Future Generation Computer Systems 79 (2018) 960–972

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Modelling and forecasting the stock market volatility of SSE Composite Index using GARCH models

Zhe Lin

School of Accounting & Finance, Xiamen University Tan Kah Kee College, Zhangzhou Fujian, 363100, China

HIGHLIGHTS

- This paper studies the econometric features of the SSE Composite Index, using GARCH type models, and compares the adaptability of these models in fitting Chinese stock market.
- Results show that the SSE Composite Index possesses significant properties of time-varying and clustering. And its series distribution presents leptokurtosis with significant ARCH and GARCH effects.
- By comparing the fitting and forecast performance of GARCH (1, 1) (symmetric) and TARCH (1, 1), EGARCH (1, 1) (asymmetric), it turns out that EGARCH (1, 1) generally outperforms the others.

ARTICLE INFO

Article history:
Received 20 June 2017
Received in revised form 29 July 2017
Accepted 16 August 2017
Available online 4 October 2017

Keywords:
Modelling
Forecasting
Stock market
SSE Composite Index
GARCH models
Stock price
Mutual conduction

ABSTRACT

The stock market is constantly changing with uncertainties. Rapid dissemination of information and fast capital flow will lead to fluctuations of stock price, and the undulating price will affect the market in return. This is a process of mutual influence and mutual conduction. China's stock market, which pertains to an emerging market, has been acutely volatile since the very beginning, and often appear radical ups and downs. This paper selects the SSE Composite Index as research object, through the application of GARCH type models to conduct empirical analysis, carving the features of this index from an econometric perspective. And on basis of the status quo of the volatility of SSE Composite Index, it offers some suggestions.

The result shows that from the time series point of view, the SSE Composite Index possesses significant properties of time-varying and clustering. Series distribution of it presents leptokurtosis with significant ARCH and GARCH effects. Moreover, by comparing the fitting and forecast performance of GARCH (1, 1) (symmetric) and TARCH (1, 1) and EGARCH (1, 1) (asymmetric), it can be concluded that EGARCH (1, 1) outperforms the others. Besides, China's securities market should strengthen its system construction, reduce excessive government intervention and advocate rational investment philosophy.

1. Introduction

1.1. Background

Since the Shanghai Stock Exchange was formally established in 1990, China's stock market has gone through a development journey of 27 years. Whereas it is now still of no high level of normalisation in terms of supervision or institution, and with a strong volatility. The unstable stock market, unscientific investments and occasional event of malignant investment make the whole market full of high-risk, and propose several challenges to institutions and individuals [1]. Many foreign institutional investors approved by the CSRC (China Securities Regulatory Commission) have joined the ranks of investors. Owing to the uncertainty and high-speed international capital flows, the environment of financing and investments has become increasingly complicated. China's securities market is experiencing a period of both opportunities and risks [2]. In this case, an accurate description of how the stock price fluctuates and how to determine the future rate of return of the stock market has become a hot issue in the academia and the investment community. Therefore, the study of volatility has significant theoretical significance and applicable value.

Traditional econometric model (such as regression analysis and time series analysis) generally assume that the variance remains unchanged in explaining the volatility of stock market return. While a large number of empirical studies show that such an assumption is not comprehensive, which cannot accurately describe the financial data objectively [3]. Financial time series data

E-mail address: zhelin@xujc.com.

https://doi.org/10.1016/j.future.2017.08.033

have unstable characteristics, greater volatilities will gather in the same period of time, and small fluctuations will gather in another period of time, this is the volatility clustering phenomenon [4]. For time series with characteristics of 'sharp peak', 'fat tail' and volatility clustering, apparently the homoscedasticity assumption is not satisfied. According to Liu and Morley [5], using traditional econometric models to conduct modelling, often would result in very serious deviation. To improve the accuracy of prediction model, economic scholars had conducted extensive to improve the traditional econometric models. In 1982, American economist, Professor Engle first proposed the ARCH (Auto-Regressive Conditional Heteroskedastic) model to modelling variance. ARCH model subverts the traditional way of thinking, and denies the linear hypothesis of the relation between risk and returns. This model uses the changing variance to establish a function which is related to the past volatilities, and it better characterises the financial data with features of 'sharp peak' and 'fat tail'. Moreover, it provides a powerful research tool for the heteroscedasticity problem, thus since it was put forward, it is much favoured by scholars. However, people gradually found that when using the ARCH model in modelling some time series, it requires a large order q to better estimate conditional heteroscedasticity. Hence, Bollerslev [6] introduced lag phase to conditional variance on basis of the ARCH model, and got the generalised ARCH model, namely GARCH (generalised autoregressive conditional heteroscedasticity) model. Thereafter, a number of scholars expanded various models similar with GARCH model, such as IGARCH, EGARCH, GARCH-M and VGARCH, forming a GARCH model family.

GARCH model family is a series of models designed to explain the regularity of fluctuations in the time series, and they are of really good ability in describing the volatility of financial data, and with a wide range of theoretical and practical value. For example, GARCH model family can be applied to test problems like efficiency of the stock market [7]; in the foreign exchange market, the GARCH model family is usually used to describe the fat tails phenomenon in the alternative swift of stationary and fluctuating period which commonly appears in the market [8]; in the stock market, GARCH model family plays an important role in risk estimation and risk management [9]; similarly, the GARCH model family can also be applied to areas like interest rate, option price, futures, and inflation rate [10]. And it is because GARCH model family takes into account of two important characteristic of financial data: excess kurtosis and volatility clustering, in this paper GARCH models are adopted to fit the variance of stock returns.

1.2. Research purpose

This paper is to use the symmetric GARCH model (GARCH (1, 1) and asymmetric GARCH model (EGARCH and TARCH model) to modelling and estimate the volatility of Shanghai Stock Exchange Composite Index. Through the empirical analysis based on these models, we can learn about the characteristics of the volatility of index. Moreover, since the volatility of SSE Composite Index contains volatility information of the whole Shanghai Stock Exchange and even the entire China's Stock Market, by studying it, we can discover some institutional problems of the stock market in China. And then we dig into their causes to offer some suggestions. Besides, by applying these three models into the estimation and forecast of SSE Index's volatility, and dividing them into symmetric and asymmetric model, I was trying to reveal the adaptability of these different models in Chinese stock index, and compare the different properties between symmetric and asymmetric GARCH models. Finally I will determine an optimal model in fitting and forecasting the SSE Composite Index, according to their statistical performance.

As mentioned above, the selection of SSE Composite Index is due to the consideration of its comprehensive coverage of volatility information of the China's stock market. The study of it not only can we know about the single index, but also have a glance at the Chinese stock market situation. In addition, the data of it is published, and easily to collect. The selection of GARCH type models is owing to their outstanding capability of estimating time series with heteroscedasticity. Furthermore, the application of asymmetric GARCH models is to better describe the asymmetric effect in the index volatility.

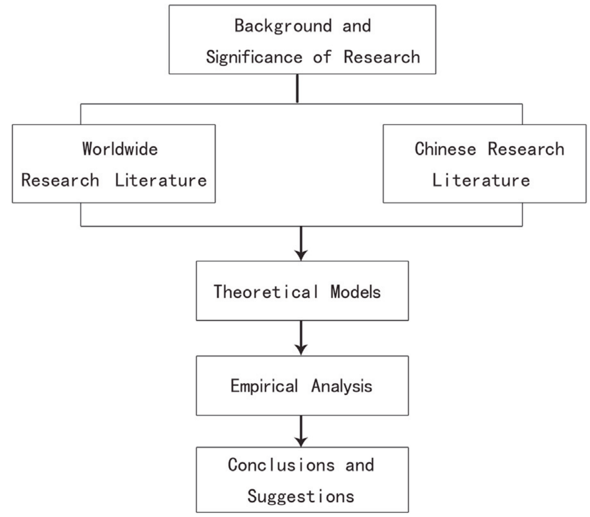

Fig. 1. Research approach.

1.3. Structure of paper

Chapter 1 introduces the research background, purpose and approaches.

Chapter 2 is the summary of relevant literature. In this section, I briefly reviewed the typical researches and corresponding empirical results in China and other countries about the volatility of stock index. In order to grasp the research status in this area and conduct research on basis of the previous studies.

Chapter 3 is the statement of theoretical models. Here I mainly describe the ARCH, GARCH model, ARCH-M, EGARCH and TARCH model, including their definition and characteristics.

Chapter 4 is the empirical analysis. I collected the daily closing price of the SSE Composite Index from January 3, 2011 to January 30, 2015, totally 996 sample observations. Applying the Eviews 8.0 and Excel to conduct empirical analysis on GARCH models, including descriptive statistic analysis, stationarity test, ARCH effect test and so on.

Chapter 5 draws conclusions and offers suggestions on basis of the volatility situation of SSE Composite Index (see Fig. 1).

2. Literature review

One of the core issue of modern financial research is the volatility in price changes of financial markets. Yan and Jiang [11] pointed out the possible reasons that people study the financial market and concern the risk of it. First, the price volatility in short-term asset makes it difficult for investors to believe the changes are due to economic fundamentals, which will reduce their enthusiasm of investment and turn to other markets. Second, investors are particularly concerned about the company in which they invest, especially the probability of default of the company. And generally the more violent the value of the company fluctuates, more possible the company is going to default. Third, an important determinant of the risk premium is the volatility, usually the greater the volatility, the greater the risk premium. Besides, the hedging financial instruments like insurance, the price of them is generally positively correlated with their volatilities. Finally, the economic and financial theory assumes that investors are generally risk averse, when the volatility of an investment project increases, the participation of its investors will be reduced. For these reasons, and in addition of the touch of 1987 US stock market crash, scholars began to pay attention and trace the methods for measuring financial market risks.

Since the analysis of this paper is about Chinese stock market, in order to integrate its specific national conditions, and in accordance with the special nature of this market, in the literature reading, I studied not only foreign literature, but also relevant Chinese literature. Therefore, for a more clear description, in this section will be divided into two parts: literature review of the world research and literature review of researches in China. The relevant research results will be given in both theoretical and empirical way. And finally there will be an evaluation of literature review.

2.1. Literature review of worldwide research

Early in 1960s, academia started to investigate the volatility of the stock market yield. Mandelbrot [4] found that stock market volatility has a feature of clustering, namely larger fluctuations will follow another larger ones. Engle [12] first proposed the ARCH model and its relevant extensions. Since the ARCH model family has a unique advantage in stimulating variance of financial time series, it has been widely applied in the researches of various financial markets, including stocks, foreign exchange, futures and currency market. In recognition of outstanding contributions of Engle and Granger in time series model, they are awarded the Nobel Prize in Economics in 2003. The so-called time series is a set of numbers formed by successive observations at different time points, in the same kind of phenomenon (for example, stock market volatility phenomenon) [13]. From a statistical perspective, this kind of chronological series is subject to various contingent factors, it therefore appears a certain randomness. Bollerslev [6] extended the factors affecting the conditional variance to two aspects: the mean square error and conditional variance of previous periods, and established the GARCH model.

On the basis of the two models mentioned above, the researchers found the presence of asymmetric information phenomenon in the fluctuations of financial time series, namely the fluctuations causing by the bad news are always much greater than the good news'. And this phenomenon is also proved by the study of Black [14], French, Schwert and Stambuahg [15], Nelson [16] and Zakoian [17]. Engle and Ng [18] studied the Japanese stock market, and they applied the EGARCH, AGARCH, NGARCH, VGARCH and GJR-GARCH model simultaneously in the research, proposing the concept of news impact curve. Moreover, they found by empirical analysis that the Japanese stock market indeed existed volatility asymmetry. And the conditional variance estimated by EGARCH model were generally larger than other models, but in the reflection of asymmetric effect, the GJR-GARCH model did the best. While Yeh and Lee [19] doubted this result, they test the GJR-GARCH model on the Chinese stock market, and found that this model generates a higher estimated conditional variance during the period of high volatility. Danielson [20] found in estimating the daily data of S&P 500 Index from 1980 to 1987, EGARCH (2, 1) model performed better than ARCH (5), GARCH (1, 2) and IGARCH (1, 1, 0). His results also shows that EGARCH model is superior to the simple SV model which uses the maximum likelihood to do estimation. Moreover, the difference of log likelihood between dynamic SV model and EGARCH model in his study was 25.5, which supports his another conclusion that SV model with 4 parameters is better than EGARCH model with 5 parameters. Cheung, He and Ng [21] discovered that stock markets in the same area often appear relatively high similarities due to geographical proximity, and therefore they inferred that common information will lead to similar heteroscedasticity in markets of the same area. Su and Fleisher [22] analysed the daily closing price of Shanghai Composite Index from 1990 to 1997, and found that Chinese government's intervention had a large impact on China's stock market. Varma [23] studied daily data of a stock index called Nifty in India from 1990 to 1998, showing that GARCH (1, 1) with generalised distribution of residual has much advantages in volatility estimation than other models. Brooks [24] selected 10 typical stock market around the world to conduct sampling data analysis, and the results shows that PARCH model can be applied to most countries in analysing the volatility of stock returns and asymmetry effect. Zakoian [17] and Singleton [25] used GARCH models to fit the yield of the stock market, proving the existence of volatility asymmetry.

Siourounis [26] used GARCH family models to estimate the Athens stock exchange market, discovering that negative shocks affect asymmetrically on the daily return series. Najand [27] estimated the S&P 500 returns data and found in stock volatility forecast, sometimes asymmetric models like EGARCH, TRARCH etc. provide more fitness than GARCH(1, 1), GARCH-M such symmetric models. By using different GARCH models to analyse the volatility of S&P 500 Index, Awartani and Corradi [28] found that under the premise of symmetric information, GARCH (1, 1) can accurately describe the volatility of the data; and when under circumstances of asymmetric information, the asymmetric GARCH models would be more suitable. Banerjee and Sarkar [24] observed that in the Indian stock market asymmetric GARCH models is more suitable than the symmetric GARCH models. Biswas [29] and Hung [30] also believed that asymmetric GARCH models perform better in the estimation of stock market volatility. Whereas, Magnus and Fosu [31] were in favour of the view that GARCH (1, 1) with an assumption of normal error distribution is superior to other conditional volatility models in forecast the day-to-day data from a stock exchange in Ghana. Study results of Pagan and Schwert [32] shows that the fitting effect and predictive ability of both GARCH and EGARCH model are preferable. Angelidis and Benos [33] chose 5 different stocks from the NYSE to conduct estimation, applying GARCH, EGARCH and TGARCH model, and compared the advantages and disadvantages of these conditional volatility models. They pointed out that the sample capacity plays an essential role in the accurate prediction of various GARCH models. Furthermore, Guidi [34] applied some of the GARCH type models to the Swiss, UK, and German stock market indexes. It turned out the EGARCH model is considered to be the optimal in conditional variance modelling and forecasting. Sabiruzzaman et al. [35] investigated the Hong Kong stock market, applying both GARCH and TGARCH model, and compared the accuracy of volatility estimation between these two kinds of models. The results revealed that TGARCH model can better estimate the leverage effect in the stock effect than GARCH model. Liu and Huang [36] also verified that in the case of asymmetric information and different distributions, the asymmetric GARCH models are of greater accuracy in predicting volatility in stock market returns, in their study of S&P 100 Index.

2.2. Literature review of research in China

The study of stock market return volatility in China started very late, while with the growth and improvement of China's capital market, relevant researches are gradually increasing. Relatively early researches are as follows. Ding [37] applied ARCH (1) and ARCH(2) to analyse the ARCH phenomenon in the A-share of

Shanghai Stock market. Wei and Zhou [38] employed the GARCH model and GJR-GARCH model to estimate the Shanghai Stock Exchange Composite Index (SSE Index) and Shenzhen Composite Index (SZ index), and compared the predictive ability of the two models. Liu [39] analysed the volatility of returns of the SSE at different time intervals, and confirmed that under the assumption of normal distribution, volatility with a long interval has a greater actual value than theoretical value. By doing empirical research on a variety of indexes, Yue [40] found there exists GARCH effects in China's stock market. Liu and Cui [41] analysed both of the SSE index and SZ index from April 3, 1991 to June 29, 2001 by GARCH model, he found that these two stock markets of China presented spillover and leverage effect of volatility. Tian and Cao [42] studied the statistical characteristics of the SSE index, and used GJR-GARCH model to prove the existence of asymmetry phenomenon in Chinese stock markets. Kong [43] proved that the SSE 180 Index and SZ 180 Index both have strong ARCH effect, and he employed EGARCH (1, 1) to fit the characteristics of these two indexes. Wang [44] compared the accuracy of estimation between GARCH model and Risk Metrics model, when analysing the indexes of Shanghai Stock Exchange, and he concluded that GARCH model appeared to be more accurate. Tong [45] adopted the ARCH conditional mean equation to dynamically analyse the CSI 300 Index, and his results indicated ARMA (1, 1)–GARCH (1, 1) model fits the long-term fluctuations of the stock market better. Chen and Han [46] took advantage of the ARCH type model to conduct empirical research on the CSI 300 Index, analysing the fluctuation characteristics of Chinese stock market. The results showed that the volatility of daily return of CSI 300 Index presented apparent clustering and continuity but with no asymmetry, and the distribution of its data series presented features of 'sharp peak' and 'fat tail'.

Chen [47] tried to describe the basic statistical characteristics of China's stock market return series. She applied EGARCH (1, 1)-M model but with three different kinds of distributions (generalised error distribution (GED), t distribution and normal distribution) to measure the VAR value of daily returns series of CSI 300 Index. Through empirical research of statistical analysis and testing, the result reveals that the EGARCH (1, 1) with GED outperformed the other two models in characterising the market risk of China's stock market. Based on this analysis result, she concluded: Chinese stock market is of intense volatility, and the entire market appears significant periodicity. In addition, the daily return of the market has significant positive correlation with its volatility which has a strong feature of clustering. It can be seen from this that the speculative behaviour is serious in stock market, investors largely prefer short-term investment. Gu and Chen [48] applied 5 different GARCH type models and 2 SV models to characterise the statistical features of returns volatility of SSE Composite Index and SZ composite Index, and compared the predictive ability of these models. They found that GARCH (1, 1) and Component GARCH (1, 1) are able to give a more accurate prediction to the market volatility, and TARCH (1, 1) and EGARCH (1, 1) can better measure the asymmetry of volatility in these two markets. Therefore they concluded that GARCH type models comparing to SV models are more suitable for China's securities market, because of their relatively accurate reflection of the stock market volatility.

Sun, Qian and Han [49] collected the total monthly turnover of one of the futures markets in China from January 5, 2009 to June 10, 2011 as the sample data, and empirically analysed them using GARCH type models. They finally considered the TARCH (1, 1) is the best choice. Zhang and He [50] used the data of Growth Enterprise Market (GEM) and Shanghai and Shenzhen main board market from year 2010 to 2011, and employed DCC-MGARCH-VAR model to study the linkage between these markets and the spillover effect of the main board markets in China. The ADF test results of market return series of the main board markets and GEM are all significant at 1% level, which means they are all stationary series. Besides, the ARCH-LM test also further confirmed the heteroscedasticity of these series, and all the market return series appear the features of leptokurtosis. Moreover, their analysis results of GARCH (1, 1) reveal that the GEM is of the largest volatility, followed by the Shenzhen stock market and then the Shanghai stock market. The reason of this may due to the relatively high overall risk of GEM, and its relatively closer association with the Shenzhen stock market than Shanghai's. Yang and Liu [51] used the 5 min high-frequency transaction data of the CSI 300 Index to study the influence of macroeconomic information at different time periods to the fluctuations of CSI 300 Index. They found among the different kinds of information, the reserve requirement ratio (RRR) had the most significant impact on the stock market. Furthermore, the macro information has an asymmetry impact on the volatility of CSI 300 Index as well. That is the fact that volatility of CSI 300 Index in bull market is significantly larger than in the bear market.

2.3. Overview of Shanghai stock exchange

Located in Shanghai Pudong new district, Shanghai Stock Exchange is China's earliest and most influential stock exchange. It was established in November 26, 1990, and formally opened on December 19 that year. By the beginning of 2010, the Shanghai Stock Exchange contains 870 public companies, 1351 listed securities, this contributes to a total market capitalisation of stock of 18.465523 trillion yuan. And by 5.5 trillion US dollars' capitalisation, it has now been the 3rd largest stock market in the world as of May 2015 [52]. A large number of mainstay enterprises in the national economy, basic industries and high-tech companies pass through audit progress to get listed here, and this not only raises funds for the firms' development, but also helps their conversion of the operating mechanism.

Shanghai Stock Exchange Composite Index referred to as SSE Composite Index was formally released in 15th July 1991, base date points was set at 100 (finance.sina.com, 2015). Its constituent stocks are all listed shares, including A and B shares, which reflects the general movements of all the listed shares in Shanghai Stock Exchange. It has a meaning of weathervane to the entire China stock market to some extent.

2.4. Chapter summary

The recent 30 years is a phase of rapid development of GARCH type models. Scholars around the world have concluded the presence of heteroscedasticity in their application of GARCH type model to analyse stock prices, stock indexes etc. From Engle and Ng [18]'s research on the Japanese market and Gallant and Danielson [20]'s research on S&P 500 Index, to Hung [30]'s research on the asymmetry phenomenon of stock market volatility, they all proved the fact that the GARCH type models have a good simulation on stock volatilities. For the China's stock markets, Liu [39] have confirmed the spillover and leverage effect of them. Tong [45], Chen [47] and Sun, Qian and Han [49] all verified the good fitting and predictive ability of GARCH type models to the volatility of market returns in China's stock markets. Through the review of literatures, I found that GARCH (1, 1) is generally adept in forecast of unconditional volatility, TARCH (1, 1) and EGARCH (1, 1) generally have a good accuracy of measuring volatility asymmetry. Hence, this is the theoretical basis of this paper. I expect these three models can have effective estimation and prediction of the volatility of SSE composite Index. I wish my research can help investors to accurately determine market risk and conduct rational investment behaviour, and offer some theoretical support to the Chinese government's reform of the financial sector.

3. Methodology

This chapter focuses on the modelling approaches of volatility, and the basic characteristics of each model. In the early researches on stock market volatility, the classic econometric models are mostly based on the premises of independence between the disturbing items of returns rate and the invariance of variances. While with the deepening of financial theory and empirical research, people discover the clustering nature of stock market volatility, namely, big fluctuations are usually accompanied by big ones, and small fluctuations are often around the same extent ones. And the variance of volatility does not remain constant, but continually changing [53]. Since the static model (like standard deviation method) has congenital defects, and these defects become more prominent especially in the study of stock market volatility. Therefore, scholars were in increasing number trying to use new methods to investigate the problem of stock market volatility. Engle proposed the ARCH model in the year 1982, to describe the characteristics of volatility's variation. This model has later been widely applied to the analysis of financial time series. In 1986, Bollerslev extended this model, and developed it to a more generalised version, the GARCH model. On basis of the GARCH model, Engle, Lilien and Robins relaxed the condition that the mean of series should be independent with the variance, extending GARCH model to the GARCH-M model. These three models are built on the premise of symmetric risks and earnings, while in the actual market, the rate of return on asset of the stock market tends to be asymmetric. That is the degree of influence of good news and bad news on the asset yields in different [54]. To address this issue, Zakoian's TARCH model and Nelson's EGARCH model provide the corresponding solutions. Both models consider the time-varying feature of volatility variance, and they can more fully reflect the characteristics of the capital market. Thus they have become the major modelling tools for the research in financial econometrics area.

3.1. ARCH model

The basic idea of ARCH model is that at a certain point, under the collection of all information in the past, the occurrence value of a noise is normally distributed [13]. The distribution has a mean of 0. The variance of it is the linear combination of the square of limited noise values in the past (reflects autoregression), and it is a time-varying amount (reflects conditional heteroscedasticity). ARCH model has a basic form as follows:

$$\begin{cases} r_t = f(t, r_{t-1}, r_{t-2}, \ldots) + \varepsilon_t, \\ \varepsilon_t = \sqrt{h_t e_t}, \\ h_t = \omega + \sum_{j=1}^{q} \lambda_j \varepsilon_{t-j}^2. \end{cases} \tag{1}$$

This is the q order autoregressive conditional heteroscedasticity model, abbreviated as ARCH (q). Where the r_t refers to daily return of stock index; r_{t-i} is the i order lagged variable ($i = 1, 2 \ldots$); f is the autoregression function of r_t; $\{\varepsilon_t\}$ is a random error series, under the collection of all information in the past, it is subjects to a normal distribution with a mean of 0 and variance h_t (conditional variance); $\{e_t\}$ is a series of white noise, each e_t is independent and identically distributed, and subject to normal distribution; λ_i is the coefficient of conditional variance equation, $\lambda_0 > 0$, $\lambda_j > 0$ ($j = 1, 2 \ldots q$), $\sum_{j=1}^{q} \lambda_j < 1$ (ensure the stationary of the ARCH process); ω is a constant. Obviously, the ARCH (q) model is to modelling the random disturbance ε_t, in order to fully extract information from the residuals, so that the residuals e_t becomes a white noise series.

3.2. Symmetric GARCH model

ARCH model provides a brief and efficient way of analysis in describing the changeable volatility of financial time series. Instead of assuming a constant variance like other traditional econometric models, ARCH treats conditional variance as transformable with time. While in practice, 'the application of ARCH model often requires a larger order, which will not only increase the number of parameters to be estimated, but also lead to other problem, such as the multicollinearity of explanatory variables' [55]. To address the existing problems in ARCH model, Bollerslev extended it by adding an autoregressive term to get the GARCH model.

The simplest GARCH model is the GARCH (1, 1):

$$R_t = a_0 + a_1 R_{t-1} + \varepsilon_t \tag{2}$$

$$\varepsilon_t \sim i.i.d.\ N\left(0, \sigma_t^2\right)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-j}^2. \tag{3}$$

Eq. (2) is a mean equation of exogenous variables and mean of error. Eq. (3) gives a function of (1) a constant, (2) volatility of the last period ε_{t-1}^2 (ARCH part), (3) prediction variance from the last period σ_{t-1}^2 (GARCH part). σ_t^2 is a prediction variance based on the information of previous periods. Besides, the first '1' in the brackets means the GARCH part with order 1, the second '1' in brackets means ARCH part with order 1. A special form of GARCH model is the ordinary ARCH model.

The higher order of GARCH model can be denoted as GARCH (p, q), where the 'p' or 'q' is the order greater than 1. The variance equation of it can be expressed as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{P} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{4}$$

where p, q are the order of GARCH part and ARCH part respectively ($p \geq 0$, $q > 0$). $\omega > 0$, $\alpha_i \geq 0$ ($i = 1, 2, \ldots, q$), $\beta_j \geq 0$ ($j = 1, 2, \ldots, p$) and $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j$ is the attenuation coefficient, which reflects the volatility persistence. When $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$, GARCH process is stationary.

In general, the GARCH (p, q) model has been able to simulate the asset yield sequence and implied volatility sequence. However, they fail to take into account the long-term memory and asymmetry of volatility, and long-term memory is widespread within the stock and foreign exchange market. Asymmetry is also an important feature in stock market volatility.

3.3. ARCH-M model

If introducing the conditional variance into Eq. (2), we can get the ARCH-M model [56]:

$$R_t = a_0 + a_1 R_{t-1} + \beta \sigma_t^2 + \varepsilon_t. \tag{5}$$

ARCH-M model is usually applied in the research about relation between expected return and expected risk of capital. The estimated coefficient of the expected risk is a trade-off between risk and return. If the conditional variance is replaced by conditional standard deviation, we can get another form of ARCH-M model.

3.4. Asymmetric GARCH model

The upward movement in asset prices is often accompanied by a stronger degree of downward movement, this is a common phenomenon in the financial markets. To explain this phenomenon, Engle and Ng [18] drew asymmetric information curves under the influence of good news and under the influence of bad news respectively. The asymmetric effect is the basic manifestation of the market's reaction to shocks. It is also known as 'leverage effect', which is an important characteristic of many financial assets. In the capital market, market analysts often find that the stock price movement also exists the asymmetric effect, which is the fact that when a stock suffers an impact of negative shocks, its volatility is much fiercer than that caused by positive shocks. Since a significant decline in the stock price reduces the interests of shareholders, increasing the risk of holding this company's stock. TARCH and EGARCH are the main two models describing such asymmetric shocks.

(1) EGARCH model

Linear GARCH model assumes that the positive shocks and negative shocks with equal absolute values cause the same degree of stock price fluctuations, namely the conditional variances are the same. However, in reality, especially in the financial markets, the positive and negative shocks with equal absolute values often cause different degrees of fluctuations. Since in the stock market, the extent of share price decline tends to be much larger than the share price rise, and the fall process seems to be more violent and volatile. Hence, the asymmetry effect in stock market volatility has been beyond the explanatory power of the linear GARCH models. Nelson [57] proposed Exponential GARCH model, namely EGARCH model, on basis of the GARCH model, he improved the model to:

$$R_t = \beta_0 + \beta_1 R_{t-1} + \varepsilon_t \tag{6}$$

$$\ln\left(\sigma_t^2\right) = \alpha_0 + \sum_{i=1}^{q}\left(\alpha_i\left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \gamma_i\left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right)\right) + \sum_{j=1}^{p}\left(\beta_j \ln \sigma_{t-j}^2\right). \tag{7}$$

The left hand side of the equation $\log \sigma_t^2$ is the logarithmic form of the conditional variance, which means the impact of leverage effect is not quadratic but exponential, so that the predicted value of the conditional variance must be non-negative. We can test for leverage effect by examining the coefficient γ. If $\gamma_1 = \gamma_2 = \ldots = 0$, then the stock price's response to news impact does not have symmetric effect; if $\gamma_j < 0$, then asymmetric effect is existed or the impact of bad news to the market is greater than the good news; if $\gamma_j > 0$, then asymmetric effect is existed, but the impact of bad news is smaller than the good ones' [55]. The advantage of this model is that the data of conditional variance σ_t^2 take the form of logarithm. Therefore, it does not require any restrictions to the other parameters of the formula, which makes the solving process more simple and flexible.

The EGARCH (1, 1) has the following form:

$$\ln\left(\sigma_t^2\right) = \alpha_0 + \beta \ln\left(\sigma_{t-1}^2\right) + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{8}$$

where, σ_{t-1}^2 = previous period's variance estimation, it scales how persistent the conditional variance of the previous period is; $\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|$ = impact of volatility from the last period; $\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ = the effect of leverage and asymmetry.

In addition, Engle and Ng [18] pointed that, when the volatility suddenly appears and then quickly disappears, the predictive ability of EGARCH model will be greatly reduced.

(2) TARCH model

Threshold GARCH model was introduced by Zakoian [17] and Glosoten, Jafanathan and Runkle [58], which is designed to detect the leverage effect in the financial market. To achieve this, simply by adding a multiplicative dummy variable into the variance equation to check that when shocks are negative whether there exists statistically significant difference. And it is under the assumption that unanticipated information shocks can affect the volatility of stock returns. The form of the model can be specified as:

$$R_t = a + bR_{t-1} + \varepsilon_t \tag{9}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}\left(\alpha_i\varepsilon_{t-i}^2\right) + \gamma\varepsilon_{t-i}^2 d_{t-1} + \sum_{j=1}^{p}\beta_i\sigma_{t-j}^2. \tag{10}$$

When $\varepsilon_t < 0$, $d_t = 1$; otherwise $d_t = 0$. The coefficient γ represents the impact of positive shocks, and the coefficient $\gamma + \alpha$ represents the impact of negative shocks. Therefore, when α is greater than 0, the impact of the 'bad news' is greater than 'good news', we get asymmetric effect. And when α equals to 0, there would be no leverage effect. Besides, TGARCH (1, 1) has the form of:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-i}^2 d_{t-1} + \beta_1\sigma_{t-1}^2. \tag{11}$$

In addition, the optimal volatility estimation model for the SSE composite index should pass the Ljung–Box Q test and the ARCH-LM test. And this is on the basis of minimum Akaike Information Criteria (AIC) and minimum Schwarz Information Criteria (SIC).

3.5. Data collection

The data for this study are the daily closing prices of SSE composite Index over the period extending from July 26, 2013 to July 28, 2017, making total observations of 978 excluding public holidays. And the last week data of July 2017 will be used to exam the prediction results of the three models. Besides these data series are collected from Wind database.

The stock market returns in this paper are accrued as compounded returns:

$$R_i = \ln\left(P_i/P_{i-1}\right) \tag{12}$$

where P_i is the daily closing index of SSE composite index at days t, and P_{i-1} is the index at days $t-1$.

4. Empirical analysis

4.1. Descriptive statistical analysis

Before processing the data, it is necessary to have an overview of the basic statistical features of the data series. Fig. 2 is a draw of the daily return of SSE Index. Besides, in this paper, the statistical software Eviews 8.0 is applied to each disposal step.

As can be seen from the figure above, we can see a slight presence of time trend in the return series. And it shows apparent features of time-varying variance and clustering. Hence, the traditional conditional variance models with assumption of homoscedasticity are no longer suitable for fitting volatility of the SSE Composite Index. Instead, the popularly used GARCH models may properly do this job, since they are capable of dealing time series with heteroscedasticity and clustering (see Fig. 3).

From these statistics we can see that the mean of return series is 0.0492%, nearly 0, which is not surprising, since stock return series usually have a regressive tendency towards a long term value. And the gap between the maximum and minimum value is 0.1447, and standard deviation has a value around 1.6%, both of these implies relatively high volatility in a stock market during sample period. The negative value of skewness may due to the negative inclination of the asymmetric tail. The kurtosis value (9.816282) is much greater than the standard normal distribution value (+3), this reveals that the distribution of SSE has characteristics of 'sharp peak' and 'fat tail'. And its J–B statistic is 2151.707, which is much higher than the J–B value of standard normal distribution (5.8825), therefore we reject the null hypothesis that return series is subject to normal distribution, namely, it obeys the skewed distribution.

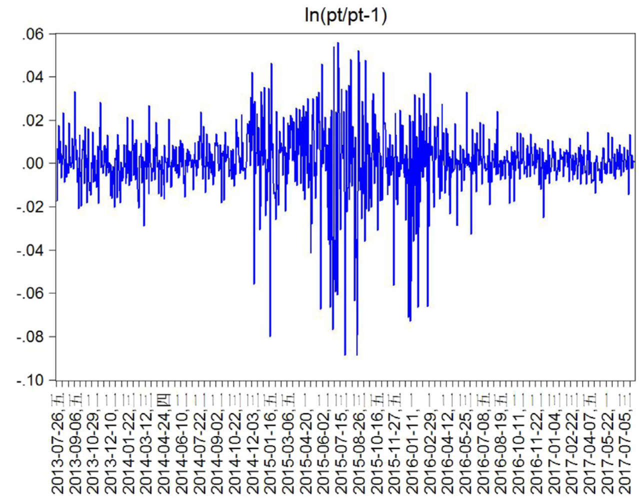

Fig. 2. Line chart of daily SSE composite index returns.

Table 1
Autocorrelation and partial correlation of return series.

Lags	AC	PAC	Q-Stat	Prob
1	0.073	0.073	5.2273	0.022
2	−0.051	−0.056	7.7636	0.021
3	−0.020	−0.012	8.1375	0.043
4	0.132	0.132	25.170	0.000
5	0.005	−0.018	25.193	0.000
6	−0.079	−0.067	31.397	0.000
7	0.019	0.037	31.759	0.000
8	0.087	0.060	39.172	0.000
9	0.043	0.032	40.975	0.000
10	−0.094	−0.077	49.784	0.000
11	−0.036	−0.025	51.080	0.000
12	0.009	−0.014	51.151	0.000
13	0.119	0.114	65.269	0.000
14	−0.110	−0.105	77.314	0.000
15	0.023	0.059	77.840	0.000
16	0.070	0.046	82.665	0.000
17	0.047	0.004	84.827	0.000
18	−0.012	0.027	84.980	0.000
19	−0.038	−0.023	86.448	0.000
20	0.121	0.098	101.08	0.000

4.2. Data processing

Next, we shall study the stationarity of the return series. And we can start from the autocorrelogram and Q statistic, which are shown in the figure below (see Table 1).

It can be seen at first glance that, the accompanying p values of every Q statistic are larger than 5% significance level, thus we can judge that the return series has no problem of autocorrelation, and then it may possess a problem of white noise, which means return series is currently no suit for building the GARCH type models. Therefore we take the first-order difference on it, and define a new series DR $= d(R)$ to describe the fluctuations. Then, by drawing its correlogram, we can further determine the stationarity of the differenced series.

From Table 2, we notice that the rate on which the autocorrelation function of series DR decays is relatively fast, it approaches 0 at period 11. Therefore, it can be initially judged that series DR is stationary, and it does not possess white noise. We can further confirm the stationarity of series DR by unit root test.

Table 2
Autocorrelation and partial correlation of series DR.

Lags	AC	PAC	Q-Stat	Prob
1	−0.433	−0.433	183.15	0.000
2	−0.084	−0.333	190.07	0.000
3	−0.064	−0.355	194.09	0.000
4	0.149	−0.149	215.98	0.000
5	−0.023	−0.087	216.48	0.000
6	−0.099	−0.173	226.10	0.000
7	0.016	−0.169	226.36	0.000
8	0.060	−0.121	229.96	0.000
9	0.050	−0.008	232.47	0.000
10	−0.105	−0.061	243.29	0.000
11	0.006	−0.067	243.33	0.000
12	−0.036	−0.180	244.61	0.000
13	0.183	0.042	277.94	0.000
14	−0.196	−0.121	315.90	0.000
15	0.047	−0.097	318.12	0.000
16	0.037	−0.049	319.49	0.000
17	0.020	−0.068	319.87	0.000
18	−0.018	−0.017	320.19	0.000
19	−0.100	−0.133	330.19	0.000
20	0.114	−0.065	343.22	0.000

4.3. Unit root test

In the tests for stationarity, the most commonly used method is the unit root test proposed by two American statistician D.A. Dickey and W.A. Fuller in the 1970s, which judges whether the autocorrelation coefficient is equal to 1. After nearly 30 years' research, this approach was eventually summarised as the Augmented Dickey–Fuller (ADF) Test (see Table 3). Here we use the ADF test to examine the stationarity of DR series and result are given below:

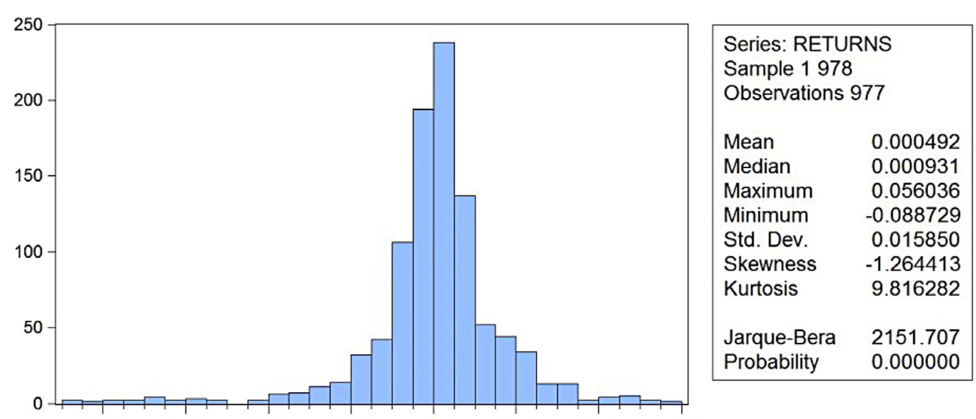

Fig. 3. Descriptive statistics of SSE composite index returns.

Table 3
Result of ADF test.

Variable	ADF	Prob.*	1% level	5% level	10% level
DR	−29.02604	0.000	−3.437	−2.864	−2.568

Notes:
* Denotes significance at 1% level.

As shown in the table, the ADF value of series DR is −29.02604, which is much smaller than 1% critical value, and the accompanying P value is 0, which also indicates a significance in 1% level. Therefore, we reject the null hypothesis that series DR has a unit root, namely, DR is stationary. Now we can use series DR to build the GARCH models.

4.4. ARCH effect test

Before processing to modelling, we need to confirm the presence of heteroscedasticity in data series, and this can be achieved by doing the ARCH effect test. Financial assets price or stock index series and other high-frequency data will often appear the feature that a large fluctuation usually followed by another large fluctuation, and a small fluctuation usually followed by another even smaller fluctuation, this is called the ARCH effect. And ARCH test refers to the procedure of examining whether the data series has the characteristics of ARCH effect. In this paper, we adopt the most common used method—the Lagrange Multiplier (LM) Test, which was proposed by Engle [12], to test ARCH effect of series DR. ARCH-LM test mainly use the output of F statistic and chi-square statistic to determine the presence of ARCH effect. The null hypothesis of ARCH-LM test is that there is no ARCH effect in the data series, if the value of F statistic or chi-square statistic is less than 5% critical value, then the null should be rejected; if the value of F statistic or chi-square statistic is greater than 5% critical value, then we accept the null.

(1) Selection of lag order and determination of mean equation

Since we employ the time series model to do modelling and estimation, we need firstly to decide the lag order of autoregression of series DR. The mean equation of the series DR takes the following form:

$$\mathrm{DR}_t = a_0 + a_i DR_{t-i} + \varepsilon_t. \tag{13}$$

We do regression on the 1, 2, and 3 lags respectively. (See Table 4).

According to AIC criteria and the judgement principle of F statistics, we choose the lag order with smallest AIC value and the largest F value, which points to lag order 1. Hence, the mean equation of market returns is transformed into:

$$\mathrm{DR}_t = a_0 + a_1 DR_{t-1} + \varepsilon_t. \tag{14}$$

Specifically, substituting the coefficient value of regression result of lag 1, the mean equation of series DR can be written as:

$$DR = 0.000475 - 0.073034 * DR(-1). \tag{15}$$

(2) Residuals autocorrelation test

Before doing the ARCH-LM test, we need to determine the lag order of it. And this can be achieved by doing the autocorrelation test on the squared residuals of series DR using ARCH (1) model. Table 5 lists the value of autocorrelation, partial correlation, Q statistic and their corresponding probability.

Table 5 reveals that the PAC value appears truncation at lag order 5 in the partial correlation function plot (partial correlation is within critical value), which indicates that the preferable lag order for ARCH-LM test is 5. Moreover, the Q statistic of the squared residuals of series DR (21.877) is far larger than the 5% critical value (7.81) and the P value of all lags are significant at 1% critical level. Therefore we reject the null hypothesis that the residuals of data series are independently distributed, and we can initially judge that the original data series has ARCH effect.

(3) ARCH-LM test

The judgement about ARCH effect above requires further confirmation. Accordingly we conduct the ARCH-LM test on residuals series of DR. (See Table 6).

In our case, the F statistic test the joint significance of all lagged squared residuals of series DR. The statistic of Obs $*R$-squared is the statistic of LM test, and it is calculated by number of observations times the value of R-squared. Given the significance level of 5% and 5 degrees of freedom, the accompanied probability of LM value is 0.0000, therefore we reject the null hypothesis. This indicates there is an obvious heteroscedasticity in the series DR, and its residuals have a strong ARCH effect. This result further illustrates that it is reasonable to use GARCH models to fit the volatility of SSE Composite Index returns.

4.5. Estimation of GARCH (1, 1) model

According to the result of Table 7, the mean equation of GARCH (1, 1) model would be:

$$\begin{aligned} DR_t &= 0.052126 DR_{t-1} + 0.000473 \\ &\quad (1.522721) \quad (1.571959). \end{aligned} \tag{16}$$

And the variance equation would be:

$$\begin{aligned} \sigma_t^2 &= 3.88\mathrm{E}-07 + 0.060276\varepsilon_{t-1}^2 + 0.937721\sigma_{t-1}^2 \\ &\quad (1.658035) \quad (8.991761) \quad (176.2655). \end{aligned} \tag{17}$$

From the estimation result, we can know that in the mean equation, the independent variable has passed the t significance test (p value is significant at 1% level, and z statistic is greater than 5% critical value 1.96). And in the variance equation, both the ARCH part and GARCH part are significant, which means series DR is of an evident feature of clustering. Moreover, the sum of the coefficients of these two parts is 0.997997, which is smaller than 1 but rather approaching 1. This is in line with constrains of the parameters, it also proves that the impact of the conditional is not ephemeral, but a persistent process. Moreover, the value of R-squared is 49.03%, which suggests that the fitting effect from a perspective of whole model is acceptably good. Besides, the significant coefficients of the GARCH part in the variance is estimated to be 0.937721, which reflects the positive correlation between income benefits and risks, proving the existence of positive risk premium.

4.6. Estimation of TARCH (1, 1) model

As introduced before in the methodology chapter, the TARCH model is a Threshold GARCH model, which uses a dummy variable to set a threshold to distinguish the impact of positive and negative on volatility. The estimation results are shown as follows.

According to the result of Table 8, the mean equation of TARCH (1, 1) model would be:

$$\begin{aligned} DR_t &= 0.048343 DR_{t-1} - 0.000621 \\ &\quad (14.08837) \quad (2.046311). \end{aligned} \tag{18}$$

And the variance equation would be:

$$\begin{aligned} \sigma_t^2 &= 2.25\mathrm{E}-07 + 0.073221\varepsilon_{t-1}^2 + 0.035982\varepsilon_{t-1}^2 * \varepsilon_{t-1} \\ &\quad + 0.946804\sigma_{t-1}^2 \\ &\quad (1.118026) \quad (7.602650) \quad (3.244216) \quad (184.2233). \end{aligned} \tag{19}$$

Table 4
Selection of autoregression lag order.

Lag order	AIC	F statistics
1	−5.453803	5.229912
2	−5.450180	2.525857
3	−5.446946	0.370768

Table 5
Correlogram of ARCH (1).

Lags	AC	PAC	Q-Stat	Prob
1	0.005	0.005	0.0216	0.883
2	−0.055	−0.055	3.0216	0.221
3	−0.024	−0.024	3.6079	0.307
4	0.134	0.132	21.877	0.000
5	0.002	−0.002	21.291	0.001
6	−0.082	−0.070	27.877	0.000
7	0.018	0.027	28.207	0.000
8	0.083	0.060	35.016	0.000
9	0.044	0.042	36.930	0.000
10	−0.094	−0.071	45.719	0.000
11	−0.0030	−0.029	46.626	0.000
12	0.002	−0.025	46.631	0.000
13	0.128	0.119	62.754	0.000
14	−0.122	−0.102	77.612	0.000
15	0.027	0.049	78.328	0.000
16	0.065	0.048	82.342	0.000
17	0.043	0.006	84.342	0.000
18	−0.013	0.028	84.519	0.000
19	−0.047	−0.029	86.738	0.000
20	0.121	0.095	101.47	0.000

Table 6
Result of ARCH-LM test.

F-statistic	36.25717	Prob. F (5965)	0.0000
Obs $*R$-squared	153.5643	Prob. Chi-square (5)	0.0000

From the estimation result, we can know that in the mean equation, the independent variable has passed the t significance test, which indicates a good fitting effect of the mean equation. And from the perspective of whole model, the R-squared value of 46.58% suggests an acceptably good fitting effect. The parameter estimates are all significant except for the cross term, which indicates an evident clustering characteristic of series DR. In addition, the term 'RESID$(-1)^2 * $(RESID$(-1) < 0$)' describes the result of leverage effect, and the coefficient of it (0.035982) is greater than 0, therefore the asymmetric impact exists in the return series. Therefore the volatility of SSE Composite Index has leverage effect, or in other words, fluctuations caused by bad news are greater than the ones caused by the same level of good news in the Shanghai Stock Exchange. When a good news shows up, namely when $\varepsilon_t > 0$, it has a shock degree of $\alpha = 0.073221$; when a bad news shows up, namely when $\varepsilon_t < 0$, it has a shock degree of $\alpha + \gamma = 0.073221 + 0.035982 = 0.109203$. Moreover, the GARCH part in the variance equation has a coefficient of 0.946804, which reflects the existence of positive risk premium.

4.7. Estimation of EGARCH (1, 1) model

EGARCH model refers to the Exponential GARCH model, which applies the ratio of disturbance over its standard deviation in the variance equation to capture the impact of positive or negative shock to the volatility. The estimation results are shown as follows (see Table 9).

According to the result of Table 9, the mean equation of EGARCH (1, 1) model would be:

$$DR_t = 0.027175 DR_{t-1} + 0.000694$$
$$(0.801512) \quad (2.378071). \tag{20}$$

And the variance equation would be:

$$\ln\left(\sigma_t^2\right) = -0.073545 + 0.097720 \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| - 0.033839 \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) + 0.999877 \ln\left(\sigma_{t-1}^2\right)$$
$$(-6.410407) \quad (9.624864) \quad (4.186649) \quad (806.3409). \tag{21}$$

From the estimation result, we can know that in the mean equation, the independent variable has passed the t significance test, which indicates a good fitting effect of the mean equation. While from the perspective of whole model, the R-squared value of 30.81% suggests that the fitting effect is not quite ideal. The parameter estimates are all significant, which indicates an evident clustering characteristic of series DR. Since the parameter γ in conditional variance equation of EGARCH (1,1) represents the degree of asymmetric effect, and the coefficient value of it is significant, we can judge from this that there is apparent leverage effect in the SSE Composite Index. Specifically, when a good news shows up, namely when $\varepsilon_t > 0$, it has a shock degree of $\alpha + \gamma = 0.097720 + (-0.033839) = 0.063881$; when a bad news shows up, namely when $\varepsilon_t < 0$, it has a shock degree of $\alpha + (-\gamma) = 0.097720 + 0.033839 = 0.131559$.

4.8. Forecast result

Based on the empirical analysis of GARCH (1, 1), TARCH (1, 1) and EGARCH (1, 1), we now employ them to forecast the last week's daily closing prices of July 2017 of SSE Composite Index, and then compare them with the actual values, in order to give an intuitive view at these three models for their predictive ability upon SSE Index (see Table 10).

As shown in the table, the forecast values of these three models are all actually quite close to the actual ones, and among them, the values from EGARCH (1, 1) are the most outstanding. Besides, the values from GARCH (1, 1) and TARCH (1, 1) are mutually close.

Table 7
Estimates of GARCH (1, 1) model.

Variable	Coefficient	Std. error	z-statistic	Prob.
C	0.000473	0.000301	1.571959	0.1160
DR(−1)	0.052126	0.034232	1.522721	0.1278
Variance equation				
C	3.88E−07	2.34E−07	1.658035	0.0973
RESID$(1)^2$	0.060276	0.006703	8.991761	0.0000
GARCH(−1)	0.937721	0.005332	176.2655	0.0000
R-squared	0.4903	Mean dependent var		0.000510
Adjusted R-squared	0.3881	S.D. dependent var		0.015848
S.E. of regression	0.015817	Akaike info criterion		−6.025824
Sum squared resid	0.243686	Schwarz criterion		−6.000806
Log likelihood	2945.602	Hannan–Quinn criter.		−6.016304
Durbin–Watson stat	1.950892			

Table 8
Estimates of TARCH (1, 1) model.

Variable	Coefficient	Std. error	z-statistic	Prob.
C	−0.000621	0.000303	2.046311	0.0407
DR(−1)	0.048343	0.034314	14.08837	0.1589
Variance equation				
C	2.25E−07	2.01E−07	1.118026	0.2636
RESID(−1)^2	0.073221	0.009631	7.602650	0.0000
RESID(−1)^2∗ (RESID(−1) <0)	0.035982	0.011091	3.244216	0.0012
GARCH(−1)	0.946804	0.005139	184.2233	0.0000
R-squared	0.4658	Mean dependent var		0.000510
Adjusted R-squared	0.3637	S.D. dependent var		0.015848
S.E. of regression	0.015819	Akaike info criterion		−6.030218
Sum squared resid	0.243746	Schwarz criterion		−6.000197
Log likelihood	2948.747	Hannan–Quinn criter.		−6018795
Durbin–Watson stat	2.332807			

Table 9
Estimates of EGARCH (1, 1) model.

Variable	Coefficient	Std. error	z-statistic	Prob.
C	0.000694	0.000292	2.378071	0.0174
DR(−1)	0.027175	0.033905	0.801512	0.4228
Variance equation				
C(3)	−0.073545	0.011473	−6.410407	0.0000
C(4)	0.097720	0.010153	9.624864	0.0000
C(5)	0.033839	0.008083	4.186649	0.0000
C(6)	0.999877	0.001240	806.3409	0.0000
R-squared	0.3081	Mean dependent var		0.000510
Adjusted R-squared	0.2057	S.D. dependent var		0.015848
S.E. of regression	0.015832	Akaike info criterion		−6.030580
Sum squared resid	0.244132	Schwarz criterion		−6.000559
Log likelihood	2948.923	Hannan–Quinn criter.		−6.019157
Durbin–Watson stat	1.903482			

Table 10
The forecast value and actual value of SSE composite index.

Date	Actual value	GARCH	TARCH	EGARCH
2017-7-24	3250.60	3143.457	3194.745	3234.753
2017-7-25	3243.69	3157.635	3193.432	3224.342
2017-7-26	3247.67	3146.247	3187.634	3250.451
2017-7-27	3249.78	3146.453	3189.324	3247.643
2017-7-28	3253.24	3157.657	3212.387	3261.498

4.9. Results compare with ARIMA model

From the recent literature, ARIMA model is widely considered to forecast asset return rates, therefore, out of comprehensiveness and prudence, this section provides the forecast results of SSE Composite Index generated by ARIMA model for a comparison and reference. The basic idea of this model is to use the sequence formed by changing the predictive variables over time as a random sequence, and then describe the random sequence with a specific mathematical model based on the autocorrelation of the time series.

The Model structure of ARIMA (p, d, q) can be expressed as:

$$\begin{cases} \Phi\ (B)\ \nabla^d x_t = \Theta\ (B)\ \varepsilon_t \\ E\ (\varepsilon_t) = 0, Var\ (\varepsilon_t) = \sigma_\varepsilon^2, E\ (\varepsilon_t \varepsilon_s) = 0, s \neq t \\ Ex_t \varepsilon_t = 0, \forall s < t \end{cases} \tag{22}$$

where, $\Phi\ (B) = 1 - \phi_1 B - \cdots \phi_p B^p$ is the autoregressive coefficient polynomials of ARIMA (p, d, q) model, and $\Theta\ (B) = 1 - \theta_1 B - \ldots \theta_q B^q$ is the moving smoothness factor polynomial.

First, it is necessary to compare the information criterion to determine the better fitting model. Table 11 is the AIC and SC value of lag order p and q of ARIMA.

As shown in Table 11, according to the AIC and SC information minimum criteria, $p = 2$ and $q = 2$ should be chosen as the lag order. Therefore, it can be initially determined the fitting model, ARIMA (2, 1, 2). Next, parameters should be estimated.

Table 11
Comparison of fitting effect of return of SSE composite index.

Model	AIC	SC
AR (1)	−5.453803	−5.443796
AR (2)	−5.454173	−5.439151
AR (3)	−5.451221	−5.431174
MA (1)	−5.454160	−5.444161
MA (2)	−5.453797	−5.438798
MA (3)	−5.453542	−5.433544
ARMA (1, 1)	−5.453365	−5.438798
ARMA (1, 2)	−5.452148	−5.438354
ARMA (1, 3)	−5.452147	−5.432116
ARMA (2, 1)	−5.452147	−5.459914
ARMA (2, 2)	−5.488389	−5.459914
ARMA (2, 3)	−5.488235	−5.459914
ARMA (3, 1)	−5.453552	−5.428493
ARMA (3, 2)	−5.484952	−5.458318
ARMA (3, 3)	−5.459460	−5.424378

Table 12
Regression Result of ARIMA (2, 1, 2).

Variable	Constant	Standard deviation	T value	P value
C	0.000501	0.000487	1.029664	0.3034
AR (1)	0.66003	0.023438	2.594627	0.0096
AR (2)	−0.946085	0.025235	−37.49056	0.0000
MA (1)	−0.080242	0.037006	−2.168345	0.0346
MA (2)	0.881241	0.036904	23.87945	0.0000

From the regression result on Table 12, it can be seen that the variable coefficients all pass the significance test, therefore the specific form of ARIMA (2, 1, 2) is:

$$R_t = 0.000501 + 0.066003 R_{t-1} - 0.946085 R_{t-2} - 0.080242 \varepsilon_{t-1} + 0.881241 \varepsilon_{t-2} + \varepsilon_t. \tag{23}$$

Apply the ARIMA (2, 1, 2) model to forecast the last week's daily closing prices of July 2017 of SSE Composite Index, it is obviously that GARCH models are much suitable in forecasting the SSE Composite Index.

Date	Actual value	ARIMA
2017-7-24	3250.60	3032.43
2017-7-25	3243.69	3053.25
2017-7-26	3247.67	3070.69
2017-7-27	3249.78	3085.97
2017-7-28	3253.24	3102.65

5. Summary

Based on the analyses above, no matter it is from the fitting effect or the degree of estimation accuracy, GARCH type models can appropriately adapt to the volatility of SSE Composite Index. Moreover, both the GARCH (1, 1) model (representing the symmetric model) and TARCH (1, 1) and EGARCH (1, 1) (representing the asymmetric model) perform good estimation in our case. While specifically, since the returns series of SSE Composite Index appeared apparent asymmetric effect as analysed above, the asymmetric GARCH models captured more details. As we can see that, though the fitting effects of these models are about the same, but the asymmetric models overall outperformed the symmetric one in forecast results. Moreover, the asymmetric models can roughly measure the impact of the positive or negative shocks, and the GARCH (1, 1) failed to achieve this point. In addition, among the two asymmetric models, EGARCH model seems to be superior to TARCH model, this is according directly to the forecast results of them. Besides, for the consideration of comprehensiveness and prudence, in the last sub-section, ARIMA model is applied to forecast the return of SSE Composite Index, which turns out that GARCH models are generally much competent. Furthermore, the SSE Composite Index reflects the overall situation of the market fluctuations of Shanghai Stock Exchange, and the Shanghai stock market as one of the two primary stock markets in China, can to some extent reflects the whole situation of China's stock markets. Therefore, the results we have got in this paper are also capable of applying to the entire Chines stock market to some extent.

6. Conclusions and suggestions

6.1. Conclusions

Through analysis all above, we can draw the following conclusions

(1) GARCH models can be applied to China's stock markets. Although these models were developed and widely used in the process of researching financial markets in developed countries, it does not obstruct the use of them in China, such an emerging stock market. Besides, we found obvious heteroscedasticity in the volatility of Shanghai Composite Index. And the GARCH models reflect the change rule of volatility in China's stock market in a high accuracy.

(2) From a perspective of time series, the volatility of Shanghai Composite Index shows features of significant time-varying and clustering. Clustering is due to some related factors, for example, the impact of information shock on the stock market shows a phenomenon that large fluctuations tend to be followed by relatively large ones, while smaller fluctuations will follow smaller ones.

(3) High liquidity is an essential reason for the high volatility of the index. In Shanghai stock market, for public investors, holding the share for five days can be seen as a long-term operation. If investors shorten the holding period is bound to make transactions frequent. A large number of participants of stock index or futures are 'day traders', they usually buy and sell frequently in the same day, and do not hold a contract overnight. Frequent trading on the SSE Composite Index makes the stock index market in a high degree liquidity, this is also one of the major contributors to the volatility of stock index.

(4) The SSE Composite Index has evident GARCH effect, indicating that returns have a positive risk premium, namely, there is positive correlation between daily returns and volatility in Shanghai stock market. Significant volatility clustering feature shows that the short-term fluctuations are dominant in China's stock market. The entire market are covered by thick 'speculative' atmosphere, short-term investment preference prevails. On the one hand, this is a manifestation of a market which is distinctly policy-oriented and without mature financial system. On the other hand, the investment consciousness of Chinese public investors are still remaining in an elementary level, which has not yet extended to 'value investment'.

(5) By establishing the TARCH and EGARCH model, we found that the returns of SSE Composite Index exists significant leverage effect. The same degree of 'bad news' has a much stronger influence on the SSE Composite Index than the good news'. The SSE Composite Index implements the T+0 trading mechanism, and investors can do short or long trade on the contracts. This is easy to arouse investors' irrational speculation desire. In theoretical analysis, it usually defines the investors as risk averse, when his/her gains are positive, the aversion may be reduced, and then a certain amount of holdings will be followed up. Correspondingly, if the gains are negative, his/her risk aversion will increase, and in order to avoid further loss, he/she will sell a certain amount of shares. Whereas if investors react slow in the process of price increase, but overreact in the down process, will further amplify the risk in the market. This may be the reason for irrational investment to cause asymmetry in market volatility. Meanwhile, since the risk management system in China is not perfect, a number of big speculators are easy to be motivated by their own interests, trying to use their power to manipulate market illegally.

6.2. Causes analysis

Through the empirical analysis of volatility SSE Composite Index, it is found that the EGARCH (1, 1) and TARCH (1, 1) models can basically fit the time series of return of SSE Composite Index. However, in the process of modelling, it has disclosed that the volatility of Shanghai stock market has a high sustainability. When the securities yield appears abnormal fluctuation when shocked by significant 'good news' or especially 'bad news', it is difficult to eliminate such impact in a short time. Hence, it can be seen from an aspect that the overall risk of China's stock market is still great. Moreover, the volatility of the Shanghai Composite Index yield has a 'leverage effect', that is, bad news has much bigger impact on market volatility than good news. This shows that Chinese investors have not formed strong philosophy of investment, their investment behaviours are vulnerable to a variety of news. These characteristics may be derived from the following facts:

- Retail investors account for 80%–90% of the turnover of China's stock market, and many of them are new investors, who have rare experience of stock investment. Even the more professional fund managers, often conduct ultra-short-term trading, since their operating performance is measured on a monthly or quarterly basis.
- China's stock market is often referred to as the 'policy market', not only refers to monetary policy, but also related to securities and banking supervision. Even the official media remarks occupy a significant place as well.

- The use of borrowed funds to buy stock is common around the world, but financing transactions have surged in the past few years in China.

6.3. Suggestions

For a healthy stock market, the normal volatility is helpful in activating the market. Through the foregoing analyses we know that there exists many problems in the securities market. Therefore, based on the actual situation, I proposed the following recommendations in policy aspect.

(1) Improve the information disclosure system. China's stock market is of an evident characteristic of volatility clustering, a lot of irrational factors are inundated in the transactions, market price fluctuates intensely. The inefficient market information makes the discovery of stock value and resource optimisation of market difficult to be realised. By improving the information disclosure system, the market can fully play the role of resource allocation. A reasonable information transmission mechanism can alleviate the speculation atmosphere.

(2) Strengthen investment philosophy education. The majority retail investors are lack of appropriate knowledge and skills in securities investment. They prefer short-term investment, and often follow suit in trade operations. This would lead to a reduction of market's scientificity. China's stock market strongly requires a promotion of rational investment philosophy, so that more investors are able to make their own analysis on the market information.

(3) Strengthen the construction of the securities market. Government intervention in China's securities market is a bit of excessive. When the stock market's bubble is in a higher level, government will release some bad news for containment; when the stock market is in a downturn period, government will publish bailout information. Such direct intervention of administrative acts distort the balance of supply and demand, having much negative influences on the volatility. Government, as administrative departments should be clear of its responsibilities, focusing on the duties of supervision and penalties for violations. In this way can we create a fair and open market order.

(4) Strengthen the risk management system. Since the stock index market has an even higher risk than general stock markets, the establishment of effective risk management system in stock index market seems to be more urgent. Specifically, in addition to the deposit system, the market should construct the some supporting system including daily limit system for ups and downs, position reporting system, position limits system, forced liquidation system and forced lighten up system.

(5) Strengthen the law construction on the securities market. The launch of stock index in China is late, relevant laws and regulations are not perfect, while speculators are vulnerable to be lured by interests to commit illegal behaviour. These behaviours will seriously disrupted the market order, distorting the price discovery function and resource allocation function of the market. Hence, the China should vigorously strengthen legal construction on her securities market, and regulate the behaviour of parties of the transaction, improving the effectiveness of management.

Acknowledgements

I wish to specially acknowledge my supervisor Professor Fabio Spagnolo, for his precious and scrupulous suggestions and comments.

I would also like to be grateful to all my module teachers (my personal tutor Professor Menelaos Karanasos, Professor John Hunter, etc.) for their earnest guidance in my Masters study. And all my dearest friends who accompanied me through every library nights and gave me support to overcome difficulties. It is all because of these lovely people, I spent a wonderful year in London at Brunel University.

References

[1] L. Zhang, et al., Benefit analysis of china's stock market based on GARCH type models, Nat. Sci. Daily 17 (2012) 2.
[2] Y. Fang, Research of Chinese Stock Market Volatility–Based on Stochastic Volatility Model, Huazhong University of Science and Technology, 2010 (Ph.D. thesis).
[3] X. Xu, Y. Chen, Empirical study on nonlinearity in China stock market, Quant. Tech. Econ. 18 (3) (2010) 110–113.
[4] B. Mandelbrot, The variation of certain speculative prices, J. Bus. 36 (4) (1963) 394–419.
[5] W. Liu, B. Morley, Volatility forecasting in the hang seng index using the GARCH approach, Asia-Pac. Financ. Mark. 16 (1) (2009) 51–63.
[6] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, J. Econ. 31 (3) (1986) 307–327.
[7] W. Schwaiger, A Note on GARCH predictable variances and stock market efficiency, J. Bank. Finance 19 (5) (1995) 949–953.
[8] In. Kang, International foreign exchange agreements and nominal exchange rate volatility: A GARCH application, N. Am. J. Econom. Finance 10 (2) (1999) 453–472.
[9] A. Lehar, et al., GARCH VS stochastic volatility: Option pricing and risk management, J. Bank. Finance 26 (2–3) (2002) 323–345.
[10] J. Duan, H. Zhang, Pricing hang seng index option around the asian financial, Finance Econ. 25 (11) (2001) 1989–2014.
[11] B. Yan, H. Jiang, Research of Stock Index and Futures: ''Throw a Sprat to Catch a Herring''- Things to Know for Public Investors, Social Sciences Academic Press, Beijing, 2008.
[12] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, Econometrica 50 (4) (1982) 987–1008.
[13] C. Wang, et al., Financial Econometrics Based on EVIEWS, Renmin University of China Press, Beijing, 2010.
[14] F. Black, Studies in stock price volatility changes, in: Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section, American Statistical Association, 1976, pp. 177–181.
[15] K. French, et al., Expected stock returns and volatility, J. Financ. Econ. 1987 (19) (1987) 3–29.
[16] D. Nelson, ARCH Models as diffusion approximations, J. Econometrics 1990 (45) (1990) 307–381.
[17] J.M. Zakoian, Threshold heteroscedastic models, J. Econom. Dynam. Control 1990 (18) (1990) 931–944.
[18] R. Engle, K. Ng, Measuring and testing the impact of news on volatility, J. Finance 48 (5) (1993) 1022–1082.
[19] Y. Yeh, T. Lee, Interaction and volatility asymmetry of unexpected returns in the greater china stock markets, Glob. Finance J. 2000 (11) (2000) 129–149.
[20] J. Danielson, Stochastic volatility in asset prices: Estimation with stimulated maximum likehood, Econometrics 1994 (64) (1994) 375–400.
[21] Y. Cheung, J. He, K. Ng, Common Predictable Components in Regional Stock Markets, University of California Santa Cruz, 1995 (Ph.D. thesis).
[22] D. Su, B. Fleishe, An empirical investigation of under-pricing in chinese IPOs, Pac.-Basin Finance J. 1999 (7) (1999) 173–202.
[23] J. Varma, Value at Risk Models in the Indian Stock Market. Working Paper. 5th July, 1999, 1999, p. 3.
[24] R. Brooks, A multi-country study of power ARCH models and national stock market returns, J. Int. Money Finance 2000 (3) (2000) 377–397.
[25] K. Singleton, Estimation of affine asset pricing model using the empirical characteristic function, J. Econometrics 102 (1) (2001) 111–141.
[26] G. Siourounis, Modelling volatility and testing for efficiency in emerging capital markets: The case of the athens stock exchange, Appl. Financ. Econ. 12 (1) (2002) 47–55.
[27] M. Najand, Forecasting stock index futures price volatility: linear vs. nonlinear models, Financ. Rev. 37 (1) (2003) 93–104.
[28] B. Awartani, V. Corradi, Predicting the volatility of the S&P 500 stock index via GARCH models: the role of asymmetries, Int. J. Forecast. 2005 (1) (2005) 167–183.
[29] G. Biswas, Trading volume and market volatility: Developed versus emerging stock markets, Financ. Rev. 2007 (42) (2007) 429–459.
[30] J. Hung, A fuzzy asymmetric? GARCH model applied to stock markets, Inform. Sci. 2009 (22) (2009) 3930–3943.
[31] F. Magnus, O. Fosu, Modelling and forecasting volatility of returns on the ghana stock exchange using GARCH models, Amer. J. Appl. Sci. 36 (4) (2006) 394–419.
[32] A. Pagan, G. Schwert, Alternative models for conditional models for conditional stock volatility, J. Econ. 2007 (45) (2007) 267–290.
[33] T. Angelidis, A. Benos, The use of GARCH model in VaR estimation, Stat. Methodol. 2009 (1) (2009) 105–128.

[34] F. Guidi, Volatility and long-term relations in equity markets: Empirical evidence from Germany, Switzerland and the UK, Icfai J. Financ. Econ. 7 (2) (2009) 7–39.
[35] M. Sabiruzzaman, Modelling and forecasting trading volume index: GARCH versus TGARCH approach, Quart. Rev. Econ. Finance 2010 (2) (2010) 141–145.
[36] H. Liu, J. Hung, Forecasting S&P 100 index volatility: the role of volatility asymmetry distributional assumption in GARCH models, Expert Syst. Appl.: Int. J. 37 (7) (2010) 4928–4934.
[37] H. Ding, ARCH phenomenon in the volatility of stock index, J. Quant. Tech. Econ. 1999 (9) (1999) 22–25.
[38] W. Wei, X. Zhou, Nonlinear GARCH forecast model in the volatility of china stock market, Forecasting 1999 (5) (1999) 47–49.
[39] G. Liu, Application of nonlinear GARCH model in the prediction of china stock market volatility, Stat. Res. 2000 (01) (2000) 22–24.
[40] C. Yue, Empirical research on shanghai stock exchange returns using nonlinear GARCH model, J. Quant. Tech. Econ. 2002 (4) (2002) 56–59.
[41] J. Liu, C. Cui, Empirical analysis of returns and volatility of shanghai and shenzhen stock market, Econ. Res. 2002 (1) (2002) 885–898.
[42] H. Tian, J. Cao, GARCH-M model in Chinese stock market return and volatility, Syst. Eng.-Theory Pract. 2003 (8) (2003) 81–86.
[43] H. Kong, Financial Market Volatility Model and Empirical Study, Capital University of Economics and Business, 2006 (Ph.D. thesis).
[44] Y. Wang, Application of GARCH model in computing the risk value of shanghai stock, Inquiry Econ. Issues 2007 (8) (2007) 153–157.
[45] Q. Tong, Study of stock market volatility–evidence from CSI 300 index, Financ. Econ. 2009 (06) (2009) 81–83.
[46] Y. Chen, L. Han, Empirical research on the volatility of CSI 300 returns, Financ. Econ. 2009 (14) (2009) 70–72.
[47] L. Chen, Stock market risk analysis based on EGARCH-M model and the CSI 300 index, Dongbei Univ. Finance Econ. J. 2010 (2) (2010) 12–18.
[48] F. Gu, Z. Chen, Study of volatility in shanghai and shenzhen stock market using GARCH and SV type model, Econ. Res. Guid 2011 (01) (2011) 4–22.
[49] D. Sun, et al., Application of GARCH type model in the forecast of china's futures market, Liaoning Normal Univ. J. 2012 (03) (2012) 4–6.
[50] J. Zhang, G. He, Time-varying and volatility spillover in China's GEM and main board stock market, Zhongnan Univ. Econ. Law J. 2012 (02) (2012) 100–106.
[51] C. Yang, X. Liu, Impact of macroeconomic information on the volatility of the CSI 300 index, Shanxi Univ. Finance J. 2012 (03) (2012) 38–44.
[52] Wikipedia. Shanghai Stock Exchange [Online] Available from: https://en.wikipedia.org/wiki/Shanghai_Stock_Exchange [Accessed 23.08.15].
[53] J. Armstrong, R. Fildes, On the selection of error measures for comparisons among forecasting methods, J. Forecast. 14 (1) (1995) 67–71.
[54] C. Brooks, S. Burke, Forecasting exchange rate volatility using conditional variance models selected by information criteria, Econom. Lett. 61 (3) (1998) 273–278;
A Banerjee, S. Sarkar, Modelling Daily Volatility of the Indian Stock Market Using Intra-Day Data. Working Paper. 8th May, 2006, 2006, p. 2..
[55] D. Asterious, S. Hall, Applied Econometrics, Palgrave Macmillan, London, 2011.
[56] R. Engle, et al., Estimating time-varing risk premia in the term structure: The ARCH-M model, Econometrica 55 (2) (1987) 391–408.
[57] D. Nelson, Conditional heteroscedasticity in asset returns: A new approach, Econometrica 59 (2) (1991) 347–370.
[58] L. Golsoten, et al., On the relation between the expected value and the volatility of the nominal excess return on stocks, J. Finance 48 (5) (1993) 1779–1801.

Zhe Lin received his M.S. degree in Finance & Investment from Brunel University in London, UK. He is currently a teaching assistant in School of Accounting& Finance, Xiamen University Tan Kah Kee College. His research interest is mainly in the area of Securities Market Modelling, Financial Engineering and Financial Risk Management. He has published several research papers in scholarly journals in the above research areas.