Investigation on two alternative model-free realized volatility estimators

Zhixin Kang
The University of North Carolina at Pembroke

Rami Cooper Maysami
The University of North Carolina at Pembroke

Aiwu Zhao
Skidmore College

ABSTRACT

In this paper the authors compare the magnitudes of the realized volatility estimators obtained from the two methods using real intra-daily high frequency return data. These two realized volatility estimators are non-parametric, model-free approaches. One estimator is the conventional realized volatility (RV) estimator by aggregating the squared returns extracted from the intra-daily high frequency return data set in each of the trading days, and the other estimator is the two-scale realized volatility estimator (TSRV) obtained using sub-grids sampling, averaging, and bias correction in intra-daily returns data in each trading day. Further, the statistical properties of these two realized volatility estimators are explored. The results show that, depending on different sub-intervals designated in a trading day, the RV estimators exhibit very different magnitudes. On the other hand, the TSRV estimator shows very stable magnitudes when the numbers of the sub-grids are big enough. In addition, these two realized volatility estimators exhibit a fractionally integration, or so-called long memory property. The results suggest that the length of a subinterval for a RV estimator and the number of the sub-grids for a TSRV estimator impose remarkable impacts on the estimation of daily realized volatility when intra-daily high frequency return data are used.

Key Words: Realized return volatility, intradaily high frequency data, GPH estimator, long memory, microstructure noise

INTRODUCTION

Return volatility, gauging the fluctuation of a financial asset's returns, is widely used in portfolio construction, option's pricing and trading, volatility-related derivatives' trading, and risk management. As such, researchers have made significant efforts in modeling the return volatility. Many different methodologies and theories have been developed in the past decades. Using low frequency daily return data, Parkinson (1980), Garman and Klass (1980), Rogers *et al*. (1994), Yang and Zhang (2000), among many other researchers, developed the methodologies for historical volatility estimation; Engle (1982), Bollerslov (1986), Nelson (1991), Baillie *et al.* (1996), Bollerslev and Mikkelsson (1996), among many others, have made significant contributions to the GARCH-class models; Taylor (1982, 1986), Ghysels *et al.* (1996), Shephard (1996), Ruiz (1994), Danielsson (1994), Andersen and Sørensen (1996), Kim *et al.* (1998), Jacquier *et al*. (1994), among many others, have significantly studied the stochastic volatility models. All these different approaches in estimating the daily return volatility employ the interday return information. Therefore, the trading dynamics during a trading day are not incorporated in these methods.

Since the late 90s, new methodologies have been advocated to estimate the daily return volatility with the availability of intraday high frequency return data. As non-parametric and model-free approaches, these methods allow one to estimate the daily return volatility by exploiting the rich information contained in the intraday return data. Andersen *et al*. (1997a, 1997b, 2001, 2003) propose to use the aggregated squared returns obtained in evenly spaced short intervals within a certain trading day to approximate the daily return volatility. Therefore, the measure of a daily return volatility based on the intraday high frequency data is related to these subintervals in a trading day. Zhang *et al.* (2005) propose a two-scale realized return volatility estimator to measure the daily volatility. This method, instead of using the returns in the subintervals within a trading day, partitions the intraday data into a certain number of sub-grids, and this is the first scale construction in the estimator. In addition, the tick-by-tick returns in the intraday data are considered in the estimator, and this is the second scale in the estimator. The daily volatility estimated in this way is closely related to the number of the sub-grids within a certain trading day.

The above-mentioned realized return volatility estimators using intraday data have imposed significant impacts on return volatility estimation and forecasting (Andersen *et al*., 1998; Aït-Sahalia *et al*., 2008), and volatility-related trading practices as well. Motivated to pursue a better understanding of these two different realized return volatility estimators, we aim to complement the existing literature by exploring their finite-sample statistical attributes. Specifically, the authors denote the realized volatility estimator proposed by Andersen *et al*. (2001, 2003) as RV, and study the impacts on the dynamics of RV estimates and on their statistical properties by using different subintervals in the estimation; the authors denote the realized volatility estimator developed by Zhang *et al*. (2005) as TSRV, and explore the impacts on the dynamics of TSRV estimates and on their statistical properties by employing different sub-grids in the estimation.

In this paper, by using the intraday high frequency data of NASDAQ 100 tracking stock in one year period from 01/02/2003 to 12/31/2003, the authors estimate the daily RVs using 30-second, 1-minute, 5-minute, 10-minute, 15-minute, 30-minute subintervals in each trading day; we estimate the daily TSRVs with 50, 100, 150, 200, 250, and 300 sub-grids in each trading day. Based on the real intraday data but different sampling scenarios we investigate the characteristics of these two different estimators by (1) comparing the dynamics in the estimates of RV and TSRV, respectively, and (2) analyzing their finite-sample statistical attributes. The empirical results in this paper confirm the existing literature that RV is significantly impacted by the choice of the subintervals, and TSRV is significantly

impacted only when the sub-grids are small. Furthermore, the finite-sample statistical properties of RV and TSRV are closely related to the different sampling scenarios.

The remainder of this paper is structured as follow. In Section 2, the authors briefly introduce the RV and TSRV, and their estimation procedures as well. Section 3 presents the data. In Section 4, the empirical results are documented and analyzed. Section 5 concludes.

RV and TSRV

Both RV and TSRV are governed by the assumption that the logarithms of a stock's trading prices follow a continuous semi-martingale. That is, the log price process y_t is modeled by the following stochastic differential equation:

$$dy_t = \mu_t \, dt + \sigma_t \, dW_t \tag{1}$$

where μ_t is the drift at time t, σ_t is the volatility of return process of y_t at time t, W_t is a standard Brownian process. Given the dynamics of σ_t at each time t in a time period $[0, T]$, one is interested in calculating the integrated variance $\int_0^T \sigma_t^2 dt$. When $T = 1$, one calculates the daily integrated variance. However, the intraday trading prices are not continuous. Therefore, one can only find the approximation for the integrated variance $\int_0^T \sigma_t^2 dt$. Both RV^2 and $TSRV^2$ are developed to serve as an estimator of $\int_0^T \sigma_t^2 dt$ with different approaches. In this study, the authors follow the literature to focus on the RV and TSRV, which are the square root of the estimated integrated variance.

RV Estimator

A daily RV is calculated by splitting a trading day into equally spaced subintervals, and then by aggregating the squared returns in these subintervals. It is proved by Andersen *et al*. (2001) and Barndorff-Nielsen and Shephard (2002) that RV is a consistent and robust estimator of the true return volatility as the subintervals approach to zero. Therefore in the estimating process, a tiny subinterval is specified to obtain the RV. The daily RV using the intraday return data is calculated as:

$$RV^2 = \sum_{j=2}^{M-1} (y_{j+1} - y_j)^2, \; j = 2,..., \; M-1 \tag{2}$$

where y_j is the log price at time j in a certain trading day, M the total number of the subintervals designated in a certain trading day. As shown in formula (2), RV is calculated using the extracted log prices at time points $j = 1, 2, 3, …, M$. Therefore this method makes use of the M observations in the intraday data. In the U.S., the stock market operates from 9:30am to 16:00pm with 6.5 hours of trading time. The authors specify the length of the subintervals in each trading day as 30 seconds, 1 minute, 5 minutes, 10 minutes, 15 minutes, and 30 minutes. Therefore, the corresponding M is 780, 390, 78, 39, 26, and 13, respectively. As is shown in formula (2), the RV is directly related to the choice of M, which is determined by the length of subintervals.

TSRV Estimator

The RV estimator only exploits a small number of the return observations in an intraday high frequency data. In contrast, the TSRV estimator employs all the observations in

an intraday data. According to Zhang *et al.* (2005), this method incorporates sub-sampling, averaging, and bias correction in the computation of the daily realized return volatility. Suppose $T = \{t_1,...,t_n\}$ is the times of the observed log prices in a certain trading day. Then T is partitioned into *K* non-overlapping sub-grids with equal number of observations. The *kth* (*k* = 1, 2,…, *K*) sub-grid extracts the observations from the whole intraday data with following times attached:

$T_k = \{t_{k-1}, t_{k-1+K},...,t_{k-1+n_k K}\}$, where n_k is the largest integer so that the $(t_{k-1+n_k K})$*th* observation is included in T_k. The *TSRV* is calculated as follow:

$$TSRV^2 = \frac{1}{K}\sum_{k=1}^{K}\sum_{t_i, t_{i+1} \in T_k}(y_{t_{i+1}} - y_{t_i})^2 - \frac{\bar{n}}{n}\sum_{t_j, t_{j+1} \in T}(y_{t_{j+1}} - y_{t_j})^2 \qquad (3)$$

where y_j is log price process, n is total observations in a intraday data, and $\bar{n} = \frac{n - K + 1}{K}$.

From the formula (3), it is shown that the estimates of the TSRV are related to the value of *K*, the number of the sub-grids partitioned in an intraday data. The first part in the formula (3) is to sample and average the squared returns across all the *K* sub-grids, and the second part takes a portion of aggregated squared returns obtained from the *total* observations. According to Zhang *et al.* (2005), the second part is introduced in the TSRV estimation to correct the microstructure noise in intraday high frequency data. The extent of the correction of the bias partly depends on $\bar{n}$, which is determined by *K* as well. In our study, in order to investigate the impacts of the *K* on the dynamics of the estimated TSRV, we specify *K* as 50, 100, 150, 200, 250, and 300, respectively.

Regarding the asymptotic properties of RV and TSRV, interested readers may refer to Andersen *et al.* (2001, 2003), Barndorff-Nielsen and Shephard (2002), and Zhang *et al.* (2005). The authors' study focus on the RV and TSRV estimates under the finite-sample circumstances with real data.

DATA

The study in this paper aims at exploring the properties of the daily RV and TSRV estimates obtained from varying subintervals and sub-grids in each trading day, respectively. This requires the intraday high frequency data be frequent enough. The NASDAQ 100 index tracking stock (QQQ) was intensively traded in each trading day in the market. Therefore the authors choose in this study QQQ's intraday high frequency data in a one-year period from 01/02/2003 to 12/31/2003 with 252 trading days. The data set is TAQ consolidated quote of QQQ downloaded from Wharton Research Data Services (WRDS). In each trading day, the trading time spans from 9:30am to 16:00pm.

Prior to the estimation of RV and TSRV, a cleaning procedure is applied to the original intra-daily high frequency data to remove the typos and unusual log price jumps.

EMPIRICAL RESULTS AND ANALYSIS

Dynamics of the RV and TSRV Estimates

Using the 30-second, 1-minute, 5-minute, 10-minute, 15-minute, and 30-minute subintervals, as explained in Section 2, we obtain the daily RV estimates. Figure 1 is the time plots of these RVs. As is shown in the figure, the estimated RVs exhibit a dynamic pattern in their magnitudes. However, among these RV estimators obtained from the different

subintervals, the overall evolutions in their paths in the one-year horizon are quite similar. The remarkable peaks and troughs almost appear at the same trading day across the different RV estimates. Furthermore, as the length of a subinterval becomes shorter, the fluctuation of the corresponding RV estimates becomes more intensive. It is observed that the RV estimates with the 30-second subinterval exhibits the highest extent in fluctuation as indicated in Figure 1 (Appendix)

Employing 50, 100, 150, 200, 250, and 300, respectively, we obtain the TSRV estimates using the intraday data. Figure 2 plots these estimates as indicated in Figure 2 (Appendix)

In Figure 2, it appears that the TSRVs exhibit quite mild change in the magnitudes, especially when the sub-grids are great 100. Moreover, as the sub-grids are greater than 200, the TSRV estimates in each trading day are almost the same for different sub-grids. Similar to the RV estimates, the TSRVs obtained from different sub-grids exhibit similar patterns of the ups and downs in the one-year trading horizon. In addition, the peaks and troughs appear at almost the same trading day.

The Statistical Properties of RV and TSRV

As a preliminary analysis of the RV and TSRV estimates, the statistical summaries are listed in Table 1 and Table 2, respectively. In addition, we adopt a standard statistical measurement of variation, the coefficient of variation (CV), to measure the extent of the dispersion existed in each of the RV and TSRV estimates. The CV is defined as:

$$Coefficient \; of \; Variation \; (CV) = 100 \times \frac{s}{\bar{x}} \qquad (4)$$

where s is the sample standard deviation (volatility of volatility) of the volatility series, and $\bar{x}$ is the sample mean of either RV or TSRV estimates as indicated in Tables 1 and 2 (Appendix)

As is shown in the two tables, the sample means of the RV estimates decrease as the length of a subinterval increases. On the other hand, the sample means of the TSRV estimates remain stable with the different sub-grids. It appears that when sub-grids are at least 100, the TSRVs present small changes in both kurtosis and skewness. The coefficients of variation associated with the TSRVs are smaller and more stable than those of RVs. For the RV estimates, when the length of a subinterval increases, the standard deviation decreases at first, and reaches the lowest at the 5-minute subinterval, then increases again. Similar pattern is observed in regard to the coefficients of the variation. Our findings confirm Andersen et al. (2001) that a 5-minute may be a best choice for a subinterval in the RV estimation. Regarding the TSRV estimates, the standard deviations only exhibit slight increases with the larger sub-grids.

To investigate the distributional dynamics of the RVs and TSRVs associated with the different sampling scenarios, we plot the kernel density curves for these two return volatility estimates in Figure 3 and Figure 4, respectively (Appendix).

As is shown in Figure 3, the kernel density curves of the RV estimates are right skewed with the exception that the curve associated with a 5-minute subinterval is close to a normal distribution. In contrast, it appears from the Figure 4 that the differences between the TSRVs' kernel density curves are not significant, especially when the sub-grids are at least 100. Moreover, the kernel density curves of the TSRVs are quite close to normal distributions (Appendix).

Long Memory Property of the RV and TSRV

One stylized fact documented in the literature is the long memory property existed in the return volatility in many financial assets. See Andersen *et al*. (2001, 2003). Furthermore, Lo (1991), Robinson (1991), Ding *et al.* (1993), Baillie *et al.* (1996), and Lobato and Savin (1998) explore the theoretical aspects of the long memory. In our study, we focus on exploring the long memory behaviors in both RVs and TSRVs under the different sampling scenarios, as explained in the previous sections. Specifically we estimate the sample autocorrelation function (ACF) associated with each of the RV and TSRV estimates. As an illustrative purpose, we choose in the Figure 5 and Figure 6 (Appendix) to present the plots of the ACFs for the RV with a 5-minute subinterval and the TSRV with 250 sub-grids, respectively (Other ACF plots exhibit similar patterns as those in these two figures, and they are available upon request). As can be seen from the plots, the estimated sample autocorrelation function of each of the volatility series fades out very slowly over a distant horizon, and exhibits strong and persistent correlations. This visual analysis presents some evidences that the RV and TSRV estimators could possess the so-called long memory, or fractionally integrated pattern, as is described in Andersen *et al*. (2001). In this case, the time series only can be differenced using a fractional differencing parameter, *d*, which is in the range of (0, 1). To test the hypothesis that both RV and TSRV estimates don't follow a long memory process, we implement an *R/S* test proposed by Mandelbrot (1972) and extended by Lo (1991) to each of the RV and TSRV estimates. In this test, we set the bandwidth of the cross variance as 10, and the significance level is specified as 5%. The critical value at 5% significance level is 1.747. Table 3 (Appendix) lists the *R/S* statistics of the RV and TSRV estimates, respectively.

As is shown in the Table 3, all the *R/S* statistics obtained for RV and TSRV estimates are greater than the critical value at the 5% significance level. The hypothesis that there is no long memory in each of the volatility series is rejected. Therefore, the test suggests that all of the RV and TSRV estimates follow a long memory, regardless of the two different estimation methods. Further, the authors estimate the fractional differencing parameter *d* for each of RV and TSRV estimates by adopting a GPH estimator proposed by Geweke and Porter-Hudak (1983). In calculating the GPH estimates, the Fourier frequency *m* is set to be *T*/2, where *T* is the total observations in the RV and TSRV series. Table 4 (Appendix) lists the estimation results for the $\hat{d}$'s.

The GPH estimates obtained from the RV and TSRV series reveal that all of the fractional differencing parameters are in the range of (0, 1). The GPH estimates $\hat{d}$'s for RVs are in a wide range, from 0.220 to 0.519. As the length of a subinterval increases, the fractional differencing parameter decreases. As indicated in the literature, when a fractional differencing parameter *d* is in the range of 0.5 and 1, the volatility series is non-stationary. Regarding TSRVs, when the number of the sub-grids is less than 150, the resulted GPH estimates are less than 0.5, implying a stationary pattern. When the number of the sub-grids is at least 200, the corresponding $\hat{d}$'s are greater than 0.5, indicating that the TSRVs follow a non-stationary pattern as the sub-grids are large. Overall, the fact that all of the estimated fractional differencing parameters are less than 1 confirms the conclusion made from the *R/S* long memory hypothesis test. Therefore, both RV and TSRV follow a long memory process regardless of the two different estimation methodologies.

CONCLUSIONS

In this paper, the authors complement the existing literature by exploring the dynamics in the daily estimates of the two influential return realized volatilities obtained using intraday high frequency, as well as their finite-sample properties. We employ the intraday data of the actively traded NASDAQ 100 tracking stock in our empirical study. Our study focuses on the realized volatility estimators with two different estimation approaches. The daily return realized volatility proposed by Ansersen *et al*. (2001, 2003), sums up the squared returns extracted in equally spaced subintervals within a certain trading day. It is denoted as RV in our study. The daily two-scale realized volatility developed by Zhang *et al*. (2005) is constructed by combining sub-sampling, averaging and bias correction in the intraday data within a certain trading day. It is denoted as TSRV in our study. Our empirical results show that the selection in the length of a subinterval in the RV estimation procedure has significant impacts on the magnitudes of the estimated RVs. In contrast, the number of the sub-grids adopted in the TSRV estimation process imposes mild effects on the magnitude of the estimated TSRV as the sub-grids are at least 100. Moreover, TSRV presents slight differences in their distribution associated with different sub-grids. And RV's distributions are more sensitive to the different subintervals. However, the RV is very close to a normal distribution as the subinterval is 5 minutes. These findings are in line with the existing literature (Andersen *et al*., 2001, 2003) even though the financial asset we selected in our study is different. In addition, the coefficients of variation (CV) indicate that the volatility of volatility embedded in the RV estimates is significantly impacted as the length of a subinterval is beyond 10 minutes. The impacts on the volatility of volatility in TSRV by the number of the sub-grids appear to be mild. Finally, our results confirm the existing literature in that both RV and TSRV follow a long memory process. However, we find that fractional differencing parameters are highly impacted by the choice of the subintervals in the RV and the choice of sub-grids in the TSRV. Specifically, the TSRVs turn out to follow a non-stationary fractional integration as the sub-grids are at least 200.

The empirical results obtained from the real intraday data may add some insights into the following aspects about the estimation of a return volatility using intraday high frequency data: (1) choosing a subinterval can have significant impacts on the estimates and distributions of the realized volatility proposed by Andersen *et al*. (2001, 2003), (2) choosing the number of sub-grids can have significant impacts on the behavior of the fractional integration in the TSRV estimator proposed by Zhang *et al*. (2005). The implication to researchers and practitioners is that the finite-sample behaviors of the return volatility obtained from intraday high frequency data under different sampling scenarios can remarkably vary when one use the alternative volatility estimation methods. Therefore, it is important to take account of these factors in the volatility-related modeling, trading, and forecasting.

APPENDIX

Table 1 The statistical summary of the RV estimates

	30-Second	1-Minute	5-Minute	10-Minute	15-Minute	30-Minute
Standard Deviation	0.00393	0.00348	0.00326	0.00336	0.00347	0.00382
Mean	0.01571	0.01428	0.01315	0.01264	0.01232	0.01159
Kurtosis	0.31966	0.61143	2.32769	4.23126	3.76897	0.36140
Skewness	0.61579	0.59184	0.77268	1.15160	1.04215	0.65356
Range	0.02036	0.02039	0.02463	0.02702	0.02832	0.02071
Coefficient of Variation	25.0075	24.3796	24.7650	26.6154	28.1793	32.9625

Table 2 The statistical summary of the TSRV estimates

	$K = 50$	$K = 100$	$K = 150$	$K = 200$	$K = 250$	$K = 300$
Standard Deviation	0.00276	0.00279	0.00284	0.00288	0.00292	0.00294
Mean	0.01171	0.01183	0.01207	0.01224	0.01238	0.01249
Kurtosis	2.89360	1.22739	1.03331	1.01424	1.02190	1.04941
Skewness	0.88800	0.49166	0.42989	0.43546	0.44581	0.46252
Range	0.02068	0.01826	0.01737	0.01774	0.01797	0.01810
Coefficient of Variation	23.6142	23.5897	23.5452	23.5516	23.5512	23.5817

Table 3 The estimated R/S statistics for RV and TSRV estimates

	RV					
	30-Second	1-Minute	5-Minute	10-Minute	15-Minute	30-Minute
S/R	2.012	1.950	1.927	1.925	1.852	1.930
	TSRV					
	$k = 50$	$k = 100$	$k = 150$	$k = 200$	$k = 250$	$k = 300$
S/R	1.807	1.878	1.902	1.912	1.920	1.924

Note: Bandwidth q = 10, Significance level = 5%, Critical value = 1.747

Table 4 Fractional differencing parameter: GPH estimators of the RV and TSRV series

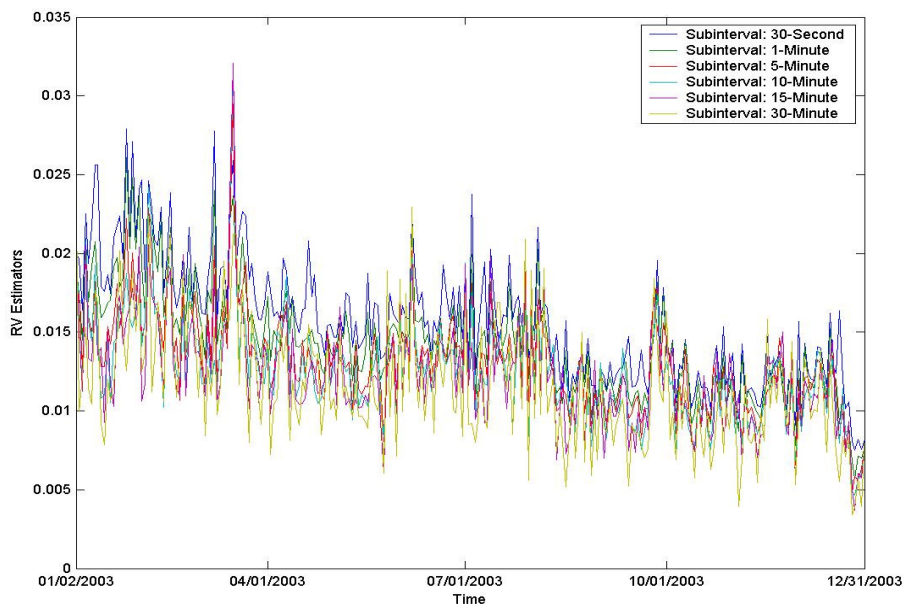| **RV (Fourier Frequency: $m = T/2$)**						
	30-Second	1-Minute	5-Minute	10-Minute	15-Minute	30-Minute
$\hat{d}$	0.519	0.477	0.439	0.389	0.288	0.220
Standard Error	(0.075)	(0.064)	(0.065)	(0.056)	(0.076)	(0.062)
TSRV (Fourier Frequency: $m = T/2$)						
	$k = 50$	$k = 100$	$k = 150$	$k = 200$	$k = 250$	$k = 300$
$\hat{d}$	0.374	0.467	0.497	0.515	0.520	0.521
Standard Error	(0.067)	(0.066)	(0.065)	(0.065)	(0.063)	(0.062)

Figure 1 The time plots of the estimated RVs with different subintervals
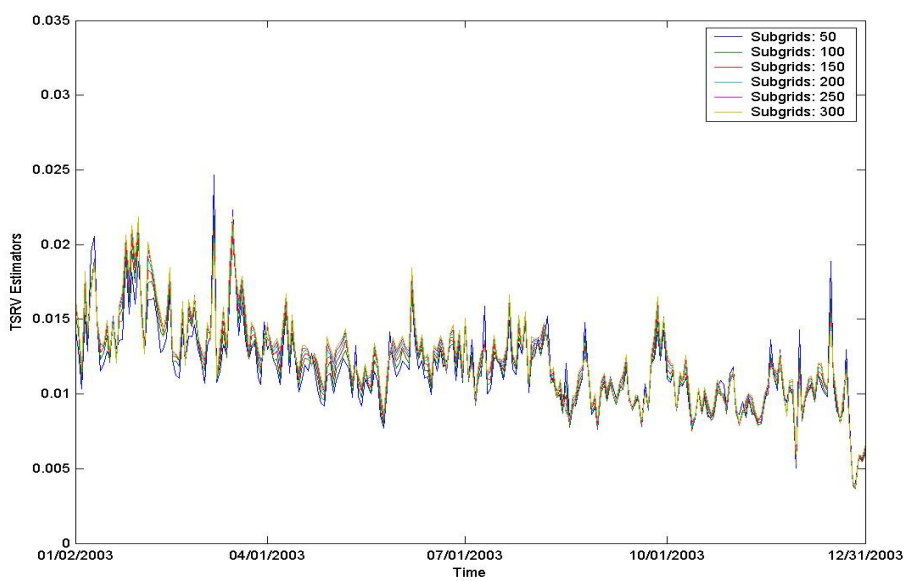

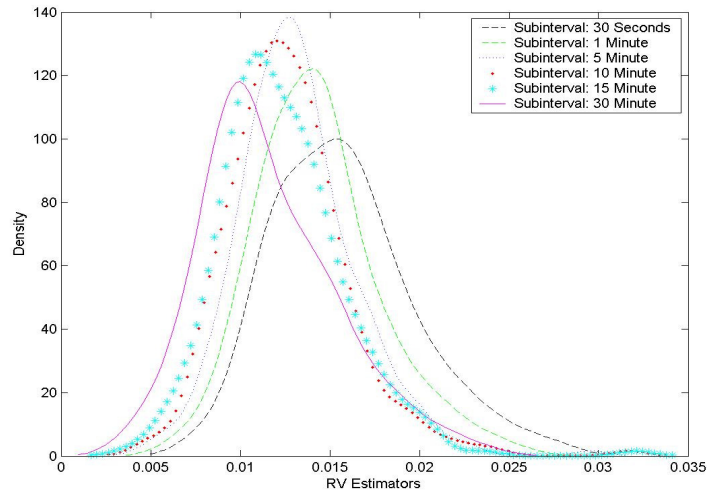**Figure 2** The time plots of the estimated TSRVs with different sub-grids

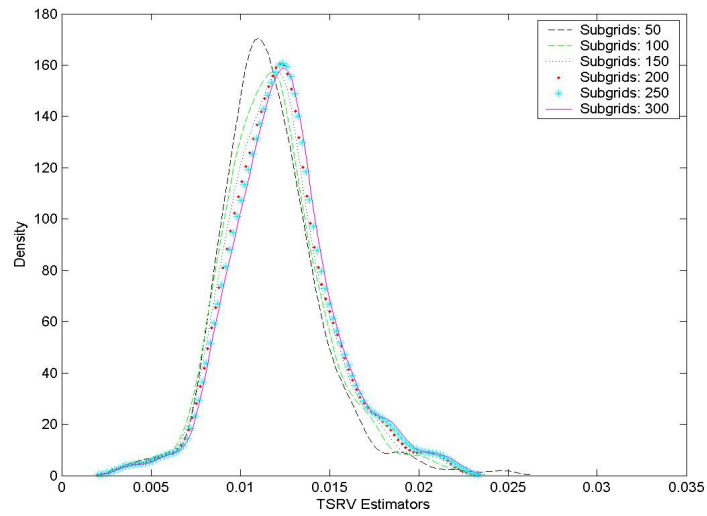**Figure 3** The kernel density curves of the RV estimates with different subintervals

Figure 4 The kernel density curves of the TSRV estimates with different sub-grids
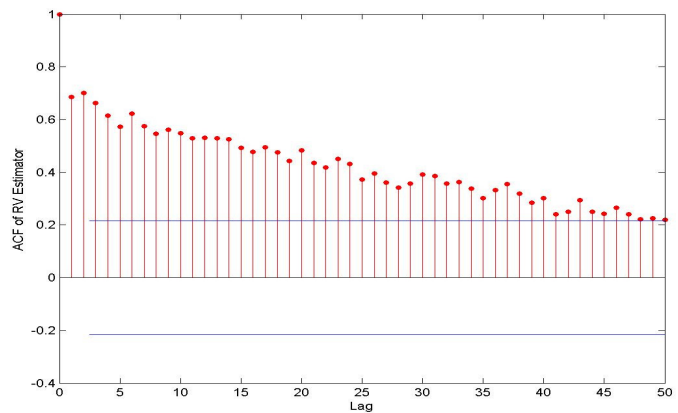
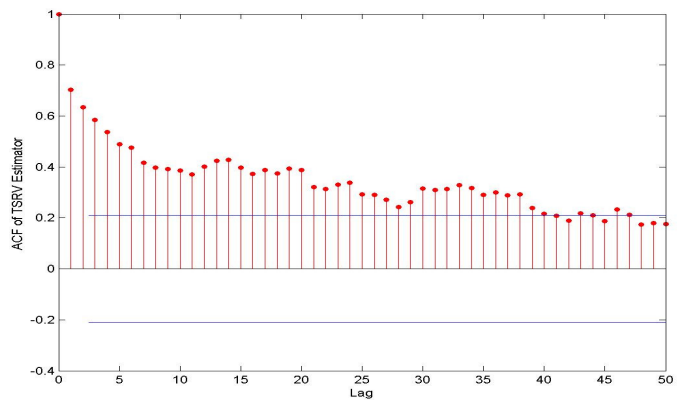**Figure 5** The ACF plot of the RV estimates with a 5-minute subinterval

Figure 6 The ACF plot of the TSRV estimates with 250 sub-grids

REFERENCES

Aït-Sahalia, Y. and L. Mancini. (2008). Out of Sample Forecasts of Quadratic Variation. Journal of Econometrics, 147, 17-33.

Andersen, T. G. and B. Sørensen. (1996). GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study. *Journal of Business and Economics Statistics*, 14, 329-52.

Andersen, T. G. and T. Bollerslev. (1997a). Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-run in High Frequency Returns. *Journal of Finance*, 52, 975-1005.

Andersen, T. G. and T. Bollerslev. (1997b). Intraday Periodicity and Volatility Persistence in Financial Markets. *Journal of Empirical Finance*, 4, 115-58.

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, F. X. Diebold, and H. Ebens. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61, 43-76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, H. Ebens and P. Labys. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71, 579-625.

Baillie, R. T. (1996). Long Memory Processes and Fractional Integration in Econometrics. *Journal of Econometrics*, 73, 5-59.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility methods. *Journal of Royal Statistical Society*, B, 64, 253-280.

Bollerslev, T. and H.O. Mikkelsen (1996). Modeling and pricing long memory in stock market volatility, *Journal of Econometrics*, 73, 151-184.

Danielsson, J. (1994). Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood. *Journal of Econometrics*, 64, 375-400.

Ding, Z., C.W.J. Granger, and R.F. Engle. (1993). A Long Memory Property of Stock Returns. *Journal of Empirical Finance*, 1, 83-106.

Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of Variance of the U.K. Inflation, *Econometrica*, 50, 987-1008.

Garman, M. B. and M. J. Klass (1980). On the estimation of security price volatility from historical data. *Journal of Business*, 53, 67-78.

Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221-238.

Ghysels, E., A. Harvey, and E. Renault. (1996). Stochastic Volatility. In Handbook of Statistics. Vol. 14: Statistical Methods in Finance, Maddala, G.S. (ed.) North-Holland, Amsterdam, pp. 119-91.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12, 371-389.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood Inference and Comparison with ARCH Models, *Review of Economic Studies*, 65, 361-93.

Lo, A. (1991). Long-term Memory in Stock Market. *Econometrica*, 59, 1279-1313.

Lobato, I.N. and N.E. Savin (1998). Real and Spurious Long-Memory Properties of Stock-Market Data. *Journal of Business & Economic Statistics*, 16, 261-68.

Mandelbrot, B. (1972). Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis," *Annals of Economic and Social Measurement*, 1, 259-290.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59, 347-370.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53, 67-78.

Rogers, L. C. G., S. E. Satchell, and Y. Yoon (1994). Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Applied Financial Economics*, 4, 241-247.

Ruiz, E. (1994). Quasi-maximum Likelihood Estimation of Stochastic Volatility Models. *Journal of Econometrics*, 63, 289-306.

Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility Models. In Time Series Models in Econometrics, Finance and Other Fields, D.R. Cox, D.V. Hinkley, and O.E. Bandorff-Nielsen (ed.) Chapman & Hall, London, pp. 1-67.

Taylor, S. J. (1982). Financial Returns Modeled by the Product of Two Stochastic Processes- A Study of Daily Sugar Prices, in *Time Series Analysis: Theory and Practice 1*, (Ed.) O. D. Anderson, North Holland, Amsterdam, pp. 203-226.

Taylor, S. J. (1986). Modeling Financial Time Series. Wiley, Chichester.

Yang, D. and Q. Zhang (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, 73, 477-491.

Zhang, L., P. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association,* 100, 1394-1411.