STOCHASTIC MODELLING OF VOLATILITY, LEVERAGE EFFECTS, LONG-MEMORY AND EXTREMAL DEPENDENCE OF FINANCIAL MARKETS

By

Richard Taiwo Abayomi Samuel (2389435)

Supervisor

Prof. Charles Chimedza

Co-Supervisor

Prof. Caston Sigauke

A Thesis Submitted to the Faculty of Science, University of the Witwatersrand, in Fulfillment of the Requirements of the Degree of Doctor of Philosophy (PhD) in Mathematical Statistics in the School of Statistics and Actuarial Science, Johannesburg, South Africa

December 14, 2024

Acknowledgement

My foremost appreciation goes to the Lord and Saviour Jesus Christ who has helped me immensely to complete this thesis. He is the unparalleled Professor in my career and to Him be all the praise. I want to thank and give my warmest gratitude to my supervisors, Prof. Charles Chimedza and Prof. Caston Sigauke for their incessant support and unreserved generosity, kindness, patience, guidance, high professional skills, and financial support. My heartfelt gratitude also goes to Dr. Rosina Mukhodobwane and her husband for their relentless support during the research period. I also acknowledge the School of Statistics and Actuarial Science, University of the Witwatersrand, and the Department of Mathematical and Computational Sciences, University of Venda, for their support with the payment of the article processing charges (APC). Finally, I am extremely grateful to my family members, whose love and support were crucial to the successful completion of this project. Thank you all.

Dedication

I would like to dedicate this research report to the Almighty God.

Abstract

This research focused on using a three-stage simulation approach verified empirically for improved volatility modelling, relevant for robust risk management. The first two stages of the study are focused on developing simulation procedures in volatility modelling using two autoregressive models that involve the family of Generalised Autoregressive Conditional Heteroscedasticity (fGARCH) and the Generalised Autoregressive Score (GAS) models. The empirical evaluations of the simulation experiments in the first two stages were carried out using the S&P SA Bond Index data. The third stage of the study is used to estimate six essential features (or stylised facts) of financial return volatility that are relevant for valuable insights into risk assessment and investment decision-making. These features include pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long-memory behaviour of volatility decomposition into long-term and short-term components. Specifically, in the first stage of the thesis, the study proposed a simulation framework using the Monte Carlo simulation (MCS) resources of the fGARCH models to determine a suitable conditional distribution for the error term to model the persistence of volatility. In the process of developing this framework, this study also proposed a new true-parameter-recovery measure which is used as a proxy of the coverage probability to accurately calculate the performance of the simulation experiment.

In the second stage of the thesis, the study built on the developed framework (in the first stage) to propose a simulation structure through the GAS model for selecting an optimal error distribution for volatility modelling. The investigations at this stage proceeded by using both the fat-tails of distributions and $\sqrt{N}$ consistency simulation experiments to show that the GAS model with a lower unconditional shape parameter (ν^*) value of 4.1 can be used to generate an appropriate simulated dataset that properly reflects the behaviour of financial returns data relevant for modelling volatility. This dynamic structure is intended to help interested users on MCS experiments utilising the GAS model for reliable volatility persistence calculations in finance and other areas. The simulation frameworks and procedures in the first two stages can be a useful guide to scientific practitioners and upcoming researchers on the relevant simulation steps to determine a suitable error distribution for volatility modelling.

In the third stage, this research comparably applied three dynamic observation-driven models consisting of the fGARCH, GAS and Beta-Skew-t-EGARCH models to estimate the

stated six essential features (or characteristics) of volatility, relevant for robust investment decisions and risk evaluation in the S&P Indian stock market. To begin with, the study comparatively used the robust fGARCH and GAS models to estimate the magnitude and dynamics of the persistence in conditional volatility using the returns from the Indian market index.

Next, the study comparatively used the one- and two-component Beta-Skew-t-EGARCH models to estimate other features of the return volatility that include leverage effect or asymmetry, skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components. Specifically, we used both the one- and two-component models to estimate leverage effects, fat-tails, and skewness in the returns. The study further used a parametric model through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle estimator, and the exact local Whittle estimator to estimate and determine the presence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns. The results of the estimations indicate that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate.

Furthermore, we used the two-component Beta-Skew-t-EGARCH model to investigate the long-memory decomposition of volatility into long-term and short-term components. Through this two-component model, the study found the existence of both long-run and short-run components of volatility in the persistence process, but the response to the effect of shocks in the short-run is higher than in the long-run volatility. This implies that higher volatility in the process is mostly due to the short-run volatility increase. Hence, through the applications of these models using the S&P Indian index, the study shows that the Indian market returns are characterised by the six volatility features. The empirical and simulation outcomes of the experiments in the third stage are used to offer both long-term and short-term suggestions to rational investors, government, and market managers for relevant assessment of the market investment risk.

Keywords: Bias, Consistency, Efficiency, Fat-tailed distributions, GARCH-type models, Heteroscedasticity, Long-memory, Score-driven models, Simulation design, Time-varying parameter estimation, Volatility persistence.

Declaration

I declare that this research report is my own work. It is being submitted for the Degree of Doctor of Philosophy to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

_____ day of ____- 20 _____ at _____

Signature of Candidate Date

Contents

1 Introduction **5**

1.1 Overview . . . 5
1.2 Background of the Study . . . 5
1.3 Motivation . . . 8
1.4 Aim and Objectives . . . 9
1.4.1 Aim . . . 9
1.4.2 Objectives . . . 9
1.5 Limitations and Assumptions . . . 10

2 Literature Review **11**

2.1 Introduction . . . 11
2.2 Applications of the GARCH Model and its Extensions . . . 11
2.3 Applications of the GAS Model and Other Score-Driven Extensions . . . 14
2.4 Applications of the Beta-Skew-t-EGARCH Models . . . 17
2.5 Applications of Fractional Integrated Models . . . 18
2.6 Summary of the Studies and Conclusion . . . 20

3 Methodology **23**

3.1 Introduction . . . 23
3.2 The GARCH Model . . . 23
3.3 The fGARCH Model . . . 25
3.4 The GAS Model . . . 26
3.4.1 Specification of the GAS Model . . . 26
3.5 Persistence and Mean Reversion in Volatility . . . 28
3.6 Test for Stationarity . . . 29
3.6.1 The Augmented Dickey Fuller Test . . . 29
3.6.2 Phillips-Perron (PP) Test . . . 30
3.6.3 The Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) Test . . . 31
3.7 Test for Serial Correlation . . . 32
3.8 Test for Heteroscedasticity or ARCH effects . . . 33

3.8.1 ARCH LM test . . . 33
3.9 Model Selection . . . 34
3.9.1 Akaike Information Criterion . . . 34
3.9.2 Bayesian Information Criterion . . . 36
3.9.3 Hannan-Quinn Information Criterion . . . 36
3.9.4 Shibata Information Criterion . . . 37
3.9.5 Model Checking through Diagnostics . . . 37
3.10 Long Memory and Short Memory Processes . . . 37
3.11 The ARIMA and ARFIMA Processes . . . 40
3.11.1 ARFIMA Model . . . 41
3.12 The FIGARCH Model . . . 42
3.13 The Log-periodogram Regression . . . 44
3.14 Geweke and Porter-Hudak Estimator . . . 44
3.15 Whittle Estimation . . . 46
3.15.1 The Whittle Likelihood . . . 46
3.15.2 The Local Whittle Estimator . . . 47
3.15.3 The Exact Local Whittle Estimator . . . 48
3.16 The Beta-Skew-t-EGARCH Model . . . 49
3.16.1 The One-Component Beta-Skew-t-EGARCH Model . . . 49
3.16.2 The Two-Component Beta-Skew-t-EGARCH Model . . . 50
3.16.3 The Skew Student's t Distribution . . . 51
3.17 The True Parameter Recovery Measure . . . 51

4 Simulation Framework to Determine Suitable Innovations for Volatility Persistence Estimation: The GARCH Approach 52
4.1 Introduction . . . 53
4.2 Materials and Methods . . . 55
4.2.1 The GARCH Model . . . 55
4.2.2 The fGARCH Model . . . 56
4.2.3 The True Parameter Recovery Measure . . . 57
4.2.4 Simulation Design . . . 58
4.2.4.1 Aim of the Simulation Study . . . 58
4.2.4.2 State the Research Questions . . . 59
4.2.4.3 Method of Implementation . . . 59
4.2.4.4 Discussion and Summary . . . 66
4.3 Results: Simulation and Empirical . . . 67
4.3.1 Practical Illustrations of the Simulation Design: Application to Bond Return Data . . . 67
4.3.1.1 The Background . . . 67

4.3.1.2 Aim of the Simulation Study . . . 68
4.3.1.3 Research Questions . . . 68
4.3.1.4 Method of Implementation . . . 69
4.3.2 Empirical Verification . . . 73
4.3.2.1 Exploratory Data Analysis . . . 73
4.3.2.2 Tests for Stationarity . . . 75
4.3.2.3 Tests for Serial Correlation and Heteroscedasticity . . . 75
4.3.2.4 Residual Diagnostic Test . . . 77
4.3.2.5 Selection of the Most Suitable Error Distribution . . . 77
4.4 The Diagnostic Plots of ARMA-fGARCH Models . . . 81
4.5 Discussion and Summarised Conclusions . . . 81
4.6 Conclusions . . . 82
4.6.1 Limitations in the Study . . . 84

5 Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model 85
5.1 Introduction . . . 86
5.2 Methods . . . 88
5.2.1 The GAS Model . . . 88
5.2.1.1 Specification of the GAS Model . . . 88
5.2.2 The True Parameter Recovery Measure . . . 90
5.2.3 The Design of the Simulation . . . 90
5.2.3.1 Aim of the Simulation Study . . . 91
5.2.3.2 State the Research Questions . . . 91
5.2.3.3 Method of Implementation . . . 92
5.2.4 Discussion and Summary . . . 95
5.3 Results: Simulation and Empirical . . . 95
5.3.1 Practical Illustrations of the Simulation Structure with Applications to Financial Returns . . . 95
5.3.2 The Background . . . 95
5.3.3 Aim of the Simulation Study . . . 99
5.3.4 Research Questions . . . 99
5.3.5 Method of Implementation . . . 101
5.3.5.1 Method of Implementation with the Student's t Error . . . 101
5.3.5.2 Method of Implementation with AST1 Error . . . 104
5.3.6 Empirical Study . . . 107
5.3.6.1 Exploratory Data Analysis . . . 108
5.3.6.2 Tests for Autocorrelation and Heteroscedasticity . . . 108
5.3.6.3 Selection of the Most Appropriate Error Assumption . . . 112

5.4 Discussion and Summarised Conclusion . . . 112
5.5 Concluding Remarks . . . 113
5.5.1 Limitations in the Study . . . 115

6 A Comparative Modelling of Essential Characteristics of Volatility 116
6.1 Introduction . . . 117
6.2 Methods . . . 119
6.2.1 The GARCH Model . . . 119
6.2.2 Persistence and Mean Reversion in Volatility . . . 120
6.2.3 The fGARCH Model . . . 121
6.2.4 The GAS Model . . . 122
6.2.5 Model Specification . . . 122
6.2.6 The Beta-Skew-t-EGARCH Model . . . 124
6.2.6.1 The One-Component Beta-Skew-t-EGARCH model . . . 124
6.2.6.2 The Two-Component Beta-Skew-t-EGARCH Model . . . 125
6.2.6.3 The Skew Student's t Distribution . . . 126
6.2.7 The True Parameter Recovery Measure . . . 126
6.3 Experimental Results: Empirical and Simulation . . . 127
6.3.1 Application of the fGARCH Model . . . 127
6.3.1.1 Exploratory Data Analysis . . . 127
6.3.1.2 Descriptive Statistics . . . 128
6.3.1.3 Tests for Stationarity . . . 128
6.3.1.4 Tests for Autocorrelation and ARCH Effects . . . 131
6.3.1.5 Residual Diagnostic Test . . . 131
6.3.1.6 Selection of an Optimal Error Distribution . . . 132
6.4 The Diagnostic Plots of ARMA-fGARCH Models . . . 132
6.4.1 Persistence in Volatility and Mean Reversion . . . 134
6.4.2 Application of the GAS Model . . . 138
6.4.3 Application of the Beta-Skew-t-EGARCH Model . . . 139
6.4.4 Simulation Study . . . 143
6.4.5 Model Comparison . . . 145
6.5 Discussion . . . 145
6.6 Conclusion . . . 147

7 Discussions Based on the Contributions of the Research Papers 149
7.1 Simulation Framework to Determine Suitable Innovations for Volatility Persistence Estimation: The GARCH Approach . . . 150
7.2 Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model . . . 152

7.3 A Comparative Modelling of Essential Characteristics of Volatility 155

8 Summary, Conclusions and Recommendations 158
8.1 Summary and Conclusion 158
8.2 Recommendations for Future Directions of the Work 160

A Outcomes of Different Patterns of Seed Values for Sets S_1 and S_2 188

B Further Visual Illustrations of S_1 and S_2 TPR Outcomes 190

C The Code for DGP through the Ugarchsim Function 191

D The $\sqrt{N}$ Consistency Outcomes of Datasets with Different Unconditional Shape Parameters 193

E R Code 195
E.1 The R Code for the Method of Implementation of the MCS Experiment in Paper 1 195
E.2 R Code for Paper 2 when the true model is GAS-Student's t 203
E.3 R Code for Paper 2 when the true model is GAS-AST1 207

List of Figures

4.1 Simulation design flowchart to determine suitable assumed innovations. 58
4.2 Panels (**A**,**B**) show the efficiency and $\sqrt{N}$ consistency in S_1 and S_2 for each seed pattern, while Panels (**C**,**D**) reveal the impacts of sample size on RMSE and SE under the assumed errors in S_1 and S_2. 62
4.3 TPR outcomes in Panels **A**(**i**) and **B**(**i**) for S_1 and S_2, respectively. The outcomes are clearly spread out in Panels **A**(**ii**) and **B**(ii) for S_1 and S_2. The dotted lines are the 95% (i.e., 0.95) nominal recovery levels. 62
4.4 The TPR outcomes against sample size are shown in Panels (**A**,**B**), while Panels (**C**,**D**) show the TPR outcomes against the MCS estimates. 63
4.5 Simulated returns (in Panel **A**) and simulated volatility (Panel **B**) of the first three replicated series. 70
4.6 The impact of sample size on RMSE, SE and bias for the fGARCH(1,1)-Student's t MCS modelling. The RMSE and SE are considerably $\sqrt{N}$ consistent, but bias is independent of N. 73
4.7 Panels (**A**,**B**) display the TPR outcomes, where the clustered outcomes in Panel (**A**) are clearly spread out in Panel (**B**). The dotted line is the 95% (i.e., 0.95) nominal recovery level. 74
4.8 EDA of price (panels **a**,**b**,**e**,**f**) and returns (panels **c**,**d**,**g**,**h**) for SA Bond Index. . 75
4.9 ARCH Portmanteau Q and Lagrange Multiplier tests. 77
4.10 The diagnostic plot of the densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plot for the ARMA-fGARCH models fitted with the skew-student's t assumed innovation. 81
4.11 The diagnostic plot of the densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plot for the ARMA-fGARCH models fitted with the GHYP assumed innovation. 82

5.1 Flowchart on simulation design to select an optimal error distribution. 91
5.2 97

5.3 Panels (A) and (B) display the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ in terms of SE and RMSE for N = {30,000; 31,000; 32,000} datasets with ν^* = 4.1, while Panels (C) and (D) display those for ν^* = 8.2, and Panels (E) and (F) for ν^* = 10.3. 98
5.4 Panels (A) and (B) display the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ in terms of SE and RMSE for N = {17,000; 18,000; 19,000} datasets with ν^* = 4.1, while Panels (C) and (D) display those for ν^* = 8.2, and Panels (E) and (F) for ν^* = 10.3. 100
5.5 103
5.6 The true parameter recovery (TPR) outlook. The dotted line denotes the 0.95 (i.e., 95%) nominal recovery level. 104
5.7 105
5.8 The true parameter recovery (TPR) outlook. The dotted line denotes the 0.95 (i.e., 95%) nominal recovery level. 107
5.9 Histograms of price and returns 108
5.10 Price and return distributions 109
5.11 Price and return densities in Panels A and B respectively, while Panels C and D display the price and return distributions with their densities. 109
5.12 Residual plot and outcomes of serial correlation test. 110

6.1 Panels (A) and (B) respectively show the plots of the price and return series, while Panels (C) and (D) display the quantile-quantile (Q-Q) and density plots, respectively, for the S&P Indian Index. Panels (E) and (F) display the ACF and PACF plots of the returns and squared returns, respectively. 129
6.2 The diagnostic plot of the densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plot for the ARMA-fGARCH models fitted with the NIG assumed innovation. 134
6.3 Conditional volatility of the fGARCH(1,1) model fitted with the NIG assumed error. 135
6.4 The decay of volatility persistence through the fGARCH(1,1)-NIG model. . . . 137
6.5 The decay of volatility persistence through the GAS-AST1 model. This decay curve follows the same trajectory as the fGARCH(1,1)-NIG decay curve in Figure 6.4 because the persistence estimates from the two models are the same (i.e., 0.9749). 139
6.6 Panels (A) and (B) present the fitted conditional standard deviations of the one- and two-component models, respectively. 141
6.7 Panels (A) and (B) present the long-term and short-term decay curves, respectively. 143

6.8 Panels (A) and (B) display the plots for one-component simulated returns and two-component simulated returns, respectively. 144

B.1 The TPR outcomes of S_1 and S_2 in Panels (**A**,**B**), respectively, where the dotted lines are the 95% (i.e., 0.95) nominal recovery levels. 190

List of Tables

4.1 Outcomes of different simulation replicates. 65
4.2 True model fGARCH(1,1)-Student's t with true parameters α = 0.0748, β = 0.9243 and $\alpha + \beta$ = 0.9991. 71
4.3 Tests for Stationarity of the Price Index and Return Series. 76
4.4 ARMA(1,1)-fGARCH(1,1) models' empirical outcomes on SA Bond return data. 78
4.5 ARMA(1,1)-apARCH(1,1) models' empirical outcomes on SA Bond data. . . . 80

5.1 Simulation outcomes. 102
5.2 Simulation outcomes. 106
5.3 ARMA(1,1) and ARMA(2,2) models' outcomes on Bond returns. 111
5.4 Engle's ARCH test results. 111
5.5 Empirical outcomes of GAS modelling on SA Bond returns data. 112

6.1 130
6.2 Tests for Stationarity of the Price Index and Return Series. 130
6.3 133
6.4 Long memory estimation through the fractional integration parameter d. 136
6.5 Empirical outcomes of the GAS modelling on the real return data. 138
6.6 Empirical and Simulation Outcomes of the Beta-Skew-t-EGARCH Model. . . 142

A.1 Outcomes of different patterns of seed values for sets $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$, with true parameter $\boldsymbol{\alpha} + \boldsymbol{\beta}$ = 0.9990. 189

D.1 The $\sqrt{N}$ consistency outcomes of datasets with different unconditional shape parameters ν^* = {4.1, 8.2, 10.3}. 194

Thesis Structure and Research Outputs

Chapter 1 presents the introduction that covers an overview of the study, with the aim and objectives of carrying out the research.

Chapter 2 presents the literature review, and it covers some of the body of works that have been conducted with the applications of time-varying models used in the thesis.

Chapter 3 presents the study's methodology that describes the models and techniques applied in the estimations and simulation framework developments.

Chapters 4, 5, and 6 collectively present the detailed studies carried out in Papers 1, 2, and 3 on simulation frameworks and the estimations of volatility features through the family GARCH, GAS, and Beta-Skew-t-EGARCH models. Other methods used in the estimations include the Autoregressive Fractionally Integrated Moving Average-Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (ARFIMA-FIGARCH) models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator. Papers 1 and 2 have been published, while Paper 3 is under review.

Chapter 7 is the amalgam of the three papers; hence, it is a blend of Chapters 4, 5 and 6. The modelling approaches used in the three papers are based on time-varying observation-driven models. The objectives of the papers are focused primarily on robust volatility modelling for effective risk management through building appropriate simulation frameworks and estimating essential features of volatility.

Chapter 8 presents the summarized conclusion of the thesis with emphasis on the novelty and contributions of each peer-reviewed paper. In addition, the chapter further discusses recommendations for the intended future direction of the work.

A List of Peer-Reviewed Articles from the Research

Published Journals

Paper 1: Richard T. A. Samuel, Charles Chimedza and Caston Sigauke (2023), Simulation Framework to Determine Suitable Innovations for Volatility Persistence Estimation: The GARCH Approach. *Journal of Risk and Financial Management*, 16(9): 392.

Paper 2: Richard T. A. Samuel, Charles Chimedza and Caston Sigauke (2024), Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model. *Statistics, Optimization & Information Computing*, 12: 1123-1148.

Manuscript under Review

Paper 3: Richard T. A. Samuel, Charles Chimedza and Caston Sigauke, A Comparative Modelling of Essential Characteristics of Volatility. *Journal of Statistics Applications & Probability*.

List of Acronyms

AVGARCH	Absolute Value GARCH
AP-GoF	Adjusted Pearson Goodness-of-Fit
AIC	Akaike information criterion
ALD	Asymmetric Laplace distribution
apARCH	Asymmetric Power ARCH
AST1	Asymmetric Student's t with one tail decay parameter
AST	Asymmetric Student's t with two tail decay parameters
ACD	Autoregressive Conditional Duration
ARCH	Autoregressive Conditional Heteroscedasticity
ACI	Autoregressive Conditional Intensity
ACM	Autoregressive Conditional Multinomial
ARIMA	Autoregressive Integrated Moving Average
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Bayesian information criterion
CGARCH	Component GARCH
DGP	Data generation process
EDA	Exploratory Data Analysis
EGARCH	Exponential GARCH
ELW	Exact Local Whittle

fGARCH	family GARCH
FIGARCH	Fractionally Integrated GARCH
GARCH	Generalised Autoregressive Conditional Heteroscedasticity
GAS	Generalised Autoregressive Score
GED	Generalised Error Distribution
GHYP	Generalised Hyperbolic
GHST	Generalised Hyperbolic Skew-Student's t
GJR GARCH	Glosten-Jagannathan-Runkle GARCH
GPH	Geweke and Porter-Hudak
HQIC	Hannan-Quinn information criterion
i.i.d.	Independent and identically distributed
JSU	Johnson's reparametrised SU
LM	Lagrange Multiplier
LW	Local Whittle
llk	log-likelihood
MLE	Maximum likelihood estimation
MSE	Mean square error
MCS	Monte Carlo simulation
MCGARCH	Multiplicative Component GARCH
NGARCH	Nonlinear ARCH
NAGARCH	Nonlinear Asymmetric GARCH
NIG	Normal Inverse Gaussian
PACF	Partial autocorrelation function
$\hat{P}$	Persistence
PQ	Portmanteau-Q
p-value	Probability value
QQ	Quantile–Quantile
QMLE	Quasi-maximum likelihood estimation
RE	Relative efficiency

RMSE	Root mean square error
SD	Score Driven
SARS	Severe Acute Respiratory Syndrome
SIC	Shibata information criterion
sGARCH	simple GARCH
SA	South Africa
SE	Standard error
S&P	Standard & Poor
TGARCH	Threshold GARCH
TPR	True Parameter Recovery
ugarchpath	Univariate GARCH Path Simulation
ugarchsim	Univariate GARCH Simulation
UniGASSim	Univariate GAS simulation
UniGASSpec	Univariate GAS specification
WLB	Weighted Ljung–Box

Chapter 1

Introduction

1.1 Overview

Volatility is a measure of the dispersion around the average return of an asset. An accurate modelling of volatility requires adequate knowledge of the underlying error distribution. However, the selection of the most suitable assumed error distribution under an appropriate volatility model, when the underlying distribution is unknown, is nontrivial. Hence, this study adopts simulation-based modelling approaches since they offer the opportunity to examine complex and interconnected processes (Leitner and Wall, 2015). The outcomes of the simulation modelling are further examined through real-data empirical modelling approach.

1.2 Background of the Study

To begin with, this proposed research study focused on developing algorithms for simulation frameworks through time-varying models for selecting an optimal conditional error distribution, relevant to modelling the persistence of volatility. Next, three robust autoregressive models will be explored to model essential features (or characteristics) of volatility for evaluating the impact of risk on financial market returns. These characteristics referred to as stylised facts include pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long-memory behaviour of volatility. The autoregressive models that are employed for modelling in this research study include the GARCH (Bollerslev, 1986), the family Generalised Autoregressive Conditional Heteroscedasticity (fGARCH) model (Hentschel, 1995), the Generalised Autoregressive Score (GAS) model (Ardia et al., 2019; Creal et al., 2013; Harvey, 2013), and the one and two components Beta-Skew-t-EGARCH models of Sucarrat (2013) and Harvey and Sucarrat (2014).

This study will generate synthetic (or simulated) datasets via simulation techniques using each of these models in implementing the proposed frameworks and estimating the stated

volatility features. The outcomes from the simulation processes will be verified through application to real financial time series data, which makes the design easily acceptable to users for modelling the persistence of volatility and for estimating essential characteristics of volatility that are pertinent to the effective management of investment risks. This study further used a parametric approach through the ARFIMA-FIGARCH models (Baillie et al., 1996) and three semiparametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (1983) and Robinson (1995a), the local Whittle estimator (Künsch, 1987; Robinson, 1995a), and the exact local Whittle estimator (Shimotsu and Phillips, 2005) to determine the presence of long memory in the return volatility.

In Chapters 4 and 5 of this thesis, a comprehensive simulation design is developed that is summarily required for a robust simulation practice in finance to determine appropriate assumed error distributions, relevant for estimating the persistence of volatility using the fGARCH and GAS models. The underlying methodology can be applied in other fields. In both chapters, the Monte Carlo simulation approaches are used to obtain the most suitable assumed errors, through which the volatility persistence is empirically estimated. The two autoregressive fGARCH and GAS models are time-varying observation-driven models. Moreover, each of them is an omnibus model that subsumes many other auto-regressive models as subclasses (Creal et al., 2011, 2013, 2014; Ghalanos, 2018; Hentschel, 1995). Observation-driven modelling occurs when time variation of parameters is presented by making parameters a function of concurrent variables, lagged dependent variables, and lagged exogenous variables (see Buccheri et al., 2021; Creal et al., 2013).

Although both fGARCH and GAS models are used to model the time variation in the conditional variance (Ardia et al., 2019; Bollerslev, 1986; Creal et al., 2013; Engle, 1982; Ghalanos, 2018), the dynamics of the time-varying process are different for each model. Whilst the fGARCH applies the dynamics of the residual to drive the conditional variance, the GAS model explores the conditional score for the process. The two proposed observation-driven fGARCH and GAS models have the benefit of giving insights into what drives the dynamics of persistence in a volatility process. The outcomes from the two models are compared to ascertain which one outperforms the other and to determine the levels of the persistence of conditional volatility when applied to real-life financial data. Simulation and data modelling approaches where the persistence of volatility is estimated using the fGARCH and GAS models comparatively under the specified error distributions are required to better comprehend the phenomenon of volatility persistence.

In Chapter 6, the study comparably applied the collaborative modelling capabilities of the fGARCH, GAS, and Beta-Skew-t-EGARCH models to estimate six key attributes of volatility for proper mitigation of the risks associated with investments. The Beta-Skew-t-EGARCH is an unrestricted version of the GAS model that was developed to model relevant stylised facts of volatility (Harvey and Sucarrat, 2014; Sucarrat, 2013), and it functions in two versions that

include the "one-component and two-component models". Both one- and two-component models can effectively model volatility features of financial returns. However, the two-component model is further capable of modelling the decomposition of volatility into the long-run and short-run components. The outcomes of the comparative estimations from the three models of the fGARCH, GAS, and Beta-Skew-t-EGARCH are used to present both short-term and long-term informed investment advice to rational investors for effective portfolio management. Furthermore, the outcomes of the volatility persistence and long-memory decomposition in the market are meant to assist investors in gaining insights into the level of risk exposures.

Stochastic modelling of volatility plays a central role in financial risk management as it is crucial in all other areas where time-varying parameter models are applicable (Ardia et al., 2019). Accurate modelling of volatility are particularly relevant in computations of parametric value-at-risk (VaR), expected shortfall (ES), derivatives pricing, trading strategies, portfolio selection, risk analysis, asset allocation and investment decisions (Leite et al., 2016; Maciel and Ballini, 2017; Wang et al., 2019), hence its relevance cannot be over-emphasized. The measurement of assets volatility is used as a proxy for risk measurement by financial institutions, and portfolio optimization modelling also requires significant measurement of volatility (Samiev, 2012). In addition, accurate measure of volatility can be used to quantify total assets' risk in finance; it can also be used for estimating market risk using VaR models, and it is a vital part of the formula for Black-Scholes (Black and Scholes, 1972) differential equation used for the derivation of traded options price (Costa, 2017). Correct modelling and forecast of volatility additionally have a significant economic impact on portfolio allocation, and volatility dynamics modelling tools have also been applied in areas like medicine, natural sciences, social sciences and more (Andersen et al., 2006). The robustness of the application of the most suitable error distribution/s will be of particular interest to model the persistence of volatility in finance and other areas.

Volatility persistence is a process where the return of today affects the future's forecast variance (Engle and Patton, 2001). The level of persistence has crucial economic implications that stem from the impact it has on the predictability of distant future value (Arago and Fernandez-Izquierdo, 2003). Risk managers, policymakers, portfolio managers, financial market participants, and other financial activities can benefit from a clear understanding of the effect of shocks on future volatility, especially with information on whether the shocks are short-lived or (highly) persistent. It is known that shocks with a permanent influence on the variance will have a greater effect on price than those with a temporary influence (Arago and Fernandez-Izquierdo, 2003). Other economic implications of the degree of persistence of a shock include its impact on dynamic hedging policies, the valuation of options and the price of securities (Arago and Fernandez-Izquierdo, 2003; Porterba and Summers, 1986).

This study extends the literature with novel contributions and brings additional insights using simulation and real-life data modelling to estimate the persistence of conditional volatility by fitting two observation-driven models, namely, the GAS and fGARCH models, to simulated

and real-life datasets under the assumptions of seven and ten error distributions, respectively. The relevance of using as many as the stated numbers of assumed error distributions is to determine the one/s that can suitably describe the financial market returns when the underlying error distribution is not known. Moreover, risk management systems are highly dependent on the underlying error distribution; hence, the identification of an assumed conditional distribution that adequately captures every aspect of the given financial data may be of great benefit to investors and risk managers (Chinhamu et al., 2017). Furthermore, since financial data are skewed and fat-tailed (Eom et al., 2019; Li, 2008), the non-Normal distributions and their skew variants in the selected error assumptions as presented in Chapters 4 and 5 are meant to adequately account for fat-tails, skewness and potential excess kurtosis in the return distribution.

To summarise, in this research, we explore the functionalities of the fGARCH and GAS models by employing simulation frameworks to effectively enhance the modelling of the persistence of volatility and illustrated using real-life financial time series data. Furthermore, the research applies the extension of the score-driven model that involves the one- and two-component Beta-Skew-t-EGARCH models to estimate key features of volatility that crucially necessitate effective management of risk. In addition, the study used the parametric ARFIMA-FIGARCH models, and three semi-parametric estimators that include the Geweke and Porter-Hudak (GPH), local Whittle (LW) and exact local Whittle (ELW) to estimate the long memory parameter in the return volatility.

1.3 Motivation

Some of the works that have been used in the literature to generally describe the simulation approach to solving volatility and other time-series related problems include Chalmers and Adkins (2020); Feng and Shi (2017); Morris et al. (2019), and Sigal and Chalmers (2016), among others, but none has directly designed a simulation structure for volatility modelling involving autoregressive models. Specifically, whilst the empirical approach of volatility modelling involving autoregressive models is pervasive in the literature, an organised simulation framework to select an optimal conditional distribution for the error term to model volatility has been rarely studied. Hence, this study is motivated by the inclination to derive a flexible but robust simulation framework for improved volatility persistence estimation using the autoregressive fGARCH and GAS models.

The presentation of the proposed framework will be carried out in two phases. The first phase is to use a step-by-step simulation procedure to obtain an optimal conditional distribution for the error term to estimate the persistence of volatility. In this case, the error distribution with the best efficiency and precision is taken as the optimal or the most suitable to describe and predict the returns for modelling the persistence of the volatility. This simulation outcome is then verified empirically through selected information criteria and/or log-likelihood

approaches. For the second phase, the volatility persistence under the identified optimal error distribution is empirically tested through stages of hypothesis testing.

This study is further motivated by the need to calculate the coverage probability of the simulation experiments. Hence, the study proposed a flexible but robust proxy measure to accurately estimate the coverage or performance of the simulation experiments involving these models.

1.4 Aim and Objectives

1.4.1 Aim

Over the years, many studies involving volatility modelling through autoregressive models have mainly been conducted via empirical approaches. A simulation-based approach is not often included because many new researchers do not have an adequate understanding of the steps involved. Specifically, little has been explored when it comes to the recent score-driven models developed by Creal et al. (2013) and Harvey (2013). Hence, this study seeks to propose an effective simulation frameworks for selecting an optimal (or the most suitable) conditional error distribution for modelling the persistence of volatility using the autoregressive fGARCH and GAS models. Following this, the study will further explore the possibilities of using the collaboration of these statistical models with the Beta-Skew-t-EGARCH model to comparably estimate essential characteristics of volatility for an effective investment decision and risk management process.

1.4.2 Objectives

Specifically, the objectives of the study include the following:

1. In paper 1, this study intends to develop a simulation framework for determining a suitable conditional distribution for volatility persistence modelling through the fGARCH model. This is intended to help users/researchers with the necessary simulation procedures relevant for volatility persistence estimation utilising this model. In addition, the study seeks to develop a true parameter recovery measure that can be used as a proxy for reliably calculating the coverage (or performance) of the simulation experiments.

2. In paper 2, the study intends to extend the simulation framework developed in paper 1 by designing a simulation structure using the GAS model. Furthermore, the paper intends to conduct relevant simulation experiments on how to generate a dataset that can adequately reflect the behaviour of financial returns data relevant for volatility modelling through the GAS model. The outcome of this may potentially guide practitioners and upcoming researchers with necessary details on such data generation.

3. In paper 3, the study intends to model six essential features of volatility, relevant for robust market risk assessments by comparably using the collaboration of the family GARCH (fGARCH), GAS and Beta-Skew-t-EGARCH models. The six volatility features comprise pronounced persistence, mean reversion, leverage effect (i.e., volatility asymmetry), conditional fat-tailedness, conditional skewness, and the long memory behaviour of volatility decomposition into long-term and short-term components. Moreover, further evidence of the presence of long memory in the volatility process will be carried out through the parametric ARFIMA-FIGARCH models, and three semi-parametric estimators namely the Geweke and Porter-Hudak, local Whittle, and exact local Whittle.

1.5 Limitations and Assumptions

Among other limitations and assumptions as detailed in Chapters 4 and 5, it is observed that the R rugarch and GAS packages used for implementing the simulation frameworks in this study do not make provisions for calculating the coverage probability. Hence, this study assumed a proxy for the coverage of the simulation experiments by developing the true-parameter-recovery (TPR) measure used for precise computations of the performance of the Monte Carlo simulation (MCS) experiments. It is also noticed that the GAS package does not have a provision to test for serial correlation in the returns, hence we assumed the outcomes of the Weighted Ljung-Box statistic through the R rugarch package (Ghalanos, 2018, 2022) and the auto.arima() function in the R forecast package (Hyndman and Khandakar, 2008) for the test.

Chapter 2

Literature Review

2.1 Introduction

Many statistical empirical applications of the time-varying GARCH/GAS models and other time-varying model extensions have been developed and applied in the literature. This study extends the literature by using both simulation and real life data empirical experiments to obtain an optimal conditional error distribution for modelling the persistence of volatility, and to estimate essential features of volatility through the fGARCH, GAS and Beta-Skew-t-EGARCH models. It is well documented that these models perform better than the classical autoregressive models, like the standard GARCH in modelling volatility because they incorporate asymmetric volatility processes, and are also robust to jumps or outliers that traditionally exist in financial time series data (see Ardia et al., 2019; Haddad et al., 2023; Harvey and Sucarrat, 2014; Takaishi, 2018). It is also believed that asymmetric GARCH models produce a more accurate forecast of equity return volatility (Mamilla et al., 2023; Sen et al., 2021). This chapter is used to present a few of the earlier studies that focussed on estimating volatility and its features using time-varying observation-driven models of the GARCH and score-driven models. The chapter also includes literature on studies involving the applications of both time-domain and frequency-domain estimations of long memory through the ARFIMA-FIGARCH models, and the Geweke and Porter-Hudak, local Whittle, and exact local Whittle estimators.

2.2 Applications of the GARCH Model and its Extensions

The volatility of the stock market's returns is usually described through the variances, and it is a perception of the risk in the market. Volatility is described as a key feature of equity returns where it is viewed as a measure of financial assets' risks. It is also used for describing the variability of market price. This description can also be seen as the variability of financial

assets' returns, or the way market activities respond "emotionally" (Tinyakova, 2012). Generally, volatility modelling has a far-reaching impact and it is very crucial for options trading, portfolio management, and risk evaluation, among others. Volatility is caused by many factors that include the behaviour of investors, the financial performance of organizations, predictable political events like democratic elections, news information, among others.

Ahmed and Naher (2021) applied the standard GARCH model and its extensions comprising the Threshold GARCH (TGARCH), Asymmetric Power ARCH model (APARCH), Exponential GARCH (EGARCH), and Integrated Generalised Autoregressive Conditional Heteroscedastic (IGARCH) models for volatility modelling and forecasting of the daily returns of the Dhaka Stock Exchange (DSE) from the period 27 January, 2013 to 6 November, 2017. The modelling was conducted using the Normal and Student's t error assumptions. The outcomes showed that the TGARCH(1,1)-Student's t was the most suitable candidate for the in-sample volatility estimation, capturing the inherent volatility asymmetry in the market returns.

Curto et al. (2009) applied the Autoregressive Moving Average-Generalised Autoregressive Conditional Heteroscedasticity (ARMA-GARCH) models fitted with each of the Normal, stable Paretian and Student's t assumed error distributions to the daily returns of the German, U.S., and Portuguese main equity market indices. The authors found that the fit of the models with the stable Paretian assumed error outperforms those of the Normal and Student's t. Using simulation and empirical modelling of the S&P 500 daily returns, Feng and Shi (2017) applied the standard GARCH model fitted with the Normal, Generalised Error Distribution (GED), Student's t and tempered stable assumed error distributions. The study revealed that the fit of the tempered stable distribution outperformed those of the other three assumed error distributions.

Gokcan (2000) modelled the monthly equity returns of seven emerging economies to compare the volatility prediction performance of GARCH(1,1) and EGARCH(1,1) models. The outcome showed that the GARCH(1,1) model performed better than the EGARCH(1,1) model. The volatility of many developed economies such as the London International Stock Exchange, Germany, Netherlands, France, and Italy has also been modelled best with the application of the GARCH models by Koutmos (1998), and Taylor and Poon (1992), among others.

Enow (2023) used the GARCH and ARCH models to model and forecast the stock market volatility of the CAC 40, Nasdaq, DAX, JSE, and Nikkei 225 financial markets from January 29, 2018 to January 29, 2023. The study revealed that the variance of the selected financial markets changes over time. Furthermore, the volatility in the CAC 40, DAX and JSE display fat-tails distributions, and it is anticipated to move by three standard deviations. Moreover, an increasing rate of volatility persistence existed in the CAC 40, DAX, and Nasdaq, with a decreasing rate of volatility persistence in the Nikkei 225 and JSE.

Mohandas (2023) applied the ARIMA and GARCH models to model and forecast the volatility in the S&P500 index from 2000 to 2022. Implied volatility (IV) was predicted using the two models to capture the shocks associated with volatility. VIX is commonly used by the entire US stock market to measure the volatility of the S&P500, while IV is the expectation of

the options market's future volatility (Mohandas, 2023). The other exogenous variables used in the modelling include trading volumes, bond yields and SP500 returns. The author's findings showed that the forecast of VIX is highly likely with the historical IVs using the ARIMA and GARCH models with improved efficiency and accuracy. Floros (2008) used the GARCH process to investigate the characteristics of volatility in the Egyptian CMA General index and the Israeli TASE-100 index over the period 1997 to 2007. The study used the daily data of these two markets and concluded that the GARCH models can describe the dynamics of daily equity returns such as leverage effects and volatility clustering.

With the application of the EGARCH model, Appiah-Kusi and Pescetto (1998) found that most of the African stock markets are characterized by changes in volatility levels, with some periods displaying extremely high volatility. The authors' findings indicated the presence of considerable asymmetries in the response of volatility to news. These asymmetries, however, are not always consistent with the conventional description of the leverage effect, since they observed that good news sometimes causes more reactions than comparable bad news. Kosapattarapim et al. (2012) applied the GARCH models fitted with six innovation distributions to three South East Asian emerging equity markets. Their findings showed that the fit of the models with non-Normal assumed innovation distributions improved the forecast than the model fitted with a Normal distribution innovation assumption.

Mamilla et al. (2023) applied the GARCH models to investigate the impact of returns volatility on nine indices of the Indian National Stock Exchange (NSE) before, during, and after the COVID-19 pandemic for periods 1 January 2018 to 31 December 2022. The indices include the NSE 100 ESG, NSE Bank, NSE 50, NSE Commodities, NSE Metal, NSE IT, NSE Realty, NSE Auto, and NSE FMCG. The study was used to examine the effect of volatility on the returns and risk of investment in the Indian stock market. The study observed that the COVID-19 period outperformed the period before COVID-19 and the overall periods. Hence, investors reaped higher returns during the period of the COVID-19 pandemic than before the crisis. Further outcomes revealed the existence of positive skewness in many of the indices under consideration; this implies that investors may be rewarded with higher returns on investments made.

The GARCH-type models that consist of the standard GARCH, EGARCH, and Glosten Jagannatan Runkle-GARCH (GJR-GARCH) were used by Ugurlu et al. (2014) to model the returns volatility of Turkey (XU100), Hungary (BUX), Poland (WIG), Czech Republic (PX), and Bulgaria (SOFIX) stock markets using daily data from January 8, 2001 to July 20, 2012. The study's finding showed no significant GARCH effect for the SOFIX index, while the GARCH, EGARCH, and GJR-GARCH effects existed in the returns of WIG, BUX, XU, and PX indices. The markets also witnessed the persistence of volatility shocks, with a significant influence of historical news on volatility.

Pandey and Kumar (2017) applied the standard GARCH model to estimate pronounced persistence in the Indian S&P CNX NIFTY 50 for the sample period January 1, 1997 to De-

cember 31, 2012, and found a high volatility persistence in the returns. The persistence of conditional volatility is a significant characteristic that describes its ability to propagate the impact of significant information or news on the volatility of an asset to future periods (Pandey and Kumar, 2017). Other studies on modelling of the persistence of volatility can be found in Bollerslev et al. (1992); Chou (1988); Engle and Bollerslev (1986); Pagan and Schwert (1990); Schwert (1989). Zarour and Siriopoulos (2008) used a Component GARCH (CGARCH) model to study the presence of volatility decomposition in the Middle East emerging markets. Their findings showed that volatility decomposition to its component is supported by the presence of permanent volatility that decays over a long period of time, and a transitory component to volatility in the Saudi Arabian, Oman, and Jordan markets. Simon and Amalia (2012) used the CGARCH model to study the daily yield behaviour of 11 European Economic and Monetary Union (EMU) nations. Their results showed that the permanent (long-run) component was described by a highly persistent and time-varying trend, whereas a strong mean-reversion to the trend was observed in the short-run component. Their findings agreed with the observations of Lee and Engle (1999). Studies associated with the decomposition of return volatility to its permanent-transitory components are increasingly found in the literature. The outcomes of these studies through the CGARCH suggest that volatility of returns can be decomposed into the long-run and short-run components (see Ane, 2006; Chen and Shen, 2000; Chiang et al., 2006; Lee and Engle, 1999; Speight et al., 2000).

Using the TGARCH model on the daily International Business Machines (IBM) stock, Tsay (2005) found asymmetry in the returns, and further observed that the asymmetric behaviour is much stronger in the TGARCH model than in the alternative GJR model. The issue of the persistence of stock returns volatility was investigated by Porterba and Summers (1986), with the claim that a long-time persistence of volatility will significantly impact stock prices. However, if the shocks to volatility are short-lived, the market will not be able to adjust the future discount rate that it requires on investments.

Engle and Patton (2001) used the GARCH(1,1) model to estimate various stylised facts of volatility in the Dow Jones Industrial Index returns from 1988 to 2000, and found a highly persistent volatility estimate with a half-life of about 73 days. Moreover, the tendency of volatility to persist over time is well documented in the literature (see Chou, 1988; Engle and Bollerslev, 1986; Pandey and Kumar, 2017, among others).

2.3 Applications of the GAS Model and Other Score-Driven Extensions

The outcomes of several studies have supported the stylised fact that the stock market returns exhibit a conditional volatility that is time-varying (e.g., Andersen et al., 2001; Bekaert and

Harvey, 1997; Bollerslev and Zhou, 2002; Brownlees and Engle, 2012; Engle, 2002; Pagan and Schwert, 1990). Consequently, such a time-based volatility directly affects investment decisions and financial modelling, such as risk management, asset pricing, portfolio optimization, and asset allocation (see Bali and Engle, 2010; Bollerslev et al., 1988; Creal et al., 2011; De Santis and Gerard, 1997; Hodrick, 1981). Hence, the search for a modelling design with the capability to capture the phenomenon of volatility clustering reliably and adequately is highly desirable. Volatility clustering is a process where large (small) changes in returns tend to follow large (small) changes of either sign.

Motivated by the short-comings of commonly adopted classic ARCH and GARCH models (see Bollerslev, 1986; Engle, 1982), which are incapable of adequately capturing the properties (or stylised facts) of conditional distribution, along with the lack of robustness, a score-driven class of volatility models was introduced in Creal et al. (2008), and Harvey and Chakravarty (2008), namely the Generalised Autoregressive Score (GAS); also known as Dynamic Conditional Score (DCS) (Ardia et al., 2019). Score-driven models are observation-driven models (Creal et al., 2008, 2013), that consist of modifications of the GARCH model, and they are proposed to mitigate the negative effects of outliers, while also capturing large changes through time (Haddad et al., 2023). One peculiar attraction of score-driven model types is that they can adequately identify distribution skewness due to the presence of outliers. Due to the dynamics of their conditional score, estimation via the maximum likelihood (ML) is easy (Ardia et al., 2019; Haddad et al., 2023). As a result of these attractive properties, many model extensions have emerged from the score-driven class. Such extensions include the asymmetric exponential score-driven model introduced in Creal et al. (2013), the dynamic models for volatility, location, and multivariate dependence for fat-tailed densities (Creal et al., 2011), and the observation-driven mixed measurement dynamic factor models introduced in Creal et al. (2014), among others (Haddad et al., 2023).

The GAS model is a score-driven model that can also be used to estimate the persistence of return volatility (Ardia et al., 2019; Creal et al., 2013; Harvey, 2013). Compared with the traditional classical GARCH model and its extensions, the literature on the GAS model and other score-driven extensions is still limited and ongoing. However, the usefulness of score-driven models has practically been demonstrated in many studies on different purposes and topics. For instance, empirical research works with the applications of the score-driven models include credit risk analysis (Creal et al., 2014), spatial econometrics (Blasques et al., 2014; Catania and Billé, 2017), dependence modelling (Harvey and Thiele, 2016; Janus et al., 2014), and credit default swap spread (Lange et al., 2017; Oh and Patton, 2018).

Other significant applications of score-driven models are in the modelling of high-frequency data (Gorgi et al., 2019; Opschoor et al., 2018), and forecasting of systemic risk (Bernardi and Catania, 2019; Eckernkemper, 2018; Oh and Patton, 2013). Specifically, Oh and Patton (2013) used the GAS model with a factor copula model for modelling system risk by studying a collection of daily credit default swap spreads on 100 U.S. firms from 2006 to 2012. The authors

found that the joint probability of distress is considerably higher during the period than in the pre-crisis period. See Ardia et al. (2019); Haddad et al. (2023); Lazar and Xue (2020); Patton et al. (2019) for more applications of the GAS and other score-driven models.

The Beta-t-EGARCH model is the first score-driven model that was presented in Harvey and Chakravarty (2008) and Creal et al. (2008), as a member of the EGARCH family (Nelson, 1991). Another score-driven EGARCH extensions using the generalised Student's t and two-component EGARCH models for the Student's t distribution were proposed in Harvey and Lange (2018, 2017). These models can incorporate skewness and asymmetry, and they have been empirically applied through analysis that explores commodity return and equity market series (Haddad et al., 2023).

Alanya-Beltran (2022) introduced an extension to the score-driven class known as the RS-Beta-t-EGARCH model for modelling time-varying financial market volatility with random shifts (RS). This model operates as an alternative method to long memory models, and it also identifies volatility clusters. The application of the model was demonstrated on equity returns in the South American emerging markets. The results of the study showed that the random shifts' estimates fit the key regime disturbance events in the studied period. The Monte Carlo simulations (MCS) experiments conducted in the study further revealed that the newly introduced model replicates the properties of the time and domain of the original series. Moreover, the new model also performed well in terms of out-sample prediction.

Another score-driven extension is the one-component Beta-t-QVAR-M model (Blazsek et al., 2022). It is an extension of the one-component Beta-t-EGARCH-M model of Harvey and Lange (2018). In the Beta-t-QVAR-M model, volatility is driven by a bivariate score-driven filter, which can be updated by score functions with regards to log-scale (i.e., the volatility that is non-linearly transformed) and location (or expected return). Haddad et al. (2023) proposed a new joint model known as the two-component Beta-t-QVAR-M-lev (quasi-vector autoregression in-mean with leverage) for modelling expected return and volatility forecasting. Leverage effects are incorporated into the score-driven volatility filter (to give Beta-t-QVAR-M-lev) in the specification of the one-component Beta-t-QVAR-M (Blazsek et al., 2022), which measures asymmetric impacts of unexpected returns on volatility (Haddad et al., 2023).

The maximum likelihood estimator for the two-component Beta-t-QVARM-lev is an extension of the one-component Beta-t-QVAR-M. The new model was demonstrated on the G20 stock market indices for the period January 2000 to April 2022, and its volatility forecasting performance outcome outperformed that of the competing two-component GARCH-M model when compared (Haddad et al., 2023). Another extension of the score-driven model is found in the work of Harvey and Lange (2018), where the one- and two-component Beta-t-EGARCH-M (i.e., Beta-t-exponential GARCH-in-mean) models are developed (Haddad et al., 2023). The models assume the concepts of the GARCH-M model of Engle et al. (1987) and the two-component GARCH model of Lee and Engle (1999) into score-driven volatility models. It is considered in two-component volatility models, e.g., Alizadeh et al. (2002); Lee and Engle

(1999), that volatility is driven by a short-run and a long-run component, where the former is used to capture temporary variation in volatility, and it incorporates leverage effects (Black, 1976), while the latter describes long-term variation in volatility. As stated in Alizadeh et al. (2002), long-memory behaviour can be captured by two-component volatility models (Haddad et al., 2023). The work of Harvey and Lange (2018) shows that the volatility prediction performance of the corresponding Beta-t-EGARCH models are improved by the one- and two-component Beta-t-EGARCH-M models.

2.4 Applications of the Beta-Skew-t-EGARCH Models

The Beta-Skew-t-EGARCH model is a skewed version of the original Beta-t-EGARCH model (Harvey and Sucarrat, 2014), and it is a dynamic and unrestricted version of the GAS model that is robust to outliers or jumps. It has performed quite well in modelling relevant features of volatility (Harvey and Sucarrat, 2014; Sucarrat, 2013), namely, conditional fat-tailedness, leverage (i.e., volatility asymmetry), and conditional skewness.

Using daily data points spanning the period 1 January 1999 to 12 October 2011 for a range of fifteen demeaned financial return series, Harvey and Sucarrat (2014) compared the fit of various Beta-t-EGARCH specifications (including the Beta-Skew-t-EGARCH model) with that of GARCH-GJR model (with leverage) developed by Glosten et al. (1993). The authors' findings showed that the Beta-t-EGARCH specifications outperformed the GARCH-GJR benchmark. The authors' findings further showed that the Beta-Skew-t-EGARCH model specification with a leverage effect, and either one or two components, gave the best outcomes overall. The findings further revealed the presence of leverage effects and negative skewness in most of the studied return series. Lastly, the comparison of the two versions of the Beta-Skew-t-EGARCH model showed that the one-component outperformed the two-component in twelve out of the fifteen studied financial return series. Sucarrat (2013) also used both the one- and two-component Beta-Skew-t-EGARCH models to estimate the long-memory behaviour and leverage effect in the financial return of the Nasdaq 100 composite index, and found that the two-component model outperformed the one-component model. Further outcomes showed that the short-run effect dominated the volatility process than the long-run effect. The author further compared the performance of the model with that of the GJR-GARCH-skewed Student's t model (Lee and Engle, 1999) and found that the Beta-Skew-t-EGARCH model performed better for both leverage effect and long-memory estimations.

Morris et al. (2019) reviewed existing works that included simulation techniques in the literature, and the authors used the findings to outline a general structured procedure for planning and reporting studies on simulation, with relevant guidance on the graphical and tabular presentation of outcomes. Chalmers and Adkins (2020) used a tutorial study to demonstrate relevant

guiding steps to writing effective computer code for MCS experiments using the SimDesign package in an R programming environment.

Several other research studies have been carried out on financial market volatility modelling and the development of reliable portfolio systems based on volatility forecasting and future equity prices with the applications of highly complex forecasting models (see Chronopoulos et al., 2018; Guha et al., 2016; Lv et al., 2018; Mamilla et al., 2023; Mehtab et al., 2021; Mehtab and Sen, 2020).

2.5 Applications of Fractional Integrated Models

A time series process that exhibits long memory is known as a fractionally integrated series (Baum et al., 2020). Granger and Joyeux (1980) and Hosking (1981) independently proposed the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model to account for a discrete-time model that shows hyperbolic decay of its autocorrelation function. The ARFIMA models avail an alternative modelling procedure to the ARIMA models. They allow time series data to display stationary ARMA attributes after being fractionally differenced (Koop et al., 1997). The ARFIMA models have been widely used by empirical researchers (see Baillie and Chung (1992); Cheung and Lai (1993), and Cheung (1993), among others). Some of the other empirical and theoretical studies on ARFIMA processes include those of Fox and Taqqu (1986), Sowell (1990, 1992), Diebold and Rudebusch (1989, 1991), Geweke and Porter-Hudak (1983), Robinson (1991), Lo (1991), Shea (1991), Yajima (1988, 1989), Porter-Hudak (1990), Li and McLeod (1986) and Cheung (2008).

Volatility persistence points to the existence of long-term memory in the second-order moment of returns and absolute returns. The long memory process can be attributed to the appearance of the interaction between many diverse information processes and, hence it is a characteristic of the returns process (Andersen and Bollerslev, 1997; Saha et al., 2020). The key attribute of a long-memory process is that its autocorrelation function fades away at a slower rate than that of a short-memory process but at a quicker rate than that of an integrated one. Various findings on the presence of long memory in time series volatility have led to the introduction of new techniques for risk aggregation and assessment, portfolio optimization, and forecasting, among others (Bobeica and Bojesteanu, 2008). Beneficial trading strategies that yield relevant gains can be created if long memory is allowed in the cross-section of asset returns (Nguyen et al., 2019; Saha et al., 2020).

Maheswaran (1990) has shown that the existence of long memory in stock returns has impactful implications for many concepts in financial economics. For instance, portfolio decisions and optimal consumption/savings may become very sensitive to investment periods if the returns are long-range dependent (Alper and Atilla, 2007). Some important explanatory factors that pertain to long memory include the speed of price adjustment (Zheng et al., 2018),

the extent of market openness (Lim and Brooks, 2010), the influence of institutional investors (Gabaix et al., 2006), bid-ask spreads, nonsynchronous trading (Campbell et al., 1998), and the institutional and economic differences between the developed and emerging markets (Liow, 2009).

Long memory models are used in biology, finance, climate, economics, and many other fields (see Beran, 1994; Beran et al., 2013; Palma, 2007; Vera-Valdés, 2021, 2024). With the stock market being a key economic indicator, researchers have given a lot of attention to examining the existence of long memory in stock returns. For instance, see Kaen and Rosenman (1986), Ding et al. (1993), Hiemstra and Jones (1997), Lo (1991), Mills (1993), Lee and Robinson (1996), and Sadique and Silvapulle (2001) on testing for the presence of long memory in stock returns.

Studies by Helms et al. (1984), Fang et al. (1994), Barkoulas et al. (1997), Booth et al. (1982), and Cheung and Lai (1993), among others, have found the presence of long memory in some kinds of foreign currency rates (Brania and Gurgul, 2015). The presence of long memory has also been found in bonds (Backus and Zin, 1993), derivatives (Barkoulas et al., 1999; Helms et al., 1984), currencies (Cheung and Lai, 1993), commodities (Cai et al., 2001; Elder and Serletis, 2008), and other specialized instruments (Madhavan and Arrawatia, 2016).

It has been noticed that there is varying evidence of the presence of long memory across markets and variables. For instance, while asset returns have been found to have weak to no sign of long memory, notably in the developed markets (Floros et al., 2007; Henry, 2002), volatility of assets has been found to present strong evidence of long memory (Bollerslev and Mikkelsen, 1996; Fleming and Kirby, 2011; Mighri and Mansouri, 2014). There are also conflicting outcomes in studies involving developing markets. While some studies found a strong presence of long memory in developing markets (Hull and McGroarty, 2014; McMillan and Thupayagale, 2009), others found that some developing markets have become more efficient than a few developed markets since the 2008 financial crisis (Mensi et al., 2019; Sensoy and Tabak, 2016).

Since the introduction of the work of Baillie et al. (1996) on testing for the fractional integration parameter d using the Fractionally Integrated GARCH (FIGARCH) approach, various studies have examined the existence of long memory in variance processes. Hiemstra and Jones (1997), for example, observed that the mean returns of many U.S. common stocks do not exhibit long-memory dependence, while the squared returns do (Sadique and Silvapulle, 2001). Bollerslev and Mikkelsen (1996) and Baillie et al. (1993) respectively found evidence of long memory in the S&P 500 stock index and the squared innovations of US$/DM exchange rate. These authors made the findings through the applications of the FIGARCH model to the respective series. More studies have shown the presence of long-memory dependencies with high persistence in the log-squared, squared, or absolute returns (Bollerslev and Wright, 2000).

Shea (1991) and Diebold and Rudebusch (1989) observed that economic fundamentals possess long memory behaviour using the methods introduced by Geweke and Porter-Hudak

(GPH) (see Geweke and Porter-Hudak, 1983). Crato and de Lima (1994) applied the GPH estimator to the squared residuals of various filtered U.S. stock returns indexes (Bobeica and Bojesteanu, 2008). Hassler and Woiters (1995) used the GPH approach to model long memory in inflation by estimating the memory parameter d (Baillie, 1996). Vera-Valdés (2024) used the semiparametric approaches involving the GPH estimator, the local Whittle estimator, and the exact local Whittle estimator to model the flow of the Nile River. Shea (1991) used the GPH approach to estimate fractional processes on interest rates, with a discussion of the long memory implication on the variance bounds tests that result from term structure. Some of the initial work that applied the GPH method appeared to find long memory in the spreads (Baillie, 1996). The Local Whittle estimators are the most stable among the general long-memory estimators (Hassler, 2011; Saha et al., 2020; Taqqu et al., 1995).

2.6 Summary of the Studies and Conclusion

Various authors like Curto et al. (2009), Feng and Shi (2017), and Ahmed and Naher (2021), among others, have empirically used the standard GARCH as described in the literature review in Section 2 to determine the most suitable error distribution among a set of selected assumed error distributions, for modelling the volatility of both the developed and developing markets. Other authors that have comparatively used the ARMA-GARCH models and their extensions to empirically model volatility and its persistence, leverage effects and other stylised facts of volatility in various developed financial markets include Gokcan (2000), Koutmos (1998), Taylor and Poon (1992), Mohandas (2023), Floros (2008), Appiah-Kusi and Pescetto (1998), Kosapattarapim et al. (2012), Ugurlu et al. (2014), Bollerslev et al. (1992), Engle and Patton (2001), Chou (1988), Pandey and Kumar (2017), Engle and Bollerslev (1986), and Mamilla et al. (2023), among other numerous authors.

Furthermore, authors like Creal et al. (2011, 2008, 2013), Harvey and Chakravarty (2008), Harvey (2013), Ardia et al. (2019), Haddad et al. (2023), Alanya-Beltran (2022), Blazsek et al. (2022), Harvey and Lange (2018), Harvey and Sucarrat (2014), Sucarrat (2013) have used the score driven GAS model and its extensions as described in Section 2 to model various stylised facts of volatility. Other authors like Catania and Billé (2017), Harvey and Thiele (2016), Janus et al. (2014), Oh and Patton (2018), Lange et al. (2017), Opschoor et al. (2018), Gorgi et al. (2019), Eckernkemper (2018), Oh and Patton (2013), Bernardi and Catania (2019), Ardia et al. (2019), Lazar and Xue (2020), Haddad et al. (2023), Lazar and Xue (2020) and Patton et al. (2019) have also used the GAS models in various aspects of time series modelling such as spatial econometrics, dependence modelling, forecasting of systemic risk, and credit default swap spread.

Granger and Joyeux (1980) and Hosking (1981) independently introduced the ARFIMA model for long memory modelling, and its applications have been widely used, mostly empiri-

cally, by authors like Fox and Taqqu (1986), Sowell (1990), Sowell (1992), Diebold and Rudebusch (1989), Diebold and Rudebusch (1991), Lo (1991), Shea (1991), Yajima (1988, 1989), Li and McLeod (1986) and Cheung (2008). Baillie et al. (1996) introduced the FIGARCH model for modelling long memory in financial returns volatility and its applications have been broadly used on various time series data by authors like Sadique and Silvapulle (2001), Bollerslev and Mikkelsen (1996), Baillie et al. (1993) and Bollerslev and Wright (2000). Also, Geweke and Porter-Hudak (1983) introduced the log periodogram frequency domain approach of long memory modelling, and its application can be found in the works of authors like Shea (1991), Diebold and Rudebusch (1989), Crato and de Lima (1994), Bobeica and Bojesteanu (2008), Hassler and Woiters (1995), and Vera-Valdés (2024).

In summary, as presented in this pool of research reviews, it is revealed that the applied ARCH/GARCH models and their variants, and the GAS model and its score-driven extensions are considerably effective in capturing and describing various features of volatility in the referenced markets. However, as observed in the reviewed literature, none of these has used the application involving the comparative collaboration of the fGARCH, GAS and Beta-Skew-t-EGARCH models as proposed in this study. Specifically, no evidence is currently available on an instructive simulation framework that can adequately guide (upcoming) researchers on the simulation steps for selecting a suitable conditional error distribution for volatility modelling through the GARCH and GAS models. This gap is more related to the GAS model. Hence, this study intends to fill this research gap with the proposed objectives (see Sections 4.3.1.2, 4.3.1.3, 5.2.3.1, 5.3.3) through tutorial studies involving these models.

In particular, this thesis intends to contribute to the nascent literature on volatility modelling by using algorithms based on the time-varying fGARCH and GAS model to design simulation approaches (that can be verified empirically) to determine an adequate conditional error distribution for volatility persistence estimation. Furthermore, the study will collaboratively apply the score-driven extension through the Beta-Skew-t-EGARCH model, with the fGARCH and GAS models to estimate crucial volatility characteristics for valuable insights on risk assessment and investment decision-making. Both the one- and two-component Beta-Skew-t-EGARCH models will be applied through empirical and simulation approaches to comparably investigate the presence of stylised facts of financial returns volatility in the stock market as detailed in Chapter 6 of this thesis.

The Beta-Skew-t-EGARCH model is particularly applied in this study because of its robust approach to modelling stylised facts of volatility. It does not only efficiently model leverage effect or volatility asymmetry, conditional skewness, and long-memory behaviour like the other score-driven extensions, but it also effectively models other attributes of volatility like the conditional fat-tailedness of financial returns, and the responses of volatility to shocks through its one and two components (see Harvey and Sucarrat, 2014; Sucarrat, 2013). Hence, its applicability is crucial for estimations in this study. Moreover, this study intends to use both the time-domain and frequency-domain approaches to determine the presence of long memory

in the volatility process. In other words, the study will use a parametric model through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle estimator, and the exact local Whittle estimator to estimate and determine the presence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns. To the best of the author's knowledge, this collaborative approach on the return, squared returns and absolute returns has not been used in previous studies.

Chapter 3

Methodology

3.1 Introduction

This chapter describes the various techniques applied in this study's modelling and simulation framework developments. The modelling is executed using R-add-on statistical software packages obtainable from CRAN repository (https://cran.r-project.org/). These packages include the rugarch (Ghalanos, 2018, 2022, 2022), SimDesign (Chalmers and Adkins, 2020), tidyverse (Wickham et al., 2019), zoo (Zeileis and Grothendieck, 2005), aTSA (Qiu, 2015), forecast (Hyndman and Khandakar, 2008), GAS (Ardia et al., 2019; Catania et al., 2020), patchwork (Pedersen, 2020), LongMemoryTS (Leschinski et al., 2019), and betategarch (Sucarrat, 2016).

3.2 The GARCH Model

The GARCH model was developed by Bollerslev (1986) as a generalisation of the Autoregressive Conditional Heteroscedasticity (ARCH) model introduced by Engle (1982). It is a classical model that is normally defined by its conditional mean and variance equations for modelling financial returns volatility (Kim et al., 2020). The mean equation is stated as

$$r_t = \mu_t + \varepsilon_t, \tag{3.1}$$

where r_t is the return series, $\varepsilon_t = z_t\sigma_t$ denotes the residual part of the return series that is random and unpredictable, where $z_t \sim N(0,1)$ are the standardised residuals which are independent and identically distributed (i.i.d.) random variables with mean 0 and variance 1 (McNeil and Frey, 2000; Smith, 2003), μ_t is the mean function that is usually stated as an Autoregressive Moving Average (ARMA) process,

$$\mu_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i}, \tag{3.2}$$

where ϕ_i (i = 1, . . ., p) and φ_i (i = 1, . . ., q) are unknown parameters. The variance equation of the GARCH(u, v) model is defined as

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_v\varepsilon_{t-v}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_u\sigma_{t-u}^2, \tag{3.3}$$

where $\omega > 0$ is the intercept term (white noise), coefficients $\alpha_i \geq 0$ (i = 1, . . ., v) and $\beta_i \geq 0$ (i = 1, . . ., u), respectively, measure the short-term and long-term effects of ε_t on the conditional variance (Maciel and Ballini, 2017). The non-negativity restrictions on the unknown parameters, α_i and β_i, are imposed for $\sigma_t^2 > 0$. Covariance stationarity of the GARCH model is ensured when $\sum_{i=1}^{\max(v,u)}(\alpha_i + \beta_i) < 1$, i.e., the statistical properties of the process do not change over time. The unconditional variance of ε_t is $\sigma^2 = E\left(\varepsilon_t^2\right) = \omega/\{1 - (\sum_{i=1}^{v}\alpha_i + \sum_{i=1}^{u}\beta_i)\}$ (Zivot, 2009), i.e., it is finite. Equation (3.3) shows that the conditional variance σ_t^2 is a linear function of past squared innovation ε_{t-i}^2 and past conditional variances σ_{t-i}^2. The GARCH model is a more parsimonious specification (Nelson, 1991; Samiev, 2012) since it is an equivalence of a certain ARCH(∞) model (Zivot, 2009). When $u = 0$ in Equation (3.3), the GARCH model changes to the ARCH model with conditional variance stated as

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_v\varepsilon_{t-v}^2. \tag{3.4}$$

GARCH(1,1) is the simplest model specification with u = 1 and v = 1 in Equation (3.3), and it is conceivably the best candidate GARCH model for several applications (Fan et al., 2014; Zivot, 2009). The volatility persistence of the GARCH$(1, 1)$ model is defined as α_1 + β_1 (see Engle and Bollerslev, 1986; Ghalanos, 2018). Volatility persistence is used to evaluate the speed of decay of shocks to volatility (Kim et al., 2020). Volatility exhibits long persistence into the future if α + $\beta \to 1$, hence the closer the sum of the coefficients is to one (zero), the greater (lesser) the persistence. However, if the sum is equal to one, then shocks to volatility persist forever and the unconditional variance is not determined by the model. This process is called integrated-GARCH (Chou, 1988; Engle and Bollerslev, 1986). If the sum is greater than one, the conditional variance process is explosive, suggesting that shocks to the conditional variance are highly persistent.

For the maximum likelihood estimation (MLE), denote the likelihood function as

$$L(\boldsymbol{\vartheta}|\boldsymbol{\varepsilon}) = \sum_{t=1}^{N} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right), \tag{3.5}$$

where $\boldsymbol{\vartheta} = (\mu, \omega, \alpha_1, ..., \alpha_v, \beta_1, ..., \beta_u)^\top$ is a vector of parameters, and $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_N)$ is a realisation of length N. The quasi-maximum likelihood estimation (QMLE) based on the Normal distribution and MLE have the same set of instructions for estimating $\hat{\boldsymbol{\vartheta}}$; the only difference, however, is in the estimation of a robust standard deviation of $\hat{\boldsymbol{\vartheta}}$ in QMLE (see Bollerslev and Wooldridge, 1992; Feng and Shi, 2017; Francq and Zakoïan, 2004).

The maximised log-likelihood function with Student's t distribution (Duda and Schmidt, 2009) is stated as

$$\ln L(\boldsymbol{\vartheta}|\boldsymbol{\varepsilon}) = \sum_{t=1}^{N} \left(\ln \left[\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}} \right] - \frac{1}{2}\ln(\sigma^2) - \left(\frac{\nu+1}{2}\right) \ln\left[1 + \frac{\varepsilon_t^2}{\sigma^2(\nu-2)}\right] \right), \tag{3.6}$$

where $\Gamma(\cdot)$ and ν are the gamma function and degree of freedom, respectively.

3.3 The fGARCH Model

The family GARCH (fGARCH) model, developed by Hentschel (1995), is an inclusive model that nests some important symmetric and asymmetric GARCH models as sub-models. The nesting includes the simple GARCH (sGARCH) model (Bollerslev, 1986), the Absolute Value GARCH (AVGARCH) model (Schwert, 1990; Taylor, 1986), the GJR GARCH (GJRGARCH) model (Glosten et al., 1993), the Threshold GARCH (TGARCH) model (Zakoian, 1994), the Nonlinear ARCH (NGARCH) model (Higgins and Bera, 1992), the Nonlinear Asymmetric GARCH (NAGARCH) model (Engle and Ng, 1993), the Exponential GARCH (EGARCH) model (Nelson, 1991), and the Asymmetric Power ARCH (apARCH) model (Ding et al., 1993). The sub-model apARCH is also a family model (but less general than the fGARCH model) that nests the sGARCH, AVGARCH, GJRGARCH, TGARCH, NGARCH models, and the Log ARCH model (Geweke, 1986; Pantula, 1986). The fGARCH(p, q) model is stated as

$$\sigma_t^{\gamma} = \omega + \sum_{j=1}^{q} \alpha_j \sigma_{t-j}^{\gamma} (|z_{t-j} - \lambda_{2j}| - \lambda_{1j}\{z_{t-j} - \lambda_{2j}\})^{\delta} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^{\gamma}. \tag{3.7}$$

Equation (3.7) is the conditional standard deviation's Box–Cox transformation, where the transformation of the absolute value function is carried out by the parameter δ, and γ determines the shape. The parameters λ_{2j} and λ_{1j} control the shifts for asymmetric small shocks and rotations for large shocks, respectively. This robust fGARCH model allows different powers for σ_t and z_t to drive how the residuals are decomposed in the conditional variance equation. The fit of the full fGARCH model can be implemented with $\gamma = \delta$ (see Ghalanos, 2018). Volatility clustering in the returns can be quantified through the model's volatility persistence stated as

$$\hat{P} = \sum_{j=1}^{p} \beta_j + \sum_{j=1}^{q} \alpha_j \varrho_j, \tag{3.8}$$

where ϱ_j, expressed in Equation (3.9), is the expected value of z_t in the absolute value asymmetry term's Box–Cox transformation. The persistence is obtained in this study through the "persistence()" function in the R rugarch package. See Ghalanos (2022); Hentschel (1995) for details on fGARCH and the nested models.

$$\varrho_j = \mathrm{E}(|z_{t-j} - \lambda_{2j}| - \lambda_{1j}(z_{t-j} - \lambda_{2j}))^{\delta} = \int_{-\infty}^{\infty} (|z - \lambda_{2j}| - \lambda_{1j}(z - \lambda_{2j}))^{\delta} f(z, 0, 1, \ldots) dz \tag{3.9}$$

3.4 The GAS Model

Volatility persistence can be effectively modelled through a Score Driven (SD) model known as the Generalised Autoregressive Score (GAS) proposed by Harvey (2013) and Creal et al. (2013) (see Ardia et al., 2019, 2018; Catania et al., 2020). The model applies the score of the conditional density function to determine the time variation in the parameters. The score functions are robust to outliers because they discount extreme values by reducing the number of extreme observations (Alanya-Beltran, 2022; Harvey, 2013; Owusu Junior and Alagidede, 2020). Thus, the model is suitable for modelling fat-tailed and skewed time series data like financial returns (Harvey, 2013; Opschoor et al., 2018). In addition, extensions to other time series dynamics, like asymmetry and long memory behaviour, are possible through the GAS process. Moreover, maximum likelihood estimation (MLE) through the model is easy to implement (Ardia et al., 2019).

Creal et al. (2011, 2013, 2014) showed that the robust GAS models incorporate many familiar observation-driven models that include the Autoregressive Conditional Heteroscedasticity (ARCH) model (Engle, 1982), the simple (standard) GARCH model (Bollerslev, 1986), the Exponential GARCH (EGARCH) model (Nelson, 1991), the Autoregressive Conditional Duration (ACD) model (Engle and Russell, 1998), the Autoregressive Conditional Intensity (ACI) model (Russell, 2001), the Beta-t-GARCH model (Harvey, 2013; Harvey and Chakravarty, 2008), the Autoregressive Conditional Multinomial (ACM) model (Rydberg and Shephard, 2003), and several related models. In other words, new time-varying parameter models of empirical interest are easy to formulate due to the flexibility of the GAS framework (see Blasques et al., 2014).

3.4.1 Specification of the GAS Model

Let an $N \times 1$ vector $\boldsymbol{y}_t$ be the dependent variable of interest, $\boldsymbol{\vartheta}_t$ the vector of time-varying parameters, $\boldsymbol{x}_t$ a vector of exogenous variables (i.e., the covariates), all at time t, and $\boldsymbol{\xi}$ a vector of static parameters. Define $\boldsymbol{Y}^t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$, $\boldsymbol{\Theta}^t = \{\boldsymbol{\vartheta}_0, \boldsymbol{\vartheta}_1, \ldots, \boldsymbol{\vartheta}_t\}$, and $\boldsymbol{X}^t = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t\}$ (Creal et al., 2012, 2013). The information set that is available at time t consists of $\{\boldsymbol{\vartheta}_t, \mathcal{F}_t\}$, where

$$\mathcal{F}_t = \{\boldsymbol{Y}^{t-1}, \boldsymbol{\Theta}^{t-1}, \boldsymbol{X}^t\}, \quad \text{for } t = 1, \ldots, n. \tag{3.10}$$

It is assumed that the generation of $\boldsymbol{y}_t$ is carried out by the observation density

$$\boldsymbol{y}_t \sim p\left(\boldsymbol{y}_t | \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi}\right). \tag{3.11}$$

It is further assumed that the mechanism required to update the time-varying parameter ϑ_t is given by the autoregressive updating equation

$$\boldsymbol{\vartheta}_{t+1} = \boldsymbol{\kappa} + \sum_{i=1}^{p} \boldsymbol{A}_i s_{t-i+1} + \sum_{j=1}^{q} \boldsymbol{B}_j \vartheta_{t-j+1}. \tag{3.12}$$

Ardia et al. (2019) presented Equation (3.12) as:

$$\boldsymbol{\vartheta}_{t+1} \equiv \boldsymbol{\kappa} + \mathbf{A}\boldsymbol{s}_t + \mathbf{B}\boldsymbol{\vartheta}_t, \tag{3.13}$$

where $\boldsymbol{\kappa}$, $\mathbf{A}$ and $\mathbf{B}$ are matrices of coefficients with appropriate dimensions and they are functions of the time-invariant parameter $\boldsymbol{\xi}$, while $\boldsymbol{s}_t$ is the scaled score, a suitable function of past data, $\boldsymbol{s}_t = \boldsymbol{s}_t(\boldsymbol{y}_t, \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi})$ (see Creal et al., 2012, 2013). The $\boldsymbol{s}_t$ indicates the direction to update the vector of parameters from $\boldsymbol{\vartheta}_t$ to $\boldsymbol{\vartheta}_{t+1}$. That is, it acts as a steepest ascent algorithm to improve the local fit of the model given the current parameter position. Vector $\boldsymbol{\kappa}$ is used to control the level of the process $\boldsymbol{\vartheta}_t$, while matrix $\mathbf{A}$ is used to determine (or control) the impact of $\boldsymbol{s}_t$ on $\boldsymbol{\vartheta}_{t+1}$, and matrix $\mathbf{B}$ is used to determine the persistence of the process (Ardia et al., 2019; Koopman et al., 2017). To implement the GAS models, matrices $\mathbf{A}$ and $\mathbf{B}$ are constrained to exist as diagonals (see Ardia et al., 2019; Oh and Patton, 2018). For instance, in a GAS model with Student's t conditional distribution, $\mathbf{A} \equiv \text{diag}(a_\mu, a_\phi, a_\nu)$ and $\mathbf{B} \equiv \text{diag}(b_\mu, b_\phi, b_\nu)$, where μ, ϕ and ν are location, scale and shape parameters, respectively with b_μ denoteing the persistence of the conditional mean (location), while b_ϕ the persistence of the conditional variance (or scale) (see Ardia et al., 2019).

The approach used in the GAS is based on the observation density in Equation (3.11) for a given parameter $\boldsymbol{\vartheta}_t$. When an $\boldsymbol{y}_t$ observation is realised, the time-varying $\boldsymbol{\vartheta}_t$ to the next period $t+1$ can be updated using Equation (3.12) with

$$\boldsymbol{s}_t = \boldsymbol{S}_t \cdot \boldsymbol{\nabla}_t, \qquad \boldsymbol{\nabla}_t = \frac{\partial \ln p\left(\boldsymbol{y}_t | \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi}\right)}{\partial \boldsymbol{\vartheta}_t}, \qquad \boldsymbol{S}_t = S\left(t, \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi}\right), \tag{3.14}$$

where $S(\cdot)$ represents a matrix function, $\boldsymbol{\nabla}_t$ is the score of Equation (3.11) evaluated at $\boldsymbol{\vartheta}_t$, ln is the natural logarithm, and $\boldsymbol{S}_t$ denotes the scaling matrix (see Ardia et al., 2019; Creal et al., 2012, 2013). The score of the GAS model depends on the complete density structure (Creal et al., 2012, 2013). Given that the driving mechanism in Equation (3.12) depends on the scaled score vector in Equation (3.14), the GAS model with orders p and q can be defined by Equations (3.11), (3.12) and (3.14). This can be referred to as the GAS(p,q) model, where the orders p and q are typically taken as $p = q = 1$ (see Creal et al., 2012, 2013). However, readers can refer to Creal et al. (2013) for details on including more lags in the GAS process (Blasques et al., 2014).

To account for the variance of $\boldsymbol{\nabla}_t$, Creal et al. (2013) suggested setting $\boldsymbol{S}_t$ to the inverse of the information matrix ($\boldsymbol{\mathfrak{I}}$) to a power $\gamma > 0$ of $\boldsymbol{\vartheta}_t$ (Ardia et al., 2019). Precisely,

$$\boldsymbol{S}_t = \boldsymbol{\mathfrak{I}}_{t|t-1}^{-\gamma}, \qquad \boldsymbol{\mathfrak{I}}_{t|t-1} = \mathrm{E}_{t-1}\left[\boldsymbol{\nabla}_t \boldsymbol{\nabla}_t^\top\right], \tag{3.15}$$

where the expectation $\mathtt{E}_{t-1}$ is taken with respect to the conditional distribution of $\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}$. The parameter γ usually takes 0, $\frac{1}{2}$, or 1 value (Ardia et al., 2019), but other choices of $\boldsymbol{S}_t$ are also possible (see Blasques et al., 2014). If $\gamma = 0$, $\boldsymbol{S}_t$ becomes an identity matrix ($\mathbf{I}$), which implies there is no scaling. If $\gamma = \frac{1}{2}$, then the conditional score $\boldsymbol{\nabla}_t$ is pre-multiplied by the square root of its covariance matrix $\boldsymbol{\mathfrak{I}}_t$, but if $\gamma = 1$, the conditional score $\boldsymbol{\nabla}_t$ is pre-multiplied by the inverse of its covariance matrix $\boldsymbol{\mathfrak{I}}_t$. The scaled score $\boldsymbol{s}_t$ is a martingale difference with respect to the distribution of $\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}$, i.e., $\mathtt{E}_{t-1}[\boldsymbol{s}_t] = 0$ for all t (see Ardia et al., 2019).

The GAS framework incorporates many familiar observation-driven models as a special case for an appropriate choice of the scaling matrix $\boldsymbol{S}_t$ (Creal et al., 2013). That is, various classes of observation-driven models can be obtained with different choices of the scaled score $\boldsymbol{s}_t$, hence these types of models are referred to as score-driven models (Creal et al., 2012, 2013; Koopman et al., 2017). In other words, Equations (3.11), (3.12) and (3.14) encompass a large set of common time series models for different choices of $\boldsymbol{s}_t$ (Blasques et al., 2014). For more details on the GAS models, see Ardia et al. (2019); Blasques et al. (2017); Catania et al. (2020); Creal et al. (2012, 2011, 2013, 2014); Harvey (2013); Koopman et al. (2017).

3.5 Persistence and Mean Reversion in Volatility

Volatility can be described as persistent if today's return produces a large effect on the prediction variance for many periods in the future (Engle and Patton, 2001). Volatility clustering means that small volatility shocks are followed up by small shocks while large volatility shocks are followed by large shocks in turn. Hence, a period of low volatility will be followed by a volatility rise, while a period of high volatility will sooner or later make way for more normal volatility (Engle and Patton, 2001). Mean reversion in volatility implies that there is a normal volatility level to which volatility will return eventually. A familiar classical measure of volatility persistence is called the "half-life" of volatility, denoted as $h2l$ (Ghalanos, 2018). This can be described as the number of days it will take the volatility to revert or move halfway back towards its unconditional mean after deviating from it (see Engle and Patton, 2001; Ghalanos, 2018) and is given by

$$h2l = \frac{\log_e \frac{1}{2}}{\log_e \hat{P}}, \tag{3.16}$$

where $\log_e$ denotes the natural logarithm, and $\hat{P}$ represents the estimate of the persistence parameter.

3.6 Test for Stationarity

The unit root tests are used to test the data series for stationarity in order to avoid a scenario where non-stationarity in the data leads to problems in statistical inference. Stationarity can be achieved by differencing the logged data. A time series process is non-stationary when the linear stochastic process has a unit root. An integrated order one data series which is non-stationary is stated as $I(1)$, while a stationary one is described as $I(0)$. The two unit root tests applied in this study to examine the stationarity of the series are the Augmented Dickey Fuller (ADF) test and Phillips-Perron (PP) test. For comparison, the study further applies stationarity tests through the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test.

3.6.1 The Augmented Dickey Fuller Test

The Augmented Dickey Fuller (ADF) test of Dickey and Fuller (1979) is used to test for stationarity of the price index and return data series. It is an augmented form of the Dickey-Fuller (DF) test for a larger and more complicated set of models of time series (than the DF test that is only valid if the series is a first-order autoregressive process, i.e., AR (1)) (Box and Jenkins, 1970) defined as:

$$y_t = \rho y_{t-1} + \boldsymbol{\delta}' \mathbf{A}_t + \varepsilon_t, \quad t = 1, ..., T., \tag{3.17}$$

where the residuals ε_t are assumed to be white noise, $\mathbf{A}_t$ is the vector of the deterministic terms, while ρ and $\boldsymbol{\delta}'$ are parameters to be estimated. Equation (3.17) can be estimated to obtain the standard DF test (in Equation 3.18) by taking the first difference where y_{t-1} is subtracted from both sides of the equation, and $\varphi = \rho - 1$.

$$\begin{aligned} y_t - y_{t-1} &= \rho y_{t-1} - y_{t-1} + \boldsymbol{\delta}' \mathbf{A}_t + \varepsilon_t \\ \Delta y_t &= (\rho - 1) y_{t-1} + \boldsymbol{\delta}' \mathbf{A}_t + \varepsilon_t \\ \Delta y_t &= r_t = (\rho - 1) r_{t-1} + \boldsymbol{\delta}' \mathbf{A}_t + \varepsilon_t \\ r_t &= \varphi r_{t-1} + \boldsymbol{\delta}' \mathbf{A}_t + \varepsilon_t, \quad \text{for, } t = 1, ..., T. \end{aligned} \tag{3.18}$$

The formulation of the ADF test permits higher order autoregressive processes since many economic and financial time series have a more complicated dynamic structure than what a simple AR(1) model can capture (Rossi, 2014). Based on this, the ADF test is an improved extension of the DF test and it is used to construct a parametric correction for higher-order correlation on the assumption that the series follows an AR(p) process (see Equation 3.19). The null and alternative hypotheses of the ADF test can be stated as:

$$H_0 : \rho = 1, \text{ data is not stationary } I(1)$$
$$\text{versus}$$
$$H_1 : |\rho| < 1, \text{ data is stationary } I(0).$$

The ADF test is implemented with the assumption that the dynamics of the data take an ARMA structure, and the test is focused on estimating the test regression:

$$r_t = \varphi r_{t-1} + \delta' \mathbf{A}_t + \sum_{j=1}^{p} \Psi_j \Delta r_{t-j} + \varepsilon_t, \tag{3.19}$$

where $\mathbf{A}_t$ is an optional vector of exogenous regressors in addition to a linear trend or a constant, or a constant and trend. The p lagged difference terms, Δr_{t-j}, are used to approximate the AR structure of the errors, with the value of p set to make the homoscedastic error serially uncorrelated. The first difference operator is represented by Δ, while $\boldsymbol{\delta}'$, Ψ, and $\varphi = (\rho - 1)$ are the parameters to estimate. The test statistic is based on the least squares estimates stated as:

$$t_{ADF} = \frac{\hat{\varphi}}{SE(\hat{\varphi})}, \tag{3.20}$$

where $SE(\hat{\varphi})$ is its standard error estimate, and $\hat{\varphi}$ is the least square estimate of φ. The null hypothesis can be rejected if the critical value is greater than the test statistic value, with the conclusion that the series is not integrated of order 1, i.e., "no unit-root is present". One practical downside of using the ADF test is in the choice of lags p. If p is too large, the power of the test will be reduced, but if p is too low, the test will be affected by the autocorrelation. Hence, Schwert (1989) proposed a procedure to choose the maximum lag $p_{max} = 12(T/100)^{1/4}$ (Arltová and Fedorová, 2016).

The limiting distribution of test statistics for the ADF models is similar to the distribution of DF test statistics; Dickey (1976) and MacKinnon (1991) tabulated this for $T \to \infty$ (see Arltová and Fedorová, 2016). The DF test statistic follows the Dickey-Fuller distribution, and critical values of the distribution (made available by a simulation) have been tabulated in Dickey (1976) and Fuller (1976). The asymptotic distributions of the unit root tests are generally non-normal and non-standard (Rossi, 2014).

3.6.2 Phillips-Perron (PP) Test

An alternative method to test for a unit root, with the serial correlation under control, was developed by Phillips and Perron (1988). This technique is known as the Phillips-Perron (PP) method and it is based on the Dickey-Fuller approach to test the null hypothesis that a time series is integrated of order 1 (i.e. H_0: $\rho = 1$) in $r_t = (\rho - 1)r_{t-1} + \delta' \mathbf{A}_t + \varepsilon_t$. The PP test also controls for a higher order serial correlation in the returns beyond the AR(1) limit of the Dickey-Fuller (DF) t-test through making a non-parametric correction to the t-test statistic. In addition,

the PP test addressed the problems frequently encountered in the selection of maximum lag p in a regression model for a unit root testing by using the standard Dickey-Fuller test with non-parametrically modified test statistics (Arltová and Fedorová, 2016).

The PP method adjusts the t-ratio of the coefficient to make sure that serial correlation has no effect on the asymptotic distribution of the test statistic. The PP test is formulated on the statistic:

$$\tilde{t}_{\varphi} = t_{\varphi}\left(\frac{\gamma_0}{f_0}\right)^{1/2} - \frac{T(f_0-\gamma_0)(SE(\hat{\varphi}))}{2f_0^{1/2}s}, \tag{3.21}$$

where the coefficient estimate standard error is represented by $SE(\hat{\varphi})$, t_{φ} is the t-ratio of φ, while $\hat{\varphi}$ is the estimate, and s denotes the standard error of the test regression. The term f_0 denotes an estimator of the residual spectrum at frequency zero, and γ_0 is a consistent estimate of the error variance in $r_t = (\rho-1)r_{t-1} + \delta'\mathbf{A}_t + \varepsilon_t$, and it is computed as $(T-k)s^2/T$, with k denoting the number of regressors (EViews, 2016). The null hypothesis that a time series is non-stationary (or integrated of order 1) can be rejected if p-value < 0.05.

The two choices to be made when performing the PP test are, first, to indicate whether the test regression should contain (or not contain) a constant, a constant and a linear time trend. Second, a method for estimating f_0 would have to be chosen, and likely choices can be based on kernel-based sum-of-covariances, or the autoregressive spectral density estimation (EViews, 2016). The asymptotic distribution of the PP modified t-ratio and that of the ADF statistic is the same (Rossi, 2014). However, the Phillips–Perron test' performance is worse in finite samples than the ADF test (Davidson and MacKinnon, 2004).

3.6.3 The Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) Test

The hypothesis supporting the KPSS (Kwiatkowski et al., 1992) test is described in contrast to those of the ADF and PP unit root tests that are focused on testing the null hypothesis that the time series y_t is non-stationary, $I(1)$, against the alternative stationarity, $I(0)$. Hence, the KPSS approach is based on testing the null hypothesis that the time series y_t is stationary, $I(0)$, around a deterministic trend (i.e., trend-stationary) against the alternative of a unit root, $I(1)$. The KPSS time series is described as the sum of deterministic trend (d_t), random walk (rw_t), and stationary random error (ε_t) (Arltová and Fedorová, 2016). It is formulated on the model:

$$\begin{aligned} y_t &= d_t + rw_t + \varepsilon_t \\ rw_t &= rw_{t-1} + u_t, \end{aligned} \tag{3.22}$$

where d_t is the deterministic part of the model, ε_t are i.i.d. $N(0,\sigma_{\varepsilon}^2)$, u_t are i.i.d. $N(0,\sigma_u^2)$, and rw_t is a random walk with variance σ_u^2.

The KPSS test is based on a Lagrange multiplier (LM) test of the hypothesis that the random walk has a variance of zero, that is, $H_0 : \sigma_u^2 = 0$, which indicates that rw_t is a constant, against the alternative hypothesis $H_1 : \sigma_u^2 > 0$. The test statistic is defined as:

$$\text{LM} = \sum_{t=1}^{T} s_t^2 / \hat{\sigma}_\varepsilon^2, \tag{3.23}$$

where $\hat{\sigma}_\varepsilon^2$ is the variance's estimate of process ε_t from the equation (3.22), and s_t is a cumulative residual function i.e., $s_t = \sum_{t=1}^{T} \hat{\varepsilon}_t, \;\; t = 1, 2, \ldots, T$. A simulation is used to derive the critical values of the LM test and they are based upon the asymptotic outputs presented in KPSS (Table 1, page 166) (EViews, 2016). The series has a unit root if the null hypothesis is rejected. Moreover, if the outcome of the KPSS test shows a stationary process, and those of the ADF and PP tests show a unit root, the user should be cautious and opt for the KPSS outcome (Rossi, 2014).

To conclude, since financial series modelling exhibits a more complex structure than can be modelled by the traditional AR(1) model-based DF test, this study adopts a robust approach by applying the three stationarity testing approaches (i.e., the unit root ADF and PP tests, and the stationarity test through the KPSS method), and compares their results.

3.7 Test for Serial Correlation

Following the test for the stationarity of the data, it is relevant to test for the existence of short-range linear dependence called autocorrelation or serial correlation in the residuals after an ARMA model is fitted, which may exist because of the relationship between a variable observation, such as financial data, and its lagged ones over various time intervals. This test can be carried out using the Weighted Ljung-Box test of Fisher and Gallagher (2012). This weighted Portmanteau test statistic can be stated as:

$$\widetilde{Q}_W = n(n+2) \sum_{k=1}^{m} \frac{(m-k+1)}{m} \frac{\hat{r}_k^2}{n-k}, \tag{3.24}$$

where m is the tested number of lags and n represents the sample size. The statistic $\widetilde{Q}_W$ is intepreted as a weighted Ljung-Box test, where $\hat{r}_k^2$ is the residual sample autocorrelation of order k. The residual at lag m is assigned the least weight, $1/m$, while the residual at lag one is assigned the most weight, 1.

The test statistic is distributed asymptotically as χ^2 random variable with $p+q$ degrees of freedom; where p and q are the autoregressive and moving average orders, respectively, in the fitted model. If p-value < 0.05, we reject the null hypothesis, i.e., H_0 : No serial correlation in the residuals.

3.8 Test for Heteroscedasticity or ARCH effects

After the short-range dependence is filtered from the data series, it is necessary to test for the presence of ARCH effects or heteroscedasticity in the data before fitting the autoregressive models to the residuals of the return series. The existence of heteroscedasticity will be tested through the ARCH Lagrange Multiplier (ARCH-LM) test as described in the section that follows.

3.8.1 ARCH LM test

The presence of ARCH effect in the residuals of the markets' returns are tested by applying the ARCH-LM test for ARCH effects proposed by Engle (1982). To test for ARCH effects in the conditional variance of returns, the testing process is in two stages: First, we fit an ARMA(p, q) model to the return series:

$$r_t = \alpha_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i} + \varepsilon_t, \tag{3.25}$$

where α_0 is the constant term, while ϕ and φ are AR and MA parameters to be estimated, respectively. The linear regression on the model will be run to obtain the residual ε_t. Following that, a regression of the squared residuals (ε_t^2) in Equation (3.26) is run on the constant term and lagged squared residuals up to k order to test for ARCH effects (heteroscedasticity).

$$\hat{\varepsilon}_t^2 = \varphi_0 + \left(\sum_{i=1}^{k} \varphi_i \hat{\varepsilon}_{t-i}^2 \right) + \psi_t, \tag{3.26}$$

where ψ_t is the error term of the regression model. The Engle's LM test statistic has asymptotic χ^2 distribution with ν degrees of freedom, where ν is the number of estimated parameters, i.e., $v = k + 1$. The test can be computed as the product of the number of observations (N) and (R^2) from the test regression, i.e., (NR^2), where R represents the sample multiple correlation coefficient (Wang et al., 2005). The existence of ARCH effect can be tested with the hypothesis:

$H_0 : \ \varphi_1 = \varphi_2 = \cdots = \varphi_k = 0$
against the alternative hypothesis
$H_1 : \ \varphi_1 \neq 0, \ \varphi_2 \neq 0, \ \cdots, \ \varphi_k \neq 0$
i.e., $\alpha_i \neq 0$ for some i.

This is the null hypothesis of "no ARCH effect" in the residual up to order k. The null hypothesis can be rejected if the critical value is less than the test statistic value. The existence of heteroscedasticity implies that volatility is time varying, which in turn validates the use autoregressive modelling.

3.9 Model Selection

Model selection in this study is carried out using information criteria (IC). This method is applied in preference to other methods because model selection using information criteria, based on information-theoretic structure, is more flexible than methods grounded in null hypothesis testing (Hamaker et al., 2011). Moreover, the IC model selection approach does not mandate the competing models to relate with one another in any particular manner, and the competing candidate models do not need to be nested. Also, the application of IC can accommodate the simultaneous comparison of any number of models, that is, the comparison of models is not limited to two at a time, nor is it mandatory to test models incrementally (like Model 1 versus Model 2, Model 2 versus Model 3, etc.). Furthermore, the designation of any model as a null model is not required for model comparison because the IC method is not based on a null hypothesis testing framework (Christensen, 2018).

Some of the frequently used IC in the literature as described in Equations (3.28), (3.31), (3.32), and (3.33) are the Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan-Quinn information criterion (HQIC), and Shibata information criterion (SIC), respectively. An information criterion is a function of the value of log-likelihood which represents the goodness of fit component, and the number of parameters in the model that represents the model complexity component. In general, the larger (smaller) the value of the log-likelihood (information criterion), the better the fit of the model.

3.9.1 Akaike Information Criterion

Akaike (1974) introduced the first information criterion called the Akaike information criterion (AIC) as stated in Equation (3.28). The logic of information criteria is rooted in information theory. The AIC is formulated on the Kullback-Leibler (K−L) distance (Kullback and Leibler, 1951), where information theory is linked to the distributions of random variables. Following the concept of the "K−L distance" given in Burnham and Anderson (2003), the discrepancy (or distance) between two models are briefly described as follows:

Let $f(x)$ represents one of the two models, known as the "true" or "data generating" model from which the actual observations exist. The second (approximating) model, represented by $g(x)$, is a model defined in relation to the theory of the researcher, about the true model. Now, using the concept of information theory, the K−L distance stated in Equation (3.27) can be defined as the information lost in the process of approximating $f(x)$ by $g(x)$ (Konishi and Kitagawa, 1996). If $f(x)$ and $g(x)$ are exactly the same, then the K−L discrepancy between the two models is zero; otherwise, the discrepancy is larger than zero (Christensen, 2018). Hence, model selection can be approached by making efforts to reduce the K−L information (Konishi and Kitagawa, 1996; Kullback and Leibler, 1951).

The K−L distance is stated as:

$$I\{g(x); f(x|\hat{\theta})\} = \int g(x)\log g(x)dx - \int g(x)\log f(x|\hat{\theta})dx, \tag{3.27}$$

which measures the discrepancy or divergence of $g(x)$ in relation to $f(x|\hat{\theta})$, where $\hat{\theta}$ is the parameter estimate of a p-dimensional vector of unknown parameters.

The AIC can be used to estimate the relative quality of statistical models from a given set of models with a collection of data. Thus, the AIC is a model selection estimator for estimating individual model's quality relative to each of the other models and is expressed as:

$$\text{AIC} = \frac{-2l}{N} + \frac{2p}{N}, \tag{3.28}$$

where l is the log-likelihood of the maximum likelihood $L(\boldsymbol{\theta})$, N is the sample size, and $\boldsymbol{\theta}$ denotes the unknown parameter vector. The $2p$ is used to penalise overparameterisation or overfitting, where p is the number of parameters to be estimated (Li et al., 2009), and it is multiplied by 2 (i.e., $2p$) to simplify the algebra (Dufour, 2007). The selected model is the one with the lowest value of the AIC.

In statistics, when a statistical model is used to represent a data generating process, such representation may never be exact because some of the information may be lost. The AIC can be used to estimate the quantity of information lost by a given model such that the higher (lower) the information lost by the model, the lower (higher) the quality of the model. The AIC is therefore used to reward goodness of fit (measured by the likelihood function) and it is also used to penalise overfitting due to overparameterisation in the model. An increase in the number of parameters of a model can increase its goodness of fit, but a penalty imposed can act as a check to prevent overparameterisation in the model.

One of the desirable feature of information criterion is efficiency, and the AIC is known to be efficient, that is, it asymptotically minimises a loss function (Chaurasia and Harel, 2013; Claeskens and Hjort, 2008). Consistency (as described in Equation 3.29) in large samples is another desirable feature of information criterion. The AIC however has been criticised for not being asymptotically consistent because, as the sample size increases, it does not select the most parsimonious true model with probability converging to one (Shibata, 1976). In other words, the AIC has a high probability of selecting a model with "too many parameters" as the sample size $N \to \infty$ (Bozdogan, 1987; Christensen, 2018; Dufour, 2007). Moreover, it is believed that this inconsistency happens because the calculation of the AIC does not directly involve sample size (Forster, 2004; Janssen and De Boeck, 1999; Li et al., 2009).

Consistency is described as:

$$Pr[I_c(M_{SC}) = I_c(M_T)] \to 1 \ \ (\text{as } N \to \infty), \tag{3.29}$$

where Pr is the probability, M_{SC} is a selected candidate model, I_c denotes the information criterion, and M_T represents the true model.

The inconsistency of the AIC is further described by Shono (2005) with the use of the relative weight of the second term on the right-hand side in Equation (3.30) (i.e., the penalty term component $2p$ of the AIC). The author revealed that this component becomes very small when compared to that of $-2l$, as the sample size is increasingly large. Based on this, the AIC is asymptotically equivalent to the maximum log-likelihood as:

$$\text{AIC} = -2l + 2p \approx -2l \ \ (\text{as } N \to \infty). \tag{3.30}$$

Therefore, the AIC tends to select the complicated model. Hence, the property of "consistency" is not satisfied by the AIC (Shibata, 1989) except in special situations where the ratio p (i.e., the number of unknown parameters)/N (sample size) is kept fixed (Shono, 2005).

3.9.2 Bayesian Information Criterion

Schwarz (1978) introduced the Bayesian information criterion (BIC) as an asymptotic approximation to a transformation of a model's Bayesian posterior probability. It can be calculated as stated in Equation (3.31) based on the empirical log-likelihood and does not need the specification of the priors (Chaurasia and Harel, 2013). The BIC shows consistency as the sample size increases, such that the criterion will select a true model of finite dimension if it is included among the candidate models (Christensen, 2018). In other words, the BIC chooses the true model with probability one (see Equation 3.29), with the assumption that the true model is among the set of selected candidate models (see Nishii, 1984; Shibata, 1981).

Using consistency property, the BIC discourages overparameterisation where it imposes more heavy penalties than the AIC on model complexity (Li et al., 2009). Hence, the BIC potentially favours highly parsimonious models which are neither overly simple nor overly complex but usually lie between the two (Breheny, 2013). It is given by

$$\text{BIC} = \frac{-2l}{N} + \frac{p \log_e(N)}{N}, \tag{3.31}$$

where $\log_e$ is the natural logarithm. For consistency, a criterion is expected to select the right model regardless of the size of the sample, which shows that such a criterion should select the same model even if the sample size increases (Buckland et al., 1997). This can be assured by increasing the penalty term as the sample size increases (Pollmann, 2015).

3.9.3 Hannan-Quinn Information Criterion

The Hannan-Quinn information criterion (HQIC) as stated in Equation 3.32 was introduced by Hannan and Quinn (1979) and Hannan (1980), and it is a model selection criterion that is formulated on the law of iterated logarithm (Javed and Mantalos, 2013). Like the BIC, the HQIC also satisfies consistency property as the sample size $N \to \infty$ (Shono, 2005).

$$\text{HQIC} = \frac{-2l}{N} + \frac{2p\log_e(\log_e(N))}{N}. \tag{3.32}$$

3.9.4 Shibata Information Criterion

The Shibata information criterion (SIC) as stated in Equation (3.33) is least frequently applied for model selection in comparison with the AIC, BIC and HQIC. However, it follows the same structure as the BIC and HQIC, where the complexity component includes the sample size N (see Ghalanos, 2024). It is stated as:

$$\text{SIC} = \frac{-2l}{N} + \log_e\left(\frac{N+2p}{N}\right) \tag{3.33}$$

The coefficients for p in the model complexity component of the four information criteria formulae depict the degree of penalty imposed on the number of model parameters. Although efficient and consistent information criteria represent different approaches to the selection of a model, they are frequently used jointly in applied work and are often presented together with each other in software outputs (Christensen, 2018). The AIC and BIC are frequently used in the literature (Chaurasia and Harel, 2013), while the HQIC and in particular the SIC (see Ghalanos, 2024) are less frequently used in practice. For model selection, the model with the lowest value of the applied information criteria is chosen as the "best" model (Christensen, 2018).

3.9.5 Model Checking through Diagnostics

To assess the quality of a fitted model, an appropriate diagnostic model checking plots can be used. Commonly used diagnostic plots include density plots and quantile plots. If the fitted model is a reasonable model for modelling the data, then the quantile plots, for instance, should display points that are approximately linear on the line $y = x$, i.e., points should lie close to the unit diagonal (Mukhodobwane et al., 2022). Substantial departures from linearity will imply a failure in the validity of the model for modelling the data (Coles, 2001).

3.10 Long Memory and Short Memory Processes

The long memory behaviour of a time series describes the correlation pattern of that series at distant lags. A series with long memory usually shows persistent temporal dependence among distant observations (Barkoulas et al., 1999). Long memory is used to describe the high-order correlation structure of a time series. A long memory time series autocorrelation function (ACF) decays hyperbolically. Such series exhibit low-frequency spectral distributions. On the other hand, the low-order correlation structure of a series is used to characterise short memory. The existence of long memory implies that the market gradually responds to information over

a long period of time. Shocks to volatility usually have long-running effects in a long-memory process, and such persistence is a vital component of derivative pricing, investment portfolios, and risk management (Brania and Gurgul, 2015). The idea of long memory was introduced by Hurst (1951), and other earlier contributions to the study of time series long memory include Mandelbrot (1971) and Mandelbrot and Van Ness (1968). These contributors formalised the empirical findings of Hurst using the cumulative river flow data (see Brania and Gurgul, 2015; Geweke and Porter-Hudak, 1983; Hosking, 1981).

In both empirical and theoretical studies, finance researchers have focused on long memory (persistence) in asset returns. Stock market researchers do not only investigate the existence of long memory in the returns but also in the return volatility using squared returns or absolute values of returns (Brania and Gurgul, 2015). The presence of long memory can be determined by measuring the fractional order of integration d of a time series (Saha et al., 2020). In other words, the differencing parameter d of fractionally differenced models characterises the long-memory identity of a time series. This parameter d can be related to the Hurst exponent (H) as $d = H - 1/2$. The Hurst exponent is a popular measure of long-memory, and it can be calculated by various methodologies that include the classical-rescaled range (R/S) analysis introduced by Hurst (1951). Other methodologies for calculating H include the modified R/S analysis and the Rescaled Variance (V/S) analysis. When $0 < H < 0.5$, the autocovariances are negative at all lags, and the time series process is termed anti-persistent. On the other hand, when $0.5 < H < 1$, the autocovariances are positive at all lags, and the process is termed persistent (see Saha et al., 2020).

The applications of the fractionally differenced models have been extensively used in finance, hydrology, econometrics, economics, telecommunication, and geophysics, among others (see Beran, 1994; Beran et al., 2013; Holan et al., 2009; Palma, 2006; Vera-Valdés, 2021). This brought to light the ARFIMA-FIGARCH models, which were developed for long memory modelling in the first and second moments (Baillie et al., 1996; Granger and Joyeux, 1980; Hosking, 1981; Rafik et al., 2014). The ARFIMA model is used for modelling long memory in the return time series (first moment), while the FIGARCH model is used to model long memory in volatility (Saha et al., 2020).

A covariance stationary stochastic process will exhibit long memory with memory parameter d under the condition that its spectral density function $f(\lambda)$ satisfies:

$$f(\lambda) \sim C\lambda^{-2d} \text{ as } \lambda \to 0^{+}, \tag{3.34}$$

where "$\sim$" indicates that the ratio of the right and left-hand sides tends to one at the limit, and C is a finite positive constant. As the process satisfies the necessary and sufficient condition in Equation (3.34) and $d > 0$, its autocorrelation function decays at a hyperbolic rate (see Beran, 1994; Brania and Gurgul, 2015; Granger and Joyeux, 1980; Hosking, 1981), i.e.,

$$\rho_{\kappa} \sim C_{\rho}\kappa^{2d-1} \text{ as } \kappa \to \infty, \tag{3.35}$$

where C_ρ is a constant. The nature of the process memory is determined by parameter d. If $d = 0$, the spectral density is bounded at zero, and the process is referred to as short memory. However, the spectral density is unbounded near the origin if $d > 0$, hence, the process shows long memory. Lastly, the process is termed antipersistent and shows negative memory when $d < 0$ because the spectral density is zero at the origin (Brania and Gurgul, 2015).

The spectral density function can also be defined in relation to the Hurst exponent (H) as:

$$f(\lambda) \sim C|\omega|^{2H-1}. \tag{3.36}$$

Furthermore, using an auto-covariance function $\gamma_y(\kappa)$, Saha et al. (2020) defined the long memory, short memory and anti-persistence processes for a second-order stationary process y_t as:

a. Long memory, if

$$\sum_{\kappa=1}^{\infty} \gamma_y(\kappa) \to \infty \tag{3.37}$$

b. Short Memory, if

$$0 < \sum_{\kappa=1}^{\infty} \gamma_y(\kappa) < \infty \tag{3.38}$$

c. Anti-persistence, if

$$\sum_{\kappa=1}^{\infty} \gamma_y(\kappa) = 0 \tag{3.39}$$

Hence, the long memory (or persistence) infers that a negative or positive movement is more likely to be followed by another movement in the same direction. This is contrary to an anti-persistent process, where a positive movement is more likely to be followed by another movement in the opposite direction (Saha et al., 2020). Also, the short-term dependence, or short memory, attribute of a time series can be used to describe the high frequency or short-run behaviour of that time series. Correlations among observations at distant lags become negligible for a short-memory time series (Barkoulas et al., 1999).

As stated in Vera-Valdés (2024), long-memory exists in a time series if:

$$\gamma_y(\kappa) \approx C_x \kappa^{2d-1} \text{ as } \kappa \to \infty, \tag{3.40}$$

where C_x is a constant. It can also be stated as:

$$f_y(\lambda) \approx C_x \lambda^{-2d} \text{ as } \lambda \to \infty, \tag{3.41}$$

where C_x is a function and $f_y(\lambda)$ denotes the spectral density function. The existence of long memory can be described from a data-oriented approach concerning the persistence of observed autocorrelation. The degree of that persistence follows a stationary process, but where the autocorrelations decay much longer than the ARMA process that fades out at an exponential rate (Baillie, 1996).

Generally, estimation of the fractional differencing parameter d to determine the existence of long memory can be carried out through several methods (see Beran, 1994). These methods include the parametric approach (e.g., through the FIGARCH model), and semi-parametric approach (e.g., the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator). Other methods used to estimate H (where $d = H - 1/2$) include the graphical techniques (like the aggregated variance method, the classic rescaled adjusted range analysis, i.e., analysis based on R/S statistic, etc.). Compared to the parametric and semi-parametric approaches, the graphical methods are sensitive to short-range autocorrelation and are generally inaccurate in their estimation of d. However, they are useful to heuristically determine the presence of long-range dependence in datasets, and can also be used to obtain a first estimate of d (Wang et al., 2007). To avoid the drawback of the graphical methods, this study applies the parametric approach through the FIGARCH model, and the semi-parametric methods via the GPH, LW, and ELW estimators to obtain the estimate of the fractional integration parameter d that reveals the existence of long-memory in the series. The semi-parametric approach is particularly attractive to users because it allows fractional differencing parameter d to be estimated without specifying the entire time series model (Barkoulas et al., 1999).

3.11 The ARIMA and ARFIMA Processes

Time series that exhibit long memory are referred to as fractionally integrated series, or $I(d)$, where d is a non-integer in the interval $-\frac{1}{2} < d < 1$, excluding 0 (Baum et al., 2020). Hence, the ARFIMA(p, d, q) model class is developed with the introduction of non-integer fractional integration parameter d. A long memory (or fractionally integrated) series is neither a nonstationary or unit root ($I[1]$) nor a stationary ($I[0]$) process; it is an $I(d)$ process, where d is a real number (Baum and Wiggins, 2024). One of the two approaches to estimating d involves the classical time-series approach that requires the full specification of the ARFIMA(p, d, q) model, where the parameters are estimated by maximum likelihood (Sowell, 1992). It has been shown by Dahlhaus (1989) and Fox and Taqqu (1986) that the maximum likelihood (ML) estimates of the ARFIMA(p, d, q) model are asymptotically unbiased. Dahlhaus (1989) has also shown that the ML estimator of d in the general ARFIMA(p, d, q) Gaussian processes is strongly consistent, asymptotically efficient, and asymptotically normally distributed in the Fisher sense (Lopes et al., 2002). The second approach is a move from the time domain (of the

first approach) to the frequency domain, where a consistent and asymptotically normal estimate of d can be obtained without the full specification of the ARMA components of the model. The estimators of d in this second approach are mainly regression-based (Baum et al., 2020).

3.11.1 ARFIMA Model

Following Granger and Joyeux (1980) and Hosking (1981), the ARFIMA(p, d, q) process can be stated as:

$$\phi(L)(1-L)^d(y_t-\mu)=\theta(L)\varepsilon_t, \tag{3.42}$$

where y_t is a time-series process for $t = 1, \ldots, T$, L is the backward shift operator, d is the fractional integration (or long memory) parameter, $\phi(L) = 1 + \phi_1 L + \cdots + \phi_p L^p$ is the autoregressive (AR) polynomial, $\theta(L) = 1 + \theta_1 L + \cdots + \theta_p L^p$ is the moving average (MA) polynomial, and $L^i y_t = y_{(t-i)}$ (Boubaker et al., 2021; Cheung, 2008; Wang et al., 2007). The AR and MA coefficients are of orders p and q, respectively, and the roots of $\phi(L)$ and $\theta(L)$ are strictly outside the unit circle to ensure stationarity and invertibility, respectively. The ε_t is white noise (Baillie, 1996), and μ is the mean of the process.
The fractional integrating operator $(1-L)^d$ for non-integer values of d is defined by the binomial expansion as:

$$(1-L)^d=\sum_{i=0}^{+\infty}\begin{bmatrix}d\\i\end{bmatrix}(-L)^i=1-dL-\frac{1}{2}d(1-d)L^2-\frac{1}{6}d(1-d)(2-d)L^3-\ldots, \tag{3.43}$$

or can be re-defined as:

$$(1-L)^d=\sum_{i=0}^{+\infty}\frac{\Gamma(i-d)L^i}{\Gamma(-d)\Gamma(i+1)}, \tag{3.44}$$

where $\Gamma(\cdot)$ is the gamma function. The fractional integration parameter d is used to describe the long-memory behavior of the process, while the $\phi(L)$ and $\theta(L)$ parameters make up the short-memory parameters and they affect only the short-run dynamics of the process.

The ARFIMA model is a general form of the standard linear ARIMA(p, d, q) processes of Box and Jenkins because it allows the degree of integration d to assume non-integer values. This generalisation is a more flexible way of studying time series data, and it enables researchers to simultaneously account for long and short-term dynamics (Fouquau and Spieser, 2014). The standard ARIMA model is obtained when d is arbitrarily restricted to integer values (Barkoulas et al., 1999). The ARFIMA class of fractional processes can be used to model data dependence that is weaker than allowed by unit root processes, but stronger than implied by stationary ARMA processes. The ARFIMA model is a parametric long-memory time series process (Bollerslev and Wright, 2000).

The ARFIMA(p, d, q) process is covariance stationary for $-0.5 < d < 0.5$, while $d < 1$ indicates mean reversion. The simplest possible model is the fractional white noise model, i.e., the ARFIMA$(0, d, 0)$, where $p = 0 = q$ (Baillie, 1996). When $d = 1$, the process reduces to ARIMA, which implies infinite memory (random walk) (Boubaker et al., 2021; Saha et al., 2020); and when $d = 0$, the process becomes the standard ARMA. If $-0.5 < d < 0$, the process shows anti-persistence. However, the process becomes stationary long memory when $0 < d < 0.5$, and it possesses shocks that decay hyperbolically; here, the autocorrelations decay to zero and will not be summable (Rafik et al., 2014). Moreover, if $0.5 \leq d < 1$, the process is a mean-reverting non-stationary, with finite impulse response weights. Hence, a broad range of low frequency behaviour can be studied and modelled when d is not confined to the integer domain (Cheung, 2008). With the assumption that $-0.5 < d < 0.5$, and $d \neq 0$, it was shown by Hosking (1981) that as $\kappa \to \infty$, the correlation function, $\rho(\cdot)$, of an ARFIMA process is κ^{2d-1}. Hence, as opposed to the faster geometric display of a stationary ARMA process, the autocorrelations of the ARFIMA process hyperbolically decay to zero as $\kappa \to \infty$. The ARFIMA model summarily comprises a short memory ARMA part (that describes the short-term behaviour of a series), and a fractional differencing of the right order that accounts for any long-term persistent property in the series (Barkoulas et al., 1999).

3.12 The FIGARCH Model

The ARFIMA-FIGARCH models are used for modelling persistence in a time series mean and volatility (Rafik et al., 2014). As a motivation from the developments of ARFIMA model types for long memory estimation, Baillie et al. (1996) introduced the FIGARCH model to capture long memory in the volatility of time series financial returns (Bobeica and Bojesteanu, 2008; Ghalanos, 2022; Saha et al., 2020). The parametric formulation of the FIGARCH model was built on the ARFIMA processes (Adenstedt, 1974; Bollerslev and Wright, 2000; Granger and Joyeux, 1980; Hosking, 1981; McLeod and Hipel). As opposed to the IGARCH model where shocks persist forever, and the standard GARCH where shocks decay at an exponential rate, the effect of shocks for the FIGARCH model is in between these two models, where shocks decay at a slower hyperbolic rate (Baillie, 1996; Ghalanos, 2022; Rafik et al., 2014).
Given the standard GARCH equation:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \tag{3.45}$$

where L is the backward shift lag operator, for which $\alpha(L) = \sum_{i=1}^{q} \alpha_i(L)^i$ and $\beta(L) = \sum_{j=1}^{p} \beta_j(L)^j$. This is re-arranged to give the ARMA in squares representation:

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)]\nu_t, \tag{3.46}$$

where $\nu_t = \varepsilon_t^2 - \sigma_t^2$ and the left-hand side is condensed as:

$$(1-L)\phi(L)\varepsilon_t^2 = \omega + [1-\beta(L)]\nu_t \tag{3.47}$$

with $\phi(L) = \sum_{i=1}^{m-1}\phi_i(L)^i$ and $m = \max\{p,q\}$. In the fractionally integrated model, $(1-L)$ is substituted by

$$(1-L)^d = \sum_{\kappa=0}^{\infty}\frac{\Gamma(d+1)}{\Gamma(\kappa+1)\Gamma(d-\kappa+1)}L^\kappa = 1-\sum_{\kappa=1}^{\infty}\pi_\kappa L^\kappa, \tag{3.48}$$

where $\pi_i = \prod_{1\leq\kappa\leq i}\frac{\kappa-1-d}{\kappa}$. The hypergeometric function expansion is normally truncated to some large number, like, 1000. The following representation of the FIGARCH model is obtained after rearranging again as:

$$\begin{aligned}\sigma_t^2 &= \omega[1-\beta(L)]^{-1} + \left\{1-[1-\beta(L)^{-1}\phi(L)(1-L)^d]\right\}\varepsilon_t^2 \\ &= \omega^* + \lambda(L)\varepsilon_t^2 \\ &= \omega^* + \sum_{j=1}^{\infty}\lambda_i L^i\varepsilon_t^2,\end{aligned} \tag{3.49}$$

where $\lambda_1 = \phi_1 - \beta_1 + d$ and $\lambda_\kappa = \beta_1\lambda_{\kappa-1} + (\frac{\kappa-1-d}{\kappa} - \phi_1)\pi_{d,\kappa-1}$. Sufficient conditions to ensure the positivity of the conditional variance for the FIGARCH$(1,d,1)$ model are $\omega > 0, \beta_1 - d \leq \phi_1 \leq \left(\frac{2-d}{2}\right)$ and $d\left(\phi_1 - \frac{1-d}{2} \leq \beta_1(\phi_1 - \beta_1 + d)\right)$. Equation (3.49) is re-written as follows, by setting $\phi(L) \equiv (1-\alpha(L))$:

$$\begin{aligned}\phi(L)(1-L)^d\varepsilon_t^2 &= \omega + (1-\beta(L)) + (\varepsilon_t^2 - \sigma_t^2) \\ \phi(L)(1-L)^d\varepsilon_t^2 &= \omega - \sigma_t^2 + \varepsilon_t^2 + \beta(L)\sigma_t^2 - \beta(L)\varepsilon_t^2 \\ \sigma_t^2 &= \omega + \varepsilon_t^2 + \beta(L)\sigma_t^2 - \beta(L)\varepsilon_t^2 - \phi(L)(1-L)^d\varepsilon_t^2 \\ \sigma_t^2 &= \omega + \left\{1-\beta(L)-\phi(L)(1-L)^d\right\}\varepsilon_t^2 + \beta(L)\sigma_t^2 \\ \sigma_t^2 &= \omega + \left\{1-\beta(L)-(1-\alpha(L))(1-L)^d\right\}\varepsilon_t^2 + \beta(L)\sigma_t^2.\end{aligned} \tag{3.50}$$

The expansion is truncated to 1000 lags and set as $(1-L)^d\varepsilon_t^2 = \varepsilon_t^2 + \left(\sum_{\kappa=1}^{1000}\pi_\kappa L^\kappa\right)\varepsilon_t^2 = \varepsilon_t^2 + \overline{\varepsilon}_t^2$, Equation (3.50) can be re-written as:

$$\begin{aligned}\sigma_t^2 &= \omega + \left\{\varepsilon_t^2 - \beta(L)\varepsilon_t^2 - (1-L)^d\varepsilon_t^2 + \alpha(L)(1-L)^d\varepsilon_t^2\right\} + \beta(L)\sigma_t^2 \\ \sigma_t^2 &= \omega + \left\{\varepsilon_t^2 - \beta(L)\varepsilon_t^2 - (\varepsilon_t^2 + \overline{\varepsilon}_t^2) + \alpha(L)(\varepsilon_t^2 + \overline{\varepsilon}_t^2)\right\} + \beta(L)\sigma_t^2 \\ \sigma_t^2 &= \omega + \varepsilon_t^2 - \beta(L)\varepsilon_t^2 - (\varepsilon_t^2 + \overline{\varepsilon}_t^2) + \alpha(L)(\varepsilon_t^2 + \overline{\varepsilon}_t^2) + \beta(L)\sigma_t^2 \\ \sigma_t^2 &= \omega - \overline{\varepsilon}_t^2 - \beta(L)\varepsilon_t^2 + \alpha(L)(\varepsilon_t^2 + \overline{\varepsilon}_t^2) + \beta(L)\sigma_t^2 \\ \sigma_t^2 &= (\omega - \overline{\varepsilon}_t^2) - \sum_{j=1}^{q}\beta_j\varepsilon_{t-j}^2 + \sum_{j=1}^{p}\alpha_j\varepsilon_{t-j}^2 + \sum_{j=1}^{p}\alpha_j\overline{\varepsilon}_{t-j}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2 \\ \sigma_t^2 &= (\omega - \overline{\varepsilon}_t^2) + \sum_{j=1}^{p}\alpha_j(\varepsilon_{t-j}^2 + \overline{\varepsilon}_{t-j}^2) + \sum_{j=1}^{q}\beta_j(\sigma_{t-j}^2 - \varepsilon_{t-j}^2).\end{aligned} \tag{3.51}$$

When $d = 1$, the FIGARCH is reduced to the IGARCH model, and when $d = 0$, to the standard GARCH model. Conrad and Haag (2006) provided a general set of sufficient conditions for the FIGARCH(p, d, q) process (Ghalanos, 2022). The long memory operator in the FIGARCH model is applied to the squared errors, i.e., the memory parameter acts on the squared errors. This contrasts the ARFIMA model that applies the long memory operator to the unconditional mean of y_t (Ghalanos, 2022; Saha et al., 2020).

3.13 The Log-periodogram Regression

Geweke and Porter-Hudak (1983) (GPH, henceforth) proposed a method for estimating the fractional differencing parameter d in the ARFIMA(p, d, q) model (see Xunyu et al., 2015). Although the long-memory dependence is described entirely by the memory parameter d, the traditional approach to estimating an ARFIMA model requires the specification of the AR and MA polynomials. However, a simpler approach to estimating d without specifying the AR and MA parts was introduced by Geweke and Porter-Hudak through a semiparametric technique (Geweke and Porter-Hudak, 1983). The log-periodogram regression estimate of d is one of the main methods used under the semiparametric approach. The proof of the limiting distribution and consistency of the GPH approach was provided by Robinson (1995a) in the Gaussian case (Bollerslev and Wright, 2000). The GPH estimator is based on the expression of the spectral density function of the ARFIMA process with frequencies around zero and it estimates only the memory parameter d (Boubaker et al., 2021; Fouquau and Spieser, 2014). Through evaluation close to the origin, the semiparametric estimators bypass the need to specify the short-term time series dynamics. Hence, the semiparametric estimators are robust to the dynamics of short run such as observational noise (Vera-Valdés, 2024).

3.14 Geweke and Porter-Hudak Estimator

Geweke and Porter-Hudak (GPH) (Geweke and Porter-Hudak, 1983) introduced a semiparametric estimator of the fractional differencing parameter, d, formulated on a regression of the log spectral density's ordinates on trigonometric function (Baillie, 1996). The estimator utilises the theory of linear filters to state the process $(1 - L)^d y_t = \mu_t$, where $\mu_t \sim I(0)$, as

$$f(\omega)_y = |1 - e^{-i\omega}|^{-2d} f(\omega)_\mu, \tag{3.52}$$

where $f(\omega)_\mu$ and $f(\omega)_y$ are the spectral densities of μ_t and y_t, respectively. Hence, Equation (3.52) can be stated as

$$\begin{aligned}\log\{f(\omega)_y\} &= \{4\sin^2(\omega/2)\}^{-d} + \log\{f(\omega)_\mu\},\\ \log\{f_y(\omega)_j\} &= \log\{f_\mu(0)\} - d\log\{4\sin^2(\omega_j/2)\} + \log[f_\mu(\omega_j)/f_\mu(0)].\end{aligned} \tag{3.53}$$

GPH proposed to estimate d from a regression that is based on Equation (3.53) with the use of spectral ordinates $\omega_1, \omega_2, \ldots, \omega_m$, from the periodogram of y_t, i.e., $I_y(\omega_j)$. Thus, for $j = 1, 2, \ldots, m$, with $m = g(T)$, where $g(T)$ is such that $\lim_{T\to\infty} g(T) = \infty, \lim_{T\to g}\{g(T)/T\} = 0, \lim_{T\to g}\{\log(T)^2/g(T)\} = 0$,

$$\log\{I_y(\omega_j)\} = a + b\log\{4\sin^2(\omega_j/2)\} + \nu_j \tag{3.54}$$

where

$$\nu_j = \log[f_\mu(\omega_j)/f_\mu(0)], \tag{3.55}$$

and on the assumption that ν_j is i.i.d with mean zero and variance $\pi^2/6$. A good estimate of d should be provided by the regression in Equation (3.54) when μ_t is white noise ε_t (Baillie, 1996). When μ_t is autocorrelated, GPH revealed that Equation (3.54) holds approximately for frequencies in the proximity of (or near) zero. If the proximity shrinks at a suitable rate with sample size, then a consistent estimator of d should be realized by the GPH procedure (Baillie, 1996).

The ordinary least squares (OLS) estimator of d in Equation (3.54) will produce the limiting distribution in Equation (3.56), if the number of ordinates m is selected.

$$\frac{(\hat{d}_{GPH} - d)}{\{\text{var}(\hat{d}_{GPH})\}^{1/2}} \Rightarrow N(0, 1). \tag{3.56}$$

The $\text{var}(\hat{d}_{GPH})$ is derived from the common regression formula, either by setting it as $\pi^2/6$ or with the use of the regression residual variance. Geweke and Porter-Hudak (1983) originally suggested the bandwidth choice of m equal $\sqrt{T}$ (Brania and Gurgul, 2015). However, it is obvious from the outcome that the GPH estimator is not $\sqrt{T}$ consistent and convergence will occur at a slower rate (Baillie, 1996). Consistency and asymptotic normality of the process were proven by Geweke and Porter-Hudak (1983) only for $d < 0$, while Robinson (1990) gave a proof of consistency for $0 < d < 0.50$; see also Künsch (1986).

While the GPH estimator is potentially robust to nonnormality and its application is simple, its potential attractiveness is reduced by the behaviour $\hat{d}_{GPH}$ where substantial autocorrelation of μ_t is present. Specifically, Agiakloglou et al. (1992) showed that it gives "serious bias" (Baillie, 1996; Boubaker et al., 2021).

Robinson (1995a) proved the asymptotical normality of the GPH estimator for $d \in (-\frac{1}{2}, \frac{1}{2})$. However, Velasco (1999) and Kim and Philips (1999) later showed its consistency for $d \in (-\frac{1}{2}, 1)$ and that it has an asymptotically normal limit distribution for $d \in (-\frac{1}{2}, \frac{3}{4})$:

$$\hat{d}_{GPH} \sim N\left(d, \frac{\pi^2}{24m}\right). \tag{3.57}$$

The Geweke and Porter-Hudak (1983) estimator is theoretically valid for $0 < d < \frac{1}{2}$. If the number of frequencies m[1] that is included in the regression is restricted in a way that $m = O(T^{4/5})$, then the asymptotic normality is obtained (see Boubaker et al., 2021; Hurvich et al., 1998).

3.15 Whittle Estimation

Estimation of the long memory parameter d of a time series with fractional integration in the frequency domain based on the maximum likelihood method is built on the approximation to a Gaussian likelihood proposed by Whittle (1951). See also Whittle (1962), Fox and Taqqu (1986), and Choudhuri et al. (2004). Unlike the time domain estimator, the frequency domain approach has gained much popularity due to the fact that its maximum-likelihood estimator is invariant to the unknown mean of the process (Baum et al., 2020; Cheung and Diebold, 1994).

3.15.1 The Whittle Likelihood

Given a sample with T observations that consist of a stationary centered process $y_1, \ldots, y_T$ that are spaced uniformly in the time domain and a sequence of m frequencies (see Baum et al., 2020),

$$\omega_j = \frac{2\pi j}{T} \quad \text{for } j = 1, 2, \ldots, m \tag{3.58}$$

These frequencies, with $m \ll T$, denote a set of angular frequencies, where each is a multiple of the fundamental frequency $2\pi/T$, since it relates to a single oscillation with period T. Equation (3.59) states the discrete Fourier transform of y_t as:

$$\hat{C}(\omega_j) = \frac{1}{\sqrt{2\pi T}} \sum_{\kappa=1}^{T} y_\kappa e^{i\omega_j \kappa}. \tag{3.59}$$

The Whittle likelihood is based on the fact that the coefficients $\hat{C}(\omega_j)$ of y_t are asymptotically independent Gaussian random variables having mean value zero and the variance is given by the spectral density of the process at that frequency. As a consequence, the likelihood function at frequency ω_j is

$$L_j = \frac{1}{\sqrt{2\pi f_y(\omega_j)}} \exp\left\{-\frac{I(\omega_j)}{2f_y(\omega_j)}\right\}, \tag{3.60}$$

[1] m is the bandwidth parameter

where $I(\omega_j)$ is the sample periodogram stated as:

$$I(\omega_j) = |C(\omega_j)|^2, \tag{3.61}$$

and $f_y(\omega_j)$ refers to the spectral density at ω_j:

$$f_y(\omega_j) = \frac{1}{2\pi} \sum_{\kappa=-\infty}^{\infty} \gamma_\kappa e^{-i\omega_j \kappa}, \tag{3.62}$$

where γ_κ represents the autocovariance at lag κ.

Estimation of the fractional differencing parameter d through the local Whittle begins when it is known that the behaviour of the spectral density of y_t at low frequencies can be defined through the condition

$$\lim_{\omega \to 0^+} \omega^{2d} f_y(\omega) = G, \tag{3.63}$$

where G is a positive quantity, whose dependence is on the parameter d. The process y_t possesses finite power when $2d < 1$, and based on that, d is used to measure the long-term duration of the memory of process y_t. The series is covariance stationary if $d \in [0, \frac{1}{2})$, but the autocorrelations decay more slowly than in the $I(0)$ state. They fade away hyperbolically to zero as opposed to the stationary ARMA process which decays faster at a geometric rate. The process is mean reverting for $d \in [\frac{1}{2}, 1)$, but it is not covariance stationary because there is no long-term effect of an innovation on future values of the process. It is shown by Granger and Joyeux (1980) that a process is nonstationary for $d \geq \frac{1}{2}$ because it has infinite variance. The process exhibits antipersistence (intermediate memory), or long-range negative dependence for $d \in (-\frac{1}{2}, 0]$. When $f_y(\omega_j)$ is substituted by its asymptotic approximation $G\omega_j^{-2d}$ from Equation (3.63), the negative of the log-likelihood at ω_j will satisfy

$$-\log L_j(G, d) = \frac{1}{2}\left\{\log 2\pi + \log G - 2d \log \omega_j + \frac{1}{G}\omega_j^{2d}|\hat{C}(\omega_j)|^2\right\}. \tag{3.64}$$

The local Whittle estimators of d are found based on this expression (Baum et al., 2020).

3.15.2 The Local Whittle Estimator

The analysis of the local Whittle estimator of d by Robinson (1995a), motivated by the findings of Künsch (1987) and based on the m lowest frequencies $\omega_1, \ldots, \omega_m$, can be obtained through the minimisation of the negative log-likelihood function

$$-\log L(G, d) = \frac{m}{2}\left\{\log 2\pi + \log G - \frac{2d}{m}\sum_{j=1}^{m} \log \omega_j + \frac{1}{G}\frac{1}{m}\sum_{j=1}^{m} \omega^{2d}|\hat{C}(\omega_j)|^2\right\}. \tag{3.65}$$

If Equation (3.65) is partially differentiated with respect to G, it can be shown that the optimal value of G for any value of d is[1]

$$\hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m} \omega_j^{2d} |\hat{C}(\omega_j)|^2. \tag{3.66}$$

Based on this outcome, it is obvious that the local Whittle estimator of d, $\hat{d}$, minimises

$$R(d) = \log\hat{G}(d) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j, \quad \hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d} I(\omega_j). \tag{3.67}$$

Robinson (1995b) has shown that the local Whittle estimator $\hat{d}$ obtained through the minimisation of Equation (3.60) is asymptotically normally distributed for $d \in (\frac{1}{2}, \frac{3}{4})$ and consistent if $d \in (\frac{1}{2}, 1)$ so that

$$\sqrt{m}(\hat{d} - d) \xrightarrow{d} N\left(0, \frac{1}{4}\right). \tag{3.68}$$

3.15.3 The Exact Local Whittle Estimator

An improvement to the local Whittle method was introduced by Shimotsu and Phillips (2005) (see Baum et al., 2020; Vera-Valdés, 2024). The authors defined the exact local Whittle (ELW) estimator as the minimiser of the function

$$R(d) = \log\hat{G}(d) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j, \quad \hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d} I_{\Delta_y^d}(\omega_j), \tag{3.69}$$

where $I_{\Delta_y^d}(\omega_\kappa)$ denotes the periodogram of the fractionally differenced series $\Delta_{y_t}^d$ and the difference operator is now described by the binomial expansion

$$\begin{aligned}\Delta_{y_t}^d = y_t - \frac{d}{1!}y_{t-1} + \frac{d(d-1)}{2!}y_{t-2} - \frac{d(d-1)(d-2)}{3!}y_{t-3} \\ + \frac{d(d-1)(d-2)(d-3)}{4!}y_{t-4} \quad \cdots\end{aligned}$$

The presence of $I_{\Delta_y^d}(\omega_j)$ in the optimisable function demands that a fractional difference of y_t be calculated for every alteration of d. Hence, the ELW estimation is computationally more demanding than the estimation of the simple local Whittle. The consistency and asymptotic normality of this estimator are the same as those of the local Whittle estimator (Baum et al., 2020; Vera-Valdés, 2024).

[1] Different authors may have varying expressions of the Fourier coefficients in Equation (3.59) by a constant factor. However, the value of d (the fractional differencing parameter) is independent of this multiplier, but the choice will affect the value of $\hat{G}$ (Baum et al., 2020).

3.16 The Beta-Skew-t-EGARCH Model

To begin with, the Beta-t-EGARCH as a model without skewness was proposed by Harvey and Chakravarty (2008), and Harvey (2013). It is an extension made to the EGARCH model where an equation that depends on the conditional score of the last observation drives the variance or scale (Creal et al., 2011, 2008). The Beta-t-EGARCH model can be regarded as an unrestricted version of the GAS model of Creal et al. (2013) (see Sucarrat, 2013). In other words, it is a dynamic model driven by a conditional score, which is a martingale difference. A martingale difference has a constant (or zero) conditional expectation (Harvey, 2013). Due to various implications of conditional skewness on asset pricing, conditional score models were extended to skewed distributions (Harvey and Sucarrat, 2014). Hence, the Beta-Skew-t-EGARCH model is a skewed version of the Beta-t-EGARCH model, where skewness can be brought in through the method introduced by Fernández and Steel (Fernández and Steel, 1998) (see Harvey and Sucarrat, 2014; Sucarrat, 2013).

A number of useful and attractive properties are attributed to this skewed model. As in the Beta-t-EGARCH model where observations that could be seen as outliers for a Normal distribution are down-weighted with the use of the conditional score (Harvey and Sucarrat, 2014), the Beta-Skew-t-EGARCH model also displays robustness to outliers or jumps. In particular, it performs quite well in modelling key stylised facts of financial returns like fat-tails, leverage effect, conditional skewness and the behaviour of volatility by decompositing into long-term and short-term components. The model has two versions, described below as "one-component and two-component models".

3.16.1 The One-Component Beta-Skew-t-EGARCH Model

The three equations below illustrate the first order Beta-Skew-t-EGARCH model's martingale difference version (Harvey and Sucarrat, 2014).

$$R_t = \exp(\lambda_t) z_t = \sigma_t z_t, \quad z_t \sim st(0, \sigma_z^2, \nu, \eta), \quad \nu > 2, \quad \eta \in (0, \infty), \tag{3.70}$$

$$\lambda_t = \omega + \lambda_t^{\dagger}, \tag{3.71}$$

$$\lambda_t^{\dagger} = \phi_1 \lambda_{t-1}^{\dagger} + \kappa_1 u_{t-1} + \kappa^* sgn(-R_{t-1})(u_{t-1} + 1), \quad |\phi_1| < 1, \tag{3.72}$$

where R_t[1] is the demeaned return, σ_t represents the volatility or conditional scale, and the conditional error z_t is not standardised i.e., the variance is not one. Equation (3.70) shows that z_t follows the skew Student's t distribution with a mean of zero, scale σ_z^2, degrees of freedom

[1]The demeaned return R_t is the same as the error term ε_t (i.e., the unpredictable part of return) in Equation (3.1), but with z_t distributed as a skew Student's t, and it is not standardised to have unit variance (see Sucarrat, 2013).

ν, and skewness parameter η. The z_t is defined as $z_t = z_t^* - \mu_{z^*}$, where z_t^* is an uncentred[1] skew Student's t variable with ν degrees of freedom, mean μ_{z^*}, and skewness parameter η. The ω denotes the log-scale intercept that is described as the long-term log-volatility, while ϕ_1 represents the persistence parameter (the bigger the persistence, the more the clustering of volatility). The κ_1 is the ARCH parameter that indicates the response to shocks (the bigger the absolute value of the parameter, the greater the response of volatility to shocks), κ^* is the leverage parameter, u_t denotes the conditional score, that is, the derivative of the log-likelihood of R_t at t with respect to λ_t, and *sgn* is the sign function. Details on this model are discussed in Sucarrat (2013).

3.16.2 The Two-Component Beta-Skew-t-EGARCH Model

In order to accommodate the long memory feature of financial return like a two-component GARCH model of Engle and Lee (1999), the two-component Beta-Skew-t-EGARCH model was introduced. The model is used to decompose volatility persistence into long-term (or long-run) and short-term (or short-run) components (Alizadeh et al., 2002). The first order two-component Beta-Skew-t-EGARCH model's martingale difference version (Harvey and Sucarrat, 2014) is given as:

$$
\begin{aligned}
R_t &= \exp(\lambda_t) z_t = \sigma_t z_t, \quad z_t \sim st(0, \sigma_z^2, \nu, \eta), \quad \nu, \eta \in (0, \infty), && (3.73)\\
\lambda_t &= \omega + \lambda_{1,t}^{\dagger} + \lambda_{2,t}^{\dagger}, && (3.74)\\
\lambda_{1,t}^{\dagger} &= \phi_1 \lambda_{1,t-1}^{\dagger} + \kappa_1 u_{t-1}, \quad |\phi_1| < 1, && (3.75)\\
\lambda_{2,t}^{\dagger} &= \phi_2 \lambda_{2,t-1}^{\dagger} + \kappa_2 u_{t-1} + \kappa^* sgn(-R_{t-1})(u_{t-1} + 1), \quad |\phi_2| < 1, \quad \phi_1 \neq \phi_2, && (3.76)
\end{aligned}
$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ represent the time-varying long-run and short-run components of log-volatility, respectively. The ϕ_1 and ϕ_2 are the long-term and short-term persistence parameters and the model is not identifiable if $\phi_1 = \phi_2$. The κ_1 and κ_2 are parameters that indicate the long-run and short-run responses to shocks, respectively. Both components are driven by the conditional score u_t, represented (see Sucarrat, 2013) as:

$$
\begin{aligned}
u_t &= \frac{\partial \text{ln} f_R(R_t)}{\partial \lambda_t} \\
&= \frac{(\nu + 1)[R_t^2 + R_t \mu_{z^*} \exp(\lambda_t)]}{\nu \exp(2\lambda_t) \eta^{2sgn(R_t + \mu_{z^*} \exp(\lambda_t))} + (R_t + \mu_{z^*} \exp(\lambda_t))^2} - 1.
\end{aligned} \tag{3.77}
$$

This can be stated more conveniently for simulation purposes as:

$$
u_t = \frac{(\nu + 1)(z_t^{*2} - \mu_{z^*} z_t^*)}{\nu \eta^{2sgn(z_t^*)} + z_t^{*2}} - 1. \tag{3.78}
$$

[1] That is, the mean is not necessarily equal to zero.

When $\eta = 1$ (i.e., a symmetric conditional distribution) in Equation (3.78), then $\frac{u_t+1}{\nu+1} \sim Beta(1/2, \nu/2)$. That explains the source of the name Beta-t-EGARCH. Leverage effect is present only in the temporary short-run component (Engle and Lee, 1999; Harvey and Sucarrat, 2014; Sucarrat, 2013).

3.16.3 The Skew Student's t Distribution

Let an ordinary (i.e., symmetric and centred) Student's t distributed variable with unit scale be denoted by ϵ^*, and let its density be denoted by $f(\epsilon^*)$. The density of an uncentred skew Student's t variable can be stated through the skewing technique of Fernández and Steel (Fernández and Steel, 1998) as:

$$f(z^*|\eta) = \frac{2}{\eta + \eta^{-1}} f\left(\frac{z^*}{\eta^{sgn(z^*)}}\right). \tag{3.79}$$

The right skewness is produced when $\eta > 1$, and the left skewness is attained when $\eta < 1$, whereas symmetry can be attained when $\eta = 1$ (Harvey and Sucarrat, 2014; Sucarrat, 2013). That is, the right (left) hand tail is heavier when $\eta > 1$ ($\eta < 1$).

3.17 The True Parameter Recovery Measure

Since the focus of MCS studies involves the ability of the estimator to recover the true parameter (see Chalmers, 2019), this study applies the "true parameter recovery (TPR)" measure in Equation (3.80) to compute the level (degree) of recovery of the true parameter through the MCS estimator. The TPR measure is a means of evaluating the performance of the MCS estimates in recovering the true parameter. That is, it is used to determine how much of the true parameter value is recovered by the MCS estimator. The TPR is expressed as

$$\text{TPR} = \left(K - \left[\frac{(\vartheta - \hat{\vartheta})}{\vartheta} \times K\right]\right)\%, \tag{3.80}$$

where $K = 0, 1, 2, \ldots, 100$ is the nominal recovery level, ϑ is the true data-generating parameter and $\hat{\vartheta}$ is its estimator from the simulated (synthetic) data. For instance, a TPR estimated value of 95% or 100% denotes that the MCS estimator recovers the complete 95% or 100% of the true parameter. This complete recovery of the true parameter ϑ can be achieved by the MCS estimator $\hat{\vartheta}$ when $\hat{\vartheta} \approx \vartheta$, where $\vartheta > 0$, such that the TPR outcome equals the given nominal recovery level K (i.e., TPR = $K\%$).

Chapter 4

Simulation Framework to Determine Suitable Innovations for Volatility Persistence Estimation: The GARCH Approach

RICHARD T. A. SAMUEL[1,*], CHARLES CHIMEDZA[1] and CASTON SIGAUKE[2]

[1] School of Statistics and Actuarial Science, University of the Witwatersrand

[2] Department of Mathematical and Computational Sciences, University of Venda

Published in: *Journal of Risk and Financial Management*, 2023, 16(9), 392.

Statement of Contributions of Joint Authorship

Richard T. A. Samuel (Candidate)
Conducted the research, writing and compilation of manuscript;

Charles Chimedza (Supervisor) & Caston Sigauke (Co-Supervisor)
Supervised, edited and coauthors of the manuscript.

This chapter is a copy of the journal paper Samuel et al. (2023) referred to above, and available at https://doi.org/10.3390/jrfm16090392**.**

ABSTRACT

This study rolls out a robust framework relevant for simulation studies through the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model using the rugarch package. The package is thoroughly investigated, and novel findings are identified for improved and effective simulations. The focus of the study is to provide necessary simulation steps to determine appropriate distributions of innovations relevant for estimating the persistence of volatility. The simulation steps involve "background (optional), defining the aim, research questions, method of implementation, and summarised conclusion". The method of implementation is a workflow that includes writing the code, setting the seed, setting the true parameters a priori, data generation process and performance assessment through meta-statistics. These novel, easy-to-understand steps are demonstrated on financial returns using illustrative Monte Carlo simulation with empirical verification using the SA Bond market index data. Among the findings, the study shows that regardless of the arrangement of the seed values, the efficiency and consistency of an estimator generally remain the same as the sample size increases. The study also derived a new and flexible true-parameter-recovery measure which can be used by researchers to determine the level of recovery of the true parameter by the MCS estimator. It is anticipated that the outcomes of this study will be broadly applicable in finance, with intuitive appeal in other areas, for volatility modelling.

4.1 Introduction

A simulation-based experiment is not often included in research because many upcoming researchers do not have an adequate understanding of the nitty-gritty details involved. Although the details involved in simulation modelling are generally inexhaustible, this study, however, unveils a crucial framework relevant to the simulation of financial time series data using the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model for volatility persistence estimation. Volatility persistence describes the effect of a shock to future expectation of the variance process (see Ding and Granger, 1996). The ultimate goal of the study is to familiarise researchers with the concepts of simulation modelling through this model. The framework utilises the robust simulating resources of the GARCH model, through set parameters, to generate data that are analysed, and the estimates from the process are then used by chosen metrics to explain the behaviour of selected statistics of interest.

Monte Carlo simulation (MCS) studies are computer-based experiments that use known probability distributions to create data by pseudo-random sampling. The data may be simulated through a parametric model or via repeated resampling (Morris et al., 2019). MCS applies the concept of imitating a real-life scenario on the computer through a certain model that can hypothetically generate the scenario. By simulating or repeating this process a considerably large number of times, it is possible to obtain outcomes that can enable precise computation of

desired statistics of concern, such as those of the possible assumed error distribution/s that can suitably describe a given stock market. In preparing for a simulation experiment, reasonably ample time is needed to organise a well-written and readable computer code and for simulated data generation. Implementation of a good simulation experiment and reporting the outcomes require adequate planning.

Series of R application software packages such as the rugarch (Ghalanos, 2022), GAS (Ardia et al., 2019), GRETL (Baiocchi and Distaso, 2003), fGARCH (Wuertz et al., 2020), SimDesign (Chalmers and Adkins, 2020), tidyverse (Wickham et al., 2019), to mention but a few, are currently available for simulation studies. This study exemplifies how the GARCH model through the rugarch package can be effectively used to improve volatility modelling through a MCS experiment, with outcomes verified empirically. The simulation steps are designed to be reasonably general with the expectation that any other related packages[1] should be able to replicate the routine. Although there are good books on simulation approaches in general (see Bratley et al., 1987; Kleijnen, 2015), up until now, to the best of our knowledge, there has not been any monograph with a direct step-by-step comprehensive layout on a simulation framework using the GARCH model. Hence, this study rolls out an inclusive simulation design that is summarily required for a robust simulation practice in finance to determine appropriate assumed innovations, relevant for estimating the persistence of volatility using this model, with the knowledge applicable in other fields. Here, the MCS approach is used to obtain the most suitable assumed innovation, through which the volatility persistence is empirically estimated.

Since the rugarch package does not make provision for calculating the coverage probability[2], this study also computes the MCS estimator's recovery levels through the "true parameter recovery (TPR)" measure as a proxy for the coverage. The results show that the MCS estimates considerably recover the true parameters. The raw data used for this study are the daily closing S&P South African sovereign bond index, abbreviated to S&P SA bond index. They are Standard & Poor data for the bond market in US dollars from Datastream (2021) for the period 4 January 2000 to 17 June 2021 with 5598 observations. The rest of the paper is organised as follows: Section 4.2 reviews the theories underpinning two heteroscedastic models, the TPR measure, and the description of the design of the simulation framework. Section 4.3 presents the practical illustration of the simulation framework, with empirical verification, on financial bond return data. Section 4.5 discusses the key findings and Section 4.6 concludes.

[1]The GRETL (Cottrell and Lucchetti, 2023; Baiocchi and Distaso, 2003), GAS (Ardia et al., 2019) and fGARCH (Pfaff, 2016; Wuertz et al., 2020) are also among the freely available and applicable software.

[2]Coverage probability is the probability that a confidence interval of estimates contains or covers the true parameter value (Hilary, 2002).

4.2 Materials and Methods

4.2.1 The GARCH Model

The GARCH model was developed by Bollerslev (1986) as a generalisation of the Autoregressive Conditional Heteroscedasticity (ARCH) model introduced by Engle (1982). It is a classical model that is normally defined by its conditional mean and variance equations for modelling financial returns volatility (Kim et al., 2020). The mean equation is stated as

$$r_t = \mu_t + \varepsilon_t, \tag{4.1}$$

where r_t is the return series, $\varepsilon_t = z_t\sigma_t$ denotes the residual part of the return series that is random and unpredictable, where $z_t \sim N(0,1)$ are the standardised residuals which are independent and identically distributed (i.i.d.) random variables with mean 0 and variance 1 (McNeil and Frey, 2000; Smith, 2003), μ_t is the mean function that is usually stated as an Autoregressive Moving Average (ARMA) process,

$$\mu_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i}, \tag{4.2}$$

where ϕ_i (i = 1, . . ., p) and φ_i (i = 1, . . ., q) are unknown parameters. The variance equation of the GARCH(u, v) model is defined as

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_v\varepsilon_{t-v}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_u\sigma_{t-u}^2, \tag{4.3}$$

where $\omega > 0$ is the intercept (white noise), coefficients $\alpha_i \geq 0$ (i = 1, . . ., v) and $\beta_i \geq 0$ (i = 1, . . ., u), respectively, measure the short-term and long-term effects of ε_t on the conditional variance (Maciel and Ballini, 2017). The non-negativity restrictions on the unknown parameters, α_i and β_i, are imposed for $\sigma_t^2 > 0$. Covariance stationarity of the GARCH model is ensured when $\sum_{i=1}^{\max(v,u)}(\alpha_i + \beta_i) < 1$, while the unconditional variance of ε_t is $\sigma^2 = E(\varepsilon_t^2) = \omega/\{1 - (\sum_{i=1}^{v}\alpha_i + \sum_{i=1}^{u}\beta_i)\}$ (Zivot, 2009). Equation (4.3) shows that the conditional variance σ_t^2 is a linear function of past squared innovation ε_{t-i}^2 and past conditional variances σ_{t-i}^2. The GARCH model is a more parsimonious specification (Nelson, 1991; Samiev, 2012) since it is an equivalence of a certain ARCH(∞) model (Zivot, 2009). When $u = 0$ in Equation (4.3), the GARCH model changes to the ARCH model with conditional variance stated as

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_v\varepsilon_{t-v}^2. \tag{4.4}$$

GARCH(1,1) is the simplest model specification with u = 1 and v = 1 in Equation (4.3), and it is conceivably the best candidate GARCH model for several applications (Fan et al., 2014; Zivot, 2009). The volatility persistence of the GARCH$(1,1)$ model is defined as α_1 + β_1

(see Engle and Bollerslev, 1986; Ghalanos, 2018). Volatility persistence is used to evaluate the speed of decay of shocks to volatility (Kim et al., 2020). Volatility exhibits long persistence into the future if $\alpha + \beta \to 1$, hence the closer the sum of the coefficients is to one (zero), the greater (lesser) the persistence. However, if the sum is equal to one, then shocks to volatility persist forever and the unconditional variance is not determined by the model. This process is called integrated-GARCH (Chou, 1988; Engle and Bollerslev, 1986). If the sum is greater than one, the conditional variance process is explosive, suggesting that shocks to the conditional variance are highly persistent.

For the maximum likelihood estimation (MLE), the log-likelihood function for maximising the likelihood of the unknown parameters given the observations is stated as

$$L(\boldsymbol{\vartheta}|\varepsilon) = \sum_{t=1}^{N} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right), \tag{4.5}$$

where $\boldsymbol{\vartheta} = (\mu, \omega, \alpha_1, ..., \alpha_v, \beta_1, ..., \beta_u)^\top$ is a vector of parameters, and $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_N)$ is a realisation of length N. The quasi-maximum likelihood estimation (QMLE) based on the Normal distribution and MLE have the same set of instructions for estimating $\hat{\vartheta}$; the only difference, however, is in the estimation of a robust standard deviation of $\hat{\vartheta}$ in QMLE (see Bollerslev and Wooldridge, 1992; Feng and Shi, 2017; Francq and Zakoïan, 2004).

The maximised log-likelihood function with Student's t distribution (Duda and Schmidt, 2009) is stated as

$$\ln L(\vartheta|\varepsilon) = \sum_{t=1}^{N} \left(\ln\left[\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}}\right] - \frac{1}{2}\ln(\sigma^2) - \left(\frac{\nu+1}{2}\right)\ln\left[1 + \frac{\varepsilon_t^2}{\sigma^2(\nu-2)}\right]\right), \tag{4.6}$$

where $\Gamma(\cdot)$ and ν are the gamma function and degree of freedom, respectively.

4.2.2 The fGARCH Model

The family GARCH (fGARCH) model, developed by Hentschel (1995), is an inclusive model that nests some important symmetric and asymmetric GARCH models as sub-models. The nesting includes the simple GARCH (sGARCH) model (Bollerslev, 1986), the Absolute Value GARCH (AVGARCH) model (Schwert, 1990; Taylor, 1986), the GJR GARCH (GJRGARCH) model (Glosten et al., 1993), the Threshold GARCH (TGARCH) model (Zakoian, 1994), the Nonlinear ARCH (NGARCH) model (Higgins and Bera, 1992), the Nonlinear Asymmetric GARCH (NAGARCH) model (Engle and Ng, 1993), the Exponential GARCH (EGARCH) model (Nelson, 1991), and the Asymmetric Power ARCH (apARCH) model (Ding et al., 1993). The sub-model apARCH is also a family model (but less general than the fGARCH model) that nests the sGARCH, AVGARCH, GJRGARCH, TGARCH, NGARCH models, and the Log

ARCH model (Geweke, 1986; Pantula, 1986). The fGARCH(u, v) model is stated as

$$\sigma_t^{\gamma} = \omega + \sum_{j=1}^{v} \alpha_j \sigma_{t-j}^{\gamma} (|z_{t-j} - \lambda_{2j}| - \lambda_{1j}\{z_{t-j} - \lambda_{2j}\})^{\delta} + \sum_{j=1}^{u} \beta_j \sigma_{t-j}^{\gamma}. \tag{4.7}$$

Equation (4.7) is the conditional standard deviation's Box–Cox transformation, where the transformation of the absolute value function is carried out by the parameter δ, and γ determines the shape. The λ_{2j} and λ_{1j} control the shifts for asymmetric small shocks and rotations for large shocks, respectively. This robust fGARCH model allows different powers for σ_t and z_t to drive how the residuals are decomposed in the conditional variance equation. The fit of the full fGARCH model can be implemented with $\gamma = \delta$ (see Ghalanos, 2018). Volatility clustering in the returns can be quantified through the model's volatility persistence stated as

$$\hat{P} = \sum_{j=1}^{u} \beta_j + \sum_{j=1}^{v} \alpha_j \varrho_j, \tag{4.8}$$

where ϱ_j, expressed in Equation (4.9), is the expected value of z_t in the absolute value asymmetry term's Box–Cox transformation. The persistence is obtained in this study through the "persistence()" function in the R rugarch package. See Ghalanos (2022); Hentschel (1995) for details on fGARCH and the nested models.

$$\varrho_j = \mathrm{E}(|z_{t-j} - \lambda_{2j}| - \lambda_{1j}(z_{t-j} - \lambda_{2j}))^{\delta} = \int_{-\infty}^{\infty} (|z - \lambda_{2j}| - \lambda_{1j}(z - \lambda_{2j}))^{\delta} f(z, 0, 1, \ldots) dz \tag{4.9}$$

4.2.3 The True Parameter Recovery Measure

Since the focus of MCS studies involves the ability of the estimator to recover the true parameter (see Chalmers, 2019), this study applies the "true parameter recovery (TPR)" measure in Equation (4.10) to compute the level (degree) of recovery of the true parameter through the MCS estimator. The TPR measure is a means of evaluating the performance of the MCS estimates in recovering the true parameter. That is, it is used to determine how much of the true parameter value is recovered by the MCS estimator. The TPR is expressed as

$$\mathrm{TPR} = \left(K - \left[\frac{(\vartheta - \hat{\vartheta})}{\vartheta} \times K \right] \right) \%, \tag{4.10}$$

where $K = 0, 1, 2, \ldots, 100$ is the nominal recovery level, ϑ is the true data-generating parameter and $\hat{\vartheta}$ is its estimator from the simulated (synthetic) data. For instance, a TPR estimated value of 95% or 100% denotes that the MCS estimator recovers the complete 95% or 100% of the true parameter. This complete recovery of the true parameter ϑ can be achieved by the MCS estimator $\hat{\vartheta}$ when $\hat{\vartheta} \approx \vartheta$, where $\vartheta > 0$, such that the TPR outcome equals the given nominal recovery level K (i.e., TPR = $K\%$).

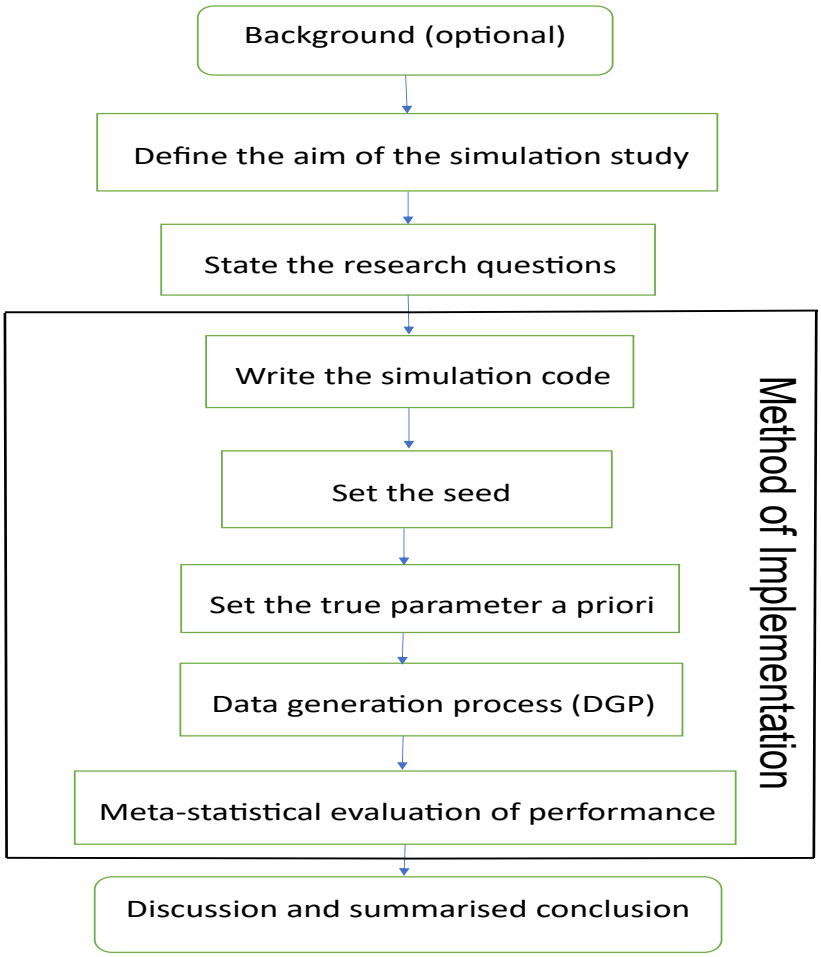

Figure 4.1: Simulation design flowchart to determine suitable assumed innovations.

4.2.4 Simulation Design

The design of the simulation framework includes "background (optional), defining the aim, research questions, method of implementation, and summarised conclusion". The method of implementation is a workflow that involves writing the code, setting the seed, setting the true parameter/s a priori, data generation process, and performance evaluation through meta-statistics. As summarised by the flowchart in Figure 4.1, these crucial steps are relevant for successful simulations through the GARCH model. The details of each design step are as follows:

4.2.4.1 Aim of the Simulation Study

After optionally stating the background that explains crucial underlying facts about the study, the next step is to define the aim of the study, and it must be clearly, concisely and unambiguously stated for the reader's understanding. The focus of MCS studies generally dwells on estimators' capabilities in recovering the true parameters ϑ, such that $\mathrm{E}(\hat{\vartheta}) = \vartheta$ for unbiasedness, $\hat{\vartheta} \to \vartheta$ as the sample size $N \to \infty$ for consistency, and root mean square error (RMSE) or standard error (SE) tends to zero as $N \to \infty$ for good efficiency or precision of the

true parameter's estimator. Hence, the aim of the study may revolve around these properties, namely, bias or unbiasedness, consistency, efficiency or precision of the estimator. The aim, then evolves from comparisons of multiple entities, such as comparing the efficiency of various error distributions, or comparing the performance of multiple models, or on improvement to an existing method.

4.2.4.2 State the Research Questions

After defining the study's aim, relevant questions concerning the purpose of the simulation should be outlined. These will be pointers to the objectives of the study. The intricacies of some statistical research questions make them better resolved via simulation approaches. Simulation provides a robust procedure for responding to a wide range of theoretical and methodological questions and can offer a flexible structure for answering specific questions pertinent to researchers' quest (Hallgren, 2013).

4.2.4.3 Method of Implementation

The simulation and empirical modelling of this study are implemented in R Statistical Software, version 4.0.3, with RStudio version 2022.12.0+353, using the rugarch (Ghalanos, 2018, 2022), SimDesign (Chalmers and Adkins, 2020), tidyverse (Wickham et al., 2019), zoo (Zeileis and Grothendieck, 2005), aTSA (Qiu, 2015) and forecast (Hyndman and Khandakar, 2008) packages. Computation is executed on Intel(R) Core(TM) i5-8265U CPU @ 1.60GHz 1.80 GHz. The method of implementing the simulation is as follows:

- Write the code: Carrying out a proper simulation experiment that mirrors real-life situations can be very demanding and computationally intensive, hence readable computer code with the right syntax must be produced. The code in this study is written to fit the true model[1] to the real data to obtain the true parameter representations[2] for the MCS. These true parameter values and other outputs from the fit are used in the code to generate simulated datasets that are analysed to obtain the MCS estimators. The standard errors of the estimates are also obtained in the process.

- Set the seed: Simulation code will generate a different sequence of random numbers each time it is run unless a seed is set (Danielsson, 2011). A set seed initialises the random number generator (Ghalanos, 2022) and ensures reproducibility, where the same result is obtained for different runs of the simulation process (Foote, 2018). The seed needs to be set only once, for each simulation, at the start of the simulation session (Ghalanos, 2022;

[1] We describe the true model as the data-generating model fitted with the true sampling distribution (see Feng and Shi, 2017).

[2] When the true model is fitted to the real data, the estimates from the fit represent the true parameters.

Morris et al., 2019), and it is better to use the same seed values throughout the process (Morris et al., 2019).

Now, through the GARCH model, this study carries out an MCS experiment to ascertain whether the seed values' pattern or arrangement affects the estimators' efficiency and consistency properties. Two sets of seeds are used for the experiment, where each set contains three different patterns of seed values. The first set is S_1 = {12345, 54321, 15243}, while the second set S_2 = {34567, 76543, 36547}. In each set, the study tries to use seed values arranged in ascending order, then reverses the order, and finally mixes up the ordered arrangement. The simulation starts by using GARCH(1,1)-Student's t, with a degree of freedom $\nu = 3$, as the true model under four assumed error distributions of a Normal, Student's t, Generalised Error Distribution (GED) and Generalised Hyperbolic (GHYP) distribution. Details on these selected error distributions can be seen in Ghalanos (2018) and Barndorff-Nielsen et al. (2013). The true parameter values used are (μ, ω, α, β) = (0.0678, 0.0867, 0.0931, 0.9059), and they are obtained by fitting GARCH(1,1)-Student's t to the SA bond return data.

Using each of the seed patterns in turn, simulated datasets of sample size N = 12000, repeated 1000 times are generated through the parameter values. However, because of the effect of initial values in the data generating process, which may lead to size distortion (Su, 2011), the first N = {11000, 10000, 9000, 8000} sets of observations are each discarded at each stage of the generated 12000 observations to circumvent such distortion. That is, only the last N = {1000, 2000, 3000, 4000} are used under each of the four assumed error distributions, as shown in Table A.1, Appendix A. These trimming steps are carried out following the simulation structure of Feng and Shi (2017)[1]. An observation-driven process such as the GARCH can be size-distorted with regards to its kurtosis, where strong size distortion may be a result of high kurtosis (Silvennoinen and Teräsvirta, 2016). The extracts of the RMSE and SE outcomes for the GARCH volatility persistence estimator $\hat{\alpha}+\hat{\beta}$ are shown in Table A.1. For S_1 in Panel A of the table, as N tends to its peak, the performance of the RMSE from the lowest to the highest under the four error distribution assumptions is Student's t, GHYP, GED and Normal in that order, while that of SE from the lowest to the highest is GHYP, Student's t, GED and Normal in that order, for the three arrangements of seed values.

For S_2 in Panel B of the table, as N reaches its peak at 4000, the performance of the RMSE from the lowest to the highest is Student's t, GHYP, GED and Normal in that order, while that of SE from the lowest to the highest is GHYP, GED, Student's t and Normal in that order, for the three S_2 patterns of seed values. Hence, efficiency and precision in terms of RMSE and SE are the same as the sample size N becomes larger under the three seeds, regardless of the

[1]This study only follows the authors' trimming steps for initial values effect. The other trimming by the authors for "simulation bias" (where some initial numbers of replications are further discarded after the initial value effect adjustment) are not used here because it is observed that it sometimes distorts the estimator's $\sqrt{N}$ consistency.

arrangement of the seed values under S_2, as also observed under S_1. In addition, the flows of $\sqrt{N}$ consistency of the estimator under the seed values in S_1 are roughly the same; this is also applicable to those of the seed values in S_2. The plotted outcomes can be visualised as displayed by the trend lines within the 95% confidence intervals in Figure 4.2 for the three seed values of sets S_1 in Panel A and S_2 in Panel B, where the efficiency and consistency outcomes are roughly the same with increase in N.

To summarise, this study observes that, as $N \to \infty$, the pattern or arrangement of the seed values does not affect the estimator's overall consistency and efficiency properties, but this may likely depend on the quality of the model used. The seed is primarily used to ensure reproducibility. The main reason for using a seed is to be able to reproduce a particular sequence of generated random numbers. Hence, the choice of a particular seed number is not to create a truly random number, but only to ensure that the same set of random numbers (or simulated data) are created or reproduced for each run of the simulation. Streams (or sequences) of random numbers for different seeds should exhibit similar statistical properties. That is, the sequence of random numbers obtainable from using a seed of 12345 to initialise a random number generator should be just as random as the ones obtainable from any other seed value (Wicklin, 2017). Panels C and D of the figure further reveal that the RMSE/SE $\to 0$ as $N \to \infty$ for the four error distributions in S_1 and S_2.

Table A.1 further shows that the MCS estimator $\hat{\alpha} + \hat{\beta}$ considerably recovers the true parameter $\alpha + \beta$ at the 95% nominal recovery level, where some of the estimates even recover the complete true value (0.9990) with TPR outcomes of 95%. These recovery outcomes can be seen in the visual plots of Figure 4.3 (or as shown in Panels A and B of Figure B.1, Appendix B), where Panels A(i) and B(i) reveal that the MCS estimates perform quite well in recovering the true parameter as shown by the closeness of the TPR outcomes to the 95% (i.e., 0.95) nominal recovery level for S_1 and S_2, respectively. The bunched up TPR outcomes in Panels A(i) and B(i) are clearly spread out as shown in Panels A(ii) and B(ii) for S_1 and S_2, respectively. From these recovery outputs, two distinct features can be observed. First, the TPR results do not depend on the sample size as shown in Panels A and B of Figure 4.4 for S_1 and S_2, which is a feature of coverage probability (see Hilary, 2002); second, the closer (farther) the MCS estimate is to zero, the smaller (larger) the TPR outcome, as revealed in Panels C and D of the figure.

- After setting the seed, the true parameter representations of the true sampling distribution (or true model) are then set a priori (Koopman et al., 2017; Mooney, 1997).

- Next, simulated observations are generated using the true sampling distribution or the true model given some sets of (or different sets of) fixed parameters. Generation of simulated datasets through the GARCH model is carried out using the R package "rugarch". Random data generation involving this package can be implemented using either of two approaches. The first approach is to carry out the data-generating simulation directly on

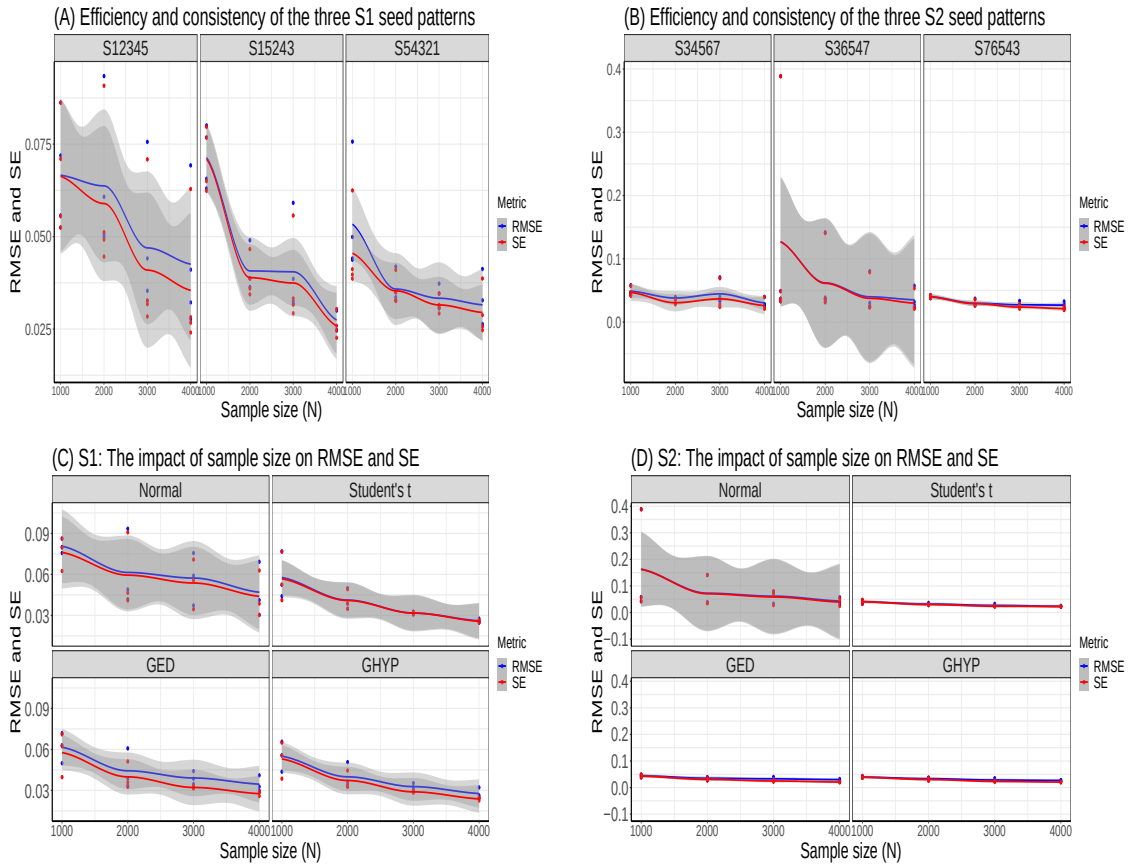

Figure 4.2: Panels (**A**,**B**) show the efficiency and $\sqrt{N}$ consistency in S_1 and S_2 for each seed pattern, while Panels (**C**,**D**) reveal the impacts of sample size on RMSE and SE under the assumed errors in S_1 and S_2.

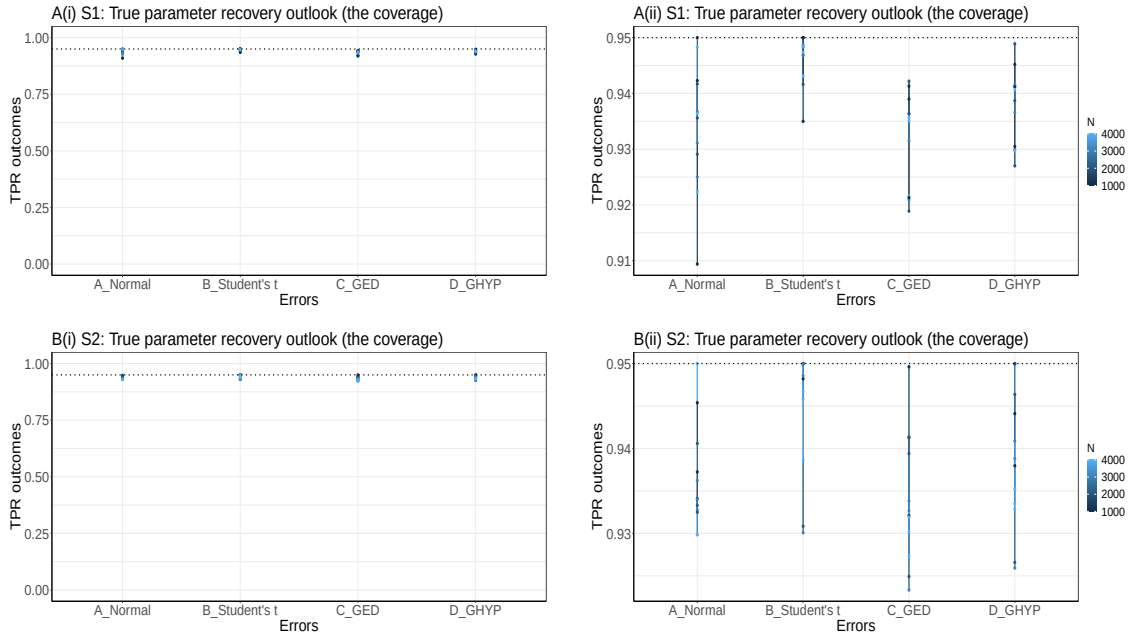

Figure 4.3: TPR outcomes in Panels **A**(**i**) and **B**(**i**) for S_1 and S_2, respectively. The outcomes are clearly spread out in Panels **A**(**ii**) and **B**(ii) for S_1 and S_2. The dotted lines are the 95% (i.e., 0.95) nominal recovery levels.

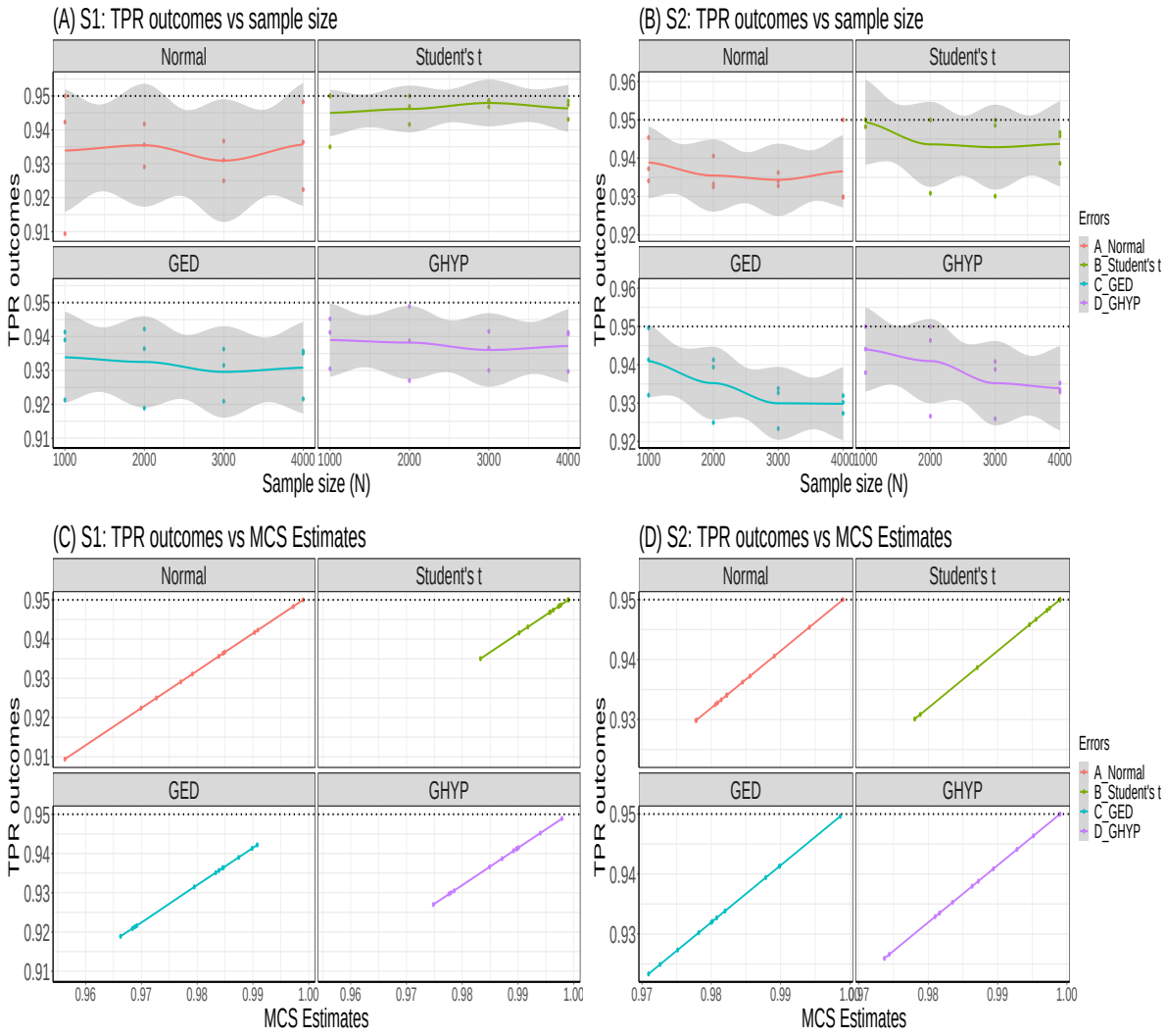

Figure 4.4: The TPR outcomes against sample size are shown in Panels (**A**,**B**), while Panels (**C**,**D**) show the TPR outcomes against the MCS estimates.

a fitted object "fit" using the ugarchsim function for the simulated random data. The second approach uses the ugarchpath function, which enables simulation of desired number of volatility paths through different parameter combinations (see Ghalanos, 2018, 2022; Pfaff, 2016 for relevant details on the two functions and their usage).

The simulation or data-generating process can be run once or replicated multiple times. This study carries out another MCS investigation through the GARCH model to determine the effect (on the outcomes) of running a given GARCH simulation once or replicating it multiple times. That is, for a given sample size and seed value, the outcome of running the simulation once is compared to that of running it with different replications such as 2500, 1000 and 300. This MCS experiment uses GARCH(1,1)-Student's t, with $\nu = 3$, as the true model under four assumed error distributions of a Normal, Student's t, GED and GHYP. However, it should be understood that any non-normal error distributions (apart from the Student's t that is used here)

can also be used with GARCH(1,1) model for the true model. The GARCH(1,1)-Student's t fitted to the SA bond return data yields the true parameter values (μ, ω, α, β) = (0.0678, 0.0867, 0.0931, 0.9059).

Using these parameter values, datasets of sample size N = 12000 are generated in each of the four distinct simulations (i.e., simulations with 1, 2500, 1000 and 300 replicates). After necessary trimmings in each simulation, to evade initial values effect, the last N = {1000, 2000, 3000} sets of observations are used at each stage of the generated 12000 observations under the four assumed innovation distributions. That is, datasets of the last three sample sizes, each simulated once, then replicated L = {2500, 1000, and 300} times are consecutively generated. From the modelling outputs, it is observed that the log-likelihood (llk), RMSE, SE and bias outcomes of $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\alpha} + \hat{\beta}$ estimators for each simulation under the four assumed errors are the same for the three sample-size datasets with the same seed value, regardless of whether the simulation is run once or replicated multiple times. For brevity, this study only displays the outcomes of the experiment under the assumed GED error for each run in Table 4.1. However, increasing the number of replications may reduce sampling uncertainty in meta-statistics (Chalmers and Adkins, 2020).

- The generated (simulated) data are analysed, and the estimates from them are evaluated using classic methods through meta-statistics to derive relevant information about the estimators. Meta-statistics (see Chalmers and Adkins, 2020) are performance measures or metrics for assessing the modelling outputs by judging the closeness between an estimate and the true parameter. A few of the frequently used meta-statistical summaries, as described below, include bias, root mean square error (RMSE) and standard error (SE). For more meta-statistics, see Chalmers and Adkins (2020); Morris et al. (2019); Sigal and Chalmers (2016).

Bias

The bias, on average, measures the tendency of the simulated estimators $\hat{\vartheta}$ to be smaller or larger than their true parameter value ϑ. It is defined as the average difference between the true (population) parameter and its estimate (Feng and Shi, 2017). The optimal value of bias is 0 (Harwell, 2018; Sigal and Chalmers, 2016). An unbiased estimator, on average, yields the correct value of the true parameter. Bias with a positive (negative) value indicates that the true parameter value is over-estimated (under-estimated). However, in absolute values, the closer the estimator is to 0, the better it is. Bias is stated mathematically as $\mathtt{E}(\tilde{\vartheta} - \vartheta)$, but can be presented in MCS (see Chalmers, 2019) as bias $= \frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \vartheta)$. The two formulae are connected as

$$\mathtt{E}(\tilde{\vartheta} - \vartheta) = \mathtt{E}(\tilde{\vartheta}) - \vartheta = \frac{1}{L}\sum_{i=1}^{L}\hat{\vartheta}_i - \vartheta = \frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \vartheta), \tag{4.11}$$

Table 4.1: Outcomes of different simulation replicates.

α	β	N	llk	$\text{RMSE}_{\hat{\alpha}}$	$\text{Bias}_{\hat{\alpha}}$	$\text{SE}_{\hat{\alpha}}$	$\text{RMSE}_{\hat{\beta}}$	$\text{Bias}_{\hat{\beta}}$	$\text{SE}_{\hat{\beta}}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{Bias}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$
Panel A: Simulation run once												
0.0931	0.9059	1000	−2020.5	0.0504	0.0328	0.0383	0.0551	−0.0443	0.0327	0.0719	−0.0115	0.0710
		2000	−3813.8	0.0246	0.0046	0.0241	0.0462	−0.0374	0.0271	0.0608	−0.0327	0.0512
		3000	−5734.2	0.0156	−0.0037	0.0152	0.0316	−0.0269	0.0166	0.0441	−0.0306	0.0317
Panel B: Simulation run with 2500 replications												
0.0931	0.9059	1000	−2020.5	0.0504	0.0328	0.0383	0.0551	−0.0443	0.0327	0.0719	−0.0115	0.0710
		2000	−3813.8	0.0246	0.0046	0.0241	0.0462	−0.0374	0.0271	0.0608	−0.0327	0.0512
		3000	−5734.2	0.0156	−0.0037	0.0152	0.0316	−0.0269	0.0166	0.0441	−0.0306	0.0317
Panel C: Simulation run with 1000 replications												
0.0931	0.9059	1000	−2020.5	0.0504	0.0328	0.0383	0.0551	−0.0443	0.0327	0.0719	−0.0115	0.0710
		2000	−3813.8	0.0246	0.0046	0.0241	0.0462	−0.0374	0.0271	0.0608	−0.0327	0.0512
		3000	−5734.2	0.0156	−0.0037	0.0152	0.0316	−0.0269	0.0166	0.0441	−0.0306	0.0317
Panel D: Simulation run with 300 replications												
0.0931	0.9059	1000	−2020.5	0.0504	0.0328	0.0383	0.0551	−0.0443	0.0327	0.0719	−0.0115	0.0710
		2000	−3813.8	0.0246	0.0046	0.0241	0.0462	−0.0374	0.0271	0.0608	−0.0327	0.0512
		3000	−5734.2	0.0156	−0.0037	0.0152	0.0316	−0.0269	0.0166	0.0441	−0.0306	0.0317

where $\tilde{\vartheta} = \hat{\vartheta}_i (i = 1, ..., L)$ is a finite ith number of the sample estimate $\hat{\vartheta}$ for the datasets, L is the number of replications, and ϑ is the true parameter.

Standard Error

Sampling variability in the estimation can be evaluated via the standard error (SE) as stated (see Chalmers, 2019; Yuan et al., 2015) in Equation 4.12. Also called Monte Carlo standard deviation, it is a measure of the efficiency or precision of the true parameter's estimator, which is used to estimate the long-run standard deviation of $\hat{\vartheta}_i$ for finite repetitions. It does not require knowing the true parameter ϑ but depends on its estimator $\hat{\vartheta}_i$ only. The smaller the sampling variability, the more the efficiency or precision of the ϑ's estimator (see Morris et al., 2019). Sampling variability decreases with increased sample size (Sigal and Chalmers, 2016).

$$\text{SE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \bar{\vartheta})^2}, \quad \text{where} \quad \bar{\vartheta} = \sum_{i=1}^{L}\hat{\vartheta}/L. \tag{4.12}$$

Root Mean Square Error

The root mean square error (RMSE) is an accuracy measure for evaluating the difference between a model's true value and its prediction. RMSE measure indicates the sampling error of an estimator when compared to the true parameter value (Sigal and Chalmers, 2016) and it is stated as

$$\text{RMSE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(\hat{\vartheta}_i - \vartheta\right)^2}. \tag{4.13}$$

Its computation involves the true parameter ϑ. An estimator with lesser RMSE is more efficient in recovering the true parameter value (Sigal and Chalmers, 2016; Yuan et al., 2015), and minimum RMSE produces maximum precision (Wang et al., 2018). Consistency of the estimator occurs when RMSE decreases such that $\hat{\vartheta} \to \vartheta$ as the sample size $N \to \infty$ (Ghalanos, 2018; Morris et al., 2019). RMSE is related to bias and sampling variability as

$$\text{RMSE} = \sqrt{\text{bias}^2 + \text{SE}^2}. \tag{4.14}$$

That is, the RMSE is an inclusive measure that combines bias and SE, such that low SE can be penalised for bias. The mean squared error (MSE) is obtained by squaring the RMSE. MCS is highly reliant on the law of large numbers, and it is expected that the distribution of an appropriately large sample should converge to that of the underlying population as the sample size increases (Gilli et al., 2019). It is also expected that the Monte Carlo sampling error should decrease as the sample size increases, but this is not always the case. That is, the sample size cannot always be sufficiently increased to limit the sampling error to a tolerable level (Gilli et al., 2019).

4.2.4.4 Discussion and Summary

After implementing the method, the last stage in the framework steps is the conclusion, which needs to reflect a summary discussion of all logical findings from the experiments, with answers to the research questions. The conclusion brings out the novelty of the research and may also include the limitations experienced and opportunities for future work. In addition, relevant information on simulation results can be conveyed through graphics, tabular presentation, or both.

4.3 Results: Simulation and Empirical

4.3.1 Practical Illustrations of the Simulation Design: Application to Bond Return Data

By way of illustration, this section practically describes how the stated steps can be applied using Monte Carlo simulations with a real data empirical verification.

4.3.1.1 The Background

It is believed that observation-driven models can appropriately estimate volatility when fitted with a suitable error distribution (Bollerslev, 1987). Observation-driven modelling exists in the presence of time-varying parameters, where parameters are functions of lagged dependent variables, concurrent variables and lagged exogenous variables (see Buccheri et al., 2021; Creal et al., 2013 for details). Data generation using the rugarch package can be performed through a variety of models that include the simple GARCH, the exponential GARCH (EGARCH), the GJR-GARCH, the Component GARCH (CGARCH) (Lee and Engle, 1999), the Multiplicative Component GARCH (MCGARCH) (Engle and Sokalska, 2012), among others, and two omnibus models apARCH and fGARCH (as described in Section 4.2.2).

The apARCH model is less robust than the fGARCH model (Ghalanos, 2018), hence the latter is used for the data generation in this study. Specifically, fGARCH(1,1) model is used as the true data-generating process (DGP) for the MC simulation because the first lag of conditional variability can considerably capture volatility clustering existing in the time series data. In other words, the dependence of volatility on recent past activities is more than on distant past activities (Javed and Mantalos, 2013). Hence, this illustrative study showcases the effectiveness of the observation-driven model fGARCH for estimating the persistence of volatility, where the outcomes of the model fitted with each of ten assumed innovation distributions of the Normal, skew-Normal, Student's t, skew-Student's t, GED, skew-GED, GHYP, Normal Inverse Gaussian (NIG), Generalised Hyperbolic Skew-Student's t (GHST) distribution and Johnson's reparametrised SU (JSU) distribution are compared. Details on the error distributions can be found in Ashour and Abdel-hameed (2010); Azzalini (1985); Azzalini and Capitanio (2003); Barndorff-Nielsen et al. (2013); Branco and Dey (2001); Eling (2014); Ghalanos (2018); Lee and Pai (2010); Pourahmadi (2007).

The DGP fGARCH(1,1) model, as stated in Equation (4.15), is used to generate simulated return observations using the Student's t error with ν = 4.1 as the true error distribution. This model deduced from Equation (4.7) is given by

$$\sigma_t^{\gamma} = \omega + \alpha_1 \sigma_{t-1}^{\gamma}(|z_{t-1} - \lambda_{21}| - \lambda_{11}\{z_{t-1} - \lambda_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma} \tag{4.15}$$

Here, a Student's t with shape parameter or degrees of freedom $\nu = 4.1$ is used to ensure that $\mathrm{E}[z_t^4] < \infty$, which enables $\sqrt{N}$ consistency of the QML estimation following the assumption of Francq and Thieu (2019) (see Hoga, 2022). Moreover, the Student's t distribution is used as the true error distribution in this study because it can suitably capture leptokurtic or fat-tailed features (Duda and Schmidt, 2009; Lin and Shen, 2006) inherent in financial data (Hentschel, 1995), and it is also assumed that stock prices appear to have a distribution much like the Student's t (Heracleous, 2007). However, based on relevance and research needs, users may choose to use any leptokurtic distributions, such as the GED or others, for their data generation. Simulation through the rugarch package can be carried out using the ugarchsim[1] and ugarchpath functions, but not all the stated data-generating models currently support the use of ugarchpath method (see Ghalanos, 2018). Hence, this illustration is implemented using the ugarchsim function through the "fit object" approach. The ugarchsim function has been used in Shahriari et al. (2023); Søfteland and Iversen (2021); Zhang (2017) as it gives the user flexibility and control.

Further background study reveals the findings of Morris et al. (2019), where the authors showed that RMSE is more applicable as a performance measure where the objective of the simulation is prediction rather than estimation. The authors discussed how more sensitive RMSE is to the choice of the number of observations used during method comparisons than when only SE or bias is used. Hence, for fairness in performance assessments, the SE is used as the key metric or measure of efficiency (precision) in this illustrative study.

It is also noticed from the outcomes of the fGARCH modelling that two sets of standard error (SE) estimates are returned. That is, the default MLE SE and the robust QMLE SE (Ghalanos, 2018; White, 1982; Zivot, 2013). This study used the robust fGARCH QMLE SEs for the simulation illustrations because they are claimed to be consistent (but not efficient) and asymptotically normally distributed if the volatility and mean equations are well specified (Bollerslev and Wooldridge, 1992; Wuertz et al., 2020).

4.3.1.2 Aim of the Simulation Study

This study aims at obtaining the most appropriate assumed error distribution for volatility persistence estimation when the underlying (true) error distribution is unknown.

4.3.1.3 Research Questions

This simulation study should result in responses to the following questions:

1. Which among the assumed error distributions is the most appropriate from the fGARCH process simulation for estimating the persistence of the volatility?

[1] See the "Synopsis of R packages" pages 125–27 in Pfaff (2016) for relevant details on rugarch package and ugarchsim function.

2. Financial data are fat-tailed (Li, 2008), i.e., non-Normal. Hence, will the combined volatility estimator $\hat{\alpha} + \hat{\beta}$ of the most suitable error assumption still be consistent under departure from Normal assumption?

3. What type (i.e., strong, weak or inconsistent) of $\sqrt{N}$ consistency, in terms of RMSE and SE, does the fGARCH estimator $\hat{\alpha} + \hat{\beta}$ exhibit?

4. How is the performance of the MCS estimator $\hat{\alpha} + \hat{\beta}$ in recovering the true parameter?

4.3.1.4 Method of Implementation

To initiate the implementation method, the written code in Appendix C is first used to fit the true model fGARCH(1,1)-Student's t to the SA bond return data (BondDataSA) through the ugarchfit function of the fGARCH fit object. Next, through the ugarchsim function, using seed 12345 in the code, the outputs from the fit are set (or used) a priori as the true parameter values (α, β) = (0.0748, 0.9243) for the simulation process as shown in Table 4.2. These parameter values with other estimates from the fit object are used directly to generate (simulate) N = 15000 sample size observations, replicated 1000 times. However, after trimming down the simulated dataset, following the simulation structure of Feng and Shi (2017), to prevent the effect of initial values, by removing the first N = {7000, 6000, 5000} sets of observations at each stage of the simulated 15000 observations, the last N = {8000, 9000, 10000} observations are processed under each of the ten assumed innovation distributions as shown in Table 4.2. For brevity, the presented code in Appendix C only shows the command lines for the first stage of the simulated data, with the trimming. This briefly illustrates how the 15000 observations are generated through the ugarchsim function and then trimmed down to 8000. The remaining two stages (i.e., N = 9000 and 10000) of the data generation and trimmings are run following this same pattern.

Figure 4.5 displays the visual outlooks of the simulated returns and volatilities for the first three series in the 1000 replicated series for N = 8000. These sampled visuals show that each of the 1000 replicated series of the simulated (synthetic) data has a unique randomness and shape that make them different from every other series. Hence, the estimate from the family GARCH simulation is the average of all the estimates from the different replicated series.

After generating simulated observations, the fGARCH(1,1) model is fitted to each simulated dataset under the ten error term assumptions as shown in the code. However, for brevity in the written code, the fGARCH(1,1) model is only fitted under the Normal error using the "distribution.model = norm" argument in the ugarchspec function. All the other error assumptions can be fitted in the same pattern by simply replacing the "norm" with the naming convention of the relevant error distribution, e.g., "snorm" for skew-Normal, "std" for Student's t, "sstd" for skew-Student's t (see Ghalanos, 2018 for details, and the complete code can be found at `https://github.com/rsamuel11`). The parsimonious ARMA(1,1) model is also used in the

code as the most suitable among the tested candidate ARMA models to remove serial correlation in the simulated observations. However, consistency can still be achieved in simulation modelling regardless of correlated sample draws. That is, the sampled variates do not need to be independent to achieve consistency (Chib, 2015).

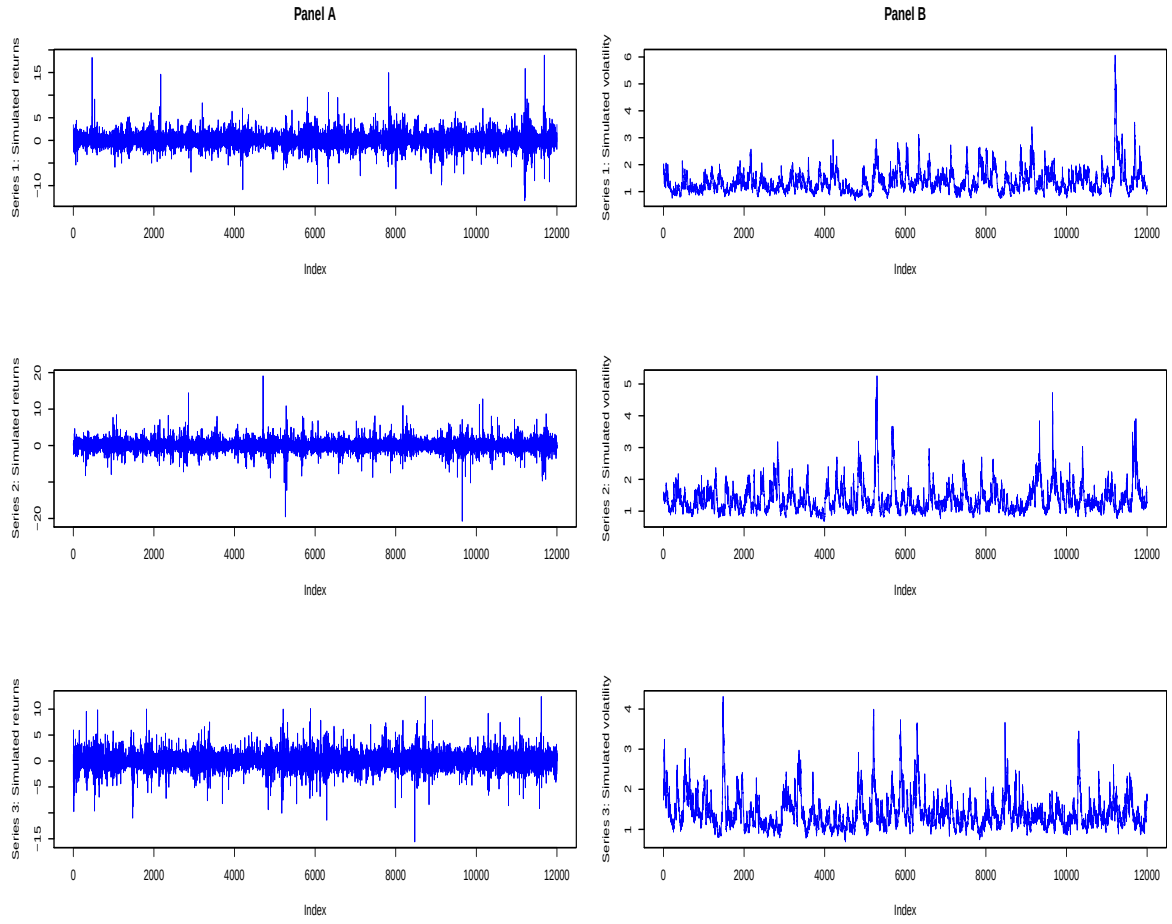

Figure 4.5: Simulated returns (in Panel **A**) and simulated volatility (Panel **B**) of the first three replicated series.

Next, the selected meta-statistics are now used to evaluate the estimators. The most suitable assumed error distribution for estimating the persistence will be obtained from the estimator/s with the best precision and efficiency from the meta-statistical comparisons made under all the selected error term assumptions. Three meta-statistical summaries that include the bias, RMSE and SE, are used in this illustration. The computation of the metrics is direct but may sometimes be nerve-racking, and manual programming may even cause unanticipated coding errors and other abrupt setbacks. To circumvent this, SimDesign statistical package (Chalmers and Adkins, 2020; Sigal and Chalmers, 2016) with in-built meta-statistical functions for computational accuracy is used in this illustration, beginning from bias estimation. The log-likelihood (llk) of the estimates, with the RMSE, bias and SE for $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\alpha} + \hat{\beta}$ es-

Table 4.2: True model fGARCH(1,1)-Student's t with true parameters α = 0.0748, β = 0.9243 and $\alpha + \beta$ = 0.9991.

	$\mathcal{N}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}+\hat{\beta}$	llk	$\text{RMSE}_{\hat{\alpha}}$	$\text{Bias}_{\hat{\alpha}}$	$\text{SE}_{\hat{\alpha}}$	$\text{RMSE}_{\hat{\beta}}$	$\text{Bias}_{\hat{\beta}}$	$\text{SE}_{\hat{\beta}}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{Bias}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	$\text{TPR}_{\hat{\alpha}+\hat{\beta}}$
															95%
Panel A	8000	0.0835	0.9234	1.0069	−13860.4	0.0390	0.0087	0.0380	0.0377	−0.0009	0.0377	0.0761	0.0078	0.0757	95.74%
Normal	9000	0.0790	0.9259	1.0049	−15490.9	0.0297	0.0042	0.0294	0.0465	0.0015	0.0464	0.0760	0.0058	0.0758	95.55%
	10000	0.0803	0.9281	1.0085	−17081.0	0.0088	0.0055	0.0069	0.0091	0.0038	0.0082	0.0178	0.0094	0.0151	95.89%
Panel B	8000	0.0834	0.9235	1.0069	−13860.4	0.0385	0.0086	0.0375	0.0372	−0.0009	0.0372	0.0752	0.0078	0.0748	95.74%
skew-	9000	0.0792	0.9262	1.0055	−15490.6	0.0138	0.0044	0.0131	0.0216	0.0019	0.0215	0.0352	0.0064	0.0346	95.61%
Normal	10000	0.0801	0.9284	1.0085	−17080.2	0.0085	0.0053	0.0066	0.0089	0.0041	0.0079	0.0173	0.0094	0.0145	95.89%
Panel C	8000	0.0736	0.9226	0.9963	−13337.1	0.0060	−0.0012	0.0058	0.0059	−0.0017	0.0056	0.0118	−0.0029	0.0115	94.73%
Student's t	9000	0.0727	0.9279	1.0006	−14912.0	0.0054	−0.0021	0.0050	0.0056	0.0036	0.0043	0.0094	0.0014	0.0093	95.14%
	10000	0.0735	0.9263	0.9999	−16428.3	0.0043	−0.0013	0.0041	0.0035	0.0020	0.0028	0.0070	0.0008	0.0069	95.07%
Panel D	8000	0.0732	0.9225	0.9957	−13337.1	0.0084	−0.0016	0.0083	0.0064	−0.0018	0.0062	0.0149	−0.0034	0.0145	94.68%
skew-	9000	0.0715	0.9262	0.9977	−14912.2	0.0061	−0.0033	0.0051	0.0040	0.0019	0.0036	0.0088	−0.0014	0.0087	94.87%
Student's t	10000	0.0743	0.9277	1.0020	−16428.4	0.0035	−0.0005	0.0034	0.0041	0.0034	0.0024	0.0065	0.0029	0.0058	95.27%
Panel E	8000	0.0770	0.9244	1.0014	−13386.3	0.0079	0.0022	0.0076	0.0076	0.0001	0.0076	0.0153	0.0023	0.0152	95.22%
GED	9000	0.0734	0.9266	1.0000	−14966.3	0.0056	−0.0014	0.0054	0.0053	0.0023	0.0048	0.0103	0.0009	0.0103	95.09%
	10000	0.0753	0.9275	1.0028	−16492.3	0.0036	0.0005	0.0035	0.0042	0.0032	0.0027	0.0073	0.0037	0.0062	95.35%
Panel F	8000	0.0750	0.9221	0.9971	−13386.2	0.0059	0.0002	0.0059	0.0059	−0.0022	0.0054	0.0115	−0.0020	0.0113	94.81%
skew-	9000	0.0734	0.9265	0.9999	−14966.0	0.0055	−0.0014	0.0054	0.0054	0.0022	0.0049	0.0103	0.0008	0.0103	95.08%
GED	10000	0.0753	0.9275	1.0028	−16492.3	0.0035	0.0006	0.0035	0.0040	0.0031	0.0025	0.0070	0.0037	0.0060	95.35%
Panel G	8000	0.0732	0.9234	0.9966	−13336.3	0.0065	−0.0016	0.0063	0.0054	−0.0009	0.0053	0.0119	−0.0025	0.0116	94.76%
GHYP	9000	0.0720	0.9279	0.9999	−14911.4	0.0057	−0.0028	0.0050	0.0056	0.0036	0.0043	0.0093	0.0008	0.0093	95.08%
	10000	0.0729	0.9265	0.9994	−16427.7	0.0045	−0.0019	0.0040	0.0035	0.0022	0.0027	0.0067	0.0003	0.0067	95.03%
Panel H	8000	0.0731	0.9229	0.9961	−13343.3	0.0058	−0.0017	0.0056	0.0057	−0.0014	0.0055	0.0115	−0.0031	0.0111	94.71%
NIG	9000	0.0719	0.9275	0.9994	−14919.7	0.0059	−0.0029	0.0052	0.0053	0.0031	0.0043	0.0095	0.0003	0.0095	95.03%
	10000	0.0729	0.9266	0.9995	−16438.1	0.0045	−0.0019	0.0041	0.0034	0.0023	0.0025	0.0066	0.0004	0.0066	95.04%
Panel I	8000	0.0711	0.9218	0.9930	−13435.0	0.0067	−0.0036	0.0056	0.0070	−0.0025	0.0065	0.0135	−0.0062	0.0121	94.42%
GHST	9000	0.0699	0.9261	0.9960	−15027.3	0.0071	−0.0049	0.0051	0.0049	0.0018	0.0046	0.0102	−0.0031	0.0097	94.71%
	10000	0.0734	0.9266	0.9999	−16569.1	0.0038	−0.0014	0.0035	0.0034	0.0022	0.0026	0.0061	0.0008	0.0061	95.08%
Panel J	8000	0.0731	0.9232	0.9963	−13337.1	0.0057	−0.0017	0.0055	0.0057	−0.0011	0.0055	0.0114	−0.0028	0.0110	94.74%
JSU	9000	0.0719	0.9277	0.9996	−14912.4	0.0057	−0.0029	0.0050	0.0053	0.0033	0.0042	0.0091	0.0005	0.0091	95.04%
	10000	0.0727	0.9264	0.9991	−16429.3	0.0045	−0.0020	0.0040	0.0033	0.0020	0.0026	0.0066	0.0000	0.0066	95.00%

timators are displayed in Table 4.2, but SE is the key performance measure for efficiency and precision.

Now comparing RMSE for $\hat{\alpha}$, the results from the table show that both the skew-GED and skew-Student's t outperform the other assumed innovations in efficiency with the least values as N tends to the peak at 10000. For $\hat{\beta}$, the JSU, followed by the GHST and NIG, outperforms the rest of the innovation assumptions in efficiency with the least RMSE value as N tends to the peak. For $\hat{\alpha} + \hat{\beta}$, the GHST followed by the skew-Student's t outperform the remaining eight innovation assumptions as N tends to the peak, but the skew-Student's t is the best as the sample size reaches the middle at N = 9000 for both $\hat{\alpha} + \hat{\beta}$ and $\hat{\beta}$.

Comparing bias for $\hat{\alpha}$, as N approaches the peak, the absolute values of biases for the GED and skew-Student's t outperform the rest. For $\hat{\beta}$, the JSU and the true innovation Student's t both take the lead as N reaches the peak. For $\hat{\alpha}+\hat{\beta}$, the JSU followed by the GHYP outperform the other innovations in absolute values of biases as N reaches the peak.

For precision and efficiency comparison in terms of the key performance metric SE, the skew-Student's t relatively outperforms the others in efficiency and precision as N tends to the peak for $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\alpha}+\hat{\beta}$, in particular, for $\hat{\beta}$ and $\hat{\alpha}+\hat{\beta}$. Finally, for the llk comparison, the GHYP outperforms the rest, with the largest estimates at the three N sample sizes. To summarise, when the true innovation is Student's t, the skew-Student's t assumed innovation distribution relatively outperforms the other nine innovation assumptions in efficiency and precision, while the GHYP performs better than the rest through the log-likelihood. It is observed here that the SEs of $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\alpha} + \hat{\beta}$ estimators for the assumed Normal innovation distribution are the largest when compared with those of the other nine assumed innovation distributions. This justifies the claim that the QMLE of the family GARCH model (with Normal innovation) is inefficient. Furthermore, it is observed from the tabulated outputs that the RMSE and SE of the estimators are considerably $\sqrt{N}$ consistent in recovering the true parameters under the assumed innovations. The visual illustrations of the consistency for the outputs of $\hat{\alpha} + \hat{\beta}$ estimator are graphically displayed in Figure 4.6. The figure shows that the closer the absolute values of biases are to zero, the closer the SE is to the RMSE. Whenever bias drifts away from zero, the gap between SE and RMSE widens, but if otherwise, then their trend lines closely follow the same trajectory. The visual plots also show that RMSE and SE decrease as N increases, but bias is independent of N.

It is also observed from the table and as shown in Panels A and B of Figure 4.7 that the MCS estimates for the estimator $\hat{\alpha} + \hat{\beta}$ considerably recover the true (volatility) parameter value of 0.9991, with TPR outcomes closely clustered around the 95% (i.e., 0.95) nominal recovery level under the ten error assumptions. This indicates a good performance of the MCS experiments with suitably valid outputs. However, the non-Normal errors perform slightly better in the recovery than the Normal and skew-Normal errors, as clearly revealed in Panel B. It can also be seen from the tabular results that the TPR outcomes are independent of N, and the closer the MCS estimate is to zero, the smaller the TPR estimate.

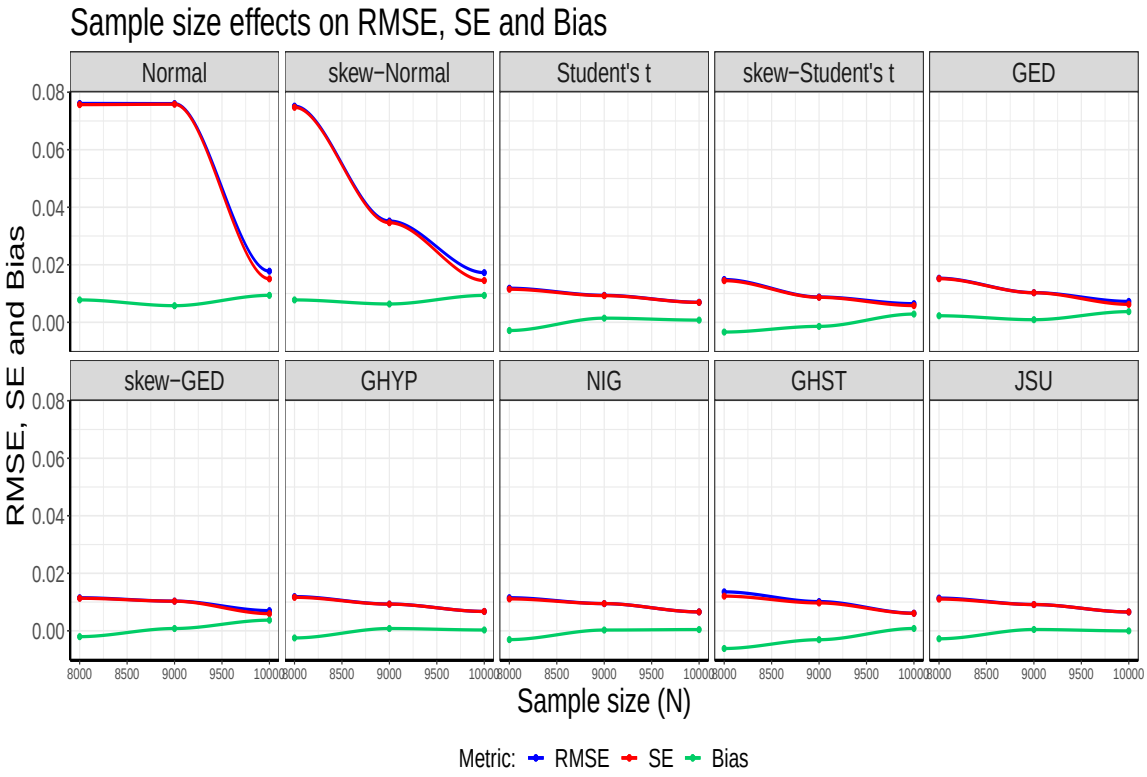

Figure 4.6: The impact of sample size on RMSE, SE and bias for the fGARCH(1,1)-Student's t MCS modelling. The RMSE and SE are considerably $\sqrt{N}$ consistent, but bias is independent of N.

4.3.2 Empirical Verification

Next, the outcomes of the MCS experiments are empirically verified using the real-life return data from the SA bond market index. Among the ten assumed error distributions, the most appropriate for the fGARCH process to estimate the volatility persistence of the bond market's returns is examined. For the market index, the price data are transformed to the log-daily returns by taking the difference of logarithms of the price, expressed in percentage as

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100. \tag{4.16}$$

The P_t and P_{t-1} are the closing bond price index at time t and the previous day's closing price at time $t-1$, respectively; r_t is the current return, and ln represents the natural logarithm.

4.3.2.1 Exploratory Data Analysis

To start with, the price index and returns are first inspected through exploratory data analysis (EDA) as displayed in Figure 4.8. The EDA visually sheds light on the content of the dataset to reveal relevant information and potential outliers. Figure 4.8 unearths some downswings or steep falls in the volatility of price (in plot a) and returns (in plot c) around the years 2002,

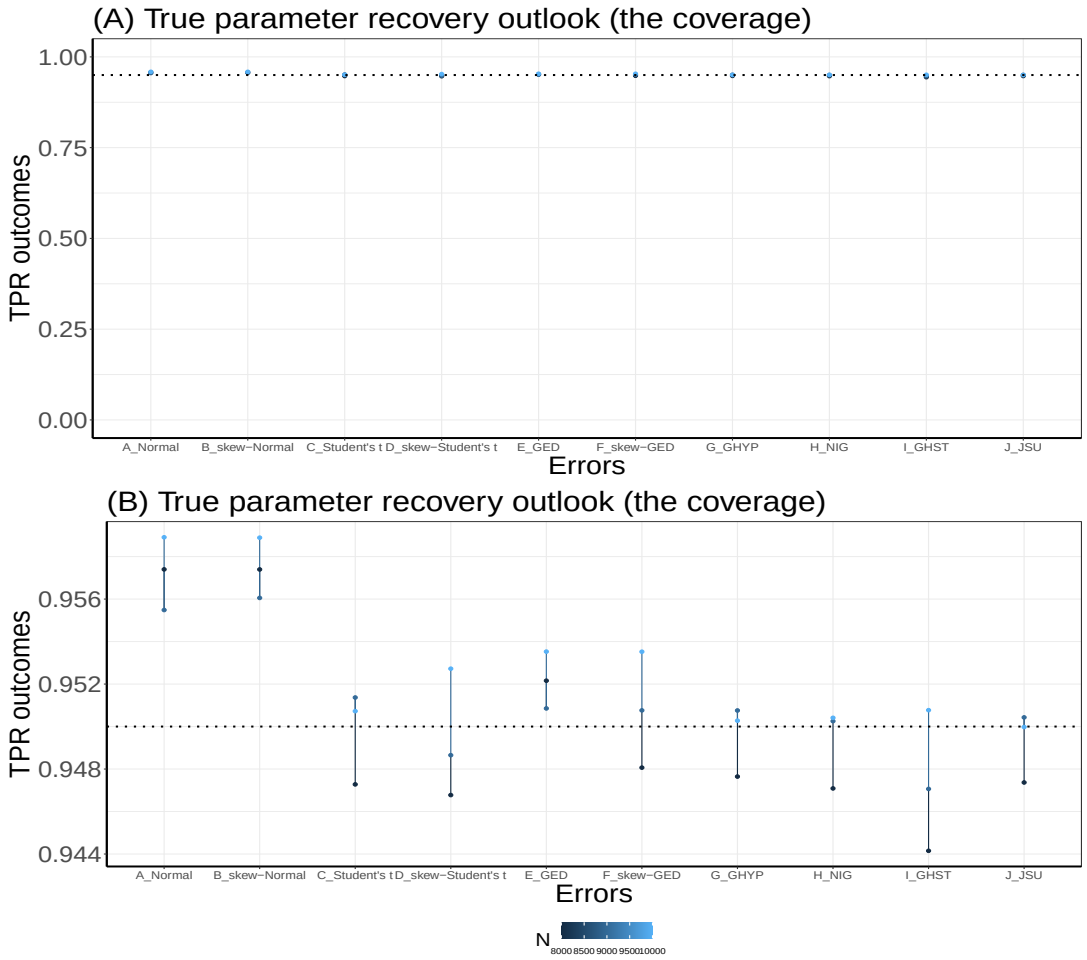

Figure 4.7: Panels (**A**,**B**) display the TPR outcomes, where the clustered outcomes in Panel (**A**) are clearly spread out in Panel (**B**). The dotted line is the 95% (i.e., 0.95) nominal recovery level.

2008, 2016 and 2020. The most recent as shown by the plots for 2020 was due to the global COVID-19 pandemic.

For further inspection, the figure is now separated into two panels: left and right. The left panels contain plots a, b, e, f for daily bond prices, while the right panels consist of plots c, d, g, h for the returns. The left panels reveal non-stationarity in the price index as observed in the price series plot, the density plot, quantile–quantile (QQ) plot and the box plot. On the other hand, the right panels show stationarity in the returns through the return series plot, the density plot, the QQ plot and the box plot. These summarily elucidate the non-stationarity in the SA daily bond prices and stationarity in the returns.

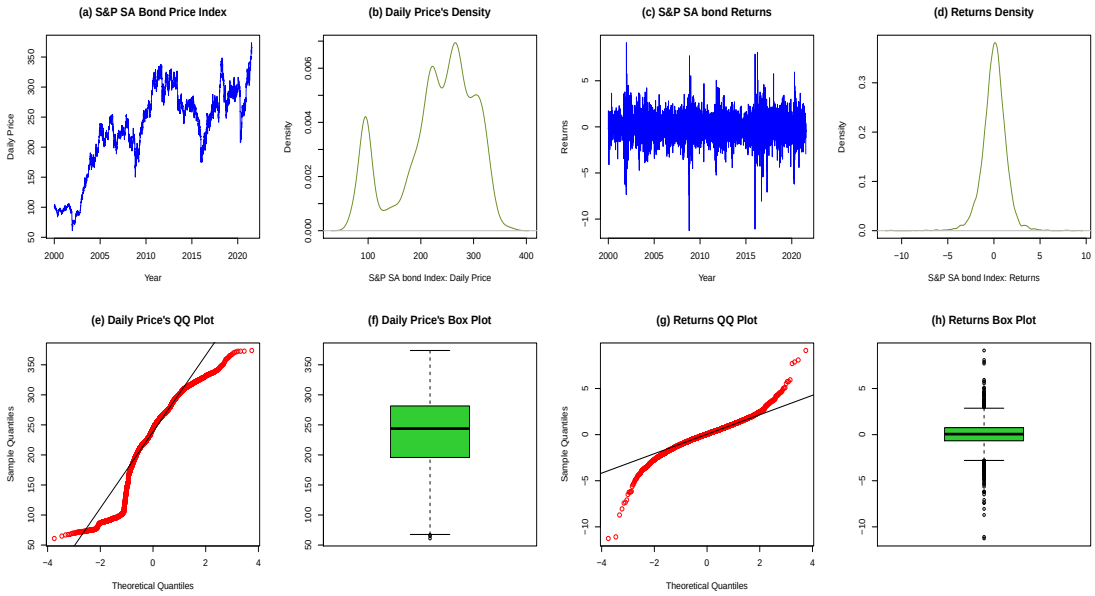

Figure 4.8: EDA of price (panels **a**,**b**,**e**,**f**) and returns (panels **c**,**d**,**g**,**h**) for SA Bond Index.

4.3.2.2 Tests for Stationarity

This study further used three stationarity tests involving the Augmented-Dickey Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) tests as described in Section 3.6 to test for stationarity in the price and return series. We start with the test for stationarity of the price index in Panel A of Table 4.3. The table shows the results of the ADF and PP unit root tests, and they indicate that the null hypothesis of non-stationarity, which implies the existence of a unit root in the price series, cannot be rejected because the p-values are greater than 0.05. Hence, the price index is non-stationary. Furthermore, the KPSS test based on the null hypothesis of stationarity is rejected because the p-value < 0.05 (see Panel A of Table 4.3). This further indicates that the price index is non-stationary.

Next, for the return series in Panel B of Table 4.3, the outcomes of the ADF and PP unit root tests indicate that the null hypothesis of non-stationarity, which points to the presence of a unit root in the returns is rejected because the p-values are less than 0.05. Hence, the return series is stationary. Also, we fail to reject the KPSS test with a null hypothesis of stationarity because the p-value > 0.05, as shown in Panel B of Table 4.3. This further implies that the return series is stationary.

4.3.2.3 Tests for Serial Correlation and Heteroscedasticity

Next, linear dependence (or serial correlation) and heteroscedasticity are filtered out by fitting ARMA-fGARCH models with each of the ten innovation distributions to the stationary return series. The ARMA(1,1) model, stated as $r_t = \phi_0 + \phi_1 r_{t-1} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t$, is found to be the most adequate among all the examined candidate ARMA(p,q) models, to remove serial correlation

Table 4.3: Tests for Stationarity of the Price Index and Return Series.

Panel A Price Index				**Panel B** Returns		
ADF Test				ADF test		
Estimate	Lag order	P-value		Estimate	Lag order	P-value
-2.5532	10	0.3441		-22.6740	10	0.0100
PP Test				PP Test		
Estimate	Lag order	P-value		Estimate	Lag order	P-value
-14.0220	10	0.3279		-5010.7	10	0.0100
KPSS Test				KPSS Test		
Estimate	Lag order	P-value		Estimate	Lag order	P-value
33.4160	10	0.0100		0.0824	10	0.1000

Note: Panel A consists of the empirical outcomes of the Augmented-Dickey Fuller (ADF) and Phillips-Perron (PP) unit root tests, and the outcomes of the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) stationarity test for the price index. Panel B contains the outcomes of the three tests for the returns.

from the SA bond market's return residuals. Table 4.4 presents the outcomes of the Weighted Ljung–Box (WLB) test (see Fisher and Gallagher, 2012 for details) for fitting the ARMA(1,1) model. The p-values of the test at lag 5 all exceed 0.05 under each error distribution. Based on this, we fail to reject the null hypothesis of "no serial correlation" in the SA bond market's returns. This means there is no evidence of autocorrelation in the return residuals.

Following the filtering of linear dependence in the return series, Engle's ARCH test (see Engle, 1982) is carried out using the Lagrange Multiplier (LM) and Portmanteau-Q (PQ) tests to check for the presence of heteroscedasticity or ARCH effects in the residuals. These tests are implemented based on the null hypothesis of homoscedasticity in the residuals of an Autoregressive Integrated Moving Average (ARIMA) model. Both tests' outcomes show highly significant p-values of 0 as shown in Figure 4.9. Hence, the null hypothesis of "no ARCH effect" in the residuals is rejected, which denotes the existence of volatility clustering. Based on this, a heteroscedastic model can be fitted to remove the ARCH effects in the series. To achieve this, the candidate robust fGARCH(u, v) models, with each of the ten error distributions, are fitted to the SA bond returns, where the fit of the parsimonious fGARCH(1,1) model as shown in Equation (4.17) is found to be the most suitable, deduced from Equation (4.7).

$$\sigma_t^{\gamma} = \omega + \alpha_1 \sigma_{t-1}^{\gamma}(|z_{t-1} - \lambda_{21}| - \lambda_{11}\{z_{t-1} - \lambda_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma}. \tag{4.17}$$

4.3.2.4 Residual Diagnostic Test

After fitting the fGARCH model to the returns, we performed residual diagnostics using the weighted ARCH LM test to ascertain if ARCH effects were filtered out. The diagnostic outcomes show that the p-value of the "ARCH LM statistic (7)" at lag 7 in Table 4.4 is greater than 5% under each of the ten innovation distributions. Hence, this indicates that heteroscedasticity is filtered out since we fail to reject the null hypothesis of "no ARCH effect" in the residuals. These outcomes show that the variance equation is well specified.

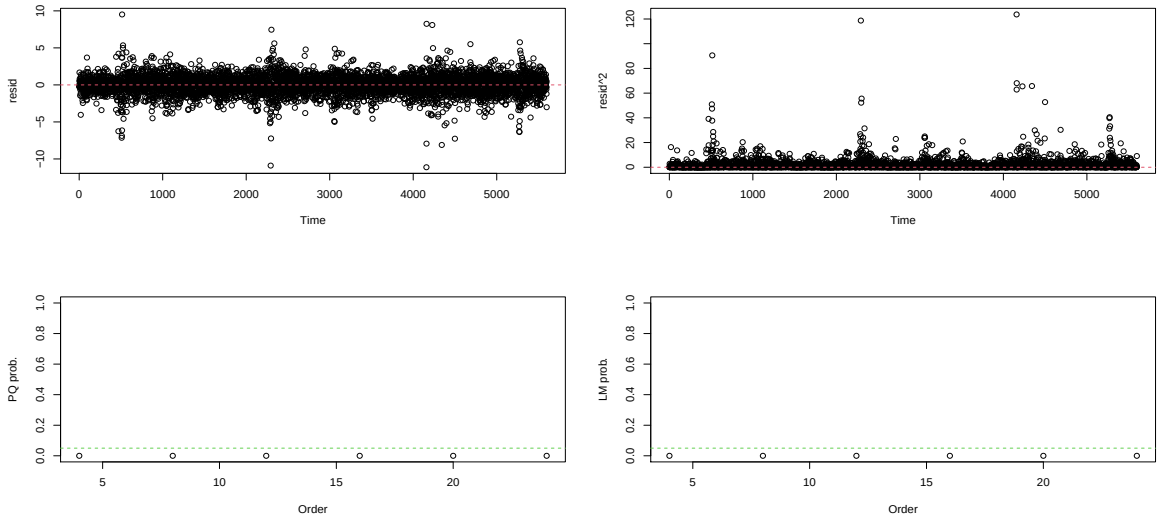

Figure 4.9: ARCH Portmanteau Q and Lagrange Multiplier tests.

4.3.2.5 Selection of the Most Suitable Error Distribution

Next, selection of the most suitable assumed error distribution to describe the market's returns, when fitted with the fGARCH model for volatility persistence estimation, is obtained from Table 4.4. It is observed from the table that all, but two, of the fGARCH volatility parameters (ω, α, β, λ_{11}, λ_{21} and γ) under the ten innovation assumptions are statistically significant at 1% level. This means that these parameters are actively needed in the model. The two exceptions are the insignificant ω for the Normal and skew-Normal, and the parameter λ_{21} that are mostly not significant or barely significant. The strongly significant λ_{11} indicates the dominance of asymmetric large shocks in the return series.

Comparisons of the error distributions are carried out using the log-likelihood and four information criteria that include the Akaike information criterion (AIC), Bayesian information

Table 4.4: ARMA(1,1)-fGARCH(1,1) models' empirical outcomes on SA Bond return data.

	Panel A Normal	Panel B skew-Normal	Panel C Student's t	Panel D skew-Student's t	Panel E GED
$\hat{\phi_0}$	0.0164^{***}	0.0078^{*}	0.0387^{*}	0.0177	0.0378^{**}
$\hat{\phi_1}$	-0.0328	-0.0760	-0.1551	-0.1884^{*}	-0.2165^{*}
$\hat{\varphi_1}$	0.1030	0.1386	0.2171^{***}	0.2459^{*}	0.2759^{*}
$\hat{\omega}$	0.0323	0.0278	0.0297^{*}	0.0270^{*}	0.0311^{*}
$\hat{\alpha}$	0.0701^{*}	0.0670^{*}	0.0690^{*}	0.0661^{*}	0.0700^{*}
$\hat{\beta}$	0.9093^{*}	0.9170^{*}	0.9188^{*}	0.9236^{*}	0.9137^{*}
$\hat{\lambda}_{11}$	0.2504^{*}	0.2344^{*}	0.3499^{*}	0.3445^{*}	0.2879^{*}
$\hat{\lambda}_{21}$	0.2245	0.2209^{***}	0.0729	0.0943	0.1445^{*}
$\hat{\gamma} = \hat{\delta}$	1.4550^{*}	1.4233^{*}	1.2362^{*}	1.2058^{*}	1.3436^{*}
Persistence ($\hat{P}$)	0.9794	0.9825	0.9764	0.9792	0.9762
WLB(5)	0.3227	0.8383	0.9103	1.6060	1.3361
p-value(5)	(1.0000)	(1.0000)	(1.0000)	(0.9955)	(0.9995)
ARCH LM statistic(7)	3.0979	3.1854	3.8897	4.1266	3.4264
p-value(7)	(0.4953)	(0.4793)	(0.3627)	(0.3287)	(0.4369)
AP − GoF	87.2	64.56	42.32	18.48	53.68
p-value	(0.0000)	(0.0000)	(0.0016)	(0.4908)	(0.0000)
Log − likelihood	-8909.189	-8886.553	-8803.012	-8790.528	-8825.745
AIC	3.1862	3.1785	3.1486	3.1445	3.1568
BIC	3.1969	3.1903	3.1605	3.1576	3.1686
SIC	3.1862	3.1785	3.1486	3.1445	3.1567
HQIC	3.1899	3.1826	3.1528	3.1491	3.1609
Run − time(seconds)	4.3245	6.6636	7.6463	11.9177	9.1407
	Panel F skew-GED	**Panel G GHYP**	**Panel H NIG**	**Panel I GHST**	**Panel J JSU**
$\hat{\phi_0}$	0.0157	0.0156	0.0155	-0.0062	0.0159
$\hat{\phi_1}$	-0.2649^{**}	-0.1890^{*}	-0.2117	-0.2302	-0.2006^{*}
$\hat{\varphi_1}$	0.3183^{*}	0.2467^{*}	0.2684	0.2833^{***}	0.2578^{*}
$\hat{\omega}$	0.0273^{*}	0.0267^{*}	0.0261^{*}	0.0251^{*}	0.0265^{*}
$\hat{\alpha}$	0.0665^{*}	0.0661^{*}	0.0657^{*}	0.0650^{*}	0.0658^{*}
$\hat{\beta}$	0.9206^{*}	0.9241^{*}	0.9246^{*}	0.9284^{*}	0.9243^{*}
$\hat{\lambda}_{11}$	0.2823^{*}	0.3370^{*}	0.3341^{*}	0.3202^{*}	0.3378^{*}
$\hat{\lambda}_{21}$	0.1592^{**}	0.0942	0.0964	0.1163^{**}	0.0949
$\hat{\gamma} = \hat{\delta}$	1.3048^{*}	1.2086^{*}	1.2171^{*}	1.1942^{*}	1.2102^{*}
Persistence($\hat{\mathbf{P}}$)	0.9797	0.9795	0.9800	0.9826	0.9796
WLB(5)	2.5350	1.5990	1.8260	2.5920	1.7170
p-value(5)	(0.7599)	(0.9957)	(0.9822)	(0.7277)	(0.9906)
ARCH LM statistic(7)	3.6331	4.0705	4.0249	4.2354	4.0750
p-value(7)	(0.4026)	(0.3365)	(0.3430)	(0.3139)	(0.3359)
AP − GoF	46.18	17.01	22.23	29.37	21.66
p-value	(0.0005)	(0.5890)	(0.2730)	(0.0604)	(0.3013)
Log − likelihood	-8810.111	-8790.079	-8793.107	-8800.387	-8791.112
AIC	3.1515	3.1447	3.1454	3.1480	3.1447
BIC	3.1646	3.1589	3.1585	3.1611	3.1578
SIC	3.1515	3.1447	3.1454	3.1480	3.1447
HQIC	3.1561	3.1497	3.1500	3.1526	3.1493
Run − time(seconds)	19.0058	49.9461	20.7803	16.8525	10.6434

Note: The "$*$", "$**$" and "$***$" are 1%, 5% and 10% significance levels, respectively. The p-values at 5% significance levels are given in the round brackets, while "(5)" and "(7)" are lags 5 and 7, respectively. The AP-GoF (for group 20) is the Adjusted Pearson Goodness-of-Fit Test, and WLB denotes the Weighted Ljung–Box test.

criterion (BIC), Hannan–Quinn information criterion (HQIC) and Shibata information criterion (SIC) (see Ghalanos, 2018 for details). The largest log-likelihood value with the smallest values of the information criteria under a given assumed innovation indicates that it is the most appropriate innovation distribution to describe the market for volatility persistence estimation.

It is observed from Table 4.4 that the values of all four information criteria are smallest under the skew-Student's t innovation distribution, with the GHYP innovation having the highest log-likelihood value. Hence, the skew-Student's t is the most suitable innovation assumption strictly based on the information criteria, while the GHYP is the most appropriate if the decision is made using the log-likelihood. The GHYP and skew-Student's t also yield better goodness of fit (GoF) outcomes when compared with the remaining eight error distributions, as shown by their large p-values in the table, which shows that they are the best fit among the ten error assumptions for the distribution of the SA bond's return residuals. Hence, the volatility persistence of the SA bond market's returns can be most suitably estimated through the ARMA(1,1)-fGARCH(1,1) model fitted with the GHYP or skew-Student's t assumed error distribution. These empirical results are consistent with the Monte Carlo simulation outcomes. The estimated volatility persistence under these most suitable error distributions are 0.9795 for the GHYP and 0.9792 for skew-Student's t. Hence, this indicates that the volatility of the SA bond market's returns is considerably highly persistent.

This study also checked the empirical outcomes of fitting the less omnibus apARCH(1,1) model to the bond return data and we arrived at the same results (of skew-Student's t through information criteria and GHYP via log-likelihood) obtained by the fGARCH(1,1) model (see Table 4.5). The table only shows the outcomes of the log-likelihood and information criteria for brevity.

For the run-time, it is observed that the skew-Student's t is about four (approximately 4.2) times faster than the GHYP for both simulation and empirical modelling. That is, it takes the GHYP about four times the computational time it takes the skew-Student's t to run the same process. Since the empirical and simulation run-times are approximately the same, we only present the empirical run-times for the ten innovations in Table 4.4 for brevity. For both simulation and empirical runs, the GHYP has the highest runtime among the ten innovations, followed by the NIG, while the Normal has the least.

From the outputs of the ARMA(1,1)-fGARCH(1,1) model in Table 4.4, the mean and variance (from the conditional standard deviation's Box–Cox transformation in Section 4.2.2) equations of the model fitted with each of the GHYP and skew-Student's t are stated as:

Table 4.5: ARMA(1,1)-apARCH(1,1) models' empirical outcomes on SA Bond data.

	Panel A	**Panel B**	**Panel C**	**Panel D**	**Panel E**
	Normal	**skew-Normal**	**Student's t**	**skew-Student's t**	**GED**
Log-likelihood	-8910.136	-8887.475	-8803.200	-8790.782	-8826.007
AIC	3.1862	3.1784	3.1483	3.1443	3.1565
BIC	3.1957	3.1891	3.1590	3.1561	3.1671
SIC	3.1862	3.1784	3.1483	3.1443	3.1565
HQIC	3.1895	3.1822	3.1521	3.1484	3.1602
	Panel F	**Panel G**	**Panel H**	**Panel I**	**Panel J**
	skew-GED	**GHYP**	**NIG**	**GHST**	**JSU**
Log-likelihood	-8810.472	-8790.315	-8793.329	-8802.039	-8791.340
AIC	3.1513	3.1444	3.1452	3.1483	3.1445
BIC	3.1631	3.1575	3.1570	3.1601	3.1563
SIC	3.1513	3.1444	3.1452	3.1483	3.1445
HQIC	3.1554	3.1490	3.1493	3.1524	3.1486

$$
\begin{aligned}
\text{With GHYP}: \quad r_t &= \mu_t + \varepsilon_t \\
&= \phi_0 + \phi_1 r_{t-1} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \\
&= 0.0156 - 0.1890 r_{t-1} + 0.2467 \varepsilon_{t-1} + \varepsilon_t \\
\sigma_t^{\gamma} &= \omega + \alpha_1 \sigma_{t-1}^{\gamma} (|z_{t-1} - \lambda_{21}| - \lambda_{11}\{z_{t-1} - \lambda_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma} \\
\sigma_t^{1.2086} &= 0.0267 + 0.0661 \sigma_{t-1}^{1.2086} (|z_{t-1} - 0.0942| - \\
&\quad 0.3370\{z_{t-1} - 0.0942\})^{1.2086} + 0.9241 \sigma_{t-1}^{1.2086}
\end{aligned}
$$

$$
\begin{aligned}
\text{With skew-Student's } t: \quad r_t &= \mu_t + \varepsilon_t \\
&= \phi_0 + \phi_1 r_{t-1} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \\
&= 0.0177 - 0.1884 r_{t-1} + 0.2459 \varepsilon_{t-1} + \varepsilon_t \\
\sigma_t^{\gamma} &= \omega + \alpha_1 \sigma_{t-1}^{\gamma} (|z_{t-1} - \lambda_{21}| - \lambda_{11}\{z_{t-1} - \lambda_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma} \\
\sigma_t^{1.2058} &= 0.0270 + 0.0661 \sigma_{t-1}^{1.2058} (|z_{t-1} - 0.0943| - \\
&\quad 0.3445\{z_{t-1} - 0.0943\})^{1.2058} + 0.9236 \sigma_{t-1}^{1.2058}
\end{aligned}
$$

4.4 The Diagnostic Plots of ARMA-fGARCH Models

Figures 4.10 and 4.11 show the diagnostic densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plots of the ARMA-fGARCH models fitted with each of the most suitable skew-student's t and GHYP assumed innovation distributions. The plots of the densities of the standardised residuals confirm that each of the most appropriate error distributions is a better fit to the SA bond market's distribution of returns than a Normal distribution (Sigauke, 2016). Furthermore, the QQ plot under each assumed error in the figures displays suitable levels of linearity since the data points are well aligned on the straight lines, except at the tails (which implies extreme movement).

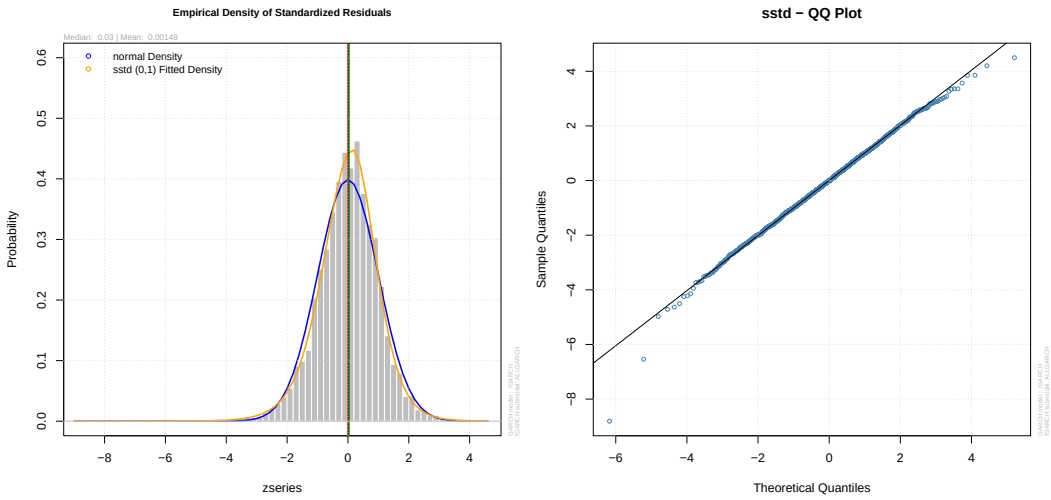

Figure 4.10: The diagnostic plot of the densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plot for the ARMA-fGARCH models fitted with the skew-student's t assumed innovation.

4.5 Discussion and Summarised Conclusions

In conclusion, it is observed that using the ugarchsim approach for the MCS illustration, the GHYP and skew-Student's t evolve as the most suitable assumed error distributions to reckon with and use with the fGARCH model for volatility persistence estimation of the SA bond returns when the underlying error distribution is unknown. These outcomes are verified empirically. The estimated persistence of volatility under these most suitable error distributions are 0.9795 for the GHYP and 0.9792 for skew-Student's t. Hence, this indicates considerably high volatility persistence in the SA bond market's returns. It is also observed that the use of the ugarchsim approach provided considerably consistent estimates (and recovery) of the true data-generating parameters.

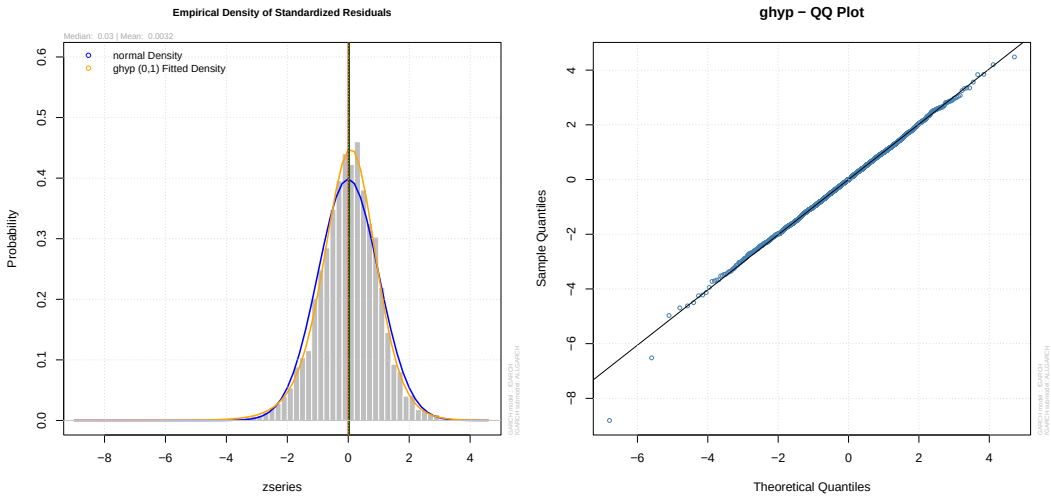

Figure 4.11: The diagnostic plot of the densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plot for the ARMA-fGARCH models fitted with the GHYP assumed innovation.

The conclusion under this section continues by providing answers to the four research questions. In this study, consistency is termed "strong" when the estimator's RMSE/SE value decreases as the sample size N increases without distortion. Otherwise, it is weak. Now, answering the questions: first, the GHYP and the skew-Student's t distributions are the most appropriate among the stated assumed error distributions from the fGARCH process simulation for the volatility persistence estimation. Second, the volatility estimator $\hat{\alpha} + \hat{\beta}$ for each of the most suitable assumed errors GHYP and skew-Student's t is strongly $\sqrt{N}$ consistent for both RMSE and SE under departures from the Normal assumption as revealed in Panels G and D of Table 4.2. Third, there are strong $\sqrt{N}$ consistency for the RMSE and SE of the fGARCH estimator $\hat{\alpha} + \hat{\beta}$ under all, but one, of the ten assumed error distributions as shown in Table 4.2. The lone exception, however, is the weak consistency in the SE of the Normal assumption. Fourth, as a proxy for the coverage of the MCS experiment, the MCS estimator $\hat{\alpha} + \hat{\beta}$ performed well in recovering the true parameter $\alpha + \beta$ through the TPR measure for the 95% nominal recovery level as revealed in Table 4.2 and Figure 4.7. The results show that the TPR outcomes are suitably close to the 95% nominal recovery level under the ten error distribution assumptions.

4.6 Conclusions

This study showcases a robust step-by-step framework for a comprehensive simulation by presenting the functionalities of the rugarch package in R for simulating and estimating time-varying parameters through the family GARCH observation-driven model. The framework

hands out an organised approach to a Monte Carlo simulation (MCS) study that involves "background (optional), defining the aim, research questions, method of implementation, and summarised conclusion". The method of implementation is a workflow that includes writing the code, setting the seed, setting the true parameters a priori, data-generation process, and performance assessment through meta-statistics.

This new, easy-to-understand framework is illustrated using financial return data; hence, users can easily use it for effective MCS studies. With the ugarchsim simulation approach involved in the modelling, the implementation method is clearly explained with relevant details. Key observations are identified, and novel findings brought to light. The framework also outlays clear coding guidelines for data generation using the package, since data generation is without a doubt an integral part of MCS studies. The key observations and novel findings in this study include, first, it is shown in the experiment that as the sample size N becomes larger, the consistency and efficiency properties of an estimator in a Monte Carlo process are generally not affected by the pattern or arrangement of the seed values, but this may depend on the quality of the model used. Hence, regardless of the arrangement of the seed values, the efficiency and consistency of an estimator generally remain the same as N tends to infinity.

Second, it is investigated and revealed in this study that the outcomes of the GARCH MCS experiments are the same regardless of whether the simulation or data generating process is run once or replicated multiple times. Third, this study derived a "true parameter recovery (TPR)" measure as a proxy for the coverage of the MCS experiment. This new (original) measure is flexible to apply and can henceforth be used by upcoming researchers to determine the level of recovery of the true parameter value by the MCS estimates. It is also observed that the volatility estimator of the used fGARCH model displays considerably strong $\sqrt{N}$ consistency.

Lastly, the outcomes of the illustrative study show that the GHYP and skew-Student's t error distributions are the most suitable among the ten assumed innovations to describe the SA bond returns for volatility persistence estimation. The fit of these two error distribution assumptions with the fGARCH model revealed considerably high volatility persistence in the returns. On a wider scale, since volatility is a practical measure of risk, the fit of the GHYP and skew-Student's t errors with a specification of the fGARCH (or, apARCH) model for a robust volatility modelling may benefit financial institutions and markets by enhancing the accuracy of their risk estimations. This could potentially lead to a significant reduction in asset losses. Moreover, it is documented that shocks with a permanent influence on the variance will have a greater effect on price than those with temporary influence (Arago and Fernandez-Izquierdo, 2003). Hence, through the fit of these innovations, policymakers and other financial market participants may benefit from a better understanding of the effects of shocks to future volatility, especially from knowing whether the effects of the shocks are transient or (highly) persistent. It is anticipated that researchers will leverage this study's novel findings and robust design for improved simulation studies in finance and other sectors.

4.6.1 Limitations in the Study

Three limitations or challenges are noticed in this study. The first is on how to obtain a sufficient sample size that can generate accurate outcomes. To tackle this using the illustrative example, the process involves testing a selected number of sample sizes, with each used in turn, until a pattern of efficiency and/or consistency starts to evolve under the stated error distributions. The most efficient error distribution in terms of the given performance measure under a particular sample size is carefully noticed. If a set of sample sizes yields the same efficiency outcomes, the outcome with the best consistency among the set can be used for a final decision on sample size determination. This is a guide to obtaining the required sample size/s.

The second is running time. It is observed that a large simulated dataset may sometimes be needed to obtain accurate computations and this may be carried out at the cost of a large computational (or running) time, depending on the model used. This can be very demanding, especially when dealing with different stages of large sample sizes. Third, since the rugarch package does not make provision for calculating the coverage probability, this study derived a proxy for the coverage using the TPR measure, and it is observed that the MCS estimates considerably recover the true parameters.

Chapter 5

Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model

RICHARD T. A. SAMUEL[1,*], CHARLES CHIMEDZA[1] and CASTON SIGAUKE[2]

[1] School of Statistics and Actuarial Science, University of the Witwatersrand

[2] Department of Mathematical and Computational Sciences, University of Venda

Published in: *Statistics, Optimization & Information Computing*, 2024, 12: 1123-1148.

Statement of Contributions of Joint Authorship

Richard T. A. Samuel (Candidate)
Conducted the research, writing and compilation of manuscript;

Charles Chimedza (Supervisor) & Caston Sigauke (Co-Supervisor)
Supervised, edited and coauthors of the manuscript.

This chapter is a copy of the journal paper referred to above.

ABSTRACT

In econometrics and finance, volatility modelling has long been a specialised field for addressing a variety of issues that pertain to the risks and uncertainties of an asset. However, volatil-

ity modelling for risk management is highly dependent on the underlying error distribution. Hence, this study presents a Monte Carlo simulation (MCS) structure for selecting an optimal or the most adequate error distribution that is relevant for modelling the persistence of volatility through the Generalised Autoregressive Score (GAS) model. The structure describes an organised approach to the MCS experiment that includes "background of the study (optional), defining the aim of the study, specifying the research questions, method of implementation, and summarised conclusion". The method of implementation is a process that consists of writing the simulation code, setting the seed, setting the true parameter a priori, data generation, and performance evaluation through meta-statistics. The simulations with empirical evaluations are implemented using the SA Bond market index data. Among the findings, the study used both fat-tails and $\sqrt{N}$ consistency experiments to show that the GAS model with a lower unconditional shape parameter value ($\hat{\nu}^* = 4.1$) can generate a dataset that adequately reflects the behaviour of financial time series data, relevant for volatility modelling. This dynamic structure is intended to help interested users on MCS experiments utilising the GAS model for reliable volatility persistence calculations in finance and other areas.

5.1 Introduction

In finance and econometrics, the modelling of volatility has long been used to address a variety of issues concerning the risks and uncertainties of an asset (Diebold and Lopez, 1995; Satchell and Knight, 2011). This study follows the simulation design steps of our earlier study on volatility persistence estimation through the family Generalised Autoregressive Conditional Heteroscedasticity (fGARCH) model created by Hentschel (1995) (see Samuel et al., 2023). That is, the family GARCH model was used in our earlier study to begin a Monte Carlo simulation (MCS) design to determine appropriate innovations for volatility persistence estimation (see Samuel et al., 2023). The Generalised Autoregressive Score (GAS) model is used in this new study to extend the robust design. As a result, this study identifies a vital simulation structure that is pertinent for selecting an optimal (or the most adequate) conditional error distribution for volatility persistence modelling using the GAS model.

The GAS is a dynamic model that is used for time-varying parameter modelling, much like the GARCH model (see Ardia et al., 2019). It is an observation-driven model that was developed to model distributional asymmetries and large changes, such as shifts or jumps, in financial time series (Haddad et al., 2023). The model has been applied by Oh and Patton (2018) for modelling systemic risk. Other broad applications are in market risk (Harvey and Sucarrat, 2014), spatial econometrics (Blasques et al., 2014; Catania and Billé, 2017), credit risk analysis (Creal et al., 2014), high-frequency data (Gorgi et al., 2019; Opschoor et al., 2018), and dependence modelling (Harvey and Thiele, 2016; Janus et al., 2014), among others (see Ardia et al., 2019; Haddad et al., 2023). Hence, this study's main contribution is to show

how the GAS model, using the GAS package (Ardia et al., 2019), can be effectively used to determine a suitable conditional distribution that is relevant for volatility persistence estimation through MC simulation with empirical verification. Moreover, simulation through any model largely revolves around generating an appropriate dataset that can produce a reasonable output approximation. Thus, one further contribution of this study is to investigate how to use the GAS model to generate (or simulate) a dataset that can adequately reflect the behaviour of financial time series data that is relevant for volatility modelling.

Volatility is related to variation in the price of an asset over time. In finance, the variance is used as a proxy for volatility (Beg and Anwar, 2014; Covarrubias, 2004). However, emphasis is usually on conditional variance because time-varying volatility clustering exists in most financial time series (Beg and Anwar, 2014). Volatility persistence is a measure of the speed of the decay of volatility following a shock. The level of persistence has crucial economic implications that stem from the impact it has on the predictability of future value (Arago and Fernandez-Izquierdo, 2003). Risk managers, policymakers, portfolio managers, financial market participants, and other financial activities can derive benefits from a clear understanding of the effect of shocks on future volatility, especially with information on whether the shocks are short-lived or (highly) persistent.

MCS studies are computer-based experiments that involve the use of known probability distributions for creating data by pseudo-random sampling. The data may be simulated through parametric models or repeated resampling (Morris et al., 2019). MCS is used for approximating real-life scenarios, where it uses random quantities (or a probabilistic approach) to provide deterministic outcomes (Johansen, 2010). Since its introduction by three scientists in the field of physics (Giovanni, 2020), MCS has become an increasingly applied technique in a wide range of fields including finance, supply chain, engineering, science (Kenton, 2021), and in all portfolio and investment types (Agarwal, 2022). One of the reasons for its wide applications is that it is based on the concept of the central limit theorem where the distribution of its estimates follows (or converges to) the Normal distribution (Kwiatkowski, 2022).

This study used the daily Standard & Poor South African sovereign bond index (S&P SA bond index, henceforth). They are the bond market's data from Datastream (2021) for the period January 4th, 2000, to June 17th, 2021, with 5598 data points. The period under study includes the COVID-19 pandemic crisis period that began in 2020. The rest of this paper is structured as follows: Section 5.2 reviews the theories underlying the GAS model, the true parameter recovery (TPR) measure, and the description of the simulation design. Section 5.3 shows how the simulation structure is illustrated using financial bond returns data, with empirical verification. Section 5.4 discusses the main findings and Section 5.5 presents concluding remarks on the novelty of the study.

5.2 Methods

5.2.1 The GAS Model

Volatility persistence can be effectively modelled through a Score Driven (SD) model known as the Generalised Autoregressive Score (GAS) proposed by Harvey (2013) and Creal et al. (2013) (see Ardia et al., 2019, 2018; Catania et al., 2020). The model applies the score of the conditional density function to determine the time variation in the parameters. The score functions are robust to outliers because they discount extreme values by reducing the number of extreme observations (Alanya-Beltran, 2022; Harvey, 2013; Owusu Junior and Alagidede, 2020), and the model is suitable for modelling fat-tailed and skewed time series data like financial returns (Harvey, 2013; Opschoor et al., 2018). In addition, extensions to other time series dynamics, like asymmetry and long memory behaviour, are possible through the GAS process. Moreover, maximum likelihood estimation (MLE) through the model is easy to implement (Ardia et al., 2019).

Creal et al. (2011, 2013, 2014) showed that the robust GAS models incorporate many familiar observation-driven models that include the Autoregressive Conditional Heteroscedasticity (ARCH) model (Engle, 1982), the simple (standard) GARCH model (Bollerslev, 1986), the Exponential GARCH (EGARCH) model (Nelson, 1991), the Autoregressive Conditional Duration (ACD) model (Engle and Russell, 1998), the Autoregressive Conditional Intensity (ACI) model (Russell, 2001), the Beta-t-GARCH model (Harvey, 2013; Harvey and Chakravarty, 2008), the Autoregressive Conditional Multinomial (ACM) model (Rydberg and Shephard, 2003), and several related models. In other words, new time-varying parameter models of empirical interest are easy to formulate due to the flexibility of the GAS framework (see Blasques et al., 2014).

5.2.1.1 Specification of the GAS Model

Let $N \times 1$ vector $\boldsymbol{y}_t$ be the dependent variable of interest, $\boldsymbol{\vartheta}_t$ the vector of time-varying parameter, $\boldsymbol{x}_t$ a vector of exogenous variables (i.e., the covariates), all at time t, and $\boldsymbol{\xi}$ a vector of static parameters. Define $\boldsymbol{Y}^t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$, $\boldsymbol{\Theta}^t = \{\boldsymbol{\vartheta}_0, \boldsymbol{\vartheta}_1, \ldots, \boldsymbol{\vartheta}_t\}$, and $\boldsymbol{X}^t = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t\}$ (Creal et al., 2012, 2013). The information set that is available at time t consists of $\{\boldsymbol{\vartheta}_t, \mathcal{F}_t\}$, where

$$\mathcal{F}_t = \{\boldsymbol{Y}^{t-1}, \boldsymbol{\Theta}^{t-1}, \boldsymbol{X}^t\}, \quad \text{for } t = 1, \ldots, n. \tag{5.1}$$

It is assumed that the generation of $\boldsymbol{y}_t$ is carried out by the observation density

$$\boldsymbol{y}_t \sim p\left(\boldsymbol{y}_t | \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi}\right). \tag{5.2}$$

It is further assumed that the mechanism required to update the time-varying parameter $\boldsymbol{\vartheta}_t$ is given by the autoregressive updating equation

$$\boldsymbol{\vartheta}_{t+1} = \boldsymbol{\kappa} + \sum_{i=1}^{p} \boldsymbol{A}_i s_{t-i+1} + \sum_{j=1}^{q} \boldsymbol{B}_j \vartheta_{t-j+1}. \tag{5.3}$$

Ardia et al. (2019) presented Equation (5.3) as:

$$\boldsymbol{\vartheta}_{t+1} \equiv \boldsymbol{\kappa} + \mathbf{A}\boldsymbol{s}_t + \mathbf{B}\boldsymbol{\vartheta}_t, \tag{5.4}$$

where $\boldsymbol{\kappa}$, $\mathbf{A}$ and $\mathbf{B}$ are matrices of coefficients with appropriate dimensions and they are functions of the time-invariant parameter $\boldsymbol{\xi}$, while $\boldsymbol{s}_t$ is the scaled score and it is a suitable function of past data, $\boldsymbol{s}_t = \boldsymbol{s}_t(\boldsymbol{y}_t, \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi})$ (see Creal et al., 2012, 2013). The $\boldsymbol{s}_t$ indicates the direction to update the vector of parameters from $\boldsymbol{\vartheta}_t$ to $\boldsymbol{\vartheta}_{t+1}$. Vector $\boldsymbol{\kappa}$ is used to control the level of the process $\boldsymbol{\vartheta}_t$, while matrix $\mathbf{A}$ is used to determine (or control) the impact of $\boldsymbol{s}_t$ on $\boldsymbol{\vartheta}_{t+1}$, and matrix $\mathbf{B}$ is used to determine the persistence of the process (Ardia et al., 2019; Koopman et al., 2017). To implement the GAS models, matrices $\mathbf{A}$ and $\mathbf{B}$ are constrained to exist as diagonals (see Ardia et al., 2019; Oh and Patton, 2018). For instance, in a GAS model with Student's t conditional distribution, $\mathbf{A} \equiv \text{diag}(a_\mu, a_\phi, a_\nu)$ and $\mathbf{B} \equiv \text{diag}(b_\mu, b_\phi, b_\nu)$, where μ, ϕ and ν are location, scale and shape parameters, respectively. Hence, b_μ denotes the persistence of the conditional mean (location), while b_ϕ refers to the persistence of the conditional variance (or scale) (see Ardia et al., 2019).

The approach used by the GAS is based on the observation density in Equation (5.2) for a given parameter $\boldsymbol{\vartheta}_t$. When an $\boldsymbol{y}_t$ observation is realised, the time-varying $\boldsymbol{\vartheta}_t$ to the next period $t+1$ can be updated using Equation (5.3) with

$$\boldsymbol{s}_t = \boldsymbol{S}_t \cdot \boldsymbol{\nabla}_t, \qquad \boldsymbol{\nabla}_t = \frac{\partial \ln p\left(\boldsymbol{y}_t | \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi}\right)}{\partial \boldsymbol{\vartheta}_t}, \qquad \boldsymbol{S}_t = S\left(t, \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\xi}\right), \tag{5.5}$$

where $S(\cdot)$ represents a matrix function, $\boldsymbol{\nabla}_t$ is the score of Equation (5.2) evaluated at $\boldsymbol{\vartheta}_t$, ln is the natural logarithm, and $\boldsymbol{S}_t$ denotes the scaling matrix (see Ardia et al., 2019; Creal et al., 2012, 2013). The score of the GAS model depends on the complete density structure (Creal et al., 2012, 2013). Given that the driving mechanism in Equation (5.3) depends on the scaled score vector in Equation (5.5), the GAS model with orders p and q can be defined by Equations (5.2), (5.3) and (5.5). This can be referred to as the GAS(p,q) model, where the orders p and q are typically taken as $p = q = 1$ (see Creal et al., 2012, 2013). However, readers can refer to Creal et al. (2013) for details on including more lags in the GAS process (Blasques et al., 2014).

To account for the variance of $\boldsymbol{\nabla}_t$, Creal et al. (2013) suggested setting $\boldsymbol{S}_t$ to the inverse of the information matrix ($\boldsymbol{\mathcal{I}}$) to a power $\gamma > 0$ of $\boldsymbol{\vartheta}_t$ (Ardia et al., 2019). Precisely,

$$\boldsymbol{S}_t = \boldsymbol{\mathcal{I}}_{t|t-1}^{-\gamma}, \qquad \boldsymbol{\mathcal{I}}_{t|t-1} = \mathrm{E}_{t-1}\left[\boldsymbol{\nabla}_t \boldsymbol{\nabla}_t^\top\right], \tag{5.6}$$

where the expectation E_{t-1} is taken with respect to the conditional distribution of $\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}$. The parameter γ usually takes 0, $\frac{1}{2}$, or 1 value (Ardia et al., 2019), but other choices of $\boldsymbol{S}_t$ are also possible (see Blasques et al., 2014). If $\gamma = 0$, $\boldsymbol{S}_t$ becomes an identity matrix ($\mathbf{I}$), which implies there is no scaling. If $\gamma = \frac{1}{2}$, then the conditional score $\boldsymbol{\nabla}_t$ is pre-multiplied by the square root of its covariance matrix $\boldsymbol{\mathfrak{I}}_t$, but if $\gamma = 1$, the conditional score $\boldsymbol{\nabla}_t$ is pre-multiplied by the inverse of its covariance matrix $\boldsymbol{\mathfrak{I}}_t$. The scaled score $\boldsymbol{s}_t$ is a martingale difference with respect to the distribution of $\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}$, i.e., $\mathrm{E}_{t-1}[\boldsymbol{s}_t] = 0$ for all t (see Ardia et al., 2019).

The GAS framework incorporates many familiar observation-driven models as a special case for an appropriate choice of the scaling matrix $\boldsymbol{S}_t$ (Creal et al., 2013). That is, various classes of observation-driven models can be obtained with different choices of the scaled score $\boldsymbol{s}_t$, hence these types of models are referred to as score-driven models (Creal et al., 2012, 2013; Koopman et al., 2017). In other words, Equations (5.2), (5.3) and (5.5) encompass a large set of common time series models for different choices of $\boldsymbol{s}_t$ (Blasques et al., 2014). For more details on the GAS models, see Ardia et al. (2019); Blasques et al. (2017); Catania et al. (2020); Creal et al. (2012, 2011, 2013, 2014); Harvey (2013); Koopman et al. (2017).

5.2.2 The True Parameter Recovery Measure

The true parameter recovery (TPR) measure was developed by Samuel et al. (2023) as a means of measuring the performance of the MCS estimator in recovering the true parameter. It is used as a proxy for the coverage of the MCS experiment to calculate the level of recovery of the true data-generating parameter by the MCS estimator. It is stated as:

$$\text{TPR} = \left(K - \left[\frac{(\vartheta - \hat{\vartheta})}{\vartheta} \times K \right] \right) \%, \tag{5.7}$$

where ϑ and $\hat{\vartheta}$ are the true parameter and estimator, respectively, while $K = 0, 1, 2, \ldots, 100$ is the nominal recovery level. A TPR outcome of 99% or 95%, for instance, denotes the recovery of 99% or 95% of the true parameter by the MCS estimator. The estimator $\hat{\vartheta}$ can recover the true parameter ϑ completely when $\hat{\vartheta} = \vartheta$, where $\vartheta > 0$. Hence, the full recovery of the true parameter occurs where the TPR estimate is equal to the specified nominal recovery level K, that is, TPR = $K\%$.

5.2.3 The Design of the Simulation

The design of the simulation structure involves "(optionally) stating the background of the study, defining the aim of the study, specifying the research questions, method of implementation, and summarised conclusion". The method of implementation is a process that consists of writing the simulation code, setting the seed, setting the true parameter a priori, data generation, and performance assessment through meta-statistics. These steps, as summarised by the

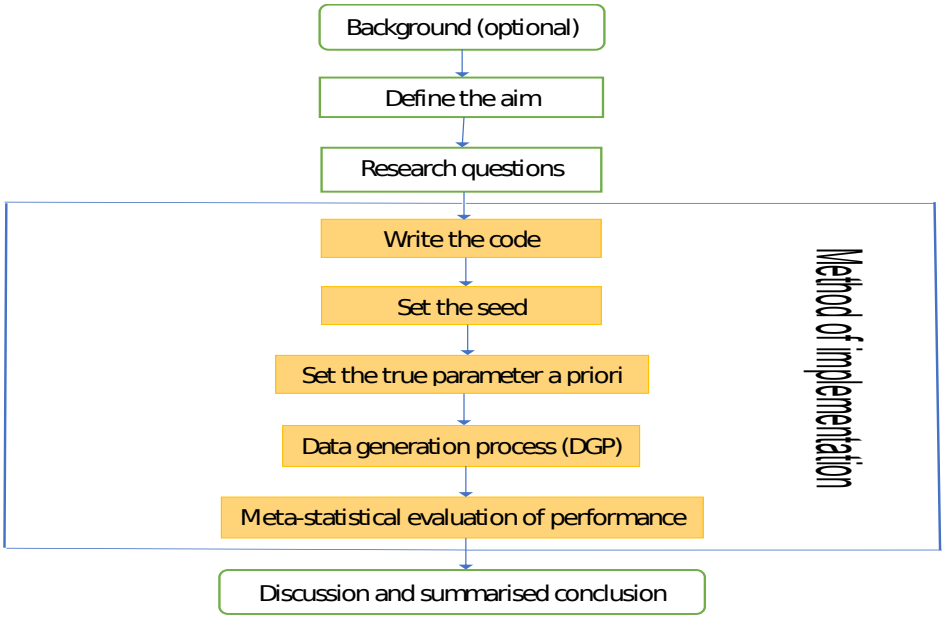

Figure 5.1: Flowchart on simulation design to select an optimal error distribution.

flowchart (Samuel et al., 2023) in Figure 5.1, are relevant for a successful simulation to select an optimal assumed innovation distribution using the GAS model. Each step of the structure is detailed as follows:

5.2.3.1 Aim of the Simulation Study

After optionally stating the background of the study, the next step in the design is to define the aim of the study, and it should be concisely stated to communicate the purpose of the study to readers. The focus of many MCS studies is generally on the estimators' abilities to recover the true parameters ϑ (Chalmers, 2019), such that $\mathrm{E}(\hat{\vartheta}) = \vartheta$ for unbiasedness, standard error (SE) or root mean square error (RMSE) tends to zero as the sample size $N \to \infty$ for good precision or efficiency of the true parameter's estimator, and $\hat{\vartheta} \to \vartheta$ as $N \to \infty$ for consistency. Hence, the study's aim could be about those features, like the consistency, bias or unbiasedness, precision or efficiency of the estimator. The aim may also focus on comparisons of multiple operations, like comparing the efficiency of various conditional distributions to obtain the most efficient, or proposing a modified version of an existing model.

5.2.3.2 State the Research Questions

Next, relevant questions that point to the objectives of the simulation should be specified and addressed in the study. The complexities of some statistical research questions may be better resolved through simulation approaches. Simulation provides a robust process for responding to many methodological and theoretical questions, and it can proffer a flexible structure for responding to certain questions that are relevant to the quest of the researchers (Hallgren, 2013).

5.2.3.3 Method of Implementation

We carry out the simulation and empirical studies of this paper in R software, version 4.0.3, with RStudio version 2023.03.1+446, using the GAS (Ardia et al., 2019; Catania et al., 2020), tidyverse (Wickham et al., 2019), patchwork (Pedersen, 2020), SimDesign (Chalmers and Adkins, 2020), rugarch (Ghalanos, 2018, 2022), forecast (Hyndman and Khandakar, 2008), and aTSA (Qiu, 2015) packages.

The method of implementing the MCS is described as follows:

- Write the simulation code: Execution of a suitable simulation experiment that reflects a real-life situation can be computationally intensive and very demanding, hence readable computer code with the right syntax must be ensured. MCS code must be well organised to avoid difficulties during debugging. Code must be efficiently and flexibly written and well arranged for easy readability.

- Set the seed: If a seed is not set, the simulation code will generate a different sequence of random numbers whenever it is run (Danielsson, 2011). A set seed is used to initialise the random number generator (Ghalanos, 2022) and to ensure reproducibility, where the same output is obtained for different runs of the simulation process (Foote, 2018). It is better to use the same seed values throughout the process (Morris et al., 2019), and this (seed) needs to be set only once, for each simulation, at the start of the simulation session (Ghalanos, 2022; Morris et al., 2019). The seed is essentially used for reproducibility. Samuel et al. (2023) used the GARCH model to show that as the sample size N increases, the arrangement (or pattern) of the seed values generally does not influence the efficiency and consistency of an estimator. This however may depend on the quality of the model used.

- Next, the true parameter value of the true sampling distribution (or true model[1]) is set a priori (Koopman et al., 2017; Mooney, 1997).

- Following this, simulated return observations are generated using the true sampling distribution or the true model given some sets of (or different sets of) parameters. Simulated data generation through the GAS model is carried out using the R GAS package. The generation of random data through this package can be implemented using either of two approaches (Ardia et al., 2019; Catania et al., 2020). The first approach is to generate the data directly from an estimated "uGASFit" object using the UniGASSim (univariate GAS simulation) function for the simulated random data. In this study, we use "fit" as the uGASFit object. The second approach involves the full specification of a GAS model that includes the selection of the conditional error distribution of the time series process,

[1]The true model is described as the data-generating model fitted with the true error distribution (see Feng and Shi, 2017).

and specifying the static parameter $\xi = (\boldsymbol{\kappa}, \mathbf{A}, \mathbf{B})$ that controls the time variation in $\boldsymbol{\vartheta}_t$, as described in Section 5.2.1. In this latter approach, the fit object will be specified as "fit = NULL" in the UniGASSim function.

The data-generating (or simulation) process can be replicated multiple times or run only once. However, data generation through the GAS package is generally designed to run only once. In their study through the autoregressive process, Samuel et al. (2023) showed that the outcomes of MCS experiments using the (family) GARCH model are the same for datasets simulated with the same seed value, regardless of whether the simulation or data-generating process is run once or replicated multiple times. Hence, the outcome of a single run is the same as the average of multiple runs or replications. Their findings further revealed that in any multiple generated (or simulated) datasets in the family GARCH model, each simulated series is distinct in randomness and shape, and it is different from every other series in the datasets.

- Next, the generated data are analysed and the estimates from the analysis are assessed using classic methods through meta-statistics to obtain relevant information about the estimator. Meta-statistics (see Chalmers and Adkins, 2020) are metrics or performance measures used for evaluating the modelling outcomes by judging the closeness between the true parameter and its estimate. The root mean square error (RMSE), bias, standard error (SE), and relative efficiency (RE) as described below are part of the familiar meta-statistical summaries used in the literature. Readers can refer to Chalmers and Adkins (2020); Morris et al. (2019); Sigal and Chalmers (2016) for more meta-statistics.

Bias

Bias is defined as the mean difference between the true parameter and its estimate (Feng and Shi, 2017). Its optimal value is 0 (Harwell, 2018; Sigal and Chalmers, 2016). Bias with a negative (positive) value indicates that the true parameter value is under-estimated (over-estimated). However, the nearer the estimator is to zero (in absolute values), the better it is. Bias can be stated in Monte Carlo simulations (see Chalmers, 2019) as:

$$\text{bias} = \frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \vartheta), \tag{5.8}$$

where $\hat{\vartheta}$ denotes the sample estimate for the ith dataset, ϑ is the true parameter, and L is the number of replications (Sigal and Chalmers, 2016).

Standard Error

The standard error (SE), as presented (see Chalmers, 2019; Yuan et al., 2015) in Equation (5.9), is used to assess the sampling variability in the estimation. It is a measure of the precision or efficiency of the true parameter's estimator. The lower the sampling variability, the more efficient or precise the true parameter's estimator (see Morris et al., 2019). As the sample size increases, the sampling variability decreases (Sigal and Chalmers, 2016).

$$\text{SE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \bar{\vartheta})^2}, \quad \text{where} \quad \bar{\vartheta} = \sum_{i=1}^{L}\hat{\vartheta}/L. \tag{5.9}$$

Root Mean Square Error

The root mean square error (RMSE) is used to assess the difference between the true value of a model and its prediction. RMSE measure denotes the sampling error of an estimator in comparison with the true parameter value (Sigal and Chalmers, 2016), and it can be stated as:

$$\text{RMSE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(\hat{\vartheta}_i - \vartheta\right)^2}. \tag{5.10}$$

An estimator with lower RMSE will recover the true parameter value more efficiently (Sigal and Chalmers, 2016; Yuan et al., 2015), and minimum RMSE yields maximum precision (Wang et al., 2018). An estimator is consistent when RMSE decreases such that $\hat{\vartheta} \to \vartheta$ as $N \to \infty$ (Ghalanos, 2018; Morris et al., 2019). RMSE, bias and sampling variability are connected as:

$$\text{RMSE} = \sqrt{\text{bias}^2 + \text{SE}^2}. \tag{5.11}$$

That is, the RMSE measure includes both bias and SE, where low SE can be penalised for bias. The mean squared error (MSE) can be derived by squaring the RMSE. It is expected that as the sample size increases, the sampling error should decrease relatively, but this is not always the outcome (Gilli et al., 2019).

Relative Efficiency

Sampling efficiency between one estimator and the other can be compared using the relative efficiency (RE) statistic by creating ratios between the estimators' RMSE values as shown in Equation (5.12). In the formula, the first RMSE or RMSE_1 is taken as the reference estimator, such that other RMSEs are assessed in relation to it. The $\text{RE} \equiv 1$ when $i = 1$, and values above (less than) 1 indicate less (greater) efficiency. Less efficiency means more variability than the statistic that is referenced, while greater efficiency indicates less variability and higher precision for recovering the true parameter ϑ (Chalmers, 2019; Sigal and Chalmers, 2016). $\text{RE} < 1$ suggests that estimator i is the preferred one, i.e., if $i = 2$ for instance, estimator 1 is less efficient in relation to estimator $i = 2$ (Davidian, 2016).

$$\text{RE} = \left(\frac{\text{RMSE}_i}{\text{RMSE}_1} \right)^2 . \tag{5.12}$$

5.2.4 Discussion and Summary

After the method of implementation, the last step in the structure is the conclusion, which should reflect a summary discussion of all relevant findings from the simulations, with answers provided to the research questions. The novelty of the work is presented in the conclusion, and this may include the limitations encountered and the future directions of the work.

5.3 Results: Simulation and Empirical

5.3.1 Practical Illustrations of the Simulation Structure with Applications to Financial Returns

This section practically illustrates how the design steps can be applied using MCS with empirical verification through the real SA bond returns data. The steps are illustrated as follows:

5.3.2 The Background

As a preamble, simulation involving the GAS model can be conducted through the UniGASSim function with the uGASFit object (e.g., using "fit" as the uGASFit object), and through the full specification of the GAS model (with "fit = NULL" in the UniGASSim function)[1]. In the latter approach which involves the full specification of the model, the coefficients of the matrices $\boldsymbol{\kappa}$, $\mathbf{A}$ and $\mathbf{B}$ in Equation (5.4) have to be specified. Both of these approaches can

[1]Readers can refer to Page 52 of Catania et al. (2020) for the usage of the UniGASSim function.

be done with a shape parameter (or degree of freedom) that is time-varying (ν_t) or constant (ν). This illustrative study focuses on ν_t. In the simulation structure, **A** and **B** matrices are constrained to exist as diagonals as described in Section 5.2.1.1. Hence, a_μ, a_ϕ, and a_ν are diagonal elements of **A**, while b_μ, b_ϕ, and b_ν are diagonal elements of **B**. In the outputs of the estimation, a1, a2, a3, b1, b2 and b3 are estimates of a_μ, a_ϕ, a_ν, b_μ, b_ϕ and b_ν, respectively. Also $\boldsymbol{\kappa} \equiv (\mathbf{I} - \mathbf{B})\Lambda^{-1}\boldsymbol{\vartheta}^*$, where **I** is an identity matrix of a suitable size, $\boldsymbol{\vartheta}^*$ is the ThetaStar object, and Λ^{-1} is a mapping function that is represented by the UniUnmapParameters() function[1]. The ThetaStar $\boldsymbol{\vartheta}^*$ is a vector of the unconditional time-varying parameters, i.e., $\boldsymbol{\vartheta}^* = (\mu^*, \phi^*, \nu^*)^\top$, where μ^*, ϕ^*, and ν^* are the unconditional location, scale, and shape parameters, respectively. The coefficients kappa1, kappa2 and kappa3 in the estimation's outputs are vector $\boldsymbol{\kappa}$'s elements, where $\boldsymbol{\kappa} \equiv (\kappa_\mu, \kappa_\phi, \kappa_\nu)^\top$(see Ardia et al., 2019 for more details).

Volatility modelling for risk management is highly dependent on the underlying error distribution. In other words, to minimize risk and adequately maximise asset returns, volatility needs to be modelled with an appropriate error assumption when the true (underlying) error distribution is unknown. Hence, this illustrative study demonstrates the dynamics of the GAS model as the true data generating process (DGP) for the simulation, where the outcomes of the model fitted with each of seven assumed error distributions are compared to obtain the most suitable for volatility persistence estimation. The assumed error distributions are the Gaussian, skew-Gaussian, Student's t, skew-Student's t, asymmetric Student's t with one tail decay parameter (AST1), asymmetric Student's t with two tail decay parameters (AST), and asymmetric Laplace distribution (ALD). Details on the error distributions can be seen in Ardia et al. (2019); Creal et al. (2011, 2013); Ghalanos (2018); Kotz et al. (2001); Zhu and Galbraith (2010) or through the function DistInfo() in R package "GAS".

For further background layout, it is known that financial data are fat-tailed (Li, 2008), i.e., non-Normal. Hence, to obtain a set of simulated observations that can closely reflect this fat-tailed behaviour of financial time series data under the GAS modelling, we carry out experiments to compare the fat-tails outcomes of three different datasets that are generated using three unconditional shape parameters or degrees of freedom ν^* = 10.3, 8.2, and 4.1. We execute this by fitting GAS-Student's t model[2] with time-varying location, scale, and shape parameters to the SA bond returns data to generate three sets of N = 30,000 simulated datasets. That is, the first dataset is obtained with ν^* = 10.3, the second and third with ν^* = 8.2 and ν^* = 4.1, respectively. For each simulated dataset, a quantile-quantile (QQ) plot is produced as displayed in Figure 5.2. It can be seen from the figure that the plot in Panel C for GAS-Student's t with ν^* = 4.1 displays the fattest tails than the other two panels with ν^* = 8.2 and ν^* = 10.3. The tails of Panel C's plot are about twice and more than twice as fat as those of Panels B and A, respectively. Hence, the GAS model fitted to a Student's t distributed error with ν^* of

[1]That is, the inverse of the mapping function Λ^{-1} in $\boldsymbol{\kappa}$ is represented by the UniUnmapParameters() function in the GAS simulation code of the R package GAS (see Ardia et al., 2019).

[2]GAS-Student's t model is the GAS model fitted with a Student's t error.

4.1 appears to be the most suitable among the three to generate the required simulated financial returns data.

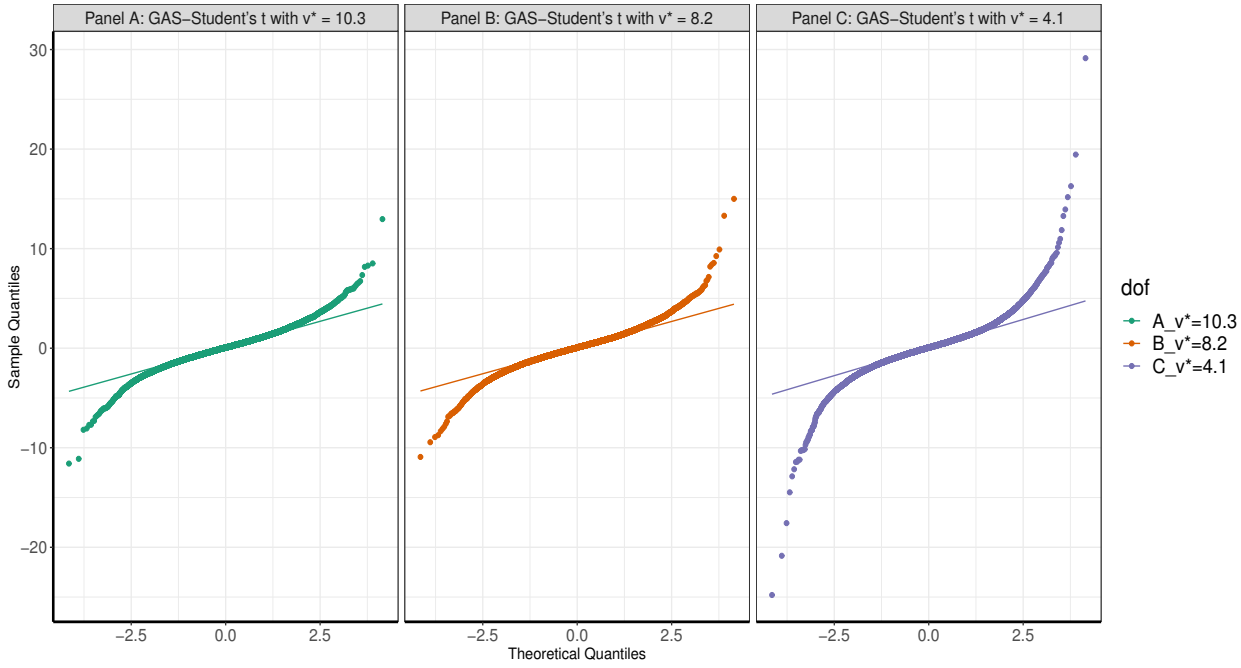

Figure 5.2: The QQ plots in Panels A, B, and C are obtained from GAS-Student's t model fitted with $\nu^* = 10.3$, $\nu^* = 8.2$, and $\nu^* = 4.1$, respectively. Each model is fitted to $N = 30{,}000$ simulated datasets.

Based on the fat-tails outcomes, we further carry out two sets of experiments to examine the $\sqrt{N}$ consistency of datasets fitted with each of these three unconditional shape parameters $\nu^* = 4.1$, 8.2 and 10.3. In the first experiment, we fit GAS-Student's t model with time-varying location, scale, and shape parameters to the SA bond returns data to obtain the parameter values that are used to generate three datasets of sample sizes $N = \{30{,}000;\ 31{,}000;\ 32{,}000\}$. That is, all the outcomes from the fit (with ν^* fixed in turn to 4.1, 8.2 and 10.3) are used to generate the three datasets. This data generation process is executed through the full specification of the GAS model. Precisely, in the first set of this first experiment, we use the outcomes from the fit with ν^* fixed to 4.1 to generate the three datasets. In the second set, we also use the outcomes from the fit with ν^* fixed to 8.2 to generate the next three datasets. In the third set, we again use the outcomes from the fit with ν^* fixed to 10.3 to generate the last three datasets.

Next, we fit the GAS model with a time-varying scale parameter to each of the three datasets under four error distributions of the Gaussian, skew-Gaussian, Student's t, and skew-Student's t. Panel A of Table D.1 in Appendix D shows the RMSE, SE and bias outcomes of the fit under each error distribution for the GAS volatility persistence estimator $\hat{b}_\phi$. It can be seen from the table that the GAS model fitted with $\nu^* = 4.1$ gives the best consistency in terms of RMSE, SE and absolute value of bias. That is, the model fitted using $\nu^* = 4.1$ outperforms those

with $\nu^* = 8.2$ and $\nu^* = 10.3$ in consistency under the four error distributions. These outcomes are visually displayed in Figure 5.3, where Panels A and B show the superior performance of $\nu^* = 4.1$ in terms of SE and RMSE[1], respectively, as compared to those of $\nu^* = 8.2$ (in Panels C and D) and $\nu^* = 10.3$ in Panels E and F. The RMSE displays some inconsistency for $\nu^* = 8.2$ and $\nu^* = 10.3$ in Panels D and F, respectively.

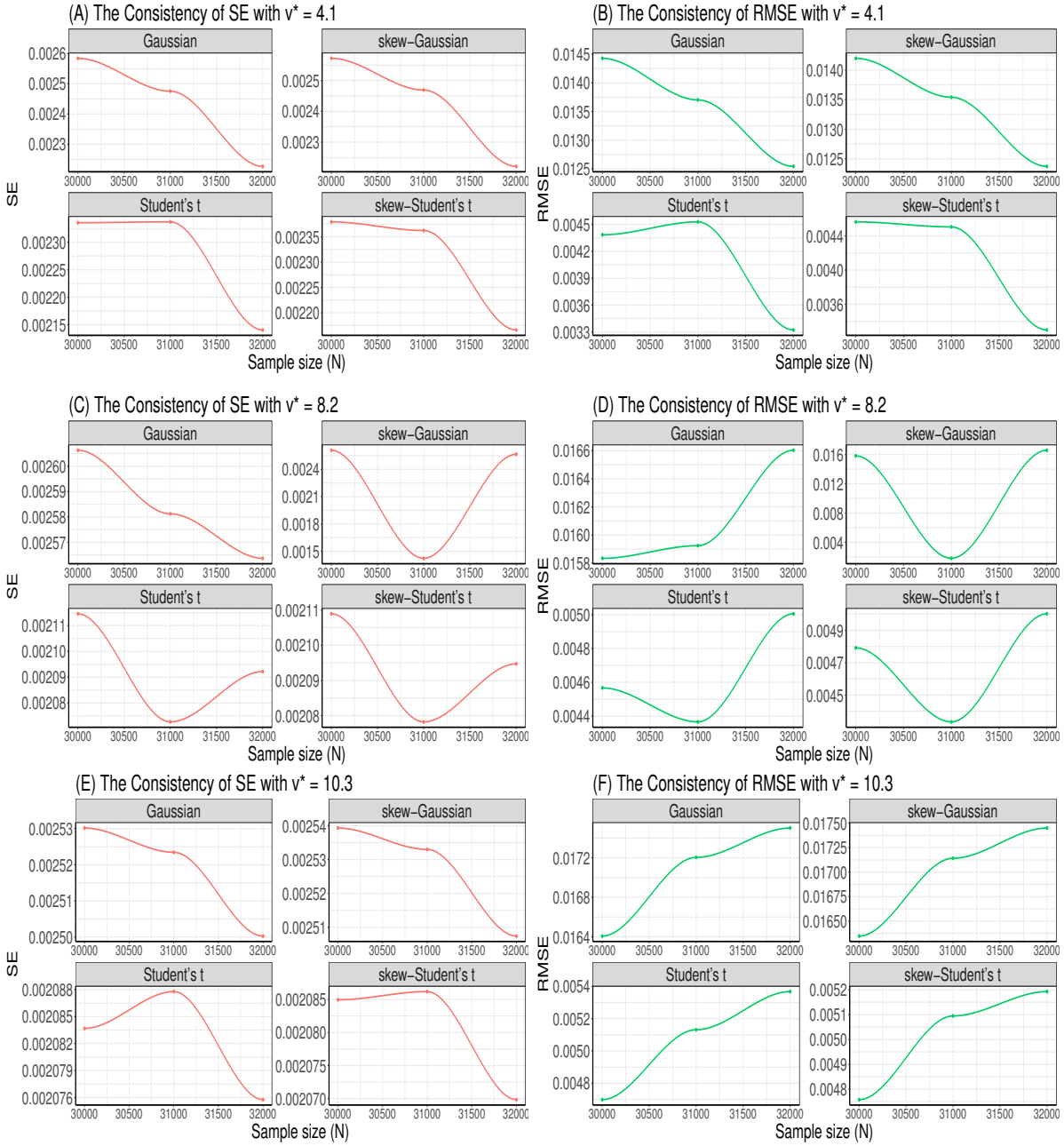

Figure 5.3: Panels (A) and (B) display the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ in terms of SE and RMSE for $N = \{30{,}000;\ 31{,}000;\ 32{,}000\}$ datasets with $\nu^* = 4.1$, while Panels (C) and (D) display those for $\nu^* = 8.2$, and Panels (E) and (F) for $\nu^* = 10.3$.

[1]We did not report the plots of the $\sqrt{N}$ consistency of the absolute value of bias for brevity reasons.

In the second experiment, we try to avoid potential initial value effects by generating three equal datasets, each of sample size N = 30,000. From the datasets, we remove the first N = {13,000; 12,000; 11,000} respectively to avoid the effect of initial values, and then use the last N = {17,000; 18,000; 19,000} under each of the four error distributions as shown in Panel B of Table D.1. These trimmings (and subsequent ones in this study) are executed following the simulation design of Feng and Shi (2017). The outcomes of the experiment are presented in Panel B of Table D.1 in the Appendix. It can be seen from the table that the GAS model fitted with ν^* = 4.1 gives the best $\sqrt{N}$ consistency in terms of RMSE, SE and absolute value of bias when compared with those of ν^* = 8.2 and ν^* = 10.3 under the four error distributions. These consistency outcomes are visually displayed in Figure 5.4, where Panels A and B show the superior performances of ν^* = 4.1 in terms of SE and RMSE, respectively, as compared to those of ν^* = 8.2 in Panels C and D, and ν^* = 10.3 in Panels E and F. Moreover, when compared with the outcomes of the first experiment, it is observed as shown in Panels A and B of Figure 5.4 that the $\sqrt{N}$ consistency of the estimator $\hat{b}_\phi$ for ν^* = 4.1 becomes better and more improved in terms of SE and RMSE after removing the initial value effects. Hence, the Student's t GAS model fitted with ν^* = 4.1 outperforms those fitted with ν^* = 8.2 and ν^* = 10.3 in consistency.

The Student's t is used as the true error distribution for this illustration because it is assumed that financial data appear to have a distribution much like it, and it is also believed that it can suitably deal with fat-tailed or leptokurtic features (Duda and Schmidt, 2009; Lin and Shen, 2006) commonly found in financial data (Hentschel, 1995). However, to buttress the illustration, we further use the AST1 as another true error distribution in addition to the Student's t, and compare their outcomes. Besides these two true sampling errors, users may choose to use any relevant non-Gaussian distributions, as stated in the DistInfo() function, for their data generation based on individual research needs. For any chosen true error distribution, the DistInfo() function reveals the parameters that must be specified as time-varying.

5.3.3 Aim of the Simulation Study

The aim of this study is to obtain the most suitable assumed error distribution to describe the returns for volatility persistence estimation when the underlying error distribution is not known.

5.3.4 Research Questions

The outcomes of this simulation study should help to respond to the following questions:

1. Which among the error assumptions is the most suitable for the GAS simulation process to estimate the volatility persistence?

2. What type (i.e., weak or strong) of $\sqrt{N}$ consistency, in terms of SE and RMSE, does the volatility persistence estimator $\hat{b}_\phi$ of the GAS process exhibit?

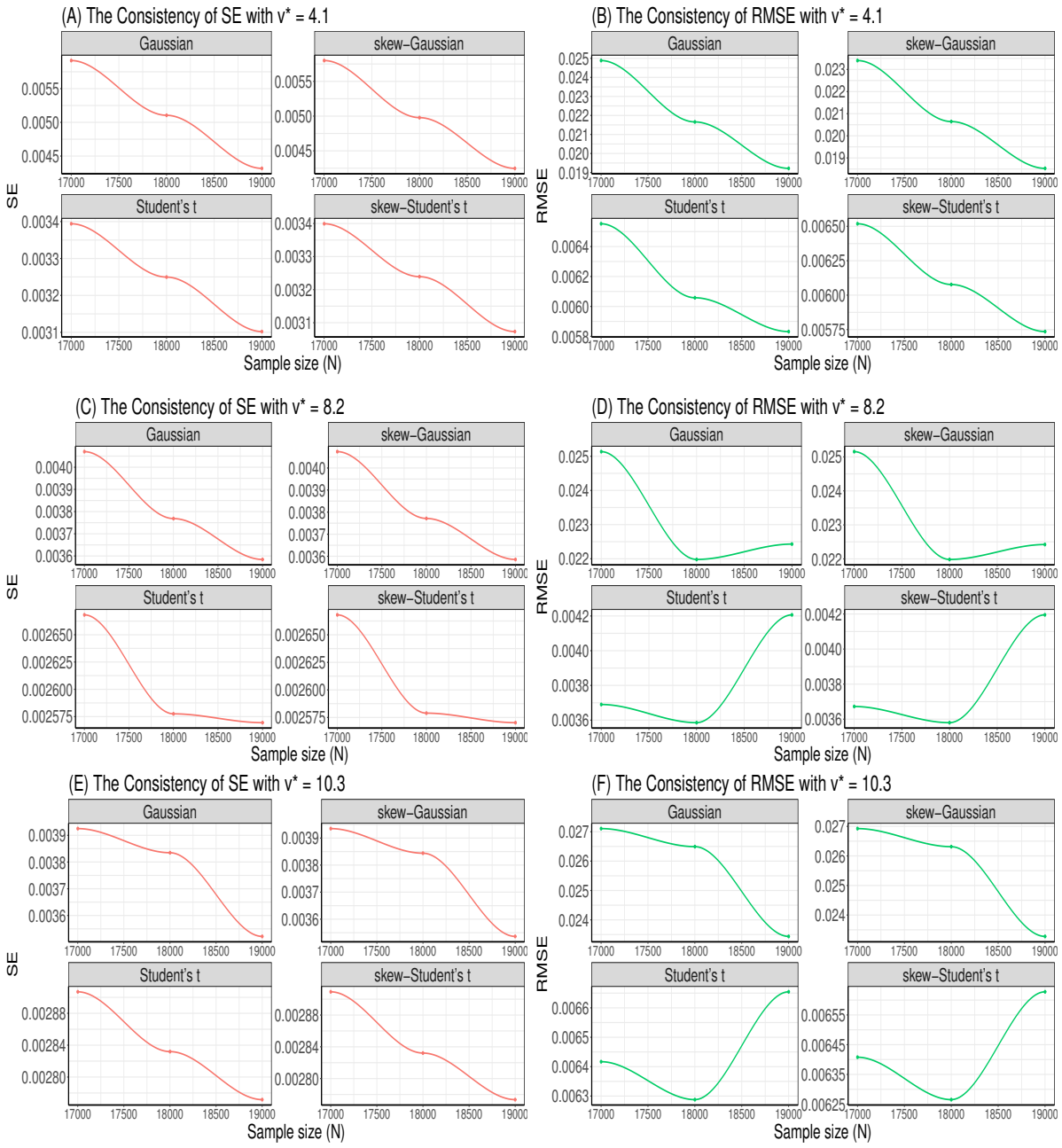

Figure 5.4: Panels (A) and (B) display the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ in terms of SE and RMSE for N = {17,000; 18,000; 19,000} datasets with ν^* = 4.1, while Panels (C) and (D) display those for ν^* = 8.2, and Panels (E) and (F) for ν^* = 10.3.

3. How does the estimator $\hat{b}_\phi$ of the most suitable assumed error distribution compare to the estimator $\hat{b}_\phi$ of competing assumed error distributions with regards to bias, and efficiency (or precision) in terms of RMSE, SE and RE?

4. How does the GAS estimator $\hat{b}_\phi$ perform in recovering the true parameter b_ϕ?

5.3.5 Method of Implementation

Since we intend to use the Student's t and AST1 as the true error distributions to generate the simulated return observations, the method of implementation is described via two Monte Carlo experiments. The implementation method is illustrated using the full specification of the GAS model in the coefficients of matrices **A**, **B** and $\boldsymbol{\kappa}$. For the two true error distributions, the scale parameters' descriptions of the conditional variance are within $0.975 < b_\phi < 1$ and $0 < a_\phi < 0.15$. The b_ϕ is the GAS volatility persistence parameter in the diagonal of matrix **B** (see Blasques et al., 2014; Creal et al., 2013). Moreover, the b_ϕ parameter of the GAS model corresponds to the $\alpha + \beta$ parameters of the GARCH model[1] (Blasques et al., 2014).

Next, the two MCS experiments to illustrate the method of implementation are carried out using each of the two true error distributions one after another as follows.

5.3.5.1 Method of Implementation with the Student's t Error

With Student's t as the true innovation distribution, the written code is first used to fit the true model GAS-Student's t with time-varying[2] location, scale and shape parameters to the SA bond returns data through the UniGASFit function, to obtain the true parameter values for the data generating process. The true parameter values are the MLE estimates obtained from fitting the GAS-Student's t model to the SA bond returns data. The time variation in the location, scale and shape parameters of the fit are graphically displayed in Figure 5.5, with the location in the top panel, followed by the scale and then the shape parameter graphics. The true parameter values used to generate the simulated returns data are $a_\mu = 0.0000$, $b_\mu = 0.3461$, $a_\phi = 0.1400$, $b_\phi = 0.9831$, $a_\nu = 9.8356$, $b_\nu = 0.5253$, $\mu^* = 0.0578$, $\phi^* = 1.0695$ (and ν^* fixed to 4.1). The a_μ, b_μ denote the location parameters, a_ϕ, b_ϕ indicate the scale parameters, and a_ν, b_ν are the shape parameters, while μ^*, ϕ^* and ν^* are the unconditional location, scale and shape parameters, respectively.

Next, through the UniGASSim function, using seed 12345 in the code, the value of the scale parameter b2 = $b_\phi = 0.9831$ from the fit is set a priori as the true parameter for the MCS process, as shown in Table 5.1. This value and all the other estimates (with the ν^* that is fixed to 4.1) from the uGASFit object are used in the code to generate three sets of N = 30,000 return observations.

Next, due to the effect of the initial values in the generation of the data, which may cause size distortion (Su, 2011), we discard the first N = {13,000; 12,000; 11,000} sets of observations at each stage of the generated 30,000 return observations to avoid the distortion. Hence,

[1] We examined this by comparing the estimate of the GAS persistence parameter $\hat{b}_\phi$, fitted with a time-varying scale parameter, and the estimate from the GARCH(1,1) model's persistence parameter $\hat{\alpha}_1 + \hat{\beta}_1$. We fitted both models to the SA bond returns data where the innovation is Gaussian, and their results yielded $\hat{b}_\phi \equiv \hat{\alpha}_1 + \hat{\beta}_1 \approx 0.98$.

[2] Time-varying parameters are specified as "TRUE" in the UniGASSpec function.

Table 5.1: Simulation outcomes.

True model: GAS-Student's t

True parameter: $b_\phi = 0.9831$

Assumed error	N	$\hat{b}_\phi$	**RMSE**$_{\hat{b}_\phi}$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RE**$_{\hat{b}_\phi}$	**TPR**$_{\hat{b}_\phi}$ **(95%)**
	17,000	0.9589	0.0249	-0.0242	0.0059	–	92.66%
Gaussian	18,000	0.9620	0.0217	-0.0210	0.0051	0.7575	92.97%
	19,000	0.9643	0.0192	-0.0187	0.0043	0.7872	93.19%
	17,000	0.9604	0.0234	-0.0227	0.0058	–	92.81%
skew	18,000	0.9630	0.0206	-0.0200	0.0050	0.7781	93.06%
Gaussian	19,000	0.9650	0.0185	-0.0180	0.0042	0.8055	93.26%
	17,000	0.9775	0.0066	-0.0056	0.0034	–	94.46%
Student's t	18,000	0.9780	0.0061	-0.0051	0.0032	0.8547	94.51%
	19,000	0.9781	0.0058	-0.0049	0.0031	0.9268	94.52%
	17,000	0.9775	0.0065	-0.0056	0.0034	–	94.46%
skew	18,000	0.9779	0.0061	-0.0051	0.0032	0.8692	94.50%
Student's t	19,000	0.9782	0.0057	-0.0048	0.0031	0.8906	94.53%
	17,000	0.9774	0.0066	-0.0057	0.0034	–	94.45%
AST	18,000	0.9777	0.0063	-0.0054	0.0033	0.9061	94.48%
	19,000	0.9782	0.0058	-0.0049	0.0031	0.8398	94.53%
	17,000	0.9775	0.0065	-0.0056	0.0034	–	94.46%
AST1	18,000	0.9780	0.0060	-0.0051	0.0032	0.8581	94.51%
	19,000	0.9788	0.0053	-0.0043	0.0031	0.7623	94.59%
	17,000	0.9768	0.0073	-0.0063	0.0037	–	94.39%
ALD	18,000	0.9775	0.0066	-0.0056	0.0035	0.8140	94.46%
	19,000	0.9778	0.0062	-0.0052	0.0033	0.8827	94.49%

Note: Simulation outcomes from the true model GAS-Student's t.

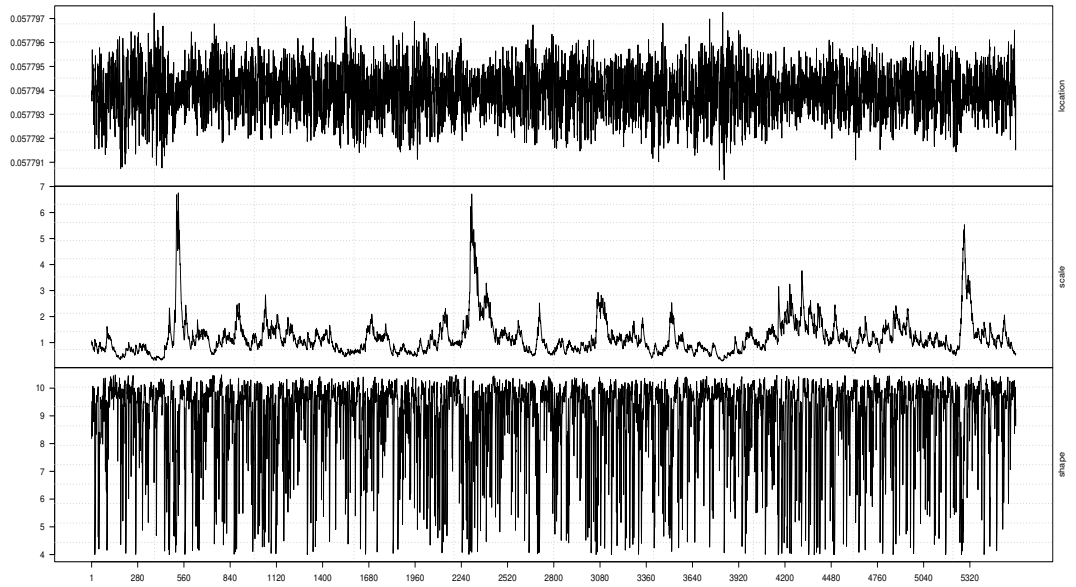

Figure 5.5: GAS-Student's t: Time-varying location (top panel), scale (middle panel), and shape (bottom panel) parameters.

we only use the last N = {17,000; 18,000; 19000} simulated returns under each of the seven assumed errors as presented in Table 5.1. After generating the simulated return datasets, the GAS model is fitted to them under each of the seven assumed error distributions. The full analytic code that shows the lines of commands for each stage of the method of implementation is available from the authors on request.

Next, comparisons through the four meta-statistical summaries comprising of the RMSE, bias, SE and RE are carried out for the volatility persistence estimator $\hat{b}_\phi$ in **B**. The most adequate assumed innovation for volatility persistence estimation will be obtained from the estimator $\hat{b}_\phi$ with the best efficiency and precision (for recovering the true parameter) in the meta-statistical comparisons executed under the seven innovation assumptions.

Now, comparing RMSE for $\hat{b}_\phi$ in Table 5.1, the AST1 assumed error distribution outperforms the other six assumed errors, including the true error Student's t, in efficiency with the least value as N tends to the peak at 19,000. For bias comparison, the AST1 also outperforms the rest of the error assumptions with the lowest absolute value of bias as N tends to the peak. Comparing SE, the Student's t, skew-Student's t, and AST1 are equal in performance for efficiency and precision in terms of SE, and they outperform the rest. The AST assumed error follows them in performance. Lastly, for RE comparison, we use each preceding RMSE as the reference estimator, i.e., RMSE_1 as the reference estimator for RMSE_2, and RMSE_2 for RMSE_3 under each error assumption. The outcomes show that the AST1 displays the highest precision for recovering the true parameter than the rest, with the smallest RE value as N tends to the peak. The RMSE and SE performance measures also showcase adequate $\sqrt{N}$ consistency under the seven assumed error distributions. To summarise, when the true error

distribution is Student's t, the AST1 assumed error distribution relatively outperforms the other six error assumptions in efficiency and precision, and in the overall performance.

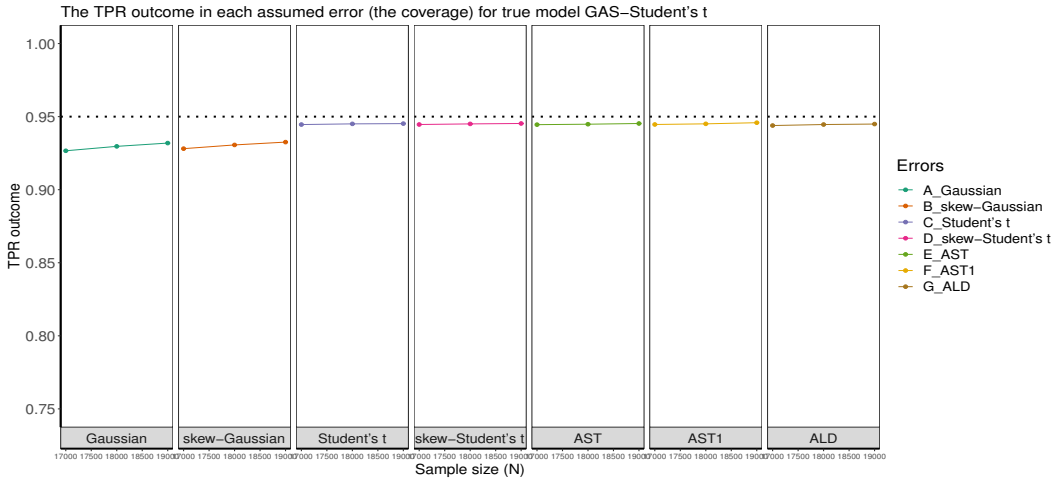

Figure 5.6: The true parameter recovery (TPR) outlook. The dotted line denotes the 0.95 (i.e., 95%) nominal recovery level.

Table 5.1 further reveals that the MCS estimator $\hat{b}_\phi$ considerably recovers the true parameter b_ϕ at the 95% (i.e., 0.95) nominal recovery level. The recovery outcomes for the seven assumed error distributions are visually displayed in Figure 5.6, where it is revealed that the MCS estimates perform considerably well in recovering the true parameter as shown by the closeness of the TPR outcomes to the 95% nominal recovery level. This indicates that the MCS experiments are well executed with suitably valid outcomes. However, the non-Gaussian error assumptions perform better in the recovery than the Gaussian and skew-Gaussian error assumptions, as shown in the plots. It is also observed from the results in the table that the MCS estimate $\hat{b}_\phi$ is directly proportional to the TPR outcome. Hence, the smaller (larger) the MCS estimate, the smaller (larger) the TPR outcome.

5.3.5.2 Method of Implementation with AST1 Error

The second illustration of the method of implementation is carried out with the AST1 as the true error distribution. The written lines of code are first used to fit the true model GAS-AST1 with time-varying location, scale, skewness, and shape parameters to the SA bond returns data through the UniGASFit function. The time variations in these parameters are graphically displayed in Figure 5.7, with the location in the top panel, followed by the scale, skewness, and shape parameter graphics downwards in subsequent panels. The MLE estimates from the fit are used as the true parameter values to generate the simulated returns for the MCS. The true parameter values used to generate the returns data are $a_\mu = 0.0000$, $b_\mu = 0.8802$, $a_\phi = 0.0420$, $b_\phi = 0.9788$, $a_\eta = 0.1186$, $b_\eta = 0.0000$, $a_\nu = 0.0001$, $b_\nu = 0.9702$, $\mu^* = 0.1548$, $\phi^* = 2.6392$, $\eta^* = 0.5329$ (and ν^* fixed to 4.1). The a_μ, b_μ are the location parameters, a_ϕ, b_ϕ denote the

scale parameters, a_η, b_η are the skewness parameters, and a_ν, b_ν are the shape parameters. The μ^*, ϕ^*, η^* and ν^* are the unconditional location, scale, skewness, and shape parameters.

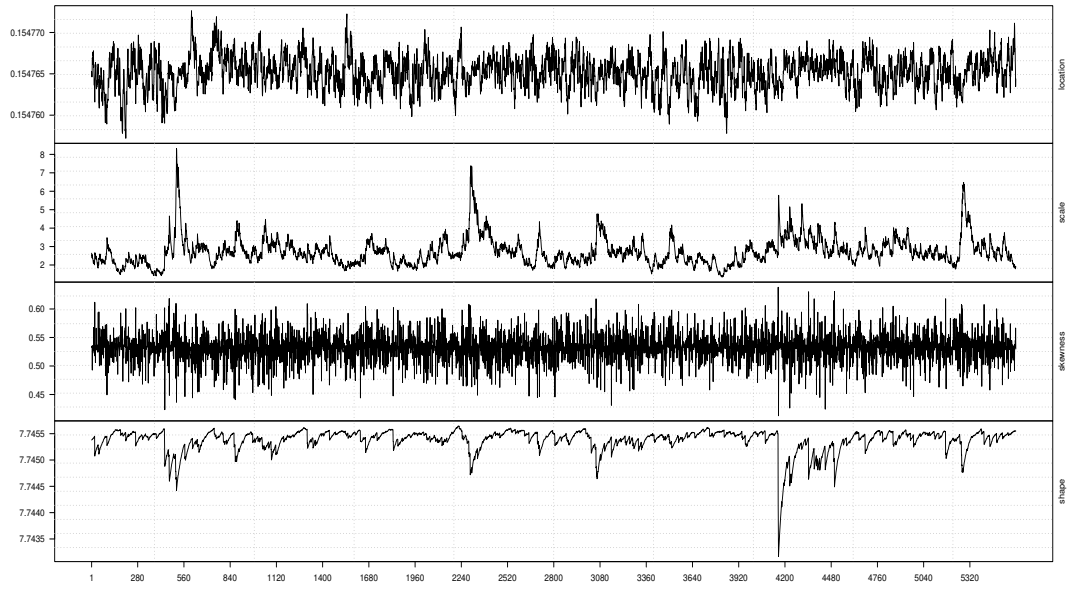

Figure 5.7: GAS-AST1: Time-varying location (top panel), scale (upper-middle panel), skewness (lower-middle panel) and shape (bottom panel) parameters.

Next, through the UniGASSim function, using seed 12345 in the code, the value of the scale parameter b2 = $b_\phi = 0.9788$ from the fit is set a priori as the true parameter for the MCS process as presented in Table 5.2. This value and all the other estimates (with the ν^* that is fixed to 4.1) from the uGASFit object are used in the code to generate three sets of N = 30,000 return observations. Next, following the previous trimming pattern, N = {17,000; 18,000; 19000} datasets are used under each of the seven assumed errors, as shown in Table 5.2. The full code is available from the authors on request.

Now, comparing RMSE and SE for $\hat{b}_\phi$ in Table 5.2, the AST1 error assumption outperforms the other six assumed errors in efficiency and precision with the least values as N tends to the peak. For bias comparison, the AST1 also outperforms the rest, with the lowest absolute value of bias as N tends to the peak. The relative efficiency (RE) comparison is not done here due to the weak $\sqrt{N}$ consistency of the RMSE under the AST and ALD assumed errors. In this study, consistency is termed "strong" ("weak") when the estimator's RMSE/SE value decreases as the sample size N increases without (with) distortion. To summarise, when the true error distribution is AST1, the AST1 assumed error distribution outperforms the other six error assumptions in efficiency and precision, and in overall performance. It is also observed from the outcomes of the $\text{SE}_{\hat{b}_\phi}$ for both true models in Tables 5.1 and 5.2 that the GAS model with the Gaussian error distribution is considerably consistent but not efficient when the true error is not Gaussian.

Table 5.2: Simulation outcomes.

True model: GAS-AST1

True parameter: b_ϕ = 0.9788

Assumed error	N	$\hat{b}_\phi$	**RMSE**$_{\hat{b}_\phi}$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**TPR**$_{\hat{b}_\phi}$ **(95%)**
	17,000	0.9444	0.0346	-0.0344	0.0036	91.66%
Gaussian	18,000	0.9488	0.0302	-0.0300	0.0031	92.09%
	19,000	0.9493	0.0300	-0.0295	0.0051	92.13%
	17,000	0.9522	0.0268	-0.0266	0.0034	92.42%
skew	18,000	0.9527	0.0264	-0.0262	0.0034	92.46%
Gaussian	19,000	0.9531	0.0259	-0.0257	0.0033	92.50%
	17,000	0.9747	0.0051	-0.0042	0.0030	94.60%
Student's t	18,000	0.9746	0.0051	-0.0042	0.0030	94.59%
	19,000	0.9752	0.0046	-0.0036	0.0028	94.65%
	17,000	0.9744	0.0053	-0.0044	0.0030	94.57%
skew	18,000	0.9744	0.0053	-0.0044	0.0030	94.57%
Student's t	19,000	0.9750	0.0047	-0.0038	0.0028	94.63%
	17,000	0.9749	0.0049	-0.0039	0.0030	94.62%
AST	18,000	0.9744	0.0053	-0.0044	0.0029	94.57%
	19,000	0.9753	0.0045	-0.0035	0.0028	94.66%
	17,000	0.9750	0.0049	-0.0039	0.0030	94.63%
AST1	18,000	0.9749	0.0049	-0.0040	0.0029	94.62%
	19,000	0.9755	0.0043	-0.0033	0.0027	94.68%
	17,000	0.9745	0.0054	-0.0044	0.0031	94.58%
ALD	18,000	0.9740	0.0057	-0.0048	0.0031	94.54%
	19,000	0.9748	0.0049	-0.0040	0.0030	94.61%

Note: Simulation outcomes from the true model GAS-AST1.

Furthermore, as revealed in Table 5.2 and as visualised in Figure 5.8, the MCS estimator $\hat{b}_\phi$ considerably recovers the true parameter b_ϕ close to the 95% nominal recovery level, for each of the seven assumed error distributions. This suggests that the MCS experiments are well carried out with suitably valid results. However, as observed under the true model GAS-Student's t, it is also observed here that the non-Gaussian error assumptions perform better in the recovery than the Gaussian and skew-Gaussian error assumptions, as shown in the plots.

It can be seen from Figures 5.6 and 5.8, and in Tables 5.1 and 5.2 that the MCS estimator $\hat{b}_\phi$ performs slightly better at recovering the true parameter b_ϕ under the Gaussian and skew-Gaussian assumed errors for the true model GAS-Student's t than it is for true model GAS-AST1. Moreover, the true model GAS-Student's t displays better $\sqrt{N}$ consistency for the RMSE and SE metrics than the GAS-AST1 as revealed in the tables. Hence, these show the preference of the Student's t error distribution for the data generation process over the AST1 error distribution.

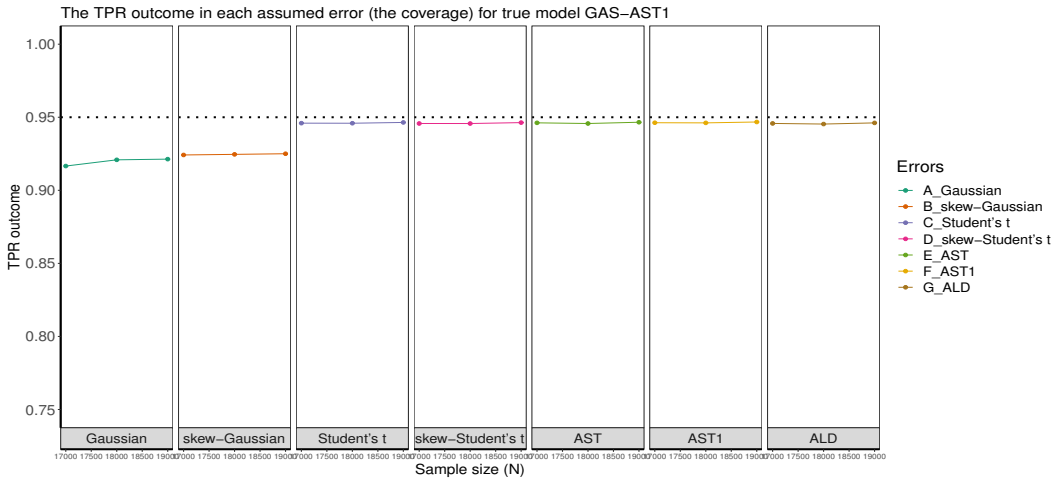

Figure 5.8: The true parameter recovery (TPR) outlook. The dotted line denotes the 0.95 (i.e., 95%) nominal recovery level.

5.3.6 Empirical Study

Here, the results of the simulation modelling are verified empirically by using the actual returns from the SA bond market index. Among the selected seven innovations, the most suitable for the GAS model to describe the market's returns for volatility persistence estimation is investigated. We transform the daily price data to the log returns by taking the log-difference of the value of the index as:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100, \tag{5.13}$$

where r_t is the current returns, ln denotes the natural logarithm, and P_t is the daily closing bond price at time t.

5.3.6.1 Exploratory Data Analysis

To begin with, the price and return series are first examined via exploratory data analysis (EDA) as shown in Figures 5.9 to 5.11. The EDA visually inspects the content of the dataset to show possible outliers and other vital information. The histogram plot in Panel A of Figure 5.9 shows a skewed non-symmetric distribution that suggests non-stationarity in the price index, while Panel B reveals approximately symmetric or stationary distribution in the return series. Panel A of Figure 5.10 also reveals the non-stationarity in the price index as observed in the price series plot, while the return series plot in Panel B shows stationarity in the returns. Figure 5.10 further reveals four obvious extreme spikes in the distributions of price and returns during the period under study. These spikes are a result of severe global events and economic downturns that led to extreme market volatilities, notably the 2002/2003 SARS epidemic outbreak (Weismer, 2020), the 2008 global financial crisis caused by the U.S. loans crisis, i.e., the defaults on consolidated mortgage-backed securities (Amadeo, 2022), the 2016 Brexit issues, and the 2020 COVID-19 pandemic outbreak. Figure 5.11 shows the price and return densities in Panels A and B respectively, while Panels C and D display the movements of the price and return distributions with their densities.

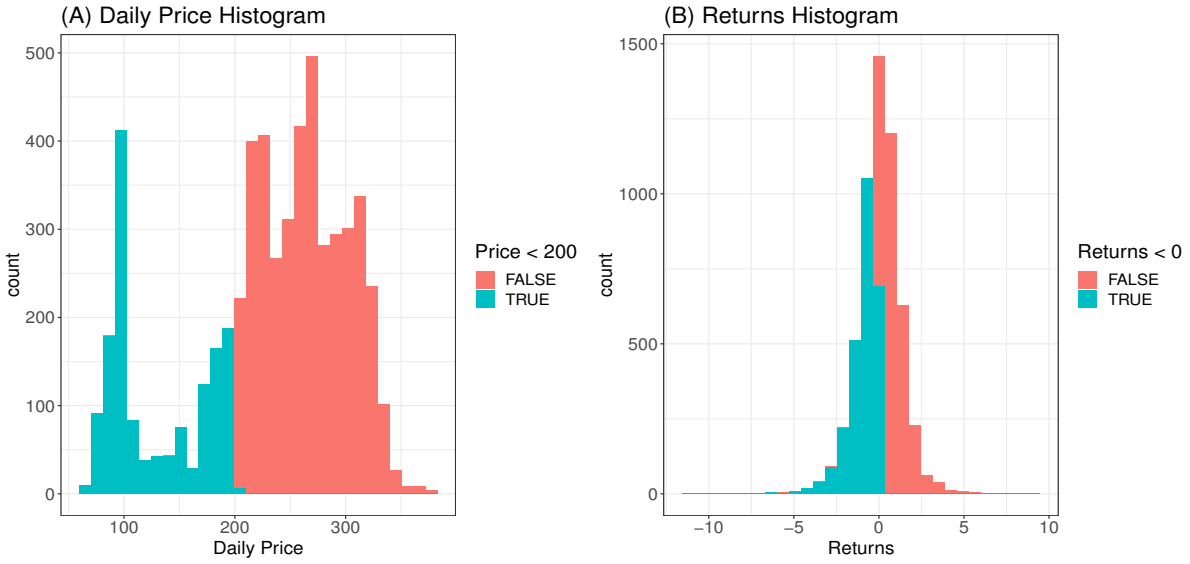

Figure 5.9: Histograms of price (in Panel A) and returns (Panel B). The two colours merely demarcate the distribution where the price is < 200 (with blue colour) and above 200 (with an indianred colour). This is also used where the returns < 0.

5.3.6.2 Tests for Autocorrelation and Heteroscedasticity

Next, we fit the Autoregressive Moving Average (ARMA(u,v)) models (see Ghalanos, 2018), with a Gaussian innovation distribution and a non-Gaussian (Student's t) innovation, to the sta-

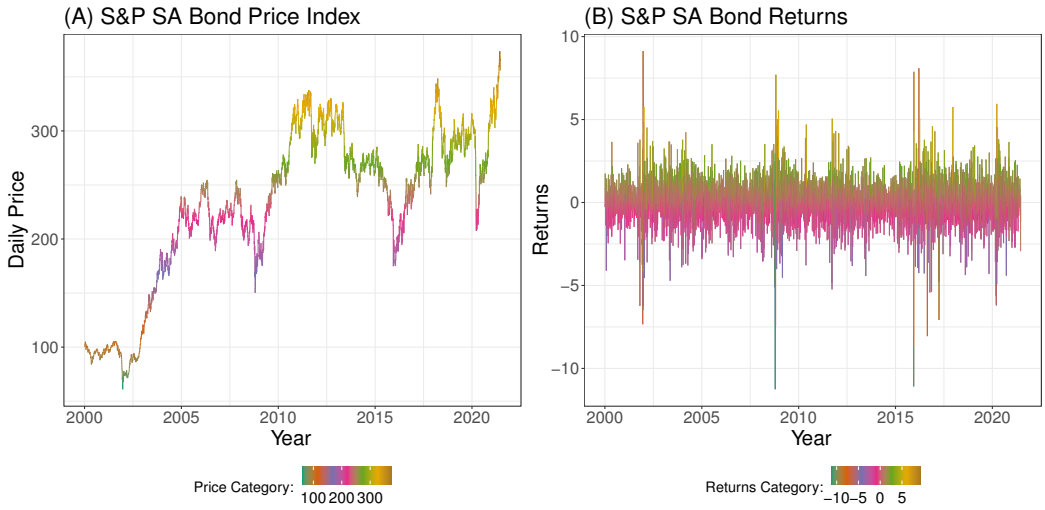

Figure 5.10: Distributions of price (in Panel A) and returns (Panel B).

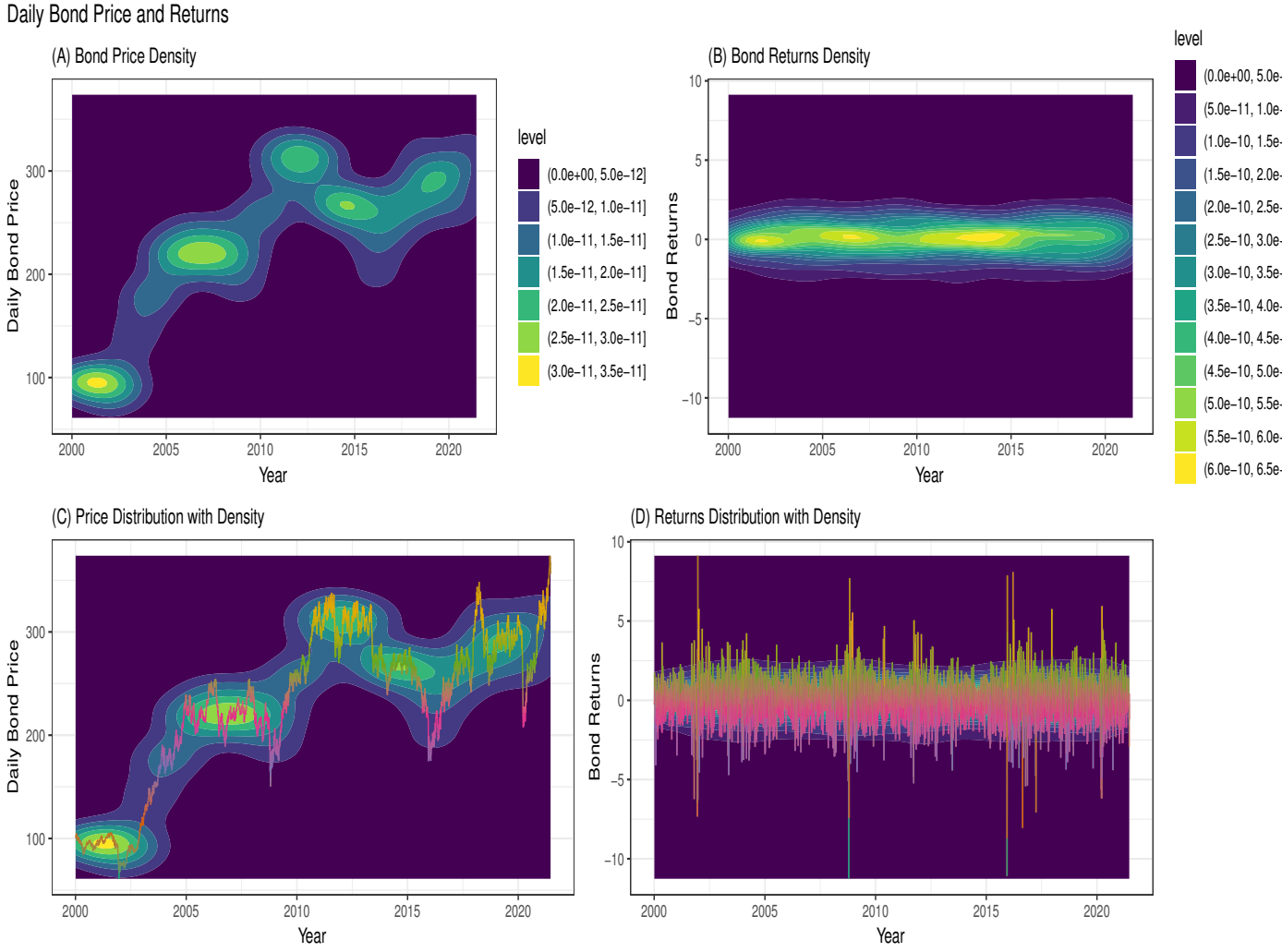

Figure 5.11: Price and return densities in Panels A and B respectively, while Panels C and D display the price and return distributions with their densities.

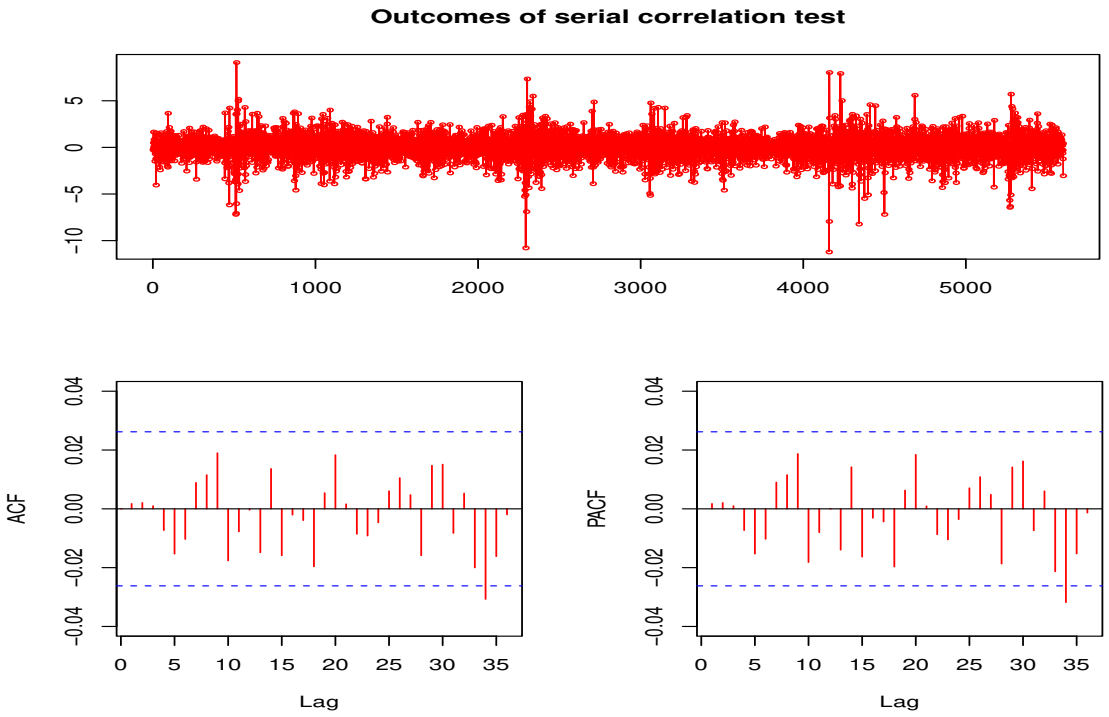

Figure 5.12: Residual plot and outcomes of serial correlation test.

tionary return series to check for the presence of autocorrelation (or serial correlation). The u and v are the orders of the AR and MA processes, respectively. Student's t innovation distribution is used because it adequately reflects the distribution of financial data (Heracleous, 2007). Among the candidate ARMA(u,v) models examined, both ARMA(1,1) and ARMA(2,2) models as stated in Equations (5.14) and (5.15) are found to be adequate to remove possible autocorrelation in the return residuals. The autocorrelation function (ACF) and partial ACF (PACF) plots of the fit of the ARMA(2,2) model are displayed in Figure 5.12[1]. Table 5.3 shows the results of the Weighted Ljung-Box (WLB) tests (see Fisher and Gallagher, 2012) on the "standardised" and "standardised squared" residuals for fitting the two models. The p-values of the tests at lag order 9 for ARMA(1,1) and lag order 19 for ARMA(2,2) are all greater than 5% under each error distribution. Hence, we cannot reject the null hypothesis of "no serial correlation" in the market's returns. This implies that serial correlation is captured in the return residuals.

$$r_t = \phi_0 + \phi_1 r_{t-1} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \tag{5.14}$$

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \varphi_2 \varepsilon_{t-2} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \tag{5.15}$$

Through the outcomes from the autocorrelation test, we carry out Engle's ARCH test (see Engle, 1982) using the Portmanteau-Q (PQ) and Lagrange Multiplier (LM) tests to examine

[1]The ACF and PACF diagrams were plotted using the auto.arima() function in the R forecast package developed by Hyndman and Khandakar (2008).

Table 5.3: ARMA(1,1) and ARMA(2,2) models' outcomes on Bond returns.

		Panel A ARMA(1,1)			Panel B ARMA(2,2)	
		Gaussian	Student's t		Gaussian	Student's t
Test on standardised residuals	$\mathbf{WLB(9)}$	**2.2258**	**3.2100**	$\mathbf{WLB(19)}$	**6.2949**	**10.3830**
	p-$\mathbf{value(9)}$	**(0.9728)**	**(0.8563)**	p-$\mathbf{value(19)}$	**(0.9573)**	**(0.4060)**
Test on standardised squared residuals	$\mathbf{WLB(9)}$	**6.5740**	**7.4880**	$\mathbf{WLB(19)}$	**6.4740**	**7.5450**
	p-$\mathbf{value(9)}$	**(0.2376)**	**(0.1617)**	p-$\mathbf{value(19)}$	**(0.2473)**	**(0.1577)**

Note: WLB denotes the Weighted Ljung-Box test, where "(9)" and "(19)" are lags 9 and 19, respectively.

the existence of heteroscedasticity in the residuals. These tests are carried out based on the null hypothesis of homoscedasticity in the residuals of an Autoregressive Integrated Moving Average (ARIMA) model. The tests are executed using the arima() function, in R package aTSA, for ARMA(1,1) and ARMA(2,2) models. The results of both tests reveal highly significant p-values of 0 from lag order 4 to 24 as presented in Panels A and B of Table 5.4. Thus, the null hypothesis of "no heteroscedasticity" in the residuals is rejected, which indicates the presence of volatility clustering. Volatility clustering implies that big (small) changes in returns tend to be followed by big (small) changes of either sign (Mandelbrot, 1963). Based on the outcomes, we fit the GAS model with each of the seven error distributions to filter out the heteroscedasticity in the returns.

Table 5.4: Engle's ARCH test results.

	Panel A $\mathbf{ARMA(1,1)}$ model				Panel B $\mathbf{ARMA(2,2)}$ model			
	PQ test		LM test		PQ test		LM test	
Lag order	**PQ**	**P-value**	**LM**	**P-value**	**PQ**	**P-value**	**LM**	**P-value**
4	904	0	4729	0	888	0	4658	0
8	1127	0	2196	0	1104	0	2184	0
12	1400	0	1361	0	1377	0	1354	0
16	1517	0	1015	0	1497	0	1010	0
20	1575	0	807	0	1556	0	803	0
24	1624	0	666	0	1602	0	663	0

Note: PQ is the Portmanteau-Q statistic, and LM is the Lagrange-Multiplier statistic.

5.3.6.3 Selection of the Most Appropriate Error Assumption

Next, we use the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (see Ghalanos, 2018; Li et al., 2009 for details) to select the optimal (or the most adequate) assumed error distribution to describe the market's returns for volatility persistence estimation through the fit of the GAS model. The error distribution with the lowest values of the information criteria is selected as the most suitable to describe the market for volatility persistence modelling.

Table 5.5 shows that the GAS parameter estimates $\hat{\kappa}_\phi$, $\hat{a}_\phi$ and $\hat{b}_\phi$ are statistically significant at 1% level under the seven assumed innovations. The only exception to these is the $\hat{\kappa}_\phi$ that is insignificant under the Student's t distribution. The comparisons of the seven assumed innovations are carried out with a time-varying scale parameter in the GAS modelling, and the empirical outcomes of the information criteria as presented in Table 5.5 shows that the values of the AIC and BIC are lowest under the AST1 innovation. Hence, the AST1 innovation assumption outperforms the other six assumed innovations. This outcome is consistent with the results of the MCS studies. The result implies that the persistence of volatility in the returns of the SA bond market can be most suitably described through the GAS model when the assumed innovation is AST1. As shown in the table, the volatility persistence estimate $\hat{b}_\phi$ in **B** under this optimal AST1 error is 0.9790, indicating a considerably high persistence in the returns.

Table 5.5: Empirical outcomes of GAS modelling on SA Bond returns data.

Panel A	Normal	skew-Normal	Student's t	skew-Student's t	AST	AST1	ALD
$\hat{\kappa}_\phi$	0.0071*	0.0033*	0.0011	0.0034*	0.0206*	0.0204*	0.0053*
$\hat{a}_\phi$	0.0872*	0.0221*	0.1653*	0.0395*	0.0412*	0.0411*	0.0587*
$\hat{b}_\phi$	0.9804*	0.9819*	0.9790*	0.9804*	0.9788*	0.9790*	0.9784*
AIC	17936.36	17900.84	17682.49	17672.13	17668.28	17666.39	17910.63
BIC	17962.88	17933.99	17715.64	17711.91	17714.69	17706.17	17943.78

Note: The "$*$" indicates 1% level of significance. The table shows the outcomes of the fit of the GAS model for time-varying scale parameters.

From the empirical outcomes of the GAS model fitted with the AST1 innovation in Table 5.5, the equation of the GAS model for the time-varying scale (or volatility) parameter can be stated as:

$$\boldsymbol{\vartheta}_{t+1} = 0.0204 + 0.0411\boldsymbol{s}_t + 0.9790\boldsymbol{\vartheta}_t. \tag{5.16}$$

5.4 Discussion and Summarised Conclusion

In conclusion, it is observed from the MCS studies that using the UniGASSim function (with full specification of the GAS model) approach, the AST1 consistently emerges as the optimal

(or the most adequate) assumed error distribution to use with the GAS model for the returns' volatility persistence estimation when the underlying distribution of errors is unknown. The result was verified empirically. The persistence of volatility under the optimal assumed innovation revealed considerably high persistence in the volatility of the SA bond market's returns.

The conclusion proceeds with answers provided to the research questions. In answer to the first question, AST1 is the most appropriate conditional distribution, among the seven assumed error distributions from the GAS process simulation, to estimate the volatility persistence. Second, the volatility persistence estimator $\hat{b}_\phi$ displays considerably strong $\sqrt{N}$ consistency. To break this down, there is strong $\sqrt{N}$ consistency for the estimator $\hat{b}_\phi$ under all the seven assumed error distributions in the Student's t true error distribution (or true model GAS-Student's t) for both RMSE and SE in Table 5.1. The AST1 true error distribution or true model GAS-AST1, in Table 5.2, also shows a considerably strong $\sqrt{N}$ consistency under the error assumptions, except the weak consistency in the AST and ALD assumed errors for RMSE metric, and under the Gaussian assumed error for SE metric. Third, when compared to competing assumed errors, the estimator $\hat{b}_\phi$ of the optimal AST1 error is the best overall in efficiency and precision in terms of RMSE, SE and RE, and in the absolute value of bias under the two true error distributions, as revealed in Tables 5.1 and 5.2.

Fourth, with the TPR measure used as a proxy for the coverage of the simulation experiments, the study's outcomes revealed a satisfactory performance of the GAS volatility persistence estimator $\hat{b}_\phi$ in recovering the true parameter b_ϕ at the 95% nominal recovery level. This indicates that the MCS experiments are well carried out with suitably valid outcomes under the two true models. However, the non-Gaussian error assumptions performed better in the recovery than the Gaussian and skew-Gaussian assumed errors. The study also observed from the outcomes of the TPR measure and the $\sqrt{N}$ consistency of the RMSE and SE, that the Student's t is a preferred error distribution for data generation than the AST1.

This illustrative study through the GAS simulation modelling focused on time-varying shape parameter or degree of freedom ν_t, where the "shape = TRUE". Interested users and readers can try the other approach with "shape = FALSE" (i.e., constant shape parameter or degree of freedom ν). Any argument specified as "FALSE" in the UniGASSpec function for the GAS specification is constrained to 0 (see Ardia et al., 2019 or use help(UniGASSpec) in the R GAS package for details).

5.5 Concluding Remarks

This novel study presents a step-by-step simulation structure for selecting an optimal assumed distribution for the error term to model volatility persistence through the observation-driven GAS model using the GAS package, with relevant findings. The structure discloses an organised approach to the Monte Carlo simulation (MCS) experiment that includes "background of

the study (optional), defining the aim of the study, specifying the research questions, method of implementation, and summarised conclusion". The method of implementation is a process that consists of writing the simulation code, setting the seed, setting the true parameter a priori, data generation, and performance evaluation through meta-statistics. This handy structure is exemplified using financial returns data, which makes it easily applicable to users for effective MCS studies.

Some of the works that have been used in the literature to generally describe the simulation approach to solving problems include Chalmers and Adkins (2020); Feng and Shi (2017); Morris et al. (2019), and Sigal and Chalmers (2016), among others, but none of these has directly designed a simulation structure for volatility modelling involving autoregressive models. Specifically, whilst the empirical approach of volatility modelling involving autoregressive models is pervasive in the literature, an organised simulation framework to select an optimal conditional distribution for the error term to model volatility has been rarely studied. Hence, this tutorial study is motivated by the inclination to derive a flexible but robust simulation framework for improved volatility persistence estimation using the autoregressive GAS model. The findings in this paper may help current practitioners and upcoming researchers in finance and other sectors involved with volatility modelling to better understand the relevant simulation steps necessary for an improved volatility persistence estimation through this model. Hence, this paper has contributed to addressing inadequate literature on the relevant simulation steps needed to select a suitable conditional error distribution for modelling the persistence of volatility.

Among the findings, this study used both the fat-tails and $\sqrt{N}$ consistency experiments to show that the GAS model with a lower unconditional shape parameter or degree of freedom value ($\hat{\nu}^* = 4.1$) can generate a dataset that adequately reflects the behaviour of financial time series data, relevant for volatility modelling. This dynamic approach, to the best of the authors' knowledge, has not been used in prior works, and the finding could be a relevant guide to users on how to obtain appropriate simulated financial time series data through the GAS model. It is also observed that to determine the sample size that is adequate for a particular simulation, focus should be put on the level of accuracy or precision (in this case, on efficiency and/or consistency) of the target statistic of interest. If a set of sample sizes yields the same efficiency outcomes, the outcome with the best consistency in the set can be used for a final decision on sample size determination.

Furthermore, in line with the findings of Gilli et al. (2019) and Feng and Shi (2017), the outcomes from this study also showed that increasing the samples or sample size does not always reduce the sampling error (RMSE). The RMSE results may be strongly or weakly $\sqrt{N}$ consistent. Moreover, it is observed that the volatility persistence estimator of the GAS model displays appreciable $\sqrt{N}$ consistency. In addition, consistent with the outcomes of other autoregressive models like the GARCH models (see Feng and Shi, 2017; Samuel et al., 2023), this study also revealed that the GAS model with the Gaussian error distribution is

consistent but not efficient when the true error distribution is not Gaussian. Also, the outcomes of the illustrative studies showed that the AST1 error assumption is the most suitable among the seven assumed errors to describe the financial returns for volatility persistence modelling. The application of this (AST1) error distribution with the GAS model revealed a considerably high persistence of volatility in the SA bond returns. In a broader context, the outcome of this robust estimation approach using the AST1 fitted with the GAS specification could be useful to investors and financial practitioners in enhancing the accuracy of their risk measurement for investment decisions and portfolio management.

Lastly, the true parameter recovery (TPR) measure was used as a proxy for the coverage of the MCS experiments. This measure is a robust but flexible presentation of the concept of the estimator's coverage, and it is easy to use for measuring the level of recovery of the true parameter value by the MCS estimates. Its outcomes revealed considerably good performance of the MCS estimator at recovering the true parameter. Hence, in the absence of a reliable and accurate means of calculating the coverage of a simulation experiment involving the GAS or any other model, this study has further largely contributed to providing for such calculation through the proxy TPR measure. It is expected that this robust structure will adequately guide interested researchers on MCS experiments using the GAS model and package for selecting an optimal assumed innovation for volatility modelling in finance and other areas.

5.5.1 Limitations in the Study

The following limitations are noticed in this study. First, MCS is somewhat complex and time-demanding, hence it requires adequate planning with critical and analytical reasoning to be properly executed. Second, the quality of the simulation modelling may depend on the quality of the model used. Third, implementing the experiments may sometimes involve large computational time. Furthermore, this study shows that a considerably large sample size may sometimes be required to obtain a reasonable output approximation. It is also noticed that the GAS package does not have a provision to test for serial correlation in the returns, hence we used the Weighted Ljung-Box (WLB) statistic through the R rugarch package and the auto.arima() function in the R forecast package to carry out the test. Lastly, since the package also does not make provision for estimating the coverage of the MCS experiments, this study used the TPR measure as a proxy for the coverage, and it is observed that the MCS estimates considerably recovered the true parameters.

Chapter 6

A Comparative Modelling of Essential Characteristics of Volatility

RICHARD T. A. SAMUEL[1,*], CHARLES CHIMEDZA[1] and CASTON SIGAUKE[2]

[1] School of Statistics and Actuarial Science, University of the Witwatersrand

[2] Department of Mathematical and Computational Sciences, University of Venda

Journal paper under review: *Journal of Statistics Applications & Probability*.

Statement of Contributions of Joint Authorship

Richard T. A. Samuel (Candidate)
Conducted the research, writing and compilation of manuscript;

Charles Chimedza (Supervisor) & Caston Sigauke (Co-Supervisor)
Supervised, edited and coauthors of the manuscript.

This chapter is a copy of the journal paper referred to above.

ABSTRACT

This study utilised the dynamics of three time-varying models to estimate six essential features of financial return volatility that are relevant for robust risk management. These features include pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long-memory behaviour of volatility decomposition into long-term and short-term components. Both simulation and empirical evidence are

provided. Through the applications of these models using the S&P Indian index, the study shows that the market returns are characterised by these volatility features. The study further used a parametric model through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle estimator, and the exact local Whittle estimator to estimate and determine the presence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns. The results of the estimations indicate that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate. However, the persistence is lower in the returns when compared with the squared returns and absolute returns. Our findings from the long-memory decomposition revealed that although the response to shocks is greater in the short-term component, it is, however, short-lived. On the contrary, despite a high degree of persistence in the long-term component, market information or unexpected news arrival only has a low long-run impact on the market. Based on this, the long-run investment risks within the Indian stock market seem to be under control. Hence, our findings suggest that rational investors should try to stay calm with the arrival of unexpected news in the market because the long-run effect of such news will not be severe, and the market will eventually return to its normal state.

6.1 Introduction

Stock market indices are prone to be characterised by features of volatility such as pronounced persistence, mean reversion, leverage effect (i.e., volatility asymmetry), conditional skewness, conditional fat-tailedness, and the long-memory behaviour of volatility decomposition into long-run and short-run components (Harvey and Sucarrat, 2014). Hence, these characteristics are usually exhibited by financial returns (see Engle and Patton, 2001; Sucarrat, 2013). Volatility persistence is a process where the return of today affects the future's forecast variance (Engle and Patton, 2001). The economic implications of the degree of persistence of a shock include its influence on dynamic hedging policies, the valuation of options and the price of securities (Porterba and Summers, 1986). Mean reversion implies that there is a normal volatility level to which volatility will eventually return. Asymmetry is a process in which positive and negative shocks have different impacts on volatility. Asymmetry in equity returns is sometimes known as the leverage effect (Engle and Patton, 2001). The leverage effect implies that volatility tends to be higher following negative returns. Conditional fat-tailedness implies that the standardised conditional return is more fat-tailed than the Normal distribution, while conditional skewness denotes that the standardised return is not symmetric. The long-memory decomposition of volatility occurs when volatility is decomposed into one short-term component and one long-term component (Sucarrat, 2013).

Volatility on its own is not directly observable, hence its measurement and the modelling of its evolution rely on some measurement methods. Two well-documented methods in the literature are the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model and the Generalised Autoregressive Score (GAS) model. These two models use the conditional variance for measuring volatility (see Ardia et al., 2019; Chou, 1988; Takaishi, 2018). Moreover, it is also well documented that the Beta-Skew-t-EGARCH model can adequately model essential features of volatility (Harvey and Sucarrat, 2014; Sucarrat, 2013) for efficient risk management. Hence, these three models can reliably capture some features of asset returns. However, it is well reported that asymmetric GARCH models perform better at modelling volatility than the simple GARCH model (see Takaishi, 2018), since the simple GARCH model is incapable of capturing asymmetries in the volatility process (Feng and Shi, 2017). Therefore, this study applies a robust extension of the GARCH model called the family GARCH (fGARCH) model (Hentschel, 1995) that incorporates the asymmetric volatility process. Based on this, the fGARCH, GAS and Beta-Skew-t-EGARCH models are used to model the stated essential features of volatility using the S&P Indian daily equity returns.

To be specific, the study comparatively applies the fGARCH and GAS models to estimate the magnitude and dynamics of the persistence (with mean reversion) in the conditional volatility of the returns. The main difference between these models is that the fGARCH model uses the dynamics of the residuals to drive the conditional variance, while the GAS model uses the dynamics of the conditional score to drive the time-varying conditional variance. Furthermore, the study comparably uses the one and two components of the Beta-Skew-t-EGARCH model to estimate the asymmetry (leverage effect), skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components.

Engle and Patton (2001) used the GARCH(1,1) model to estimate various stylised facts of volatility, namely, pronounced persistence, mean-reversion and asymmetry in the Dow Jones Industrial Index returns from 1988 to 2000, and found a highly persistence volatility estimate with a half-life of about 73 days. Moreover, the tendency of volatility to persist over time is well documented in the literature (see Chou, 1988; Engle and Bollerslev, 1986; Pandey and Kumar, 2017 among others). Oh and Patton (2018) used the GAS model with a factor copula model to study systemic risk. Other GAS modelling applications are in credit risk analysis (Creal et al., 2014), spatial econometrics (Blasques et al., 2014; Catania and Billé, 2017), and high-frequency data (Gorgi et al., 2019; Opschoor et al., 2018), among others. See Ardia et al. (2019); Haddad et al. (2023) for more dynamic applications of the GAS model. The applications of the Beta-Skew-t-EGARCH model for modelling relevant features of volatility are well studied in Sucarrat (2013), and Harvey and Sucarrat (2014).

Our study should result in responses to the following questions. First, which assumed innovations are the most adequate from the fGARCH and GAS modelling to estimate the persistence of the volatility of the returns? Second, how persistent is the market returns volatility? Third, how are the models compared in terms of performance? Fourth, what volatility features

characterise the returns and what are the implications on investment in the market? This study used the daily S&P Indian index data, obtained from Thomson Reuters Datastream (2021) from January 4th, 2010, to June 18th, 2021, with a total of 2990 observations. We chose these periods so as to include the periods of both relative calm and the recent volatility spike (market turmoil) caused by the global COVID-19 pandemic. Moreover, with a vast population, a rapidly growing economy, and admirable investment opportunities, India is one of the greatest investment destinations in the world. It is projected to have the largest population globally before 2030 (Hertog et al., 2022), and its financial markets are increasingly attracting many foreign and domestic investors. The rest of the paper is structured as follows. In Section 6.2, the applied research theories and methodologies are presented. Section 6.3 presents the empirical and simulation evidence of the estimations. Section 6.5 discusses the novel findings, and Section 6.6 concludes.

6.2 Methods

6.2.1 The GARCH Model

The GARCH model (Bollerslev, 1986) is an extension of the autoregressive conditional heteroscedasticity model proposed by Engle (1982) for volatility modelling. It is usually described by its conditional mean and variance equations. The mean equation can be stated as:

$$r_t = \mu_t + \varepsilon_t, \tag{6.1}$$

where r_t denotes the returns, $\varepsilon_t = z_t\sigma_t$ is the random or unpredictable residual. The z_t is the standardized residual returns ($z_t = \varepsilon_t/\sigma_t$) which are independent, identically distributed (i.i.d.) random variables with zero mean and unit variance. The μ_t in Equation (6.1) represents the mean function and it is usually expressed as an Autoregressive Moving Average (ARMA(m, n)) process,

$$\mu_t = \sum_{j=1}^{m} \varsigma_j r_{t-j} + \sum_{j=1}^{n} \psi_j \varepsilon_{t-j}, \tag{6.2}$$

where ς_j and ψ_j are unknown parameters. The conditional variance equation of the general GARCH(p, q) model can be defined as:

$$\sigma_t^2 = \underbrace{\underbrace{\omega + \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2}_{\text{ARCH}} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2}_{\text{GARCH}}, \tag{6.3}$$

where $\alpha_j \geq 0$ and $\beta_j \geq 0$ are the ARCH and GARCH coefficients, respectively, while $\omega > 0$ is the intercept. The first-order GARCH(1,1) is possibly the best candidate and the most widely used GARCH model for modelling volatility (Harvey, 2013). The rate of decay of shocks to volatility in the conditional variance of the GARCH process is measured by summing the coefficients (α, β). This refers to the volatility persistence of the GARCH models and it indicates the speed of the decay of volatility after a shock. If the sum of the coefficients equals one, then shocks to the volatility do not decrease over time, hence the persistence is felt forever, and the unconditional variance of the process does not exist. Such a situation is called integrated GARCH (IGARCH) (Chou, 1988).

A further extension of the IGARCH process known as the fractional IGARCH (FIGARCH) was introduced by Baillie et al. (1996), where the volatility persistence is shorter than an IGARCH but longer than the standard GARCH (Baillie, 1996; Ghalanos, 2022; Rafik et al., 2014). Shocks to volatility show long persistence into the future when the sum is close to one. This produces a mean-reversion system in which the variance process (volatility), be it high or low, eventually returns very slowly to the mean (normal) state. Lastly, the process of shocks to the conditional variance shows high persistence when the sum is greater than one, which implies explosive volatility forecasts.

6.2.2 Persistence and Mean Reversion in Volatility

Volatility can be described as persistent if today's return produces a large effect on the prediction variance for many periods in the future (Engle and Patton, 2001). Volatility clustering means that small volatility shocks are followed up by small shocks while large volatility shocks are followed by large shock in turn. Hence, a period of low volatility will be followed by a volatility rise, while a period of high volatility will sooner or later make way for more normal volatility (Engle and Patton, 2001). Mean reversion in volatility implies that there is a normal volatility level to which volatility will return eventually. A familiar classical measure of volatility persistence is called the "half-life" of volatility, denoted as $h2l$ (Ghalanos, 2018). This can be described as the number of days it will take the volatility to revert or move halfway back towards its unconditional mean after deviating from it (see Engle and Patton, 2001; Ghalanos, 2018) and is given by

$$h2l = \frac{\log_e \frac{1}{2}}{\log_e \hat{P}}, \tag{6.4}$$

where $\log_e$ denotes the natural logarithm, and $\hat{P}$ represents the estimate of the persistence parameter.

6.2.3 The fGARCH Model

The family GARCH abbreviated fGARCH (Hentschel, 1995) is an omnibus model that subsumes some familiar asymmetric and symmetric GARCH models as sub-classes (Ghalanos, 2018). These sub-classes include the standard GARCH (sGARCH) model (Bollerslev, 1986), the Threshold GARCH (TGARCH) model (Zakoian, 1994), the Nonlinear Asymmetric GARCH (NAGARCH) model (Engle and Ng, 1993), the Absolute Value GARCH (AVGARCH) model (Schwert, 1990; Taylor, 1986), the Nonlinear ARCH (NGARCH) model (Higgins and Bera, 1992), the Exponential GARCH (EGARCH) model (Nelson, 1991), the GJR GARCH (GJR-GARCH) model (Glosten et al., 1993), and the Asymmetric Power ARCH (apARCH) model (Ding et al., 1993). The fGARCH(p, q) model can be stated as:

$$\sigma_t^{\gamma} = \omega + \sum_{j=1}^{p} \alpha_j \sigma_{t-j}^{\gamma} (|z_{t-j} - \zeta_{2j}| - \zeta_{1j}\{z_{t-j} - \zeta_{2j}\})^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{\gamma}. \tag{6.5}$$

As discussed in Ghalanos (2018), Equation (6.5) is related to the Box-Cox transformation of the conditional standard deviation, where the absolute value function can be transformed by the parameter δ, while the shape is determined by γ. This omnibus model allows the decomposition of the residuals to be driven by different powers for z_t and σ_t in the conditional variance equation. The model also allows for both rotations and shifts in the news impact curve, where the rotation drives large volatility shocks while the shift is the main origin of asymmetry for small volatility shocks. The parameter δ is subject to shifts and rotations through the ζ_{2j} and ζ_{1j}, respectively. The full specification of the family GARCH model can be fitted when δ = γ (see Ghalanos, 2018). The volatility persistence of the fGARCH model can be obtained through the parameter estimate $\hat{P}$, stated as:

$$\hat{P} = \sum_{j=1}^{q} \beta_j + \sum_{j=1}^{p} \alpha_j \varrho_j, \tag{6.6}$$

where ϱ_j, as stated in Equation (6.7), denotes the expectation of z_t below the Box-Cox transformation that is associated with the absolute value asymmetry term.

$$\varrho_j = \mathrm{E}(|z_{t-j} - \zeta_{2j}| - \zeta_{1j}(z_{t-j} - \zeta_{2j}))^{\delta} = \int_{-\infty}^{\infty} (|z - \zeta_{2j}| - \zeta_{1j}(z - \zeta_{2j}))^{\delta} f(z, 0, 1, \ldots) dz \tag{6.7}$$

This study used the "persistence()" function from the rugarch package of R software to estimate the persistence. The unconditional variance of the fGARCH model, as related to the persistence, is $\hat{\sigma}^2 = \hat{\omega}/(1 - \hat{P})^{2/\gamma}$ (Ghalanos, 2018). Readers can refer to Ghalanos (2018, 2022); Hentschel (1995) for more information on the nested models and the fGARCH model.

6.2.4 The GAS Model

An alternative method to the family GARCH model for modelling volatility can be found in a Score Driven (SD) model known as the Generalised Autoregressive Score (GAS) model, introduced by Harvey (2013) and Creal et al. (2013) (see Ardia et al., 2019; Catania et al., 2020). The model uses the score of the conditional density function to determine the time variation in the parameters. The score functions are robust to outliers, and the model is quite suitable for modelling skewed or fat-tailed time series data like financial returns (Alanya-Beltran, 2022; Haddad et al., 2023; Harvey, 2013; Opschoor et al., 2018). Moreover, like the other observation-driven models, extensions to long memory behaviour, asymmetric, and other time series dynamics are possible. Furthermore, likelihood estimation using the GAS model is simple and direct (Ardia et al., 2019).

6.2.5 Model Specification

Let $N \times 1$ vector $\boldsymbol{r}_t$ imply the dependent variable of interest, $\boldsymbol{\vartheta}_t$ the vector of time-varying parameter, $\boldsymbol{x}_t$ a vector of exogenous variables (i.e., the covariates), all at time t, and $\boldsymbol{\zeta}$ a vector of time-invariant parameters. Define $\boldsymbol{R}^t = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_t\}$, $\boldsymbol{\Theta}^t = \{\boldsymbol{\vartheta}_0, \boldsymbol{\vartheta}_1, \ldots, \boldsymbol{\vartheta}_t\}$, and $\boldsymbol{X}^t = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t\}$. The information set that is available at time t consists of $\{\boldsymbol{\vartheta}_t, \mathcal{F}_t\}$, where

$$\mathcal{F}_t = \{\boldsymbol{R}^{t-1}, \boldsymbol{\Theta}^{t-1}, \boldsymbol{X}^t\}, \quad \text{for } t = 1, \ldots, n. \tag{6.8}$$

It is assumed that the generation of $\boldsymbol{r}_t$ is through the observation density (Creal et al., 2012, 2013)

$$\boldsymbol{r}_t \sim p\left(\boldsymbol{r}_t | \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\zeta}\right). \tag{6.9}$$

It is further assumed that the mechanism for updating $\boldsymbol{\vartheta}_t$ (i.e., the time-varying parameter) is given by the autoregressive updating equation:

$$\boldsymbol{\vartheta}_{t+1} = \boldsymbol{\kappa} + \sum_{i=1}^{p} \boldsymbol{A}_i s_{t-i+1} + \sum_{j=1}^{q} \boldsymbol{B}_j \boldsymbol{\vartheta}_{t-j+1}. \tag{6.10}$$

Equation (6.10) is presented in Ardia et al. (2019) as:

$$\boldsymbol{\vartheta}_{t+1} \equiv \boldsymbol{\kappa} + \mathbf{A}\boldsymbol{s}_t + \mathbf{B}\boldsymbol{\vartheta}_t, \tag{6.11}$$

where $\boldsymbol{\kappa}$, $\mathbf{A}$ and $\mathbf{B}$ are matrices of coefficients with appropriate dimensions and they are functions of the static parameter $\boldsymbol{\zeta}$, while the scaled score $\boldsymbol{s}_t$ is a suitable function of past data, $\boldsymbol{s}_t = \boldsymbol{s}_t(\boldsymbol{r}_t, \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\zeta})$ (see Creal et al., 2012, 2013). Vector $\boldsymbol{\kappa}$ controls the level of the process $\boldsymbol{\vartheta}_t$, the matrix of coefficients $\mathbf{A}$ controls (or determines) the impact of $\boldsymbol{s}_t$ on $\boldsymbol{\vartheta}_{t+1}$, while matrix $\mathbf{B}$ determines the persistence of the process (Ardia et al., 2019; Koopman et al., 2017). In particular, $\boldsymbol{s}_t$ denotes the direction for updating the vector of parameters from $\boldsymbol{\vartheta}_t$ to $\boldsymbol{\vartheta}_{t+1}$,

hence, $\mathbf{A}$ can be described as the step of the update. In other words, $\boldsymbol{s}_t$ acts as a steepest ascent algorithm to improve the local fit of the model given the current parameter position (see Ardia et al., 2019). To implement the GAS models[1], the $\mathbf{A}$ and $\mathbf{B}$ matrices are constrained to exist as diagonals (see Ardia et al., 2019; Oh and Patton, 2018), e.g., for a GAS model with Student's t error distribution, $\mathbf{A} \equiv \text{diag}(a_\mu, a_\sigma, a_\nu)$ and $\mathbf{B} \equiv \text{diag}(b_\mu, b_\sigma, b_\nu)$, where μ, σ and ν are location, scale and shape parameters, respectively. Hence, b_μ refers to the persistence of the conditional mean (location), while b_σ is the persistence of the conditional variance (or scale) (see Ardia et al., 2019). This persistence parameter b_σ of the GAS model coincides with the persistence parameters $\alpha + \beta$ of the standard GARCH model[2] of Bollerslev (Ardia et al., 2019; Blasques et al., 2014; Creal et al., 2013; Oh and Patton, 2018).

The GAS approach depends on the observation density in Equation (6.9) for a given parameter $\boldsymbol{\vartheta}_t$. When an observation r_t is realised, the time-varying $\boldsymbol{\vartheta}_t$ to the next period $t+1$ can be updated using Equation (6.10) with

$$\boldsymbol{s}_t = \boldsymbol{S}_t \cdot \boldsymbol{\nabla}_t, \qquad \boldsymbol{\nabla}_t = \frac{\partial \ln p\left(\boldsymbol{r}_t | \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\zeta}\right)}{\partial \boldsymbol{\vartheta}_t}, \qquad \boldsymbol{S}_t = S\left(t, \boldsymbol{\vartheta}_t, \mathcal{F}_t; \boldsymbol{\zeta}\right), \tag{6.12}$$

where $S(\cdot)$ represents a matrix function, ln denotes the natural logarithm, $\boldsymbol{\nabla}_t$ is the score of Equation (6.9) evaluated at $\boldsymbol{\vartheta}_t$, and $\boldsymbol{S}_t$ is the scaling matrix (see Ardia et al., 2019; Creal et al., 2012, 2013; Haddad et al., 2023). Given the dependence of the stated driving mechanism in Equation (6.10) on the scaled score vector in Equation (6.12), the GAS model with orders p and q can be defined by Equations (6.9), (6.10) and (6.12). The model can be referred to as GAS (p, q) and the orders p and q are typically taken as $p = q = 1$ (see Creal et al., 2012, 2013). However, for details on including more lags in the GAS process, see Blasques et al. (2014); Creal et al. (2013).

As reported by Ardia et al. (2019), the authors Creal et al. (2013) suggested setting $\boldsymbol{S}_t$ to the inverse of the information matrix ($\mathcal{I}$) to a power $\gamma > 0$ of $\boldsymbol{\vartheta}_t$ to account for the variance of $\boldsymbol{\nabla}_t$. To be precise,

$$\boldsymbol{S}_t = \mathcal{I}_{t|t-1}^{-\gamma}, \qquad \mathcal{I}_{t|t-1} = \mathrm{E}_{t-1}\left[\boldsymbol{\nabla}_t \boldsymbol{\nabla}_t^\top\right], \tag{6.13}$$

where the expectation E_{t-1} is taken with respect to the conditional distribution of $\boldsymbol{r}_t | \boldsymbol{r}_{1:t-1}$. The parameter γ normally takes value in the set $\{0, \frac{1}{2}, 1\}$. However, other choices of S_t are possible as well (see Blasques et al., 2014). When $\gamma = 0$, $\boldsymbol{S}_t = \mathbf{I}$ (identity matrix), which means there

[1]The GAS model is implemented in this study with the use of the R package GAS developed by Ardia et al. (2019).

[2]The GAS model with assumed Normal distribution coincides with the standard GARCH(1,1) model of Bollerslev (Bollerslev, 1986) (see Ardia et al., 2019; Blasques et al., 2014; Creal et al., 2013; Oh and Patton, 2018). Hence, we investigated this by comparing the estimate of the persistence $\hat{b}_\sigma$ from the GAS model fitted with a time-varying scale parameter, and the estimate $\hat{\alpha}_1 + \hat{\beta}_1$ from the GARCH(1,1) model. Both models were fitted to the real return S&P Indian stock data under the Normal error, and their outcomes yielded $\hat{b}_\sigma \equiv \hat{\alpha}_1 + \hat{\beta}_1 \approx 0.97$.

is no scaling. If $\gamma = \frac{1}{2}$ ($\gamma = 1$), then the conditional score $\boldsymbol{\nabla}_t$ is pre-multiplied by the square root of (the inverse of) its covariance matrix $\mathcal{I}_t$. However, whatever the choice of γ, $\boldsymbol{s}_t$ is a martingale difference with respect to the distribution of $\boldsymbol{r}_t|\boldsymbol{r}_{1:t-1}$, i.e., $\mathtt{E}_{t-1}[\boldsymbol{s}_t] = 0$ for all t (see Ardia et al., 2019). The GAS framework embodies many available observation-driven models in the literature for a suitable choice of the scaling matrix $\boldsymbol{S}_t$ (Creal et al., 2013; Haddad et al., 2023). Readers can refer to Ardia et al. (2019); Blasques et al. (2017); Catania et al. (2020); Creal et al. (2012, 2011, 2013); Haddad et al. (2023); Harvey (2013); Koopman et al. (2017) for more details on the GAS model.

6.2.6 The Beta-Skew-t-EGARCH Model

To begin with, the Beta-t-EGARCH as a model without skewness was proposed by Harvey and Chakravarty (2008), and Harvey (2013). It is an extension made to the EGARCH model where an equation that depends on the conditional score of the last observation drives the variance or scale (Creal et al., 2011, 2008). The Beta-t-EGARCH model can be regarded as an unrestricted version of the GAS model of Creal et al. (2013) (see Sucarrat, 2013). In other words, it is a dynamic model driven by a conditional score, which is a martingale difference. A martingale difference has a constant (or zero) conditional expectation (Harvey, 2013). Due to various implications of conditional skewness on asset pricing, conditional score models were extended to skew distributions (Harvey and Sucarrat, 2014). Hence, the Beta-Skew-t-EGARCH model is a skewed version of the Beta-t-EGARCH model, where skewness can be brought in through the method introduced by Fernández and Steel (Fernández and Steel, 1998) (see Harvey and Sucarrat, 2014; Sucarrat, 2013).

A number of useful and attractive properties are attributed to this skewed model. As in the Beta-t-EGARCH model where observations that could be seen as outliers for a Normal distribution are down-weighted with the use of the conditional score (Harvey and Sucarrat, 2014), the Beta-Skew-t-EGARCH model also displays robustness to outliers or jumps. In particular, it performs quite well in modelling key stylised facts of financial returns like fat-tails, leverage effect, conditional skewness and the long-memory behaviour of volatility decomposition into long-term and short-term components. The model has two versions, described below as "one-component and two-component models".

6.2.6.1 The One-Component Beta-Skew-t-EGARCH model

The three equations below illustrate the first order Beta-Skew-t-EGARCH model's martingale difference version (Harvey and Sucarrat, 2014).

$$R_t = \exp(\lambda_t) z_t = \sigma_t z_t, \quad z_t \sim st(0, \sigma_z^2, \nu, \eta), \quad \nu > 2, \quad \eta \in (0, \infty), \tag{6.14}$$

$$\lambda_t = \omega + \lambda_t^{\dagger}, \tag{6.15}$$

$$\lambda_t^{\dagger} = \phi_1 \lambda_{t-1}^{\dagger} + \kappa_1 u_{t-1} + \kappa^* sgn(-R_{t-1})(u_{t-1} + 1), \quad |\phi_1| < 1, \tag{6.16}$$

where R_t[1] is the demeaned return, σ_t represents the volatility or conditional scale, and the conditional error z_t is not standardised i.e., the variance is not one. Equation (6.14) shows that z_t follows the skew Student's t distribution with a mean of zero, scale σ_z^2, degrees of freedom ν, and skewness parameter η. The z_t is defined as $z_t = z_t^* - \mu_{z^*}$, where z_t^* is an uncentred[2] skew Student's t variable with ν degrees of freedom, mean μ_{z^*}, and skewness parameter η. The ω denotes the log-scale intercept that is described as the long-term log-volatility, while ϕ_1 represents the persistence parameter (the bigger the persistence, the more the clustering of volatility). The κ_1 is the ARCH parameter that indicates the response to shocks (the bigger the absolute value of the parameter, the greater the response of volatility to shocks), κ^* is the leverage parameter, u_t denotes the conditional score, that is, the derivative of the log-likelihood of R_t at t with respect to λ_t, and *sgn* is the sign function. Details on this model are discussed in Sucarrat (2013).

6.2.6.2 The Two-Component Beta-Skew-t-EGARCH Model

In order to accommodate the long memory feature of financial return like a two-component GARCH model of Engle and Lee (1999), the two-component Beta-Skew-t-EGARCH model was introduced. The model is used to decompose volatility persistence into long-term (or long-run) and short-term (or short-run) components (Alizadeh et al., 2002). The first order two-component Beta-Skew-t-EGARCH model's martingale difference version (Harvey and Sucarrat, 2014) is given as:

$$R_t = \exp(\lambda_t) z_t = \sigma_t z_t, \quad z_t \sim st(0, \sigma_z^2, \nu, \eta), \quad \nu, \eta \in (0, \infty), \tag{6.17}$$

$$\lambda_t = \omega + \lambda_{1,t}^{\dagger} + \lambda_{2,t}^{\dagger}, \tag{6.18}$$

$$\lambda_{1,t}^{\dagger} = \phi_1 \lambda_{1,t-1}^{\dagger} + \kappa_1 u_{t-1}, \quad |\phi_1| < 1, \tag{6.19}$$

$$\lambda_{2,t}^{\dagger} = \phi_2 \lambda_{2,t-1}^{\dagger} + \kappa_2 u_{t-1} + \kappa^* sgn(-R_{t-1})(u_{t-1} + 1), \quad |\phi_2| < 1, \quad \phi_1 \neq \phi_2, \tag{6.20}$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ represent the time-varying long-run and short-run components of log-volatility, respectively. Both components are driven by the conditional score u_t. The ϕ_1 and

[1]The demeaned return R_t is the same as the error term ε_t (i.e., the unpredictable part of return) in Equation (6.1), but with z_t distributed as a skew Student's t, and it is not standardised to have unit variance (see Sucarrat, 2013).

[2]That is, the mean is not necessarily equal to zero.

ϕ_2 are the long-term and short-term persistence parameters and the model is not identifiable if $\phi_1 = \phi_2$. The κ_1 and κ_2 are parameters that indicate the long-run and short-run responses to shocks, respectively.

6.2.6.3 The Skew Student's t Distribution

Let an ordinary (i.e., symmetric and centred) Student's t distributed variable with unit scale be denoted by ϵ^*, and let its density be denoted by $f(\epsilon^*)$. The density of an uncentred skew Student's t variable can be stated through the skewing technique of Fernández and Steel (Fernández and Steel, 1998) as:

$$f(z^*|\eta) = \frac{2}{\eta + \eta^{-1}} f\left(\frac{z^*}{\eta^{sgn(z^*)}}\right). \tag{6.21}$$

The right skewness is produced when $\eta > 1$, and the left skewness is attained when $\eta < 1$, whereas symmetry can be attained when $\eta = 1$ (Harvey and Sucarrat, 2014; Sucarrat, 2013). That is, the right (left) hand tail is heavier when $\eta > 1$ ($\eta < 1$).

6.2.7 The True Parameter Recovery Measure

Monte Carlo simulation (MCS) studies largely focus on the estimator's ability to recover the true data-generating parameter (see Chalmers, 2019). Hence, the True Parameter Recovery (TPR) measure was introduced by Samuel et al. (2023) as a way of measuring how the MCS estimator performs at recovering the true parameter. The measure is used as a proxy for the coverage[1] of the MCS experiment to calculate the level of recovery of the true parameter by the MCS estimator. It can be stated (see Samuel et al., 2023) as:

$$\text{TPR} = \left(K - \left[\frac{(\vartheta - \hat{\vartheta})}{\vartheta} \times K\right]\right)\%, \tag{6.22}$$

where ϑ is the true data-generating parameter, $\hat{\vartheta}$ denotes the estimator, and $K = 0, 1, 2, \ldots, 100$ represents the nominal recovery level. A TPR outcome of 90% or 95%, for instance, implies that the MCS estimator is able to recover 90% or 95% of the true parameter. The MCS estimator $\hat{\vartheta}$ will fully recover the true parameter $\vartheta > 0$ when $\hat{\vartheta} = \vartheta$, such that the outcome of the TPR is equal to the specified nominal recovery level K (that is, $K\%$ = TPR) (see Samuel et al., 2023).

[1] The probability that the true parameter is contained within a confidence interval of estimates is known as the coverage probability (Hilary, 2002).

6.3 Experimental Results: Empirical and Simulation

This section presents the outcomes of the estimations through the applications of the fGARCH, GAS and Beta-Skew-t-EGARCH models. The three applied models can be estimated through the Maximum Likelihood Estimation (MLE) once a distribution for the innovations is specified.

6.3.1 Application of the fGARCH Model

In this section, we present the outcomes of the estimations involving the fGARCH model. It is believed that observation-driven[1] models like the (family) GARCH can yield efficient outcomes when fitted with the true (or appropriate) innovation distribution (Bollerslev, 1987; Feng and Shi, 2017). Hence, among ten selected innovations, we use this study to determine an optimal or the most suitable assumed innovation distribution that is relevant for volatility persistence estimation through the fGARCH model. The ten selected assumed innovations are the Gaussian (or Normal), Student's t, Generalised Error Distribution (GED), skew-Normal, skew-Student's t, skew-GED, Johnson's reparametrised SU (JSU) distribution, Generalised Hyperbolic (GHYP) distribution, Generalised Hyperbolic Skew-Student's t (GHST) distribution, and Normal Inverse Gaussian (NIG). See Ashour and Abdel-hameed (2010); Azzalini (1985); Azzalini and Capitanio (2003); Barndorff-Nielsen et al. (2013); Branco and Dey (2001); Eling (2014); Ghalanos (2018); Lee and Pai (2010); Pourahmadi (2007) for details on the assumed innovations.

The estimation is carried out using the actual return data from the S&P Indian stock index. To transform the daily closing price data to the log returns, we take the log-difference of the value of the index as:

$$r_t = \ln\left(\frac{P_{t+1}}{P_t}\right) \times 100, \tag{6.23}$$

where P_t is the daily closing equity price at time t, ln is the natural logarithm and r_t is the current returns.

6.3.1.1 Exploratory Data Analysis

We begin with the visual inspections of the price level and the index returns over the sample period using exploratory data analysis (EDA) as revealed in Panels (A $-$ F) of Figure 6.1. The EDA gives relevant insights into the dataset to disclose vital information such as detecting possible outliers. The EDA shows that the market is characterised by time-varying volatility, with a steep plunge in volatility of price and return (see Panels A and B) in 2020 due to the emergence of the global COVID-19 pandemic crisis. Volatility clustering is apparent in the return series, where large (small) changes tend to follow large (small) changes of either sign. The quantile-quantile (QQ) plot in Panel C shows that the returns are non-Normally distributed,

[1] Observation-driven models are developed to model large changes (which may occur in the form of jumps or shifts) and distributional asymmetries that often exist in financial time series (Haddad et al., 2023).

while the density plot in Panel D denotes stability (stationarity) in the return series. The return series plot in Panel B also reveals stationarity in the returns. Panel E displays the correlogram of the returns, showing the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. The plots indicate weak dependence in the mean of the series. Hence, we assume a constant conditional mean (see Engle and Patton, 2001). As for the squared returns, the correlogram as displayed in Panel F shows some moderate dependence at the initial lags, hence we fit an ARMA model as presented in Section 6.3.1.4 to capture this.

6.3.1.2 Descriptive Statistics

Panel A of Table 6.1 presents some descriptive statistics on the return data. The table shows that the mean of the daily return is close to zero. The daily variance is 1.6383, which implies an average annualised volatility of 20.32% [1]. The skewness coefficient shows that the distribution of returns is significantly negatively skewed, a familiar feature of stock returns. This outcome reveals the impact of information arrival in the market, and it shows that investors and other market participants tend to react more to bad news than they do to good news. The kurtosis coefficient is very high. Kurtosis is a measure of the thickness of the tails of the return distribution, and the result shows evidence of leptokurtosis with a value greater than three. That is, it surpasses the kurtosis of the Normal distribution, which is three, and that suggests a fat-tailed distribution for describing the return series. Lastly, the outcome of the Shapiro-Wilk test rejects the assumption of normality with a p-value = 0. Hence, this implies that the assumption of a Normal distribution for modelling the volatility persistence of the returns is not realistic.

6.3.1.3 Tests for Stationarity

This study further used three stationarity tests involving the Augmented-Dickey Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) tests to test for stationarity in the price and return series. We begin with the test for stationarity of the price index in Panel A of Table 6.2. The table displays the outcomes of the ADF and PP unit root tests, and they indicate that the null hypothesis of non-stationarity, which implies the presence of a unit root in the price series, cannot be rejected because the p-values are greater than 0.05. Hence, the price index is non-stationary. Furthermore, the KPSS test that is based on the null hypothesis of stationarity is rejected because the p-value < 0.05 (see Panel A of Table 6.2). This further indicates that the price index is non-stationary.

Next, for the return series in Panel B of Table 6.2, the results of the ADF and PP unit root tests indicate that the null hypothesis of non-stationarity, which implies the presence of a unit root in the returns is rejected because the p-values are less than 0.05. Hence, the return series is

[1] Annualised volatility for 252 trading days (i.e., one year) $= \sqrt{252} \times \sqrt{\text{variance}} = 20.32\%$ (see Engle and Patton, 2001).

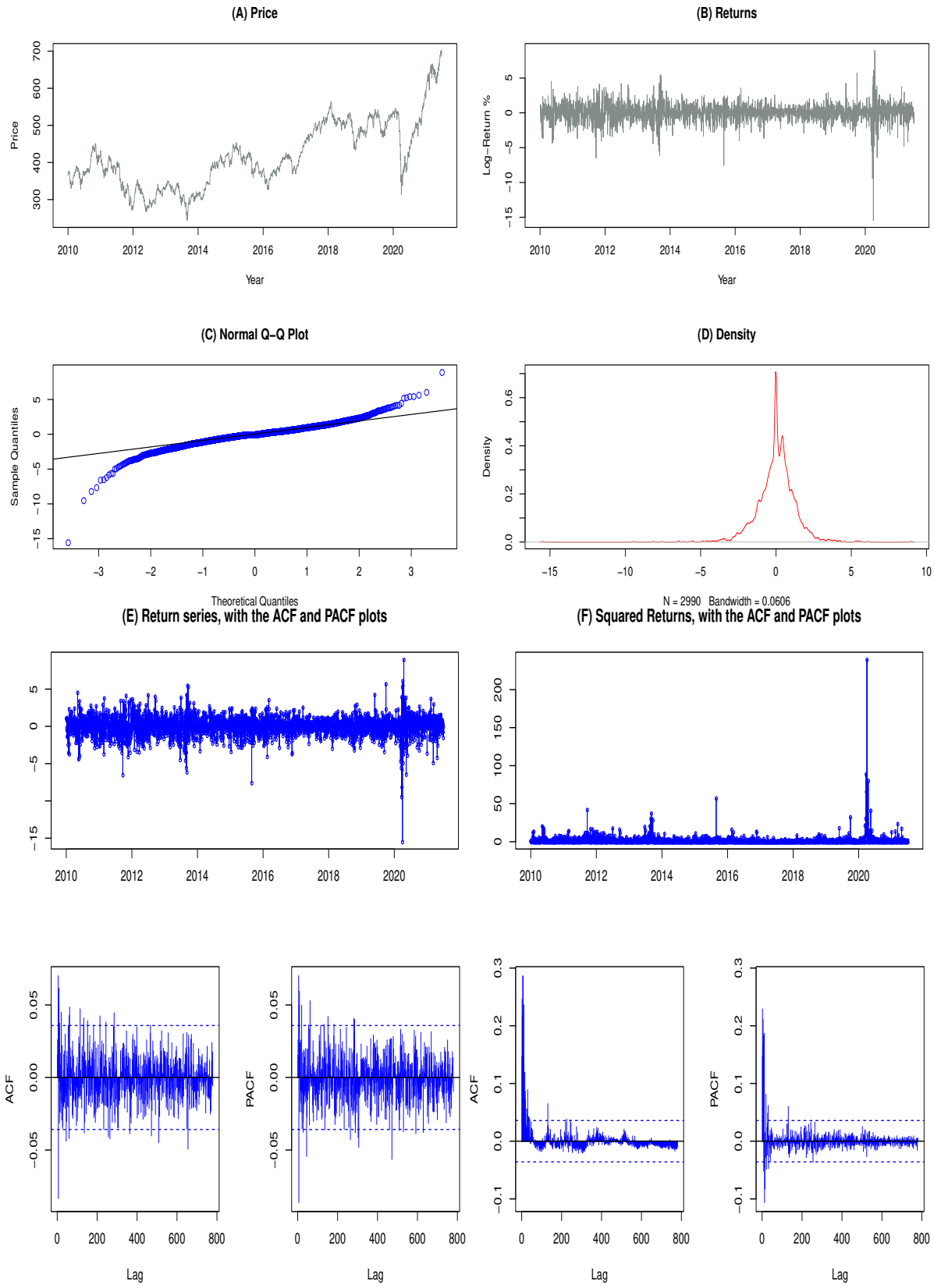

Figure 6.1: Panels (A) and (B) respectively show the plots of the price and return series, while Panels (C) and (D) display the quantile-quantile (Q-Q) and density plots, respectively, for the S&P Indian Index. Panels (E) and (F) display the ACF and PACF plots of the returns and squared returns, respectively.

Table 6.1:

Panel A Returns Descriptive Statistics			**Panel B** Engle's ARCH Test Outcomes				
				PQ test		LM test	
			Lag order	PQ	P-value	LM	P-value
Mean	0.0215		4	452	0	3446	0
Variance	1.6383		8	1043	0	1130	0
Skewness	-0.9525		12	1361	0	641	0
Kurtosis	11.6371		16	1434	0	443	0
Shapiro-Wilk	0.9222		20	1487	0	343	0
p-value	0.0000		24	1502	0	283	0

Note: LM is the Lagrange-Multiplier statistic, while PQ is the Portmanteau-Q statistic in Panel B.

stationary. Also, we cannot reject the KPSS test with a null hypothesis of stationarity because the p-value > 0.05, as presented in Panel B of Table 6.2. This further implies that the return series is stationary.

Table 6.2: Tests for Stationarity of the Price Index and Return Series.

Panel A Price Index				**Panel B** Returns		
Estimate	Lag order	P-value		Estimate	Lag order	P-value
ADF Test				ADF test		
-2.4231	9	0.3992		-15.6340	9	0.0100
PP Test				PP Test		
-10.9580	9	0.4986		-3072.1	9	0.0100
KPSS Test				KPSS Test		
19.6090	9	0.0100		0.1052	9	0.1000

Note: Panel A comprises of the outcomes of the ADF and PP unit root tests, and the outcomes of the KPSS stationarity test for the price index. Panel B contains the outcomes of the three tests for the returns.

6.3.1.4 Tests for Autocorrelation and ARCH Effects

Next, autocorrelation and ARCH effects (or heteroscedasticity) are removed by fitting ARMA-fGARCH models with each of the ten innovations to the stationary returns. ARMA(1,1) model as shown in Equation (6.24) is used as the best, among the candidates' ARMA(m,n) models, to capture autocorrelation in the residuals. Table 6.3 shows the results of the Weighted Ljung-Box test (Fisher and Gallagher, 2012) on both the standardised residuals (denoted as WLB SRs) and the standardised squared residuals (WLB SSRs) from the fit of the ARMA(1,1) model. The table shows that for both instances, the p-values of the test at lag order 5 are greater than 5% under the ten innovation distributions. Based on this, the null hypothesis of "no serial correlation" in the returns cannot be rejected. Hence, serial correlation is removed in the residuals.

$$r_t = \varsigma_0 + \varsigma_1 r_{t-1} + \psi_1 \varepsilon_{t-1} + \varepsilon_t \tag{6.24}$$

After removing serial correlation in the returns, we carry out Engle's ARCH test (Engle, 1982) using the Portmanteau-Q (PQ) and Lagrange Multiplier (LM) tests to examine the existence of ARCH effects or heteroscedasticity in the residuals. Both tests are carried out based on the null hypothesis of homoscedasticity in the residuals of an ARIMA[1] model. The results from the two tests reveal highly significant p-values of 0 from lag order 4 to 24 as shown in Panel B of Table 6.1. Therefore, we reject the null hypothesis of "no ARCH effects" in the residuals, which implies the presence of volatility clustering in the returns. Based on this, we fit candidates' fGARCH(p, q) models, with each of the ten distributed errors, to capture the ARCH effects in the returns. Hence, following Engle and Patton (2001), we use the Bayesian Information Criterion (BIC) (Li et al., 2009) and found that the best candidate model, to capture heteroscedasticity in the returns, among the fGARCH(p, q) class for $p \in$ [1,5] and $q \in$ [1,2] is the parsimonious fGARCH(1,1) (see Equation 6.25).

$$\sigma_t^\gamma = \omega + \alpha_1 \sigma_{t-1}^\gamma (|z_{t-1} - \zeta_{21}| - \zeta_{11}\{z_{t-1} - \zeta_{21}\})^\delta + \beta_1 \sigma_{t-1}^\gamma. \tag{6.25}$$

6.3.1.5 Residual Diagnostic Test

Following the fit of the fGARCH(1,1) model to the returns, we carry out some diagnostic checks using the weighted ARCH LM test to ascertain if heteroscedasticity has been removed. Based on the diagnostic results, it can be seen from Table 6.3 that the p-value of the "ARCH LM (5)" statistic at lag order 5 exceeds 5% under each of the ten assumed errors. Hence, this implies that the ARCH effect is captured since we cannot reject the null hypothesis of "no ARCH effects" in the residuals.

[1] ARIMA is Autoregressive Integrated Moving Average.

6.3.1.6 Selection of an Optimal Error Distribution

Next, the selection of the most adequate innovation distribution (among the ten innovation assumptions) that can be used to describe the returns for estimating the volatility persistence is carried out in Table 6.3. Models' comparison and selection are made using four information criteria that consist of the Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), Shibata Information Criterion (SIC), and Hannan-Quinn Information Criterion (HQIC) (see Ghalanos, 2018; Li et al., 2009). The assumed innovation with the minimum (or lowest) value of information criteria will be the most adequate innovation distribution required to estimate the persistence. To begin with, Table 6.3 shows that all, but one, of the fGARCH volatility parameter estimates ($\hat{\omega}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, $\hat{\zeta}_{11}$, and $\hat{\zeta}_{21}$) under the ten assumed errors are highly significant at 1% level. To be precise, the only exclusions to the 1% significance are the $\hat{\omega}$ which is significant at 5% under the Student's t, and the $\hat{\zeta}_{11}$ which is nearly insignificant across the board. These highly significant outcomes in the parameters reveal the presence of volatility clustering in the conditional variance. Moreover, the strongly significant $\hat{\zeta}_{21}$ indicates that the market is driven more by short volatility shocks. In other words, the effects of short volatility shocks are more pronounced in the market than those of large shocks.

For model selection, Table 6.3 shows that the four information criteria have their lowest values under the NIG innovation distribution. Therefore, the NIG innovation (fitted with the fGARCH(1,1) model) is the most suitable to describe the market returns when the underlying error distribution is unknown for volatility modelling. This result is evident in the claim that the NIG distribution is analytically tractable, and it can adequately model the skewness of financial market variables, such as equity prices, exchange rates, and interest rates (Aas and Haff, 2006). Moreover, it has been applied many times for financial applications both as the unconditional return distribution and as the conditional distribution of a GARCH model (see Aas and Haff, 2006). Hence, this makes the distribution attractive for financial modelling applications.

6.4 The Diagnostic Plots of ARMA-fGARCH Models

Figure 6.2 displays the diagnostic densities of the standardised residuals and the diagnostic QQ plot of the ARMA-fGARCH models fitted with the most suitable NIG assumed innovation distribution. The QQ plot shows a suitable level of linearity since the observations are considerably well aligned on the straight line, except at the tails (which suggests extreme movement). Furthermore, the plots of the densities of the standardised residuals confirm that this most appropriate NIG error distribution is a better fit to the Indian market's returns distribution than a Normal distribution (Sigauke, 2016).

Table 6.3:

	(A) Normal	(B) skew-Normal	(C) Student's t	(D) skew-Student's t	(E) GED
$\hat{\phi}_0$	0.0045	-0.0073	0.0462**	0.0110	0.0389**
$\hat{\phi}_1$	0.5358*	0.5260**	0.3417	0.3545	0.2327*
$\hat{\varphi}_1$	-0.4668**	-0.4612***	-0.2724	-0.2872	-0.1767*
$\hat{\omega}$	0.0482*	0.0453*	0.0441**	0.0447*	0.0460*
$\hat{\alpha}$	0.0769*	0.0844*	0.0871*	0.0903*	0.0812*
$\hat{\beta}$	0.8046*	0.8168*	0.8054*	0.8075*	0.8093*
$\hat{\zeta}_{11}$	-0.1721	-0.1842***	-0.1682	-0.1893	-0.1725
$\hat{\zeta}_{21}$	1.3867*	1.3282*	1.3281*	1.3301*	1.3288*
$\hat{\gamma} = \hat{\delta}$	2.1253*	1.9995*	1.9723*	2.0131*	2.0521*
Persistence ($\hat{P}$)	0.9726	0.9735	0.9728	0.9760	0.9711
WLB SRs(5)	0.7405	0.6924	0.7014	0.8040	1.5361
p-value(5)	(1.0000)	(1.0000)	(1.0000)	(1.0000)	(0.9973)
WLB SSRs(5)	2.927	3.039	2.761	3.094	3.034
p-value(5)	(0.4205)	(0.4000)	(0.4524)	(0.3902)	(0.4009)
ARCH LM (5)	3.9170	3.8687	3.4943	3.5942	3.7311
p-value(5)	(0.1816)	(0.1862)	(0.2257)	(0.2144)	(0.1999)
AIC	3.0541	3.0426	3.0085	3.0006	3.0052
BIC	3.0722	3.0627	3.0286	3.0227	3.0253
SIC	3.0541	3.0426	3.0085	3.0006	3.0052
HQIC	3.0606	3.0499	3.0157	3.0086	3.0124
	(F) skew-GED	**(G) GHYP**	**(H) NIG**	**(I) GHST**	**(J) JSU**
$\hat{\phi}_0$	0.0064	0.0108	**0.0106**	-0.0027	0.0097
$\hat{\phi}_1$	0.2684**	0.3145	**0.3208**	0.3712	0.3312
$\hat{\varphi}_1$	-0.2017***	-0.2501	**−0.2556**	-0.3052	-0.2651
$\hat{\omega}$	0.0458*	0.0447*	**0.0443***	0.0434*	0.0442*
$\hat{\alpha}$	0.0831*	0.0878*	**0.0900***	0.0963*	0.0916*
$\hat{\beta}$	0.8116*	0.8104*	**0.8113***	0.8187*	0.8112*
$\hat{\zeta}_{11}$	-0.1796	-0.1910	**−0.1919**	-0.1998***	-0.1918
$\hat{\zeta}_{21}$	1.3298*	1.3263*	**1.3191***	1.2903*	1.3151*
$\hat{\gamma} = \hat{\delta}$	2.0496*	2.0335*	**1.9993***	1.8975*	1.9821*
Persistence ($\hat{P}$)	0.9735	0.9748	**0.9749**	0.9756	0.9754
WLB SRs(5)	1.1550	1.0025	**0.9421**	0.8057	0.8857
p-value(5)	(0.9999)	(1.0000)	**(1.0000)**	(1.0000)	(1.0000)
WLB SSRs(5)	3.146	3.171	**3.127**	3.029	3.100
p-value(5)	(0.381)	(0.3767)	**(0.3843)**	(0.4017)	(0.3891)
ARCH LM (5)	3.8738	3.6558	**3.6329**	3.7242	3.6114
p-value(5)	(0.1857)	(0.2078)	**(0.2102)**	(0.2006)	(0.2125)
AIC	3.0008	2.9995	**2.9990**	3.0030	2.9992
BIC	3.0229	3.0236	**3.0210**	3.0251	3.0213
SIC	3.0008	2.9994	**2.9989**	3.0030	2.9992
HQIC	3.0087	3.0081	**3.0069**	3.0110	3.0072

Note: WLB SRs (SSRs) denote the Weighted Ljung-Box test for standardised residuals (standardised squared residuals), where "(5)" is lag order 5. The p-value at 5% level of significance is presented in the round bracket for each error. The "$*$", "$**$" and "$***$" are 1%, 5%, and 10% significance levels, respectively. The information criteria are computed as AIC = $-2L/N + 2p/N$, BIC = $-2L/N + p\log_e(N)/N$, HQIC = $-2L/N + 2p\log_e(\log_e(N))/N$, and SIC = $-2L/N + \log_e((N+2p)/N)$, where N denotes the sample size, L is the log-likelihood of the maximum likelihood of unknown parameter vector $L(\Theta)$, and p denotes the number of estimated parameters (Ghalanos, 2018; Li et al., 2009).

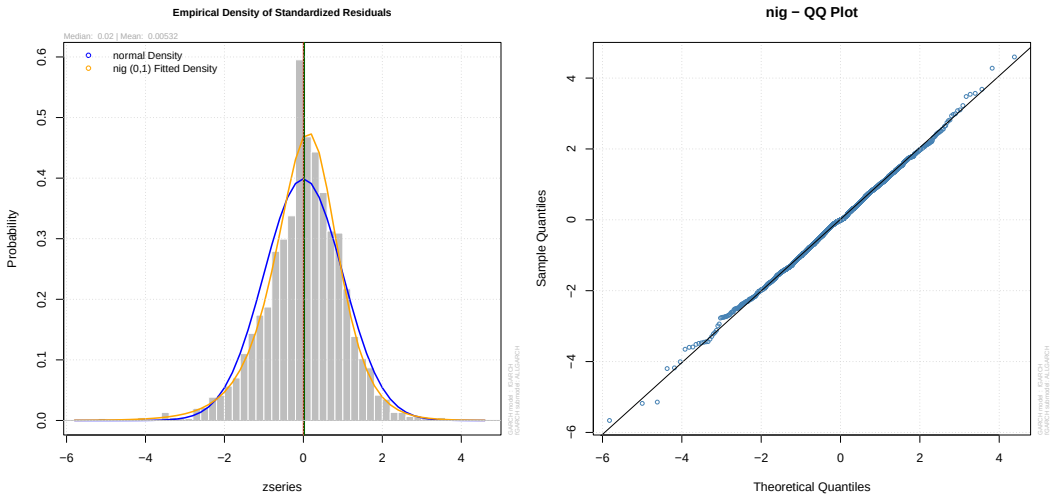

Figure 6.2: The diagnostic plot of the densities of the standardised residuals and the diagnostic quantile-quantile (QQ) plot for the ARMA-fGARCH models fitted with the NIG assumed innovation.

6.4.1 Persistence in Volatility and Mean Reversion

The estimated volatility persistence under this optimal NIG error distribution is 0.9749. This indicates that the volatility of the S&P Indian equity market's returns displays considerably high persistence. To determine if a model has adequately captured all the persistence present in the variance of returns, Engle and Patton (2001) suggested that the standardised squared residuals (SSRs, henceforth) should be serially uncorrelated. Hence, from our outcomes in Table 6.3, the p-value of the Weighted Ljung-Box test of the SSRs is greater than 5% under the NIG innovation (and generally under all the remaining nine assumed innovation distributions)[1], implying that the SSRs are serially uncorrelated. This indicates that the fGARCH(1,1) model fitted with the NIG assumed innovation has adequately captured all of the persistence present in the variance of returns.

Broadly speaking, the outcomes of our study show that the mean and variance equations are well specified, with no evidence of correlation in the standardised residuals (SRs) and SSRs of the model. The outcomes of the calculated ARCH LM test also show that the ARCH effect is filtered out in the residuals since we cannot reject the null hypothesis of "no ARCH effects" based on the ARCH LM results. Thus, the estimated fGARCH(1,1) model in this study is adequately specified and correctly fitted, and the persistence is also adequately captured by the model. Figure 6.3 shows the conditional volatility of the fGARCH(1,1) model fitted with the NIG assumed distribution, and it displays alternating phases of lower and higher volatility in the returns.

[1]These outcomes are observed up to the ninth lags for both SRs and SSRs under the ten error assumptions.

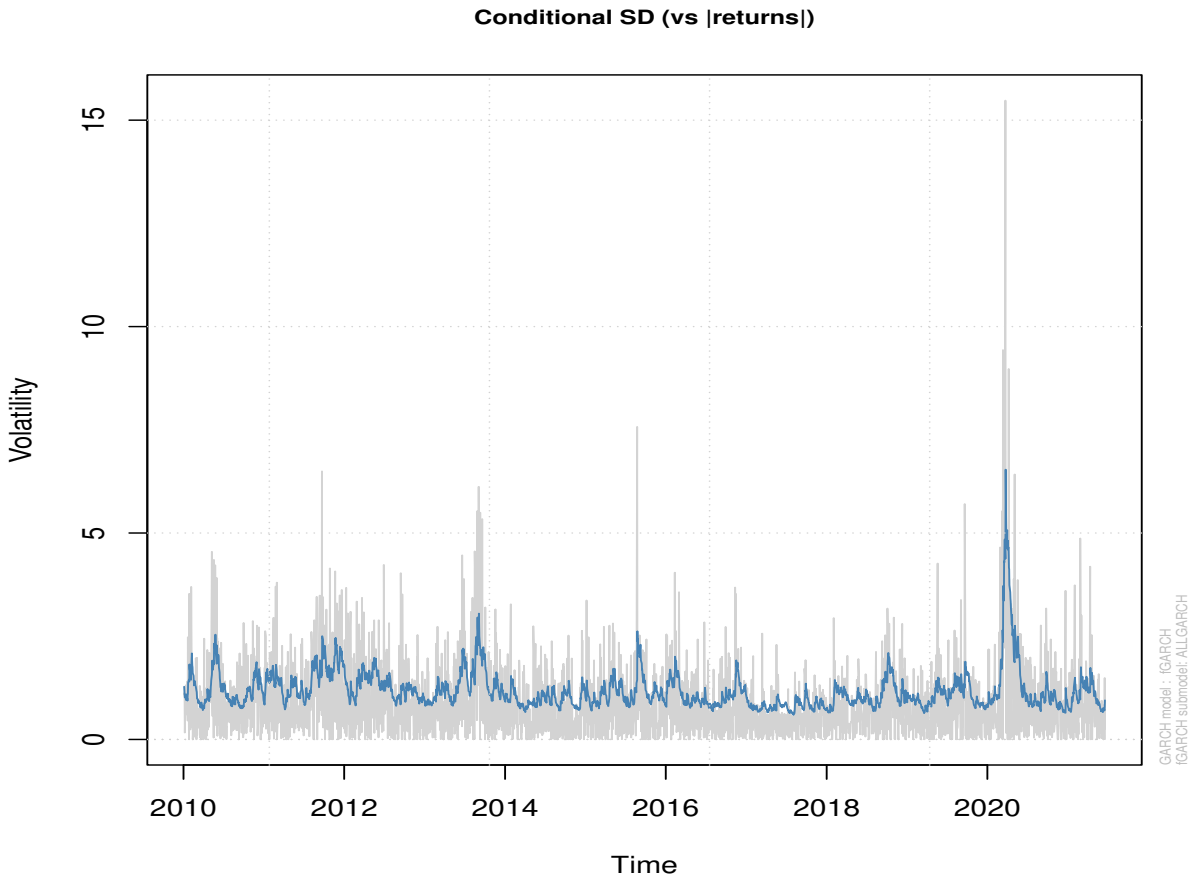

Figure 6.3: Conditional volatility of the fGARCH(1,1) model fitted with the NIG assumed error.

The estimated outcome of 0.9749 is reasonably close to one, which is a sign of considerable persistence in the volatility. The existence of long memory in financial assets' volatility relates with the persistence of shocks in returns (Bobeica and Bojesteanu, 2008). However, a crucial test to determine the presence of a long memory feature is to examine it in the return and volatility of the series (Rafik et al., 2014). Hence, this study used the semiparametric and parametric approaches to estimate the fractional differencing parameter d to determine the existence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns (Brania and Gurgul, 2015). Following the steps of Bobeica and Bojesteanu (2008), we used a parametric approach through the FIGARCH model, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (1983) (GPH, henceforth) and Robinson (1995a), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator. The long memory outputs are obtained by calculating the fractional differencing parameter d in the returns and volatility using these estimators. We decided to use the LW and ELW methods as improvements to the GPH technique due to the acclaimed bias in the estimation of the GPH estimator (see Agiakloglou et al., 1992; Baillie, 1996; Boubaker et al., 2021).

The parametric FIGARCH model is applied with fractional white noise ARFIMA$(0, d, 0)$

specification (Cheung, 2008). For the semiparametric approach, we follow Geweke and Porter-Hudak (1983) to use the sample selection function $m = g(T) = T^{\alpha}$, $0 < \alpha < 1$ (see Shea, 1991). Hence, we used the optimal rate $m = T^{4/5}$ for the number of frequencies, with a power equal to 0.8 (see Andrews and Guggenberger, 2003; Boubaker et al., 2021; Hurvich et al., 1998; Leschinski et al., 2019). This is also the default bandwidth applied in Vera-Valdés (2024), where m is the bandwidth that specifies the number of Fourier frequencies and T is the sample size (see Leschinski et al., 2019).

The outcomes of the estimation of d in the ARFIMA process by the GPH, LW, ELW, and FIGARCH methods for the returns, absolute returns, and squared returns are reported in Table 6.4. The three semiparametric estimators yield closely related results for d within 0 and 0.5, indicating the presence of long memory in the series. The outcomes of d through the parametric ARFIMA-FIGARCH method are also within 0 and 1, indicating reasonable evidence of long memory in the conditional variance of the series. These results suggest that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate. However, the persistence is lower in the returns compared to the squared and absolute returns.

Table 6.4: Long memory estimation through the fractional integration parameter d.

	Returns	Absolute returns	Squared returns
Panel A: Semiparametric estimation methods			
GPH estimator	0.0389	0.2828	0.2749
LW estimator	0.0347	0.2611	0.2461
ELW estimator	0.0504	0.2779	0.2628
Panel B: Parametric estimation method			
ARFIMA-FIGARCH process	–	0.1178	0.1278

Note: Estimated outcomes of the fractional integration parameter d for the returns, absolute returns, and squared returns (conditional volatility) using the semiparametric (in Panel A) and parametric (in Panel B) approaches.

The obtained outcome of 0.9749 indicates that the market's returns volatility displays considerable persistence, with a volatility half-life of about 27 days. This suggests the presence of considerable long memory in the return's volatility, but it is still mean reverting since the outcome is significantly less than one. This implies that even if it takes a while, the volatility process does go back to its mean (see Engle and Patton, 2001). With this decay rate of

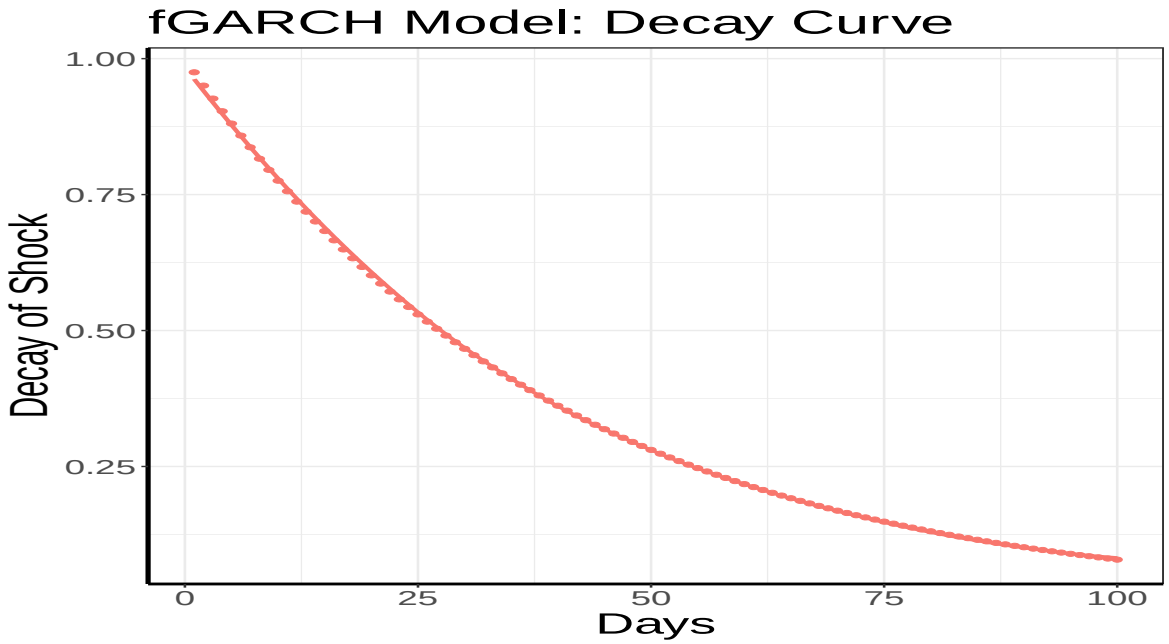

Figure 6.4: The decay of volatility persistence through the fGARCH(1,1)-NIG model.

0.9749, we plot the decay process using the $(0.9749)^{\text{day}}$ following the method of Chen and Shen (2000), Chiang et al. (2006), and Chou (1988) for the first 100 days as shown in Figure 6.4. The figure displays the decay of shock impacts through the fGARCH(1,1)-NIG model, and it shows that the impact decays slowly (see Engle and Patton, 2001), dropping to half the intensity in about 27 days. The limit of the decay sequence is zero (Engle and Patton, 2001) as shown by the decline towards zero. This confirms that the volatility process is mean reverting. The unconditional variance of the fGARCH(1,1)-NIG model for the S&P Indian market over the sample period is 1.7659, implying a mean annualised volatility of 21.10%, which is quite close to the sample estimate of the unconditional volatility presented in Section 6.3.1.2.

From the outcomes of the ARMA(1,1)-fGARCH(1,1) models in Table 6.3, the mean and variance equations of the fGARCH(1,1) model when fitted with the assumed NIG innovation can be stated as:

$$
\begin{aligned}
r_t &= \mu_t + \varepsilon_t \\
&= \phi_0 + \phi_1 r_{t-1} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \\
&= 0.0106 + 0.3208 r_{t-1} - 0.2556 \varepsilon_{t-1} + \varepsilon_t \\
\sigma_t^{\gamma} &= \omega + \alpha_1 \sigma_{t-1}^{\gamma} (|z_{t-1} - \zeta_{21}| - \zeta_{11}\{z_{t-1} - \zeta_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma} \\
\sigma_t^{1.9993} &= 0.0443 + 0.0900 \sigma_{t-1}^{1.9993} (|z_{t-1} - 1.3191| + 0.1919\{z_{t-1} - 1.3191\})^{1.9993} + 0.8113 \sigma_{t-1}^{1.9993}
\end{aligned}
$$

6.4.2 Application of the GAS Model

Next, we present the outcomes of the estimations involving the observation-driven GAS model. Hence, among seven selected assumed innovations, we use this study to determine an optimal or the most adequate assumed innovation distribution that captures volatility persistence estimation through the GAS model. The seven assumed innovations are the Normal, skew-Normal, Student's t, skew-Student's t, asymmetric Student's t with two tail decay parameters (AST), asymmetric Student's t with one tail decay parameter (AST1), asymmetric Laplace distribution (ALD). Details on these error distributions can be found in Ardia et al. (2019); Creal et al. (2011, 2013); Ghalanos (2018); Kotz et al. (2001); Zhu and Galbraith (2010).

Model selection and comparisons are carried out using two information criteria AIC and BIC as shown in Table 6.5. It is observed from the table that the GAS parameter estimates $\hat{\kappa}_\sigma$, $\hat{a}_\sigma$ and $\hat{b}_\sigma$ are statistically significant at 1% level under the seven innovation distributions, except for the $\hat{\kappa}_\sigma$ that is insignificant under the Student's t distribution and is 5% significant under the skew-Student's t distribution. The table's results further show that both the AIC and BIC have their lowest values under the AST1 innovation. Hence, the volatility persistence in this market's returns can be most adequately described through the GAS model fitted with the AST1 assumed innovation.

Table 6.5: Empirical outcomes of the GAS modelling on the real return data.

	Normal	**skew-Normal**	**Student's t**	**skew-Student's t**	**AST**	**AST1**	**ALD**
$\hat{\kappa}_\sigma$	0.0080*	0.0039*	−0.0030	0.0036**	0.0227*	0.0226*	0.0048**
$\hat{a}_\sigma$	0.1239*	0.0305*	0.2048*	0.0499*	0.0509*	0.0508*	0.0666*
$\hat{b}_\sigma$	0.9689*	0.9696*	0.9743*	0.9746*	0.9747*	0.9749*	0.9724*
AIC	9275.222	9259.497	9109.057	9103.281	9094.407	9092.654	9110.983
BIC	9299.234	9289.512	9139.072	9139.299	9136.429	9128.672	9140.998

Note: The empirical outcomes of the fit of the GAS model. The "*" and "**" denote 1% and 5% levels of significance, respectively.

The estimated volatility persistence $\hat{b}_\sigma$ under this optimal AST1 distributed error is 0.9749. The outcome indicates that the GAS parameter $\hat{b}_\sigma$ in $\mathbf{B}$ is estimated close to unity with a volatility half-life of about 27 days, which implies a considerable persistent dynamic process for $\boldsymbol{\vartheta}_t$. This persistence outcome of 0.9749 coincides with the persistence outcome obtained from the best model fGARCH(1,1)-NIG (i.e., the fGARCH(1,1) model fitted with the NIG assumed distribution) in Section 6.4.1. The high persistence outcome is consistent with the findings of Pandey and Kumar (2017), who used the GARCH(1,1) model on the Indian S&P CNX NIFTY 50 for the sample period 1997 to 2012 and found high persistence in the volatility process. However, this study used a more robust approach that involved a comparative use of the omnibus fGARCH and GAS models to estimate the persistence of the return volatility.

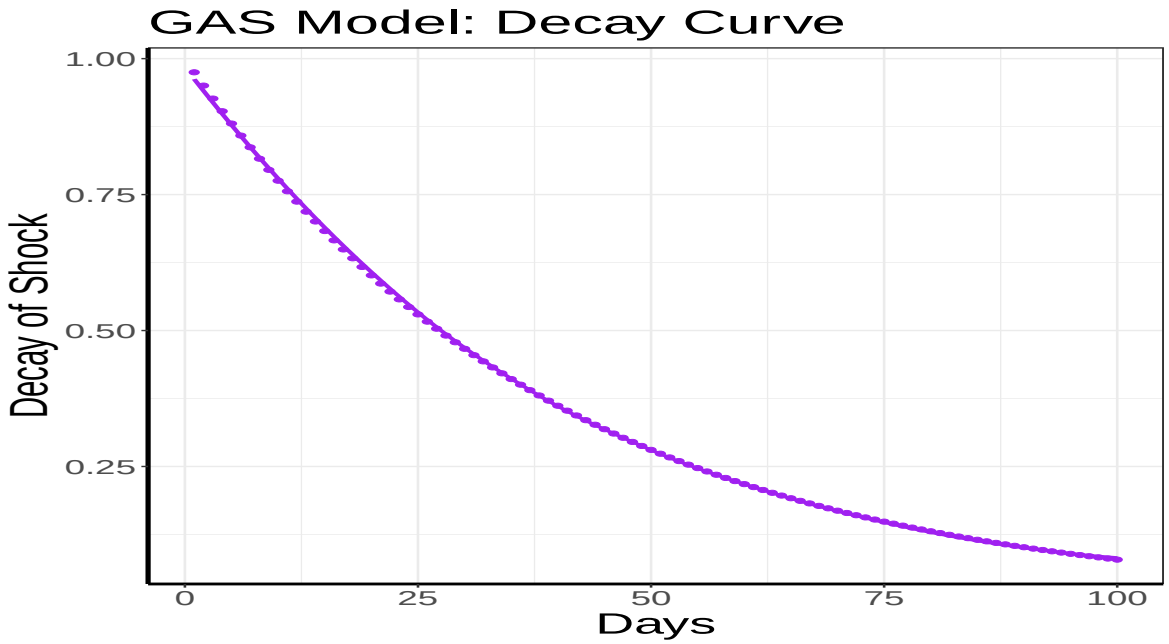

Figure 6.5: The decay of volatility persistence through the GAS-AST1 model. This decay curve follows the same trajectory as the fGARCH(1,1)-NIG decay curve in Figure 6.4 because the persistence estimates from the two models are the same (i.e., 0.9749).

Figure 6.5 shows that the impact of shocks through the GAS-AST1 model decays geometrically toward zero mean reversion, where the persistence impact dropped by half in about 27 days. That is, it follows the same trajectory as in fGARCH(1,1)-NIG model in Figure 6.4, thus both models have the same persistence outcomes. The outcome further suggests the existence of long memory in the volatility of the returns. In addition, the estimated value of the unconditional scale (volatility or variance) under the GAS-AST1 model for the S&P Indian market over the sample period is 2.4605, implying a mean annualised volatility of 24.90%. To summarise, this study shows that when the underlying true innovation distribution is unknown, the fGARCH model fitted with the NIG assumed innovation and the GAS model fitted with the AST1 are the most suitable to describe the returns for volatility persistence estimation in the S&P Indian market.

From the outputs of the GAS-AST1 model in Table 6.5, the equation of the GAS model for the time-varying scale (or volatility) parameter can be stated as:

$$\boldsymbol{\vartheta}_{t+1} = 0.0226 + 0.0508\boldsymbol{s}_t + 0.9749\boldsymbol{\vartheta}_t. \tag{6.26}$$

6.4.3 Application of the Beta-Skew-t-EGARCH Model

These estimates of the persistence from both the fGARCH and GAS models indicate that the returns volatility exhibits considerable long memory since the persistence tends towards (or is

close to) one. Moreover, their mean reversions (from the half-life computations) are somehow slow, which implies that the volatility of returns approaches the average or long-run volatility slowly. Based on this, we carry out further investigations on the long-memory behaviour of volatility decomposition into long-term and short-term components, and possible asymmetry (or leverage effect), fat-tails and skewness in the process using the one- and two-component Beta-Skew-t-EGARCH models, and then compare their outcomes. To be precise, we use both the one- and two-component models to estimate leverage effects, fat-tails, and skewness in the returns. We further use the two-component model to investigate the long-memory decomposition of volatility. That is, the two-component model is used to determine if the persistence of volatility can be decomposed into long-term and short-term processes. The results of fitting these one- and two-component Beta-Skew-t-EGARCH models are then compared to determine a superior fit. The conditional error of the Beta-Skew-t-EGARCH model is distributed as a skew Student's t (see Harvey and Sucarrat, 2014; Sucarrat, 2013). Skewness η in the two models is applied through the method proposed by Fernández and Steel (1998) (see Ghalanos, 2018; Harvey and Sucarrat, 2014; Sucarrat, 2013). Hence, $\eta < 1 (\eta > 1)$ implies left (right) skewness.

We begin with the one-component modelling. Panel A of Figure 6.6 displays the plot of the fitted conditional standard deviations for the one-component model. The plot shows that the return series is characterised by time-varying volatility with a strong volatility spike caused by the global COVID-19 pandemic crisis in 2020. The outcomes of estimation through the one-component Beta-Skew-t-EGARCH model are presented in Panel A of Table 6.6. From the table (in Panel A), the estimated degrees of freedom $\hat{\nu}$ in the skew Student's t innovation is 6.2102, which is a reasonably fat-tailed conditional Student's t density. The skewness estimate $\hat{\eta}$ is about 0.8709, which relates to pronounced negative skewness ($\eta < 1$) in the residuals z_t.

The leverage effect estimate $\hat{\kappa}^*$ is positive, which indicates that large negative returns are being followed by higher volatility. This implies that negative shocks or bad news will impact future volatility more than positive shocks or good news of the same size. In other words, the outcome shows an asymmetric effect with a stronger impact from negative shocks. In the event of asymmetric effects, Mishra (2010) reported that investors' attention becomes more short-term focused. Investors tend to constantly review their investment portfolios for liquidity and performance, even if such investments are purchased with a long-term view. This could have an adverse influence on economic growth and business investment spending because investors tend to move their funds to more liquid and less risky assets as a result of such effects.

Next is the two-component modelling. Panel B of Figure 6.6 displays the plot of the fitted conditional standard deviations for the two-component model. Panel A of Table 6.6 presents the outcomes of the estimations involving the two-component model. From the table, the degrees of freedom estimate in the skew Student's t innovation is 6.3534, suggesting fat-tails in the conditional Student's t density. The estimated skewness is about 0.8742, which indicates a pronounced negative skewness. This implies that the risk of a large negative demeaned stock

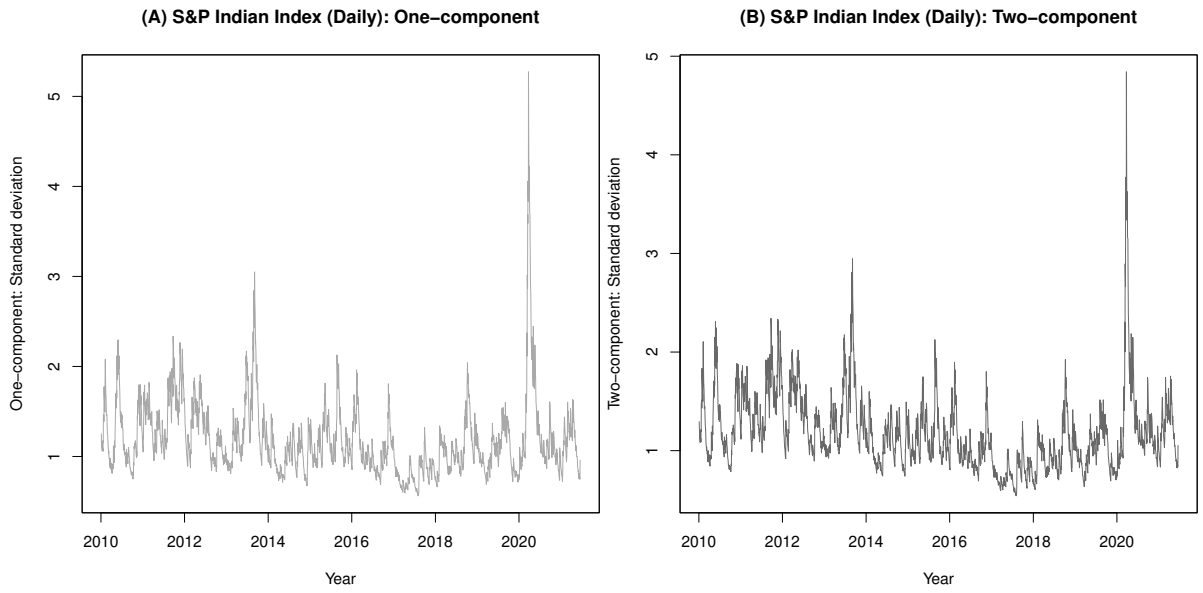

Figure 6.6: Panels (A) and (B) present the fitted conditional standard deviations of the one- and two-component models, respectively.

return is greater than that of a large positive demeaned stock return. The persistence of shocks in the long-run component $\hat{\phi}_1$ is very high at 0.9988, with a mean-reversion half-life ($h2l_{\phi_1}$) of about 577 days (i.e., about a year and seven months). This long-run half-life outcome suggests that the long-run effect of the volatility, which was partly caused by the 2020/2021 global COVID-19 pandemic crisis, would persist for as long as about a year and seven months before returning halfway back to the normal state. This indicates that even if the volatility of returns appears to have quite a long memory, it will still mean revert since the persistence estimate is less than one (Chiang et al., 2006; Engle and Patton, 2001). This implies that even though it takes a long time to revert, the volatility process does go back to its mean (Chiang et al., 2006; Engle and Patton, 2001). Due to the persistence caused by the pandemic outbreak, the Indian economy contracted by 6.6 percent during the fiscal year 2021 but staged a mild recovery in the fiscal year 2022 when it grew 8.7 percent (CRISIL, 2022).

The persistence of shocks in the short-run component $\hat{\phi}_2$ is 0.9550, with a half-life ($h2l_{\phi_2}$) outcome of about 15 days. Hence, the short-run component decays much faster than the long-run component. In other words, it can be seen from the persistence outcomes that $\hat{\phi}_2 < \hat{\phi}_1 < 1$ and from the half-life results that $h2l_{\phi_2} < h2l_{\phi_1}$, hence the short-run component decays more quickly than the long-run component that dominates the volatility persistence process. Panel A (Panel B) of Figure 6.7 shows that the decay of persistence impact in the long-term (short-term) reached half in about 577 days (15 days) and it continues geometrically towards zero mean reversion.

The parameter estimate $\hat{\kappa}_1$ that indicates the long-run response of volatility to shocks is 0.0076, while the estimate $\hat{\kappa}_2$ for the short-run response to shocks is 0.0301. These estimates

Table 6.6: Empirical and Simulation Outcomes of the Beta-Skew-t-EGARCH Model.

	Panel A		**Panel B**	
	Empirical Outcomes		**Simulation Outcomes**	
	Beta-Skew-t-EGARCH Model		**Beta-Skew-t-EGARCH Model**	
	One-Component	**Two-Component**	**One-Component**	**Two-Component**
$\hat{\omega}$ **[se]**	0.0142 [0.0482]	0.0529 [0.1293]	0.0212 [0.0430]	0.0461 [0.1106]
$\hat{\phi}_1$ **[se]**	0.9721 [0.0056]	0.9988 [0.0015]	0.9771 [0.0034]	0.9979 [0.0012]
$h2l_{\phi_1}$ **(Days)**		577		
$\hat{\phi}_2$ **[se]**	–	0.9550 [0.0092]	–	0.9529 [0.0078]
$h2l_{\phi_2}$ **(Days)**		15		
$\hat{\kappa}_1$ **[se]**	0.0402 [0.0054]	0.0076 [0.0030]	0.0413 [0.0041]	0.0151 [0.0038]
$\hat{\kappa}_2$ **[se]**	–	0.0301 [0.0061]	–	0.0238 [0.0062]
$\hat{\kappa}^*$ **[se]**	0.0337 [0.0039]	0.0379 [0.0044]	0.0324 [0.0030]	0.0414 [0.0035]
$\hat{\eta}$ **[se]**	0.8709 [0.0207]	0.8742 [0.0208]	0.8611 [0.0163]	0.8613 [0.0162]
$\hat{\nu}$ **[se]**	6.2102 [0.6855]	6.3534 [0.7171]	6.0746 [0.4639]	6.3577 [0.5032]
BIC	3.0255	3.0266		

Note: $\hat{\omega}$ estimates the long-term log-volatility, $\hat{\phi}_1$ and $\hat{\phi}_2$ are estimators for the long-term and short-term persistence parameters, respectively, and $\hat{\kappa}_1$ ($\hat{\kappa}_2$) estimates the long-run (short-run) response to shocks. $h2l_{\phi_1}$ and $h2l_{\phi_2}$ are measures of the half-life for $\hat{\phi}_1$ and $\hat{\phi}_2$, respectively. $\hat{\kappa}^*$ estimates the leverage parameter, while $\hat{\eta}$ and $\hat{\nu}$ are estimates of the skewness and degrees of freedom, respectively. [se] in square bracket is the standard error of the estimated parameter.

show a significant discrepancy, where the outcomes reveal that the unexpected arrival of news influences the short-run component considerably more than the long-run component. In other words, the response to the effect of shocks in the short-run is higher than in the long-run volatility. More precisely, the long-run component displays smaller proportional effects of about 20 percent response to volatility shocks as compared to 80 percent from the short-run component[1]. However, even though the short-run component has a stronger shock effect, it is short-lived (see Chen and Shen, 2000; Chiang et al., 2006; Engle and Lee, 1999 for related outcomes).

To summarise, this study finds the existence of both short-term and long-term volatility in the persistence process, where the response to the effect of shocks in the short-run is much higher than in the long-run volatility. This infers that higher volatility in the process is mostly

[1]The total displayed response to shocks via $\hat{\kappa}_1$ and $\hat{\kappa}_2$ from the model as shown in Table 6.6 is 0.0377 (i.e., 0.0076 + 0.0301). Hence, there are 20.16% $\approx$ 20% and 79.84% $\approx$ 80% responses to volatility shocks by the long-term and short-term components, respectively.

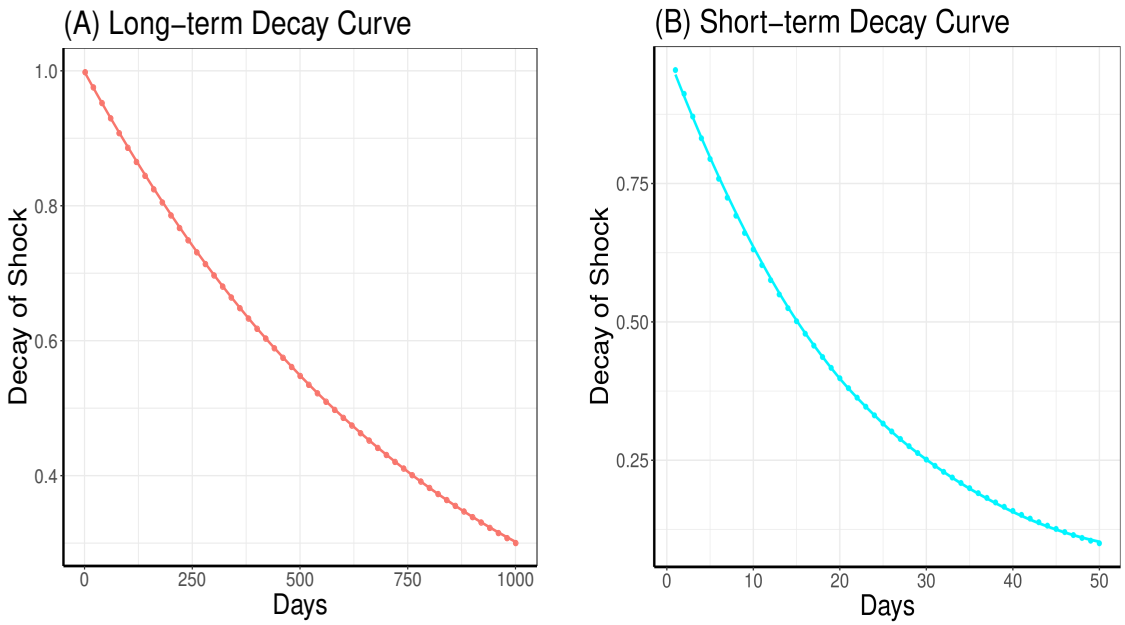

Figure 6.7: Panels (A) and (B) present the long-term and short-term decay curves, respectively.

due to the short-run volatility increase. However, the short-run volatility fluctuation is brief, while the long-run mean-reversion of volatility persistence will dominate thereafter. Precisely, the short-run effect is big but short-lived. Although the long-run component displays a pronounced longer persistence into the future, its response to volatility shocks is much lower than that of the transient short-run component. This means that investment and other market risks in the long term seem to be considerably under control in the market.

6.4.4 Simulation Study

Following the empirical outcomes, we run a set of Monte Carlo experiments using the true parameter outcomes from the Beta-Skew-t-EGARCH specification to further ascertain the validity of both the one- and two-component Beta-Skew-t-EGARCH model's results.

For the one-component simulation, we use sample size N = 5,000 simulated returns, generated from the true parameters $\omega = 0.0142$, $\phi_1 = 0.9721$, $\kappa_1 = 0.0402$, $\kappa^* = 0.0337$, $\eta = 0.8709$, and $\nu = 6.2102$. These true parameter values are empirical outcomes (i.e., MLE estimates) from fitting the first order one-component Beta-Skew-t-EGARCH model to the real Indian returns data. We use seed value 12345 for the simulation. Next, we fit the one-component Beta-Skew-t-EGARCH model to the simulated dataset and obtained the outcomes as presented in Panel B of Table 6.6 under the "One-Component" model. From the table, the estimated degrees of freedom $\hat{\nu}$ in the skew Student's t is 6.0746, the leverage parameter estimate $\hat{\kappa}^*$ is 0.0324, while the skewness estimate $\hat{\eta}$ is about 0.8611. These outcomes are quite close to the empirical results in Panel A of the table.

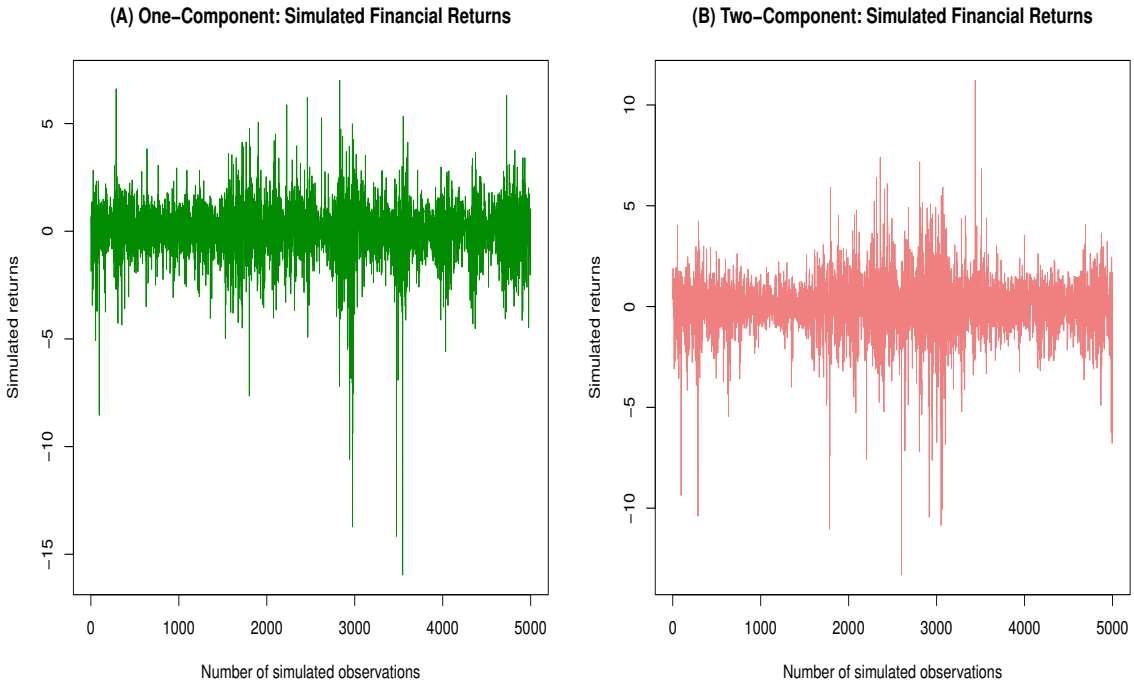

Figure 6.8: Panels (A) and (B) display the plots for one-component simulated returns and two-component simulated returns, respectively.

For the two-component simulation, we follow the same steps by using N = 5,000 simulated returns, generated from the true parameters $\omega = 0.0529$, $\phi_1 = 0.9988$, $\phi_2 = 0.9550$, $\kappa_1 = 0.0076$, $\kappa_2 = 0.0301$, $\kappa^* = 0.0379$, $\eta = 0.8742$, and $\nu = 6.3534$. We use seed value 12345 for the simulation. The true parameter values are obtained by fitting the first order two-component Beta-Skew-t-EGARCH model to the real Indian returns data. Next, we fit the two-component Beta-Skew-t-EGARCH model to the simulated dataset and obtained the outcomes as presented in Panel B of Table 6.6 under the "Two-Component" model. The estimates of the long-run $\hat{\phi}_1$ and short-run $\hat{\phi}_2$ are 0.9979 and 0.9529, respectively, and they are quite close and consistent with the empirical outcomes in Panel A of the table.

Moreover, given a 95% nominal recovery level, the TPR outcome for the leverage parameter estimate $\hat{\kappa}^*$ in the one-component model is 91.34%, while the TPR outcome for the long-run (short-run) estimate is 94.91% (94.79%) in the two-component model. These outcomes indicate a good performance of the simulation experiments. Hence, the simulation experiments performed considerably well with suitably valid outcomes. Thus, the leverage effect is evident, with long-memory decomposition into the long-run and short-run components. The response to volatility shocks is more pronounced in the short-lived short-run component than it is in the long-run component. Panels A and B of Figure 6.8 display the plots for one- and two-component simulated returns, respectively. Both plots show that the simulation return series is characterised by time-varying volatility.

6.4.5 Model Comparison

Next, we compare the results from the one- and two-component Beta-Skew-t-EGARCH model through the BIC[1] values, and the outcomes suggest that the one-component model outperforms the two-component model. This outcome is consistent with that of Harvey and Sucarrat (2014), where the one-component model also performed better than the two-component model; but the outcome is in contrast with the findings of Sucarrat (2013), where the two-component model outperformed the one-component model. Moreover, as reported by Harvey and Sucarrat (2014), the use of the two-component model does not always give a better fit.

6.5 Discussion

In this study, we used three autoregressive models comprising the fGARCH, GAS, and Beta-Skew-t-EGARCH models to estimate six features of return volatility that are relevant for robust risk management in the S&P Indian market index. These features are stylised facts that characterise the market, and they include volatility persistence, mean-reversion, asymmetry (or leverage effect), skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components . The ability of these models to capture these stylised facts provided a high degree of robustness for volatility modelling in the market's returns. Both simulations and empirical evidence were used to show the accuracy of the estimations.

To begin with, the study comparatively used the robust fGARCH and GAS models to estimate the magnitude and dynamics of the persistence in conditional volatility using the returns from the market index. The outcomes of the estimations found the NIG and AST1 assumed errors as the most adequate (or optimal) error distributions to use with the fGARCH and GAS models, respectively, for volatility modelling when the underlying error distribution is unknown in the returns. Hence, the NIG (AST1) distribution may be widely used with the fGARCH (GAS) model to improve the accuracy of volatility modelling for risk measures in finance and other areas. Appropriate risk management with proper economic policy implementation could create a channel for profit maximization by financial institutions and individual investors (Dralle, 2011).

[1] We used BIC for model selection because it shows consistency as the sample size increases, such that the criterion will select a true model of finite dimension if it is included among the candidate models (Christensen, 2018). In other words, the BIC chooses the true model with probability one (see Equation 3.29), with the assumption that the true model is among the set of selected candidate models (see Nishii, 1984; Shibata, 1981). Moreover, using consistency property, the BIC discourages overparameterisation where it imposes more heavy penalties than the AIC on model complexity (Li et al., 2009). Hence, the BIC potentially favours highly parsimonious models which are neither overly simple nor overly complex but usually lie between the two (Breheny, 2013).

By fitting each of the optimal assumed innovations to their respective model, the study found considerably high volatility persistence in the market returns. The high persistence suggests the presence of volatility clustering in the returns. Knowledge about the clustering of volatility allows market agents to adopt dynamic and flexible trading strategies that are suitable either for high-volatility or low-volatility regimes (Pandey and Kumar, 2017). The study further showed that the fGARCH and GAS models performed equally well in the volatility persistence (and mean-reversion) estimations when fitted with their respective optimal assumed errors. Next, we comparatively used the one- and two-component Beta-Skew-t-EGARCH models to estimate other features of the return volatility that include leverage effect or asymmetry, skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components. Specifically, we used both the one- and two-component models to estimate leverage effects, fat-tails, and skewness in the returns. Through the one- and two-component models, our findings show that negative skewness and leverage effects are pronounced, with considerable fat-tails in the conditional density. The leverage estimate is positive, which indicates that large negative returns are being followed by higher volatility. The pronounced negative skewness estimate implies that the risk of a large negative stock return is greater than that of a large positive stock return.

Also, we used a parametric approach through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator to estimate and determine the presence of long memory in the returns, absolute returns, and squared returns. The results of the estimations indicate that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate. However, the persistence is lower in the returns when compared to the squared and absolute returns.

Furthermore, we used the two-component Beta-Skew-t-EGARCH model to investigate the long-memory decomposition of volatility. Through this two-component model, the study found the existence of both long-run and short-run components of volatility in the persistence process, but the response to the effect of shocks in the short-run is higher than in the long-run volatility. This response to shock effects is also part of the findings of the fGARCH modelling. This implies that higher volatility in the process is mostly due to the short-run volatility increase. Further findings through the two-component Beta-Skew-t-EGARCH model using the half-life estimation showed that the short-run volatility fluctuation reverts much faster to the mean or normal volatility state than the long-run volatility persistence. Consequently, these results show that with the arrival of news in the stock market, the long-run component displays a much lower response to the effect of shocks (or change to volatility) than the transient short-run component. In summary, the short-run effect is big but short-lived, while the long-run effect is much lower but persists into the future. This means that investment and other market risks in the long term seem to be considerably under control in the market.

Lastly, a comparison of the two versions of the Beta-Skew-t-EGARCH model showed that the one-component model outperformed the two-component model. These discussed outcomes summarily answered the four research questions, and the study shows that the market returns are characterised by the six stated volatility features, namely, pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long memory behaviour of volatility decomposition into long-term and short-term components.

6.6 Conclusion

In conclusion, this study largely contributes to the literature by comparing the fit of two robust models involving the fGARCH and GAS. The fGARCH model uses the dynamics of the residuals to drive the conditional volatility while the GAS model uses the dynamics of the conditional score to drive the time-varying parameters of the time series process. To the best of the authors' knowledge, the comparison of these two robust models has not been carried out in previous studies for modelling the dynamics of volatility. Risk management systems are highly dependent on the underlying assumed distribution, and the identification of a distribution that adequately captures every aspect of the given financial data. This may be of great benefit to investors and risk managers (Chinhamu et al., 2017). Hence, our findings showed that to determine a reliable volatility (risk) modelling approach in a financial time series with unknown underlying error distribution, the NIG (AST1) assumed error could be recommended for use with the fGARCH (GAS) model. The study further revealed considerably high volatility persistence, negative skewness, leverage effect and fat-tails in the S&P Indian financial market returns.

Lastly, our findings from the long-memory behaviour of volatility decomposition revealed that although the response to shocks is greater in the short-term component, it is however short-lived. On the contrary, despite a high degree of persistence in the long-term component, market information or unexpected news arrival only has a low long-run impact on the market. Based on this, the long-run investment risks within the Indian stock market seem to be under control. Hence, our findings suggest that rational investors should try to stay calm with the arrival of unexpected news or unforeseen events in the Indian stock market because the long-run effect of such news will not be severe, and the market will eventually return to its normal state. With the presence of short-term and long-term components and their impacts on the market, this study also suggests that market managers and government should make efforts to understand the implications of changes in their system of trading and policies implemented. Such moves (or actions) will facilitate improvement in the market activities and further enable them to better control risks in the market.

For future studies, the authors intend to explore the functionalities of other robust time-varying models for volatility modelling and forecasting in multiple market indices. Specifically, the authors intend to use the apARCH model (Ding et al., 1993) as an extension of the GARCH model, and the score-driven extensions involving the two-component Beta-t-QVAR-M-lev model proposed by Haddad et al. (2023), and the Beta-t-EGARCH model with random shifts (RS-Beta-t-EGARCH model) developed by Alanya-Beltran (2022).

Chapter 7

Discussions Based on the Contributions of the Research Papers

This chapter presents the underlying contributions of carrying out this research by summarising the contributions of the research objectives through the peer-reviewed Papers 1 to 3. The main aims and objectives of this research study were achieved by addressing the aims and objectives of each specific chapter as described in the peer-reviewed research articles presented in Chapters 4 to 6. Furthermore, each paper was used to address some key questions pertinent to the objectives of the research study.

The primary focus of Papers 1 and 2 in Chapters 4 and 5 was to develop a simulation framework/structure through the family GARCH (fGARCH) and GAS models, respectively, to determine a suitable conditional distribution for the error term to model the persistence of volatility. The two papers were intended to familiarise practitioners and upcoming researchers with well-guided concepts of simulation modelling through these models. Hence the papers disclosed a vital framework relevant to the simulation of financial time series data using the referred models. In the process of developing the referred framework, the study further developed a "True Parameter Recovery" measure which is now formally used as a proxy for the coverage probability to measure the performance of the MCS estimator in recovering the true parameter. This development has crucially contributed to the statistical body and the measure can potentially guide upcoming researchers and users/practitioners through the process of estimating the performance of their MCS experiments.

In Paper 3, Chapter 6, the study comparatively applied these two models (i.e., the fGARCH and GAS), and the one- and two-component Beta-Skew-t-EGARCH models to the real financial time series data to model six essential characteristics of volatility for robust risk management. The six modelled volatility features are stylised facts that include pronounced persistence, leverage effect (i.e., volatility asymmetry), mean reversion, conditional fat-tailedness, conditional skewness, and the long memory behaviour of volatility decomposition into long-

term and short-term components. The Beta-Skew-t-EGARCH model is an unrestricted version of the GAS model (Creal et al., 2013; Sucarrat, 2013). The study further used a parametric approach through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator to estimate the existence of long memory in the returns volatility. The contributions from each paper based on several conducted studies are further discussed in the following sections.

7.1 Simulation Framework to Determine Suitable Innovations for Volatility Persistence Estimation: The GARCH Approach

The objectives of Paper 1 (Chapter 4): To develop a simulation framework that is relevant to determine suitable conditional error distributions for modelling volatility persistence using the GARCH model. In addition, the paper was also used to develop a "true parameter recovery" measure that is used as a proxy of the coverage probability to evaluate the performance of the simulation experiment.

The research carried out in Chapter 4 (Paper 1) has been able to present a comprehensive step-by-step simulation framework through the GARCH model. Our goal in undertaking this study was to provide a means of improving volatility modelling using the GARCH model through Monte Carlo simulation. Specifically, the framework outlined necessary simulation steps that are relevant to determine adequate error distribution for estimating the persistence of volatility. However, it is well reported that asymmetric GARCH models perform better at modelling volatility than the simple classical GARCH model (see Takaishi, 2018), since the classical GARCH model is not capable of capturing asymmetries in the volatility process (Feng and Shi, 2017). Based on that, this study applied a robust extension of the GARCH model called the family GARCH (fGARCH) model (Hentschel, 1995) that incorporates the asymmetric volatility process.

Moreover, apart from the attractive property of incorporating an asymmetric volatility process, the fGARCH model also has the advantage of using the rotations and shifts in the news impact curve to drive both large and small volatility shocks, respectively (Ghalanos, 2018). Hence the fGARCH model was used to demonstrate the simulation steps in this paper. The framework also presented clear coding guidelines that explained the algorithms for data generation through the R rugarch package that was used for the illustration.

As discussed in Chapter 4, the robust simulation design steps were illustrated using financial time series data, with empirical verification. The empirical modelling necessitated some stages of hypothesis testing involving the tests for serial correlations and ARCH effects using

the ARMA-fGARCH models, the weighted ARCH LM test, and Engle's ARCH test (see Engle, 1982) that comprises the Lagrange Multiplier (LM) and Portmanteau-Q (PQ) tests. Also, the study used relevant diagnostics, like residual diagnostics using the weighted ARCH LM test, and other diagnostic model checking plots, like the density and QQ plots to evaluate the quality of the fitted ARMA-fGARCH models. Details can be seen in Chapter 4.

For a practical illustration of the simulation design, the study comparatively used ten error distributions comprising the Normal, skew-Normal, Student's t, skew-Student's t, GED, skew-GED, GHYP, Normal Inverse Gaussian (NIG), Generalised Hyperbolic Skew-Student's t (GHST) distribution and Johnson's reparametrised SU (JSU) distribution. The outcome of the simulation with empirical verification revealed that the GHYP was the most suitable assumed error distribution to describe the returns for volatility persistence estimation when the log-likelihood was used as the basis for model selection. However, when the decision for model selection was made using the AIC, BIC, SIC, and HQIC information criteria, the skew-Student's t emerged as the most appropriate assumed distribution to describe the error for volatility persistence estimation.

Furthermore, among the findings, this study developed a new, flexible but robust "true parameter recovery (TPR)" measure, which is a good proxy for the existing coverage probability measure. The TPR measure is a means of assessing the performance of the MCS estimates in recovering the true parameter. That is, it is used to determine how much of the true parameter value is recovered by the MCS estimator. Moreover, it is a robust technique that precisely reports over-estimation, under-estimation and full recovery of the true parameter value by the MCS estimator. One further reason backing the development of this new recovery measure was because the R rugarch package used for executing the simulation design study does not make provision for calculating the coverage of the simulation experiment. Hence, this study proposed this initiative to develop the referred proxy measure. This new recovery measure was introduced in this published paper, and we strongly believe it can henceforth guide researchers on how to calculate the performance of Monte Carlo simulation experiments. Hence, the development of this recovery measure has created an avenue for further explorative study by researchers.

One of the required steps under the method of implementing the simulation design is to "set the seed", which enables the reproducibility of the simulation outcomes. Based on that, this study further used the GARCH model to show that as the sample size N increases, the pattern or arrangement of the seed values does not "generally" impact the $\sqrt{N}$ consistency and efficiency outcomes of an estimator. However, this outcome may depend on the quality of the model used. Hence, the study was able to show that the use of the seed value is basically for reproducibility.

A further significant contribution from this paper is that it used the GARCH model to show that, for a given sample size and seed value, the outcome of running a simulation process once is the same as replicating it multiple times. This experiment was carried out in the study

in Chapter 4 by using the GARCH-Student's t as the true model on three sample-size datasets with the same seed value. From these, four distinct simulations were conducted/executed, i.e., simulation run once, and simulations run with 300, 1,000 and 2,500 replications. The results of the executions showed that the log-likelihood, RMSE, SE and bias outcomes of the $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\alpha} + \hat{\beta}$ estimators for each simulation are the same for the three sample-size datasets with the same seed value, regardless of whether the simulation is run once or replicated multiple times. Hence, the outcome of a single run is the same as the average of multiple runs or replications.

The findings in this paper of the thesis may help current practitioners and upcoming researchers in finance and other sectors involved with volatility modelling to better understand the relevant simulation steps that are necessary for an improved volatility persistence estimation through this model. Hence, this paper has contributed to addressing the inadequate literature on the relevant simulation steps needed to select a suitable conditional distribution for the error term to model the persistence of volatility. Moreover, in the absence of a reliable means of calculating the coverage of a simulation experiment, this study has further largely contributed to providing for such calculation through the proxy TPR measure.

7.2 Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model

The objectives of Paper 2 (Chapter 5): To use the GAS model to further extend the simulation design proposed in Paper 1 (see Samuel et al., 2023). That is, this study follows these simulation design steps by utilising the GAS model in this instance. Hence, the main objective of the paper was to illustrate how the GAS model, using the R GAS package (Ardia et al., 2019), can be used to select an appropriate assumed error distribution relevant for estimating the persistence of volatility through MC simulation, with empirical verification. Moreover, simulation through any model largely involves generating an appropriate dataset that can produce a reasonable output approximation. Therefore, one further objective of this paper was to experimentally show how to use the GAS model to generate or simulate a dataset that can adequately reflect the behaviour of financial time series data relevant for modelling volatility.

Volatility modelling for risk management is highly dependent on the underlying error distribution. So, if the true (underlying) error distribution is not known, then volatility will need to be modelled with a suitable assumed error distribution. Hence, this study used the dynamics of the GAS model to illustrate how to obtain an optimal error distribution for volatility persistence estimation through MC simulation by comparing the outcomes of the model fitted with each of the seven assumed error distributions. The assumed error distributions are the Gaussian,

skew-Gaussian, Student's t, skew-Student's t, asymmetric Student's t with one tail decay parameter (AST1), asymmetric Student's t with two tail decay parameters (AST), and asymmetric Laplace distribution (ALD). As discussed in Paper 2, the Student's t and AST1 were used as the true error distributions to generate the simulated return observations.

The Student's t was particularly used as the true error distribution for the illustration because it is assumed that financial data appear to have a distribution much like it, and it is also believed that it can adequately deal with fat-tailed or leptokurtic features (Duda and Schmidt, 2009; Lin and Shen, 2006) commonly found in financial data (Hentschel, 1995). The AST1 was used as the second non-Gaussian true error distribution to compare its outcomes with those of the Student's t conditional distribution. In other words, the GAS-Student's t and GAS-AST1 were used as the true models for the data generation, where the comparisons of the seven assumed error distributions were evaluated through four meta-statistical summaries comprising of the RMSE, bias, SE and RE. Under both simulations involving the two true models, the AST1 consistently emerged as the optimal or the most suitable assumed error. The outcome was further verified empirically using the information criteria as detailed in Paper 2. Moreover, the empirical verification also passed through stages of hypothesis testing for linear dependency and heteroscedasticity as discussed in Chapter 5. The modelling outcome implies that the persistence of volatility in the returns can be most adequately described through the GAS model when the assumed error is AST1. The estimate of the persistence of volatility in the returns under the optimal AST1 assumed error distribution was found to be considerably high. In real-life applications, the outcome of this robust estimation approach using the AST1 fitted with the GAS specification could be useful to investors and financial practitioners in enhancing the accuracy of their risk measurement for investment decisions and portfolio management.

The second crucial objective of this paper was to illustrate how to use the GAS process to simulate an appropriate dataset that can adequately imitate the behaviour of financial time series data required to estimate the persistence of volatility. To begin with, it is known that financial data are fat-tailed (Li, 2008), i.e., non-Gaussian. Hence, to obtain the like of a simulated dataset that can closely reflect this fat-tailed behaviour of financial time series data under the GAS modelling, this study further conducted two different experiments to compare the fat-tails and $\sqrt{N}$ consistency outcomes of three different datasets that were generated using three unconditional shape parameters or degrees of freedom $\nu^* = \{10.3, 8.2, 4.1\}$. For both experiments, the GAS-Student's t was used as the true model for the data generation process. In the first experiment that involved the fat-tails comparisons, the GAS model fitted to a Student's t distributed error with ν^* of 4.1 appeared as the most suitable among the three unconditional shape parameters to generate the required simulated financial returns data.

As for the second experiment that involved comparisons based on $\sqrt{N}$ consistency, this experiment was carried out in two stages. In the first stage, the used data were generated without removing potential initial value effects, while in the second stage, necessary trimmings were carried out to remove the effects of initial values. For both stages, it was observed that

the GAS model fitted with ν^* = 4.1 gave the best $\sqrt{N}$ consistency in terms of RMSE, SE and absolute value of bias. In other words, the model fitted using ν^* = 4.1 outperformed those with ν^* = 8.2 and ν^* = 10.3 in $\sqrt{N}$ consistency under the selected four error distributions of the Gaussian, skew-Gaussian, Student's t, and skew-Student's t. However, when the $\sqrt{N}$ consistency outcomes of the two stages (i.e., the data generation with and without removing the effects of initial values) were compared, it was observed that the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ for ν^* = 4.1 became more improved in terms of RMSE, SE and absolute value of bias when the effects of initial values were addressed.

Hence, to generate the required datasets for the two true models GAS-Student's t and GAS-AST1 that this paper used to illustrate the method of implementation under the simulation structure, the unconditional shape parameter ν^* was fixed to 4.1 for both instances. As detailed in Paper 2 in Chapter 5, it was observed that datasets generated using this process adequately yielded outcomes that reflected those of financial data relevant for modelling volatility. The outcomes of the two simulations conducted through the two true models using the datasets generated with ν^* fixed to 4.1 were regularly consistent with the outcomes of the empirical verifications of the experiments. Hence, this study has contributed to overcoming some of the potential hurdles practitioners and researchers may face while generating or simulating an adequate financial returns dataset through the GAS model. This finding is a significant contribution to the existing literature, which could be a relevant guide to users on how to obtain appropriate simulated financial time series data through the GAS model.

Although there are good books/articles on simulation approaches in general (see Bratley et al., 1987; Chalmers and Adkins, 2020; Ferrer-Comalat et al., 2020; Jones et al., 2014; Kleijnen, 2015; Morris et al., 2019; Mota and Mota, 2017; Wickham et al., 2019), however, before this study and as far as the author is aware, there was no existing study in the literature with adequate details as presented in this study with regards to simulation process/steps involving the GAS model for selecting an optimal conditional distribution required to model volatility. Hence, the main contribution of this paper is summarily in two folds. One, it has presented how the GAS model through the GAS package can be used to select an optimal conditional distribution for the error term to model volatility through MC simulation with empirical verification. Second, it has experimentally shown how to adequately generate a suitable financial dataset for a Monte Carlo simulation through the GAS model.
The following are other significant findings from the paper:

i) The GAS model (like the GARCH/fGARCH model, as observed in Paper 1, Chapter 4) with the Gaussian error distribution is considerably consistent but not efficient when the true error is not Gaussian.

ii) In line with the findings of Gilli et al. (2019) and Feng and Shi (2017), the outcomes from this study also showed that increasing the samples or sample size does not always

reduce the sampling error (RMSE). The RMSE results may be strongly or weakly $\sqrt{N}$ consistent.

iii) It is observed that the volatility persistence estimator of the GAS model displays appreciable $\sqrt{N}$ consistency.

iv) With flexible lines of code, this study used the R rugarch and GAS packages to show that the persistence parameter b_ϕ of the GAS model coincides with the persistence parameters $\alpha + \beta$ of the classical GARCH model of Bollerslev as adduced by Ardia et al. (2019); Blasques et al. (2014); Creal et al. (2013); Oh and Patton (2018).

Furthermore, since the GAS package used for the execution of the simulation design does not have any provision for estimating the coverage of the MCS experiments, this study used the TPR measure, developed by Samuel et al. (2023), as a proxy for the coverage. This measure was used to calculate the estimator's coverage, and it is easy to use for measuring the level of recovery of the true parameter value by the MCS estimates. Its outcomes revealed considerably good performance of the MCS estimator at recovering the true parameter. The outcomes therefore suggested that the MCS experiments were well carried out with suitably valid results. Moreover, this robust easy-to-use framework was illustrated using real-life financial returns data, which makes it easily applicable to users for effective MCS studies. Hence, by using this structure, this study has formally presented relevant simulation steps necessary for selecting an adequate assumed conditional distribution for volatility modelling through the GAS model.

7.3 A Comparative Modelling of Essential Characteristics of Volatility

The objectives of Paper 3 (Chapter 6): To use the collaboration of the family GARCH (fGARCH) (Hentschel, 1995), GAS (Ardia et al., 2019) and Beta-Skew-t-EGARCH (Harvey and Sucarrat, 2014; Sucarrat, 2013) models to comparably estimate six essential features of financial returns volatility, relevant for robust risk management.

Based on the objectives, this chapter aimed to showcase a robust volatility modelling approach to risk management and investment decisions. The estimated volatility features include pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long memory behaviour of volatility decomposition into long-term and short-term components. Specifically, the study comparably used the collaboration of the fGARCH and GAS models, on one hand, to model the persistence and mean reversion of the volatility of returns, while the one- and two-component Beta-Skew-t-EGARCH models were comparably used on the other hand to estimate the remaining four features of returns volatility. The study further used a parametric approach through the ARFIMA-FIGARCH

models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator to estimate and determine the presence of long memory in the returns, absolute returns, and squared returns. Hence, the ability of these models to capture these six stylised facts of volatility provided a high degree of robustness for volatility modelling in the market's returns.

This thesis has shown through Papers 1 and 2 (in Chapters 4 and 5) how to use simulation design with empirical verifications to select a suitable error distribution for volatility persistence estimation through the robust observation-driven fGARCH and GAS models. The simulation frameworks and procedures in these first two papers can be a useful guide to practitioners and upcoming researchers in science concerning the relevant simulation steps to determine a suitable error distribution for volatility modelling. Following these, Paper 3 in Chapter 6 was used to compare the modelling outcomes of the two robust models (fGARCH and GAS) for estimating the persistence of conditional volatility, and its mean-reversion. The main difference between these models is that the fGARCH model uses the dynamics of the residuals to drive the conditional variance, while the GAS model uses the dynamics of the conditional score to drive the time-varying conditional variance. The study showed that the two models performed equally well in the volatility persistence and mean-reversion estimations when fitted with their respective optimal assumed error distributions. Paper 3 went further to use an extension of the score-driven model called the one- and two-component Beta-Skew-t-EGARCH model to estimate the other features of volatility that include leverage effect or volatility asymmetry, conditional skewness, conditional fat-tails, and long memory behaviour of volatility decomposition.

Following the empirical outcomes, simulation experiments were run to further ascertain the validity of both the one- and two-component Beta-Skew-t-EGARCH model's results. That is, the one- and two-component models were used in turn to generate the needed simulated datasets, where the estimated outcomes of the simulation process were very close to the outcomes of the empirical modelling when compared. This study also used the TPR measure as a proxy for the coverage of the simulation study to calculate the performance of the MCS experiments, and its outcomes gave a precise estimation of the performance of the simulation experiments. Specifically, the TPR outcomes further attested to the closeness between the estimated outcomes of both the empirical and simulation modelling.

The main goal in undertaking this study was to model the volatility features that characterize the Indian stock market returns and investigate their risk impacts on investment in the market. It is worth knowing that, with a massive population and rapidly growing economy, India is one of the greatest investment destinations in the world. Its financial market is increasingly attracting many domestic and foreign investors. Hence, this study utilised the dynamics of the three stated time-varying models, and the time and frequency domain ARFIMA models to estimate the referred essential volatility features relevant for robust market risk management.

Through the applications of these models, the study showed that the Indian market returns are characterised by these volatility features.

The study also applied the R tidyverse package developed by Wickham et al. (2019) to visually plot the "decay curves" following the method of Chen and Shen (2000), Chiang et al. (2006) and Chou (1988). These curves visually revealed the trajectories of the impacts of the decay of shocks through the fGARCH and GAS models when fitted with their optimal assumed error distributions. The curves were also used in the Beta-Skew-t-EGARCH modelling to visually display the long-term and short-term decay of shocks, with their half-life mean-reversions in the market. The decay curves summarily showed that the volatility process is mean-reverting.

Several authors have used each of the models GARCH, GAS and Beta-Skew-t-EGARCH individually to model volatility and its features. For instance, authors like Engle and Patton (2001) used the GARCH(1,1) model to estimate various stylised facts of volatility in the Dow Jones Industrial Index returns from 1988 to 2000, and found a highly persistence volatility estimate with a half-life mean-reversion of about 73 days. Oh and Patton (2018) used the GAS model with a factor copula model to study systemic risk. Other applications of the GAS model are in credit risk analysis (Creal et al., 2014), high-frequency data (Gorgi et al., 2019; Opschoor et al., 2018), and spatial econometrics (Blasques et al., 2014; Catania and Billé, 2017), among others. The applications of the Beta-Skew-t-EGARCH model for modelling relevant features of volatility are well studied in Sucarrat (2013), and Harvey and Sucarrat (2014). However, to the author's knowledge, there was no earlier study conducted before this study to comparably apply the collaboration of the three (i.e., the family GARCH, GAS, and Beta-Skew-t-EGARCH) models to the Indian stock market to estimate the stated essential characteristics of volatility. Moreover, as far as the author is aware, this study is the first to use the collaboration of the parametric technique through the ARFIMA-FIGARCH models, and three semi-parametric approaches involving the GPH, LW, and ELW estimators to determine the presence of long memory in the Indian market's returns and return volatility, i.e., squared returns and absolute values of returns.

It is well-reported that risk management systems are highly dependent on the underlying error distribution, and the identification of a distribution that adequately captures every aspect of a given financial data may be of great benefit to investors and risk managers (Chinhamu et al., 2017). Hence among the findings, this Paper 3 study obtained error distributions that could be widely used to enhance the accuracy of volatility modelling for risk measures in finance and other areas. The findings further revealed that investment and other market risks in the long term are considerably under control and relatively safe for rational investors. Hence, based on the study's outcomes as detailed in Chapter 6, this thesis further contributes to the extant literature by proffering well-guided recommendations to the government, market managers and investors on both short-term and long-term investment decisions and policy implementation in the Indian financial market.

Chapter 8

Summary, Conclusions and Recommendations

8.1 Summary and Conclusion

Stochastic modelling of volatility and its characteristics play a central role in empirical finance and financial risk management. In summary, this study has used simulation approaches, whose outcomes were verified by real-life financial data, to present a framework for determining an appropriate error distribution relevant to estimating the persistence of volatility through two autoregressive models of the fGARCH and GAS. The framework utilises the robust simulating resources of these models, through set parameters, to generate data that are analysed, and the estimates from the process are then used by chosen metrics (or performance measures) to explain some attributes like the consistency and efficiency or precision of the estimator. The steps involved in the framework approaches are well detailed in Chapters 4 and 5, and they are further summarised in Chapter 7 of this thesis.

Furthermore, in Chapter 6, the study used a collaboration of the two models (fGARCH and GAS) and the Beta-Skew-t-EGARCH model to estimate six essential features of volatility for risk evaluations. In addition, both the time domain and frequency domain methods of ARFIMA modelling were further used to determine the presence of long memory in the returns. The main focus of this thesis is to help practitioners and (upcoming) researchers on MCS experiments utilising these models for reliable estimation of volatility persistence in finance and other areas. The focus further extends to estimating relevant features of volatility for improved risk management processes and investment strategies.

The aims and objectives of the thesis were achieved through the following research outcomes and contributions:

1. **The development of a simulation framework through the family GARCH model**

The development of this robust user-friendly framework presented the algorithms for determining an appropriate conditional distribution for the error term to estimate the persistence of volatility. Its applicability was demonstrated using financial returns data, which makes it easily applicable to users. Hence, we highly recommend the simulation framework as a reliable guide to users for effective simulation studies.

2. **The development of the true-parameter-recovery (TPR) measure to calculate the performance of the MCS experiment**

   The development of the TPR measure is a key contribution as it is used as a proxy of the coverage probability to accurately compute the underestimation, overestimation, and full recovery of the true parameter by the MCS estimator. Hence, the measure can be used to precisely evaluate the performance of MCS experiments. The presentations of the outcomes of the TPR measure were further conveyed through clearly displayed visual graphics in Figures 4.3, 4.4, 4.7, 5.6, 5.8, and B.1 (in Appendix B) using the R tidyverse package of Wickham et al. (2019). To the best of our awareness, this measure marks the first developed proxy technique for estimating the performance of MCS estimator of the GAS simulation process. Hence, we highly recommend its use to researchers and practitioners for accurate computation of the performance of MCS experiments.

3. **The development of a simulation structure through the score-driven GAS model**

   This structure presents an algorithm that showcases steps that are relevant for simulation modelling through the GAS model using the R GAS package. Quite a number of papers have reported the applications of the GAS models for various statistical estimations, but none, as far as we know, has presented a simulation structure approach to selecting a suitable conditional distribution for the error term as conducted in this study. Thus, this study has demonstrated this simulation approach for volatility persistence estimation through the GAS model as detailed in Chapter 5. It is highly expected that this robust structure will adequately guide researchers in finance and other areas with relevant steps on MCS experiments using the GAS model and package.

4. **To show how to use the GAS model to generate a dataset that adequately reflects the behaviour of financial time series data**

   The outcome of this objective is a significant contribution to the body of statistics through the GAS model. Moreover, the outcomes of the dataset generated (or simulated) by fixing the unconditional shape parameter or degrees of freedom ν^* to 4.1 consistently agreed with the empirical outcomes. Hence, we recommend this approach as a reliable guide to generate an appropriate financial returns dataset for simulation modelling through the GAS model.

5. **To estimate essential characteristics of volatility by comparably using the collaboration of the fGARCH, GAS and Beta-Skew-t-EGARCH models**

   To the best of our knowledge, this comparative modelling through the collaboration of the three robust observation-driven models fGARCH, GAS, and Beta-Skew-t-EGARCH has not been applied in previous studies for modelling the dynamic features of volatility in the Indian stock market. The study also comparatively applied the parametric and semi-parametric methods of ARFIMA modelling through the FIGARCH, GPH, LW and ELW models to determine the presence of long memory in the Indian market returns, absolute returns, and squared returns, which to the author's awareness has not been done in previous studies. As detailed in both empirical and simulation modelling, the study's outcomes were used to proffer relevant long-term and short-term recommendations/suggestions to the rational investors, government and market managers for relevant investment decisions and risk control in the market.

8.2 Recommendations for Future Directions of the Work

This thesis has mainly been used to develop guided frameworks and methods to enhance volatility modelling by practitioners and researchers. However, the study is an ongoing research, and there are still areas to cover as follows:

i) The research intends to further use the ugarchpath function of the rugarch package through any of the models that support its use for the simulation framework illustration in Paper 1, Chapter 4,

ii) Extension of the simulation framework ideas to modelling and estimating multivariate processes,

iii) Developing a simulation design for volatility forecasting and portfolio management,

iv) Exploring the functionalities of other robust time-varying models for volatility modelling and forecasting of multiple market indices. Specifically, the study intends to use the apARCH model (Ding et al., 1993) as an extension of the GARCH model, and the score-driven extensions involving the two-component Beta-t-QVAR-M-lev model proposed by Haddad et al. (2023), and the Beta-t-EGARCH model with random shifts (RS-Beta-t-EGARCH model) developed by Alanya-Beltran (2022),

v) Extension into simulation and empirical extremal (asymptotic) dependence modelling between financial time series variables, using the conditional multivariate extreme value (CMEV), bivariate point process, and extreme value copula models.

References

Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew student's t-distribution. *Journal of Financial Econometrics*, 4, 275–309. 132

Adenstedt, R. (1974). On large-sample estimation for the mean of a stationary random sequence. *Annals of Statistics*, 2, 1095–1107. 42

Agarwal, K. (2022). The Monte Carlo simulation: Understanding the basics. https://tinyurl.com/3vepmtv3 (accessed on 25 June 2024). 87

Agiakloglou, C., Newbold, P. and M., W. (1992). Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, 14, 235–246. 45, 135

Ahmed, M. and Naher, N. (2021). Modelling and forecasting volatility of daily stock returns using GARCH models: Evidence from Dhaka stock exchange. *The Asian Institute of Research, Economics and Business Quarterly Reviews*, 4, 74–89. 12, 20

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction on Automatic Control*, 19, 716–723. 34

Alanya-Beltran, W. (2022). Modelling stock returns volatility with dynamic conditional score models and random shifts. *Finance Research Letters*, 45, 102121. 16, 20, 26, 88, 122, 148, 160

Alizadeh, S., Brandt, M. and Diebold, F. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57, 1047–1092. 16, 17, 50, 125

Alper, O. and Atilla, C. (2007). Modeling long-term memory effect in stock prices: A comparative analysis with GPH test and Daubechies wavelets. *Studies in Economics and Finance*, 25, 1–13. 18

Amadeo, K. (2022). The stock market crash of 2008. https://www.thebalance.com/stock-market-crash-of-2008-3305535 (accessed on 25 June 2024). 108

Andersen, T. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance*, 52, 975–1005. 18

Andersen, T. G., Bollerslev, T., Christoffersen, P. F. and Diebold, F. X. (2006). Volatility and correlation forecasting. *Handbook of Economic Forecasting*, 1, 778–878. 7

Andersen, T., Bollerslev, T., Diebold, F. and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61, 43–76. 14

Andrews, D. and Guggenberger, P. (2003). A bias-reduced log periodogram regression estimator for the long-memory parameter. *Econometrica*, 71, 675–712. 136

Ane, T. (2006). Short- and long-term components of volatility in Hong Kong stock returns. *Applied Financial Economics*, 16, 439–460. 14

Appiah-Kusi, J. and Pescetto, G. (1998). Volatility and volatility spill-overs in emerging markets: The case of the African stock markets. *Ekonomia*, 2, 171–186. 13, 20

Arago, V. and Fernandez-Izquierdo, A. (2003). GARCH models with changes in variance: An approximation to risk measurements . *Journal of Asset Management*, 4, 277–287. 7, 83, 87

Ardia, D., Boudt, K. and Catania, L. (2019). Generalized autoregressive score models in R: The GAS package. *Journal of Statistical Software*, 88, 1–28. 5, 6, 7, 11, 15, 16, 20, 23, 26, 27, 28, 54, 86, 87, 88, 89, 90, 92, 96, 113, 118, 122, 123, 124, 138, 152, 155

Ardia, D., Boudt, K. and Catania, L. (2018). Downside risk evaluation with the R package GAS. *The R Journal*, 10, 410–421. 26, 88

Arltová, M. and Fedorová, D. (2016). Selection of unit root test on the basis of length of the time series and value of AR(1) parameter. *STATISTIKA*, 96, 47–64. 30, 31

Ashour, S. K. and Abdel-hameed, M. A. (2010). Approximate skew normal distribution. *Journal of Advanced Research*, 1, 341–350. 67, 127

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171–178. 67, 127

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 65, 367–389, [0911.2342]. 67, 127

Backus, D. and Zin, S. (1993). Long-memory inflation uncertainty: Evidence from the term structure of interest rates. *Journal of Money, Credit and Banking*, 25, 681–700. 19

Baillie, R. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73, 5–59. 20, 40, 41, 42, 44, 45, 120, 135

Baillie, R., Bollerslev, T. and Mikkelsen, H. (1993). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Working Paper No. 168, Department of Finance, Northwestern University*. 19, 21

Baillie, R., Bollerslev, T. and Mikkelsen, H. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74, 3–30. 6, 19, 21, 38, 42, 120

Baillie, R. and Chung, C. (1992). Estimation of fractionally integrated processes with ARCH innovations. *Manuscript*. 18

Bali, T. and Engle, R. (2010). The intertemporal capital asset pricing model with dynamic conditional correlations. *Journal of Monetary Economics*, 57, 377–390. 15

Barkoulas, J., Labys, W. and Onochie, J. (1997). Fractional dynamics in international commodity prices. *Journal of Futures Markets*, 17, 161–189. 19

Barkoulas, J., Labys, W. and Onochie, J. (1999). Long memory in futures prices. *Financial Review*, 34, 91–100. 19, 37, 39, 40, 41, 42

Barndorff-Nielsen, O. E., Mikosch, T. and Resnick S. I. (2013). *Levy processes: Theory and applications*. Birkhauser, Boston; Springer Science+Business Media; New York. 60, 67, 127

Baum, C., Hurn, S. and Lindsay, K. (2020). Local Whittle estimation of the long-memory parameter. *The Stata Journal*, 20, 565–583. 18, 40, 41, 46, 47, 48

Baum, C. and Wiggins, V. (2024). Estimate long memory in a time series via modified log-periodogram regression. http://fmwww.bc.edu/RePEc/bocode/m/modlpr.html(accessedon5August2024). 40

Beg, R. and Anwar, S. (2014). Detecting volatility persistence in GARCH models in the presence of the leverage effect. *Quantitative Finance*, 14, 2205–2213. 87

Bekaert, G. and Harvey, C. (1997). Emerging equity market volatility. *Journal of Financial Economics*, 74, 29–77. 14

Beran, J. (1994). Statistics for long-memory processes. Chapman and Hall. https://api.semanticscholar.org/CorpusID:62495143 (accessed on 25 August 2024). 19, 38, 40

Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-memory processes: Probabilistic properties and statistical methods*. SpringerLink : Bücher. Springer Berlin Heidelberg. https://books.google.co.za/books?id=KUZBAAAAQBAJ (accessed on 25 August 2024). 19, 38

Bernardi, M. and Catania, L. (2019). Switching generalized autoregressive score copula models with application to systemic risk. *Journal of Applied Econometrics*, 34, 43–65. 15, 20

Black, F. (1976). Studies of stock market volatility changes. *In: 1976 Proceedings of the American Statistical Association Business and Economic Statistics Section*. 17

Black, F. and Scholes, M. (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance, American Finance Association* 27(2), 399–417. 7

Blasques, F., Creal, D. D., Janus, P., Koopman, S. J., Lucas, A., Scharth, M. and Schwaab, B. (2014). Generalized autoregressive score models for time-varying parameters: new models and applications. https://www.gasmodel.com/2014GAS_ISFrotterdam.pdf (accessed on 26 June 2024). 26, 27, 28, 88, 89, 90, 123, 155

Blasques, F., Gorgi, P. and Koopman, S. J. (2017). Accelerating GARCH and score-driven models: Optimality, estimation and forecasting. *Estimation and Forecasting*, 1–37. 28, 90, 124

Blasques, F., Koopman, S. and Lucas, A. (2014). Maximum likelihood estimation for generalized autoregressive score models. *TI 2014-029/III, Tinbergen Institute Discussion Paper*, 1–65. 15, 86, 118, 157

Blasques, F., Koopman, S. J. and Lucas, A. (2014). Stationarity and ergodicity of univariate generalized autoregressive score. *Electronic Journal of Statistics*, 8, 1088–1112. 101

Blazsek, S., Escribano, A. and Licht, A. (2022). Score-driven location plus scale models: asymptotic theory and an application to forecasting Dow Jones volatility. *Studies in Nonlinear Dynamics & Econometrics*, 28, 61–82. 16, 20

Bobeica, G. and Bojesteanu, E. (2008). Long memory in volatility. An investigation on the central and eastern European exchange rates. *European Research Studies*, 11, 7–18. 18, 20, 21, 42, 135

Bollerslev, T. (1987). Conditionally heteroskedasticity time series model for speculative prices and rates of returns. *The Review of Economic and Statistics*, 69, 542–547. 67, 127

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedastic. *Journal of Econometrics*, 31, 307–327. 5, 6, 15, 23, 25, 26, 55, 56, 88, 119, 121, 123

Bollerslev, T. (1987). Conditionally heteroskedasticity time series model for speculative prices and rates of returns. *The Review of Economic and Statistics* 69, 542–47. 67, 127

Bollerslev, T. and Mikkelsen, H. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73, 151–184. 19, 21

Bollerslev, T. and Wright, J. (2000). Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. *Journal of Econometrics*, 98, 81–106. 19, 21, 41, 42, 44

Bollerslev, T., Chou, R. and Kroner, K. (1992). ARCH modeling in finance. *Journal of Econometrics*, 52, 5–59. 14, 20

Bollerslev, T., Engle, R. and Wooldridge, J. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96, 116–131. 15

Bollerslev, T. and Mikkelsen, H. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73, 151–184. 19, 21

Bollerslev, T. and Wooldridge J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11, 143–72. 24, 56, 68

Bollerslev, T. and Wright, J. (2000). Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. *Journal of Econometrics*, 98, 81–106. 19, 21, 41, 42, 44

Bollerslev, T. and Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109, 33–65. 15

Booth, G., Kaen, F. and Koveos, P. (1982). R/S analysis of foreign exchange markets under two international monetary regimes. *Journal of Monetary Economics*, 10, 407–415. 19

Boubaker, H., Canarella, G., Gupta, R. and Miller, S. (2021). Long-memory modeling and forecasting: evidence from the U.S. historical series of inflation. *Studies in Nonlinear Dynamics and Econometrics*, 25, 289–310. 41, 42, 44, 45, 46, 135, 136

Box, G. and Jenkins, G. (1970). Time series analysis: Forecasting and control. *Holden-Day, San Francisco*. 29

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52, 345–370. 35

Branco, M.D. and Dey D.K (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis* 79, 99–113. 67, 127

Brania, K. and Gurgul, H. (2015). The impact of estimation methods and data frequency on the results of long memory assessment. *Managerial Economics*, 16, 7–37. 19, 38, 39, 45, 135

Bratley, P., Fox, B. L. and Schrage, L. E. (1987). *A guide to simulation*, second ed.; Springer Science \& Business Media: New York, 1–397. 54, 154

Breheny, P. (2013). Model selection I, BST 760: Advanced regression. 1–32. https://myweb.uiowa.edu/pbreheny/uk/teaching/760-s11/notes/2-17.pdf (accessed on 23 August 2024). 36, 145

Brownlees, C. and Engle, R. (2012). Volatility, correlation and tails for systemic risk measurement. *SSRN Working Paper*. 15

Buccheri, G., Bormetti G., Corsi F. and Lillo F. (2021). A score-driven conditional correlation model for noisy and asynchronous data: An application to high-frequency covariance dynamics. *Journal of Business and Economic Statistics* 39, 920–36. 6, 67

Buckland, S., Burnham, K. and Augustin, N. (1997). Model selection: An integral part of inference. *Biometrics*, 53, 603–618. 36

Burnham, K. and Anderson, D. (2003). *Model selection and multimodel inference: A practical information-theoretic approach (2nd ed.)*. New York, NY: Springer. 34

Cai, J., Cheung, Y. and Wong, M. (2001). What moves the gold market? *Journal of FuTures Markets: Futures, Options, and Other Derivative Products*, 21, 257–278. 19

Campbell, J., Lo, A., MacKinlay, A. and R.F., W. (1998). The econometrics of financial market. *Macroeconomic Dynamics*, 2, 559–562. 19

Catania, L., Ardia, D. and Boudt, K. (2020). Generalized autoregressive score models. Package 'GAS' Version 0.3.3. https://cran.r-project.org/web/packages/GAS/GAS.pdf, (accessed on 26 June 2024). 23, 26, 28, 88, 90, 92, 95, 122, 124

Catania, L. and Billé, A. (2017). Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Applied Econometrics*, 32, 1178–1196. 15, 20, 86, 118, 157

Chalmers, P. (2019). Introduction to Monte Carlo simulations with applications in R using the SimDesign package, 1–46. https://philchalmers.github.io/SimDesign/pres.pdf, (accessed on 26 June 2024). 51, 57, 64, 65, 91, 93, 94, 95, 126

Chalmers, R. P. and Adkins, M. C. (2020). Writing effective and reliable Monte Carlo simulations with the SimDesign package. *The Quantitative Methods for Psychology*, 16, 248–280. 8, 17, 23, 54, 59, 64, 70, 92, 93, 114, 154

Chaurasia, A. and Harel, O. (2013). Model selection rates of information based criteria. *Electronic Journal of Statistics*, 7, 2762–2793. 35, 36, 37

Chen, S. and Shen, C. (2000). GARCH, jumps, and permanent and transitory components of volatility: The case of the Taiwan exchange rate. *Mathematics and Computers in Simulation*, 67, 201–216. 14, 137, 142, 157

Cheung, Y. (1993). Long memory in foreign-exchange rates. *Journal of Business and Economic Statistics*, 11, 93–102. 18

Cheung, Y. (2008). Test for fractional integration: A Monte Carlo investigation. *Journal of Time Series Analysis*, 14, 331–345. 18, 21, 41, 42, 136

Cheung, Y. and Diebold, F. (1994). On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. *Journal of Econometrics*, 62, 301–316. 46

Cheung, Y. and Lai, K. (1993). A fractional cointegration analysis of purchasing power parity. *Journal of Business and Economic Statistics*, 1, 103–112. 18, 19

Chiang, S. M., Yeh, C. P. and Chiu, C. L. (2006). Permanent and transitory components in the Chinese stock market: The ARJI-Trend model. *Taylor and Francis*, 45, 35–55. 14, 137, 141, 142, 157

Chib, S. (2015). Monte Carlo methods and Bayesian computation: Overview. *Elsevier* , 763–767. 70

Chinhamu, K., Huang, C. K. and Chikobvu, D. (2017). Evaluating risk in gold prices with generalized hyperbolic and stable distributions. *Proceedings of the 57th Annual Conference of SASA*, 17–24. 8, 147, 157

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using GARCH. *Journal of Applied Econometrics*, 3, 279–294. 14, 20, 24, 56, 118, 120, 137, 157

Choudhuri, N., Ghosal, S. and Roy, A. (2004). Contiguity of the Whittle measure for a Gaussian time series. *Biometrika*, 91, 211–218. 46

Christensen, W. (2018). Model selection using information criteria (Made Easy in SAS®). *University of California, Los Angeles, wchristensen@ucla.edu.* `https://github.com/wendychristensen`. 34, 35, 36, 37, 145

Chronopoulos, D., Fotios, I. and Nikolaos, V. (2018). Information demand and stock return predictability. *Journal of International Money and Finance*, 80, 59–74. 18

Claeskens, G. and Hjort, N. (2008). *Model selection and model averaging*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511790485(accessedon27August2024). 35

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. Springer. https://books.google.co.za/books?id=2nugUEaKqFEC. 37

Conrad, C. and Haag, B. (2006). Inequality constraints in the fractionally integrated GARCH model. *Journal of Financial Econometrics*, 4, 413–449. 44

Costa F. J. M. (2017). Forecasting volatility using GARCH models. Master in finance, University of Minho, Portugal. 7

Cottrell, A and Lucchetti R. J. (2023). Gnu regression, econometrics, and time-series library. *Gretl User's Guide*, 1–486. 54

Covarrubias, G. (2004). Persistence, sudden changes, and modeling volatility of financial time series. Doctor of philosophy, Texas Tech University. 87

Crato, N. and de Lima, P. (1994). Long-range dependence in the conditional variance of stock returns. *Economics Letters*, 45, 281–285. 20, 21

Creal, D., Koopman, S. J. and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics*, 29, 552–563. 6, 15, 20, 26, 28, 49, 88, 90, 96, 124, 138

Creal, D., Koopman, S. J. and Lucas, A. (2012). Generalized autoregressive score models. *Aenorm*, 20, 37–41. 26, 27, 28, 88, 89, 90, 122, 123, 124

Creal, D., Koopman, S. J. and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics*, 29, 552–563. 6, 15, 20, 26, 28, 49, 88, 90, 96, 124, 138

Creal, D., Koopman, S. J. and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28, 777–795. 5, 6, 9, 15, 20, 26, 27, 28, 49, 67, 88, 89, 90, 96, 101, 122, 123, 124, 138, 150, 155

Creal, D., Koopman, S. and Lucas, A. (2008). A general framework for observation driven time-varying parameter models. *Tinbergen Institute Discussion Paper, Amsterdam, TI*, 2008–108/4. 15, 16, 20, 49, 124

Creal, D., Koopman S. J. and Lucas A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28, 777–95. 5, 6, 9, 15, 20, 26, 27, 28, 49, 67, 88, 89, 90, 96, 101, 122, 123, 124, 138, 150, 155

Creal, D., Schwaab, B., Koopman, S. and Lucas, A. (2014). Observation-driven mixed-measurement dynamic factor models with an application to credit risk. *The Review of Economics and Statistics*, 96, 898–915. 6, 15, 26, 28, 86, 88, 90, 118, 157

CRISIL. (2022). Indian Economy: States of aftermath. *CRISIL: An S&P Global Economy*. https://shorturl.at/2WGtE (accessed on 26 June 2024). 141

Curto, J., Pinto, J. C. and Tavares, G. N. (2009). Modeling stock markets' volatility using GARCH models with normal, Student's t and stable Paretian distributions. *Statistical Papers*, 50, 311–321. 12, 20

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Annals of Statistics*, 17, 1749–1766. 40

Danielsson, J. (2011). *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*; John Wiley & Sons: Chichester, West Sussex, 1–274. 59, 92

Datastream. (2021). Thomson reuters datastream. https://solutions.refinitiv.com/datastream-macroeconomic-analysis? (accessed on 17 June 2021). 54, 87, 119

Davidian, M. (2016). Simulation studies in statistics. North Carolina State University, North Carolina, 1–6, 2016 http://homepage.divms.uiowa.edu/~kcowles/STAT5400_2016/lect09.pdf (accessed on 25 June 2024). 95

Davidson, R. and MacKinnon, J. (2004). *Econometric theory and methods*. 613. New York, Oxford University. 31

De Santis, G. and Gerard, B. (1997). International asset pricing and portfolio diversification with time-varying risk. *The Journal of Finance*, 52, 1881–1912. 15

Dickey, D. (1976). *Estimation and hypothesis testing in nonstationary time series*. Phd dissertation, Iowa State University. 30

Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series With a unit root. *Journal of the American Statistical Association*, 74, 427–431. 29

Diebold, F. X. and Lopez, J. A. (1995). Modeling volatility dynamics. In *Macroeconometrics: Developments, Tensions, and Prospects*; Hoover, K.D., Ed.; Springer Netherlands: Dordrecht, 427–472. 86

Diebold, F. and Rudebusch, G. (1989). Long memory and persistence in aggregate output. *Journal of Monetary Economics*, 24, 189–209. 18, 19, 21

Diebold, F. and Rudebusch, G. (1991). Is consumption too smooth? Long memory and the Deaton paradox. *Reviews for Economics Letters*, 35, 155–60. 18, 21

Ding, Z., Granger, C. and Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, 83–106. 19, 25, 56, 121, 148, 160

Ding Z. and Granger C. W. J. (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73, 185–215. 53

Dralle, B. (2011). Modelling volatility in financial time series. Master of science in statistics, University of KwaZulu-Natal. 145

Duda, M. and Schmidt, H. (2009). Evaluation of various approaches to value at risk. Master Thesis, Lund University, Lund, Sweden. https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=1436923&fileOId=1646971 (accessed on 30 August 2023). 25, 56, 68, 99, 153

Dufour, J. (2007). Model selection. *Université de Montréal*, 1–18. https://jeanmariedufour.research.mcgill.ca/Dufour_2006_Palgrave_ModelSelection_W.pdf (accessed on 27 August 2024). 35

Eckernkemper, T. (2018). Modeling systemic risk: time-varying tail dependence when forecasting marginal expected shortfall. *Journal of Financial Econometrics*, 16, 63–117. 15, 20

Elder, J. and Serletis, A. (2008). Long memory in energy futures prices. *Review of Financial Economics*, 17, 146–155. 19

Eling, M. (2014). Fitting asset returns to skewed distributions: Are the skew-normal and skew-student good models? *Insurance: Mathematics and Economics*, 59, 45–56. 67, 127

Engle, R. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17, 425–446. 15

Engle, R. F. (1982). Autoregressive conditional heteroscedacity with estimates of variance of United Kingdom inflation. *Econometrica*, 50, 987–1008. 6, 15, 23, 26, 33, 55, 76, 88, 110, 119, 131, 151

Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5, 1–50. 14, 20, 24, 56, 118

Engle, R. and Lee, G. (1999). A permanent and transitory component model of stock return volatility, in R.F. Engle and H. White (eds.), cointegration, causality, and forecasting: A festschrift in honor of Clive W.J. Granger. *Oxford University Press, New York*, 475–497. 50, 51, 125, 142

Engle, R., Lilien, D. and Robins, R. (1987). Estimating time-varying risk premia in the term structure: the ARCH-M model. *Econometrica*, 55, 391–407. 16

Engle, R. F. and Patton, A. J. (2001). What good is a volatility model?. *Quantitative Finance*, 1, 237–245. 7, 14, 20, 28, 117, 118, 120, 128, 131, 134, 136, 137, 141, 157

Engle, R. and Russell, J. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66, 1127–1162. 26, 88

Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5, 1–50. 14, 20, 24, 56, 118

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of Finance*, 48, 17749–1778. 25, 56, 121

Engle, R. F. and Sokalska M. E. (2012). Forecasting intraday volatility in the US equity market. Multiplicative component GARCH. *Journal of Financial Econometrics*, 10, 54–83. 67

Enow, S. (2023). Forecasting volatility in international financial markets. *International Journal of Research*, 12, 197–203. 12

Eom C., Kaizoji T. and Scalas E. (2019). Fat tails in financial return distributions revisited: Evidence from the Korean stock market. *Physica A: Statistical Mechanics and its Applications*, 526(3), 121055. 8

EViews (2016). *EViews 9 users guide II*. IHS Global Inc. 31, 32

Fan, J., Qi L. and Xiu D. (2014). Quasi-maximum likelihood estimation of GARCH models with heavy-tailed likelihoods. *Journal of Business and Economic Statistics*, 32, 178–91. 24, 55

Fang, H., Lai, K. and Lai, M. (1994). Fractal structure in currency futures prices. *Journal of Futures Markets*, 14, 169–181. 19

Feng, L. and Shi, Y. (2017). A simulation study on the distributions of disturbances in the GARCH model. *Cogent Economics and Finance*, 5, 1355503. 8, 12, 20, 24, 56, 59, 60, 64, 69, 92, 93, 99, 114, 118, 127, 150, 154

Fernández, C. and Steel, M. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, *93*, 359–371. 49, 51, 124, 126, 140

Ferrer-Comalat, J. C., Linares-Mustarós, S., Merigó, J. M. and Kacprzyk, J. (2020). *Modelling and simulation in management sciences*, first ed.; Vol. 894, Springer Nature Switzerland AG, 1–218. 154

Fisher, T. J. and Gallagher, C. M. (2012). New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Association*, 107, 777–787. 32, 76, 110, 131

Fleming, J. and Kirby, C. (2011). Long memory in volatility and trading volume. *Journal of Banking and Finance*, 35, 1714–1726. 19

Floros, C. (2008). Modelling volatility using GARCH models: Evidence from Egypt and Israel. *Middle Eastern Finance and Economics*, 2, 31–41. 13, 20

Floros, C., Jaffry, S. and Lima, G. (2007). Long memory in the Portuguese stock market. *Studies in Economics and Finance*, 24, 220–232. 19

Foote, W. G. (2018). *Financial engineering analytics: A practice manual using R*. https://bookdown.org/wfoote01/faur/ (accessed on 25 June 2024). 59, 92

Forster, M. (2004). Simplicity and unification in model selection. *University of Wisconsin Madison*. https://philpapers.org/rec/FORCS (accessed on 23 August 2024). 35

Fouquau, J. and Spieser, P. (2014). Stock returns memories: a "stardust" memory? *Dans Finance*, 35, 57–85. 41, 44

Fox, R. and Taqqu, M. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics*, 14, 517–532. 18, 21, 40, 46

Francq, C. and Thieu, L. (2019). QML inference for volatility models with covariates. *Econometric Theory*, 35, 37–72. 68

Francq, C. and Zakoïan, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10, 605–637. 24, 56

Fuller, W. (1976). *Introduction to statistical time series*. Wiley, New York. 30

Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. (2006). Institutional investors and stock market volatility. *The Quarterly Journal of Economics*, 121, 461–504. 19

Geweke, J. (1986). Comment on: Modelling the persistence of conditional variances. *Econometric Reviews*, 5, 57–61. 25, 57

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221–238. 6, 18, 20, 21, 38, 44, 45, 46, 135, 136

Ghalanos, A. (2018). Introduction to the Rugarch package. (Version 1.3-8). mirrors.nic.cz/R/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf (accessed on 30 August 2023). 6, 10, 23, 24, 25, 28, 56, 57, 59, 60, 63, 66, 67, 68, 69, 79, 92, 94, 96, 108, 112, 120, 121, 127, 132, 133, 138, 140, 150

Ghalanos, A. (2022). Rugarch: Univariate GARCH models. R Package Version 1.4-7. https://cran.r-project.org/web/packages/rugarch/rugarch.pdf (accessed on 30 August 2023). 10, 23, 25, 42, 44, 54, 57, 59, 63, 92, 120, 121

Ghalanos, A. (2024). Introduction to the rugarch package (version 1.4-3). 1–50. https://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf (accessed on 27 August 2024). 37

Gilli, M., Maringer, D. and Schumann, E. (2019). Generating random numbers. In *Numerical Methods and Optimization in Finance*; Janco, C., Ed.; Academic Press, 103–132. 66, 94, 114, 154

Giovanni, B. and Distaso, W. (2003). 'GRETL : Econometric software for the GNU generation. *Journal of applied econometrics*, 18, 105–1102. 54

Giovanni, R. (2020). Essential Monte Carlo analysis. In *Elements of Numerical Mathematical Economics with Excel*; Romer, B., Ed.; Elsevier Academic Press, 763–795. 87

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48, 1779–1801. 17, 25, 56, 121

Gokcan, S. (2000). Forecasting volatility of emerging stock markets: Linear versus non-linear GARCH models. *Journal of Forecasting*, 19, 499–504. 12, 20

Gorgi, P., Hansen, P., Janus, P. and Koopman, S. (2019). Realized Wishart-GARCH: a score-driven multi-asset volatility model. *Journal of Financial Econometrics*, 17, 1–32. 15, 20, 86, 118, 157

Granger, C. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–29. 18, 20, 38, 41, 42, 47

Guha, B., Avijan, D. and Gautam, B. (2016). Measurement of risk vs return of Indian sectoral indices. *Journal of Advanced Management Science*, 4, 106–111. 18

Haddad, M. F. C., Blazsek, S., Arestis, P., Fuerst, F. and Sheng, H. H. (2023). The two-component Beta-t-QVAR-M-lev: a new forecasting model. *Financial Markets and Portfolio Management*, 37, 379–401. 11, 15, 16, 17, 20, 86, 118, 122, 123, 124, 127, 148, 160

Hallgren, K. A. (2013). Conducting simulation studies in the R programming environment. *Tutorials in Quantitative Methods for Psychology*, 9, 43–60. 59, 91

Hamaker, E.L., van Hattum, P., Kuiper, R. M. and Hoijtink, H. (2011). Model selection based on information criteria in multilevel modeling. *In J. J. Hox and J. K. Roberts (Eds.), Handbook for advanced multilevel analysis, Routledge/Taylor and Francis Group*, 231–255. 34

Hannan, E. (1980). The estimation of the order of an ARMA process. *Annals of Statistics*, 8, 1071–1081. 36

Hannan, E. and Quinn, B. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B (Methodological)*, 41, 190–195. http://www.jstor.org/stable/2985032 (accessed on 27 August 2024). 36

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails: With applications to financial and economic time series*; Cambridge University Press: New York, USA, 1–261. 5, 9, 15, 20, 26, 28, 49, 88, 90, 120, 122, 124

Harvey, A. and Chakravarty, T. (2008). Beta-t-EGARCH. Cambridge working papers in economics 0840. *Faculty of Economics, University of Cambridge*, 1–41. https://econpapers.repec.org/paper/camcamdae/0840.htm (accessed on 25 June 2024). 15, 16, 20, 26, 49, 88, 124

Harvey, A. and Lange, R. (2018). Modeling the interactions between volatility and returns using EGARCH-M. *Journal of Time Series Analysis*, 39, 909–919. 16, 17, 20

Harvey, A. and Lange, R. (2017). Volatility modeling with a generalized t distribution. *Journal of Time Series Analysis*, 38, 175–190. 16

Harvey, A. and Sucarrat, G. (2014). EGARCH models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis*, 26, 320–338. 5, 6, 11, 17, 20, 21, 49, 50, 51, 86, 117, 118, 124, 125, 126, 140, 145, 155, 157

Harvey, A. and Thiele, S. (2016). Testing against changing correlation. *Journal of Empirical Finance*, 38, 575–589. 15, 20, 86

Harwell, M. (2018). A strategy for using bias and RMSE as outcomes in Monte Carlo studies in statistics. *Journal of Modern Applied Statistical Methods*, 17, 1–16. 64, 93

Hassler, U. (2011). Estimation of fractional integration under temporal aggregation. *Journal of Econometrics*, 162, 240–247. 20

Hassler, U. and Woiters, J. (1995). Long memory in inflation rates: International evidence. *Journal of Business and Economic Statistics*, 13, 37–45. 20, 21

Helms, B., Kaen, F. and Rosenman, R. (1984). Memory in commodity futures contracts. *Journal of Futures Market*, 4, 559–567. 19

Henry, O. (2002). Long memory in stock returns: Some international evidence. *Applied Financial Economics*, 12, 725–729. 19

Hentschel, L. (2007). All in the family nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, 39, 71–104. 5, 6, 25, 56, 57, 68, 86, 99, 118, 121, 150, 153, 155

Heracleous, M. S. (2007). Sample kurtosis, GARCH-t and the degrees of freedom issue. *EUR Working Papers*, 1–22. https://cadmus.eui.eu/handle/1814/7693 (accessed on 25 June 2024). 68, 110

Hertog, S., Gerland, P. and Wilmoth, J. (2022). India overtakes China as the world's most populous country. *UN DESA Population Division, World Population Prospects: Summary of Results*, 1–5. 119

Henry, O. (2002). Long memory in stock returns: Some international evidence. *Applied Financial Economics*, 12, 725–729. 19

Hiemstra, C. and Jones, J. (1997). Another look at long memory in common stock returns. *Journal of Empirical Finance*, 4, 373–401. 19

Higgins, M. L. and Bera, A. K. (1992). A class of nonlinear Arch models. *International Economic Review*, 33, 137–158. 25, 56, 121

Hilary, T. (2002). Descriptive statistics for research. https://www.stats.ox.ac.uk/pub/bdr/IAUL/Course1Notes2.pdf (accessed on 31 August 2023). 54, 61, 126

Hodrick, R. (1981). International asset pricing with time-varying risk premia. *Journal of International Economics*, 11, 573–587. 15

Hoga, Y. (2022). Extremal dependence-based specification testing of time series. *Journal of Business and Economic Statistics*, 41, 1274–1287. 68

Holan, S., McElroy, T. and Chakraborty, S. (2009). A Bayesian approach to estimating the long memory parameter. *Bayesian Analysis*, 4, 159–190. 38

Hosking, J. (1981). Fractional differencing. *Biometrika*, 68, 165–176. 18, 20, 38, 41, 42

Hull, M. and McGroarty, F. (2014). Do emerging markets become more efficient as they develop? Long memory persistence in equity indices. *Emerging Markets Review*, 18, 45–61. 19

Hurst, H. (1951). Long term storage capacity of reservoirs. *Proceedings of the American Society of Civil Engineers*, 11, 770–99. 38

Hurvich, C., Deo, R. and Brodsky, J. (1998). The mean squared error of geweke and porter-hudak's estimator of the long memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19, 1095–1112. 46, 136

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27, 1–22. 10, 23, 59, 92, 110

Janssen, R. and De Boeck, P. (1999). Confirmatory analyses of componential test structure using multidimensional item response theory. *Multivariate Behavioral Research*, 34, 245–268. 35

Janus, P., Koopman, S. J. and Lucas, A. (2014). Long memory dynamics for multivariate dependence under heavy tails. *Journal of Empirical Finance*, 29, 187–206. 15, 20, 86

Javed, F. and Mantalos, P. (2013). GARCH-type models and performance of information criteria. *Communications in Statistics: Simulation and Computation*, 42, 1917–1933. 36, 67

Johansen, A. (2010). Monte Carlo methods. In *International Encyclopedia of Education*; Peterson, P.; Baker, E.; MCGaw, B., Eds.; Elsevier Academic Press, 296–303. 87

Jones, O., Maillardet, R. and Robinson, A. (2014). *Introduction to scientific programming and simulation using R*, second ed.; Chapman & Hall/CRC: Florida, 1–573. 154

Kaen, F. and Rosenman, R. (1986). Predictable behaviour in financial markets: some evidence in support of Heiner's hypothesis. *American Economic Review*, 76, 212–220. 19

Kenton, W. (2021). Monte Carlo simulation. `https://ln.run/c22uo` (accessed on 25 June 2024). 87

Kim, S.Y., Huh, D., Zhou, Z. and Mun, E. Y. (2020). A comparison of Bayesian to maximum likelihood estimation for latent growth models in the presence of a binary outcome. *International Journal of Behavioral Development*, 44, 447–57. 23, 24, 55, 56

Kim, C. and Philips, P. (1999). Log periodogram regression in the nonstationary case. *Mimeo, Yale University*. 45

Kleijnen, J. P. C. (2015). *Design and analysis of simulation experiments*, 2nd ed. New York: Springer, vol. 230. 54, 154

Konishi, S. and Kitagawa, G. (1996). Generalised information criteria in model selection. *Biometrika*, 83, 875–890. 34

Koop, G., Ley, E., Osiewalski, J. and Steel, M. (1997). Bayesian analysis of long memory and persistence using ARFIMA models. *Journal of Econometricss*, 76, 149–169. 18

Koopman, S. J., Lucas, A. and Zamojski, M. (2017). Dynamic term structure models with score driven time varying parameters: estimation and forecasting. *NBP Working Paper*, 3–80. 27, 28, 61, 89, 90, 92, 122, 124

Kosapattarapim, C., Lin, Y. and Mccrae, M. (2012). Evaluating the volatility forecasting performance of best fitting GARCH models in emerging Asian stock markets. *International Journal of Mathematics and Statistics*, 12, 1–15. 13, 20

Kotz, S., Kozubowski, T. and Podgorski, K. (2001). *The laplace distribution and generalizations*; Springer-Verlag: New York, 1–441. 96, 138

Koutmos, G. (1998). Asymmetries in the conditional mean and the conditional variance: Evidence from nine stock markets. *Journal of Economics and Business*, 50, 277–90. 12, 20

Kullback, S. and Leibler, R. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22, 79–86. 34

Künsch, H. (1986) Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability*, **23**, 1025–1030. 45

Künsch, H. (1987) Statistical aspects of self-similar processes. *in Prokhorov, Y. and Sazanov, V.V. (eds), Proceedings of the First World Congress of the Bernoulli Society. Utrecht, VNU Science Press*, 67–74. 6, 47

Kwiatkowski, R. (2022). Monte Carlo simulation — a practical guide. `https://shorturl.at/31I0z` (accessed on 25 June 2024). 87

Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159–178. 31

Lange, R., Lucas, A. and Siegmann, A. (2017). Score-driven systemic risk signaling for European sovereign bond yields and CDS spreads. *In: Bilio, M., Pelizzon, L., Savona, R. (eds.) Systemic Risk Tomography: Signals. Measurement and Transmission Channels. Elsevier, Amsterdam*. 15, 20

Lazar, E. and Xue, X. (2020). Forecasting risk measures using intraday data in a generalized autoregressive score framework. *International Journal of Forecasting*, 36, 1057–1072. 16, 20

Lee, G. and Engle, R. (1999). A permanent and transitory component model of stock return volatility. In Cointegration Causality and Forecasting A Festschrift in Honor of Clive WJ Granger. *Oxford University Press*, 475–497. 14, 16, 17, 67

Lee, Y. H. and Pai, T. Y. (2010). REIT volatility prediction for skew-GED distribution of the GARCH model. *Expert Systems with Applications*, 37, 4737–4741. 67, 127

Lee, D. and Robinson, P. (1996). Semiparametric exploration of long memory in stock market prices. *Journal of Statistical Planning and Inference*, 50, 155–174. 19

Leite, A. L., Pinto, A. C. F and Klotzle, M. C. (2016). Efeitos da Volatilidade Idiossincrática na Precificação de Ativos. *Revista Contabilidade e Financas*, 27(70), 98–112. 7

Leitner, S. and Wall, F. (2015). Simulation-based research in management accounting and control: an illustrative overview. *Journal of Management Control*, 26(2-3), 105–129. 5

Leschinski, C., Voges, M. and Wenger, K. (2019). Long memory time series. Package 'LongMemoryTS' Version 0.1.0. https://cran.r-project.org/web/packages/GAS/GAS.pdf (accessed on 22 July 2024). 23, 136

Li, F., Cohen, A., Kim, S. and Cho, S. (2009). Model selection methods for mixture dichotomous IRT models. *Applied Psychological Measurement*, 33, 353–373. 35, 36, 112, 131, 132, 133, 145

Li, Q. (2008). Three essays on stock market volatility. All Graduate Theses and Dissertations, Spring 1920 to Summer 2023. 308. http://digitalcommons.usu.edu/etd/308/ (accessed on 30 August 2023). 8, 69, 96, 153

Li, W. and McLeod, A. (1986). Fractional time series modelling. *Biometrika*, 73, 217–21. 18, 21

Lim, K. and Brooks, R. (2010). Why do emerging stock markets experience more persistent price deviations from a random walk over time? A country-level analysis. *Macroeconomic Dynamics*, 14, 3–41. 19

Lin, C. H. and Shen, S. S. (2006). Can the student-t distribution provide accurate value at risk? *Journal of Risk Finance*, 7, 292–300. 68, 99, 153

Liow, K. (2009). Long-term memory in volatility: Some evidence from international securitized real estate markets. *The Journal of Real Estate Finance and Economics*, 39, 415–438. 19

Lo, A. (1991). Long-term memory in stock market prices. *Econometrica*, 59, 1279–314. 18, 19, 21

Lopes, S., Olbermann, B. and Reisen, V. (2002). Non-stationary Gaussian ARFIMA processes: Estimation and application. *Brazilian Review of Econometrics Rio de Janeiro*, 22, 103–126. 40

Lv, Q., Han, L., Wan, Y. and Yin, L. (2018). Stock net entropy: Evidence from the Chinese growth enterprise market. *Entropy*, 20, 805. 18

Maciel, L. D. S. and Ballini, R. (2017). Value-at-risk modeling and forecasting with range-based volatility models: Empirical evidence. *Revista Contabilidade e Financas*, 28, 361–376. 7, 24, 55

MacKinnon, J. (1991). Critical values for cointegration tests. *in R.F. Engle and C.W.J. Granger (eds), Long-run Economic Relationships: Readings in Cointegration, Oxford University Press, Oxford*, 267–276. 30

Madhavan, V. and Arrawatia, R. (2016). Relative efficiency of G8 sovereign credit default swaps and bond scrips: An adaptive market hypothesis perspective. *Studies in Microeconomics*, 4, 127–150. 19

Maheswaran, S. (1990). Predictable short-term variation in asset prices: Theory and evidence. *Working Paper, Carlson School of Management, University of Minnesota*. 18

Mamilla, R., Kathiravan, C., Salamzadeh, A., Dana, L. and Elheddad, M. (2023). COVID-19 pandemic and indices volatility: Evidence from GARCH models. *Journal of Risk and Financial Management*, 16, 447. 11, 13, 18, 20

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36, 394–419. 111

Mandelbrot, B. (1971) A fast fractional Gaussian noise generator. *Water Resources Research*, **7**, 543–553. 38

Mandelbrot, B. and Van Ness, J. (1968) Fractional Brownian motion, fractional noises and applications. *SIAM Review*, **10**, 422–437. 38

McLeod, A. and Hipel, K. (1978). Preservation of the rescaled adjusted range. *Water Resources Research*, 14, 491–518. 42

McMillan, D. and Thupayagale, P. (2009). The efficiency of african equity markets. *Studies in Economics and Finance*, 26, 275–292. 19

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7, 271–300. 23, 55

Mehtab, S., Jaydip, S. and Abhishek, D. (2021). Stock price prediction using machine learning and LSTM-based deep learning models. *In Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, October 14–17, 2020, Revised Selected Papers 2. Singapore: Springer*, 88–106. 18

Mehtab, S. and Sen, J. (2020). Stock price prediction using convolutional neural network on a multivariate time series. *Paper presented at the 3rd National Conference on Machine Learning and Artificial Intelligence (NCMLAI'20), New Delhi, India*. 18

Mensi, W., Tiwari, A. and Al-Yahyaee, K. (2019). An analysis of the weak form efficiency, multi-fractality and long memory of global, regional and European stock markets. *The Quarterly Review of Economics and Finance*, 72, 168–177. 19

Mighri, Z. and Mansouri, F. (2014). Modeling international stock market contagion using multivariate fractionally integrated aparch approach. *Cogent Economics and Finance*, 2, 963632. 19

Mills, T. (1993). Is there long memory in UK stock returns? *Applied Financial Economics*, 3, 303–306. 19

Mishra, P. K. (2010). A GARCH model approach to capital market volatility: The case of India. *Indian Journal of Economics & Business*, 9, 631–641. 140

Mohandas, G. (2023). Volatility modelling and forecasting in financial markets. *International Research Journal of Modernization in Engineering Technology and Science*, 5. 12, 13, 20

Mooney, C. Z. (1997). *Monte Carlo simulation*, 1st ed. Thousand Oaks: SAGE Publications, 116, 1–103. 61, 92

Morris, T. P., White, I. R. and Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38, 2074–2102. 8, 17, 53, 60, 64, 65, 66, 68, 87, 92, 93, 94, 114, 154

Mota, M. M. and Mota, I. F. D. L. (2017). *Applied simulation and optimization 2*, first ed.; Springer International Publishing AG: Cham, Switzerland, 2017; 1–282. 154

Mukhodobwane, R. M., Sigauke, C., Chagwiza, W. and Garira, W. (2022). Stochastic modelling of the BRICS equity markets' risks. *Journal of Statistics Applications and Probability*, 11, 215–239. 37

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59, 347–70. 16, 24, 25, 26, 55, 56, 88, 121

Nguyen, D., Prokopczuk, M. and Sibbertsen, P. (2019). The memory of stock return volatility: Asset pricing implications. *Journal of Financial Markets*, 47, 100487. 18

Nishii, R. (1984). Asymptotic properties of criteria for selection of variables in multiple regression. *Annals of Statistics*, 12, 758–765. 36, 145

Oh, D. and Patton, A. (2013). Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association*, 108, 689–700. 15, 20

Oh, D. and Patton, A. (2018). Time-varying systemic risk: evidence from a dynamic copula model of CDS spreads. *Journal of Business & Economic Statistics*, 36, 181–195. 15, 20, 27, 86, 89, 118, 123, 155, 157

Opschoor, A., Janus, P., Lucas, A. and Van Dijk, D. (2018). New HEAVY models for fat-tailed realized covariances and returns. *Journal of Business and Economic Statistics*, 36, 643–657. 15, 20, 26, 86, 88, 118, 122, 157

Owusu-Junior, P. and Alagidede, I. (2020). Risks in emerging markets equities: Time-varying versus spatial risk analysis. *Physica A: Statistical Mechanics and its Applications*, 542, 123474. 26, 88

Pagan, A. and Schwert, G. (1990). Testing for covariance stationarity in stock market data. *Economics Letters*, 33, 165–170. 14

Pagan, A. and Schwert, G. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, 45, 267–290. 15

Palma, W. (2006). *Long-memory time series: Theory and methods*. Wiley Series in Probability and Statistics. Wiley. 38

Palma, W. (2007). *Long-memory time series: Theory and methods*. Wiley Series in Probability and Statistics. Wiley. https://books.google.co.za/books?id=HhGa8CcUsWIC (accessed on 25 June 2024). 19

Pandey, R. and Kumar, A. (2017). Modelling persistence in conditional volatility of asset returns. *Afro-Asian Journal Finance and Accounting*, 7, 16–34. 13, 14, 20, 118, 138, 146

Pantula, S. G. (1986). Comment: Modelling the persistence of conditional variances. *Econometric Reviews*, 5, 71–74. 25, 57

Patton, A., Ziegel, J. and Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value-atrisk). *Journal of Econometrics*, 211, 388–413. 16, 20

Pedersen, T. L. (2020). patchwork: The composer of plots. https://CRAN.R-project.org/package=patchwork (accessed on 25 June 2024). 23, 92

Pfaff, B. (2016). *Modelling volatility*. John Wiley & Sons, Ltd., 116–32. 54, 63, 68

Phillips, P. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75, 335–346. 30

Pollmann, M. (2015). Are you sure you are using the correct model? Model selection and averaging of impulse responses. *Marble series: Quantitative methods in Business and Economics*, 1, 1–45. 36

Porterba, J. and Summers, L. (1986). The persistence of volatility and stock market fluctuations. *American Economic Review*, 76, 1143–1151. 7, 14, 117

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association*, 85, 338–44. 18

Pourahmadi, M. (2007). Construction of skew-normal random variables: Are they linear combinations of normal and half-normal? *Journal of Statistical Theory and Applications*, 3, 314–328. 67, 127

Qiu, D. (2015). aTSA: Alternative time series analysis. `https://cran.r-project.org/web/packages/aTSA/aTSA.pdf` (accessed on 30 August 2023). 23, 59, 92

Rafik, N., Esmaeil, N., Nadiya, G. and Ashkan, A. (2014). Long memory analysis: An empirical investigation. *International Journal of Economics and Financial Issues*, 4, 16–26. 38, 42, 120, 135

Robinson, P. (1990). Time series with strong dependence. *Advances in econometrics, 6th world congress (Cambridge University Press, Cambridge)*. 45

Robinson, P. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, 47, 67–84. 18

Robinson, P. (1995a). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23, 1630–1661. 6, 44, 45, 47, 135

Robinson, P. (1995b). Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, 23, 1048–1072. 48

Rossi, R. (2014). Unit roots tests. *Financial Econometrics*, 10, 1–40. 29, 30, 31, 32

Russell, J. (2001). Econometric modeling of multivariate irregularly-spaced high-frequency data. *University of Chicago, Graduate School of Business* 2001. 26, 88

Rydberg, T. and Shephard, N. (2003). Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics*, 1, 2. 26, 88

Sadique, S. and Silvapulle, P. (2001). Long-term memory in stock market returns: International evidence. *International Journal of Finance and Economics*, 6, 59–67. 19, 21

Saha, K., Madhavan, V. and Chandrashekhar, G. (2020). Pitfalls in long memory research. *Cogent Economics and Finance*, 8, 1733280. 18, 20, 38, 39, 42, 44

Samiev, S. (2012). *GARCH (1,1) with exogenous covariate for EUR/SEK exchange rate volatility: On the effects of global volatility shock on volatility*. Master's thesis, Umea University, Umea, Sweden. https://www.diva-portal.org/smash/get/diva2:676106/FULLTEXT01.pdf (accessed on 31 August 2023). 7, 24, 55

Samuel, T., Chimedza, C. and Sigauke, C. (2023). Simulation framework to determine suitable innovations for volatility persistence estimation: The GARCH approach. *Journal of Risk and Financial Management*, 16, 392. 52, 86, 90, 91, 92, 93, 114, 126, 152, 155

Satchell, S. and Knight, J. (2011). *Forecasting volatility in the financial markets*; Elsevier. 86

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464. 36

Schwert, G. (1989). Why does stock market volatility change over time? *Journal of Finance*, 44, 1115–1153. 14, 30

Schwert, G. W. (1990). Stock volatility and the crash of '87. *The Review of Financial Studies*, 3, 77–102. 25, 56, 121

Sen, J., Sidra, M. and Abhishe, D. (2021). Volatility modeling of stocks from selected sectors of the Indian economy using GARCH. *Paper presented at the 2021 Asian Conference on Innovation in Technology (ASIANCON), Pune, India, August 27–29*, 1–9. 11

Sensoy, A. and Tabak, B. (2016). Dynamic efficiency of stock markets and exchange rates. *International Review of Financial Analysis*, 47, 353–371. 19

Shahriari, S., Sisson, S. A. and Rashidi, T. (2023). Copula ARMA-GARCH modelling of spatially and temporally correlated time series data for transportation planning use. *Transportation Research Part C*, 146, 103969. 68

Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*, 63, 117–126. 35

Shibata, R. (1981). An optimal selection of regression variables. *Biometrika*, 68, 45–54. 36, 145

Shibata, R. (1989). Statistical aspects of model selection. *IIASA International Institute for Applied Systems Analysis, IIASA Working Paper*, 1–31. 36

Shimotsu, K. and Phillips, P. (2005). Exact local Whittle estimation of fractional integration. *Annals of Statistics*, 33, 1890–1933. 6, 48

Shea, G. (1991). Uncertainty and implied variance bounds in long-memory models of the interest rate term structure. *Empirical Economics*, 16, 287–312. 18, 19, 20, 21, 136

Shono, H. (2005). Is model selection using Akaike's information criterion appropriate for catch per unit effort standardization in large samples? *Fisheries science*, 71, 978–986. 36

Sigal, M. J. and Chalmers, P. R. (2016). Play it again: Teaching statistics with Monte Carlo simulation. *Journal of Statistics Education*, 24, 136–156. 8, 64, 65, 66, 70, 93, 94, 95, 114

Sigauke, C. (2016). Volatility modeling of the JSE all share index and risk estimation using the Bayesian and Frequentist approaches. *Economics, Management, and Financial Markets*, 11, 33–48. 81, 132

Silvennoinen, A. and Teräsvirta, T. (2016). Testing constancy of unconditional variance in volatility models by misspecification and specification tests. *Studies in Nonlinear Dynamics and Econometrics*, 20, 347–64. 60

Simon, S. and Amalia, M. (2012). Volatility in EMU sovereign bond yields: permanent and transitory components. *Applied Financial Economics*, 22, 1453–1464. 14

Smith, R. L. (2003). Statistics of extremes, with applications in environment, insurance, and finance. In *Extreme Values in Finance, Telecommunications, and the Environment*, 20–97. Chapman & Hall/CRC, Florida. 23, 55

Søfteland, A. and Iversen, G. S. (2021). *Applying GARCH-EVT-Copula forecasting in active portfolio management*. Master's thesis, NTNU, Trondheim, Norway. `https://no.ntnu_inspera_82752696_84801861.pdf` (accessed on 13 August 2023). 68

Sowell, F. (1990). Fractional unit root distribution. *Econometrica*, 50, 495–505. 18, 21

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53, 165–188. 18, 21, 40

Speight, A., McGillan, D. and Gwilym, O. A. (2000). Intraday volatility components in FTSE-100 stock index futures. *Journal of Futures Markets*, 20, 425–444. 14

Su, J. J. (2011). On the oversized problem of Dickey-Fuller-type tests with GARCH errors. *Communications in Statistics: Simulation and Computation*, 40, 1364–1372. 60, 101

Sucarrat, G. (2013). betategarch: Simulation, estimation and forecasting of first-order beta-skew-t-EGARCH models. *The R Journal*, 5, 137–147. 5, 6, 17, 20, 21, 49, 50, 51, 117, 118, 124, 125, 126, 140, 145, 150, 155, 157

Sucarrat, G. (2016). *betategarch: Simulation, estimation and forecasting of Beta-Skew-t-EGARCH Models*. `https://CRAN.R-project.org/package=betategarch`. R package version 3.3. 23

Takaishi, T. (2018). Volatility estimation using a rational GARCH model. *Quantitative Finance and Economics*, 2, 127–136. 11, 118, 150

Taqqu, M., Teverovsky, V. and Willinger, W. (1995). Estimators for long-range dependence: An empirical study. *Fractals*, 3, 785–798. 20

Taylor, S. J. (1986). *Modelling financial time series*, second ed.; World Scientific Publishing Co. Pte. Ltd., 1–297. 25, 56, 121

Taylor, S.and Poon, S. (1992). Stock returns and volatility: An empirical study of the UK stock market. *Journal of Banking and Finance*, 16, 37–59. 12, 20

Tinyakova, V. (2012). The new approaches in econometric research of financial markets. Distributed volatility. *Review of Applied Socio-Economic Research*, 4, 247–255. 12

Tsay, R. S. (2005). *Analysis of financial time series*; John Wiley and Sons: London, 154–205. 14

Ugurlu, E., Thalassinos, E. and Muratoglu, Y. (2014). Modeling volatility in the stock markets using GARCH models: European emerging economies and Turkey. *International Journal in Economics and Business Administration*, 2, 72–87. 13, 20

Velasco, C. (1999). Non-stationary log-periodogram regression. *Journal of Econometrics*, 91, 325–371. 45

Vera-Valdés, J. (2021). Temperature anomalies, long memory, and aggregation. *Econometrics*, 9, 1–22. 19, 38

Vera-Valdés, J. (2024). LongMemory.jl: Generating, estimating, and forecasting long memory models in Julia. *arXiv:2401.14077v1*. 19, 20, 21, 39, 44, 48, 136

Wang, C., Chen, Q. and Gerlach, R. (2019). Bayesian realized-GARCH models for financial tail risk forecasting incorporating the two-sided Weibull distribution. *Quantitative Finance*, 19(6), 1017–1042. 7

Wang, C., Gerlach, R. and Chen Q. (2018). A semi-parametric realized joint value-at-risk and expected shortfall regression framework. *arXiv*, arXiv:1807.02422. 66, 94

Wang, W., Van Gelder, P., Vrijling, J. and Ma, J. (2005). Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes. *Nonlinear process. Geophysics European Geosciences Union*, 12, 55–66. 33

Wang, W., Van Gelder, P., Vrijling, J. and Chen, X. (2007). Detecting long-memory: Monte Carlo simulations and application to daily streamflow processes. *Hydrology and Earth System Sciences*, 11, 851–862. 40, 41

Weismer, J. (2020). History as guide : Lessons from the 2002-2003 SARS outbreak. https://shorturl.at/okCsz (accessed on 18 June 2024). 108

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50, 1–25. 68

Whittle, P. (1951). Hypothesis testing in time series analysis. *Uppsala: Almqvist and Wiksells*. https://search.worldcat.org/title/Hypothesis-testing-in-time-series-analysis/oclc/22153644 (accessed on 27 August 2024). 46

Whittle, P. (1962). Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute*, 39, 105–129. 46

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4, 1–6. 23, 54, 59, 92, 154, 157, 159

Wicklin, R. (2017). How to choose a seed for generating random numbers. Making statistics accessible and machine learning understandable. https://shorturl.at/xTLrf (accessed on 18 June 2024). 61

Wuertz, D., Setz, T., Chalabi Y., Boudt C., Chausse P. and Miklovac M. (2020). *fGarch: Rmetrics—Autoregressive conditional heteroskedastic modelling*. R Package Version 3042.83.2.. https://cran.r-project.org/web/packages/fGarch/fGarch.pdf (accessed on 30 August 2023). 54, 68

Xunyu, Y., Ping, G. and Handong, L. (2015). Improving estimation of the fractionally differencing parameter in the SARFIMA model using tapered periodogram. *Economic Modelling*, 46, 167–179. 44

Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Annals of Statistics*, 16, 791–801. 18, 21

Yajima, Y. (1989). A central limit theorem of Fourier transforms of strongly dependent stationary processes. *Journal of Time Series Analysis*, 10, 373–83. 18, 21

Yuan, K. H., Tong, X. and Zhang, Z. (2015). Bias and efficiency for SEM with missing data and auxiliary variables: Two-stage robust method versus two-stage ML. *Structural Equation Modeling: A Multidisciplinary Journal*, 22, 178–192. 65, 66, 94

Zakoian, J. M. (1994). Threshold heteroscedastic models. *Journal of Economic Dynamics and Control*, 18, 931–955. 25, 56, 121

Zarour, B. and Siriopoulos, C. (2008). Transitory and Permanent volatility components: The case of the Middle East stock markets. *Review of Middle East Economics and Finance*, 4, 1–14. 14

Zeileis, A. and Grothendieck, G. (2005). zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 14, 1–27. 23, 59

Zhang, Y. T. (2017). *Volatility forecasting with the multifractal model of asset returns*. University of Florida, 1–24. 68

Zheng, M., Liu, R. and Li, Y. (2018). Long memory in financial markets: A heterogeneous agent model perspective. *International Review of Financial Analysis*, 58, 38–51. 18

Zhu, D. and Galbraith, J. (2010). A generalized asymmetric Student-t distribution with application to financial econometrics. *Journal of Econometrics*, 157, 297–305. 96, 138

Zivot, E. (2009). *Practical issues in the analysis of univariate GARCH models*. Berlin and Heidelberg: Springer, 113–55. 24, 55

Zivot, E. (2013). Univariate GARCH. `https://faculty.washington.edu/ezivot/econ589/univariateGarch2012powerpoint.pdf` (accessed on 30 August 2023). 68

Appendix A

Outcomes of Different Patterns of Seed Values for Sets S_1 and S_2

Table A.1: Outcomes of different patterns of seed values for sets S_1 and S_2, with true parameter $\alpha + \beta = 0.9990$.

Panel A (S_1)		Seed: 12345				Seed: 54321				Seed: 15243			
	N	$\hat{\alpha}+\hat{\beta}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	TPR (95%)	$\hat{\alpha}+\hat{\beta}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	TPR (95%)	$\hat{\alpha}+\hat{\beta}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	TPR (95%)
Normal	1000	0.9990	0.0862	0.0862	95.00%	0.9563	0.0757	0.0625	90.94%	0.9909	0.0801	0.0796	94.23%
	2000	0.9771	0.0934	0.0908	92.91%	0.9903	0.0419	0.0410	94.17%	0.9839	0.0490	0.0466	93.56%
	3000	0.9727	0.0756	0.0709	92.50%	0.9850	0.0373	0.0345	93.67%	0.9791	0.0591	0.0557	93.11%
	4000	0.9700	0.0693	0.0629	92.24%	0.9846	0.0412	0.0387	93.64%	0.9972	0.0304	0.0303	94.83%
Student's t	1000	0.9990	0.0525	0.0525	95.00%	0.9833	0.0441	0.0412	93.50%	0.9990	0.0768	0.0768	95.00%
	2000	0.9902	0.0500	0.0492	94.16%	0.9990	0.0349	0.0349	95.00%	0.9958	0.0388	0.0386	94.69%
	3000	0.9973	0.0327	0.0327	94.83%	0.9977	0.0308	0.0308	94.87%	0.9956	0.0318	0.0316	94.68%
	4000	0.9918	0.0277	0.0267	94.31%	0.9974	0.0258	0.0258	94.85%	0.9963	0.0247	0.0246	94.74%
GED	1000	0.9875	0.0719	0.0710	93.90%	0.9688	0.0499	0.0397	92.13%	0.9899	0.0630	0.0624	94.13%
	2000	0.9663	0.0608	0.0512	91.89%	0.9908	0.0336	0.0326	94.22%	0.9847	0.0387	0.0360	93.64%
	3000	0.9684	0.0441	0.0317	92.09%	0.9846	0.0347	0.0315	93.63%	0.9795	0.0385	0.0333	93.15%
	4000	0.9692	0.0410	0.0282	92.16%	0.9833	0.0328	0.0288	93.51%	0.9839	0.0300	0.0259	93.57%
GHYP	1000	0.9940	0.0557	0.0555	94.52%	0.9785	0.0437	0.0386	93.05%	0.9897	0.0657	0.0650	94.12%
	2000	0.9748	0.0507	0.0446	92.70%	0.9979	0.0328	0.0328	94.89%	0.9871	0.0364	0.0344	93.87%
	3000	0.9780	0.0353	0.0284	93.00%	0.9901	0.0305	0.0292	94.15%	0.9849	0.0325	0.0293	93.66%
	4000	0.9776	0.0322	0.0241	92.97%	0.9898	0.0263	0.0247	94.12%	0.9892	0.0247	0.0226	94.07%
Panel B (S_2)		**Seed: 34567**				**Seed: 76543**				**Seed: 36547**			
	N	$\hat{\alpha}+\hat{\beta}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	TPR (95%)	$\hat{\alpha}+\hat{\beta}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	TPR (95%)	$\hat{\alpha}+\hat{\beta}$	$\text{RMSE}_{\hat{\alpha}+\hat{\beta}}$	$\text{SE}_{\hat{\alpha}+\hat{\beta}}$	TPR (95%)
Normal	1000	0.9856	0.0583	0.0568	93.72%	0.9942	0.0424	0.0421	94.54%	0.9823	0.3888	0.3884	93.41%
	2000	0.9814	0.0396	0.0354	93.33%	0.9891	0.0370	0.0357	94.06%	0.9806	0.1419	0.1407	93.25%
	3000	0.9845	0.0708	0.0693	93.62%	0.9809	0.0334	0.0281	93.28%	0.9822	0.0805	0.0787	93.40%
	4000	0.9990	0.0397	0.0397	95.00%	0.9778	0.0326	0.0248	92.98%	0.9779	0.0575	0.0535	92.99%
Student's t	1000	0.9971	0.0474	0.0474	94.82%	0.9990	0.0422	0.0422	95.00%	0.9990	0.0329	0.0329	95.00%
	2000	0.9789	0.0364	0.0303	93.08%	0.9990	0.0281	0.0281	95.00%	0.9990	0.0315	0.0315	95.00%
	3000	0.9781	0.0326	0.0249	93.01%	0.9975	0.0237	0.0236	94.86%	0.9990	0.0234	0.0234	95.00%
	4000	0.9871	0.0253	0.0223	93.87%	0.9955	0.0218	0.0215	94.67%	0.9946	0.0238	0.0234	94.58%
GED	1000	0.9802	0.0463	0.0423	93.21%	0.9899	0.0389	0.0378	94.13%	0.9986	0.0490	0.0490	94.96%
	2000	0.9726	0.0386	0.0282	92.49%	0.9898	0.0280	0.0265	94.13%	0.9879	0.0388	0.0371	93.94%
	3000	0.9710	0.0398	0.0282	92.33%	0.9820	0.0276	0.0218	93.38%	0.9808	0.0303	0.0243	93.27%
	4000	0.9800	0.0285	0.0213	93.19%	0.9782	0.0284	0.0194	93.02%	0.9752	0.0321	0.0215	92.73%
GHYP	1000	0.9863	0.0436	0.0417	93.80%	0.9928	0.0383	0.0378	94.41%	0.9990	0.0370	0.0370	95.00%
	2000	0.9744	0.0377	0.0285	92.66%	0.9952	0.0265	0.0262	94.64%	0.9990	0.0358	0.0358	95.00%
	3000	0.9737	0.0351	0.0242	92.59%	0.9872	0.0243	0.0213	93.88%	0.9894	0.0256	0.0237	94.09%
	4000	0.9810	0.0278	0.0212	93.29%	0.9835	0.0246	0.0192	93.53%	0.9816	0.0275	0.0213	93.35%

Appendix B

Further Visual Illustrations of S_1 and S_2 TPR Outcomes

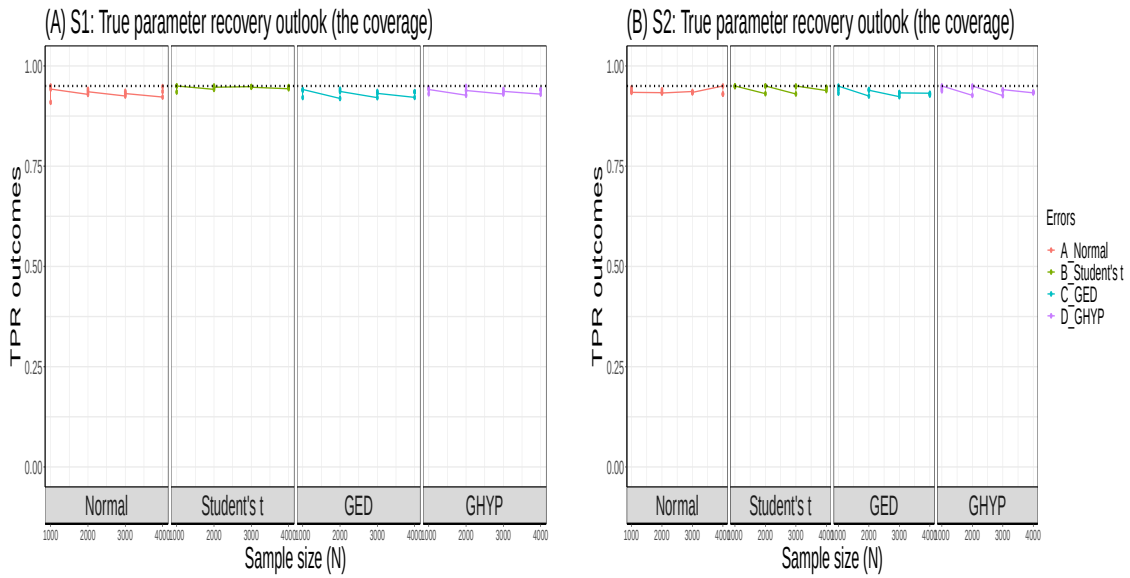

Figure B.1: The TPR outcomes of S_1 and S_2 in Panels (**A**,**B**), respectively, where the dotted lines are the 95% (i.e., 0.95) nominal recovery levels.

Appendix C

The Code for DGP through the Ugarchsim Function

The Code for the Method of Implementation of the MCS Experiment

```
library(rugarch)
attach(BondDataSA)
BondDataSA<-as.data.frame(BondDataSA)
spec = ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),
                                        submodel = "ALLGARCH"),
                  mean.model = list(armaOrder = c(1,1),
                  include.mean = TRUE),
                  distribution.model = "std",
                  fixed.pars=list(shape=4.1))
fit = ugarchfit(data = BondDataSA[,4,drop=FALSE], spec = spec)
fit
coef(fit)

# simulate for N=8000
sim = ugarchsim(fit, n.sim=15000, n.start=1, m.sim=1000,
                 rseed = 12345, startMethod= "sample")
simGARCH <- fitted(sim)
simGARCH
simGARCH <- as.data.frame(simGARCH)
simGARCH

#Remove the first 7000 for initial values effect
R_simGARCH <- simGARCH[-c(1:7000), ]
R_simGARCH
```

```
#Fit ARMA(1,1)-fGARCH(1,1) to the simulated dataset R_simGARCH
#For Normal
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),
                                         submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "norm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)
```

Appendix D

The $\sqrt{N}$ Consistency Outcomes of Datasets with Different Unconditional Shape Parameters

Table D.1: The $\sqrt{N}$ consistency outcomes of datasets with different unconditional shape parameters ν^* = {4.1, 8.2, 10.3}.

True parameter: $b_\phi = 0.9831$

Panel (A)		$\nu^* = 4.1$				$\nu^* = 8.2$				$\nu^* = 10.3$			
	N	$\hat{b}_\phi$	$\text{Bias}_{\hat{b}_\phi}$	$\text{SE}_{\hat{b}_\phi}$	$\text{RMSE}_{\hat{b}_\phi}$	$\hat{b}_\phi$	$\text{Bias}_{\hat{b}_\phi}$	$\text{SE}_{\hat{b}_\phi}$	$\text{RMSE}_{\hat{b}_\phi}$	$\hat{b}_\phi$	$\text{Bias}_{\hat{b}_\phi}$	$\text{SE}_{\hat{b}_\phi}$	$\text{RMSE}_{\hat{b}_\phi}$
	30,000	**0.9689**	**−0.0142**	**0.0026**	**0.0144**	0.9674	-0.0156	0.0026	0.0158	0.9669	-0.0162	0.0025	0.0164
Gaussian	31,000	**0.9696**	**−0.0135**	**0.0025**	**0.0137**	0.9674	-0.0157	0.0026	0.0159	0.9660	-0.0170	0.0025	0.0172
	32,000	**0.9707**	**−0.0123**	**0.0022**	**0.0125**	0.9667	-0.0164	0.0026	0.0166	0.9657	-0.0173	0.0025	0.0175
	30,000	**0.9691**	**−0.0140**	**0.0026**	**0.0142**	0.9675	-0.0156	0.0026	0.0158	0.9669	-0.0161	0.0025	0.0163
skew	31,000	**0.9698**	**−0.0133**	**0.0025**	**0.0135**	0.9842	0.0011	0.0014	0.0018	0.9661	-0.0170	0.0025	0.0171
Gaussian	32,000	**0.9709**	**−0.0122**	**0.0022**	**0.0124**	0.9667	-0.0164	0.0026	0.0166	0.9658	-0.0173	0.0025	0.0174
	30,000	**0.9794**	**−0.0037**	**0.0023**	**0.0044**	0.9790	-0.0040	0.0021	0.0046	0.9789	-0.0042	0.0021	0.0047
Student's t	31,000	**0.9792**	**−0.0039**	**0.0023**	**0.0045**	0.9792	-0.0038	0.0021	0.0044	0.9784	-0.0047	0.0021	0.0051
	32,000	**0.9805**	**−0.0025**	**0.0021**	**0.0033**	0.9785	-0.0045	0.0021	0.0050	0.9781	-0.0050	0.0021	0.0054
	30,000	**0.9792**	**−0.0039**	**0.0024**	**0.0046**	0.9788	-0.0043	0.0021	0.0048	0.9788	-0.0043	0.0021	0.0048
skew	31,000	**0.9792**	**−0.0038**	**0.0024**	**0.0045**	0.9793	-0.0038	0.0021	0.0043	0.9784	-0.0046	0.0021	0.0051
Student's t	32,000	**0.9806**	**−0.0025**	**0.0022**	**0.0033**	0.9785	-0.0045	0.0021	0.0050	0.9783	-0.0048	0.0021	0.0052

Panel (B)		$\nu^* = 4.1$				$\nu^* = 8.2$				$\nu^* = 10.3$			
	N	$\hat{b}_\phi$	$\text{Bias}_{\hat{b}_\phi}$	$\text{SE}_{\hat{b}_\phi}$	$\text{RMSE}_{\hat{b}_\phi}$	$\hat{b}_\phi$	$\text{Bias}_{\hat{b}_\phi}$	$\text{SE}_{\hat{b}_\phi}$	$\text{RMSE}_{\hat{b}_\phi}$	$\hat{b}_\phi$	$\text{Bias}_{\hat{b}_\phi}$	$\text{SE}_{\hat{b}_\phi}$	$\text{RMSE}_{\hat{b}_\phi}$
	17,000	**0.9589**	**−0.0242**	**0.0059**	**0.0249**	0.9583	-0.0248	0.0041	0.0251	0.9562	-0.0268	0.0039	0.0271
Gaussian	18,000	**0.9620**	**−0.0210**	**0.0051**	**0.0217**	0.9614	-0.0217	0.0038	0.0220	0.9569	-0.0262	0.0038	0.0265
	19,000	**0.9643**	**−0.0187**	**0.0043**	**0.0192**	0.9609	-0.0221	0.0036	0.0224	0.9599	-0.0232	0.0035	0.0234
	17,000	**0.9604**	**−0.0227**	**0.0058**	**0.0234**	0.9582	-0.0248	0.0041	0.0252	0.9564	-0.0266	0.0039	0.0269
skew	18,000	**0.9630**	**−0.0200**	**0.0050**	**0.0206**	0.9614	-0.0217	0.0038	0.0220	0.9570	-0.0260	0.0038	0.0263
Gaussian	19,000	**0.9650**	**−0.0180**	**0.0042**	**0.0185**	0.9609	-0.0221	0.0036	0.0224	0.9601	-0.0230	0.0035	0.0233
	17,000	**0.9775**	**−0.0056**	**0.0034**	**0.0066**	0.9805	-0.0025	0.0027	0.0037	0.9773	-0.0057	0.0029	0.0064
Student's t	18,000	**0.9780**	**−0.0051**	**0.0032**	**0.0061**	0.9806	-0.0025	0.0026	0.0036	0.9775	-0.0056	0.0028	0.0063
	19,000	**0.9781**	**−0.0049**	**0.0031**	**0.0058**	0.9797	-0.0033	0.0026	0.0042	0.9770	-0.0060	0.0028	0.0067
	17,000	**0.9775**	**−0.0056**	**0.0034**	**0.0065**	0.9805	-0.0025	0.0027	0.0037	0.9774	-0.0057	0.0029	0.0064
skew	18,000	**0.9779**	**−0.0051**	**0.0032**	**0.0061**	0.9806	-0.0025	0.0026	0.0036	0.9775	-0.0056	0.0028	0.0063
Student's t	19,000	**0.9782**	**−0.0048**	**0.0031**	**0.0057**	0.9797	-0.0033	0.0026	0.0042	0.9770	-0.0060	0.0028	0.0066

Note: The $\sqrt{N}$ consistency outcomes of datasets with different unconditional shape parameters or degrees of freedom ν^* = {4.1, 8.2, 10.3} for sample sizes N = {30,000; 31,000; 32,000} in Panel A, and N = {17,000; 18,000; 19,000} in Panel B.

Appendix E

R Code

E.1 The R Code for the Method of Implementation of the MCS Experiment in Paper 1

```
#################N=8000#############################################################

library(rugarch)
attach(BondDataSA)
BondDataSA<-as.data.frame(BondDataSA)
head(BondDataSA)
spec = ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                  mean.model = list(armaOrder = c(1,1), include.mean = TRUE),
                  distribution.model = "std",fixed.pars=list(shape=4.1))

fit = ugarchfit(data = BondDataSA[,4,drop=FALSE], spec = spec)
fit
coef(fit)

# simulate for N=8000

sim = ugarchsim(fit, n.sim=15000, n.start=1, m.sim=1000, rseed = 12345,
                startMethod="sample")

simGARCH <- fitted(sim)
simGARCH

simGARCH <- as.data.frame(simGARCH)
simGARCH

###Remove the first 7000 for initial values effect
```

```
R_simGARCH <- simGARCH[-c(1:7000), ]
R_simGARCH

###Fit ARMA(1,1)-fGARCH(1,1) with "norm" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "norm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "snorm" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "snorm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "std" to the simulated dataset R_simGARCH######
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "std")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "sstd" to the simulated dataset R_simGARCH######
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "sstd")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ged" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "ged")
```

```
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "sged" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "sged")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ghyp" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "ghyp")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "nig" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "nig")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ghst" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "ghst")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "jsu" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
```

```
                    distribution.model = "jsu")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

################N=9000#############################################################

library(rugarch)
attach(BondDataSA)
BondDataSA<-as.data.frame(BondDataSA)
head(BondDataSA)
spec = ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                  mean.model = list(armaOrder = c(1,1), include.mean = TRUE),
                  distribution.model = "std",fixed.pars=list(shape=4.1))

fit = ugarchfit(data = BondDataSA[,4,drop=FALSE], spec = spec)
fit
coef(fit)

# simulate for N=9000

sim = ugarchsim(fit, n.sim=15000, n.start=1, m.sim=1000, rseed = 12345,
                startMethod="sample")

simGARCH <- fitted(sim)
simGARCH

simGARCH <- as.data.frame(simGARCH)
simGARCH

###Remove the first 6000 for initial values effect

R_simGARCH <- simGARCH[-c(1:6000), ]
R_simGARCH

###Fit ARMA(1,1)-fGARCH(1,1) with "norm" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "norm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "snorm" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
```

```
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "snorm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "std" to the simulated dataset R_simGARCH######
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "std")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "sstd" to the simulated dataset R_simGARCH######
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "sstd")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ged" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "ged")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "sged" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "sged")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ghyp" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
```

```
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "ghyp")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "nig" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "nig")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ghst" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "ghst")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "jsu" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                    mean.model = list(armaOrder = c(1,1),
                                      include.mean = TRUE),
                    distribution.model = "jsu")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

################N=10000#############################################################

library(rugarch)
attach(BondDataSA)
BondDataSA<-as.data.frame(BondDataSA)
head(BondDataSA)
spec = ugarchspec(variance.model = list(model = "fGARCH",
                                        garchOrder = c(1,1),submodel = "ALLGARCH"),
                  mean.model = list(armaOrder = c(1,1), include.mean = TRUE),
                  distribution.model = "std",fixed.pars=list(shape=4.1))
```

```
fit = ugarchfit(data = BondDataSA[,4,drop=FALSE], spec = spec)
fit
coef(fit)

# simulate for N=10000

sim = ugarchsim(fit, n.sim=15000, n.start=1, m.sim=1000, rseed = 12345,
                startMethod="sample")

simGARCH <- fitted(sim)
simGARCH

simGARCH <- as.data.frame(simGARCH)
simGARCH

###Remove the first 5000 for initial values effect

R_simGARCH <- simGARCH[-c(1:5000), ]
R_simGARCH

###Fit ARMA(1,1)-fGARCH(1,1) with "norm" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "norm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "snorm" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "snorm")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "std" to the simulated dataset R_simGARCH######
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "std")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)
```

```
###Fit ARMA(1,1)-fGARCH(1,1) with "sstd" to the simulated dataset R_simGARCH######
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "sstd")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ged" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "ged")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "sged" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "sged")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ghyp" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "ghyp")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "nig" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "nig")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
```

```
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "ghst" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "ghst")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)

###Fit ARMA(1,1)-fGARCH(1,1) with "jsu" to the simulated dataset R_simGARCH#####
spec <- ugarchspec(variance.model = list(model = "fGARCH",
                                         garchOrder = c(1,1),submodel = "ALLGARCH"),
                   mean.model = list(armaOrder = c(1,1),
                                     include.mean = TRUE),
                   distribution.model = "jsu")
fit = ugarchfit(data = R_simGARCH, spec = spec)
show(fit)
coef(fit)
```

E.2 R Code for Paper 2 when the true model is GAS-Student's t

```
library(GAS)
attach(BondDataSA)
BondDataSA<- as.data.frame(BondDataSA)
head(BondDataSA)
GASSpec = UniGASSpec(Dist = "std", ScalingType = "Identity",
                     GASPar = list(location = TRUE,
                                   scale = TRUE, shape = TRUE))
fit = UniGASFit(GASSpec, BondDataSA[ ,4])
fit

coef(fit)

################FOR N = 17000#####################################################

# simulate for N = 17000

set.seed(12345)
A <- diag(c(1.013270e-06, 1.400355e-01, 9.835611e+00))
```

```
B <- diag(c(3.460703e-01, 9.830649e-01, 5.253297e-01))
ThetaStar <- c(0.05779397, 1.06953983, 4.1)
kappa <- (diag(3) - B) %*% UniUnmapParameters(ThetaStar, "std")
sim <- UniGASSim(fit = NULL, T.sim = 30000, kappa = kappa, A = A, B = B,
                 Dist = "std", ScalingType = "Identity")
sim
getObs(sim)

simGAS <- getObs(sim)
simGAS

simGAS <- as.data.frame(simGAS)
simGAS

###Remove the first 13000 for initial values effect

R_simGAS <- simGAS[-c(1:13000), ]
R_simGAS


###Fit GAS model with each of the error distributions to the simulated dataset R_simGAS


#For Gaussian
GASSpec_norm <- UniGASSpec(Dist = "norm", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_norm, data = R_simGAS)
Fit

#For skew-Gaussian
GASSpec_snorm <- UniGASSpec(Dist = "snorm", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_snorm, data = R_simGAS)
Fit

#For Student-t
GASSpec_std <- UniGASSpec(Dist = "std", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_std, data = R_simGAS)
Fit

#For skew-Student-t
GASSpec_sstd <- UniGASSpec(Dist = "sstd", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_sstd, data = R_simGAS)
Fit

#For Asymmetric Student-t with two tail decay parameters
GASSpec_ast <- UniGASSpec(Dist = "ast", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ast, data = R_simGAS)
Fit

#For Asymmetric Student-t with one tail decay parameter
```

```
GASSpec_ast1 <- UniGASSpec(Dist = "ast1", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ast1, data = R_simGAS)
Fit

#For Asymmetric Laplace Distribution
GASSpec_ald <- UniGASSpec(Dist = "ald", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ald, data = R_simGAS)
Fit

################FOR N = 18000################################################

# simulate for N = 18000

set.seed(12345)
A <- diag(c(1.013270e-06, 1.400355e-01, 9.835611e+00))
B <- diag(c(3.460703e-01, 9.830649e-01, 5.253297e-01))
ThetaStar <- c(0.05779397, 1.06953983, 4.1)
kappa <- (diag(3) - B) %*% UniUnmapParameters(ThetaStar, "std")
sim <- UniGASSim(fit = NULL, T.sim = 30000, kappa = kappa, A = A, B = B,
                 Dist = "std", ScalingType = "Identity")
sim
getObs(sim)

simGAS <- getObs(sim)
simGAS

simGAS <- as.data.frame(simGAS)
simGAS

###Remove the first 12000 for initial values effect

R_simGAS <- simGAS[-c(1:12000), ]
R_simGAS

###Fit GAS model with each of the error distributions to the simulated dataset R_simGAS

#For Gaussian
GASSpec_norm <- UniGASSpec(Dist = "norm", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_norm, data = R_simGAS)
Fit

#For skew-Gaussian
GASSpec_snorm <- UniGASSpec(Dist = "snorm", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_snorm, data = R_simGAS)
Fit

#For Student-t
GASSpec_std <- UniGASSpec(Dist = "std", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_std, data = R_simGAS)
```

```
Fit

#For skew-Student-t
GASSpec_sstd <- UniGASSpec(Dist = "sstd", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_sstd, data = R_simGAS)
Fit

#For Asymmetric Student-t with two tail decay parameters
GASSpec_ast <- UniGASSpec(Dist = "ast", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ast, data = R_simGAS)
Fit

#For Asymmetric Student-t with one tail decay parameter
GASSpec_ast1 <- UniGASSpec(Dist = "ast1", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ast1, data = R_simGAS)
Fit

#For Asymmetric Laplace Distribution
GASSpec_ald <- UniGASSpec(Dist = "ald", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ald, data = R_simGAS)
Fit

################FOR N = 19000#################################################

# simulate for N = 19000

set.seed(12345)
A <- diag(c(1.013270e-06, 1.400355e-01, 9.835611e+00))
B <- diag(c(3.460703e-01, 9.830649e-01, 5.253297e-01))
ThetaStar <- c(0.05779397, 1.06953983, 4.1)
kappa <- (diag(3) - B) %*% UniUnmapParameters(ThetaStar, "std")
sim <- UniGASSim(fit = NULL, T.sim = 30000, kappa = kappa, A = A, B = B,
                 Dist = "std", ScalingType = "Identity")
sim
getObs(sim)

simGAS <- getObs(sim)
simGAS

simGAS <- as.data.frame(simGAS)
simGAS

###Remove the first 11000 for initial values effect

R_simGAS <- simGAS[-c(1:11000), ]
R_simGAS

#####Fit GAS model with each of the error distributions to the simulated dataset R_
    simGAS
```

```
#For Gaussian
GASSpec_norm <- UniGASSpec(Dist = "norm", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_norm, data = R_simGAS)
Fit

#For skew-Gaussian
GASSpec_snorm <- UniGASSpec(Dist = "snorm", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_snorm, data = R_simGAS)
Fit

#For Student-t
GASSpec_std <- UniGASSpec(Dist = "std", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_std, data = R_simGAS)
Fit

#For skew-Student-t
GASSpec_sstd <- UniGASSpec(Dist = "sstd", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_sstd, data = R_simGAS)
Fit

#For Asymmetric Student-t with two tail decay parameters
GASSpec_ast <- UniGASSpec(Dist = "ast", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ast, data = R_simGAS)
Fit

#For Asymmetric Student-t with one tail decay parameter
GASSpec_ast1 <- UniGASSpec(Dist = "ast1", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ast1, data = R_simGAS)
Fit

#For Asymmetric Laplace Distribution
GASSpec_ald <- UniGASSpec(Dist = "ald", GASPar = list(scale = TRUE))
Fit = UniGASFit(GASSpec_ald, data = R_simGAS)
Fit
```

E.3 R Code for Paper 2 when the true model is GAS-AST1

```
library(GAS)
attach(BondDataSA)
BondDataSA <- as.data.frame(BondDataSA)
head(BondDataSA)
GASSpec = UniGASSpec(Dist = "ast1", ScalingType = "Identity",
                     GASPar = list(location = TRUE, skewness = TRUE,
                                   scale = TRUE, shape = TRUE))
```

```
fit = UniGASFit(GASSpec, BondDataSA[ ,4])
fit

coef(fit)

################FOR N = 17000#####################################################

# simulate for N = 17000

set.seed(12345)
A <- diag(c(1.007662e-06, 4.204045e-02, 1.185833e-01, 1.400373e-04))
B <- diag(c(8.802454e-01, 9.788202e-01, 2.053495e-05, 9.702012e-01))
ThetaStar <- c(0.1547653, 2.6391596, 0.5328714, 4.1)
kappa <- (diag(4) - B) %*% UniUnmapParameters(ThetaStar, "ast1")
sim <- UniGASSim(fit = NULL, T.sim = 30000, kappa = kappa, A = A, B = B,
                 Dist = "ast1", ScalingType = "Identity")
sim
getObs(sim)

simGAS <- getObs(sim)
simGAS

simGAS <- as.data.frame(simGAS)
simGAS

###Remove the first 13000 rows for initial values effect

R_simGAS <- simGAS[-c(1:13000), ]
R_simGAS

###Fit GAS model with each of the error distributions to the simulated dataset R_simGAS

#For Gaussian
GASSpec_norm <- UniGASSpec(Dist = "norm", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_norm, data = R_simGAS)
fit

#For skew-Gaussian
GASSpec_snorm <- UniGASSpec(Dist = "snorm", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_snorm, data = R_simGAS)
fit

#For Student-t
GASSpec_std <- UniGASSpec(Dist = "std", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_std, data = R_simGAS)
fit

#For skew-Student-t
GASSpec_sstd <- UniGASSpec(Dist = "sstd", GASPar = list(scale = TRUE))
```

```
fit = UniGASFit(GASSpec_sstd, data = R_simGAS)
fit

#For Asymmetric Student-t with two tail decay parameters
GASSpec_ast <- UniGASSpec(Dist = "ast", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ast, data = R_simGAS)
fit

#For Asymmetric Student-t with one tail decay parameter
GASSpec_ast1 <- UniGASSpec(Dist = "ast1", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ast1, data = R_simGAS)
fit

#For Asymmetric Laplace Distribution
GASSpec_ald <- UniGASSpec(Dist = "ald", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ald, data = R_simGAS)
fit

################FOR N = 18000################################################

# simulate for N = 18000

set.seed(12345)
A <- diag(c(1.007662e-06, 4.204045e-02, 1.185833e-01, 1.400373e-04))
B <- diag(c(8.802454e-01, 9.788202e-01, 2.053495e-05, 9.702012e-01))
ThetaStar <- c(0.1547653, 2.6391596, 0.5328714, 4.1)
kappa <- (diag(4) - B) %*% UniUnmapParameters(ThetaStar, "ast1")
sim <- UniGASSim(fit = NULL, T.sim = 30000, kappa = kappa, A = A, B = B,
                 Dist = "ast1", ScalingType = "Identity")
sim
getObs(sim)

simGAS <- getObs(sim)
simGAS

simGAS <- as.data.frame(simGAS)
simGAS

###Remove the first 12000 for initial values effect

R_simGAS <- simGAS[-c(1:12000), ]
R_simGAS

###Fit GAS model with each of the error distributions to the simulated dataset R_simGAS

#For Gaussian
GASSpec_norm <- UniGASSpec(Dist = "norm", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_norm, data = R_simGAS)
fit
```

```
#For skew-Gaussian
GASSpec_snorm <- UniGASSpec(Dist = "snorm", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_snorm, data = R_simGAS)
fit

#For Student-t
GASSpec_std <- UniGASSpec(Dist = "std", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_std, data = R_simGAS)
fit

#For skew-Student-t
GASSpec_sstd <- UniGASSpec(Dist = "sstd", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_sstd, data = R_simGAS)
fit

#For Asymmetric Student-t with two tail decay parameters
GASSpec_ast <- UniGASSpec(Dist = "ast", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ast, data = R_simGAS)
fit

#For Asymmetric Student-t with one tail decay parameter
GASSpec_ast1 <- UniGASSpec(Dist = "ast1", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ast1, data = R_simGAS)
fit

#For Asymmetric Laplace Distribution
GASSpec_ald <- UniGASSpec(Dist = "ald", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ald, data = R_simGAS)
fit

################FOR N = 19000#################################################

# simulate for N = 19000

set.seed(12345)
A <- diag(c(1.007662e-06, 4.204045e-02, 1.185833e-01, 1.400373e-04))
B <- diag(c(8.802454e-01, 9.788202e-01, 2.053495e-05, 9.702012e-01))
ThetaStar <- c(0.1547653, 2.6391596, 0.5328714, 4.1)
kappa <- (diag(4) - B) %*% UniUnmapParameters(ThetaStar, "ast1")
sim <- UniGASSim(fit = NULL, T.sim = 30000, kappa = kappa, A = A, B = B,
                 Dist = "ast1", ScalingType = "Identity")
sim
getObs(sim)

simGAS <- getObs(sim)
simGAS

simGAS <- as.data.frame(simGAS)
```

```
simGAS

###Remove the first 11000 rows for initial values effect

R_simGAS <- simGAS[-c(1:11000), ]
R_simGAS

###Fit GAS model with each of the error distributions to the simulated dataset R_simGAS


#For Gaussian
GASSpec_norm <- UniGASSpec(Dist = "norm", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_norm, data = R_simGAS)
fit

#For skew-Gaussian
GASSpec_snorm <- UniGASSpec(Dist = "snorm", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_snorm, data = R_simGAS)
fit

#For Student-t
GASSpec_std <- UniGASSpec(Dist = "std", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_std, data = R_simGAS)
fit

#For skew-Student-t
GASSpec_sstd <- UniGASSpec(Dist = "sstd", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_sstd, data = R_simGAS)
fit

#For Asymmetric Student-t with two tail decay parameters
GASSpec_ast <- UniGASSpec(Dist = "ast", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ast, data = R_simGAS)
fit

#For Asymmetric Student-t with one tail decay parameter
GASSpec_ast1 <- UniGASSpec(Dist = "ast1", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ast1, data = R_simGAS)
fit

#For Asymmetric Laplace Distribution
GASSpec_ald <- UniGASSpec(Dist = "ald", GASPar = list(scale = TRUE))
fit = UniGASFit(GASSpec_ald, data = R_simGAS)
fit
```