分类号 O212.1　　　　学校代码 14325

U D C 519.2　　　　密　　级 公开

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

本科生毕业设计（论文）

题　　目： **基于 ARMA-GARCH 模型的中国股票指数实证分析**

姓　　名： **黄秋逸**

学　　号： **11810708**

系　　别： **统计与数据科学系**

专　　业： **统计学**

指导教师： **蒋学军 副教授**

2022 年 5 月 26 日

CLC O212.1 School code 14325

UDC 519.2 Available for reference ☐Yes ☐No

Undergraduate Thesis

Thesis Title: **An empirical analysis of the Chinese stock market indexes by ARMA-GARCH models**

Student Name: **Qiuyi HAUNG**

Student ID: **11810708**

Department: **Department of Statistics and Data Science**

Program: **Statistics**

Thesis Advisor: **Associate Professor Xuejun JIANG**

Date: May 26, 2022

诚信承诺书

1. 本人郑重承诺所呈交的毕业设计（论文），是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料均真实可靠。

2. 除文中已经注明引用的内容外，本论文不包含任何其他人或集体已经发表或撰写过的作品或成果。对本论文的研究作出重要贡献的个人和集体，均已在文中以明确的方式标明。

3. 本人承诺在毕业论文（设计）选题和研究内容过程中没有抄袭他人研究成果和伪造相关数据等行为。

4. 在毕业论文（设计）中对侵犯任何方面知识产权的行为，由本人承担相应的法律责任。

作者签名:

________年____月____日

COMMITMENT OF HONESTY

1. I solemnly promise that the paper presented comes from my independent research work under my supervisor's supervision. All statistics and images are real and reliable.
2. Except for the annotated reference, the paper contents no other published work or achievement by person or group. All people making important contributions to the study of the paper have been indicated clearly in the paper.
3. I promise that I did not plagiarize other people's research achievement or forge related data in the process of designing topic and research content.
4. If there is violation of any intellectual property right, I will take legal responsibility myself.

Signature:

Date:

基于 ARMA-GARCH 模型的中国股票指数实证分析

黄秋逸

（统计与数据科学系　指导教师：蒋学军）

[摘要]： ARMA-GARCH 模型刻画了时间序列的条件均值和条件方差，且具有良好的可解释性和可扩展性，因而被广泛应用于计量经济学、天文学、水文学、交通运输等领域。长期以来，利用 ARMA-GARCH 模型对股指波动率进行建模一直是计量经济学的热点话题。ARMA-GARCH 模型比较流行的扩展主要基于对 GARCH 模型的两种修改: 使用其他 GARCH 族模型代替 GARCH 模型，如 IGARCH 模型和 TGARCH 模型；使用重尾误差分布代替正态误差分布，如学生氏分布和广义误差分布。然而，上述 ARMA-GARCH 模型的拓展在中国股市的有效性和准确性还需要更多的实证研究加以证明。此外，ARMA-GARCH 模型的另一个应用上的限制是其不合理的模型建立过程。

针对上述问题，本研究设计了五种方案，对模型建立过程、GARCH 族模型和误差分布进行了比较。详细地说，第一种模型建立方法由 Tsay(2005)[1] 提出，首先分别确定 ARMA 模型和 GARCH 模型的阶数，然后联合估计参数，称为“Tsay 方法”。而第二种方法直接对搜索范围内的所有模型进行估计，并选取 AIC 值最小的模型作为最佳模型，称为“全可能回归方法”。与 Tsay 方法相比，全可能回归方法计算量大，但可以对不同 GARCH 模型和误差分布进行比较，并且保证了最终模型在 AIC 下的最优性。基于这些良好的性质，本研究应用全可能回归方法比较了 IGARCH 和 TGARCH 模型在中国股票市场上的有效性，然后讨论了使用学生氏分布、广义误差分布及它们的偏分布的必要性。

对于中国股票市场，实证分析的主要结论如下。(1) 全可能回归方法在 ARMA-GARCH 模型的阶数识别方面显示出更好的性能。(2) 建议采用

IGARCH (1,1) 模型作为条件方差模型。(3) 在非偏分布和偏分布假设下，广义误差分布和偏广义误差分布分别适合作为大数据集 (本例为 n>1000) 的误差分布。(4) GARCH (1,1) 模型和正态分布是更适合用于小数据集 (n<1000)，以避免过拟合。(5) 为降低计算复杂度，全可能回归方法的搜索范围的选择可以遵循上述规则。

[关键词]：ARMA；GARCH；中国股票市场；波动率；建模方法

[**ABSTRACT**]: The ARMA-GARCH models sketch the conditional mean and conditional variance of time series and have good interpretability and extensibility, so they are well-known and widely used in econometrics, astronomy, hydrology, transportation and other areas. Modeling the volatility of stock indexes by ARMA-GARCH models has become a hot topic in econometrics for a long time. The more popular extensions of ARMA-GARCH models are mainly based on two modifications of the GARCH model: replacing the GARCH model with other GARCH family models like IGARCH model and TGARCH model and replacing Normal distribution with heavy-tailed innovation distributions such as Student's t distribution and Generalized error distribution. However, for the Chinese stock market, more empirical evidence on the validity and accuracy of the above ARMA-GARCH model extensions needs to be collected. Moreover, another fatal limitation of the application of ARMA-GARCH models is the unreasonable model establishment process.

To address the above issues, five schemes are arranged to compare the model establishment process, GARCH family models and innovation distributions. In detail, the first method proposed by Tsay (2005)[1] determines the order of the ARMA-GARCH model sequentially then estimates the parameters jointly, referred to as the Tsay's method. The second method directly estimates all the models within the search scope and selects the best model with the lowest AIC, named the All possible regression method. Compared to the Tsay's method, the All possible regression method is computational intensive but feasible for the comparison of different GARCH models and innovation distributions and its optimality is guaranteed under AIC. Based on these nice properties, this study applies the All possible regression method to evaluate the validity of the IGARCH and TGARCH model, then discuss the necessity of the Student's t distribution and Generalized error distribution. The skewed innovation distributions are also discussed if the skewed assumption is valid.

For the Chinese stock market, the major findings from the empirical analyses are the following. (1) The All possible regression method achieves a better performance in identifying orders of the ARMA-GARCH model. (2) The IGARCH (1,1) model is recommended to be the conditional variance model. (3) The Generalized error distribution and skewed Generalized error distribution are suitable for large datasets (n>1000 in our case) under non-skewed and skewed assumption respectively. (4) The GARCH (1,1) model and Normal distribution is preferred for small datasets (n<1000 in our case) to avoid overfitting. (5) To reduce computational complexity, the choice of the searching scope of the All possible regression method can follow the rules above.

[**Key words**]: ARMA; GARCH; Chinese stock market; Volatility; Model establishment method

目录

1. Introduction

1.1 Research background and literature review

In modern capital markets, financial instruments are growing rapidly, and widely distributed in various areas including stocks, options, features and so on. As the key measure of financial instruments, the volatility plays an important role in the theory and practice of asset pricing, asset allocation and risk management. For example, volatility models can estimate the VaR in risk management and determine the asset allocation in the Markowitz model. Moreover, the Volatility Index (VIX), constructed by the Chicago Board Options Exchange (CBOE), has been traded since 2004. Therefore, the study of fluctuations of financial instrument yields has been an enduring subject of econometrics.

As a representative of financial instruments, the stocks has been concerned by researchers all over the world for a long time. China's current stock market was launched in 1986. In 1991, the Shanghai Stock Exchange and the Shenzhen Stock Exchange released the Shanghai Stock Exchange Composite Index (SSEC) and the Shenzhen Component Index (SZI) respectively. SSEC is a stock market index of all stocks that are traded on the Shanghai Stock Exchange. SZI is an index of 500 stocks that are traded on the Shenzhen Stock Exchange. These two weighted indexes are officially compiled by the exchanges, which sketch the volatility of the Chinese stock market to the maximum extent. Thus, the SSEC and the SZI are the best targets to study the Chinese stock market.

Time is not short since scientists use statistical tools to study time series data. For homoscedastic time series, Box and Jenkins (1970)[2] proposed the ARIMA model, which extended the ARMA model to time series with unit roots. However, Fama (1965)[3] found that financial yields are volatility clustering, known as conditional heteroscedasticity and the Markowitz portfolio theory, in its original design, is based on conditional variance. Therefore, the practice deficiency calls for the prediction of the conditional variance.

A series of conditional heteroscedasticity models emerged, which can be divided into two categories: the first category models the conditional variance by a random equation, like stochastic volatility models[1]. The second category determines the conditional variance by a function, such as ARCH models. This thesis mainly refers to the second category. Granger and Andersen (1978)[4] put forward a bilinear model whose conditional variance is proportional to the square of the past yields. This model is conditional heteroscedastic but the unconditional variance is either zero or infinite. To tackle this problem, Engle (1982)[5] proposed the autoregression conditional heteroscedasticity (ARCH) model based on the bilinear model and applied it to analyse the British inflation rate. The conditional variance of the ARCH model is a linear combination of the squared historical yields, thus the unconditional variance becomes a constant. Since the ARCH model only introduces a limited amount of past observations, it is only capable for short term prediction[6] and requires a large number of parameters. The generalized autoregression conditional heteroscedasticity (GARCH) model proposed by Bollerslev (1986)[7] further considers historical conditional variance. The GARCH models require fewer parameters since it is equivalent to an ARCH (∞) model with decrease weight parameters. Further improvements of GARCH models include three directions.

First, the sum of parameters is constrained to be less than one in GARCH models. But Engle and Bollerslev (1986)[8] found that the sum of coefficients is close to one in the research of exchange rate between the U.S. dollar and the Swiss franc, which means the unconditional variance tends to infinity. To better portray the long memory volatility of exchange rate series, they restrict the sum of the parameters to one in the GARCH model, known as the integrated GARCH model today.

Second, the GARCH model responds equally to positive and negative fluctuations, while Black (1976)[9] found that the financial market tends to react differently to good and bad news. Zakoian (1990)[10] proposed the threshold GARCH (TGARCH) model to measure the asymmetric reactions of the market by means of indicative variables. More references in asymmetric GARCH model can be found in, e.g., Nelson (1991)[11], Glosten, Jagannathan and Runkel (1993)[12], and so on.

Third, the tail probability of the GARCH model is too small to tolerate outliers[1] in practice. Although the likelihood function can be complicated, a series of thick tail distributions are successively introduced into the GARCH model, which include but are not limited to Student's t distribution by Bollerslev (1987)[13], Generalized error distribution by Nelson (1991)[11], Skewed student's t distribution by Hansen (1994)[14], etc.

It should be noted that the expectations of the actual return series are not constant, which is different from the assumption of the GARCH model. To deal with this problem, Engle, Lilien and Robins (1987)[15] proposed the GARCH-mean (GARCH-M) model. Additionally, an ARMA model can be introduced to capture the conditional mean first, then we apply the GARCH model for conditional variance left in residuals, which is the ARMA-GARCH model used in this paper.

Whereas, researchers have not reached an agreement on the model establishment method of the ARMA-GARCH model. Some literatures directly adopt the GARCH (1,1) model for its simplicity, sound performance and avoiding the tedious model comparison process, see Ying SHI[16]. Some other literatures search the best model within a certain range based on certain criteria like Akaike information criteria (AIC), see, e.g. Yuling WANG et al. (2021)[17] and Huidan GU (2017)[18]. Besides, Tsay (2005)[1] proposed a model establishment method, which will be introduced in the Section 3.3.

1.2 Research objective and framework

To assess the credibility of model establishment methods and the performance of ARMA-GARCH model extensions, this paper incorporates theoretical research and empirical analysis to analyze the return rate of Shanghai Composite Index and Shenzhen Component Index based on ARMA-GARCH models. Five modeling schemes are arranged for each dataset and analyses on residuals and predictions are also presented.

The rest of this paper is organized as follows. In Section 2, we give an outline of financial time series and GARCH family models. In Section 3, we perform a positive analysis to show the excellent performance of the All possible regression method. Based on this method, the best models of different modeling schemes are reported and the residual and predictive analyses are carried out. In Section 4, we discuss problems we met during the research and draw the final conclusion.

2. Prerequsite knowledge

2.1 GARCH family model

2.1.1 GARCH model

Bollerslev (1986)[7] introduced the past conditional variance into the ARCH model, referred to as the GARCH model

$$\begin{gathered} a_t = \sigma_{t|t-1}\varepsilon_t, \\ \sigma^2_{t|t-1} = \omega + \textstyle\sum_{i=1}^{q} \alpha_i a^2_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j|t-j-1}, \end{gathered}$$

where $\{a_t\}$ is a zero mean time series, $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i > 0$, $j > 0$. $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean zero and variance one. Note that the order of the GARCH part is p and the order of the ARCH part is q.

Here are some important properties of GARCH model:

1. GARCH (p, q) for a_t implies ARMA (max(p, q), p) for a_t^2.
2. Unconditional variance of a_t is $\omega/(1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j)$.
3. Under the assumption of finite variance, $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$.

For further references in proofs and explanations, see Tsay (2005)[1].

2.1.2 IGARCH model

From the second property of the GARCH model, if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j \to 1$, then the unconditional variance $\to \infty$. Engle and Bollerslev (1986)[8] proposed the integrated GARCH model which constrain $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j = 1$. Under this restriction, the GARCH (1,1) model can be rewritten as

$$\begin{gathered} a_t = \sigma_{t|t-1}\varepsilon_t, \\ \sigma^2_{t|t-1} = \omega + \beta\sigma^2_{t-1|t-2} + (1-\beta)a^2_{t-1}, \end{gathered}$$

where $\{a_t\}$ is a zero mean time series, $\omega > 0$, $\beta \geq 0$. The unconditional variance of a_t is undefined under the IGARCH model.

2.1.3 TGARCH model

To capture the leverage effects, a series of asymmetric models have been developed. The threshold GARCH model, as a representative, introduces a leverage parameter into the

GARCH model

$$
\begin{gathered}
a_t = \sigma_{t|t-1}\varepsilon_t, \\
\sigma^2_{t|t-1} = \omega + \textstyle\sum_{i=1}^{q} \left(\alpha_i + \gamma_i I_{t-i}\right) a^2_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j|t-j-1}, \\
I_{t-i} = \begin{cases} 1 & \text{if } a_{t-i} < 0 \\ 0 & \text{if } a_{t-i} \geq 0 \end{cases},
\end{gathered}
$$

where $\{a_t\}$ is a zero mean time series, $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i > 0, j > 0$.

From the model, a positive a_{t-i} contributes $\alpha_i a^2_{t-i}$ to $\sigma^2_{t|t-1}$, whereas a negative a_{t-i} has a larger impact $(\alpha_i + \gamma_i)a^2_{t-i}$ when $\gamma_i > 0$. Through the leverage parameter, we can capture the asymmetric effect of good and bad news in the market.

2.2 Commonly used innovation distribution

The leptokurtosis and negative biased features of the return rate series suggest the Normal assumption might not be adequate for the GARCH model. This thesis further adopts the Student's t distribution, Generalized error distribution and their skewed versions as innovation distributions.

The density of random variable $Z \sim GED(\kappa)$ with $\kappa \in (0, \infty)$ is

$$
f(z) = \frac{\kappa e^{-0.5|z|^{\kappa}}}{2^{1+\kappa^{-1}} \beta \Gamma\left(\kappa^{-1}\right)},
$$

where κ is the shape parameter. When $\kappa = 2$, Z follows a standard normal distribution. For $\kappa < 2$, the tail of Z is thicker than the standard normal distribution.

The skewed distribution transformation chosen in this study was proposed by Fernandez and Steel (1998)[19]. For a univariate pdf $f(x)$ following $f(x) = f(|x|)$, the skewed distribution family of $f(x)$ is given by

$$
f(x|\gamma) = \frac{2}{\gamma + \frac{1}{\gamma}} \left((f(\frac{x}{\gamma}) I_{[0,\infty)}(x) + f(\gamma x) I_{(-\infty,0)}(x) \right),
$$

where $\gamma \in (0, \infty)$. Clearly, when $\gamma = 1$, the distribution of x is not skewed. For $\gamma > 1$, the distribution is positive skewed. For $\gamma < 1$, the distribution is negative skewed.

2.3 Characteristics of return rate time series

As a special time series, financial time series contains the following three characteristics.

First and foremost, the series $\{r_t\}$ is either serial uncorrelated or only with low serial correlation, but the squared series $\{r_t^2\}$ is high serial correlated. Thus the ARMA-GARCH model, possessing the property of capturing the serial correlation of $\{r_t\}$ and $\{r_t^2\}$ at the same time, is suitable for fitting stock return series.

Secondly, the volatility is clustering and asymmetric. Volatility is clustering since large fluctuations often follow with large fluctuations and small fluctuations often follow with small fluctuations in the stock market. Most investors are risk aversion, they respond heavier to a negative fluctuation than a positive one, which leads to the asymmetric volatility. To model the asymmetric volatility, the TGARCH model or skewed innovation distributions are required.

Thirdly, compared to the Normal distribution, the density of the volatility is leptokurtosis and fat-tailed. Therefore, distributions other than normal should be used to model return rates.

3. Empirical study

In the following empirical study, we apply the ARMA-GARCH models to the daily closed price series of two major Chinese stock market indexes: Shanghai Stock Exchange Composite Index (SSEC) and Shenzhen Component Index (SZI).

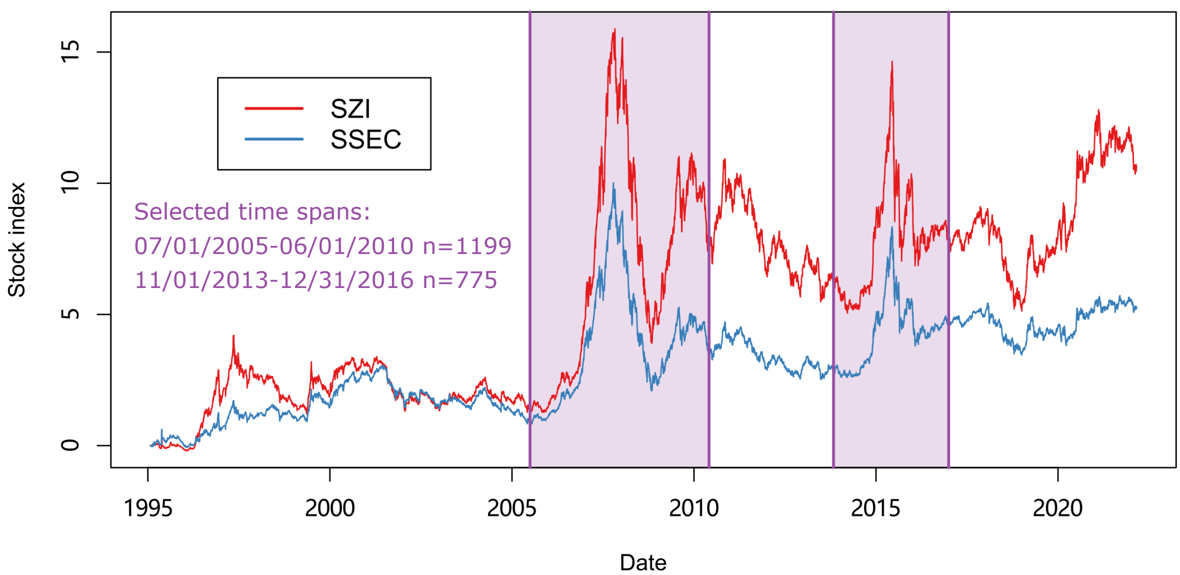
To save space, in this paper we use the following abbreviations:

SSEC	is the Shanghai Stock Exchange Composite Index,
SZI	is the Shenzhen Component Index,
N	is the Normal distribution,
T	is the Student's t distribution,
GED	is the Generalized error distribution,
SN	is the skewed Normal distribution,
ST	is the skewed Student's t distribution,
SGED	is the skewed Generalized error distribution.

3.1 Data acquisition and processing

In order to capture the complete dynamic of the stock index, time spans containing the bull, bear and consolidation period are chosen. The bull and bear periods are referred to the continuous time span in which the stock index are rising and falling respectively. The consolidation period is the continuous time span in which the stock index does not maintain a rising or falling pattern. The selected time spans are illustrated in Figure 1, in which the Y-axis is transformed to exhibit both of the indexes. The non-trading days are omitted in the datasets.

Figure 1 Closing price series of SSEC and SZI and selected time spans



In this thesis, the analyses of return rate are based on the continuous compound interest, which is given by

$$r_t = \ln P_t - \ln P_{t-1},$$

3.2 Basic statistics and tests

Unless particularly stated, the analyses below applies to the return rate series.

3.2.1 Summary of datasets

First and foremost, the basic statistics of four datasets are provided in Table 1. From the table, we can make the following points: (1) The sample size of the second time span is less than two-thirds that of the first time span. (2) In both time spans, the indexes are left-skewed and leptokurtic.

Table 1 Basic statistics of return rates

Statistics	07/01/2005-06/01/2010		11/01/2013-12/31/2016	
	SSEC	SZI	SSEC	SZI
n	1199	1199	775	775
Minimum rate	-0.092562	-0.097500	-0.088729	-0.086036
Maximum rate	0.090343	0.091615	0.056036	0.062542
SD	0.020425	0.022452	0.017404	0.019781
Skewness	-0.434222	-0.432816	-1.207604	-0.966616
Kurtosis	2.336674	1.757114	5.448543	3.513512

From Figure A.1, the two indexes share the similar trend in the same time span and the bull, bear and consolidation period are included in each dataset, which meets the need of capturing the complete dynamic of the stock index. Both the return rate and squared return rate illustrate remarkable volatility clustering phenomenon, that is, the volatility is moderate during the consolidation and becomes severe during the bull and bear. Thus it is reasonable for us to apply GARCH models to capture the ARCH effect.

In Figure A.2, the red lines are the kernel density estimates of the histogram, which can be treated as the non-standardized empirical density function. The blue lines are the Normal distributions with means and variances calculated from the return rates. Compared to the blue line, the peak is stronger and the tail is wider for the red line.

From Figure A.3, scatter points higher than qq-lines on the left-hand side and lower on the right-hand side, illustrating a fat-tailed pattern. In addition, the distribution is left skewed visually, since the deviation is much more severe on the left-hand side.

Generally speaking, both Figure A.2 and Figure A.3 show that the distribution of the return rate are leptokurtosis, heavy-tailed and left-skewed compared to the Normal distributions. Thus the heavy-tailed innovation distribution assumption is plausible.

3.2.2 Test of normality

Apart from visual analyses, several statistical tests are also performed on the return rate. From Table 2, refer to all three normality tests, we can reject the null hypothesis that return rates follow Normal distribution under 99% of confidence level among four datasets.

Table 2 P-values of normality, unit root and ARCH effect test of return rates

Statistics tests	07/01/2005-06/01/2010		11/01/2013-12/31/2016	
	SSEC	SZI	SSEC	SZI
Kolmogorov-Smirnov test	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
Jarque - Bera Normalality Test	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
Shapiro - Wilk Normality Test	<2.2e-16	2.1e-14	<2.2e-16	<2.2e-16
Augmented Dickey-Fuller Test	<0.01	<0.01	<0.01	<0.01
Engle's Lagrange Multiplier Test	3.1e-5	1.3e-4	2.6e-10	2.0e-10

3.2.3 Test of unit root

Before establishing the ARMA-GARCH model, we employ the augmented Dickey-Fuller test to confirm if there exists unit roots, since the ARMA model is only applicable for time series without unit roots. From Table 2, we can reject the null hypothesis that unit roots exist in the return rate series under 99% of confidence level.

3.2.4 Test of serial correlation

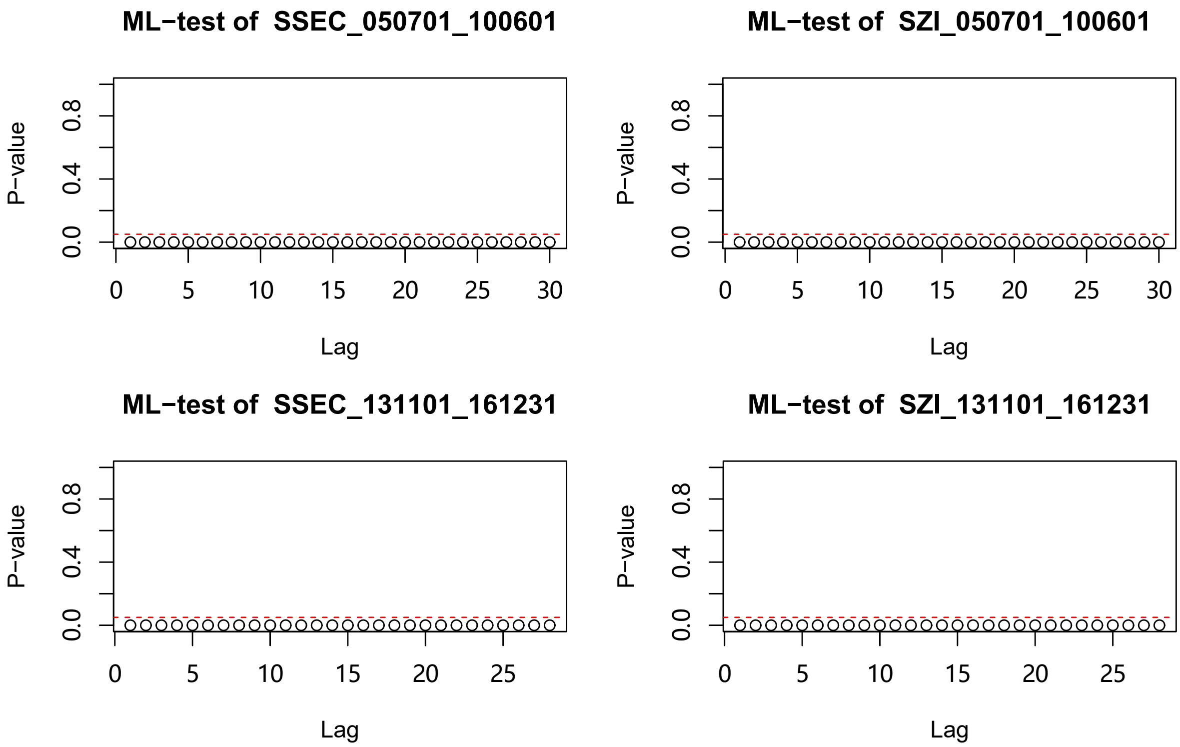
According to Figure A.4, the ACF and PACF of four datasets are significant in some lags. This characteristic suggests an ARMA model is needed. Moreover, the ACF and the PACF of squared series are remarkably significant in most lags, showing evidence of the ARCH effect.

3.2.5 Test of ARCH effect

In order to test the ARCH effect of the return rate, we apply the Engle's Lagrange Multiplier Test and the McLeod-Li test. From Table 2, according to the Engle's Lagrange Multiplier test, the ARCH effect is significant under 99% of confidence level. In Figure 2, the ARCH effect is significant among all lags under 95% of confidence level. Thus an ARCH model with infinite order, which is equivalent to a GARCH model, is needed.

From the time series plot, normality plot and qq-plot, the series are volatility clustering, left-skewed and leptokurtic. Since the ADF test suggests the series are stationary, the slight correlation of the series implies an ARMA model should be established as the conditional mean model. The squared series are highly correlated, confirmed a significant ARCH effect,

Figure 2 McLeod-Li test



which is also spotted by the McLeod-Li test and the Engle's Lagrange Multiplier Test. Thus the GARCH family model is appropriate to be the conditional variance model. The time series are skewed and fail to pass normal tests, so the heavy-tailed innovation distributions are needed. To sum up, ARMA-GARCH family models are adequate for modeling the volatility of the return rate referred to in this thesis and the heavy-tailed innovation distributions are required.

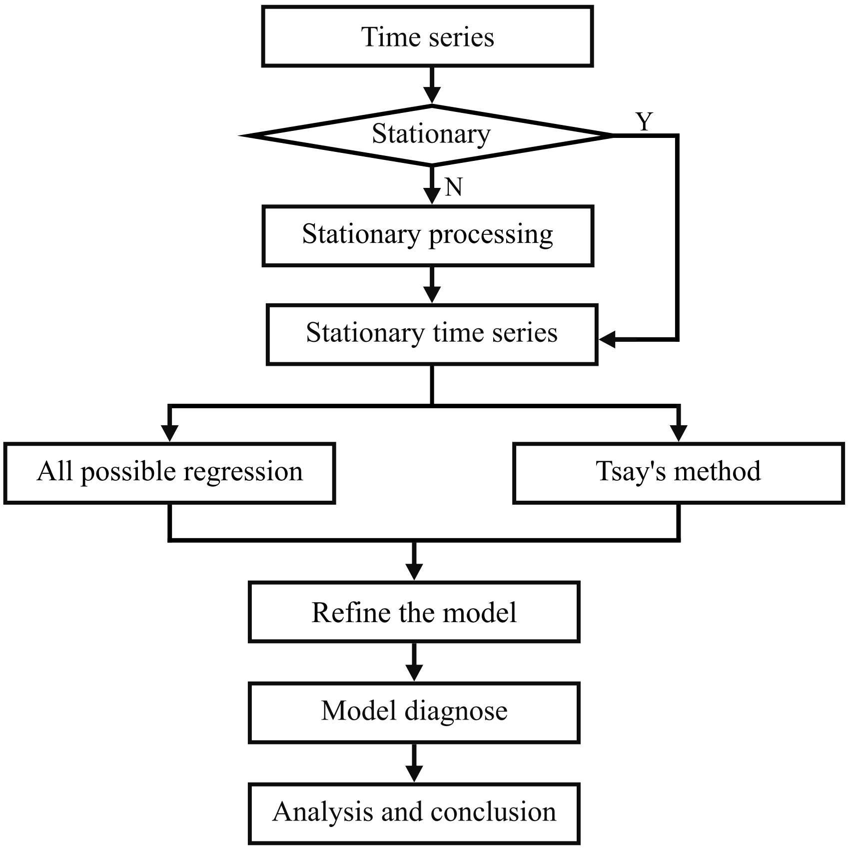
3.3 Model establishment process

In literature, GARCH (1,1) model is commonly used, which is characterized by simplicity, sound performance and avoiding the tedious model comparison process. But the reliability of the GARCH (1,1) model is, however, questionable. In this paper, two model selection methods are introduced to establish ARMA-GARCH models. To streamline our introduction, the flow chart illustrating the model establishment and analyses process is provided in Figure 3.

The right path in Figure 3 was introduced by Tsay (2005)[1], named the Tsay's method. This approach can be separated into five steps:

1. Specify a mean equation like ARMA models for the time series by the EACF plot.
2. Use the residuals of the ARMA model to test for ARCH effects.
3. Specify a volatility model like ARCH models by the ARMA effect of the squared

Figure 3 The flow chart of model establishment and analyses process



residuals, if the ARCH effect is statistically significant. The ARMA order is also determined by the EACF plot.

4. Perform a joint estimation of the ARMA-GARCH model.
5. Check the fitted model and refine it if necessary.

In general, the Tsay's method determines the order of ARMA and GARCH model separately, then estimates the parameters of the model jointly. However, this method is not capable of comparison of different GARCH models and different innovation distributions. Thus, the scope of the Tsay's method is restricted in the GARCH model and Normal distribution.

The left path in Figure 3 is the All possible regression method, which can be divided into three steps:

1. Set the significant level, model establishment scheme and criterion like AIC, BIC, etc.
2. Perform the regression for all the ARMA-GARCH models within the model establishment scheme and refine those models until all the parameters are significant under the predefined significant level.
3. Report the best model according to the predefined criterion.

This method will report the best model under certain criterion within the given scope. Yet, fitting and refining thousands of models required large-scale computational resources.

The detailed discussion and comparison of the two models will be given in the Section 4.1.

3.4 Scope of model establishment and objective

According to the research objective stated in Section 1.2, five model establishment schemes are constructed and summarized in Table 3. In scheme 1 and 2, we can compare the credibility of two model establishment methods. Armed with the All possible regression method, we assess the performance of ARMA-GARCH model extensions in scheme 3-5. Specifically, in scheme 3, we check the validity of IGARCH models and TGARCH models. Hereafter, in scheme 4, the capability of the Normal, Student's t and GED innovation distributions are studied. In scheme 5, the corresponding skewed derivatives are thoroughly compared under the skewed assumption.

The scheme 3 and 4, as the subsets of scheme 5, rest on the Normal and non-skewed assumption respectively, through which we can not only evaluate local best models but also find the right balance between model complexity and computational efficiency for the All possible regression method.

Table 3 Scopes of the model establishment for different schemes

Scheme	ARMA	Conditional Var model		Innov. dist.	Model selection method
		Model	Scope		
1	(0,0)-(4,4)	GARCH	(0,0)-(3,3)	N	Tsay's method
2	(0,0)-(4,4)	GARCH	(0,0)-(3,3)	N	All possible regression
3	(0,0)-(4,4)	GARCH TGARCH IGARCH	(0,0)-(3,3)	N	All possible regression
4	(0,0)-(4,4)	GARCH TGARCH IGARCH	(0,0)-(3,3)	N, T, GED	All possible regression
5	(0,0)-(4,4)	GARCH TGARCH IGARCH	(0,0)-(3,3)	N, T, GED, SN, ST, SGED	All possible regression

For the All possible regression method, the best models of are selected under the AIC and the searching scope of ARMA models and GARCH models are constrained subject to the limitation of computational capability.

3.5 Estimation results and analysis

Following the model establishment process in section 3.3, for four datasets, we obtain the best model of each model establishment schemes in Table 3. These best models are reported in Table A.1 and the estimate parameters of the best models are reported in Table A.2, parameters are left blank if they don't belong to the model.

According to scheme 1 and 2, in Table A.1, the All possible regression method acquires models with AIC lower than the Tsay's method in three datasets and reports the same model in dataset SSEC 13-16. Moreover, as what mentioned in Section 2.1.2, $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j$ represented the persistence of the volatility of GARCH models. In Table A.2, the sum of alpha1 and beta1 in the GARCH models are close to one, indicating a long-lasting interplay between past volatility and future volatility.

From scheme 3 in Table A.1, the IGARCH(1,1) model outperforms other models in contributing a lower AIC in all four datasets. It is confirmed with scheme 2 where the sum of alpha1 and beta1 is close to one. This fact suggests us to constrain the searching scope of the GARCH order in the All possible regression method, which will remarkably reduce the computational amount without loss of accuracy. Moreover, TGARCH models are not significant and the probable causes will be discussed further in Section 4.2.

In scheme 4, the AIC reduce larger under the GED innovation distribution assumption than in scheme 3. Since the GED assumption introduces one more variable into the model, this reduction indicates that the model fits the data set well under the GED hypothesis. Consisting with what we mentioned in the normality test, the shape parameter of the GED range between 1.0196 and 1.2997, which means the tail of the corresponding GED is thick.

From scheme 5, the skewed Generalized error distribution transcends all six innovation distribution candidates in AIC. But the reduction is modest compared with the GED. The skewness parameter, varying from 0.8817 to 0.9445, shows a slightly negative bias. Even though this inconsiderable bias is not significant under the TGARCH model.

To sum up, (1) the All possible regression method performs better in the ARMA-GARCH model establishment. (2) For the conditional variance model, the IGARCH (1,1) is recommended. (3) For heavy-tailed innovation distributions, the GED and SGED are recommended based on non-skewed and skewed innovation distribution assumptions respectively.

3.6 Residual analysis

Figure A.5 presentes the qq-plots for the residuals of our final model. Under the Normal innovation distribution assumption (scheme 1, 2 and 3), scatter points deviate from qq-lines on both sides, especially on the left-hand side. This phenomenon is similar to the qq-plots for the original time series. When we assume the innovation distribution follows GED (scheme 4), the scatter points converge towards the qq-lines but shows a left bias (thick tail on the left-hand side and thin tail on the right-hand side). As we assume the skewed-GED innovation distribution (scheme 5), the scatter points fit the qq-lines quite well even for outliers. In summary, the GED and SGED innovation distribution assumptions are quite plausible.

Furthermore, the BDS test is introduced to assess the independence of residuals. Under the null hypothesis, the test assumes that the time series is independent. From Table 4, residuals are independent under 95% of confidence level for all final models.

It is noteworthy that all four datasets show the same results and we only take dataset SSEC 05-10 as an example here in consideration of the length of the paper.

Table 4 P-value of the BDS-test of five schemes on dataset SSEC 05-10

Scheme 1			Scheme 2			Scheme 3		
Epsilon[a]	Dimension		Epsilon	Dimension		Epsilon	Dimension	
	2	3		2	3		2	3
0.5242	0.4254	0.8847	0.5257	0.8087	0.7048	0.5195	0.4289	0.8976
1.0485	0.4971	0.7706	1.0514	0.7030	0.6845	1.0390	0.5051	0.7764
1.5727	0.7334	0.7656	1.5770	0.9842	0.8764	1.5585	0.7313	0.7663
2.0970	0.7099	0.5504	2.1027	0.9893	0.7069	2.0780	0.7105	0.5532

Scheme 4			Scheme 5		
Epsilon	Dimension		Epsilon	Dimension	
	2	3		2	3
0.5227	0.2198	0.7374	0.5239	0.2360	0.8244
1.0455	0.5713	0.9199	1.0477	0.6801	0.9720
1.5682	0.8706	0.7211	1.5716	0.9777	0.7516
2.0909	0.8784	0.8950	2.0955	0.8756	0.8657

[a] The epsilon distance is chosen automatically.

3.7 Prediction

Apart from the residual analysis, we also evaluate the prediction performance of different models by ten-day forecasting. Figure 4 shows forecasting return rates and 95% confidence intervals for different schemes. The red horizontal lines represent zero return rates and the colored dash lines are the corresponding 95% confidence intervals. The confidence bands generate under different schemes are generally the same and tolerate unpredictable outliers like May 24 in SSEC 05-10 and Dec 21 in SSEC 13-16.

The loss functions for different schemes are provided in Table 5. The following analyses are based on the MSE loss function. Compare scheme 1 with 2, the best models of the All possible regression method yield better prediction performance in three datasets. According to scheme 3, 4 and 5, the IGARCH model with heavy-tailed innovation distributions achieves less loss in time span 05-10 but the GARCH model with Normal innovation distribution are better in time span 13-16. The complicated model yields a worse prediction performance in time span 13-16 might result from the overfitting, since the sample size in time span 13-16 is much less than that of time span 05-10.

Figure 4 10 step ahead predictions and corresponding confidence intervals

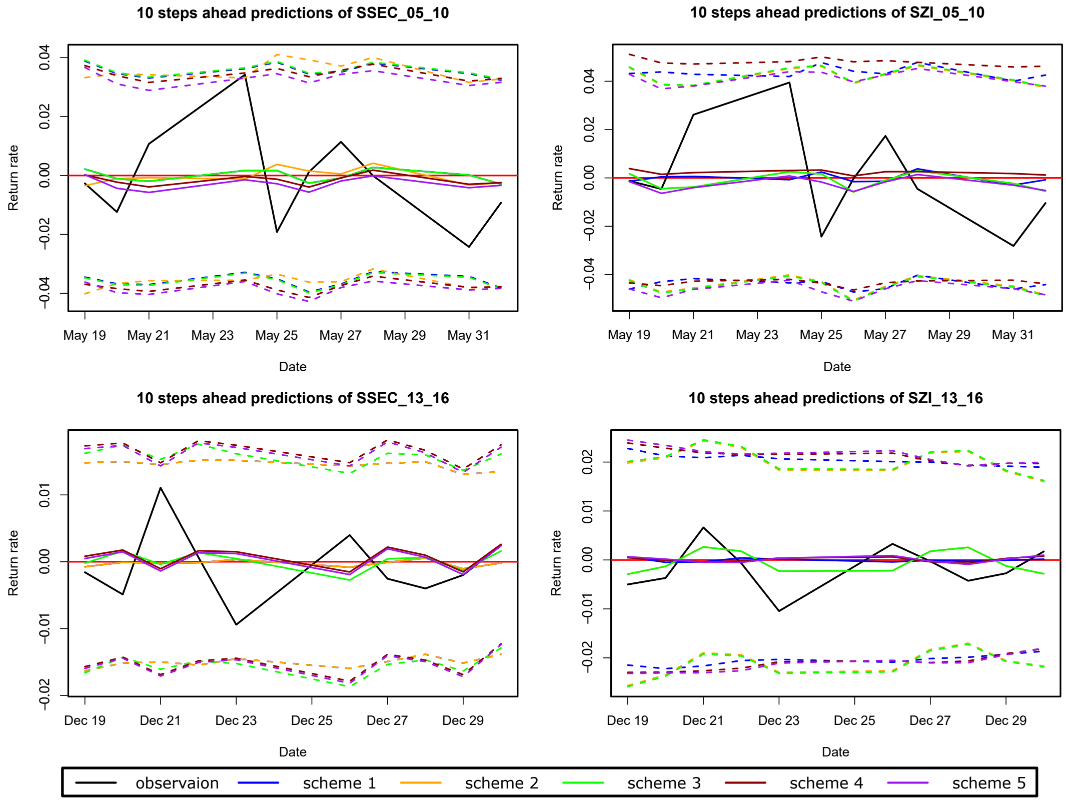


Table 5 Loss functions ($\times$1000) of predictions

	Scheme 1	Scheme 2	Scheme 3	Scheme 4	Scheme 5
			SSEC 05-10		
MSE	0.261	0.260	0.261	0.253	0.252
MAE	13.188	12.517	13.184	12.742	12.184
			SZI 05-10		
MSE	0.416	0.406	0.406	0.410	0.400
MAE	16.076	15.808	15.805	15.580	15.409
			SSEC 13-16		
MSE	0.030	0.030	0.035	0.039	0.039
MAE	4.282	4.282	4.567	4.889	4.737
			SZI 13-16		
MSE	0.024	0.020	0.020	0.024	0.024
MAE	3.967	3.943	3.949	3.769	3.700

4. Discussion and conclusion

4.1 Comparison of two model selection methods

From the point of view of computational requirement, only two EACF plots, one ARMA model estimation and one ARMA-GARCH model estimation are needed for Tsay's method. However, the All possible regression method requires a huge amount of computation to fit, refine and sort the potential models before reporting the best one. Take the scheme 5 as an example, 28800 models in total need to be checked from eight parameter dimensions (AR, MA, ARCH, GARCH, GARCH model, Innovation distribution, ω, and μ). Yet the total model will be reduced to $1/96$ if the order of the GARCH model and the innovation distribution are restricted to (1,1) and GED or SGED respectively, which is verified reasonable hereinbefore. Under this circumstances, the computational intensity will be quite acceptable.

From the standpoint of expansibility, the Tsay's method doesn't consider the comparison of GARCH models and innovation distributions, whereas the All possible regression method is born to make comparison among them. Thus the All possible regression method is recommended when establishing ARMA-GARCH models with different GARCH models and innovation distributions.

From the perspective of model establishing procedure, the Tsay's method performs the parameter estimation of the ARMA and GARCH model simultaneously, since the conditional mean model and conditional variance model will influence each other when parameters are estimated successively. But the Tsay's method determines the orders of ARMA and GARCH sequentially, that is, the order of the GARCH model is determined under the residual of the ARMA model, which may lead to ordering error. The All possible regression method, in a more plausible way, performs the order determination and the parameter estimation at the same time.

From the user's point of view, for the All possible regression method, no statistical discrimination from the user is required during the order determination procedure. The Tsay's method, on the contrary, requires users to determine the ARMA and GARCH orders from the EACF plot twice at least, which may be technical and complicated, especially when the order determination result illustrated by the sample EACF plot is vague. In a general sense, possessing automatic operation and explicit criterion, the method proposed in this paper is more appropriate for inexperienced users.

From the objective of the modeling process, the Tsay's method merely finds a significant model, but the optimality isn't checked. The All possible regression method reports a significant model outperforming other models under certain criterion. Thus, the optimality of our method is automatically guaranteed.

Last but not least, we now make some comments about the All possible regression method. The criterion we choose will fundamentally decide the best model we get. Normally, the Akaike information criterion (AIC) is good enough. But the Bayesian information criterion (BIC) normally suggests a smaller model, which might be helpful to avoid overfitting. What's more, additional experiments might be needed if the best model can not be directly determined in the predefined scope, especially when the best model lies in the boundary of the searching scope.

To sum up, the All possible regression method is the more reliable method at a higher computational cost. Under appropriate searching scope, it can strike the balance among broad application, predictive accuracy and computational efficiency.

4.2 The significance of TGARCH model

Unlike the asymmetric property mentioned in the introduction section, the leverage parameter of the TGARCH model is not significant among all four datasets. This phenomenon needs further investigation.

YingSHI (2017)[16] found the asymmetric property of Shanghai exchange 180 index is not significant under the EGARCH (1,1) model and the GJR-GARCH (1,1) model. According to the author, the return rate series includes a complete bull-bear transition, which may offset some of the asymmetrical characteristics of volatility. Therefore, the overall leverage effect is not significant. It is reasonable that the TGARCH model is not significant because our datasets also includes the bull, bear and consolidation period.

In conclusion, although negative skewness of the yield series can be observed statistically, significant asymmetry can not be captured by the TGARCH model.

4.3 Conclusion

Based on the results of the study, this paper puts forward suggestions below.

(1) The All possible regression method is recommended to establish the ARMA-GARCH model. (2) The IGARCH (1,1) model are recommended to be the conditional variance model. (3) The stock market behavior differently in different time spans, counteracting the leverage effect. Therefore, TGARCH models are not suitable for datasets consist of more than one market states. (4) The Generalized error distribution and skewed Generalized error distribution are suggested to be the innovation distributions under non-skewed and skewed assumption respectively. (5) The IGARCH (1,1) model and heavy-tailed innovation distributions performs worse than GARCH (1,1) model and Normal distribution in prediction when the sample size is small. If the sample size isn't large enough ($N<1000$ in our case), one should take care of overfitting. (6) To reduce computational complexity, the choice of the searching scope of the All possible regression method can follow the rules above.

4.4 Further Study

In this thesis, the ARMA-GARCH models have been proved to be a suitable and powerful tool for modeling the volatility of the Chinese stock market. Yet, some open questions need further study. First, the asymmetric volatility indeed exists but not significant. To capture the leverage property, the time span should be trimed to include only one market states. Furthermore, new criterion such as BIC need to be introduced for the All possible regression method to tackle the overfitting.

References

[1] TSAY R S. Analysis of financial time series[M]. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2005.

[2] GEORGE E. P. BOX G M J. Time Series Analysis: Forecasting and Control[M]. 5th ed. London: Palgrave Macmillan UK, 1970.

[3] FAMA E F. The Behavior of Stock-Market Prices[J]. The Journal of Business, 1965, 38(1): 34-105.

[4] GRANGER C, ANDERSEN A. An Introduction to Bilinear Time Series Models[M]. Gottingen: Vandenhoeck und Rupreckt, 1978.

[5] ENGLE R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation[J]. Econometrica, 1982, 50(4): 987-1007.

[6] 玄海燕. 几类时间序列模型统计推断及其应用[D]. 大连: 大连理工大学, 2019.

[7] BOLLERSLEV T. Generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics, 1986, 31(3): 307-327.

[8] ENGLE R F, BOLLERSLEV T. Modelling the persistence of conditional variances[J]. Econometric Reviews, 1986, 5(1): 1-50.

[9] FISHER B. Studies in stock price volatility changes[J]. Proceedings of the 1976 Business meeting of the business and economics statistics section, 1976: 177-181.

[10] ZAKOIAN J M. Threshold heteroskedastic models[J]. Journal of Economic Dynamics and Control, 1994, 18(5): 931-955.

[11] NELSON D B. Conditional Heteroskedasticity in Asset Returns: A New Approach[J]. Econometrica, 1991, 59(2): 347-370.

[12] GLOSTEN, R. L, JAGANNATHAN R, et al. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks[J]. The Journal of Finance, 1993, 48(5): 1779-1801.

[13] BOLLERSLEV T. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return[J]. The Review of Economics and Statistics, 1987, 69(3): 542-547.

[14] HANSEN B E. Autoregressive Conditional Density Estimation[J]. International Economic Review, 1994, 35(3): 705-730.

[15] ENGLE R F, LILIEN D M, ROBINS R P. Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model[J]. Econometrica, 1987, 55(2): 391-407.

[16] 石颖. 基于 GARCH 模型的上证 180 指数收益率波动研究[D]. 长沙: 中南财经政法大学, 2019.

[17] WANG Y, XIANG Y, LEI X, et al. Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market[J]. Economic Research-Ekonomska Istraživanja, 2021, 1(1): 1-25.

[18] 顾慧旦. 基于 G-ARMA-GARCH 族模型的沪深指数日收益率序列模型研究[D]. 南京: 南京大学, 2017.

[19] FERNANDEZ C, STEEL M F J. On Bayesian Modeling of Fat Tails and Skewness[J]. Journal of the American Statistical Association, 1998, 93(441): 359-371.

Appendix A

Figure A.1 Time series of indexes, return rates and squared return rates for four datasets

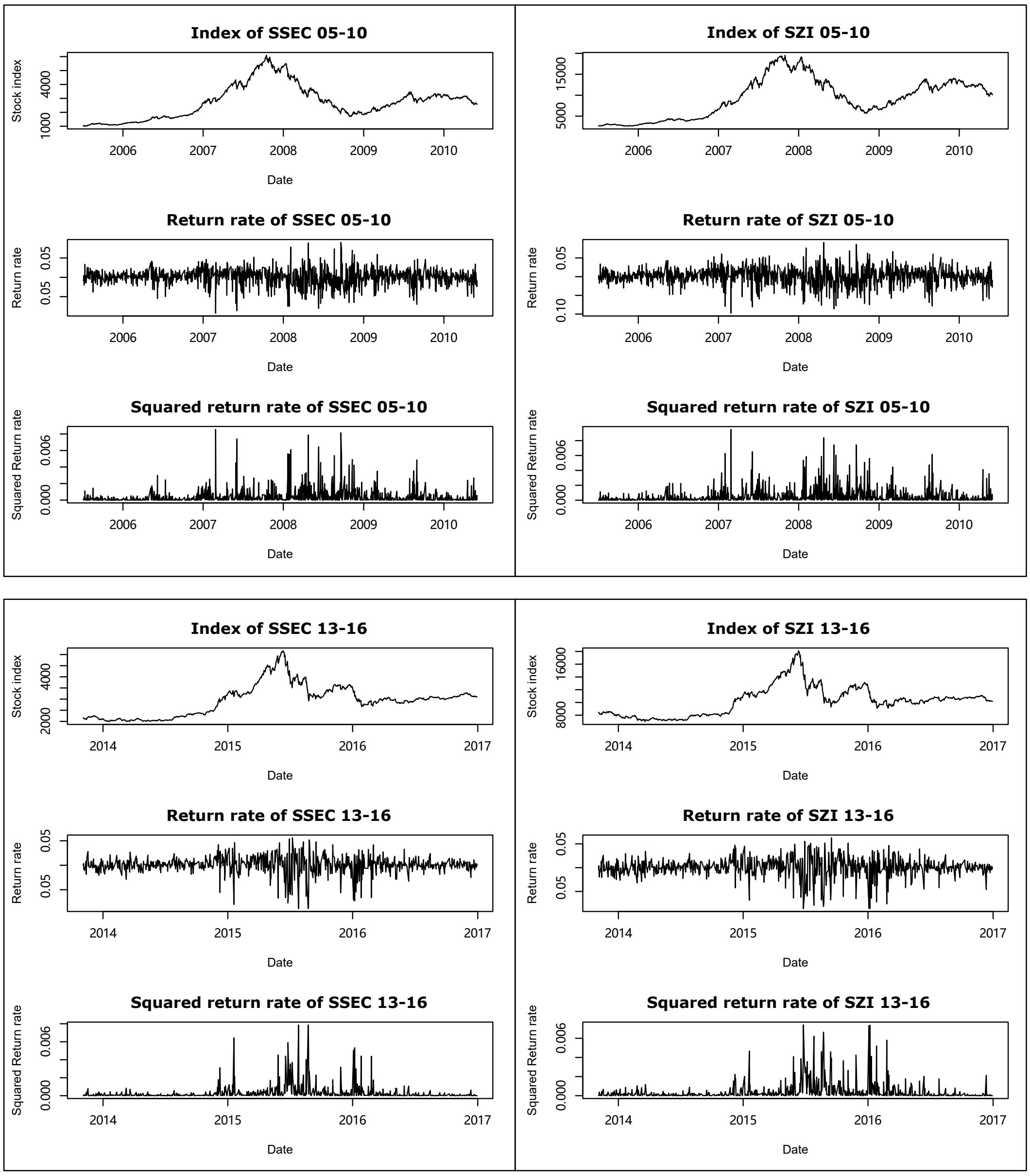


Figure A.2 Densities of return rates and their normal approximations

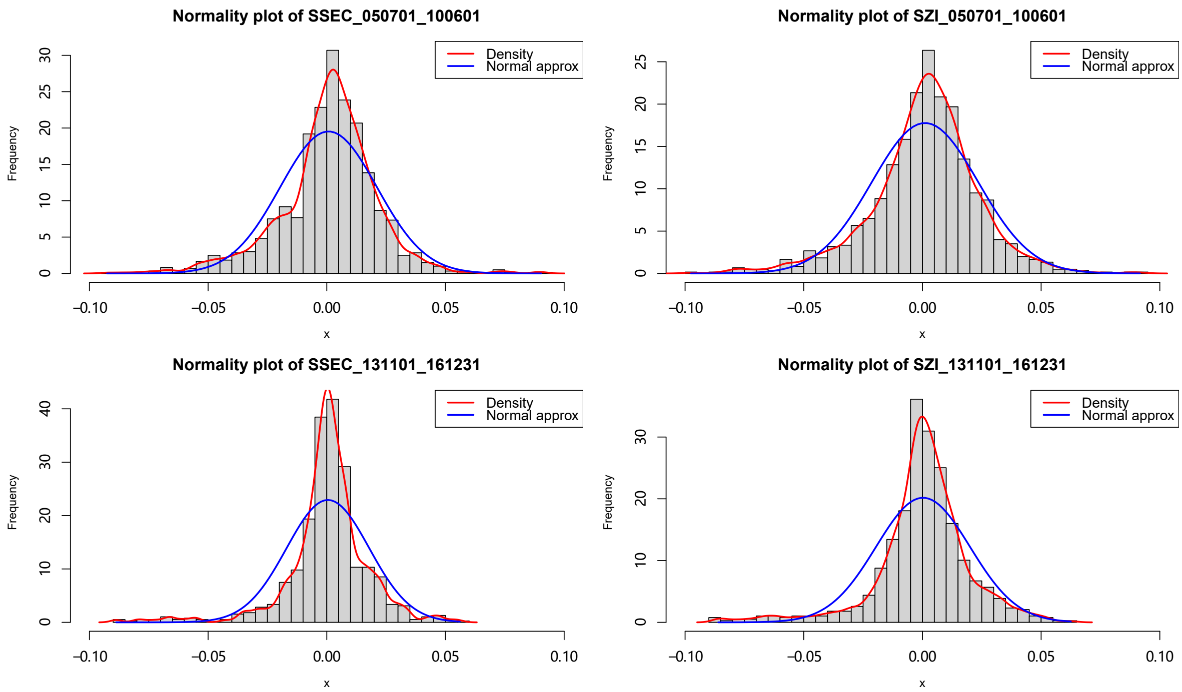


Figure A.3 QQ-plots of return rates

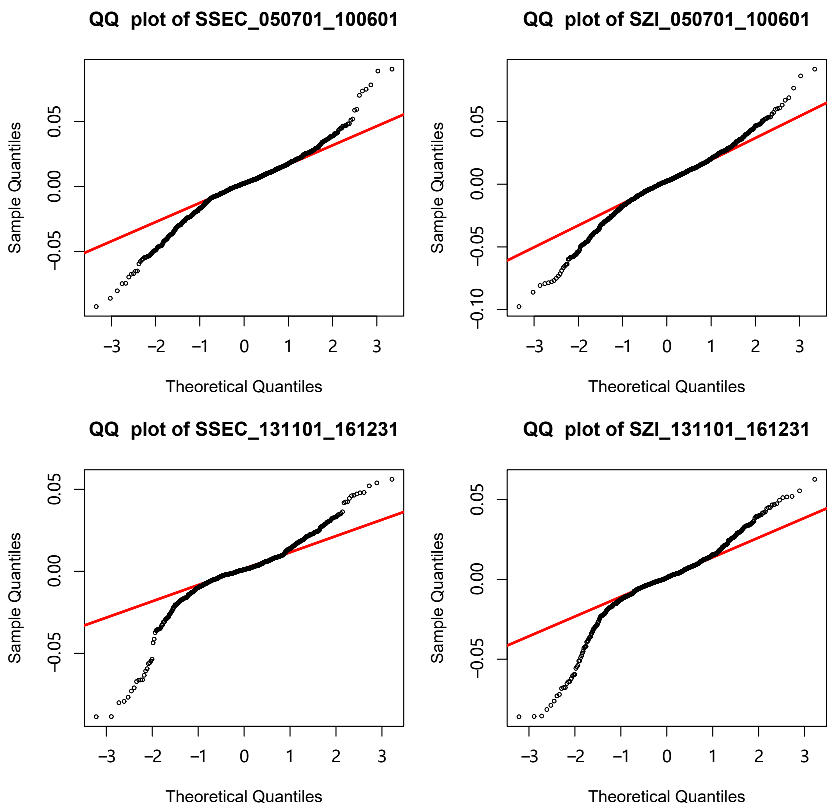


Figure A.4 ACF and PACF of return rates and squared return rates

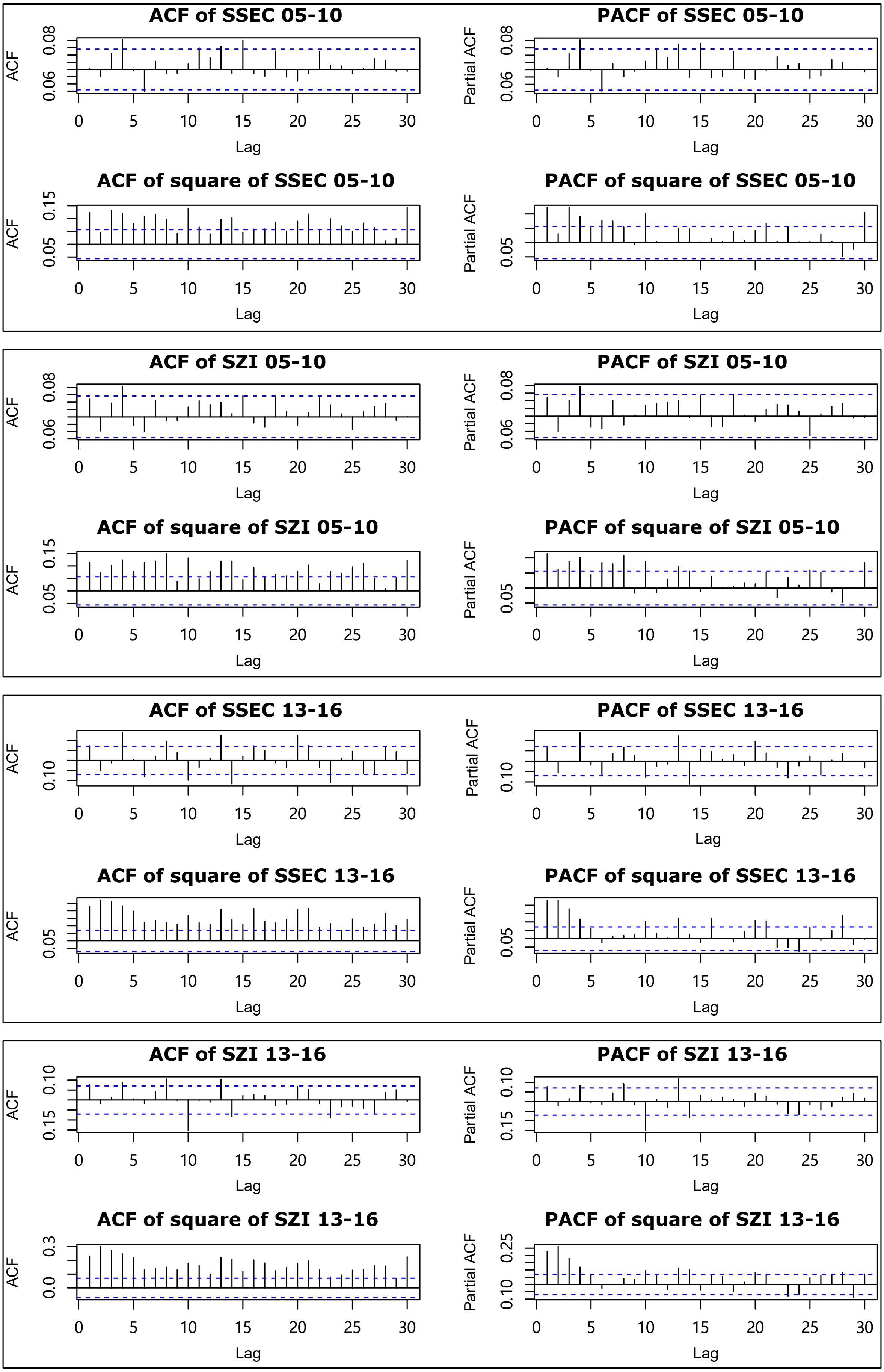


Figure A.5 QQ-plots of five schemes on dataset SSEC 05-10

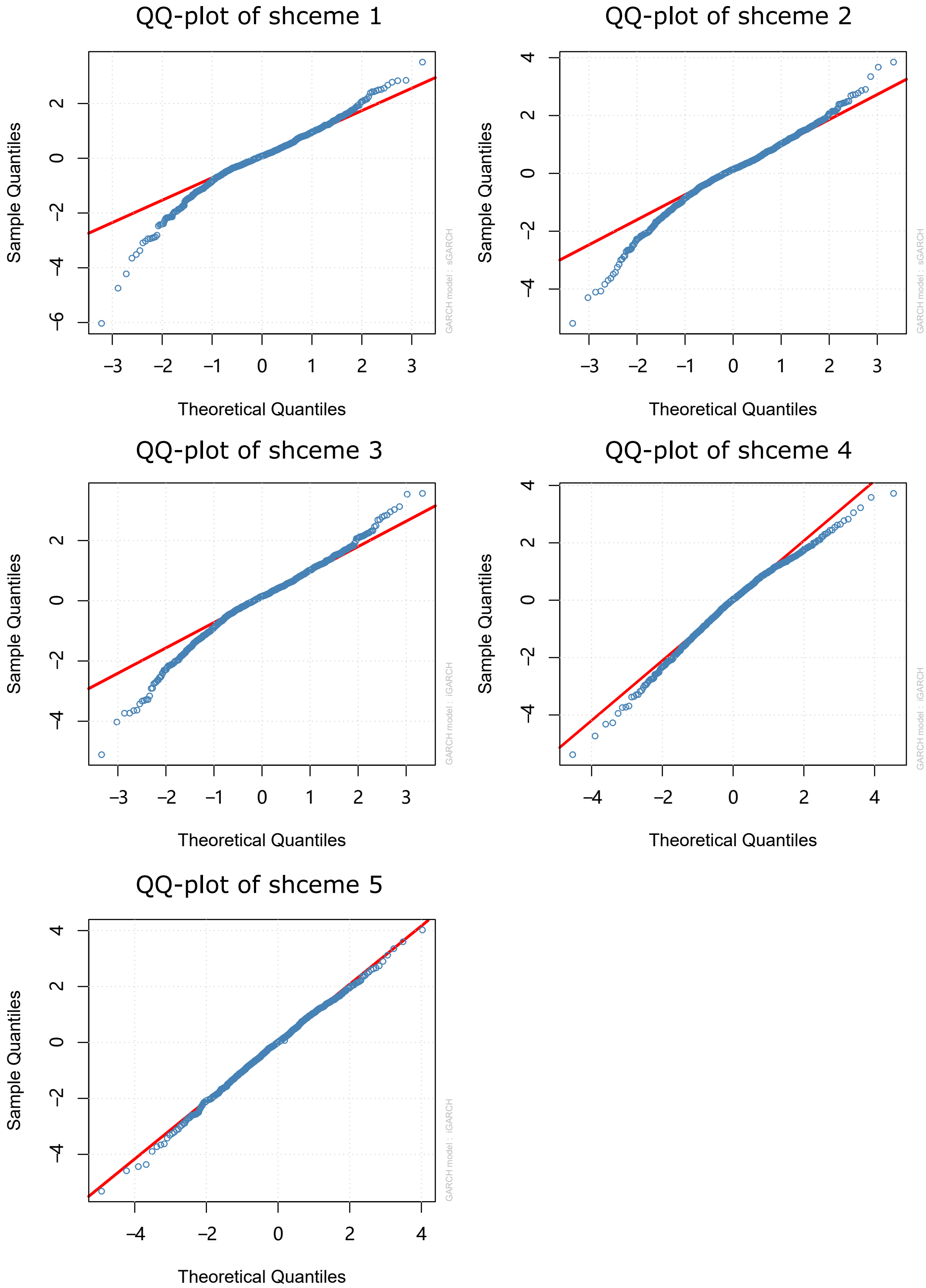


Table A.1 The best models and corresponding AIC

Dataset		Scheme	ARMA-GARCH	μ	ω	Innov. Dist.	AIC
05-10	SSEC	1	(2,2)-(1,1)	0	0	N	-5.1093
		2	(4,4)-(1,1)	0	0	N	-5.1117
		3	(2,2)-I(1,1)[a]	0	0	N	-5.1128
		4	(4,3)-I(1,1)		0	GED	-5.1917
		5	(4,3)-I(1,1)	0	0	SGED	-5.2052
	SZI	1	(2,2)-(1,1)	0	0	N	-4.8907
		2	(4,3)-(1,1)	0	0	N	-4.8941
		3	(4,3)-I(1,1)	0	0	N	-4.8977
		4	(1,1)-I(1,1)	0	0	GED	-4.9485
		5	(3,3)-I(1,1)	0	0	SGED	-4.9639
13-16	SSEC	1	(3,3)-(1,1)	0	0	N	-5.7360
		2	(3,3)-(1,1)	0	0	N	-5.7360
		3	(5,4)-I(1,1)	0	0	N	-5.7457
		4	(3,3)-I(1,1)		0	GED	-5.8465
		5	(3,3)-I(1,1)	0	0	SGED	-5.8472
	SZI	1	(3,3)-(1,1)	0	0	N	-5.3460
		2	(2,2)-(1,1)	0	0	N	-5.3526
		3	(2,2)-I(1,1)	0	0	N	-5.3568
		4	(2,2)-I(1,1)	0	0	GED	-5.4564
		5	(2,2)-I(1,1)	0	0	SGED	-5.4659

[a] I stands for the IGARCH model.

Table A.2 Detailed estimations of the best model

Param.		ARMA										GARCH		Innov. Dist.			AIC
		μ	ar1	ar2	ar3	ar4	ar5	ma1	ma2	ma3	ma4	alpha1	beta1	dist.	skew	shape	
SSEC_05_10	Scheme 1: ARMA(2,2)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		-0.190	-0.988				0.196	0.973			0.055	0.944	N			-5.109
	Std. Error		0.008	0.005				0.012	0.003			0.011	0.013				
	Scheme 2: ARMA(4,4)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		1.357	-1.541	1.061	-0.756		-1.348	1.496	-0.959	0.709	0.057	0.942	N			-5.112
	Std. Error		0.030	0.036	0.034	0.026		0.018	0.008	0.004	0.016	0.012	0.013				
	Scheme 3: ARMA(2,2)-IGARCH(1,1)-μ=0-ω=0-N																
	Estimate		-0.190	-0.988				0.196	0.973			0.055	0.945	N			-5.113
	Std. Error		0.007	0.005				0.012	0.003			0.013	NA				
	Scheme 4: ARMA(4,3)-IGARCH(1,1)-ω=0-GED																
	Estimate	0.004	0.486	-0.515	0.962	0.054		-0.515	0.550	-0.977		0.053	0.947	GED		1.176	-5.192
	Std. Error	0.001	0.037	0.007	0.008	0.009		0.008	0.003	0.001		0.015	NA			0.062	
	Scheme 5: ARMA(4,3)-IGARCH(1,1)-μ=0-ω=0-SGED																
	Estimate		0.726	-0.516	0.757	0.035		-0.746	0.501	-0.714		0.055	0.945	SGED	0.882	1.183	-5.205
	Std. Error		0.008	0.008	0.011	0.008		0.000	0.000	0.000		0.015	NA		0.018	0.069	
SZI_05_10	Scheme 1: ARMA(2,2)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		-0.966	-0.798				1.029	0.828			0.051	0.948	N			-4.891
	Std. Error		0.167	0.090				0.165	0.087			0.011	0.013				
	Scheme 2: ARMA(4,4)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		0.867	-0.872	1.034	-0.055		-0.804	0.829	-0.974		0.052	0.947	N			-4.894
	Std. Error		0.002	0.004	0.002	0.005		0.000	0.001	0.000		0.011	0.013				
	Scheme 3: ARMA(4,4)-IGARCH(1,1)-μ=0-ω=0-N																
	Estimate		0.866	-0.871	1.033	-0.054		-0.804	0.829	-0.974		0.052	0.948	N			-4.898
	Std. Error		0.002	0.003	0.002	0.004		0.000	0.001	0.001		0.013	NA				
	Scheme 4: ARMA(3,3)-IGARCH(1,1)-ω=0-GED																
	Estimate		1.000					-0.981				0.050	0.950	GED		1.297	-4.949
	Std. Error		0.001					0.001				0.013	NA			0.070	
	Scheme 5: ARMA(3,3)-IGARCH(1,1)-μ=0-ω=0-SGED																
	Estimate		0.819	-0.820	0.996			-0.806	0.823	-0.975		0.052	0.948	SGED	0.882	1.298	-4.964
	Std. Error		0.005	0.002	0.008			0.006	0.012	0.004		0.014	NA		0.029	0.072	
SSEC_13_16	Scheme 1: ARMA(3,3)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		-0.063	-0.145	0.840			0.109	0.163	-0.855		0.063	0.936	N			-5.736
	Std. Error		0.029	0.026	0.026			0.005	0.002	0.001		0.013	0.014				
	Scheme 2: ARMA(3,3)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		-0.063	-0.145	0.840			0.109	0.163	-0.855		0.063	0.936	N			-5.736
	Std. Error		0.029	0.026	0.026			0.005	0.002	0.001		0.013	0.014				
	Scheme 3: ARMA(5,4)-IGARCH(1,1)-μ=0-ω=0-N																
	Estimate		-1.599	-1.482	-1.545	-0.849	0.067	1.678	1.599	1.653	0.985	0.066	0.934	N			-5.746
	Std. Error		0.004	0.004	0.006	0.007	0.005	0.001	0.001	0.001	0.001	0.014	NA				
	Scheme 4: ARMA(3,3)-IGARCH(1,1)-ω=0-GED																
	Estimate	-0.001	-0.213	0.240	0.981			0.225	-0.218	-0.994		0.066	0.934	GED		1.020	-5.847
	Std. Error	0.000	0.003	0.002	0.004			0.001	0.001	0.000		0.013	NA			0.070	
	Scheme 5: ARMA(3,3)-IGARCH(1,1)-μ=0-ω=0-SGED																
	Estimate		-0.164	0.201	0.971			0.193	-0.189	-0.991		0.066	0.934	SGED	0.945	1.068	-5.847
	Std. Error		0.004	0.006	0.005			0.001	0.002	0.001		0.011	NA		0.030	0.078	
SZI_13_16	Scheme 1: ARMA(3,3)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		0.459	-1.029	0.261			-0.423	1.019	-0.204		0.045	0.954	N			-5.346
	Std. Error		0.040	0.020	0.048			0.014	0.001	0.022		0.012	0.013				
	Scheme 2: ARMA(2,2)-GARCH(1,1)-μ=0-ω=0-N																
	Estimate		0.197	-1.000				-0.206	1.004			0.047	0.952	N			-5.353
	Std. Error		0.004	0.005				0.002	0.001			0.013	0.013				
	Scheme 3: ARMA(2,2)-IGARCH(1,1)-μ=0-ω=0-N																
	Estimate		0.197	-1.000				-0.206	1.004			0.048	0.952	N			-5.357
	Std. Error		0.004	0.005				0.002	0.001			0.014	NA				
	Scheme 4: ARMA(2,2)-IGARCH(1,1)-μ=0-ω=0-GED																
	Estimate		0.220	-0.969				-0.245	0.976			0.058	0.942	GED		1.130	-5.456
	Std. Error		0.012	0.004				0.017	0.028			0.014	NA			0.084	
	Scheme 5: ARMA(2,2)-IGARCH(1,1)-μ=0-ω=0-SGED																
	Estimate		0.192	-0.953				-0.223	0.951			0.061	0.939	SGED	0.905	1.163	-5.466
	Std. Error		0.005	0.009				0.006	0.008			0.010	NA		0.026	0.086	

[a] Since the IGARCH model estimate beta1 directly from 1-alpha1, the Std. Error , t Value and Pr(>|t|) are NAs.
[b] The ω is not significant for all best models, so it's not included in the table.
[c] The parameters are all significant under 99% confidence level.

Acknowledgement

When I first came to the Southern University of Science and Technology, I wanted to major in a basic subject that is closely related to practical applications. I chose statistics at first, not because I love statistics itself, but because it is widely distributed in various disciplines, which can be helpful when I switch to another major. However, in the later specialized courses, I learned a series of principles of statistics, such as a large proportion of real-world datasets following the central limit theorem and the influence of prior knowledge on decision-making can be depicted by the Bayesian statistics. The concise theorems from textbooks mixed with the vivid observations in daily life make me fascinated. One day, I read the *Statistics and Truth* written by C.R.Rao, whose title page read:

All knowledge is, in final analysis, history.
All sciences are, in the abstract, mathematics.
All judgements are, in their rationale, statistics.

Yes! Statistics is the very subject I dream to study, which has profound theoretical foundation and wide application.

Here, I would like to express my warmest gratitude to my thesis supervisor, Professor Xuejun JIANG, for his great guidance, patience and encouragement. Only the excessive amount of time he took the trouble to talk to me and to review my drafts made this thesis possible. I would also like to thank Hongxia JIN for her valuable insight on my manuscript. All remaining errors are my own.

Meanwhile, I want to thank my parents for their tolerance that I neglected them during the thesis preparation and their financial support. I would also like to thank my friends, including my roommate, Haodong GAO and friends from the Department of Statistics and Data science Qiyu HUANG, Xuelin ZHU, Xingran TAO, and so on. Their encouragement and enthusiasm have always been crucial in making me move forward.

Finally, I would like to quote the last line from Alfred Tennyson's 'Ulysses': 'To strive, to seek, to find, and not to yield'. These words inspired me during my most difficult time.