Long Memory Realized Volatility Modeling

Uri Miron
(Advisor: Ron DiRusso, Hyman Beck & Co.)

1 Introduction

Volatility modeling has been an extremely active area of research in quantitative finance over the last quarter century. The popularity of the subject may be seen as a testament both to the explosion of new mathematical and econometric techniques, including ARCH/GARCH and stochastic volatility, and to its applicability to option trading strategies. Such strategies have grown easier for a wide variety of investors to pursue as variance and volatility swaps have become widely available and transaction costs for vanilla options have decreased.

A relatively recent trend in the volatility modeling literature has been the use of *realized volatility* (RV) estimated from high frequency intraday returns[1]. As argued in, for example, Andersen, Bollerslev, Diebold and Labys (2001) [3], we can view the ex-post estimate of RV as an observed, if noisy, variable, rather than a latent variable characterized only by distributional measures.

One notable characteristic of realized volatility is the fact that its autocorrelation tends to decay very slowly. While the exact shape of this decay may vary by asset and period, it is not uncommon to see statistically significant autocorrelation for a period of months or even years (see figure 1). This observation has led to a class of models known as long-memory models, including a number of fractionally integrated models such as the ARFIMA model of log-RV proposed in Andersen, Bollerslev, Diebold and Labys (2003) [4].

A different approach to long memory RV modeling is provided by the Heterogenous Autoregressive Model of Realized Volatility (HAR-RV) of Corsi (2004) [5], which is structured as a factor model, with factors corresponding to realized volatilities over different aggregation periods. One advantage of this approach is that it provides an intuitive mapping between the factors driving volatility and the different types of investors active in the market.

A natural question in volatility modeling, which has received relatively scant attention in the volatility literature, is whether factors that are exogenous to the price process may be effectively incorporated into the model to enhance forecasting power. For example, while much has been written about the volatility of macroeconomic activity, its connection to market volatility has not been widely studied. A recent paper by Engle, Ghysels and Sohn (2008) [6] is one of the few works that have examined this question.

[1] The term realized volatility is used in a number of different ways in the volatility literature. The definition followed in this work, as detailed in section 2, is also referred to as realized variance in some sources.

A different example of an (at least technically) exogenous factor that clearly has significant predictive power is implied volatility, which represents the outlook of (options) market participants. In fact, a number of studies argue that implied volatility is a better predictor of volatility than many price-driven models (see [7]).

Since potential exogenous factors are typically much lower-frequency than price data, it is natural to investigate their potential value in volatility prediction in the context of a long memory model. HAR-RV is especially convenient for this purpose as it already takes the form of a factor model, while incorporating the high-frequency data whose use has become increasingly common in recent years among a wide swath of market participants.

2 Ex-Post Realized Volatility Estimation

If we assume an asset follows the log-price process $dX_t = \mu_t dt + \sigma_t dW_t$, its integrated volatility (IV) is:

$$\langle X, X \rangle_T = \int_0^T \sigma_t^2 dt \tag{1}$$

The discrete RV estimator of the IV is:

$$[X, X]_T = \sum_{i=1}^{n} (X_{t_{i+1}} - X_t)^2 \tag{2}$$

One challenge associated with any RV model is to properly estimate the true ex-post realized volatility in the presence of market microstructure noise (including bid-ask bounce but also effects such as data inaccuracies). The simplest approach to estimating realized volatility under these conditions is to decrease the sampling frequency to one in which microstructure effects become negligible (typically, prices are sampled every five to thirty minutes). Often, the mid bid-ask quote is used instead of the last trade price to further mitigate microstructure effects.

A number of other approaches to RV estimation have been proposed in recent years. One of them is the Two Scales Realized Volatility (TSRV) estimator, developed in Ait-Sahalia et al. (2006) [1]. TSRV uses the standard RV estimators at two different time scales in order to construct an estimate that explicitly adjusts for the bias associated with microstructure. It is based on the idea that at a very high frequency, the sum of the squared returns is an estimator not of the "true" volatility but of the microstructure noise. In addition, the efficiency of the slow scale estimator is increased by subsampling. For example, if the fast time scale is five seconds and the slow time scale is five minutes, the slow scale estimate is the mean of sixty RV estimates corresponding to a grid of initial sampling times spaced five seconds apart. The TSRV estimator formula is:

$$\widehat{\langle X, X \rangle}_T = \left([Y, Y]_T^{(K)} - \frac{\bar{n}_K}{\bar{n}_J} [Y, Y]_T^{(J)} \right) \frac{n}{(K - J)\bar{n}_K} \tag{3}$$

where X_t is the true price; $Y_t = X_t + \epsilon_t$ is the observed price with microstructure noise; J and K are, respectively, the fast and slow time scales (in seconds); and $\bar{n}_J = (n - J + 1)/J$ and $\bar{n}_K = (n - K + 1)/K$ (see Ait-Sahalia et al. (2006) [1] for the derivation).

The optimal value for the slow time scale is $K = O(n^{2/3})$, while the choice of the fast time scale depends on the level of serial correlation in ϵ_t (if the noise is uncorrelated, $J = 1$ is optimal). The estimator is extremely robust to the choice of J and K, however, as illustrated in figure 2.

3 Realized Volatility Forecasting with HAR-RV

The Heterogenous Autoregressive Realized Volatility (HAR-RV) model of Corsi (2004) [5] models realized volatility as a process driven by past realized volatilities over different aggregation periods. The relationship between volatilities at different periods is asymmetric – short-term RV is influenced by long-term RV, but the opposite is not the case. The asymmetry of the volatilities at different time horizons corresponds to the behavior of different market participants. The idea is that while short-term traders are concerned about long-term volatility and adjust their trading strategies accordingly, the strategies pursued by long-term investors, such as pension funds, are not influenced by short-term volatility.

As in the Corsi paper, we consider three frequencies: daily, weekly and monthly. The dynamics at the highest (daily) frequency are:

$$IV_{t+1}^{(d)} = \beta_0 + \beta_d RV_t^{(d)} + \beta_w RV_t^{(w)} + \beta_m RV_t^{(m)} + \omega_{t+1}^{(d)} \tag{4}$$

where $\omega_{t+1}^{(d)}$ is a zero-mean serially independent innovation.

The dynamics at the weekly and monthly frequencies have similar forms, except in each case only the lower frequency components come into play.

The model parameters can be estimated using linear regression. Since the error terms for multi-day forecasts are both heteroskedastic and autocorrelated, a Newey-West covariance matrix must be used to calculate confidence levels for the parameters.

To produce an n-day volatility forecast, the three factors must be updated at each daily step, so that the (direct) contribution of known volatilities to each of the lower frequency factors gradually decreases to zero, while the contribution of forecasted volatilities gradually increases until reaching 100% (after a week for the weekly factor and a month for the monthly factor). The daily volatility factor, of course, is comprised of a projected realized volatility from day $t + 1$ on. All HAR-RV forecasts calculated are out of sample, and the parameters are re-estimated every day.

Following the common practice in most of the realized volatility literature, volatility forecasts are evaluated using Mincer-Zarnowitz regressions of the estimated realized volatilities over the "true" realized volatilities, aggregated over the appropriate number of days (these are of course estimated values as well, but we assume that the error in the ex-post estimates is marginal compared to the errors in the forecasts). The Mincer-Zarnowitz regression equation for the h-day ahead forecast is thus:

$$\sum_{j=1}^{h} RV_{t+j}^{(d)} = b_0 + b_1 \sum_{j=1}^{h} E_t \left(RV_{t+j}^{(d)} \right) + error \tag{5}$$

One advantage of the HAR-RV model is that while it can explain long memory dynamics, it is not constrained by its parametric to long memory only. It can be shown that the process has long memory under some conditions for the beta values, but short memory for others (see Corsi (2004) [5]). Therefore, as illustrated by the empirical results, the model is capable of adjusting quite well to periods where market dynamics are better described as short-memory, such as times of temporary distress or fundamental regime change.

4 Connection to Exogenous Factors

In the spirit of Engle, Ghysels and Sohn (2008) [6], we examine the connection between low-frequency exogenous factors (both macroeconomic and asset-specific) and realized volatility. As in the aforementioned paper, the basic idea is to use a high-frequency price-driven model as a baseline, and examine whether the incorporation of the exogenous factors yields meaningful contribution.

Though the approaches are similar, the time scales of the data examined in the two instances are vastly different. In the case of the Engle et. al paper, the highest frequency data is daily close prices, and the overall period examined is well over a century. In the case of this work, tick data is used for the construction of the baseline model, and the period examined is less than a decade.

The contribution of the exogenous factors is explored using in and out of sample regressions. For consistency, the in sample regressions are limited to those periods available for out of sample regression as well.

5 Empirical Results

5.1 The Data

Two types of data are used, high-frequency stock prices, and low-frequency exogenous factors time series.

5.1.1 Stock Prices

We use TAQ trade data for two exchange traded funds tracking broad U.S. equity indices, SPY (S&P 500) and QQQ (Nasdaq 100). The period examined is 03/10/1999 – 07/31/2008, for a total of 2364 trading days.

For each day, 23,400 one-second logarithmic returns, corresponding to the 6.5 hour regular trading day, are calculated in order to allow for maximum flexibility in choosing the short time scale (for seconds during which no trades occurred, flat interpolation is used to determine the price). The data is filtered based on correction flags provided in the TAQ trades file, as well as for obvious data errors.

Though it is likely that the optimal TSRV time scales change for both ETFs during the period covered (as a result of changes in trading volume and liquidity), a single estimator is used for the entire period, with a fast scale of five seconds and a slow scale of five minutes. As discussed previously, the estimator's robustness ensures this is of little practical concern, which has been verified empirically for a selection of sub-periods for both stocks (see figure 2).

5.1.2 Exogenous Factors

The exogenous factors examined are:

- Industrial production growth (level and volatility).
- Yield curve measures (differences between three and one year constant maturity treasuries, and between ten and three years).
- Fama-French factors (market, high-minus-low, big-minus-small and momentum).
- Implied volatility (VIX for SPY, VXN for QQQ).
- Trading activity measures (trading volume and a state variable indicating whether the day's return was negative, as in TGARCH).

The industrial production data is of monthly frequency, while the other factors are daily.

5.2 HAR-RV Forecasts

out of sample HAR-RV 1, 5, 10, 15 and 20 day ahead forecasts are calculated for 1575 days (from early April 2002 through early July, 2008) for each ETF, using the procedure described in section 3. A rolling window of 750 days is used to calculate the parameters of the HAR-RV equation (4). The results are shown in tables 1 and 2. Tables 3 and 4 show the results of the same Mincer-Zarnowitz regression when only the (relatively) high volatility year of 2007 is considered. We see that while the R^2 values are somewhat lower, they are still quite high. This robustness is consistent throughout the evaluation period.

Perhaps the most convincing evidence of the model's efficacy, however, comes from the plots of the actual and forecasted realized volatilities. Looking at the same 2007 time frame, it is clear that the model maintains reasonable predictive power during both low and high volatility sub-periods, and adjusts quickly to changes from one to the other (figures 3 and 4).

From the scatter plots of the actual and forecasted values (figures 5 and 6), we see that the predictive power of the model seems to peak for forecasts of 5-10 days. While this is not the case for all periods, it seems to be the most common scenario. A reasonable explanation is that these forecasts enjoy an aggregation effect not felt by the one day forecast, while at the same time suffering less from the propagation of forecasting errors than the longer-period forecasts (which enjoy an even greater, but diminishing, aggregation effect).

5.3 Model Extensions

Both in sample and out of sample regressions were performed to investigate the explanatory and predictive power of the exogenous factors described in section 5.1.2. The evaluation period consists of 825 days (from late March 2005 through early July, 2008).

As with the HAR-RV parameter estimation, a rolling window of 750 days is used for the out of sample regressions. The HAR-RV factors are not re-estimated in the presence of the exogenous factors, however, but rather the forecast itself is viewed as a single factor, and we investigate the potential contribution of the exogenous factors to this benchmark. The Mincer-Zarnowitz regressions (5) are identical in form to those used for the evaluation of the HAR-RV forecasts.

In terms of predictive power, the R^2 values obtained for the forecasts based on HAR-RV and each exogenous factor group were, for all forecasting periods except one day, lower than those achieved for the same periods using HAR-RV only. Plots of the forecasted and actual realized volatilities also indicate a relative weakness of the extended models (although the differences are not very large, as the exogenous factor weighting tend to be fairly small).

In the case of the 1-day ahead forecasts, the R^2 values did show a small improvement for all factors examined. Given the size of the improvement, however, it is unlikely this reflects anything more than a spurious regression effect. Since the one day forecast is much more sensitive to outliers than the other forecasts, an improved fit at even a small number of crucial points could easily have a greater impact on the R^2 than a weaker fit for a much larger part of the sample.

For the in sample regressions, regression coefficients were calculated as in equation (5), with the addition of the exogenous factors on the right hand side. The factor weightings for some of the factors do have t-values high enough to suggest that the factors may at least have some explanatory power (especially for 1, 5, and 10 day periods). These include implied volatility, the state variable indicating whether the current day's return is negative, and the Fama-French market and momentum factors. Nevertheless, the out of sample results suggest that whatever information is present in these external factors is already well expressed in the high-frequency price volatility.

6 Conclusions

The performance of the baseline model for the two exchange traded funds examined is quite strong. Both the HAR-RV model and the TSRV estimator play an important role in these results. The results in Ait-Sahalia (2008) [2], which uses an AR(1) model on $\log(IV^{1/2})$ would appear to suggest that, for one day forecasts, an approach that uses high frequency data with effective treatment of market microstructure yields strong forecasts even with a very simple model. While it is unlikely, given the long memory characteristics observed in the realized volatilities, that such a model would yield comparable results for longer period forecasts, it is nevertheless worth noting how dramatic an impact the treatment of ex-post data may have for realized volatility forecasting.

The HAR-RV results are particularly impressive in light of the fact that the model was implemented with little in the way of extra refinement, other than the use of TSRV for RV estimation (a different filtering technique is applied in the paper itself). Potential avenues for refinement that could be explored for the model include adding additional volatility factors (both lower or higher frequency, e.g. a one hour RV factor), as well as employing advanced regression techniques (the former is explicitly suggested in the paper [5] itself).

Given the strength of the baseline forecasts for the stocks examined, it is not too surprising that the use of exogenous factors did not improve the forecasts. Part of the difficulty is the sheer number of potential factors to examine, which raises the risk of data mining. Naturally, this risk is exacerbated by fact that the time series in question are low frequency. Ultimately, even if one finds a set of signals that appears to work, the very process of finding it may end up blurring the line between systematic and fundamental analysis. From a practical perspective, this may or may not be a problem, but it does raise a potential issue that is perhaps more insidious than blatant data mining.

	b_0	b_1	b_0 (t)	b_1 (t)	R^2
1d	-2.10e-03	1.08e+00	-1.44e+00	1.10e+01	6.62e-01
5d	-4.64e-03	1.20e+00	-1.55e+00	6.46e+00	8.41e-01
10d	-4.77e-03	1.19e+00	-1.55e+00	6.47e+00	8.88e-01
15d	-4.49e-03	1.18e+00	-2.43e+00	1.15e+01	9.13e-01
20d	-3.86e-03	1.14e+00	-2.99e+00	1.83e+01	9.10e-01

Table 1: Mincer-Zarnowitz regressions for SPY

	b_0	b_1	b_0 (t)	b_1 (t)	R^2
1d	-1.18e-03	9.22e-01	-4.82e-01	1.30e+01	6.93e-01
5d	-2.12e-03	9.07e-01	-5.19e-01	7.94e+00	8.36e-01
10d	-1.67e-03	8.87e-01	-2.42e-01	4.73e+00	8.71e-01
15d	-1.47e-03	8.79e-01	-1.93e-01	4.25e+00	8.93e-01
20d	-1.13e-03	8.71e-01	-1.60e-01	4.42e+00	9.06e-01

Table 2: Mincer-Zarnowitz regressions for QQQ

	b_0	b_1	b_0 (t)	b_1 (t)	R^2
1d	1.50e-03	9.70e-01	1.59e+00	1.44e+01	4.80e-01
5d	-7.83e-04	1.16e+00	-1.02e+00	2.12e+01	9.13e-01
10d	8.83e-04	1.06e+00	6.89e-01	2.04e+01	8.58e-01
15d	2.63e-03	9.32e-01	1.16e+00	1.44e+01	7.32e-01
20d	4.17e-03	8.12e-01	1.27e+00	9.57e+00	6.15e-01

Table 3: Mincer-Zarnowitz regressions for SPY (2007)

	b_0	b_1	b_0 (t)	b_1 (t)	R^2
1d	-8.61e-05	1.05e+00	-5.45e-02	1.41e+01	3.66e-01
5d	-7.81e-03	1.41e+00	-6.73e+00	1.91e+01	8.41e-01
10d	-6.34e-03	1.35e+00	-6.29e+00	3.24e+01	8.91e-01
15d	-3.40e-03	1.22e+00	-2.05e+00	2.00e+01	8.20e-01
20d	-7.54e-04	1.10e+00	-3.65e-01	1.43e+01	7.49e-01

Table 4: Mincer-Zarnowitz regressions for QQQ (2007)

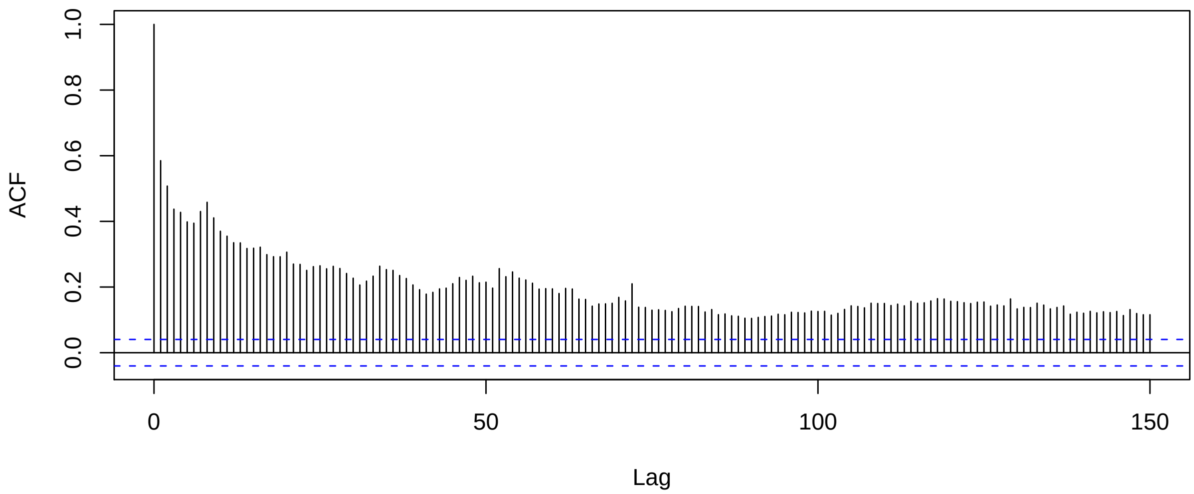

Figure 1: SPY autocorrelation function.

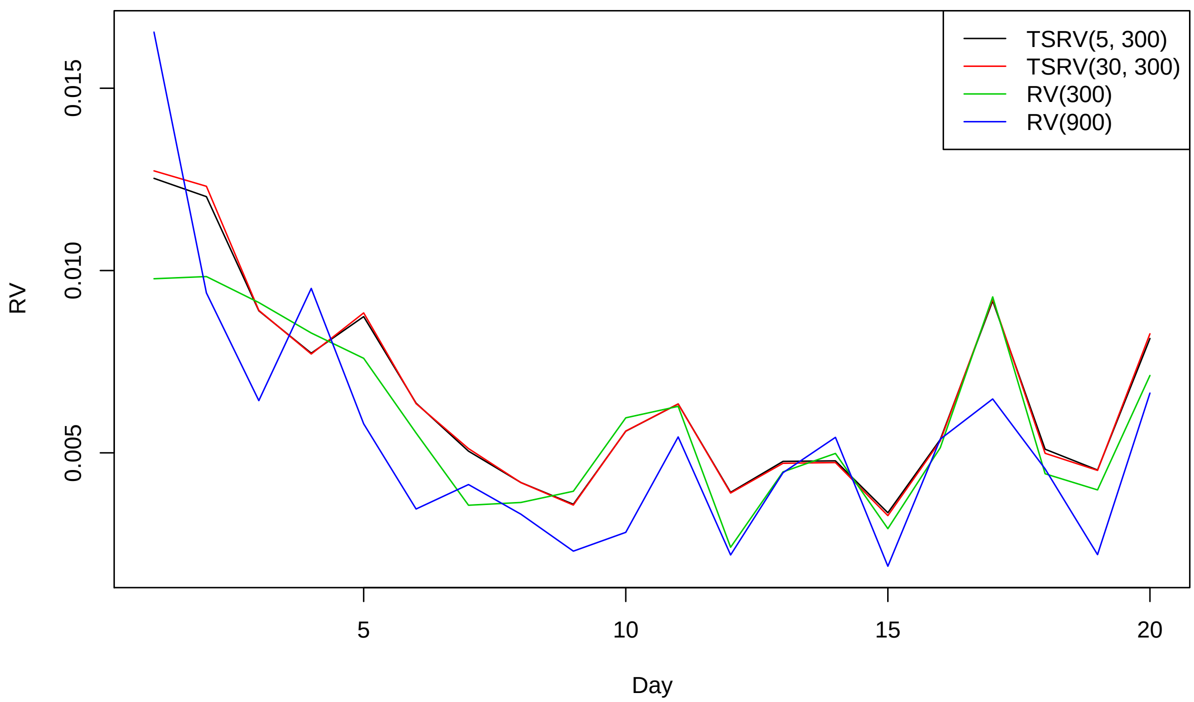

Figure 2: Annualized TSRV and RV estimates of realized volatilities for SPY during January 2007 (all time scales are in seconds).

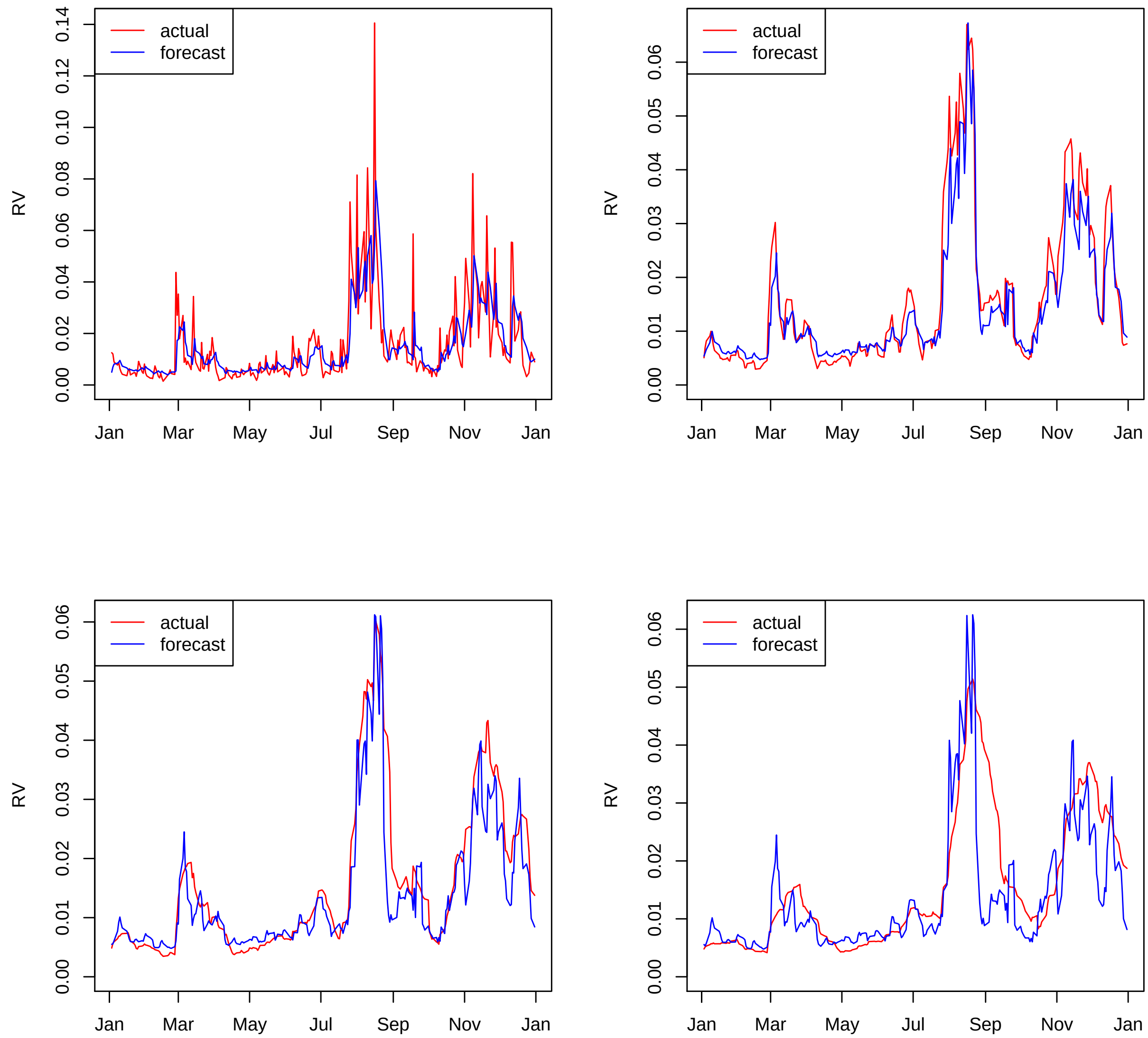

Figure 3: Annualized HAR-RV forecasts and actual values for 1, 5, 10 and 20 days (by row, from left to right) for SPY.

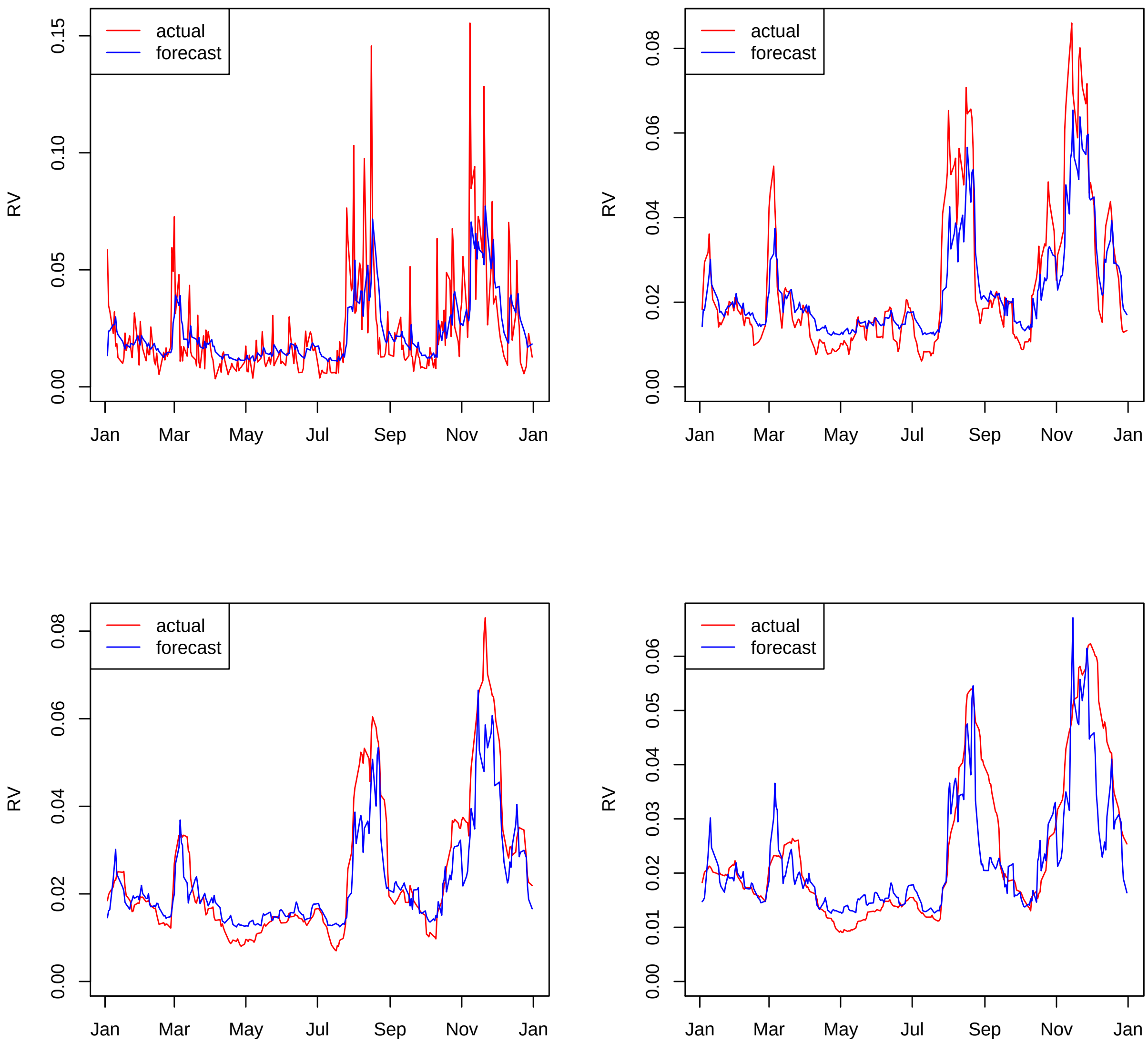

Figure 4: Annualized HAR-RV forecasts and actual values for 1, 5, 10 and 20 days (by row, from left to right) for QQQ.

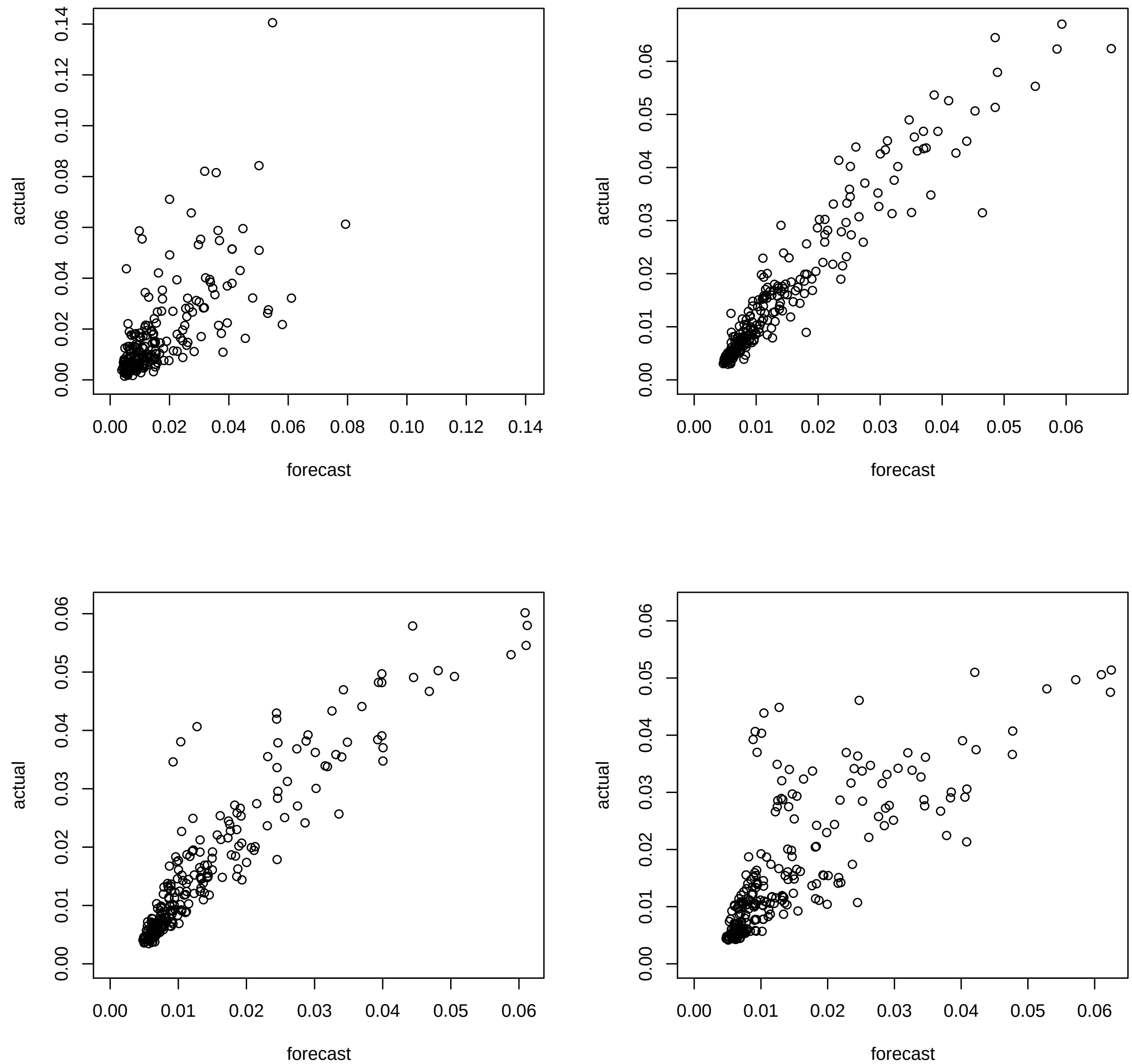

Figure 5: Scatter plots of actual and forecasted realized volatilities for 1, 5, 10 and 20 days (by row, from left to right) for SPY.

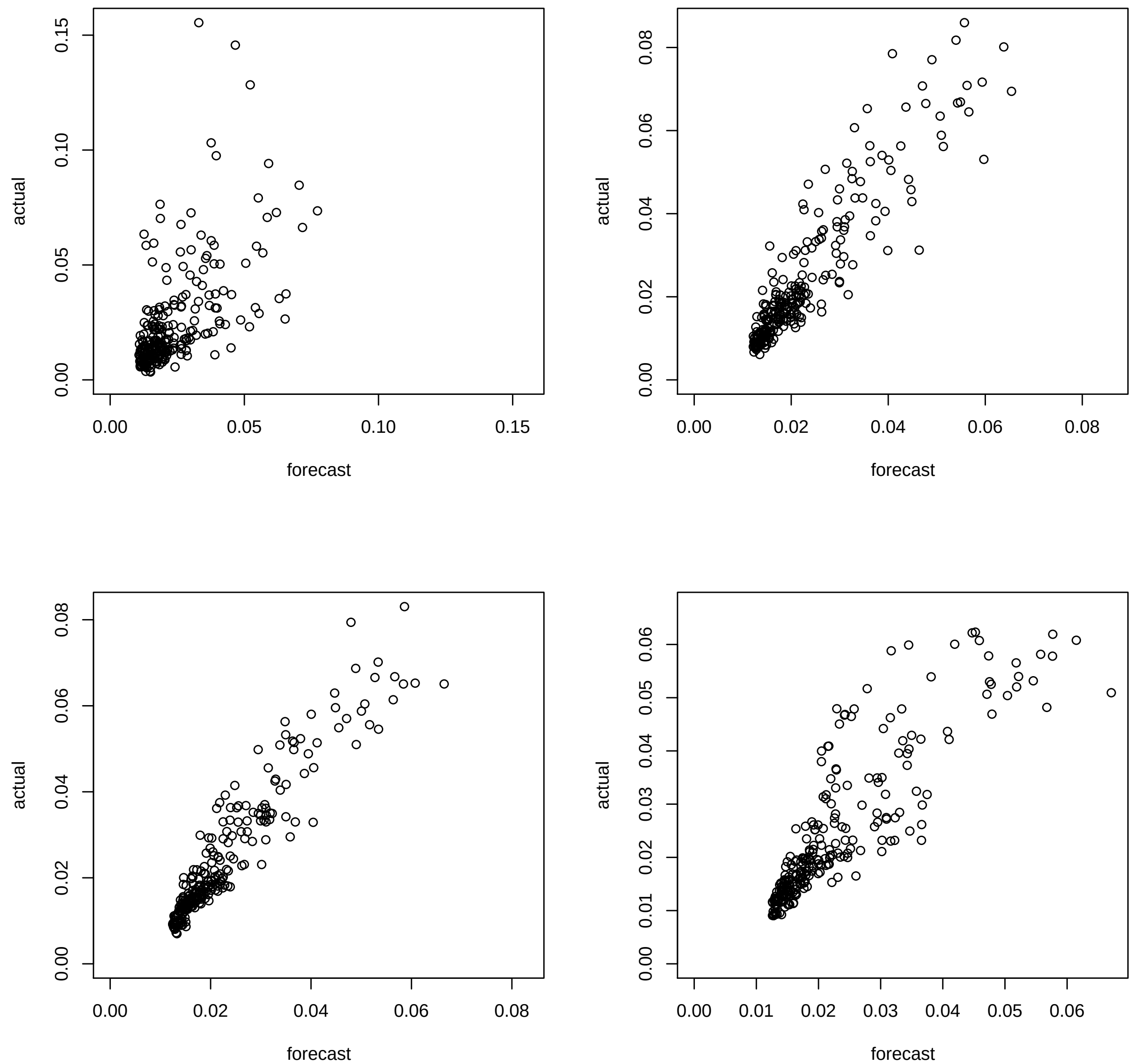

Figure 6: Scatter plots of actual and forecasted realized volatilities for 1, 5, 10 and 20 days (by row, from left to right) for QQQ.

References

[1] Ait-Sahalia, Y., Mykland, P., Zhang, L., 2006. Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise. Working Paper.

[2] Ait-Sahalia, Y., Mancini, M., 2008. Out of Sample Forecasts of Quadratic Variation. Journal of Econometrics 147, 17-33.

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

[4] Andersen, T. G., Bollerslev, T., Diebold F. X., Labys, P, 2004. Modeling and Forecasting Realized Volatility. Econometrica 71, 579-625.

[5] Corsi, F., 2004. A Simple Long Memory Model of Realized Volatility. Tech. rep., University of Lugano.

[6] Engle, R., Ghysels, E., Sohn. B., 2008. On The Economic Sources of Stock Market Volatility. Working Paper.

[7] Poon, S.-H. and C.W.J. Granger, 2003. Forecasting volatility in nancial markets. Journal of Economic Literature 41, 478-539.