Prediction on Shanghai Composite Index Volatility with Day-of-week Effect, Volume and Turnover Based on HAR Model

Yingfa Zhang[1,*]

[1] *College of economic and business, Australian National University, Canberra ACT 2601,*
[*]*Corresponding author. Email: yingfa.zhang@anu.edu.au*

ABSTRACT

The Shanghai Composite Index is the earliest and one of the most important indexes in Chinese stock market, which is calculated by capitalization market weighted average for all the stocks listed on Shanghai Stock Exchange. However, previous researches mainly use low- frequency-data-based GARCH-type model to predict the volatility of the Shanghai Composite Index without considering the day-of-week effects and the impacts of volume and turnover. In this paper, the HAR-RV model is established primarily based on heterogeneous autoregressive (HAR) theory and five-minute middle-frequency data. Then, trading volume, turnover and day-of-week effects are taken into consideration, respectively, i.e., the HAR-RV-VT model and HAR-RV-W model are constructed. Finally, a mixed HAR-RV-VT-W model is obtained by using the above three factors simultaneously. According to the result, the day-of-week effect and turnover have significant negative impacts on volatility of Shanghai Composite Index while volume has a positive influence. In general, more useful information will be provided based on our mixed model combining volume, turnover and day-of-week effect, which pave a better path to predict the volatility of Shanghai Composite Index.

Keywords: Shanghai Composite Index, HAR-type model, Volatility prediction

1. INTRODUCTION

The capital market of mainland China ranked the first in Asia and forth in the world according to Credit Suisse Global Investment Return Yearbook 2020, which has more than 59 trillion yuan capitalization now. There are four major indexes in China stock market: Shanghai Securities Composite Index, Shenzhen Stock Component Index, Small and Medium-sized Board Index and Growth Enterprise Board Index. Among all four indexes, the Shanghai Composite Index is the earliest and most important one. Shanghai Composite Index, whose full name is Shanghai Stock Exchange (SSE) Stock Price Composite Index, is a stock index compiled by the Shanghai Stock Exchange. It was officially released on December 19, 1990, also the earliest index published in China. The sample of the stock index is all stocks listed on the Shanghai Stock Exchange, among which the newly listed stocks are included in the calculation of the stock index on the second day of listing [1-3]. The Shanghai Composite Index reflects the overall changes of the China stock market, from which trader can get general information they need (e.g., variation trend and overall performance of the market). Contemporarily, there are a great number of literatures focus on the prediction of the Shanghai Composite Index. Meanwhile, many new methods are also used to forecasting Shanghai Composite Index [4].

Investors can judge the approximate trend of stock prices according to the rise and fall of the index. With accurate prediction of volatility of for Shanghai Composite Index, one can better avoid risks and take the lead in investment. Therefore, it is of great significant to improve the volatility prediction, which will have a great help on hedging, risk avoiding and transfer, returns maximizing, price discovery etc. This paper mainly involves the related research on the volatility prediction of Shanghai Composite Index.

During the past 20 years, the market has witnessed the rapid development of the SSE, which becomes the most important stock exchanges in China [5]. As the earliest stock index in China, the Shanghai Composite Index takes all stocks listed on the Shanghai Stock Exchange into the calculation range and calculates index according to the weight distribution of stock issuance. It has great

importance since it roughly reflects the comprehensive performance of the overall stock markets, allowing people to understand the overall change trends of the stock market. On account of the importance of the Shanghai Composite Index, there are many literatures on it including the prediction of its volatility, return etc. In return analysis, the fractality of the time series is a popular method, which reflects the scaling behavior of the data [6-9]. Wen et al. [10] pointed out that the logarithm-mic return series of Shanghai Composite Index is about 0.6298 by the method of R/S analysis, which indicates that the Shanghai Composite Index exhibits fractal features and has a long memory. As for volatility, there are some factors should be focused on: structural breaks, asymmetry, long-memory and day-of-week effect [11-12]. Many studies now use ARCH, GARCH or EGARCH models to predict the volatility of the Shanghai Composite Index[13-14]. Differently, HAR-type model is used in this paper.

In recent years, HAR-type model attracted a large number of scholars to conduct meticulous and in-depth research on owing to its accuracy. As a result, HAR-type developed rapidly these years. Using the high-frequency based realized volatility as a variable to predict related finance market problem was first raised in 1998 by Andersen and Bollerslev[15]. Besides, Andersen et al. pointed out that the model with RV has a better performance than GARCH and SV model in prediction of finance market in 2003 [16]. Subsequently, the HAR-RV model, a heterogeneous autoregressive model with RV, was proposed by Corsi in 2009 with obvious better performance than traditional GARCH model or SV model [17]. Now, there are many studies based on the HAR-type model in forecasting the future volatility of different finance market (e.g., future market [18-19]).

In this paper, the volatility prediction of Shanghai Composite Index is presented via several kinds of HAR-type model. Past data including volatility, volume and turnover is utilized and the day-of week effect is taken into consideration to investigate the relationship between these factors and Shanghai Composite Index. Firstly, the original HAR-RV model is constructed. Then, model with volume and turnover (HAR-RV-VT model) and model with day-of-week (HAR-RV-W model) are introduced separately. Afterwards, the mixed model that integrates all factors (HAR-RV-VT-W model) is demonstrated. The regression results indicates that these new models perform better compared with original model. Specifically, the models with volume and turnover as independent variables perform much better in medium-term and long-term volatility prediction, but not so obvious in short-term. Differently, models consider day-of-week effect have an improvement in short-term volatility prediction, which shows a strong relationship between volatility and day-of-week effect, but no obvious change in medium-term and long-term. In detail, the regression result indicates that volume has a positive effect on volatility of Shanghai Composite Index, especially in short-term and medium-term prediction. Moreover, the influence of day-of-week effect (negative) is most obvious in short-term prediction, which will be eliminated in medium-term and long-term. Meanwhile, it is found that the models integrate all factors only have the highest accuracy in short-term prediction, performing not as well as HAR-RV-VT model in medium-term and long-term. Therefore, compare with all models, taking day-of-week effect into consideration will lead to a good performance in short-term prediction while taking volume and turnover into consideration will result in a good performance in medium-term and long-term prediction.

This paper mainly concludes the following contributions compared to the existing studies based on stock and finance market volatility prediction. First, volume and turnover are regarded as independent variables in HAR-type, which differs from existing studies considering these factors in GARCH-type model [20]. Secondly, the day-of-week effect is included in the model, which is seldom used in other literatures based on HAR model [21-22]. Since day-of-week effect is an extremely common phenomenon that can't be ignored, take it into account will no doubt lead to an improvement in accuracy of prediction. Finally, the results are based on 5-minute medium-frequency data as our data source in HAR-type model. On the contrary, most previous studies related to the volatility of Shanghai Composite Index prediction mainly used low-frequency-data-based GARCH-type model [23], ARCH-type model[24], VAR-type model[25] and SV model[26]. In general, a new method for predict the volatility of Shanghai Composite Index is presented in this paper.

2. DATA DESCRIPTION

In this paper, 5 minutes stock market data is collected to convey the impact of market information, which possesses many advantages. The interval of sampling has great influence on the volatility forecasting in stock market. There will be microscopic noise if the sampling interval is too small while losing too much information if the sampling interval is too large. According to relevant research in recent years (Gong et al. [27]), 5 minutes is proper to be chosen as the sampling interval to measure the volatility of the stock market. The sample contains the Shanghai Composite Index closing price with the volume and turnover from Wind. The sample period starts from the June 1st 2019 and ends on March 22nd 2021, with 21073 items during 441 trading days. Finally, 439 trading days' data is obtained after weeding out vacancies and calculating RV, which is shown in Figure. 1.
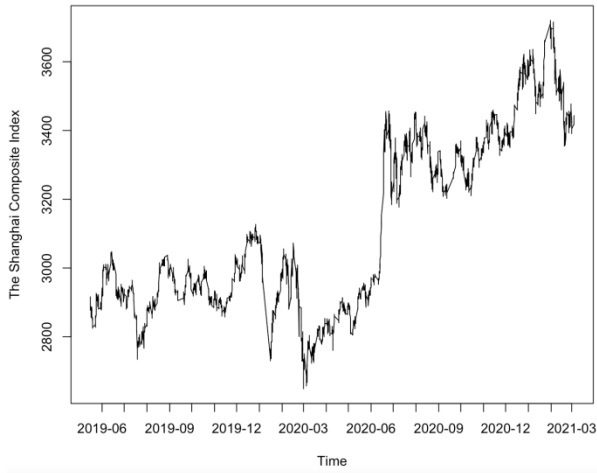
Figure 1 Shanghai Composite Index from June, 2019 to March, 2021.

Figure 2 Price daily volatility of Shanghai Composite Index

As shown in Figure 1, the Shanghai Composite Index fluctuated between 2700 and 3100 from June 2019 to July 2020. There are three significant drops happened on July 30th 2019, January 23rd 2020 and March 11th 2020. During 15 days after July 1st, the index experienced a surge, from 3005 to 3422 and then continued to increase to 3700 with fluctuation by February 22nd 2020. Finally dropped to 3372 at the end of the sampling period.

The daily volatility of Shanghai Composite Index is illustrated in Figure 2, which has a high degree of aggregation and asymmetry. The volatility of the index fifers significantly over time, with sometimes obviously higher than other times. When a major event occurs, the index often reacts violently, such as on January 23rd 2020, which is Chinese new year and the time COVID-19 broke out in China. Besides, on March 22nd 2020, the time COVID-19 broke out all over the world, the index both experienced a plunge. In addition, the surge in July 2020 was because the world's major economies were back to work and the investment value of stock market had become increasingly. Moreover, one sees that the days before and after the day with high volatility are usually also with a high volatility but not so significant.

Many scholars use HAR model to research the volatility of the particular finance market. Therefore, the HAR model is applied in this study to make in-sample prediction in Shanghai Composite Index in next section.
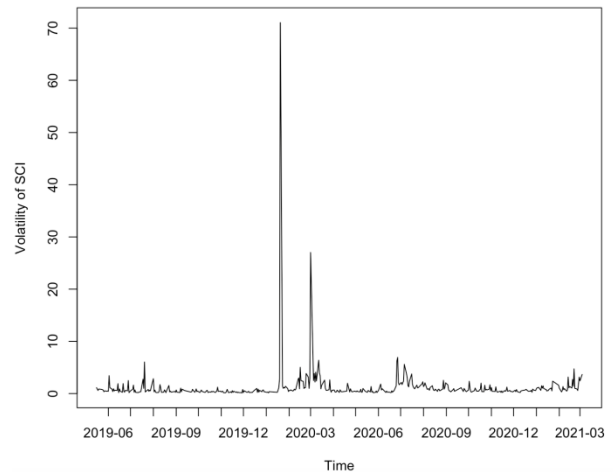
3. ECONOMETRIC MODELS

Four types of models will be discussed here to forecast the volatility of the Shanghai Composite Index one day, one week and one month on average in the future including HAR-RV, HAR-RV-VT, HAR-RV-SW and HAR-RV-VT-SW.

Firstly, the simplest model is HAR-RV model, which only considers the last one day, one week and one month's volatility as independent variables, in order to test whether the past volatility can provide information. The task is to forecast the future volatility in the future financial market. In addition, the HAR-RV-VT model tests the role volume and turnover played in forecasting the future volatility. Because stocking trading takes one week (Monday to Friday) as a cycle, the day-of-the-week effect is taken into account on the basis of two models mentioned above. With this in mind, two more models are established, RV-SW and RV-VT-W, which contain week factors as variables.

3.1. HAR-RV model

According to Andersen and Bollerslev [16] for calculating RV, each trading day is divided into 5-minute intervals and denote i^{th} interval's closing price in day t as $P_{t,i}$. The RV of trading day t is denoted as RV_t^d, which has the expression as following equation:

$$RV_t^d = \sum_{i=1}^{N} r_{t,i}^2 \qquad (1)$$

where $r_{t,i}$ denotes i^{th} the ith ($i = 1, 2, \cdots, N$) logarithmic rate of return in given day t with following formula:

$$r_{t,i} = 100 \cdot \left(\ln P_{t,i} - \ln P_{t,i-1} \right) \qquad (2)$$

In the above formula, the value is magnified a hundred times for easy observation.

Then, RV_t^w and RV_t^m are used to denote the last week's realized volatility and last month's realized volatility respectively, with the expressions below:

$$RV_t^w = \frac{1}{5}\sum_{j=0}^{4} RV_{t-j}^d \qquad (3)$$

$$RV_t^m = \frac{1}{22}\sum_{j=0}^{21} RV_{t-j}^d \qquad (4)$$

The average future volatility for H days, $\overline{RV_{t+H}}$, can be calculated by the following formula:

$$\overline{RV_{t+H}} = \frac{1}{H}\sum_{i=1}^{H} RV_{t+i}^d \qquad (5)$$

However, Andersen et al. [28] pointed out that in HAR model, logarithmic form has better performance than linear form, i.e., the same form is applied in the model to predict the future volatility of the Shanghai Composite Index. The logarithmic form of HAR model has the expression as following:

$$\overline{RV_{t+H}} = \beta_0 + \beta_1 \ln(RV_t^d) + \beta_2 \ln(RV_t^w) + \beta_3 \ln(RV_t^m) \qquad (6)$$

3.2. HAR-RV-VT model

According to Li and Dong [29], there is a significant positive relationship between transaction volume, turnover and Shanghai Composite Index's volatility. In other words, large (small) volatility is accompanied by large (small) trading volume. Besides, this relationship gradually strengthens with the increase of volatility with the characteristics of asymmetry. Therefore, in this model, the transaction volume and turnover will be implanted into the independent variables of the model. and the new model, HAR-RV-VT model, has the expression as following:

$$\overline{RV_{t+H}} = \beta_0 + \beta_1 \ln(RV_t^d) + \beta_2 \ln(RV_t^w) + \beta_3 \ln(RV_t^m) + \beta_4 \ln turnover + \beta_5 \ln volume \qquad (7)$$

In order to reduce the undesirable effects on estimation caused by magnitude difference, logarithmic form of turnover and volume is utilize. In the expression above, $\ln turnover$ denotes the logarithm of the turnover and $\ln volume$ denotes the logarithm of the volume respectively.

3.3. HAR-RV-W model

Many studies have shown that day-of week effect indeed have influences in the stock market. This effect has been extensively studied in recent years as an unavoidable factor. So far, many studies have investigated the impacts on finance market (e.g., South African's Johannesburg Stock Exchange [30], Taiwan stock market [31]). All of these studies show significant differences in stock market returns and volatility over the course of each trading day of the week. Furthermore, Chen and Zhou [32] applied the modified Levene test of unconditional volatility and the GARCH model of conditional volatility to study the Shanghai Composite

Index. The results indicate that the Shanghai Composite Index does have a significant weekly effect: Monday high volatility phenomenon.

On the basis, the day-of-week effect is added into the consideration of the HAR-RV model. The HAR-RV-SW model has the expression as the following:

$$\overline{RV_{t+H}} = \beta_0 + \beta_1 \ln(RV_t^d) + \beta_2 \ln(RV_t^w) + \beta_3 \ln(RV_t^m) + \gamma_1 Mon + \gamma_2 Tue + \gamma_3 Wed + \gamma_4 Thu \qquad (8)$$

where Mon, Tue, Wed and Thu are the dummy variables represent Monday, Tuesday, Wednesday and Thursday, respectively. Friday is set as the basis, which means when trading day is Friday, all dummy variables equal to 0. And when trading day isn't Friday, for example, Monday, then $Mon = 1$ and the other dummy variables equal to 0. Given all above, the day-of-week effect can be analyzed by comparing the coefficients of each dummy variable calculated by the model. For example, if γ_1 is significant, then Monday has a significant influence on the volatility.

3.4. HAR-RV-VT-W model

In this section, all possible factors mentioned above are taken into consideration to evaluate the Shanghai Composite Index, testing whether this full model can provide a better estimation. The logarithmic form of HAR-RV-VT-SW model has expression as following:

$$\overline{RV_{t+H}} = \beta_0 + \beta_1 \ln(RV_t^d) + \beta_2 \ln(RV_t^w) + \beta_3 \ln(RV_t^m) + \beta_4 \ln turnover + \beta_5 \ln volume + \gamma_1 Mon + \gamma_2 Tue + \gamma_3 Wed + \gamma_4 Thu \qquad (9)$$

4. RESULT ANALYSIS

In this section, the variables will be summarized with characteristics decriptions and then use OLS method [17] is utilized subsequently to estimate the parameters of 4 kinds of HAR-type models (HAR-RV model, HAR-RV-VT model, HAR-RV-W model and HAR-RV-VT-W mode)l. After getting the estimated parameters and significance level, the impact of past volatility, volume, turnover and day-of-week effect on future volatility will be discussed. Based on the results, volatility, volume and turnover do contain the needed information to forecast the future volatility. Moreover, day-of-week effect indeed has great influence on future volatility especially on Monday.

4.1. Description Analysis

Table 1. Description statistic of variable

Variable	Mean	Std.dev.	Min	Median	Max
RV	1.1501	3.9223	0.1135	0.5690	71.0574
$\ln turnover$	23.86	0.3558	23.18	23.84	24.91
$\ln volume$	12.53	0.3985	11.72	12.52	13.58

According to the above table of descriptive analysis of main variables, one finds that the volatility of Shanghai Composite Index has a large range, from 0.1135 to 71.0574, which indicates the index volatility's instability. As shown in the Figure 3, most RV concentrate between 0 and 2, with some have extremely large value, which leads to the high Std. dev.
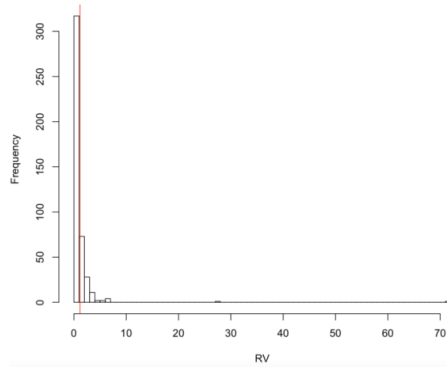

Figure 3 Frequency histogram of daily RV from June, 2019 to March, 2020.

According to Table 1 and Figure 4, $\ln turnover$ basically follows a right skewed normal distribution with mean 23.86 and Std. dev. 0.3558. Compared to RV, $\ln turnover$ is relatively stable.
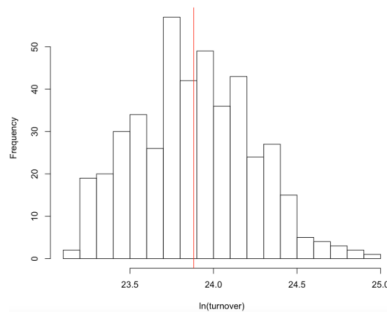

Figure 4 Frequency histogram of daily $\ln turnover$ from June, 2019 to March, 2021

Seen from the Figure 5, hat $\ln volume$ also approximately follows a right skewed normal distribution, with the skewness smaller than that of $\ln turnover$. The mean and Std. dev. equal to 12.53 and 0.3985 respectively. Similar to $\ln turnover$, $\ln volume$ is more stable than RV, which may play an important role in forecasting the future RV.
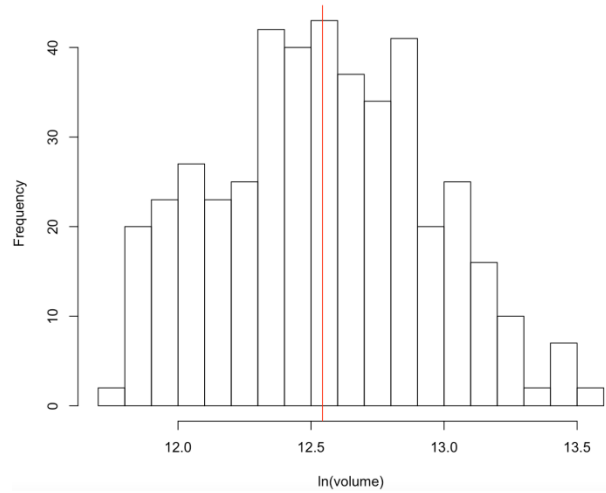

Figure 5 Frequency histogram of daily $\ln volume$ from June, 2019 to March, 2021

4.2. Regression Result

Based on the above HAR-type models, the different impacts of daily volatility, weekly volatility and monthly volatility on future's volatility forecasting are evaluated. The related results are summarized in Table 2.

Table 2. Parameter Estimation Results of HAR-RV-type Models of Shanghai Composite Index

	HAR-RV			HAR-RV-VT			HAR-RV-W			HAR-RV-VT-W		
	1-day	1-week	1-month	1-day	1-week	1-month	1-day	1-week	1-month	1-day	1-week	1-month
β_0	-0.213*** (-5.323)	-0.098** (-2.670)	-0.020 (-0.518)	-1.133 (-0.055)	-2.220 (-0.120)	-88.986*** (-4.905)	-0.044 (-0.547)	-0.121 (1.609)	-0.051 (-0.646)	-2.477 (-0.124)	-2.018 (-0.108)	-88.946*** (-4.874)
$lnRV_t^d$	0.308*** (5.120)	0.283*** (5.107)	0.136* (2.359)	0.241*** (3.900)	0.200*** (3.572)	-0.013 (-0.229)	0.325*** (5.332)	0.302*** (5.252)	0.149* (2.491)	0.252*** (4.037)	0.217*** (3.726)	-0.004 (-0.077)
$lnRV_t^w$	0.248** (2.957)	0.095 (1.225)	0.101 (1.253)	0.227** (2.742)	0.068 (0.911)	0.042 (0.578)	0.226** (2.721)	0.080 (1.020)	0.091 (1.112)	0.205* (2.512)	0.056 (0.730)	0.036 (0.480)
$lnRV_t^m$	0.164* (2.182)	0.216** (3.117)	0.119 . (1.657)	0.142 . (1.894)	0.188** (2.770)	0.034 (0.519)	0.171* (2.337)	0.215** (3.101)	0.119 (1.641)	0.148* (2.031)	0.188** (2.764)	0.035 (0.518)
lnt				-5.758 (-0.593)	-6.669 (-0.758)	30.989*** (3.600)				-5.315 (-0.559)	-6.727 (-0.761)	30.983*** (3.579)
lnv				7.574 . (1.724)	9.189* ()2.307	-3.732 (-0.958)				7.621 . (1.773)	9.179* (2.295)	-3.738 (-0.955)
Mon							-0.362** (-3.272)	-0.041 (0.479)	-0.008 (-0.077)	-0.375*** (-3.458)	-0.057 (-564)	-0.045 (-0.456)
Tue							-0.322** (-2.914)	0.050 (0.479)	0.063 (0.583)	-0.345** (-3.181)	0.023 (0.226)	0.008 (0.082)
Wed							-0.136 (-1.223)	0.085 (0.808)	0.066 (0.598)	-0.160 (-1.464)	0.057 (0.558)	0.012 (0.115)
Thu							-0.004 (-0.038)	0.040 (0.383)	0.047 (0.431)	-0.014 (-0.131)	0.028 (0.281)	0.009 (0.089)
$Adj-R^2$	0.3439	0.2887	0.1128	0.3664	0.3356	0.2673	0.3693	0.2844	0.1054	0.3944	0.3310	0.2599

Note: t statistics in parentheses. $p <0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2 summarizes the parameter estimators of all four kinds of HAR-RV-type models (HAR-RV, HAR-RV-VT, HAR-RV-W and HAR-RV-VT-W) when forecasting Shanghai Composite Index's average volatility in three different future periods (one day, one week and one month). Compared the coefficients of different independent variables, all past data contain the necessary information in forecasting the future volatility of the Shanghai Composite Index although the same data may have different usefulness in different models. Secondly, volume and turnover have some influence in future volatility, but not so significant. Finally, day-of-week effect indeed is a factor which will influence the future volatility.

4.2.1. HAR-RV Model

In HAR-RV model, one sees that one-day model has the best performance, with all three independent variables ($\ln(RV_t^d)$, $\ln(RV_t^w)$ and $\ln(RV_t^m)$) are significantly positive, while in one-week and one-month models, with only two and one independent variables significantly positive respectively ($\ln(RV_t^d)$ and $\ln(RV_t^w)$ for one-week and $\ln(RV_t^d)$ for one-month). Through $adjusted - R^2$ (one-day model has the highest and one-month model has the lowest), predicting the long-term performance of the Shanghai Composite Index is not as accurate as predicting the short-term performance.

4.2.2. HAR-RV-VT Model

In this model, volume and turnover are added into consideration. From the Table 2, it is found that in one-day model, $\ln volume$ is significant positive. Besides, with the two new independent variables, the $adjusted - R^2$ increases compare with original model, which means HAR-RV-VT model has better performance in predicting one day volatility due to introducing $\ln volume$ and $\ln turnover$. As for one-week model, the improvement brought by the introduction of new variables is much more obvious (p-value of $\ln volume$ is smaller), increasing $adjusted - R^2$ from 0.2887 to 0.3356. This result is consistent with Li and Dong[29] that large (small) volatility is accompanied by large (small) trading volume. However, in one-month model, $\ln turnover$ is the only significant independent variable. It's weird because this model ignores all the information contained in past volatility and volume. Furthermore, the $adjusted - R^2$ is too small to make it a reasonable model.

4.2.3. HAR-RV-W Model

This is the model that considers the influence of day-of week effect, which has the strongest impact on short-term prediction while is not obvious on medium-term and long-term prediction. One easily notices that in one-day

model, Monday and Tuesday are both significantly negative, with the three original variables keeping the same significance level. It can be partly explained by this theory: traders accumulated a lot of information, which will have a negative impact on volatility at the beginning of the next week. However, this impact will be eliminated in one-week and one-month model for the reason that each 5 days' trading cycle contains a Monday and Tuesday, and a 22 days' trading cycle has 4 Mondays and 4 Tuesdays. In this case, the day-of-week effect will only affects the short-term model while be offset when the prediction period in the model exceeds five days.

4.2.4. HAR-RV-VT-W Model

HAR-RV-VT-W model is the full model that contains all possible variables. The result of this model has some characteristics from previous models. According to the one-day model results in Table 2, $\ln(RV_t^d)$, $\ln(RV_t^w)$, $\ln(RV_t^m)$ and $\ln voume$ are all significantly positive with different significance level and Monday and Tuesday are significantly negative. Besides, it has the highest $adjusted - R^2$ (0.3944) compared to all other models. With regard to one-week and one-month model, there are no obvious differences in the significance level of variables between HAR-RV-VT-W model and HAR-RV-VT model, while the $adjusted - R^2$s of former is a little smaller than those of latter. The reason is that the influence of day-of-week effect is eliminated by trading cycle.

Compared all the models, HAR-RV-VT-W model has the best performance in short-term prediction since it considers the most comprehensive factors that have economic significance. However, in medium-term and long-term prediction, HAR-RV-VT model has the best performance because it excluded irrelevant or non-significant factors such as day-of-week effects, which will be eliminated due to trading cycle.

5. CONCLUSION

The Shanghai Composite Index is an important judgment index for the general investment, i.e., the volatility prediction is particularly crucial. In this paper, a predictive study is carried out on the Shanghai Composite Index by using 5-minutes-frequency data. The basic predictive model was HAR-RV-type model. Besides, trading volume, turnover and day-of-week effects are added into basic model in order to improve the models' prediction accuracy.

Firstly, original HAR-RV models are established which contains only past volatility data as independent variables. It is found that the short-term model has the best prediction performance while worse in the medium-term and long-term. Secondly, trading volume and turnover are taken into consideration forming HAR-RV-VT model. According to $adjusted - R^2$, the volume

plays a significant role in one-day and one-week prediction, which can be explained by some existing theory. Thirdly, day-of-week effect is introduced getting the HAR-RV-W model, which will have a significant influence on Monday's and Tuesday's volatilities. Finally, all variables are composited into the HAR-RV-VT-W model and some conclusions are proposed accordingly.

Compared all short-term model, HAR-RV-VT-W model has the biggest $adjusted - R^2$, which indicates that it has the best performance. That makes sense because it takes all significant factors (volume, turnover and day-of-week effect) into consideration. As for medium-term and long-term model, in which day-of-week effects are eliminated, HAR-RV-VT model has better performance. With regard to prediction performance of the same model for different time periods, it is easily find that short-term model is the most accuracy. This is a reasonable conclusion because it's always more difficult to predict the volatility in the further future.

Predicting the volatility of Shanghai Composite Index is a challenge task because there are so many factors will have great influence on it. Moreover, its inherently randomness makes prediction much more difficult. In order to get a more accurate prediction, there are still some factors should be taken into consideration. According to Li et al. [33] and previous analysis, structural breaks is a factor that can't be ignored. Whereas, only the influence of past average data is considered with different periods on the future forecasts in HAR-type model. In further research, the method of time series analysis should be presented, which process the past data at intervals of days instead as a kind of correction. In addition, Shanghai Composite index is not the only index in the stock market of the finance market. Each index in the same market and even different markets has a certain degree of interaction according to Mutlu and Arık [34]. Thus, investigating on the relationship between Shanghai Composite Index and indexes in other market is also a direction to improve the accuracy of the prediction.

REFERENCES

[1]. Fethi M D, Pasiouras F. Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey[J]. European journal of operational research, 2010, 204(2): 189-198

[2]. Applications of artificial intelligence in finance and economics[M]. Emerald Group Publishing Limited, 2004.

[3]. Krollner B, Vanstone B J, Finnie G R. Financial time series forecasting with machine learning techniques: a survey[C]//ESANN. 2010.

[4]. Wang Z. (2021). Predicting the rise and fall of shanghai composite index based on artificial intelligence. *EDP sciences*, *235*

[5]. Wang, J., Shao, W., Ma, C., Chen, W., & Kim, J. (2021). Co-movements between shanghai composite index and some fund sectors in china. *Physica A*, 125981.

[6]. Zhi-Qiang Jiang and Wei-Xing Zhou, "Scale invariant distribution and multifractality of volatility multipliers in stock markets", *Physica A 381*, pp. 343-350, 2007.

[7]. Yuan Ying, Zhuang Xin-tian and Jin Xiu, "Measuring multifractality of stock price fluctuation using multifractal detrended fluctuation analysis", *Physica A 388*, pp. 2189-2197, 2009.

[8]. Zhi-Qiang Jiang and Wei-Xing Zhou, "Multifractal analysis of Chinese stock volatilities based on the partition function approach", *Physica A 387*, pp. 4881-4888, 2008.

[9]. Zhanliang, J., & Handong, L. (2010). Multifractal analysis of realized range-based volatility in shanghai stock exchange composite index. *International Conference on Logistics Systems and Intelligent Management*, 2985-988.

[10]. WEN, F., LI, Z., XIE, C., & SHAW, D. (2012). study on the fractal and chaotic features of the shanghai composite index. *Fractals (Singapore), 20*(2), 133-140.

[11]. Yang, K., & Chen, L. (2014). Realized volatility forecast: Structural breaks, long memory, asymmetry, and Day-of-the-Week effect. *International Review of Finance, 14*(3), 345-392.

[12]. Niu, H., & Wang, J. (2013). Volatility clustering and long memory of financial time series and financial price model. *Digital Signal Processing, 23*(2), 489-498.

[13]. Cheng Cheng. (2014). Analysis on Volatility Characteristics of Shanghai Stock Exchange Index Based on GARCH Model. (Required Research dissertation, Shandong University).

[14]. Cheng, X.L. (2013). Empirical analysis of Shanghai stock index volatility based on GARCH model family. Commercial Economics, 000(014), 29-31.

[15]. Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review, 39*(4), 885–905.

[16]. Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica, 71*(2), 579–625.

[17]. Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics, 7*(2), 174–196.

[18]. Gong, X., & Lin, B. (2018b). Structural breaks and volatility forecasting in the copper futures market. Journal of Futures Markets, 38(3), 290–339.

[19]. Gong, X., & Lin, B. (2018a). The incremental information content of investor fear gauge for volatility forecasting in the crude oil futures market. Energy Economics, 74, 370–386.

[20]. Wang, Y.H., & Wang, K.T., (2004). The impact of stock trading volume on return volatility: An empirical analysis of individual stocks in Shenzhen. Financial Research, 000(012), 81-88.

[21]. Wen, F., Gong, X., & Cai, S. (2016). Forecasting the volatility of crude oil futures using HAR-type models with structural breaks. Energy Economics, 59, 400–413.

[22]. Gong, X., & Lin, B. (2017). Forecasting the good and bad uncertainties of crude oil prices using a HAR framework. Energy Economics, 67, 315–327.

[23]. Zheng, Z. (2004). A comparative study on the volatility prediction ability of the Shanghai Composite Index based on GARCH models with different distribution assumptions. Value Engineering (03), 76-78.

[24]. Ying Y, Zhang Y, & Wu RH. (2005). The characteristics and volatility analysis of the return rate of Shanghai Stock Exchange Index. Forum on Statistics and Information (02), 63-67.

[25]. Hu, J., & Geng, J., (2012). A comparison of volatility models based on Fit and Forecast VAR -- Empirical analysis of Shanghai Stock Exchange Index. The finance.

[26]. Wang Jinchao. (2008). Analysis and comparison of GARCH model and SV model on China's stock market volatility -- a case study of Shanghai Composite Index. Science and Management (2), 25-27.

[27]. Huang, C., Gong, X., Chen, X., & Wen, F. (2013). Measuring and forecasting volatility in Chinese stock market using HAR-CJ-M model. *Abstract and Applied Analysis, 2013*, 1-13.

[28]. Andersen, T. G., Bollerslev, T., & Huang, X. (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics, 160*(1), 176–189.

[29]. Li, D., & Dong, L. (2008). A Study on the Dynamic Relationship between Volatility and Trading Volume in China's Stock Market -- Based on the Angle of Quantile Regression. Journal of Shanxi University of Finance and Economics, 30(7), 76-76.

[30]. du Toit, E., Hall, J.H. & Pradhan, R.P. 2018, "The day-of-the-week effect: South African stock market indices", *African journal of economic and management studies,* vol. 9, no. 2, pp. 197-212.

[31]. Liu, Y. & Chen, Y. 2011, "The day-of-the-week volatility effect in Taiwan's stock index futures market is puzzle", *Journal of interdisciplinary mathematics,* vol. 14, no. 5-6, pp. 619-625.

[32]. Zhou, S., & Chen, L., (2004). The Sunday Effect of Shanghai Stock Market Volatility.] Mathematical Statistics and Management (03), 58-61.

[33]. Li, W., Cheng, Y., & Fang, Q. (2020). Forecast on silver futures linked with structural breaks and day-of-the-week effect. *The North American Journal of Economics and Finance, 53*, 101192.

[34]. Mutlu, E., & Arık, E. (2015). Interaction between single-stock futures and the underlying securities: A cross-country analysis. *Emerging Markets Finance & Trade, 51*(3), 647-657.