

Volatility Forecasting of China Silver Futures: the Contributions of Chinese Investor Sentiment and CBOE Gold and Silver ETF Volatility Indices

Wenbin HU¹

¹The University of Melbourne, Melbourne, Australia;

Abstract. This paper is to detect the role of CBOE gold ETF volatility index (GVZ), CBOE silver ETF volatility index (VXSLV), and constructed Chinese investor sentiment (CnSENT) on the volatility forecasting of China silver futures over daily, weekly and monthly horizons. Different types of HAR models and ridge regression models are utilized to do the analysis, and the out-of-sample R-square statistics and different rolling window sizes are used to ensure the robustness of the conclusion. The empirical results suggest that GVZ and VXSLV have the explanatory power on the China silver futures. Particularly, VXSLV has a better performance than GVZ. However, the predictive power of CnSENT is doubtful as some results indicate that it cannot improve the prediction accuracy. Additionally, the ridge regression method does not achieve a better result than all types of HAR models.

1 Introduction

Volatility forecasting is crucial in asset pricing and risk management, which has attracted many researchers' attention [1-6]. Due to the unobservable attribute of true volatility, many financial researches have focused on the implied volatility and historical volatility [7].

Silver, as a safe-haven asset, has played a crucial role in the commodity futures market [8]. The investors usually used it for the purpose of risk hedging and asset allocation optimization. During the economic downturn, the demand for silver increases [9].

The growing literatures focus on the volatility forecasting of futures prices. [10] suggested that the CBOE gold ETF volatility index (GVZ) and CBOE silver ETF volatility index (VXSLV) played a crucial role in volatility forecasting of China gold futures market. [11] indicated that VXSLV could help to predict the volatility of the China silver futures market. However, no study to date has examined the joint contribution of GVZ and VXSLV to the volatility forecasting of China silver futures market. Furthermore, [12] discovered that the FEARS index as a proxy for investor sentiment had the explanatory ability on the crude oil futures volatility. However, there has been less previous evidence for the explanatory power of Chinese investor sentiment (CnSENT) on the China silver futures volatility.

The Heterogeneous autoregressive model (HAR) proposed by [13] has been extensively used. The standard HAR model includes three components, which are daily, weekly and monthly variables. It can capture 'stylized facts' of long memory related to the volatility. The empirical findings suggest that high-frequency HAR

Corresponding author e-mail:wenbinhu@unimelb.edu.au

models can achieve superior performance than GARCH models and stochastic volatility models [14]. However, one of the disadvantages of HAR models is that it cannot solve the issues of overfitting and multicollinearity [10]. Ridge regression model, as a popular shrinkage model, can solve these issues.

Based on the issues above, Chinese silver futures market is analyzed thereby. And whether CBOE gold ETF volatility index, CBOE silver ETF volatility index and Chinese investor sentiment, i.e., GVZ, VXSLV, and CnSENT, have the explanatory power on the China silver futures volatility is investigated by using different types of HAR models. Moreover, the ridge regression method is also used to solve the overfitting issue and to see whether this method can achieve superior performance results than other models. Additionally, out-of-sample R-square statistics and different window sizes are utilized for robustness checks.

This paper has the following contributions to the existing literature. Firstly, [10] considered the joint contribution of GVZ and VXSLV on the China gold futures market. However, this research considers the predictive power of GVZ and VXSLV on the volatility forecasting of the China silver futures market. Secondly, [11] only used VXSLV to detect the predictive power on the China silver futures volatility. However, in this paper, the joint contribution of GVZ, VXSLV, and CnSENT to the China silver futures volatility is explored. Finally, existing studies used the FEARS index as a measure for the investor sentiment. However, the FEARS index may not represent the investors' sentiment in the context of the Chinese market. Another way is used to construct the CnSENT and to detect its explanatory power on the

Chinese silver futures market.

The rest of this paper is organized as follows. Section 2 relates to the data description. Section 3 presents the methodology. Section 4 discusses the results. Finally, Section 5 concludes.

2 The Data

Data has been obtained from Bloomberg, Wind, and Choice databases between February 03, 2017, to June 24, 2020. The data frequency for CnSENT is a daily level. The

data frequency for China silver futures prices, GVZ, and VXSLV, are high frequency.

Figure 1 shows the time series of all variables' daily RV on the whole sample period. Compared with GVZ, VXSLV, and CnSENT daily RV, China silver futures' RV is not very dynamic.

Table 1 is a summary of the descriptive statistics of all variables. Kolmogorov-Smirnov test and Jarque-Bera test indicate that all variables are not normally distributed. The Ljung-Box Q tests indicate that the autocorrelation between them is high. The augmented Dickey-fuller test indicates that all variables passed the unit root test, and there is no need for further data transformation.

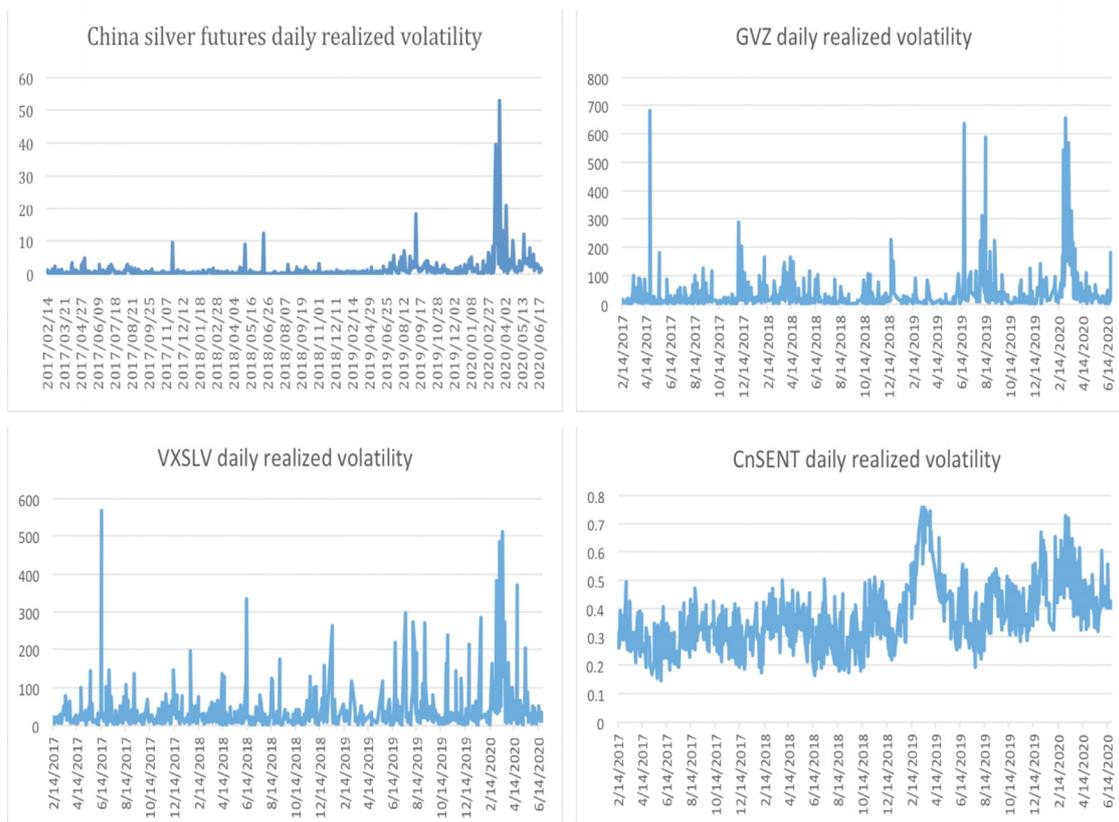


Figure 1. Time series of daily realized volatility of GVZ, VXSLV, CnSENT and China silver futures

Table 1. Descriptive statistics of China silver futures, GVZ, VXSLV, and CnSENT

	China silver futures	GVZ	VXSLV	CnSENT
Observations	702	702	702	702
Mean	1.4048	37.3772	36.7861	0.3608
Minimum	0.0628	0.9059	0.0024	0.1462
Maximum	52.8080	682.0110	568.3978	0.7583
Std. dev	3.5462	71.2868	58.8067	0.1099
Skewness	8.736	5.493	4.560	0.856
Kurtosis	99.85	40.95	29.76	3.904
Jarque-Bera	283270***	45656***	23380***	109.5***
K-S	0.353***	0.304***	0.266***	0.073***
Q(5)	560.49***	324.16***	115.24***	1214.9***
Q(22)	948.89***	549.66***	195.96***	3404.5***
ADF	-14.332***	-19.255***	-21.273***	-13.266***

Notes: The Kolmogorov-Smirnov and Jarque-Bera methods are used to test whether the sample follows the normal distribution. The Ljung-Box Q is used to test the serial autocorrelation of a time series based on a number of lags. The augmented Dickey-fuller test is used to determine whether a unit root exists in a sample. * indicates significance at the 10% level. ** indicates significance at the 5% level. *** indicates significance at the 1% level

3 Methodology

3.1 Investor sentiment construction

Based on the paper written by [15], several indicators are used to construct CnSENT as they can better capture the investors' emotions. These indicators are stock high opening rate, stock growth rate, stock effective growth rate, stock limit up rate, stock limit up closing rate, quantity relative ratio, and stock turnover rate. The stock high opening rate is calculated based on the Shanghai

(securities) composite index, and others are calculated based on China's A-share market. Due to the relative sensitivities of stock growth rate, stock limit up closing rate, quantity relative ratio and stock turnover rate, the values of these indicators are based on a 20-day average. Table 2 describes the notations and definitions related to these indicators.

Table 2. Key notation

Notation	Definition
SHOR	Stock high opening rate
HOD 60 days	High opening days within 60 days
SGR	Stock growth rate
SR	Number of stocks that rose on day t
SF	Number of stocks that fell on day t
SEGR	Stock effective growth rate
SRE	Number of stocks that rose effectively on day t
TNSS	Total number of stocks in the stock market
SLUR	Stock limit up rate
SKY	Number of stocks with a suspension of business in case of skyrocketing of stock prices
SLU	Number of stocks with a suspension of business in case of slump of stock prices
SLUCR	Stock limit up closing rate
CLU	Number of stocks that successfully closed limit upon day t
QRR	Quantity relative ratio
STR	Stock turnover rate
SV	Stock volume on day t
TNSI	Total number of stocks issued

Notes: these notations will be used to describe the following formulas of these indicators.

The calculation of these indicators is described as follows:

$$SHOR = \frac{HOD\ 60\ days}{60} \quad (1)$$

$$SGR = \frac{SR}{(SR+SF)} \quad (2)$$

$$SEGR = \frac{SRE}{TNSS} \quad (3)$$

A stock price increases by 5% means that this is an effective increase.

$$SLUR = \frac{SKY}{(SKY+SLU)} \quad (4)$$

$$SLUCR = \frac{CLU}{SKY} \quad (5)$$

If the stock price reached the daily limit and continued at the daily limit until the close of the stock market, it means that this stock successfully closed limit up on day t.

QRR can be directly obtained from the Wind database without any further calculation.

$$STR = \frac{SV}{TNSI} \times 100\% \quad (6)$$

As the correlation between these seven indicators is not very strong, they can be used in the following calculation and modeling. According to the method used by [16], the min-max normalization method is used to process the data. The entropy weight method (EWM) is an important model that has been widely used in the research [17]. Based on that method, their weights are calculated and then each indicator is weighted to construct the CnSENT.

3.2 Standard HAR model and its extensions

Following [18], the daily realized volatility (RV) can be defined as the sum of squared intraday returns:

$$RV_t = \sum_{i=1}^M r_{t,i}^2 \quad (7)$$

Where M represents the sampling frequency, $r_{t,i}$ represents the i th intraday returns on day t denoted by $r_{t,i} = 100 \times (\ln P_{t,i} - \ln P_{t,i-1})$. [13] indicated that the heterogeneous autoregressive model for realized volatility (HAR-RV) was an attractive model that has been widely studied and practiced. Many researchers used it as their benchmark model. In addition, compared with other models, this model can predict the performance more accurately [19]. The standard HAR-RV model can be specified as follows:

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \omega_{t+1} \quad (8)$$

Where ω_{t+1} is an unexpected error term, $RV_t^{(d)}$ is lagged daily RV. $RV_t^{(w)}$ is lagged weekly RV and $RV_t^{(m)}$ is lagged monthly RV. The weekly and monthly RV formulas can be specified as follows:

$$RV_t^{(w)} = \frac{1}{5}(RV_t^{(d)} + RV_{t-1}^{(d)} + RV_{t-2}^{(d)} + RV_{t-3}^{(d)} + RV_{t-4}^{(d)}) \quad (9)$$

$$RV_t^{(m)} = \frac{1}{4}(RV_t^{(w)} + RV_{t-1}^{(w)} + RV_{t-2}^{(w)} + RV_{t-3}^{(w)}) \quad (10)$$

The standard HAR-RV model are used as the benchmark model for comparison and analysis purpose. [11] noted that VXSLV contained the information for the prediction of RV in China silver futures market. [10] indicated that GVZ and VXSLV were helpful for the RV

forecasting in China gold futures market. What is the most interesting is that except VXSLV, whether GVZ and CnSENT can improve the performance of RV forecasting in China silver futures market. Thus, additional HAR extensive models can be specified as follows:

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{GVZ}^{(d)}GVZ_t^{(d)} + \beta_{GVZ}^{(w)}GVZ_t^{(w)} + \beta_{GVZ}^{(m)}GVZ_t^{(m)} + \omega_{t+1} \quad (11)$$

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{VXSLV}^{(d)}VXSLV_t^{(d)} + \beta_{VXSLV}^{(w)}VXSLV_t^{(w)} + \beta_{VXSLV}^{(m)}VXSLV_t^{(m)} + \omega_{t+1} \quad (12)$$

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{CnSENT}^{(d)}CnSENT_t^{(d)} + \omega_{t+1} \quad (13)$$

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{GVZ}^{(d)}GVZ_t^{(d)} + \beta_{GVZ}^{(w)}GVZ_t^{(w)} + \beta_{GVZ}^{(m)}GVZ_t^{(m)} + \beta_{VXSLV}^{(d)}VXSLV_t^{(d)} + \beta_{VXSLV}^{(w)}VXSLV_t^{(w)} + \beta_{VXSLV}^{(m)}VXSLV_t^{(m)} + \beta_{CnSENT}^{(d)}CnSENT_t^{(d)} + \omega_{t+1} \quad (14)$$

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{GVZ}^{(d)}GVZ_t^{(d)} + \beta_{GVZ}^{(w)}GVZ_t^{(w)} + \beta_{GVZ}^{(m)}GVZ_t^{(m)} + \beta_{VXSLV}^{(d)}VXSLV_t^{(d)} + \beta_{VXSLV}^{(w)}VXSLV_t^{(w)} + \beta_{VXSLV}^{(m)}VXSLV_t^{(m)} + \omega_{t+1} \quad (15)$$

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{GVZ}^{(d)}GVZ_t^{(d)} + \beta_{GVZ}^{(w)}GVZ_t^{(w)} + \beta_{GVZ}^{(m)}GVZ_t^{(m)} + \beta_{CnSENT}^{(d)}CnSENT_t^{(d)} + \omega_{t+1} \quad (16)$$

$$RV_{t+1} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \beta_{VXSLV}^{(d)}VXSLV_t^{(d)} + \beta_{VXSLV}^{(w)}VXSLV_t^{(w)} + \beta_{VXSLV}^{(m)}VXSLV_t^{(m)} + \beta_{CnSENT}^{(d)}CnSENT_t^{(d)} + \omega_{t+1} \quad (17)$$

3.3 Ridge regression

Overfitting issues refer to a model that contains more variables than are needed, which violates parsimony [20]. Ridge regression method has been extensively used to solve this problem [21]. This method is used to see whether this method can better improve the forecast performance. Based on the method proposed by [22] and used by [10], it has been calculated as:

$$\widehat{RV}_{t+1} = X_t' \widehat{\beta}^{Ridge} \quad (18)$$

where $\widehat{\beta}^{Ridge}$ is the estimator of regression coefficients, $X_t = (1, RV_t^{(d)}, RV_t^{(w)}, RV_t^{(m)}, GVZ_t^{(d)}, GVZ_t^{(w)}, GVZ_t^{(m)}, VXSLV_t^{(d)}, VXSLV_t^{(w)}, VXSLV_t^{(m)}, CnSENT_t^{(d)})$. $\widehat{\beta}^{Ridge}$ can be computed as:

$$\widehat{\beta}_i = (X'X + kI)^{-1}X'Y \quad (19)$$

Where X stands for (t-1)*10 matrix, Y represents the vector of the dependent variable, I represents the identity matrix, and K represents the ridge parameter.

3.4 Out-of-sample forecasting method

In order to further explore the out-of-sample predictive performance of these variables, [23]'s method is followed and R-squared statistics is computed as follows:

$$R_{OS}^2 = 1 - \frac{\sum_{k=1}^q (RV_{m+k} - \widehat{RV}_{m+k})^2}{\sum_{k=1}^q (RV_{m+k} - \widehat{RV}_{bench})^2} \quad (20)$$

Where RV_{m+k} represents actual RV, \widehat{RV}_{m+k} represents forecasted RV and \widehat{RV}_{bench} represents the RV predicted by the HAR-RV model on day_{m+k}. The results of other models are compared with the result of the benchmark model. If ΔR_{OS}^2 is positive, it means that this model has a better forecasting performance than the benchmark model. Moreover, [24] method is followed to do the hypothesis test to see whether the mean squared forecast error (MSFE) of the HAR-RV model is less than or equal to the MSFE of other models. The p-value is based on [24] test.

4 Results

In this section, the regression is done to get the in-sample estimation results of all models. Then, the out-of-sample R-squared statistics is used to see the out-of-sample predictive performance of these variables. Finally, different forecasting window sizes are used to do the robustness check in order to further confirm the results.

A. In-sample estimations

Table 3 is the summary of day_{t+1} in-sample estimation results of all models on the whole sample period from February 03, 2017, to June 24, 2020, by Ordinary Least Square (OLS) method. ΔR^2 represents the rise for the adjusted R² of other models relative to the adjusted R² of the benchmark model.

The result for HAR- RV model shows that $\beta^{(d)}$ and $\beta^{(m)}$ are all significant at 1% level and $\beta^{(w)}$ are not significant, suggesting that short term and long term lagged RV play an important role in the prediction of China silver futures volatility. The coefficients in the HAR-GVZ model indicate that $\beta_{GVZ}^{(d)}$ and $\beta_{GVZ}^{(w)}$ are significant at least at 5% significant level. However, $\beta_{GVZ}^{(m)}$ has a negative influence on the forecasting of China silver futures volatility, while its $\beta_{GVZ}^{(w)}$ has a positive impact on it. The coefficients estimate in the HAR-VXSLV model suggest that $\beta_{VXSLV}^{(w)}$ and $\beta_{VXSLV}^{(m)}$ are significant at least at 5% significant level. However, contrary to the findings of [11], it cannot be found that $\beta_{VXSLV}^{(d)}$ could contribute to the prediction of China silver futures volatility. This is interesting to note that the CnSENT in the Chinese stock market may not improve the prediction of China silver futures volatility as its coefficient is not significant.

With regards to the overall fitting evaluation, the results indicate that GVZ and VXSLV have predictive power for the forecast of China silver futures future volatility, which is consistent with the observations in [10] for China gold futures. VXSLV can improve volatility forecasts, which is also consistent with the results in [11] for China silver futures.

The ΔR^2 for HAR-GVZ model and HAR-VXSLV model are 10.1% and 21.9%, respectively, suggesting that VXSLV can contribute more in forecasting China silver futures volatility. The results show that CnSENT does not seem to improve the forecast of China silver futures volatility as the ΔR^2 for HAR-CnSENT model has

decreased. By only focusing on R-square, it is found that HAR-GVZ-VXSLV-CnSENT and HAR-GVZ-VXSLV are the two best models for prediction among all models. However, by concentrating on R^2 and ΔR^2 , it suggests that HAR-GVZ-VXSLV model is the best model that generates the highest ΔR^2 . This result can be explained by

the negative influence of this investor sentiment variable. Overall, all in-sample results point out that U.S. based gold and silver volatility indices can be very helpful for the forecast of China silver futures volatility, while the investor sentiment cannot improve the performance of the model.

Table 3. In-sample results of all models

	<i>HAR-RV</i>	<i>HAR-GVZ</i>	<i>HAR-VXSLV</i>	<i>HAR-CnSENT</i>	<i>HAR-GVZ-VXSLV-CnSENT</i>
$\beta^{(d)}$	0.374***	0.366***	0.434***	0.374***	0.441***
	(9.93)	(9.66)	(10.58)	(9.89)	(10.71)
$\beta^{(w)}$	-0.100	-0.288***	-0.913***	-0.100	-0.946***
	(-1.05)	(-2.93)	(-7.02)	(-1.05)	(-7.26)
$\beta^{(m)}$	0.410***	0.486***	0.857***	0.409***	0.877***
	(4.45)	(5.22)	(7.03)	(4.44)	(7.17)
$\beta_{GVZ}^{(d)}$		-0.005**			-0.006***
		(-2.40)			(-3.10)
$\beta_{GVZ}^{(w)}$		0.015***			0.014***
		(3.26)			(3.13)
$\beta_{GVZ}^{(m)}$		0.000			-0.006
		(0.10)			(-1.37)
$\beta_{VXSLV}^{(d)}$			-0.003		-0.001
			(-1.22)		(-0.43)
$\beta_{VXSLV}^{(w)}$			0.047***		0.042***
			(6.30)		(5.57)
$\beta_{VXSLV}^{(m)}$			-0.017**		-0.015*
			(-2.17)		(-1.87)
β_{CnSENT}				0.077	-0.054
				(0.08)	(-0.06)
β_0	0.412***	0.140	-0.186	0.386	-0.188
	(3.50)	(1.10)	(-1.38)	(1.05)	(-0.53)
R^2	0.368	0.407	0.449	0.368	0.462
<i>adj. R</i> ²	0.365	0.402	0.445	0.364	0.454
<i>F</i>	135.076	79.343	94.358	101.164	59.316
<i>N</i>	701	701	701	701	701
ΔR^2		0.101	0.219	-0.003	0.244

	<i>HAR-GVZ-VXSLV</i>	<i>HAR-GVZ-CnSENT</i>	<i>HAR-VXSLV-CnSENT</i>
$\beta^{(d)}$	0.441***	0.367***	0.434***
	(10.74)	(9.65)	(10.55)
$\beta^{(w)}$	-0.946***	-0.288***	-0.914***
	(-7.27)	(-2.93)	(-7.02)
$\beta^{(m)}$	0.877***	0.487***	0.859***
	(7.20)	(5.22)	(7.02)
$\beta_{GVZ}^{(d)}$	-0.006***	-0.005**	
	(-3.13)	(-2.36)	
$\beta_{GVZ}^{(w)}$	0.014***	0.015***	
	(3.13)	(3.26)	
$\beta_{GVZ}^{(m)}$	-0.006	0.000	
	(-1.37)	(0.10)	
$\beta_{VXSLV}^{(d)}$	-0.001		-0.003
	(-0.43)		(-1.21)
$\beta_{VXSLV}^{(w)}$	0.042***		0.048***
	(5.57)		(6.29)

$\beta_{VXSLV^{(m)}}$	-0.015*		-0.017**
	(-1.88)		(-2.18)
β_{CnSENT}		-0.228	-0.167
		(-0.23)	(-0.18)
β_0	-0.207	0.218	-0.128
	(-1.53)	(0.61)	(-0.36)
R^2	0.462	0.407	0.449
$adj. R^2$	0.455	0.401	0.444
F	66.002	67.923	80.770
N	701	701	701
ΔR^2	0.247	0.099	0.216

Notes: The table shows day_{t+1} in-sample estimation results of all models over the whole sample period from February 03, 2017, to June 24, 2020. ΔR^2 means that the adjusted R-square increases or decreases of other models relative to the adjusted R-square of the benchmark model. * indicates significance at the 10% level. ** indicates significance at the 5% level. *** indicates significance at the 1% level

B. Out-of-sample evaluations

According to [10], the in-sample test was used to confirm the accuracy of the model for past patterns. The out-of-sample test was used to ensure that the model has the ability to forecast future performance. From the perspective of investors, these researchers paid more attention to the estimation results of the out-of-sample test. In this paper, the out-of-sample R-squared statistics is used to get a deep understanding of the predictive power of models for future trends. Particularly, whether GVZ, VXSLV, and CnSENT contain the information for forecasting the volatility of China silver futures is examined.

Figure 2 shows the fluctuation of actual RV and the benchmark model' forecasted RV based on a 350-day rolling window. It indicates that the forecasting results generated by the benchmark model can capture some movements of realized volatility in the China silver futures market.

Figure 3 shows the fluctuation of actual RV and all other models' forecasted RV based on a 350-day rolling window. There are two ridge regression models where the X_t in M8 contains all variables except CnSENT and the X_t in M9 contains all variables. Compared with the forecasting results produced by the benchmark model, other HAR type models perform better as they can generate more sensitive forecasts. This means that U.S. based information is very useful for the prediction of China silver futures volatility. Furthermore, the results show that models based on the ridge regression method can predict the realized volatility more precisely than the benchmark model.

A summary of out-of-sample results based on a 350-day rolling window is given in table 4. The ΔR_{OS}^2 represents the variation of out-of-sample R^2 of other models relative to the out-of-sample R^2 of the benchmark model. The p-value is based on [24] test.

The results show that they are all significant at least at 5% significant level. By incorporating GVZ into the benchmark model, the ΔR_{OS}^2 has increased to 7.76%, suggesting that GVZ can improve forecasting performance. The ΔR_{OS}^2 has increased to 17.43% by including VXSLV, indicating that VXSLV can contribute more to the predictive accuracy than GVZ, which is consistent with the in-sample results. However, the contribution of CnSENT is still doubtful as some model shows a positive influence on volatility forecasting and HAR-VXSLV-CnSENT model shows an opposite result. The HAR-GVZ-VXSLV-CnSENT model generates the

maximum ΔR_{OS}^2 (21.35%). Moreover, the HAR-GVZ-VXSLV model still remains a very good forecasting performance, indicating that U.S. based GVZ and VXSLV have significant predictive powers for the future volatility of China silver futures.

For the out-of-sample test, ridge regression method is used to examine whether this method has superior predictive performance than all other models by solving the overfitting issue. However, the results reveal that the models based on ridge regression method do not perform better than other models. By including the variable of CnSENT, its result is in accord with the result of HAR-GVZ-VXSLV-CnSENT model. Without the incorporation of CnSENT, its prediction performance is the same as the performance of HAR-GVZ-VXSLV model.

Overall, the usefulness of GVZ and VXSLV for the volatility forecast of China silver futures can be discovered. However, the forecasting ability of CnSENT is questionable.

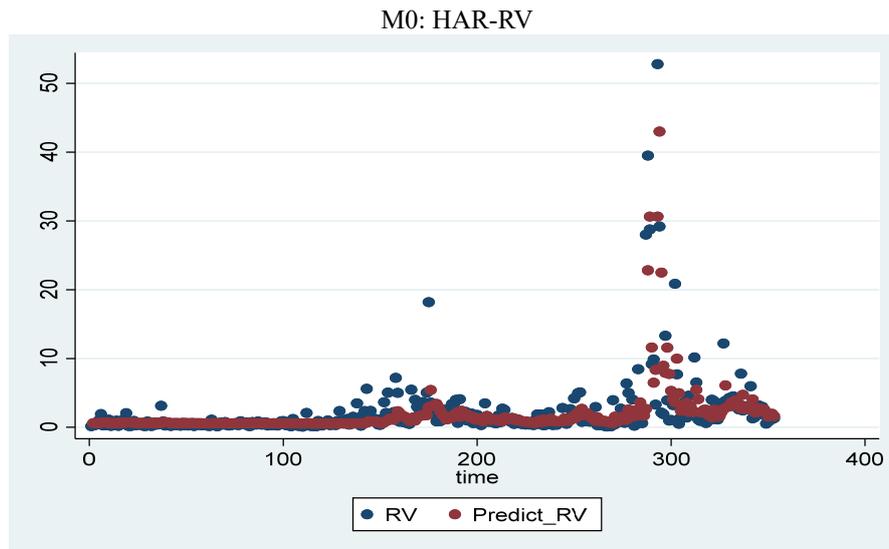


Figure 2. The variation of actual RV and the benchmark model' forecasted RV



Figure 3. The variation of actual RV and all other models' forecasted RV

Table 4. Out-of-sample estimation results

Forecasting models	ΔR_{OS}^2 (%)	MSPE-adjusted	p-value
M1: HAR-GVZ	7.7616***	0.9478	0.0061
M2: HAR-VXSLV	17.4334**	3.2967	0.0121
M3: HAR-CnSENT	0.2613**	0.0371	0.0165

M4: HAR-GVZ-VXSLV	21.1983***	3.9214	0.0087
M5: HAR-GVZ-CnSENT	7.9909***	0.9833	0.0054
M6: HAR-VXSLV-CnSENT	14.6438**	2.9205	0.0114
M7: HAR-GVZ-VXSLV-CnSENT	21.3475***	3.9450	0.0085
M8: Ridge (not includes CnSENT)	21.1983***	3.9214	0.0087
M9: Ridge (includes CnSENT)	21.3475***	3.9450	0.0085

Notes: The results are based on 350-day rolling window and day_{t+1} . The Significance test for ΔR_{05}^2 is based on the p-value of MSPE-adjusted statistics. Hypothesis test relates to whether the mean squared forecast error (MSFE) of the HAR-RV model is less than or equal to the MSFE of other models [24]. * indicates significance at the 10% level. ** indicates significance at the 5% level. *** indicates significance at the 1% level.

C. Robustness check with a different rolling window

Using a different rolling window size may help to further determine whether the conclusion is robust or not. How to select the estimation window size has attracted the attention of many researchers as using a different window size may result in different empirical conclusions [25]. Thus, the method of [10], using a 250-day rolling window, is applied to redo the process and generate the estimation results again.

Table 5 is the summary of the out-of-sample results based on different window sizes. The empirical results are

Table 5. Out-of-sample estimation results based on a 250-day rolling window

Forecasting models	ΔR_{05}^2 (%)	MSPE-adjusted	p-value
M1: HAR-GVZ	7.5098***	0.7408	0.0047
M2: HAR-VXSLV	17.4009**	2.7343	0.0108
M3: HAR-CnSENT	0.5508**	0.0580	0.0155
M4: HAR-GVZ-VXSLV	21.3842***	3.2947	0.0076
M5: HAR-GVZ-CnSENT	7.9957***	0.7886	0.0036
M6: HAR-VXSLV-CnSENT	15.4677***	2.4764	0.0086
M7: HAR-GVZ-VXSLV-CnSENT	21.8969***	3.3401	0.0071
M8: Ridge (not includes CnSENT)	21.3842***	3.2947	0.0076
M9: Ridge (includes CnSENT)	21.8969***	3.3401	0.0071

Notes: The results are based on 250-day rolling window and day_{t+1} . The significance test for ΔR_{05}^2 is based on the p-value of MSPE-adjusted statistics. Hypothesis test relates to whether the mean squared forecast error (MSFE) of the HAR-RV model is less than or equal to the MSFE of other models [24]. * indicates significance at the 10% level. ** indicates significance at the 5% level. *** indicates significance at the 1% level.

5 Conclusions

This paper focuses on the volatility prediction of China's silver futures market. Specifically, the roles of CBOE gold ETF volatility index, and CBOE silver ETF volatility index are investigated, and Chinese investor sentiment is constructed based on the Chinese stock market from both the in-sample and out-of-sample perspectives.

The results reveal that GVZ and VXSLV can significantly improve the prediction accuracy of China silver futures. VXSLV plays a more important role than GVZ. Based on different types of HAR models, the results report that GVZ at the daily horizons and VXSLV at daily and monthly horizons have a negative impact on the volatility prediction. However, based on the in-sample and out-of-sample results, it cannot be concluded that CnSENT plays a crucial role in volatility forecasting, as the contribution of this variable is unstable. This may be due to the illiquidity in the Chinese silver futures market. Furthermore, although the ridge regression method is used to deal with the overfitting issue in the model, its forecasting performance is not better than all other HAR-type models.

revealing that GVZ and VXSLV can be utilized to predict the volatility of China silver futures by incorporating them into the benchmark model, respectively. Compared with GVZ, VXSLV has better predictive power. HAR-GVZ-VXSLV-CnSENT is the best model, and HAR-GVZ-VXSLV still maintains a good prediction performance. The role of the CnSENT in the volatility forecasting is still doubtful. Two types of ridge regression models cannot perform better than all other models. The results based on a 250-day rolling window are consistent with the conclusions based on a 350-day rolling window.

It would be interesting to further examine whether CnSENT can be utilized to assist in volatility forecasting of China silver futures by using other models or other indicators to construct a new Chinese investor sentiment. In addition, currently, many practitioners also take the influence of leverage effects and geopolitical risks into consideration when exploring the volatility prediction. Nevertheless, these issues are left for further study.

References

1. Christoffersen, P., 2014. The economic value of realized volatility: using high-frequency returns for option valuation. *J. Financ. Quant. Anal.* 49.
2. Clements, M.P., Galvão, A.B., Kim, J.H., 2008. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *J. Empir. Financ.* 15, 729–750.
3. Corsi, F., Fusari, N., Vecchia, D.L., 2013. Realizing smiles: options pricing with realized volatility. *J. Financ. Econ.* 107, 284–304.
4. Maheu, J.M., Mccurdy, T.H., 2011. Do high-frequency measures of volatility improve forecasts of return distributions? *Ssrn Electron. J.* 160, 69–76.

5. Seo, S.W., Kim, J.S., 2015. The information content of option-implied information for volatility forecasting with investor sentiment. *J. Bank. Financ.* 50, 106–120.
6. Wen, F., Xiao, J., Huang, C., Xia, X., 2018. Interaction between oil and US dollar exchange rate: nonlinear causality, time-varying influence and structural breaks in volatility. *Appl. Energy* 50, 319–334.
7. Noh, J. and Kim, T.H., 2006. Forecasting volatility of futures market: the S&P 500 and FTSE 100 futures using high frequency returns and implied volatility. *Applied Economics*, 38(4), pp.395-413.
8. Huang, J., Tan, N., & Zhong, M. (2014). Incorporating Overconfidence into Real Option Decision-Making Model of Metal Mineral Resources Mining Project 2014. DOI:10.1155/2014/232516.
9. Li, W., Cheng, Y. and Fang, Q., 2020. Forecast on silver futures linked with structural breaks and day-of-the-week effect. *The North American Journal of Economics and Finance*, p.101192.
10. Wei, Y., Liang, C., Li, Y., Zhang, X. and Wei, G., 2020. Can CBOE gold and silver implied volatility help to forecast gold futures volatility in China? Evidence based on HAR and Ridge regression models. *Finance Research Letters*, 35, p.101287.
11. Luo, X. and Ye, Z., 2015. Predicting volatility of the Shanghai silver futures market: What is the role of the US options market?. *Finance Research Letters*, 15, pp.68-77.
12. Yang, C., Gong, X. and Zhang, H., 2019. Volatility forecasting of crude oil futures: The role of investor sentiment and leverage effect. *Resources Policy*, 61, pp.548-563.
13. Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *J. Financ. Econometr.* 7 (2), 174–196.
14. Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2002. Modeling and forecasting realized volatility. *Econometrica* 71, 579–625.
15. Beer, F. and Zouaoui, M., 2013. Measuring stock market investor sentiment. *Journal of Applied Business Research (JABR)*, 29(1), pp.51-68.
16. Al Shalabi, L. and Shaaban, Z., 2006, May. Normalization as a preprocessing engine for data mining and the approach of preference matrix. In *2006 International conference on dependability of computer systems* (pp. 207-214). IEEE.
17. Zhu, Y., Tian, D. and Yan, F., 2020. Effectiveness of Entropy Weight Method in Decision-Making. *Mathematical Problems in Engineering*, 2020.
18. Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *Int. Econ. Rev.* 39 (4), 885–905.
19. Qiao, G., Teng, Y., Li, W. and Liu, W., 2019. Improving volatility forecasting based on Chinese volatility index information: Evidence from CSI 300 index and futures markets. *The North American Journal of Economics and Finance*, 49, pp.133-151.
20. Hawkins, D.M., 2004. The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1), pp.1-12.
21. Exterkate, P., Groenen, P.J., Heij, C., van Dijk, D., 2016. Nonlinear forecasting with many predictors using kernel ridge regression. *Int. J. Forecast.* 32 (3), 736–753.
22. [22].Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12 (1), 55–67.
23. [23].Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? *Rev. Financial Stud.* 21 (4), 1509–1531.
24. [24].Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *J. Econometr.* 138 (1), 291–311.
25. [25].Rossi, B. and Inoue, A., 2012. Out-of-sample forecast tests robust to the choice of window size. *Journal of Business & Economic Statistics*, 30(3), pp.432-453.