

Research on Prediction and Early Warning of A-Share Market Volatility Based on HAR-Type Models

Zhaohao WEI

*School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190,
China*

E-mail: 18706785168@163.com

Jichang DONG

*School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190,
China*

E-mail: jcdonglc@ucas.ac.cn

Zhi DONG*

*School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190,
China*

E-mail: dongzhi@ucas.ac.cn

Abstract Based on the different premium volatility characteristics of various systematic factors in the A-share market, this paper constructs six representative high-frequency volatility prediction models that consider multiple complex risk structures. On this basis, a detailed comparative analysis of the differences in volatility characteristics among various factors is conducted, and the optimal prediction and early warning framework for the A-share market is proposed. Research shows that: 1) The volatility research results only for individual market indexes are not universally representative. 2) The fluctuation characteristics among different systematic factors and their respective optimal prediction model frameworks generally have significant differences, that is, there is no single fixed combination of model parameters. 3) Complex risk characteristics such as long memory, measurement errors, and high-frequency jump fluctuations obviously exist in the A-share market. The optimal forecast and early warning framework for the A-share market can be constructed by a combination of models that consider one or more of the above risk characteristics. The above conclusions have important practical reference value for the risk warning and prevention of the A-share market and the formulation of related policies.

Keywords volatility prediction; high-frequency data; factor model; HAR-type model

1 Introduction

According to the APT (Arbitrage Pricing Theory), the expected return of an asset can be explained by the systematic risk it bears. The factor load part explains the reason for the difference in risk return when each asset faces the same risk, and the factor premium part

Received March 21, 2023, accepted May 20, 2023

* Corresponding author

explains the risk compensation benchmark under the unit load exposure. With the continuous change of the factor premium level, the fair risk compensation return level required by each asset is also constantly changing. That is, facing unit factor loads, the price benchmarks and systematic risk levels of each asset will exhibit volatility characteristics consistent with the trend of each systematic factor premium. Therefore, how to scientifically identify and manage the fluctuation level of each factor premium will be of great significance for the optimization of the corresponding investment portfolio and risk management.

Volatility forecasting has always been a core issue in the field of empirical asset pricing and risk management. The traditional volatility forecasting method is represented by the GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) model, which effectively depicts the volatility clustering effect and heteroscedasticity characteristics widely existing in various assets in the financial market^[1]. On this basis, a large number of scholars have successively constructed multiple derivative GARCH models represented by IGARCH, AVGARCH, EGARCH, NGARCH, APARCH, NAGARCH, TGARCH, and csGARCH, taking into account complex structures such as nonlinearity, asymmetry, and differences in long-term and short-term volatility, and have demonstrated their effectiveness in depicting volatility in different scenarios^[2–12]. However, with the gradual improvement in the availability of financial data, Andersen and Bollerslev further proposed the realized volatility index RV (realized volatility) using more high-frequency data information, and demonstrated its consistent estimation results of fitting true volatility under the general semi-martingale hypothesis^[13]. On this basis, Crosi further proposed a heterogeneous autoregressive model HAR (Heterogeneous Autoregressive model of Realized Volatility) based on the realized volatility, thereby avoiding the dependence on implicit volatility prediction structures in previous GARCH models, and effectively utilizing more high-frequency information about volatility prediction^[14]. A large number of previous studies have shown that the HAR model, which considers both the long memory characteristics of volatility and the heterogeneity characteristics of investors, can more effectively explain and predict the future trend and volatility of relevant indicators, and is widely applicable to various representative financial markets and products such as stock markets, bonds, crude oil, and precious metals^[15–23].

Considering the different risk characteristics that may exist, some scholars have further optimized and developed traditional HAR models through more diverse feature introduction or feature construction methods based on high-frequency prediction models represented by heterogeneous autoregressive. Specifically, first of all, Andersen, et al. proposed HAR-J and HAR-CJ models that can incorporate jump characteristics for the first time^[24]. That is, by identifying and separating high-frequency jump components, it effectively proves the differential performance of continuous components and jump components in the prediction process of realized volatility. On this basis, a large number of scholars have further carried out research on high-frequency jump fluctuation components and their predictive effectiveness, but have obtained diametrically opposite research conclusions. Among them, some scholars believe that high-frequency jumping volatility components can effectively predict the future trend of asset prices or other financial indicators. Taking the Chinese stock market as an example, Qiao, et al. constructed a TVC-HAR-CJ-iVX model that can simultaneously consider high-frequency

jump components and time-varying coefficients^[25]; Ma, et al. constructed a MS-HAR-TJI model that can distinguish high-frequency jump intensity and jump components^[26]; Yuan, et al. constructed a LHAR-CJ jump volatility prediction model based on multifractal theory^[27]. The above representative studies all prove the effectiveness of jumping volatility components in high-frequency forecasting. Similarly, Guo and Ma^[28], Ye, et al.^[29], Ma, et al.^[30] and Zeng, et al.^[31] further extended this part of the research to precious metals, ETFs, and G7 stocks market, as well as various representative financial markets around the world, and have demonstrated the necessity of considering the impact of jump components in high-frequency forecasting. However, another part of scholars believe that in the volatility prediction of related investment targets, the high-frequency jump component has only a weak predictive ability, that is, its contribution is relatively limited^[32–34]. Especially for the Chinese stock market, this phenomenon may become more obvious^[35–37]. In addition, Bunic and Gisler^[38] even pointed out directly through comparative research on 18 representative stock markets, such as the United States, Japan, Germany, Hong Kong, South Korea, and Singapore, that the jumping component cannot provide information about future fluctuations in corresponding indexes or investment targets and achieve effective prediction results. Overall, there is no unified conclusion yet on the discussion of the predictive ability of high-frequency jump components.

In addition to the above research on high-frequency jump components, Bollerslev, et al.^[39] further improved the original HAR model by considering measurement errors. Facts have proven that measurement errors are widespread in research related to high-frequency prediction. By constructing a HARQ model that can introduce error correction terms, the prediction ability of the benchmark HAR model will be greatly improved^[40, 41]. At the same time, Wang, et al.^[42] also confirmed this view in the Chinese stock market through research on the Shanghai and Shenzhen 300 Index. Song and Wang constructed a HARQ(F)-CJ model with higher prediction accuracy for the volatility characteristics of China's stock market by simultaneously considering the impact of jump components and measurement errors^[43]. In addition, Wang, et al.^[44], Zhao and Xiao^[45], Wan and Tian^[46], Liu, et al.^[47], Xiao, et al.^[48], and Zhang, et al.^[49] have also improved the benchmark HAR model to varying degrees by introducing external factors such as trade friction, noise factors, option trading, price gap, U.S. stock markets, and global stock markets, and further improved the prediction accuracy of the HAR model in the actual application of the Chinese stock market in different scenarios. Similarly, Clements and Preve^[50], Plakandaras, et al.^[51], and Li, et al.^[52] have further expanded this research idea to various major global financial markets represented by the United States stock market, and have comprehensively confirmed the application potential of HAR-type models in high-frequency forecasting and financial transactions through modifications to influencing factors and estimation methods.

To sum up, relying on the benchmark HAR model, many scholars have formed rich research results in the field of high-frequency volatility prediction in recent years. Especially for the discussion of long-memory characteristics, jumping components, and high-frequency measurement errors, many scholars have modified the benchmark HAR from different perspectives and further improved the applicability and prediction accuracy of the model. However, it is worth noting that in previous studies, the selection of research targets often only focused on repre-

sentative indexes of individual markets or individual stocks in special component sectors, that is, only considering the impact of special industry sectors or market factors, while ignoring key considerations for other systematic factors. Therefore, the relevant research conclusions or the constructed optimal HAR model are only limited to the ability to explain and predict a portion of the factor premium, and do not have a truly universal representation. Previous research results have shown that in addition to market factors, there are also various types of systematic factors in China's A-share market, and market factors cannot represent all systematic risks. Therefore, based on the high-frequency RP-PCA factor combination that relies on the implicit factor structure and can capture more asset pricing information, this article further improves and constructs a prediction and early warning framework for the A-share market, thereby providing a more comprehensive and reliable volatility measurement result on asset pricing and risk management in the A-share market from a global perspective. Compared to previous studies, the main contributions of this paper are as follows: 1) Compared with previous studies on index volatility forecasting, which are limited to the overall trend of A-share market such as SSE, SZSE or CSI 300, this paper constructs and extracts the systematic components of the A-share market from a global perspective and systematically studies the volatility characteristics of each factor premium, thus giving a more general and representative high-frequency forecasting framework and forecasting results that are suitable for the A-share market. 2) The effects of long memory, jump structures, and measurement errors that may exist in predicting the volatility characteristics of each factor premium are comprehensively investigated. By using different jump volatility separation methods, we also corroborate the application prospects of HAR-type models in the A-share market and their optimal forecasting model architecture from multiple perspectives. At the same time, this paper also comprehensively confirms the serious shortcomings of early warning studies using a single parameter combination or a single prediction model in previous studies by comparing the differences among the optimal prediction architectures of various factors. Overall, in this study, we not only highlight the key shortcomings of considering only market factors, but also provide a diversified prediction framework and more generally representative prediction results and early warning mechanisms in relevant high-frequency volatility prediction methods.

The remainder of the article is organized as follows: Section 2 is the theoretical model setting, which describes in detail the theoretical framework and specific measurement methods of the factor model relied on in this article, as well as the various types of high-frequency volatility prediction models constructed in this article; Section 3 is empirical research, which comprehensively compares the volatility characteristics of various factors, as well as their prediction errors and performance under different volatility prediction models; Section 4 is the research conclusion.

2 Methodology

2.1 Factor Model

In order to avoid the problems of model uncertainty and omission of variables in the previous factor combination construction process^[53], this paper constructs a systematic factor combination based on the RP-PCA implicit factor extraction method as shown in Equation (1) with

reference to the study of Lettau and Pelger^[54].

$$\hat{F}, \hat{\Lambda} = \operatorname{argmin}_{\Lambda, F} \frac{1}{NM} \sum_{i=1}^N \sum_{j=0}^M (X_{ij} - F_j \Lambda_i^\top)^2 + \gamma \frac{1}{N} \sum_{i=1}^N (\bar{X}_i - \bar{F} \bar{\Lambda}_i^\top)^2 \quad (1)$$

where γ measures the weight ratio between cross-sectional pricing error and time-series pricing error and has $\gamma \geq -1$. F and Λ are systematic factors and factor loadings, respectively. N and M represent the number of assets and the frequency of high-frequency observations in a trading day, respectively. Previous studies have shown that compared to traditional methods represented by the Fama-French multifactor model, the factor construction method based on RP-PCA can not only better avoid the parameter error caused by subjective setting bias, but also obtain factor combinations with stronger asset pricing ability^[55, 56]. In the specific factor identification process, the perturbation eigenvalue ratio criterion as shown in Equation (2) is used to ensure the effective separation of the systematic spectrum for the case of strong and weak factor mixture.

$$\hat{k}(\iota) = \max\{k \leq N-1 : \text{ER}_k > 1 + \iota\}, \quad (2)$$

where, k is the number of factors and ι is the truncation parameter. ER is the ratio of perturbed eigenvalues and has $\text{ER}_k = \hat{\lambda}_k / \hat{\lambda}_{k+1}$, where $\hat{\lambda}_k = \lambda_k + \sqrt{N} * \text{median}\{\lambda_1, \dots, \lambda_N\}$. Referring to Pelger's study^[55], in this paper, the parameters $\iota = 0.1$ and $k = 5$ are set to ensure for the effective identification and extraction of the systematic spectrum.

2.2 High Frequency Volatility Forecasting Model

In this study, the definitions of the realized volatility RV, Bipower variation BV, and the modified high-frequency continuous component CV are shown in Equations (3)~(5), respectively. Among them, in the calculation of CV, each factor premium is adjusted by TOD (Time-of-Day) as shown in Equation (6) to eliminate the effect of cyclical characteristics in the return series.

$$\text{RV}_t^{(i)} = \sum_{j=tn+1}^{tn+n} \left| \Delta_j^n p^{(i)} \right|^2, \quad (3)$$

$$\text{BV}_t^{(i)} = \frac{\pi}{2} \sum_{j=tn+2}^{tn+n} \left| \Delta_j^n p^{(i)} \right| \left| \Delta_{j-1}^n p^{(i)} \right|, \quad (4)$$

$$\text{CV}_t^{(i)} = \sum_{j=tn+1}^{tn+n} \left| \Delta_j^n p^{(i)} \right|^2 \mathbb{1} \left(\left| \Delta_j^n p^{(i)} \right| \leq \alpha_j^{(i)} n^{-\varpi} \right), \quad (5)$$

$$\text{TOD}_j^{(i)} = \frac{n \sum_{t=1}^T \left| \Delta_{jt}^n p \right|^2 \mathbb{1} \left(\left| \Delta_{jt}^n p^{(i)} \right| \leq \tau \sqrt{\text{BV}_t^{(i)} \wedge \text{RV}_t^{(i)}} n^{-\varpi} \right)}{\sum_{s=1}^{nT} \left| \Delta_s^n p \right|^2 \mathbb{1} \left(\left| \Delta_s^n p^{(i)} \right| \leq \tau \sqrt{\text{BV}_{[s/n]}^{(i)} \wedge \text{RV}_{[s/n]}^{(i)}} n^{-\varpi} \right)}, \quad (6)$$

where $\mathbb{1}$ is the indicator function, and $\alpha_j^{(i)} = \tau \sqrt{\left(\text{BV}_{[j/n]}^{(i)} \wedge \text{RV}_{[j/n]}^{(i)} \right) * \text{TOD}_{j-[j/n]n}^{(i)}}$. Referring to the results of extensive numerical simulation analysis by Bollerslev, et al.^[57], in this paper, threshold parameters $\varpi = 0.49$ and $\tau = 3.5$ are set respectively to ensure effective separation of

high-frequency jump components. On this basis, this paper has constructed six types of high-frequency prediction models, represented by HAR, HAR-CJ, HAR-BVJ, HARQ, HARQ-CJ, and HARQ-BVJ, which respectively consider different volatility characteristics. The theoretical models are shown in Equations (7)~(12), respectively.

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + \varepsilon_t, \quad (7)$$

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + \alpha_1 CJ_{t-1} + \alpha_2 CJ_{t-1|t-5} + \alpha_3 CJ_{t-1|t-22} + \varepsilon_t, \quad (8)$$

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + \alpha_1 BVJ_{t-1} + \alpha_2 BVJ_{t-1|t-5} + \alpha_3 BVJ_{t-1|t-22} + \varepsilon_t, \quad (9)$$

$$RV_t = \beta_0 + (\beta_1 + \gamma_1 RQ_{t-1}^{1/2}) RV_{t-1} + (\beta_2 + \gamma_2 RQ_{t-1|t-5}^{1/2}) RV_{t-1|t-5} + (\beta_3 + \gamma_3 RQ_{t-1|t-22}^{1/2}) RV_{t-1|t-22} + \varepsilon_t, \quad (10)$$

$$RV_t = \beta_0 + (\beta_1 + \gamma_1 RQ_{t-1}^{1/2}) RV_{t-1} + (\beta_2 + \gamma_2 RQ_{t-1|t-5}^{1/2}) RV_{t-1|t-5} + (\beta_3 + \gamma_3 RQ_{t-1|t-22}^{1/2}) RV_{t-1|t-22} + \alpha_1 CJ_{t-1} + \alpha_2 CJ_{t-1|t-5} + \alpha_3 CJ_{t-1|t-22} + \varepsilon_t, \quad (11)$$

$$RV_t = \beta_0 + (\beta_1 + \gamma_1 RQ_{t-1}^{1/2}) RV_{t-1} + (\beta_2 + \gamma_2 RQ_{t-1|t-5}^{1/2}) RV_{t-1|t-5} + (\beta_3 + \gamma_3 RQ_{t-1|t-22}^{1/2}) RV_{t-1|t-22} + \alpha_1 BVJ_{t-1} + \alpha_2 BVJ_{t-1|t-5} + \alpha_3 BVJ_{t-1|t-22} + \varepsilon_t, \quad (12)$$

where $RV_{t-1|t-5}$ and $RV_{t-1|t-22}$ are the weekly and monthly realized volatility, respectively, and their calculation methods are shown in Equations (13) and (14). The calculation of the jump component and the realized quality component is the same. β_0 , β_1 , β_2 and β_3 are the intercept terms and the estimated coefficients of daily, weekly, and monthly realized volatility, respectively. α_1 , α_2 and α_3 are estimated coefficients for daily, weekly, and monthly jump volatility, respectively. Among them, there are two methods for dividing high-frequency jump components based on BV and CV. γ_1 , γ_2 and γ_3 represent the estimated coefficients of the daily, weekly, and monthly error correction terms based on the realized quality, respectively. The calculation method of RQ is shown in Equation (15).

$$RV_{t-1|t-5} = \frac{1}{5} \sum_{i=1}^5 RV_{t-i}, \quad (13)$$

$$RV_{t-1|t-22} = \frac{1}{22} \sum_{i=1}^{22} RV_{t-i}, \quad (14)$$

$$RQ_t = \frac{M}{3} \sum_{i=1}^M r_{t,i}^4, \quad (15)$$

where r is the factor premium. Using the above diverse settings, we can examine the high-frequency volatility prediction architecture applicable to different factors, and conduct a specific

comparison of prediction results and model selection process by combining the root mean square error of Equation (16) and the general prediction error evaluation criteria represented by the quasi likelihood function type loss function of Equation (17).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\text{RV}_t - \widehat{\text{RV}}_t)^2}, \quad (16)$$

$$\text{QLIKE} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\text{RV}_t}{\widehat{\text{RV}}_t} - \ln \left(\frac{\text{RV}_t}{\widehat{\text{RV}}_t} \right) - 1 \right), \quad (17)$$

where, $\widehat{\text{RV}}_t$ is the forecast result of realized volatility. Overall, the root mean square error provides a universal prediction error criterion based on a symmetric loss mechanism, while QLIKE provides a more robust prediction error criterion by constructing an asymmetric loss function that imposes a greater penalty mechanism on volatility undervaluation.

3 Empirical Results

3.1 Data and Descriptive Statistics

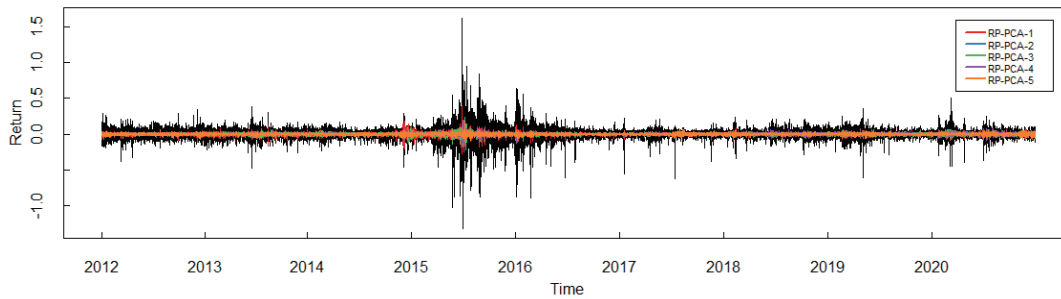
In this article, we have selected all stocks represented by the CSI All Share Index as the research object to avoid sample bias caused by the fact that individual characteristic indexes can only cover individual industry sectors. The sample interval is selected from January 1, 2012 to December 31, 2020, with a total of 2188 trading days during the sample period. The high-frequency observation interval is chosen to be 5 minutes to avoid the adverse effects of excessive microstructure noise caused by higher observation frequencies. The high-frequency observation time interval is from 9:30 am to 11:30 am and from 1:00 pm to 3:00 pm every day, with a total of approximately 76,037,000 intraday high-frequency trading observations during the sample period. The risk-free rate is the one-year time deposit rate announced by the People's Bank of China. All raw data are obtained from the Tinysoft database. The descriptive statistics of the premiums of each factor are shown in Table 1. The average five-minute return and standard deviation of each factor are annualized based on 251 trading days per year.

Table 1 shows that, first of all, RP-PCA-1 has the highest risk premium level and corresponding volatility characteristics, which again confirms the conclusion that most of the systematic risk found in previous studies has been explained in the market factor. Secondly, observing the kurtosis and skewness of each factor, we can see that the premium levels of each systematic factor have obvious left skewness and spikes with thick tails. Combining the results of the Q -statistic with the lagged 22 periods and Figure 1, it can be further seen that the factors have obvious long time series memory and volatility clustering characteristics. That is, in various time windows, there is often a trend characteristic of large fluctuations followed by large fluctuations, and small fluctuations followed by small fluctuations, especially during the period of surge and crash of A-shares in 2015 and 2016. Similar conclusions on volatility characteristics have also been reflected in the studies of Chen, et al.^[58], Tang and Zhu^[59], and Liu and He^[60]. Finally, from the ADF test results, it can be seen that the return series of each factor is a stationary series, which satisfies the prerequisite requirements for further modeling.

Table 1 Descriptive statistical results of each factor premium

	RP-PCA-1	RP-PCA-2	RP-PCA-3	RP-PCA-4	RP-PCA-5
Mean	0.22	0.02	0.01	0.02	0.01
Std.Dev	0.84	0.19	0.14	0.12	0.11
Min	-1.33	-0.27	-0.22	-0.25	-0.10
Max	1.62	0.40	0.23	0.19	0.19
Skewness	-0.42	-0.30	-0.05	-0.07	-0.49
Kurtosis	40.06	50.08	29.64	51.78	18.11
$Q(22)$	527.18*** (0.00)	2291.20*** (0.00)	1743.80*** (0.00)	2075.00*** (0.00)	3255.30*** (0.00)
ADF	-44.49**	-43.02**	-39.58**	-43.11**	-46.09**
Observations	105024	105024	105024	105024	105024

Notes: 1) *** and ** denote statistical significance at 1% and 5% significance levels, respectively.
 2) $Q(22)$ indicates the result of Ljung-Box test to test serial autocorrelation.

**Figure 1** The premium sequence of each factor based on the five-minute high-frequency observation window

3.2 Volatility Estimation Results

As mentioned above, the premium volatility level of each factor is estimated by six types of HAR family models represented by HAR, HAR-CJ, HAR-BVJ, HARQ, HARQ-CJ and HARQ-BVJ. In order to compare the differences in the volatility characteristics of each factor, Table 2 shows the estimation results of each factor based on different model settings in the full sample interval.

Table 2 Parameter estimation results based on the full sample interval

	β_0	β_1	β_2	β_3	α_1	α_2	α_3	γ_1	γ_2	γ_3
RP-PCA-1 Factor										
HAR	0.01* (0.08)	0.31*** (0.00)	0.30*** (0.00)	0.31*** (0.00)						
HAR-CJ	0.01* (0.08)	0.35*** (0.00)	0.26*** (0.00)	0.32*** (0.00)	-0.35*** (0.00)	0.35 (0.18)	-0.10 (0.86)			
HAR-BVJ	0.01 (0.22)	0.36*** (0.00)	0.24*** (0.00)	0.78*** (0.00)	-0.56*** (0.00)	0.15 (0.68)	-3.54*** (0.00)			
HARQ	-0.00 (0.77)	0.54*** (0.00)	0.21** (0.01)	0.33*** (0.00)				-0.04*** (0.00)	0.01 (0.28)	-0.03 (0.15)
HARQ-CJ	-0.01 (0.34)	0.53*** (0.00)	0.24*** (0.00)	0.28*** (0.00)	-0.15 (0.13)	0.37 (0.17)	1.31* (0.06)	-0.04*** (0.00)	0.00 (0.84)	-0.06** (0.02)
HARQ-BVJ	-0.02** (0.01)	0.56*** (0.00)	0.19** (0.03)	1.01*** (0.00)	-0.40** (0.01)	-0.05 (0.90)	-4.74*** (0.00)	-0.04*** (0.00)	0.01 (0.54)	-0.09*** (0.00)
RP-PCA-2 Factor										
HAR	0.00*** (0.00)	0.35*** (0.00)	0.46*** (0.00)	0.03 (0.45)						
HAR-CJ	0.00*** (0.00)	0.53*** (0.00)	0.30*** (0.00)	0.12** (0.01)	-1.36*** (0.00)	1.17*** (0.00)	-1.56*** (0.00)			
HAR-BVJ	0.00** (0.02)	0.34*** (0.00)	0.43*** (0.00)	0.26*** (0.00)	-0.03 (0.81)	0.24 (0.28)	-1.74*** (0.00)			
HARQ	-0.00 (0.48)	0.57*** (0.00)	0.38*** (0.00)	0.20** (0.01)				-0.46*** (0.00)	0.07 (0.68)	-1.36*** (0.00)
HARQ-CJ	0.01 (0.74)	0.46*** (0.00)	0.42*** (0.00)	0.25*** (0.00)	-1.45*** (0.00)	1.36*** (0.00)	-0.91* (0.07)	0.14 (0.12)	-0.36* (0.09)	-1.18*** (0.00)
HARQ-BVJ	-0.00 (0.48)	0.59*** (0.00)	0.36*** (0.00)	0.17** (0.04)	0.61*** (0.00)	-0.80** (0.02)	-0.95 (0.10)	-0.84*** (0.00)	-0.25 (0.36)	-0.62 (0.23)
RP-PCA-3 Factor										
HAR	0.00* (0.08)	0.06** (0.03)	0.73*** (0.00)	0.06 (0.13)						
HAR-CJ	0.00** (0.01)	-0.00 (0.91)	0.61*** (0.00)	0.42*** (0.00)	1.63*** (0.00)	-0.15 (0.68)	-4.28*** (0.00)			
HAR-BVJ	0.00 (0.18)	0.03 (0.30)	0.60*** (0.00)	0.08 (0.21)	2.22*** (0.00)	1.48** (0.02)	-0.75 (0.45)			
HARQ	-0.00*** (0.00)	0.08* (0.09)	0.63** (0.01)	0.66*** (0.00)				-0.26 (0.27)	0.41 (0.45)	-6.82*** (0.00)
HARQ-CJ	-0.00*** (0.00)	0.01 (0.82)	0.61*** (0.00)	0.78*** (0.00)	1.58*** (0.00)	0.09 (0.81)	-3.19*** (0.00)	-0.16 (0.49)	-0.27 (0.62)	-5.30*** (0.00)
HARQ-BVJ	0.00*** (0.00)	-0.03 (0.61)	0.53*** (0.00)	0.80*** (0.00)	2.23*** (0.00)	0.79 (0.22)	-2.95*** (0.00)	0.17 (0.46)	0.52 (0.33)	-6.67*** (0.00)

Table 2 (continued)

	β_0	β_1	β_2	β_3	α_1	α_2	α_3	γ_1	γ_2	γ_3
RP-PCA-4 Factor										
HAR	0.00*** (0.00)	0.71*** (0.00)	0.07** (0.03)	0.03 (0.27)						
HAR-CJ	0.00*** (0.00)	0.72*** (0.00)	0.07 (0.11)	0.42 (0.27)	-0.04 (0.76)	0.04 (0.92)	1.65* (0.06)			
HAR-BVJ	0.00*** (0.00)	0.69*** (0.00)	-0.38*** (0.00)	0.40*** (0.00)	0.62*** (0.00)	3.75*** (0.00)	-3.20*** (0.00)			
HARQ	-0.00** (0.02)	0.92*** (0.00)	0.85*** (0.00)	-0.53*** (0.00)				-1.22*** (0.00)	-4.17*** (0.00)	3.01*** (0.00)
HARQ-CJ	-0.00** (0.01)	0.90*** (0.00)	1.02*** (0.00)	-0.60*** (0.00)	0.05 (0.68)	-1.55*** (0.00)	0.13 (0.88)	-1.32*** (0.00)	-4.47*** (0.00)	3.66*** (0.00)
HARQ-BVJ	-0.00** (0.03)	0.77*** (0.00)	0.66*** (0.00)	-0.22* (0.05)	-0.67*** (0.00)	-4.59*** (0.00)	2.94*** (0.00)	-0.67*** (0.00)	-4.59*** (0.00)	2.94*** (0.00)
RP-PCA-5 Factor										
HAR	0.00*** (0.00)	0.09*** (0.00)	0.57*** (0.00)	0.14*** (0.00)						
HAR-CJ	0.00** (0.00)	0.08*** (0.00)	0.59*** (0.00)	0.09* (0.09)	0.06 (0.59)	-0.47* (0.06)	0.99** (0.02)			
HAR-BVJ	0.00*** (0.00)	0.12*** (0.00)	0.59*** (0.00)	0.13 (0.11)	-0.27 (0.12)	-0.30 (0.48)	0.09 (0.90)			
HARQ	0.00 (0.71)	0.22*** (0.00)	0.67*** (0.00)	0.15* (0.05)				-1.15*** (0.00)	-3.57*** (0.00)	-4.43** (0.04)
HARQ-CJ	0.00 (0.68)	0.24*** (0.00)	0.74*** (0.00)	0.03 (0.75)	-0.09 (0.39)	0.80*** (0.00)	1.06** (0.01)	-1.26*** (0.00)	-4.68*** (0.00)	-2.11 (0.35)
HARQ-BVJ	0.00 (0.64)	0.24*** (0.00)	0.69*** (0.00)	0.13 (0.16)	-0.24 (0.15)	-0.27 (0.52)	0.14 (0.84)	-0.18*** (0.00)	-3.67*** (0.00)	-4.52** (0.04)

Notes: 1) ***, **, * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

2) Affected by precision limitations and approximate values, 0.00 does not mean that the actual value of the estimated coefficient is 0.

Specifically, first of all, the estimation results of RP-PCA-1 factors representing the overall trend of the market indicate that their factor premiums have significant long-term volatility memory characteristics. That is, the realized volatility estimation coefficients β_1 , β_2 and β_3 representing daily, weekly, and monthly volatility respectively have significant statistical significance. This conclusion remains stable even after the separation of jump components or the introduction of correction terms. Secondly, from the results after separating the jump compo-

nents, when relying on the BV classification criteria, more jump components that can provide effective prediction information are mainly concentrated in daily and monthly periods, while weekly jumps do not contribute significantly to the estimation of realized volatility. That is to say, investors who anchor market premiums focus more on the measurement and management of near and distal risk, and less on the medium-term changes in factor premiums. However, for dividing jump volatility by CV criteria, it is more focused on proximal risk management. That is, after correcting the jump volatility undervaluation phenomenon, the volatility level of the daily realized volatility has further improved, prompting more investors to focus more on the current investment portfolio. Finally, from the perspective of error correction terms, after adding realized quality to the benchmark HAR, a significantly negative error correction term coefficient exhibits its own significant mean-reversion process. That is, with the gradual reduction of the measurement error, the corresponding prediction information about RV is further improved, and its prediction accuracy is continuously improved through the error correction process. At the same time, comparing the error correction coefficients under different term structures, it can be found that more error processes are mainly concentrated in the proximal risk represented by the daily degree, which is also consistent with the research conclusions of Bollerslev, et al.^[39]. Moreover, it can be seen that when the jump volatility component is further introduced based on error correction, the prediction contribution of its continuous component is still significant. However, under the CV-based jump stripping method, the contribution of its proximal jump volatility gradually disappears, but only has a certain impact in the remote period represented by the monthly time interval. However, for the BV-based volatility jump identification method, its research conclusions on the contribution of jump volatility are still robust. In terms of absolute impact, after adding the error correction term, the contribution of each jump component further increases at the far end. Correspondingly, the error correction process at the remote end also presents a trend of synchronous improvement.

Secondly, from the estimated results of RP-PCA-2 factors, it shows a significant difference in volatility characteristics from RP-PCA-1. Firstly, comparing the baseline HAR model, it can be found that the significance of the realized volatility coefficient at the far end is changed after adding the jump component or the influence of the error correction process. That is, after considering the impact of jumps or measurement errors, investors' attention to the process of long-term risk compensation is further enhanced, thereby further changing their measurement structure for realized volatility. Secondly, after separating the components of jumping volatility, the classification criteria based on CV will consider all the characteristics of jumping volatility in the short-term, mid-term and long-term, while in the division standard based on BV, this phenomenon only exists in the consideration of long-term fluctuations. The significantly different estimation results suggest that the strong assumption of simultaneous jumps in successive five-minute windows does not exist and that this assumption tends to result in significant underestimation of volatility, which is not applicable in the estimation of premium volatility for factor 2. Finally, from the perspective of the error correction process, the significantly negative error correction coefficient satisfies the initial assumptions of the theoretical model, which further proves the rationality of the framework setting for the relevant volatility measurement model. Comparing the results of further introducing the impact of measurement errors based on the

separation of jump components, it can be found that the HARQ-BVJ, HARQ, and HARQ-CJ model combinations sequentially increase the dependence on daily volatility, daily and monthly volatility, and weekly and monthly volatility, respectively. Correspondingly, the correction process of measurement errors also shows a synchronous pattern. In terms of the degree of error correction, compared to the RP-PCA-1 factor, the premium volatility measurement process of the RP-PCA-2 factor further improves the degree of correction for its own measurement error, thus accelerating the mean-reversion process of its RV and showing a significantly different model parameter structure from that of factor 1.

Thirdly, the estimation results from the RP-PCA-3 factor show a greater focus on constructing the potential impact of short-term volatility trends and medium-term volatility trends on RV in the benchmark HAR model, while discarding the combination of parameters that continue to incorporate distal volatility features, i.e., excluding longer-term volatility memory features. This phenomenon is not significantly improved even after the introduction of jump volatility or the effect of measurement error. Specifically, first of all, in the model combination of HAR-CJ with CV as the classification criterion, although it further enhances the impact of distal volatility on RV, it loses the information of near-term continuous volatility. In the model architecture of HAR-BVJ based on BV indicators, it not only fails to effectively incorporate the impact of remote volatility characteristics, but also directly abandons the consideration of daily continuous volatility components, and indeed only incorporates the medium-term impact represented by weekly time intervals. Secondly, this phenomenon remains evident after further introducing the effects of measurement error. That is, it only increases the proportion of medium-term and long-term parameter weights, but fails to take into account the impact of near-term fluctuations. Correspondingly, in the entire process of dynamic error correction, it only pays more attention to the detailed consideration of monthly measurement errors. Overall, compared with different model architectures, the introduction of jump components and measurement errors in the volatility measurement architecture of RP-PCA-3 factors does not seem to be able to provide it with a more effective and comprehensive description of volatility characterization process.

Fourthly, from the estimation results of RP-PCA-4, although it has a relatively similar benchmark HAR architecture to RP-PCA-3, it achieves more effective retention of original premium volatility information. Specifically, first of all, after only incorporating the impact of jump fluctuations, the HAR-BVJ based model framework exhibits a more comprehensive parameter architecture, which comprehensively examines the differential volatility characteristics in different time intervals of daily, weekly and monthly. However, in terms of the estimated coefficients, the volatility impact of the jump component is much larger than its continuous volatility component, which is significantly different from the results of the other factors measured under the HAR-BVJ-based framework. Secondly, after further supplementing the effect of error correction terms, compared to the HARQ model that only considers measurement errors and the HARQ-CJ hybrid model based on the CV division standard, the model architecture of HARQ-BVJ based on the BV standard shows the best fitting results. That is, compared to the HARQ and HARQ-CJ models, HARQ-BVJ retains more information about continuous volatility components, jump volatility components, and error correction processes in various short and long periods, thereby providing a more comprehensive parameter framework for characterizing

RP-PCA-4 factor premium volatility. However, it is important to note that even after adding an error correction term, the jump component still maintains a larger influence coefficient compared to the continuous fluctuation component, and simultaneously improves the measurement error correction level in different periods.

Finally, the estimation results of RP-PCA-5 factors indicate that the benchmark HAR model can better fit the volatility characteristics of its factor premium. After further introducing the impact of jump volatility components and error correction terms, its volatility information about factor premiums still emphasizes the measurement of short-term and medium-term characteristics, with a certain degree of model robustness. In the actual stock market, it is reflected in the fact that there are more short-term investors focused on capturing risk compensation after exposure to the RP-PCA-5 factor.

Overall, the above research results show the fitting results of different factors under different volatility characteristics assumptions. By comparing the calculation results, it can be found that each factor has significant characteristic differences in both the term structure of premium volatility characteristics and the impact of jump components and measurement errors. Therefore, premium characteristics for different systematic factors should be combined with different volatility models to provide more accurate volatility forecasting results, rather than relying solely on a single fixed parameter structure.

3.3 Volatility Forecast Results

As mentioned above, the premium volatility level of each factor is estimated by six types of HAR family models represented by HAR, HAR-CJ, HAR-BVJ, HARQ, HARQ-CJ and HARQ-BVJ. In order to compare the differences in the volatility characteristics of each factor, Table 2 shows the estimation results of each factor based on different model settings in the full sample interval.

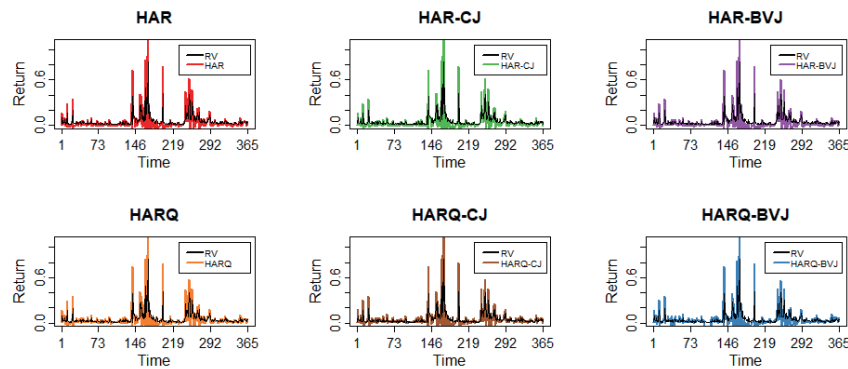
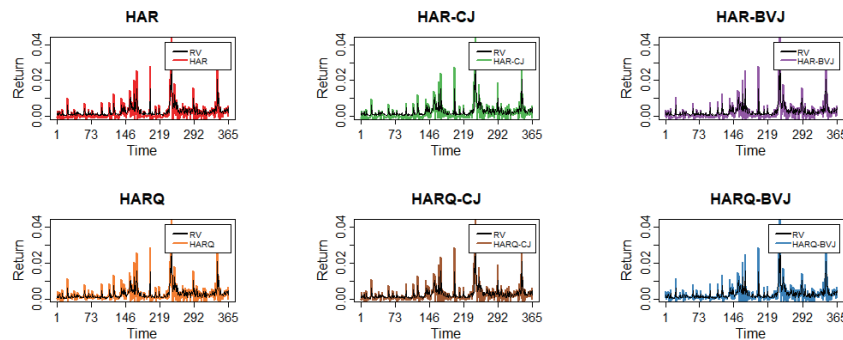
Based on the above research results, in this part of the study, we further comprehensively compared the out-of-sample prediction results of each systematic factor under the assumption of different volatility characteristic structures. Specifically, this paper uses a time window with an interval length of 1823 to gradually update the model parameters and predict the out-of-sample realized volatility for the next 365 trading days. The average forecast error of each model at each out-of-sample time point is compared and analyzed to give the optimal high-frequency volatility forecasting architecture for each RP-PCA factor. The specific calculation results are shown in Table 3.

From Table 3 and Figures 2~6, it can be seen that, taking into account different prediction error evaluation criteria, the above models better depict the premium volatility characteristics of different factors and provide accurate out-of-sample prediction results. Specifically, first of all, for the RP-PCA-1 factor, there does not seem to be a unique optimal prediction architecture under different prediction error criteria. Comparatively speaking, both the benchmark HAR model and the HAR-Q model considering the impact of measurement errors can better predict the expected volatility of the RP-PCA-1 factor premium. However, further comparing the prediction results of the HAR model and the HAR-Q model under the error criteria of QLIKE and RMSE, it can be found that compared to the optimal prediction results under the same prediction error index, the prediction errors of the HAR model and the HAR-Q model are

Table 3 Out-of-sample prediction results under time-varying parameters

		HAR	HAR-CJ	HAR-BVJ	HARQ	HARQ-CJ	HARQ-BVJ
RP-PCA-1	RMSE	0.0667(1)	0.0669(2)	0.0696(5)	0.0686(3)	0.0689(4)	0.0780(6)
	QLIKE	0.2236(2)	0.2250(3)	0.2369(4)	0.2142(1)	0.2676(5)	0.8631(6)
RP-PCA-2	RMSE	0.0027(2)	0.0027(4)	0.0026(1)	0.0027(5)	0.0027(3)	0.0027(6)
	QLIKE	0.2178(3)	0.3564(6)	0.2049(2)	0.2007(1)	0.2380(5)	0.2255(4)
RP-PCA-3	RMSE	0.0021(1)	0.0023(2)	0.0027(5)	0.0024(3)	0.0026(4)	0.0028(6)
	QLIKE	0.0997(1)	0.1119(2)	0.1420(5)	0.1224(3)	0.1252(4)	0.1471(6)
RP-PCA-4	RMSE	0.0017(1)	0.0017(2)	0.0019(3)	0.0022(5)	0.0023(6)	0.0021(4)
	QLIKE	0.1176(1)	0.1221(2)	0.1849(3)	1.0559(6)	0.5597(5)	0.4969(4)
RP-PCA-5	RMSE	0.0015(6)	0.0015(5)	0.0015(4)	0.0014(3)	0.0014(1)	0.0014(2)
	QLIKE	0.1964(5)	0.1983(6)	0.1953(4)	0.1866(3)	0.1865(2)	0.1854(1)

Notes: 1) The numbers in () represent the specific ranking of the prediction accuracy of each high-frequency volatility model under the specified error evaluation criteria. 2) Due to the limitation of precision and the influence of approximate value, some forecast error results with the same value on the surface are not exactly equal in fact.

**Figure 2** Out-of-sample prediction results of RP-PCA-1**Figure 3** Out-of-sample prediction results of RP-PCA-2

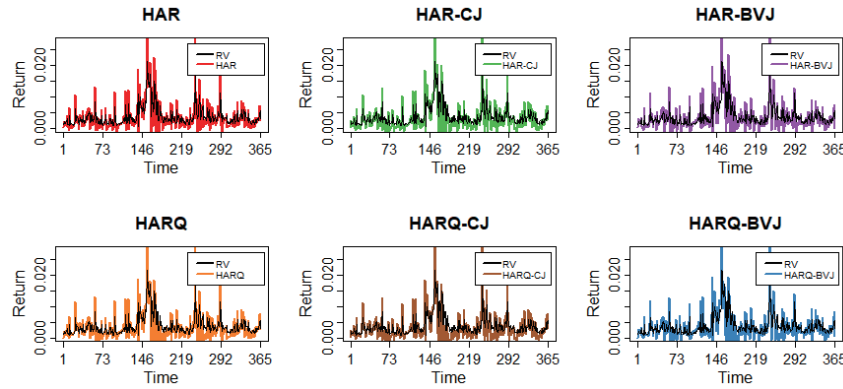


Figure 4 Out-of-sample prediction results of RP-PCA-3

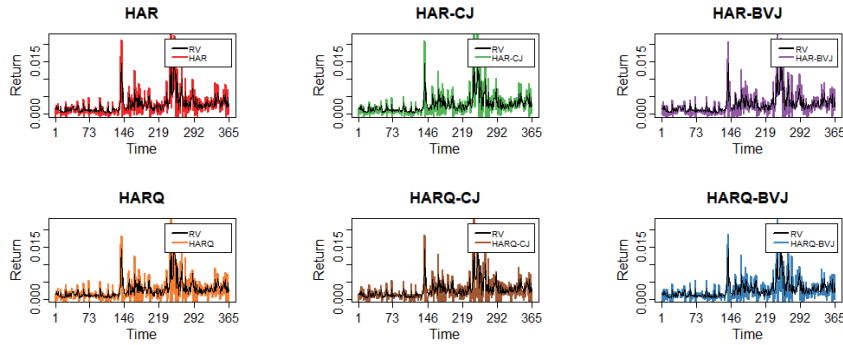


Figure 5 Out-of-sample prediction results of RP-PCA-4

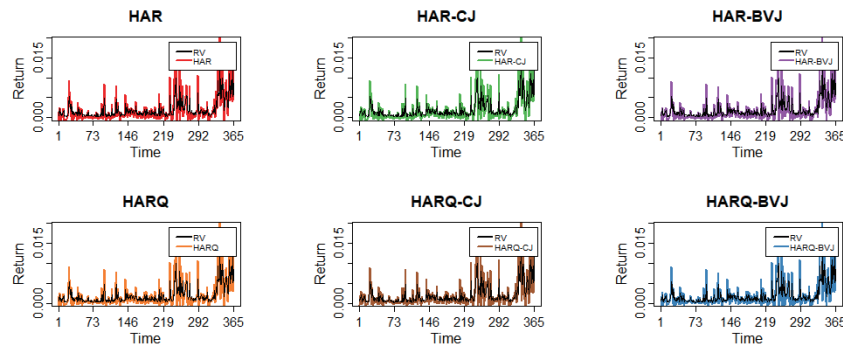


Figure 6 Out-of-sample prediction results of RP-PCA-5

respectively increased by 4.39% and 2.85%. That is, overall, the HAR-Q based model architecture provides more accurate prediction results for RP-PCA-1 factor.

Secondly, the prediction results of RP-PCA-2 show that both the HAR-BVJ model considering BV as the dividing criterion for volatility jump components and the HAR-Q model considering the error correction process can provide volatility prediction results with small prediction errors. In contrast, HAR-BVJ exhibits optimal prediction performance under the RMSE criteria. However, under the evaluation criteria of QLIKE, HAR-Q should be chosen as the optimal prediction model. However, a further comparison of the prediction performance of HAR-BVJ and HAR-Q under QLIKE and RMSE standards can be found that HAR-Q has

expanded the prediction error by 3.85% compared to the best prediction results under the same prediction error standard, while under the same conditions, HAR-BVJ has only expanded the prediction error by 2.09%. That is, HAR-BVJ, which takes into account jump fluctuations, better depicts and predicts the volatility of the factor premium of RP-PCA-2.

Thirdly, from the volatility prediction results of RP-PCA-3, it can be seen that using the benchmark HAR model is sufficient to fully simulate the volatility characteristics of its factor premium and provide the corresponding optimal prediction results. In contrast, after introducing jump components or error correction terms, the prediction accuracy of the model has not been further improved, and even further expanded the prediction error. This phenomenon indicates that there are not too many extreme returns in the factor premium series of RP-PCA-3. The relatively stable factor premium process ensures the prediction effectiveness and robustness of continuous volatility components, rather than focusing more on the consideration of individual jump characteristics. This factor premium characteristic of RP-PCA-3 and related research conclusions can also be reflected and confirmed in the descriptive statistical results of its samples and the fitting results within the samples. Similarly, RP-PCA-4 also exhibits a relatively consistent estimation result with RP-PCA-3, that is, relying solely on the benchmark HAR model can better depict accurate prediction information about its future volatility trend.

Finally, from the estimation results of RP-PCA-5, different from the above factors, its optimal volatility prediction framework relies on the HARQ-CJ model and the HARQ-BVJ model that consider both jump characteristics and error correction items. The prediction performance of the two types of models is relatively similar. In contrast, HARQ-BVJ and HARQ-CJ respectively expand the prediction error by 0.55% and 0.59% compared with the best prediction results under the RMSE and QLIKE standards. Overall, for the volatility prediction process of RP-PCA-5 factor premium, HARQ-BVJ is a better choice.

In addition, a further comparison of the prediction performance of each volatility model on different factors shows that the benchmark HAR model provides the most number of best prediction results, followed by the HARQ model that considers the error correction process. In contrast, although the jump model based on the BV classification criteria shows the best prediction performance for some individual factors, a comprehensive comparison of the prediction results of all factors shows that the jump structure divided based on the CV index has a smaller prediction error compared to the BV on the whole. However, for the volatility prediction model that further consider both the jump structure and the error correction term, its best prediction performance only exists in some individual factors. That is, the inclusion of more complex volatility features often does not mean a significant improvement in the corresponding prediction performance, but rather increases the redundancy of the entire prediction architecture and reduces the generalization ability of its model outside the sample.

Overall, the above research results indicate that there is no uniform and consistent optimal prediction architecture for different systematic factors. From the above research, it can be seen that in addition to RP-PCA-3 and RP-PCA-4, RP-PCA-1, RP-PCA-2, and RP-PCA-5 can also further improve the accuracy of their factor premium volatility prediction by incorporating measurement errors, jump fluctuations, and jump and error correction terms on the basis of the benchmark HAR model, respectively. At the same time, the above results also indicate

that the research conclusions obtained in the study of volatility characteristics and prediction structures represented by the market factor are not generally representative. Therefore, the findings relying only on the RP-PCA-1 factor will not provide a comprehensive picture of the risk compensation characteristics of the A-share market and will not allow for the systematic construction of a corresponding volatility risk warning mechanism. In addition, the controversy related to the volatility predictive power of the jump component in previous studies is also better explained in our paper. That is, excluding the impact on the selection of market indexes or sample intervals, it can be found that the volatility prediction ability of jump components only exists in some individual factors, rather than being completely effective or ineffective. To sum up, this article constructs and provides a more comprehensive prediction and early warning framework for the overall premium risk and volatility characteristics of the A-share market.

4 Conclusion

In the above study, this paper constructs a more representative RP-PCA five-factor combination based on the frontier statistical inference theory. Based on this, we construct six types of representative high-frequency forecasting models of HAR, HAR-CJ, HAR-BVJ, HARQ, HARQ-CJ, and HARQ-BVJ that can consider different volatility characteristics, and compare their respective forecasting performance for each factor premium fluctuation, so as to comprehensively analyze the differences of volatility characteristics of each factor premium, and construct and optimize the prediction and early warning framework and model combination for the Chinese stock market, and give a more general representative research conclusion.

The research results show that, first of all, factors representing five different types of systematic risk have significantly different volatility characteristics. In addition to RP-PCA-3 and RP-PCA-4, the prediction of factor premium volatility for RP-PCA-1, RP-PCA-2, and RP-PCA-5 requires further consideration of measurement error and the impact of jump volatility components to obtain and construct a more accurate prediction and warning structure. Secondly, from the estimation results of RP-PCA-1, it can be found that the optimal prediction structure for the market factor and its own premium volatility characteristics are not universally representative. This finding overturns the previous research idea of relying solely on the research of individual representative market indexes to construct the prediction and early warning framework for the overall A-share market. Finally, the results of this paper show that for each premium factor, the introduction of more complex volatility features does not always mean higher forecasting accuracy. On the contrary, in most cases, the benchmark high-frequency forecasting framework is sufficient to fully describe the premium fluctuation characteristics of each factor. In general, this paper constructs a variety of volatility prediction and early warning models for the A-share market, and presents the optimal model architecture and parameter selection scheme for each systematic factor. The relevant research conclusions are not only beneficial for investors to more accurately and rationally identify the risk characteristics of the A-share market, thereby further better managing investment portfolio risks, but also for financial regulatory agencies to further improve their regulatory measures, thereby effectively ensuring the long-term stable development of the A-share market while preventing financial risks and improving market efficiency, which has important practical significance.

References

- [1] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 1986, 31(3): 307–327.
- [2] Engle R F, Bollerslev T. Modeling the persistence of conditional variances. *Econometric Reviews*, 1986, 5: 1–50.
- [3] Schwert G W. Why does stock market volatility change over time? *The Journal of Finance*, 1989, 44: 1115–1153.
- [4] Nelson, Daniel B. Conditional Heteroskedasticity in asset returns: A new approach. *Modelling Stock Market Volatility*, 1991, 59(2): 347–370.
- [5] Higgins M L, Bera A K. A class of nonlinear ARCH models. *International Economic Review*, 1992, 33: 137–158.
- [6] Ding Z, Granger C W J, Engle R F. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1993, 1: 83–106.
- [7] Engle R F, Ng V K. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 1993, 48: 1749–1778.
- [8] Jean-Michel, Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics & Control*, 1994, 18: 931–955.
- [9] Engle R F, Lee G G J. A permanent and transitory component model of stock return volatility. *SSRN* 5848, 1999: 475–497.
- [10] Wang R, Wang H. Correlation analysis of stock market and fund market based on M-Copula-EGARCH-M-GED model. *Journal of Systems Science and Information*, 2020, 8(3): 240–252.
- [11] Roni B, Abbas G, Wang S. Return and volatility spillovers effects: Study of Asian emerging stock markets. *Journal of Systems Science and Information*, 2018, 6(2): 97–119.
- [12] Roni B, Wu C, Jewel R K, et al. A study on the volatility of the Bangladesh stock market-based on GARCH type models. *Journal of Systems Science and Information*, 2017, 5(3): 193–215.
- [13] Andersen T G, Bollerslev T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 1998, 39(4): 885–905.
- [14] Crosi F. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 2004, 2: 174–196.
- [15] Zhu X, Zou J, Han F, et al. Forecasting and modeling of China's nonferrous metal futures market volatility based on the introduction of external shocks: Taking Copper as example. *Chinese Journal of Management Science*, 2018, 26(9): 52–61.
- [16] Zhang Y J, Zhang H. Volatility forecasting of crude oil futures market: Which structural change-based HAR models have better performance? *International Review of Financial Analysis*, 2023, 85: 102454.
- [17] Lyócsa S, Stašek D. Improving stock market volatility forecasts with complete subset linear and quantile HAR models. *Expert Systems with Applications*, 2021, 183(30): 115416.
- [18] Niu Z, Ma F, Zhang H. The role of uncertainty measures in volatility forecasting of the crude oil futures market before and during the COVID-19 pandemic. *Energy Economics*, 2022, 112: 106120.
- [19] Chen W, Ma F, Wei Y, et al. Forecasting oil price volatility using high-frequency data: New evidence. *International Review of Economics & Finance*, 2020, 66: 1–12.
- [20] Özbekler A G, Kontonikas A, Triantafyllou A. Volatility forecasting in European government bond markets. *International Journal of Forecasting*, 2021, 37(4): 1691–1709.
- [21] Li W, Cheng Y, Fang Q. Forecast on silver futures linked with structural breaks and day-of-the-week effect. *The North American Journal of Economics and Finance*, 2020, 53: 101192.
- [22] Jayawardena N I, Todorova N, Li B, et al. Volatility forecasting using related markets' information for the Tokyo stock exchange. *Economic Modelling*, 2020, 90: 143–158.
- [23] Wilms I, Rombouts J, Crous C. Multivariate volatility forecasts for stock market indices. *International Journal of Forecasting*, 2021, 37(2): 484–499.
- [24] Andersen T G, Bollerslev T, Diebold F X. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Review of Economics & Statistics*, 2007, 89(4): 701–720.
- [25] Qiao G, Teng Y, Li W, et al. Improving volatility forecasting based on Chinese volatility index information: Evidence from CSI 300 index and futures markets. *The North American Journal of Economics and Finance*, 2019, 49: 133–151.

- [26] Ma F, Wang J, Guo Y, et al. Forecasting the realized volatility based on jump, jump intensity and regime switching in stock market. *Systems Engineering — Theory & Practice*, 2023, 43(2): 371–382.
- [27] Yuan Y, Zhang T, Zhuang X. Multifractal volatility forecast of Chinese stock market. *Systems Engineering-Theory & Practice*, 2020, 40(9): 2269–2281.
- [28] Guo Y, Ma F. Forecasting the Chinese gold futures market volatility using Markov-switching regime and mixed data sampling model. *Chinese Journal of Management Science*, 2022, DOI: 10.16381/j.cnki.issn1003-207x.2021.0213.
- [29] Ye W, Xia W, Wu B, et al. Using implied volatility jumps for realized volatility forecasting: Evidence from the Chinese market. *International Review of Financial Analysis*, 2022, 83: 102277.
- [30] Ma F, Wahab M I M, Zhang Y. Forecasting the U.S. stock volatility: An aligned jump index from G7 stock markets. *Pacific-Basin Finance Journal*, 2019, 54: 132–146.
- [31] Zeng Q, Lu X, Li T, et al. Jumps and stock market variance during the COVID-19 pandemic: Evidence from international stock markets. *Finance Research Letters*, 2020, 48: 102896.
- [32] Degiannakis S, Filis G, Klein T, et al. Forecasting realized volatility of agricultural commodities. *International Journal of Forecasting*, 2022, 38(1): 74–96.
- [33] Bu R, Hizmeri R, Izzeldin M, et al. The contribution of jump signs and activity to forecasting stock price volatility. *Journal of Empirical Finance*, 2023, 70: 144–164.
- [34] Zhang C, Chen H, Jiang P. Is market microstructure noise really just “noise”? Evidence from the empirical study for the Chinese A stock market. *Journal of Applied Statistics and Management*, 2021, 40(5): 932–950.
- [35] Gong X, Cao J, Wen F, et al. The HAR-type models with leverage and structural breaks and their applications to the volatility forecasting of stock market. *Systems Engineering — Theory & Practice*, 2020, 40(5): 1113–1133.
- [36] Liang C, Li Y, Ma F, et al. Global equity market volatilities forecasting: A comparison of leverage effects, jumps, and overnight information. *International Review of Financial Analysis*, 2021, 75: 101750.
- [37] Shen L, Lu X, Huynh T L D, et al. Air quality index and the Chinese stock market volatility: Evidence from both market and sector indices. *International Review of Economics & Finance*, 2023, 84: 224–239.
- [38] Buncic D, Gislis K I M. The role of jumps and leverage in forecasting volatility in international equity markets. *Journal of International Money and Finance*, 2017, 79: 1–19.
- [39] Bollerslev T, Patton A J, Quaedvlieg R. Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 2016, 192(1): 1–18.
- [40] Cipollini F, Gallo G M, Otranto E. Realized volatility forecasting: Robustness to measurement errors. *International Journal of Forecasting*, 2021, 37(1): 44–57.
- [41] Li J, Wang S, Wei Y. Volatility forecasting models based on volatility measurement error. *Systems Engineering — Theory & Practice*, 2018, 38(8): 1905–1918.
- [42] Wang Y, Liang F, Wang T, et al. Does measurement error matter in volatility forecasting? Empirical evidence from the Chinese stock market. *Economic Modelling*, 2020, 87: 148–157.
- [43] Song Y, Wang X. Modeling and forecasting volatility of Chinese stock market based on dynamic estimation errors. *Chinese Journal of Management Science*, 2017, 25(9): 19–27.
- [44] Wang R, Li W, Huang Y. Trade friction, intraday jump and stock index volatility — Evidence from high-frequency data on Chinese stock indices. *Studies of International Finance*, 2019, 12: 63–73.
- [45] Zhao H, Xiao J. Volatility forecasting in the presence of microstructure noise and measurement error. *Chinese Journal of Management Science*, 2020, 28(4): 48–60.
- [46] Wan D, Tian Y. Can option trading volume forecast volatility? Evidence from Shanghai Stock Exchange 50ETF option. *Systems Engineering — Theory & Practice*, 2023, 43(3): 755–771.
- [47] Liu Y, Qu J, Dong X, et al. The application of signed range to forecasting the volatility of financial. *Systems Engineering — Theory & Practice*, 2021, 41(9): 2256–2270.
- [48] Xiao J, Wen F, Zhao Y, et al. The role of US implied volatility index in forecasting Chinese stock market volatility: Evidence from HAR models. *International Review of Economics & Finance*, 2021, 74: 311–333.
- [49] Zhang Y, Lei L, Wei Y. Forecasting the Chinese stock market volatility with international market volatilities: The role of regime switching. *The North American Journal of Economics and Finance*, 2020, 52: 101145.
- [50] Clements A, Preve D P A. A practical guide to harnessing the HAR volatility model. *Journal of Banking & Finance*, 2021, 133: 106285.
- [51] Plakandaras V, Gupta R, Balcilar M, et al. Evolving United States stock market volatility: The role

- of conventional and unconventional monetary policies. *The North American Journal of Economics and Finance*, 2022, 60: 101666.
- [52] Li Z C, Xie C, Zeng Z J, et al. Forecasting global stock market volatilities in an uncertain world. *International Review of Financial Analysis*, 2023, 85: 102463.
- [53] Giannone D, Lenza M, Primiceri G E. Economic predictions with big data: The illusion of sparsity. *Econometrica*, 2021, 89(5): 2409–2437.
- [54] Lettau M, Pelger M. Factors that fit the time series and cross-section of stock returns. *The Review of Financial Studies*, 2020, 33(5): 2274–2325.
- [55] Pelger M. Understanding systematic risk: A high-frequency approach. *Journal of Finance*, 2020, 75(4): 2179–2220.
- [56] Lettau M, Pelger M. Estimating latent asset-pricing factors. *Journal of Econometrics*, 2020, 218(1): 1–31.
- [57] Bollerslev T, Todorov V, Li S Z. Jump tails, extreme dependencies, and the distribution of stock returns. *Journal of Econometrics*, 2013, 172(2): 307–324.
- [58] Chen W, Wei Y, Ma F, et al. A new method of stock market volatility forecasting in high-frequency perspective: HARFIMA model. *Journal of Management Sciences in China*, 2020, 23(11): 103–116.
- [59] Tang Y, Zhu P. Research of long memory, risk and efficiency of bull and bear based on CSI300 index futures: From the perspective of multifractality. *Management Review*, 2019, 31(8): 59–70.
- [60] Liu S, He X. Analysis of the relationship between the volatility of stock index futures and spot based on long memory model. *Mathematics in Practice and Theory*, 2018, 48(20): 58–64.