Received: 12 January 2020 | Revised: 16 August 2020 | Accepted: 28 November 2020

DOI: 10.1002/for.2748

RESEARCH ARTICLE

WILEY

Can night trading sessions improve forecasting performance of gold futures' volatility in China?

Xuan Yao[1] | **Xiaofeng Hui**[1] | **Kaican Kang**[2]

[1]School of Management, Harbin Institute of Technology, China

[2]Credit Research Initiative, National University of Singapore, Singapore

Correspondence
Yao Xuan, Harbin Institute of Technology, Harbin 150001, China.
Email: yaoxuan@hit.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 71532004, 71773024

Abstract

We use heterogeneous autoregression (HAR) and two related HAR extension models to examine volatility forecasting performances before and after the launch of night trading sessions in the Shanghai Futures Exchange (SHFE) gold futures market. To capture fluctuations from external information and volatility of realized volatility (RV), we incorporate the trading volume and jumping into the HAR-V-J model in the first place and then incorporate a GARCH specification into the HAR-GARCH model. Results showed that there were large fluctuations in SHFE gold futures market before the launch of night trading sessions and mostly stemmed from overnight fluctuation in the international gold futures market. After the launch of night trading sessions, the realized volatility has a clear trend of moderation. In the in-sample estimation, both jump and external information are found to have significant explanatory power with the HAR-V-J model. Additionally, the volatility clustering and high persistence of the realized volatility were confirmed by the GARCH coefficients. Last but not the least, night trading sessions have significantly improved the out-of-sample forecasting performances of realized volatility models. Among them, the HAR-V-J model is the best-performing model. This conclusion holds for various prediction horizons and has great practical values for investors and policymakers.

KEYWORDS

gold futures markets, high frequency, night trading, realized volatility, volatility forecasting

1 | INTRODUCTION

With significant worldwide altercations across political, economic, and financial domains, the fluctuation frequency and range of commodity futures prices are getting larger, especially for the gold futures, a product that contains the attributes of commodity and financial assets. The function of a commodity futures exchange is to standardize and promote futures trading for as many participants as possible. İncreasing trading frequencies have promoted more and more innovations over the years, driving increased participation through electronic networks. Extended trading hours is one of the most important innovations, which has been implemented across countries. However, empirical findings of the effect of extra trading sessions on volatility and contributions to risk management in commodity futures markets lack consensus. Some argued that the intra-volatility increases when extended trading hours cover the release of key government supply and demand reports (Kao & Fung, 2012; Kauffman, 2013). The researches coverage thus far, the pre-market and after-hours trading sessions, are mostly outside of regular hours with limited liquidity. Proponents argued for greater flexibility when the extended trading hours make information transfer across markets more smoothly, therefore leading to higher liquidity and less volatility (Hui, Yao, & Ma, 2020; Klein & Todorova, 2018). These studies zoom into China's exchanges to examine the effects of introducing the night trading sessions at Shanghai Futures Markets (SHFE).

Prior to July 2013, gold and silver futures in the SHFE could only be traded during the regular daytime, unlike American and European exchanges availing trading hours around the clock. Limited trading hours in the Chinese futures market obstruct the transmission of information. The former operation was not conducive for investors who participate in the Chinese market to manage risks caused by the market fluctuations at night. Until 5 July 2013, the night trading sessions also referred to as the Continuous Trading Program, of gold and silver futures were launched by SHFE. The program starts from 21:00 to 2:30 of the following day, availing an additional 5.5 h on top of the original intermittent 4 h trading time. The night trading sessions have largely covered the active trading hours of major international precious metal futures exchanges, for example, COMEX and LME. Since then, the SHFE has become the second-largest gold futures exchange worldwide in terms of trading volume—next to New York—and experience a growing demand for by international investors. Given that the information from international markets can be immediately reflected in the price of the Chinese market, the market volatility in China has been significantly reduced. This is especially true for the overnight volatility. Meanwhile, the risk management willingness of Chinese investors is growing.

Volatility can be considered as a measure of risk or fluctuation on the returns of an asset. The expected future volatility not only greatly facilitates the investment decision-making process but also helps financial practitioners to anticipate financial risks. Broadly speaking, the volatility measurements can be classified into three categories: (1) historical volatility; (2) implied volatility; and (3) realized volatility. In the historical volatility models, a time-varying volatility process is extracted from financial returns data. Most volatility models can be regarded as variants of the generalised autoregressive conditional heteroskedasticity (GARCH) models; see Bollerslev (1986) and Bollerslev, Engle, and Nelson (1994) for a review. A rival class of volatility models is associated with the stochastic volatility (SV) model; such as Taylor (1986) and Harvey, Ruiz, and Shephard (1994). The main drawback of the historical volatility models is that they cannot accommodate the intraday information and generally fail to capture the longer daily volatility movements sufficiently. Implied volatility, derived from option pricing data in combination with an option pricing model, can be regarded as a good predictor of future volatility. Implied volatility is forward-looking as opposed to historical-based methods, which are backwards-looking. Implied volatility is often applied to price options contracts. For those securities without corresponding market option prices, however, the implied volatility can hardly be obtained. Andersen, Bollerslev, Diebold, and Ebens (2001) construct estimates of ex post realized volatility (RV) by summing squares and cross-products of intraday high-frequency returns. There are several advantages to using RV to measure volatility for actively traded assets. First, it is model-free. Second, as the sampling frequency of the returns approaches infinity, it is barely harmed by the measurement error theoretically (Andersen et al., 2001). Lastly, in the absence of market microstructure noise, it is also an unbiased and efficient proxy of the integrated volatility; see Huang and Tauchen (2005) and Andersen, Bollerslev, and Meddahi (2004).

Night trading sessions in SHFE provide investors with a platform to avoid overnight position-holding risks. Furthermore, a reduction in intraday volatility would help enterprises reduce the exposure in risks. Accurately forecasting volatility is the key to risk management for traders, portfolio managers, policymakers, and other market participants alike. However, a volatility forecasting method can be ineffective even in the presence of time series with low volatility, but only if the series is characterized by regularities (Tsakalozos, Drakakis, & Rickard, 2011). As such, in order to examine the contribution of night trading to volatility forecasting, we apply a group of HAR models to capture the effects of night trading.

A set of volatility models based on RV has been developed over the years. Among them, the heterogeneous autoregressive (HAR) realized volatility model put forth by Corsi (2009) is widely used today. The main idea of the HAR model is that agents with different time horizons perceive, react to, and affect different types of volatility components, which in essence is a basic linear model following an autoregressive structure. Despite its simplicity, the HAR model can reproduce the same volatility persistence observed in the empirical data as well as many of the other main characteristics of financial data (fat tails, self-similarity, etc.). As for the forecasting performance of future volatility, the HAR model incurs smaller forecasting error than daily standard deviation and those based on parametric assumptions. Later, Corsi et al. find the residuals of commonly used time series models for realized volatility exhibit non-Gaussianity and volatility clustering (Corsi, Mittnik, Pigorsch, & Pigorsch, 2008). They put forth the HAR-GARCH model that allows for time-varying volatility of realized volatility and found that it substantially improves the model's fitness as well as predictive performance. Currently, many HAR-GARCH-type models have been looked into as possible means for modelling the volatility experienced in the commodity futures markets (Qu, Duan, & Niu, 2018). In this paper, both the HAR model and HAR-GARCH model are used to investigate the volatility performance.

Other variables that have potential effects on volatility are added to the HAR model, such as market liquidity captured by trading volume and leverage effects captured by negative returns (Fleming, Kirby, & Ostdiek, 2003; Todorova, 2015). Empirical results show that including market variables, such as volume, bid-ask spread, and the slope of the futures curve, helps with predictive power and significantly improves daily, weekly, and monthly volatility forecasts (Haugom, Langeland, Molnár, & Westgaard, 2014). Previous studies focus primarily on mature futures markets, such as the London Metals Exchange (LME) and the Chicago Mercantile Exchange (CME) with nearly 24 trading hours, which could provide continuous trading data. As for emerging markets, however, such as China commodity futures markets, intra-data with limited trading hours may cause bias. Evidence has shown that the information accumulated during non-trading hours can contribute substantially to the overall risk in China commodity futures markets (Liu & An, 2014; Tseng, Lai, & Lin, 2012). Therefore, overnight returns during non-trading hours were added to the HAR models when studying the volatility in China's commodity futures markets. Researchers found it helpful as the inclusion improved the models' goodness of fit and predictive power (Wang, Wu, & Xu, 2015). However, these papers did not account for the night trading sessions into the volatility models in China commodity futures markets. In our paper, besides considering overnight volatility into the RV, we also include the trading volume (V) measuring external information and a jump component to the HAR-V-J model.

We aim to contribute to the research on commodity futures markets in the following aspects. First, we used an extensive set of high-frequency data to obtain precise gold futures daily realized volatility. The daily realized volatility fell markedly since the night trading sessions were launched in July 2013. Further decomposing the daily realized volatility into the night, overnight, and day realized volatility, we found that the additional night trading sessions absorb news arrivals from international markets, which decrease the overnight jumps. Second, we extended the HAR model by incorporating jump and external information measured by volume in HAR-V-J model and incorporated volatility clustering into HAR-GARCH model by combining a GARCH specification. Third, we took advantage of HAR and the two HAR extension models to conduct a comparative study of volatility forecasting performances before and after the launch of night trading sessions in SHFE gold futures market. The findings have significant practical values for investors and policymakers.

The remainder of this paper is conducted as follows. Section 2 describes the realized volatility measurements and the realized volatility models. Section 3 provides summarized data and compares the in-sample estimation and out-of-sample forecasting performances before and after the launch of night trading sessions. Section 4 concludes and offers corresponding suggestions.

2 | METHODOLOGY

2.1 | Realized volatility

Following Corsi (2009), we assumed that under a risk-neutral market, the logarithmic $N \times 1$ vector price of a financial asset, p_t, follows a standard continuous-time stochastic volatility process,

$$dp_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t \tag{1}$$

where μ_t follows a continuous and locally bounded variance process, the process for the $N \times N$ positive definite diffusion matrix, σ_t, is strictly stationary. W_t denotes a standard N-dimensional Brownian motion, q_t is the counting process with a time-varying intensity, $\lambda(t)$, satisfying $p(dq_t = 1) = \lambda(t)dt$, $\kappa_t \equiv p_t - p_{t-}$, where κ_t stands for the size of discrete jump components under the process of logarithmic price.

We normalized the unit time interval, or $t = 1$, to represent one trading day. For integer j, we defined the within-day geometric returns as

$$r_{t,j} = p_{t-1+\frac{j}{M}} - p_{t-1+\frac{j-1}{M}},\ j = 1, 2, \ldots, M \tag{2}$$

where M is the sample frequency. Specifically, the daily returns of day t, r_t is made of three components, r^n_{t-1}, r^o_{t-1}, and r^d_t, representing night trading returns of day $t-1$, overnight returns of day $t-1$, and returns during daytime trading of day t, respectively. This arrangement is in line with the time share and daily records of SHFE and other Chinese futures exchanges. We choose to neglect lunch break hours or break hours between daytime and night trading sessions to better focus on the continuous intraday data (Figure 1).

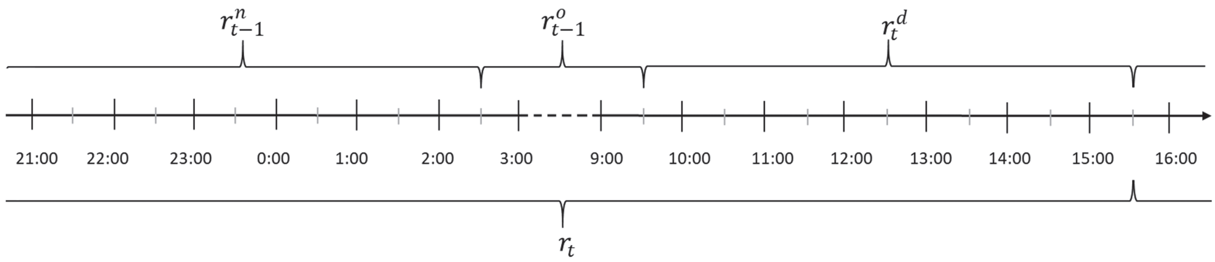


FIGURE 1 breakdown of the daily return of day t

$$r_t = r_{t-1}^n + r_{t-1}^o + r_t^d \tag{3}$$

Following Barndorff-Nielsen and Shephard's study of two general measures of realized within-day price variance, we have the classic realized variance, $\sum_{j=1}^{M} r_{t,j}^2$, and the realized bi-power variation (BPV) (Barndorff-Nielsen & Shephard, 2004),

$$BPV_t = \frac{1}{2\pi}\left(\frac{M}{M-1}\right)\sum_{j=1}^{M}\left|r_{t,j-1}\right|\left|r_{t,j}\right| \tag{4}$$

where $1/2\pi$ denotes the expected mean according to a standard normal distributed random variable.

Barndorff-Nielsen and Shephard (2004) proved that as the sample frequency increases, the realized variance

$$\sum_{j=1}^{M} r_{t,j}^2 \to \int_0^h \sigma_{t+\tau} d_\tau + \sum_{j=1}^{K_t} \kappa_{t,j}^2 \tag{5}$$

where K_t is the number of jumps within day t. And Barndorff-Nielsen and Shephard (2004) give that

$$BPV_t \to \int_0^h \sigma_{t+\tau} d_\tau \tag{6}$$

almost surely for all t. Thus, BPV_t provides a consistent estimator of the integrated variance unaffected by jumps.

In the following content, we follow Corsi (2009) to take the square root of the realized variance as ex post realized volatility for the integrated latent volatilities that are asymptotically free of measurement error.

$$RV_t = \sqrt{\sum_{j=1}^{M} r_{t,j}^2} \tag{7}$$

We took the truncated difference $J_t = \max\left(0, RV_t - \sqrt{BPV_t}\right)$ as a consistent estimator of the pure jump. In the gold futures market, prices jump due to unanticipated events happened globally or fluctuations from other financial markets.

2.2 | HAR and HAR extension models

The standard HAR model proposed by Corsi (2009) is an additive cascade model of volatility components defined over different time horizons as follows:

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_t^W + \beta_3 RV_t^M + \varepsilon_{t+1} \tag{8}$$

where RV_t^W is the lagged weekly realized volatility and RV_t^M is the lagged monthly realized volatility. The volatility components over different time horizons reflect the actions of different types of market participants. Institutional investors, such as insurance companies and pension funds who trade much less frequently and possibly for larger amounts, play a major role in the lagged monthly realized volatility. Meanwhile, the trading activities from short-term traders, such as dealers, market makers, retail investors, and intraday speculators, have a significant effect on the lagged weekly or daily realized volatility. They are calculated according to the following equations.

$$RV_t^W = \frac{1}{5}\left(RV_t + RV_{t-1} + \ldots + RV_{t-4}\right) \tag{9}$$

$$RV_t^M = \frac{1}{20}\left(RV_t + RV_{t-1} + \ldots + RV_{t-19}\right) \tag{10}$$

According to our definition of daily returns, as after-hour information, the squared overnight returns have been accounted for in the daily realized volatility. Specifically,

$$RV_t = \sqrt{\sum_{j=1}^{M_1}\left(r_{t-1,j}^n\right)^2 + \left(r_{t-1}^o\right)^2 + \sum_{j=1}^{M_2}\left(r_{t,j}^d\right)^2} \tag{11}$$

$$\sqrt{BPV_t} = \sqrt{\frac{1}{2\pi}\left(\frac{M_1}{M_1-1}\right)\sum_{j=1}^{M_1}\left|r_{t-1,j-1}^n\right|\left|r_{t-1,j}^n\right| + 0 + \frac{1}{2\pi}\left(\frac{M_2}{M_2-1}\right)\sum_{j=1}^{M_2}\left|r_{t,j-1}^d\right|\left|r_{t,j}^d\right|} \tag{12}$$

where M_1 is the number of night trading intraday prices, M_2 is the number of daytime intraday prices, $M_1+M_2 = M$. Overnight BPV_t is specified as zero. This means we divide overnight returns into a jump component.

Clark (2012) stated that both traded volume and volatility are driven by the same underlying "news" variable and will, therefore, be positively correlated. Studies confirmed a shift in volume to the new period after each extension (Asem & Kaul, 2008). Since the launch of night trading sessions, the SHFE has become the second-largest gold futures exchange worldwide in terms of the trading volume. Together considering potential threat arising from jumps—a stochastic process arising due to unanticipated news which significantly affects futures price—we added trading volume and jump components into the standard HAR model to get the HAR-V-J model, which is given as follows:

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_t^W + \beta_3 RV_t^M + \beta_4 Vol_t + \beta_5 J_t + \varepsilon_{t+1} \tag{13}$$

To keep the model concise, we only considered the daily trading volume and daily jump in the model. In this paper, the trading volume is standardized to eliminate the effect of the quantity of data.

The volatility clustering in financial markets is widely confirmed. In this paper, we follow Corsi et al. (2008) to establish the HAR-GARCH model. This allows for time-varying volatility by including a GARCH (p, q) specification. For conciseness, we adopted the GARCH (1, 1) process for the conditional variance of realized volatility. Hence, the proposed model is represented as

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_t^W + \beta_3 RV_t^M + \varepsilon_{t+1} \tag{14}$$

$$\varepsilon_{t+1} = \sqrt{h_{t+1}}e_{t+1} \tag{15}$$

$$h_{t+1} = \omega + \alpha\varepsilon_t^2 + \beta h_t \tag{16}$$

$$e_{t+1} \mid \Omega_t(0\tilde{,}1) \tag{17}$$

where Ω_t is the σ-field generated by all the information available up to time t. The residual ε_{t+1} follows a conditional density with time-varying variance.

3 | EMPIRICAL ANALYSIS

3.1 | Data and descriptive statistics

Our dataset covers 11 years of 5-min-frequency data (widely adopted as optimal sampling interval) from the main gold futures contracts of gold, from 9 January 2008, the launch date of the gold futures contract, to 9 January 2019. To make a comparative analysis of the volatility forecasting performances, we have divided the sample into two periods by the launch date of the night trading session, namely, 5 July 2013. Therefore, sample period I is from 9 January 2008, 9:00:00 CST to 5 July 2013, 15:00:00 CST; sample period II is from 5 July 2013, 21:00:00 CST to 9 January 2019, 15:00:00 CST. There are 48 5-min intervals in the daytime trading and 66 of the night trading, which generated 58,643 price observations in the first period (1,328 days) and 147,721 price observations in the second one (1,318 days). The data are from Wind dataset.

To obtain the daily realized volatility, we calculated the 5-min logarithmic returns and overnight logarithmic returns, whose descriptive statistics are reported in Table 1. Sample period I exhibits larger extreme values and higher standard deviations relative to Sample period II, which shows the volatility of the gold futures in SHFE significantly decreases after the launch of night trading sessions. Both sample periods are leptokurtic and have fat-tail distributions. In particular, the launch of night trading sessions changed the returns' skewness from negative to positive, which has significantly reduced the downside risk.[1]

Figure 2 displays the returns of periods I and II. Daily returns of the period I fluctuate more markedly than that in period II, particularly during the financial crisis, namely, between 2008 and 2009. Decomposing daily returns into overnight and day (time) returns before night trading session was launched, it was found that the volatilities came mainly from overnight fluctuation. After the introduction of night trading sessions, fluctuation of daily returns has a clear trend of moderation, which is mostly absorbed by night trading sessions. Figure 3 shows the realized volatility between two sample periods, providing more direct evidence. A similar conclusion can be obtained.

Table 2 presents descriptive statistics for the realized volatility. The Augmented Dickey-Fuller test indicates that the realized volatility series are stationary. The

[1]Empirical evidence shows that low or negative skewness implicates higher downside risks of asset returns. Reference on Amaya, D., P. Christoffersen, K. Jacobs and A. Vasquez, 2015, Does Realized Skewness Predict the Cross-Section of Equity Returns? *Journal of Financial Economics*, 118(1): 135–167.

TABLE 1 Descriptive statistics for returns on SHFE gold futures

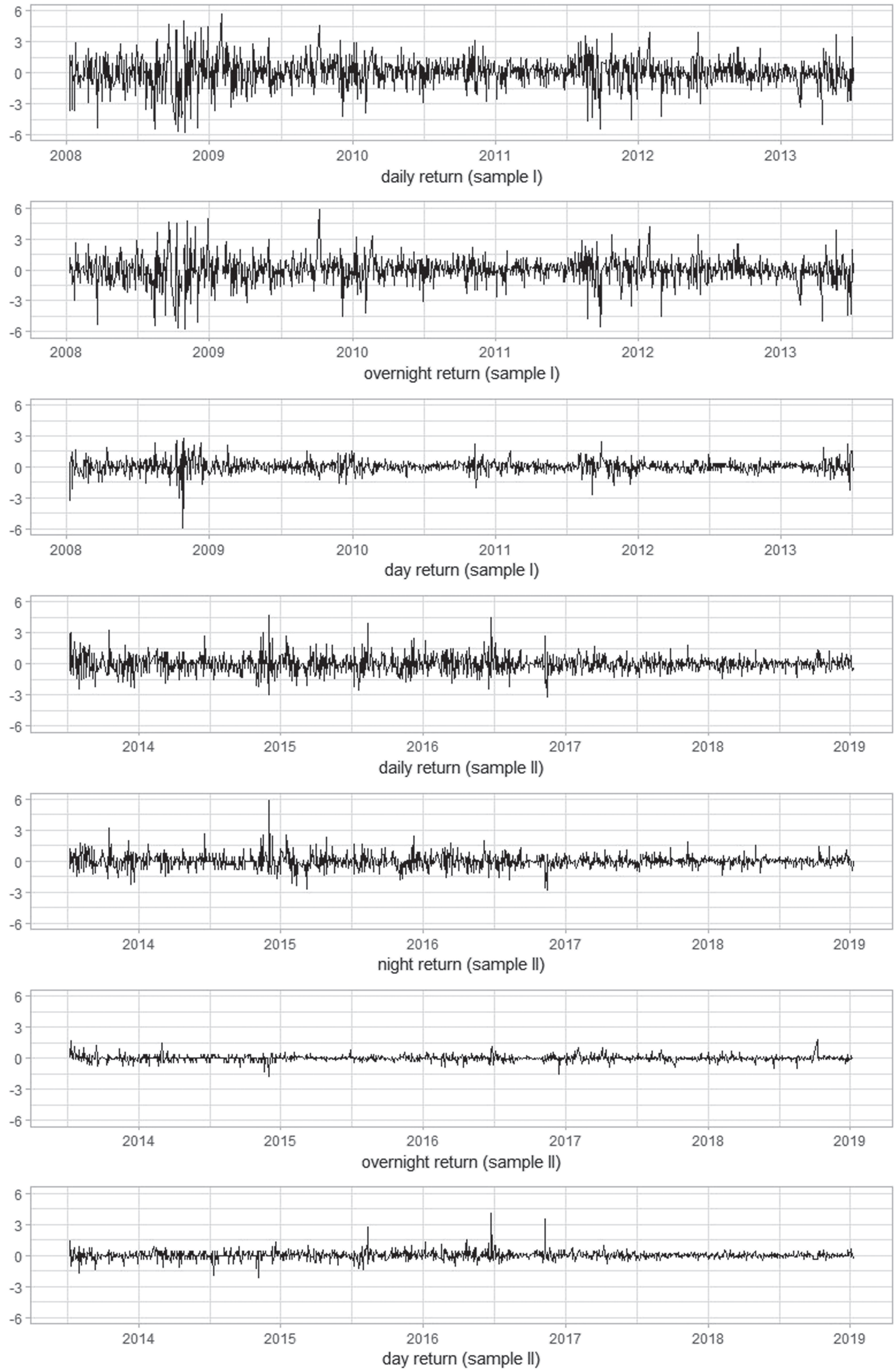
Variable	Period I Daily returns	Period I Overnight returns	Period I Daytime returns	Period I Volume	Period II Daily returns	Period II Night returns	Period II Overnight returns	Period II Daytime returns	Period II Volume
Obs.	1,328	1,328	1,328	1,328	1,318	1,318	1,318	1,318	1,318
Mean	0.008	−0.036	0.044	38,879	0.006	−0.01	0.003	0.013	316,387
Median	0.044	0	0.029	27,335	0	0	0	0	271,970
Std.	1.381	1.253	0.594	36,780	0.798	0.631	0.259	0.381	177,982
Min	−7.843	−8.333	−5.932	4	−3.13	−2.796	−1.732	−2.198	52,946
Max	5.76	7.377	2.746	320,426	4.692	5.934	1.777	4.071	1,365,032
Skew	−0.505	−0.417	−0.826	3.2	0.538	0.834	0.239	1.398	1.900
Ex-Kurt	3.324	6.325	11.235	14.318	3.14	8.214	8.926	18.579	5.409
JB	661***	2,262***	7,162***	−3.692**	604***	3,442***	4,406***	19,453***	−5.238***

Note: Values of returns are percentage logarithmic returns. Ex-Kurt represents excess kurtosis. JB is the Jarque–Bera test for normality. Period I is from 9 January 2008, 9:00:00 CST to 5 July 2013, 15:00:00 CST; Period II is from 5 July 2013, 21:00:00 CST to 9 January 2019, 15:00:00 CST.

**Significance at the 5% level.

***Statistical significance at the 1% level.

FIGURE 2 Comparison of two sample periods of returns series



Ljung-Box Q-statistics up to the 20th order confirms a high degree of serial correlation, except for the lag 5 Overnight RV from sample period II. Thus, the HAR-type models and their extensions can be applied.

3.2 | In-sample analysis

Table 3 lists the in-sample estimation results for three HAR and extension HAR models from two sample periods.

Generally, comparing the goodness of fitting of two periods, the adjusted R square, it has improved remarkably in the second sample period than that of the first sample period, when the night trading session was not launched.

It is also clear that all coefficients in the HAR models from both periods are highly significant and their following influences increase with the lagged horizon of realized volatility as shown by the changes of coefficients. Lagged monthly realized volatility has the strongest explanatory power. Before the launch of night trading

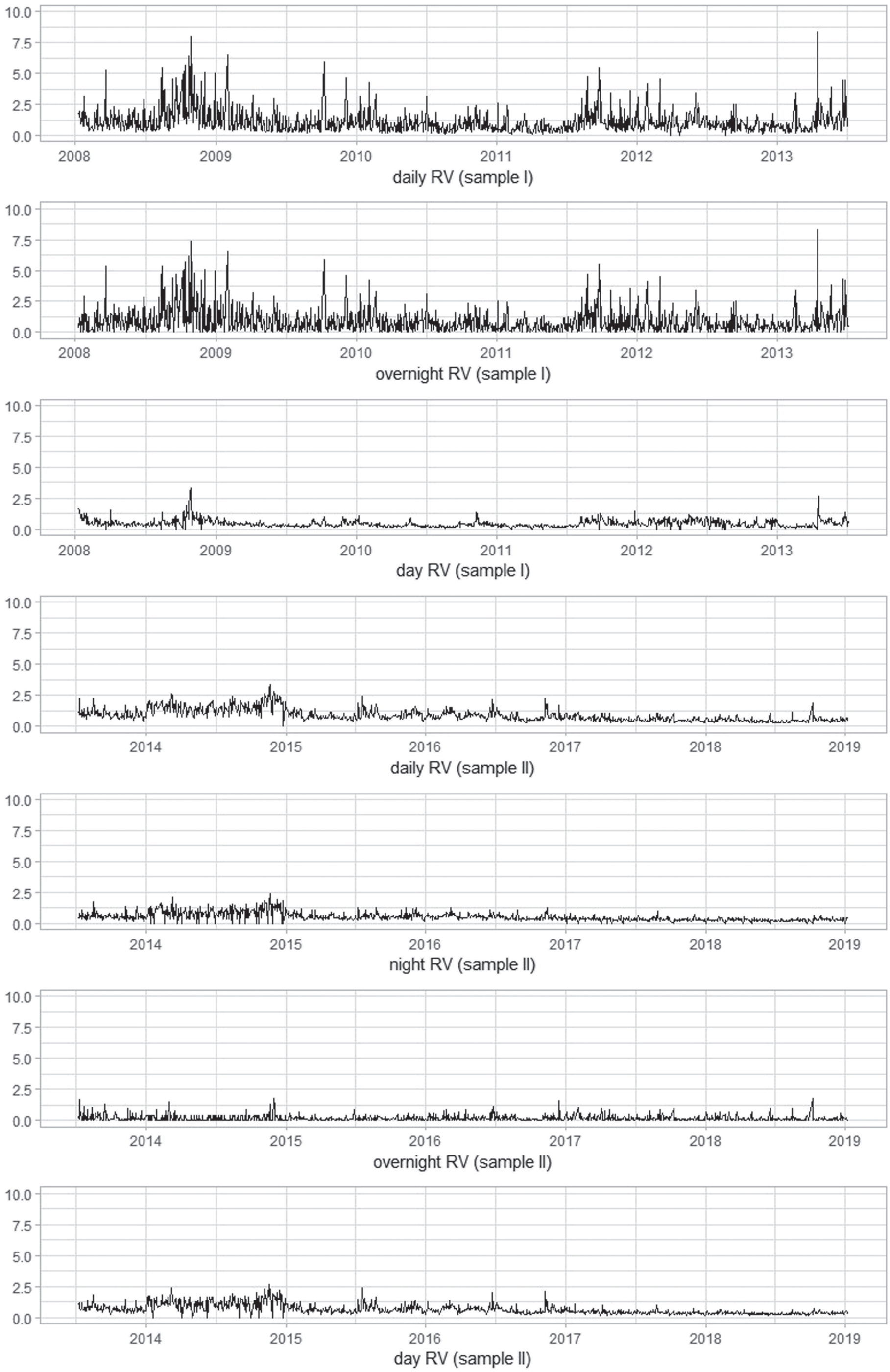


FIGURE 3 Comparison of two sample periods of RV series

sessions, the 1-day volatility has only a slight influence on the next day's volatility; the week-lagged volatility has a tenfold impact; The month-lagged volatility has more than 20-fold impact. When it came to the second sample period with night trading sessions, 1-day volatility largely played an increasing role, but still not as important as the week- and month-volatility. This illustrates that institutional investors, such as insurance companies and pension funds who trade much less frequently and possibly for larger amounts, play a major role in the SHFE gold futures throughout. Meanwhile, the trading activities from short-term traders, such as dealers, market makers, and intraday speculators, are getting more and more active.

Upon their inclusion, both jump and trading volume in HAR-V-J model are statistically significant. But the fitting degree is smaller in magnitude compared with the HAR model. The 1-day lagged trading volume in both sample periods are positively related to the next day's volatility, which is consistent with our expectation that

TABLE 2 Descriptive statistics for realized volatilities of SHFE gold futures

	Period I			Period II			
Variable	Daily RV	Overnight RV	Daytime RV	Daily RV	Night RV	Overnight RV	Daytime RV
Obs.	1,328	1,328	1,328	1,318	1,318	1,318	1,318
Mean	1.029	0.829	0.463	0.694	0.514	0.155	0.362
Median	0.753	0.548	0.388	0.567	0.414	0.084	0.274
Std.	0.905	0.940	0.302	0.407	0.332	0.208	0.273
Max	8.378	8.333	3.286	2.974	2.363	1.777	2.244
Skew	2.957	2.812	2.873	1.430	1.575	2.980	1.914
Ex-Kurt	12.724	11.397	17.356	2.149	2.931	13.421	4.698
ADF	−6.415***	−7.271***	−5.551***	−4.813***	−5.645***	−9.376***	−6.426***
LB(5)	343.2***	112.8***	1760.2***	2296.6***	1669.7***	4.296	1127.0***
LB(10)	580.9***	214.4***	2734.1***	4365.7***	3206.8***	29.8***	1997.0***
LB(20)	983.9***	390.2***	3884.2***	7987.7***	5903.6***	68.3***	3599.0***

Note: Ex-Kurt represents excess kurtosis. ADF is the Augmented Dickey-Fuller test for stationary. LB (·) stands for the Ljung-Box *Q*-statistics. Period I is from 9 January 2008, 9:00:00 CST to 5 July 2013, 15:00:00 CST; Period II is from 5 July 2013, 21:00:00 CST to 9 January 2019, 15:00:00 CST.
***Statistical significance at the 1% level.

TABLE 3 In-sample fit comparison of the HAR, HAR-V-J and HAR-GARCH models

	Period I: before the launch of night sessions			Period II: after the launch of night sessions		
Variables	HAR	HAR-V-J	HAR-GARCH	HAR	HAR-V-J	HAR-GARCH
Intercept	0.206***	0.220***	0.183***	0.038***	0.064***	0.051***
RV_t	0.022***	1.749***	−0.039	0.165***	−0.071**	0.117***
RV_t^W	0.226***	0.116***	0.309***	0.271***	0.265***	0.381***
RV_t^M	0.552***	0.453***	0.544***	0.506***	0.503***	0.416***
Vol_t		0.035***			0.017***	
J_t		−1.740***			0.277***	
ω			0.011***			0.002***
α			0.056***			0.054***
β			0.932***			0.918***
Adjusted R^2	0.16	0.18	0.16	0.549	0.55	0.547
AIC	3232.3	3202.3	2995.4	319.5	319.3	70.8
BIC	3258.2	3238.5	3031.63	345.3	355.5	107.0

Note: *t* statistics are calculated by Newey-West covariance correction. AIC/BIC represents the Akaike/Bayesian Information Criterion. Period I is from 9 January 2008, 9:00:00 CST to 5 July 2013, 15:00:00 CST; Period II is from 5 July 2013, 21:00:00 CST to 9 January 2019, 15:00:00 CST.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

when more information arises, there is greater volatility. It is interesting to note that before night trading sessions were launched, the lagged 1-day jump greatly reduced the coming day's volatility. A possible explanation is that before the launch of night trading sessions, investors tend to overreact to the information, leading to excess volatility of the market to some extent.

When considering the volatility clustering features of realized volatility in the HAR-GARCH model, the coefficients of both ARCH terms and GARCH terms are positive and significant. In addition, the sizes of these coefficients are almost unchanged in the two sample periods, which indicates that the introduction of night trading sessions has not changed the features of volatility

TABLE 4 Relative forecasting power test for the out-of-sample volatility forecasts

	N.forward	Period I HAR	Period I HAR-V-J	Period I HAR-GARCH	Period II HAR	Period II HAR-V-J	Period II HAR-GARCH
MAE	1	0.454	0.464	0.820	0.092[a]	0.097	0.630
	5	2.551	0.477	0.813	0.532	0.122[a]	0.638
	10	5.061	0.486	0.828	1.069	0.149[a]	0.651
MSE	1	0.628	0.629	1.215	0.023[a]	0.023[a]	0.455
	5	3.507	0.654	1.214	0.127	0.028[a]	0.468
	10	7.033	0.674	1.242	0.251	0.036[a]	0.489

[a]The best performance among all HAR-type models.

clustering and high persistence of the realized volatility in the SHFE gold futures market.

3.3 | Out-of-sample forecasting comparison

Out-of-sample predictions are used to make direct comparisons of the forecasting performance of the models before and after the introduction of night trading sessions. We used the period from 9 April 2012 to 5 July 2013 in Period I and from 19 October 2017 to 9 January 2019 in Period II to evaluate the out-of-sample forecast accuracy of the HAR model, HAR-V-J model, and HAR-GARCH model. Out-of-sample includes 300 days. One day, 1 week (5 days), and 2 weeks (10 days) forward forecasts are generated with the rolling window method. The first estimation windows are respectively, from 9 January 2008 to 8 April 2012 in Period I, and from 6 July 2013 to 18 October 2017 in Period II.

To assess and compare the accuracy of the HAR models forecasts, we adopted the loss functions, which include the mean absolute error (MAE) and the mean square error (MSE). They are defined as follows:

$$MAE = N^{-1}\sum_{i=1}^{N}\left|RV_{t+i} - RV_{t+i,pred}\right| \quad (18)$$

$$MSE = N^{-1}\sum_{i=1}^{N}\left(RV_{t+i} - RV_{t+i,\ pred}\right)^2 \quad (19)$$

where RV_{t+i}, $RV_{t+i,pred}$ denote the actual and i-day-forward forecasts of daily volatility from models. N is the number of sample days observed.

Table 4 shows the MAE and MSE comparisons for predictions made by three model specifications when using three different rolling windows in two sample periods. All comparisons indicated that the forecasting performances of volatility models showed great improvement in the period after night trading sessions were launched. Looking into the specific models in the second sample period, when predicting 1-day forward volatility, the classical HAR model performs slightly better than HAR-V-J model. But over the predicting horizon, the performance of the HAR model deteriorates sharply, while HAR-V-J model keeps up a high level of prediction power. As for HAR-GARCH model, the performance is not satisfactory for all prediction time horizons.

In general, the HAR-V-J model, which takes into account the endogenous volatility and the exogenous information, is the best performing model.

4 | CONCLUSION

In order to explore whether the introduction of night trading sessions in SHFE gold futures market increased volatility forecasting performances, we leveraged HAR and another two HAR extension models to capture and compare the effects. Our results show that there were large fluctuations in SHFE gold futures market before the launch of night trading sessions, which are mostly driven by the overnight fluctuation in the international gold market. After the launch of night trading sessions, realized volatility has a clear trend of moderation. In the in-sample estimation, both jump and external information have significant explanatory power in HAR-V-J model. Besides, the volatility clustering and high persistence of the realized volatility are confirmed by the GARCH coefficients. Finally, the study found that the night trading sessions have significantly improved the out-of-sample forecasting performances of the realized volatility models. Among them, the HAR-V-J model has revealed to be the best performing model. This conclusion holds for different prediction horizon and has significant practical values for investors and policymakers.

Future research can opt to explore different pathways. To note, it is possible to expand the covered research sample to more commodity futures that have introduced the night trading sessions in China's future markets. Further studies on commodity futures that are closely connected to the international futures markets are recommended. These include but are not limited to nonferrous metals futures and agriculture futures. Moreover, the suspension of night trading sessions due to unexpected events, such as the COVID-19 pandemic, could further validate the impact of night trading sessions.

ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (grant nos. 71532004 and 71773024).

ORCID

Xuan Yao https://orcid.org/0000-0002-9368-176X
Xiaofeng Hui https://orcid.org/0000-0002-7085-6176
Kaican Kang https://orcid.org/0000-0001-9082-3788

REFERENCES

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*, 43–76. https://doi.org/10.1016/S0304-405X(01)00055-1

Andersen, T. G., Bollerslev, T., & Meddahi, N. (2004). Analytical evaluation of volatility forecasts. *International Economic Review*, *45*(4), 1079–1110. https://doi.org/10.1111/j.0020-6598.2004.00298.x

Asem, E., & Kaul, A. (2008). Trading time and trading activity: Evidence from extensions of the NYSE trading day. *European Journal of Finance*, *14*(3), 225–242.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37. https://doi.org/10.1093/jjfinec/nbh001

Bollerslev, T. (1986). Bollerslev, generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T., Engle, R., & Nelson, D. B. (1994). ARCH models. *Handbook of Econometrics*, *IV*(313), 2959–3038.

Clark, P. K. (2012). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, *41*(1), 135–155.

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

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, *27*(1–3), 46–78. https://doi.org/10.1080/07474930701853616

Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics*, *67*(3), 473–509. https://doi.org/10.1016/S0304-405X(02)00259-3

Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, *61*(61), 247–264.

Haugom, E., Langeland, H., Molnár, P., & Westgaard, S. (2014). Forecasting volatility of the U.S. oil market. *Journal of Banking and Finance*, *47*(1), 1–14.

Huang, X., & Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, *3*(4), 456–499.

Hui, X., Yao, X., & Ma, Y. (2020). Research on the effect of night trading on the market of precious metal futures. *Operations Research Management Science*, *29*(5), 207–217.

Kao, E. H., & Fung, H. G. (2012). Intraday trading activities and volatility in round-the-clock futures markets. *International Review of Economics and Finance*, *21*(1), 195–209. https://doi.org/10.1016/j.iref.2011.06.003

Kauffman, N. (2013). Have extended trading hours made agricultural commodity markets riskier? *Economic Review (Kansas City)*, *93*(3), 67–94.

Klein, T., & Todorova, N. (2018). The impact of SHFE's night trading session on volume and realized volatility of aluminum and copper futures markets. *SSRN Electronic Journal*, (September), 1–29. https://doi.org/10.2139/ssrn.3249598

Liu, Q., & An, Y. (2014). Risk contributions of trading and non-trading hours: Evidence from Chinese commodity futures markets. *Pacific-Basin Finance Journal*, *30*, 17–29. https://doi.org/10.1016/j.pacfin.2014.07.005

Qu, H., Duan, Q., & Niu, M. (2018). Modeling the volatility of realized volatility to improve volatility forecasts in electricity markets. *Energy Economics*, *74*, 767–776. https://doi.org/10.1016/j.eneco.2018.07.033

Taylor, S. J. (1986). Modelling price volatility. In *Modelling financial time series* (pp. 62–96). New York: World Scientific.

Todorova, N. (2015). The course of realized volatility in the LME non-ferrous metal market. *Economic Modelling*, *51*, 1–12. https://doi.org/10.1016/j.econmod.2015.07.005

Tsakalozos, N., Drakakis, K., & Rickard, S. (2011). Signal extrapolation using empirical mode decomposition with financial applications. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 5744–5747.

Tseng, T.-C., Lai, H. C., & Lin, C. F. (2012). The impact of overnight returns on realized volatility. *Applied Financial Economics*, *22*(5), 357–364. https://doi.org/10.1080/09603107.2011.613760

Wang, X., Wu, C., & Xu, W. (2015). Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects. *International Journal of Forecasting*, *31*(3), 609–619.

AUTHOR BIOGRAPHIES

Xuan Yao Ph.D student in Harbin Institue of Technology Research Interests: Risk Management, Financial Engineering.

Prof Xiaofeng Hui Professor in School of management, Harbin Institute of Technology, China.

Research Interests: Financial crisis contagion mechanism, Financial market and financial risk control.

Dr Kaican Kang Research Associate in Credit Research Initiative, National University of Singapore, Singapore. Research Interests: Risk management.

How to cite this article: Yao X, Hui X, Kang K. Can night trading sessions improve forecasting performance of gold futures' volatility in China? *Journal of Forecasting*. 2021;40:849–860. https://doi.org/10.1002/for.2748