

Rebuild Normality in Chinese Financial Markets with Range-Based Bipower Variance

Lianqian Yin¹, Lixia Zheng¹ & Yang Zhong¹

¹ Finance Department of International Business School, Jinan University, Zhuhai, Guangdong Province, China

Correspondence: Lianqian Yin, Finance Department of International Business School, Jinan University, Qianshan road 206#, Zhuhai City, Guangdong Province, China. E-mail: lianqian@jnu.edu.cn

Received: August 29, 2014

Accepted: October 21, 2014

Online Published: November 22, 2014

doi:10.5539/ijbm.v9n12p252

URL: <http://dx.doi.org/10.5539/ijbm.v9n12p252>

Abstract

Realized range and bipower variance are two important improvements for increasingly popular high-frequency realized estimators in financial markets. This paper verifies a new type of estimator named realized range-based bipower variance and carries out its empirical research with high frequency data from Shanghai Composite Index (SHCI) and Shenzhen Synthesis Index (SZSI). The results show that 1) this estimator combines merits of both realized range and realized bipower variance. It is as efficient as realized range estimator and at the same time remains a consistent estimation of integrated variance of Chinese financial markets' fluctuation; 2) After standardized by realized range based bipower variance, the distributions of SHCI and SZSI's daily returns are neither skew nor with high kurtosis anymore. The fat tail and high peak of daily returns are basically eliminated by this high-frequency variance estimator, and the standardized distributions of these returns are nearly normal. 3) Comparative studies show that among four types of realized volatility estimators, the range based bipower variance is the best one to rebuild normality in Chinese financial markets. These findings mean when measuring volatility or fluctuations of financial assets, the usage of this new estimator will increase the performance of many financial practices like pricing or risk management. One feasible way to extend this paper is to consider co-estimators of related assets and detect their impacts to the dynamics of volatilities.

Keywords: volatility, range, realized bipower variance, Chinese financial markets, stock index

1. Introduction and Literature Review

Literatures of using range as the proxy of volatility are prolific. According to the definition of Chou (2005), range is the distance between the highest and the lowest prices of assets in some fixed sampling period. Mandelbrot (1963) uses range to test the long-run dependence characteristic of asset prices. Parkinson (1980) argues the natural logarithm of stock prices roughly follow Random Walk process with a constant diffusion parameter which equals the variation of returns. He also compared variations of the range estimator with variations of the traditional return estimator and found that the range estimator is five times more efficient than the return estimator. Beckers (1983) incorporates the information of close prices into the range estimator and adjusted the estimator by Parkinson (1980), his empirical study of 208 kinds of stocks and options proves that the performance of range estimator is much superior to the traditional return estimator. Intuition tells us that ranges contain more information than close prices. Wiggins (1991) shows that compared with the return estimator, the range estimator have a problem of downward bias. And its efficiency is damaged a lot by the outliers. Once outliers get removed from the sample, the range estimator remains high efficiency. Furthermore, Andersen and Bollerslev (2001) find that range estimator provides a higher R-square value than the traditional return estimator in Mincer-Zarnowitz (MZ) regression when using realized volatility as the proxy of true underlying volatility.

2. Method

2.1 Theory with High Frequency Data in Financial Markets

The standardization of the sample space interval is set to be one for convenience, and divide the sample interval $[t, t+1]$ on average for an interval of $i=1,2,3\dots M$ at the length of Δ . The sampling frequency of is $M=[1/\Delta]$. Within the period of Δ , the last price, the highest price and lowest price are defined sequentially as $C_{t,i} = P_{t-1+i\Delta}$, $H_{t,i} = \sup_{(i-1)\Delta < j < i\Delta} P_{t-1+j}$ and $L_{t,i} = \inf_{(i-1)\Delta < j < i\Delta} P_{t-1+j}$.

Andersen and Bollerslev (2002) defined realized variance as the sum of squared returns on financial assets:

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

Where $r_{t,i}^2$ is sum of squared return $r_{t,i}$ ($i = 1, 2, \dots, I$) calculated by the closing price $C_{t,i}$ of each interval. Christensen and Podolskij (2007) defined realized range as:

$$RR_t = \frac{1}{\lambda_{2,I}} \sum_{i=1}^I (\ln H_{t,i} - \ln L_{t,i})^2 \quad (2)$$

Where $H_{t,i}$ represents the highest price and $L_{t,i}$ represents lowest price of interval i . $\lambda_{r,I}$ is the r -order moment of range calculated by I observations of a standard Brownian motion during unit interval. Even though there is no specific formula for it, when the sampling frequency tends to infinity, the statistic will converge to the following expression:

$$\lambda_{r,I} \xrightarrow{\Delta \rightarrow 0, I \rightarrow \infty} \lambda_r = 4 / \sqrt{\pi} (1 - 4/2^r) 2^{r/2} \Gamma((r+1)/2) \zeta(r-1), r \geq 1$$

Where $\Gamma(x/2)$ and $\zeta(x)$ is respectively Gamma and Riemann's Zeta function. In particular, $\lambda_1 = \sqrt{8/\pi}$, $\lambda_2 = 4 \ln 2$ when $r=1,2$ according to Parkinson(1980). While Barndorff-Nielsen and Shephard (2004) defined bipower variation as :

$$RB_t = (\mu_t)^{-2} \sum_{i=1}^I |r_{t,i}| * |r_{t,i+1}| \quad (3)$$

Where $\mu_r = 2^{r/2} \Gamma((r+1)/2) / \Gamma(1/2)$, $r \in (0,2)$.

For obtaining the theoretical properties of the above three estimators, logarithmic price of assets $p(t) = \ln P_t$ is subjected to a continuous random process with limited jumps:

$$dp(t) = \mu(t)dt + \sigma(t)dB(t) + \kappa(t)dq(t) \quad (4)$$

Where $\mu(t)$ is a drifting process with continuous and locally bounded variance. $\sigma(t)$ is a random volatility process with strictly positive values and in line with Caglad natures, and $B(t)$ is a standard Brownian motion. $\kappa(t)$ is a random jump processes, and $q(t)$ indicates the occurring time of stochastic jumps with density $\lambda(t)$. Asymptotic features of realized variance were addressed by a series of papers on quadratic variance, such as ABDE (2001), ABL (2002) as well as Barndorff and Shephard (2005). Their researches concluded that formula (1) converges to the quadratic variance of the above procedure in probability as the sampling frequency M approaches infinity:

$$RV_t \xrightarrow{\Delta \rightarrow 0, p} \int_{t-1}^t \sigma^2(s)ds + \sum_{t-1 < s \leq t} \kappa^2(s) \quad (5)$$

Realized variance converges to the integral variance of random process assets' logarithmic price according to formula (5). If jumps exist, realized variance will not be a consistent estimator of integral variance. Further, Dijk and Martens (2007) pointed out that the variance of formula (2) is about 5 times smaller than formula (1). That is to say realized range is more efficient than realized variance. However, Christensen and Podolskij (2006) concluded that if assets' logarithmic prices subject to formula (4), formula (2) will in probability converge to:

$$RR_t \xrightarrow{\Delta \rightarrow 0, p} \int_{t-1}^t \sigma^2(s)ds + \frac{1}{\lambda_{2,I}} \sum_{t-1 < s \leq t} \kappa^2(s) \quad (6)$$

As a result, realized range isn't the robust estimator of financial markets' volatility in the presence of jumps even if it is effective. On the other hand, B-S and Shephard (2004) build a large sample theory for realized power volatility based on formula (4). Bipower variation of formula (3) converges to the continuous part variance of formula (4) in probability as the sampling frequency approaches infinity:

$$RB_t \xrightarrow{\Delta \rightarrow 0, p} \int_{t-1}^t \sigma^2(s)ds \quad (7)$$

According to formula (7), bipower variance is a consistent estimator of integral variance of asset prices and

robust to the presence of jumps and other noises.

2.2 Realized Range Based Volatility Estimator and Its Consistency

To improve efficiency of formula (4), realized range and bipower variation are combined to design a new kind of volatility estimation named range-based realized bipower variation based on the research of Christensen and Podolskij (2007). Firstly price range is employed instead of returns as the proxy of assets' volatility, and then bipower variation is established to keep the consistency. This estimator is called range-based realized bipower and it is more efficient than realized variance as well it keeps consistency of integrated variance of asset prices. Range-based realized bipower variance is defined as:

$$RBV_t = \left(\frac{1}{\lambda_{1,T}}\right)^2 \sum_{i=1}^{T-1} |\ln H_{t,i} - \ln L_{t,i}| |\ln H_{t,i+1} - \ln L_{t,i+1}| \tag{8}$$

Like the range estimator proposed by Christensen and Podolskij (2006) The above formal converges to integrated variance of formula (4) in probability as the sampling frequency approaches infinity:

$$RBV_t \xrightarrow{\Delta \rightarrow 0, p} \int_{t-1}^t \sigma^2(s) ds \tag{9}$$

Thus, compared with the other two improvements of realized variance, formula (8) is not only more effective but also a consistent estimator of financial markets prices' variance.

3. Results

The estimator based on high-frequency data will be seriously affected by the market micro-structure noises. Therefore how to select data frequency is very important in the performance of realized variance. ABDE (2001) proposes spare sampling technique to minimize the errors caused by micro-structure noises. Although this approach can reduce errors, it does not use the market information efficiently. Hence the optimal sampling frequency which can minimize the errors and also exhaust market information is needed to be selected. Bandi and Russell (2006) as well as ZMA (2006) use the optimal sampling frequency based on the minimization of the mean squared errors. They compared this approach with equidistant sampling intervals which contain dependency micro-structure noises. They find that five minutes' sampling frequency is the most optimum to avoid market microstructure noise. Therefore the five minutes' sampling frequency is commonly used in high frequency data researches. This paper selects 5 minute frequency data of Shanghai Composite Index (SHCI) and Shenzhen Synthesis Index (SZSI). The sample period is form June 02, 2010 to May 26, 2013, a total of 720 trading days, an interval of 4 hours a day with 48 intervals. $I = 48$, $T = 720$ mean for each index the data sets contain 34560 intervals and each interval data has its own opening price, the highest price, the lowest price as well as the closing price. These are used to construct estimators in equation (1), (2), (3) and (8).

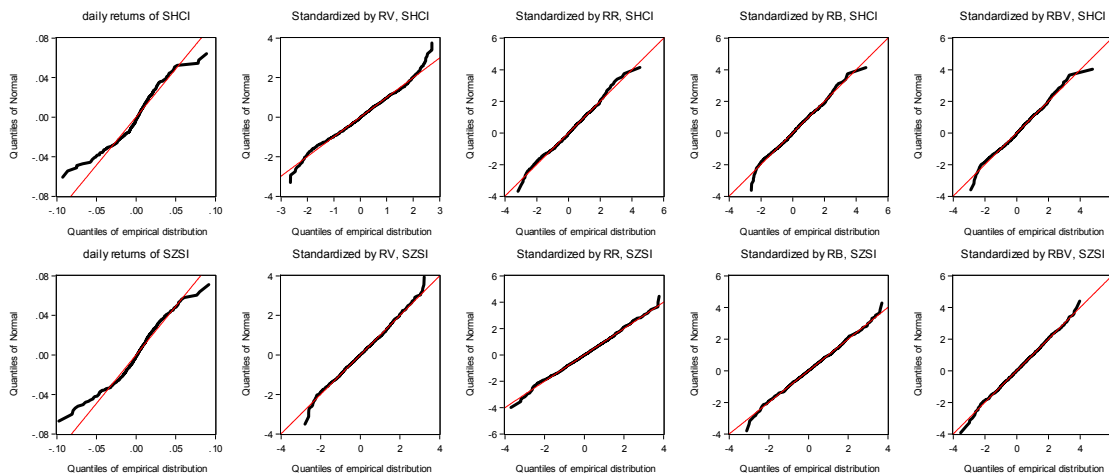


Figure 1. QQ plots of series in Table 1 against with normal distributions

In the evaluation and comparison of volatility estimation, according to Andersen and Bollerslev (1998) realized variance usually acts as basic evaluation criteria to evaluate the estimation accuracy of models with different parameters. However Dijk and Martens(2007) suppose that standardize returns' unconditional distributions are always the basic evaluation criteria in comparison of different kinds of realized variance estimators. Generally speaking, realized variance should eliminate the non-normal characteristics of original returns as far as possible and turn the standardize returns' unconditional distribution closely enough to normal distribution. Figure 1 presents QQ plots of series in Table 1 against with normal distributions of two main indexes, SHCI and SZSI, while table 1 presents their descriptive statistics of unconditional raw and standardized returns.

Table 1. Descriptive statistics of unconditional raw and standardized returns

	Mean	S.D.	Skewness	Kurtosis	J-B statistics	Q(12)
SHCI						
r	1.673	19.553	-0.576	6.077	323.5(0.00)	26.342
r/sqrt(rv)	0.227	1.103	-0.148	2.582	2.582(0.02)	31.927
r/sqrt(rr)	0.230	1.225	-0.143	2.907	2.726(0.26)	29.717
r/sqrt(rb)	0.256	1.212	0.019	2.818	1.035(0.60)	33.739
r/sqrt(rbv)	0.225	1.193	-0.074	2.956	0.712(0.70)	30.381
SZSI						
r	2.060	21.604	-0.416	5.279	176.3(0.00)	18.903
r/sqrt(rv)	0.215	1.164	-0.061	2.663	3.847(0.15)	27.272
r/sqrt(rr)	0.230	1.317	-0.080	2.827	1.658(0.44)	24.580
r/sqrt(rb)	0.240	1.262	-0.002	2.768	1.616(0.45)	28.478
r/sqrt(rbv)	0.240	1.302	-0.005	2.832	0.849(0.65)	24.253

Note. The mean and standard deviation in table 1 is 1000 times of the original value, and J-B is Jarque-Bera statistics and in parentheses is P value.

Some conclusions can be drawn from table 1: (1) daily returns of SHCI and SZSI are left-skewed with high peak and fat tails. These are consistent with direct effect shown in figure 1. After the standardized by variance, the returns are approximately 0.23 and the standard deviations are around 1.2; (2) After standardized by realized range-based bipower variance, the returns' unconditional distribution of SHCI has kurtosis very close to 3(2.956), which is closely enough to normal distribution; (3) In eliminating the skewness of the original daily returns, realized bipower variation RB and realized range-based bipower variance RBV are much better than realized variance RV and realized range RR no matter for SHCI or SZSI. The skewness of unconditional returns standardized by RB and RBV are significantly smaller than those with RV and RR. In addition, the skewness of RB and RBV of SZSI's standardized returns are almost 0(-0.002 and -0.005) which means their skewness is totally eliminated; (4) The Jarque-Bera statistics are used to judge whether a distribution is close enough to normal distribution. The Jarque-Bera statistics of original daily returns of SHCI or SZSI reject the null hypothesis that they subject to normal distribution with the P value of 0. The P value of Jarque-Bera statistics of RBV standardized returns are the highest in both two financial markets which indicates that RBV has better comprehensive properties of "Effective & consistent" than other realized volatilities; (5) Ljung-Box 12-order lag statistic in the table also shows that while dynamics in volatility get rid of returns' high kurtosis and skewness, but the dynamics of raw returns are all remained for further modeling. Appropriate structural models like ARFIMA are still useful in capturing returns' dynamics.

4. Discussions

Volatility is one of the core problems in many financial practices but due to the unobservable it is often modeled as latent variable and its dynamics is regarded as one important source of high kurtosis and skewness of asset returns' unconditional distributions. Rebuild the normality of asset returns' unconditional distributions is a very important to solve many financial problems concisely.

In this paper, the range-based realized variance is constructed with sparse sampling techniques to filter jumps and noises in high frequency data and remain its high effectiveness in estimation of assets volatilities. Then the empirical study with high frequency data from Chinese Shanghai and Shenzhen markets chooses two main indexes, Shanghai Composite Index (SHCI) and Shenzhen Synthesis Index (SZSI), as main study objects. The results show that range-based realized bipower variation does combine the merits of bipower variation and

realized range and it can eliminate the non-normal characteristics of original returns and turn the standardized returns' unconditional distribution closely enough to normal distribution.

These findings mean when measuring volatility or fluctuations of financial assets, the usage of our new estimator will increase the performance of many financial practices like pricing or risk management. With the development of the information technology, high frequency data are more and more available. Developing an accurate and robust estimation of assets volatility is more and more important. One feasible way to extend this paper is to consider co-estimators of related assets and detect their impacts to the dynamics of volatilities.

Acknowledgments

While remaining responsible for any errors in this paper, the authors would like to thank comments from the 18th International Conference on Macroeconomics Analysis and International Finance. We are also particularly grateful to "Application of Open Source R QuantLib in Financial Engineering Study" program of Jinan University and Social Science Association of Zhuhai city in Guangdong province. Thanks Luoming Fan and Qiuli Gao for their personal assistance in manuscript preparation.

References

- Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57(3), 1047–1092. <http://onlinelibrary.wiley.com/doi/10.1111/1540-6261.00454/abstract>
- Andersen, T. G., Bollerslev, T. (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association*, 96(453), 42–55.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1), 43–76. [http://dx.doi.org/10.1016/S0304-405X\(01\)00055-1](http://dx.doi.org/10.1016/S0304-405X(01)00055-1)
- Bandi, F. M., & Russell, J. R. (2006). Separating market microstructure noise from volatility. *Journal of Financial Economics*, 79(3), 655–692.
- Barndorff-Nielsen & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B* 64(2), 253–280. <http://dx.doi.org/10.1111/1467-9868.00336>
- Barndorff-Nielsen, & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1), 1–48. <http://dx.doi.org/10.1093/jffinec/nbh001>
- Barndorff-Nielsen, & Shephard, N. (2005). How accurate is the asymptotic approximation to the distribution of realized variance? In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and Inference for Econometrics Models* (pp. 306–331). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511614491.014>
- Barndorff-Nielsen, S. E., Graversen, J. J., Podolskij, M., & Shephard, N. (2006). A central limit theorem for realized power and bipower variations of continuous semimartingales. In Y. Kabanov, R. Lipster and J. Stoyanov (Eds.), *From Stochastic Calculus to Mathematical Finance* (pp. 33–68). The Shiryayev Festschrift, Springer, Berlin. <http://dx.doi.org/10.1007/978-3-540-30788-43>
- Beckers, S. (1983). Variances of security price returns based on high, low, and closing prices. *The Journal of Business*, 56(1), 97–112. <http://dx.doi.org/10.1086/296188>
- Chou, R. (2005). Forecasting financial volatilities with extreme values: the Conditional Auto Regressive Range (CARR) Model. *Journal of Money Credit and Banking*, 37(3), 561–582. <http://dx.doi.org/10.1353/mcb.2005.0027>
- Christensen, & Podolskij. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2), 323–349. <http://dx.doi.org/10.1016/j.jeconom.2006.06.012>
- Dijk & Martens. (2007). Measuring volatility with the realized range. *Journal of Econometrics*, 138, 118–207.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394–419. <http://dx.doi.org/10.1086/294632>
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53, 61–65. <http://dx.doi.org/10.1086/296071>
- Wiggins, J. B. (1991). Empirical tests of the bias and efficiency of the extreme-value variance estimator of

common stocks. *The Journal of Business*, 64(3), 417–432. [http://dx.doi.org/10.1016/0167-6687\(93\)91040-2](http://dx.doi.org/10.1016/0167-6687(93)91040-2)
Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association*, 100(472), 1394–1411. <http://dx.doi.org/10.1198/016214505000000169>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).