tests are superior to the other alternative cointegration tests. However, the power of the Johansen tests are slightly more powerful when $(\alpha + \beta)$ is larger.

Consider the studies of the power of Johansen tests under all six different types of error distributions, a GARCH(1,1) with SSTD provides the better performance than the others error distributions. The changing of the scale of the shape parameter in SSTD for the Johansen tests do not yield much difference on the power of the Johansen tests. It also can be seen that the Johansen tests with $\lambda_{max}$ statistic outperforms the $\lambda_{trace}$ statistic with 100 sample sizes.
CHAPTER 5. COINTEGRATION TESTS WITH NON-NORMAL GARCH ERROR DISTRIBUTIONS

5.7 Conclusion

The performance of different cointegration tests, when the cointegrating error following a GARCH(1,1) model with six different types of error distributions are examined. The size of Dickey-Fuller test with $T(\hat{\rho} - 1)$ statistic is lower when compared with other cointegration tests and this test tends to be smaller for the errors of a GARCH(1,1) model with skewed generalized error distribution. It indicates that the Dickey-Fuller tests yields the best performance in terms of the size of the test. However, the power of the Dickey-Fuller tests are very low compared with the Johansen tests.

The power of Johansen tests provides the best performance in all different type of parameters compared with other cointegration tests. The power of Johansen tests with the $\lambda_{\text{max}}$ statistic is slightly better than the $\lambda_{\text{trace}}$ statistic with 100 sample sizes. Furthermore, the power of Johansen tests slightly increases when the errors of GARCH(1,1) model is given by the skewed Student-t error distribution.
Table 5.7: The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 100$ at 5% level

<table>
<thead>
<tr>
<th>Parameter</th>
<th>100 observations</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\omega = 0.1, \alpha = 0.3, \beta = 0.6)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Trace($\lambda_{trace}$)</td>
<td></td>
<td>0.0930</td>
<td>0.0909</td>
<td>0.0894</td>
<td>0.0891</td>
<td>0.0862</td>
</tr>
<tr>
<td>Maxeigen($\lambda_{max}$)</td>
<td></td>
<td>0.0905</td>
<td>0.0940</td>
<td>0.0891</td>
<td>0.0899</td>
<td>0.0884</td>
</tr>
<tr>
<td>CRDW</td>
<td></td>
<td>0.0730</td>
<td>0.0731</td>
<td>0.0719</td>
<td>0.0722</td>
<td>0.0695</td>
</tr>
<tr>
<td>DF ($t_{\rho}$)</td>
<td></td>
<td>0.0762</td>
<td>0.0800</td>
<td>0.0717</td>
<td>0.0743</td>
<td>0.0758</td>
</tr>
<tr>
<td>DF $T(\hat{\rho} - 1)$</td>
<td></td>
<td>0.0665</td>
<td>0.0673</td>
<td>0.0649</td>
<td>0.0640</td>
<td>0.0618</td>
</tr>
<tr>
<td>WB</td>
<td></td>
<td>0.0777</td>
<td>0.0797</td>
<td>0.0719</td>
<td>0.0744</td>
<td>0.0749</td>
</tr>
<tr>
<td>(\omega = 0.1, \alpha = 0.3, \beta = 0.6)</td>
<td></td>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Trace($\lambda_{trace}$)</td>
<td></td>
<td>0.1070</td>
<td>0.0989</td>
<td>0.1029</td>
<td>0.1006</td>
<td>0.1057</td>
</tr>
<tr>
<td>Maxeigen($\lambda_{max}$)</td>
<td></td>
<td>0.1065</td>
<td>0.1027</td>
<td>0.1043</td>
<td>0.1035</td>
<td>0.1066</td>
</tr>
<tr>
<td>CRDW</td>
<td></td>
<td>0.0822</td>
<td>0.0778</td>
<td>0.0903</td>
<td>0.0865</td>
<td>0.0877</td>
</tr>
<tr>
<td>DF ($t_{\rho}$)</td>
<td></td>
<td>0.0854</td>
<td>0.0808</td>
<td>0.0908</td>
<td>0.0879</td>
<td>0.0893</td>
</tr>
<tr>
<td>DF $T(\hat{\rho} - 1)$</td>
<td></td>
<td>0.0761</td>
<td>0.0710</td>
<td>0.0824</td>
<td>0.0799</td>
<td>0.0798</td>
</tr>
<tr>
<td>WB</td>
<td></td>
<td>0.0849</td>
<td>0.0790</td>
<td>0.0927</td>
<td>0.0879</td>
<td>0.0882</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
### Table 5.8: The size of the test for GARCH(1,1) with \( \omega = 0.1 \), \( \alpha = 0.65 \), \( \beta = 0.05 \) and \( T = 100 \) at 5% level

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Shape</th>
<th>STD</th>
<th>STDP</th>
<th>SSTD</th>
<th>STD</th>
<th>STDP</th>
<th>SSTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>0.5</td>
<td>0.0838 &amp; 0.0800 &amp; 0.0680 &amp; 0.0690 &amp; 0.0690 &amp; 0.0670 &amp; 0.0670 &amp; 0.0670</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.0318 &amp; 0.0318 &amp; 0.0318 &amp; 0.0318 &amp; 0.0318 &amp; 0.0318 &amp; 0.0318 &amp; 0.0318</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.9680 &amp; 0.9680 &amp; 0.9680 &amp; 0.9680 &amp; 0.9680 &amp; 0.9680 &amp; 0.9680 &amp; 0.9680</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.9160 &amp; 0.9160 &amp; 0.9160 &amp; 0.9160 &amp; 0.9160 &amp; 0.9160 &amp; 0.9160 &amp; 0.9160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.0813 &amp; 0.0813 &amp; 0.0813 &amp; 0.0813 &amp; 0.0813 &amp; 0.0813 &amp; 0.0813 &amp; 0.0813</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.0880 &amp; 0.0880 &amp; 0.0880 &amp; 0.0880 &amp; 0.0880 &amp; 0.0880 &amp; 0.0880 &amp; 0.0880</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960 &amp; 0.0960</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800 &amp; 0.0800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
Table 5.9: The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 1,000$ at 5% level

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1,000 observations</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(SN)</td>
<td>GED</td>
<td>SGED</td>
<td>(SN)</td>
<td>GED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Trace($\lambda_{trace}$)</td>
<td>0.0719</td>
<td>0.0753</td>
<td>0.0736</td>
<td>0.0746</td>
<td>0.0646</td>
</tr>
<tr>
<td>Maxeigen($\lambda_{max}$)</td>
<td>0.0759</td>
<td>0.0761</td>
<td>0.0748</td>
<td>0.0735</td>
<td>0.0696</td>
</tr>
<tr>
<td>CRDW</td>
<td>0.0676</td>
<td>0.0689</td>
<td>0.0702</td>
<td>0.0753</td>
<td>0.0648</td>
</tr>
<tr>
<td>DF ($t_\rho$)</td>
<td>0.0660</td>
<td>0.0651</td>
<td>0.0708</td>
<td>0.0738</td>
<td>0.0614</td>
</tr>
<tr>
<td>DF $T(\hat{\rho} - 1)$</td>
<td>0.0583</td>
<td>0.0596</td>
<td>0.0630</td>
<td>0.0653</td>
<td>0.0564</td>
</tr>
<tr>
<td>WB</td>
<td>0.0597</td>
<td>0.0679</td>
<td>0.0686</td>
<td>0.0724</td>
<td>0.0698</td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Trace($\lambda_{trace}$)</td>
<td>0.0839</td>
<td>0.0764</td>
<td>0.0955</td>
<td>0.0998</td>
<td>0.0969</td>
</tr>
<tr>
<td>Maxeigen($\lambda_{max}$)</td>
<td>0.0855</td>
<td>0.0761</td>
<td>0.0988</td>
<td>0.0998</td>
<td>0.1005</td>
</tr>
<tr>
<td>CRDW</td>
<td>0.0822</td>
<td>0.0721</td>
<td>0.1050</td>
<td>0.1010</td>
<td>0.1008</td>
</tr>
<tr>
<td>DF ($t_\rho$)</td>
<td>0.0785</td>
<td>0.0701</td>
<td>0.1024</td>
<td>0.0978</td>
<td>0.0967</td>
</tr>
<tr>
<td>DF $T(\hat{\rho} - 1)$</td>
<td>0.0724</td>
<td>0.0640</td>
<td>0.0936</td>
<td>0.0909</td>
<td>0.0915</td>
</tr>
<tr>
<td>WB</td>
<td>0.0796</td>
<td>0.0681</td>
<td>0.0941</td>
<td>0.0889</td>
<td>0.0921</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests
Table 5.10: The size of the test for GARACH(1,1) with $\omega = 0.1$, $\alpha = 0.65$, $\beta = 0.05$ and $T = 1,000$ at 5% level

<table>
<thead>
<tr>
<th>Shape</th>
<th>0</th>
<th>0.5</th>
<th>0.95</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>WN</td>
<td>0.0694</td>
<td>0.0755</td>
<td>0.0705</td>
<td>0.0733</td>
</tr>
<tr>
<td>CRDW</td>
<td>0.0667</td>
<td>0.0729</td>
<td>0.0726</td>
<td>0.0677</td>
</tr>
<tr>
<td>DF</td>
<td>0.0615</td>
<td>0.0646</td>
<td>0.0619</td>
<td>0.0677</td>
</tr>
<tr>
<td>WB</td>
<td>0.0614</td>
<td>0.0681</td>
<td>0.0643</td>
<td>0.0669</td>
</tr>
<tr>
<td>STD</td>
<td>0.0607</td>
<td>0.0607</td>
<td>0.0607</td>
<td>0.0607</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
Table 5.11: The power of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 100$ at 5% level

<table>
<thead>
<tr>
<th>Parameter</th>
<th>100 observations</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\omega = 0.1, \alpha = 0.3, \beta = 0.6)$</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
<td>Shape</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1 0.5 3</td>
<td>0.1 0.5 3</td>
<td>0.1 0.5 3</td>
<td>0.1 0.5 3</td>
<td>Trace($\lambda_{trace}$)</td>
<td>0.8027 0.8030 0.8015 0.8033 0.7978 0.8014 0.8029 0.8002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maxeigen($\lambda_{max}$)</td>
<td>0.8254 0.8278 0.8266 0.8270 0.8225 0.8206 0.8267 0.8225</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CRDW</td>
<td>0.6731 0.7044 0.6951 0.6988 0.6856 0.6934 0.6868 0.6913</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DF ($t_\rho$)</td>
<td>0.5614 0.5902 0.5781 0.5834 0.5681 0.5777 0.5796 0.5786</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DF $T(\hat{\rho} - 1)$</td>
<td>0.6166 0.6464 0.6413 0.6422 0.6269 0.6339 0.6363 0.6325</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>0.5750 0.6023 0.5987 0.5994 0.5843 0.5952 0.5933 0.5940</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\omega = 0.1, \alpha = 0.3, \beta = 0.6)$</td>
<td>STD</td>
<td>SSTD</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1 0.5 3</td>
<td>0.1 0.5 3</td>
<td>0.1 0.5 3</td>
<td>0.1 0.5 3</td>
<td>Trace($\lambda_{trace}$)</td>
<td>0.8000 0.8010 0.8135 0.8113 0.8098 0.8145 0.8063 0.8109</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maxeigen($\lambda_{max}$)</td>
<td>0.8243 0.8227 0.8362 0.8344 0.8351 0.8284 0.8278 0.8301</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CRDW</td>
<td>0.6731 0.6738 0.7359 0.7163 0.7264 0.7160 0.6994 0.7134</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DF ($t_\rho$)</td>
<td>0.5672 0.5628 0.6234 0.6177 0.6142 0.6027 0.5883 0.6044</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DF $T(\hat{\rho} - 1)$</td>
<td>0.6170 0.6181 0.6811 0.6656 0.6741 0.6625 0.6435 0.6624</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>0.5825 0.5834 0.6224 0.6201 0.6224 0.6199 0.6020 0.6209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
Table 5.12: The size of the test for GARCH(1,1) with non-normal GARCH error distributions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Shape</th>
<th>Skew</th>
<th>Trace(λ)</th>
<th>Maxeigen(λ)</th>
<th>CRDW</th>
<th>DF(\hat{\rho})</th>
<th>DF(\hat{\rho}−1)</th>
<th>WB</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SGED</td>
<td>0.8078</td>
<td>0.8076</td>
<td>0.6726</td>
<td>0.6167</td>
<td>0.6148</td>
<td>0.5792</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SGED</td>
<td>0.8090</td>
<td>0.8080</td>
<td>0.7094</td>
<td>0.6051</td>
<td>0.6095</td>
<td>0.5955</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SN</td>
<td>0.8123</td>
<td>0.8123</td>
<td>0.6996</td>
<td>0.5967</td>
<td>0.5967</td>
<td>0.6097</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>GEED</td>
<td>0.8071</td>
<td>0.8090</td>
<td>0.7031</td>
<td>0.6061</td>
<td>0.6046</td>
<td>0.5742</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SGED</td>
<td>0.8045</td>
<td>0.8045</td>
<td>0.6863</td>
<td>0.5783</td>
<td>0.5763</td>
<td>0.5545</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SGED</td>
<td>0.8061</td>
<td>0.8061</td>
<td>0.6945</td>
<td>0.5921</td>
<td>0.5921</td>
<td>0.5906</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SN</td>
<td>0.8076</td>
<td>0.8076</td>
<td>0.6936</td>
<td>0.5873</td>
<td>0.5873</td>
<td>0.5873</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>GEED</td>
<td>0.8039</td>
<td>0.8039</td>
<td>0.6936</td>
<td>0.5807</td>
<td>0.5807</td>
<td>0.5807</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SGED</td>
<td>0.8076</td>
<td>0.8076</td>
<td>0.6936</td>
<td>0.5807</td>
<td>0.5807</td>
<td>0.5807</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SGED</td>
<td>0.8026</td>
<td>0.8026</td>
<td>0.6936</td>
<td>0.5807</td>
<td>0.5807</td>
<td>0.5807</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>SN</td>
<td>0.8076</td>
<td>0.8076</td>
<td>0.6936</td>
<td>0.5807</td>
<td>0.5807</td>
<td>0.5807</td>
</tr>
<tr>
<td>100 obs.</td>
<td>N</td>
<td>GEED</td>
<td>0.8039</td>
<td>0.8039</td>
<td>0.6936</td>
<td>0.5807</td>
<td>0.5807</td>
<td>0.5807</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
Table 5.13: The size of the test for GARACH(1,1) with $\omega = 0.1, \alpha = 0.3, \beta = 0.6$ and $T = 1,000$ at 5% level

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1,000 observations</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N)</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Trace($\lambda_{trace}$)</td>
<td>1.0000</td>
<td>0.9999</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Maxeigen($\lambda_{max}$)</td>
<td>1.0000</td>
<td>0.9999</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>CRDW</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>DF ($t_\beta$)</td>
<td>1.0000</td>
<td>0.9999</td>
<td>1.0000</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
<tr>
<td>DF $T(\hat{\rho} - 1)$</td>
<td>1.0000</td>
<td>0.9999</td>
<td>1.0000</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
<tr>
<td>WB</td>
<td>1.0000</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

| (N = 0.1, $\alpha = 0.3, \beta = 0.6$) | STD |          |          |          |          |
| Shape                      | 4.5 | 9        | 4.5      | 9        |          |
| Skew                       | 0.1 | 0.5      | 3        | 0.1      | 0.5      | 3        |
| Trace($\lambda_{trace}$)   | 1.0000 | 1.0000 | 1.0000   | 1.0000   | 0.9999   | 0.9999   | 1.0000   | 1.0000   |
| Maxeigen($\lambda_{max}$)  | 1.0000 | 1.0000 | 1.0000   | 1.0000   | 1.0000   | 0.9999   | 1.0000   | 0.9999   |
| CRDW                       | 1.0000 | 1.0000 | 1.0000   | 1.0000   | 1.0000   | 1.0000   | 1.0000   | 1.0000   |
| DF ($t_\beta$)             | 1.0000 | 1.0000 | 1.0000   | 1.0000   | 0.9999   | 0.9998   | 1.0000   | 0.9999   |
| DF $T(\hat{\rho} - 1)$     | 1.0000 | 1.0000 | 1.0000   | 1.0000   | 1.0000   | 0.9999   | 0.9999   | 1.0000   |
| WB                         | 1.0000 | 1.0000 | 1.0000   | 1.0000   | 0.9999   | 0.9999   | 0.9999   | 0.9999   |

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
Table 5.14: The size of the test for GARACH(1,1) with $\omega=0.1$, $\alpha=0.65$, $\beta=0.05$ and $T=1,000$ at 5% level

<table>
<thead>
<tr>
<th>Parameter</th>
<th>STD</th>
<th>SSTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
<td>$0.000$</td>
<td>$1.000$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$0.000$</td>
<td>$1.000$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$0.000$</td>
<td>$1.000$</td>
</tr>
<tr>
<td>$T$</td>
<td>$1,000$</td>
<td>$0.000$</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best performance of the test comparing among alternative cointegration tests.
Chapter 6

Improving Volatility Forecasting Based on Cointegration Information

Making an accurate volatility forecast is very important for both financial investors and policy makers. Asset portfolio selections and trading techniques rely on accurate volatility forecasts. The more accurate of the forecast volatility, the more efficient of the financial risk management.

In Chapter 4, the volatilities of underlying financial time series of stock indices in three emerging stock exchanges were individually modeled using GARCH(p,q) models with alternative types of error distributions. The best fitting GARCH model with an appropriate error distribution for each stock exchange index was used to forecast the future volatility. It was clear that GARCH(p,q) models with non-normal error distributions were superior to GARCH(p,q) models with a normal error distribution for those three stock markets.

Future volatility forecasting can be improved using symmetric GARCH(p,q) models with alternative assumptions of GARCH error terms. It is one of the approaches used to improve the volatility forecasting of underlying financial time series by changing GARCH error terms with non-normal distributions. Other alternative approaches used to improve the performance of volatility forecasting have been proposed, such as using asymmetric GARCH models, applying non-
normal error distributions into asymmetric GARCH error terms, and adopting some statistical techniques for achieving better accuracy in prediction. Numerous studies on the improvement of volatility forecasting performance using alternative modifications of GARCH models are briefly reviewed in the following section.

6.1 Alternative Approaches Employed to Improve the Performance of Volatility Forecasting

In order to improve volatility forecasting, numerous modifications of symmetric and asymmetric GARCH models have been developed and these modifications lead to better forecasting performance. Liu and Hung (2010) investigated the performance of volatility forecasting for the Standard & Poor’s 100 stock index series by using six different types of GARCH models, including the symmetric GARCH model with four alternative types of error distributions (standard normal, standard Student-t, heavy-tailed and skewed generalized-t distributions) and two asymmetric GARCH models (GJR-GARCH and EGARCH). Comparative analysis of six competing models for fitting and forecasting volatility were carried out. Their empirical results showed that the GJR-GARCH model provided the most accurate volatility forecast, followed by the EGARCH model. According to their results, they claimed that an asymmetric GARCH model yielded better performance in volatility forecasting than a symmetric GARCH model with alternative error distributions. Franses and van Dijk (1996) examined the volatility forecasting performance of two non-linear modifications of GARCH models for the weekly stock market. They compared the performance of the Quadratic GARCH (QGARCH) model with the GJR-GARCH model and found that the QGARCH model significantly outperformed the GJR-GARCH model when the underlying sample did not contain extreme observations.
In terms of improving volatility forecasting performance, the studies relating to the modifications of GARCH models have been conducted by incorporating some traditional statistical techniques into the models, such as the principle of Regime-Switching, the Artificial Neural Networks (ANN), the moving average and the exponential smooth transition method. Due to the high persistence of individual shocks related to GARCH estimating and forecasting, Klaassen (2002) developed the Markov regime-switching GARCH model by distinguishing two regimes with different volatility levels. A GARCH model was formed by using two regimes with different levels of volatility and regime-specific. GARCH formulas were used to describe the variance within the regimes. The out-of-sample results revealed that the regime-switching GARCH model provided significantly better volatility forecasting performance than the standard GARCH model.

Taylor (2004) adopted the new smooth transition exponential smoothing approach with different types of GARCH(1,1) models for volatility forecasting and considered the 1-step ahead volatility prediction. The results showed that the 1-step ahead volatility forecasting of GARCH(1,1) models using the new smooth transition exponential smoothing method performed better than the standard symmetric GARCH(1,1) model and asymmetric GARCH(1,1) (IGARCH, GJR-GARCH) models.

Anwar and Mikami (2007) examined the accuracy of volatility forecasting by comparing the Artificial Neural Networks (ANN), Multiple Linear Regressions (MLR) and a GARCH model. All their empirical results demonstrated that the ANN model outperformed the MLR and the GARCH model in explaining the Rate of Return (RR) of Indonesian Islamic banks.

Xu and Liu (2011) investigated the volatility forecasting of the two Chinese stock indices, the Shanghai and Shenzhen stock markets, by comparing the forecasting performance of the Empirical Mode Decomposition and Neural Network
learning paradigm (EMD-NN), the GARCH(1,1) family (GARCH, EGARCH and GJR-GARCH) models, the Neural Network (NN) method and the moving average method. The out-of-sample forecasts showed that the EMD-NN provided better volatility forecasting performance than other models while the moving average was the lowest performer in the Shenzhen stock market based on the MAE criterion. The GJR-GARCH model performed the worst in both stock markets and the Neural Network was superior to the GARCH(1,1) family models.

It is apparent from previous studies on improving the volatility forecasting of underlying financial time series that the traditional statistic methods can be applied to the standard symmetric and asymmetric GARCH models. This is done by incorporating certain features of those traditional statistic methods into the different types of standard GARCH models. All new modifications of GARCH models in the literature reviews yield better volatility forecasting performance than the conventional GARCH models.

There is an effort in this thesis to improve the performance of volatility forecasting by considering the cointegration relationships among time series. If underlying time series are found to be cointegrated and each of these time series can be well fitted by univariate GARCH models individually then the knowledge of cointegration relationships among cointegrated time series can be used to improve the performance of volatility forecasting comparing with the GARCH model built based on individual time series. This issue will be examined in this chapter.

McCrae et al. (2002) investigated the performance of short-term and medium-term time horizons in terms of mean forecasting of univariate ARIMA model against the cointegration models. The univariate ARIMA model is the model which referred only to integration, while the cointegration models contain both integration and cointegration information. The authors compared the forecasting accuracy for a system of five cointegrated Asian exchange rate time series. The
multiple out-of-sample forecasts from one step-ahead to forty steps-ahead were carried out. Their results revealed that the ARIMA model forecasts performed more accurately for short-term horizons (up to five days) while cointegration models based on ECM performed better for medium-term time horizons (from six to forty days). McCrae et al. (2002) studied comparisons of the short-term and medium-term time horizons in mean forecasting performance between the Box-Jenkins type ARIMA model and cointegration-based ECM under the normal distribution assumption.

This thesis extends the idea of these authors’ work by comparing the performance in terms of the volatility forecasting between individual univariate GARCH(p,q) models and cointegration-based ECM by taking into account alternative non-normal distribution assumptions.

The objective of this chapter is to investigate the following two issues:

1. Whether information about the cointegration relationships among underlying financial time series can make a contribution to the accuracy of volatility forecasting?

2. According to the empirical results from Chapter 4, return series exhibit non-normal innovations. A large volume of research has shown that GARCH models with non-normal distributions can further improve the volatility forecasting performance. It is crucial to examine the improvement of volatility forecasting performance when the knowledge of cointegration is considered in the presence of the symmetric GARCH(1,1) model by taking into account the six different types of error distribution used in Chapter 4.

Chapter 5 shows that the Johansen tests yield the best performance for detect-
CHAPTER 6. IMPROVING VOLATILITY FORECASTING BASED ON COINTEGRATION INFORMATION

ing cointegration relationships of underlying time series. The Johansen approach was more powerful than the other three cointegration tests (DF, CRDW and WB tests). This chapter will employ the Johansen approach to examine the cointegration relationships among the underlying financial time series.

To investigate whether the knowledge of cointegration is beneficial to the improvement of volatility forecasting for financial time series, the Monte Carlo simulations are carried out. A comparison between the best fitting univariate GARCH(p,q) model and the cointegration based on Error Correction Model (ECM) is considered when evaluating their performance in volatility forecasting. This chapter also applies real observations from two emerging stock indices to compare the volatility forecasting performance between univariate GARCH models and cointegration based on error correction models.

6.2 Simulation Study for Evaluation of Volatility Forecasting: A Comparison between a GARCH Model and Cointegration based on ECM

Underlying financial time series can individually be fitted by a GARCH(p,q) models and the best fitting GARCH(p,q) models can be used to predict the volatility for each series. The performance of volatility forecasting can also be improved by assuming an alternative error distribution in a fitted GARCH(p,q) model.

This section focuses on how to further improve volatility forecasting of underlying financial time series by considering the cointegration relationships among time series. If the underlying financial time series can be individually fitted by GARCH(p,q) models and are also found to be cointegrated among these series, then the knowledge of cointegration relationships might be used to improve the
To compare the performance of volatility forecasting for these two approaches, the Monte Carlo simulations are conducted in the case of two non-stationary time series \( x_{1t} \) and \( x_{2t} \) by following the work of Lee and Tse (1996). The cointegrated system used in these simulations is the same system as used in Chapter 5 (see Equations (5.31) and (5.32). The cointegrated system is defined as follows:

\[
\Delta x_{1t} = -0.2(x_{1,t-1} - x_{2,t-1}) + e_{1t}, \quad (6.1)
\]
\[
\Delta x_{2t} = e_{2t}. \quad (6.2)
\]

Two time series \( x_{1t} \) and \( x_{2t} \) are generated from the cointegrated system in Equations (6.1) and (6.2). The error terms \( \{e_{1t}\} \) and \( \{e_{2t}\} \) are independently simulated by following a GARCH(1,1) model with different sets of parameters and different types of error distributions with their different parameters for shape and skewness.

After obtaining the two cointegrated time series \( x_{1t} \) and \( x_{2t} \), there are two possible ways to model the return volatility of these time series. One is using a GARCH\((p,q)\) model. The other is using ECM where the cointegrating errors follow a GARCH\((p,q)\) model.

Time series \( x_{1t} \) and \( x_{2t} \) are initially produced from ECM in Equations (6.1) and (6.2). In order to compare the performance of volatility forecasting between a univariate GARCH\((p,q)\) model and ECM, time series \( x_{1t} \) is fitted by using a GARCH\((p,q)\) model. The model in GARCH\((p,q)\) form is denoted as “Model A” and is defined as below:
The best fitting univariate GARCH(p,q) model for $x_{1t}$ is

$$\Delta x_{1t} = \mu + \varepsilon_t,$$

$$\varepsilon_t = \eta_t \sqrt{h_t},$$

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}.$$ (6.5)

The other model is ECM which is denoted as “Model B”. The cointegrating error terms in ECM are allowed to follow a GARCH(1,1) model. Model B is defined as below:

The ECM where the cointegrating errors $\{e_{1t}\}$ follow a GARCH(1,1) model is

$$\Delta x_{1t} = -0.2(x_{1,t-1} - x_{2,t-1}) + e_{1t},$$

$$e_{1t} = \eta_{1t} \sqrt{h_{1t}},$$

$$h_{1t} = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}.$$ (6.8)

Time series $x_{1t}$ is divided into two parts. The first part is called in-sample data in which the first 5,000 observations are used to build up a volatility forecasting model. The second part is called out-sample data in which the remaining 5,000 observations are used to investigate the performance of volatility forecasting (10,000 samples are simulated).

To carry out these simulation studies, four different sets of parameters applied to the GARCH(1,1) model from Equations (6.7) and (6.8); $(\omega = 0.1, \alpha = 0.2, \beta = 0.3),(\omega = 0.1, \alpha = 0.5, \beta = 0.2),(\omega = 0.1, \alpha = 0.3, \beta = 0.6)$ and $(\omega = 0.1, \alpha = 0.1, \beta = 0.8)$ are considered. The six alternative error distributions (N, SN, STD, SSTD, GED and SGED) are considered for the error terms in the GARCH(1,1) model. The skewness values of the SN, SSTD and SGED distributions are chosen as 0.1, 0.5 and 3 for each distribution. The degrees of freedom $(\nu)$ for STD are 5 and 8. The kurtosis of the STD was given by $3(\nu - 2)/(\nu - 4)$ for $\nu > 4$; it is
9 for $\nu = 5$ and is 4.5 for $\nu = 8$. The shape parameter for the skewed normal, generalized error and skewed generalized error distributions are equal to 3.

Combining the four different sets of parameters, the six different types of error distributions and their shape and skewness parameters into the GARCH(1,1) model, sixty four sets of data series are simulated so as to examine the performance of volatility forecasting. The 10,000 independent samples are simulated for sixty four data series in each Model (Model A and B).

Time series $x_{1t}$ is generated from the cointegrated system in (6.1) and (6.2): the true ECM for $x_{1t}$ is considered as a model which contains the cointegration information (Model B).

Model A refers to the best fitting univariate GARCH(p,q) model of time series $x_{1t}$. Sixty four data series of $x_{1t}$ generate from the cointegration system are individually fitted by a GARCH(p,q) model so as to identify the best fitting GARCH(p,q) model for each data series. To determine the order of the best fitting GARCH(p,q) models under the non-normal error distributions, the procedure described in Chapter 3 is used. The results of order determination for Model A are reported in Tables 6.1 - 6.4 using an AIC criterion. By regarding the minimum AIC values and significance of coefficient parameters of these GARCH(1,1) models, Tables 6.1 - 6.4 show that the order determinations for sixty four data sets are the GARCH(1,1) model.

To find the best fitting GARCH(1,1) model with different types of error distributions in Model A, the error terms of each GARCH(1,1) model are changed by crossing with the six competing types of error distributions (N, SN, STD, SSTD, GED and SGED). An appropriate error distribution in each GARCH(1,1) model is examined based on the smallest AIC values. Tables 6.5 - 6.8 show the AIC values of the sixty four data sets fitted by the GARCH(1,1) model with six different types of error distributions. The smallest AIC value is considered as the
criterion for the best fitting GARCH model in Model A.

The results in Tables 6.1 - 6.4 and Tables 6.5 - 6.8 demonstrate the sixty four data sets \( \{ x_{1t} \} \) which are generated from Equations (6.3) to (6.5) with the six alternative error distributions. The GARCH error terms \( \{ \eta_t \} \) in Equation (6.4) can be individually fitted well by the GARCH(1,1) model with one of the six types of error distributions. The best fitting GARCH(1,1) models of each of the sixty four models (Model A) are used to compare the volatility forecasting performance with the data set from the true ECM (Model B).

The out-of-sample volatility forecasting performance between the two different models are used to examine whether the forecasting can be further improved when the knowledge of cointegration relationships is taken into account. To evaluate the volatility forecasting performance between Models A and B, Mean Square Error (MSE) in Equation (3.54) and Mean Absolute Error (MAE) in Equation (3.55) are employed.

### 6.3 Simulation Results

The measurement of volatility forecasting performance between Models A and B, using the MSE(k) and MAE(k) based on out-of-sample data, are reported in Tables 6.9 - 6.16. All the simulation results show that the models which contain the cointegrated information (Model B) produce the smaller values of MSE and MAE compared with the best fitting univariate GARCH model (Model A).

Tables 6.9 - 6.12 show the volatility forecasting performance based on MSE criterion. Considering the performances of volatility forecasting in the same set parameter of each GARCH model, Model B with GARCH-SSTD model outperforms other alternative error distributions for short forecasts (1- and 2- steps ahead) when regarding the smallest MSE values. There is only one case of Model B (GARCH(1,1) with the parameter \((0.1,0.2,0.3)\)) when the GARCH-GED mod-
Table 6.1: AIC values for specifying the order of the GARCH(p,q) models used to fit $x_{1t}$, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1$, $\alpha = 0.2$, $\beta = 0.3$

<table>
<thead>
<tr>
<th>Competing GARCH(p,q) models of Model A</th>
<th>The true probability distribution of $\eta_{ht}$ in the ECM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>1.3824</td>
</tr>
<tr>
<td>GARCH(1,2)</td>
<td>1.3827</td>
</tr>
<tr>
<td>GARCH(1,3)</td>
<td>1.3826</td>
</tr>
<tr>
<td>GARCH(2,1)</td>
<td>1.3828</td>
</tr>
<tr>
<td>GARCH(2,2)</td>
<td>1.3826</td>
</tr>
<tr>
<td>Std</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>1.2692</td>
</tr>
<tr>
<td>GARCH(1,2)</td>
<td>1.2697</td>
</tr>
<tr>
<td>GARCH(1,3)</td>
<td>1.2700</td>
</tr>
<tr>
<td>GARCH(2,1)</td>
<td>1.2696</td>
</tr>
<tr>
<td>GARCH(2,2)</td>
<td>1.2698</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best GARCH(p,q) model. All coefficient of parameter in GARCH(1,1) models for all types of error distributions are significant.
CHAPTER 6. IMPROVING VOLATILITY FORECASTING BASED ON COINTEGRATION INFORMATION

### Table 6.2: AIC values for specifying the order of the GARCH(p,q) models used to fit $x_t$, where $x_t$ was simulated from Equation (6.6) with $\omega = 0$, $\alpha = 0.5$, $\beta = 0.2$

<table>
<thead>
<tr>
<th>Shape</th>
<th>Skew</th>
<th>4.5</th>
<th>9</th>
<th>4.5</th>
<th>9</th>
<th>4.5</th>
<th>9</th>
<th>4.5</th>
<th>9</th>
<th>4.5</th>
<th>9</th>
<th>4.5</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>STD</td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>1.4895</td>
<td>1.6283</td>
<td>1.1419</td>
<td>1.2910</td>
<td>1.1848</td>
<td>1.4588</td>
<td>1.5434</td>
<td>1.4671</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(1,2)</td>
<td>1.4899</td>
<td>1.6285</td>
<td>1.1425</td>
<td>1.2914</td>
<td>1.1854</td>
<td>1.4589</td>
<td>1.5438</td>
<td>1.4699</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(1,3)</td>
<td>1.4903</td>
<td>1.6284</td>
<td>1.1431</td>
<td>1.2921</td>
<td>1.1859</td>
<td>1.4594</td>
<td>1.5442</td>
<td>1.4674</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(2,1)</td>
<td>1.4899</td>
<td>1.6288</td>
<td>1.1421</td>
<td>1.2915</td>
<td>1.1853</td>
<td>1.4595</td>
<td>1.5440</td>
<td>1.4677</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(2,2)</td>
<td>1.4902</td>
<td>1.6287</td>
<td>1.1425</td>
<td>1.2918</td>
<td>1.1857</td>
<td>1.4592</td>
<td>1.5442</td>
<td>1.4672</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The true probability distribution of $\eta_t$ in the ECM models is $N(0, \Sigma)$. For all types of error distributions, the highlighted cells in the table indicate the best GARCH(p,q) model. All coefficient of parameters in GARCH(1,1) models are significant.
Table 6.3: AIC values for specifying the order of the GARCH(p,q) models used to fit $x_{1t}$, where $x_{1t}$ was simulated from Equation (6.6) with $\omega = 0.1$, $\alpha = 0.3$, $\beta = 0.6$

<table>
<thead>
<tr>
<th>Competing GARCH(p,q) models of Model A</th>
<th>The true probability distribution of $\eta_{1t}$ in the ECM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Shape</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>2.7579</td>
</tr>
<tr>
<td>GARCH(1,2)</td>
<td>2.7580</td>
</tr>
<tr>
<td>GARCH(1,3)</td>
<td>2.7580</td>
</tr>
<tr>
<td>GARCH(2,1)</td>
<td>2.7583</td>
</tr>
<tr>
<td>GARCH(2,2)</td>
<td>2.7585</td>
</tr>
<tr>
<td>STD</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>2.5264</td>
</tr>
<tr>
<td>GARCH(1,2)</td>
<td>2.5269</td>
</tr>
<tr>
<td>GARCH(1,3)</td>
<td>2.5273</td>
</tr>
<tr>
<td>GARCH(2,1)</td>
<td>2.5266</td>
</tr>
<tr>
<td>GARCH(2,2)</td>
<td>2.5269</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best GARCH(p,q) model. All coefficient of parameter in GARCH(1,1) models for all types of error distributions are significant.
CHAPTER 6. IMPROVING VOLATILITY FORECASTING BASED ON COINTEGRATION INFORMATION

Table 6.4: AIC values for specifying the order of the GARCH(p,q) models used to fit $x_t$, where $x_t$ was simulated from Equation (6.6) with $\omega = 0$, $\alpha = 0.5$, $\beta = 0.8$.

<table>
<thead>
<tr>
<th>Shape</th>
<th>$\text{Shape}$</th>
<th>$\text{Skew}$</th>
<th>$\text{GARCH}(1,1)$</th>
<th>$\text{GARCH}(1,2)$</th>
<th>$\text{GARCH}(1,3)$</th>
<th>$\text{GARCH}(2,1)$</th>
<th>$\text{GARCH}(2,2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.9662 2.9666</td>
<td>2.9652 2.9656</td>
<td>2.9652 2.9656</td>
<td>2.9651 2.9651</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
</tr>
<tr>
<td>0.5</td>
<td>2.9662 2.9666</td>
<td>2.9652 2.9656</td>
<td>2.9652 2.9656</td>
<td>2.9651 2.9651</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
</tr>
<tr>
<td>0.1</td>
<td>2.9662 2.9666</td>
<td>2.9652 2.9656</td>
<td>2.9652 2.9656</td>
<td>2.9651 2.9651</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
</tr>
<tr>
<td>0</td>
<td>2.9662 2.9666</td>
<td>2.9652 2.9656</td>
<td>2.9652 2.9656</td>
<td>2.9651 2.9651</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
<td>2.9652 2.9652</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the best GARCH(p,q) model. All coefficients of parameter in GARCH(1,1) models for all types of error distributions are significant.
Table 6.5: AIC values for identifying appropriate error distributions in Model A, where $x_{lt}$ was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.2, \beta = 0.3$

<table>
<thead>
<tr>
<th>Competing error distributions in GARCH(1,1) model of Model A</th>
<th>The true probability distribution of $\eta_{lt}$ in the ECM</th>
<th>N</th>
<th>SN</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>N</td>
<td>1.3824</td>
<td>1.3934</td>
<td>1.3977</td>
<td>1.4172</td>
<td>1.4454</td>
</tr>
<tr>
<td>SN</td>
<td>1.3828</td>
<td>1.3209</td>
<td>1.3289</td>
<td>1.4171</td>
<td>1.3882</td>
</tr>
<tr>
<td>STD</td>
<td>1.3903</td>
<td>1.3897</td>
<td>1.3669</td>
<td>1.4359</td>
<td>1.4568</td>
</tr>
<tr>
<td>SSTD</td>
<td>1.3907</td>
<td>1.3257</td>
<td>1.2802</td>
<td>1.4359</td>
<td>1.4010</td>
</tr>
<tr>
<td>GED</td>
<td>1.3825</td>
<td>1.3919</td>
<td>1.3732</td>
<td>1.4155</td>
<td>1.4452</td>
</tr>
<tr>
<td>SGED</td>
<td>1.3828</td>
<td>1.3202</td>
<td>1.2764</td>
<td>1.4158</td>
<td>1.3886</td>
</tr>
<tr>
<td>STD</td>
<td>1.2692</td>
<td>1.3548</td>
<td>0.9601</td>
<td>1.0819</td>
<td>0.9971</td>
</tr>
<tr>
<td>SSTD</td>
<td>1.2694</td>
<td>1.3548</td>
<td>0.9601</td>
<td>1.0819</td>
<td>0.9971</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the smallest AIC values from among six competing error distributions in the same column.
## AIC Values for Identifying Appropriate Error Distributions in Model A

The true probability distribution of $\eta_{1t}$ in the ECM in GARCH(1,1) model of Model A was simulated from Equation (6.6) with $\omega = 0.1, \alpha = 0.05, \beta = 0.2$.

### Competing Error Distributions

<table>
<thead>
<tr>
<th>Shape</th>
<th>SN</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.7070</td>
<td>1.7085</td>
<td>1.7084</td>
<td>1.7087</td>
</tr>
<tr>
<td>0.1</td>
<td>1.7146</td>
<td>1.6429</td>
<td>1.7088</td>
<td>1.7089</td>
</tr>
<tr>
<td>0.5</td>
<td>1.6989</td>
<td>1.5826</td>
<td>1.6932</td>
<td>1.6937</td>
</tr>
<tr>
<td>3</td>
<td>1.8190</td>
<td>1.7656</td>
<td>1.8142</td>
<td>1.8146</td>
</tr>
<tr>
<td>0.1</td>
<td>1.7414</td>
<td>1.6435</td>
<td>1.7440</td>
<td>1.7445</td>
</tr>
<tr>
<td>0.5</td>
<td>1.7731</td>
<td>1.6435</td>
<td>1.7716</td>
<td>1.7721</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the smallest AIC values from among six competing error distributions in the same column.
Table 6.7: AIC values for identifying appropriate error distributions in Model A, where \( x_{1t} \) was simulated from Equation (6.6) with \( (\omega = 0.1, \alpha = 0.3, \beta = 0.6) \)

<table>
<thead>
<tr>
<th>Competing error distributions in GARCH(1,1) model of Model A</th>
<th>The true probability distribution of ( \eta_{1t} ) in the ECM</th>
</tr>
</thead>
<tbody>
<tr>
<td>N ( \quad ) SN ( \quad ) GED ( \quad ) SGED ( \quad )</td>
<td>N ( \quad ) SN ( \quad ) GED ( \quad ) SGED ( \quad )</td>
</tr>
<tr>
<td>Shape ( \quad ) 3 ( \quad ) 0.1 ( \quad ) 0.5 ( \quad ) 3 ( \quad ) 0.1 ( \quad ) 0.5 ( \quad ) 3 ( \quad )</td>
<td></td>
</tr>
<tr>
<td>Skew ( \quad ) 2.7579 ( \quad ) 2.7500 ( \quad ) 2.7526 ( \quad ) 2.7234 ( \quad ) 2.8558 ( \quad ) 2.8678 ( \quad ) 2.8209 ( \quad ) 2.8767 ( \quad )</td>
<td></td>
</tr>
<tr>
<td>SN ( \quad ) 2.7582 ( \quad ) 2.6387 ( \quad ) 2.6876 ( \quad ) 2.6321 ( \quad ) 2.8556 ( \quad ) 2.8169 ( \quad ) 2.7958 ( \quad ) 2.8431 ( \quad )</td>
<td></td>
</tr>
<tr>
<td>STD ( \quad ) 2.7574 ( \quad ) 2.7318 ( \quad ) 2.7439 ( \quad ) 2.7058 ( \quad ) 2.8672 ( \quad ) 2.8720 ( \quad ) 2.8284 ( \quad ) 2.8806 ( \quad )</td>
<td></td>
</tr>
<tr>
<td>SSTD ( \quad ) 2.7578 ( \quad ) 2.6307 ( \quad ) 2.6853 ( \quad ) 2.6266 ( \quad ) 2.8673 ( \quad ) 2.8225 ( \quad ) 2.8033 ( \quad ) 2.8405 ( \quad )</td>
<td></td>
</tr>
<tr>
<td>GED ( \quad ) 2.7552 ( \quad ) 2.7411 ( \quad ) 2.7490 ( \quad ) 2.7157 ( \quad ) 2.8556 ( \quad ) 2.8681 ( \quad ) 2.8213 ( \quad ) 2.8769 ( \quad )</td>
<td></td>
</tr>
<tr>
<td>SGED ( \quad ) 2.7555 ( \quad ) 2.6294 ( \quad ) 2.6842 ( \quad ) 2.6258 ( \quad ) 2.8557 ( \quad ) 2.8167 ( \quad ) 2.7952 ( \quad ) 2.8337 ( \quad )</td>
<td></td>
</tr>
</tbody>
</table>

| STD \( \quad \) SSTD \( \quad \) | 4.5 \( \quad \) 9 \( \quad \) 4.5 \( \quad \) 9 \( \quad \) | 0.1 \( \quad \) 0.5 \( \quad \) 3 \( \quad \) 0.1 \( \quad \) 0.5 \( \quad \) 3 \( \quad \) |
| Shape \( \quad \) 2.5822 \( \quad \) 2.6975 \( \quad \) 2.4486 \( \quad \) 2.5046 \( \quad \) 2.4195 \( \quad \) 2.6495 \( \quad \) 2.6776 \( \quad \) 2.6264 \( \quad \) |
| Skew \( \quad \) 2.5821 \( \quad \) 2.6971 \( \quad \) 2.2091 \( \quad \) 2.3640 \( \quad \) 2.2247 \( \quad \) 2.5020 \( \quad \) 2.5858 \( \quad \) 2.5044 \( \quad \) |
| STD \( \quad \) 2.5264 \( \quad \) 2.6774 \( \quad \) 2.2492 \( \quad \) 2.3546 \( \quad \) 2.2702 \( \quad \) 2.5897 \( \quad \) 2.6357 \( \quad \) 2.5799 \( \quad \) |
| SSTD \( \quad \) 2.5265 \( \quad \) 2.6775 \( \quad \) 2.1042 \( \quad \) 2.2711 \( \quad \) 2.1480 \( \quad \) 2.4725 \( \quad \) 2.5642 \( \quad \) 2.4805 \( \quad \) |
| GED \( \quad \) 2.5354 \( \quad \) 2.6812 \( \quad \) 2.2040 \( \quad \) 2.2853 \( \quad \) 2.1681 \( \quad \) 2.6117 \( \quad \) 2.6491 \( \quad \) 2.5965 \( \quad \) |
| SGED \( \quad \) 2.5351 \( \quad \) 2.6813 \( \quad \) 2.2043 \( \quad \) 2.2867 \( \quad \) 2.1719 \( \quad \) 2.4765 \( \quad \) 2.5661 \( \quad \) 2.4845 \( \quad \) |

Note: Highlighted cells in the table indicate the smallest AIC values from among six competing error distributions in the same column.
CHAPTER 6. IMPROVING VOLATILITY FORECASTING BASED ON COINTEGRATION INFORMATION

Table 6.8: AIC values for identifying appropriate error distributions in Model A, where $x_t$ was simulated from Equation (6.6) with $\alpha = 0.1, \omega = 0.1, \beta = 0.8$.

<table>
<thead>
<tr>
<th>Shape</th>
<th>GED</th>
<th>SGED</th>
<th>STD</th>
<th>SSTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the smallest AIC values from among six competing error distributions in the same column.
eled yield better results than the GARCH-SSTD model (See Table 6.10). For long forecasts (10- and 15-step ahead), Model B with GARCH-GED provides the best performance on volatility forecasting. Tables 6.13 - 6.16 show the same tendency of the volatility forecasting performance of Model B for short and long horizontal forecasts based on MAE criterion.

Considering the smallest values of MSE and MAE with the same GARCH parameters setting in Tables 6.9 - 6.12 and Tables 6.13 - 6.16 respectively, it can be seen that the short horizontal forecasts tend to provide better accuracy than the long horizons (see the highlighted cells in Tables 6.9 - 6.16).

To confirm whether these conclusions are still valid for real life data, empirical studies will be demonstrated in the next section.
Table 6.9: The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 1-step ahead forecasts

<table>
<thead>
<tr>
<th>$\left(\omega, \alpha, \beta\right)$</th>
<th>1-step ahead forecasts</th>
<th>Shape</th>
<th>SN</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\left(0.1, 0.2, 0.3\right)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.0795</td>
<td>0.1558</td>
<td>0.1334</td>
<td>0.0802</td>
<td>0.2008</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0382</td>
<td>0.1148</td>
<td>0.0963</td>
<td>0.0560</td>
<td>0.0938</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.1555</td>
<td>0.4822</td>
<td>0.0521</td>
<td>0.0608</td>
<td>0.0573</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1431</td>
<td>0.3853</td>
<td>0.0292</td>
<td>0.0340</td>
<td>0.0317</td>
</tr>
<tr>
<td>$\left(0.1, 0.5, 0.2\right)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.1544</td>
<td>0.3038</td>
<td>0.2717</td>
<td>0.1789</td>
<td>0.4813</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0757</td>
<td>0.2085</td>
<td>0.1699</td>
<td>0.1050</td>
<td>0.3276</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2557</td>
<td>1.7672</td>
<td>0.0619</td>
<td>0.0873</td>
<td>0.0680</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2450</td>
<td>1.5570</td>
<td>0.0389</td>
<td>0.0531</td>
<td>0.0442</td>
</tr>
<tr>
<td>$\left(0.1, 0.3, 0.6\right)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.8156</td>
<td>1.6607</td>
<td>1.7850</td>
<td>0.9190</td>
<td>0.8549</td>
</tr>
<tr>
<td>Model B</td>
<td>0.4290</td>
<td>1.3253</td>
<td>1.2263</td>
<td>0.6834</td>
<td>0.7101</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.3087</td>
<td>2.1826</td>
<td>0.2435</td>
<td>0.3850</td>
<td>0.2740</td>
</tr>
<tr>
<td>Model B</td>
<td>1.1827</td>
<td>1.6391</td>
<td>0.1573</td>
<td>0.2328</td>
<td>0.1774</td>
</tr>
<tr>
<td>$\left(0.1, 0.1, 0.8\right)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.8509</td>
<td>0.9755</td>
<td>0.9908</td>
<td>1.6587</td>
<td>1.8941</td>
</tr>
<tr>
<td>Model B</td>
<td>0.8383</td>
<td>0.9275</td>
<td>0.9476</td>
<td>1.2690</td>
<td>1.8460</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.8119</td>
<td>1.0321</td>
<td>0.7298</td>
<td>0.9478</td>
<td>0.8300</td>
</tr>
<tr>
<td>Model B</td>
<td>1.5880</td>
<td>1.0128</td>
<td>0.4402</td>
<td>0.5526</td>
<td>0.4810</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MSE value for the same parameter of each GARCH model.
Table 6.10: The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 2-step ahead forecasts

<table>
<thead>
<tr>
<th>$(\omega, \alpha, \beta)$</th>
<th>2-step ahead forecasts</th>
<th>Shape</th>
<th>SN</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0.1, 0.2, 0.3)$</td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.0733</td>
<td>0.1589</td>
<td>0.1481</td>
<td>0.0841</td>
<td>0.0556</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0370</td>
<td>0.1259</td>
<td>0.1015</td>
<td>0.0557</td>
<td>0.0323</td>
</tr>
<tr>
<td>STD</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.1477</td>
<td>0.4712</td>
<td>0.0565</td>
<td>0.0640</td>
<td>0.0605</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1331</td>
<td>0.3456</td>
<td>0.0328</td>
<td>0.0367</td>
<td>0.0353</td>
</tr>
<tr>
<td>STD</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.1456</td>
<td>0.3306</td>
<td>0.3294</td>
<td>0.1863</td>
<td>0.3887</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0703</td>
<td>0.2651</td>
<td>0.1905</td>
<td>0.1057</td>
<td>0.2220</td>
</tr>
<tr>
<td>STD</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.2427</td>
<td>1.8124</td>
<td>0.0732</td>
<td>0.0946</td>
<td>0.0755</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2298</td>
<td>1.2860</td>
<td>0.0487</td>
<td>0.0594</td>
<td>0.0532</td>
</tr>
<tr>
<td>STD</td>
<td>1.1920</td>
<td>1.8536</td>
<td>1.0282</td>
<td>1.0416</td>
<td>1.0234</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.8124</td>
<td>1.7325</td>
<td>2.1756</td>
<td>0.9620</td>
<td>0.6718</td>
</tr>
<tr>
<td>Model B</td>
<td>0.4230</td>
<td>1.5262</td>
<td>1.3986</td>
<td>0.6960</td>
<td>0.6133</td>
</tr>
<tr>
<td>STD</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>1.2840</td>
<td>1.6147</td>
<td>0.2749</td>
<td>0.4080</td>
<td>0.2941</td>
</tr>
<tr>
<td>Model B</td>
<td>1.1442</td>
<td>1.5060</td>
<td>0.1844</td>
<td>0.2527</td>
<td>0.2031</td>
</tr>
<tr>
<td>STD</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>9.7868</td>
<td>0.9331</td>
<td>0.9653</td>
<td>1.7410</td>
<td>1.4691</td>
</tr>
<tr>
<td>Model B</td>
<td>0.8278</td>
<td>0.9109</td>
<td>0.9252</td>
<td>1.2820</td>
<td>0.9820</td>
</tr>
<tr>
<td>STD</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>2.7802</td>
<td>1.3679</td>
<td>0.7648</td>
<td>0.9710</td>
<td>0.8513</td>
</tr>
<tr>
<td>Model B</td>
<td>2.4707</td>
<td>0.9935</td>
<td>0.4626</td>
<td>0.5677</td>
<td>0.5030</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MSE value for the same parameter of each GARCH model.
Table 6.11: The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 10-step ahead forecasts

<table>
<thead>
<tr>
<th>$(\omega, \alpha, \beta)$</th>
<th>10-step ahead forecasts</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td><strong>Shape</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.0714</td>
<td>0.2080</td>
<td>0.2158</td>
<td>0.1333</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0472</td>
<td>0.2032</td>
<td>0.1698</td>
<td>0.0970</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.1456</td>
<td>0.4606</td>
<td>0.0868</td>
<td>0.0875</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1329</td>
<td>0.3368</td>
<td>0.0553</td>
<td>0.0601</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.1241</td>
<td>0.4534</td>
<td>0.5993</td>
<td>0.5336</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0916</td>
<td>0.3168</td>
<td>0.3516</td>
<td>0.2819</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2362</td>
<td>2.2967</td>
<td>0.1666</td>
<td>0.1586</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2261</td>
<td>1.3308</td>
<td>0.1029</td>
<td>0.1067</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.7573</td>
<td>0.7932</td>
<td>0.6213</td>
<td>0.6089</td>
</tr>
<tr>
<td>Model B</td>
<td>0.5328</td>
<td>0.6159</td>
<td>0.6007</td>
<td>0.5010</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.2766</td>
<td>0.8818</td>
<td>0.8723</td>
<td>0.7938</td>
</tr>
<tr>
<td>Model B</td>
<td>1.1417</td>
<td>0.7902</td>
<td>0.5922</td>
<td>0.5092</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.0566</td>
<td>1.4351</td>
<td>1.5184</td>
<td>1.6961</td>
</tr>
<tr>
<td>Model B</td>
<td>0.9685</td>
<td>1.3822</td>
<td>1.1187</td>
<td>1.5556</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.7117</td>
<td>1.9744</td>
<td>0.9429</td>
<td>1.4374</td>
</tr>
<tr>
<td>Model B</td>
<td>1.6389</td>
<td>1.9279</td>
<td>0.7271</td>
<td>0.9894</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MSE value for the same parameter of each GARCH model.
Table 6.12: The comparison of volatility forecasting performance between Model A and Model B evaluated by MSE for 15-step ahead forecasts

<table>
<thead>
<tr>
<th>$(\omega, \alpha, \beta)$</th>
<th>15-step ahead forecasts</th>
<th>(N)</th>
<th>(SN)</th>
<th>(GED)</th>
<th>(SGED)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.1, 0.2, 0.3)</td>
<td></td>
<td>(N)</td>
<td>(SN)</td>
<td>(GED)</td>
<td>(SGED)</td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>0.0843</td>
<td>0.1933</td>
<td>0.2105</td>
<td>0.1224</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.0473</td>
<td>0.1901</td>
<td>0.1571</td>
<td>0.0842</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>0.0923</td>
<td>0.4588</td>
<td>0.0786</td>
<td>0.0873</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.0875</td>
<td>0.3363</td>
<td>0.0557</td>
<td>0.0608</td>
</tr>
<tr>
<td>(0.1, 0.5, 0.2)</td>
<td></td>
<td>(N)</td>
<td>(SN)</td>
<td>(GED)</td>
<td>(SGED)</td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>0.1425</td>
<td>0.4369</td>
<td>0.5900</td>
<td>0.5312</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.0929</td>
<td>0.3012</td>
<td>0.3095</td>
<td>0.2760</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>0.1772</td>
<td>2.3416</td>
<td>0.1382</td>
<td>0.1646</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.1732</td>
<td>1.3468</td>
<td>0.1070</td>
<td>0.1116</td>
</tr>
<tr>
<td>(0.1, 0.3, 0.6)</td>
<td></td>
<td>(N)</td>
<td>(SN)</td>
<td>(GED)</td>
<td>(SGED)</td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>0.9739</td>
<td>0.8529</td>
<td>0.6340</td>
<td>0.6128</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.5835</td>
<td>0.7321</td>
<td>0.6081</td>
<td>0.4059</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>1.0322</td>
<td>0.8611</td>
<td>0.8143</td>
<td>0.8373</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.9241</td>
<td>0.7150</td>
<td>0.6075</td>
<td>0.6055</td>
</tr>
<tr>
<td>(0.1, 0.1, 0.8)</td>
<td></td>
<td>(N)</td>
<td>(SN)</td>
<td>(GED)</td>
<td>(SGED)</td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>1.0513</td>
<td>1.3314</td>
<td>1.1580</td>
<td>1.8444</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>0.9720</td>
<td>1.3500</td>
<td>1.1386</td>
<td>1.7262</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td></td>
<td>1.7946</td>
<td>1.9572</td>
<td>1.2787</td>
<td>1.4508</td>
</tr>
<tr>
<td>Model B</td>
<td></td>
<td>1.6520</td>
<td>1.9225</td>
<td>0.9448</td>
<td>1.0389</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MSE value for the same parameter of each GARCH model.
CHAPTER 6. IMPROVING VOLATILITY FORECASTING BASED ON COINTEGRATION INFORMATION

Table 6.13: The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 1-step ahead

<table>
<thead>
<tr>
<th>$(\omega, \alpha, \beta)$</th>
<th>1-step ahead forecasts</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.1, 0.2, 0.3)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2185</td>
<td>0.2301</td>
<td>0.2312</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1476</td>
<td>0.1923</td>
<td>0.1873</td>
</tr>
<tr>
<td>STD</td>
<td>SSTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2455</td>
<td>0.3314</td>
<td>0.1578</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2320</td>
<td>0.3071</td>
<td>0.1244</td>
</tr>
<tr>
<td>(0.1, 0.5, 0.2)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2961</td>
<td>0.3211</td>
<td>0.3322</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1945</td>
<td>0.2663</td>
<td>0.2605</td>
</tr>
<tr>
<td>STD</td>
<td>SSTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.3213</td>
<td>0.6317</td>
<td>0.1780</td>
</tr>
<tr>
<td>Model B</td>
<td>0.3071</td>
<td>0.6106</td>
<td>0.1368</td>
</tr>
<tr>
<td>(0.1, 0.3, 0.6)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.7002</td>
<td>0.8255</td>
<td>0.9011</td>
</tr>
<tr>
<td>Model B</td>
<td>0.4986</td>
<td>0.7143</td>
<td>0.7366</td>
</tr>
<tr>
<td>STD</td>
<td>SSTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.7613</td>
<td>1.8890</td>
<td>0.3607</td>
</tr>
<tr>
<td>Model B</td>
<td>0.7362</td>
<td>1.8438</td>
<td>0.2984</td>
</tr>
<tr>
<td>(0.1, 0.1, 0.8)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>0.9668</td>
<td>1.0460</td>
<td>1.0668</td>
</tr>
<tr>
<td>Model B</td>
<td>0.7067</td>
<td>0.9275</td>
<td>0.9697</td>
</tr>
<tr>
<td>STD</td>
<td>SSTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Model A</td>
<td>1.0888</td>
<td>1.8487</td>
<td>0.6033</td>
</tr>
<tr>
<td>Model B</td>
<td>1.0578</td>
<td>1.7928</td>
<td>0.5096</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MAE value for the same parameter of each GARCH model.
Table 6.14: The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 2-step ahead forecasts.

<table>
<thead>
<tr>
<th>$(\omega, \alpha, \beta)$</th>
<th>N</th>
<th>SN</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shape</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.2195</td>
<td>0.2375</td>
<td>0.2415</td>
<td>0.2116</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1468</td>
<td>0.2025</td>
<td>0.1911</td>
<td>0.1727</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.2401</td>
<td>0.3244</td>
<td>0.1755</td>
<td>0.1777</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2254</td>
<td>0.2785</td>
<td>0.1411</td>
<td>0.1393</td>
</tr>
<tr>
<td><strong>(0.1, 0.5, 0.2)</strong></td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td><strong>Shape</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.3047</td>
<td>0.3542</td>
<td>0.3548</td>
<td>0.2881</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1978</td>
<td>0.2938</td>
<td>0.2759</td>
<td>0.2382</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.3234</td>
<td>0.6124</td>
<td>0.2199</td>
<td>0.2294</td>
</tr>
<tr>
<td>Model B</td>
<td>0.3105</td>
<td>0.5352</td>
<td>0.1806</td>
<td>0.1824</td>
</tr>
<tr>
<td><strong>(0.1, 0.3, 0.6)</strong></td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td><strong>Shape</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.7117</td>
<td>0.8610</td>
<td>0.9673</td>
<td>0.6922</td>
</tr>
<tr>
<td>Model B</td>
<td>0.5038</td>
<td>0.7623</td>
<td>0.7777</td>
<td>0.5869</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.7703</td>
<td>1.8075</td>
<td>0.4133</td>
<td>0.4761</td>
</tr>
<tr>
<td>Model B</td>
<td>0.7421</td>
<td>1.7050</td>
<td>0.3513</td>
<td>0.3788</td>
</tr>
<tr>
<td><strong>(0.1, 0.1, 0.8)</strong></td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td><strong>Shape</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>0.9706</td>
<td>1.2029</td>
<td>1.2132</td>
<td>0.9378</td>
</tr>
<tr>
<td>Model B</td>
<td>0.7020</td>
<td>1.0974</td>
<td>0.9880</td>
<td>0.8004</td>
</tr>
<tr>
<td><strong>STD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Model A</td>
<td>1.0724</td>
<td>1.7688</td>
<td>0.6422</td>
<td>0.6872</td>
</tr>
<tr>
<td>Model B</td>
<td>1.0329</td>
<td>1.6828</td>
<td>0.5363</td>
<td>0.5559</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MAE value for the same parameter of each GARCH model.
Table 6.15: The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 10-step ahead forecasts

<table>
<thead>
<tr>
<th>(ω, α, β)</th>
<th>10-step ahead forecasts</th>
<th>Shape</th>
<th>SN</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.1, 0.2, 0.3)</td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2138</td>
<td>0.2786</td>
<td>0.2862</td>
<td>0.2434</td>
<td>0.1701</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1624</td>
<td>0.2451</td>
<td>0.2328</td>
<td>0.2028</td>
<td>0.1350</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2400</td>
<td>0.3218</td>
<td>0.2445</td>
<td>0.2226</td>
<td>0.2308</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2267</td>
<td>0.2609</td>
<td>0.1732</td>
<td>0.1674</td>
<td>0.1746</td>
</tr>
<tr>
<td>(0.1, 0.5, 0.2)</td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.3164</td>
<td>0.8778</td>
<td>0.8834</td>
<td>0.5405</td>
<td>0.2858</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2487</td>
<td>0.8439</td>
<td>0.7546</td>
<td>0.4782</td>
<td>0.2307</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.3493</td>
<td>0.6811</td>
<td>0.3883</td>
<td>0.3516</td>
<td>0.3543</td>
</tr>
<tr>
<td>Model B</td>
<td>0.3417</td>
<td>0.5074</td>
<td>0.2917</td>
<td>0.2813</td>
<td>0.2903</td>
</tr>
<tr>
<td>(0.1, 0.3, 0.6)</td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.7636</td>
<td>1.4868</td>
<td>1.8845</td>
<td>1.5158</td>
<td>0.7332</td>
</tr>
<tr>
<td>Model B</td>
<td>0.6247</td>
<td>1.4044</td>
<td>1.5749</td>
<td>1.0459</td>
<td>0.5904</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.8638</td>
<td>1.7179</td>
<td>0.8957</td>
<td>0.8041</td>
<td>0.8482</td>
</tr>
<tr>
<td>Model B</td>
<td>0.8406</td>
<td>1.4307</td>
<td>0.6568</td>
<td>0.6353</td>
<td>0.6576</td>
</tr>
<tr>
<td>(0.1, 0.1, 0.8)</td>
<td>N</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>0.9518</td>
<td>1.2441</td>
<td>1.4304</td>
<td>1.1134</td>
<td>0.7954</td>
</tr>
<tr>
<td>Model B</td>
<td>0.7566</td>
<td>1.1412</td>
<td>1.1872</td>
<td>0.9482</td>
<td>0.6435</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSTD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Model A</td>
<td>1.0916</td>
<td>1.6348</td>
<td>1.0402</td>
<td>0.9403</td>
<td>0.9852</td>
</tr>
<tr>
<td>Model B</td>
<td>1.0308</td>
<td>1.4501</td>
<td>0.9271</td>
<td>0.7114</td>
<td>0.7369</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MAE value for the same parameter of each GARCH model.
Table 6.16: The comparison of volatility forecasting performance between Model A and Model B evaluated by MAE for 15-step ahead forecasts

<table>
<thead>
<tr>
<th>(ω, α, β)</th>
<th>N</th>
<th>SN</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.1, 0.2, 0.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2222</td>
<td>0.2610</td>
<td>0.2808</td>
<td>0.1739</td>
</tr>
<tr>
<td>Model B</td>
<td>0.1626</td>
<td>0.2365</td>
<td>0.2247</td>
<td>0.1937</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.2150</td>
<td>0.3148</td>
<td>0.2179</td>
<td>0.1997</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2028</td>
<td>0.2600</td>
<td>0.1745</td>
<td>0.1756</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.5, 0.2)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.3230</td>
<td>0.8839</td>
<td>0.8921</td>
<td>0.5391</td>
</tr>
<tr>
<td>Model B</td>
<td>0.2502</td>
<td>0.8488</td>
<td>0.7586</td>
<td>0.4785</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.3480</td>
<td>0.6830</td>
<td>0.3398</td>
<td>0.3555</td>
</tr>
<tr>
<td>Model B</td>
<td>0.3187</td>
<td>0.5170</td>
<td>0.2972</td>
<td>0.2801</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.3, 0.6)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.8380</td>
<td>1.5160</td>
<td>1.9880</td>
<td>1.1281</td>
</tr>
<tr>
<td>Model B</td>
<td>0.6730</td>
<td>1.4840</td>
<td>1.6285</td>
<td>1.0434</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>0.9014</td>
<td>1.7245</td>
<td>0.8660</td>
<td>0.8619</td>
</tr>
<tr>
<td>Model B</td>
<td>0.8230</td>
<td>1.4018</td>
<td>0.7428</td>
<td>0.7106</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.1, 0.8)</td>
<td>N</td>
<td>SN</td>
<td>GED</td>
<td>SGED</td>
</tr>
<tr>
<td>Shape</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.0161</td>
<td>1.2029</td>
<td>1.4241</td>
<td>1.0563</td>
</tr>
<tr>
<td>Model B</td>
<td>0.7769</td>
<td>1.0974</td>
<td>1.1542</td>
<td>0.9067</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>4.5</td>
<td>9</td>
<td>4.5</td>
<td>9</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1</td>
<td>0.5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Model A</td>
<td>1.0053</td>
<td>1.5820</td>
<td>0.9568</td>
<td>0.9515</td>
</tr>
<tr>
<td>Model B</td>
<td>0.9471</td>
<td>1.4075</td>
<td>0.7805</td>
<td>0.7530</td>
</tr>
</tbody>
</table>

Note: Highlighted cells in the table indicate the minimum MAE value for the same parameter of each GARCH model.
CHAPTER 6. IMPROVING VOLATILITY FORECASTING BASED ON COINTEGRATION INFORMATION

6.4 Empirical Study: Volatility Forecasting Performance using Emerging Stock Indices

Two emerging stock indices, Thailand (SET) and Malaysia (KLCI) are employed in this empirical study. The results of the cointegration analysis in Chapter 5 found that these two indices were cointegrated spanning the same period from 1/07/1998 to 31/12/2002. Each index (SET and KLCI) can be individually fitted into an univariate GARCH model. Therefore, the volatility of each index can be forecasted through its best fitting GARCH model.

These two stock indices are first examined using the best fitting GARCH(p,q) models by following the same procedure used in the simulation studies in Section 6.2. It is found that the GARCH(1,1) model is the most appropriate order of GARCH for SET and KLCI for their cointegrated time period from 1/07/1998 to 31/12/2002. The AIC values are reported in Table 6.17 for determining the appropriate error distribution to be used in the GARCH(1,1) model.

Table 6.17: The AIC values given by models with different error distributions

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLCI</td>
<td>-5.1151</td>
<td>-5.1144</td>
<td>-5.1706</td>
<td><strong>-5.1733</strong></td>
<td>-5.1666</td>
<td>-5.1707</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.17 shows that the GARCH(1,1) model with a skewed student-t distribution is the best fitting model for both SET and KLCI based on AIC criterion. The estimated parameters and diagnostic are reported in Tables 6.18 - 6.19.
Table 6.18: Estimated parameters by individual model fitting and the diagnostics from the GARCH(1,1)-SSTD model for SET for data spanning from 1/07/1998 to 31/12/2002

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SSTD</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>4.622\times 10^{-4}</td>
<td>0.6531</td>
</tr>
<tr>
<td>$\omega$</td>
<td>3.615\times 10^{-4}</td>
<td>5.63\times 10^{-5}***</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>2.488\times 10^{-1}</td>
<td>0.0038**</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>6.251\times 10^{-1}</td>
<td>5.164\times 10^{-13}***</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.2270</td>
<td>&lt; 2\times 10^{-16}***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>7.8290</td>
<td>0.0031**</td>
</tr>
<tr>
<td>ARCH(1)-LM test</td>
<td>5.4388</td>
<td>0.9416</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>5.9064</td>
<td>0.9812</td>
</tr>
</tbody>
</table>

Note: (*),(**) and (***) denote significance at 5%, 1% and 0.1 %, respectively.

Table 6.19: Estimated parameters by individual model fitting and diagnostics from the GARCH(1,1)-SSTD model for KLCI for data spanning from 1/07/1998 to 31/12/2002

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SSTD</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>1.390\times 10^{-3}</td>
<td>0.0897</td>
</tr>
<tr>
<td>$\omega$</td>
<td>5.136\times 10^{-5}</td>
<td>0.0109*</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>2.733\times 10^{-1}</td>
<td>0.0008***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>6.326\times 10^{-1}</td>
<td>6.290\times 10^{-13}***</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.1230</td>
<td>&lt; 2\times 10^{-16}***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>5.0350</td>
<td>2.430\times 10^{-5}**</td>
</tr>
<tr>
<td>ARCH(1)-LM test</td>
<td>7.0949</td>
<td>0.8512</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>9.2954</td>
<td>0.8615</td>
</tr>
</tbody>
</table>

Note: (*),(**) and (***) denote significance at 5%, 1% and 0.1 %, respectively.

Figure 5.1 in Chapter 5 shows the co-movement of these two emerging stock indices. They have similarities in their movements and trends. This implies that SET and KLCI have a close relationship with each other. In Chapter 5, the Johansen tests were employed to analyze the cointegration relationship between these two indices. Table 5.2 shows the results of unit root tests and Table 5.3 shows the results of autocorrelation tests for their residuals. From these tests, it can be concluded that both stock indices were integrated in the same order of
one \((I(1))\). The autocorrelation tests suggested that the appropriate lag used in VECM was 8 \((VECM(8))\). Table 5.4 shows the results for Johansen cointegration analysis. There is only one cointegrating vector for LSET and LKLCI. The cointegrating vector was \(\beta' = (1.0000, -0.7532)\). The normalized cointegration equation with respect to LSET is

\[
LSET = 0.7532LKLCI + \delta_t.
\]  

(6.9)

The normalized cointegration equation with respect to LKLCI is

\[
LKLCI = 1.3276LSET + \delta^*_t.
\]  

(6.10)

where \(\delta_t\) and \(\delta^*_t\) are the cointegrating errors from the normalized cointegration equations with respect to LSET and LKLCI, respectively.

The number of lags used in the VECM is eight (i.e., \(VECM(8))\). The residuals of the \(VECM(8)\) are saved and checked for model fitting. Tables 5.5 - 5.6 in Chapter 5 show the results of the best fitting GARCH model for the series of residuals, the results suggest that the GARCH(1,1) model with skewed Student-t distribution is the best fitting model for the residuals in the \(VECM(8)\). The first differences for SET are defined as \(\Delta SET_t\) and \(\Delta KLCI_t\) for KLCI at time \(t\). The \(VECM(8)\) of \(\Delta SET_t\) and \(\Delta KLCI_t\) are shown in the following equations (Equations (6.11) and (6.12), respectively).

\[
\Delta SET_t = -0.025\Delta SET_{t-1} - 0.007\Delta SET_{t-2} - 0.020\Delta SET_{t-3} + 0.032\Delta SET_{t-4} - 0.056\Delta SET_{t-5} + 0.026\Delta SET_{t-6} + 0.049\Delta SET_{t-7} + 0.019\Delta KLCI_{t-1} - 0.011\Delta KLCI_{t-2} + 0.028\Delta KLCI_{t-3} - 0.001\Delta KLCI_{t-4} + 0.031\Delta KLCI_{t-5} - 0.029\Delta KLCI_{t-6} + 0.061\Delta KLCI_{t-7} - 0.037(SET_{t-8} - 0.7532KLCI_{t-8}) + \xi_t,
\]  

(6.11)
\[
\Delta KLCI_t = -0.030\Delta KLCI_{t-1} + 0.023\Delta KLCI_{t-2} + 0.008\Delta KLCI_{t-3} \\
+0.079\Delta KLCI_{t-4} - 0.183\Delta KLCI_{t-5} + 0.060\Delta KLCI_{t-6} \\
+0.052\Delta KLCI_{t-7} + 0.029\Delta SET_{t-1} + 0.007\Delta SET_{t-2} \\
-0.044\Delta SET_{t-3} + 0.022\Delta SET_{t-4} + 0.021\Delta SET_{t-5} \\
-0.017\Delta SET_{t-6} + 0.006\Delta SET_{t-7} \\
+0.025(KLCI_{t-8} - 1.3276SET_{t-8}) + \xi^*_t.
\] (6.12)

Similar to Section 6.2, the evaluation of the volatility forecasting performance between Model A and Model B is carried out. The data set of Model B for SET and KLCI are obtained from Equations (6.11) and (6.12) respectively.

To compare the performance of the volatility forecasting between Model A and Model B, \(\Delta SET_t\) and \(\Delta KLCI_t\) are fitted by a GARCH model. By following the same procedure as used in the simulation studies in Section 6.2, it was found that GARCH(1,1) was the most appropriate order of GARCH model for both SET and KLCI in VECM(8). To determine the type of distributions to be used in the model, six types of error distributions are considered. The AIC values are reported in Table 6.20, which are used to determine an appropriate error distribution for the GARCH(1,1) model.

Table 6.20: The AIC values given by VECM(8) with different error distributions for SET and KLCI

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>N</th>
<th>SN</th>
<th>STD</th>
<th>SSTD</th>
<th>GED</th>
<th>SGED</th>
</tr>
</thead>
<tbody>
<tr>
<td>SET</td>
<td>-5.8476</td>
<td>-5.8461</td>
<td><strong>-5.8522</strong></td>
<td>-5.8490</td>
<td>-5.8517</td>
<td>-5.8487</td>
<td></td>
</tr>
</tbody>
</table>

From Table 6.20, it can be seen that GARCH(1,1) models with Student-t and generalized error distributions are the best fitting models in VECM(8) for SET
and KLCI respectively. The estimated parameters and diagnostics for the two stock indices are reported in Tables 6.21 - 6.22.

Table 6.21: Estimated parameters and diagnostics of GARCH(1,1)-STD model in VECM(8) for SET, the data spanning from 1/07/1998 to 31/12/2002

<table>
<thead>
<tr>
<th>Parameter</th>
<th>STD</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>-0.0001</td>
<td>0.1373</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.0001</td>
<td>0.0001***</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.0075</td>
<td>0.0052**</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.0877</td>
<td>$&lt; 2\times 10^{-16}$***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.0000</td>
<td>0.0043***</td>
</tr>
<tr>
<td>ARCH(1)-LM test</td>
<td>6.7360</td>
<td>0.9597</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>8.0711</td>
<td>0.8692</td>
</tr>
</tbody>
</table>

Note: (*),(**) and (***) denote significance at 5%, 1% and 0.1 %, respectively.

Table 6.22: Estimated parameters and diagnostics of GARCH(1,1)-GED model in VECM(8) for KLCI, the data spanning from 1/07/1998 to 31/12/2002

<table>
<thead>
<tr>
<th>Parameter</th>
<th>GED</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>-0.0004</td>
<td>0.1862</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.0002</td>
<td>0.0003***</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.0129</td>
<td>0.0013**</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.0391</td>
<td>$&lt; 2\times 10^{-16}$***</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.4250</td>
<td>$&lt; 2\times 10^{-16}$***</td>
</tr>
<tr>
<td>ARCH(1)-LM test</td>
<td>5.3702</td>
<td>0.9444</td>
</tr>
<tr>
<td>$Q^2(15)$</td>
<td>8.5895</td>
<td>0.8979</td>
</tr>
</tbody>
</table>

Note: (*),(**) and (***) denote significance at 5%, 1% and 0.1 %, respectively.

The volatility forecasting for 1-, 2-, 10- and 15- steps ahead are carried out. The empirical results for forecasting comparisons between Model A and B are reported in Table 6.23.

Table 6.23 shows that the values of MSE and MAE for Model B are lower than those for Model A in both stock indices. This indicates that their volatility forecasts can be further improved if the knowledge of the cointegration relationship between the stock indices is taken into account. Furthermore, the short
Table 6.23: Comparisons of volatility forecasting performance of Model A and B for emerging stock indices SET and KLCI covering the period 1/07/1998 to 31/12/2002

<table>
<thead>
<tr>
<th></th>
<th>SET 1-step</th>
<th>2-step</th>
<th>10 step</th>
<th>15 step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>MAE</td>
<td>MSE</td>
<td>MAE</td>
<td>MSE</td>
</tr>
<tr>
<td>A</td>
<td>0.2851</td>
<td>0.0849</td>
<td>0.2883</td>
<td>0.0864</td>
</tr>
<tr>
<td>B</td>
<td>0.2464</td>
<td>0.0642</td>
<td>0.2507</td>
<td>0.0661</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>KLCI 1-step</th>
<th>2-step</th>
<th>10 step</th>
<th>15 step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>MAE</td>
<td>MSE</td>
<td>MAE</td>
<td>MSE</td>
</tr>
<tr>
<td>A</td>
<td>0.2999</td>
<td>0.0769</td>
<td>0.2810</td>
<td>0.0735</td>
</tr>
<tr>
<td>B</td>
<td>0.2194</td>
<td>0.0517</td>
<td>0.2258</td>
<td>0.0546</td>
</tr>
</tbody>
</table>

horizontal forecasts provided better accuracy of volatility forecasts.

The conclusions drawn from the empirical results in this section are the same as those from the simulated results in the Section 6.2. As previously shown, allowing a GARCH model to be used with a non-normal error distribution might provide a greater chance to improve the performance of volatility forecasting. Furthermore, if the underlying financial time series are found to be cointegrated while each time series can be individually modeled by the best fitting GARCH mode, then the knowledge of cointegration among these financial time series can be beneficial to the volatility forecasting performance.

6.5 Conclusion

The simulation results revealed that Model B, which contains the knowledge of cointegrated information, could further improve volatility forecasting performance and provide better forecasts than the best fitting univariate GARCH model. The empirical results also confirmed that knowledge of cointegration relationships benefited the performance of volatility forecasting of the underlying financial time series.

Considering the alternative error distributions in the GARCH model for Model
B, a model with a non-normal error distribution tended to outperform a model with the normal error distribution.

Therefore, cointegration relationships among the underlying financial time series have been shown to be beneficial in achieving an accurate volatility forecasting performance. Furthermore, the non-normal error distributions such as skewed Student-t and generalized error distributions in a GARCH model have demonstrated an ability to enhance the accuracy of volatility forecasting.
Chapter 7

Estimation of Value-at-Risk for Emerging Stock Markets

Value at Risk (VaR) plays a crucial role in the risk management as a tool for financial institutions. It is used to determine the downside risk of financial positions. The risk measurement is practically relevant to the volatility forecasting of underlying financial data. The accuracy of volatility predictions are very important in controlling the risk in investment. The more accurate performance of volatility forecasts, the more successful VaR estimates to achieve. Therefore, estimations of VaR rely on the accuracy of volatility forecasting performance.

As previously mentioned, underlying financial time series often exhibited heavy tails, excess kurtosis and skewness. The modification of GARCH models has been very successful in capturing these properties. GARCH models can be used to significantly improve the volatility predictions for these financial time series. In Chapter 4, the results of studies clearly confirm that GARCH models with various types of non-normal error distributions can benefit the performance of volatility forecasting. Their predictive abilities of volatility were more accurate than the traditional GARCH models with normal error distribution.

Furthermore, the results in Chapter 6 reveal that the abilities of volatility forecasting can be further improved by taking into account the cointegration information among underlying time series. The performance of volatility fore-
casting given by models that contain cointegration relationships (Model B) pro-
vide better results than the individually univariate GARCH models (Model A).
A model with the non-normal error distribution in the cointegrating errors also
yields better performance of volatility forecasting than the model with normal
error distribution.

This chapter further investigates the comparisons of VaR estimates based on
Models A and B. The data used in this study are the daily closing prices for SET
and KLCI. Backtesting methods are used to evaluate which VaR estimates are
more accurate and reliable to use in risk management.

7.1 Definition of Value at Risk

Tsay (2010) stated that the VaR can be defined as the maximum loss of a fi-
nancial position over a given period of time horizon with a certain probability \( c \).
Alternatively, VaR can provide the definition with a sense of the minimum loss
as well. Both definitions will be presented in this section.

Firstly, VaR is defined under a probabilistic framework which is related to the
loss function \( L(k) \), where \( k \) is the number of periods \( (k = 1, 2, \ldots) \). The change
in value in the risk for the underlying returns of the financial position for the
next \( k \) periods from the initial time period \( t \) to \( t + k \) can be denoted as \( \Delta V(k) \).
The Cumulative Distribution Function (CDF) of \( L(k) \) can be defined by \( F_k(x) \)
or its quantiles. The financial position of VaR over the time horizon \( k \) can be
considered with the right tail probability \( c \) of the loss function \( L(k) \) as follows:

\[
c = P[L(k) \geq VaR]
\]  

(7.1)

Alternatively, VaR can be considered under a probabilistic function of un-
derlying financial returns. Let \( r_t \) be the financial return at time \( t \). The change
in value for the next \( k \) periods from time period \( t \) to \( t + k \) can be written as
\[ \Delta V(k) = r(t + k) - r(k) = \Delta r. \] In terms of the return distribution, \( L(k) \) is negative function of \( \Delta V(k) \) and the financial position of VaR over the time horizon \( k \) can be considered with the left tail probability \( c \) of the distribution of financial returns as follows:

\[
c = P[\Delta V(k) \leq VaR], \quad (7.2)
\]
\[
= P[\Delta r \leq VaR]. \quad (7.3)
\]

In this chapter, the VaR estimates are defined by taking into consideration the distribution of financial returns. This distribution of return is used to calculate the VaR estimates by concerning with the left tail quantile of the distribution with probability \( c \). As a result, VaR estimates are essentially determined by two components: the time horizon \( k \) and the probability of interest \( c \). Studies dealing with VaR calculation are usually applied to the left tail of financial return distribution. When the return distribution is typically normal distribution, VaR can be denoted with a certain confidence level \((1 - c)\%\) as follows:

\[
VaR_t = Z_c \hat{\sigma}_t + \mu \quad (7.4)
\]

where \( VaR_t \) is the estimated VaR at time \( t \) for the confidence level \( 100 \times (1 - c)\% \), \( Z_c \) is denoted as \( P[Z < Z_c] = c \) which is the left quantile at probability \( c \) of a standard normal distribution \( N(\mu, \sigma^2) \) and \( \hat{\sigma}_t \) refers to the estimated standard deviation at time \( t \).

As previously discussed, the assumption of financial returns is clearly not normally distributed which exhibit a heavy tail, excess kurtosis and sometimes skewness. A GARCH model can cope with these properties well. Füss et al. (2007) stated that the traditional assumption of standard VaR estimates are based on normal distribution. Therefore, the normality assumption might be inadequate for the VaR estimates of financial returns. In order to improve the VaR estimates,
there were some empirical studies in VaR estimates associated with GARCH models that take into account the alternative non-normal error distributions. So and Yu (2006), Hung et al. (2008) and Angelides et al. (2004) examined the improvement of the precision of VaR calculations by allowing the use of GARCH model with non-normal error distributions. Their results revealed that adoption of GARCH models with non-normal error distributions (Student-t and heavy-tailed distributions) improved the accuracy of VaR estimates. However, there are no empirical studies that introduced the cointegration-based ECM into the conventional VaR framework for time varying volatility. None of them discussed whether the cointegration information still contributes the improvement in VaR estimates. Therefore, it is of interest to employ models that provided better volatility forecasting performances to improve the VaR estimates associated with GARCH model and the cointegration relationships. Different six types of error distributions are also considered as the error terms in GARCH models.

The uses of GARCH models with alternative non-normal error distributions are applied into VaR estimates comparing with the models that contain the cointegration information. Backtesting is employed to evaluate the accuracy of VaR estimates given by Model A and B.

### 7.2 Estimation of VaR

A GARCH(p,q) model for the conditional variance of returns can be written as follows:

\[
    r_t = \mu + \varepsilon_t \tag{7.5}
\]

\[
    \varepsilon_t = \eta \sqrt{h_t} \tag{7.6}
\]

\[
    h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}. \tag{7.7}
\]
where \( r_t \) denotes the return series, \( \omega, \alpha_i \) and \( \beta_j \) are non-negative constants with restriction of \( \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1 \) in order to ensure the conditional variance \( (h_t) \) is positive as well as stationary. The \( \eta_t \) in GARCH(p,q) model is usually assumed to be a standard normal distribution. By the definition of VaR in Jorion (2001), the VaR based on GARCH model with normal error distribution can be calculated as follows:

\[
VaR^N_t = Z^N_c \sqrt{h_t} + \mu
\]  

(7.8)

where \( VaR^N_t \) denotes the VaR estimation based on a GARCH model with normal error distribution. \( Z^N_c \) denotes the left quantile at \( c \) for the return distribution estimated from GARCH model with normal error distribution, \( \sqrt{h_t} \) is the estimated volatility forecast from GARCH model with normal error distribution at time \( t \) and \( \mu \) is the mean of GARCH model.

In order to apply other alternative error distributions into GARCH model, the equations of VaR calculation can be derived as follows:

\[
VaR^D_t = Z^D_c \sqrt{h_t} + \mu
\]  

(7.9)

where \( VaR^D_t \) refers to the VaR estimates based on GARCH models with other alternative non-normal error distributions (D). \( VaR^D_t \) can be calculated by using the simulations to obtain return distributions when the errors are not normally distributed. The other five alternative error distributions in Chapter 3 can be applied into GARCH models for calculating the VaR estimations. \( Z^D_c \) denotes the left quantile at \( c \) for the return distribution corresponding to GARCH model with its error distribution, \( \sqrt{h_t} \) is the estimated volatility forecast at time \( t \) by using the GARCH model with each alternative error distribution and \( \mu \) is the mean of GARCH model.

For example, if the GARCH model with skewed normal error distribution are
used for calculating VaR estimates then the VaR can be calculated as follows:

\[ VaR_t^{SN} = Z_{c}^{SN} \sqrt{h_t} + \mu \]  

(7.10)

where \( VaR_t^{SN} \) denotes the VaR estimates based on GARCH model with skewed normal error distribution, \( Z_{c}^{SN} \) denotes the left quantile at \( c \) for the return distribution estimated from a GARCH model with a skewed normal error distribution, \( \sqrt{h_t} \) is the estimated volatility forecast from a GARCH model with a skewed normal error distribution at time \( t \) and \( \mu \) is the mean of GARCH model.

According to the empirical results from Tables 6.17 and 6.20 in Chapter 6, two types of models to describe the return on two stock indices (SET and KLCI) are taken into consideration. The best fitting univariate GARCH model (Model A) for both indices were the GARCH(1,1) model with skewed Student-t error distribution. Another model were that the models based on ECM with GARCH models (Model B) when the error terms were Student-t and generalized error distributions for SET and KLCI, respectively. These two different types of models are used to calculate VaR estimates for 1- and 2- steps ahead forecasting with 95% confidence level in this thesis.

When VaR estimations are calculated based on the GARCH model with skewed Student-t error distribution, it can be written as follows:

\[ VaR_t^{STD} = Z_{c}^{STD} \sqrt{h_t} + \mu \]  

(7.11)

where \( VaR_t^{STD} \) denotes the VaR estimates based on a GARCH model with skewed Student-t error distribution, \( Z_{c}^{STD} \) denotes the left quantile at \( c \) for the return distribution estimated from GARCH model with skewed Student-t error distribution, \( \sqrt{h_t} \) is the estimated volatility forecast from GARCH model with skewed Student-t error distribution at time \( t \) and \( \mu \) is the mean of GARCH model.

The distributions of return and estimated parameters in the distribution play
an important role in VaR estimates. Parameters are estimated given by the historical data observations. If the parameters in return distribution are inappropriately estimated, it can lead to a poor estimation of return distribution function. Consequently, inappropriate return distributions will affect the estimate of VaR. It can lead to overestimation or underestimation of the value of VaR. In this chapter, returns are modeled by two approaches. One approach is called Model A approach and the other is called Model B approach. For each approach, Monte Carlo method is used to estimate the VaR of forecasting returns. The estimates of parameter $\mu$ in Tables 6.18 - 6.19 and Tables 6.21 - 6.22 are close to zero. In this chapter, all VaR estimates are calculated by assuming that $\mu$ is zero. In the following, GARCH(1,1) model with SSTD error distribution is used as an example to demonstrate how the VaR of 1- and 2- steps ahead of returns are estimated for Model A in both SET and KLCI. For Model B, GARCH(1,1) with STD error distribution is used for SET and GARCH(1,1) with GED error distribution for KLCI, respectively. For example, GARCH(1,1) model with SSTD error distribution in Model A for both SET and KLCI can be rewritten as following equations:

$$r_t = \varepsilon_t; \quad (7.12)$$

$$\varepsilon_t = \eta_t \sqrt{h_t}; \quad (7.13)$$

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \quad (7.14)$$

where $\eta_t$ has SSTD distribution with mean 0 and variance 1.

Given information updated to $t-1$, the estimate of the 1-step ahead volatility forecasting of GARCH(1,1) is given by the following equation:

$$h_{t-1}(1) = \hat{\omega}_0 + \hat{\alpha}_1 \hat{r}_{t-1}^2 + \hat{\beta}_1 h_{t-1} \quad (7.15)$$

where $\hat{\omega}_0$, $\hat{\alpha}_1$ and $\hat{\beta}_1$ are the estimates of $\omega$, $\alpha$ and $\beta$. 
The estimation of 2-step ahead volatility forecasting given information updated to $t - 1$ is given by the following equation:

$$h_{t-1}(2) = \hat{\omega}_0 + (\hat{\alpha}_1 + \hat{\beta}_1)h_{t-1}(1)$$ (7.16)

After obtaining volatility forecasts for 1-step ahead ($h_{t-1}(1)$) and 2-step ahead ($h_{t-1}(2)$), the returns for 1- and 2- steps ahead are simulated based on Equation (7.13). Then, using the empirical distribution of return to estimate the VaR of corresponding return.

When the VaR estimates for both Model A and model B are obtained, it is necessary to evaluate the VaR estimates. The Backtesting methods are employed to evaluate the VaR estimations in next section.

### 7.3 Evaluation of VaR Estimates by Backtesting Methods

As was discussed earlier, the knowledge of cointegration was able to improve the volatility forecasting performance. The Model B which contained cointegration information performed more accurately than the best fitting univariate GARCH model (Model A). It is of interest to determine whether the impact of the approaches of Model A and Model B on VaR estimate is significant or not. In the following study, the VaR estimates based on Models A and B for both stock indices (SET and KLCI) are calculated with 95% confidence level for 1- and 2-steps ahead.

VaR value is useful to predict the accuracy of the future risks of financial time series but it is meaningless to use VaR estimates without evaluation. Accordingly, it is necessary to evaluate the precision of VaR estimates after the VaR calculating which one is adequate to be used for predicting the future risks. In order to evaluate the adequacy of the VaR estimates, Backtesting methods are adopted to
evaluate VaR performances. Backtesting methods are a statistical procedure to evaluate the precision of VaR estimates by systematically counting the number of VaR violations, how often that the true return series are less than VaR estimates. If the values of true return series are less than VaR estimates, it is called a violation or exception.

This section aims to introduce the Backtesting for evaluating VaR estimates. There are two different types of the VaR Backtesting methods. One is called the unconditional coverage, the other one is the conditional coverage.

The VaR Backtesting have been developed to examine the accuracy of VaR estimates. The first Backtesting method under the unconditional coverage test is referred to Kupiec test (Kupiec, 1995) and also developed another test at the same time, called TUFF-test (Time Until First Failure).

Christoffersen (1998) has developed the Backtesting method under conditional coverage test which is a joint test of the unconditional coverage and independence. Christoffersen and Pelletier (2004) proposed the Backtesting based on a duration approach and Haas (2001) developed another type of Backtesting method by mixing the Duration-Based Approach of Christoffersen and Pelletier (2004) with TUFF-test.

The Backtesting method for unconditional coverage test used in this chapter is the Kupiec test by Kupiec (1995) and the Backtesting method for conditional coverage test used in this chapter is the interval forecast test by Christoffersen (1998).

7.3.1 The Statistical Framework of VaR Backtesting

A variety of VaR Backtesting tests have been proposed that can be used to test the accuracy of VaR estimates. Consider realization of the return series over a fixed time interval $r_{t,t+1}$ where the VaR estimate at time $t$ with probability $c$ is
CHAPTER 7. ESTIMATION OF VALUE-AT-RISK FOR EMERGING STOCK MARKETS

defined as $VaR_t(c)$. The “hit” function can be defined as follows:

$$I_{t+1}(c) = \begin{cases} 
1 & \text{if } r_{t,t+1} \leq VaR_t(c) \text{ or violation occurs} \\
0 & \text{if } r_{t,t+1} > VaR_t(c) \text{ or no violation occurs} 
\end{cases} \quad (7.17)$$

An accurate VaR estimate can be determined whether the hit sequence satisfies two properties as below (Christoffersen, 1998):

1. Unconditional Coverage Property - The probability of realizing return less than the $VaR_t(c)$ must be precisely $c \times 100\%$. If $r_t$ series occurs the violation more frequently than $c \times 100\%$ of the time then this would suggest that the VaR estimate systematically understates the actual VaR. Alternatively, if there are too few VaR violation, that means the estimation of VaR is overestimate.

2. Independence Property - Any two elements of the hit sequence, $I_{t+i}(c)$ and $I_{t+j}(c)$, where $i < j$, must be independent from each other. If the VaR estimate is accurate, then the violation at time $t + j$ should not depend on whether or not the violation occurred on the previous time $t + i$.

The hit sequence is identically and independently distributed as a Bernoulli random variable with probability $c$ or $I_t(c) \sim i.i.d.\ Bernoulli(c)$. If VaR estimate produces a hit sequence that satisfies the unconditional coverage and independence properties that this VaR estimate is an accurate VaR to reliably use in risk management. As a result, a variety of VaR Backtesting have been developed to examine whether VaR estimate can produce the hit sequence satisfies one or both of these properties. First, the Kupiec test of the unconditional coverage property by Kupiec (1995) will be introduced, then follow by the interval forecast test of Christoffersen (1998) which is the joint test for both unconditional coverage and independence properties.
7.3.2 The Kupiec Test

Kupiec (1995) test of unconditional coverage is the most well-known to test the VaR evaluation based on the failure rates. It is also known as the proportion of failures test (POF). It is used to test whether the observed failure rate is significantly different from the failure rate suggested by the confidence level. The hypothesis for Kupiec test is defined as follows:

\[ H_0 : p = \hat{p} \]
\[ H_1 : p > \hat{p} \]

where \( \hat{p} \) is the observed failure rate which \( \hat{p} = \frac{x}{T} \), where \( p \) is the failure rate suggested by the confidence level, \( 0 \leq p \leq 1 \) and \( x \) is the number of violations while \( T \) is the total number of observations. The number of violations \( x \) are approximately binomially distributed. The density function of binomial distribution is defined as below:

\[
f(x) = \binom{T}{x} p^x (1 - p)^{T-x} \tag{7.18}
\]

Kupiec (1995) proposed the likelihood ratio (\( LR_{uc} \)) test statistics for the Kupiec test which is given by:

\[
LR_{uc} = -2 \log \left( \frac{(1 - p)^{T-x} \hat{p}^x}{[1 - (\frac{x}{T})]^{T-x}(\frac{x}{T})^x} \right) \sim \chi^2(1) \tag{7.19}
\]

where \( x/T \) is the Maximum Likelihood estimator of \( p \), and \( x \) denotes a binomial random variable referring the number of VaR violations. Kupiec test is asymptotically chi-square distributed with one degree of freedom \( \chi^2(1) \). The hypothesis test is used for determining accuracy of VaR estimates. If the null hypothesis can be accepted, it means the observed failure rate is not different from the failure rate suggested by the confidence level. The VaR estimates are acceptably accu-
rate and are reliable to use for predicting the future risks. On the contrary, if the null hypothesis is rejected, it refers to that VaR estimates are not appropriate to use for predicting the future risks.

However, Campbell (2005) mentioned that the Kupiec test had low power of the test and should not rely on tests of unconditional coverage. This chapter also uses the Backtesting of conditional coverage presented in the next following subsection.

### 7.3.3 Christoffersen’s Interval Forecast Test

Kupiec test is one of straightforward unconditional coverage tests which focuses on the unconditional coverage property of an adequate VaR estimates but do not examine whether the independent property is satisfied. As a result, this test might fail to detect VaR estimates which an accurate VaR estimates must exhibit both of the unconditional coverage and independence properties.

To solve this problem, Christoffersen (1998) developed a conditional coverage test, namely Christoffersen’s interval forecast test. This test is used to examine whether VaR estimates exhibit both correct unconditional coverage and serial independence properties of the hit sequence. In other words, this test investigates two properties of VaR estimates whether the number of violations is equal to the expectation, and also examine whether the VaR estimates can produce the independent distribution of hit sequence.

The Christoffersen’s interval forecast test (Christoffersen, 1998) has the likelihood ratio test statistic \( LR_{cc} \) which is a joint test of both unconditional coverage and independence properties of VaR estimates mentioned before. By combining the independence test \( LR_{ind} \) with the Kupiec test \( LR_{uc} \), the likelihood ratio test statistic of the conditional coverage test \( LR_{cc} \) can be written as follows:

\[
LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2)
\]  \hfill (7.20)
where $LR_{uc}$ and $LR_{ind}$ are the likelihood ratio test statistic of unconditional coverage and independence tests, respectively. The likelihood ratio of the Christoffersen’s interval forecast test ($LR_{cc}$) has an asymptotic chi-square distribution with two degrees of freedom ($\chi^2(2)$). The likelihood ratio test of unconditional coverage test was mentioned in the previous subsection. The independence test will be presented in the following.

Let $I_t$ be the indicator function of violations which is defined as below:

$$I_t = \begin{cases} 
1 & \text{if } r_t \leq VaR_t(c) \text{ or violation occurs} \\
0 & \text{if } r_t > VaR_t(c) \text{ or no violation occurs}
\end{cases} \quad (7.21)$$

where $r_t$ is realization of the return series at time $t$ and $VaR_t(c)$ is the estimation of VaR at time $t$ with the probability $c$.

Christoffersen (1998) test for serial independence in $I_t$ is used to test whether the probability of a violation at time $t$ given by that a violation occurred at the previous time $t-1$, denotes as $p_{11}$, is equal to the probability of a violation at time $t$ given by that no violation occurred at the previous time $t-1$, denotes as $p_{01}$. The contingency table of possible outcome is shown as below:

<table>
<thead>
<tr>
<th>$I_{t-1} = 0$</th>
<th>$I_{t-1} = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_t = 0$</td>
<td>$n_{00}$</td>
</tr>
<tr>
<td>$I_t = 1$</td>
<td>$n_{01}$</td>
</tr>
<tr>
<td>$n_{00} + n_{01}$</td>
<td>$n_{10} + n_{11}$</td>
</tr>
</tbody>
</table>

where,

$n_{00}$ denotes as no VaR violation at time $t$ and on $t-1$

$n_{10}$ denotes as no VaR violation at time $t$ but there was VaR violation on $t-1$

$n_{01}$ denotes as there is VaR violation at time $t$ but no VaR violation at time $t-1$
CHAPTER 7. ESTIMATION OF VALUE-AT-RISK FOR EMERGING
STOCK MARKETS

$n_{11}$ denotes as there is VaR violation at time $t$ and time $t - 1$.

The number of observations in stat $j$ after having been in state $i$ at the
previous time denotes as $n_{ij}$. In general, let $p_{ij}$ be the probability of conditional
observed violation in state $j$ given previous state $i$. Then $p_{01}$ and $p_{11}$ can be
calculated as follows:

$$p_{01} = \frac{n_{01}}{n_{00} + n_{01}},$$

$$p_{11} = \frac{n_{11}}{n_{10} + n_{11}}.$$

where

$$\hat{p} = \frac{n_{01} + n_{11}}{n_{00} + n_{01} + n_{10} + n_{11}}$$

and $N = n_{00} + n_{01} + n_{10} + n_{11}$.

Under the null hypothesis of independence test, the violation at time $t$ should
not depend on the violation occurred on the previous time $(t - 1)$. Accordingly,
the probabilities $p_{01}$, $p_{11}$ and $\hat{p}$ should be equal as below:

$$p_{01} = p_{11} = \hat{p}$$

The relevant likelihood ratio test statistic for independence of violation ($LR_{ind}$)
was defined as follows:

$$LR_{ind} = -2 \log \left( \frac{(1 - \hat{p})^{n_{00} + n_{10}} \hat{p}^{n_{01} + n_{11}}}{(1 - p_{01})^{n_{00} n_{01}} p_{01}^{n_{10}} n_{11} p_{11}} \right) \sim \chi^2(1)$$

where an asymptotical distribution of the test is chi-square distribution with one
degree of freedom ($\chi^2(1)$).

In the Christoffersen’s interval forecast test (Christoffersen, 1998), $LR_{uc}$ and
$LR_{ind}$ are separately calculated and then combined together. Both test statistics are chi-square distribution with one degree of freedom. Due to the $LR_{cc}$ is a joint test of these two test statistics ($LR_{uc}$ and $LR_{ind}$), thus the $LR_{cc}$ can also be tested by using chi-square distribution with two degrees of freedom.

The hypothesis for the Christoffersen’s interval forecast test is the same as the Kupiec test but considers the independence test as well. The null and alternative hypothesis are defined as follows:

$$H_0 : p = \hat{p}$$
$$H_1 : p > \hat{p}$$

If the null hypothesis can be accepted, then the VaR estimates are acceptably accurate and are reliable to use for predicting the future risks. On the other hand, if the null hypothesis is rejected, it refers to that VaR estimates are not appropriate to use.

### 7.4 Empirical Results and Evaluation of VaR Estimates

The VaR estimates of SET and KLCI given by Model A and B are denoted as $SET_A$, $KLCI_A$, $SET_B$ and $KLCI_B$, respectively. The comparisons of VaR estimates for 1- and 2- steps ahead are carried out in this section.

The data used to calculate the VaR estimates are the daily closing prices for SET and KLCI over the period from 1/07/1998 to 31/12/2002 with a total of 1,133 observations. To calculate VaR estimates based on Model A and Model B, the first 500 observations are used as in-sample, the remaining 633 observations are used as the out-of-sample. All GARCH models are estimated based on 500 daily closing observations, and the length of rolling window is 500 observations.

After obtaining VaR estimates by using the rolling window, the number of
violations are counted using Equation (7.21). The Kupiec test statistic ($LR_{uc}$) in Equation (7.19) for the unconditional coverage, and the Christoffersen’s interval forecast test statistic ($LR_{cc}$) in Equation (7.20) for conditional coverage are calculated to test the hypothesis at 95% confidence level.

In theory, the number of expected violation with 95% confidence level for 1- and 2- steps ahead are 31.65 and 31.6 (5% of 633 and 632), respectively. The results of Kupiec test for 1- and 2- steps VaR estimates given by Model A and Model B for both SET and KLCI are reported in Table 7.2.

Table 7.2: Out-of-sample Value-at-Risk for 1- and 2- steps ahead given by the Kupiec test at 95% confidence level

<table>
<thead>
<tr>
<th>VaR Model</th>
<th>1-step ahead forecasts</th>
<th>2-step ahead forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean VaR estimate</td>
<td>Number of violations</td>
</tr>
<tr>
<td>SET$_A$</td>
<td>-0.0497</td>
<td>43</td>
</tr>
<tr>
<td>KLCI$_A$</td>
<td>-0.0313</td>
<td>46</td>
</tr>
<tr>
<td>SET$_B$</td>
<td>-0.0536</td>
<td>38</td>
</tr>
<tr>
<td>KLCI$_B$</td>
<td>-0.0374</td>
<td>41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VaR Model</th>
<th>Mean VaR estimate</th>
<th>Number of violations</th>
<th>Test statistic $LR_{uc}$</th>
<th>Critical Value $\chi^2(1)$</th>
<th>Test Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>SET$_A$</td>
<td>-0.0137</td>
<td>52</td>
<td>0.0823</td>
<td>11.7021*</td>
<td>Reject</td>
</tr>
<tr>
<td>KLCI$_A$</td>
<td>-0.0152</td>
<td>56</td>
<td>0.0886</td>
<td>16.2911*</td>
<td>Reject</td>
</tr>
<tr>
<td>SET$_B$</td>
<td>-0.0458</td>
<td>50</td>
<td>0.0791</td>
<td>9.6563*</td>
<td>Reject</td>
</tr>
<tr>
<td>KLCI$_B$</td>
<td>-0.0527</td>
<td>55</td>
<td>0.0870</td>
<td>15.0834*</td>
<td>Reject</td>
</tr>
</tbody>
</table>

Note: (*) indicated significance at the 5% level.

The Table 7.2 shows the results for each VaR estimates at 95% confidence level evaluating by Kupiec test. The critical value of chi-square with one degree of freedom is 3.84 at 95% confidence level. By comparing the four VaR estimates ($SET_A$, $KLCI_A$, $SET_B$ and $SET_B$), the 1-step ahead of VaR estimate of $SET_B$ and $KLCI_B$ are not significant and the null hypothesis cannot be rejected. This result indicates that the 1-step ahead VaR estimates given by Model B are sig-
nificantly acceptable. However, the rest of VaR estimates fail to pass the Kupiec test, thus indicating that VaR estimates are inaccurate at 95% confidence level.

The results in Chapter 6 indicate that a knowledge of cointegration can further improve the performance of volatility forecasting comparing with the individual best fitting univariate GARCH model. According to the results of Kupiec test, the VaR estimates by taking into consideration the cointegration information of underlying time series can also provide an more accurate and reliable estimations of VaR.

Conditional coverage test \((LR_{cc})\) developed by Christoffersen (1998), is also now examined. This conditional coverage test is the joint test of two tests, the testing independence of violation \((LR_{ind})\) and the Kupiec test \((LR_{uc})\). The test results of Kupiec test for four VaR estimates are shown in Table 7.2. The test results for independence of violation \((LR_{ind})\) are calculated, which can be used to obtain the conditional test \((LR_{cc})\) for all VaR estimates.

Input data for calculating the Backtesting independence test \((LR_{ind})\) of 1- and 2- steps ahead at 95% confidence level are reported in Table 7.3.

Table 7.3: Input data for calculating the Backtesting independence test

<table>
<thead>
<tr>
<th>VaR</th>
<th>n₀₀</th>
<th>n₁₀</th>
<th>n₀₁</th>
<th>n₁₁</th>
<th>p₀₁</th>
<th>p₁₁</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SET_A)</td>
<td>548</td>
<td>42</td>
<td>42</td>
<td>1</td>
<td>0.0711</td>
<td>0.0232</td>
<td>0.0679</td>
</tr>
<tr>
<td>(KLCI_A)</td>
<td>542</td>
<td>45</td>
<td>45</td>
<td>1</td>
<td>0.0766</td>
<td>0.0217</td>
<td>0.0726</td>
</tr>
<tr>
<td>(SET_B)</td>
<td>558</td>
<td>37</td>
<td>37</td>
<td>1</td>
<td>0.0621</td>
<td>0.0263</td>
<td>0.0600</td>
</tr>
<tr>
<td>(KLCI_B)</td>
<td>564</td>
<td>34</td>
<td>34</td>
<td>1</td>
<td>0.0568</td>
<td>0.0285</td>
<td>0.0552</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VaR</th>
<th>n₀₀</th>
<th>n₁₀</th>
<th>n₀₁</th>
<th>n₁₁</th>
<th>p₀₁</th>
<th>p₁₁</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SET_A)</td>
<td>537</td>
<td>43</td>
<td>43</td>
<td>9</td>
<td>0.0741</td>
<td>0.1730</td>
<td>0.0822</td>
</tr>
<tr>
<td>(KLCI_A)</td>
<td>534</td>
<td>45</td>
<td>45</td>
<td>8</td>
<td>0.0777</td>
<td>0.1509</td>
<td>0.0838</td>
</tr>
<tr>
<td>(SET_B)</td>
<td>540</td>
<td>42</td>
<td>42</td>
<td>8</td>
<td>0.0721</td>
<td>0.1600</td>
<td>0.0791</td>
</tr>
<tr>
<td>(KLCI_B)</td>
<td>537</td>
<td>40</td>
<td>40</td>
<td>15</td>
<td>0.0693</td>
<td>0.2727</td>
<td>0.0870</td>
</tr>
</tbody>
</table>

The critical value of the \(\chi^2\) distribution at 95% confidence level with two
degrees of freedom is 5.99, if the test statistics values are lower than the critical values then the VaR model is significantly acceptable.

Table 7.4 shows the unconditional coverage of Kupiec test ($LR_{uc}$) results, Christoffersen’s independence test ($LR_{ind}$) results and the conditional coverage joint test ($LR_{cc}$) of unconditional coverage and independence test results of 1- and 2- steps ahead at 95% confidence level. The $LR_{cc}$ values from Equation (7.20) are calculated and obtained all outcomes of hypothesis testing in Table 7.4.

Table 7.4: Conditional coverage Backtesting results for 1- and 2- steps ahead at 95% confidence level

<table>
<thead>
<tr>
<th>VaR</th>
<th>1-step ahead forecasts</th>
<th>2-step ahead forecasts</th>
<th>Test outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$LR_{uc}$</td>
<td>$LR_{ind}$</td>
<td>$LR_{cc}$</td>
</tr>
<tr>
<td>$SET_A$</td>
<td>3.8713</td>
<td>1.8861</td>
<td>5.7574</td>
</tr>
<tr>
<td>$KLCI_A$</td>
<td>6.0442</td>
<td>2.5348</td>
<td>8.5791*</td>
</tr>
<tr>
<td>$SET_B$</td>
<td>1.2637</td>
<td>1.0067</td>
<td>2.2705</td>
</tr>
<tr>
<td>$KLCI_B$</td>
<td>2.6705</td>
<td>0.6039</td>
<td>3.2744</td>
</tr>
<tr>
<td>$SET_A$</td>
<td>11.7021</td>
<td>4.9507</td>
<td>16.6528*</td>
</tr>
<tr>
<td>$KLCI_A$</td>
<td>16.2911</td>
<td>2.8508</td>
<td>19.1419*</td>
</tr>
<tr>
<td>$SET_B$</td>
<td>9.6563</td>
<td>3.9426</td>
<td>13.5989*</td>
</tr>
<tr>
<td>$KLCI_B$</td>
<td>15.0834</td>
<td>3.5071</td>
<td>18.5905*</td>
</tr>
</tbody>
</table>

Note: (*) indicated significance at the 5% level.

The critical value of the $\chi^2$ distribution at 95% confidence level with two degree of freedom is 5.99. Three VaR estimates ($SET_A$, $SET_B$ and $KLCI_B$) pass the coverage test in 1-step ahead forecasting while $KLCI_A$ fails to pass this test at 5% level. For VaR estimates of 2-step ahead forecasts, all VaR estimates also fail this test at 5% level.

From the Backtesting results, the knowledge of cointegration of underlying financial time series can provide more accurate in the performance of volatility forecasting, particularly for 1 step-ahead VaR forecasting in ($SET_B$) and ($KLCI_B$). The 2-step ahead of VaR estimates are rejected for all models but the
test statistics of Model B for SET and KLCI tend to be pass this test slightly more than Model A.

7.5 Conclusion

In this chapter, VaR estimates are calculated using the best fitting GARCH model with skewed Student-t distribution (Model A) and the models which contain the knowledge of cointegration (Model B) between two stock markets (SET and KLCI).

A comparison of VaR estimates between Model A and Model B for SET and KLCI are investigated and which VaR model is more reliable to use is studied. Two different types of Backtesting, the unconditional coverage ($LR_{uc}$) and conditional coverage ($LR_{cc}$) are used in evaluating the VaR estimates.

The results of both Backtesting methods show that VaR estimates calculated based on the knowledge of cointegration information (Model B) can produce adequate VaR forecasts for 1-step ahead for both SET and KLCI while the rest of VaR estimates cannot fully pass these tests. It indicates that the VaR estimates calculated based on Model B are more accurate and reliable than Model A to use for 1-step ahead forecasting at 95% confidence level. While the 2-step ahead of VaR estimates cannot pass these tests for all models but values of test statistics $LR_{cc}$ for Model B are smaller than that Model A for both SET and KLCI.
Chapter 8

Conclusion and Further Research

8.1 Conclusion

Improving the performance of volatility forecasting for the underlying financial time series has received attention of financial statisticians, mathematicians and financial investors. Volatility models now play an important role in the area of econometric financial time series analysis. Unfortunately, in the initial ARCH model, there were some restrictions on volatility modeling when the model had a higher order degree. In these cases, ARCH models sometimes required estimates of a large number of parameters and violated the requirement for non-negative coefficient estimators (Zakaria and Abdalla, 2012). To overcome this difficulty, Bollerslev (1986) developed the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model.

The traditional GARCH model is based on the assumption of normality in the error term distribution. Whereas financial returns are typically non-normally distributed. In this thesis, GARCH models with non-normal distributions are examined to see if this improves volatility forecasting performance. The error distributions in GARCH model include six types of distributions: normal, skewed normal, Student-t, skewed Student-t, generalized error and skewed generalized error distributions. The order of GARCH model considered in this thesis is not only GARCH(1,1) but a higher order GARCH(p,q) model is also used. Along
with other factors that make contribute to non-normality, some stylized fact of financial returns and the probability density functions of these six distributions are presented in Chapter 2 and 3, respectively.

In Chapter 3, the approach how to identify order determination of GARCH\((p,q)\) model under the non-normality error assumption is suggested. AIC criterion is used for model selection in this thesis. Firstly, AIC criterion is used to select an appropriate order among competing GARCH\((p,q)\) models by assuming the error terms are normally distributed. Secondly, the error distribution in the selected GARCH\((p,q)\) model will be changed by using the six types of error distributions to identify an appropriate error distribution for the GARCH\((p,q)\) model. Then, diagnostics tests are carried out by using Ljung-Box and the Lagrange multiplier test (LM). This suggestion for order determination of the GARCH\((p,q)\) model under non-normal error distribution is acceptable according to simulation studies.

Shamiri and Isa (2009) stated that the best fitting model based on the AIC does not necessarily provide the best forecast of volatility in terms of MSE and MAE. Their conclusion was based on the study of the Kuala Lumpur Composite Index (KLCI) from Malaysia fitted to various types of GARCH models. To examine whether Shamiri and Isa’s statement is still acceptable when the underlying financial time series are fitted by GARCH\((p,q)\) model with higher orders, simulation studies on this problem are carried out in Chapter 3. Using the pair t-test, the simulated results showed that the values of MSE (MAE) given by the best fitted model and the best forecast performance model are not significantly different.

Many researchers have investigated the performance of GARCH\((1,1)\) models with alternative non-normal error distributions in mature stock markets. In Chapter 4, the aim of this study is to examine and compare the volatility forecasting performance among competing GARCH\((p,q)\) models with six different types
of error distributions using the three emerging stock indices of South East Asia, Thailand, Malaysia and Singapore. Furthermore, this chapter also investigates whether the best fitting model, in terms of AIC still provides the best performance of volatility forecasting of the underlying series measured by the Mean Squared Error (MSE) and the Mean Absolute Error (MAE) criteria. The results for identifying the best fitting of each stock markets show that the GARCH(1,3) model with skewed Student-t error distribution is the best fitting model for SET, the GARCH(1,1) model with generalized error distribution for KLCI and the GARCH(2,1) model with Student-t for STI. The results reveal that the best fitting models in each stock index does not necessarily provide the best performance of volatility forecasting.

To evaluate the performance of volatility forecasting using the Percent Error (PE), the means of MSE (MAE) between the best fitting model and the best performance model are not much difference. The empirical results of evaluation for the performance of volatility forecasting between these two model in Chapter 4 is the same as the simulated results in Chapter 3. The results of volatility forecasting performance for each stock indices show that GARCH models with non-normal error distributions tend to provide better out-of-sample forecasting performance than GARCH models with normal error distribution evaluated by MSE and MAE.

In Chapter 5, the results of the size and power for cointegration tests are investigated. From the results of Chapter 5, it showed that the size of Dickey-Fuller test with $T(\hat{\phi} - 1)$ statistic is lower when compared with other cointegration tests. The size of Dickey-Fuller test tends to be smaller when the errors of a GARCH(1,1) model is skewed generalized error distribution. It indicates that the Dickey-Fuller tests yield the best performance in terms of the size of the test. However, the power of the Dickey-Fuller tests are very low comparing with the
Johansen tests. The power of Johansen tests provide the best performance in all different type of parameters compared with other cointegration tests, particularly the $\lambda_{max}$ statistic is slightly better than the $\lambda_{trace}$ statistic. Furthermore, the power of Johansen tests slightly increase when the errors of GARCH(1,1) model is given by the skewed Student-t error distribution. The Johansen approach is used for examining the performance of volatility forecasting when underlying financial time series are cointegrated in Chapter 6.

In Chapter 6, the comparison of volatility forecasting performance between the individual univariate GARCH(p,q) model and the model that contains the cointegration information are carried out. Simulation study on the two emerging stock indices of Thailand(SET) and Malaysia(KLCI) are employed in the study. The simulation results reveal that the model which contains the knowledge of cointegration information can further improve the volatility forecasting performance and provide better forecasts than the best fitting univariate GARCH model. The empirical results also confirms the knowledge of cointegration relationships can be beneficial to the performance of volatility forecasting of the underlying financial time series. Considering the alternative error distributions in GARCH model for the model that contain cointegration information (Model B), a model B with the non-normal error distributions tends to outperform a model with normal error distribution. Therefore, it reveals that cointegration relationships among the underlying financial time series can provide more accuracy volatility forecasting performance when these time series are cointegrated. Furthermore, the non-normal error distributions such as skewed Student-t and generalized error distributions in a GARCH model can enhance an accuracy of volatility forecasting as well.

Chapters 4 - 6 provide the contribution for improving the the volatility forecasting performance of underlying financial time series. A GARCH(p,q) model
with the non-normal error distributions can well capture some stylized facts of return series and perform better than a normal distribution. If time series are cointegrated, the knowledge of cointegrated information can also further improve the performance of volatility predicting. The model that contains the cointegration information outperforms the individual univariate GARCH(p,q) model by taking into account the non-normal error distribution in GARCH(p,q) model.

Chapter 7 investigates the comparisons of the VaR estimations between univariate GARCH model and the cointegrated ECM model by using the daily closing prices for SET and KLCI. The Backtesting methods are employed to evaluate which VaR estimations is more accurate and reliable to use for the underlying financial time series. The empirical results show that the VaR estimates are more accurate and reliable in 1-step ahead for both SET and KLCI if the knowledge of cointegration information is considered. The null hypothesis of Kupiec and Christoffersen tests for VaR models which contain the knowledge of cointegration (Model B) can be accepted at 95% confidence level for both SET and KLCI at 1-step ahead VaR forecasting. In other words, if time series are cointegrated then the knowledge of cointegration information provides more information on volatility forecasting and benefit to obtain the more accurate and reliable VaR estimate.

The study carried in this thesis provides necessary ground work for further research in developing GARCH model and other models for improving volatility forecasting.

8.2 Further Research

In Chapter 4, symmetric GARCH(p,q) models are used to forecast the volatility of financial returns. As mentioned, symmetric GARCH(p,q) models can capture some stylized characteristic of returns but cannot deal with the leverage
effect. Asymmetric GARCH\((p,q)\) model can better capture this return property than symmetric model. Currently, the EGARCH\((1,1)\) model becomes a popular model for volatility forecasting. Thus, further research can adopt the EGARCH model with higher orders for improving the volatility forecasting performance. Furthermore, there are various types of non-normal distributions can be applied into GARCH error terms and compare the performance of volatility forecasting.

In Chapter 5, two underlying time series are considered to evaluate the size and power of cointegration tests. Further investigations can be carried out in case of more than two underlying time series by taking into account asymmetric GARCH models with alternative non-normal distributions in the cointegrating error terms.

In Chapters 6 - 7, the symmetric GARCH\((p,q)\) models are used to compare the volatility forecasting performance with the the error correction model. Further research can also explore the comparisons of performance when EGARCH\((p,q)\) with the non-normal error distributions are employed.
Appendix A

Program Files

A.1 Programs used in Chapter 3

R program for simulation studies on the order determination in Table 3.1.

```r
set.seed(1)

num = 3000
rt=c()
alpha0 = 0.01
alpha1 = 0.3
alpha2 = 0.01
beta1 = 0.5
beta2 = 0.02
beta3 = 0.01

rt=garch.sim(alpha=c(alpha0,alpha1,alpha2),beta=c(beta1,beta2),
n=num,rnd=rstd)
rt=garch.sim(alpha=c(alpha0,alpha1,alpha2),beta=c(beta1,beta2),
n=num,rnd=rsstd)
rt=garch.sim(alpha=c(alpha0,alpha1,alpha2),beta=c(beta1,beta2),
n=num,rnd=rged)

#################
# Fit rt 1:1500 #
#################

fit1= garchFit(rt~arma(0,0)+garch(2,2),data=rt[1:1500],cond.dist="norm")
summary(fit1)

fit2= garchFit(rt~arma(0,0)+garch(1,2),data=rt[1:1500],cond.dist="norm")
summary(fit2)
```
fit3= garchFit(rt~arma(0,0)+garch(1,3),data=rt[1:1500],cond.dist="norm")
summary(fit3)

fit4= garchFit(rt~arma(0,0)+garch(2,1),data=rt[1:1500],cond.dist="norm")
summary(fit4)

fit5= garchFit(rt~arma(0,0)+garch(2,2),data=rt[1:1500],cond.dist="norm")
summary(fit5)

fit6= garchFit(rt~arma(0,0)+garch(2,3),data=rt[1:1500],cond.dist="norm")
summary(fit6)

R program for simulation studies on the performance of volatility forecasting comparing the best fitting model and the best performance model in Tables 3.2, 3.3, 3.4 and 3.5.

A.1.1 Simulations for the True GARCH(1,3) and GARCH(2,1) Models

set.seed(1)
numcase=6536

rt=c()
at=c()
at[0]=0
at_sq=0
h_es=c()
h_es[0]=0
h_true=c()
h_true[0]=0
omega=0
mu=0
alpha1=0
beta_1=0
beta_2=0
beta_3=0
alpha0=0.00007
alpha1=0.02354
beta1=0.05387
beta2=0.00127
APPENDIX A. PROGRAM FILES

beta3=0.18574

MAE=0
MSE=0
MSE_r = c()
MAE_r = c()
aic = c()
Mean_MSE = 0
Mean_MAE = 0
Mean_aic = 0
SD_MSE = 0
SD_MAE = 0
SD_aic = 0

#########################
# Start Looping #
#########################

no=1
for (rep in 1:no) {

rt=garch.sim(alpha=c(alpha0,alpha1),beta=c(beta1,beta2,beta3)
   ,n=numcase,rnd=rsged)

at=rt-mean(rt)

h_true[1]<-alpha0+alpha1*0+beta1*0+beta2*0+beta3*0
h_true[1]

for (i in 2:numcase) {
   h_true[i]=alpha0+alpha1*at[i-1]^2+beta1*h_true[i-1]
      +beta2*h_true[i-2]+beta3*h_true[i-3]
}

h_true
fit_1=garchFit(rt~arma(0,0)+garch(1,3),data=rt[1:3535],cond.dist="norm")
# cond.dist="norm, snorm, std, sstd, ged, sged
summary(fit_1)
fit_1@fit$ics[1]

para=coef(garchFit(rt~arma(0,0)+garch(1,3),data=rt[1:3535],cond.dist="norm"))
para
mu<-para[1];mu
omega<-para[2];omega
alpa1<-para[3];alpa1
beta_1<-para[4];beta_1
beta_2<-para[5];beta_2
beta_3<-para[6];beta_3

h_es[1]<- omega+alpa1*0+beta1*0+beta2*0+beta3*0
h_es[1]

for (k in 2:numcase) {
    h_es[k]=omega+alpa1*at[k-1]^2+beta_1*h_es[k-1]
            +beta_2*h_es[i-2]+beta_3*h_es[i-3]
}
h_es

garch13<-function(forstep,omega,alpa1,beta_1,beta_2,beta_3,
currenth,currenth_1,currenth_2)
{
    vol<-matrix(c(0),nrow=(forstep+3),ncol=1)
    vol[1]<-currenth_2
    vol[2]<-currenth_1
    vol[3]<-currenth
    for ( i in 4:(forstep+3)) {
        vol[i]= omega+(alpa1+beta_1)*vol[i-1]+beta_2*vol[i-2]+beta_3*vol[i-3]
    }
    vol[forstep+3,]
garch13(10, para[2], para[3], para[4], para[5], para[6], h_es[3535], h_es[3534], h_es[3533])

for (j in 1:100) {
    MAE=MAE+abs(h_true[3535+(j-10)+1]-garch13(10, omega, alpa1, beta_1, beta_2, beta_3, h_es[3535+j-1], h_es[3535+j-2], h_es[3535+j-3]))
    MSE=MSE+(h_true[3535+(j-10)+1]-garch13(10, omega, alpa1, beta_1, beta_2, beta_3, h_es[3535+j-1], h_es[3535+j-2], h_es[3535+j-3]))^2
}

MAE=MAE/100
MSE=MSE/100

aic[rep] = fit_1@fit$ics[1]
MAE_r[rep] = MAE
MSE_r[rep] = MSE

MSE_r
MAE_r

set.seed(1)
num=5407

rt=c()
at=c()
at[0]=0
at_sq=0
h_es=c()
h_es[0]=0
h_true=c()
h_true[0]=0
omega=0
mu=0
alpa1=0
alpa2=0
beta_1=0
alpha0=0.00008
alpha1=0.05334
alpha2=0.06147
beta1=0.08599
MAE=0
MSE=0
MSE_r = c()
MAE_r = c()
aic = c()
Mean_MSE = 0
Mean_MAE = 0
Mean_aic = 0
SD_MSE = 0
SD_MAE = 0
SD_aic = 0

#################
# Start Looping #
#################

no=100
for (rep in 1:no) {

rt=garch.sim(alpha=c(alpha0,alpha1,alpha2),beta=c(beta1),
n=num,rnd=rsged)
# cond.dist="norm, snorm, std, sstd, ged, sged

at=rt-mean(rt)
at_sq=at^2
h_true[1]<-alpha0+alpha1*0+alpha2*0+beta1*0
h_true[1]

for (i in 2:num) {
    h_true[i]=alpha0+alpha1*at[i-1]^2
    +alpha2*at[i-2]^2+beta1*h_true[i-1]
}
h_true

#################
# Fit rt 1:2500 #
fit_1 = garchFit(rt~arma(0,0)+garch(2,1),data = rt[1:2500],cond.dist="sged")
summary(fit_1)
fit_1@fit$ics[1] # for AIC
para<-coef(garchFit(rt~arma(0,0)+garch(2,1),data = rt[1:2500],
               cond.dist="sged"))
para
mu<-para[1];mu
omega<-para[2];omega
alpa1<-para[3];alpa1
alpa2<-para[4];alpa1
beta_1<-para[5];beta_1
h_es[1]<- omega+alpa1*0+alpa2*0+beta1*0
h_es[1]
for (k in 2:num) {
  h_es[k]=omega+alpa1*at[k-1]^2
    +alpa2*at[k-2]^2+beta_1*h_es[k-1]
}
h_es

garch21<-function(forstep,omega,alpa1,alpa2,beta_1,currenth,currenth_1,
currenth_2)
{
  vol<-matrix(c(0),nrow=(forstep+3),ncol=1)
  vol[1]<-currenth_1
  vol[2]<-currenth
  for ( i in 3:(forstep+2))
  {
    vol[i]= omega+(alpa1+beta_1)*vol[i-1]+alpa2*vol[i-2]
  }
  vol[forstep+2,]
}
garch21(1, para[2], para[3], para[4], para[5], h_es[2500], h_es[2499])

for (j in 1:100) {
  MAE = MAE + abs(h_true[2500+(j-1)+1] - garch21(1, omega, alpa1, alpa2, beta_1,
                        h_es[2500+j-1], h_es[2500+j-2]))
  MSE = MSE + (h_true[2500+(j-1)+1] - garch21(1, omega, alpa1, alpa2, beta_1,
                        h_es[2500+j-1], h_es[2500+j-2]))^2
}
MAE = MAE/100
MSE = MSE/100

R program for simulation studies on the pair t-test in Table 3.6.

# y<-read.csv("C:/garch(1,3)_n.csv",header=TRUE)
# y<-read.csv("C:/garch(1,3)_sn.csv",header=TRUE)
# y<-read.csv("C:/garch(1,3)_std.csv",header=TRUE)
# y<-read.csv("C:/garch(1,3)_sstd.csv",header=TRUE)
y<-read.csv("C:/garch(1,3)_ged.csv",header=TRUE)
# y<-read.csv("C:/garch(1,3)_sged.csv",header=TRUE)
y<-as.matrix(y)

MSE1 = t.test(y[,22],y[,2],paired = TRUE,alternative="greater")#n vs sn
MSE2 = t.test(y[,22],y[,6],paired = TRUE,alternative="greater")#n vs std
MSE3 = t.test(y[,22],y[,10],paired = TRUE,alternative="greater")#n vs sstd
MSE4 = t.test(y[,22],y[,14],paired = TRUE,alternative="greater")#n vs ged
MSE5 = t.test(y[,22],y[,18],paired = TRUE,alternative="greater")#n vs sged

MAE1 = t.test(y[,24],y[,4],paired = TRUE,alternative="greater")#n vs sn
MAE2 = t.test(y[,24],y[,8],paired = TRUE,alternative="greater")#n vs std
MAE3 = t.test(y[,24],y[,12],paired = TRUE,alternative="greater")#n vs sstd
MAE4 = t.test(y[,24],y[,16],paired = TRUE,alternative="greater")#n vs ged
MAE5 = t.test(y[,24],y[,20],paired = TRUE,alternative="greater")#n vs sged
A.2  Programs used in Chapter 4

R program for modeling and forecasting the volatility in Tables 4.2 - 4.7.

A.2.1  Modeling and Forecasting for SET

```r
h=c()
h[0]=0
at=c()
at[0]=0
numcase<-6536
predh=c()
MAE=0
MSE=0
MAPE=0

y<-read.csv("C:/setall_new.csv",header=TRUE);y
y<-as.matrix(y)

fit_1= garchFit( y~arma(0,0)+garch(1,3), data = y[1:3535], cond.dist="norm")
summary(fit_1)
# cond.dist="norm, snorm, std, sstd, ged, sged

para<-coef(garchFit(y~arma(0,0)+garch(1,3), data = y[1:3535],
cond.dist="norm"))
para
mu<-para[1];mu
omega<-para[2];omega
alpha1<-para[3];alpha1
```
beta1<-para[4];beta1
beta2<-para[5];beta2
beta3<-para[6];beta3
at<-y-mu;at

h[1]<-omega+alpha1*0^2+beta1*0+beta2*0+beta3*0
h[1]

#############################################################################
# Calculated volatility #
#############################################################################

for (i in 2:numcase) {
  h[i]=omega+alpha1*at[i-1]^2+beta1*h[i-1]
  +beta2*h[i-2]+beta3*h[i-3]
}
h

#############################################################################
# Forecast Volatility k-step ahead #
#############################################################################
garch13<-function(forstep,omega,alpha1,beta1,beta2,beta3,currenth,currenth_1,currenth_2) {
  vol<-matrix(c(0),nrow=(forstep+3),ncol=1)
  vol[1]<-currenth_2
  vol[2]<-currenth_1
  vol[3]<-currenth
  for (i in 4:(forstep+3)) {
    vol[i]= omega+(alpha1+beta1)*vol[i-1]+beta2*vol[i-2]+beta3*vol[i-3]
  }
  vol[forstep+3,]
}
garch13(15,para[2],para[3],para[4],para[5],para[6],h[3535],h[3534],h[3533])

for (j in 1:100) {
  MAE=MAE+abs(at[3535+(j-1)+15]^2-garch13(15,omega,alpha1,beta1,beta2,beta3 ,h[3535+j-1],h[3535+j-2],h[3535+j-3]))
  MSE=MSE+(at[3535+(j-1)+15]^2-garch13(15,omega,alpha1,beta1,beta2,beta3 ,h[3535+j-1],h[3535+j-2],h[3535+j-3]))^2
}
MAE=MAE/100
A.2.2 Modeling and Forecasting for KLCI

```r
h=c()
h[0]<-0
at=c()
at[0]<-0
dif=c()
diff=c()
numcase<-3880
predh=c()
MAE=0
MSE=0
MAPE=0

y<-read.csv("C:/Kualaluampur.csv",header=TRUE);y
y<-as.matrix(y)

fit_1= garchFit( y~arma(0,0)+garch(1,1), data = y[1:1500], cond.dist="std")
summary(fit_1)
# cond.dist="norm, snorm, std, sstd, ged, sged

para<-coef(garchFit(y~arma(0,0)+garch(1,1), data = y[1:1500],
cond.dist="std"))
para

mu<-para[1];mu
omega<-para[2];omega
alpha1<-para[3];alpha1
beta1<-para[4];beta1
at<-y-mu ;at

h[1]<- omega+alpha1*0+beta1*0
h[1]

# calculated volatility #

for (i in 2:numcase) {
  h[i]=omega+alpha1*at[i-1]^2+beta1*h[i-1]
}```
predh=function(current,forstep,omega,alpha1,beta1,currentat,currenth) {
  sig1=omega+alpha1*currentat^2+beta1*currenth
  sig2=(omega*(1-(alpha1+beta1)^(forstep-1)))/(1-alpha1-beta1)+(alpha1+beta1)^^(forstep-1)*sig1
  sig2
}

for (j in 1:100) {
  MAE= MAE+ abs(at[1500+j-1+1]^2-predh(1500+j-1,1,omega,alpha1,beta1,at[1500+j-1],h[1500+j-1]))
  MSE= MSE+ (at[1500+j-1+1]^2-predh(1500+j-1,1,omega,alpha1,beta1,at[1500+j-1],h[1500+j-1]))^2
}

MAE=MAE/100
MSE=MSE/100

MAE
MSE

A.2.3 Modeling and Forecasting for STI

h=c()
h[0]<-0
at=c()
at[0]<-0
numcase<-5407
predh=c()
MAE=0
MSE=0
MAPE=0

y<-read.csv("C:/Singapore.csv",header=TRUE);y
y<-as.matrix(y)
APPENDIX A. PROGRAM FILES

183

fit_1= garchFit(y~arma(0,0)+garch(2,1),data = y[1:2500],cond.dist="norm")
summary(fit_1)
# cond.dist="norm, snorm, std, sstd, ged, sged
para<-coef(garchFit(y~arma(0,0)+garch(2,1),data=y[1:2500],cond.dist="norm"))
para
mu<-para[1];mu
omega<-para[2];omega
alpha1<-para[3];alpha1
alpha2<-para[4];alpha2
beta1<-para[5] ;beta1
at<-y-mu ;at
h[1]<-omega+alpha1*0+alpha2*0+beta1*0
h[1]
###########################
# calculated volatility #
###########################
for (i in 2:numcase) {
h[i]=omega+alpha1*at[i-1]^2
+alpha2*at[i-2]^2+beta1*h[i-1]
}
h
#####################################
# Forecast Volatility k-step ahead #
#####################################
garch21<-function(forstep,omega,alpha1,alpha2,beta1,currenth,currenth_1)
{
vol<-matrix(c(0),nrow=(forstep+2),ncol=1)
vol[1]<-currenth_1
vol[2]<-currenth
for ( i in 3:(forstep+2))
{
vol[i]= omega+(alpha1+beta1)*vol[i-1]+alpha2*vol[i-2]
}
vol[forstep+2,]
}
garch12(15,para[2],para[3],para[4],para[5],h[2500],h[2499])
for (j in 1:100) {
MAE= MAE+ abs(at[2500+j-1+1]^2-garch21(1,omega,alpha1,alpha2,beta1,




```r
h[2500+j-1],h[2500+j-2]))
MSE= MSE+(at[2500+j-1+1]^2-garch21(1,omega,alpha1,alpha2,beta1,
h[2500+j-1],h[2500+j-2]))^2
}
MAE=MAE/100
MSE=MSE/100

A.3 Programs used in Chapter 5

A.3.1 Unit Root Tests and Cointegration Analysis among SET, KLCI and STI

R program for checking unit root tests and cointegration among SET, KLCI and STI in Tables 5.1 - 5.4.

```
returnx.adf = ur.df(returnx, type = "none", lags = 10)
summary(logx.adf)
summary(returnx.adf)

logy.adf = ur.df(logy, type = "none", lags = 10)
returny.adf = ur.df(returny, type = "none", lags = 9)
summary(logy.adf)
summary(returny.adf)

logz.adf = ur.df(logz, type = "none", lags = 1)
returnz.adf = ur.df(returnz, type = "none", lags = 12)
summary(logz.adf)
summary(returnz.adf)

# Cointegrate Johansen #

VARselect(data1, lag.max = 10, type = "none")
VARselect(data2, lag.max = 10, type = "none")
VARselect(data3, lag.max = 10, type = "none")
VARselect(data4, lag.max = 10, type = "none")

p1ct = VAR(data1, p=1, type = "none")
serial.test(p1ct, lags.pt = 15, type = "PT.asymptotic")
normality.test(p1ct)
arch.test(p1ct, lags.multi=1)
plot(stability(p1ct), nc=2)

H1.trace = ca.jo(data1, type=’trace’, ecdet=’none’, K=10, spec=’longrun’)
H1.eigen = ca.jo(data1, type=’eigen’, ecdet=’none’, K=5, spec=’longrun’)
H2.trace = ca.jo(data2, type=’trace’, ecdet=’none’, K=2, spec=’longrun’)
H2.eigen = ca.jo(data2, type=’eigen’, ecdet=’none’, K=2, spec=’longrun’)
H3.trace = ca.jo(data3, type=’trace’, ecdet=’none’, K=2, spec=’longrun’)
H3.eigen = ca.jo(data3, type=’eigen’, ecdet=’none’, K=2, spec=’longrun’)
H4.trace = ca.jo(data4, type=’trace’, ecdet=’none’, K=2, spec=’longrun’)
H4.eigen = ca.jo(data4, type=’eigen’, ecdet=’none’, K=2, spec=’longrun’)

H1.trace@teststat
H1.eigen@teststat
summary(H1.trace)
summary(H1.eigen)
summary(H2.trace)
summary(H2.eigen)
summary(H3.trace)
summary(H3.eigen)
summary(H4.trace)
summary(H4.eigen)

A.3.2 Simulation Studies on the Size of the Test for GARCH(1,1) Model

R program for simulation studies on the size of the test for 100 observations model in Tables 5.7 - 5.10.

set.seed(1)
x=c()
y=c()
t=c()
Ktest_t=c()
ttest_t=c()
whitetest=c()
dwteststat=c()
trace=c()
eigen=c()
trend=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()
Ktcri=c()} })
num=100
Ktcri_100 = -20.7
Ktcri_100 = -3.45
trace_cri = 25.32 #critical for trend
eigen_cri = 18.96 #critical for trend
trace_cri = 17.95 #critical for none
eigen_cri = 14.90 #critical for none
# trace_cri = 19.96 # critical for const
# eigen_cri = 15.67 # critical for const

ye = c()
x = c()

# ex <- garch.sim(alpha = c(0.3, 0.65), beta = c(0.05), n = 600, rnd = rnorm)
# ey <- garch.sim(alpha = c(0.3, 0.65), beta = c(0.05), n = 600, rnd = rnorm)

spec = garchSpec(model = list(omega = 0.1, alpha = 0.3, beta = 0.6),
cond.dist = "norm") # normal

spec = garchSpec(model = list(omega = 0.3, alpha = 0.65, beta = 0.05, shape = 3,
skew = 0.1), cond.dist = "snorm") # skew normal

spec = garchSpec(model = list(omega = 0.1, alpha = 0.3, beta = 0.6, shape = 9),
cond.dist = "std") # student-t

spec = garchSpec(model = list(omega = 0.3, alpha = 0.65, beta = 0.05, shape = 4.5,
skew = 3), cond.dist = "sstd") # skew student-t

spec = garchSpec(model = list(omega = 0.3, alpha = 0.65, beta = 0.05, shape = 3),
cond.dist = "ged") # GED

spec = garchSpec(model = list(omega = 0.3, alpha = 0.65, beta = 0.05, shape = 3,
skew = 3), cond.dist = "sged") # skew GED

b = garchSim(spec, n = 600)
c = garchSim(spec, n = 600)
a = t(b)
d = t(c)
ex = a[1,]
ey = d[1,]

#########################
## Calculate x ##
#########################
x[1] = 0 + ex[1]  # diffx = et

for (i in 2:600) {
    x[i] = x[i-1] + ex[i]
}

#########################
## Calculate y ##
#########################
y[1] = 0 + ey[1]  # diffy = et

for (i in 2:600) {

y[i] = y[i-1]+ey[i]

### Linear Comb of x,y use OLS [501-600] Obs ###
ordd=y[501:600]~x[501:600]
eer=residuals(lm(ordd))
eer_sq <- eer^2

# Dickey Fuller K-test and t-test #
#ye=diff(e) #delta e
for( i in 1:num-1) {
  xe[i]=eer[i]  #one lag of residual
  ye[i]=eer[i+1]
  t[i]=i+1
}
roll.lm_t = lm(ye~0+t+xe) #delta_et = trend+(Roll*e[t-1]+u
roll_t = summary(roll.lm_t)
roll_t_es = summary(roll.lm_t)$coefficients

K_test_t = num*(roll_t_es[2] - 1) ;K_test_t
t_test_t = (roll_t_es[2] - 1)/roll_t_es[4] ;t_test_t

Ktest_t[rep]=K_test_t
ttest_t[rep]=t_test_t
trend[rep]=roll_t_es[7]
drift[rep]=roll_t_es[10]

# White’test the DF-t test with white correction #
xn <- x[501:600]
yn <- y[501:600]

#r_sq = summary(whi)
#r_sq$r.squared
# Cointegrate Regression Durbin Watson test #
dw <- dwtest(lm(ordd))
dw$statistic

dwteststat[rep]=dw$statistic

# Cointegrate Johansen #
data1=cbind(xn, yn)
H1.trace=ca.jo(data1, type='trace', ecdet='none', K=2, spec='longrun')
H1.eigen=ca.jo(data1, type='eigen', ecdet='none', K=2, spec='longrun')
H1.trace@teststat
H1.eigen@teststat
trace[rep]=H1.trace@teststat[2]
eigen[rep]=H1.eigen@teststat[2]

Ktest_t
ttest_t
dwteststat
trace
eigen

# find proportion for Ktest#
Kcount_t_Reject=0
for (i in 1:length(Ktest_t))
  if (Ktest_t[i] < Kt_cri_100) Kcount_t_Reject=Kcount_t_Reject+1

# find proportion for t-test#
tcount_t_Reject=0
for (i in 1:length(ttest_t))
  if (ttest_t[i] < tt_cri_100) tcount_t_Reject=tcount_t_Reject+1

# find proportion for Ktest#
# find proportion for DW test#
# find proportion for DW test#
DW_Reject=0
for (i in 1:length(dwteststat))
  if (dwteststat[i] > 0.38) DW_Reject=DW_Reject+1

# find proportion for Trace Johansen#
# find proportion for Trace Johansen#
trace_Reject=0
for (i in 1:length(trace))
  if (trace[i] > trace_cri) trace_Reject=trace_Reject+1

# find proportion for Eigen Johansen#  
# find proportion for Eigen Johansen#
Eigen_Reject=0
for (i in 1:length(eigen))
  if (eigen[i] > eigen_cri) Eigen_Reject=Eigen_Reject+1

trace_Reject
Eigen_Reject
DW_Reject
Kcount_t_Reject
tcount_t_Reject
#trend_Reject
#drift_Reject
#Whitep_Reject
#Whitewithout_Reject

R program for simulation studies on the size of the test for 1,000 observations
model in Tables 5.7 - 5.10.

set.seed(1)
x=c()
y=c()
t=c()
Ktest_t=c(}
ttest_t=c()
whitetest=c()
dwteststat=c()
trend=c()
drift=c()
	no=10000
    for (rep in 1:no) {
        num=1000

        Kt_cri_1000 = -8.1
        tt_cri_1000 = -1.95

        #trace_cri = 25.32 #critical for trend
        #eigen_cri = 18.96 #critical for trend

        trace_cri = 17.95 #critical for none
        eigen_cri = 14.90 #critical for none

        #trace_cri = 19.96 #critical for const
        #eigen_cri = 15.67 #critical for const

        ye=c()
        xe=c()

        #ex <-garch.sim(alpha=c(0.3,0.65),beta=c(0.05),n=1500,rnd=rnorm)
        #ey <-garch.sim(alpha=c(0.3,0.65),beta=c(0.05),n=1500,rnd=rnorm)

        spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05),
                        cond.dist = "norm")#normal
        spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=3,skew=0.5),
                        cond.dist = "snorm")#skew normal
        spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=9),
                        cond.dist = "std")#student-t
        spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=4.5,skew=3),
                        cond.dist = "sstd")#skew student-t
        spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=3),
                        cond.dist = "ged")#GED
        spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=3,skew=0.1),
                        cond.dist = "sged")#skew GED

        b=garchSim(spec, n = 1500)
        c=garchSim(spec, n = 1500)
a=t(b)  
d=t(c)  
ex=a[1,]  
ey=d[1,]  

###############################  
## Calculate x ##  
###############################  
x[1]=0+ex[1]  # diffx = et  
for (i in 2:1500) {  
    x[i] = x[i-1]+ex[i]  
}  

###############################  
## Calculate y ##  
###############################  
y[1]=0+ey[1]  #diffy = et  
for (i in 2:1500) {  
    y[i] = y[i-1]+ey[i]  
}  

###########################################################  
## Linear Comb of x,y use OLS [501-600] Obs ##  
###########################################################  
ordd=y[501:1500]~x[501:1500]  
eer=residuals(lm(ordd))  
eer_sq <- eer^2  

###########################################################  
# Dickey Fuller  K-test and t-test #  
###########################################################  
#ye=diff(e)  #delta e  
for( i in 1:num-1) {  
    xe[i]=eer[i]  #one lag of residual  
    ye[i]=eer[i+1]  
    t[i]=i+1  
}  

roll.lm_t = lm(ye~0+xe)  #delta_et = trend+(Roll*e[t-1]+u  
roll_t = summary(roll.lm_t)  
roll_t_es = summary(roll.lm_t)$coefficients
K_test_t = num*(roll_t_es[1] - 1) ;K_test_t

Ktest_t[rep]=K_test_t
ttest_t[rep]=t_test_t
trend[rep]=roll_t_es[7]
drift[rep]=roll_t_es[10]

# Cointegrate Regression Durbiin Watson test #
dw <- dwtest(lm(ordd))
dw$statistic

dwteststat[rep]=dw$statistic

# Cointegrate Johansen #
data1=cbind(xn, yn)
H1.trace=ca.jo(data1, type='trace', ecdet='none', K=2, spec='longrun')
H1.eigen=ca.jo(data1, type='eigen', ecdet='none', K=2, spec='longrun')
H1.trace@teststat
H1.eigen@teststat
trace[rep]=H1.trace@teststat[2]
eigen[rep]=H1.eigen@teststat[2]

# find proportion for Ktest #
Kcount_t_Reject=0

for (i in 1:length(Ktest_t))
    if (Ktest_t[i] < Kt_cri_1000) Kcount_t_Reject=Kcount_t_Reject+1
# find proportion for ttest#
tcount_t_Reject=0
for (i in 1:length(ttest_t))
  if (ttest_t[i] < tt_cri_1000) tcount_t_Reject=tcount_t_Reject+1

# find proportion for DW test#
DW_Reject=0
for (i in 1:length(dwteststat))
  if (dwteststat[i] > 0.04) DW_Reject=DW_Reject+1

# find proportion for Trace Johansen#
trace_Reject=0
for (i in 1:length(trace))
  if (trace[i] > trace_cri) trace_Reject=trace_Reject+1

# find proportion for Eigen Johansen#
Eigen_Reject=0
for (i in 1:length(eigen))
  if (eigen[i] > eigen_cri) Eigen_Reject=Eigen_Reject+1

trend_Reject=0
for (i in 1:length(trend))
  if (trend[i] <= 0.05 ) trend_Reject=trend_Reject+1

drift_Reject=0
for (i in 1:length(drift))
  if (drift[i] <= 0.05 ) drift_Reject=drift_Reject+1

trace_Reject
Eigen_Reject
R program for simulation studies on the power of the test for 100 observations model in Tables 5.11 - 5.14.

```r
set.seed(1)
x=c()
y=c()
t=c()
trend=c()
Ktest=c()
Ktest_t=c()
ttest=c()
ttest_t=c()
whitetest=c()
dwteststat=c()
trace=c()
eigen=c()

#################
# Start Looping #
#################

no=10000
  for (rep in 1:no) {
    num=100

    Kt_cri_100 = -7.9
    tt_cri_100 = -1.95

    #trace_cri = 25.32 #critical for trend
    #eigen_cri = 18.96 #critical for trend

    trace_cri = 17.95 #critical for none
    eigen_cri = 14.90 #critical for none

    #trace_cri = 19.96 #critical for const
    #eigen_cri = 15.67 #critical for const

    ye=c()
```
xe=c()
#ex <-garch.sim(alpha=c(0.3,0.65),beta=c(0.05),n=600,rnd=rnorm)
#ey <-garch.sim(alpha=c(0.3,0.65),beta=c(0.05),n=600,rnd=rnorm)

#spec = garchSpec(model = list(omega = 0.3, alpha = 0.65, beta = 0.05),
cond.dist = "norm")#normal
#spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=3,skew=3),
cond.dist = "snorm")#skew normal
#spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=4.5),
cond.dist="std")#student-t
#spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=4.5,skew=3),
cond.dist = "sstd")#skew student-t
#spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=3),
cond.dist="ged")#GED
spec=garchSpec(model=list(omega=0.3,alpha=0.65,beta=0.05,shape=3,skew=3),
cond.dist="sged")#skew GED

b=garchSim(spec, n = 600)
c=garchSim(spec, n = 600)
a=t(b)
d=t(c)
ex=a[1,]
ey=d[1,]

# Calculate x
for (i in 2:600) {
    x[i] = x[i-1]+ex[i]
}

# Calculate y
for (i in 2:600) {
    y[i] = 0.8*y[i-1]+0.2*x[i-1]+ey[i]
}

# Linear Comb of x,y use OLS [501-600] Obs #
### Appendix A. Program Files

```r
# Program code

ordd = y[501:600] - x[501:600]
eer = residuals(lm(ordd))
eer_sq <- eer^2

# Dickey Fuller K-test and t-test#

# ye = diff(e) # delta e
for( i in 1:num-1) {
  xe[i] = eer[i] # one lag of residual
  ye[i] = eer[i+1]
  t[i] = i+1
}

roll.lm_t = lm(ye~0+xe) # delta et = trend+(Roll*e[t-1]+u
roll_t = summary(roll.lm_t)
roll_t_es = summary(roll.lm_t)$coefficients

K_test_t = num*(roll_t_es[1] - 1) ; K_test_t
T_test_t = (roll_t_es[1] - 1)/roll_t_es[2] ; t_test_t

trend[rep]=roll_t_es[7]
Ktest_t[rep]=K_test_t
ttest_t[rep]=t_test_t

# Cointegrate Regression Durbin Watson test#

dw <- dwtest(lm(ordd))
dw$statistic
dwteststat[rep]=dw$statistic

# Cointegrate Johansen#

data1 = cbind(xn, yn)
H1.trace=ca.jo(data1, type='trace', ecdet='none', K=2, spec='longrun')
H1.eigen=ca.jo(data1, type='eigen', ecdet='none', K=2, spec='longrun')
H1.trace@teststat
H1.eigen@teststat
trace[rep]=H1.trace@teststat[2]
eigen[rep]=H1.eigen@teststat[2]
```
}  

# find proportion for Ktest #
Kcount_t_Reject=0

for (i in 1:length(Ktest_t))
  if (Ktest_t[i] < Kt_cri_100) Kcount_t_Reject=Kcount_t_Reject+1

# find proportion for ttest #
tcount_t_Reject=0

for (i in 1:length(ttest_t))
  if (ttest_t[i] < tt_cri_100) tcount_t_Reject=tcount_t_Reject+1

# find proportion for Trace Johansen #
trace_Reject=0

for (i in 1:length(trace))
  if (trace[i] > trace_cri) trace_Reject=trace_Reject+1

trace_Reject

# find proportion for Eigen Johansen #
Eigen_Reject=0

for (i in 1:length(eigen))
  if (eigen[i] > eigen_cri) Eigen_Reject=Eigen_Reject+1

Eigen_Reject

# find proportion for DW test #
DW_Reject=0

for (i in 1:length(dwteststat))
  if (dwteststat[i] > 0.38) DW_Reject=DW_Reject+1
APPENDIX A. PROGRAM FILES

DW_Reject

#############################
# find proportion for roll #
#############################
trend_Reject=0

for (i in 1:length(trend))
  if (trend[i] < 0.05) trend_Reject=trend_Reject+1

trace_Reject
Eigen_Reject
DW_Reject
Kcount_t_Reject
tcount_t_Reject

R program for simulation studies on the power of the test for 1,000 observations model in Tables 5.11 - 5.14.

set.seed(1)
x=c()
y=c()
t=c()
trend=c()
Ktest=c()
Ktest_t=c()
ttest=c()
ttest_t=c()
whitetest=c()
dwteststat=c()
trace=c()
eigen=c()

#############################
# Start Looping #
#############################

no=10000
  for (rep in 1:no) {
    num=1000

    Kt_cri_1000 = -21.8
  }
tt_cri_1000 = -3.41

#trace_cri = 25.32 #critical for trend
#eigen_cri = 18.96 #critical for trend

trace_cri = 17.95 #critical for none
eigen_cri = 14.90 #critical for none

#trace_cri = 19.96 #critical for const
#eigen_cri = 15.67 #critical for const

ye=c()
xz=c()
# Calculate y #

```r
y[1] = (0.8*0) + (0.2*0) + ey[1]

for (i in 2:1500) {
  y[i] = 0.8*y[i-1] + 0.2*x[i-1] + ey[i]
}
```

# Linear Comb of x,y use OLS [501-600] Obs #

```r
ordd = y[501:1500] ~ x[501:1500]
eer = residuals(lm(ordd))
eer_sq <- eer^2
```

# Dickey Fuller K-test and t-test #

```r
# ye = diff(e) # delta e
for (i in 1:num-1) {
  xe[i] = eer[i] # one lag of residual
  ye[i] = eer[i+1]
  t[i] = i+1
}

roll.lm_t = lm(ye~0+t+xe) # delta et = trend + (Roll*e[t-1]+u
roll_t = summary(roll.lm_t)
roll_t_es = summary(roll.lm_t)$coefficients

K_test_t = num*(roll_t_es[2] - 1) ; K_test_t
t_test_t = (roll_t_es[2] - 1)/roll_t_es[4] ; t_test_t

trend[rep] = roll_t_es[7]
Ktest_t[rep] = K_test_t
ttest_t[rep] = t_test_t
```

# Cointegrate Regression Durbin Watson test #

```r
dw <- dwtest(lm(ordd))
dw$statistic
```
# Cointegrate Johansen #

data1=cbind(xn, yn)
H1.trace=ca.jo(data1, type='trace', ecdet='none', K=2, spec='longrun')
H1.eigen=ca.jo(data1, type='eigen', ecdet='none', K=2, spec='longrun')
H1.trace@teststat
H1.eigen@teststat
trace[rep]=H1.trace@teststat[2]
eigen[rep]=H1.eigen@teststat[2]

# find proportion for Ktest #
Kcount_t_Reject=0
for (i in 1:length(Ktest_t))
  if (Ktest_t[i] < Kt_cri_1000) Kcount_t_Reject=Kcount_t_Reject+1

# find proportion for ttest #
tcount_t_Reject=0
for (i in 1:length(ttest_t))
  if (ttest_t[i] < tt_cri_1000) tcount_t_Reject=tcount_t_Reject+1

# find proportion for Trace Johansen #
trace_Reject=0
for (i in 1:length(trace))
  if (trace[i] > trace_cri) trace_Reject=trace_Reject+1

trace_Reject

# find proportion for Eigen Johansen #
Eigen_Reject=0

for (i in 1:length(eigen))
  if (eigen[i] > eigen_cri) Eigen_Reject=Eigen_Reject+1

Eigen_Reject

########################################################################
# find proportion for DW test #
########################################################################
DW_Reject=0

for (i in 1:length(dwteststat))
  if (dwteststat[i] > 0.04) DW_Reject=DW_Reject+1

DW_Reject

########################################################################
# find proportion for White test with White correction #
########################################################################
Whitep_Reject=0

for (i in 1:length(whitetest))
  if (whitetest[i] > 5.99) Whitep_Reject=Whitep_Reject+1

########################################################################
# find proportion for roll #
########################################################################
trend_Reject=0

for (i in 1:length(trend))
  if (trend[i] < 0.05) trend_Reject=trend_Reject+1

trace_Reject
Eigen_Reject
DW_Reject
Kcount_t_Reject
tcount_t_Reject

R program for simulation studies on the size of the test for 100 and 1,000
observations model using Wild Bootstrap approach in Tables 5.7 - 5.10.

set.seed(1)
x=c()
y=c()
ep=c()

aa=matrix(0,1000,number)
eer=c()
eer_sq=c()
ex=c()
ey=c()

num=1000

#sumeer=c()
hterm=c()
eer_sta=c()
eer_stasq=c()
pro=c()
al_pro=c()
prosta=c()
al_prosta=c()
tao_sta=c()
tao_r=c()

LL=c()
RR=c()

#Mac_cri_0.05 = -3.39827 #T=100
Mac_cri_0.05 = -3.34368 #T=1000

###########
# Looping #
###########

number = 10000

for (re in 1:number) {

#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6),
    cond.dist="norm")#normal
spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=3, skew=3),
    cond.dist="snorm")#skew normal
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=4.5),
    cond.dist="std")#student-t
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=9, skew=3),
    cond.dist="sstd")#skew student-t
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=3),
    cond.dist="ged")#GED
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=3, skew=0.1),
    cond.dist="sged")#skew GED

b=garchSim(spec, n = 1500)
c = garchSim(spec, n = 1500)
a = t(b)
d = t(c)
ex = a[1,]
ey = d[1,]

# Calculate x
for (i in 2:1500) {
    x[i] = x[i-1] + ex[i]
}

# Calculate y
for (i in 2:1500) {
    y[i] = y[i-1] + ey[i]
}

# Linear Comb of x, y use OLS [501-600] Obs
ordd = y[501:1500] ~ x[501:1500]
aa[, re] = residuals(lm(ordd))

# number = 5
for (re in 1:number) {
    eer = aa[, re]
    eer_sq <- eer^2
    ep[1] = eer[1]
    for (i in 2:num) {
        ep[i] = eer[i] - eer[i-1]
    }
}
total_p=0 #
total_x=0 #

for (i in 2:num) {
  pro[i] = eer[i]*eer[i-1]
  total_p = total_p + pro[i]
  total_x = total_x + eer_sq[i-1]
}

alpha = total_p/total_x

total_alp=0 #

for (i in 2:num) {
  al_pro[i] = (eer[i]-alpha*eer[i-1])^2
  total_alp = total_alp + al_pro[i]
}

S_sq = (1/(num-2))*total_alp
S = sqrt(S_sq)

tao = ((alpha-1)/S)*(total_x)^(1/2)
tao_r[re] = tao

# for (j in 1:num) {
#  sumeer[j] = sum(eer_sq[1:j])
#}

hterm[1] = 0

for (i in 2:num) {
  hterm[i] = eer[i]^2/total_x
}

no=1500
for (rep in 1:no) {

#################################
## calculate hterm ##
#################################

e_st = rnorm(num,0,1)

#################################
# Calculate x_sta ##
#################################


for (k in 2:num) {
    eer_st[k] = eer_st[k-1] + ((abs(ep[k])*e_st[k])/sqrt(1-hterm[k]))
eer_stsq[k] = eer_st[k]^2
}

#################################
# Calculate Tao_sta #
#################################

total_psta=0 #
total_xsta=0 #

for (i in 2:num) {
    prosta[i] = eer_st[i]*eer_st[i-1]
total_psta = total_psta + prosta[i]
total_xsta = total_xsta + eer_stsq[i-1]
}

alpha_st = total_psta/total_xsta

total_alpsta=0 #

for (i in 2:num) {
    al_prosta[i] = (eer_st[i]-alpha_st*eer_st[i-1])^2
    total_alpsta = total_alpsta + al_prosta[i]
}

S_sqsta = (1/(num-2))*total_alpsta
S_st = sqrt(S_sqsta)
taosta = ((alpha_sta-1)/S_sta)*(total_xsta)^(1/2)

tao_sta[rep] = taosta

}  
#LL[re] = quantile(tao_sta, probs=.05)
LL[re] = quantile(tao_sta, probs=.025)

}

############################################################
# find proportion #
############################################################
Bcount_Reject = 0
for (i in 1:length(tao_r)) {
  if ((tao_r[i] < Mac_cri_0.05)&&(tao_r[i]< LL[i]))
    #if ((tao_r[i] < Mac_cri_0.05)||(tao_r[i]< LL[i]))
      Bcount_Reject = Bcount_Reject+1
}

Bcount_Reject

R program for simulation studies on the power of the test for 100 and 1,000 observations model using Wild Bootstrap approach in Tables 5.11 - 5.14.

set.seed(1)
x=c()
y=c()
ep=c()
aa=matrix(0,100,number)
eer=c()
eer_sq=c()
ex=c()
ey=c()
num=100
#sumeer=c()
hterm=c()
eer_sta=c()
eer_stasq=c()
pro=c()
al_pro=c()
prosta=c()
al_prosta=c()
tao_sta=c()
tao_r=c()
LL=c()
RR=c()

#Mac_cri_0.05 = -3.39827 #T=100
Mac_cri_0.05 = -3.34368 #T=1000

###########
# Looping #
###########
number = 10000

for (re in 1:number) {

#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6),
    #cond.dist="norm")#normal
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=3, skew=0.5),
    #cond.dist="snorm")#skew normal
spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=9),
    #cond.dist="std")#student-t
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=4.5, skew=3),
    #cond.dist="sstd")#skew student-t
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=3),
    #cond.dist="ged")#GED
#spec=garchSpec(model=list(omega=0.1, alpha=0.3, beta=0.6, shape=3, skew=3),
    #cond.dist="sged")#skew GED

b=garchSim(spec, n = 600)
c=garchSim(spec, n = 600)
a=t(b)
d=t(c)
ex=a[1,]
ey=d[1,]

#################
## Calculate x ##
#################
x[1]=0+ex[1] # diffx = et

for (i in 2:600) {
    x[i] = x[i-1]+ex[i]
}
## Calculate y ##

\[ y[1] = (0.8*0) + (0.2*0) + e[y][1] \]

for (i in 2:600) {
    \[ y[i] = 0.8*y[i-1] + 0.2*x[i-1] + e[y][i] \]
}

## Linear Comb of x, y use OLS [501-600] Obs ##

ordd = y[501:600] ~ x[501:600]

aa[, re] = residuals(lm(ordd))

# aa

# aa = read.csv(file="eer-5.csv")

# number=5

for (re in 1:number) {
    eer = aa[, re]
    eer_sq <- eer^2
    ep[1] = eer[1]
    for (i in 2:num) {
        ep[i] = eer[i] - eer[i-1]
    }
}

# Calculate Tao #

total_p = 0 #

for (i in 2:num) {
    pro[i] = eer[i] * eer[i-1]
    total_p = total_p + pro[i]
    total_x = total_x + eer_sq[i-1]
}

alpha = total_p / total_x
total_alp=0 #
for (i in 2:num) {
    al_pro[i] = (eer[i]-alpha*eer[i-1])^2
    total_alp = total_alp + al_pro[i]
}
S_sq = (1/(num-2))*total_alp
S = sqrt(S_sq)
tao = ((alpha-1)/S)*(total_x)^(1/2)
tao_r[re] = tao

#for (j in 1:num) {
#sumeer[j] = sum(eer_sq[1:j])
#}
hterm[1] = 0
for (i in 2:num) {
    hterm[i] = eer[i]^2/total_x
}

##############
# Looping #
##############
no=1000
for (rep in 1:no) {

    ##########################
    ## calculate hterm ##
    ##########################
    e_sta = rnorm(num,0,1)

    ##########################
    # Calculate x_sta ##
    ##########################
for (k in 2:num) {
    eer_stasq[k] = eer_sta[k-1]^2
    eer_sta[k] = eer_sta[k] + ((abs(ep[k])*e_sta[k])/sqrt(1-hterm[k]))
}

#Calculate Tao_stas#

for (i in 2:num) {
    prosta[i] = eer_sta[i]*eer_sta[i-1]
    total_psta = total_psta + prosta[i]
    total_xsta = total_xsta + eer_stasq[i-1]
}

alpha_sta = total_psta/total_xsta

for (i in 2:num) {
    al_prosta[i] = (eer_sta[i]-alpha_sta*eer_sta[i-1])^2
    total_alpsta = total_alpsta + al_prosta[i]
}

S_sqsta = (1/(num-2))*total_alpsta
S_sta = sqrt(S_sqsta)

taosta = ((alpha_sta-1)/S_sta)*(total_xsta)^{1/2}

tao_sta[rep] = taosta

#LL[re] = quantile(tao_sta, probs=.05)
LL[re] = quantile(tao_sta, probs=.025)
APPENDIX A. PROGRAM FILES

for (i in 1:length(tao_r)) {
    if ((tao_r[i] < Mac_cri_0.05) && (tao_r[i] < LL[i]))
        #if ((tao_r[i] < Mac_cri_0.05) || (tao_r[i] < LL[i]))
            Bcount_Reject = Bcount_Reject + 1
    }
}
Bcount_Reject

A.4 Programs used in Chapter 6

A.4.1 Unit Root Tests and Cointegration Analysis among SET, KLCI and STI

R program for simulation studies on the performance of volatility forecasting comparisons between Model A and Model B in Tables 6.1 - 6.16.

set.seed(1000)
num = 10000
x = c()
y = c()
ex = c()
ey = c()
h = c()
h1 = c()
rt1 = c()
rt2 = c()
predh = c()
MAE1 = 0
MSE1 = 0
MAE2 = 0
MSE2 = 0
s = 10
a10 = 0.1
a11 = 0.1
be1 = 0.8
#ex = garch.sim(alpha = c(a10, a11), beta = c(be1), n = 10000, rnd = rnorm)
#ey = garch.sim(alpha = c(a10, a11), beta = c(be1), n = 10000, rnd = rnorm)
#spec = garchSpec(model = list(omega = a10, alpha = a11, beta = be1),
cond.dist="norm")#normal
#spec=garchSpec(model=list(omega=al0,alpha=al1,beta=be1,shape=3,skew=3),
cond.dist="snorm")#skew normal
#spec=garchSpec(model=list(omega=al0,alpha=al1,beta=be1,shape=9),
cond.dist="std")#student-t
#spec=garchSpec(model=list(omega=al0,alpha=al1,beta=be1,shape=9,skew=3),
cond.dist="sstd")#skew student-t
#spec=garchSpec(model=list(omega=al0,alpha=al1,beta=be1,shape=3),
cond.dist="ged")#GED
spec=garchSpec(model=list(omega=al0,alpha=al1,beta=be1,shape=3,skew=3),
cond.dist="sged")#skew GED

b=garchSim(spec, n = 10000)
c=garchSim(spec, n = 10000)
a=t(b)
d=t(c)
ex=a[1,]
ey=d[1,]

#################
## Generate x ##
#################
x[1]=0+ex[1]  # diffx = et
for (i in 2:10000) {
   x[i] = x[i-1]+ex[i]
}

#################
## Generate y ##
#################
y[1]=(0.8*0)+(0.2*0)+ey[1]
for (i in 2:10000) {
   y[i] = 0.8*y[i-1]+0.2*x[i-1]+ey[i]
}
df = diff(y)

#par(mfrow=c(2,2))
#plot(diff(x), type='l')
#plot(diff(y), type='l')
#plot(df, type='l')
#plot(y, type='l')
### Fit y by GARCH ###

```r
fitt = garchFit( df~arma(0,0)+garch(1,1), data = df[1:5000], cond.dist="sged")
summary(fitt)
```

```r
para = coef(garchFit(df~arma(0,0)+garch(1,1), data = df[1:5000],
cond.dist="sged"))
mu = para[1]
omega = para[2]
alpha1 = para[3]
beta1 = para[4]
```

### calculated volatility ###

```r
h[1] = omega+alpha1*0+beta1*0
h1[1] = al0+al1*0+be1*0
rt1[1] = y[1]-0.8*0-0.2*0
rt2[1] = y[1]-0
```

```r
for (i in 2:10000) {
  h[i]=omega+alpha1*ey[i-1]^2+beta1*h[i-1]
  h1[i]=al0+al1*ey[i-1]^2+be1*h1[i-1]
  rt1[i]=y[i]-0.8*y[i-1]-0.2*x[i-1]
  rt2[i]=y[i]-y[i-1]
}
```

### Forecast Volatility k-step ahead of y ###

```r
predh=function(current,forstep,al0,al1,be1,currentepy,currentth) {
```
sig11 = al0 + al1 * currentepy^2 + be1 * currenth
sig21 = (al0*(1-(al1+be1)^((forstep-1))/(1-al1-be1))
        + (al1+be1)^(forstep-1)*sig11
}

for (i in 1:100) {
    MAE1 = MAE1 + abs((rt1[5000+i+s]-mean(rt1))^2-
                     predh(5000+i-1+1,s,al0,al1,be1,rt1[5000+i-1+s],h1[5000+i-1+s]))
    MSE1 = MSE1 + ((rt1[5000+i+s]-mean(rt1))^2-
                     predh(5000+i-1+1,s,al0,al1,be1,rt1[5000+i-1+s],h1[5000+i-1+s]))^2
}

MAE1 = MAE1/100
MSE1 = MSE1/100

#########################################
# Forecast Volatility k-step ahead of y #
#########################################

predh = function(current, forstep, omega, alpha1, beta1, currentepy, currenth){
    sig12 = omega + alpha1 * currentepy^2 + beta1 * currenth
    sig22 = (omega*(1-(alpha1+beta1)^(forstep-1))/(1-alpha1-beta1))
            + (alpha1+beta1)^(forstep-1)*sig12
    sig22
}

for (i in 1:100) {
    MAE2 = MAE2 + abs((rt2[5000+i+s]-mean(rt2))^2-
                     predh(5000+i-1+1,s,omega,alpha1,beta1,rt2[5000+i-1+s],h[5000+i-1+s]))
    MSE2 = MSE2 + ((rt2[5000+i+s]-mean(rt2))^2-
                     predh(5000+i-1+1,s,omega,alpha1,beta1,rt2[5000+i-1+s],h[5000+i-1+s]))^2
}

MAE2 = MAE2/100
MSE2 = MSE2/100

set.seed(1000)
X=c()
Y=c()
ex=c()
ey=c()
h=c()
rt1=c()
rt2=c()
predh=c()
s=15
# s = 1, 2, 10, 15
MAE1=0
MSE1=0
MAE2=0
MSE2=0
SET=c()
KLCI=c()
x<-read.csv("C:/KLCI daily.csv",header=TRUE);x
y<-read.csv("C:/SET daily.csv",header=TRUE);y
x<-as.matrix(x)
y<-as.matrix(y)

X = x[,3] #X=KLCI
Y = y[,3] #Y=SET

# Fit rt 1:1500 #

fit_1= garchFit( Y~arma(0,0)+garch(1,1), data = Y[1:500], cond.dist="sstd")
summary(fit_1)

para<-coef(garchFit(Y~arma(0,0)+garch(1,1), data = rt[1:500],
cond.dist="sged"))
para
mu<-para[1];mu
ome<-para[2];ome
alpha<-para[3];alpha
beta1<-para[4];beta1
shape<-para[5];sha
skew<-para[6];skewed

spec=garchSpec(model=list(omega=ome,alpha=alpha,beta=beta1,shape=sha,
,skew=skewed),cond.dist="sstd")
b=garchSim(spec, n = 1133)
c=garchSim(spec, n = 1133)
a=t(b)
d=t(c)
ex=a[1,]
ey=d[1,]

# calculated volatility #

h[1]= omega+alpha1*0+beta1*0

rt1[1]= X[1]-0.9699*0-0.0532*0+0.0143*0-0.0711*0+0.2629*0-0.2430*0
       +0.0601*0-1.0293*0+0.0217*0-0.0515*0+0.0661*0-0.0003*0
       -0.0395*0+0.0178*0-0.0257*(0-(1.3276*0))
rt1[2]= X[2]-0.9699*X[1]-0.0532*0+0.0143*0-0.0711*0+0.2629*0
       -0.2430*0+0.0601*0-1.0293*Y[1]+0.0217*0-0.0515*0+0.0661*0
       -0.0003*0-0.0395*0+0.0178*0-0.0257*(0-(1.3276*0))
rt1[3]= X[3]-0.9699*X[2]-0.0532*X[1]+0.0143*X[1]-0.0711*0+0.2629*0
       -0.2430*0+0.0601*0-1.0293*Y[2]+0.0217*Y[1]-0.0515*0+0.0661*0
       -0.0003*0-0.0395*0+0.0178*0-0.0257*(0-(1.3276*0))
rt1[4]= X[4]-0.9699*X[3]-0.0532*X[2]+0.0143*X[2]-0.0711*0+0.2629*0
       -0.2430*0+0.0601*0-1.0293*Y[3]+0.0217*Y[2]-0.0515*Y[1]+0.0661*0
       -0.0003*0-0.0395*0+0.0178*0-0.0257*(0-(1.3276*0))
rt1[5]= X[5]-0.9699*X[4]-0.0532*X[3]+0.0143*X[3]-0.0711*X[1]+0.2629*0
       -0.2430*0+0.0601*0-1.0293*Y[4]+0.0217*Y[3]-0.0515*Y[2]+0.0661*Y[1]
       -0.0003*0-0.0395*0+0.0178*0-0.0257*(0-(1.3276*0))
rt1[6]= X[6]-0.9699*X[5]-0.0532*X[4]+0.0143*X[4]-0.0711*X[2]+0.2629*X[1]
       -0.2430*0+0.0601*0-1.0293*Y[5]+0.0217*Y[4]-0.0515*Y[3]+0.0661*Y[2]
       -0.0003*Y[1]-0.0395*0+0.0178*0-0.0257*(0-(1.3276*0))
rt1[7]= X[7]-0.9699*X[6]-0.0532*X[5]+0.0143*X[5]-0.0711*X[3]+0.2629*X[2]
       -0.2430*X[1]+0.0601*0-1.0293*Y[6]+0.0217*Y[5]-0.0515*Y[4]+0.0661*Y[3]
       -0.0003*Y[2]-0.0395*Y[1]+0.0178*0-0.0257*(0-(1.3276*0))
rt1[8]= X[8]-0.9699*X[7]-0.0532*X[6]+0.0143*X[6]-0.0711*X[4]+0.2629*X[3]
       -0.2430*X[2]+0.0601*X[1]-1.0293*Y[7]+0.0217*Y[6]-0.0515*Y[5]
       +0.0661*Y[4]-0.0003*Y[3]-0.0395*Y[2]+0.0178*Y[1]
       -0.0257*(0-(1.3276*0))
rt2[1] = X[1]-0

for (j in 9:1133) {
  h[j] = omega+alpha1*ex[j-1]^2+beta1*h[j-1]
  rt1[j] = X[j]-0.9699*X[j-1]-0.0532*X[j-2]+0.0143*X[j-3]-0.0711*X[j-4]
           +0.2629*X[j-5]-0.2430*X[j-6]+0.0601*X[j-7]+0.0217*Y[j-1]
           +0.0217*Y[j-2]-0.0515*Y[j-3]+0.0661*Y[j-4]-0.0003*Y[j-5]
           -0.0395*Y[j-6]+0.0178*Y[j-7]-0.0257*(X[j-8]-1.3276*Y[j-8])
  rt2[j] = X[j]-X[j-1]
}

# Forecast Volatility k-step ahead of y #

predh=function(current,forstep,omega,alpha1,beta1,currentepy,currenth) {
  sig1=omega+alpha1*currentepy^2+beta1*currenth
  sig2=(omega*(1-(alpha1+beta1)^(forstep-1))/(1-alpha1-beta1))
       +(alpha1+beta1)^(forstep-1)*sig1
  sig2
}

for (i in 1:100) {
  MAE1= MAE1+ abs((rt1[500+i+s]-mean(rt1))^2
                 -predh(500+i-1+s,omega,alpha1,beta1,rt1[500+i-1+s],h[500+i-1+s]))
  MSE1= MSE1+ ((rt1[500+i+s]-mean(rt1))^2
              -predh(500+i-1+s,omega,alpha1,beta1,rt1[500+i-1+s],h[500+i-1+s]))^2
}
MAE1=MAE1/100
MSE1=MSE1/100

# Forecast Volatility k-step ahead of y #

predh=function(current,forstep,omega,alpha1,beta1,currentepy,currenth) {
\[
sig_1 = \omega + \alpha_1 \text{current} \times \text{current}^2 + \beta_1 \text{current} \\
\sig_2 = (\omega \times (1 - (\alpha_1 + \beta_1)^{\text{forstep}-1}) / (1 - \alpha_1 - \beta_1)) \\
\text{ } + (\alpha_1 + \beta_1)^{\text{forstep}-1} \times \sig_1 \\
\]

\[
\text{for } (i \text{ in } 1:100) \{ \\
\MAE_2 = \MAE_2 + \text{abs}((\text{rt2}[500+i+s] - \text{mean(\text{rt2}))^2 \\
\text{ } - \text{predh}(500+i-1+s, \omega, \alpha_1, \beta_1, \text{rt2}[500+i-1+s], \text{h}[500+i-1+s])) \\
\MSE_2 = \MSE_2 + ((\text{rt2}[500+i+s] - \text{mean(\text{rt2}))^2 \\
\text{ } - \text{predh}(500+i-1+s, \omega, \alpha_1, \beta_1, \text{rt2}[500+i-1+s], \text{h}[500+i-1+s]))^2 \\
\}
\]

\[
\MAE_2 = \MAE_2 / 100 \\
\MSE_2 = \MSE_2 / 100 \\
\]

\section*{A.5 Programs used in Chapter 7}

\subsection*{A.5.1 Value at Risk Estimations}

R program for results of VaR estimations for \textit{SET} and \textit{KLCI} at 1- and 2-steps ahead in Tables 7.2.

\[
\text{x<-read.csv("C:/3STOCK_LOGP.csv",header=TRUE)} \\
\text{SET = x[,4]} \\
\text{KLCI = x[,5]} \\
\text{VaR_SET1 = c()} \\
\text{VaR_KLCI1 = c()} \\
\text{RepVaR_SET2 = c()} \\
\text{RepVaR_KLCI2 = c()} \\
\text{for (k in 1:632) \{ \\
\text{fit_SET= garchFit(SET~arma(0,0)+garch(1,1), data = SET[k:499+k], cond.dist="sstd")} \\
\text{summary(fit_SET)} \\
\text{para_SET<-coef(garchFit(SET~arma(0,0)+garch(1,1), data = SET[k:499+k],} \\
\]}
cond.dist="sstd")

para_SET
mu_SET<-para_SET[1];mu_SET
omega_SET<-para_SET[2]
alpha_SET<-para_SET[3]
beta1_SET<-para_SET[4]
skew_SET<-para_SET[5]
shape_SET<-para_SET[6]

fit_KLCI= garchFit(KLCI~arma(0,0)+garch(1,1), data = KLCI[k:499+k],
cond.dist="sstd")
summary(fit_KLCI)
para_KLCI<-coef(garchFit(KLCI~arma(0,0)+garch(1,1), data = KLCI[k:499+k],
cond.dist="sstd"))

para_KLCI
mu_KLCI<-para_KLCI[1];mu_KLCI
omega_KLCI<-para_KLCI[2]
alpha_KLCI<-para_KLCI[3]
beta1_KLCI<-para_KLCI[4]
skew_KLCI<-para_KLCI[5]
shape_KLCI<-para_KLCI[6]

# 1 step forecast #
set.seed(10)
hes1_SET = c()
hes1_KLCI= c()
hes2_SET = c()
hes2_KLCI= c()
h1_SET = c()
h1_KLCI = c()
h2_SET = c()
h2_KLCI = c()

R_SET1 = c()
R_KLCI1 = c()
R_SET2 = c()
R_KLCI2 = c()

h1_SET = volatility(fit_SET, type="h")
h1_KLCI = volatility(fit_KLCI, type="h")
# Start Looping #

no = 10000
eta1_SET = rsstd(no)
eta2_SET = rsstd(no)
eta1_KLCI = rsstd(no)
eta2_KLCI = rsstd(no)


for (i in 1:no) {

R_SET1[i] = eta12_SET * sqrt(hes1_SET)
R_KLCI1[i] = eta12_KLCI * sqrt(hes1_KLCI)
}

return_SET1 = sort(R_SET1)
VaR_SET = quantile(return_SET1,0.05)
VaR_SET
return_KLCI1 = sort(R_KLCI1)
VaR_KLCI = quantile(return_KLCI1,0.05)
VaR_KLCI

VaR_SET1[k] = VaR_SET
VaR_KLCI1[k] = VaR_KLCI

# 2 step forecast #

eta22_SET = rsstd(no)
eta22_KLCI = rsstd(no)


for (i in 1:no) {
RSET2[i] = eta22_SET * sqrt(hes2_SET)
RKLCI2[i] = eta22_KLCI * sqrt(hes2_KLCI)
}

return_SET2 = sort(R_SET2)
VaR_SET2 = quantile(return_SET2,0.05)
VaR_SET2

return_KLCI2 = sort(R_KLCI2)
VaR_KLCI2 = quantile(return_KLCI2,0.05)
VaR_KLCI2

RepVaR_SET2[k] = VaR_SET2
RepVaR_KLCI2[k] = VaR_KLCI2

}##End looping

VaR_SET1
VaR_KLCI1
RepVaR_SET2
RepVaR_KLCI2

R program for results of VaR estimations for SET_B and KLCI_B at 1- and 2-
steps ahead in Tables 7.2.

x<-read.csv("C:/Residual of SET and KLCI.csv",header=TRUE)
Res_SET = x[,1]
Res_KLCI = x[,2]

y<-read.csv("C:/3STOCK_LOGP.csv",header=TRUE)

SET1 = y[,4]
KLCI1= y[,5]
LSET = y[,2]
LKLCI = y[,3]

VaR_ReturnSET1 = c()
VaR_ReturnKLCI1 = c()
VaR_ReturnSET2 = c()
VaR_ReturnKLCI2 = c()

no = 10000
for (k in 1:632) {

  fit_Res_SET = garchFit(Res_SET~arma(0,0)+garch(1,1),
                         data = Res_SET[k:499+k], cond.dist="std")
  summary(fit_Res_SET)
  para_Res_SET<-coef(garchFit(Res_SET~arma(0,0)+garch(1,1),
                             data = Res_SET[k:499+k], cond.dist="std"))
  para_Res_SET
  mu_Res_SET<-para_Res_SET[1]; mu_Res_SET
  alpa1_Res_SET<-para_Res_SET[3]

  fit_Res_KLCI= garchFit(Res_KLCI~arma(0,0)+garch(1,1),
                         data = Res_KLCI[k:499+k], cond.dist="ged")
  summary(fit_Res_KLCI)
  para_Res_KLCI<-coef(garchFit(Res_KLCI~arma(0,0)+garch(1,1),
                             data = Res_KLCI[k:499+k], cond.dist="ged"))
  para_Res_KLCI
  mu_Res_KLCI<-para_Res_KLCI[1]; mu_Res_KLCI
  alpa1_Res_KLCI<-para_Res_KLCI[3]

}  #####################################################################

# 1 step forecast #
# 1 step forecast#
set.seed(10)
hes1_Res_SET = c()
hes1_Res_KLCI= c()
hes2_Res_SET = c()
hes2_Res_KLCI= c()
h1_Res_SET = c()
h1_Res_KLCI = c()
h2_Res_SET = c()
h2_Res_KLCI = c()

ESET1_ECM = c()
EKLCI1_ECM = c()
EKLCI2_ECM = c()
ESET2_ECM = c()

h1_SET = volatility(fit_SET, type="h")
h1_KLCI = volatility(fit_KLCI, type="h")

#################
# Start Looping #
#################

eta1_Res_SET = rsstd (no)
eta12_Res_SET = rsstd (no)
eta1_Res_KLCI = rsstd (no)
eta12_Res_KLCI = rsstd (no)


for (i in 1:no) {
  ESET1_ECM[i] = eta12_SET[i]*sqrt(hes1_Res_SET)
  EKLCI1_ECM[i] = eta12_KLCI[i]*sqrt(hes1_Res_KLCI)
}

ReSET1 = c()
ReKLCI1 = c()

ReSET1[i] = (-0.025*SET1[499+k-1])-(0.007*SET1[499+k-2])
    -(0.020*SET1[499+k-3])+(0.032*SET1[499+k-4])
    -(0.056*SET1[499+k-5])+(0.026*SET1[499+k-6])
    +(0.049*SET1[499+k-7])+(0.019*KLCI1[499+k-1])
    -(0.011*KLCI1[499+k-2])+(0.028*KLCI1[499+k-3])
\begin{align*}
-0.001 \cdot KLCI1[499+k-4] &+ 0.031 \cdot KLCI1[499+k-5] \\
-0.029 \cdot KLCI1[499+k-6] &+ 0.061 \cdot KLCI1[499+k-7] \\
-0.037 \cdot (LSET[499+k-1] - 0.7532 \cdot LKLCI[499+k-1]) &+ ESET1\_ECM[499+k] \\
\text{ReKLCI1}[i] &= (-0.030 \cdot KLCI1[499+k-1] &+ 0.023 \cdot KLCI1[499+k-2] \\
&+ 0.008 \cdot KLCI1[499+k-3] &+ 0.079 \cdot KLCI1[499+k-4] \\
&- 0.183 \cdot KLCI1[499+k-5] &+ 0.060 \cdot KLCI1[499+k-6] \\
&+ 0.052 \cdot KLCI1[499+k-7] &+ 0.029 \cdot SET1[499+k-1] \\
&+ 0.007 \cdot SET1[499+k-2] &- 0.044 \cdot SET1[499+k-3] \\
&+ 0.022 \cdot SET1[499+k-4] &+ 0.021 \cdot SET1[499+k-5] \\
&- 0.017 \cdot SET1[499+k-6] &+ 0.006 \cdot SET1[499+k-7] \\
&+ 0.028 \cdot (LKLCI[499+k-1] - 1.3276 \cdot LSET[499+k-1]) &+ EKLCI1\_ECM[499+k] \\
\text{ReturnSET1} &= \text{sort}(\text{ReSET1}) \\
\text{VaR}\_\text{ReSET} &= \text{quantile}(\text{ReturnSET1}, 0.05) ; \text{VaR}\_\text{ReSET} \\
\text{ReturnKLCI1} &= \text{sort}(\text{ReKLCI1}) \\
\text{VaR}\_\text{ReKLCI} &= \text{quantile}(\text{ReturnKLCI1}, 0.05) ; \text{VaR}\_\text{ReKLCI} \\
\text{VaR}\_\text{ReturnSET1}[k] &= \text{VaR}\_\text{ReSET} \\
\text{VaR}\_\text{ReturnKLCI1}[k] &= \text{VaR}\_\text{ReKLCI} \\
\end{align*}

\textbf{APPENDIX A. PROGRAM FILES}

\vspace{1cm}

\begin{verbatim}
ReturnSET1 = sort(ReSET1)
VaR_ReSET = quantile(ReturnSET1, 0.05) ; VaR_ReSET
ReturnKLCI1 = sort(ReKLCI1)
VaR_ReKLCI = quantile(ReturnKLCI1, 0.05) ; VaR_ReKLCI

VaR_ReturnSET1[k] = VaR_ReSET
VaR_ReturnKLCI1[k] = VaR_ReKLCI

# Calculate 2 step-ahead for return Dis in ECM #

ReSET2 = c()
ReKLCI2 = c()

eta2_Res_SET = rsstd (no)
eta22_Res_SET = rsstd(no)
eta2_Res_KLCI = rsstd (no)
eta22_Res_KLCI = rsstd (no)

for (i in 1:no) {

}
\end{verbatim}
# Calculate Error correcting term in in ECM used for 2 step-ahead #

```
for (i in 1:no) {

ESET2_ECM[i] = eta22_Res_SET * sqrt(hes2_Res_SET[i])
EKLCI2_ECM[i] = eta22_Res_KLCI * sqrt(hes2_Res_KLCI[i])
}

ReSET2[i] = (-0.025*ReSET1[499+k-1])-(0.007*SET[499+k-1])
-(0.020*SET[499+k-2])+(0.032*SET[499+k-3])
-(0.056*SET[499+k-4])+(0.026*SET[499+k-5])
+(0.049*SET[499+k-6])+(0.019*ReKLCI1[499+k-1])
-(0.011*KLCI[499+k-1])+(0.028*KLCI[499+k-2])
-(0.001*KLCI[499+k-3])+(0.031*KLCI[499+k-4])
-(0.029*KLCI[499+k-5])+(0.061*KLCI[499+k-6])
-0.037*(LSET[499+k-7]-0.7532*LKLCI[499+k-7])+ESET1_ECM[499+k]

ReKLCI2[i] = (-0.030*ReKLCI1[499+k-1])+(0.023*KLCI[499+k-1])
+(0.008*KLCI[499+k-2])+(0.079*KLCI[499+k-3])
-(0.183*KLCI[499+k-4])+(0.060*KLCI[499+k-5])
+(0.052*KLCI[499+k-6])+(0.029*ReSET1[499+k-1])
+(0.007*SET[499+k-1])-(0.044*SET[499+k-2])
+(0.022*SET[499+k-3])+(0.021*SET[499+k-4])
-(0.017*SET[499+k-5])+(0.006*SET[499+k-6])
+0.025*(LSET[499+k-7]-1.3276*LKLCI[499+k-7])+EKLCI1_ECM[499+k]

ReturnSET2 = sort(ReSET2)
VaR_ReSET2 = quantile(ReturnSET2,0.05) ; VaR_ReSET2
ReturnKLCI2 = sort(ReKLCI2)
VaR_ReKLCI2 = quantile(ReturnKLCI2,0.05) ; VaR_ReKLCI2

VaR_ReturnSET2[k] = VaR_ReSET2
VaR_ReturnKLCI2[k] = VaR_ReKLCI2
}
```

} # end Looping

VaR_ReturnSET1
VaR_ReturnKLCI1
VaR_ReturnSET2
VaR_ReturnKLCI2
Bibliography


