

Forecasting return volatility in the presence of microstructure noise*

ZHIXIN KANG[†], LAN ZHANG AND RONG CHEN

Measuring and forecasting volatility of asset returns is very important for asset trading and risk management. There are various forms of volatility estimates, including implied volatility, realized volatility and volatility assumed under stochastic volatility models and GARCH models. Research has shown that these different methods are closely related but have different perspectives, strengths and weaknesses. In order to exploit their connections and take advantage of their different strengths, in this paper, we propose to jointly model them with a vector fractionally integrated autoregressive and moving average (VARFIMA) model. The model is also used for forecasting purpose. In addition, we investigate the impacts of the two realized volatility estimators obtained from intra-daily high frequency data on the forecasts of return volatility. Our methods are applied to five individual stocks and forecasting performances are compared with those from a GARCH(1,1) model and a basic stochastic volatility (SV) model and their extended versions. The proposed VARFIMA model outperforms other volatility forecasting models in this study. Our results show that including the two different realized volatility estimators obtained from the intra-daily high frequency data in the VARFIMA model imposes significant impacts on the forecasting precision for return volatility.

KEYWORDS AND PHRASES: Intra-daily high frequency data, Microstructure noise, Return volatility forecasting, Vector ARFIMA model.

1. INTRODUCTION

Modeling and forecasting the return volatility of financial assets have drawn significant attention from both academia and the financial industry due to its importance in asset pricing, volatility-related derivative trading, and risk management. However, volatility cannot be directly measured and has to be inferred from the returns of an underlying asset or its option prices observed in the market.

*We thank the AE and two anonymous referees for their helpful comments and suggestions which greatly improved the paper. Chen's research was partially supported by NSF grant DMS-0905763, DMS-0915139 and DMS-0800183. Zhang's research was partially supported by the Oxford-Man Institute at the University of Oxford and ICFD at the University of Illinois at Chicago.

[†]Corresponding author.

Various models and methods have been developed for measuring volatility, based on available data and assumptions. Among them, there are four major types of measures and their extensions. *Implied volatility* (IV) is the volatility implied by the observed option prices of the asset, based on a theoretical option pricing model, for example, the seminal Black-Scholes-Merton model (Black and Scholes, 1973; Merton, 1973) or its various extensions including Black (1976), Cox et al. (1979), and Hull and White (1987), among many others.

Realized volatility (RV) uses intra-daily high frequency data to directly measure the volatility under a general semimartingale model setting, using different subsampling methods (Andersen and Bollerslev, 1998; Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2002; Dacorogna et al., 2001; Zhang et al., 2005; Zhang, 2006; Barndorff-Nielsen et al., 2008).

The Autoregressive Conditional Heteroscedasticity model (ARCH) by Engle (1982) and the *generalized ARCH model* (GARCH) by Bollerslev (1986) assess the latent volatility process based on the return series of a financial asset, assuming a deterministic relationship between the current volatility with its past and other variables. *The stochastic volatility model* (SV) extends the ARCH/GARCH model by including randomness in the inter-temporal relationship of the volatility process. For a sample of literature on this topic, see Hull and White (1987), Scott (1987), and Wiggins (1987). In addition, Bollerslev et al. (1994), Ghysels et al. (1995), and Shephard (1996) provide reviews on ARCH/GARCH-type and stochastic volatility models.

The aforementioned approaches provide closely related but different volatility measures. Each approach has their strength and weakness. On the one hand, both ARCH/GARCH-type and SV models successfully capture the temporal dependence in the volatility process. However, they cannot accommodate the intra-daily variability in the asset returns and tend to have poor forecast for ex-post squared returns over a day or longer time horizon. In contrast, by construction, daily realized volatility naturally contains the information about the intra-day variations which ARCH/GARCH lacks. The realized volatility by itself, however, cannot tell us the inter-temporal dependence of the volatility process across days or longer horizon. Finally, although implied volatility cannot directly measure the variability of underlying asset returns, it does reflect, to

some degree, the (options) market’s expectations on the asset volatility. In order to exploit the relationship of different volatility measures and take advantage of their distinctive strength, in this paper we first investigate the characteristics of the volatility measures from the four approaches. Our study shows that the relationship between the measures are closer than previously recognized in the literature, with no apparent leading terms. Such an observation indicates a joint vector modeling approach, instead of a transfer function type of modeling in which one variable is the output and others are input. In addition, all of these four volatility measures display certain long memory characteristic.

Our preliminary findings provide the motivation of modeling the volatility measures jointly using a vector time series model. We consider two groups of measures: one including GARCH(1,1) volatility, realized volatility and implied volatility (Class I), and the other including the SV volatility, realized volatility and implied volatility (Class II). To capture the long memory characteristics of the volatility processes, a vector fractional integrated ARMA model (VARFIMA) is used. Forecasting performance comparison is then carried out with real data. It shows that the proposed model indeed produces improved forecasting performances. It is noted that the long memory behavior of the volatility process can also be modeled by a regime switching process (Hidalgo and Robinson, 1996) but it is beyond the scope of this paper.

It is important to mention that there are studies that combine different volatility measures for better modeling and forecasting performance. For example, Andersen et al. (2003) found that by incorporating the realized volatility measure based on five-minute returns, volatility forecast improved over the conventional GARCH forecast. Our current paper adopts an improved measure of realized volatility, called *two-scale realized volatility* (TSRV, Zhang et al. (2005)). TSRV is computed from tick-by-tick intra-day returns – a much denser and richer returns series – and it corrects the bias from the market microstructure noise which is typically present in the high frequency data. The enhanced accuracy in TSRV over conventional RV should provide further improvement in a volatility forecast. Also in the literature, by introducing lagged realized volatility and implied volatility in the basic GARCH and SV models, Koopman et al. (2005) found that the inclusion of realized volatility in GARCH improved the forecasting of the daily return volatility, whereas the incorporation of implied volatility in GARCH and the SV model helped very little. Different from Koopman et al. (2005), we model the volatility measures jointly and thus are able to capture the long memory characteristics of the process.

The rest of the paper is organized as follows. Section 2 provides some preliminaries, including details of the volatility measures used in the paper, a description of the data set used in our study and some findings on the structures

and relationships of the measures. Section 3 introduces a VARFIMA model for the volatility measures and provides details on the model estimation approach. Section 4 compares the one-day and five-day ahead out-of-sample return volatility forecasts using the proposed model with some existing volatility forecasting models. Section 5 contains a brief conclusion and remarks.

2. PRELIMINARIES

2.1 Volatility measurements

Our study focuses on four different daily volatility measures, namely implied volatility, realized volatility, volatility based on a GARCH model and that based on a stochastic volatility (SV) model. Details are as follows.

(i) Implied Volatility

Implied volatility (IV) of an underlying asset is the volatility implied from its option prices observed in the market. It is typically derived from calibrating a theoretical option pricing formula against the market price of the option. Because an option with a different strike price (or expiration date) can yield a different IV, an IV index is often calculated from a weighted average of IVs of various options and serves as a representative IV measure in practice. We used IV index provided by IVolatility.com, where the weighting scheme takes into account the delta and vega of each participating option. For basic concepts in options pricing, we refer to Hull (2008).

(ii) Realized Volatility

Realized volatility (RV), different from the implied volatility that conveys the market’s assessment of future volatility, measures the market’s historical volatility in the past. They are constructed by using intra-daily high frequency data. In this study we use two different versions with the intention to exploit their differences in forming volatility forecasts. Both assume that the logarithmic (efficient) prices of a financial asset follow a semi-martingale process. This rather general assumption is required by the no-arbitrage law in financial theory. The difference between RV and TSRV is that the former assumes one observes the efficient prices precisely whereas in TSRV construction, one considers a hidden semi-martingale setting, namely, one observes efficient prices (modeled as semi-martingale) plus noise.

Specifically, let \mathcal{G} be a complete collection of the trading times in a day, $\mathcal{G} = \{t_0, t_1, \dots, t_n\}$, with $t_0 = 0$ and $t_n = T$. Let $\{y_{t_j}\}$ be the logarithmic price of a financial asset observed at time t_j , $t_j \in \mathcal{G}$. Also let \mathcal{H} be a subset of \mathcal{G} , with sample size $n_{\text{sparse}}, n_{\text{sparse}} \leq n$. The standard RV^2 , realized variance, is then calculated as the sum of the squared returns within that day:

$$(1) \quad RV^2 = \sum_{t_j, t_{j,+} \in \mathcal{H}} (y_{t_{j,+}} - y_{t_j})^2,$$

where t_j and $t_{j,+}$ are the adjacent elements in \mathcal{H} , with $t_j < t_{j,+}$. We obtain the standard RV by taking the square root of that in (1).

In the absence of market microstructure in the data, conventional RV in (1) is a consistent estimator of the daily variation of returns, as the sampling interval shrinks (Jacod and Protter, 1998). However, empirical studies suggest that market microstructure noise is prevalent in high frequency data (Andersen and Bollerslev, 1998; Dacorogna et al., 2001). As prices are sampled at a finer interval, microstructure noise becomes progressively dominant and as a consequence, RV becomes increasingly unreliable with a bias inversely proportional to the sampling interval length (Zhang et al., 2005). In the empirical finance literature, the sampling period is typically equal to or larger than 5 minutes in order to reduce the impact of microstructure noise. Following this literature, we choose five-minute sampling intervals to compute the daily RV measure using the intra-daily high frequency data¹.

Zhang et al. (2005) proposed an approach to correct the microstructure bias by combining the RV estimators from two different time scales, resulting in two-scale realized volatility (TSRV). Specifically, TSRV is obtained by taking the square root of the two-scale realized variance, which is calculated as:

$$(2) \quad \text{TSRV}^2 = \left(1 - \frac{\bar{n}}{n}\right)^{-1} \left(RV_K^2 - \frac{\bar{n}}{n}RV_1^2\right)$$

where $\bar{n} = \frac{n-K+1}{K}$ and

$$(3) \quad RV_K^2 = \frac{1}{K} \sum_{t_j, t_{j+K} \in \mathcal{G}} (y_{t_{j+K}} - y_{t_j})^2.$$

RV_1^2 is a special case of (3) with $K = 1$. Note that RV_K^2 is the realized variance based on sampling every K -th price while RV_1^2 is the RV based on all available prices in \mathcal{G} .

Our preliminary analysis show that the estimated TSRVs are fairly robust to the choice of K , especially when K is equal to or greater than 200.

(iii) GARCH Model and Its Extensions

The generalized autoregressive conditional heteroscedasticity (GARCH) model was proposed by Bollerslev (1986). A GARCH(p,q) model assumes a form of:

$$(4) \quad \begin{aligned} y_t &= \sigma_t \varepsilon_t, \quad t = 1, \dots, T \\ \sigma_t^2 &= \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_p y_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \end{aligned}$$

where y_t is the daily de-measured returns of a financial asset, σ_t the instantaneous volatility of the return process at time t , p the order of the ARCH term, q the order of the GARCH

term. This model successfully describes most of the recognized stylized features in asset return series, as mentioned in Section 1.

A GARCH model can be extended by including realized volatility (RV or TSRV) and implied volatility (IV) in the variance equation, as follows,

$$(5) \quad \begin{aligned} y_t &= \sigma_t \varepsilon_t, \quad t = 1, \dots, T \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \sum_{k=1}^M \phi_k RV_{t-k}^2 \\ &\quad + \dots + \sum_{l=1}^N \gamma_l IV_{t-l}^2. \end{aligned}$$

Estimation of the above models can be done through maximum likelihood estimation (Doornik and Ooms, 2003; Laurent and Peters, 2006).

In Section 4 where forecasting performance is evaluated, we also consider a different type of extension to GARCH(1,1), namely, the fractional integrated GARCH(1,1) (FIGARCH(1,1)) model (Baillie et al., 1996). This model is able to capture long memory property in the return volatility. The general form of a FIGARCH model can be written as:

$$(6) \quad \begin{aligned} y_t &= \sigma_t \varepsilon_t, \quad t = 1, \dots, T \\ \sigma_t^2 &= \alpha_0 \left[1 - \sum_{j=1}^q \beta_j B^j\right]^{-1} \\ &\quad + \left[1 - \left[1 - \sum_{j=1}^q \beta_j B^j\right]^{-1} \sum_{i=1}^p \alpha_i B^i (1 - B)^d\right] \varepsilon_t^2 \end{aligned}$$

where y_t is the demeaned returns, ε_t follows an *i.i.d.* standard normal distribution and B is the backshift operators defined as: $Bx_t = x_{t-1}$. The fractional differencing parameter d is a non-integer real number. Similar to GARCH model and its extensions, a FIGARCH model can be estimated using the maximum likelihood method, and the G@rch package (Laurent and Peters, 2006) in Ox software is employed for estimation and forecasting procedures in this study.

(iv) Stochastic Volatility Model and Its Extensions

A basic stochastic volatility (SV) model (Taylor, 1986) is in a form of:

$$(7) \quad y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \exp(h_t), \quad h_t = \mu + \varphi h_{t-1} + \sigma_\eta \eta_t,$$

where y_t and σ_t^2 are the de-measured returns of a financial asset and its instantaneous variance, respectively, at time t . The noise processes ε_t and η_t are independent and follow *i.i.d.* standard normal distributions. The logarithm of the instantaneous variance h_t has a persistence parameter φ , which is positive and takes a value less than 1.

¹Note that RV is not a sufficient statistic whereas TSRV is.

Similar to [Koopman et al. \(2005\)](#), we extend SV model to include realized volatility and/or implied volatility as follows:

$$(8) \quad \begin{aligned} y_t &= \sigma_t \varepsilon_t, \quad \sigma_t^2 = \exp(h_t), \\ h_t &= \mu + \varphi h_{t-1} + \sum_{i=1}^M p_i \log(RV_{t-i}^2) \\ &\quad + \sum_{j=1}^N q_j \log(IV_{t-j}^2) + \sigma_\eta \eta_t, \end{aligned}$$

where M and N represent the maximum lags of realized volatility (RV or TSRV) and implied volatility (IV), respectively.

Gibbs sampler is used for estimating the SV-type models, as well as for obtaining predictions. In each case, 50,000 samples are generated, with 2,000 burn-in samples.

2.2 Data

For our empirical study, we use five individual stocks traded in the New York Stock Exchange (NYSE), namely Microsoft (MSFT), Citi (C), Disney (DIS), Pfizer (PFE), and General Electric (GE). All of them are highly liquid, representing five different industries. The time period considered is from January 2, 2001 to December 31, 2003, with 752 daily observations in total for each series. Intra-daily high frequency return data are also obtained in this period.

The daily return data set and the intra-day high frequency data set are downloaded from Wharton Research Database Services (WRDS). The intra-daily high frequency data is the consolidated trades in the NYSE's TAQ database. When constructing the daily realized volatility (RV and TSRV) with the intra-day high frequency data set, we remove those prices with more than 1% bounceback, defined as $|y_{t_i} - y_{t_{i-1}}| > 1\%$ and $|y_{t_{i+1}} - y_{t_i}| > 1\%$ in addition to the conditions that the consecutive returns $y_{t_i} - y_{t_{i-1}}$ and $y_{t_{i+1}} - y_{t_i}$ hold opposite signs. Such incidents are often due to data recording errors.

For the five stocks considered in the period of January 2, 2001 to December 31, 2003, the average daily observation frequencies in the high frequency data are summarized in [Table 1](#). As is shown in [Table 1](#), the ranges of daily observations differ from stock to stock. We use $k = 200$ for TSRV calculation, as explained in [Section 2.1 \(ii\)](#).

Table 1. Summary of average daily observation frequency

Series	Avg. Obs.
Citi	7,634
Disney	4,612
GE	12,910
Microsoft	43,954
Pfizer	8,530

2.3 Characteristics and relationship of the volatility measures

When exploring the characteristics of different volatility measures, we used the implied volatility series published in [www.Ivolatility.com](#), the estimated daily realized volatilities RV and TSRV, and the instantaneous daily volatility measures under GARCH(1,1) and basic SV model for the five stocks. Our findings for the five stocks are similar, so we only report that of Microsoft. [Figure 1](#) shows the autocorrelation functions of four MSFT volatility series. The ACF plot of RV is omitted since it is quite similar to TSRV. From [Figure 1](#), all the volatility measures have strong and persistent autocorrelations, an evidence for the volatility clustering phenomenon.

In order to investigate the relationship between the four different volatility measures, [Figure 2](#) presents their cross-correlation functions. It is evident that strong cross-correlation exists in each pair of the volatility measures. Within 20 leads/lags, the cross-correlations between any two volatility measures are at least 0.4. The maximum correlation between TSRV and other measures does not occur at lag zero, instead, a lagged TSRV seems to have strong cross-correlation with other volatility measures.

Both theoretical and empirical literature have documented long-memory volatility property. Early work in this area includes [Robinson \(1991\)](#) and [Ding et al. \(1993\)](#).

This long-range dependence in volatility process can often be characterized by certain fractional integration models where a fractional integrated process x_t of order d can be transformed into a stationary process via fractional differencing. As its name suggests, the differencing parameter d is a non-integer real number.

We apply the long memory R/S test proposed by [Lo \(1991\)](#) to each of the volatility processes, using 20 as the bandwidth for the cross variance. We also obtain the GPH fractional differencing estimator d using the methodology proposed by [Geweke and Porter-Hudak \(1983\)](#). [Tables 2 and 3](#) present the R/S statistics and the GPH estimates, respectively, for each of the five stocks' volatility measures.

[Table 2](#) shows that all of the R/S statistics are significant at 0.05 level, with critical value 1.747, thus suggesting all four volatility measures exhibit fractional integration characteristics. In [Table 3](#), the GPH estimates of the fractional differencing parameters are listed, where the values in the parentheses are standard deviations corresponding to each of the GPH estimates. The GPH estimates show that, although the different volatility measures are of long memory, the extent of fractional integration varies for different volatility measures. The GPH estimates for GARCH(1,1) volatility and basic SV volatility are larger than those of realized volatilities (RV and TSRV), while the estimate for implied volatility lies in between. This pattern holds for all the five individual stocks.

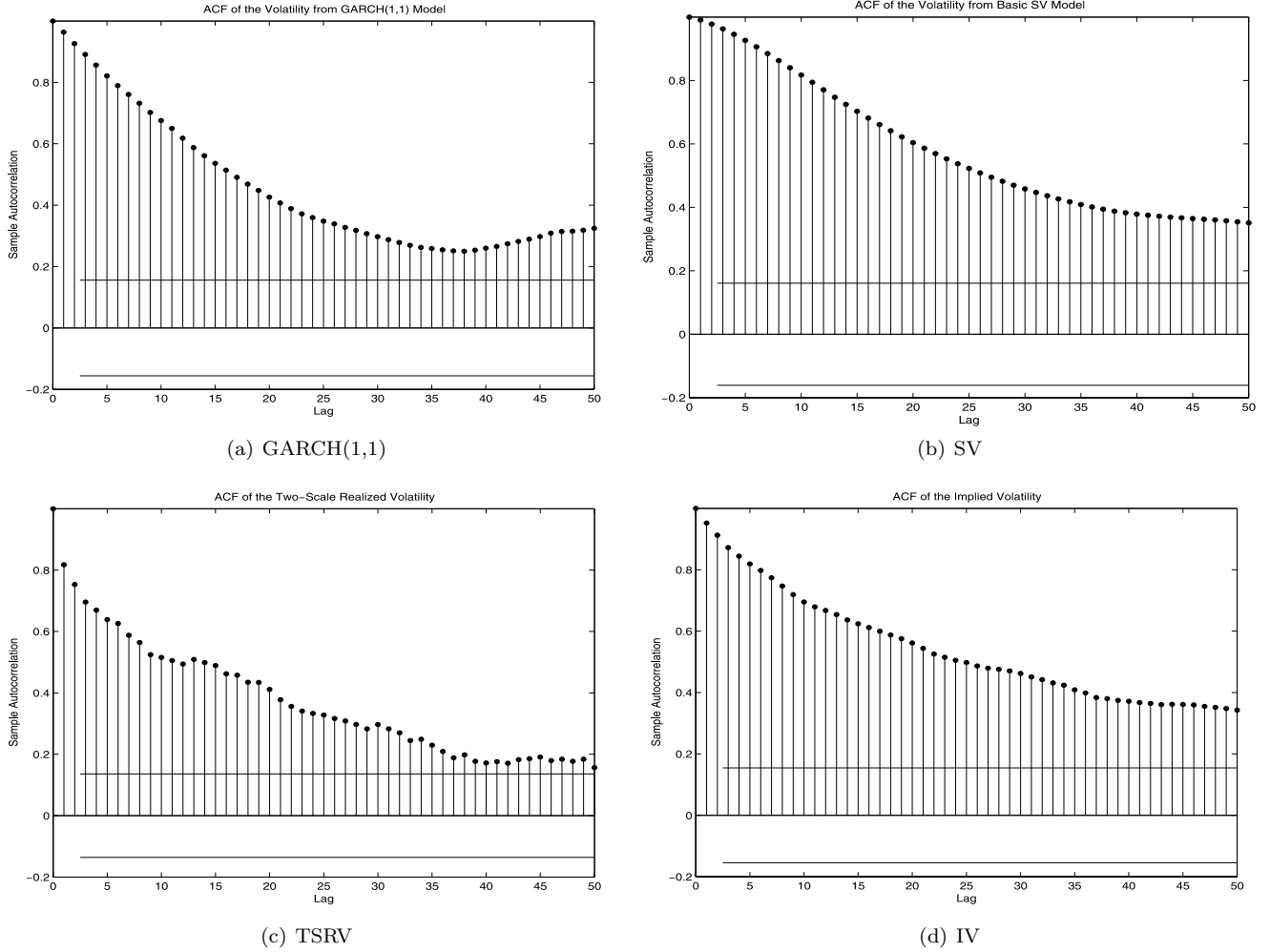


Figure 1. Estimated autocorrelations of the four volatility measures of Microsoft. Time period: 01/02/2001-12/31/2003.

3. A VECTOR ARFIMA(p,d,q) MODEL FOR VOLATILITY SERIES

The empirical evidence in Section 2 suggests that fractional integration patterns exist in all the volatility series and the volatility measures have strong short-term and long-term cross-correlations. The extensions of GARCH and SV models in the literature, as cited in Section 1, either do not incorporate the rich information from the ultra-high-frequency returns with microstructure noise correction, or fail to capture the long memory property in the return volatilities.

As a consequence, the forecasting performance from these models is limited. We propose to model GARCH and SV return volatility jointly with realized volatility and implied volatility, using a vector fractionally integrated autoregressive and moving average model (VARFIMA). This model is able to capture the long memory characteristic property of the individual volatility measure, as well as the interactive relationship between each other. We will allow different

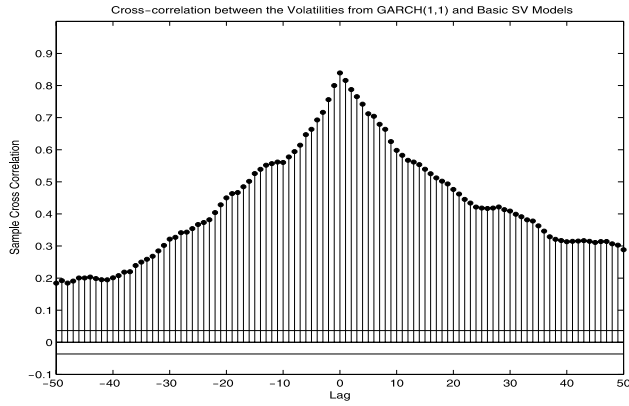
differencing parameter d 's to reflect different degrees of fractional integration among the volatility measures.

Without losing generality, a VARFIMA(p,d,q) model may be expressed as:

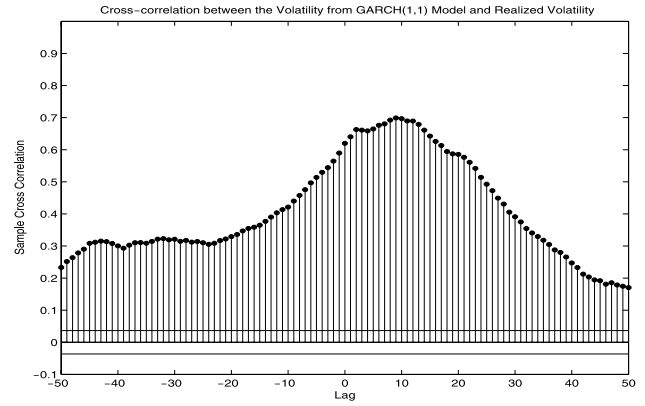
$$(9) \quad (\mathbf{I} - \Phi_1 B - \dots - \Phi_p B^p) \mathbf{M}(B) \mathbf{y}_t = (\mathbf{I} - \Theta_1 B - \dots - \Theta_q B^q) \boldsymbol{\varepsilon}_t$$

where \mathbf{y}_t is a de-meaned $k \times 1$ vector consisting of k time series at time t . Here we assume $\mathbf{M}(B)$ to be a $k \times k$ diagonal matrix with diagonal elements being $(1 - B)^{d_1}, (1 - B)^{d_2}, \dots, (1 - B)^{d_k}$ where d_i is the fractional differencing parameter of i^{th} dimension. The noise vector $\boldsymbol{\varepsilon}_t$ is assumed to be Gaussian with $\boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma})$. Stationarity and other properties of such processes are similar to that of univariate ARFIMA model, established by (Dahlhaus, 1988, 1989; Fox and Taqqu, 1986, 1987; Li and McLeod, 1986; Yajima, 1985).

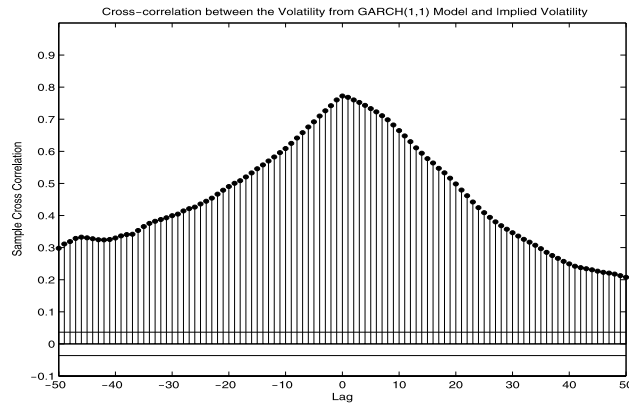
We implement two classes of VARFIMA in this paper with, both of three dimensional ($k = 3$):



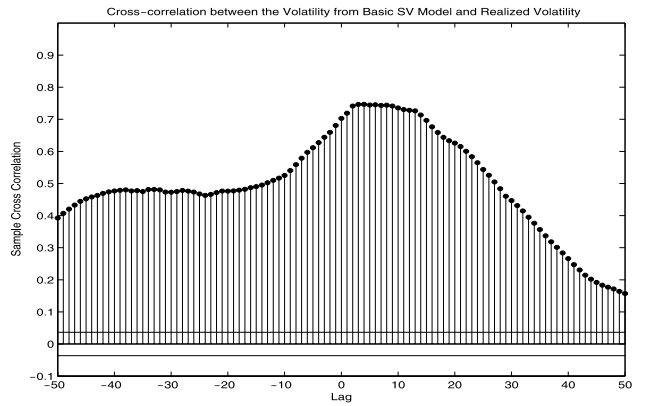
(a) GARCH(1,1) vs. SV



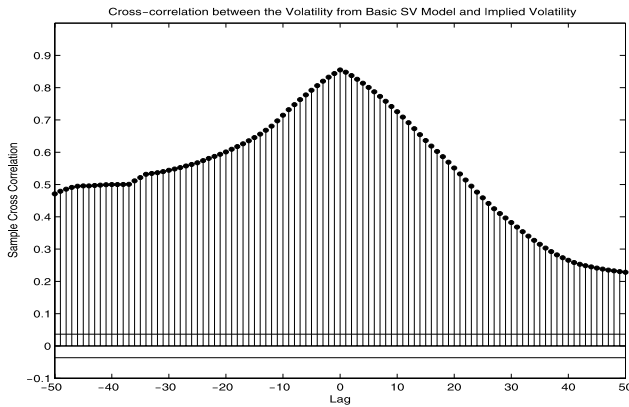
(b) GARCH(1,1) vs. TSRV



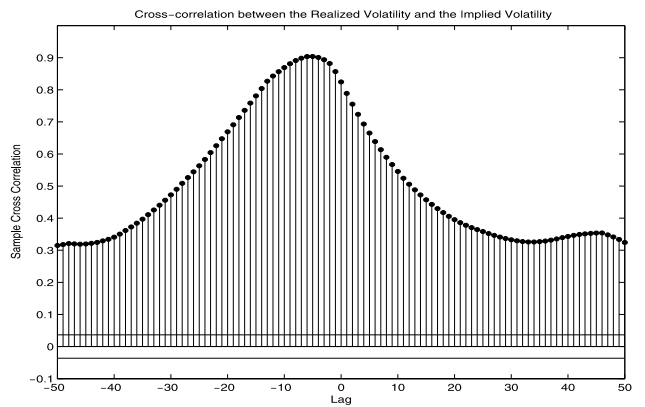
(c) GARCH(1,1) vs. IV



(d) SV vs. TSRV



(e) SV vs. IV



(f) TSRV vs. IV

Figure 2. Estimated cross-correlations of the four volatility measures of Microsoft. In each figure, the first volatility measure leads the second one when the lags are positive; the second volatility measure leads the first one when the lags are negative. Time period: 01/02/2001-12/31/2003.

VARFIMA I:

In this class of models we use the three dimensional time series $\mathbf{y}_t = (V_{GARCH,t}, V_{Realized,t}, V_{Implied,t})'$ in model (9). where $V_{GARCH,t}$, $V_{Realized,t}$, and $V_{Implied,t}$ represent the GARCH volatility, realized volatility, and implied volatility at time t , respectively. All three are

treated as observed, either directly or estimated as described in Section 2. The AR coefficients $\Phi_{I,1}, \dots, \Phi_{I,p}$ and MA coefficients $\Theta_{I,1}, \dots, \Theta_{I,q}$ are 3 by 3 matrices, and ε_t is a 3 by 1 vector. The AR and MA orders (p, q) will be determined with model selection criteria.

Table 2. R/S statistics of the volatility measures

Series	GARCHVol	SVVol	RVol	TSRVol	ImpVol
Citi	1.8990	1.9504	2.0244	1.9799	1.9231
Disney	1.9303	1.9084	2.0403	2.0138	1.9872
GE	1.8226	2.0640	1.9264	1.9407	1.8845
Microsoft	1.7735	1.9458	1.9229	1.8358	1.8714
Pfizer	1.8776	1.9794	1.8238	1.9628	1.8765

VARFIMA II:

Here we use $\mathbf{y}_t = (V_{Stochastic,t}, V_{Realized,t}, V_{Implied,t})'$ in model (9), where $V_{Stochastic,t}$, $V_{Realized,t}$, and $V_{Implied,t}$ represent the stochastic volatility, realized volatility, and implied volatility at time t , respectively.

In both cases, we use the standard RV or the TSRV for realized volatility $V_{Realized}$ and compare the impacts of these two realized volatility estimators on the volatility forecasts. The corresponding models are hence labelled as Class I-RV and I-TSRV and Class II-RV and II-TSRV, respectively.

Maximum likelihood method is employed for model estimation. Theoretical properties of the estimators are direct extensions of the results obtained for univariate ARFIMA models in (Beran, 1995; Chung, 1996; Dahlhaus, 1989; Li and McLeod, 1986; Robinson, 2001; Sowell, 1992; Yajima, 1985). The estimation is computationally intensive due to the nonlinearity in the fractional integration parameters d_1, d_2 and d_3 and the large number of parameters in coefficient matrices $\Phi(B)$ and $\Theta(B)$. Specifically we use a grid search for d_1, d_2 , and d_3 around the initial value of the GPH estimates of each individual univariate series, shown in Table 2. For each combination of grid values of (d_1, d_2, d_3) , its likelihood function value is obtained by estimating the corresponding VARFIMA(p,(d1,d2,d3),q) model. The optimal values of d_1, d_2 , and d_3 are those which generate the maximum likelihood value for the specified VARFIMA(p,q) model. Given optimal d_1, d_2 , and d_3 values, the rest of the parameters are estimated through a standard estimation procedure for vector ARMA model with refinements.

Here we only report the results of Microsoft stock. Findings for other stocks show similar features and are omitted to avoid redundancy.

Table 4 lists the estimates of d 's and the corresponding maximum likelihood values with different specifications of p and q as well as their corresponding AIC (Akaike, 1969) values. In Table 4, $d_{I,k}, k = 1, 2, 3$, are the estimated fractional differencing parameter corresponding to GARCH volatility, TSRV and implied volatility in the Class I, respectively. Similarly, $d_{II,k}, k = 1, 2, 3$ are those in the Class II, respectively.

According to the AIC values in Table 4, the best model in Class I is VARFIMA(2,(1.09, 0.68, 0.90),1) and the best model in Class II is VARFIMA(2,(0.98, 0.65, 0.91),1). We can see that the orders p and q do have certain effect on the fractional integration order.

Table 3. GPH estimators of fractional differencing parameters

Series	GARCHVol	SVVol	RVol	TSVol	ImpVol
Citi	1.0159 (0.0467)	1.0935 (0.0328)	0.5367 (0.0479)	0.6441 (0.0570)	0.7538 (0.0510)
Disney	0.7923 (0.0540)	1.0165 (0.0285)	0.4770 (0.0464)	0.5180 (0.0527)	0.7029 (0.0503)
GE	0.8664 (0.0492)	0.9730 (0.0253)	0.5023 (0.0438)	0.5173 (0.0494)	0.6589 (0.0473)
Microsoft	0.9299 (0.0497)	0.9934 (0.0227)	0.5538 (0.0429)	0.6082 (0.0465)	0.7864 (0.0506)
Pfizer	0.9519 (0.0512)	1.0916 (0.0282)	0.4922 (0.0502)	0.5221 (0.0474)	0.7970 (0.0512)

To obtain more accurate coefficient estimation and to avoid model ambiguity often encountered in vector ARMA models (Tiao and Tsay, 1989; Tsay, 1989), we employ a coefficient-refining procedure with fixed d_i 's to zero-out the insignificant coefficients in the AR and MA coefficient matrices. Specifically, with fixed d_i 's, a backward elimination procedure is used to remove insignificant coefficients one at a time with models re-estimated iteratively, until all remaining coefficients are significant at 5% level.

The estimated AR and MA coefficient matrices with their standard errors, as well as the estimated covariance matrix for the error term in our VARFIMA model, from the two best models in Classes I and II are given below. As shown in the estimated coefficient matrices, the estimated significant coefficients do reflect the inter- and cross- correlation among the three different volatility measures in both models. Specifically, for the best model in Class I, the non-zero lag-1 AR and MA coefficients are only related to realized volatility, except the MA(1) coefficient of the realized volatility itself. Hence it seems that the past observations of the realized volatility has more influences on all three volatility measures under this model. In addition, the estimated noise covariance matrix shows that the noises in realized volatility series and implied volatility series have higher correlation than the other two pairs. The structure of the coefficient matrices in the best model of Class II is very different from that in Class I, showing the difference between volatility measures based on GARCH models and SV models. However, the noises in realized volatility series and implied volatility series still have higher correlation than other two pairs in the best model in Class II, showing a constant pattern.

$$\hat{\Phi}_{I,1} = \begin{pmatrix} 0 & 0.21(0.06) & 0 \\ 0 & 0 & 0 \\ 0 & -0.55(0.14) & 0 \end{pmatrix};$$

$$\hat{\Phi}_{I,2} = \begin{pmatrix} -0.08(0.03) & 0.08(0.02) & 0 \\ 0 & 0 & 0.18(0.05) \\ 0 & -0.13(0.03) & 0.29(0.06) \end{pmatrix};$$

Table 4. Estimation results of the VARFIMA(p,(d₁, d₂, d₃),q) models

Model	$\hat{d}_{I,1}$	$\hat{d}_{I,2}$	$\hat{d}_{I,3}$	ML_I	AIC_I	$\hat{d}_{II,1}$	$\hat{d}_{II,2}$	$\hat{d}_{II,3}$	ML_{II}	AIC_{II}
VARFIMA(1,d _i ,0)	0.91	0.59	0.81	10,364.00	-36.79	0.96	0.63	0.86	10,733.43	-37.92
VARFIMA(2,d _i ,0)	0.90	0.55	0.84	10,368.24	-36.93	0.92	0.59	0.85	10,738.27	-37.99
VARFIMA(3,d _i ,0)	0.98	0.54	0.85	10,373.11	-37.03	0.96	0.57	0.89	10,741.03	-37.99
VARFIMA(4,d _i ,0)	0.92	0.57	0.83	10,376.89	-36.86	0.95	0.61	0.83	10,743.95	-37.96
VARFIMA(5,d _i ,0)	1.02	0.58	0.86	10,378.20	-36.92	0.99	0.59	0.88	10,747.21	-38.03
VARFIMA(1,d _i ,1)	0.88	0.61	0.76	10,373.90	-36.57	0.93	0.57	0.89	10,744.95	-37.95
VARFIMA(2,d _i ,1)	1.09	0.68	0.90	10,393.29	-37.18	0.98	0.65	0.91	10,753.92	-38.26
VARFIMA(3,d _i ,1)	0.96	0.53	0.77	10,381.09	-37.05	0.96	0.57	0.82	10,749.10	-37.98
VARFIMA(1,d _i ,2)	1.03	0.67	0.86	10,378.27	-36.86	0.99	0.63	0.85	10,746.73	-37.96
VARFIMA(2,d _i ,2)	0.89	0.52	0.73	10,376.39	-36.83	0.92	0.51	0.76	10,745.39	-37.82
VARFIMA(3,d _i ,2)	0.93	0.46	0.72	10,391.55	-37.04	0.95	0.55	0.68	10,749.51	-37.99
VARFIMA(1,d _i ,3)	0.90	0.44	0.71	10,383.09	-36.98	0.92	0.47	0.77	10,746.25	-37.93
VARFIMA(2,d _i ,3)	0.97	0.46	0.75	10,396.42	-37.13	0.97	0.49	0.71	10,750.44	-37.93
VARFIMA(3,d _i ,3)	1.01	0.55	0.81	10,393.87	-36.99	1.03	0.57	0.85	10,748.99	-37.90

$$\hat{\Theta}_{I,1} = \begin{pmatrix} 0 & 0.22(0.06) & 0 \\ 0 & 0.28(0.04) & -0.50(0.07) \\ 0 & -0.59(0.14) & 0 \end{pmatrix};$$

$$\hat{\Sigma}_I = \begin{pmatrix} 0.20E-05 & 0.43E-06 & 0.86E-07 \\ 0.43E-06 & 0.12E-04 & 0.20E-05 \\ 0.86E-07 & 0.20E-05 & 0.30E-05 \end{pmatrix};$$

$$\hat{\Phi}_{II,1} = \begin{pmatrix} 0.85(0.02) & 0 & 0 \\ 0 & 0 & -0.82(0.35) \\ 1.02(0.24) & -0.38(0.10) & 0 \end{pmatrix};$$

$$\hat{\Phi}_{II,2} = \begin{pmatrix} 0 & -0.005(0.002) & 0 \\ 1.39(0.36) & 0 & 0.26(0.07) \\ 0 & -0.07(0.03) & 0.18(0.06) \end{pmatrix};$$

$$\hat{\Theta}_{II,1} = \begin{pmatrix} 0 & 0 & -0.017(0.005) \\ 0 & 0.25(0.04) & -1.31(0.35) \\ 0.88(0.33) & -0.41(0.10) & 0 \end{pmatrix};$$

$$\hat{\Sigma}_{II} = \begin{pmatrix} 0.50E-07 & -0.25E-07 & 0.12E-07 \\ -0.25E-07 & 0.12E-04 & 0.20E-05 \\ 0.12E-07 & 0.20E-05 & 0.30E-05 \end{pmatrix}.$$

4. VOLATILITY FORECASTING

The practical relevance of sophisticated volatility modeling to a large extent hinges on its forecasting performance. In practice, volatility forecasting is very important due to its close relation to asset pricing, derivatives' pricing and trading, and risk management. For example, in derivative trading, volatility swap, volatility corridor and variance swap are traded in the over-the-counter market every day. Better forecasts of an asset's return volatility can help practitioners gauge a market trend and make a more intelligent trading decision. In risk management, reliable and long-horizon volatility forecasts make risk assessment and man-

agement feasible, both from regulators and financial institutions' viewpoints.

Based on an estimated VARFIMA(p,d,q) model for the volatility series, we can obtain volatility forecast through standard methods. Specifically, based on model (9), we have

$$(10) \quad \mathbf{M}(B)\mathbf{y}_{t+\ell} = \Phi_1\mathbf{M}(B)\mathbf{y}_{t-1+\ell} + \dots + \Phi_p\mathbf{M}(B)\mathbf{y}_{t-p+\ell} + (\mathbf{I} - \Theta_1B - \dots - \Theta_qB^q)\boldsymbol{\varepsilon}_{t+\ell}$$

Hence,

$$(11) \quad \mathbf{y}_{t+\ell} = (\mathbf{I} - \mathbf{M}(B))\mathbf{y}_{t+\ell} + \Phi_1\mathbf{M}(B)\mathbf{y}_{t-1+\ell} + \dots + \Phi_p\mathbf{M}(B)\mathbf{y}_{t-p+\ell} + (\mathbf{I} - \Theta_1B - \dots - \Theta_qB^q)\boldsymbol{\varepsilon}_{t+\ell}$$

where \mathbf{I} is a $k \times k$ identity matrix. Since the expression $(\mathbf{I} - \mathbf{M}(B))\mathbf{y}_{t+\ell}$ does not involve $\mathbf{y}_{t+\ell}$, the above equation can be used for prediction. By taking expectations on both sides, the predictor $\hat{\mathbf{y}}_{t+\ell}$ can be easily obtained (Box et al., 2008) through the following rules: $E(\mathbf{y}_{t+j}) = \mathbf{y}_{t+j}$ for $j \leq 0$, $E(\mathbf{y}_{t+j}) = \hat{\mathbf{y}}_t(j)$ for $0 < j \leq \ell$, where $\hat{\mathbf{y}}_t(j)$ represents the j -step ahead forecasted $\hat{\mathbf{y}}_t$ at time t ; $E(\boldsymbol{\varepsilon}_{t+j}) = \mathbf{0}$ for $j > 0$, and $E(\boldsymbol{\varepsilon}_{t+j}) = \hat{\boldsymbol{\varepsilon}}_{t+j}$ for $j \leq 0$ (the estimated residuals), and $(1 - B)^d = 1 + \sum_{k=1}^{\infty} \frac{\Gamma(-d+k)}{\Gamma(-d)\Gamma(k+1)}B^k$ is approximated by the finite summation using only available data.

We shall compare the forecasting power of different volatility models, including VARFIMA I, VARFIMA II (the number of dimension is $k = 3$), as well as the GARCH(1,1) and basic SV model together with their extensions as specified in Section 2. Note that for Class VARFIMA I, we evaluate the forecasting performance of the GARCH volatility, and for Class VARFIMA II, we evaluate the forecasting performance of the SV volatility.

It is well known that the microstructure noise is prominent in the intra-daily high frequency data, hence presents a

concern in estimating the daily realized volatility. Since the TSRV is constructed to remove the effect of the microstructure noise on volatility estimation, we shall be particularly interested in the impacts of including RV *versus* TSRV on our VARFIMA(p,d,q) forecasts.

We focus on one-day and five-day forecasts of return volatility, which are widely followed in practice, to investigate the forecasting power of our approach. Specifically, we perform out-of-sample rolling forecasting with a total of 251 one-day and five-day ahead daily forecasts, where the out-of-sample period is from 12/31/2002 to 12/31/2003. In the forecasting procedure, we fix the fractional differencing parameters d_i where $i = 1, 2, 3$ by using the estimated d_i 's in our VARFIMA Classes I and II from the procedures introduced in Section 3, and then exact maximum likelihood estimates on all AR and MA parameters in the model are obtained in each of the rolling windows, and the one-day and five-day ahead forecasts for the return volatility are made based on the fitted model accordingly. Following the results in Section 3, we employ a VARFIMA(2, d, 1) model for both Class VARFIMA I and VARFIMA II here. Therefore, a more explicit version of equation (11) may be written as:

$$(12) \quad \mathbf{y}_{t+\ell} = (\mathbf{I} - \mathbf{M}(B))\mathbf{y}_{t+\ell} + \Phi_1\mathbf{M}(B)\mathbf{y}_{t-1+\ell} + \Phi_2\mathbf{M}(B)\mathbf{y}_{t-2+\ell} + (\mathbf{I} - \Theta_1 B)\boldsymbol{\varepsilon}_{t+\ell}$$

where Φ_1 and Φ_2 are the AR(2) coefficients matrices, and Θ_1 is the MA(1) coefficients matrix. $\mathbf{M}(B)$ follows the same definition as that in model (9) with $k = 3$.

In addition, we also obtained out-of-sample one-day ahead and five-day ahead volatility forecasts using the basic GARCH(1,1) and SV model, as well as the extended GARCH and SV models expressed in (2) and (4) with lag-1 realized volatility and implied volatility included in the variance and log-variance equations, respectively. Furthermore, we derived the volatility forecasts from the FIGARCH(1,1) model. Whenever realized volatility is involved, both RV and TSRV are considered.

Following Andersen and Bollerslev (1998), we treat the daily realized volatility obtained from the intra-daily high frequency data as the true return volatility. Specifically we use the daily TSRV, instead of the standard RV, as the benchmark of the daily return volatility, as TSRV is a more precise estimator (Zhang et al., 2005).

To evaluate the forecasting performance of different models, we used three criteria: regression R^2 , heteroscedasticity-adjusted root mean squared error (HRMSE), and heteroscedasticity-adjusted mean absolute error (HMAE). Specifically, Goodness-of-fit measured by R^2 is obtained by regressing the volatility benchmark $\tilde{\sigma}_t^2$ – in our case, TSRV – against the volatility forecast $\hat{\sigma}_t^2$ within the same time horizon.

The R^2 is obtained using the OLS approach:

$$(13) \quad \tilde{\sigma}_t^2 = \alpha + \beta\hat{\sigma}_t^2 + \epsilon_t.$$

A higher R^2 suggests a higher proportion of the variation in the benchmark can be explained by the volatility forecast.

HRMSE is computed as:

$$(14) \quad HRMSE = \sqrt{\frac{1}{M-\ell} \sum_{j=1}^M \left(1 - \frac{\hat{\sigma}_j^2}{\tilde{\sigma}_j^2}\right)^2}$$

where M is the total number of the forecasts, $\ell = 1$ or 5 depending on whether the forecast is one day or five days ahead. Again, $\tilde{\sigma}_t$ and $\hat{\sigma}_t$ are the benchmark TSRV and volatility forecast on day t , respectively.

Different from the R^2 measure, HRMSE measures the local fluctuations of the forecasted return volatility from the benchmark. HMAE is similar to HRMSE except that it uses mean absolute error. It is defined as:

$$(15) \quad HMAE = \frac{1}{M-\ell} \sum_{j=1}^M \left|1 - \frac{\hat{\sigma}_j^2}{\tilde{\sigma}_j^2}\right|$$

Tables 5 and 6 compare the one-day-ahead and five-day-ahead forecasting performance, using all five stocks. Column A indicates whether standard RV or TSRV is used as a representative for realized volatility in different forecasting models (Columns C-I). Column B is about different forecasting evaluation criteria.

For both one-day forecasts (Table 5) and five-day forecasts (Table 6), the VARFIMA models outperform the other models, yielding the highest R^2 's and lowest HRMSE's and HMAE's for most of the stocks under consideration. In particular, one-day-ahead VARFIMA volatility forecasts can explain 55.1%–66.8% of the variation (R^2) in the benchmark $\tilde{\sigma}_t^2$, while five-day-ahead forecasts explain 35.1%–40.4% of such variation, across stocks. For each evaluation criteria in Column B, the incremental improvement in the forecast performance has a clear pattern across models. First, extended GARCH forecast (Column G) outperforms basic GARCH forecast (Column E), indicating that realized volatility and implied volatility bring in additional information in the GARCH forecast. Similarly, extended SV (Column H) has better forecasting power than basic SV (Column F). Second, the capability of capturing long-memory characteristic in volatility measure helps the forecasting. This is evident from the superior forecast of FIGARCH (Column I) over extended GARCH model (Column G). Also, the VARFIMA II (Column D) forecasts perform much better than the extended SV forecast. Third, overall enhancement in the forecast performance from FIGARCH to VARFIMA I (Column C) seems to indicate that joint modeling of the dynamic relations between different volatility measures has a gain. And finally, for any given volatility model, using the TSRV instead of the standard RV as the realized volatility estimator consistently improves the forecasting. This could be caused by two reasons: one is that TSRV is constructed from a much richer return series (i.e. tick-by-tick data) whereas the

Table 5. Evaluation of one-day ahead return volatility forecasting performance

A	B	C	D	E	F	G	H	I
		<i>VARFIMA I</i>	<i>VARFIMA II</i>	<i>GARCH</i>	<i>SV</i>	<i>GARCH + RV + IV</i>	<i>SV + RV + IV</i>	<i>FIGARCH</i>
Citi: <i>TSRV</i>	R^2	0.642	0.652	0.449	0.471	0.535	0.540	0.573
	<i>HRMSE</i>	0.315	0.310	0.393	0.385	0.341	0.335	0.321
	<i>HMAE</i>	0.293	0.289	0.376	0.371	0.327	0.323	0.307
Citi: <i>RV</i>	R^2	0.561	0.579	0.388	0.392	0.441	0.445	0.509
	<i>HRMSE</i>	0.357	0.348	0.425	0.409	0.376	0.382	0.361
	<i>HMAE</i>	0.322	0.318	0.390	0.384	0.350	0.345	0.333
Disney: <i>TSRV</i>	R^2	0.625	0.618	0.437	0.441	0.530	0.537	0.562
	<i>HRMSE</i>	0.332	0.337	0.409	0.402	0.347	0.338	0.334
	<i>HMAE</i>	0.309	0.319	0.383	0.391	0.332	0.326	0.319
Disney: <i>RV</i>	R^2	0.557	0.551	0.359	0.355	0.438	0.440	0.500
	<i>HRMSE</i>	0.348	0.353	0.436	0.443	0.385	0.380	0.366
	<i>HMAE</i>	0.335	0.338	0.398	0.403	0.357	0.353	0.331
GE: <i>TSRV</i>	R^2	0.626	0.633	0.442	0.450	0.537	0.543	0.573
	<i>HRMSE</i>	0.331	0.323	0.394	0.389	0.340	0.334	0.328
	<i>HMAE</i>	0.304	0.297	0.378	0.370	0.328	0.319	0.311
GE: <i>RV</i>	R^2	0.565	0.578	0.370	0.383	0.443	0.450	0.507
	<i>HRMSE</i>	0.353	0.348	0.430	0.425	0.379	0.371	0.355
	<i>HMAE</i>	0.333	0.319	0.387	0.381	0.350	0.345	0.325
Microsoft: <i>TSRV</i>	R^2	0.623	0.631	0.451	0.459	0.543	0.549	0.585
	<i>HRMSE</i>	0.339	0.328	0.388	0.379	0.332	0.328	0.319
	<i>HMAE</i>	0.309	0.302	0.367	0.359	0.317	0.311	0.307
Microsoft: <i>RV</i>	R^2	0.570	0.581	0.382	0.389	0.448	0.453	0.519
	<i>HRMSE</i>	0.348	0.339	0.419	0.413	0.368	0.360	0.348
	<i>HMAE</i>	0.327	0.311	0.380	0.373	0.340	0.338	0.316
Pfizer: <i>TSRV</i>	R^2	0.660	0.668	0.463	0.471	0.557	0.563	0.593
	<i>HRMSE</i>	0.325	0.309	0.379	0.370	0.325	0.318	0.308
	<i>HMAE</i>	0.298	0.297	0.350	0.343	0.304	0.299	0.290
Pfizer: <i>RV</i>	R^2	0.577	0.585	0.398	0.402	0.469	0.481	0.535
	<i>HRMSE</i>	0.337	0.331	0.405	0.400	0.357	0.350	0.331
	<i>HMAE</i>	0.317	0.308	0.368	0.360	0.330	0.323	0.301

Table 6. Evaluation of five-day ahead return volatility forecasting performance

A	B	C	D	E	F	G	H	I
		<i>VARFIMA I</i>	<i>VARFIMA II</i>	<i>GARCH</i>	<i>SV</i>	<i>GARCH + RV + IV</i>	<i>SV + RV + IV</i>	<i>FIGARCH</i>
Citi: <i>TSRV</i>	R^2	0.395	0.392	0.201	0.189	0.225	0.231	0.297
	<i>HRMSE</i>	0.497	0.500	0.639	0.660	0.579	0.568	0.525
	<i>HMAE</i>	0.462	0.469	0.582	0.594	0.560	0.552	0.507
Citi: <i>RV</i>	R^2	0.371	0.374	0.177	0.170	0.202	0.206	0.271
	<i>HRMSE</i>	0.543	0.536	0.650	0.673	0.585	0.579	0.548
	<i>HMAE</i>	0.504	0.496	0.610	0.618	0.549	0.544	0.513
Disney: <i>TSRV</i>	R^2	0.380	0.400	0.194	0.199	0.233	0.239	0.313
	<i>HRMSE</i>	0.509	0.508	0.643	0.638	0.570	0.563	0.517
	<i>HMAE</i>	0.468	0.471	0.595	0.590	0.552	0.543	0.498
Disney: <i>RV</i>	R^2	0.352	0.351	0.181	0.186	0.207	0.214	0.296
	<i>HRMSE</i>	0.535	0.529	0.649	0.645	0.591	0.580	0.523
	<i>HMAE</i>	0.510	0.513	0.603	0.597	0.573	0.565	0.501
GE: <i>TSRV</i>	R^2	0.399	0.404	0.206	0.213	0.241	0.249	0.324
	<i>HRMSE</i>	0.497	0.499	0.633	0.625	0.563	0.558	0.503
	<i>HMAE</i>	0.459	0.461	0.580	0.572	0.544	0.538	0.489
GE: <i>RV</i>	R^2	0.360	0.369	0.186	0.190	0.212	0.215	0.305
	<i>HRMSE</i>	0.529	0.518	0.640	0.640	0.579	0.573	0.511
	<i>HMAE</i>	0.491	0.492	0.594	0.588	0.560	0.553	0.493
Microsoft: <i>TSRV</i>	R^2	0.390	0.385	0.196	0.193	0.233	0.228	0.319
	<i>HRMSE</i>	0.520	0.523	0.637	0.643	0.581	0.589	0.519
	<i>HMAE</i>	0.470	0.471	0.606	0.613	0.552	0.548	0.497
Microsoft: <i>RV</i>	R^2	0.365	0.378	0.182	0.177	0.207	0.201	0.301
	<i>HRMSE</i>	0.545	0.535	0.653	0.664	0.589	0.594	0.523
	<i>HMAE</i>	0.511	0.503	0.608	0.619	0.573	0.577	0.505
Pfizer: <i>TSRV</i>	R^2	0.397	0.402	0.207	0.213	0.238	0.245	0.334
	<i>HRMSE</i>	0.503	0.498	0.624	0.616	0.575	0.569	0.504
	<i>HMAE</i>	0.468	0.463	0.592	0.589	0.550	0.541	0.490
Pfizer: <i>RV</i>	R^2	0.370	0.377	0.193	0.200	0.215	0.219	0.315
	<i>HRMSE</i>	0.538	0.531	0.638	0.633	0.581	0.573	0.509
	<i>HMAE</i>	0.512	0.512	0.613	0.602	0.568	0.560	0.498

standard RV is calculated from a sparse return series (i.e. five-minute returns), the other reason is that TSRV is a more precise volatility measure in the sense that it corrects the bias from the microstructure noise whereas standard RV is vulnerable to the microstructure noise in the high frequency data.

5. CONCLUSIONS

In this study, we proposed a vector ARFIMA model to capture the long memory and cross-correlation of different volatility measures of a financial return series. The volatility measures involved in our VARFIMA model are the volatility generated from the GARCH(1,1) model, the volatility generated from the basic SV model, two realized volatility estimators (RV and TSRV) constructed from intra-daily high frequency data set, and the implied volatility.

In an out-of-sample forecasting comparison, the proposed vector model outperforms some existing volatility models in the literature, including GARCH(1,1) and its extension, basic SV and its extension, as well as FIGARCH(1,1). Our VARFIMA model has three attractive features: (a) it successfully captures the long memory properties in the volatility process; (b) it incorporates richer information from options market (through implied volatility) and from intra-daily tick-by-tick data, meanwhile it is shielded from the microstructure noise in the intra-daily data (through TSRV); and (c) it jointly models the dynamic inter-day relations between different volatility measures. Our data analysis suggests that all above features contribute to a better volatility forecast.

Received 24 September 2009

REFERENCES

- AKAIKE, H. (1969). Fitting autoregression models for prediction. *Annals of the Institute of Statistical Mathematics*, 21:243–247. [MR0246476](#)
- 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., DIEBOLD, F. X., and LABYS, P. (2001). The distribution of exchange rate realized volatility. *Journal of the American Statistical Association*, 96:42–55. [MR1952727](#)
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X., and LABYS, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625. [MR1958138](#)
- BAILLIE, R. T., BOLLERSLEV, T., and MIKKELSEN, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74:3–30. [MR1409033](#)
- BARNDORFF-NIELSEN, O. E., HANSEN, P. R., LUNDE, A., and SHEPHARD, N. (2008). Designing realized kernels to measure ex-post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536. [MR2468558](#)
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility methods. *Journal of Royal Statistical Society, B*, 64:253–280. [MR1904704](#)
- BERAN, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and longmemory autoregressive integrated moving average models. *Journal of Royal Statistical Society, Series B*, 57:659–672. [MR1354073](#)
- BLACK, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179.
- BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.
- BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327. [MR0853051](#)
- BOLLERSLEV, T., ENGLE, R., and NELSON, D. B. (1994). Arch models. In Engle, R. F. and McFadden, D., editors, *Handbook of Econometrics*, volume IV. North Holland Press, Amsterdam. [MR1315984](#)
- BOX, G., JENKINS, G., and REINSEL, G. (2008). *Time series analysis: Forecasting and control*. Wiley, New York, fourth edition. [MR2419724](#)
- CHUNG, C. F. (1996). Estimating a generalized long memory process. *Journal of Econometrics*, 73:237–259. [MR1410006](#)
- COX, J., ROSS, S., and RUBINSTEIN, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263.
- DACOROGNA, M. M., GENÇAY, R., MÜLLER, U., OLSEN, R. B., and PICTET, O. V. (2001). *An Introduction to High-Frequency Finance*. Academic Press, San Diego.
- DAHLHAUS, R. (1988). Small sample effects in time series analysis: A new asymptotic theory and a new estimator. *Annals of Statistics*, 16:808–841. [MR0947580](#)
- DAHLHAUS, R. (1989). Efficient parameter estimation for self-similar processes. *Annals of Statistics*, 17:1749–1766. [MR1026311](#)
- DAHLHAUS, R. (2006). Correction note: Efficient parameter estimation for self-similar processes. *Annals of Statistics*, 34:1045–1047. [MR2283403](#)
- DING, Z., GRANGER, C. W. J., and ENGLE, R. F. (1993). A long memory property of stock market returns and a new model markets. *Journal of Empirical Finance*, 1:83–106.
- DOORNIK, J. A. and OOMS, M. (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Computational Statistics and Data Analysis*, 42:333–348. [MR2005400](#)
- ENGLE, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of the u.k. inflation. *Econometrica*, 50:987–1008. [MR0666121](#)
- FAN, J. and WANG, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association*, 102:1349–1362. [MR2372538](#)
- FOX, R. and TAQQU, M. S. (1986). Large sample properties of parameter estimates for strongly dependent stationary gaussian time series. *Annals of Statistics*, 14:517–532. [MR0840512](#)
- FOX, R. and TAQQU, M. S. (1987). Central limit theorems for quadratic forms in random variables having long-range dependence. *Probability Theory and Related Fields*, 74:213–240. [MR0871252](#)
- GEWEKE, J. and PORTER-HUDAK, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4:221–238. [MR0738585](#)
- GHYSELS, E., HARVEY, A., and RENAULT, E. (1995). Stochastic volatility. In Maddala, G. S. and Rao, C. R., editors, *Handbook of Statistics 14, Statistical Methods in Finance*. North Holland Press, Amsterdam. [MR1602124](#)
- HIDALGO, J. and ROBINSON, P. M. (1996). Testing for structural change in a long-memory environment. *Journal of Econometrics*, 70:159–174. [MR1378772](#)
- HULL, J. (2008). *Options, Futures, and Other Derivatives*. Prentice Hall, India, seventh edition.
- HULL, J. and WHITE, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300.
- JACOD, J. and PROTTER, P. (1998). Asymptotic error distributions for the euler method for stochastic differential equations. *Annals of Probability*, 26:267–307. [MR1617049](#)
- KOOPMAN, S. J., JUNGBACKER, B., and HOL, E. (2005). Forecasting daily variability of the S & P 100 stock index using historical, re-

- alized and implied volatility measurements. *Journal of Empirical Finance*, 12.
- LAURENT, S. and PETERS, J. P. (2006). *GARCH 4.2, Estimating and Forecasting ARCH Models*. Timberlake Consultants Press, London.
- LI, W. K. and MCLEOD, A. I. (1986). Fractional time series modeling. *Biometrika*, 73:217–221. [MR0836451](#)
- LO, A. (1991). Long-term memory in stock market prices. *Econometrica*, 59:1279–1313.
- MERTON, R. C. (1973). The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183. [MR0496534](#)
- ROBINSON, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, 47:67–84. [MR1087207](#)
- ROBINSON, P. M. (2001). The memory of stochastic volatility models. *Journal of Econometrics*, 101:195–218. [MR1820250](#)
- SCOTT, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative Analysis*, 22:419–438.
- SHEPHARD, N. (1996). Statistical aspects of arch and stochastic volatility. In Cox, D. R., Hinkley, D. V., and Barndorff-Nielsen, O. E., editors, *Likelihood, Time Series with Econometric and Other Applications*. Chapman and Hall, London. [MR1398214](#)
- SOWELL, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53:165–188. [MR1185210](#)
- TAYLOR, S. J. (1986). *Modelling Financial Time Series*. John Wiley, Chichester.
- TIAO, G. C. and TSAY, R. S. (1989). Model specification in multivariate time series. *Journal of the Royal Statistical Society, B*, 51:157–213. [MR1007452](#)
- TSAY, R. S. (1989). Parsimonious parameterization of vector autoregressive moving average models. *Journal of Business & Economic Statistics*, 7:327–341.
- WIGGINS, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19:351–372.
- YAJIMA, Y. (1985). On estimation of long-memory time series models. *Australian Journal of Statistics*, 27:303–320. [MR0836185](#)
- ZHANG, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12:1019–1043. [MR2274854](#)
- ZHANG, L., MYKLAND, P. A., and AÏT-SAHALIA, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100:1394–1411. [MR2236450](#)

Zhixin Kang
 Department of Economics, Finance, and Decision Sciences
 School of Business
 University of North Carolina at Pembroke
 One University Drive
 Pembroke, NC 28372
 E-mail address: zhixin.kang@uncp.edu

Lan Zhang
 Department of Finance
 College of Business Administration
 University of Illinois at Chicago
 601 South Morgan Street, MC 168
 Chicago, IL 60607

Oxford-Man Institute
 University of Oxford, UK
 E-mail address: lanzhang@uic.edu

Rong Chen
 Department of Statistics
 Rutgers University
 Piscataway, NJ 08854

Department of Business Statistics and Econometrics
 Peking University, Beijing, China
 E-mail address: rongchen@stat.rutgers.edu