Physica A 499 (2018) 288–297

Contents lists available at ScienceDirect

Physica A

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

Forecasting the value-at-risk of Chinese stock market using the HARQ model and extreme value theory

Guangqiang Liu [a], Yu Wei [b,*], Yongfei Chen [c], Jiang Yu [a], Yang Hu [a]

[a] School of Economics and Management, Southwest Jiaotong University, Chengdu, China
[b] School of Finance, Yunnan University of Finance and Economics, 237 Longquan Road, Kunming, Yunnan, China
[c] School of Economics, Yunnan University of Finance and Economics, Kunming, China

HIGHLIGHTS

- HAR-type models are used to model the realized volatility of CSI300 index.
- The one-day-ahead VaR of CSI300 index is forecasted using the HARQ, HARQF, HAR, HARJ, CHAR, SHAR and EVT theory.
- The HARQ-type models are superior to several traditional HAR-type models in forecasting the VaR of Chinese stock market.

ARTICLE INFO

Article history:
Received 8 September 2017
Received in revised form 25 November 2017
Available online 15 February 2018

Keywords:
Realized volatility
HARQ
Extreme value theory
VaR

ABSTRACT

Using intraday data of the CSI300 index, this paper discusses value-at-risk (VaR) forecasting of the Chinese stock market from the perspective of high-frequency volatility models. First, we measure the realized volatility (RV) with 5-minute high-frequency returns of the CSI300 index and then model it with the newly introduced heterogeneous autoregressive quarticity (HARQ) model, which can handle the time-varying coefficients of the HAR model. Second, we forecast the out-of-sample VaR of the CSI300 index by combining the HARQ model and extreme value theory (EVT). Finally, using several popular backtesting methods, we compare the VaR forecasting accuracy of HARQ model with other traditional HAR-type models, such as HAR, HAR-J, CHAR, and SHAR. The empirical results show that the novel HARQ model can beat other HAR-type models in forecasting the VaR of the Chinese stock market at various risk levels.

1. Introduction

In recent years, there has been keen research interest in using intraday high-frequency transaction data to examine the volatility and value-at-risk of financial assets, with many of the contributions coming from the field of econophysics [1–6]. Realized volatility (RV), proposed by Anderson and Bollerslev [7] and by Barndorff-Nielsen and Shephard [8], lays an important foundation for the study of high-frequency volatility modeling. The simple and easy-to-estimate heterogeneous autoregressive (HAR) model of Corsi [9] has arguably emerged as the preferred specification for RV-based modeling. The HAR model, which overcomes deficiencies related to the computational complexity of the ARFIMA model, has become an important idea for further RV research [10,11].

By dividing the total volatility into continuous and discontinuous volatility (jumps) and using the bi-power variation measure of Barndorff-Nielsen and Shephard [12,13], HAR-with-jumps (HAR-J) and continuous-HAR (CHAR) are derived from

* Corresponding author.
E-mail address: weiyusy@126.com (Y. Wei).

https://doi.org/10.1016/j.physa.2018.02.033
0378-4371/

the HAR model. Moreover, Patton and Sheppard [14] propose semi-variance-HAR (SHAR) by separating the total volatility into positive and negative volatility on the basis of semi-variance measurement. RV is affected by the sum of integrated volatility (IV), and although high-frequency data is used in the modeling process due to IV discontinuity, the measurement method leads to errors-in-variables. The traditional method is to take measurement error variance as constant so as to reduce errors-in-variables. Barndorff-Nielsen and Shephard [13] propose a progressive distribution theory that considers instantaneous volatility of errors-in-variables in RV modeling. Perceptually, the smaller the measurement error variance is, the more precise the forecasting value will be. Addressing measurement error variance, Bollerslev et al. [15] propose the HAR quarticity (HARQ) model, which considers the time-varying characteristics of the autoregressive parameter during the process of modeling. Their results show that HARQ has remarkable advantages over the traditional HAR models.

This paper forecasts the volatility of the Chinese stock market with the newly introduced HARQ model and compares its performance with that of traditional HAR-RV-type ones. Our motivation to investigate the Chinese stock market is twofold. First, China is the second largest economy in the world, and its stock market is becoming increasingly influential in the global financial system. Second, the HARQ-type models proposed by Bollerslev et al. [15] are applied mainly in developed stock markets, and to our knowledge, no study has yet used them in the Chinese stock market.

In addition, although financial market volatility is a type of risk for market participants, in practice it is not directly observable or measurable for investors. It is thus necessary to use other indicators to measure the market risk of financial assets. Since the Basel Accord II, value-at-risk (VaR) has become the benchmark of market risk measurement in both academia and industry. VaR is a statistical technique used to measure and quantify the level of market risk within a firm or investment portfolio over a specific time horizon. In recent years, some studies have begun to focus on the measurement and forecasting of VaR using realized volatility. For example, Brownless and Gallo [16] analyze the VaR forecasting effect of five main high-frequency volatility measurement methods: RV, BPV, TSRV, PKV, and realized range-based volatility. Giot and Laurent [17] forecast the one-day-ahead VaR with a daily ARCH-type model based on daily realized volatility. Jung and Maderitsch [18] investigate the volatility transmission between stock markets in Hong Kong, Europe, and the United States using intraday data covering the period 2000 to 2011. Bedowska-Sojka [19] calculates the one-day-ahead forecasts of VaR for the WIG20 index quoted on the Warsaw Stock Exchange with HAR-RV, HAR-RV-J, and ARFIMA models.

With the development of related theories of high-frequency volatility, it makes sense to apply these innovative methods to the Chinese financial market. This study takes high-frequency 5-minute data of the CSI300 index as the example and introduces the new HARQ model into the market risk (VaR) measurement of the Chinese stock market. First, this paper models and forecasts the volatility of the CSI300 index using the HARQ model, which can take RV measurement error variance into account and allow the parameters of the models to vary explicitly with the (estimated) model uncertainty. This represents an important extension of traditional HAR-type models. Second, a combination of the HARQ model and extreme value theory (EVT) is used to measure and forecast the VaR of the CSI300 index. EVT relates to the asymptotic behavior of extreme observations of a random variable. It provides the fundamentals for the statistical modeling of rare events, and is used to compute tail-related measures. Lastly, the forecasting performances of various HAR-type models are quantitatively evaluated with the unconditional coverage test, independence test, and conditional coverage test. The empirical results show that the HARQ model can produce more accurate VaR forecasts than several traditional HAR-type models, illustrating the importance of considering RV measurement error variance when modeling and forecasting future volatility or market risk.

The rest of this paper is structured as follows. Section 2 provides RV, HAR, EVT, and VaR theories. Section 3 reports the empirical results, and Section 4 concludes.

2. Methodology

2.1. RV and the basic HAR-RV model

Consider a single asset for which the price process p_t is determined by the stochastic differential equation:

$$d\log(p_t) = \mu_t d_t + \sigma_t dW_t, \tag{1}$$

where μ_t and σ_t denote the drift and the instantaneous volatility processes, respectively, and W_t is a standard Brownian motion. Situations involving price jumps are not discussed here for convenient elaboration, and discontinuous price fluctuation (jumps) re further analyzed in the following part. Following the vast realized volatility literature, the main aim of the study is to forecast the latent integrated variance (IV) over a daily horizon. Specifically, normalizing the unit time interval to a day, the one-day IV is formally defined by:

$$IV_t = \int_{t-1}^{t} \sigma_s^2 ds. \tag{2}$$

The integrated variance is not directly observable. However, the realized variance or realized volatility (RV), defined by the summation of high-frequency returns, is observable as:

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2, \tag{3}$$

where $M = 1/\Delta$, 1 represents the trading time of day t, Δ represents the sum number of transaction times of trading frequency, and Δ-period intraday return is defined by $r_{t,i} = \log\left(P_{t-1+i\Delta}\right) - \log\left(P_{t-1+(i-1)\Delta}\right)$. According to the progressive distribution theory proposed by Barndorff-Nielsen and Shephard [13], when $\Delta \to 0$, the estimation error of RV_t can be defined as

$$RV_t = IV_t + \eta_t, \quad \eta_t \sim MN(0, 2\Delta IQ_t), \tag{4}$$

where $IQ_t = \int_{t-1}^{t} \sigma_s^4 ds$ denotes the integrated quarticity (IQ), and MN stands for mixed normal distribution, i.e., a normal distribution conditional on the realization of IQ_t . In parallel to the integrated variance, the IQ may consistently be estimated by the realized quarticity (RQ),

$$RQ_t = \frac{M}{3} \sum_{i=1}^{M} R_{t,i}^4. \tag{5}$$

The basis of recent RV research has been the heterogeneous autoregressive model of realized volatility proposed by Corsi [9]:

$$RV_t = \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t, \tag{6}$$

where $RV_{t-j|t-h} = \frac{1}{h+1-j} \sum_{i=j}^{h} RV_{t-i}, j \leq h$. The corresponding daily, weekly and monthly realized volatility can be calculated by selecting lagged RV at one day, one week, or one month.

2.2. Other traditional HAR-RV-type models

HAR-with-jumps (HAR-J) and continuous-HAR (CHAR) are derived from the HAR model by dividing the total volatility into continuous and discontinuous volatility. This decomposition is most commonly implemented using the bi-power variation (BPV) measure of Barndorff-Nielsen and Shephard [12,13], which affords a consistent estimate of the continuous variation in the presence of jumps. The HAR-J model, in particular, includes a measure of the jump variation as an additional explanatory variable in the standard HAR model:

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + \beta_J J_{t-1} + u_t, \tag{7}$$

where $J_t = \max\left[RV_t - BPV_t, 0\right]$, and BPV is defined as

$$BPV_t = u_1^{-2} \sum_{i=1}^{M-1} |r_t, i||r_t, i+1|, \tag{8}$$

where $u_1 = \sqrt{2/\pi} = E\left(|Z|\right)$, and Z is a standard normal random distribution. Empirically, the jump component has typically been found to be largely unpredictable. This motivates the alternative CHAR model, which only includes measures of the continuous variation on the right-hand side of Eq. (8):

$$RV_t = \beta_0 + \beta_1 BRV_{t-1} + \beta_2 BRV_{t-1|t-5} + \beta_3 BRV_{t-1|t-22} + u_t. \tag{9}$$

Several empirical studies have documented that the HAR-J and CHAR models often perform (slightly) better than the basic HAR model. Patton and Sheppard [14] propose a semi-variance HAR model (SHAR), the main idea being to decompose the total volatility into positive and negative volatility based on the semi-variance measures. In particular, letting $RV_t^- = \sum_{i=1}^{M} r_{t,i}^2 I\{r_{t,i} < 0\}$ and $RV_t^+ = \sum_{i=1}^{M} r_{t,i}^2 I\{r_{t,i} > 0\}$, the SHAR model is then defined as:

$$RV_t = \beta_0 + \beta_1^+ RV_{t-1}^+ + \beta_1^- RV_{t-1}^- + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t. \tag{10}$$

2.3. HARQ and HARQF models

On the basis of the standard HAR, Bollerslev et al. [15] propose the following model by allowing the parameters to vary explicitly with the variance of RV measurement errors:

$$\begin{aligned} RV_t = {} & \beta_0 + \underbrace{\left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right)}_{\beta_{1,t}} RV_{t-1} + \underbrace{\left(\beta_2 + \beta_{2Q} RQ_{t-1|t-5}^{1/2}\right)}_{\beta_{2,t}} RV_{t-1|t-5} \\ & + \underbrace{\left(\beta_3 + \beta_{3Q} RQ_{t-1|t-22}^{1/2}\right)}_{\beta_{3,t}} RV_{t-1|t-22} + u_t, \end{aligned} \tag{11}$$

where $RQ_{t-1|t-k} = \frac{1}{k} \sum_{j=1}^{k} RQ_{t-j}$. Of course, the magnitude of the (normalized) measurement errors in the realized volatilities will generally decrease with the horizon k as the errors are averaged out. This suggests that adjusting for measurement errors in the daily lagged realized volatilities is likely to play a more important role than adjusting for the weekly and monthly coefficients. The model in Eq. (11), which allows all of the parameters to vary with an estimate of the measurement error variance, is referred to as the "full HARQ" model, or HARQF, by Bollerslev et al. [15]. As the simplified version of the model in Eq. (11), the HARQ model only allows the coefficient on the daily lagged RV to vary as a function of $RQ^{1/2}$ as shown by Eq. (12):

$$RV_t = \beta_0 + \underbrace{\left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right)}_{\beta_{1,t}} RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t. \tag{12}$$

2.4. Extreme value theory

Extreme value theory, which originates from the field of hydrology, has been widely applied to the modeling of extreme financial movements, with good empirical results [20,21]. In brief, extreme value theory is a concrete method for modeling the distribution of the maximum values over a threshold u. Unitary extreme value theory includes BMM (Block Maxima Model) and POT (Peaks over Threshold). According to Pickands [22], giving a sufficiently large threshold u of a sequence variable X, the series' distribution over the threshold is similar to the generalized Pareto distribution (GPD). With the increase of u, the distribution of $F_u(x)$ can be fitted through the GPD, which is defined as follows:

$$F(x) = \begin{cases} 1 - \left(1 + \xi \dfrac{x}{\beta}\right)^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - e^{-\frac{x}{\xi}}, & \xi = 0. \end{cases} \tag{13}$$

There is no unanimous agreement on how to determine u. DuMouchel [23] suggests determining u and carrying out the estimation by selecting the top and bottom 10% of observations as the extreme values. Therefore, when the GPD is used to model and analyze the CSI300 index, the highest 10% and lowest 10% of observations are selected as the extreme values.

2.5. Forecasting methodology of VaR

Value-at-risk (VaR) is a widely used risk measurement tool in the financial field. It is specified as the maximum loss faced by financial assets in a future horizon under a certain quantile level. If the one-day-ahead VaR under the 1% quantile level is Q, this indicates that the asset loss will not exceed Q in the next day with a probability of 99%. As stated above, the congruent relationship between the daily return and the standardized return is:

$$r_t = \sqrt{h_t} * \vartheta_t, \vartheta_t \sim F, \tag{14}$$

where r_t is the daily return, ϑ_t is the standardized return, h_t is the volatility that can be estimated by different HAR-type models, and F is the probability distribution function for standardized returns. The VaR under the quantile of $100(1-p)\%$ forecasted with the information of t-1 day can be expressed as:

$$VaR_{t|t-1}^{p} = F_{(1-p)}^{-1} \sqrt{h_{t|t-1}}, \tag{15}$$

where $F_{(1-p)}^{-1}$ is the (1-p) quantile of F distribution. Moreover, the concrete distribution of standardized return ϑ_t needs to be discussed, and extreme value theory is used in this paper. Therefore, forecasting the VaR of the CSI300 index has two main steps: (1) RV is forecasted by making use of HAR-type models; and (2) extreme value theory (GPD) is used to depict the standardized return. After that, VaR can be forecasted. Based on the earlier discussion of extreme value theory and the estimation results of EVT, the VaR under the quantile level of p is as follows:

$$VaR_{t|t-1}^{p} = F_{(1-p)}^{-1} \sqrt{\hat{h}_{t|t-1}} = \left\{ u + \frac{\hat{\beta}}{\hat{\xi}} \left[\left(\frac{1-p}{n_u/n} \right)^{-\hat{\xi}} - 1 \right] \right\} \sqrt{\hat{h}_{t|t-1}}, \tag{16}$$

where $\hat{\beta}$ and $\hat{\xi}$ are the estimated parameters of a GPD, and $\hat{h}_{t|t-1}$ is the volatility forecasted by various HAR-type models discussed in Sections 2.1 to 2.3.

2.6. VaR backtesting

It is essential to inspect the accuracy of the VaR forecasting results with related quantitative methods. The VaR backtesting methods include the unconditional coverage test, independent test, and conditional coverage test. Unconditional coverage (UC) judges the forecasting accuracy of the model by the statistical consistency of the predicted failure rate with the theoretical failure rate. Kupiec [24] proposes the UC test by judging the difference of the two failure rates. First, define the hit sequence as:

$$S_t = \begin{cases} 1, & if \ \ r_t < VaR_t, \\ 0, & if \ \ r_t \geq VaR_t. \end{cases} \tag{17}$$

The upper formulation means that the S_t sequence is determined by the comparison of the return at t and the VaR forecasted with the relevant method. According to the definition of VaR (a negative value represents losses) in this paper, we assign the value of S_t as 1 when r_t is less than VaR_t or 0 otherwise. If T is the total number of VaR series, then it is used to inspect the overall length of the sequence; T_1 is the number of $S_t = 1$, and T_0 is the number of $S_t = 0$. The null hypothesis of UC backtesting of VaR forecasted under the quantile level α is:

$$H_0 : (T_1/T) = \alpha. \tag{18}$$

The likelihood ratio (LR) is defined as:

$$LR = -2\ln\left\{(1-\alpha)^{T_0}\alpha^{T_1}/\left[(1-T_1/T)^{T_0}(T_1/T)^{T_1}\right]\right\} \sim \chi^2(1). \tag{19}$$

Therefore, the validity of VaR forecasting can be judged by inspecting whether the distribution of LR follows the distribution of $\chi^2(1)$. More specifically, the p-value of the UC test is used to consider whether to reject the null hypothesis H_0. The bigger the p-value is, the less likely to be rejected is the null hypothesis and the more reliable is the corresponding model used in the forecasting of VaR.

However, validation of VaR forecasting only through a check of the difference between the empirical failure rate and the theoretical failure rate is insufficient. If, for example, the theoretical failure rate is 10% for 1000 VaR forecasting observations and there are also 100 empirical failures, *but* all 100 occur consecutively, it is easy to conclude that the forecasting model is not proper, even though it technically passes UC backtesting. To solve this problem, some studies propose checking whether failures occur independently of each other. This backtesting method is called independent backtesting (IND). Another method, labeled conditional coverage backtesting (CC), is used to test the UC and IND jointly. Furthermore, Candelon et al. [25] propose a GMM duration-based method to backtest the IND null hypothesis. First, d_i as the duration between two consecutive violations is defined as: $d_i = t_i$-t_{i-1}.

Definition. The orthonormal polynomials associated to a geometric distribution with a success probability β are defined by the following recursive relationship,$\forall d \in N^*$:

$$M_{j+1}(d;\beta) = \frac{(1-\beta)(2j+1)+\beta(j-d+1)}{(j+1)\sqrt{1-\beta}} M_j(d;\beta) - \left(\frac{j}{j+1}\right) M_{j-1}(d;\beta). \tag{20}$$

For any order $j \in N^*$, with $M_{-1}(d;\beta) = 0$ and $M_0(d;\beta) = 1$, if the true distribution is a geometric distribution with a success probability β, then it follows that:

$$E\left[M_j(d;\beta)\right] = 0, \forall j, d \in N^*. \tag{21}$$

Assume that $\{d_1; d_2; \ldots; d_N\}$ are the durations between VaR violations, computed from the sequence of the hit variables $\{S_t(\alpha)\}_{t=1}^T$. Under the CC assumption, the durations d_i, $i = 1, 2, \ldots, N$, are i.i.d. and have a geometric distribution with a success probability equal to the coverage rate α. Hence, the null hypothesis of CC can be expressed as follows:

$$H_{0,CC} : E\left[M_j(d_i;\alpha)\right] = 0, j = \{1, 2, \ldots, p\}, \tag{22}$$

where p denotes the number of moment conditions. A separate test for the IND hypothesis can also be derived. It consists in testing the hypothesis of a geometric distribution (implying the absence of dependence) with a success probability equal to β, where β denotes the true violation rate that is not necessarily equal to the coverage rate α. This independence assumption can be expressed as the following moment condition:

$$H_{0,IND} : E\left[M_j(d_i;\beta)\right] = 0, j = \{1, 2, \ldots, p\}. \tag{23}$$

In this case, the expectation of the duration variable is equal to $1/\beta$ as soon as the first polynomial $M_1(d;\beta)$ is included in the set of moment conditions. Therefore, under $H_{0,IND}$, the durations between two consecutive violations have a geometric distribution. According to Candelon et al. [25], the statistic for the CC test is:

$$J_{CC}(p) = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} M(d_i;\alpha)\right)^T \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} M(d_i;\alpha)\right) \frac{d}{N\to\infty} \chi^2(p), \tag{24}$$

where $M(d_i;\alpha)$ denotes a $(p, 1)$vector whose components are the orthonormal polynomials $M_j(d_i;\alpha)$, for $j = \{1, 2, \ldots, p\}$ and α denotes the coverage rate. Similarly, the statistic for IND is:

$$J_{IND}(p) = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} M\left(d_i;\hat{\beta}\right)\right)^T \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} M\left(d_i;\hat{\beta}\right)\right) \frac{d}{N\to\infty} \chi^2(p-1). \tag{25}$$

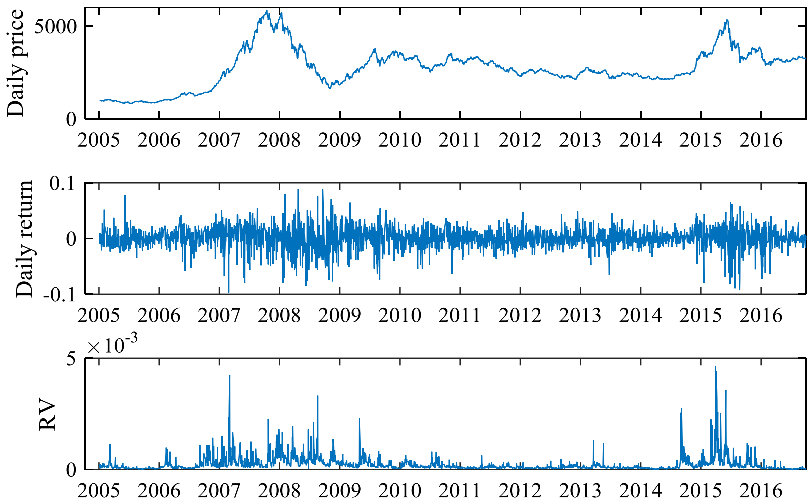

Fig. 1. The time evolutions of daily price, return and RV of CSI300 index.

Table 1
The descriptive statistics of the daily return of CSI300 index.

Mean	Std. Dev.	Skewness	Kurtosis	Jarque–Bera	ADF
0.0004	0.0185	-0.5225^{***}	3.3366^{***}	1453.8^{***}	-51.8077^{***}
0.0002	0.0003	5.3326^{***}	47.1052^{***}	244851.6^{***}	-24.3606^{***}

Note: The Jarque–Bera statistic tests for the null hypothesis of normality in the sample distribution. ADF is the statistic of the augmented Dickey–Fuller test. *** indicates significance at the 1% level. ADF is the statistic of the augmented Dickey–Fuller test.

In our empirical study, the null hypothesis is tested through IND, UC, and CC statistics. The criterion is based on the statistical p value. The larger the p value is, the more reliable the null hypothesis is. In other words, a larger p value in one backtesting indicates a more accurate VaR forecasting result, implying that the corresponding volatility model is more reliable.

3. Empirical analysis

3.1. Data and descriptive statistics

As the sample, this paper takes the CSI300 index, which contains the market information of two stock exchanges in China, i.e., the Shanghai Stock Exchange and the Shenzhen Stock Exchange. The 5-minute high-frequency data are taken from the CSI300 index initially issued, covering the period January 4, 2005 to September 29, 2016 (total of 2854 trading days). There are 48 high-frequency (5-minute) data with 4 h of trading each day. We obtain the data from the CCER-Xenophon stock market high-frequency database. Using this 5-minute high-frequency data sample, we can calculate the logarithm intraday returns, daily returns, and daily realized volatilities of the CSI300 index. Fig. 1 presents the time evolutions of daily price, return, and RV of the CSI300 index. We can see obvious volatility clustering existing in the return and RV series.

Table 1 shows the descriptive statistics of the daily return and RV of the CSI300 index. We find that the skewness and kurtosis are significantly different from zero for the two series, and the Jarque–Bera statistics also indicate the rejection of a normal distribution for the CSI300 return. The augmented Dickey–Fuller tests support the rejection of the null hypothesis of a unit root at the 1% significance level, implying that the two series are stationary and may be modeled directly without further transforms.

3.2. In-sample estimation of the HAR-type models

To evaluate the overall goodness-of-fit of various HAR-type models, we estimate these models based on the whole data sample. Table 2 reports the in-sample estimation results of the six HAR-type models in this paper.

Table 2 indicates, first, that most of the coefficients of these HAR-type models are significant, implying that the HAR-type models may model the volatilities of the CSI300 index well. Second, the coefficient β_1^- in the SHAR model is significantly positive, while β_1^+ is not, indicating that negative returns may contribute more to future volatility than positive returns, which is well known as the leverage effect in volatilities. Lastly, the adjusted R^2 of various HAR-type models are all larger than 0.45, revealing that the overall goodness-of-fit of these models is acceptable. It is also worth noting that the novel HARQF and HARQ models get the largest and the second largest adjusted R^2.

Table 2
In-sample estimation results for the six HAR-type models.

	HAR	HAR-J	CHAR	SHAR	HARQ	HARQF
β_1	0.3947*** (0.0843)	0.4473*** (0.1431)	0.4838*** (0.1228)		0.5618*** (0.0689)	0.4328*** (0.0730)
β_2	0.2063*** (0.0900)	0.2036*** (0.0893)	0.2712*** (0.1285)	0.2443*** (0.0858)	0.1939*** (0.0896)	0.5405*** (0.1074)
β_3	0.2863*** (0.0990)	0.2977*** (0.0945)	0.4520*** (0.1403)	0.2744*** (0.0903)	0.2102*** (0.0757)	0.0964 (0.0918)
β_J		−0.2210 (0.3771)				
β_1^+				−0.2109 (0.1644)		
β_1^-				0.9438*** (0.1285)		
β_{1Q}					−281.6967*** (122.4045)	−99.3611 (112.6335)
β_{2Q}						−741.1066*** (189.0760)
β_{3Q}						45.8383 (319.7285)
Adj. R^2	0.4627	0.4641	0.4593	0.4727	0.4754	0.4857

Note: The bracketed numbers are the standard errors of the estimations. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

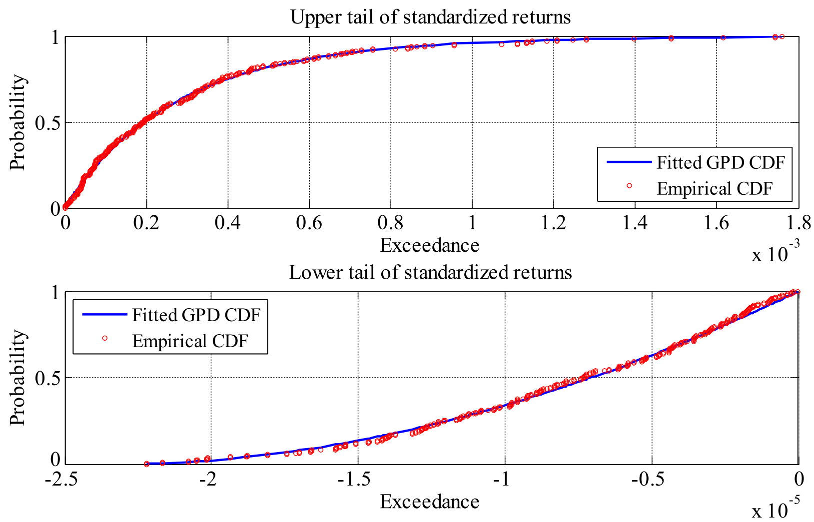

Fig. 2. The empirical and fitted GDP CDFs of CSI300 standardized returns.

3.3. *Modeling the return tails using EVT*

Using the volatilities estimated by different HAR-type models, we can get the standardized returns defined in Eq. (14). Then, we can use the GPD distribution to fit the tails of the standardized returns. Fig. 2 shows the estimated upper and lower tails of the standardized returns based on the volatilities estimated from the HARQ model.

Fig. 2 shows that although the empirical upper and lower tails of the standardized returns are not symmetric, the GPD fits the two tails quite well. This supports the rationality of using the extreme value method to describe the large risk of the CSI300 index.

3.4. *VaR forecasting results and backtesting*

In this section, we use the rolling window method to forecast the one-day-ahead VaR of the CSI300 index based on various HAR-type models. The rolling window is fixed to contain 1854 observations, and the last 1000 observations of the whole sample are chosen as the out-of-sample forecasting horizon. To get robust conclusions, we forecast the VaR of the CSI300 index at 5 different quantiles in the lower and upper tails, respectively, i.e., 0.005, 0.01, 0.025, 0.05 and 0.1. To be clear, Fig. 3

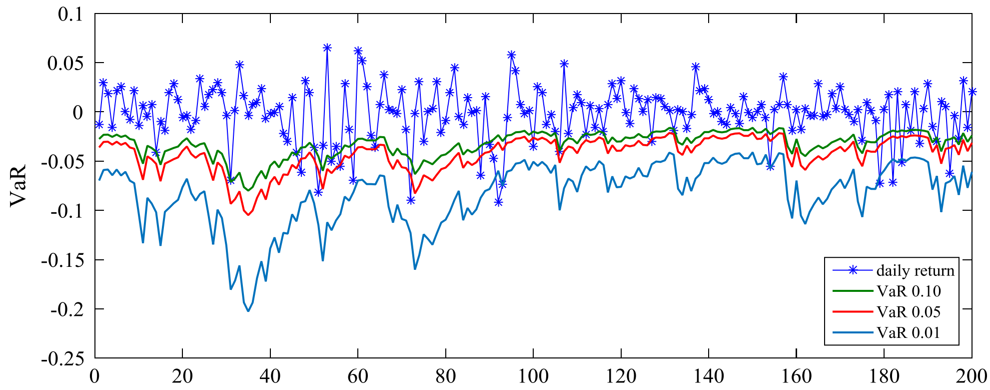

Fig. 3. The out-of-sample VaR forecasts of CSI300 index based on HARQ model.

Table 3
The backtesting results of UC and IND tests for the lower-tail VaR.

Test	Quantiles	Models					
		HAR	HAR-J	CHAR	SHAR	HARQ	HARQF
UC	0.10	<u>0.0220</u>	<u>0.0220</u>	<u>0.0295</u>	<u>0.0085</u>	0.0657	**0.1315**
	0.05	0.3746	0.3746	0.2332	0.1783	0.6603	**0.8843**
	0.025	0.5352	0.5352	0.4050	0.4050	**0.8384**	0.6892
	0.01	**0.7465**	0.7465	**0.7465**	0.3136	**0.7465**	**0.7465**
	0.005	0.6422	0.6422	0.6422	**1.0000**	0.6422	0.6422
IND(1)	0.10	<u>0.0124</u>	<u>0.0119</u>	<u>0.0155</u>	<u>0.0059</u>	**0.5767**	0.0586
	0.05	0.9074	0.9232	**0.9457**	0.1462	0.3151	0.2314
	0.025	**0.4816**	0.4535	0.3758	0.1312	0.3142	0.1518
	0.01	0.1918	0.1885	0.1177	0.0831	**0.4267**	0.4333
	0.005	0.3175	0.3175	<u>0.0225</u>	**0.4727**	0.3530	0.3940
IND(2)	0.10	0.0796	0.0794	<u>0.0132</u>	0.1809	**0.3379**	0.3121
	0.05	0.2547	0.2530	0.2125	0.2020	0.3390	**0.6092**
	0.025	0.0743	0.0722	0.0756	0.1408	0.2945	**0.6540**
	0.01	0.0582	0.0538	<u>0.0343</u>	0.0844	0.1640	**0.6565**
	0.005	0.4487	**0.4522**	0.4418	0.1848	0.2184	0.4131

Note: The numbers in the table are p-values of the backtesting results. The larger the p value, the higher the accuracy of the VaR predicted by the model. Underlining indicates rejection of the null hypothesis of the backtesting at the 5% significant level, indicating poor VaR forecasting accuracy of the corresponding model. Bold refers to the highest-accuracy model for VaR forecasting under a specific quantile level. IND(n) refers to n moment conditions, which are shown in detail in Eqs. (24)–(25).

only presents the first 200 out-of-sample lower-tail VaR forecasts at 3 quantiles based on the HARQ model. However, to evaluate quantitatively the forecasting accuracy of different VaR models, we depend on the results of UC, IND, and CC tests, which are reported in Tables 3 and 4.

Table 3 indicates that, with regard to UC backtesting results, under the five different quantiles, the HARQ and HARQF models achieve the 4 largest p-values, implying they can achieve better VaR forecasting accuracy than traditional HAR-models. In addition, all p-values of the HARQ and HARQF models are larger than 0.05. On the other hand, the HAR, HAR-J, CHAR, and SHAR models all get p-values smaller than 0.05, indicating their unstable VaR forecasting power for the Chinese stock market. In terms of IND(1) and IND(2) backtesting, the empirical results also show that all p-values under the HARQ and HARQF models are larger than 0.05. In 6 out of 10 quantile cases, HARQ and HARQF get the largest p-values among all rival HAR-type models. Again, the four traditional HAR-type models, HAR, CHAR, SHAR, and HAR-J, get p-values smaller than 0.05.

To jointly test the failure rate and independence accuracy, we refer to the results of CC backtesting reported in Table 4. The superior performance of the HARQ and HARQF models is very clear. Under the CC(1) test, all of the 5 largest p-values are achieved by the HARQ and HARQF models, while the other four models obtain p-values smaller than 0.05. Similarly, under the CC(2) test, the HARQ and HARQF models get 4 out of the 5 largest (and one second largest) p-values. We can conclude that the novel HARQ and HARQF models, which are capable of depicting the time-varying coefficients of HAR models, can make more accurate VaR forecasts than some traditional HAR-type models.

For robustness, Table 5 reports the UC, IND, and CC testing results for the upper-tail VaR.

Table 5 shows, first, that the performance of the HARQ and HARQF models is superior to that of traditional HAR-type models in forecasting the upper-tail VaR of the CSI300. In 14 out of 20 cases of the backtesting results, the HARQ and HARQF models produce the highest p-values, implying good predictive power across various evaluation criteria. Second, we also find

Table 4
The backtesting results of CC test for the lower-tail VaR.

Test	Quantiles	Models					
		HAR	HAR-J	CHAR	SHAR	HARQ	HARQF
CC (1)	0.10	<u>0.0197</u>	<u>0.0197</u>	<u>0.0289</u>	<u>0.0081</u>	**0.5767**	0.1533
	0.05	0.0608	<u>0.0606</u>	0.0683	<u>0.0041</u>	**0.3151**	0.0723
	0.025	<u>0.0377</u>	<u>0.0330</u>	<u>0.0362</u>	<u>0.0120</u>	**0.3142**	0.0758
	0.01	0.4390	0.4376	0.2748	0.2171	0.4267	**0.9444**
	0.005	0.3064	0.3057	<u>0.0154</u>	0.1748	0.3530	**0.7224**
CC(2)	0.10	0.0840	0.0860	<u>0.0155</u>	0.0830	0.3379	**0.5802**
	0.05	0.6130	0.6117	0.5178	0.4764	0.3390	**0.5815**
	0.025	0.1656	0.1699	0.1064	0.1598	0.2945	**0.7967**
	0.01	0.1043	0.1030	0.0522	0.1071	0.1640	**0.8431**
	0.005	0.9357	0.9268	**0.9446**	0.5093	0.2184	0.9423

Note: The numbers in the table are *p*-values of the backtesting results. The larger the *p* value, the higher the accuracy of the VaR predicted by the model. Underlining indicates rejection of the null hypothesis of the backtesting at the 5% significant level, indicating poor VaR forecasting accuracy of the corresponding model. Bold refers to the highest-accuracy model for VaR forecasting under a specific quantile level. CC(n) refers to n moment conditions, which are shown in detail in Eqs. (24)–(25).

Table 5
The backtesting results of the upper-tail VaR forecasts.

Test	Quantiles	Models					
		HAR	HAR-J	CHAR	SHAR	HARQ	HARQF
UC	0.90	0.9162	**1.0000**	0.8325	0.9159	0.5307	0.2988
	0.95	**0.8843**	0.7702	0.6603	0.5566	**0.8843**	0.3926
	0.975	0.1359	0.0857	0.0512	<u>0.0152</u>	0.2943	**0.4050**
	0.99	0.5102	0.7465	0.5102	<u>0.0301</u>	0.7465	**0.7544**
	0.995	<u>0.0285</u>	<u>0.0285</u>	<u>0.0015</u>	0.1258	0.1258	**0.3325**
IND(1)	0.90	<u>0.0207</u>	<u>0.0307</u>	<u>0.0347</u>	<u>0.0335</u>	**0.0714**	0.0663
	0.95	0.1915	0.2747	0.2007	0.1736	0.3854	**0.4198**
	0.975	0.0634	0.0630	0.0752	<u>0.0443</u>	0.1144	**0.2902**
	0.99	0.6408	0.6905	0.7926	0.6494	0.6859	**0.7986**
	0.995	0.7814	0.8654	**0.8711**	0.8403	0.7650	0.8195
IND(2)	0.90	**0.9072**	0.8698	0.7862	0.3467	0.7395	0.6797
	0.95	**0.5194**	0.4776	0.5131	0.3601	0.2693	0.1836
	0.975	0.4690	0.4880	0.4865	**0.5969**	0.4729	0.5387
	0.99	0.7781	0.8799	**0.9357**	0.1512	0.9527	0.9186
	0.995	0.9674	**0.9820**	0.9816	0.9707	<u>0.0176</u>	0.8699
CC(1)	0.90	<u>0.0383</u>	0.0668	0.0886	0.0716	0.1783	**0.1839**
	0.95	0.4292	0.6446	0.5280	0.4471	0.9346	**0.9484**
	0.975	0.1558	0.1552	0.1644	0.0790	0.2310	**0.6653**
	0.99	**0.5694**	0.3874	0.3190	0.5534	0.4069	0.2892
	0.995	0.2882	0.1312	0.1591	0.2876	0.2594	**0.3010**
CC(2)	0.90	0.0839	0.1373	0.1669	0.0376	0.3026	**0.2817**
	0.95	0.4789	**0.6373**	0.5720	0.1042	0.5471	0.4619
	0.975	0.0900	0.0914	0.0970	<u>0.0385</u>	0.1747	**0.6786**
	0.99	**0.8020**	0.6118	0.3596	0.4383	0.5863	0.3965
	0.995	0.3227	0.1684	0.1672	**0.4434**	0.2168	0.3460

Note: The numbers in the table are *p*-values of the backtesting results. The larger the *p* value, the higher the accuracy of the VaR predicted by the model. Underlining indicates rejection of the null hypothesis of the backtesting at the 5% significant level, indicating poor VaR forecasting accuracy of the corresponding model. Bold indicates the highest-accuracy model for VaR forecasting under a specific quantile level. IND(n) and CC(n) refer to n moment conditions, shown in detail in Eqs. (24)–(25).

that all of the *p*-values under the HARQ and HARQF models are larger than 10%, with the exception of one *p*-value under the IND(2) test. Finally, according to the CC tests, which jointly test the failure rate and independence accuracy, HARQF achieves the best forecasting accuracy in 6 out of 10 cases, further proving the validity of introducing time-varying coefficients into traditional HAR models.

4. Conclusions

Using high-frequency data to study volatility has attracted keen interest in recent years. The HAR model is commonly used in volatility measures using high-frequency data and is the basis of other correlated models in RV research, such as the SHAR, CHAR, and HAR-J models. In addition, the HARQ model recently proposed by Bollerslev et al. [15] is an important extension of traditional HAR-type models, taking time-varying autoregressive parameters into account on the basis of HAR. However, the extant literature on the Chinese stock market lacks any discussion of RV modeling with HARQ family models.

Against this backdrop, this paper discusses VaR forecasting using the HARQ model combined with extreme value theory. The VaR of the CSI300 index is forecasted with the HAR, HAR-J, CHAR, SHAR, HARQ, and HARQF models for comparison, and the prediction results of each model are backtested by the UC, IND, and CC methods. The empirical results show that the new HARQF and HARQ model can beat other HAR-type models in forecasting the VaR of the Chinese stock market. Thus, we think that research using the HARQ model and extreme value theory has important practical implications for risk management in the Chinese stock market.

Acknowledgments

We would like to show our sincere gratitude to five anonymous reviewers, as well as Professor Eugene Stanley, Editor of Physica A, whose comments and suggestions greatly improved the quality of this paper. The authors are also grateful for the financial support from the National Natural Science Foundation of China (71371157, 71671145), the humanities and social science fund of ministry of education, China (16YJA790062, 17YJA790015, 17XJA790002), and the young scholar fund of science & technology department of Sichuan province, China (2015JQO010), Fundamental research funds for the central universities, China (26816WCX02).

References

[1] R. Chen, C. Li, W. Wang, Z. Wang, Empirical analysis on future-cash arbitrage risk with portfolio VaR, Physica A 398 (2014) 210–216.
[2] F. Petroni, M. Serva, Observability of market daily volatility, Physica A 444 (2016) 838–842.
[3] S. Lahmiri, On fractality and chaos in Moroccan family business stock returns and volatility, Physica A 473 (2017) 29–39.
[4] Z. Liu, Y. Ye, F. Ma, J. Liu, Can economic policy uncertainty help to forecast the volatility: A multifractal perspective, Physica A 482 (2017) 181–188.
[5] D. Mei, J. Liu, F. Ma, W. Chen, Forecasting stock market volatility: Do realized skewness and kurtosis help? Physica A 481 (2017) 153–159.
[6] 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.
[7] T.G. Andersen, T.M. Bollerslev, DM-Dollar volatility: Intraday activity patterns, macroeconomic announcements, and longerrun dependencies, J. Finance 53 (1998) 219–265.
[8] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (2002) 253–280.
[9] F. Corsi, A simple approximate long-memory model of realized volatility, Scien. Elec. Publishing 7 (2009) 174–196.
[10] K. Choi, W.C. Yu, E. Zivot, Long memory versus structural breaks in modeling and forecasting realized volatility, J. Internat. Money & Finance 29 (2010) 857–875.
[11] G.M. Gallo, E. Otranto, Forecasting realized volatility with changing average levels, Int. J. Forecast. 31 (2015) 620–634.
[12] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realized co-variation: High frequency based covariance regression and correlation in financial economics, Economics 72 (2004) 885–925.
[13] O.E. Barndorff-Nielsen, N. Shephard, Power and bi-power variation with stochastic volatility and jumps, J. Econ. Finance 2 (2004) 1–37.
[14] A.J. Patton, K. Sheppard, Good volatility, bad volatility: Signed jumps and the persistence of volatility, Revi. Econ. Statist. 97 (2015) 683–697.
[15] T. Bollerslev, A.J. Patton, R. Quaedvlieg, Exploiting the errors: A simple approach for improved volatility forecasting, J. Econ. 192 (2016) 1–18.
[16] C.T. Brownless, G.M. Gallo, Comparison of volatility measures: A risk management perspective, J. Financ. Econ. 8 (2009) 29–56.
[17] P. Giot, S. Laurent, Modelling daily value-at-risk using realized volatility and ARCH type models, J. Empi. Financ. 11 (2004) 379–398.
[18] R.C. Jung, R. Maderitsch, Structural breaks in volatility spillovers between international financial markets: Contagion or mere interdependence? J. Bank. Financ. 47 (2014) 331–342.
[19] B. Bedowska-Sojka, Daily VaR forecasts with realized volatility and GARCH models, Argu. Oecon. 1 (2015) 157–173.
[20] A.J. Mcneil, Estimating the tails of loss severity distributions using extreme value theory, Astin Bull. 27 (1997) 117–137.
[21] R. Gençay, F. Selçuk, Extreme value theory and value-at-risk: Relative performance in emerging markets, Int. J. Forecast. 20 (2004) 287–303.
[22] J. Pickands III, Statistical inference using extreme order statistics, Ann. Statist. 3 (1975) 119–131.
[23] W.H. DuMouchel, Estimating the stable index α in order to measure tail thickness: A critique, Ann. Statist. 11 (1983) 1019–1031.
[24] P.H. Kupiec, Techniques for verifying the accuracy of risk measurement models, J. Derivatives 3 (1995) 73–84.
[25] B. Candelon, G. Colletaz, C. Hurlin, et al., Backtesting value-at-risk: A GMM duration-based test, J. Financ. Econ. 9 (2011) 314–343.