

Prediction of S&P 500 Index Using HAR-RV Models with Structural Breaks, Day-of-the-Week Effect and Trading Volume

Xuanrong Gao^{1,*}

¹Department of Computer Science and Engineering in the School of Engineering and Department of Mathematics in the School of Science, Hong Kong University of Science and Technology, Hong Kong, China

*Corresponding author: xgaobb@connect.ust.hk

Abstract. Economic trends are one the crucial research topics. Standard and Poor's 500 index (S&P500) is an indicator of economic change with diversification. In this article, the time series models of forecasting realized volatility, regarding the heterogeneous autoregressive theory, are evaluated. For the day, week, and month predictions, heterogeneous autoregressive models with realized volatility (HAR-RV) type models take structural breaks, the day-of-the-week impact, and trade volume into account. Structure breaks and day-of-the-week effects have positive effects on in-sample prediction, while there is not enough evidence to show that trading volume is significant for in-sample prediction. For out-sample prediction, simple loss functions and the Diebold Mariano test are employed to compare the capability of HAR-RV type models in this article for out-sample performance. The HAR-RV, HAR-RV-VOL, and HAR-RV-WV models are regarded to be accurate for short- and mid-term out-sample prediction. The HAR-RV, HAR-RV-SV, and HAR-RV-SWV models are thought to be significant over the long run.

Keywords: Standard and Poor's 500 indexes; HAR-RV model; structural break; day-of-the-week effect; trading volume.

1. Introduction

Standard and Poor's 500 Index, a capitalization-weighted indicator, monitors the stock performance of 500 large-cap stocks listed on American exchanges. It gauges the performance of the USA economy by tracking market value fluctuations of these equities that represent the leading industries, where stocks included are selected using eight main criteria to reflect the value of comparable businesses [1]. Risk is more varied and can reflect a variety of market developments when there are many enterprises and a wide range of sectors. The diversification of the S&P 500 enables traders to monitor the performance of both the index itself and companies in the same industries. With these characteristics, S&P 500 is chosen in this article.

Realized Volatility (RV) is the assessment of variation in returns for an investment product, based on its historical returns over a specified period. RV, with the advantage of the accuracy of the volatility in prediction, contains more information about the market than traditional volatility based on low-frequency models, for example, GARCH.

Structural break refers to the abruptly altering point in a time series at a certain moment in time. Time series data always include significant inaccuracy, and structural flaws can lead to significant forecasting errors and overall model unreliability. Time series data also usually have ambiguous patterns or features. Understanding the real mechanisms behind collected data may help us comprehend, forecast, and make decisions. This is made possible by spotting structural flaws in models.

The day-of-the-week effect refers to the stock market's yield fluctuation within a week. It describes the propensity of stocks to deliver comparatively high returns on a particular day, compared with other days in the week. For Standard & Poor's 500 indexes, a rolling sample test together with the GARCH model is employed, showing that Tuesday anomalies are prominent [2].

In-depth studies have been focused on RV-type models, modeling realized volatility using various patterns to increase its accuracy. For the first time, concerns with forecasting were studied using high-frequency data-based RV as a surrogate measure of volatility [3]. In 2009, Corsi [4] raised the HAR-

RV model, demonstrating that it is more capable of predicting volatility than the GARCH model. Based on the HAR-RV model, the HAR-CJ model is built in 2003 [5], together with the HAR-RV model with jump variance. In these two models, RV is divided into continuous sample path variance and jump variance. Later, the HAR-RV model considering asymmetry and memory over the long run was proposed in 2012, which provides evidence that long-term memory contains information on volatility predictions [6]. The purpose of the HAR-CJN model, which combines the HAR model with continuous sample route variance, jump variance, and overnight return variation, was to examine the importance of overnight information [7].

In this article, the HAR-RV models are employed to forecast realized volatility. Throughout the forecasting process, models would be built to examine the factor's impacts on the Standard & Poor's 500 Index prediction accuracy. Factors examined in this paper include structural fractures, day-of-the-week effect, and trading volume. To locate the structural fractures, the Iterated Cumulative Sums of Squares Algorithm is utilized which finds 17 structural breaks with a threshold of 1.628. There is a dispute over the importance of structural breaks for in-sample prediction. Depending on the p-value, certain structural fractures are regarded as precisely important, while others are not. For mid- and long-term predictions, Thursday is important, and Wednesday and Friday have an impact on 1-day predictions. The fact that the day-of-the-week effect provides a negative influence on forecasting the S&P 500 is supported by the reality that most of its coefficients are negative. Adj. R^2 is used to determine the in-sample test's accuracy. In contrast to the other two, which both improve prediction accuracy, trading volume scarcely changes. For out-sample prediction, MAE, MSE and MAPE, and the Diebold Mariano test (DM test) are used to evaluate and there is not enough evidence to show that dummy variables should not be ignored in out-sample prediction. The DM test's results show that the HAR-RV, HAR-RV-VOL, and HAR-RV-WV models are significant for 1-day and 1-week prediction, but the DM test's p-value indicates that it is impossible to detect the accuracy relationship between any of them. The accuracy correlations between the HAR-RV, HAR-RV-SV, and HAR-RV-SWV models, each of which owns accuracy and significance, are underdetermined for 1-month prediction.

There are 4 sections to the paper. A brief summary of this work is provided in the Session 1 Introduction as stated above. The methodology session is covered in Section 2, along with data processing for realized volatility, dummy variables of three factors about structure, weekday, and volume per day, and HAR-type models that take the aforementioned elements into account. The result analysis session for the in- and out-sample is in Section 3.

2. Methodology

2.1 Data

Based on the closure price of the 5-minute high-frequency data set, realized volatility data is calculated. The Iterated Cumulative Sums of Squares Algorithm (ICSS Algorithm) is used to determine the structural breaks of the S&P 500. Dummy variables are used to calculate the impact factors mentioned before. The conclusions are provided in the last section. For all the tables in this paper, the number in parentheses is t statistics and * means p-value < 0.1, ** means p-value < 0.05 and when p-value < 0.01, one more star is added.

2.1.1 Realized volatility

The financial markets' ability to forecast volatility is significantly influenced by sample frequency. Microscopic noise will result from an excessive sample frequency. If the frequency is low, there will be some information loss about the daily volatility. Based on the previous research [8, 9], In this paper, data for a 5-minute frequency is selected to measure RV of S&P 500, which comes from Dukascopy Swiss Banking Group from 1st January 2017 to 21st August 2022, 592704 items of raw data. Based on the definition of RV, 1453 daily realized volatility data is available after eliminating the vacancy value from 31st January 2017 to 19th August 2022.

For daily RV of HAR-RV model on the i^{th} trading day, it is denoted by $RV_{i,d}$. The expression of it is as:

$$RV_{i,d} = \sum_{j=1}^M r_{j,i}, \quad (1)$$

where $r_{j,i}$ is return rate in the j^{th} 5-minute trading period of the i^{th} trading day.

$$r_{j,i} = (\ln P_{close,j} - \ln P_{close,j-1}) \times 100. \quad (2)$$

Because the original return is always too small for easy observation and measurement, which is a logarithmic difference between the close of the $j - 1^{th}$ 5-minute trading period and the j^{th} 's. For easy observation, the logarithmic rate of return is multiplied by a hundred as shown in the following formula. As there are five trading days per week and the number of the annual trading day is around 258. Based on the trading period of the S&P 500, the weekly and monthly RV $RV_{i,w}$ and $RV_{i,m}$ for the i^{th} trading day can be expressed by the formula below. For $RV_{i,w}$, D is 5. For $RV_{i,m}$, D is 21.

$$\overline{RV_{i+D}} = \frac{1}{D} \sum_{k=1}^D RV_{i+k,d} \quad (3)$$

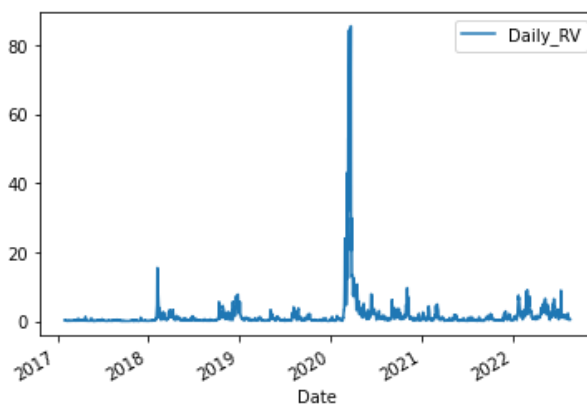


Fig. 1 S&P 500 Daily RV from 31st January 2017 to 19th August 2022

Fig. 1 plots the daily realized volatility, from which structural breaks can be found between 2017 and 2022. Major alterations in economic conditions, such as shifts in policy, calamities, and wars, are referred to as structural breaks. These frequently alter the specifications of economic variables. In this paper, structural breaks would be considered while predicting the realized volatility.

2.1.2 Structural breaks

To identify the structural breaks, ICSS Algorithm is employed. Suppose there are T sample observations. Based on the process of employing the ICSS algorithm [10], the cumulative sum of squares (CSS) for the residuals of all possible structural break observations from 1 to k is defined as

$$C_k = \sum_{t=1}^k r_t^2 \quad (4)$$

where $k = 1, 2, \dots, T$ and C_T represent the CSS residuals covering the entire sample period. σ_i^2 is the unconditional variance for the i^{th} interval, where $i = 1, 2, \dots, M$. M is the count of variance differences in observations. For the daily realized volatility of day t, which is between the T_i and T_{i+1} period, the unconditional variance of it is σ_i^2 . T_i represents the day of the i^{th} structural point.

$$D_k = \frac{C_k}{C_T} - \frac{k}{T} \quad (5)$$

is used to test whether there exists a structural break, particularly $D_0 = D_T = 0$. Set the null hypothesis $H_0: \sigma_i^2 = \sigma_{i+1}^2$ and the alternative hypothesis $H_1: \sigma_i^2 \neq \sigma_{i+1}^2$ to test the difference of constant unconditional variances between two independent sample intervals. The CSS is normalized and centered by D_k . No variance structural break occurs in a time series for a given k when it oscillates about 0, but the variance changes when D_k deviates from zero.

$$D_{k^*} = \max_k |D_k| \quad (6)$$

is defined to find the maximum absolute value of the potential structural break. The k^* represents the location in the series of $\max_k |D_k|$.

The null hypothesis claimed may be rejected if D_{k^*} is greater than the default threshold [8]. In this case, k^* is set to be the value of k when selecting $\max_k |D_k|$. Then there exists a structural break point near k , if $\max_k |D_k|$ is greater than the default threshold. In terms of asymptotic behavior, $\frac{D_k}{\sqrt{\frac{T}{2}}}$ resembles a Brownian bridge due to its variance homogeneity [10]. Based on research by Gong and Lin [11] at the asymptotic 99th percentile of $k = \max_k \left| \frac{D_k}{\sqrt{\frac{T}{2}}} \right|$, this paper selects 1.628 as the default threshold to estimate. When D_k is greater than 1.628, it's determined to be a break. The test is held in both the forward and backward order to ensure that the possible turning point is precisely a structural break.

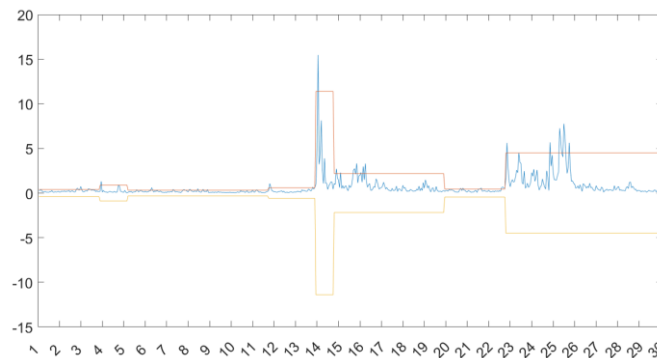


Fig. 2 Daily realized volatility and the ± 3 s.d. bands around structural breaks

Fig. 2 plot the result of structural breaks using the ICSS algorithm. The period from 31st January 2017 to 19th August 2022 is divided into 18 segments by 17 structural breaks. The exact period is listed in Table 1. Dummy variables SB_j ($j = 1, 2, \dots, 17$) are set to track how structural breaks affect the data. These 17 dummy variables are set to 0 at the time before the first structural breakdowns, which is chosen as the basis. For SB_i between the j^{th} and the $j + 1^{th}$ structural breaks, SB_j is set to 1, and others are set to 0.

Table 1. Structural breaks of the daily RV of S&P 500.

Serial number	Time period	Standard deviation
1	from 31/1/2017 to 4/24/2017	0.1298
2	from 4/24/2017 to 5/31/2017	0.2920
3	from 5/31/2017 to 12/1/2017	0.1063
4	from 12/1/2017 to 2/5/2018	0.2239
5	from 2/5/2018 to 2/28/2018	3.6815
6	from 2/28/2018 to 7/24/2018	0.7264
7	from 7/24/2018 to 10/11/2018	0.2418
8	from 10/11/2018 to 5/6/2019	1.4975
9	from 5/6/2019 to 5/28/2019	0.7963
10	from 5/28/2019 to 8/1/2019	0.2689
11	from 8/1/2019 to 9/20/2019	1.0282
12	from 9/20/2019 to 2/20/2020	0.3831
13	from 2/20/2020 to 3/24/2020	24.8815
14	from 3/24/2020 to 6/10/2020	5.9201
15	from 6/10/2020 to 11/10/2020	1.6925
16	from 11/10/2020 to 11/26/2021	0.7120
17	from 11/26/2021 to 1/21/2022	0.7056
18	from 1/21/2022 to 19/8/2022	1.8120

2.1.3 Day-of-the-week effect and trading volume

S&P 500 trades from Monday to Friday each week. In this article, Monday is used as the foundation, while Tuesday through Friday are represented by the four dummy variables WD_i , where $i = 1, 2, 3$, and 4. When Tuesday is the trading day, WD_1 is set to 1 and the other variables are set to zero. The same procedure is followed for Wednesdays, Thursdays, and Fridays. Four of these dummy variables had their values set to zero on Mondays.

Each 5-minute high-frequency trading volume with a unit of 10 million is added together to create the daily trading volume. The daily trading volatility of the i^{th} trading day is Vol_i with its formula as followed.

$$Vol_i = \sum_{j=1}^M vol_{j,i} \quad (7)$$

where $vol_{j,i}$ is the trading volume in the j^{th} 5-minute trading period of the i^{th} trading day.

2.2 HAR-RV models

In this article, the logarithmic HAR-RV model is chosen for prediction. As the influence of dummy variables in modeling is unknown, models of the combination of these variables are formed and the outcomes would be compared, shown as in Table 2.

Table 2. HAR-RV models with factors and their combination forms

abbreviation	Description	formula
HAR – RV	Logarithmic HAR-RV model	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m}$
HAR – RV – SB	Logarithmic HAR-RV model with structural breaks	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \sum_{j=1}^{17} \beta_j SB_j$
HAR – RV – VW	Logarithmic HAR-RV model with the day-of-the-week effect	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \sum_{k=1}^4 \gamma_k WD_k$
HAR – RV – VOL	Logarithmic HAR-RV model with trading volume	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \theta \ln Vol_i$
HAR – RV – SW	Logarithmic HAR-RV model with structural breaks and day-of-the-week effect	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \sum_{j=1}^{17} \beta_j SB_j + \sum_{k=1}^4 \gamma_k WD_k$
HAR – RV – WV	Logarithmic HAR-RV model with day-of-the-week effect and trading volume	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \sum_{k=1}^4 \gamma_k WD_k + \theta \ln Vol_i$
HAR – RV – SV	Logarithmic HAR-RV model with structural breaks and trading volume	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \sum_{j=1}^{17} \beta_j SB_j + \theta \ln Vol_i$
HAR – RV – SWV	Logarithmic HAR-RV model with structural breaks, the day-of-the-week effect and trading volume	$\ln \overline{RV}_{t+D} = \alpha_0 + \alpha_1 \ln RV_{i,d} + \alpha_2 \ln RV_{i,w} + \alpha_3 \ln RV_{i,m} + \sum_{j=1}^{17} \beta_j SB_j + \sum_{k=1}^4 \gamma_k WD_k + \theta \ln Vol_i$

3. Results

3.1 Descriptive statistical analysis

The mean of $RV_{i,d}$, $RV_{i,w}$ and $RV_{i,m}$ as shown in Table 3. below of the descriptive statistical analysis of describable variables, are on the decline. The maximum value of these three variables exhibits the same tendency, which is the inverse of the trend of the minimum as well as standard deviation. The definition of $RV_{i,w}$ and $RV_{i,m}$ leads to these conclusions of maximum and minimum. The differences between maximum and minimum also follow the decreasing tendency. Though setting the unit of Vol_i to be 10 million, both the mean and sd are much greater than the other three variables. Also, the minimum of Vol_i is the smallest, while its maximum is the greatest.

Table 3. Descriptive data regarding the variables

variable	mean	standard deviation	minimum	maximum
$RV_{i,d}$	85.5366	1.5460	4.9327	0.0199
$RV_{i,w}$	54.2981	1.5454	4.3792	0.0469
$RV_{i,m}$	32.8460	1.5411	3.6518	0.0736
Vol_i	118.8843	10.6593	13.0631	0.0119

3.2 In-sample analysis

From table 4., for structural breaks, the significance of SB_j is polarizing. SB_{12} is the most important for the prediction of any period. SB_7 , SB_{13} , and SB_{14} are significant for 1-day and 1-week prediction. SB_{10} is more significant for 1-month prediction than 1-week and considered to be insignificant for 1-day prediction. The significant distribution of SB_4 is similar as SB_{10} . In each model relevant to structural breaks, the significance of SB_i varies, and for all of them, the number of significant SB_i is less than or equal to half of the total number of SB_i .

For the impact of weekdays, WD_3 is of importance for mid-term and long-term prediction with a negative coefficient. Also, from the HAR-RV results of mid- and long-term prediction, the coefficients of which are mostly negative, especially for those considered to be significant. The significant WD_k of short-term is WD_2 and WD_4 . The coefficients of WD_2 in 1-day prediction are positive, while those of WD_4 are always negative.

For all the models shown above and for any period of prediction, trading volume is insignificant.

Table 4. Three Scheme comparing

	HAR-RV(log)			HAR-RV-SB			HAR-RV-WD			HAR-RV-SW			HAR-RV-VOL			HAR-RV-SV			HAR-RV-WV			HAR-RV-SWV		
	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth	1 day	1 wee k	1 mo nth
α_0	-			-	-	-	-		0.01	-	-		0.1		0.01	-	-	-	0.0	0.01	-	-	-	0.00
	0.1 039 ***	0.0 175 ***	0.0 141 ***	0.3 101 ***	0.0 559 **	0.0 017	0.1 029 **	0.0 376 ***	94* ***	0.3 173 ***	0.0 366	0.00 3	0.1 217 ***	0.098	44* **	0.3 412 ***	0.0 629 **	0.000 59	0.1 233 **	0.0 296 **	0.01 97* **	0.3 527 ***	0.0 439	-
	(- 5.0 481)	(- 2.8 283)	(- 6.1 339)	(- 3.3 192)	(- 1.9 752)	(- 0.1 644)	(- 2.5 770)	(- 1.90)	(3.1 507)	(4.3 917)	(- 3.1 118)	(0.2 751)	(- 3.2 723)	(0.8 719)	(3.4 665)	(- 3.3 903)	(- 2.0 604)	(- 0.52 370)	(- 2.4 365)	(1.9 717)	(3.4 186)	(- 3.3 586)	(- 1.3 22)	(- 0.11 0.02)
	0.5 298 ***	0.1 975 ***	0.0 330 ***	0.4 858 ***	0.1 875 ***	0.0 295 ***	0.5 330 ***	0.2 006 ***	0.03 39*	0.4 883 ***	0.1 905 ***	0.03 03*	0.5 245 ***	0.1 953 ***	0.03 31*	0.4 763 ***	0.1 854 ***	0.02 82*	0.5 270 ***	0.1 982 ***	0.03 39*	0.4 775 ***	0.1 882 ***	0.02 90*
	$\ln RV_i$	(- 15. 339 9)	(- 19. 000 4)	(- 8.5 711 2)	(- 14. 079 2)	(- 17. 929 83)	(- 7.7 83)	(15. 19. 8)	(19. 398 7)	(8.7 382)	(14. 125 3)	(18. 113)	(7.9 431)	(14. 677 0)	(18. 152 5)	(8.3 096)	(13. 816 2)	(16. 096 2)	(7.0 664)	(14. 721 2)	(18. 345 1)	(8.4 639)	(13. 114 5)	(16. 993 3)
$\ln RV_i$	0.2 828 ***	0.8 073 ***	0.0 401 ***	0.2 491 ***	0.7 992 ***	0.0 308 ***	0.2 795 ***	0.8 043 ***	0.03 92*	0.2 457 ***	0.7 964 ***	0.03 00*	0.8 819 ***	0.04 068 ***	0.2 01*	0.7 461 ***	0.7 985 ***	0.2 04*	0.8 785 ***	0.8 039 ***	0.03 92*	0.2 423 ***	0.7 957 ***	0.02 96*
	(- 5.7 345)	(- 54. 371)	(- 7.2 757)	(- 5.0 176)	(- 5.3 115)	(- 5.6 439)	(5.6 734)	(7.1 083)	(4.9 037)	(5.2 855)	(5.4 872)	(5.7 110)	(5.4 307)	(7.2 716)	(4.9 442)	(5.2 919)	(5.5 568)	(5.6 476)	(5.4 020)	(7.0 990)	(4.8 788)	(5.2 987)	(5.3 987)	
	0.1 253 ***	- 186 *	0.9 290 ***	0.0 651 ***	0.0 392 ***	0.9 329 ***	0.1 253 ***	- 185 *	0.92 91*	0.0 652 ***	0.0 388 ***	0.93 30*	0.1 229 ***	0.0 196 ***	0.92 90*	0.0 664 ***	0.0 389 ***	0.93 31*	0.1 227 ***	- 195 *	0.92 91*	0.0 667 ***	0.0 385 ***	0.93 32*
	(- 3.4 131)	(- 1.6 811)	(- 2.26 08)	(- 1.4 299)	(- 2.8 390)	(- 1.8 41)	(3.4 286)	(- 1.6 762)	(2.2 6.94 22)	(1.4 386)	(- 2.8 163)	(18 6.68 87)	(3.3 274)	(- 1.7 625)	(2.2 5.18 20)	(1.4 585)	(- 2.8 147)	(18 6.33 23)	(3.3 339)	(- 1.7 601)	(2.2 5.45 61)	(1.4 711)	(- 2.7 911)	(18 6.61 50)
	$\ln Vol$												0.0 101)	0.0 044)	- 0.00 02	0.0 179)	0.0 04)	0.0 24)	0.0 115)	0.0 045)	- 0.00 01	0.0 202)	0.0 042)	0.0 25)
SB_1				- 553 *	- 0.0 041	- 0.0 076			0.2 496 *	- 0.0 043	- 0.00 73					0.2 624 *	- 0.0 057	- 0.00 85			0.2 575 *	- 0.0 06	- 0.00 82	
				(- 1.7 892)	(- 0.0 951)	(- 0.4 838)			(- 1.7 571)	(- 0.1 001)	(- 0.46 44)					1.8 354)	0.1 313)	(- 0.54 32)			(- 1.8 099)	0.1 377)	(- 0.52 55)	
				- 0.1 096)	- 0.0 187)	- 0.0 029)			- 0.1 09)	- 0.0 186)	- 0.00 3					0.0 857)	0.0 133)	0.00 03			0.0 819)	0.0 131)	0.0 03	
				(- 1.1 466)	(- 0.6 444)	(- 0.2 802)			(- 1.1 451)	(- 0.6 435)	(- 0.28 61)					0.8 577)	0.4 405)	(0.0 236)			(- 0.8 241)	0.4 320)	(0.0 273)	
				0.0 107)	0.0 069)	0.0 116)			0.0 144)	0.0 078)	0.01 19					0.0 112)	0.0 07)	0.01 17			0.0 149)	0.0 079)	0.01 19	
SB_3	(- 0.0 891)	(- 0.1 893)	(- 0.8 792)						(0.1 198)	(0.2 145)	(0.8 979)					(0.0 929)	(0.1 921)	(0.8 838)			(0.1 243)	(0.2 175)	(0.9 029)	
SB_4				0.2 722 ***	0.1 145 **	0.0 973 ***			0.2 122 ***	0.1 72*	0.09 **					0.2 849 ***	0.1 174 **	0.09 **			0.2 893 ***	0.1 152 **	0.09 **	
	(- 1.5 044)	(- 2.0 88)	(- 4.8 936)						(1.5 265)	(2.0 480)	(4.8 908)					(1.5 687)	(2.1 314)	(4.9 613)			(1.6 003)	(2.0 938)	(4.9 614)	

Although the significance and the parameters of the dummy variables have been discussed in the last session, their effects on accuracy are still ambiguous. To determine whether these dummy variables have a beneficial influence on the accuracy of in-sample prediction, adjusted R^2 is used to quantify the accuracy of prediction models. From the last line of Table 4, all the models improve the accuracy of 1-day prediction, except for the HAR-RV-Vol model. Considering both structure and the impact of weekdays improve the accuracy of modeling, especially for structural breaks. In mid-term prediction, HAR-RV-SW outperforms the others, and the accuracy of HAR-RV-Vol, which is close to the one of logarithmic HAR-RV, has the lowest accuracy among models with dummy variables. For 1-month prediction, the accuracy of the HAR-RV model is 0.9962, and the models considering structure, weekdays and their combination improve the models' accuracy to 0.9965.

3.3 out-sample analysis

3.3.1 Loss functions

In this session, the loss functions used to assess the prediction models are MAE, MSE, and MAPE. From Table 5., for 1-day prediction, the HAR-RV-Vol model only outperforms the others evaluated using MAE and all the models with dummy variables perform worse than the HAR-RV model measured by MSE. From MAPE, HAR-RV-SB, and HAR-RV-Vol together with their combination models outperform the HAR-RV form. For 1-week prediction, only HAR-RV-WD and HAR-RV-WV outperform measured by MAE and the MSE of HAR-RV is the most well-performed compared with other models when increasing significant digits and none of the models containing dummy variables surpasses HAR-RV model for both mid- and long-term based on MAPE. For 1-month prediction, only the HAR-RV-Vol model was slightly well-performed than the HAR-RV models and none of them perform better on the dimension of MAE.

Table 5. Data of loss functions

	HAR – RV			HAR – RV – SB			HAR – RV – WD			HAR – RV – SW		
	short-term	mid-term	long-term	short-term	mid-term	long-term	short-term	mid-term	long-term	short-term	mid-term	long-term
MAE	0.61	0.11	0.03	0.62	0.13	0.04	0.63	0.11	0.03	0.63	0.13	0.04
E	94	99	87	55	88	06	29	98	91	48	83	09
MSE	0.73	0.02	0.00	0.74	0.03	0.00	0.74	0.02	0.00	0.75	0.03	0.00
	90	62	28	85	29	31	84	62	29	69	27	31
MAP	1.72	0.57	0.32	1.60	0.61	0.38	2.07	0.59	0.35	1.92	0.63	0.42
E	27	49	13	61	42	16	98	65	93	79	79	17
	HAR – RV – Vol			HAR – RV – SV			HAR – RV – WV			HAR – RV – SWV		
	short-term	mid-term	long-term	short-term	mid-term	long-term	short-term	mid-term	long-term	short-term	mid-term	long-term
MAE	0.61	0.12	0.03	0.62	0.13	0.03	0.62	0.11	0.03	0.63	0.13	0.04
E	45	00	87	64	69	99	78	98	91	53	65	01
MSE	0.74	0.02	0.00	0.75	0.03	0.00	0.75	0.02	0.00	0.75	0.03	0.00
	98	61	28	00	22	30	79	62	29	70	21	30
MAP	1.67	0.59	0.32	1.61	0.60	0.36	2.04	0.61	0.36	1.93	0.62	0.40
E	88	30	26	39	20	82	09	52	02	64	55	83

3.3.2 Diebold Mariano test

DM test [12] is employed to assess the model's correctness in more precision and detail. Its null hypothesis is that the predicting accuracy of the two models is equivalent. In this paper, differences in accuracy are assessed using MAE and MSE.

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\Sigma}(d_i)}{T}}} \quad (8)$$

where $\bar{d}$ is the mean of d_i , $d_i = e_{i,1} - e_{i,2}$ is the difference between two models. e_i is the residual between each prediction value and actual value and the expressions of e_i are chosen to be the same as MAE and MSE. $\hat{\Sigma}(d_i)$ is the robust estimate of the autocovariance for d_i .

The null hypothesis is accepted when the p-value > 0.05 , indicating that the two models perform similarly. The null hypothesis is disproved when the p-value < 0.05 , indicating that the effects of tested models vary. DM surpassing 0 indicates that the model surpasses its benchmark when the p-value is less than 0.05. The benchmark performs better on the contrary.

The HAR-RV model is intended to serve as a baseline for determining if forecasting accuracy has considerably improved for the HAR-RV-SB, HAR-RV-DW, and HAR-RV-Vol models. Then, use each of HAR-RV-SB, HAR-RV-DW, and HAR-RV-Vol as a benchmark to assess how well they work along with one of the factors mentioned above. Finally, two-factor models are utilized to compare to the model made up of all the elements mentioned before.

Table 6. Diebold Mariano test

benchmarks		MAE			MSE		
		1 day	1 week	1 month	1 day	1 week	1 month
	HAR	-	-		-	-	
	– RV	3.9937***	4.5425***	-1.9750**	4.7889***	4.5554***	-1.9865**
	– SB						
HAR – RV	HAR						
	– RV	-1.1153	-0.6900	-1.9341*	-1.0572	-0.7698	-1.9177*
	– DW						
	HAR						
	– RV	0.7691	-0.4173	-0.5946	1.5590	0.4153	0.0255
	– VOL						
	HAR						
	– RV	-0.6622	0.2857	-1.4854	-1.1862	-0.2099	-1.8211*
HAR – RV	– SW						
– SB	HAR						
	– RV	6.6670***	8.5047***	5.0536***	7.7769***	7.8636***	4.5850***
	– SV						
	HAR						
	– RV	0.5271	0.2097	-0.2092	1.4335	0.3500	0.1059
HAR – RV	– WV						
– DW	HAR						
	– RV	3.8950***	4.3660***	-1.8192*	4.8065***	4.4824***	-1.9592*
	– SW						
	HAR						
	– RV	-1.1282	-0.4602	-1.917*	-1.0263	-0.7850	-1.9168*
HAR – RV	– WV						
– VOL	HAR						
	– RV	3.6381***	4.0814***	-1.3900	4.4298***	4.2069***	-1.5600
	– SV						
	HAR						
	– RV	6.7690***	8.2477***	4.8401***	7.7211***	7.7305***	4.4477***
HAR – RV	– SWV						
– SW	HAR						
	– RV	-0.3347	0.3967	-1.4582	-0.8884	-0.0759	-1.7763*
	– SWV						
	HAR						
	– RV	-	-	-1.2250	-	-	-1.5209
HAR – RV	– WV	3.4291***	3.9263***		4.3784***	4.1083***	
– WV	– SWV						

According to Table 6, the HAR-RV model outperforms the HAR-RV-SB model for any period of prediction, the HAR-RV-SV model exceeds the HAR-RV-SB model and the HAR-RV-SWV model is more accurate than the HAR-RV-SW model. HAR-RV-SB model performs better than HAR-RV-SW for 1 day and week prediction, the HAR-RV-VOL model outperforms HAR-RV-SV, and the HAR-RV-WV model is more accurate than the HAR-RV-SWV model. Using ' to represent the accuracy of the model.

Three inequations can be formed to show the accurate relationship for 1 day and week:

$$\begin{cases} HAR - RV' > HAR - RV - SB' > HAR - RV - SW' \\ HAR - RV - VOL' > HAR - RV - SV' > HAR - RV - SB' \\ HAR - RV - WV' > HAR - RV - SWV' > HAR - RV - SW' \end{cases} \quad (9)$$

Three inequations can be formed to show the accurate relationship in the long-term:

$$\begin{cases} HAR - RV' > HAR - RV - SB' \\ HAR - RV - SV' > HAR - RV - SB' \\ HAR - RV - SWV' > HAR - RV - SW' \end{cases} \quad (10)$$

Table 6., shows no substantial difference between the HAR-RV and HAR-RV-VOL models, which makes it impossible to determine with certainty which model is more accurate. The associations between HAR-RV and HAR-RV-WV models are untested over the short and mid-term, HAR-RV and HAR-RV-SV, and HAR-RV-SWV models are not tested over the long term.

From Table 7. and Table 8., there is no clear relationship between HAR-RV and HAR-RV-WV models for the short- and mid-term and between HAR-RV, HAR-RV-SV, and HAR-RV-SWV models over the long-term.

Table 7. Additional short-term and mid-term Diebold Mariano test

benchmarks		peroid	MAE	MSE
HAR – RV	HAR – RV – WV	short-term	-0.5632	-0.639
		mid-term	-0.9892	-0.8023

Table 8. Additional long-term DM test f

benchmarks		MAE	MSE
	HAR – RV – SV	-1.3877	-1.5495
HAR – RV	HAR – RV – SWV	-1.8255*	-2.1706**
HAR – RV – SV	HAR – RV – SWV	-1.4582	-1.7763*

4. Conclusion

In this paper, RV forecasting of the S&P 500 by using a dataset with a frequency of five minutes. Employing HAR-RV type modes, the influence of structure, weekdays, and trading volume are evaluated for S&P500. From the definition of the HAR-RV model, three types of RV are calculated. To find the structural breaks, the ICSS algorithm has been used. Dummy variables have been constructed to measure the impacts of three factors. In this article, HAR-RV, HAR-RV-SB, HAR-RV-VW, HAR-RV-Vol, HAR-RV-SW, HAR-RV-WV, HAR-RV-SV, and HAR-RV-SWV models are applied in both in- and out-sample prediction.

Analyzing in-sample prediction, structural breaks and the day-of-the-week impact are crucial to increase the model's accuracy, indicating that significant information to help improve the prediction is contained, while the trading volume with few contributions. However, for out-sample prediction evaluated by MSE, MAE, and MAPE, HAR-RV, HAR-RV-Vol, HAR-RV-SB, and HAR-RV-WD models outperform in some loss functions of different forecasting periods. To further measure the

accuracy of the model prediction, the DM test is employed with MAE and MSE separately. HAR-RV, HAR-RV-Vol, and HAR-RV-WV models are well performed over short-term and mid-term out-sample prediction compared with the others. For the long-term, HAR-RV, HAR-RV-SV, and HAR-RV-SWV models outperform. Since enough evidence to judge the relationships between well-performed models for 1-day and 1-week as well as between well-performed models over the long run by DM test and these models covering all factors, none of the dummy variables should be neglected in out-sample prediction to form these models.

A more efficient method may be used to measure the performance of out-sample prediction in follow-up research and more models and dummy variables may be involved for better out-sample prediction of the RV of S&P500.

References

- [1] Bebchuk, L., Hirst, S., & Rhee, J. (2013). Towards the declassification of S&P 500 boards. *Harv. Bus. L. Rev.*, 3, 157.
- [2] Zhang, J., Lai, Y., & Lin, J. (2017). The day-of-the-week effects of stock markets in different countries. *Finance Research Letters*, 20, 47-62.
- [3] Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885–905.
- [4] Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174–196.
- [5] 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.
- [6] Asai, M., McAleer, M., & Medeiros, M. C. (2012). Asymmetry and Long Memory in Volatility Modeling. *Journal of Financial Econometrics*, 10(3), 495–512.
- [7] Zhang, J., Lai, Y., & Lin, J. (2017). The day-of-the-Week effects of stock markets in different countries. *Finance Research Letters*, 20, 47–62.
- [8] 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.
- [9] Xiong, Q. (2021). Forecast on S&P 500 Index Based on HAR-RV Model With VIX and Day-of-the-Week Effect. *Proceedings of the 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021)*, 1333-1338.
- [10] Inclan, C., & C. Tiao, G. (1994). Use of Cumulative Sum of Squares for Retrospective Detection of Changes in Variance (Vol. 89).
- [11] Gong, X., & Lin, B. (2018). The incremental information content of investor fear gauge for volatility forecasting in the crude oil futures market. *Energy Economics*, 74, 370-386.
- [12] Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1), 134-144.