Financial volatility forecasting with range-based autoregressive volatility model

Hongquan Li\textsuperscript{a,b,}\textsuperscript{*}, Yongmiao Hong\textsuperscript{b,c}

\textsuperscript{a}School of Business, Hunan Normal University, Changsha 410081, PR China
\textsuperscript{b}Department of Economics, Cornell University, Ithaca, NY 14853, USA
\textsuperscript{c}Wang Yanan Institute for Studies in Economics and MOE Key Laboratory of Econometrics, Xiamen University, Xiamen, Fujian 361005, PR China

\textbf{A R T I C L E I N F O}

\textbf{Article history:}
Received 9 May 2010
Accepted 1 December 2010
Available online xxxx

\textbf{JEL classification:}
G32
C01
C53

\textbf{Keywords:}
Volatility modeling
Price range
Forecasting performance
Intraday information
GARCH

\textbf{A B S T R A C T}

The classical volatility models, such as GARCH, are return-based models, which are constructed with the data of closing prices. It might neglect the important intraday information of the price movement, and will lead to loss of information and efficiency. This study introduces and extends the range-based autoregressive volatility model to make up for these weaknesses. The empirical results consistently show that the new model successfully captures the dynamics of the volatility and gains good performance relative to GARCH model.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Volatility plays a very important role in finance, whether in asset pricing, portfolio selection, or risk management. The interest in modeling and forecasting of volatility has steadily increased during the last decade (for details refer to the survey by Poon and Granger (2003)). Volatility was traditionally assumed constant volatility and estimated as the sample standard deviation of returns for a period (based on the closing prices) called historical volatility. However, it is now well known that volatility is time-varying. This fact has been uncovered in three ways: by estimating parametric time series

\textsuperscript{*} Corresponding author at: School of Business, Hunan Normal University, Changsha 410081, PR China.
E-mail addresses: hl662@cornell.edu, Lhquan2000@126.com (H. Li).

1544-6123/$ - see front matter © 2010 Elsevier Inc. All rights reserved.

models like GARCH and Stochastic Volatility, from option price implied volatilities, and from direct measures, such as the realized volatility. Among them, the GARCH model is most-adopted for modeling the time-varying conditional volatility. GARCH models the time-varying variance as a function of lagged squared residuals and lagged conditional variance. The strength of the GARCH model lies in its flexible adaptation of the dynamics of volatilities and its ease of estimation when compared to the other models.

Essentially, the GARCH model is return-based model, which is constructed with the data of closing prices. Hence, though the GARCH model is a useful tool to model changing variance in time series, and provides acceptable forecasting performance, it might neglect the important intraday information of the price movement. For example, when today’s closing price equals to last day’s closing price, the price return will be zero, but the price variation during the today might be turbulent. However, the return-based GARCH model cannot catch it. Using the intraday GARCH, some studies try to remedy the limit of the traditional GARCH. An alternative way to model the intraday price variation is adopting the price range instead.

The range, defined as the difference between the highest and lowest log prices over a fixed sampling intervals (e.g. 1-day or 1-week), has a long, colorful and distinguished history of use as a volatility estimator. Compared to the historical volatility, range-based volatility estimators are claimed to be 5–14 times more efficient (e.g. Garman and Klass, 1980; Parkinson, 1980; Rogers and Satchell, 1991; Yang and Zhang, 2000). They are easy to implement as they only require the readily available high, low, opening and closing prices. In fact, the range has been reported for many years in major business newspapers through so-called “candlestick plots”. Despite these advantages, the range-based volatility estimators have not attracted enough attention in the estimation and forecasting of volatility. This could be due to their poor performance in empirical studies. Chou (2005) conjectures that the fundamental reason is that they cannot well capture the dynamics of volatilities. By properly modeling the dynamic process, price range volatility would retain its superiority in forecasting volatility.

This paper aims to fill in this gap by introducing range-based autoregressive volatility (AV) model and investigating the ability and superiority of AV estimators to forecast the future volatility through comparing with GARCH volatility. Previous works (Beckers, 1983; Wiggins, 1992) examined the forecasting ability of price range estimators using only historical volatility as the benchmark. The GARCH volatility measure adopted here is a significantly improved benchmark.

The rest of the paper is organized as follows: Section 2 introduces the price range estimators and gives a brief description of volatility models, focusing on the range-based AV models definition and estimation. Section 3 presents the result of volatility model estimation on S&P500 index. Section 4 focuses on out-of-sample volatility forecast comparison; the approach adopted for evaluating the performance of different volatility forecasting methods is also detailed. The final section provides conclusions.

2. Volatility models

Modeling the behavior of speculative asset returns has been a central theme in the scopes of financial economics and econometrics. The easiest assumption to model daily returns is a zero-mean normal random variable.

$$r_t = \sigma_t \varepsilon_t \quad \varepsilon_t \sim i.i.d.(0, 1)$$  \hspace{1cm} (1)

In Eq. (1) $\varepsilon_t$ is a zero-mean white noise often assumed to be normal and $\sigma_t$ is the time-varying volatility. Assuming that $\varepsilon_t$ is a normal white noise, the returns conditional on $\sigma_t$ are normal. While the normality is often assumed for the conditional distribution, by modeling $\sigma_t$ as being time-varying the unconditional distribution is leptokurtic. Different specifications for $\sigma_t$ define different volatility models.

2.1. ARCH-type model

ARCH-type models have been widely used to describe conditional heteroskedasticity and are deemed to closely resemble the typical behavior of speculative markets, among which one of the most popular is the GARCH(1,1), originally proposed by Bollerslev (1986).
Among several modifications to the standard GARCH models, Nelson (1991) developed a very successful asymmetric GARCH model, the Exponential GARCH (EGARCH), which accounts for asymmetric impact of returns on conditional variance. For an EGARCH(1,1), Eq. (2b) is modified to

$$\ln{\sigma_t^2} = \gamma + \beta \ln{\sigma_{t-1}^2} + \alpha \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \omega \frac{\epsilon_{t-1}}{\sigma_{t-1}}$$  \hspace{1cm} (3)

The parameter $\omega$ quantifies the asymmetry. The logarithmic formulation ensures a positive conditional variance.

2.2. Range-based autoregressive volatility model

The range reveals more information than the traditional volatility which only using closed prices, because the extremes are formed from the entire price process. Range estimators are also proved to be highly efficient in contrast to classical volatility proxy based on the daily return. Beckers (1983) empirically showed that volatility estimators can be significantly improved by incorporating high and low prices, along with closing prices. The more recent studies (e.g. Alizadeh et al., 2002; Bali and Weinbaum, 2005; Shu and Zhang, 2006) found a strong support for range estimators using realized volatility as the benchmark. Particularly, Alizadeh et al. (2002) and Shu and Zhang (2006) found that the range estimators are not significantly biased and are robust to microstructure errors like bid-ask spread. Despite the fact that the range is a less efficient volatility proxy than realized volatility under ideal conditions (e.g. Andersen and Bollerslev, 1998; Andersen et al., 2001), it may Nevertheless prove superior in real-world situations in which market microstructure biases contaminate high-frequency prices and returns (Alizadeh et al., 2002). The relative efficiency and simplicity of range estimators make a strong case for evaluating their performance further. The classical range estimator is introduced by Parkinson (1980). His volatility estimator is given below:

$$\hat{\sigma}_{RNG}^2 = \left( \frac{1}{4 \ln 2} \right) (\ln H_t - \ln L_t)^2$$  \hspace{1cm} (4)

where $H_t$ and $L_t$ are the daily (or weekly) high and low prices, respectively. This range volatility estimator is based on the assumption that the asset price follows a driftless geometric Brownian motion and is theoretically shown by Parkinson to be 5.2 times more efficient than the classical estimator based on closing prices. Garman and Klass (1980), Beckers (1983), Wiggins (1992), Rogers and Satchell (1991), Kunitomo (1992), and Yang and Zhang (2000) further extend the range estimator to incorporate information about the opening and closing prices and the treatment of a time-varying drift, as well as other considerations.

Despite the elegant theory and the support of simulation results, the range-based estimator has performed poorly in empirical studies. The reason for this is its failure to capture the dynamic evolution of volatilities. In order to solve the problem, the range-based autoregressive volatility model is used to uncover the volatility process in the paper.

The autoregressive volatility (AV) model introduced by Hsieh (1991, 1993, 1995) is much better able to capture the dynamics in volatility. The AV model is given below:

$$r_t = \sigma_{RNG,t} \epsilon_t \quad \epsilon_t \sim i.i.d. N(0, 1)$$  \hspace{1cm} (5a)

$$\ln{\sigma_{RNG,t}^2} = \alpha + \sum \beta_i \ln{\sigma_{RNG,t-i}^2} + \nu_t \quad \nu_t \sim i.i.d.(0, \sigma_v^2)$$  \hspace{1cm} (5b)

where $\sigma_{RNG,t}$ is the range-based volatility estimator. The AV model is motivated by the fact that the volatility is highly autocorrelated. The ex ante volatility can be recovered, as follows. Regress $\ln{\hat{\sigma}_{RNG,t}^2}$ on its own lags and a constant term using ordinary least squares (OLS). For simplicity, $\nu_t$ and $\epsilon_t$ are assumed to be independent.

The equation of the conditional variance can in general be easily extended to incorporate other explanatory variables. For example, we can easily modify Eq. (5b) to Eq. (6) to incorporate asymmetric impact of returns on conditional variance.

\[
\ln \sigma_{\text{RNG},t}^2 = \alpha_0 + \beta_1 \ln \sigma_{\text{RNG},t-1}^2 + \alpha_1 \left| \frac{r_{t-1}}{\sigma_{\text{RNG},t-1}} \right| + \alpha_2 \frac{r_{t-1}}{\sigma_{\text{RNG},t-1}} + \nu_t \quad \nu_t \sim \text{i.i.d.}(0, \sigma^2_{\nu})
\]  

It is an asymmetric autoregression volatility model (henceforth, AV-\( \alpha \) model). The parameter \( \alpha_2 \) quantifies the asymmetry.

The AV-type models and the popular GARCH-type models differ in three important respects: (1) The AV model has found much less volatility persistence than the GARCH model, (2) the GARCH model has been estimated using the maximum likelihood method, which requires a specific distributional assumption on the error term \( e_t \). The AV model does not require any distributional assumptions, and (3) the AV model includes a stochastic term in the variance equation, which make it more general and flexible.

3. Volatility Model estimation on S&P500 index

3.1. The data

We employ weekly (5-trading days) high, low, opening and closing prices of the S&P500 index. Our data consists of about fourteen years of daily S&P500 index from May 27, 1994 to April 22, 2008, consisting 700 weekly data points. The total sample is divided into two parts. The first 600 data have been taken as the estimation sample, while the last 100 data from April 26, 2006 to April 22, 2008 have been used as out-of-sample period for volatility forecasting.

Fig. 1 shows weekly returns (5-trading days’ return) series of the S&P500 index. It suggests that the returns are moving around an approximately zero-mean with time-varying clustering volatility. Fig. 2 presents the plot of the weekly range volatility series and Table 1 reports the statistics of volatility under different measures. The daily volatility has a mean of about 2% corresponding to an annualized volatility of 14%. Its standard deviation (about 1%) indicates significant variation in the volatility of S&P500. It is interesting to observe the difference in the values of the ACF’s and of the Ljung–Box Q statistics for the absolute return and the range series. The Q statistics are 1664.70 for the range and 215.26 for the absolute returns indicating a much stronger degree of persistence in volatility for the range than for the absolute return series. This fact partly stimulates us to employ AV model in volatility forecasting.

Four volatility models, a GARCH(1,1), an EGARCH(1,1), and two range-based AV models, will be estimated and compared in the next two sections. GARCH(1,1) and EGARCH(1,1) are the most popular ARCH-type models used in application, which are ideal benchmarks for volatility forecast comparison.
3.2. Model estimation

Table 2 presents the estimates of AV model and AV-\(\alpha\) model using range volatility estimator. The number of lags in AV models is determined by the Schwarz criterion. The persistence of volatility is measured by the sum of the \(b\) coefficients, which are 0.80 for AV model and 0.89 for AV-\(\alpha\) model. They are less than 1 in two cases, indicating that log volatility is strictly stationary. When compared to the ARCH-type models in which persistence coefficient equal to 0.99 (The estimation results of ARCH-type models are available from the author upon requests), the AV model has much less persistence for S&P500. Our finding is thus consistent with the results reported by Hsieh (1995) that the popular ARCH-type models have a tendency to put too much persistence into volatility and the AV model is much better able to capture the dynamics of volatility which includes volatility clustering and mean reversion behavior. The AV model's good performance is validated by the empirical findings (see Tables 3 and 4 and the next section). Both LB Q-statistic and ARCH tests (see Table 3) prove the ability of the AV models in capturing nonlinear dependence: as in the case of GARCH, the squared standardized returns are not autocorrelated and there are no residual ARCH effects.

In order to compare GARCH and AV forecasting performance in-sample, we report the squared correlation, \(R^2\), from the regression.

Table 1
Descriptive statistics of the volatility measured by different estimators.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Max.</th>
<th>Min.</th>
<th>Std. dev.</th>
<th>ACF(1)</th>
<th>ACF(15)</th>
<th>Q(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(</td>
<td>r_t</td>
<td>)</td>
<td>1.67</td>
<td>8.67</td>
<td>0.00</td>
<td>1.40</td>
<td>0.185</td>
</tr>
<tr>
<td>(\sigma_{\text{RNG}})</td>
<td>1.95</td>
<td>7.27</td>
<td>0.42</td>
<td>1.06</td>
<td>0.609</td>
<td>0.247</td>
<td>1664.79</td>
</tr>
</tbody>
</table>

Notes: \(\sigma_{\text{RNG}}\) is range-based estimator using weekly highest and lowest prices; Q(15) statistics represent the Ljung-Box Q statistics for autocorrelation of volatility series.

Table 2
Estimates of AV and AV-\(\alpha\) model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate(^1)</th>
<th>p-value(^1)</th>
<th>Estimate(^2)</th>
<th>p-value(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_0)</td>
<td>(-1.5762)</td>
<td>0.0000</td>
<td>(-1.4708)</td>
<td>0.0000</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>0.3277</td>
<td>0.0000</td>
<td>0.5166</td>
<td>0.0000</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>0.1923</td>
<td>0.0000</td>
<td>0.1844</td>
<td>0.0000</td>
</tr>
<tr>
<td>(\beta_3)</td>
<td>0.1981</td>
<td>0.0000</td>
<td>0.1388</td>
<td>0.0000</td>
</tr>
<tr>
<td>(\beta_4)</td>
<td>0.0878</td>
<td>0.0178</td>
<td>0.0541</td>
<td>0.0609</td>
</tr>
<tr>
<td>(x_1)</td>
<td>0.6405</td>
<td>0.0000</td>
<td>0.6000</td>
<td>0.0000</td>
</tr>
<tr>
<td>(x_2)</td>
<td>(-0.0835)</td>
<td>0.0000</td>
<td>(-0.0835)</td>
<td>0.0000</td>
</tr>
<tr>
<td>(\sum \beta_i)</td>
<td>0.80</td>
<td></td>
<td>0.89</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Estimate\(^1\) and p-value\(^1\) represent values for AV model, and Estimate\(^2\) and p-value\(^2\) for AV-\(\alpha\) model.

3.2. Model estimation

The number of lags in AV models is determined by the Schwarz criterion. The persistence of volatility is measured by the sum of the \(\beta\) coefficients, which are 0.80 for AV model and 0.89 for AV-\(\alpha\) model. They are less than 1 in two cases, indicating that log volatility is strictly stationary. When compared to the ARCH-type models in which persistence coefficient equal to 0.99 (The estimation results of ARCH-type models are available from the author upon requests), the AV model has much less persistence for S&P500. Our finding is thus consistent with the results reported by Hsieh (1995) that the popular ARCH-type models have a tendency to put too much persistence into volatility and the AV model is much better able to capture the dynamics of volatility which includes volatility clustering and mean reversion behavior. The AV model's good performance is validated by the empirical findings (see Tables 3 and 4 and the next section). Both LB Q-statistic and ARCH tests (see Table 3) prove the ability of the AV models in capturing nonlinear dependence: as in the case of GARCH, the squared standardized returns are not autocorrelated and there are no residual ARCH effects.

In order to compare GARCH and AV forecasting performance in-sample, we report the squared correlation, \(R^2\), from the regression.
which provides the proportion of realized volatility \( \sigma_t \) explained by the volatility estimate \( \hat{\sigma}_t \) from volatility forecasting model. Table 4 presents the regression test results. All evidence clearly demonstrates the superiority of the range-based AV models. First, in every case, the \( R^2 \) of the range-based AV forecasts is higher than that of the return-based GARCH forecasts. Second, when compared to the GARCH-type models, the AV models have much less deviation from the unbiasedness condition that \( c_0 = 0 \) and \( c_1 = 1 \), thus providing little bias. In addition, models that incorporate some form of asymmetry offer significant advantages over the corresponding symmetric models (e.g., GARCH vs. EGARCH, AV vs. AV-\( \alpha \)). Among all models, the AV-\( \alpha \) model performs best.

4. Out-of-sample volatility forecast comparison

Ultimately, the usefulness of volatility models depends on their ability to accurately forecast future volatility. Therefore, we perform a variety of out-of-sample forecasting exercises to determine which specification performs best by this criterion.

To evaluate forecast accuracy, four popular measures are used, namely, the root mean square error (RMSE), the mean absolute error (MAE), the Theil-U statistic and the regression’ \( R^2 \) statistic. They are defined by

\[
\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\sigma_t - \hat{\sigma}_t)^2}
\]

\[
\text{MAE} = \frac{1}{T} \sum_{t=1}^{T} |\sigma_t - \hat{\sigma}_t|
\]

\[
\text{Theil-U} = \frac{\sum_{t=1}^{T} (\hat{\sigma}_t - \hat{\sigma}_t)^2}{\sum_{t=1}^{T} (\hat{\sigma}_t - \hat{\sigma}_t)^2}
\]

Table 4 provides a comparison of in-sample volatility forecasts using a regression method. For panel A, realized volatility is measured by \( r_{RNG} \), which is the range-based estimator using weekly highest and lowest prices. For panel B, realized volatility is measured by \( r_{SSDR} \), which is the sum of squared daily returns within each week.

Notes: \textit{T}-statistics computed using Newey-West standard errors are in parentheses. The realized volatility measure \( \sigma_{RNG} \) is range-based estimator using weekly highest and lowest prices (see Eq. (4)). SSDR is the sum of squared daily returns within each week, and the square root of SSDR is denoted by \( \sigma_{SSDR} \).

where $\sigma_i$ is the realized weekly volatility measured by $\sigma_{RNG}$ defined by Eq. (4) or $\sigma_{SSDR}$ calculated from the sum of squared daily returns within each week.

The RMSE and MAE are two of the most popular measures to test the forecasting power of a model. Despite their mathematical simplicity, however, both of them are not invariant to scale transformations. The Theil-U-statistic is a desirable measure to evaluate the accuracy of various forecasting methods (see Armstrong and Fildes, 1995). In the Theil-U statistic, the error of prediction is standardized by the error from the random walk forecast. For the random walk model, which can be treated as the benchmark model, the Theil-U statistic equals 1. Table 5 presents the examination results.

All results in Table 5 unequivocally support the conclusion that the range-based AV models provide more accurate forecasts of realized volatility than the corresponding GARCH models (e.g., AV vs. GARCH, AV-$\alpha$ vs. EGARCH) under every evaluation criteria. The RMSE and MAE statistics indicate that AV models yield smaller error than that of GARCH models. A closer examination of the evaluation reveals that the differences in the performance of the two-type models are more obvious when $\sigma_{RNG}$ is used for the realized volatility. Given the fact that price range use more information (intra-daily) than SSDR (daily information), it is not surprising that range-based volatility estimator contains less noise and will yield more precise pictures in forecast comparisons.

Under the Theil-U statistic, all models perform better than the random walk model. They all have the Theil-U statistic less than 1. The best performer is again the AV-$\alpha$ model with the U statistic of 0.54 and 0.68. The result of the regression-based comparison is in consistent with the previous evidence. AV models dominate GARCH models in producing higher $R^2$ values.

### 5. Conclusion

This paper examines and demonstrates the ability and superiority of price range estimators to forecast the future volatility through comparing with the GARCH volatility. In order to properly model the dynamics of volatility process, the autoregressive volatility model is adopted. Two types of volatility models are discussed and estimated: return-based GARCH model and range-based AV model. The comparison study includes out-of-sample forecasting performance as well as in-sample comparison. The results from both in-sample and out-of-sample forecasts consistently show that the range-based AV model successfully captures the dynamics of the volatility and gains good performance relative to GARCH model. Furthermore, we find that the inclusion of the lagged return can significantly improve the forecasting ability of the AV model. Our empirical results also suggest the existence of a leverage effect in the US stock markets (Baillie and Bollerslev, 1989; Engle, 1982).

The AV model provides a simple, yet effective framework for forecasting the volatility dynamics. It would be interesting to explore whether alternative choices of volatility measures, such as the realized variance (RV, see, e.g., Andersen et al., 2001b) and realized range-based variance (RRV, see Christensen and Podolskij, 2007), fit the class of the AV models. Generally, the empirical results of this article provide strong support for the application of the AV model in the stock markets that will be of great

---

Table 5
Out-of-sample forecast performance of competing models.

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
<th>Theil-U</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Realized volatility measured by $\sigma_{RNG}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>7.27</td>
<td>5.83</td>
<td>0.93</td>
<td>0.38</td>
</tr>
<tr>
<td>EGARCH(1,1)</td>
<td>6.90</td>
<td>5.51</td>
<td>0.83</td>
<td>0.48</td>
</tr>
<tr>
<td>AV</td>
<td>6.77</td>
<td>5.00</td>
<td>0.80</td>
<td>0.45</td>
</tr>
<tr>
<td>AV-$\alpha$</td>
<td>5.57</td>
<td>3.68</td>
<td>0.54</td>
<td>0.65</td>
</tr>
<tr>
<td><strong>Panel B. Realized volatility measured by $\sigma_{SSDR}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>8.85</td>
<td>6.93</td>
<td>0.86</td>
<td>0.35</td>
</tr>
<tr>
<td>EGARCH(1,1)</td>
<td>8.20</td>
<td>6.54</td>
<td>0.74</td>
<td>0.44</td>
</tr>
<tr>
<td>AV</td>
<td>8.68</td>
<td>6.46</td>
<td>0.83</td>
<td>0.42</td>
</tr>
<tr>
<td>AV-$\alpha$</td>
<td>7.90</td>
<td>5.83</td>
<td>0.68</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Notes: RMSE and MAE have been multiplied by 10^3.
interest to academics and practitioners, particularly those involved in making international risk management decisions.

Acknowledgments

We would like to thank the editor and anonymous referee for their helpful comments. H.Q. Li thanks National Natural Science Foundation of China (No. 71001036), Social Science Research Fund under Ministry of Education of China (No. 09YJC790084) and Scientific Research Fund of Hunan Provincial Education Department for financial support. We also would like to thank seminar participants at Advanced Econometrics, Cornell University for their comments.

References


