北京大学中国经济研究中心
China Center for Economic Research

讨论稿系列
Working Paper Series

E2018019 2018-09-16

Does Measurement Error Matter in Volatility Forecasting? Empirical Evidence from the Chinese Stock Market

Yajing Wang, Zhuo Huang, Fang Liang, Tianyi Wang*

Abstract: Based on methods from Bollerslev et al. (2016), we explicitly account for the heteroskedasticity in the measurement errors and high volatility in Chinese stock prices and propose a new realized volatility forecasting model, LogHARQ, to forecast the realized volatility of Chinese stock index futures and options. 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. In an economic sense, using the LogHARQ model for volatility forecasting leads to significant utility gains for investors.

Keywords: Realized volatility, Measurement errors, Volatility forecasting, Chinese stock market

1. Introduction

The fast growth in the capitalization of the Chinese stock market has led to academic interest. The Chinese stock market is highly volatile due to features of the types of listed firms and of investors. Individual investors, who are more likely to be noise traders, play a major role in driving Chinese stock price movements, while the lack of security supplies makes the market vulnerable to speculation, further worsening the situation. The highly volatile nature of the Chinese stock market demands a suitable econometric specification to model and forecast market volatility.

Studies of volatility in the Chinese stock market use various econometric models. GARCH class models are frequently used in modeling and forecasting the volatility of stock index futures and options (e.g., see Yang et al., 2012; Hou and Li, 2014; So and Tse, 2004; Fabozzi et al., 2004). Chen et al. (2012), who study the effect of index futures trading on spot volatility in the Chinese stock market, adopt a panel data evaluation approach to avoid the potential omitted variable bias. Wei et

---

* Yajing Wang is at the School of Banking and Finance, University of International Business and Economics, Beijing, China. Zhuo Huang is at the National School of Development, Peking University, Beijing, China. Fang Liang is at the National School of Development, Peking University, Beijing, China. Tianyi Wang is at the School of Banking and Finance, University of International Business and Economics, Beijing, China. The authors claim financial support from the National Natural Science Foundation of China (71671004).

* Corresponding author: Tianyi Wang. School of Banking and Finance, University of International Business and Economics, Beijing 100029, P.R. China, Tel:86-10-64492513, Email: tianyiwang@uibe.edu.cn.

al. (2011) compare the hedging effectiveness of the copula-MFV model with copula-GARCH models using the prices of Chinese stock index spots and futures. They show that the copula-MFV model has a better hedging effectiveness than the copula-GARCH models. Resent literature focused on the importance of realized measures and variety of models are proposed such as the Realized GARCH model (Hansen et al., 2012, Hansen and Huang, 2016), the HEAVY model (Shephard and Sheppard, 2010) etc.

The HAR model, proposed by Corsi (2009), captures the long-memory characteristics of financial data and is parsimonious and easy to estimate. The HAR model has been widely used for volatility forecasting in the Chinese stock market (e.g., see Wang and Huang, 2012; Ma et al., 2014, Huang et al., 2016).

However, Bollerslev et al. (2016) argue that the HAR model ignores time variability in the magnitude of the realized volatility measurement errors and so suffers from an errors-in-variables problem. The errors are proved to attenuate the parameters of the model. As a solution, they propose the linear HARQ model, which allows the parameters of the model to change with the magnitude of measurement error, and show that the linear HARQ model outperforms the forecasts from the HAR model. Similar to the HARQ model, the HARS model put forth by Bekierman and Manner (2018) captures the effect of measurement errors by including a time-varying state variable.

Following Bollerslev et al. (2016), we derive a logarithmic version of the linear HARQ model using infill large sample theory and the asymptotic distribution of realized volatility. Inspired by the spirit of the EGARCH model (Nelson, 1991) and empirical evidence showing the superior volatility forecasting performance of the EGARCH model compared with the GARCH model in volatile times, we assume log-linear measurement errors to derive the LogHARQ model. Compared with the linear HARQ model, our model is better at forecasting realized volatility when the realized quarticity is large. In particular, the LogHARQ model is suitable for forecasting the volatility of Chinese stock index futures and options because the prices of these assets are highly volatile and the measurement errors are relatively large. We use the China Securities Index (hereafter CSI) 300 stock index and SSE 50ETF to represent Chinese stock index futures and options, respectively, and predict their volatility. Our empirical findings suggest that the LogHARQ model significantly improves on 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 evaluate the economic benefit of using the LogHARQ model as a volatility forecasting model for investment decisions. Fleming et al. (2001) develop a framework of assessing the economic value of volatility timing strategies. They consider a risk-averse investor who has mean-variance preferences and allocates her wealth across different assets. Based on Fleming et al. (2001), Fleming et al. (2003) use the realized volatility to form estimates of the conditional covariance matrix of asset returns and find that volatility timing performance can be improved by using high frequency data. Marquering and Verbeek (2004) propose a framework for evaluating the economic value of volatility timing strategies when allocating between two assets, one risky and the other risk-free. Building on Marquering and Verbeek (2004), Nolte and Xu (2015) include realized jumps in the information set and find that a risk-averse investor can significantly improve her portfolio performance by incorporating realized jumps into her volatility timing strategy.

Following Marquering and Verbeek (2004), we investigate the economic benefits of using the LogHARQ model as the volatility forecast model in a volatility timing based portfolio allocation strategy. We use the CSI 300 stock index and SSE 50ETF as risky assets and a one-year fixed deposit as the risk-free asset. Using the HAR, HARQ, and LogHAR models as benchmarks, we show that an investor would be willing to pay a fee to use the LogHARQ model as a forecast model, which indicates that the LogHARQ model is better than the HAR, HARQ, and LogHAR models when forecasting Chinese stock market volatility.

The remainder of the paper is organized as follows. Section 2 provides the notation and derives the LogHARQ model. Section 3 describes our dataset and reports the volatility forecasting accuracy of LogHARQ model and other benchmark models. Section 4 discusses the economic value of using the LogHARQ model as a volatility forecast model. Section 5 presents the results of robustness checks and the conclusion comes in Section 6.

2.Models

The stylized fact that volatility clustering implies autocorrelation between future volatility and current or past volatilities. One simple way to model such a volatility process is a Log-AR (1) model of integrated variance:

$$\ln IV_t = \theta_0 + \theta_1 \ln IV_{t-1} + \epsilon_t$$

IV is defined as the integral of the instantaneous variance of the return process over a given period:

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

Andersen and Bollerslev (1998) show that

$$E\left(IV_{\mathrm{t}} \mid F_{t-1}\right) = \operatorname{var}\left(r_{\mathrm{t}} \mid F_{t-1}\right)$$

where F_{t-1} denotes the information set at time (t-1).

Because IV is positive and highly right-skewed, we take its logarithm to reduce the possibility of parameter distortion from extreme IV values. Directly measuring IV with data is not possible because we do not know the exact process of the instantaneous variance. Andersen and Bollerslev (1998) propose an approximate measure which uses a summation of high-frequency returns, named the "realized variance":

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

$\mathrm{r}_{t,i} \equiv \ln P_{t+(i+1)\Delta} - \ln P_{t+i\Delta}$ defines the Δ-period intraday return, where Δ is defined as one over the number of samples for each trading day (M).

Bandorff-Nielsen and Shephard (2006) suggest that when M tends to infinity, $\ln RV_t$ approaches $\ln IV_t$ plus a mean zero distribution:

$$\ln RV_t = \ln IV_t + \eta_t \qquad \eta_t : \; MN(0, \mathrm{c}\, IQ_t IV_t^{-2})$$

$$E(\eta_t) = 0 \qquad Var(\eta_t) = \mathrm{c}\, IQ_t IV_t^{-2}$$

where $c = (\mu_1^{-4} + 2\mu_1^{-2} - 5)/M$ and $\mu_1 = \sqrt{2/M}$. This implies that the autoregression estimated with $\ln RV_t$,

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

suffers from a measurement error problem and the corresponding parameter is biased toward zero. Assuming independence between (η_t, ϵ_t), the bias can be quantified as

$$\hat{\beta}_1 \rightarrow \frac{Var(\ln IV_t)}{Var(\ln IV_t) + Var(\eta_t)} \theta_1 = \frac{Var(\ln IV_t)}{Var(\ln IV_t) + cIQ_t IV_t^{-2}} \theta_1$$

where we have used the approximation given above. IQ is called the integrated quarticity:

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

A feasible estimate of IQ, as suggested by Bollerslev et al. (2016), is the "realized quarticity":

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

Taking the inverse of the equation gives the relationship between the correct parameter (θ_1) and the OLS estimated parameter ($\hat{\beta}_1$):

$$\theta_1 = \hat{\beta}_1 \left(1 + cIQ_t / (IV_t^2 Var(\ln IV_t))\right)$$

If we assume $Var(\ln IV_t)$ is time invariant and collect terms with respect to their time variability, we have

$$\theta_1 = \hat{\beta}_1 + \hat{\beta}_{1Q} \left(IQ_t / IV_t^2\right)$$

Using the estimated versions of IQ and IV leads to a feasible correction equation:

$$\theta_1 = \hat{\beta}_1 + \hat{\beta}_{1Q} \left(RQ_t / RV_t^2\right)$$

Inserting the correction equation into the Log-AR (1) model, we have the corrected version and name it Log-ARQ (1):

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

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

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

The idea behind this is that weekly and monthly averaged variances have smaller measurement error due to their sample average nature. Therefore, we can leave them uncorrected and save two parameters. The second version corrects all lagged variances (referred to as LogHARQ-F):

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

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

For the linear HARQ model, Bollerslev et al. (2016) propose the following structure based on the approximation of IV. They consider the effect of the measurement error and then adjust the coefficients of RV_{t-1} to get the equation

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

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.

3. Modeling and Forecasting Volatility

3.1 Data

We focus our empirical studies on the CSI 300 stock index and SSE 50ETF, which play leading roles in stock index futures and options. Intraday prices at 5-minute intervals are obtained from the RESSET database. Our sample starts on January 4, 2007 and ends on December 30, 2016 for a total of 2407 observation days.[1] We split the sample into two subsamples, one covering 2007-2011 and the other covering 2012-2016.

Table 1 reports descriptive statistics of the daily realized volatilities and logarithmic realized volatilities for the CSI 300 stock index and SSE 50ETF in the two subsamples. Results of ADF tests reject the null hypothesis of a unit root for every single series at the 1% level.

Table 1

Summary Statistics

	SSE 50ETF				CSI 300			
	2007-2011		2012-2016		2007-2011		2012-2016	
	RV	lnRV	RV	lnRV	RV	lnRV	RV	lnRV
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

[1] 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.

Std. Dev.	3.11E-04	0.817	4.29E-04	1.024	3.73E-04	0.876	3.80E-04	0.989
Skewness	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 1% level.

Figures 1 and 2 present the time variation of log RV for the CSI 300 stock index and SSE 50ETF, respectively. During the 2015 Chinese stock market crash, log RV for both assets experienced a significant increase, which motivates investigating volatility forecasting performance in the two subsamples separately.

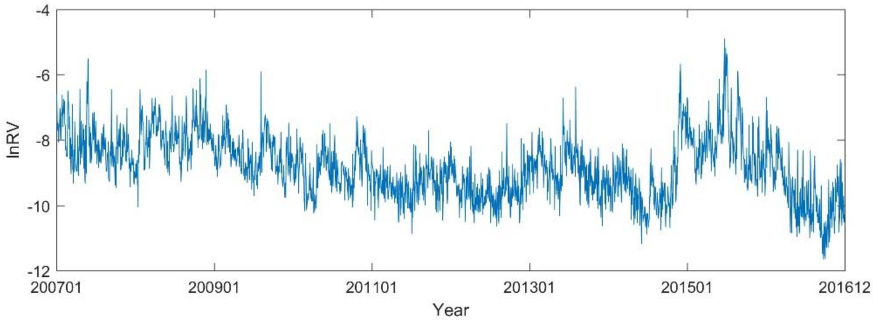

Fig. 1. Time variation of logarithmic realized volatility of SSE 50ETF.

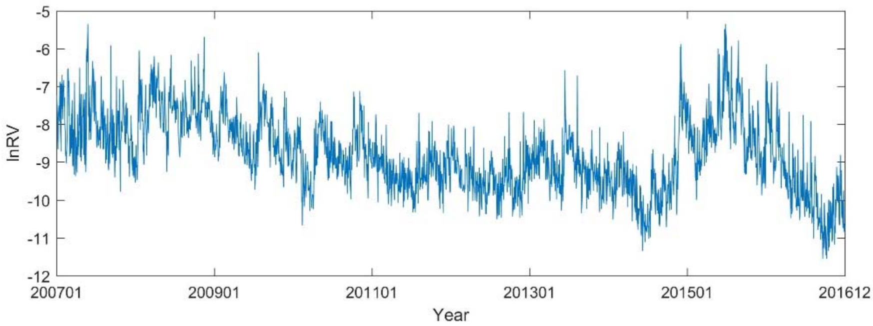

Fig. 2. Time variation of logarithmic realized volatility of the CSI 300 stock index.

3.2 In-sample estimation results

Table 2 presents the parameter estimates for the full sample for the LogARQ and LogHARQ models, together with the benchmark LogAR and LogHAR models. We report the adjusted R-squares for comparison between different models.

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

Table 2

Estimation results for the full sample

	SSE 50ETF				CSI 300			
	LogAR	LogHAR	LogARQ	LogHARQ	LogAR	LogHAR	LogARQ	LogHARQ
β_0	-2.098***	-0.517***	-1.558***	0.032	-1.703***	-0.425***	-1.577***	-0.135
s.e.	(0.136)	(0.137)	(0.214)	0.212	(0.115)	(0.116)	(0.171)	(0.166)
β_1	0.763***	0.294***	0.829***	0.364***	0.807***	0.353***	0.822***	0.386***
s.e.	(0.015)	(0.032)	(0.025)	(0.034)	(0.013)	(0.027)	(0.020)	(0.031)
β_2		0.417***		0.399***		0.381***		0.374***
s.e.		(0.047)		(0.045)		(0.042)		(0.043)
β_3		0.231***		0.246***		0.219***		0.228***
s.e.		(0.037)		(0.037)		(0.035)		(0.035)
β_{1Q}			2.13E-03	2.11E-03***			4.99E-04	1.08E-03***
s.e.			(7.04E-04)	(5.77E-04)			(4.36E-04)	(4.08E-04)
Adj.R2	0.471	0.479	0.482	0.496	0.456	0.473	0.456	0.479

Note: *** denotes significance at the 1% level. Robust standard errors are reported in parentheses.

3.3 Out-of-sample forecast results

We assess the one-day-ahead forecast performance for the realized volatilities of the CSI 300 stock index and SSE 50ETF. The forecast series are obtained by estimating the parameters of the models with a fixed length rolling window comprised of the previous 1000 observations. The recursive estimation method is also used and the results are reported for comparison.

To compare with the results obtained from the linear HAR model, we assume that the residuals of the LogHAR model and LogHARQ models are normally distributed, so that the forecast of the LogHARQ model can be expressed as

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

Consistent with the literature (e.g., see Bollerslev et al., 2016), we use a standard MSE measure and the QLIKE loss to evaluate the out-of-sample performance of the different models:

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

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

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

The MSE and QLIKE for the full sample are reported in Table 3. Panel A shows the ratios of the losses of the different linear models relative to the losses of the HAR model. Panel B presents the loss ratios for the logarithmic models relative to the LogHAR model.

Table 3

Out-of-sample forecast losses of different models for the full sample

		SSE 50ETF				CSI 300			
Panel A		AR	HAR	ARQ	HARQ	AR	HAR	ARQ	HARQ
MSE	RW	1.050	1.000	0.952	1.035	1.107	1.000	1.148	1.101
	IW	1.077	1.000	0.997	0.966	1.150	1.000	1.093	1.002
QLIKE	RW	1.349	1.000	1.160	1.006	1.439	1.000	1.321	1.035
	IW	1.406	1.000	1.175	0.963	1.551	1.000	1.150	0.939
Panel B		LogAR	LogHAR	LogARQ	LogHAR	LogAR	LogHAR	LogAR	LogHARQ
MSE	RW	1.132	1.000	1.072	0.986	1.186	1.000	1.128	0.989
	IW	1.029	1.000	0.998	0.986	1.074	1.000	1.046	0.995
QLIKE	RW	1.228	1.000	1.141	0.990	1.224	1.000	1.144	0.981
	IW	1.151	1.000	1.089	0.990	1.139	1.000	1.080	0.983

Note: The table reports the ratio of the losses for the linear and log models. Panel A shows the ratio of losses for the linear models relative to the losses of the HAR model. Panel B presents the loss ratios for the logarithmic models relative to the LogHAR model. Both MSE and the QLIKE losses are adopted to evaluate the out-of-sample performance of the different models. The forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. We report the loss ratios for all loss functions and window lengths combinations.

Table 3 provides evidence that the LogHARQ model performs better than all other logarithmic models for all loss functions and window lengths combinations. The LogHARQ model gains nearly 2% in forecast accuracy relative to the benchmark LogHAR model as measured by QLIKE.

We compare the out-of-sample forecast performance of the different models in the two subsamples and report the results in Table 4. Table 4 presents the ratio of the losses for the different models relative to the losses of the HAR model (columns 3-5 and 7-9) and the LogHAR model (columns 6 and 10). Panel A shows the results for the 2007-2011 sample period; Panel B reports the loss ratios for 2012-2016.

Table 4

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

		SSE 50ETF				CSI 300			
		HARQ /HAR	LogHAR /HAR	LogHARQ /HAR	LogHARQ /LogHAR	HARQ /HAR	LogHAR /HAR	LogHARQ /HAR	LogHARQ /LogHAR
Panel A		2007-2011							
MSE	RW	1.030	0.817	0.802	0.981	0.976	0.793	0.789	0.995
	IW	0.983	0.748	0.733	0.979	0.971	0.743	0.735	0.989
QLIKE	RW	0.984	0.899	0.870	0.968	0.921	0.853	0.861	1.009
	IW	0.960	0.855	0.827	0.967	0.902	0.809	0.807	0.998
Panel B		2012-2016							
MSE	RW	1.106	0.789	0.774	0.981	1.046	0.882	0.854	0.967
	IW	1.146	0.818	0.803	0.982	1.053	0.892	0.862	0.967
QLIKE	RW	0.961	0.779	0.754	0.968	0.895	0.823	0.814	0.989
	IW	1.029	0.822	0.798	0.971	0.919	0.858	0.847	0.988

Note: This table shows the ratios of the losses for the different models relative to the losses of the HAR model (columns 3-5 and 7-9) and the LogHAR model (columns 6 and 10). Panel A shows the results for 2007-2011. Panel B reports the loss ratios for 2012-2016. Both MSE and the QLIKE losses are adopted to evaluate the out-of-sample performance of the different models. The forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. We report the loss ratios for all loss functions and window lengths combinations.

Table 4 shows that for each asset, the loss ratios HARQ/HAR, LogHAR/HAR, and LogHARQ/HAR are decreasing, which is a supportive evidence for the superior forecasting performance of LogHARQ model compared to HARQ, LogHAR, and the basis HAR model. The results are consistent across Panels A and B, so the higher forecast accuracy holds for both sample periods. The improvements in the forecast accuracy of the LogHARQ model relative to the HAR model range from 12% to 26%.

Table 5 reports the loss ratios stratified according to RQ. After further splitting the forecasting results in Table 4 into forecasts for days when the previous day's RQ was very high and forecasts for the rest of the sample, we find that the LogHARQ model achieves even better forecast performance when RQ is high. This verifies our hypotheses that the LogHARQ model has more predictive power for realized volatility when volatility is highly volatile. To sum up, by explicitly accounting for the heteroskedasticity in the measurement errors and high volatility in Chinese stock prices, LogHARQ model performs better than existing logarithmic and linear models, particularly when RQ is large. The LogHARQ model improves the forecasting accuracy of realized volatility in the Chinese stock market.

Table 5

RQ Stratified out-of-sample forecast losses

		SSE 50ETF				CSI 300			
		HARQ /HAR	LogHAR /HAR	LogHARQ /HAR	LogHARQ /LogHAR	HARQ /HAR	LogHAR /HAR	LogHARQ /HAR	LogHARQ /LogHAR
Panel A		Bottom 95% RQ （2007-2011）							
MSE	RW	0.968	0.834	0.820	0.983	0.902	0.814	0.820	1.000
	IW	0.956	0.790	0.777	0.984	0.887	0.779	0.780	0.994
QLIKE	RW	0.969	0.907	0.888	0.979	0.896	0.863	0.884	1.015
	IW	0.953	0.871	0.853	0.979	0.875	0.822	0.831	1.003
Panel B		Top 5% RQ (2007-2011)							
MSE	RW	1.552	0.671	0.647	0.964	1.468	0.655	0.585	0.950
	IW	1.149	0.492	0.460	0.935	1.447	0.543	0.487	0.947
QLIKE	RW	1.214	0.767	0.593	0.773	1.208	0.739	0.596	0.907
	IW	1.058	0.653	0.490	0.751	1.176	0.668	0.555	0.909
Panel C		Bottom 95% RQ (2012-2016)							
MSE	RW	0.945	0.815	0.795	0.975	0.965	0.893	0.881	0.982
	IW	0.986	0.851	0.830	0.976	0.977	0.908	0.896	0.981
QLIKE	RW	0.945	0.787	0.763	0.969	0.880	0.823	0.821	0.996
	IW	1.015	0.833	0.810	0.972	0.905	0.860	0.857	0.995
Panel D		Top 5% RQ (2012-2016)							
MSE	RW	1.793	0.679	0.686	1.011	1.141	0.871	0.822	0.950
	IW	1.807	0.684	0.693	1.013	1.141	0.872	0.824	0.951
QLIKE	RW	1.311	0.606	0.569	0.939	1.120	0.815	0.702	0.883
	IW	1.323	0.606	0.574	0.947	1.121	0.819	0.707	0.885

Note: The table reports the loss ratios stratified according to the relative magnitude of the realized quarticity. Panels B and D show the ratios for days following a day with an RQ value in the top 5% in sample covering 2007-2011 and the 2012-2016 sample, respectively. Panels A and C present the results for the remaining 95% of the days in the two subsamples. Both MSE and the QLIKE losses are adopted to evaluate the out-of-sample performance of the different models. The forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. We report the loss ratios for all loss functions and window lengths combinations.

4. Economic Value Test

We now focus on the economic value of the model. We measure the economic value of the LogHARQ model as the cost an investor is willing to pay to use the LogHARQ model as a volatility forecasting model instead of other models.

To evaluate the economic value, we first construct volatility timing based portfolio allocation strategies. We assume the investor is risk averse. Her portfolio contains risky assets and risk-free assets. A one-year fixed deposit is used as the risk-free asset; the risky assets are the CSI 300 stock index and SSE 50ETF. Daily returns for these assets are used as the basis of the portfolio allocation. The economic intuition for the strategy is quite simple. Given the expected return, the investor puts more weight on the risky asset when the volatility of the risky asset is low and she turns to the risk-free asset when the volatility is high. The investor maximizes her utility:

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

where $E_t(r_{p,t+1})$ is the conditional expected return of the portfolio, $Var_t(r_{p,t+1})$ denotes the conditional variance, and ω_t is the optimal weight on the risky asset. The expected return is

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

where $E_t(r_{m,t+1})$ is the conditional expected return of the risky assets and $r_{f,t+1}$ is the risk-free return.

Although we could use more sophisticated utility functions, we choose simple mean-variance preferences because our primary interest is whether the improvements in volatility forecast accuracy provided by the LogHARQ model gain an additional economic value. The mean-variance utility function is

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

Thus, the expression of the optimal weight on the risky asset is

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

where γ is the investor's risk aversion coefficient.

We compute the conditional variance of the portfolio as

$$Var_t(r_{m,t+1}) = BCF \widehat{R}V_{t+1}$$

$$Var_t(r_{P,t+1}) = \omega^2 Var_t(r_{\mathrm{m},t+1}) = \omega^2 BCF \times \widehat{R}V_{t+1}\,.$$

$\widehat{R}V_{t+1}$ is the realized volatility forecast obtained from the predictive models. BCF (Nolte and Xu, 2015) is used to match the realized volatility of 6.5 hours' high frequency trading to daily variance:

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

We estimate the optimal weight on the risky assets ω_t based on the daily return and realized volatility and then compute the average realized utility of the portfolio as

$$\bar{U}(R_p) = \frac{1}{T}\sum_{t=0}^{T-1}[r_{p,t+1} - \frac{\gamma}{2}Var_t(r_{p,t+1})]\,,$$

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

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

Table 6 reports the performance fee using different models as the benchmark model. We use the HAR, HARQ, and LogHAR models as benchmarks and report the results in Panels A, B, and C, respectively. We use the CSI 300 stock index and SSE 50ETF as risky assets and construct the portfolio with risk-free assets individually. Two subsamples are used, one covering 2007-2011 and the other covering 2012-2016. Risk aversion coefficients of 1, 2, and 3 are used and we present the results in Table 6.

Table 6

Daily performance fee

	2007-2011		2012-2016	
	SSE 50ETF	CSI 300	SSE 50ETF	CSI 300
Panel A	LogHARQ-HAR			
γ=1	0.00084	0.00110	0.00189	0.00225
γ=2	0.00042	0.00055	0.00094	0.00112
γ=3	0.00028	0.00037	0.00063	0.00075
Panel B	LogHARQ-HARQ			
γ=1	0.00042	0.00049	0.00146	0.00057
γ=2	0.00021	0.00025	0.00073	0.00029
γ=3	0.00014	0.00016	0.00049	0.00019
Panel C	LogHARQ-LogHAR			
γ=1	0.00002	0.00011	0.0003	0.0005
γ=2	0.00001	0.00005	0.0001	0.0002
γ=3	0.00001	0.00004	0.0001	0.0002

Note: The table reports the performance fee the investor is willing to pay to use LogHARQ as a forecasting model instead of other models (e.g., HAR model). We use the HAR, HARQ, and LogHAR models as benchmarks and report the results in Panels A, B, and C, respectively. SSE 50ETF and the CSI 300 stock index are each used as the risky asset. Results of two subsamples are presented. γ is the risk aversion coefficient.

Table 6 shows that the performance fee is positive for every asset, sample period, and risk aversion coefficient combination, indicating that the investor is willing to pay a fee 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.

5. Robustness Checks

To investigate the sensitivity of our results to the sampling frequency of prices, we consider 10- and 15-minute RVs as alternative RV measures and compare the out-of-sample performance of the LogHARQ model using these different RV measures. We consider a 1-minute realized kernel as a proxy for the true volatility series. The results are reported in Table 7. The left panel shows the loss ratios for the LogHARQ model relative to the HARQ model using 5-, 10-, and 15-minute RVs. The right panel presents the ratios of the losses of the LogHARQ model using different RV measures to the losses of using the 5-minute RV.

The LogHARQ model based on the 5-minute RV outperforms the LogHARQ model based on the alternative RV measures, as indicated by the lowest ratio in the "RV5" column in the right panel. Despite the superior performance of the 5-minute RV, the LogHARQ models based on these alternative RV measures still offer significant forecast improvements relative to the HARQ models based on the same RV measures. The results hold for both assets considered.

Table 7

Alternative RV measures

		LogHARQ/HARQ						LogHARQ/LogHARQ(RV5)					
		SSE 50ETF			CSI 300			SSE 50ETF			CSI 300		
		RV5	RV10	RV15	RV5	RV10	RV15	RV5	RV10	RV15	RV5	RV10	RV15
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: The table reports the loss ratios for the LogHARQ model relative to the HARQ model using 5-, 10-, and 15-minute RVs (in the left panel), and the ratios of the losses for the LogHARQ model using different RV measures to the losses of using the 5-minute RV (in the right panel). "RV5," "RV10," and "RV15" stand for 5-, 10-, and 15-minute RVs, respectively. Both MSE and the QLIKE losses are adopted to evaluate the out-of-sample performance. The forecast series are obtained using both a rolling window (RW) estimation and an increasing window (IW) estimation. We report the loss ratios for all loss functions and window length combinations.

6. Conclusions

We propose a new realized volatility forecasting model. Based on methods in Bollerslev et al. (2016), we explicitly account for the heteroskedasticity in the measurement errors and high volatility in Chinese stock prices to derive the LogHARQ model for forecasting the realized volatility of stock index futures and options. Our LogHARQ model performs better than other logarithmic models and linear models, particularly when RQ is large. Importantly, using the LogHARQ model as volatility forecasting model leads to significant gains for investors compared to other forecasting models. The

LogHARQ model improves the accuracy of forecasting realized volatility, benefitting risk management and portfolio allocation decisions.

References

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review. 39(4), 885-905.

Barndorffnielsen, O.E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics. 4(1), 1-30.

Bekierman, J., Manner, H., 2018. Forecasting realized variance measures using time-varying coefficient models. International Journal of Forecasting. 34(2), 276-287.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2016. Exploiting the errors: A simple approach for improved volatility forecasting. Journal of Econometrics. 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. Journal of Futures Markets. 33(12), 1167-1190.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics. 7(2), 174-196.

Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics. 159(2), 276-288.

Fabozzi, F. J., Tunaru, R., Wu, T., 2010. Modeling volatility for the Chinese equity markets. Annals of Economics & Finance, 5(1), 79-92.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. Journal of Finance. 56, 329-352.

Fleming, J., Kirby, C., Ostdiek, B.,2003. The economic value of volatility timing using realized volatility. Journal of Financial Economics. 67, 473-509.

Hansen, P.R., Huang, Z., 2016. Exponential GARCH modeling with realized measures of volatility. Journal of Business & Economic Statistics. 34(2), 269-287.

Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: A joint model for returns and realized measures of volatility. Journal of Applied Econometrics. 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. International Review of Economics & Finance. 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. Economic Modelling. 52, 812-821.

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 & Its Applications. 405, 171-180.

Marquering, W., Verbeek, M., 2004. The economic value of predicting stock index returns and volatility. Journal of Financial & Quantitative Analysis. 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. Journal of Empirical Finance. 34, 45-59.

Shephard, N., Sheppard, K., 2010. Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics. 25(2), 197-231.

So, R. W., Tse, Y., 2010. Price discovery in the Hang Seng index markets: Index, futures, and the tracker fund. Journal of Futures Markets. 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. Annals of Economics & 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 & 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. Journal of Futures Markets. 32(2), 99-121.