Economic Modelling 87 (2020) 148–157

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

Does measurement error matter in volatility forecasting? Empirical evidence from the Chinese stock market[☆]

Yajing Wang [a], Fang Liang [b], Tianyi Wang [a,*], Zhuo Huang [b]

[a] School of Banking and Finance, University of International Business and Economics, Beijing, China
[b] National School of Development, Peking University, Beijing, China

ARTICLE INFO

JEL classification:
C22
C51
C53
C58

Keywords:
Realized volatility
Measurement errors
Volatility forecasting
Chinese stock market

ABSTRACT

Based on methods developed by Bollerslev et al. (2016), we explicitly accounted for the heteroskedasticity in the measurement errors and for the high volatility of Chinese stock prices; we proposed a new model, the LogHARQ model, as a way to appropriately forecast the realized volatility of the Chinese stock market. Out-of-sample findings suggest that the LogHARQ model performs better than existing logarithmic and linear forecast models, particularly when the realized quarticity is large. The better performance is also confirmed by the utility based economic value test through volatility timing.

1. Introduction

The Chinese stock market is highly volatile due to features such as the various types of listed firms and investors. Individual investors, who are more likely to be noise traders, play a major role in driving the movement of Chinese stock prices, while the lack of security supplies leaves the market vulnerable to speculation, further worsening the situation. The highly volatile nature of the Chinese stock market demands a suitable econometric specification in order to appropriately model and forecast market volatility.

Studies of volatility within the Chinese stock market have so far used various econometric models. The classic choice is GARCH (Generalized Autoregressive Conditional Heteroskedasticity) class models that are frequently used to model and forecast the volatility of stock index futures and options (e.g., Fabozzi et al., 2004; Hou and Li, 2014; So and Tse, 2004; Yang et al., 2011). For example, Chen et al. (2013) studied the effect of index futures trading on spot volatility in the Chinese stock market and adopted a panel data evaluation approach to avoid a potentially omitted variable bias. Wei et al. (2011) highlighted the hedging effectiveness of the copula-MFV (MFV: Multifractal Volatility) model over the copula-GARCH models using the prices of the Chinese stock index spots and futures. The recent development of realized measures based models such as the Realized GARCH model (Hansen et al., 2012; Hansen and Huang, 2016) or the HEAVY (High Frequency Based Volatility) model (Shephard and Sheppard, 2010) and reduced-form models such as the HAR (Heterogeneous Autoregressive) model (Corsi, 2009) have also attracted considerable attention. Among them, the parsimonious and easy to estimate HAR model has been widely used for volatility forecasting in the Chinese stock market (e.g., Wang and Huang, 2012; Ma et al., 2014; Huang et al., 2016, etc).

However, Bollerslev et al. (2016) argued that the HAR model ignores time variability with regard to the magnitude of the realized volatility measurement errors, and thus suffers from an errors-in-variables problem. The errors have been shown to attenuate the parameters of the model, which is why they should be accounted for. As a solution, Bollerslev et al. (2016) proposed the linear HARQ model, which allows the model parameters to change with the magnitude of the measurement error. In their study, they showed that the linear HARQ model can outperform the HAR model. Similar to the HARQ model, the HARS model developed by Bekierman and Manner (2018) captures the effect of

☆ We would like to thank Prof. Paresh Narayan (Editor), Prof. Joakim Westerlund (Associate Editor) and an anonymous referee for valuable suggestions that greatly improved this paper.
* Corresponding author.
E-mail address: tianyiwang@uibe.edu.cn (T. Wang).

https://doi.org/10.1016/j.econmod.2019.07.014
Received 21 September 2018; Received in revised form 5 July 2019; Accepted 18 July 2019
Available online 22 July 2019

measurement errors by including a time-varying state variable.

Other than models that use the "raw" realized volatility[1], transformations of realized variance with nonlinear functions such as square root (Corsi, 2009), logarithm (Andersen et al., 2003, etc.), or the more general Box-Cox transformation (Gonçalves and Meddahi, 2011) are common in the literature. Among them, taking the logarithm is widely used in practice (such as in Andersen et al., 2007; Bollerslev et al., 2009; Bekaert and Hoerova, 2014; Andersen et al., 2019, etc.) and can at least be dated back to Nelson (1991) in volatility modeling literature. There are several advantages to working with log-linear volatility models. First, realized volatility is highly skewed with extreme values and changes rapidly, which presents a challenge in parameter estimations. As shown in Andersen et al. (2001) and others, the logarithmic transformation can ease both issues and lead the (logged) realized variance to an approximately normal distribution[2], which provides better statistical properties for estimation. Second, the log-transformed models generally produce better out-of-sample forecasts compared to linear versions[3] as mentioned in abundant literature (Ma et al., 2014; Hansen and Huang, 2016; Audrino and Hu, 2016, etc.). Third, log-linear models automatically ensure positive volatility and make it easier to add additional features such as leverage effects (Corsi and Reno, 2009), signed jump variations (Patton and Sheppard, 2009), etc. Thus, results based on log-linear models have wider applications.

Motivated by these, we build our model based on Bollerslev et al. (2016) but focused on logarithmic realized variance. We derive a logarithmic version of the linear HARQ model using the infill large sample theory and the asymptotic distribution of (log) realized volatility. Compared with the linear HARQ model, our model is better at forecasting the realized volatility, specifically when the realized variance is in rapid change. This means the LogHARQ model is suitable for forecasting the volatility of highly volatile markets which are characterized by considerable measurement errors. Our results are based on the China Securities Index (CSI) 300 and the Shanghai Stock Exchange (SSE) 50 ETF, both of which underlie China's index futures and options, respectively. Our empirical findings suggest that the LogHARQ model significantly improves out-of-sample forecasting accuracy relative to several commonly used volatility prediction models. The improvement is more pronounced when the realized quarticity is large.

In addition to the out-of-sample forecast improvements, we evaluated the economic benefit of using the LogHARQ model as a volatility forecasting model for investment decisions. Fleming et al. (2001) developed a framework for assessing the economic value of volatility timing strategies. They considered a risk-averse investor who has mean-variance preferences and allocates her wealth across different assets. Based on this, Fleming et al. (2003) used the realized volatility to form estimates for the conditional covariance matrix of asset returns; they found that the performance of volatility timing can be improved by using high-frequency data. Similarly, Marquering and Verbeek (2004) proposed a framework for evaluating the economic value of volatility timing strategies when allocating between two different assets, one of which is a risk asset while the other one remains risk-free. The framework is also applied by Nolte and Xu (2015) to show the inclusion of realized jumps in volatility timing can improve a risk-averse investor's portfolio performance.

Following Marquering and Verbeek (2004), we investigated the economic benefits of using the LogHARQ model as a volatility forecast model within a volatility timing–based portfolio allocation strategy. We used the CSI 300 and the SSE 50ETF as the risk assets and a one-year fixed deposit as the risk-free asset. Using the HAR, HARQ, and LogHAR models as benchmarks, we showed that an investor is willing to pay a fee to use the LogHARQ model as a forecast model that further supports the usage of LogHARQ model in forecasting Chinese stock market volatility. Several robustness checks indicated that our results are robust to alternative realized measures and model specifications, and that they can be applied to a large set of Chinese stock indices[4].

The remainder of this paper is organized as follows. Section 2 presents the notations and derives the LogHARQ model. Section 3 describes our dataset and reports on the volatility forecasting accuracy of the LogHARQ model as well as of the other benchmark models. Section 4 discusses the economic value of using the LogHARQ model through volatility timing. Section 5 presents the results of the robustness checks. The conclusion is presented in Section 6.

2. Model specification

Assuming that the volatility of the return process follows a continuous process with an instantaneous variance σ_t^2, the daily volatility is therefore defined as the integral of the instantaneous variance over a given period, namely:

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

Andersen and Bollerslev (1998) showed that the conditional expectation of IV_t equals the conditional variance of the daily return across the informational set at time t-1 and proposed the realized variance as the total sum of the squared high-frequency returns to estimate IV_t:

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

where $r_{t,i} = \ln(P_{t+(i+1)\Delta}/P_{t+i\Delta})$ defines the Δ -period intraday return, and Δ is defined as 1 divided by the number of observations (M) for each trading day. Due to data limitations[5], RV_t approximates IV_t with a measurement error, and such error will affect parameters of regression-based volatility models. Bollerslev et al. (2016) proposed a corrected version of linear HAR using the structure of the error term.

In practice, both IV and its approximation, RV, are highly right-skewed, especially in the case of emerging markets where volatility is generally higher and can undergo significant changes. The occasionally extreme values distort the parameters of linear volatility models. As mentioned above, modeling volatility in logarithm form is a popular choice. However, as taking the logarithm is a nonlinear transformation, the asymptotic distribution (when M tends to infinity) of $\ln RV_t$ is distorted too, as shown in Barndorff-Nielsen and Shephard (2006):

$$\ln RV_t = \ln IV_t + \eta_t \quad \eta_t \sim MN\left(0, cIQ_t IV_t^{-2}\right) \tag{3}$$

where $c = (\mu_1^{-4} + 2\mu_1^{-2} - 5)/M$ and $\mu_1 = \sqrt{2/M}$. IQ_t is the integrated quarticity, defined as:

$$IQ_t = \int_{t-1}^{t} \sigma_s^4 ds \tag{4}$$

[1] The sum of intraday squared returns.

[2] In certain applications where linear models are essential, some studies trim extreme realized volatility (e.g., Majewski et al., 2015; Huang et al., 2019) before estimation. Such trimming is generally not needed (thus their information can be used in estimation) for log-linear models as logging automatically reduces the magnitude of extreme realized variance.

[3] Evidence from our empirical investigation indicates that the log-linear HAR model works better than the linear HAR model, especially with rolling window settings. In some cases, both with Chinese data and US data (the exact dataset used by Bollerslev et al. (2016); see appendix), the log-linear HAR model (even without measurement error correction) outperforms the linear HARQ model in Bollerslev et al. (2016), indicating that taking the logarithm is not a trivial consideration in regression-based models.

[4] In addition, we tested our new specification with the 27 Dow Jones Constituents, as described by Bollerslev et al. (2016) in the appendix, and found supportive results of LogHARQ over HAR, HARQ, and LogHAR.

[5] As Δ cannot be arbitrarily small.

As the variance of measurement error changes form as compared with the linear case[6], the correction term used by Bollerslev et al. (2016) is no longer applicable here.

To obtain the correction term for the log-linear model, we started by assuming that the logarithm of IV follows an AR(1) process motivated by the stylized volatility clustering feature, documented in the literature for illustration[7]:

$$\ln IV_t = \theta_0 + \theta_1 \ln IV_{t-1} + \varepsilon_t \tag{5}$$

Here, we refer to θ_1 as the "actual" parameter that measures the contribution of lagged volatility, assuming independence between ε_t and η_t. As lnIV cannot be directly observed, the empirical application of equation (5) often relies on regression based on lnRV instead of lnIV:

$$\ln RV_t = \beta_0 + \beta_1 \ln RV_{t-1} + e_t \tag{6}$$

where e_t is simply the residue of the regression and is assumed to be independent of η_t, as in Bollerslev et al. (2016). Therefore, the OLS (Ordinary Least Square) estimator of β_1 remains consistent. In the presence of measurement error η_t, we can rewrite equation (6) by inserting equation (3):

$$\ln IV_t = \beta_0 + \beta_1(\ln IV_{t-1} + \eta_{t-1}) + (e_t - \eta_t) \tag{7}$$

As η_{t-1} cannot be observed, the "estimated" parameter β_1 will not be equal to the "actual" parameter θ_1. Instead, from equation (7), we have:

$$\widehat{\beta}_1 = \frac{\sum_t (\ln \widetilde{IV}_{t-1} + \eta_{t-1}) \ln \widetilde{IV}_t}{\sum_t (\ln \widetilde{IV}_{t-1} + \eta_{t-1})^2} = \frac{\left(\sum_t \ln \widetilde{IV}_{t-1} \ln \widetilde{IV}_t\right)/T + \left(\sum_t \eta_{t-1} \ln \widetilde{IV}_t\right)/T}{\left(\sum_t (\ln \widetilde{IV}_{t-1})^2\right)/T + \left(\sum_t \eta_{t-1}^2\right)/T + 2\left(\sum_t \eta_{t-1} \ln \widetilde{IV}_{t-1}\right)/T} \rightarrow \frac{Cov(\ln IV_t, \ln IV_{t-1})}{Var(\ln IV) + Var(\eta)} = \beta_1 \tag{8}$$

where $\ln \widetilde{IV}_t$ is the demeaned version of $\ln IV_t$ (i.e. $\ln IV_t - \overline{\ln IV_t}$) and the limit is taken when $T \rightarrow \infty$. Following Bollerslev et al. (2016), additional assumptions that η_t is identically distributed (i.e., $IQ_t IV_t^{-2}$ is constant) and $Var(\ln IV)$ is constant are imposed. Also from equation (5), we have:

$$\widehat{\theta}_1 = \frac{\sum_t \ln \widetilde{IV}_{t-1} \ln \widetilde{IV}_t}{\sum_t (\ln \widetilde{IV}_{t-1})^2} = \frac{\left(\sum_t \ln \widetilde{IV}_{t-1} \ln \widetilde{IV}_t\right)/T}{\left(\sum_t (\ln \widetilde{IV}_{t-1})^2\right)/T} \rightarrow \frac{Cov(\ln IV_t, \ln IV_{t-1})}{Var(\ln IV)} = \theta_1 \tag{9}$$

Combining (8) and (9), we have

$$\beta_1 = \frac{Var(\ln IV)}{Var(\ln IV) + Var(\eta)} \theta_1 = \frac{Var(\ln IV)}{Var(\ln IV) + cIQIV^{-2}} \theta_1 \tag{10}$$

Equation (10) shows the well-known "bias toward zero" effect induced by measurement error and suggests that if the variance of the measurement error is not a constant, the "estimated" parameter should be a time-varying parameter even if the "actual" parameter is time invariant. To recover θ_1, we inverse equation (10):

$$\theta_1 = \beta_1 \left(1 + cIQ/\left(IV^2 Var(\ln IV)\right)\right) \tag{11}$$

In practice, taking the time-varying variance of the measurement error into account, following Bollerslev et al. (2016), we relax the constant assumption of $IQIV^{-2}$ and obtain

$$\theta_1 = \beta_1 + \beta_{1Q}\left(IQ_{t-1}/IV_{t-1}^2\right) \tag{12}$$

where $\beta_{1Q} = \beta_1 c / Var(\ln IV)$. The $t-1$ is selected because θ_1 is the coefficient of $\ln RV_{t-1}$. The integrated quarticity can be estimated by the realized quarticity, which is defined as:

$$RQ_t = \frac{M}{3} \sum_{i=1}^{M} r_{t,i}^4 \tag{13}$$

Together with equation (12), a feasible corrected equation can be written as:

$$\theta_1 = \beta_1 + \beta_{1Q}\left(RQ_{t-1}/RV_{t-1}^2\right) \tag{14}$$

Inserting the corrected equation into the LogAR (1) model (replacing β_1 with the recovered θ_1), we get:

$$\ln RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \frac{RQ_{t-1}}{RV_{t-1}^2}\right) \ln RV_{t-1} + e_t \tag{15}$$

In line with Bollerslev et al. (2016), we define the LogARQ (1) model using the following square root form in order to increase the robustness of the model[8]:

$$\ln RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \frac{RQ_{t-1}^{1/2}}{RV_{t-1}}\right) \ln RV_{t-1} + e_t \tag{16}$$

For the HAR-type model, two corrected versions are proposed. The first version only corrects the lagged daily variance (referred to as LogHARQ), resulting in:

$$\ln RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \frac{RQ_{t-1}^{1/2}}{RV_{t-1}}\right) \ln RV_{t-1} + \beta_2 \ln RV_{t-1|t-5} + \beta_3 \ln RV_{t-1|t-22} + e_t \tag{17}$$

where $\ln RV_{t-j|t-h} = \sum_{i=j}^{h} \ln RV_{t-i}/(h+1-j)$.

The reasoning behind this is that both weekly and monthly averaged variances have smaller measurement errors due to their sample average nature[9]. Therefore, we can leave them uncorrected and thus retain two of our parameters. The second version (referred to as LogHARQ-F) corrects all lagged variances, resulting in:

$$\ln RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \frac{RQ_{t-1}^{1/2}}{RV_{t-1}}\right) \ln RV_{t-1} + \left(\beta_2 + \beta_{2Q} \frac{RQ_{t-1|t-5}^{1/2}}{RV_{t-1|t-5}}\right) \ln RV_{t-1|t-5} + \left(\beta_3 + \beta_{3Q} \frac{RQ_{t-1|t-22}^{1/2}}{RV_{t-1|t-22}}\right) \ln RV_{t-1|t-22} + e_t \tag{18}$$

We use the first version in the current paper. The results from the second version are similar, and are available upon request.

For the linear HARQ model, Bollerslev et al. (2016) proposed the

[6] The asymptotic of RV_t is $RV_t - IV_t \sim MN(0, 2\Delta IQ_t)$, as shown in Barndorff-Nielsen and Shephard (2006).

[7] Volatility clustering implies a strong autocorrelation between future volatility and current or past volatilities. One can view AR(1) as a simplified version of HAR.

[8] Alternative specifications of correction terms such as RQ/RV^2 or $\log(RQ^{1/2}/RV)$ are also tested; we found no significant improvements in the model when these specifications are used. More detailed results are available upon request.

[9] As mentioned below, this setting is in line with Bollerslev et al. (2016).

Table 1
Summary statistics.

	SSE 50ETF				CSI 300			
	2007–2011		2012–2016		2007–2011		2012–2016	
	RV	ln*RV*	*RV*	ln*RV*	*RV*	ln*RV*	*RV*	ln*RV*
Mean	2.76E-04	−8.555	2.09E-04	−9.141	3.34E-04	−8.408	1.96E-04	−9.162
Median	1.84E-04	−8.600	9.46E-05	−9.266	2.11E-04	−8.465	9.31E-05	−9.282
Maximum	4.09E-03	−5.498	4.87E-03	−5.324	4.76E-03	−5.348	4.77E-03	−5.345
Minimum	1.91E-05	−10.864	6.89E-06	−11.885	2.32E-05	−10.673	9.55E-06	−11.559
Std. Dev.	3.11E-04	0.817	4.29E-04	1.024	3.73E-04	0.876	3.80E-04	0.989
Skewness[a]	4.918	0.284	6.502	0.653	3.956	0.254	6.552	0.680
Kurtosis	43.177	3.007	54.849	3.770	30.172	2.656	58.578	3.834
ADF Stat.	−8.765***	−5.491***	−7.540***	−4.347***	−8.472***	−5.598***	−7.540***	−4.347***

Note: The ADF tests use 5 lags. *** denotes significance at the 0.01 level.
[a] Kurtosis is defined as excess kurtosis.

following structure based on the approximation of *IV* and the resulting equation is[10]:

$$RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right) RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t \tag{19}$$

where $RV_{t-j|t-h} = \sum_{i=j}^{h} RV_{t-i}/(h+1-j)$.

It is worth noting that the functional transformation of the integrated variance has a non-trivial effect on the correction process of the model. Similar non-trivial effects can also be found when other realized measures are used to approximate *IV*, because different realized measures may have different asymptotic distributions.

Since variables (including the correction terms) are observable, all proposed models above can be estimated using the ordinary least squares method, in which the current volatility is regressed on the lagged volatilities and their corresponding correction terms. In order for the parameters of the correction terms to be comparable across different models, the correction terms are standardized to zero mean and unit variance. The statistical significance of the parameters is evaluated based on the HAC (Heterogeneous and Autocorrelation Consistent) robust standard errors.

3. Data and volatility forecasting

3.1. Data

We focused our empirical study on the CSI 300 and the SSE 50ETF as they underlie the index futures and options accordingly; we also extended our analysis to a larger set of Chinese stock indices for robustness check. Our sample is obtained from the RESSET database[11], and covers a total of 2407 days with 5-min intervals of intraday prices between January 4, 2007 and December 30, 2016.[12] We also evenly split the sample into two subsamples at the end of 2011 for subsample investigation.

Table 1 reports the descriptive statistics on the daily realized volatilities and logarithmic realized volatility for the CSI 300 and the SSE 50ETF in the two included subsamples. The results of the ADF tests reject the null hypothesis on the presence of a unit root for every series at the 0.01 level.

Fig. 1a and b presents the time variation of the logarithmic *RV* for the CSI 300 and the SSE 50ETF, respectively. During the 2015 Chinese stock market crash, the log *RV* for both assets experienced a significant increase, which reinforces the necessity to investigate the volatility forecasting performance with the two subsamples separately.

3.2. In-sample estimation results

Table 2 presents the full-sample parameter estimates for the LogARQ and LogHARQ models, together with the benchmark LogAR and LogHAR models. We standardized the correction form, $RQ^{1/2}/RV$, to be zero mean and unit variance, which makes it easier to compare the coefficients of β_1 across models. We also report on the adjusted R-squares for cross model comparison.

Table 2 shows that the measurement error plays an important role in forecasting the realized volatility for both assets, as indicated by the significance of β_{1Q}. When the time-varying measurement error is considered, the LogARQ and the LogHARQ models assign a greater weight to the daily lag, in line with the indications of Bollerslev et al. (2016). Consistent with previous studies investigating HAR models (e.g., Corsi, 2009; Corsi et al., 2010), the estimates β_1, β_2, and β_3 are also statistically significant in the LogHARQ model.

3.3. Out-of-sample forecast results

The one-day-ahead forecast series are obtained by estimating the parameters of the models with a fixed length rolling window comprised of the previous 1000 observations. The increasing window method is also used[13], and the results of this analysis are reported for comparison.

To compare these results with the results obtained using the linear HAR model, we assumed that the residuals of the LogHAR and LogHARQ models are normally distributed[14], so that the forecast of the LogHARQ model can be expressed as follows:

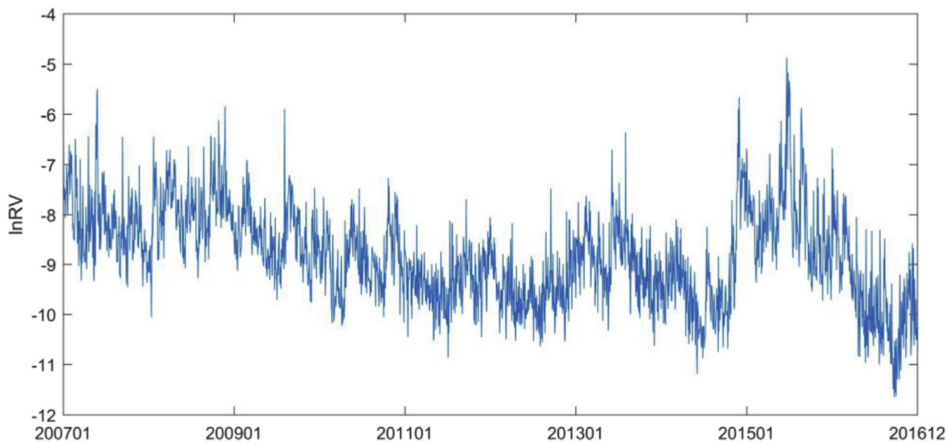
$$F_t = exp\left(\beta_0 + \left(\beta_1 + \beta_{1Q}\frac{RQ_{t-1}^{1/2}}{RV_{t-1}}\right)\ln RV_{t-1} + \beta_2 \ln RV_{t-1|t-5} + \beta_3 \ln RV_{t-1|t-22} + \frac{\sigma_e^2}{2}\right) \tag{20}$$

where σ_t is the standard deviation of error e_t from equation (6).

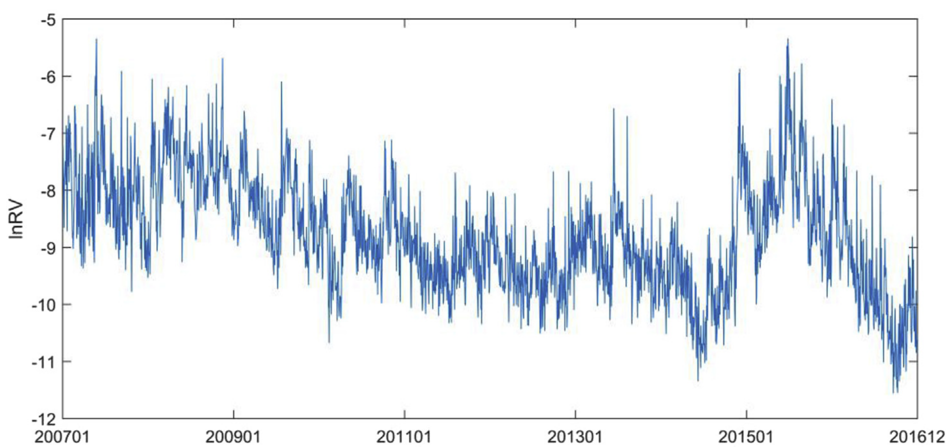
Consistent with previous literature (e.g., Bollerslev et al., 2016), we used a standard MSE (Mean Square Error) measure and the QLIKE (Quasi-likelihood) loss function to evaluate the out-of-sample performance of the different models, such that:

$$MSE = \sum_{t=\tau}^{T}\left(RV_t - F_t\right)^2 \tag{21}$$

[10] A fully corrected version is also discussed in their paper and results are similar to this partial corrected one.
[11] Website: http://www.resset.cn/databases.
[12] The Chinese stock market experienced a market breakdown on January 4 and 7, 2016, which triggered a circuit breaker. Data for these two days are excluded from our sample.
[13] With initial window length set as 1000 observations as well.
[14] See Bekaert and Hoerova (2014) etc. for similar assumption.



(a) SSE 50ETF



(b) CSI 300

Fig. 1. The logarithmic realized volatility of the stock indices.

Table 2
Estimation results for the full sample.

	SSE 50ETF				CSI 300			
	LogAR	LogHAR	LogARQ	LogHARQ	LogAR	LogHAR	LogARQ	LogHARQ
β_0	-1.957^{***}	-0.496^{***}	-1.814^{***}	-0.579^{***}	-1.703^{***}	-0.425^{***}	-1.717^{***}	-0.549^{***}
	(0.123)	(0.132)	(0.118)	(0.132)	(0.115)	(0.116)	(0.111)	(0.118)
β_1	0.779^{***}	0.318^{***}	0.795^{***}	0.377^{***}	0.807^{***}	0.353^{***}	0.805^{***}	0.394^{***}
	(0.014)	(0.027)	(0.013)	(0.028)	(0.013)	(0.027)	(0.012)	(0.028)
β_2		0.405^{***}		0.355^{***}		0.381^{***}		0.345^{***}
		(0.043)		(0.043)		(0.042)		(0.042)
β_3		0.221^{***}		0.203^{***}		0.219^{***}		0.199^{***}
		(0.036)		(0.036)		(0.035)		(0.035)
β_{1Q}			0.018^{***}	0.009^{***}			0.016^{***}	0.009^{***}
			(0.001)	(0.001)			(0.001)	(0.001)
Adj.	0.452	0.454	0.468	0.464	0.456	0.473	0.473	0.480

Note: *** denotes significance at the 0.01 level. Robust standard errors are reported in parentheses.

$$QLIKE = \sum_{t=\tau}^{T}\left[\frac{RV_t}{F_t} - \ln\left(\frac{RV_t}{F_t}\right) - 1\right] \tag{22}$$

where F_t refers to the one-step-ahead forecasts, and RV_t denotes the true realized volatilities.

The MSE and QLIKE loss ratios are reported in Table 3. All loss ratios are computed relative to the benchmark HAR model and therefore comparable across different models.

Table 3 provides evidence that the LogHARQ model performs better

Table 3
Out-of-sample forecast losses of different models for the full sample.

		AR	HAR	ARQ	HARQ	LogAR	LogHAR	LogARQ	LogHARQ
Panel A		SSE 50ETF							
MSE	RW	1.050	1.000	0.952	1.035	1.097	0.969	1.038	0.956
		(0.295)	(0.295)	(1.000)	(0.295)	(0.295)	(0.295)	(0.295)	(0.929)
	IW	1.077	1.000	0.997	**0.966**	1.078	1.048	1.045	1.033
		(0.258)	(0.658)	(0.658)	(1.000)	(0.258)	(0.370)	(0.658)	(0.658)
QLIKE	RW	1.349	1.000	1.160	1.006	1.215	0.990	1.129	**0.980**
		(0.000)	(0.576)	(0.000)	(0.576)	(0.000)	(0.576)	(0.000)	(1.000)
	IW	1.406	1.000	1.175	0.963	1.100	0.956	1.042	**0.947**
		(0.000)	(0.165)	(0.000)	(0.501)	(0.000)	(0.501)	(0.002)	(1.000)
Panel B		CSI 300							
MSE	RW	1.107	1.000	1.148	1.101	1.174	0.990	1.117	**0.979**
		(0.064)	(0.392)	(0.392)	(0.392)	(0.064)	(0.392)	(0.064)	(1.000)
	IW	1.150	**1.000**	1.093	1.002	1.093	1.018	1.065	1.014
		(0.026)	(1.000)	(0.141)	(0.950)	(0.041)	(0.813)	(0.274)	(0.881)
QLIKE	RW	1.439	1.000	1.321	1.035	1.196	0.978	1.119	**0.959**
		(0.000)	(0.216)	(0.002)	(0.216)	(0.001)	(0.216)	(0.002)	(1.000)
	IW	1.551	1.000	1.150	0.939	1.055	0.927	1.001	**0.911**
		(0.000)	(0.036)	(0.000)	(0.137)	(0.000)	(0.137)	(0.001)	(1.000)

Note: This table reports the ratio of the losses for different models relative to the benchmark HAR model. Both the MSE and the QLIKE loss functions are adopted, and the forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. The value in bold is associated with the lowest ratio in each row. The value in parentheses represents the probability of inclusion within the model confidence set (Hansen et al., 2011) with its R statistics (Hansen et al., 2003). Values greater than 0.1 denote the models which were included at the confidence interval of 90% (i.e., those models are statistically equivalent to the best set of models in terms of corresponding loss functions), and we emphasize those models by underlining their ratios.

Table 4
Out-of-sample forecast losses of different models for the two subsamples.

		SSE 50ETF			CSI 300		
		HARQ	LogHAR	LogHARQ	HARQ	LogHAR	LogHARQ
Panel A		2007–2011					
MSE	RW	1.030	0.817	**0.802**	0.976	0.793	**0.789**
	IW	0.983	0.748	**0.733**	0.971	0.743	**0.735**
QLIKE	RW	0.984	0.899	**0.870**	0.921	**0.853**	0.861
	IW	0.960	0.855	**0.827**	0.902	0.809	**0.807**
Panel B		2012–2016					
MSE	RW	1.106	0.789	**0.774**	1.046	0.882	**0.854**
	IW	1.146	0.818	**0.803**	1.053	0.892	**0.862**
QLIKE	RW	0.961	0.779	**0.754**	0.895	0.823	**0.814**
	IW	1.029	0.822	**0.798**	0.919	0.858	**0.847**

Note: This table reports the ratio of the losses for different models relative to the benchmark HAR model. Panel A shows the loss ratios for the 2007–2011 period, while Panel B reports the loss ratios for the 2012–2016 period. Both MSE and QLIKE loss functions are adopted; the forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. The value in bold is associated with the lowest ratio in each row. We emphasize the models that are included in the model confidence set (Hansen et al., 2011) and their R statistics (Hansen et al., 2003) by underlining their ratios.

than other linear and logarithmic models in most cases. The LogHARQ model improves by at least 4% in its forecasting accuracy relative to the benchmark HAR model, as measured by the QLIKE function. Lee et al. (2014) also modeled the volatility of China's stock exchanges, and found that the long-memory feature of the Chinese market volatility suggests a possibility of constructing nonlinear models in order to improve forecasting performance.

We used the model confidence set (MCS) proposed by Hansen et al. (2011) to evaluate whether the performance difference is statistically significant[15]. This method assigns a *p*-value[16] to each model, which denotes the probability of the given model to represent one of the best set of models in terms of forecasting performance under the given loss functions.

For both indexes, the HARQ, LogHAR, and LogHARQ models are all included in the confidence set regardless of the value of their loss functions and window lengths. The AR, LogAR, ARQ, LogARQ, and the standard HAR models underperform when it comes to out-of-sample forecasting, since they are all excluded from the model's confidence set for a given loss function. Although the HARQ, LogHAR, and LogHARQ are all included, the *p*-values for the LogHARQ are higher in most cases, indicating the superior forecasting power of our proposed model over the existing linear and logarithmic forecasting models.

We compared the out-of-sample forecast performance of the different models for the two subsamples and reported the results in Table 4. Panel A shows the loss ratios for the 2007–2011 period, while Panel B reports the loss ratios for the 2012–2016 period[17].

For each asset, we can observe that the loss ratios for the HARQ, LogHAR, and LogHARQ models decrease, evidence which supports the superior forecasting performance of the LogHARQ model. The results are consistent across Panels A and B, so the higher forecast accuracy holds for both sample periods investigated. Improvements in the forecast accuracy of the LogHARQ model relative to the HAR model range from 13% to

[15] The R statistics (Hansen et al., 2003) is used here; the alternative SQ statistic provides similar results.

[16] The increase of p-value indicates that the prediction accuracy of the model increases.

[17] To save space, we only focus on the three models included in the MCS from Table 3 and only report the ratio of the loss function when compared with the HAR model; we only emphasize the models in the MCS for the 90% confidence interval.

Table 5
Stratified RQ out-of-sample forecast losses.

		SSE 50ETF			CSI 300		
		HARQ	LogHAR	LogHARQ	HARQ	LogHAR	LogHARQ
Panel A		Bottom 95% RQ (2007–2011)					
MSE	RW	0.968	0.834	**0.820**	0.902	**0.814**	0.820
	IW	0.956	0.790	**0.777**	0.887	**0.779**	0.780
QLIKE	RW	0.969	0.907	**0.888**	0.896	**0.863**	0.884
	IW	0.953	0.871	**0.853**	0.875	**0.822**	0.831
Panel B		Top 5% RQ (2007–2011)					
MSE	RW	1.552	0.671	**0.647**	1.468	0.655	**0.585**
	IW	1.149	0.492	**0.460**	1.447	0.543	**0.487**
QLIKE	RW	1.214	0.767	**0.593**	1.208	0.739	**0.596**
	IW	1.058	0.653	**0.490**	1.176	0.668	**0.555**
Panel C		Bottom 95% RQ (2012–2016)					
MSE	RW	0.945	0.815	**0.795**	0.965	0.893	**0.881**
	IW	0.986	0.851	**0.830**	0.977	0.908	**0.896**
QLIKE	RW	0.945	0.787	**0.763**	0.880	0.823	**0.821**
	IW	1.015	0.833	**0.809**	0.905	0.860	**0.857**
Panel D		Top 5% RQ (2012–2016)					
MSE	RW	1.793	**0.679**	0.686	1.141	0.871	**0.822**
	IW	1.807	**0.684**	0.693	1.141	0.872	**0.824**
QLIKE	RW	1.311	0.606	**0.569**	1.120	0.815	**0.702**
	IW	1.323	0.606	**0.574**	1.121	0.819	**0.707**

26%. In terms of the MCS, LogHARQ is the only model which is included for both series with regard to all subsamples[18].

Table 5 reports the stratified loss ratios based on their realized quarticity. By further splitting the sample from Table 4, we reported the results for forecasts on days when the previous day's RQ was very high (i.e., Top 5% RQ) and for forecasts concerning the rest of the sample (i.e., Bottom 95% RQ). We found that the LogHARQ model achieves an even better predictive performance when the RQ is high. This verifies our hypothesis that the LogHARQ model has a stronger predictive power for realized volatility when the market is highly volatile.

To sum up, by explicitly accounting for the heteroskedasticity in the measurement errors and for the high-volatility features of Chinese stock prices, the LogHARQ model performs better than existing logarithmic and linear forecasting models, particularly when the RQ is large. The LogHARQ model increases the accuracy of the forecasting of realized volatility within the Chinese stock market.

Note: This table reports the ratio of the losses for different models relative to the benchmark HAR model. Panels B and D show the ratios for the days following a day with an RQ value in the top 5%, in the sample covering the 2007–2011 period and the 2012–2016 period, respectively. Panels A and C present the results for the remaining 95% of the days in the two subsamples. Both the MSE and the QLIKE loss functions are adopted; the forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. The value in bold is associated with the lowest ratio in each row. We emphasize the models which are included in the model confidence set (Hansen et al., 2011) and their R statistics (Hansen et al., 2003) by underlining their ratios.

4. Economic value test

We now focus on the economic value of the model. We define the economic value of the LogHARQ model as the cost that an investor is willing to pay in order to use the LogHARQ model as their preferred volatility forecasting model (i.e., instead of other models).

To evaluate the economic value, we first construct a series of volatility timing–based portfolio allocation strategies. We assume that the investor is risk-averse and thus form a portfolio containing both risk assets (i.e., the CSI 300 and the SSE 50 ETF) and risk-free assets (i.e., a one-year fixed deposit). The daily returns of these assets are used as the basis for portfolio allocation. The economic intuition of this strategy is quite simple. Given the expected return, the investor places more weight on a risk asset when its volatility is low, while preferring the risk-free asset when the risk asset's volatility is high. The investor maximizes their utility as follows:

$$\underset{w_t}{Max} U\left[E_t\left(r_{p,t+1}\right), Var_t\left(r_{p,t+1}\right)\right] \tag{23}$$

where $E_t(r_{p,t+1})$ represents the conditional expected return of the portfolio, $Var_t(r_{p,t+1})$ denotes the conditional variance, and ω_t represents the optimal weight of the risk asset. The expected return is calculated as follows:

$$E_t\left(r_{p,t+1}\right) = r_{f,t+1} + \omega_t\left(E_t\left(r_{m,t+1}\right) - r_{f,t+1}\right) \tag{24}$$

where $E_t(r_{m,t+1})$ represents the conditional expected return of the risk asset and $r_{f,t+1}$ represents the risk-free return[19].

Although we could have utilized more sophisticated utility functions, we chose to use the simple mean-variance preferences because our primary interest is whether the improvements in the accuracy of volatility forecasting, as produced by the LogHARQ model, could gain an additional economic value. The mean-variance utility function is as follows:

$$U\left[E_t\left(r_{p,t+1}\right), Var_t\left(r_{p,t+1}\right)\right] = E_t\left(r_{p,t+1}\right) - \frac{\gamma}{2} Var_t\left(r_{p,t+1}\right) \tag{25}$$

Thus, the equation for the optimal weight of the risk asset is:

$$\omega_t = \frac{E_t\left(r_{m,t+1}\right) - r_{f,t+1}}{\gamma Var_t\left(r_{m,t+1}\right)} \tag{26}$$

where γ represents the investor's risk aversion coefficient.

We computed the conditional variance of the portfolio as follows:

$$Var_t(r_{m,t+1}) = BCF \times \widetilde{RV_{t+1}} \tag{27}$$

[18] For most cases where both the LogHARQ and the LogHAR model are included, the p-values (omitted to save space, available upon request) are lower for the LogHAR than for the LogHARQ model.

[19] We have not imposed the assumption of zero expected returns in this analysis. Following the indications of Fleming et al. (2003) and Nolte and Xu (2015), we used the average daily return across the sample period as a proxy for the expected return.

Table 6
Daily performance fee.

		2007–2011			2012–2016		
		LogHARQ-HAR	LogHARQ-HARQ	LogHARQ-LogHAR	LogHARQ-HAR	LogHARQ-HARQ	LogHARQ-LogHAR
Panel A		SSE 50ETF					
$\gamma = 1$	RW	0.00296	0.00146	0.00020	0.00111	0.00108	0.00047
	IW	0.00330	0.00186	0.00044	0.00107	0.00106	0.00045
$\gamma = 2$	RW	0.00148	0.00073	0.00010	0.00055	0.00054	0.00023
	IW	0.00165	0.00093	0.00022	0.00054	0.00053	0.00022
$\gamma = 3$	RW	0.00099	0.00049	0.00007	0.00037	0.00036	0.00016
	IW	0.00110	0.00062	0.00015	0.00036	0.00035	0.00015
Panel B		CSI 300					
$\gamma = 1$	RW	0.00169	0.00105	0.00002	0.00161	0.00056	0.00027
	IW	0.00182	0.00129	0.00005	0.00153	0.00049	0.00025
$\gamma = 2$	RW	0.00084	0.00052	0.00001	0.00081	0.00028	0.00013
	IW	0.00091	0.00065	0.00002	0.00077	0.00025	0.00013
$\gamma = 3$	RW	0.00056	0.00035	0.00001	0.00054	0.00019	0.00009
	IW	0.00061	0.00043	0.00002	0.00051	0.00016	0.00008

Note: This table reports the performance fee that an investor is willing to pay in order to use LogHARQ as a forecasting model instead of other models (e.g., the HAR model). We used the HAR, HARQ, and LogHAR models as benchmarks and report the results for all three models. The SSE 50ETF and the CSI 300 are each used as the risk assets. The forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. The results of the two subsamples are presented, where γ represents the risk aversion coefficient.

Table 7
Alternative *RV* measures.

		LogHARQ/HARQ						LogHARQ/LogHARQ(*RV*5)					
		SSE 50ETF			CSI 300			SSE 50ETF			CSI 300		
		*RV*5	*RV*10	*RV*15	*RV*5	*RV*10	*RV*15	*RV*5	*RV*10	*RV*15	*RV*5	*RV*10	*RV*15
MSE	RW	0.694	0.689	0.654	0.796	0.850	0.831	1.000	1.016	1.041	1.000	1.052	1.054
	IW	0.694	0.689	0.688	0.799	0.854	0.840	1.000	1.015	1.040	1.000	1.052	1.054
QLIKE	RW	0.837	0.816	0.674	0.975	0.932	0.860	1.000	1.034	1.028	1.000	0.996	0.957
	IW	0.831	0.808	0.711	0.981	0.951	0.893	1.000	1.035	1.030	1.000	0.995	0.958

Note: This table reports the loss ratios for the LogHARQ model relative to the HARQ model when using 5-, 10-, and 15-min *RVs* (left panel), and the ratios of the losses for the LogHARQ model when using different *RV* measures as opposed to the losses when using the 5-min *RV* (right panel). *RV5*, *RV10*, and *RV15* represent the 5-, 10-, and 15-min *RVs*, respectively. Both the MSE and the QLIKE loss functions are adopted; the forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation.

$$Var_t\left(r_{P,t+1}\right) = \omega^2 Var_t\left(r_{m,t+1}\right) = \omega^2 BCF \times \widetilde{RV}_{t+1} \tag{28}$$

where $\widetilde{RV}_{t+1}$ is the realized volatility forecast obtained from the predictive models. The BCF (Nolte and Xu, 2015) is used to match the realized volatility of the 6.5 h high-frequency trading to the daily variance, namely:

$$BCF = \frac{1/n\sum_{t=1}^{n} r_t^2}{1/n\sum_{t=1}^{n} RV_t} \tag{29}$$

We estimated the optimal weight of the risk assets ω_t based on the daily return and the realized volatility; then we computed the average realized utility of the portfolio as follows:

$$\overline{U}\left(R_p\right) E_t\left(r_{p,t+1}\right) - \frac{\gamma}{2} Var_t\left(r_{P,t+1}\right) \quad \overline{U}\left(R_p\right) = \frac{1}{T}\sum_{t=0}^{T-1}\left[r_{p,t+1} - \frac{\gamma}{2} Var_t\left(r_{p,t+1}\right)\right] \tag{30}$$

which measures the average realized utility of the portfolio using the LogHARQ model to forecast *RV*. To evaluate the economic value of LogHARQ as a forecasting model, we computed a performance fee which represents the cost that an investor is willing to pay in order to use LogHARQ as a forecasting model instead of other models (e.g., the HAR model). We computed the performance fee (denoted by Δ) by solving the following equation:

$$\frac{1}{T}\sum_{t=0}^{T-1}\left[\left(r_{p,t+1} - \Delta\right) - \frac{\gamma}{2} Var_t\left(r_{p,t+1}\right)\right] = \frac{1}{T}\sum_{t=0}^{T-1}\left[r_{bm,t+1} - \frac{\gamma}{2} Var_t\left(r_{bm,t+1}\right)\right] \tag{31}$$

Table 6 reports the performance fee after using different models as the benchmark model. We used the HAR, HARQ, and LogHAR models as benchmarks and report the results for each model in Panels A, B, and C, respectively. We used the CSI 300 and the SSE 50ETF as risk assets, and constructed the portfolio with the risk-free assets individually. Two subsamples were used, one covering the 2007–2011 period and the other covering the 2012–2016 period. Risk aversion coefficients of 1, 2, and 3 are used, and the results are presented in Table 6.

Table 6 shows that the performance fee remains positive for every combination of assets, sample periods, and risk aversion coefficients, indicating that an investor would indeed be willing to pay a fee in order to use the LogHARQ model. As a forecast model, LogHARQ is superior to the HAR, HARQ, and LogHAR models when forecasting Chinese stock market volatility. The results are consistent across both sample periods investigated.

5. Robustness checks

5.1. Alternative RV measures

To investigate whether our results are sensitive to the sampling frequency of prices, we considered 10- and 15-min *RVs* as alternative *RV* measures and compared the out-of-sample performance of the LogHARQ model using these different *RV* measures. We used a 1-min realized kernel (*RK*) as a proxy for the true volatility series. The results are reported in Table 7. The left panel presents the loss ratios for the LogHARQ model relative to the HARQ model using 5-, 10-, and 15-min *RVs*. The right panel presents the ratios of the losses of the LogHARQ model when using different *RV* measures and when comparing them to the losses of

Table 8
Out-of-sample forecast losses for alternative stock indexes.

			AR	HAR	ARQ	HARQ	LogAR	LogHAR	LogARQ	LogHARQ
Panel A										
Avg. ratio Med. ratio	MSE	RW	1.269	1.000	1.317	1.085	1.112	0.938	1.044	**0.927**
			1.249	1.000	1.282	1.101	1.123	0.926	1.025	**0.919**
		IW	1.239	1.000	1.314	1.093	1.105	0.942	1.037	**0.931**
			1.220	1.000	1.280	1.102	1.086	0.932	1.024	**0.922**
	QLIKE	RW	1.545	1.000	1.068	0.995	0.995	**0.907**	0.996	**0.907**
			1.555	1.000	1.074	0.998	0.990	0.894	0.985	**0.890**
		IW	1.498	1.000	1.071	1.048	1.016	0.933	1.016	**0.932**
			1.495	1.000	1.074	1.037	1.018	0.924	1.014	**0.920**
Panel B										
% in MCS (avg. *p*-val)	MSE	RW	0.00%	42.86%	0.00%	42.86%	14.29%	100.00%	28.57%	100.00%
			(0.004)	(0.151)	(0.034)	(0.151)	(0.069)	(0.486)	(0.091)	(0.844)
		IW	0.00%	42.86%	0.00%	42.86%	28.57%	100.00%	28.57%	100.00%
			(0.009)	(0.166)	(0.039)	(0.166)	(0.076)	(0.488)	(0.099)	(0.829)
	QLIKE	RW	0.00%	85.71%	14.29%	42.86%	28.57%	100.00%	28.57%	100.00%
			(0.006)	(0.274)	(0.081)	(0.144)	(0.172)	(0.714)	(0.148)	(0.860)
		IW	0.00%	100.00%	28.57%	42.86%	28.57%	100.00%	28.57%	100.00%
			(0.016)	(0.361)	(0.069)	(0.113)	(0.183)	(0.724)	(0.145)	(0.819)

Note: Panel A reports the ratio of the losses for different forecasting models relative to the losses of the standard HAR model. The average and median loss ratios across seven stock indexes are reported. The lowest ratio is given in bold. Panel B reports the frequency of inclusion in the MCS using the R statistics. The corresponding average *p*-values are reported in parentheses.

the 5-min *RV*.

It is clear that for both series the LogHARQ outperforms the HARQ, and that this result holds for *RVs* sampled at different frequencies. This highlights the importance of conducting a log-linear specification instead of a linear specification, at least with regard to our dataset[20]. Within the LogHARQ model, *RV* with a lower sampling frequency generally provides a better performance (especially under the MSE loss function), which is consistent with the current literature on volatility forecasting for high-frequency data.

5.2. Alternative stock indexes

In some of our previous empirical works, we focused on the CSI 300 and the SSE 50ETF as underlying assets of the Chinese index futures and options. We now utilize a larger set of Chinese stock indexes as a robustness check for our out-of-sample results. The extended set includes seven stock indexes, namely the SSE A Share Index, the SSE B Share Index, the SSE Composite Index, the SZSE Component Index, the CSI 500 Index, the SZSE SME Price Index, and the SZSE CHINEXT Price Index. Due to our reliance on the availability of data, the data used covers the period from January 4, 2012 to December 30, 2016, which includes 1214 trading days.

Following the indications of Bollerslev et al. (2016), we report the average and median loss ratios across the above-mentioned stock indexes in Table 8. For comparison, all of the loss ratios are relative to the losses computed from the HAR model. Regardless of the loss functions and window lengths, the LogHARQ model consistently obtains the lowest average and median loss ratios, which represents strong evidence for its superior out-of-sample performance in forecasting Chinese stock market volatility.

Similar to the results of the two series, the LogHAR and LogHARQ models are both included within the confidence set regardless of their series, loss functions, and window lengths. The AR, LogAR, ARQ, LogARQ, HARQ, and standard HAR models underperform when it comes to out-of-sample forecasting, since they are excluded from the model confidence set under certain conditions, such as certain loss function and particular series (i.e., they are not always 100% included within the MCS). Although the LogHAR and the LogHARQ models are both included, the *p*-values of LogHARQ are higher in most cases, indicating the better forecasting power of our proposed model over the existing linear and logarithmic forecasting models.

6. Conclusions

Based on the methods used by Bollerslev et al. (2016), we explicitly accounted for the heteroskedasticity of measurement errors and for the high volatility of Chinese stock prices, and proposed a new model, the LogHARQ model, as a way to forecast the realized volatility of the Chinese stock market. Our LogHARQ model performed better than other logarithmic and linear models, particularly when the realized quarticity was large. The better performance is also confirmed by the utility based economic value test through volatility timing. This suggesting potential benefits of using LogHARQ for risk management and portfolio allocation decisions.

Acknowledgements

The authors claim financial support from the National Natural Science Foundation of China (71871060, 71671004).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.econmod.2019.07.014.

References

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int. Econ. Rev. 39 (4), 885–905.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001. The distribution of realized stock return volatility. J. Financ. Econ. 61 (1), 43–76.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.
Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. Rev. Econ. Stat. 89 (4), 701–720.
Andersen, T.G., Fusari, N., Todorov, v., 2019. The pricing of tail risk and the equity premium: evidence from international option markets. J. Bus. Econ. Stat. (Forthcoming).
Audrino, F., Hu, Y., 2016. Volatility forecasting: downside risk, jumps and leverage effect. Econometrics 4, 8.
Barndorff-Nielsen, O.E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. J. Financ. Econom. 4 (1), 1–30.
Bekaert, G., Hoerova, M., 2014. The VIX, the variance premium and stock market volatility. J. Econom. 183 (2), 181–192.

[20] We also tested our model with a similar dataset obtained from Bollerslev et al. (2016) and found evidence that favors the log-linear specification to some extent.

Bekierman, J., Manner, H., 2018. Forecasting realized variance measures using time-varying coefficient models. Int. J. Forecast. 34 (2), 276–287.
Bollerslev, T., Kretschmer, U., Pigorsch, C., Tauchen, G., 2009. A discrete-time model for daily S & P500 returns and realized variations: jumps and leverage effects. J. Econom. 150 (2), 151–166.
Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2016. Exploiting the errors: a simple approach for improved volatility forecasting. J. Econom. 192 (1), 1–18.
Chen, H., Han, Q., Li, Y., Wu, K., 2013. Does index futures trading reduce volatility in the Chinese stock market? A panel data evaluation approach. J. Futures Mark. 33 (12), 1167–1190.
Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econom. 7 (2), 174–196.
Corsi, F., Reno, R., 2009. HAR Volatility Modelling with Heterogeneous Leverage and Jumps (Working paper).
Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. J. Econom. 159 (2), 276–288.
Fabozzi, F.J., Tunaru, R., Wu, T., 2004. Modeling volatility for the Chinese equity markets. Ann. Econ. Finance 5 (1), 79–92.
Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. J. Financ. 56, 329–352.
Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using realized volatility. J. Financ. Econ. 67, 473–509.
Gonçalves, S., Meddahi, N., 2011. Box–cox transforms for realized volatility. J. Econom. 160 (1), 129–144.
Hansen, P.R., Huang, Z., 2016. Exponential GARCH modeling with realized measures of volatility. J. Bus. Econ. Stat. 34 (2), 269–287.
Hansen, P.R., Lunde, A., Nason, J.M., 2003. Choosing the best volatility models: the model confidence set approach. Oxf. Bull. Econ. Stat. 65, 839–861.
Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79 (2), 453–497.
Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: a joint model for returns and realized measures of volatility. J. Appl. Econom. 27 (6), 877–906.
Hou, Y., Li, S., 2014. The impact of the CSI 300 stock index futures: positive feedback trading and autocorrelation of stock returns. Int. Rev. Econ. Financ. 33 (3), 319–337.
Huang, Z., Liu, H., Wang, T., 2016. Modeling long memory volatility using realized measures of volatility: a realized HAR GARCH model. Econ. Modell. 52, 812–821.
Huang, Z., Chen, T., Tianyi, W., 2019. VIX term structure and VIX futures pricing with realized volatility. J. Futures Mark. 39, 72–93.
Lee, C.F., Chen, G.M., Rui, O.M., 2014. Stock returns and volatility on China's stock markets. J. Financ. Res. 24, 523–543.
Ma, F., Wei, Y., Huang, D., Chen, Y., 2014. Which is the better forecasting model? A comparison between HAR-*RV* and multifractality volatility. Physica A Statistical Mechanics and its Applications 405, 171–180.
Majewski, A.A., Bormetti, G., Corsi, F., 2015. Smile from the past: a general option pricing framework with multiple volatility and leverage components. J. Econom. 187, 521–531.
Marquering, W., Verbeek, M., 2004. The economic value of predicting stock index returns and volatility. J. Financ. Quant. Anal. 39 (2), 407–429.
Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347–370.
Nolte, I., Xu, Q., 2015. The economic value of volatility timing with realized jumps. J. Empir. Financ. 34, 45–59.
Patton, A.J., Sheppard, K., 2009. Optimal combinations of realised volatility estimators. Int. J. Forecast. 25 (2), 218–238.
Shephard, N., Sheppard, K., 2010. Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. J. Appl. Econom. 25 (2), 197–231.
So, R.W., Tse, Y., 2004. Price discovery in the Hang Seng index markets: index, futures, and the tracker fund. J. Futures Mark. 24 (9), 887–907.
Wang, T., Huang, Z., 2012. The relationship between volatility and trading volume in the Chinese stock market: a volatility decomposition perspective. Ann. Econ. Finance 13 (1), 217–242.
Wei, Y., Wang, Y., Huang, D., 2011. A copula–multifractal volatility hedging model for CSI 300 index futures. Physica A Statistical Mechanics and its Applications 390 (23–24), 4260–4272.
Yang, J., Yang, Z., Zhou, Y., 2011. Intraday price discovery and volatility transmission in stock index and stock index futures markets: evidence from China. J. Futures Mark. 32 (2), 99–121.