Title of paper: **Volatility Forecasting Performance of Two-Scaled Realized Volatility**

Authors: **Sonia Jindal[a] and Vipul[b]**

Affiliations:

[a] Doctoral Student
Indian Institute of Management Lucknow
Prabandh Nagar, Near Sitapur Road
Lucknow 226013
India
Email: sonia_jindal@iiml.ac.in

[b] Professor of Finance
Indian Institute of Management Lucknow
Prabandh Nagar, Near Sitapur Road
Lucknow 226013
India
Email: vipul@iiml.ac.in

Volatility Forecasting Performance of Two-Scaled Realized Volatility

ABSTRACT

This paper examines the forecasting performance of two-scale realized volatility (TSRV), in comparison to the conventional sparse-sampled realized volatility (SSRV) measure, using selected volatility forecasting models. TSRV time series, though stationary, is highly persistent and follows a long-memory process similar to SSRV time series. There is evidence that the forecasts based on TSRV are more efficient and less biased, as compared to those based on SSRV, for all volatility forecasting models employed. This implies that the quality of forecast predominantly depends on the quality of estimate, and not on the forecasting model used. EWMA model dominates on account of efficiency and bias for daily forecasts with TSRV. Random walk model dominates for weekly and monthly forecasts with TSRV.

Keywords: Volatility forecasting, Realized volatility, two-scale realized volatility (TSRV), sparse-sampled realized volatility (SSRV), random walk, EWMA

JEL Classification Codes: G1, G17

Volatility Forecasting Performance of Two-Scaled Realized Volatility

INTRODUCTION

Volatility of the underlying is central to the theory and practice of option pricing and risk management. Regardless of its wide application, volatility is still an ambiguous term because it is unobservable directly and there is no unique universally accepted definition for it. Thus volatility estimation and forecasting have generated a significant amount of discussion in the financial literature for the last few decades. The existing literature proposes various measures for volatility estimation and various models for volatility forecasting. But many of these measures and models, in addition to being complex, do not provide unbiased and efficient forecasts. Inaccurate forecasts could acutely affect the precision of option pricing and the effectiveness of trade risk management.

Volatility measures can be derived from the past time series of prices (historical volatility) or the prices of a market traded option (implied volatility) of an asset. Since options are not available on all assets and for all time horizons, historical volatility is generally used for volatility estimation and forecasting. Historical volatility estimation has developed in tandem with more and more availability and/ or use of data on past time series of asset prices, bringing in more accuracy with each development. Conventionally, volatility was estimated as a constant value, being the standard deviation of close-to-close returns of the asset over a specified period of time. To further improve the accuracy of the volatility estimate, Parkinson (1980) and Garman and Klass (1980) suggested the use of the other readily available asset prices (open, high and low). This led to the development of various Range Based volatility estimators. Many empirical studies[1] have found that the range based volatility estimates exhibit better estimation and forecasting performance than the traditional daily close-to-close returns.

In the last fifteen years, using high-frequency intraday data, researchers[2] have found an opportunity to further improve the volatility estimate to give a true measure of ex-post volatility (realized volatility). This has improved the quality of benchmarks for comparison of forecasts generated from competing models. Andersen and Bollerslev (1998) found that, realized volatility also enhanced the performance of volatility forecasts based on conventional volatility forecasting models. Additionally, new volatility forecasting models have also been devised based on the empirical properties of historical realized volatility.

Andersen and Bollerslev (1998) contend that theoretically, as the observation frequency increases to infinitesimally small intervals, the cumulative squared returns converge to an unbiased measure of actual historical volatility. However, practically, the presence of market microstructure noise due to non-synchronous trading, discrete price observations and bid-ask spreads makes sampling at very high-frequency undesirable. Andersen et al. (2001) suggested sparse sampling like five-minute interval cumulative squared returns, to minimize market microstructure noise. However, simple sparse sampling led to discarding intermediate data. For using complete data and eliminating noise, Zhang et al. (2005) suggested two refinements. The first being, sparse sampling of data over sub-grids of observations and averaging the results obtained across those sub-grids. The second being, the two-scale realized volatility (TSRV). TSRV is a combination of averages of realized volatilities. The realized volatilities are estimated

[1] Li and Weinbaum (2000), Bali and Weinbaum (2005), Shu and Zhang (2006), Vipul and Jacob (2007)
[2] Andersen and Bollerslev (1998), Andersen et al. (2001), Pong et al. (2004)

over sub-grids on a slow time scale, and also with all the data, for providing bias-correction. Zhang et al. argue that TSRV is the best estimator of realized volatility. TSRV, although considered a better estimate of volatility than the simple sparse-sampled realized volatility (SSRV), has not been explored for volatility forecasting.

Traditionally, simple techniques like random walk, simple average and moving average were used for volatility forecasting. These methods assigned equal weights to all the past volatility estimates included, and zero weights to the past estimates not included, whereas, one would expect the more recent events to be more relevant and therefore, to have higher weights. This is accommodated by models like Exponentially Weighted Moving Average (EWMA). Common time-series forecasting models like AR, MA, ARMA and ARIMA have also been used for volatility forecasting. Most of the volatility forecasting literature is dominated by GARCH class models which are based on the daily close-to-close returns. These models try to capture the empirical behavior of the daily close-to-close returns namely, time-variation, clustering and mean reversion. But with the availability of high-frequency intraday data, the scenario has changed. Andersen et al. (2001) studied the distribution of realized volatility and found it to be quite different from that of the traditional daily close-to-close measure. They suggested that GARCH class models may not be able to capture the empirical properties of the high-frequency intraday realized volatility.

Literature suggests that GARCH class models fare poorly as volatility forecasting models. Andersen et al. (2003) found that GARCH based volatility models are inferior to the time series models, based on high-frequency based realized volatility, in the foreign exchange market.

Superiority of realized volatility measures for volatility estimation and forecasting has been well recognized in the literature. Ghysels et al. (2006) showed that volatility measures based on high-frequency, data rather than daily returns or price ranges, predict future volatility more efficiently. Empirical literature indicates that realized volatility time series displays long memory with a fractional difference parameter d of around 0.4. Andersen et al. (2001) found that high-frequency volatility follows a long memory process and when it is studied as fractionally integrated series, the transformation induces normality. Researchers[3] have extensively employed Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) to model realized volatility.

Corsi (2003) suggested a simpler volatility forecasting model, Heterogeneous Auto-Regressive (HAR) model, to take care of the long memory process. HAR model is based on heterogeneous market hypothesis; it models the heterogeneity of market agents with respect to their time horizons. Andersen et al. (2007) extended the HAR model by dividing the past realized volatility into continuous and jump components with the help of Realized Bi-power variation measure proposed by Barndorff-Nielsen and Shephard (2004, 2006). They found that separation of continuous and jump components enhanced the volatility forecasting performance of the simple HAR model. Many studies[4] on different markets have also found that continuous and jump components of realized volatility, when separated, led to better volatility forecasting. All these studies, based on high-frequency data, have used SSRV measure for volatility forecasting. To our knowledge, no study has explored TSRV measure for volatility forecasting.

[3] Andersen et al. (2003), Pong et al. (2004) and Martens and Zein (2004)

[4] Chung et al. (2008) on Taiwan Stock Exchange; Liao (2011) on three individual Chinese stocks; Kumar (2010) on Indian financial markets (S&P CNX Nifty index)

With this background, the paper builds on the existing literature on three counts. First, it examines the volatility forecasting performance of TSRV vis-à-vis the most commonly used five-minute SSRV measure. Second, it assesses the performance of TSRV in volatility forecasting using conventional (random walk and EWMA), time series (ARIMA and ARFIMA) and HAR (HAR and HAR-J) volatility forecasting models. Third, it evaluates the forecasting performance based on both efficiency and bias. The forecasts based on TSRV, are found to be significantly more efficient and less biased, than those based on SSRV. In terms of efficiency, the performance of forecasts based on TSRV is comparable across random walk (not for daily forecasts), EWMA, ARIMA, ARFIMA and HAR volatility forecasting models. In terms of bias, the performance of forecasts based on TSRV worsens with the complexity of volatility forecasting model.

The remaining article is organized as follows. The next section provides the data and data sources. The third section describes the methodology used to assess the forecasting performance of TSRV measure in the paper. The fourth section presents the results and analysis. The fifth section concludes the paper.

DATA

The study uses S&P CNX Nifty (Nifty) index data, from 2nd January 2001 to 30th June 2011, to assess the forecasting performance of TSRV measure. Nifty is the leading index of National Stock Exchange of India (NSE). NSE, established in 1994, is the largest Indian stock exchange in terms of trading volumes. Nifty is a value-weighted stock index of NSE, derived from the prices of 50 largest capitalization and most liquid stocks. Tick-by-tick data of Nifty is used to calculate the TSRV and SSRV measures. This data was provided by the Indian Institute of Management Ahmedabad Library database. Trading data is missing for four trading days in the ten and half year data set. It is expected that in such a big data set, the errors due to missing data will be minimal. The data is filtered for special trading sessions on weekends and Diwali[5], as these sessions may not reflect the interaction of all the market players. This results in a data set of 2596 trading days.

METHODOLOGY

Volatility Estimation Measures

SSRV measure is taken as the cumulative returns at five-minute interval, similar to other such studies[6]. For each trading day t, this measure is calculated by summing up five minute close-to-close squared returns as follows.

$$SSRV_t(\Delta) = \sigma_t{}^2 = \sum_{j=1}^{T/\Delta} r_{j\Delta,\Delta}^2 \tag{1}$$

where Δ is the sampling interval equal to five minutes, r is the return equal to the log closing price relative for the sampling interval, T is the open market time period in minutes, and σ^2 is the measure of variance. TSRV measure, suggested by Zhang et al. (2005), is calculated as in Vipul and Jacob (2007). It is as follows.

$$TSRV_t = \sigma_t^2 = \frac{N}{(N-\bar{n})}[\bar{\sigma}^2_{\text{low\_frequecy, t}} - \frac{\bar{n}}{N}\sigma^2_{\text{high\_frequency, t}}] \tag{2}$$

where $\bar{n}$ is the average number of returns across all the subsamples at the low frequency and N is the total number of returns at the high frequency. This is under the assumption that the price

[5] Diwali is an annual Indian festival on which special trading sessions are organized by Indian Exchanges.
[6] Andersen et al. (2001, 2007), Ghysels et al. (2006), Koopman et al. (2005) and Kumar (2010), for instance

process is independent of the noise. In this study, TSRV values are calculated with high frequency as one second and low frequency as five minutes. The SSRV and TSRV measures calculate the open market volatility. Volatility forecasting models require realized volatility measures for the entire day, including the closed-market period. Therefore, SSRV and TSRV measures are scaled up by the ratio of daily close-to-close to open-to-close historical variances. This scaling factor ρ, similar to those employed by Koopman et al. (2005) and Vipul and Jacob (2007), is as follows.

$$\rho = \sum_{t=1}^{T} r_{cc}^2 \Big/ \sum_{t=1}^{T} r_{oc}^2 \tag{3}$$

where r_{cc} is the daily close-to-close return, r_{oc} is the daily open-to-close return, and T is the total number of trading days included in the study. The scaled daily realized variances are added over the relevant period to estimate weekly and monthly realized variances. A week is considered to be five trading days and a month is considered to be twenty-two trading days.

These estimators of realized volatility consider the price process to follow a path of continuous diffusion, and do not consider jumps in the price process. Barndorff-Nielsen and Shephard (2004, 2006) allow for separate (non-parametric) identification of the continuous and jump components of the quadratic variation process. For this purpose, they define the standardized realized bi-power variation (BPV), for each trading day t, as follows.

$$BPV_t(\Delta) = \mu_1^{-2} \sum_{j=2}^{T/\Delta} \left| r_{j\Delta,\Delta} \right| \left| r_{(j-1)\Delta,\Delta} \right| \tag{4}$$

where $\mu_1 \equiv \sqrt{(2/\pi)} = E(|Z|)$ denotes the mean of the absolute value of standard normally distributed random variable Z, Δ is the sampling interval equal to five minutes, r is the return equal to the log closing price relative for the sampling interval, and T is the open market time period in minutes. Barndorff-Nielsen and Shephard show that realized BPV represents the continuous component of the quadratic variation process. Hence, the contribution of jumps to the quadratic variation process may be consistently estimated by the difference between realized variance and BPV as follows.

$$J_t(\Delta) = \max[RV_t(\Delta) - BPV_t(\Delta), 0] \qquad (5)$$

where Δ is the sampling interval equal to five minutes, J_t is the jump contribution, RV_t is the realized variance (SSRV), and BPV_t is the bi-power variation for trading day t. Since the estimates of squared jumps in this approach could have a negative value, the left hand side measurements are truncated at zero. Jump measures are calculated with the help of equation (5), in line with Andersen et al. (2007), considering the markets to be continuous.

Volatility Forecasting Models

For forecasting volatility, standard deviation is preferred to variance as the volatility measure because the latter would involve the fourth moments (Poon and Granger, 2003). The forecasting models employed in this study include the conventional models: random walk and EWMA; the time series models: ARIMA and ARFIMA; and the HAR models: HAR and HAR-J. They are described in the following text.

Random Walk $$\hat{\sigma}_t = \sigma_{t-1} \qquad (6)$$

where σ_{t-1} is the estimate of volatility on day t-1, and $\hat{\sigma}_t$ is the forecast of volatility on day t, volatility measured as standard deviation.

EWMA $$\hat{\sigma}_t = (1-\lambda)\sigma_{t-1} + (\lambda)\sigma_{t-1}; 0 \le \lambda \le 1 \qquad (7)$$

where σ_{t-1} is the estimate of volatility on day t-1, $\hat{\sigma}_t$ and σ_{t-1} is the forecast of volatility on day t and day t-1 respectively, and λ is the smoothing parameter. λ is estimated from the data. Some researchers argue that the smoothing parameter should be allowed to change over time, in order to adapt to the latest characteristics of the time series. Others argue that this leads to unstable forecasts (Fildes, 1979). Makridakis et al. (1982) found that adaptive smoothing parameter is less successful than a constant optimized smoothing parameter. Accordingly, in this study, a constant smoothing parameter is used. Gardner (1985) recommends that the smoothing parameter should be found by minimizing the sum of ex-post 1-step-ahead forecast errors. In line with his recommendation, in this study, smoothing parameter is found by minimizing the root mean square errors of the forecasts.

Amongst the time series models, the study uses the generalized Auto-Regressive Moving Average (ARMA) model in the form of Auto-Regressive Integrated Moving Average (ARIMA) model to forecast volatility. The model is generally referred to as an ARIMA (p, d, q) model where p, d, and q are non-negative integers that refer to the order of autoregressive, integrated, and moving average parts of the model respectively. It can be written as follows.

ARIMA (p, d, q) $$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d \sigma_t = (1+\sum_{i=1}^{q}\theta_i L^i)\varepsilon_t \qquad (8)$$

where L is the lag operator, ϕ_i are the parameters of the autoregressive part, θ_i are the parameters of the moving average part, and ε_t is the error term.

Pong et al. (2004), Koopam et al. (2005) and many others have found that realized volatility can be modeled as an ARFIMA time series. Auto-Regressive Fractionally Integrated Moving

Average (ARFIMA) model generalizes ARIMA model by allowing non-integer values of the differencing parameter and is useful in modeling time series with long memory. This model is especially advantageous when a time series, though stationary (i.e. series does not have a unit root or d is not equal to 1), exhibits persistence. It can be written as follows.

ARFIMA (p, d, q) $$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d \sigma_t = (1+\sum_{i=1}^{q}\theta_i L^i)\varepsilon_t \text{ ; d is a non-integer} \tag{9}$$

Koopam et al. (2005) found that most parsimonious and effective description of the dynamics in the S&P 100 realized volatility is provided by the ARFIMA (1, d, 0) model. Following their approach, this study tries to forecast the realized volatility of Nifty with the ARFIMA (1, d, 0) model. Additionally, the study attempts to forecast realized volatility with a simpler and more parsimonious form of ARFIMA model i.e. ARFIMA (0, d, 0). In this model, autoregressive and moving average terms are absent and differencing parameter d is a non-integer. This form is helpful when a time series can be explained with just the differencing parameter. It can be written as follows.

ARFIMA (0, d, 0) $$(1-L)^d \sigma_t = \varepsilon_t \text{ ; d is a non-integer} \tag{10}$$

Best ARIMA model is chosen according to AIC value. Both ARIMA and ARFIMA models are estimated using the maximum likelihood estimation and the innovations and their variance are found by a Kalman filter.

Figlewski (1997) suggests that, while making weekly or monthly predictions, forecasts constructed from weekly and monthly data should be used. Thus for the random walk, EWMA, ARIMA and ARFIMA models, non overlapping data is used in the estimation of weekly and monthly volatility measures as in Martens and Zein (2004). In the EWMA, ARIMA and ARFIMA models the estimation period for the out-of-sample forecasts is 1250 days for daily forecasts, 250 weeks for weekly forecasts and 60 months for monthly forecasts. Each of these estimation sets corresponds to a period of about 5 years.

The HAR model, proposed by Corsi (2003), is specified as a multi-component volatility model. The model has an additive hierarchical structure, such that the volatility is specified as a sum of components over different horizons. The model comprises of three volatility components: daily, weekly and monthly. These represent (1) the short-term daily activity of speculators; (2) the medium-term activity typically due to portfolio managers, who rebalance their positions weekly, and (3) the long-term activity with a characteristic time of one or more months. The model can be written as follows.

HAR $$\left(RV_{t,t+H}\right)^{1/2} = c + \beta_{dH}\left(RV_t\right)^{1/2} + \beta_{wH}\left(RV_{t-5,t}\right)^{1/2} + \beta_{mH}\left(RV_{t-22,t}\right)^{1/2} + \varepsilon_{t,t+H} \quad (11)$$

where $RV_{t,t+H}$ is the H day ahead ex-post measure of daily realized volatility, H being the forecasting horizon. RV_t , $RV_{t-5,t}$ and $RV_{t-22,t}$ are the measured contemporaneous daily, weekly and monthly realized volatility, and $\varepsilon_{t,t+H}$ represents the volatility measurement and estimation errors. The consistency of the above requires the realized volatility measures to be unbiased.

Andersen et al. (2007) introduced HAR-J model to account for jump components in volatility forecasting. It can be written as follows.

HAR-J $$\left(RV_{t,t+H}\right)^{1/2} = c + \beta_{dH}\left(RV_t\right)^{1/2} + \beta_{wH}\left(RV_{t-5,t}\right)^{1/2} + \beta_{mH}\left(RV_{t-22,t}\right)^{1/2} + \beta_{jH}\left(J_t\right)^{1/2} + \varepsilon_{t,t+H} \quad (12)$$

where J_t is the contemporaneous jump component. For the HAR models, overlapping data is used in the estimation of weekly or monthly volatility measures, as in Andersen et al. (2007). HAR models are estimated using ordinary least square (OLS) estimation.

Evaluation of Performance

The study uses both bias and efficiency for evaluation of the competing volatility measures and forecasting methods. The efficiency is measured by the mean squared error (MSE), and the mean absolute error (MAE); and the bias, by the mean bias and the mean relative bias (MRB). These loss functions are as follows.

$$\text{MSE} = E(\hat{\sigma}_t - \sigma_t)^2 \quad (14)$$

$$\text{MAE} = E\left|\hat{\sigma}_t - \sigma_t\right| \quad (15)$$

$$\text{Mean Bias} = E\left(\hat{\sigma}_t - \sigma_t\right) \quad (16)$$

$$\text{MRB} = E[\left(\hat{\sigma}_t - \sigma_t\right) / \sigma_t] \quad (17)$$

The efficiency and bias of the forecasts, based on the two estimation measures (TSRV and SSRV), and the various forecasting models (random walk, EWMA, ARIMA, ARFIMA, HAR and HAR-J), are compared to find the most appropriate estimator and forecasting model. For this purpose, Wilcoxon signed ranks test (recommended by Diebold & Mariano, 1995) is used. In this test, the null hypothesis is that the two forecasts selected for comparison are equally efficient (or biased). This is one of the most powerful non-parametric tests for comparing related samples, and is a useful alternative, when normality assumptions do not hold. Its power efficiency approaches 95.5 percent as compared to t-test, as sample size increases, and is close to 95 percent even for small-sized samples (Siegel and Castellan, 1988). It is used because most of the financial time series data are known to exhibit non-normality. For all the volatility forecasting models, the ex-post benchmark volatility measure is taken as TSRV.

RESULTS

Volatility Estimation

The average variance of the microstructure noise, estimated by the TSRV measure is very low, compared to the variance of the sparsely sampled (over sub-grids) and averaged five minute returns. It is about 0.0000273 whereas the average variance of TSRV is about 0.000177. This indicates that the estimation of realized volatility is quite accurate, even with sparse sampling (over sub-grids) and averaging only. The explicit bias correction of TSRV, further improves it. The close-to-open variance is about 4.93% of the open-to-close variance.[7] Accordingly, a scaling factor of 1.0493 is used for converting the TSRV and SSRV measures for the open market period to those for an entire day. Table[8] I lists the sample statistics for TSRV and SSRV time series. The table shows that, though SSRV is upward biased and exhibits higher variance, both the series have similar properties. Both are non-normal, leptokurtic and positively skewed as found by Vipul and Jacob (2007). Both the series are stationary, but the fractional parameter d calculated by Geweke and Porter-Hudak method is significantly different from zero (it is between 0.3-0.5) for both. This shows presence of long memory behavior in the two series. These results are similar to those found by Andersen et al. (2001), for high-frequency data on deutsche-mark and yen returns against the US dollar.

Volatility Forecasting

Table II lists the performance of forecasts, based on TSRV and SSRV, using the random walk model. The table shows that forecasts based on TSRV are more efficient and less biased vis-à-vis those based on SSRV. The MRB of random walk forecasts based on SSRV is in the range of 24-

---

[7] This is based on the average ratio of close-to-close to open-to-close historical volatility of Nifty for the 2596 trading days covered by the study.

[8] All tables are provided at the end of the article

28% whereas that based on TSRV is in the range of 3-5%. This implies that previous day SSRV is a highly biased forecast for the next day.

Table III lists the performance of in-sample and out-of-sample forecasts, based on TSRV and SSRV, using the EWMA model. In the out-of-sample data sets, an attempt was made to find an adaptive smoothing parameter, which changes over time, but it gave unstable forecasts. Thus a constant smoothing parameter is found by minimizing the root mean square errors (RMSE) of forecasts for the full period, as suggested by Gardner (1985). This is done for daily, weekly and monthly, in-sample and out-of-sample, data sets. The smoothing parameter λ, based on TSRV and SSRV is in the range 0.4-0.48 and 0.69-0.82 respectively. These values are very different from the ones suggested by J. P. Morgan's *RiskMetrics*™ model (λ= 0.94 for daily data). To reconcile this difference, λ for daily in-sample close-to-close returns is estimated. Its value (0.79), that minimizes the RMSE with TSRV as benchmark is quite close to 0.94. The lower values of λ for TSRV imply that forecasts, based on TSRV, are more dependent on the recent past component than the smoothed component. That probably is the reason, why random walk forecasts based on TSRV perform so well. Table III shows that the in-sample EWMA forecasts, based on TSRV, have marginally higher efficiency and the out-of-sample forecasts have marginally lower efficiency, as compared to random walk model. However, the bias for EWMA is higher than that for the random walk. The forecasts based on SSRV follow a similar pattern, with the efficiency of EWMA model not consistently better than that for the random walk. However, on both efficiency and bias criteria, TSRV forecasts are better than SSRV forecasts, as in the case of random walk model.

Table IV lists the performance of in-sample and out-of-sample forecasts, based on TSRV and SSRV, using ARIMA model. The table shows that forecasts, using ARIMA model, based on TSRV outperform those based on SSRV, both in terms of efficiency and bias, as in the earlier models. Table V lists the ranges of order of autoregressive (p), integration (d) and moving average (q) parts in the estimated ARIMA models. Tables IV and V indicate that, d in all the estimated ARIMA models, for both in-sample and out-of-sample, is either zero or one. Since ARIMA model assumes the d to be an integer, it does not take care of fractional integration. Table I shows that, both TSRV and SSRV time series, in their sigma form, did not have a unit root, and possessed long memory. This indicates presence of fractional integration. Pong et al. (2004), Koopam et al. (2005) and many others have found that realized volatility can be effectively modeled as an ARFIMA series. This calls for the use of ARFIMA model in volatility forecasting of Nifty.

Table VI lists the performance of in-sample and out-of-sample forecasts, based on TSRV and SSRV, using ARFIMA (0, d, 0) model. The table shows that forecasts, using ARFIMA (0, d, 0) model, based on both TSRV and SSRV, are similar to their ARIMA forecasts. Thus implying that, although Nifty can be modeled with ARIMA model well ARFIMA (0, d, 0) provides a more parsimonious alternative. Table VI also shows that forecasts based on SSRV, are more efficient with ARFIMA (0, d, 0) model than with the random walk model. For instance, the MSE of daily out-of-sample forecasts, based on SSRV, are 0.5899 and 0.7024 with ARFIMA (0, d, 0) and the random walk model respectively. Similar improvements are seen in weekly and monthly forecasts. On the other hand, forecasts based on TSRV, do not show significant improvements in efficiency with the use of ARFIMA (0, d, 0) model. The MSE of daily out-of-sample forecasts, based on TSRV, are 0.4468 and 0.4485 with ARFIMA (0, d, 0) and the random walk model respectively. Weekly and monthly forecasts become less efficient for ARFIMA (0, d, 0). Table V also indicates that forecasts, based on both TSRV and SSRV, are more biased with the ARFIMA (0, d, 0) model than their random walk forecasts. ARFIMA (0, d, 0) model enhances efficiency of forecasts based on SSRV, but also increases their bias. The enhanced efficiency of forecasts, based on SSRV is still less than the efficiency of forecasts based on TSRV.

Table VII lists the performance of in-sample and out-of-sample forecasts, based on TSRV and SSRV, using ARFIMA (1, d, 0) model. The table shows that forecasts, using ARFIMA (1, d, 0) model, based on both TSRV and SSRV, do not show any improvement in efficiency and bias over those with ARFIMA (0, d, 0) model. Therefore, TSRV and SSRV time series, can be better modeled by a simpler and more parsimonious ARFIMA (0, d, 0) model. These results contrast with the results of Koopam et al. (2005), who found that the ARFIMA (1, d, 0) model explains the dynamics in the realized volatility of S&P 100 most effectively and parsimoniously.

Table VIII lists the in-sample OLS regression results of the HAR and the HAR-J models, based on both TSRV and SSRV. These results indicate that the adjusted R-squares of the regressions are the highest for daily forecasts, and the lowest for monthly forecasts. This pattern is repeated across the two models and the two measures. Coefficients of historical daily, weekly and monthly realized volatility are significant for all forecasting horizons, in both HAR and HAR-J model regressions, based on both TSRV and SSRV. But the coefficient of the jump component in HAR-J model is significant only in the regressions based on SSRV. It is not significant in the regressions based on TSRV. Consequently, R-squares of the HAR-J model improve over that of the HAR model, for forecasts based on SSRV. Andersen et al. (2007) used the SSRV measure for realized volatility and found similar results for S&P 500. They had therefore credited HAR-J model with accounting for the jump component in the HAR model. Our results indicate that the inclusion of jump component improves the estimation of the HAR model only for SSRV and not for TSRV. The TSRV measure is robust enough to incorporate the effect of jumps in the historical realized volatility series.

Table IX and X list the performance of in-sample and out-of-sample forecasts, based on TSRV and SSRV, using HAR model and HAR-J model respectively. The tables show that efficiency and bias of forecasts, based on TSRV, are very similar for both the HAR and HAR-J model. The efficiency and bias of forecasts, based on SSRV, improve with the inclusion of jump component in the HAR model. Despite this improvement, forecasts based on SSRV still lag behind forecasts based on TSRV measure, both in terms of efficiency and bias. Table XI lists the performance of out-of-sample forecasts, based on TSRV as compared to those based on SSRV, using all the forecasting models. It represents the relative performance of the TSRV measure over the SSRV measure. The table shows that for all forecasting models, and for all forecasting horizons, the forecasts based on TSRV are more efficient and less biased than those based on SSRV. The p-values of the Wilcoxon signed-ranks test confirm this fact.

Table XII lists the comparative performance of out-of-sample forecasts, based on TSRV, in terms of efficiency. The table compares the performance of all the forecasting models for daily, weekly and monthly forecasts. Panel A of the table indicates that ARFIMA (0, d, 0) model forecasts daily volatility more efficiently than random walk, EWMA and ARIMA. The p-values show that all the models provide, significantly more efficient daily forecasts, than the random walk model. There is no significant difference in the efficiency of daily forecasts provided by the long memory models (ARFIMA and HAR). Except for ARFIMA (0, d, 0) performing better than EWMA and ARIMA, there is no significant difference in performance of EWMA, ARIMA, ARFIMA and HAR models.

Panel B of table XII indicates that the HAR models forecast weekly volatility more efficiently than the other models. The p-values show that HAR models provide significantly more efficient weekly forecasts than the ARFIMA (1, d, 0) model. Other than that, there is no significant difference in efficiency of weekly forecasts amongst the various models. Panel C of table XI indicates that random walk model forecasts monthly volatility as well as the other models. The p-values show that there is no significant difference in efficiency of monthly forecasts provided by random walk, EWMA, ARIMA, ARFIMA and HAR models.

Table XIII represents the comparison of out-of-sample relative bias, relative to zero, based on TSRV, for all forecasting models and all forecasting horizons. The table shows that all forecasts are positively biased. At 1% level, the bias in random walk forecasts is not significantly different from zero for all forecasting horizons. For monthly forecasts, the bias in forecasts, from all models except ARIMA model, is not significantly different from zero, at 1% level. But, at 5% level, only the random walk model has a bias not significantly different from zero. The Z-statistics indicate that the bias worsens with the complexity of forecasting model. Based on TSRV, EWMA model provides a good trade-off between efficiency and bias for daily forecasts. For weekly and monthly forecasts random walk is the best option on these criteria.

CONCLUSIONS

The study examines the forecasting performance of two-scale realized volatility (TSRV) in comparison to the conventional sparse sampled realized volatility (SSRV) measure using various forecasting models. TSRV suggested by Zhang et al. (2005) and SSRV suggested by Andersen et al. (2001) are both measures of realized volatility (ex-post true volatility) calculated using high frequency intra-day data. TSRV combines average values of realized volatility measured at two frequencies, using complete data. SSRV calculates cumulative squared returns at a lower frequency, like five minutes. Both these approaches are followed to minimize market microstructure noise.

Most studies have used SSRV, as the measure of realized volatility, to forecast volatility. TSRV, although considered a better estimate of volatility than SSRV with its more robust bias-correction technique, has not been explored in volatility forecasting. The study attempts to fill this gap by comparing the performance of forecasts based on TSRV relative to those based on SSRV, in terms of both efficiency and bias. The study employs conventional (random walk and EWMA), time series (ARIMA and ARFIMA) and HAR (HAR and HAR-J) volatility forecasting models for this purpose.

There is evidence that the forecasts, based on TSRV, are significantly more efficient and significantly less biased than those based on SSRV. These results are universal, regardless of the volatility forecasting model or the volatility forecasting horizon. Thus quality of forecast primarily depends on the quality of estimate and not so much on the forecasting model used. This implies that better estimation leads to better forecasting. These results are similar to those of Vipul and Jacob (2007), where the authors found that the range based estimators, which estimated the historical volatility better, also led to better forecasts regardless of volatility forecasting model or horizon.

Both TSRV and SSRV time series are stationary, but exhibit persistence with a fractional integrating parameter between 0.3-0.5. There is evidence that ARFIMA (0, d, 0) model provides significantly more efficient daily forecasts than ARIMA model with TSRV. The ARFIMA (0, d, 0) and the ARFIMA (1, d, 0) models, provide almost equally efficient and equally biased forecasts with both TSRV and SSRV, for all time horizons. These results indicate that the dynamics of the realized volatility of Nifty is better captured by the fractional differencing parameter. HAR-J model provides more efficient forecasts than the HAR model, with SSRV. However, with TSRV measure, both these models give similar forecasts implying that the inclusion of jump component in the HAR model is not required, for forecasts based on TSRV.

The study also assesses the competing forecasting models, to find out the model, which provides the most efficient and least biased forecasts, with TSRV. For daily forecasting horizons, the efficiency of forecasts is comparable across EWMA, ARIMA, ARFIMA and HAR models. For weekly and monthly forecasting horizons, it is comparable across random walk, EWMA, ARIMA ARFIMA and HAR models. Bias of forecasts, worsens with the complexity of model. It is not significantly different from zero for random walk forecasts for all horizons. Based on TSRV, EWMA model provides a good trade-off between efficiency and bias for daily forecasts. Random walk model is the best option for weekly and monthly forecasts on these criteria.

REFERENCES

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

Andersen, T. G., Bollerslev, T., and Diebold, F.X. (2007). Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and Forecasting Realized Volatility. Econometrica, 71, 579–625.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The Distribution of Realized Exchange Rate Volatility. Journal of the American Statistical Association, 96(453), 42-55.

Bali, T.G. and Weinbaum, D. (2005). A Comparative Study of Extreme Value Volatility Estimators. Journal of Futures Markets, 25(9), 873-92.

Barndorff-Nielsen, O.E., and Shephard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. Journal of Financial Econometrics, 2, 1-37.

Barndorff-Nielsen, O.E., and Shephard, N. (2006). Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation. Journal of Financial Econometrics, 4, 1-30.

Chung, H., Huang, C.S., and Tseng, T.C. (2008). Modeling and Forecasting of Realized Volatility Based on High-Frequency Data: Evidence from Taiwan. International Research Journal of Finance and Economics, 22, 178-191.

Corsi, F. (2003). A Simple Long Memory Model of Realized Volatility. Working Paper, University of Southern Switzerland.

Diebold, F.X., and Mariano, R.S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13, 253–263.

Fildes, R. (1979). Quantitative forecasting – the state of the art: extrapolative models. Journal of the Operational Research Society, 30, 691-710.

Figlewski, S. (1997). Forecasting volatility. Financial Markets, Institutions and Instruments, 6(1), 1–88.

Gardner, E.S. Jr. (1985). Exponential smoothing: the state of the art. Journal of Forecasting, 4, 1-28.

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

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

Kumar, M. (2010). Improving the accuracy: volatility modeling and forecasting using high-frequency data and the variational component. Journal of Industrial Engineering and Management, 3(1), 199-220.

Koopman, S.J., Jungbacker, B. and Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurements. Journal of Empirical Finance, 12, 445-475.

Li, K. and Weinbaum, D. (2000). The Empirical Performance of Alternative Extreme Value Volatility Estimators. Working Paper, Stern School of Business, New York.

Liao, Y. (2011). Does Decomposing Realized Volatility Help in Risk Prediction: Evidence from Chinese Mainland Stocks. Working Paper, Australian National University, Australia.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., and Winkler, R. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. Journal of Forecasting, 1, 111-153.

Martens, M., and Zein, J. (2004). Predicting financial volatility: high frequency time series forecasts vis-à-vis implied volatility. Journal of Future Markets, 24(11), 1005-1028.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. Journal of Business, 53, 61–65.

Poon, S., and Granger, C.W.J. (2003). Forecasting volatility in financial markets: A Review. Journal of Economic Literature, 41, 478–539.

Pong, S., Shackleton, M. B., Taylor, S. J., and Xu, X. (2004). Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models. Journal of Banking & Finance, 28, 2541–2563.

Shu, J., and Zhang, J.E. (2006). Testing range estimators of historical volatility. Journal of Futures Markets, 26, 297–313.

Siegel, S., and Castellan, N.J. (1988). The Case of One Sample, Two Measures or Paired Replicates. In: Siegel, S. and Castellan, N.J. (Second Ed.), Nonparametric Statistics for the Behavioral Sciences. Mc-Graw Hill, Singapore, 87-95.

Vipul and Jacob, J. (2007). Forecasting Performance of Extreme-Value Volatility estimators. Journal of Futures Markets, 27(11), 1085-1105.

Zhang, L., Mykland, P. A. and Aı¨t-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.

Table I: Descriptive Statistics of TSRV and SSRV

	TSRV		SSRV	
	σ^2	σ	σ^2	σ
mean	1.772	1.156	2.788	1.391
sd	0.035	0.660	0.080	0.924
skewness	12.16	3.58	19.81	4.86
kurtosis	219.15	27.81	534.58	51.19
Jarque bera test (p value)	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
adf test (p value)	< 0.01	< 0.01	< 0.01	< 0.01
GPH (d)	0.38	0.44	0.34	0.43
No. of observations	2596	2596	2596	2596

Note: σ^2 represents the variance and σ represents the standard deviation. The figures of σ^2 and σ are given in percent square and percent respectively. Jarque bera test has a null hypothesis of normality. Adf test has an alternate hypothesis of stationarity. GPH (d) represents the fractional parameter d estimated by the method of Geweke and Porter-Hudak.

Table II: Forecasting performance using Random Walk model

Random walk	**RMSE**	**MAE**	**Mean Bias**	**MRB**
Panel A: TSRV				
Daily	0.4485	0.2736	0.0002	4.1453
Weekly	1.0687	0.5944	0.0023	3.8280
Monthly	2.2141	1.5163	0.0152	4.9409
Panel B: SSRV				
Daily	0.7024	0.3947	0.2351	24.2384
Weekly	1.5320	0.9144	0.5821	25.2559
Monthly	3.1230	2.1258	1.3304	28.3470

Note: RMSE, MAE, Mean Bias and MRB are given in percentages.

Table III: Forecasting performance of TSRV and SSRV using EWMA model

	RMSE	**MAE**	**Mean Bias**	**MRB**
Panel A: TSRV in-sample				
Daily (λ=0.42)	0.4275	0.2550	0.0003	5.0062
Weekly (λ=0.47)	0.9912	0.5772	0.0040	5.2506
Monthly (λ=0.45)	2.0970	1.4555	0.0231	6.7795
Panel B: TSRV out-of-sample				
Daily (λ=0.40)	0.4644	0.2800	0.0004	5.1652
Weekly (λ=0.46)	1.0868	0.6433	0.0015	5.3011
Monthly (λ=0.48)	2.5214	1.7848	0.0396	8.8016
Panel C: SSRV in-sample				
Daily (λ=0.81)	0.5612	0.3790	0.2356	27.8712
Weekly (λ=0.75)	1.2847	0.9339	0.5957	30.2466
Monthly (λ=0.69)	2.6915	2.2112	1.3784	33.6028
Panel C: SSRV out-of-sample				
Daily (λ=0.82)	0.6253	0.4308	0.2773	29.5649
Weekly (λ=0.74)	1.4387	1.0636	0.6843	31.5841
Monthly (λ=0.74)	3.1395	2.6919	1.6347	38.2520

Note: λ is chosen by minimizing the RMSE of the forecasts. RMSE, MAE, Mean Bias and MRB are given in percentages.

Table IV: Forecasting performance of TSRV and SSRV using ARIMA model

	RMSE	MAE	Mean Bias	MRB
Panel A: TSRV in-sample				
Daily (p=1, d=1, q=1)	0.4103	0.2468	0.0008	6.0445
Weekly (p=1, d=1, q=2)	0.9617	0.5620	0.0143	7.2331
Monthly (p=1, d=0, q=0)	1.9821	1.3378	-0.0001	8.6953
Panel B: TSRV out-of-sample				
Daily ($\bar{p}$=1.86, $\bar{d}$=0.83, $\bar{q}$=1.72)	0.4511	0.2723	-0.0074	5.9132
Weekly ($\bar{p}$=1.18, $\bar{d}$=0.65, $\bar{q}$=1.35)	1.0875	0.6308	-0.0435	5.9104
Monthly ($\bar{p}$=0.83, $\bar{d}$=0.29, $\bar{q}$=0.50)	2.5482	1.7310	0.0460	11.4205
Panel C: SSRV in-sample				
Daily (p=1, d=1, q=1)	0.5301	0.3676	0.2389	28.5850
Weekly (p=1, d=1, q=1)	1.2307	0.9039	0.6054	30.5340
Monthly (p=1, d=1, q=1)	2.4127	1.9933	1.3243	33.3463
Panel C: SSRV out-of-sample				
Daily ($\bar{p}$=2.18, $\bar{d}$=0.83, $\bar{q}$=2.82)	0.6085	0.4202	0.2783	29.7245
Weekly ($\bar{p}$=1.46, $\bar{d}$=0.57, $\bar{q}$=1.32)	1.3478	0.9924	0.6144	29.9383
Monthly ($\bar{p}$=0.83, $\bar{d}$=0.38, $\bar{q}$=0.66)	2.9690	2.5444	1.4226	34.9357

Note: ARIMA model is estimated using maximum likelihood estimation. Best ARIMA model is chosen according to AIC value with a maximum value of order of autoregressive and moving average as five. RMSE, MAE, Mean Bias and MRB are given in percentages.

Table V: Ranges of p, d, q of ARIMA model used in out-of-sample forecasting

	Maximum			Minimum		
	p	**d**	**q**	**p**	**d**	**q**
Panel A: TSRV						
Daily	5	1	5	1	0	1
Weekly	5	1	4	0	0	0
Monthly	1	1	2	0	0	0
Panel B: SSRV						
Daily	5	1	5	1	0	1
Weekly	4	1	5	0	0	0
Monthly	1	1	2	0	0	0

Table VI: Forecasting performance of TSRV and SSRV using ARFIMA (0, d, 0) model

	RMSE	MAE	Mean Bias	MRB
Panel A: TSRV in-sample				
Daily (d=0.498)	0.4115	0.2435	0.0011	7.0876
Weekly (d=0.483)	0.9616	0.5595	0.0058	7.5623
Monthly (d=0.453)	2.0123	1.3707	0.0272	9.2070
Panel B: TSRV out-of-sample				
Daily ($\bar{d}$ =0.496)	0.4468	0.2669	-0.0007	7.3578
Weekly ($\bar{d}$ =0.454)	1.0754	0.6311	-0.0176	7.7337
Monthly ($\bar{d}$ =0.396)	2.4800	1.6767	-0.0730	10.4142
Panel C: SSRV in-sample				
Daily (d=0.428)	0.5211	0.3649	0.2373	29.0786
Weekly (d=0.461)	1.1976	0.8865	0.5897	30.6631
Monthly (d=0.416)	2.4861	2.1046	1.3538	34.4527
Panel C: SSRV out-of-sample				
Daily ($\bar{d}$ =0.448)	0.5899	0.4140	0.2782	30.5474
Weekly ($\bar{d}$ =0.437)	1.3217	0.9912	0.6484	31.9492
Monthly ($\bar{d}$ =0.354)	2.8190	2.4607	1.3628	35.8523

Note: ARFIMA (0, d, 0) model is estimated using maximum likelihood estimation. RMSE, MAE, Mean Bias and MRB are given in percentages.

Table VII: Forecasting performance of TSRV and SSRV using ARFIMA (1, d, 0) model

	RMSE	**MAE**	**Mean Bias**	**MRB**
Panel A: TSRV in-sample				
Daily (d=0.425)	0.4086	0.2449	0.0011	7.0962
Weekly (d=0.325)	0.9643	0.5574	0.0050	7.9474
Monthly (d=0.000)	1.9846	1.3354	0.0101	8.4288
Panel B: TSRV out-of-sample				
Daily ($\bar{d}$ =0.336)	0.4477	0.2697	-0.0065	7.2595
Weekly ($\bar{d}$ =0.015)	1.1107	0.6414	-0.0731	6.8989
Monthly ($\bar{d}$ =0.000)	2.4587	1.6784	-0.1031	8.1277
Panel C: SSRV in-sample				
Daily (d=0.448)	0.5223	0.3649	0.2372	28.9223
Weekly (d=0.346)	1.1917	0.8811	0.5891	31.2163
Monthly (d=0.000)	2.4669	2.0235	1.3202	33.7337
Panel C: SSRV out-of-sample				
Daily ($\bar{d}$ =0.390)	0.5902	0.4091	0.2697	30.1667
Weekly ($\bar{d}$ =0.174)	1.3300	0.9526	0.5869	30.6010
Monthly ($\bar{d}$ =0.000)	2.8176	2.3608	1.2555	33.1045

Note: ARFIMA (1, d, 0) model is estimated using maximum likelihood estimation. RMSE, MAE, Mean Bias and MRB are given in percentages.

Table VIII: In sample regression results of HAR and HAR-J model

	HAR (TSRV)			HAR (SSRV)			HAR-J (TSRV)			HAR-J (SSRV)		
H	1	5	22	1	5	22	1	5	22	1	5	22
c	0.001	0.006	0.022	0.002	0.008	0.031	0.001	0.006	0.022	0.002	0.008	0.030
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_d	0.574	0.806	0.903	0.376	0.682	0.769	0.575	0.799	0.884	0.562	0.936	1.023
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_w	0.063	0.194	0.416	0.115	0.194	0.383	0.063	0.191	0.409	0.105	0.180	0.369
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_m	0.034	0.114	0.234	0.036	0.116	0.225	0.034	0.114	0.235	0.034	0.112	0.221
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_j							-0.002	0.026	0.066	-0.307	-0.420	-0.418
							(0.884)	(0.373)	(0.299)	(0.000)	(0.000)	(0.000)
Adj. R^2	0.615	0.545	0.404	0.474	0.459	0.327	0.615	0.545	0.404	0.501	0.470	0.330

Note: The table reports the in-sample OLS estimates for daily (H=1) and overlapping weekly (H=5), and monthly (H=22) HAR and HAR-J volatility forecast regressions based on TSRV and SSRV volatility measures.

Table IX: Forecasting performance of TSRV and SSRV using HAR model

	RMSE	MAE	Mean Bias	MRB
Panel A: TSRV in-sample				
Daily	0.4109	0.2458	0.0000	7.1574
Weekly	0.9164	0.5364	0.0000	6.8959
Monthly	1.9741	1.2885	0.0000	8.3083
Panel B: TSRV out-of-sample				
Daily	0.4478	0.2685	-0.0164	6.4092
Weekly	1.0020	0.5853	-0.0677	5.4597
Monthly	2.2388	1.4471	-0.2137	6.5165
Panel C: SSRV in-sample				
Daily	0.5299	0.3647	0.2347	29.5341
Weekly	1.1708	0.8541	0.5758	30.3313
Monthly	2.4592	1.9847	1.3003	33.4223
Panel C: SSRV out-of-sample				
Daily	0.5795	0.3924	0.2426	28.9341
Weekly	1.2249	0.8946	0.5480	28.6742
Monthly	2.5384	2.0493	1.1285	30.7959

Note: HAR model is estimated using OLS estimation. RMSE, MAE, Mean Bias and MRB are given in percentages.

Table X: Forecasting performance of TSRV and SSRV using HAR-J model

	RMSE	**MAE**	**Mean Bias**	**MRB**
Panel A: TSRV in-sample				
Daily	0.4109	0.2459	0.0000	7.1576
Weekly	0.9163	0.5360	0.0000	6.8979
Monthly	1.9737	1.2877	0.0000	8.3101
Panel B: TSRV out-of-sample				
Daily	0.4500	0.2686	-0.0128	6.6437
Weekly	1.0051	0.5844	-0.0581	5.7576
Monthly	2.2378	1.4477	-0.2170	6.4529
Panel C: SSRV in-sample				
Daily	0.5192	0.3599	0.2347	28.8628
Weekly	1.1590	0.8461	0.5758	29.9589
Monthly	2.4506	1.9800	1.3003	33.2990
Panel C: SSRV out-of-sample				
Daily	0.5530	0.3755	0.2185	26.5974
Weekly	1.1933	0.8713	0.5196	27.4439
Monthly	2.7493	2.2582	1.4427	36.0195

Note: HAR-J model is estimated using OLS estimation. RMSE, MAE, Mean Bias and MRB are given in percentages.

Table XI: Forecasting performance of TSRV relative to SSRV

Forecasting Method	% change in efficiency (MSE)	p-value	% change in relative bias (MRB)	p-value
Panel A: Daily				
Random Walk	-59.23	0.000	-82.90	0.000
EWMA	-44.86	0.000	-82.53	0.000
ARIMA	-45.06	0.000	-80.11	0.000
ARFIMA (0, d, 0)	-42.63	0.000	-75.91	0.000
ARFIMA (1, d, 0)	-42.47	0.000	-75.94	0.000
HAR	-40.29	0.000	-77.85	0.000
HAR-J	-33.77	0.000	-75.02	0.000
Panel B: Weekly				
Random Walk	-51.34	0.000	-84.84	0.000
EWMA	-42.94	0.000	-83.22	0.000
ARIMA	-34.89	0.000	-80.26	0.000
ARFIMA (0, d, 0)	-33.81	0.000	-75.79	0.000
ARFIMA (1, d, 0)	-30.26	0.000	-77.46	0.000
HAR	-33.09	0.000	-80.96	0.000
HAR-J	-29.05	0.000	-79.02	0.000
Panel C: Monthly				
Random Walk	-49.74	0.000	-82.57	0.000
EWMA	-35.50	0.000	-76.99	0.000
ARIMA	-26.34	0.000	-67.31	0.000
ARFIMA (0, d, 0)	-22.61	0.000	-70.95	0.000
ARFIMA (1, d, 0)	-23.85	0.002	-75.45	0.000
HAR	-22.21	0.000	-78.84	0.000
HAR-J	-33.75	0.000	-82.09	0.000

Note: Change in efficiency (relative bias) represents percentage change in MSE (MRB) of the out-of-sample forecasts based on TSRV measure as compared to the forecasts based on SSRV measure. The statistical significance of this difference is indicated by the p-values of Wilcoxon signed-ranks test.

Table XII: Comparative forecasting efficiency of various models based on TSRV measure

		Random Walk	**EWMA**	**ARIMA**	**ARFIMA (0, d, 0)**	**ARFIMA (1, d, 0)**	**HAR**
Panel A: Daily							
% increase in MSE	**EWMA**	7.190					
Z statistic		5.151					
p-value		**0.000**					
% increase in MSE	**ARIMA**	1.138	-5.645				
Z statistic		5.756	0.327				
p-value		**0.000**	0.744				
% increase in MSE	**ARFIMA**	-0.755	-7.412	-1.872			
Z statistic	**(0, d, 0)**	5.788	2.611	3.001			
p-value		**0.000**	**0.009**	**0.003**			
% increase in MSE	**ARFIMA**	-0.379	-7.061	-1.500	0.379		
Z statistic	**(1, d, 0)**	6.588	1.283	0.888	-1.772		
p-value		**0.000**	0.200	0.375	0.076		
% increase in MSE	**HAR**	-0.321	-7.007	-1.443	0.438	0.059	
Z statistic		7.149	2.364	1.120	-1.326	1.582	
p-value		**0.000**	0.018	0.263	0.185	0.114	
% increase in MSE	**HAR-J**	0.676	-6.076	-0.457	1.442	1.059	1.000
Z statistic		7.084	2.553	1.404	-1.036	2.088	1.863
p-value		**0.000**	0.011	0.160	0.300	0.037	0.062

Table XII (continued): Comparative forecasting efficiency of various models based on TSRV measure

		Random Walk	EWMA	ARIMA	ARFIMA (0, d, 0)	ARFIMA (1, d, 0)	HAR
Panel B:Weekly							
% increase in MSE	**EWMA**	3.418					
Z statistic		-0.428					
p-value		0.668					
% increase in MSE	**ARIMA**	3.558	0.135				
Z statistic		-0.043	0.712				
p-value		0.966	0.476				
% increase in MSE	**ARFIMA**	1.252	-2.095	-2.227			
Z statistic	**(0, d, 0)**	-0.134	0.343	0.955			
p-value		0.893	0.731	0.340			
% increase in MSE	**ARFIMA**	8.012	4.441	4.300	6.676		
Z statistic	**(1, d, 0)**	-0.282	-1.105	-0.416	-0.968		
p-value		0.778	0.269	0.678	0.333		
% increase in MSE	**HAR**	-12.095	-15.000	-15.115	-13.182	-18.615	
Z statistic		1.066	1.677	2.179	1.498	2.566	
p-value		0.286	0.093	0.029	0.134	**0.010**	
% increase in MSE	**HAR-J**	-11.540	-14.464	-14.579	-12.634	-18.101	0.631
Z statistic		1.105	1.753	2.180	1.671	3.006	1.695
p-value		0.269	0.080	0.029	0.950	**0.003**	0.090

Table XII (continued): Comparative forecasting efficiency of various models based on TSRV measure

		Random Walk	EWMA	ARIMA	ARFIMA (0, d, 0)	ARFIMA (1, d, 0)	HAR
Panel C: Monthly							
% increase in MSE	**EWMA**	29.690					
Z statistic		0.646					
p-value		0.518					
% increase in MSE	**ARIMA**	32.460	2.136				
Z statistic		0.693	0.089				
p-value		0.488	0.929				
% increase in MSE	**ARFIMA**	25.465	-3.258	-5.281			
Z statistic	**(0, d, 0)**	1.227	0.484	1.212			
p-value		0.220	0.628	0.226			
% increase in MSE	**ARFIMA**	23.324	-4.908	-6.897	-1.706		
Z statistic	**(1, d, 0)**	1.878	0.050	0.646	-0.314		
p-value		0.060	0.960	0.518	0.754		
% increase in MSE	**HAR**	2.249	-21.159	-22.808	-18.504	-17.090	
Z statistic		1.545	1.761	2.381	2.330	2.063	
p-value		0.122	0.078	0.017	0.020	0.039	
% increase in MSE	**HAR-J**	2.157	-21.230	-22.877	-18.577	-17.164	-0.090
Z statistic		1.444	1.583	2.094	2.164	1.947	-1.597
p-value		0.149	0.113	0.036	0.030	0.052	0.110

Note: Change in efficiency is indicated by percentage change in MSE of the out-of-sample forecasts, based on TSRV measure. The % increase in MSE is computed for the forecasting method mentioned in the second column with respect to the method mentioned in the header row. The statistical significance of this difference is indicated by the Z statistic and p-values of Wilcoxon signed-ranks test. A positive sign of the Z statistic indicates that the 'second column' forecasting model is more efficient than the 'header row' forecasting model.

Table XIII: Comparative forecasting bias (relative to zero) of various models based on TSRV measure

		Random Walk	EWMA	ARIMA	ARFIMA (0, d, 0)	ARFIMA (1, d, 0)	HAR	HAR-J
Panel A: TSRV								
Daily	Z statistic	-2.451	-4.651	-6.186	-8.705	-8.017	-6.959	-7.315
	p-value	0.014	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Weekly	Z statistic	-1.729	-3.441	-3.997	-5.178	-4.005	-4.302	-4.379
	p-value	0.084	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Monthly	Z statistic	-1.397	-2.025	-2.884	-2.520	-2.195	-2.272	-2.350
	p-value	0.162	0.043	**0.004**	0.012	0.028	0.023	0.019

Note: The table represents the Z statistic and p-values of Wicoxon signed-rank tests for statistical difference in relative bias compared with zero for various forecasting models. A negative sign of the Z statistic indicates that the forecasting model is positively biased.