THE INFORMATION CONTENT OF OPEN INTEREST FOR THE REALIZED RANGE-BASED VOLATILITY: EVIDENCE FROM CHINESE FUTURES MARKET

Tseng-Chan Tseng[1], Hung-Cheng Lai[2], Conghua Wen[3]

[1] Zhaoqing University, Economics and Management College, Zhaoqing, China.
   2018013044@zqu.edu.cn, ORCID: 0000-0001-5259-817X

[2] Overseas Chinese University, Department of Finance, Taichung, Taiwan.
   hclai@ocu.edu.tw, ORCID: 0000-0002-3370-4579

[3] Xi'an Jiaotong-Liverpool University, Department of Mathematical Sciences, Suzhou, China.
   Conghua.wen@xjtlu.edu.cn, ORCID: 0000-0002-3938-8108

To cite this document

ABSTRACT

Purpose - The paper studies the impact of the infroamtion content of open interst on the realized range-based vaolatility of Chinese futures markets.
Methodology- We employ a hybrid range-based estimator to measure the integrated variance in the heterogeneous autoregressive (HAR) model, which also incorporates the variable of open interest into the HAR model on index futures prices of China Securities Index (CSI) 300.
Findings- Our findings demonstrate that the variable of open interest has a significant explanatory power with regard to the future realized volatility of the CSI 300 index futures.
Conclusion- The modified model enhances volatility forecasting performance, thereby indicating it has more accurate predictive power. Our results provide supports for the implication of the sequential information arrival hypothesis.

Keywords: Volatility forecasting, HAR-RRV model, realized range-based tripower variance, CSI 300 futures
JEL Codes: C32; G13; Q40

1. INTRODUCTION

Over recent decades, there have been considerable studies dedicated to explore the relationship between the volatility of futures contracts, their trading volume and open interest. According to the sequential information arrival (SIA) hypothesis, originally proposed by Copeland (1976), and subsequently complemented by Jennings et al. (1981), traders receive new information in a sequential and random fashion. Thus, from an initial position of equilibrium where all traders possess the same set of information, the arrival of new information on the market leads traders to revise their expectations accordingly (Darrat et al., 2003). As such, the SIA hypothesis suggests that lagged volume or lagged open interest may have the ability to predict current volatility, and vice versa.[1]

The recent studies demonstrated that trading volume was positively related to volatility, suggesting that trading volume can serve as a proxy measure for the unobservable information flowing into the market. (Lee and Rui, 2002; Chan et al., 2004 ; Kumar and Pandy, 2010; Boonvorachote and Lakmas, 2016). However, open interest is also an important variable in the futures market. It is defined as the number of contracts existing in a futures market that has not yet been closed out (Ripple and Moosa, 2009). Open interest is different from trading volume which captures the number of contracts traded during a specific time period. Ferris et al. (2002) suggested that the open interest can serve as a proxy measure for the information

---

[1] Both Lamoureux and Lastrapes (1990) and Andersen (1996) suggest that trading volume can serve as a proxy measure for the information flowing into the market.

flowing into and out of the futures contracts. Ripple and Moosa (2009) found that the open interest does provide explanation of futures volatility for the crude oil contract.[2]

More recently, Yen and Chen (2010) found that significant in-sample relationships exist amongst the future's daily volatilities, the lagged total volume, and the lagged total open interest. Shakeel and Ashraf (2012) also demonstrated that significant relationships between return volatility, volume and open interest in the Indian futures markets. Souček (2013) investigated the link between the demands for hedging on various markets proxied by the open interest. Boonvorachote and Lakmas (2016) showed a positive contemporaneous relationship between expected and unexpected trading volume and volatility, however open interest mitigates volatility.

The heterogeneous autoregressive realized volatility (HAR-RV) model, proposed by Corsi (2009), a simple approximate long-memory dynamics model of realized volatility, has been proved to achieve an excellent forecasting performance, and also provides a valuable economic interpretation. However, Christensen and Podolskij (2007) and Martens and Dijk (2007) claimed that the estimation of realized range-based volatility (RRV) showed more efficient and sufficient than realized return-based volatility sampled at fixed intervals, since the extremes were formed from the entire price process. More recently, Christensen and Podolskij (2012) proposed the realized range-based multipower variation and presented how it can be used to estimate the ex-post quadratic return variation and conduct jump-robust inference about integrated variance.

Even though a lot of studies in the forecasting of realized volatility have focused on the equity markets of developed economics (Ghysels et al., 2006; Andersen et al., 2007; Forsberg and Ghysels, 2007; Christoffersen et al., 2010), as well as foreign exchange markets (Andresen et al., 2001; Beine et al., 2007; Lyócsa et al., 2016) and commodities (Haugom et al., 2014; Lyócsa and Molnár, 2016), less studies are devoted to the equity futures markets of developing economies, especially using the range-based estimation. We therefore adopt the SIA hypothesis in this study with an attempt to achieve improvements in the overall forecasting of realized range-based volatility (RRV), whilst also incorporating lagged and open interest into the HAR-RRV model to facilitate our examination of China Securities Index (CSI) 300 index futures. Particularly, we use the realized range-based tripower variance (RTV), proposed by Christensen and Podolskij (2012), to measure the integrated variance in this study. We denote this modified forecasting model as the HAR-RRV-RTV-OI model. The main implication of this study is the information content of prior open interest can be used to forecast future volatility.

Despite being the second largest economy in the world, the derivatives markets are just emerging in China, especially the stock index futures markets. In recent years, however, the trading volume in the index futures market has jumped significantly in China. Moreover, active participating in the global market, China is now playing a fundamentally critical role in the world economy. Therefore, it is of great values to study the futures markets of China.

Ripple and Moosa (2009) argued that the futures markets differ from equity markets in many respects and suggested that the open interest provides additional information because the relationship between open interest and trading volume is quite complex. Thus, this empirical analysis based upon futures market complements the literature on the interconnection of non-parametric historical volatility measurements using model-based forecasting with intraday high-low statistics, as well as under the SIA hypothesis for open interest. Moreover, our empirical analysis extends recent study on the forecasting capabilities of realized range-based multipower variations (Christensen and Podolskij, 2012). To our best knowledge, this study is the first paper to show a model using realized range-based tripower variance as a regressor and attempts to explore the impact of open interest on realized volatility for the Chinese futures market, whilst assessing the forecasting performance of the modify model.

The remainder of the paper proceeds as follows. Methodology is described in the Section 2. The penultimate section presents the empirical findings, with the final section offers the concluding remarks for this study.

2. METHODOLOGY

2.1. Realized Volatility Measures

The realized volatility (RV) measure based on the sum of squared intraday returns, originally proposed by Andersen and Bollerslev (1998), is the availability of estimator of integrated variance that facilitates the formulation of a more accurate and meaningful inter-daily measurement of ex-post volatility. Subsequently, a series of studies address the modelling and forecasting of realized volatility. (See, for example, Barndorff-Nielsen and Shephard, 2006; Andersen et al., 2007; Forsberg and Ghysels, 2007; Corsi, 2009; Curci and Corsi, 2012; Louzis et al., 2012; Vortelinos and Thomakos, 2012; Xu, 2012; Souček and Todorova, 2013; Christoffersen et al., 2014; Iliescu and Dutta, 2016). Furthermore, the HAR-RV model has been

---

[2] Also see, for example, Bessembinder and Seguin (1993); Chen et al. (1995); Liew and Brooks (1998); Girma and Mougoue (2002); Motladiile and Smit (2003); Serletis and Shahmoradi (2006).

demonstrated that it could successfully capture the effects of volatility persistence, proposed by Corsi (2009), and also been shown to achieve better forecasting performance than other models (Andersen et al., 2007; McAleer and Medeiros, 2008; Jain, 2011; Todorova and Souček, 2014; Pu et al., 2016).

More recently, Christensen and Podolskij (2007) and Martens and Dijk (2007) replaced each squared intra-day return by the high–low range for that period to create a novel and more efficient estimator called the realized range. They suggested that the intraday range-based estimation of volatility was a more accurate and efficient method to predict volatility. Todorova (2012) developed the realized range-based volatility, an estimator obtained from intraday ranges, which is more efficient and less biased than the daily ranges volatility. Todorova and Husmann (2012) also demonstrated that realized range-based estimator performed better in terms of both bias and efficiency, although its performance still suffered from discrete trading.

2.2. A Hybrid Range-Based Estimator

Following Christensen and Podolskij (2007) and Martens and van Dijk (2007), we consider a scalar log-price p(t) at time t evolving within continuous time jump-diffusion semi-martingales. The quadratic variation (QV) for the cumulative return process, which, for any sequence of partitions 0 = t0 < t1 < …< tn = t with $\max_{1\leq j\leq n}|t_j - t_{j-1}| \to 0$ for n → ∞, is defined as:

$$QV_t = \int_0^t \sigma^2(t)dt + \sum_{0<t_j<t} \kappa^2(t)$$

(1)

where the integrated volatility (IV) of the continuous component, $\int_0^t \sigma^2(t)dt$ and the jump component, $\sum_{0<t_j<t} \kappa^2(t)$.

Using the measures of realized range-based variance used by Christensen and Podolskij (2012), we define realized range-based variance (RRV), realized range-based bipower variation (RBV), and realized range-based tripower variation (RTV) as follows:

$$RRV_t^b = \frac{1}{\lambda_2} \sum_{j=1}^n (s_{t,j})^2$$

(2)

$$RBV_t = \frac{1}{\lambda_1} \frac{1}{\lambda_1} \sum_{i=2}^n s_{t,j}\, s_{t,j-1}$$

(3)

$$RTV_t = \frac{n}{n-2} (\frac{1}{\lambda_{2/3}})^3 \sum_{j=3}^n (s_{t,j})^{2/3} (s_{t,j-1})^{2/3} (s_{t,j-2})^{2/3}$$

(4)

where $s_{t,j} = \max_{(j-1)/n\leq l,h\leq j/n}\{p_h - p_l\}$ for i = 1, 2,…, n, with n is sampling frequency in a trading day, and ph and pl are the respective highest and lowest prices during each five-minute period, and $\lambda_r = 4/\sqrt{\pi}(1-4/2^r)2^{r/2}\Gamma((r+1)/2)\varsigma(r-1)$. The superscript b appearing in the definition of RRV_t^b indicates that it will be biased in the presence of jumps (Christensen and Podolskij, 2006).

We denote that range-based bi-power variation as RBV_t and range-based tripower variation as RTV_t, which both provide an alternative means of drawing inferences on the continuous path component of QV_t.

$$RBV_t \xrightarrow{p} \int_0^t \sigma^2(t)dt \quad \text{and} \quad RTV_t \xrightarrow{p} \int_0^t \sigma^2(t)dt$$

(5)

However, Christensen and Podolskij (2012) suggested that the realized range-based tripower variance (RTV) was not only a more robust estimator of the integrated variance but also the most efficient so far. They proposed a hybrid range estimator, which has the form of a linear combination of the original range-statistic and a jump-robust measure of integrated variance. Based on this manipulation, we consider a purely range-based estimator that is consistent for the quadratic variation of the jump-diffusion semi-martingale defined as:

$$RRV_t \equiv \lambda_2 RRV_t^b + (1-\lambda_2)RTV_t \xrightarrow{p} \int_0^t \sigma^2(t)dt + \sum_{0<t_j<t} \kappa^2(t)$$

(6)

2.3. Volatility Modeling

Traditionally, many relationships in finance are intrinsically nonlinear, and the most popular nonlinear models are the ARCH or GARCH models. GARCH-type processes are very popular since they capture various aspects of the volatility behaviour. However, on the downsides of GARCH-type models, the implied persistence of their estimates is too high to be reconciled with the observed behaviour of squared returns (Gallo and Pacini, 2000). More recently, Andersen et al. (2011) argued that simple linear models can often predict future volatility more accurately than the sophisticated models such as FIGARCH and ARFIMA.

The HAR-RV model, proposed by Corsi (2009), not only successfully captures long-memory feature of realized volatility in a parsimonious linear manner, but also provides a reasonable and meaningful economic interpretation. Thus, this paper uses an efficient range-based estimator to describe the dynamic volatility in the HAR mode. We adopt RTV as the regressor for prediction of realized range-based volatility, similar to Tseng et al. (2012) and Todorova and Souček (2014), with the following specification of the HAR-RRV-RTV model:

$$RRV_{t,t+H} = \beta_0 + \beta_D RTV_{t-1,t} + \beta_W RTV_{t-5,t} + \beta_M RTV_{t-20,t} + \varepsilon_{t,t+H} \tag{7}$$

where H denotes the prediction horizons. This study considers five prediction horizons, one-day, and one- to four-week periods, for H=1, 5, 10, 15 and 20. The measures of multi-period realized volatility, comprising of $RTV_{t-5,t}$ and $RTV_{t-20,t}$, are the mean of the daily volatilities over the relevant time period.

Following Haugom et al. (2014), we Define $OI^+ = \max(OI, \text{median}(OI))$, $OI^- = \min(OI, \text{median}(OI))$ and consider various possibilities for the relationship between OI and volatility. We also incorporate OI⁺ and OI⁻ into the HAR–RRV–RTV model, and attempt to understand the effect of "normal" and "unusual" OI in volatility. However, the main purpose of this study is to attempt to improve the forecasting performance of future volatility. We find the incorporated factors do not improve the out-of-sample forecasting performance of future volatility (the results for incorporating OI⁺ and OI⁻ into the HAR–RRV–RTV model are not reported here to save space).

The recent study by Ripple and Moosa (2009) showed the importance of open interest in determining volatility, exerting a significant negative effect for the crude oil contract. Meanwhile, Shakeel and Ashraf (2012) also demonstrated that expected open interest mitigated volatility in the Indian futures markets. Furthermore, the sequential information arrival (SIA) hypothesis would hold if volatility is dependent upon the lagged volume or lagged open interest. Open interest, defined as the total number of contracts in place or open at the end of the trading day, proxies the demand for futures as hedging instruments and is often used as a proxy for market trading activity (Souček, 2013). Ripple and Moosa (2009) suggested that open interest provides additional important information, essentially as a result of the very complex relationship that exists between open interest and trading volume.

The fundamental premise of the SIA hypothesis is that the lagged trading volume or lagged open interest have not been completely assimilated into the prior change in price. In such a case, this study may expect that prior open interest, which contains relevant information not included in prior volatility, can be used for the prediction of future volatility. Thus, adopting the SIA hypothesis, we incorporate lagged open interest into the HAR–RRV–RTV model, and propose the following specification to obtain our socalled HAR–RRV–RTV-OI model:

$$RRV_{t,t+H} = \beta_0 + \beta_D RTV_{t-1,t} + \beta_W RTV_{t-5,t} + \beta_M RTV_{t-20,t} + \beta_{OD} OI_{t-1,t} + \varepsilon_{t,t+H} \tag{8}$$

where $OI_{t-1,t}$ denotes the augmenter of logarithm open interest of near-month contracts. The prediction horizons (H) of RRVt are also considered in line with the HAR–RRV–RTV model.

3. EMPIRICAL RESULTS

3.1. Data Source and Empirical Results

The China Securities Index (CSI) 300 index futures contract was launched on April 16, 2010 on the China Financial Futures Exchange (CFFEX). This study obtained intraday data on the Win Chinese Financial data and Information, with the sample period covering from April 21, 2010 to May 31, 2016, thereby providing 48 data periods of 5-minute duration within a single trading day, for a total of 1485 trading days for CSI 300 index futures. Intraday data on CSI 300 index futures are employed in this study to predict the volatility in futures prices based upon the market trading hours from 9:30am to 11:30am and 13:00pm to 15:00pm. Table 1 reports the descriptive statistics of the realized range-based volatility levels. The last column of Table 1 reports the descriptive statistics of the open interest of near-month contracts.

The last row of Table 1 reports the Ljung-Box test statistic for up to tenth-order serial correlation (LB10). From the results of Table 1, we noted that RTV appears to have a higher degree of serial correlation than RRV in the LB statistic, which implies

that RTV will be better at predicting future volatility. Although the LB statistic of RTV is slightly less than that of RBV, the standard deviation of RTV is less than that of RBV. It is noted that the range-based tripower extension is more efficient than its bipower companion. The final column of Table 1 denoting the log($\triangle$OI), reveals that the augmenter of logarithm open interest of near-month contracts has a high serial correlation.

Table 1: Descriptive Statistics of Realized Range-Based Volatility Levels for CSI 300 Futures

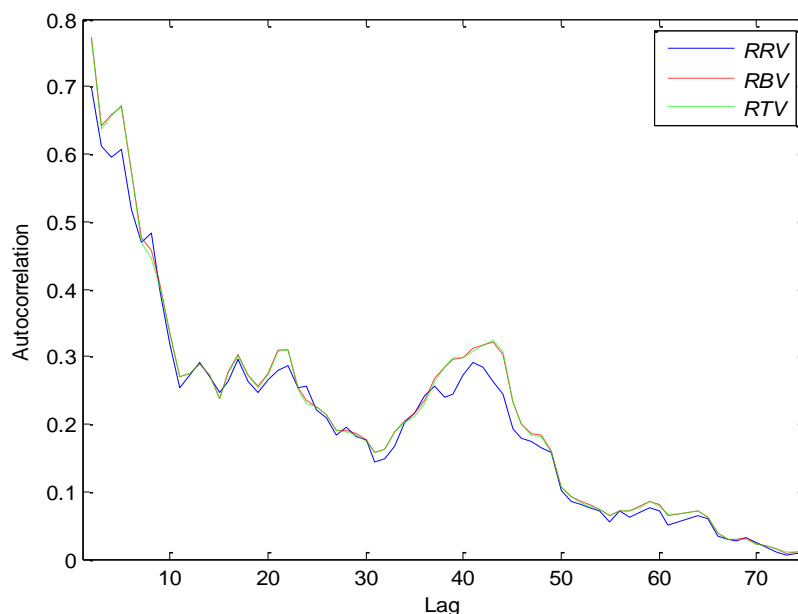| CSI 300 index futures	Variables[a]			
	RRV_t	RBV_t	RTV_t	$log(\triangle OI_t)$
Mean	2.917	2.190	2.131	0.002
Std. Dev.	6.231	4.877	4.795	0.222
Skew	9.138	9.262	9.551	2.254
Kurtosis	118.354	120.155	128.721	19.771
Min.	0.172	0.138	0.137	-1.262
Max.	111.222	91.382	92.186	1.769
LB$_{(10)}$[b]	4794.051	5388.055	5331.586	79.410

Notes:

[a] RRV_t denotes realized range-based variance; RBV_t denotes realized range-based bipower variance; RTV_t denotes realized range-based tripower variance; log($\triangle OI_t$) denotes the augmenter of logarithm open interest of near-month contracts.

[b] The critical value of the test statistic for LB$_{(10)}$ was 18

The autocorrelation function (ACF) of RRV, RBV and RTV, based upon range-based variance, is illustrated in Figure 1. From the inclusion of the first 75 lags, we find that the autocorrelations of RRV, RBV and RTV are positive, starting at about 0.7-0.8 and ending at about 0. It is interesting to note that the decay pattern for all of the series is on a rapidly decreasing trend for the first 30 lags. This situation of the autocorrelations of RBV and RTV both are slightly higher than that of RRV is emphasized in Table 1.

Figure 1: Autocorrelation Function for Realized Range-Based Volatility Measures of RRV, RBV and RTV in CSI 300 Index Futures, 2011-2016

The time-series plots of the annualized of RRV, RBV, and RTV elements covering the sample period from 2011 to 2016 are illustrated in Figure 2. The three annualized variances exhibit a strong own serial dependence, reflecting the volatility clustering in the data. We noted that both sample paths of RBV and RTV are less volatile than RRV, which is consistent with the finding of Christensen and Podolskij (2012). In particular, these variances reveal greater volatilities around June-August 2015, which sees the crash of Chinese Stock Markets. Intuitively, RTV is an efficient and persistent process, since it is jump-robust estimates of integrated variance.

Table 2 reports the in-sample estimation results of the HAR-RRV-RTV and HAR-RRV-RTV-OI models for CSI 300 index futures, with the results of the HAR-RRV-RTV model being presented in Panel A, and the results of the HAR-RRV-RTV-OI model being presented in Panel B. We focus on five different prediction horizons, one-day and one- to four-week periods, corresponding to RRV$_{t,\,t+H}$, for H = 1, 5, 10, 15 and 20, respectively. The estimates of β$_W$ for the 1-day prediction horizon and the estimates of β$_D$ and β$_W$ for one-week prediction horizon in the HAR–RRV–RTV model in Panel A are found to be statistically significant, at least at the 5% level. Moreover, With the exception of β$_W$ for the two- to four-week prediction horizons, all the other estimates of β$_D$ and β$_M$ in the HAR–RRV–RTV model in Panel A are found to be statistically significant at least at the 5% level.

Figure 2: Time series plot of annualized RRV,RBV and RTV for CSI 300 index futures, 2011-2016
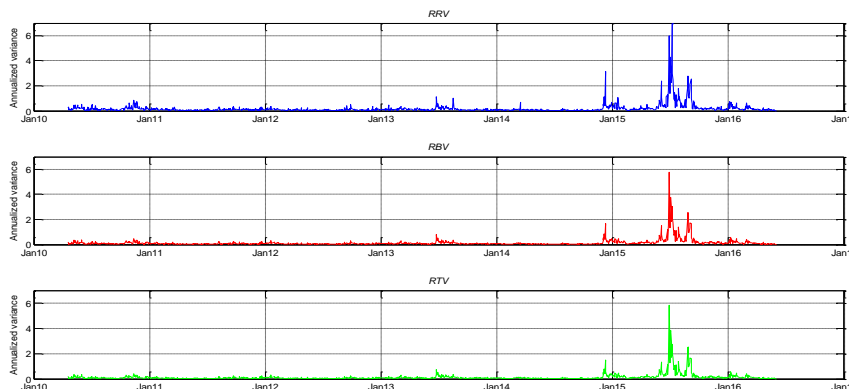

Table 2: In-sample estimates for CSI 300 index futures using the HAR-RRV-RTV and HAR-RRV-RTV-OI models [a]

Variables	Horizons									
	1 day		1 week		2 weeks		3 weeks		4 weeks	
	Coeff.[b]	S.E.	Coeff.[b]	S.E.	Coeff.[b]	S.E.	Coeff.[b]	S.E.	Coeff.[b]	S.E.
Panel A: *HAR-RRV-RTV* Model : $(RRV_{t,t+H} = \beta_0 + \beta_D RTV_t + \beta_W RTV_{t-5,t} + \beta_M RTV_{t-20,t} + \varepsilon_{t,t+H})$										
β$_0$	0.488 ***	0.125	0.714***	0.166	1.000 ***	0.166	1.198***	0.179	1.310***	0.195
β$_D$	0.320	0.198	0.436***	0.096	0.472***	0.069	0.345***	0.067	0.276***	0.064
β$_W$	0.815***	0.270	0.527**	0.210	0.155	0.146	0.139	0.141	0.154	0.150
β$_M$	0.001	0.101	0.065	0.096	0.266 **	0.124	0.316**	0.159	0.316**	0.155
Adj-R²	0.590		0.666		0.508		0.427		0.400	
MSE	16.052		9.379		12.018		12.310		11.653	
Panel B: *HAR-RRV-RTV-OI* Model : $(RRV_{t,t+H} = \beta_0 + \beta_D RTV_t + \beta_W RTV_{t-5,t} + \beta_M RTV_{t-20,t} + \beta_{OD} OI_t + \varepsilon_{t,t+H})$										
β$_0$	0.482 ***	0.123	0.709***	0.163	0.996 ***	0.165	1.194 ***	0.178	1.308 ***	0.195
β$_D$	0.322	0.199	0.437***	0.097	0.473 ***	0.070	0.346***	0.067	0.277 ***	0.064
β$_W$	0.818 ***	0.268	0.530**	0.209	0.157	0.146	0.140	0.141	0.155	0.150
β$_M$	-0.001	0.100	0.063	0.095	0.264 **	0.124	0.315 **	0.159	0.315 **	0.155
β$_{OD}$	1.020 *	0.532	0.981***	0.310	0.812 **	0.338	0.611 *	0.338	0.365	0.305
Adj-R²	0.591		0.667		0.510		0.428		0.401	
MSE	16.001		9.331		11.986		12.291		11.646	

Notes:

a   The table presents the adjusted R2 and mean squared error (MSE) for one-day and one- to four-week in-sample predictions of RRV for IF300 futures. The S.E. values are based upon Newey-West HAC standard errors. The dependent variables for the HAR-RRV-RTV and HAR-RRV-RTV-OI models for all horizons are standardized realized variance: RRVt,t+H/H.

b   *** indicates significance at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level.

It is interesting to note that the larger coefficients of β_W are found in the one-day and one-week prediction horizons, but the β_W coefficients are found to have less importance in the longer-run horizons. On the contrary, the β_M coefficients are found to have greater importance in the longer-run horizons. Meanwhile, the results reveal that the adjusted R^2, which is 0.590 for the daily prediction horizon, ranges between 0.666 and 0.400 for the one- to four-week horizons, with the MSE ranging between 16.052 and 11.653 across all five prediction horizons.

Several of the prior studies suggest that open interest has significant explanatory power with regard to the volatility of futures prices. For example, Bessembinder and Seguin (1993) and Chan et al. (2004) found that volatility was negatively related to open interest. In contrast, however, Liew and Brooks (1998) and Serletis and Shahmoradi (2006) suggested that a positive and statistically significant relationship existed between volatility and open interest (see also Yen and Chen 2010; Ripple and Moosa 2009; Shakeel and Ashraf 2012; Boonvorachote and Lakmas 2016). In the present study, we focus on the impact of open interest on realized volatility, in an attempt to improve futures volatility forecasting performance.

The in-sample estimation results of the HAR-RRV-RTV-OI model on CSI 300 index futures contracts are reported in Panel B of Table 2, from which we can see that a similar volatility pattern of HAR-RRV-RTV model is also discernible in the HAR-RRV-RBV-OI model. Subsequently, we focus on the estimates of β_{OD} for the five prediction horizons. With the exception of β_{OD} for the four-week prediction horizon, all the other estimates of β_{OD} in the HAR– RRV–RTV-OI model are found to be positively significant, at least at the 10% level, thereby indicating that lagged open interest containing the relevant information (which is not embedded in prior volatility) does have some effect on future realized volatility. As noted by Girma and Mougoue (2002) and Yen and Chen (2010), the lagged open interest reflects the relevant market information.

In other words, our findings indicate that lagged open interest does have some significant explanatory power with regard to futures realized volatility. This is also emphasized by the adjusted R^2 and MSE terms. Across all prediction horizons, the adjusted R^2 (MSE) is higher (lower) in the HAR-RRV-RTV-OI models than in the HAR-RRV-RTV models.

The economic interpretation of the HAR-RRV-RTV-OI model is that the future volatility expectations can be effectively composed by various components of lagged volatility and lagged open interest, most of which contain relevant information which has not been completely assimilated into the prior change in price.

3.2. Robustness Checks

The in-sample results imply that the modified model has more accurate predictive ability than the HAR-RRV-RTV model. However, the primary aim of this study is to improve overall volatility forecasting performance. In order to verify the modified model has accurate predictive capability, we conducted the modified Diebold–Mariano (DM) test (Harvey et al. 1997) to assess forecasting accuracy and evaluate the performance of the out-of-sample forecasting. Table 3 presents the out-of-sample results of HAR-RRV models for CSI 300 index futures, with Panel A reporting the MSE and Panel B reporting the t-statistic of the modified DM test.

Table 3: Out-of-sample forecasts performance for CSI 300 index futures

Horizon	Models	
	HAR-RRV-RTV	*HAR-RRV-RTV-OI*
Panel A: *HAR-RRV Model* [a]		
1 day	105.147	105.085
1 week	75.201	74.815
2 weeks	135.684	135.489
3 weeks	182.101	181.900
4 weeks	137.644	137.702
Panel B: Modified DM test [b]		
1 day	–	0.269
1 week	–	2.150**
2 weeks	–	2.046**
3 weeks	–	1.832**
4 weeks	–	-0.936

Notes:

[a] The table presents the out-of-sample forecasts for CSI 300 futures using five-minute intraday data covering the period from 31 May 2015 to 31 May 2016, with the data from 21 April 2010 to 31 May 2015 being used to estimate the HAR-RRV model parameters. The

[a] dependent variable is *RRV*, and the table entries represent the mean standard error (MSE) for the out-of-sample predictions based upon one-day, and one- to four-week out-of-sample *RRV* prediction horizons.

[b] Modified Diebold-Mariano (DM) test statistic p-values are presented for one-day and one- to four-week out-of-sample *RRV* predictions. The benchmark model for all horizons is the HAR-RRV-RTV model. If the test-statistic is positive (negative), it indicates that the examined model is more (less) accurate than the benchmark model.

[c] *** indicates significance at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level.

From the results reported in Panel A of Table 3, we can see that for CSI 300 index futures across all prediction horizons, the MSE results are lower in the HAR-RRV-RTV-OI model than those in the HAR-RRV-RTV model. Panel B of Table 3 shows the results of modified DM test, revealing that for one- to three-week prediction horizons, each of the HAR–RRV–RTV–OI models is statistically significant at the 5% level, as compared to the HAR–RRV–RTV model. The results demonstrate that the modified model improves volatility forecasting performance, thereby indicating lagged open interest does provide significant explanatory power in the volatility for CSI 300 index futures. Our results are found to hold for both the in-sample and out-of-sample predictions.

In summary, we incorporate an open interest variable into the HAR-RRV-RTV model in the present study, and claim that the modified model not only provides additional and useful information on volatility dynamics, but also better forecasting performance than the HAR-RRV-RTV model in terms of predicting the volatility for CSI 300 index futures. Finally, our findings are in line with Girma and Mougoue (2002) and Yen and Chen (2010), and support the implication of the SIA hypothesis.

4. CONCLUSIONS

This is the first study to show a linkage between realized range-based tripower variance and open interest in CSI 300 index futures contracts with the resultant volatility forecasting performance. Our primary aim is to incorporate the open interest variable into the HAR-RRV model with an attempt to achieve improvements in the overall forecasting of realized range-based volatility.

The values of this study are threefold. First of all, our empirical analysis reveals that open interest has significant explanatory power with regard to futures realized volatility for CSI 300 index futures. Secondly, our findings suggest that the modified model, our so-called HAR-RRV-RTV-OI model, does enhance volatility forecasting performance, indicating that the HAR-RRV-RBV-OI model has more accurate predicting power than the HAR-RRV-RBV model. Finally, our results not only provide supports for the implication of the SIA hypothesis, but also complement the literatures on the forecasting capabilities of realized range-based multipower variations. For our further study, we would like to investigate larger data sets, in particular equity futures from other countries, for example other Asia Pacific, and European markets.

REFERENCES

Andersen T.G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. Journal of Finance, 51, 169-204.

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

Andersen, T.G., T. Bollerslev, and F.X. Diebold (2007). Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics, 89, 701-20.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2001). The Distribution of Realized Exchange Rate Volatility. Journal of the American Statistical Association, 96: 42-55.

Andersen, T.G., T. Bollerslev, and X. Huang (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. Journal of Econometrics, 160, 176-189.

Barndorff-Nielsen, O.E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bi-power variation. Journal of Financial Econometrics, 4, 1-30.

Beine, M., J. Lahaye, S. Laurent, C. J. Neely and F. C. Palm (2007). Central bank intervention and exchange rate volatility, its continuous and jump components. International Journal of Finance & Economics, 12, 201-223.

Bessembinder, H., and P.J. Seguin (1993). Price volatility, trading volume, and market depth: Evidence from futures markets. Journal of Financial and Quantitative Analysis, 28, 21-29.

Boonvorachote, T. and K. Lakmas (2016). Price volatility, trading volume, and market depth in Asian commodity futures exchanges. Kasetsart Journal of Social Sciences, 37, 53-58.

Chan K.C., H.-G. Fung, and W.K. Leung (2004). Daily volatility behavior in Chinese futures markets. Journal of International Financial Markets, institutions and Money, 14, 491-505.

Chen N.F., C.J. Cuny, and R.A. Haugen (1995). Stock volatility and the levels of the basis and open interest in futures contracts. Journal of Finance, 25.281-300.

Christensen, K., and M. Podolskij (2006). Range-based estimation of quadratic variation. Working paper, Aarhus School of Business.

Christensen, K., and M. Podolskij (2007). Realized range-based estimation of integrated variance. Journal of Econometrics, 141, 323-349.

Christensen, K. and M. Podolskij (2012). Asymptotic theory of range-based multipower variation. Journal of Financial Econometrics, 10, 417-456.

Christoffersen, P., Jacobs, K., and Mimouni, K. (2010). Volatility dynamics for the S&P500: evidence from realized volatility, daily returns, and option prices. Review of Financial Studies, 23(8), 3141-3189.

Christoffersen, P., B. Feunou, K. Jacobs, and N. Meddahi (2014). The economic value of realized volatility: Using high-frequency returens for option valuation. Journal of Financial and Quantitative Analysis, 49, 663-697.

Copeland, T. (1976). A model of asset trading under the assumption of sequential information arrival. Journal of Finance, 31, 1149-1168.

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

Curci, G. and F. Corsi (2012). Discrete sine transform for multi-scale realized volatility measures. Quantitative Finance, 12, 263-279.

Darrat, A.F., S. Rahman, and M. Zhong (2003). Intraday trading volume and return volatility of the DJIA stocks: A note. Journal of Banking and Finance, 27, 2035-2043.

Ferris, S.P., H.Y. Park, and K. Park (2002). Volatility, open interest, volume, and arbitrage: Evidence from the S&P 500 futures market. Applied Economics Letters, 9, 369-372.

Forsberg, L., and E. Ghysels. 2007. 'Why Do Absolute Returns Predict Volatility So Well?. Journal of Financial Econometrics, 5, 31-67.

Gallo, G.M., and B. Pacini (2000). The effects of trading activity on market volatility. European Journal of Finance, 6, 163-75.

Ghysels, E., P. Santa-Clara and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. Journal of Econometrics, 131, 59-95.

Girma, P.B. and M. Mougoue (2002). An empirical examination of the relation between futures spreads volatility, volume, and open interest. Journal of Futures Markets, 22, 1083-1102.

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

Iliescu, N., and S. Dutta (2016). The information content of implied volatility in developed versus developing FX markets. Applied Economics, 48, 5396-5404.

Jain, K. (2011). Time-varying beta: the heterogeneous autoregressive beta model. Manuscript, Duke University.

Jennings, R., L. Starks, and J. Fellingham (1981). An equilibrium model of asset trading with sequential information arrival. Journal of Finance, 36, 143-161.

Kumar, B., and A. Pandy (2010). Price volatility, trading volume and open interest: Evidence from Indian commodity futures markets. Working paper. India, Indian Institute and Management.

Lamoureux, C., and W. Lastrapes (1990). Heteroscedasticity in stock return data: volume versus GARCH effects. Journal of Finance, 45, 221-229.

Lee, B.-S., and O. Rui (2002). The dynamic relationship between stock returns and trading volume: domestic and cross-country evidence. Journal of Banking and Finance, 26, 51-78.

Liew, K.Y., and R.D. Brooks (1998). Returns and volatility in the Kuala Lumpur crude palm oil futures market. The Journal of Futures Markets, 18, 985–999.

Louzis, D.P., S. Xanthopoulos-Sisinis, and A.P. Refenes (2012). Stock index realized volatility forecasting in the presence of heterogeneous leverage effects and long range dependence in the volatility of realized volatility. Applied Economics, 44, 3533-3550.

Lyócsa, Š., and P. Molnár (2016). Volatility forecasting of strategically linked commodity ETFs: gold-silver. Quantitative Finance, 1-14.

Lyócsa, Š., P.Molnár, and I. Fedorko (2016). Forecasting Exchange Rate Volatility: The Case of the Czech Republic, Hungary and Poland. Finance a Uver: Czech Journal of Economics & Finance, 66(5).

Martens, M., and D. van Dijk (2007). Measuring volatility with the realized range. Journal of Econometrics, 138, 181-207.

McAleer, M., and M. Medeiros (2008). Realized Volatility: A Review. Econometric Review, 27, 10-45.

Motladiile, B. and E.v.d.M. Smit (2003). Relationship between share index volatility, basis and open interest in futures contracts: The South African experience. South African Journal of Business Management, 34, 41-50.

Pu, W., Y. Chen, and F. Ma (2016). Forecasting the realized volatility in the Chinese stock market: further evidence. Applied Economics, 48, 3116-3130.

Ripple, R.D., and I.A. Moosa (2009). The effect of maturity, trading volume, and open interest on crude oil futures price range-based volatility. Global Finance Journal, 20, 209-219.

Serletis, A., and A. Shahmoradi (2006). Returns and volatility in the NYMEX Henry Hub natural gas futures market. OPEC Review: Energy Economics & Related Issues, 30, 171-186.

Shakeel, M., and S. Ashraf (2012). Empirical relationship between index futures prices, volume and open interest: evidence from Indian futures market. The IUP Journal of Applied Finance, 18, 48-66.

Souček, M. (2013). Crude oil, equity and gold futures open interest co-movements. Energy Economics, 40, 306-315.

Souček, M., and N. Todorova (2013). Realized volatility transmission between crude oil and equity futures markets: A multivariate HAR approach. Energy Economics, 40, 586-597.

Tseng, T.-C., H.-C. Lai, and C.-F. Lin (2012). The impact of overnight returns on realized volatility. Applied Financial Economics, 22, 357–64.

Todorova, N. (2012). Volatility estimators based on daily price ranges versus the realized range. Applied Financial Economics, 22, 215-229.

Todorova, N. and M. Souček (2014). The impact of trading volume, number of trades and overnight returns on forecasting the daily realized range. Economic Modelling, 36, 332-340.

Todorova, N., and S. Husmann (2012). A comparaive study of range-based stock return volatility estimators for the German market. Journal of Futures Markets, 32, 560-586.

Vortelinos, D.I., and D.D. Thomakos (2012). Realized volatility and jumps in the Athens Stock Exchange. Applied Financial Economics, 22, 97-112.

Xu, D. (2012). Examining realized volatility regimes under a threshold stochastic volatility model. International Journal of Finance & Economics, 17, 373-389.

Yen, S.M., and M.-H. Chen (2010). "Open interest, volume, and volatility: evidence from Taiwan futures markets," Journal of Economics and Finance, 34, 113-141.