Physica A 516 (2019) 194–211

Contents lists available at ScienceDirect

Physica A

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

Modeling stock market volatility using new HAR-type models

Xu Gong, Boqiang Lin *

School of Management, China Institute for Studies in Energy Policy, Collaborative Innovation Center for Energy Econ. and Energy Policy, Xiamen University, Xiamen, 361005, China

HIGHLIGHTS

- We propose two new heterogeneous autoregressive models.
- Our models exhibit better performance than other heterogeneous autoregressive models.
- The low-frequency volatility contains much predictive information for future volatility.
- The leverage contains much in-sample prediction information for future volatility.

ARTICLE INFO

Article history:
Received 7 May 2018
Received in revised form 8 August 2018
Available online 19 October 2018

Keywords:
Volatility forecasting
Realized volatility
HAR-RV model
EEMD

ABSTRACT

Modeling volatility with reasonable accuracy is essential in asset allocation, asset pricing, and risk management. In this paper we use the ensemble empirical mode decomposition method and Zhang et al. (2008, 2009)'s method to decompose realized volatility into different volatility components. Then, we propose two new heterogeneous autoregressive (HAR) models by combining with the volatility components and leverage effect. Finally, we use high-frequency data for the S&P 500 as the study sample and perform parameter estimations on eight HAR-type models (including two new models). The results indicate that our models that are used to model 1-day, 1-week and 1-month future volatilities have an advantage over other existing HAR-type models. This advantage is substantial in the case of 1-month future volatility. In addition, the leverage contains significant in-sample prediction information for future volatility.

1. Introduction

Modeling and forecasting volatility is a fundamental aspect of option pricing [1,2], risk management [3–7], portfolio optimization [8,9], and hedge strategy [10–12], and it is closely related to the valuation of derivative instruments and the profits of traders.

To model and forecast volatility, Engle [13] proposed the ARCH model based on the "clustering" and "persistency" of volatility in financial markets. Bollerslev [14] further proposed the GARCH model based on the ARCH model. Taylor [15] subsequently proposed the SV model. Many studies thereafter extended the ARCH, GARCH and SV models, resulting in the GARCH-type and SV-type models [16–20]. The GARCH-type and SV-type models have exhibited well performance in financial markets, but they do not sufficiently describe whole-day volatility information, as they use low-frequency data to measure volatility. Thus, there are flaws in these models. In recent years, computers have greatly reduced the cost of recording and storing high-frequency data; as a result, this type of data has become an important tool in the study of financial market volatility. Andersen and Bollerslev [21] first used high-frequency data to propose a new method of measuring volatility,

* Corresponding author.
E-mail address: bqlin@xmu.edu.cn (B. Lin).

https://doi.org/10.1016/j.physa.2018.10.013

namely, realized volatility (RV). Compared with GARCH-type and SV-type models, RV has many advantages. For example, no model is necessary, and it is more accurate. It has therefore contributed greatly to advances in volatility research.

Since Andersen and Bollerslev [21], volatility models based on high-frequency data have developed rapidly. Blair et al. [22] added the exogenous variable (RV) to the GARCH model and built the GARCH-RV model. They found that the GARCH-RV model has better performance than that of the GARCH model. Then, Andersen et al. [23] developed the VAR-RV model and found this model to be superior to the GARCH model. Koopman et al. [24] proposed the GARCH-RV, SV-RV and ARFIMA-RV models by adding the RV to the GARCH, SV and ARFIMA models, respectively. Their results show that the new models are better able to forecast volatility than traditional volatility models, such as the GARCH and SV models. Kotkatvuori-Örnberg [25] found that the AR(FI)MA model for realized volatility reveal well performance when forecasting the volatility of S&P 500 index (SPX) daily returns.

Volatility models based on RV have led to substantial research achievements. On the basis of the heterogeneous market hypothesis of Müller et al. [26], Corsi [27] divided RV into short-term, medium-term and long-term volatility and proposed the heterogeneous autoregressive with RV (HAR-RV) model; the HAR-RV model proved to be markedly better than the GARCH and ARFIMA-RV models and has advanced research substantially. A great deal of research (such as Andersen et al. [28]; Ma et al. [29]) proved that the performance of the HAR-RV model is superior to the performance of the GARCH-type, SV-type, VAR-RV and ARFIMA-RV models. Furthermore, many studies, based on the HAR-RV model, have developed new volatility models to improve forecasting performance further. On the basis of the HAR-RV model, Andersen et al. [30] decomposed RV into continuous sample path variation and discontinuous jump variation and built the HAR-RV-J and HAR-RV-CJ models, thereby greatly improving model accuracy. Andersen et al. [28] discovered that the overnight return variance played a large role in the volatility of assets and developed the HAR-RV-CJN model. These authors found that the HAR-RV-CJN model demonstrated better forecasting performance than that of the GARCH and HAR-RV models. Huang et al. [31] compared the forecasting performance of the HAR-RV and HAR-RV-CJ models and concluded that the HAR-RV-CJ model performed better than the HAR-RV model. Çelik and Ergin [32] used the GARCH and HAR-RV models to forecast the volatility of the Turkish stock market and found that the volatility forecasting model that employed high-frequency data performed better than the traditional GARCH model. Additionally, Bekierman and Manner [33], Dong and Feng [34], Duong and Swanson [35], Gong and Lin [36], Peng et al. [37], Pu et al. [38], and Wang et al. [39] also showed that the HAR-type volatility models perform well in predicting future volatility of the stock market.

The above research has shown that the RV-type models based on high-frequency data performed better than traditional GARCH and SV models and that the HAR-type volatility models top the list. Although the existing HAR-type volatility models possess good forecasting power, higher accuracy is more favorable to the analysis of asset allocation, asset pricing, and risk management in financial markets. Therefore, it is necessary to improve the performance of the model further. We noticed the HAR-RV-CJ model, in which RV is divided into continuous sample path variation and discontinuous jump variation, is better able to predict volatility than the HAR-RV model.

However, RV is a complex, nonlinear time series. If it is simply decomposed into continuous sample path variation and discontinuous jump variation, it is difficult to understand its change rule, which prevents us from analyzing and modeling volatility. Therefore, our objective is to find a better way to decompose RV. Wu and Huang [40,41] proposed the ensemble EMD method (EEMD method) on the basis of the empirical mode decomposition method (EMD method; Huang et al., [42]). Compared with Fourier decomposition, wavelet decomposition and other data processing methods, the EEMD method has the following advantages. First, it is relatively easy to understand and implement. In addition, the fluctuations within a time series are automatically and adaptively selected from the time series. Finally, it is robust for nonlinear time series decomposition. The EEMD method can adaptively decompose a time series into several meaningful intrinsic mode function (IMF) components and one non-oscillatory trend component. The IMFs and trend component show the nature of the time series well. Therefore, the EEMD method is an effective decomposition tool.

At present, the EEMD method is widely used in physics, biomedical and earthquake prediction and other natural science fields, but its application in finance is less widespread. Zhang et al. [43,44] used the EEMD method to decompose oil prices into several intrinsic mode functions (IMFs) and one residual and synthesized these IMFs into two parts. They found that the price sequence after the decomposition-synthetic is logical in an economic sense. Then, Feng et al. [45] used the EEMD method and Zhang et al. [43,44]'s synthetic method to analyze the price of carbon. Combining the EEMD method, the LSSVR model and the ADD method, Li et al. [46] proposed a new hybrid forecasting model. These authors used this model to forecast the country risk of petroleum-exporting countries and found that it exhibits better predictive power than the ARIMA, GM (1, 1), FNN and LSSVR models. Xu et al. [47] found that the EEMD method performs better on the orthogonality of IMFs than the EMD and that the EEMD is more stable for the series of nine stock markets (i.e., CAC, DAX, FTSE, DJI, NQCI, SP, HSI, SC and SZ). The above studies show that the EEMD method for processing data is effective in the field of finance. However, the financial asset price is typically analyzed in the literature, while financial asset volatility is overlooked. Thus, this paper attempts to explore whether this method can effectively decompose financial asset volatility and thus increase the power to forecast future volatility.

We use the EEMD method to decompose RV into several IMFs and one trend volatility (T) and then use Zhang et al. [43,44]'s synthetic method to synthesize the high-frequency volatility (HF) and low-frequency (LF) volatility from the IMFs. In addition, we propose two new models to improve volatility forecasting. The rest of the paper is organized as follows. In the next section, we introduce RV, which includes the measurement and decomposition of RV. Section 3 proposes four volatility models without the leverage effect (HAR-RV, HAR-RV-J, HAR-RV-CJ and HAR-RV-HLT) and four volatility models with the leverage effect (LHAR-RV, LHAR-RV-J, LHAR-RV-CJ, and LHAR-RV-HLT). In Section 4, we use high-frequency data for the S&P 500 as the study sample and perform parameter estimation on the volatility models. Section 5 concludes.

2. Realized volatility

2.1. Data

The choice of sampling frequency of intraday high-frequency data greatly influences on the realized volatility measure. On the one hand, low sampling frequency cannot well contain the volatility information of that day. On the other hand, high sampling frequency may lead to microstructure noise. Liu et al. [48] study the accuracy of over 400 different estimators of asset return volatility constructed from high-frequency data. They find that other estimators are difficult to significantly beat five-minute realized volatility. Therefore, following [49–54], we take both the influences into consideration, and choose 5-minute high-frequency data from January 2, 1996 to June 5, 2013 (including 340964 data) for the S&P 500 index.

2.2. Measurement of realized volatility

As proposed by Andersen and Bollerslev [21], realized volatility RV_t^0 can be written as

$$RV_t^0 = \sum_{\mathrm{i}=1}^{M} r_{t,i}^2 \tag{1}$$

where $r_{t,i}$ is the return of the ith (i = 1, . . ., M) on trading day t, namely, $r_{t,i} = 100(\ln P_{\mathrm{t,i}} - \ln P_{t,i-1})$. $P_{\mathrm{t,i}}$ is the ith closing price of the trading day t.

In addition, Andersen and Bollerslev [21] studied realized volatility in the exchange market. Trades are not made on a 24 h basis in the stock market as they are in the exchange market, so the realized volatility calculated by Eq. (1) can only reflect market volatility for trading periods and not for information regarding market volatility in periods in which no trades are made (namely, the market volatility caused by overnight information—the overnight return variance from the closing of the previous day to the opening of that day). In addition, Hansen and Lunde [55] found that only when the overnight return variance and realized volatility were combined could they achieve consistent estimation of integrated volatility. Andersen et al. [28] showed that the overnight return variance in the SP and US stock market were 16.0% and 16.5% of the total return volatility, respectively. We also find that the overnight return variance is 11.3% of the entire volatility when data from January 2, 1996 to June 5, 2013 are used to calculate the volatility of the S&P 500 index. Therefore, according to Huang, et al. [31] and Gong et al. [56], considering the overnight return variance, we adjust the realized volatility as

$$RV_t = RV_t^0 + r_{t,n}^2 = \sum_{\mathrm{j}=1}^{N} r_{t,j}^2 \tag{2}$$

where $r_{t,1}$ and $r_{t,n}$ are the overnight return, $r_{t,1} = r_{t,n} = 100 * (\ln P_{t,o} - \ln P_{t-1,c})$. $P_{t,o}$ represents the opening price of phase t, and $P_{t-1,c}$ denotes the closing price of phase $t-1$; $r_{t,2}$ is the 1st return at phase t, $r_{t,2} = 100(\ln P_{t,1} - \ln P_{t,o})$, $P_{t,1}$ is the 1st closing price at phase t; $r_{t,3}$ denotes the 2nd return at phase t, $r_{t,3} = 100(\ln P_{t,2} - \ln P_{t,1})$; $\cdots$; $r_{t,M}$ is the $(N-1)$th return at phase t; and $r_{t,M} = 100(\ln P_{t,N-1} - \ln P_{t,N-2})$.

2.3. Decomposition of realized volatility

2.3.1. Continuous and jump volatilities

(1) Decomposition method

We divide the realized volatility into the continuous sample path variation and discontinuous jump variation using the method of Barndorff-Nielsen and Shephard [57,58]. We suppose that the return on the financial asset is a continuous process. We use quadratic variation (QV) to measure the total variation of the return volatility and integrated variation (IV) to measure the continuous component of the total variation. Thus, the jump component is the difference between the quadratic variation and integrated variation. In fact, the observed data in financial markets are discrete, so many scholars use the realized volatility to denote quadratic variation and the realized bipower variation (RBV) to denote integrated variation. Then, Barndorff-Nielsen and Shephard [57,58] decomposed the realized volatility into the continuous sample path variation and discontinuous jump variation on the basis of quadratic variation theory.

We assume that the logarithmic price of the financial asset ($p_t = \ln(P_t)$) within the trading day follows a standard jump–diffusion process

$$\mathrm{d}p_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t, 0 \le t \le T \tag{3}$$

where μ_t is the drift term with a continuous and locally finite variation sample path. σ_t denotes a strictly positive stochastic volatility process. W_t is a standard Brownian motion. $\kappa_t dq_t$ denotes the pure jump part.

For the discrete prices process, the log return volatility at time t includes jump volatility and is not an unbiased estimator of integrated volatility. The log return from t−1 to t is the quadratic variation

$$QV_t = [\mathrm{r}, \mathrm{r}]_{\mathrm{t}} = \int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1<s\le t} \kappa_s^2 \tag{4}$$

where $\int_{t-1}^{t} \sigma_s^2 ds < \infty$ is an integrated variation and denotes the continuous component of the total variation. $\sum_{t-1<s\leq t} \kappa_s^2$ is a jump volatility and denotes the cumulative jump variation in [t − 1, t].

Andersen and Bollerslev [21] found that the quadratic variation that cannot be observed directly can be estimated by the discrete data. When $M \to \infty$, the realized volatility RV_t can be used as a consistent estimator of QV_t.

$$RV_t \overset{M\to\infty}{\longrightarrow} QV_t = \int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1<s\leq t} \kappa_s^2 \tag{5}$$

In addition, the integrated volatility IV_t can be estimated by the realized bipower variation RBV_t (Barndorff-Nielsen and Shephard [57,58]). When $M \to \infty$, the realized volatility RV_t can be used as a consistent estimator of QV_t. RBV_t can be used as a consistent estimator of the continuous sample path variation.

$$RBV_t = z_1^{-2} \frac{M}{M-2} \sum_{j-3}^{M} \left| r_{t,j-2} \right| \left| r_{t,j} \right| \tag{6}$$

where $z_1 = E(Z_t) = \sqrt{\pi/2}$, Z_t is a random variable that follows a standard normal distribution. M/(M − 2) denotes an adjustment for the sample size. According to [59–62], we apply Z-statistics to identify the discontinuous jump variation

$$Z_t = \frac{(RV_t - RBV_t)RV_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5)\frac{1}{M}\max(1, \frac{RTQ_t}{RBV_t^2})}} \to N(0, 1) \tag{7}$$

where $\mu_1 = \sqrt{2/\pi}$, RTQ_t is realized tri-power quarticity, $RTQ_t = M\mu_{4/3}^{-3}(\frac{M}{M-4}) \sum_{j=4}^{M} \left| r_{t,j-4} \right|^{4/3} \left| r_{t,j-2} \right|^{4/3} \left| r_{t,j} \right|^{4/3}$, and $(\mu_{4/3} = E(|Z_T|^{4/3}) = 2^{2/3}\Gamma(7/6)\Gamma(1/2)^{-1})$.

According to Andersen et al. [28,30], the discontinuous jump variation J_t can be measured by

$$J_t = I(Z_t > \phi_\alpha)(RV_t - RBV_t) \tag{8}$$

And the continuous sample path variation C_t can be defined by

$$C_t = I(Z_t \leq \phi_\alpha)RV_t + I(Z_t > \phi_\alpha)RBV_t \tag{9}$$

where $I(\cdot)$ is an indicator function. According to Andersen et al. [28,30], α equals 0.99.

(2) Decomposition results

The real financial market is impacted by large amounts of information and by investor irrationality. In addition, the volatility of the market is discontinuous and includes a few jumps. Thus, using the above method, we decompose realized volatility RV_t^d into the continuous sample path variation C_t and discontinuous jump variation J_t^d.

Fig. 1 shows the continuous sample path variation C_t, discontinuous jump variation J_t^d and realized volatility RV_t^d for the S&P 500 and distinctly indicates the dynamic dependencies in the different components. The variation of the continuous sample path variation time series is similar to that of the realized volatility time series, and they are more predictable than the discontinuous jump variation time series. Observing the discontinuous jump variation time series in Fig. 1, we find the discontinuous jump variation mainly concentrated in the period from 1999 to 2003 and the period from 2008 to 2013. The period from 1999 to 2003 has a number of jumps because it was during this time that the U.S. stock market felt the effects of the bursting of the dotcom bubble as well as the terrorist attacks on September 11th. For the period from 2008 to 2013, the jumps are due to the global financial crisis of 2008 and the European debt crisis of 2012.

2.3.2. The high-frequency, low-frequency and trend volatilities

(1) EEMD method

The ensemble empirical mode decomposition (EEMD) is an intuitive and self-adaptive data processing method for nonlinear and non-stationary data. It was proposed by Wu and Huang [28,29] based on the empirical mode decomposition (EMD) method. The EMD method proposed by Huang et al. [42] can extract meaningful IMFs and one non-oscillatory trend from a complex series. Following Huang et al. [42], the data series $x(t)$ (t = 1, 2, …, n) can be decomposed by the EMD method according to the following procedure.

(a) To find out all local maximum and minimum values of the original series $x(t)$, use the cubic spline function of the upper envelope $e_{\max}(t)$ and lower envelope $e_{\min}(t)$ of the series. The mean of the upper envelope and lower envelope denotes the average envelope of the original series $m(t)$, where $m(t) = (e_{\max}(t) + e_{\min}(t))/2$.

(b) The original time series $x(t)$ minus $m(t)$ obtains a new sequence $d(t)$, which removes a low-frequency component, namely, $d_1^1(t) = x(t) - m(t)$.

(c) Repeating the above process enables $d_1^n(t)$ to meet the definition of the IMF and obtains the first-order IMF component of the original series $x(t)$, namely, $C_1(t) = IMF_1(t) = d_1^n(t)$.

(d) $x(t)$ minus $c_1(t)$ obtains a new sequence $r_1(t)$, which removes a high-frequency component, namely, $r_1(t) = x(t) - c_1(t)$.

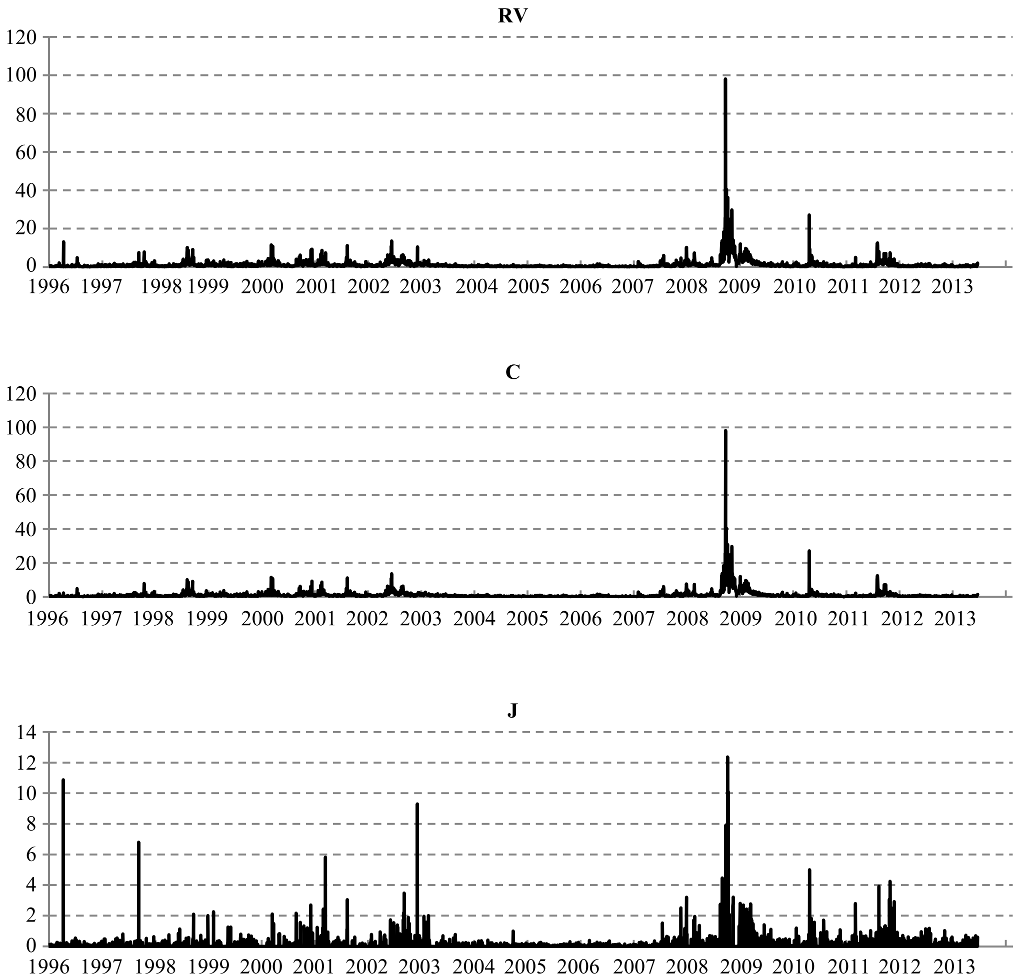

Fig. 1. Volatility components for S&P 500.

(e) $r_1(t)$ is used as a primitive sequence. Repeating the above process from Step (a) to Step (d) obtains the residual $r_n(t)$. When $r_n(t)$ becomes an oscillation function, monotonic function or constant, the processing procedure of the EMD method ends.

The original series $x(t)$ is decomposed into n IMFs and one non oscillatory trend by the process involving Step (a) to Step (e), namely, $x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t)$.

The EMD method for time series decomposition may bring about the mode mixing problem (namely, an intrinsic mode function contains different frequency modes), which reduces the physical meaning of the IMFs. To solve this problem, Wu and Huang [40,41] proposed the EEMD method. The processing procedure of the EEMD method is as follows.

(a) Composite the original series $x(t)$ and a white noise with finite amplitude $a(t)$ to get a new series $x'(t)$, namely, $x'(t) = x(t) + a(t)$.

(b) Use the processing procedure of the EMD method to decompose the new series $x'(t)$.

(c) Choose different white noise, and repeat Step (a) and Step (b).

(d) Calculate the mean of all IMFs to obtain the final real IMF_i and, at the same time, the oscillation $r_n(t)$.

(2) Decomposition results

The realized volatility of the stock market is a complex, nonlinear time series. To analyze the characteristic of realized volatility, we use the EEMD method to decompose the daily realized volatility RV_t^d, weekly realized volatility RV_t^w and monthly realized volatility RV_t^m, where $RV_t^w = (RV_t^d + RV_{t-1}^d + \cdots + RV_{t-4}^d)/5$ and $RV_t^m = (RV_t^d + RV_{t-1}^d + \cdots + RV_{t-21}^d)/22$.

The daily realized volatility, RV_t^d, weekly realized volatility RV_t^w and monthly realized volatility RV_t^m of the S&P 500 index can be decomposed into 11 different IMFs IMF_i and a residual $r_n(t)$. The decomposition results of RV_t^d are shown in Fig. 2a; the decomposition results of RV_t^w are shown in Fig. 2b; and the decomposition results of RV_t^m are shown in Fig. 2a.

In Figs. 2a–2c, RVD, RVW and RVM denote RV_t^d, RV_t^w and RV_t^m of the S&P 500 index, respectively. IMF 1 $\sim$ IMF 11 represent different frequency IMFs, IMF_i of RV_t^d, RV_t^w and RV_t^m. RESIDUAL denotes the trend of RV_t^d, RV_t^w and RV_t^m. These figures show

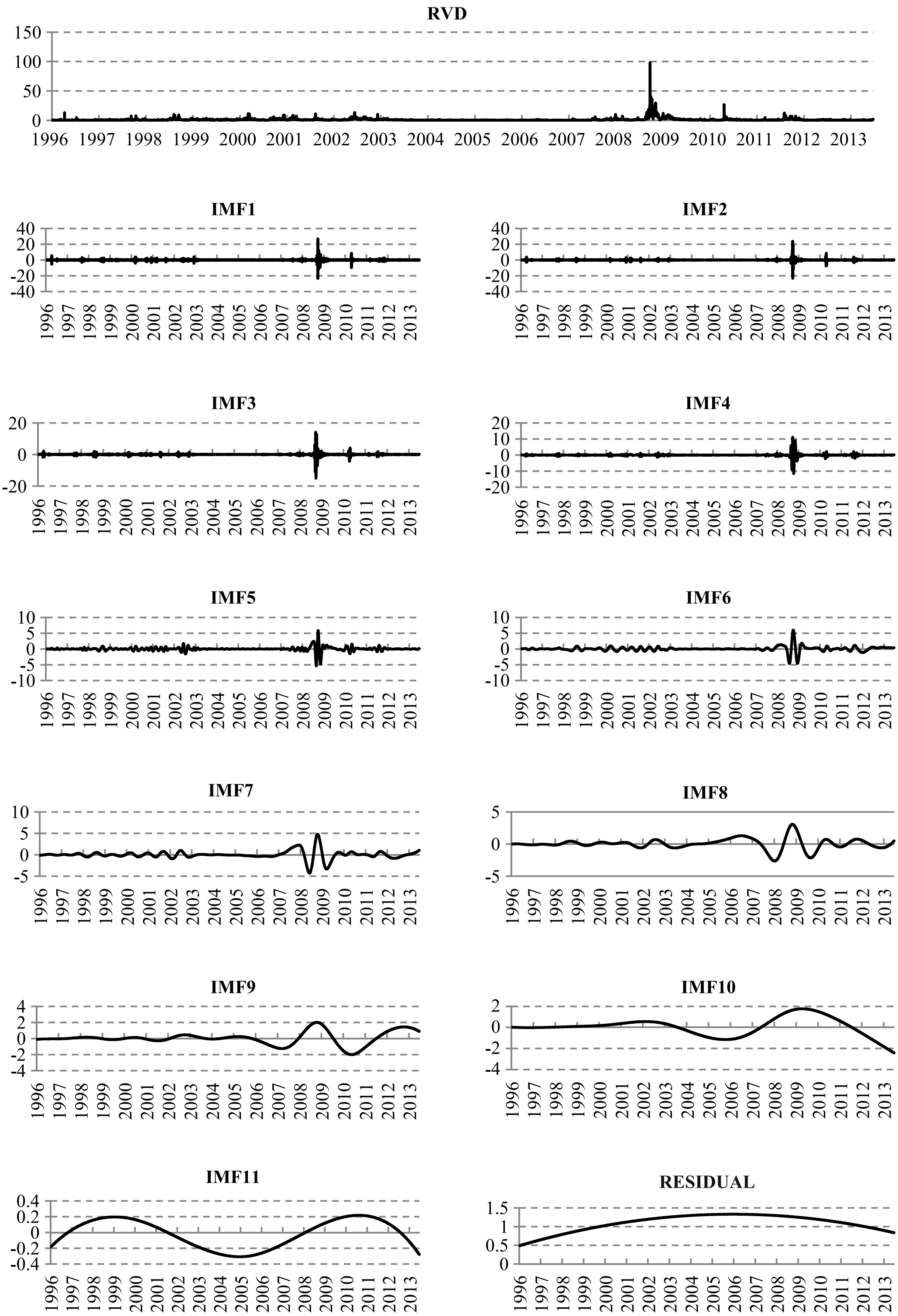

Fig. 2a. The IMFs and residual for the daily realized volatility through EEMD.

that RV_t^d, RV_t^w and RV_t^m are three complex, nonlinear time series. If we do not decompose these time series, it is not possible to find their inner nature and regularity. We use the EEMD method to decompose the realized volatility into different IMF_i and $r_n(t)$, and find that the fluctuation frequency of IMF_i in the front is very high and the fluctuation frequency of IMF_i at the

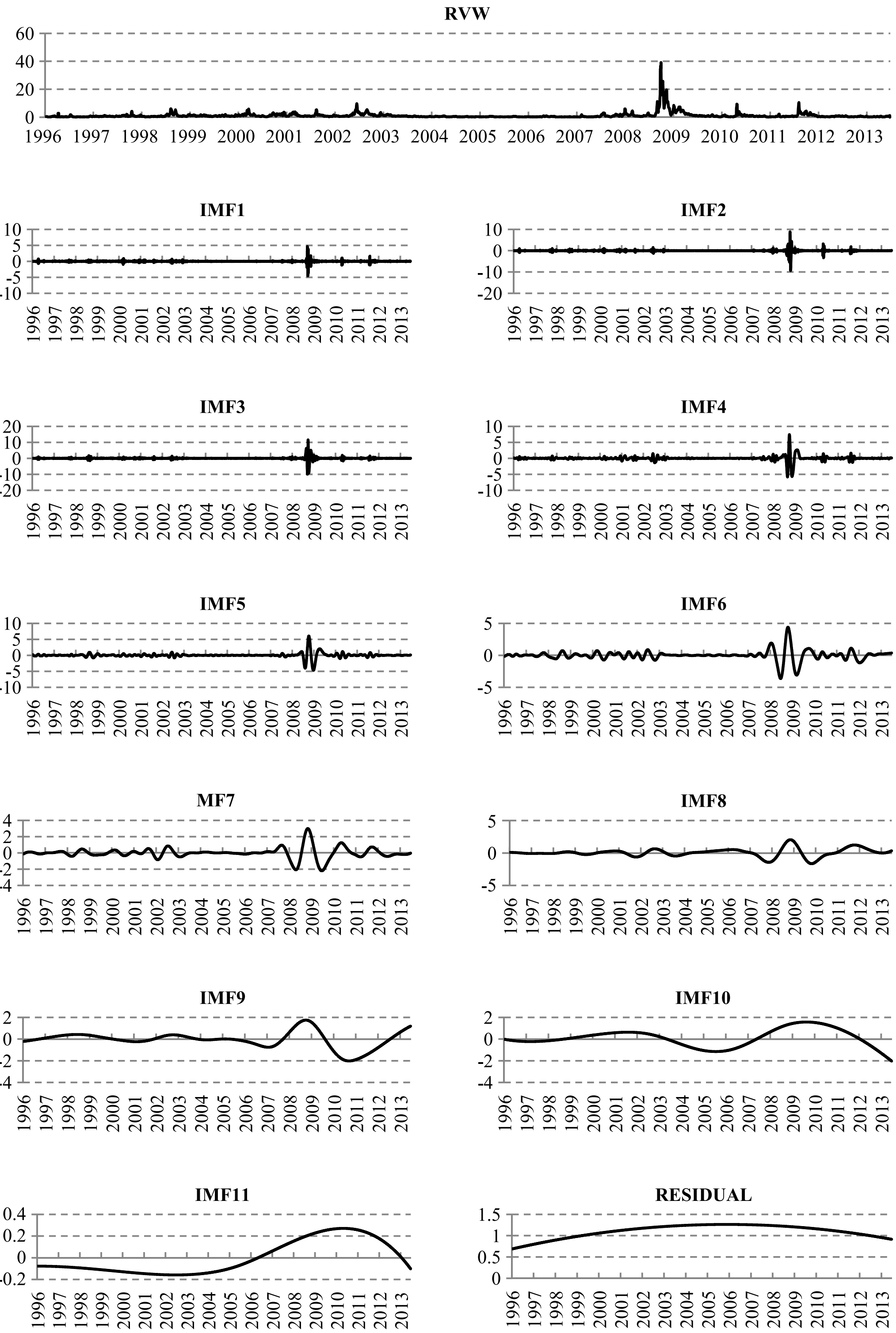

Fig. 2b. The IMFs and residual for the weekly realized volatility through EEMD.

back is low; we also find that $r_n(t)$ has an obvious trend. This approach affords a great deal of convenience in analyzing and forecasting realized volatility in the stock market.

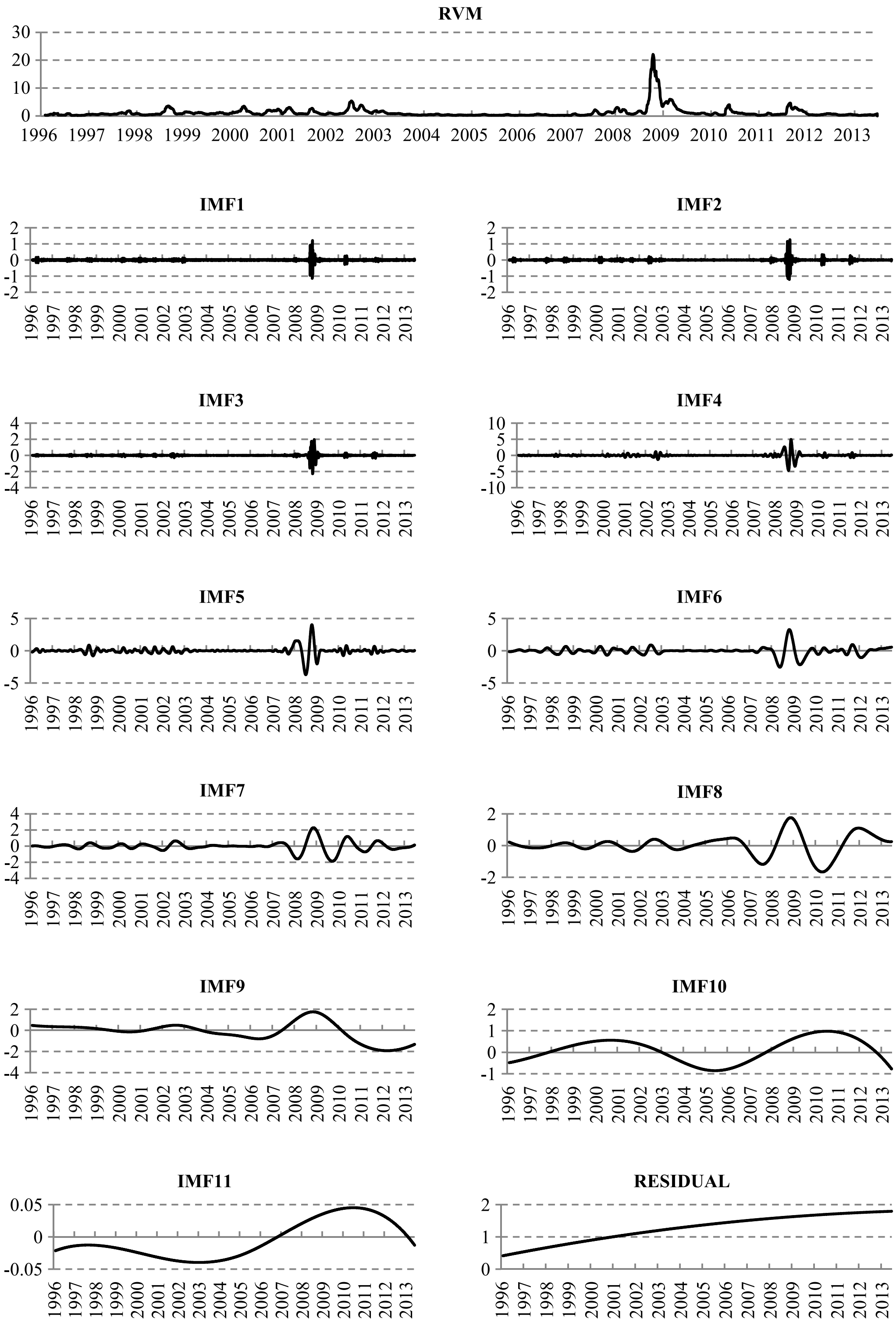

Fig. 2c. The IMFs and residual for the monthly realized volatility through EEMD.

(3) Synthesis of IMFs

To better analyze and forecast volatility, we further address each component of realized volatility. Table 1a lists the descriptive statistics of IMFs and RESIDUAL, which are obtained by using the EEMD method to decompose the daily

Table 1a
Descriptive statistics for the daily IMFs and residual.

	Mean	Correlation coefficient	Variance	Percentage of variance	Cumulative percentage of variance
IMF1	−0.0211	0.3789***	1.2099	15.63%	15.63%
IMF2	−0.0060	0.3275***	0.9613	12.42%	28.05%
IMF3	−0.0096	0.2542***	0.7058	9.120%	37.17%
IMF4	−0.0041	0.2572***	0.8053	10.40%	47.57%
IMF5	0.0148	0.1843***	0.4181	5.400%	52.97%
IMF6	0.0298**	0.3689***	0.6120	7.910%	60.88%
IMF7	−0.0107	0.3377***	0.8824	11.40%	72.28%
IMF8	0.0359***	0.3105***	0.6903	8.920%	81.20%
IMF9	0.0334***	0.2541***	0.6267	8.100%	89.29%
IMF10	0.0578***	0.2899***	0.7496	9.680%	98.98%
IMF11	−0.0110***	0.1521***	0.0305	0.390%	99.37%
RESIDUAL	1.0983***	0.0859***	0.0488	0.630%	100.0%

Notes: Asterisks in Column 2 indicate statistical significance at the 1% (***) or 5% (**) level where we refuse the null hypothesis that the value is equal to zero. "Correlation coefficient" indicates the correlation coefficient between the variable and daily realized volatility. Asterisks in Column 3 indicate statistical significance at the 1% level (***) where the variable and daily realized volatility have correlation properties.

Table 1b
Descriptive statistics for the weekly IMFs and residual.

	Mean	Correlation coefficient	Variance	Percentage of variance	Cumulative percentage of variance
IMF1	0.0026	0.1000***	0.0504	1.070%	1.070%
IMF2	0.0072	0.1999***	0.2890	6.120%	7.180%
IMF3	0.0027	0.2149***	0.6118	12.95%	20.13%
IMF4	0.0009	0.2345***	0.5669	12.00%	32.13%
IMF5	0.0118	0.4292***	0.5385	11.40%	43.53%
IMF6	0.0023	0.4211***	0.6130	12.97%	56.50%
IMF7	0.0078	0.4705***	0.4282	9.060%	65.56%
IMF8	0.0671***	0.3732***	0.3813	8.070%	73.63%
IMF9	0.0626***	0.3000***	0.5604	11.86%	85.49%
IMF10	0.1019***	0.3337***	0.6400	13.54%	99.04%
IMF11	0.0020	0.1795***	0.0228	0.48%	99.52%
RESIDUAL	1.1044***	0.1031***	0.0227	0.48%	100.00%

realized volatility RV_t^d; Table 1b lists the descriptive statistics of IMFs and RESIDUAL from RV_t^w; and Table 1c lists the descriptive statistics of IMFs and RESIDUAL from RV_t^m. In these tables, we find that several high-frequency IMFs in the front do not negate the null hypothesis that the mean is zero, but the remaining low-frequency IMFs negate the null hypothesis. Thus, the high-frequency IMFs for the most part fluctuate near zero, and the low-frequency IMFs deviate from zero. Analyzing the correlation coefficients between IMFs (or RESIDUAL) and realized volatility shows that all IMFs (or RESIDUAL) except IMF_1 in RV_t^w are related to their corresponding realized volatility, indicating that it is reasonable to use the EEMD method to decompose the realized volatility.

According to Zhang et al. [43,44], we obtain the whole volatility by compositing the IMFs whose means are equal to zero and define this volatility as high-frequency (HF) volatility. We get the low-frequency (LF) volatility by compositing the IMFs whose means do not equal zero. In addition, we define $r_n(t)$ as trend (T) volatility. In Table 1a, we add $IMF_1 \sim IMF_5$ of RV_t^d to get the daily high-frequency volatility HF_t^d, synthesize $IMF_6 \sim IMF_{11}$ to get the daily low-frequency volatility LF_t^d, and define $r_n(t)$ as the daily trend volatility T_t^d. Similarly, in Table 1b, we add $IMF_1 \sim IMF_7$ of RV_t^w to get the weekly high-frequency volatility HF_t^w, synthesize $IMF_8 \sim IMF_{11}$ to get the weekly low-frequency volatility LF_t^w, and define $r_n(t)$ as the weekly trend volatility T_t^w. In Table 1c, we add $IMF_1 \sim IMF_5$ of RV_t^m to get the monthly high-frequency volatility HF_t^m, synthesize $IMF_6 \sim IMF_{11}$ to get the monthly low-frequency volatility LF_t^m, and define $r_n(t)$ as the monthly trend volatility T_t^m.

To intuitively show characteristics of the high-frequency, low-frequency and trend volatilities, we list their changing process in Figs. 3–5. Fig. 3 shows the daily high-frequency volatility HF_t^d, daily low-frequency volatility LF_t^d and daily trend volatility T_t^d. Fig. 4 shows the weekly high-frequency volatility HF_t^w, weekly low-frequency volatility LF_t^w and weekly trend volatility T_t^w. Fig. 5 shows the monthly high-frequency volatility HF_t^m, monthly low-frequency volatility LF_t^m and monthly trend volatility T_t^m. In three figures, the left vertical axis indicates the values of LF_t^d, LF_t^w and LF_t^m, and the right vertical axis indicates the values of HF_t^d, HF_t^w, HF_t^m, T_t^d, T_t^w and T_t^m. Based on these figures, we find that HF_t^d, HF_t^w and HF_t^m are high-frequency and complex series, and it is difficult to identify their changing regularity. LF_t^d, LF_t^w and LF_t^m are low-frequency and simple series. T_t^d, T_t^w and T_t^m show an obvious trend and have strong predictability.

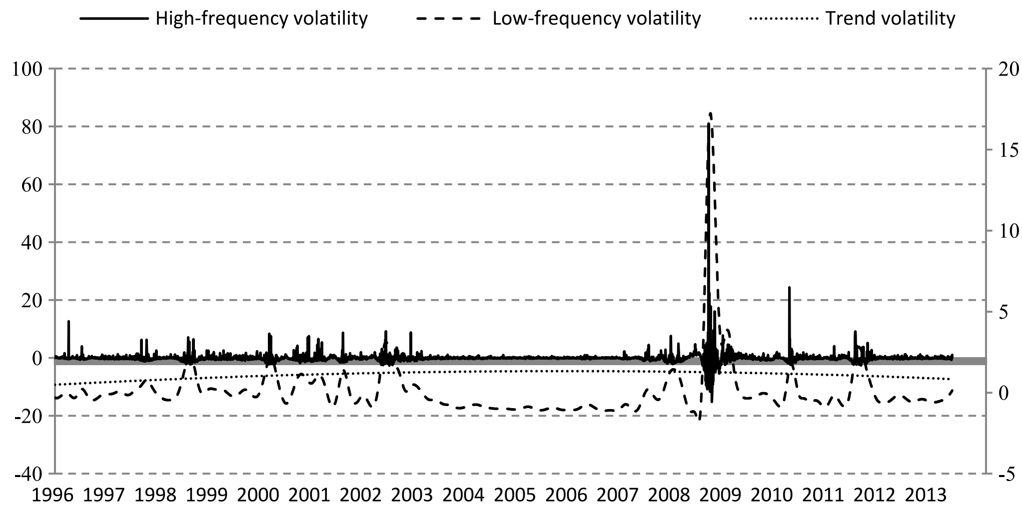

Fig. 3. The daily high-frequency, low-frequency and trend volatilities.

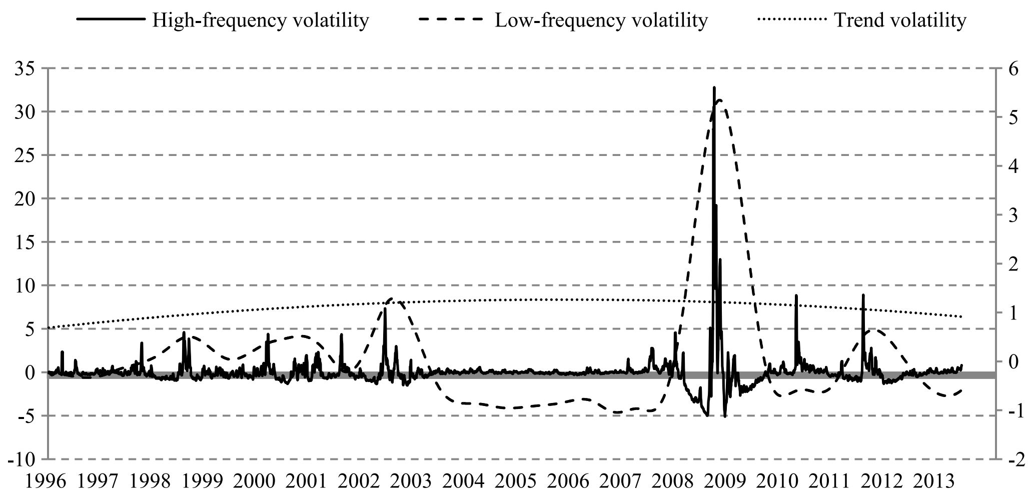

Fig. 4. The weekly high-frequency, low-frequency and trend volatilities.

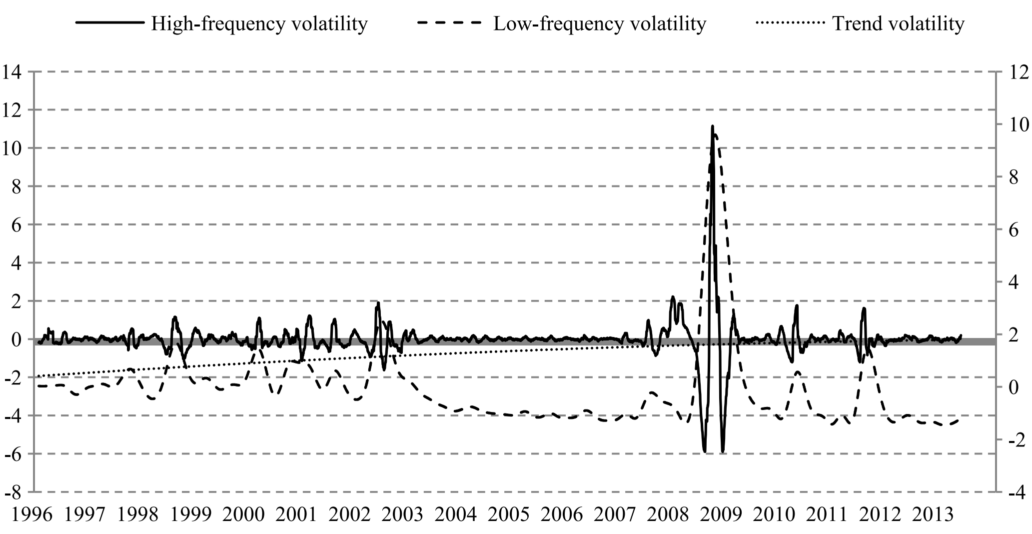

Fig. 5. The monthly high-frequency, low-frequency and trend volatilities.

Table 1c
Descriptive statistics for the monthly IMFs and residual.

	Mean	Correlation coefficient	Variance	Percentage of variance	Cumulative percentage of variance
IMF1	0.0000	0.0162	0.0029	0.090%	0.090%
IMF2	0.0001	0.0763***	0.0071	0.230%	0.320%
IMF3	0.0002	0.1473***	0.0321	1.040%	1.360%
IMF4	0.0087	0.2904***	0.3455	11.16%	12.52%
IMF5	0.0035	0.5149***	0.3525	11.39%	23.91%
IMF6	0.0353***	0.5147***	0.4027	13.01%	36.92%
IMF7	0.0232***	0.5328***	0.3376	10.91%	47.83%
IMF8	0.0375***	0.3749***	0.3989	12.89%	60.72%
IMF9	0.1213***	0.3764***	0.7587	24.51%	85.23%
IMF10	0.1043***	0.2933***	0.2927	9.460%	94.69%
IMF11	0.0048***	0.1934***	0.0008	0.030%	94.71%
RESIDUAL	1.2650***	0.1038***	0.1636	5.290%	100.0%

3. Econometric model

3.1. The HAR-RV model

According to the heterogeneous market hypothesis proposed by Müller et al. [26], Corsi [27] argued different traders in a heterogeneous market may satisfy different prices, so the market experiences trade and volatility. Corsi classified volatility as short-term, medium-term and long-term. Short-term volatility is mainly created by short-term investors who trade financial assets based on daily data. Medium-term volatility is mainly created by medium-term investors who trade financial assets based on weekly data. Long-term volatility is mainly created by long-term investors who trade financial assets based on monthly data. Based on this premise, Corsi proposed a new volatility forecasting model (HAR-RV model). It was defined as

$$RV_{t+H}^{d} = c + \alpha_1 RV_t^d + \alpha_2 RV_t^w + \alpha_3 RV_t^m + \varepsilon_{t+H} \tag{10}$$

where $RV_{t+H}^d = (RV_{t+1}^d + RV_{t+2}^d + \cdots + RV_{t+H}^d)/H$ denotes the realized volatility of the future H days. RV_t^d is the daily realized volatility. $RV_t^w = (RV_t^d + RV_{t-1}^d + \cdots + RV_{t-4}^d)/5$ represents the weekly realized volatility. $RV_t^m = (RV_t^d + RV_{t-1}^d + \cdots + RV_{t-21}^d)/22$ is the monthly realized volatility. This model mainly shows that volatility in the financial market is a complex mix that includes different volatility components, which is created by short-term, medium-term and long-term investors' trading behavior.

3.2. The HAR-RV-J model

To improve the prediction ability of the volatility model, Andersen et al. [30] added discontinuous jump variation to the HAR-RV model and proposed the HAR-RV-J model.

$$RV_{t+H}^{d} = c + \alpha_1 RV_t^d + \alpha_2 RV_t^w + \alpha_3 RV_t^m + \beta_1 J_t^d + \varepsilon_{t+H} \tag{11}$$

where J_t^d denotes the discontinuous jump variation. It is obtained by the decomposition method in Section 2.3.1.

3.3. The HAR-RV-CJ model

Andersen et al. [30] decomposed realized volatility into continuous sample path variation and discontinuous jump variation and built the HAR-RV-CJ model based on the HAR-RV model. The HAR-RV-CJ model can be written as

$$RV_{t+H}^{d} = c + \alpha_1 C_t^d + \alpha_2 C_t^w + \alpha_3 C_t^m + \beta_1 J_t^d + \beta_2 J_t^w + \beta_3 J_t^m + \varepsilon_{t+H} \tag{12}$$

where C_t^d is the daily continuous sample path variation. $C_t^w = (C_t^d + C_{t-1}^d + \cdots + C_{t-4}^d)/5$ denotes the weekly continuous sample path variation. $C_t^m = (C_t^d + C_{t-1}^d + \cdots + C_{t-21}^d)/22$ indicates the monthly continuous sample path variation. $J_t^w = (J_t^d + J_{t-1}^d + \cdots + J_{t-4}^d)/5$ denotes the weekly discontinuous jump variation. $J_t^m = (J_t^d + J_{t-1}^d + \cdots + J_{t-21}^d)/22$ indicates the monthly discontinuous jump variation.

3.4. The HAR-RV-HLT model

With respect to the HAR-RV and HAR-RV-CJ models, we will decompose realized volatility into high-frequency, low-frequency and trend volatilities and propose a new volatility forecasting model (HAR-RV-HLT). The modeling process of the HAR-RV-HLT model can be written as follows.

According to Corsi [27], we assume short-term traders are influenced by short-term, medium-term and long-term volatilities, medium-term traders are influenced by medium-term and long-term volatilities, and long-term investors are influenced solely by long-term volatility. We define a latent partial volatility $\widetilde{\sigma}_t^{\cdot}$. $\widetilde{\sigma}_t^d$ is the short-term volatility component, $\widetilde{\sigma}_t^{\mathrm{w}}$ denotes the medium-term volatility component, and $\widetilde{\sigma}_t^{\mathrm{m}}$ indicates the long-term volatility component. $\widetilde{\sigma}_t^d$, $\widetilde{\sigma}_t^{\mathrm{w}}$ and $\widetilde{\sigma}_t^{\mathrm{m}}$ can be expressed, respectively, as

$$\widetilde{\sigma}_{\mathrm{t+1m}}^{\mathrm{m}} = \mathrm{c_m} + \phi_{\mathrm{m}}\mathrm{RV}_{\mathrm{t}}^{\mathrm{m}} + \widetilde{\varepsilon}_{\mathrm{t+1m}}^{\mathrm{m}} \tag{13a}$$

$$\widetilde{\sigma}_{\mathrm{t+1w}}^{\mathrm{w}} = \mathrm{c_w} + \phi_{\mathrm{w}}\mathrm{RV}_{\mathrm{t}}^{\mathrm{w}} + \gamma_{\mathrm{w}}\mathrm{E}(\widetilde{\sigma}_{\mathrm{t+1m}}^{\mathrm{m}}) + \widetilde{\varepsilon}_{\mathrm{t+1w}}^{\mathrm{w}} \tag{13b}$$

$$\widetilde{\sigma}_{\mathrm{t+1d}}^{d} = \mathrm{c_d} + \phi_{\mathrm{d}}\mathrm{RV}_{\mathrm{t}}^{\mathrm{d}} + \gamma_{\mathrm{d}}\mathrm{E}(\widetilde{\sigma}_{\mathrm{t+1w}}^{\mathrm{w}}) + \widetilde{\varepsilon}_{\mathrm{t+1d}}^{\mathrm{d}} \tag{13c}$$

We use the EEMD method to decompose realized volatility into several IMFs and a residual and use Zhang et al. [43,44]'s method to synthesize three volatility components. Decomposing and synthesizing the monthly realized volatility RV_t^m obtains the monthly high-frequency volatility HF_t^m, monthly low-frequency volatility LF_t^m and monthly trend volatility T_t^m. In the same way, we get the weekly high-frequency volatility HF_t^w, weekly low-frequency volatility LF_t^w and weekly trend volatility T_t^w from the weekly RV, RV_t^w, and the daily high-frequency volatility HF_t^d, daily low-frequency volatility LF_t^d and daily trend volatility T_t^d from the daily RV, RV_t^d. Then, we use HF_t^m, LF_t^m and T_t^m to replace RV_t^m in Model (13a) and use HF_t^w, LF_t^w and T_t^w to replace RV_t^w in Model (13b); we use HF_t^d, LF_t^d and T_t^d to replace RV_t^d in Model (13c).

$$\widetilde{\sigma}_{\mathrm{t+1m}}^{\mathrm{m}} = c_m + \phi_{m1}{}^*HF_t^m + \phi_{m2}{}^*LF_t^m + \phi_{m3}{}^*LF_t^m + \widetilde{\varepsilon}_{\mathrm{t+1m}}^{\mathrm{m}} \tag{14a}$$

$$\widetilde{\sigma}_{\mathrm{t+1w}}^{\mathrm{w}} = \mathrm{c_w} + \phi_{w1}{}^*HF_t^w + \phi_{w2}{}^*LF_t^w + \phi_{w3}{}^*T_t^w + \lambda_w E(\widetilde{\sigma}_{\mathrm{t+1m}}^{\mathrm{m}}) + \widetilde{\varepsilon}_{\mathrm{t+1w}}^{\mathrm{w}} \tag{14b}$$

$$\widetilde{\sigma}_{\mathrm{t+1d}}^{d} = c_d + \phi_{d1}{}^*HF_t^d + \phi_{d2}{}^*LF_t^d + \phi_{d3}{}^*T_t^d + \gamma_d E(\widetilde{\sigma}_{\mathrm{t+1w}}^{\mathrm{w}}) + \widetilde{\varepsilon}_{\mathrm{t+1d}}^{\mathrm{d}} \tag{14c}$$

where $\widetilde{\varepsilon}_{\mathrm{t+1m}}^{\mathrm{m}}$, $\widetilde{\varepsilon}_{\mathrm{t+1w}}^{\mathrm{w}}$ and $\widetilde{\varepsilon}_{\mathrm{t+1d}}^{\mathrm{d}}$ are contemporaneously and serially independent zero-mean nuisance variates. HF_t^m, LF_t^m and T_t^m have an effect on the trade behavior of short-term, medium-term and long-term investors and affect short-term, medium-term and long-term volatilities in the financial market. HF_t^w, LF_t^w and T_t^w have an effect on the trade behavior of short-term and medium-term investors and affect short-term and medium-term volatilities. HF_t^d, LF_t^d and T_t^d have an effect on the trade behavior of short-term investors and affect short-term volatility.

Integrating Model (14a), (14b) and (14c), $\widetilde{\sigma}_{\mathrm{t+1d}}^{d}$ can be written as

$$\begin{aligned}\widetilde{\sigma}_{\mathrm{t+1d}}^{d} &= c + \alpha_1{}^*HF_t^d + \alpha_2{}^*HF_t^w + \alpha_3{}^*HF_t^m + \beta_1{}^*LF_t^d + \beta_2{}^*LF_t^w + \beta_3{}^*LF_t^m \\ &\quad + \gamma_1{}^*T_t^d + \gamma_2{}^*T_t^w + \gamma_3{}^*T_t^m + \widetilde{\varepsilon}_{\mathrm{t+1d}}^{\mathrm{d}}\end{aligned} \tag{15}$$

where $\widetilde{\sigma}_{\mathrm{t+1d}}^{d}$ can be expressed as $\widetilde{\sigma}_{\mathrm{t+1d}}^{d} = RV_{\mathrm{t+1d}}^d + \varepsilon_{\mathrm{t+1d}}^d$, and thus

$$\begin{aligned}RV_{t+1d}^d &= c + \alpha_1{}^*HF_t^d + \alpha_2{}^*HF_t^w + \alpha_3{}^*HF_t^m + \beta_1{}^*LF_t^d + \beta_2{}^*LF_t^w + \beta_3{}^*LF_t^m \\ &\quad + \gamma_1{}^*T_t^d + \gamma_2{}^*T_t^w + \gamma_3{}^*T_t^m + \varepsilon_{t+1d}\end{aligned} \tag{16}$$

where $\varepsilon_{t+1d} = \widetilde{\varepsilon}_{\mathrm{t+1d}}^{\mathrm{d}} - \varepsilon_{\mathrm{t+1d}}^d$.

Extending the forecast period of Eq. (16), we get the heterogeneous autoregressive realized volatility model with high-frequency, low-frequency and trend volatilities (HAR-RV-HLT):

$$\begin{aligned}RV_{t+H}^d &= c + \alpha_1{}^*HF_t^d + \alpha_2{}^*HF_t^w + \alpha_3{}^*HF_t^m + \beta_1{}^*LF_t^d + \beta_2{}^*LF_t^w + \beta_3{}^*LF_t^m \\ &\quad + \gamma_1{}^*T_t^d + \gamma_2{}^*T_t^w + \gamma_3{}^*T_t^m + \varepsilon_{t+H}\end{aligned} \tag{17}$$

3.5. Leverage effect

To better measure and forecast volatility, we add the leverage effect to the HAR-RV, HAR-RV-J, HAR-RV-CJ and HAR-RV-HLT models and obtain the LHAR-RV, LHAR-RV-J, LHAR-RV-CJ and LHAR-RV-HLT models, respectively. They can be written as

$$\begin{aligned}RV_{t+H}^d &= c + \alpha_1 RV_t^d + \alpha_2 RV_t^w + \alpha_3 RV_t^m \\ &\quad + \lambda_1 r_t^{d-} + \lambda_2 r_t^{w-} + \lambda_3 r_t^{m-} + \varepsilon_{t+H}\end{aligned} \tag{18}$$

$$\begin{aligned}RV_{t+H}^d &= c + \alpha_1 RV_t^d + \alpha_2 RV_t^w + \alpha_3 RV_t^m + \beta_1 J_t^d \\ &\quad + \lambda_1 r_t^{d-} + \lambda_2 r_t^{w-} + \lambda_3 r_t^{m-} + \varepsilon_{t+H}\end{aligned} \tag{19}$$

$$\begin{aligned}RV_{t+H}^d &= c + \alpha_1 C_t^d + \alpha_2 C_t^w + \alpha_3 C_t^m + \beta_1 J_t^d + \beta_2 J_t^w + \beta_3 J_t^m \\ &\quad + \lambda_1 r_t^{d-} + \lambda_2 r_t^{w-} + \lambda_3 r_t^{m-} + \varepsilon_{t+H}\end{aligned} \tag{20}$$

$$\begin{aligned}RV_{t+H}^d &= c + \alpha_1{}^*HF_t^d + \alpha_2{}^*HF_t^w + \alpha_3{}^*HF_t^m + \beta_1{}^*LF_t^d + \beta_2{}^*LF_t^w + \beta_3{}^*LF_t^m \\ &\quad + \gamma_1{}^*T_t^d + \gamma_2{}^*T_t^w + \gamma_1{}^*T_t^m + \lambda_1 r_t^{d-} + \lambda_2 r_t^{w-} + \lambda_3 r_t^{m-} + \varepsilon_{t+H}\end{aligned} \tag{21}$$

Table 2
Summary statistics for all variables.

	Mean	Std.Dev.	Skewness	Kurtosis	Q(5)	Q(10)	Q(15)	t statistic
RV_t^d	1.2077	2.6623	15.549	448.34	7104.7***	11 883***	15 654***	−5.6892***
RV_t^w	1.2114	2.1955	8.3057	102.21	16 526***	27 077***	35 435***	−6.5443***
RV_t^m	1.2105	1.9169	6.3270	54.086	20 818***	39 471***	55 240***	−7.2477***
C_t^d	1.0936	2.5446	17.137	531.01	6836.5***	11 309***	14 741***	−8.6021***
C_t^w	1.0970	2.0772	8.5438	107.48	16 626***	27 016***	34 954***	−6.6678***
C_t^w	1.0963	1.8123	6.4497	55.765	20 659***	39 146***	54 676***	−7.5511***
J_t^d	0.1141	0.4979	12.776	236.40	118.20***	187.40***	237.78***	−17.041***
J_t^w	0.1144	0.2481	6.5297	71.706	7086.3***	7914.6***	8560.4***	−6.8287***
J_t^m	0.1142	0.1543	3.7995	25.730	17 167***	30 494***	39 698***	−6.2779***
HF_t^d	−0.0268	1.8693	20.382	846.23	584.95***	701.87***	829.63***	−16.001***
HF_t^w	−0.0004	1.7414	8.1271	119.52	13 996***	19 919***	23 452***	−9.9763***
HF_t^m	−0.0123	0.9805	2.1998	47.377	19 213***	32 097***	38 462***	−14.352***
LF_t^d	0.1375	1.8988	5.6551	43.727	21 367***	41 615***	59 872***	−3.8891***
LF_t^w	0.1057	1.2392	2.4592	9.7040	21 566***	43 079***	64 477***	−3.5223***
LF_t^m	−0.0428	1.6375	3.4642	18.059	21 437***	42 610***	63 278***	−3.0852**
T_t^d	1.1013	0.2174	−0.9302	2.9300	21 469***	42 736***	63 803***	−26.676***
T_t^w	1.1060	0.1489	−0.9072	2.8754	21 447***	42 691***	63 732***	−21.642***
T_t^m	1.2650	0.4044	−0.4686	1.9811	21 408***	42 688***	63 841***	−0.0324
r_t^{d-}	−0.4276	0.8167	−3.4997	21.798	294.10***	620.03***	852.84***	−10.828***
r_t^{w-}	−0.1811	0.3438	−3.5093	23.081	4493.3***	4856.4***	5215.1***	−18.263***
r_t^{m-}	−0.0784	0.1612	−3.4062	18.827	15 221***	23 895***	28 255***	−8.7222***

Note: Asterisks indicate statistical significance at the 1% (***), 5% (**) or 10% (*) level.

where r_t^{d-}, r_t^{w-} and r_t^{m-} indicate the daily, weekly, and monthly leverages, respectively, namely, $r_t^{d-} = r_t I\{r_t < 0\}$, $r_t^{w-} = (r_t + r_{t-1} + \cdots + r_{t-4}) I\{(r_t + r_{t-1} + \cdots + r_{t-4})/5 < 0\}$, $r_t^{m-} = (r_t + r_{t-1} + \cdots + r_{t-21}) I\{(r_t + r_{t-1} + \cdots + r_{t-21})/22 < 0\}$, where r_t is return at period t and $I\{\cdot\}$ is an indicator function.

4. Empirical analysis

4.1. Summary statistics

Table 2 lists summary statistics results of all variables. The table shows that the means of daily, weekly and monthly variables are approximately equal. The realized volatility, continuous sample path variation, discontinuous jump variation, high-frequency volatility and low-frequency volatility are "right skewed" and exhibit "excess kurtosis and a fat tail". However, the trend volatility is "left skewed" and exhibits "low kurtosis". The leverage is "left skewed" and exhibits "excess kurtosis and a fat tail". In addition, the Q (5), Q (10) and Q (15) of all variables are significant under the 1% significance level, which shows these variables have significant correlation. According to the unit root test (ADF test), we find that all variables except the monthly trend volatility T_t^m refuse the null hypothesis that it has a unit root.

Fig. 6a shows the lagged correlation function between future realized volatility RV_{t+H} and X_t as a function of H, with X_t being RV_t^d itself, continuous sample path variation and discontinuous jump variation. In this figure, RVD, RVW, RVM, RVCD, RVCW, RVCM, RVJD, RVJW and RVJM represent RV_t^d, RV_t^w, RV_t^m, C_t^d, C_t^w, C_t^m, J_t^d, J_t^w and J_t^m, respectively. The results of RV_t^d, RV_t^w and RV_t^m show that realized volatility has a long memory, which indicates realized volatility (RV_t^d, RV_t^w or RV_t^m) has a certain prediction power for future realized volatility, RV_{t+H}. In addition, the correlation between continuous sample path variation (C_t^d, C_t^w or C_t^m) and future realized volatility is high, but the correlation between the discontinuous jump variation (J_t^d, J_t^w or J_t^m) and future realized volatility is low. Thus, the continuous sample path variation exhibits much forecasting power for future realized volatility, but the discontinuous jump variation shows little forecasting power for future realized volatility.

Fig. 6b shows the lagged correlation function between future realized volatility RV_{t+H}, and X_t as a function of H, with X_t being high-frequency volatility, low-frequency volatility, trend volatility and leverage. In Fig. 6b, HFD, HFW, HFM, LFD, LFW, LFM, TD, TW, TM, RD, RW and RM denote HF_t^d, HF_t^w, HF_t^m, LF_t^d, LF_t^w, LF_t^m, T_t^d, T_t^w, T_t^m, $\left|r_t^{d-}\right|$, $\left|r_t^{w-}\right|$ and $\left|r_t^{m-}\right|$, respectively. The figure shows that high-frequency volatility (HF_t^d, HF_t^w or HF_t^m) and trend volatility (T_t^d, T_t^w or T_t^m) are minimally related to future realized volatility, but low-frequency volatility (LF_t^d, LF_t^w or LF_t^m) is closely related to future realized volatility. Thus, high-frequency volatility and trend volatility demonstrate little forecasting power for future realized volatility, but low-frequency volatility demonstrates a great deal of forecasting power for future realized volatility. In addition, the correlation

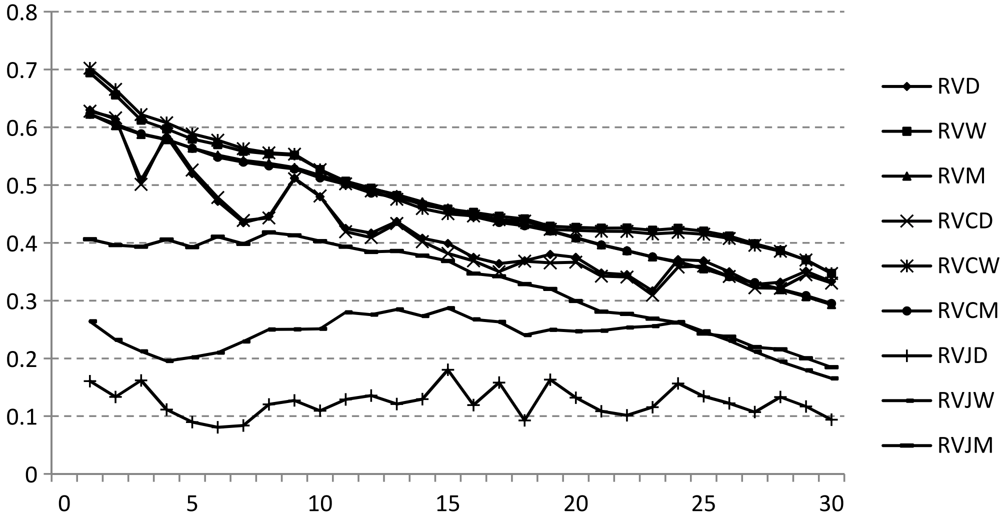

Fig. 6a. Lagged correlation function between future realized volatility and other variables.

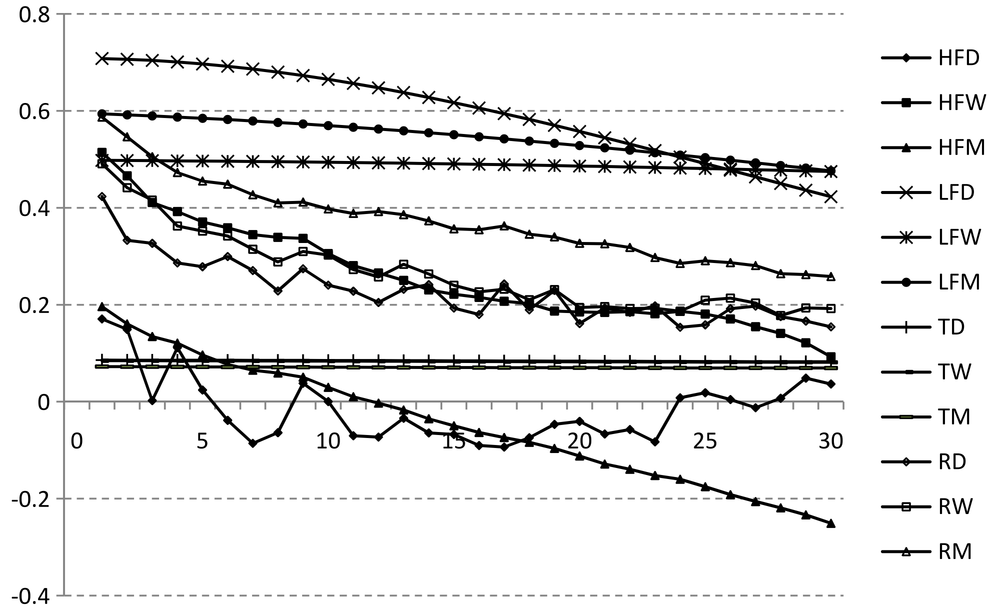

Fig. 6b. Lagged correlation function between future realized volatility and other variables.

between leverage ($\left|r_t^{d-}\right|$, $\left|r_t^{w-}\right|$ or $\left|r_t^{m-}\right|$) and future realized volatility is high, , which donates the volatility in American stock market exists leverage effect.

The above results prove the rationality of our proposed models, in addition to preliminarily judging the forecasting power of each variable for future realized volatility. The results show that realized volatility, continuous sample path variation, low-frequency volatility and leverage exhibit good forecasting power, but other variables offer poor prediction power.

4.2. Parameter estimation

In this paper, we compare forecasting power for 1-day, 1-week and 1-month future volatilities, so H in the HAR-RV, HAR-RV-J, HAR-RV-CJ, HAR-RV-HLT, LHAR-RV, LHAR-RV-J, LHAR-RV-CJ and LHAR-RV-HLT models is chosen as 1, 5 and 22. That is to say, RV_{t+1}^d, RV_{t+5}^d and RV_{t+22}^d are the 1-day, 1-week and 1-month future volatilities. Because realized volatility, continuous sample path variation, discontinuous jump variation, high-frequency volatility, low-frequency volatility, trend volatility and leverage in these models are overlap variables, we use the OLS with Newey–West to estimate the parameters.

Table 3
Estimation results for forecasting 1-day future volatility.

	Without leverage				With leverage			
	HAR-RV	HAR-RV-J	HAR-RV-CJ	HAR-RV-HLT	LHAR-RV	LHAR-RV-J	LHAR-RV-CJ	LHAR-RV-HLT
c	0.1230**	0.1343***	0.1502***	95.937	−0.1667	−0.1556	−0.1344	219.31**
	(2.4136)	(2.8035)	(3.6528)	(1.0762)	(−1.2468)	(−1.1603)	(−1.1277)	(2.1302)
α_1	0.2040	0.2208	0.1803	0.1544	0.0752	0.0917	0.0572	0.0580
	(1.5923)	(1.5085)	(1.2114)	(1.1964)	(0.5614)	(0.6088)	(0.3760)	(0.4481)
α_2	0.4954***	0.4979***	0.6070***	0.2949***	0.4368***	0.4407***	0.5388	0.2867**
	(3.1525)	(3.0134)	(3.1606)	(2.8019)	(2.9198)	(2.8311)	(3.0972)	(2.3961)
α_3	0.1999**	0.2000**	0.1590	−0.4136	0.1996**	0.1975**	0.1610**	−0.2743
	(2.2238)	(2.2028)	(1.5324)	(−1.5850)	(2.4742)	(2.4495)	(1.7570)	(−1.3100)
β_1		−0.3044**	0.1612*	1.0568***		−0.2612**	0.0524	0.7759***
		(−2.4384)	(1.7670)	(4.3004)		(−2.5122)	(0.6062)	(4.2838)
β_2			−0.8883	0.4395***			−0.7968**	0.3817***
			(−1.6448)	(3.0554)			(−1.9233)	(2.6400)
β_3			0.9381*	−0.4672			0.8186*	−0.3302
			(1.7380)	(−1.6400)			(1.7061)	(−1.3664)
γ_1				191.15				435.20**
				(1.0812)				(2.1321)
γ_2				−272.09				−622.61**
				(−1.0748)				(−2.1291)
γ_1				−3.5665				−7.3916**
				(−1.1557)				(−2.1105)
λ_1					−0.4470***	−0.4661***	−0.4566***	−0.3869***
					(−2.8312)	(−3.0111)	(−3.0379)	(−2.7238)
λ_2					−1.2251**	−1.1712**	−1.1490**	−0.9998*
					(−1.9504)	(−1.8029)	(−1.8253)	(−1.9454)
λ_3					−1.3327***	−1.3088***	−1.3399***	−1.1484***
					(−2.8766)	(−2.7819)	(−2.6581)	(−2.6191)
Adj.R^2	0.5006	0.5034	0.5117	0.5556	0.5684	0.5704	0.5768	0.6006

Notes: The parentheses in the table reports t-statistics.

Table 3 lists the results of all models which predict 1-day future volatility. In this section, we first analyze the results of four HAR-type models without leverage (HAR-RV, HAR-RV-J, HAR-RV-CJ and HAR-RV-HLT). According to the result of the HAR-RV model, we find that the coefficients of RV_t^w and RV_t^m are significantly positive and that the coefficient of RV_t^d is approximately significant, which shows that volatility in the American stock market has a long memory; in addition, historical volatility contains in-sample forecast information for 1-day future volatility. At the same time, this result indicates investors in the U.S. stock market are characterized by heterogeneity, in conformity with heterogeneous market hypothesis theory. Analyzing the result of the HAR-RV-J model, we find that J_t^d is significantly negative, which shows that discontinuous jump variation has in-sample forecasting power for 1-day future volatility. The result of the HAR-RV-CJ model shows that the coefficients of C_t^w, J_t^d and J_t^m are statistically significant and that other coefficients are not significant. Furthermore, only C_t^w, J_t^d and J_t^m in the HAR-RV-CJ model can forecast 1-day future volatility. The result of the HAR-RV-HLT model shows that low-frequency volatility exhibits much in-sample forecasting power for 1-day future volatility, but high-frequency volatility and trend volatility show little in-sample forecasting power. This result agrees with the result of the correlation analysis in Section 3.1. In addition, comparing the adjusted R-squared (Adj.R^2) of the HAR-RV, HAR-RV-J, HAR-RV-CJ and HAR-RV-HLT models, we find that the Adj.R^2 of the HAR-RV-HLT model is highest, and this model exhibits the strongest in-sample forecasting ability for 1-day future volatility.

In this section, we analyze the results of four HAR-type models with leverage (LHAR-RV, LHAR-RV-J, LHAR-RV-CJ and LHAR-RV-HLT). The coefficients of r_t^{d-}, r_t^{d-} and r_t^{m-} in these models are all significantly negative, which the volatility of American stock market exists leverage effect, and historical negative rates can increase future volatility. Comparing the Adj.R^2s of the LHAR-RV, LHAR-RV-J, LHAR-RV-CJ and LHAR-RV-HLT models, we find that the Adj.$R^2$of the LHAR-RV-HLT model is highest, and this model exhibits the strongest in-sample forecasting ability for 1-day future volatility. In addition, comparing comparison of the Adj.R^2s of HAR-type models with leverage and HAR-type models without leverage indicates the in-sample forecasting power of HAR-type models with leverage is higher than that of HAR-type models without leverage.

Tables 4 and 5 list the estimation results for 1-week and 1-month future volatility, respectively. Compared with the estimation results in Table 3, the coefficients in Tables 4 and 5 are more significant, and the Adj.R^2s are larger; thus, the variables in these models include more in-sample prediction information for 1-week and 1-month future volatilities. A comparison of the Adj.R^2s of HAR-type models with leverage and HAR-type models without leverage indicates that the forecasting power of HAR-type models with leverage for 1-week and 1-month future volatilities is higher than that of the HAR-type models without leverage. With respect to Table 4, we compare the Adj.R^2s of HAR-type models without leverage and HAR-type models with leverage and find that the Adj.$R^2$of the HAR-RV-HLT model is obviously higher than that of the HAR-RV, HAR-RV-J and HAR-RV-CJ models and that the Adj.R^2 of the LHAR-RV-HLT model is obviously higher than that of the LHAR-RV, LHAR-RV-J and LHAR-RV-CJ models. Thus, the HAR-RV-HLT and LHAR-RV-HLT models exhibit much stronger

Table 4
Estimation results for forecasting 1-week future volatility.

	Without leverage				With leverage			
	HAR-RV	HAR-RV-J	HAR-RV-CJ	HAR-RV-HLT	LHAR-RV	LHAR-RV-J	LHAR-RV-CJ	LHAR-RV-HLT
c	0.1837***	0.1971***	0.1956***	145.25	−0.0160	−0.0028	0.0021	214.28**
	(2.6953)	(3.0723)	(3.6043)	(1.4306)	(−0.1576)	(−0.0284)	(0.0863)	(2.2068)
α_1	0.1942***	0.2141***	0.1676***	0.1228***	0.0977**	0.1174**	0.0759***	0.0697**
	(3.2992)	(2.9507)	(2.6538)	(3.2513)	(2.1556)	(1.9656)	(5.7654)	(2.3495)
α_2	0.3384***	0.3415***	0.4675***	0.0467	0.2856***	0.2903***	0.4085***	0.0316
	(4.8637)	(4.5269)	(5.5894)	(0.8127)	(5.2040)	(4.8890)	(18.130)	(0.5924)
α_3	0.3166***	0.3167***	0.2510**	−0.5716**	0.3271***	0.3246***	0.2605***	−0.4925**
	(3.1834)	(3.1662)	(2.1315)	(−2.3376)	(3.9522)	(3.8883)	(11.762)	(−2.2038)
β_1		−0.3609***	0.1290	1.4288***		−0.3097***	0.0707	1.2795***
		(−3.0824)	(1.5250)	(4.5389)		(−3.3593)	(1.6188)	(4.4366)
β_2			−1.3137**	0.2694***			−1.2754***	0.2290***
			(−2.0227)	(3.4833)			(−12.107)	(3.2469)
β_3			1.5775**	−0.6556**			1.5225***	−0.5739**
			(2.1934)	(−2.4326)			(8.2843)	(−2.2857)
γ_1				289.23				425.74**
				(1.4363)				(2.2107)
γ_2				−412.33				−608.40**
				(−1.4299)				(−2.2060)
γ_1				−5.3560				−7.4872**
				(−1.5461)				(−2.2572)
λ_1					−0.2042***	−0.2268***	−0.2131***	−0.0985***
					(−4.4693)	(−4.7068)	(−7.7315)	(−2.7652)
λ_2					−1.0597**	−0.9958**	−0.9629***	−0.6618**
					(−2.2476)	(−2.0899)	(−12.565)	(−2.3946)
λ_3					−1.1381**	−1.1097**	−1.1746***	−0.7933**
					(−2.1401)	(−2.0941)	(−6.6614)	(−2.0580)
Adj.R^2	0.6204	0.6264	0.6450	0.7922	0.6710	0.6753	0.6917	0.8096

in-sample forecasting power for 1-week future volatility than the other models. In particular, when all models forecast 1-week future volatility, Adj.R^2 of the HAR-RV-HLT model is 72.8%, 72.3% and 70.6% higher than that of the HAR-RV, HAR-RV-J and HAR-RV-CJ models, respectively, and Adj.$R^2$of the LHAR-RV-HLT model is 65.1%, 64.9% and 63.4% higher than that of the LHAR-RV, LHAR-RV-J and LHAR-RV-CJ models, respectively, in Table 5. Therefore, the HAR-RV-HLT and LHAR-RV-HLT models exhibit rather stronger in-sample forecasting power for 1-month future volatility than other models.

5. Conclusion

We first analyze different volatility components, which are obtained by Barndorff-Nielsen & Shephard [57,58]'s method or the EEMD method and Zhang et al. [43,44]'s method. Then, we propose the HAR-RV-J, HAR-RV-CJ, HAR-RV-HLT, LHAR-RV, LHAR-RV-J, LHAR-RV-CJ and LHAR-RV-HLT models based on the HAR-RV model. In addition, we use high-frequency data for the S&P 500 to perform parameter estimation on all models.

Analysis of different volatility components shows enormous jumps in the period from 1999 to 2003 and the period from 2008 to 2013; thus, the volatility of the two periods is greater than that of other periods in the American stock market. In addition, continuous sample path variation and low-frequency volatility contain much predictive information for future volatility, but discontinuous jump variation, high-frequency volatility and trend volatility contain little predictive information.

According to the estimation results in Section 4, we know that the volatility in the American stock market has a long memory, and historical volatility contains forecast information for future volatility. Moreover, the HAR-RV-HLT and LHAR-RV-HLT models exhibit better performance for future volatility (especially 1-month future volatility) than other HAR-type models.

The above results show that the HAR-RV-HLT and LHAR-RV-HLT models are good models, and they are more favorable to the study of practical problems such as asset allocation, asset pricing, and risk management.

Acknowledgments

The paper is supported by the Social Science Planning Project of Fujian Province, China (No. FJ2017C075), National Natural Science Foundation of China (No. 71701176), Project funded by China Postdoctoral Science Foundation (No. 2018T110642, 2017M612121), Report Series from Ministry of Education of China (No. 10JBG013), and National Social Science Foundation of China (No. 17AZD013).

Table 5
Estimation results for forecasting 1-month future volatility.

	Without leverage				With leverage			
	HAR-RV	HAR-RV-J	HAR-RV-CJ	HAR-RV-HLT	LHAR-RV	LHAR-RV-J	LHAR-RV-CJ	LHAR-RV-HLT
c	0.3500***	0.3558***	0.3096***	211.22***	0.2275***	0.2330***	0.1887**	229.87***
	(6.1510)	(6.4083)	(4.6582)	(3.4793)	(3.2877)	(3.4525)	(2.2671)	(4.0190)
α_1	0.0970**	0.1056**	0.0945**	0.0015	0.0407	0.0489	0.0410	0.0023
	(2.5410)	(2.3618)	(2.0977)	(0.1781)	(1.3202)	(1.3471)	(1.0872)	(0.3045)
α_2	0.2971***	0.2984***	0.3492***	−0.0624**	0.2533***	0.2553***	0.3001***	−0.0867***
	(3.0663)	(3.0494)	(3.3016)	(−1.9932)	(2.5946)	(2.5951)	(2.9751)	(−2.6039)
α_3	0.3185***	0.3185***	0.2239	−0.9116***	0.3213***	0.3203***	0.2239*	−0.9110***
	(3.1273)	(3.1103)	(1.6397)	(−9.1739)	(3.3526)	(3.3289)	(1.7097)	(−8.7337)
β_1		−0.1566*	−0.0089	1.5831***		−0.1298*	−0.0481	1.5695***
		(−1.8423)	(−0.1473)	(16.815)		(−1.7885)	(−0.7586)	(15.685)
β_2			−0.2054	0.2952***			−0.1978	0.2637***
			(−0.4870)	(5.3018)			(−0.5046)	(4.8307)
β_3			1.7136*	−0.8386***			1.7064*	−0.8211***
			(1.7672)	(−7.9756)			(1.8351)	(−7.7260)
γ_1				420.11***				456.93***
				(3.4927)				(4.0317)
γ_2				−599.65***				−652.53***
				(−3.4788)				(−4.0180)
γ_1				−7.6997***				−8.2801***
				(−3.8532)				(−4.3871)
λ_1					−0.1552**	−0.1647***	−0.1620***	0.0062
					(−2.5246)	(−2.6566)	(−2.7072)	(0.4409)
λ_2					−0.5145**	−0.4878*	−0.4515*	0.1147
					(−2.0066)	(−1.9057)	(−1.8216)	(1.5399)
λ_3					−1.0342*	−1.0224*	−1.1098**	−0.7188***
					(−1.9482)	(−1.9187)	(−1.9732)	(−3.8267)
Adj.R^2	0.5498	0.5512	0.5569	0.9499	0.5761	0.5770	0.5823	0.9514

References

[1] Z. Huang, T. Wang, P.R. Hansen, Option pricing with the Realized GARCH model: an analytical approximation approach, J. Futures Mark. 37 (4) (2017) 328–358.
[2] H.Y.X.Pan.J. Song, Volatility cones and volatility arbitrage strategies—empirical study based on SSE ETF option, China Financ. Rev. Int. 7 (2017) 203–227.
[3] T. Zhang, Z. Liu, Fireworks algorithm for mean-VaR/CVaR models, Physica A 483 (2017) 1–8.
[4] X. Gong, F. Wen, X.H. Xia, et al., Investigating the risk-return trade-off for crude oil futures using high-frequency data, Appl. Energy 196 (2017) 152–161.
[5] J. Ma, X. Xiong, F. He, W. Zhang, Volatility measurement with directional change in Chinese stock market: statistical property and investment strategy, Physica A 471 (2017) 169–180.
[6] Z. Huang, C. Tong, T. Wang, VIX term structure and VIX futures pricing with realized volatility, J. Futures Mark. (2018) 1–22, http://dx.doi.org/10.1002/fut.21955.
[7] C. Chiarella, X.Z. He, L. Wei, Learning, information processing and order submission in limit order Markets, J. Econom. Dynam. Control 61 (2015) 245–268.
[8] Z. Dai, F. Wen, Some improved sparse and stable portfolio optimization problems, Financ. Res. Lett. (2018) http://dx.doi.org/10.1016/j.frl.2018.02.026.
[9] Z. Dai, F. Wen, A generalized approach to sparse and stable portfolio optimization problem, J. Ind. Manag. Optim. 14 (4) (2018) 1651–1666.
[10] Y. Wang, C. Wu, L. Yang, Hedging with futures: does anything beat the naïve hedging strategy?, Manag. Sci. 61 (2015) 2870–2889.
[11] F. Wen, J. Xiao, C. Huang, X. Xia, Interaction between oil and us dollar exchange rate: nonlinear causality, time-varying influence and structural breaks in volatility, Appl. Econ. 50 (3) (2018) 319–334.
[12] J. Xiao, M. Zhou, F. Wen, F. Wen, Asymmetric impacts of oil price uncertainty on Chinese stock returns under different market conditions: Evidence from oil volatility index, Energy Econ. 74 (2018) 777–786.
[13] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50 (1982) 987–1007.
[14] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, J. Econometrics 31 (1986) 307–327.
[15] S.J. Taylor, Modeling Financial Time Series, John Wiley, Chichester, 1986.
[16] T. Takaishi, Rational GARCH model: an empirical test for stock returns, Physica A 473 (2017) 451–460.
[17] S.R. Bentes, K.A. Dawson, J.O. Indekeu, et al., Is stock market volatility asymmetric? a multi-period analysis for five countries, Physica A 499 (2018) 258–265.
[18] Z. Pan, L. Liu, Forecasting stock return volatility: a comparison between the roles of short-term and long-term leverage effects, Physica A 492 (2018) 168–180.
[19] S. He, Y. Wang, Revisiting the multifractality in stock returns and its modeling implications, Physica A 467 (2017) 11–20.
[20] J. Kim, Y.J. Park, D. Ryu, Testing CEV stochastic volatility models using implied volatility index data, Physica A 499 (2018) 224–232.
[21] T.G. Andersen, T. Bollerslev, Answering the critics: Yes, ARCH models do provide good volatility forecasts, Int. Econ. Rev. 4 (1998) 885–905.
[22] B.J. Blair, S.H. Poon, S.J. Taylor, Forcasting S & P100 volatility: The incremental information content of implied volatility and high frequency index returns, J. Econometrics 45 (2001) 195–213.
[23] T.G. Andersen, T. Bollerslev, F.X. Diebold, et al., Modeling and forecasting realized volatility, Econometrica 71 (2003) 579–625.
[24] S.J. Koopman, B. Jungbacker, E. Hol, Forecasting daily variability of the S & P 100 stock index using historical, realised and implied volatility measurements, J. Empir. Financ. 12 (2005) 445–475.

[25] J. Kotkatvuori-Örnberg, Measuring actual daily volatility from high frequency intraday returns of the S & P futures and index observations, Expert Syst. Appl. 43 (2016) 213–222.
[26] U.A. Müller, M.M. Dacorogna, R.D. Davé, et al., Fractals and Intrinsic Time: A Challenge to Econometricians, Olsen & Associates, Zürich, 1993, (Unpublished manuscript).
[27] F. Corsi, A simple approximate long-memory model of realized-volatility, J. Financ. Economet 7 (2009) 174–196.
[28] T.G. Andersen, T. Bollerslev, X. Huang, A reduced form framework for modeling volatility of speculative prices based on realized variation measures, J. Econometrics 160 (2011) 176–189.
[29] F. Ma, Y. Wei, D. Huang, et al., Which is the better forecasting model? A comparison between HAR-RV and multifractality volatility, Physica A 405 (2014) 171–180.
[30] T.G. Andersen, T. Bollerslev, F.X. Diebold, Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility, Rev. Econ. Statist. 89 (2007) 701–720.
[31] C. Huang, X. Gong, X. Chen, et al., Measuring and forecasting volatility in Chinese stock market using HAR-CJ-M model, Abstr. Appl. Anal. (2013).
[32] S. Çelik, H. Ergin, Volatility forecasting using high frequency data: Evidence from stock markets, Econ. Model. 36 (2014) 176–190.
[33] J. Bekierman, H. Manner, Forecasting realized variance measures using time-varying coefficient models, Int. J. Forecast. 34 (2) (2018) 276–287.
[34] S. Dong, Y. Feng, Does index futures trading cause market fluctuations? China Financ. Rev. Int. 8 (2018) 173–198.
[35] D. Duong, N.R. Swanson, Empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction, J. Econometrics 187 (2015) 606–621.
[36] X. Gong, B. Lin, Structural changes and out-of-sample prediction of realized range-based variance in the stock market, Physica A 494 (2018) 27–39.
[37] H. Peng, R. Chen, D. Mei, et al., Forecasting the realized volatility of the Chinese stock market: Do the G7 stock markets help? Physica A 501 (2018) 78–85.
[38] W. Pu, Y. Chen, F. Ma, Forecasting the realized volatility in the Chinese stock market: further evidence, Appl. Econ. 48 (2016) 3116–3130.
[39] X. Wang, C. Wu, W. Xu, Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects, Int. J. Forecast. 31 (2015) 609–619.
[40] Z. Wu, N.E. Huang, A study of the characteristics of white noise using the empirical mode decomposition method, Proc. R. Soc. Lond. A: Math. Phys. Eng. Sci. 460 (2004) 1597–1611.
[41] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (2009) 1–41.
[42] N.E. Huang, Z. Shen, S.R. Long, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, Process R. Soc. Lond. A 454 (1998) 903–995.
[43] X. Zhang, K.K. Lai, S.Y. Wang, A new approach for crude oil price analysis based on empirical mode decomposition, Energy Econ. 30 (2008) 905–918.
[44] X. Zhang, L. Yu, S.Y. Wang, K.K. Lai, Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method, Energy Econ. 31 (2009) 768–778.
[45] Z.H. Feng, C.F. Liu, Y.M. Wei, How does carbon price change? Evidences from EU ETS, Int. J. Glob. Energy Issues 35 (2011) 132–144.
[46] J.P. Li, L. Tang, X.L. Sun, et al., Country risk forecasting for major oil exporting countries: a decomposition hybrid approach, Comput. Ind. Eng. 63 (2012) 641–651.
[47] M. Xu, P. Shang, A. Lin, Cross-correlation analysis of stock markets using EMD and EEMD, Physica A 442 (2016) 82–90.
[48] L.Y. Liu, A.J. Patton, K. Sheppard, Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes, J. Econometrics 187 (2015) 293–311.
[49] D. Amaya, P. Christoffersen, K. Jacobs, et al., Does realized skewness predict the cross-section of equity returns? J. Financ. Econ. 118 (2015) 135–167.
[50] T.G. Andersen, D. Dobrev, E. Schaumburg, Jump-robust volatility estimation using nearest neighbor truncation, J. Econometrics 169 (2012) 75–93.
[51] T. Bollerslev, D. Osterrieder, N. Sizova, et al., Risk and return: Long-run relations, fractional cointegration, and return predictability, J. Financ. Econ. 108 (2013) 409–424.
[52] X. Gong, B. Lin, Forecasting the good and bad uncertainties of crude oil prices using a HAR framework, Energy Econ. 67 (2017) 315–327.
[53] Y. Wang, F. Ma, Y. Wei, et al., Forecasting realized volatility in a changing world: A dynamic model averaging approach, J. Bank. Financ. 64 (2016) 136–149.
[54] F. Wen, X. Gong, S. Cai, Forecasting the volatility of crude oil futures using HAR-type models with structural breaks, Energy Econ. 59 (2016) 400–413.
[55] P.R. Hansen, A. Lunde, Consistent ranking of volatility models, J. Econometrics 131 (2006) 97–121.
[56] X. Gong, Z. He, P. Li, et al., Forecasting return volatility of the CSI 300 index using the stochastic volatility model with continuous volatility and jumps, Discrete Dyn. Nat. Soc. 2014 (2014) 1–10.
[57] O.E. Barndorff-Nielsen, N. Shephard, Power and bipower variation with stochastic volatility and jumps, J. Financ. Econometrics 2 (2004) 1–37.
[58] O.E. Barndorff-Nielsen, N. Shephard, Econometrics of testing for jumps in financial economics using bipower variation, J. Financ. Econometrics 4 (2006) 1–30.
[59] X. Gong, B. Lin, The incremental information content of investor fear gauge for volatility forecasting in the crude oil futures market, Energy Econ. 74 (2018) 370–386.
[60] X. Gong, B. Lin, Structural breaks and volatility forecasting in the copper futures market, J. Futures Mark. 38 (3) (2018) 290–339.
[61] X. Huang, G. Tauchen, The relative contribution of jumps to total price variance, J. Financ. Econometrics 3 (2005) 456–499.
[62] Q. Ji, D. Zhang, China's crude oil futures: introduction and some stylized facts, Financ. Res. Lett. (2018) http://dx.doi.org/10.1016/j.frl.2018.06.005.

Xu Gong is an associate professor. He works in School of Management, China Institute for Studies in Energy Policy, Collaborative Innovation Center for Energy Economics and Energy Policy, Xiamen University, Xiamen, China.

His research centers on volatility forecasting, assets pricing, and energy finance. Email: xugong@xmu.edu.cn

Boqiang Lin is a professor. He also works in School of Management, China Institute for Studies in Energy Policy, Collaborative Innovation Center for Energy Economics and Energy Policy, Xiamen University, Xiamen, China.

His research centers on financial risk management and energy finance. Email: bqlin@xmu.edu.cn