Research Article

GARCH-Type Model with Continuous and Jump Variation for Stock Volatility and Its Empirical Study in China

Huannan Zhang and Qiujun Lan

Business School of Hunan University, Changsha 410082, China

Correspondence should be addressed to Qiujun Lan; lanqiujun@hnu.edu.cn

Received 29 October 2013; Accepted 19 December 2013; Published 12 January 2014

Academic Editor: Fenghua Wen

Copyright © 2014 H. Zhang and Q. Lan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

On the basis of GARCH-RV-type model, we decomposed the realized volatility into continuous sample path variation and discontinuous jump variation, then proposed a new volatility model which we call the GARCH-type model with continuous and jump variation (GARCH-CJ-type model). By using the 5-minute high frequency data of HUSHEN 300 index in China, we estimated parameters of the GARCH-type model, the GARCH-RV-type model, and the GARCH-CJ-type model and compared the three types of models’ predictive power to the future volatility. The results show that the realized volatility and the continuous sample path variation have certain predictive power for future volatility, but the discontinuous jump variation does not have that kind of function. What is more, the GARCH-CJ-type model has a more power to predict the future volatility than the other two types of models. Therefore, the GARCH-CJ-type model is much more useful for the research on the capital assets pricing, the derivative security valuation, and so on.

1. Introduction

The research on asset volatility in financial market is the foundation of finance, such as capital assets pricing, financial derivatives pricing, and financial risk measurement. The premise of quantitative financial analysis is to accurately measure and predict asset volatility. Therefore, the measurement and prediction of asset volatility are a hotspot of research all the time.

To measure and predict asset volatility accurately, Engle [1], in view of “clustering” and “persistence” of volatility, proposed an autoregressive conditional heteroscedastic (ARCH) model; Bollerslev [2] built a generalized ARCH (GARCH) model based on the ARCH model. Then, GARCH model was extended; Nelson [3] found that the asset volatility is “asymmetric.” He modified the GARCH model and built an EGARCH model; Glosten et al. [4] also examined the “asymmetry” and built a TGARCH model (also called GJR model). The above models (called GARCH-style model in this paper) have been proved to have strong power to predict the future volatility of assets [5].

Admittedly, GARCH-type models have fairly strong predictive power, but there is room for improvement, as the accuracy pursuit for future volatility prediction is endless in financial operations, such as financial asset pricing, financial derivative pricing, and financial risk management. Therefore, it is necessary to improve the predictive power of the models. In order to perfect the accuracy of predictions, the realized volatility (RV) as an exogenous variable has been introduced by Koopman et al. [6] into the volatility equation of GARCH model. They built a GARCH-RV model and found that the GARCH-RV model has stronger predictive power than the GARCH model. Fuertes et al. and Frijns et al. [7, 8] also showed that the GARCH-RV model has stronger power to predict the asset volatility than the GARCH model. But in realistic financial markets, the asset volatility is a continuous process with some jump components. When Andersen et al. and Huang et al. [9, 10] studied the HAR-type RV model, they found that model built with continuous sample path variation and discontinuous jump variation that decomposed from RV has stronger power than the undecomposed HAR-RV model in measuring and predicting the asset volatility. For this reason, in studying the GARCH model with an introduction of an endogenous variable RV, it is more reasonable to decompose RV into C and J and introduce the two parts into the volatility equation of the GARCH model. On the basis
of the GARCH model, this paper decomposes RV into two parts, \( C \) and \( J \), and constructs a GRACH-CJ model in an attempt to further improve the predictive power for the future volatility. Similarly, this paper will also extend the EGARCH model and GJR model to EGARCH-RV model, GJR-RV model, EGARCH-CJ model, and GJR-CJ model. After that, we estimated parameters of the above models and compared their predictive power for the future volatility, respectively, to identify the volatility model with stronger power for the asset volatility measurement and prediction, using the 5-minute high frequency data of HUSHEN 300 index in China.

The remainder of the paper is organized as follows. The GARCH-CJ-type model construction will be introduced in Section 2. The empirical evidence and predictive power of the models will be presented in Section 3. The last part, Section 4, is the conclusion.

2. Model Construction

2.1. GARCH-CJ Model

2.1.1. GARCH-RV Model Construction. Stock return volatility cannot be observed directly but can be measured in the asset return series. The return volatility is "clustering" and "persistent." The ARCH model proposed in Engle [1] can well capture the volatility clustering of the return series, but the model is rather complicated when the regression order gets bigger. On the basis of the ARCH model, Bollerslev [2] proposed the GARCH model to overcome the defect. GARCH(1,1) is expressed as follows:

\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot u_t, \quad h_t = \omega + \alpha \epsilon_{t-1}^2 + \theta h_{t-1},
\]

where \( R_t \) is the return, \( \mu_{t-1} \) denotes the conditional mean of \( R_t \) based on all available information, \( h_t \) is the volatility, \( u_t \) is the white noise disturbance, and \( \omega, \alpha, \) and \( \theta \) are parameters to be estimated.

In order to improve the measurement of volatility and the accuracy of the prediction of the model, Koopman et al. [6] introduced the realized volatility (RV) as an exogenous variable into the volatility equation of GARCH(1,1) model to build a GARCH-RV model

\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot u_t, \quad h_t = \omega + \alpha \epsilon_{t-1}^2 + \theta h_{t-1} + \lambda RV_{t-1},
\]

where \( \lambda \) is also a parameter to be estimated as \( \omega, \alpha, \theta, \) and \( RV_{t-1} \) is the realized volatility at \( t - 1 \) period, which is defined according Martens [11] and Koopman et al. [6]. With overnight return variance, realized volatility can be expressed as

\[
RV_t = \sum_{i=1}^{N} r_{t,i}^2 + \sum_{j=1}^{M} r_{t,n,j}^2, \quad M = N + 1,
\]

where \( N \) is the number of equally divided parts of a trading day; \( r_{t,1} \) denotes the first return after the opening quotation at Day \( t \), \( r_{t,1} = 100(\ln P_{t,1} - \ln P_{t,0}) \), and \( P_{t,0} \) is the first closing price at Day \( t \), \( r_{t,2} \) is the opening price at Day \( t \), \( r_{t,2} = 100(\ln P_{t,2} - \ln P_{t,1}) \) \( \ldots \) so, \( r_{t,N} \) expresses the \( N \)th return at Day \( t \) after opening, \( r_{t,N} = 100(\ln P_{t,N} - \ln P_{t,N-1}) \), \( r_{t,n} \) and \( r_{t,M} \) refer to overnight return variance, \( r_{t,n} = r_{t,M} = 100(\ln P_{t,0} - \ln P_{t-1,i}) \), and \( P_{t-1,i} \) is the closing price in Day \( t - 1 \).

2.1.2. GARCH-CJ Model Construction. The real financial market reveals evidently nonlinear features [12] and the financial asset price volatility is not continuous but shows jump volatility, since the market is subject to the impact of some big information shocks and investors’ irrational factors. Andersen et al. [9] showed that it has more power to predict the future volatility by decomposing the realized volatility into continuous sample path variation and discontinuous jump variation. In order to improve the predictive power of the model, we will introduce the continuous sample path variation \( C_t \) and the discontinuous jump variation \( J_t \) decomposed from the realized volatility into model (2).

To decompose the realized volatility (RV), Barndorff-Nielsen and Shephard [13, 14] proposed Realized Bipower Variation (RV); that is,

\[
RV_t^{[r,s]} = \left( \frac{h}{M} \right)^{1-(r+s)/2} \sum_{j=1}^{M-1} \left| r_{j,s} \right|^{r+1} \left| r_{j+1,s} \right|^{s},
\]

where \( h > 0 \) is a fix time interval, \( r, s \geq 0 \) are constant (usually, \( 1 \) is given), and \( M \) is the sample frequency within interval \( h \). According to Barndorff-Nielsen and Shephard’s research, when \( M \to \infty \), the difference between \( RV_t \) and \( RV_t^{[r,s]} \) is equivalent to a consistent estimator for discontinuous jump variation \( J_t \):

\[
RV_t - RV_t^{[r,s]} \xrightarrow{M \to \infty} J_t.
\]

With a limited sample size, \( J_t \) calculated from (5) may not always be nonnegative. In order for \( J_t \) to be always nonnegative, we will treat \( J_t \) in the following way:

\[
J_t = \max \left[ RV_t - RV_t^{[r,s]}, 0 \right].
\]

In calculating discontinuous jump variation \( J_t \), sampling intraday data at unequal frequency will result in calculation error. In order to improve the calculation accuracy of \( J_t \), it is necessary to introduce some statistic to test the significance of \( J_t \). This paper adopts \( Z_t \) statistic proposed by Barndorff-Nielsen and Shephard [13, 14] based on the bipower variation theory to test \( J_t \). \( Z_t \) is expressed as follows:

\[
Z_t = \frac{(RV_t - RV_t^{[r,s]}) RV_t^{-1}}{\sqrt{((\pi/2)^2 + \pi - 5) (1/M) \max (1, RTQ_t/RBV_t^2)}} \rightarrow N (0, 1),
\]
where
\[
RTQ_t = M \mu_{\beta}^3 \left( \frac{M}{M - 4} \right) \sum_{j=4}^{M} \left[ r_{t,j-4} r_{t,j-2} r_{t,j} \right]^{-\frac{1}{4}},
\]
\[
\left( \mu_{\beta}^3 = E \left( |Z_t|^\frac{1}{4} \right) = 2^{\frac{2}{3}} \Gamma \left( \frac{7}{6} \right) \Gamma \left( \frac{1}{2} \right)^{-1} \right) .
\]

The classic RBV calculation is closely related to the sampling frequency of the intraday data. With the increase in the sampling frequency, the RBV estimate cannot converge to integral volatility because of the influence of factors, such as the market microstructure. So using RBV as the robust estimator for \( J_t \) is biased, and this paper adopts MedRV, proposed by Andersen et al. [15] as a robust estimator instead. MedRV can be expressed as follows:
\[
\text{MedRV}_t = \frac{\pi}{6 - 4 \sqrt[3]{\pi} + \pi} \left( \frac{M}{M - 2} \right)
\times \sum_{i=2}^{M-1} \text{Med}(|r_{t,j-1},|r_{t,j},|r_{t,j+1}|)^2.
\]

Accordingly, RTQ_{t,\alpha}, the statistic for Z_t in (6), is replaced by MedRTQ_t, which is expressed as follows:
\[
\text{MedRTQ}_t = \frac{3\pi M}{9 \pi + 72 - 52 \sqrt[3]{\pi}} \left( \frac{M}{M - 2} \right)
\times \sum_{i=2}^{M-1} \text{Med}(|r_{t,j-1},|r_{t,j},|r_{t,j+1}|)^4.
\]

After replacing RBV_t with MedRV_t and replacing RTQ_t with MedRTQ_t in formula (7), we calculate the statistic Z_t with (7) and get the estimator for discontinuous jump variation at the \( 1 - \alpha \) significance level:
\[
J_t = I \left( Z_t > \phi_\alpha \right) (RV_t - \text{MedRV}_t).
\]

Accordingly, the continuous sample path variation estimator is
\[
C_t = I \left( Z_t \leq \phi_\alpha \right) RV_t + I \left( Z_t > \phi_\alpha \right) \text{MedRV}_t.
\]

In actual calculation, we need to select a suitable confidence level \( \alpha \). Drawing on previous research, we choose 0.99 as the confidence level \( \alpha \) in this paper. In addition, with the test of statistic \( Z_t \) and relevant bipower variation theory, we can get the estimators for the continuous sample path variation \( C_t \) and discontinuous jump variation \( J_t \) of the log return volatility.

According to above RV decomposition method, we decompose RV_{t-1} of the model (2) into \( C_{t-1} \) and \( J_{t-1} \). Here is the GARCH-CJ model
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
h_t = \omega + \alpha \epsilon_{t-1}^2 + \theta h_{t-1} + \lambda C_{t-1} + \gamma J_{t-1}.
\]

### 2.2. EGARCH-CJ Modeling Building

In view of the asymmetric effect of good and bad news on volatility, Nelson et al. [3] constructed an EGARCH model on the basis of the GARCH model. Later, researchers built more EGARCH-type models, among which a commonly used EGARCH-type can be expressed as
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
\ln(h_t) = \omega + \alpha \left( \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \beta \left( \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \theta \ln(h_{t-1}).
\]

Using the method discussed in Section 2.1, we take the log of the last period's realized volatility (RV_{t-1}) and introduce the log value as an exogenous variable into EGARCH(1,1) and thus get EGARCH-RV
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
\ln(h_t) = \omega + \alpha \left( \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \beta \left( \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \theta \ln(RV_{t-1}).
\]

We decompose RV_{t-1} into \( C_{t-1} \) and \( J_{t-1} \), take the log of \( C_{t-1} \) and \( J_{t-1} \), and thus obtain the EGRACH-CJ model
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
\ln(h_t) = \omega + \alpha \left( \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \beta \left( \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \theta \ln(h_{t-1}) + \gamma \ln(J_{t-1}) + \lambda \ln(C_{t-1}).
\]

### 2.3. GJR-CJ Model Construction

On the basis of the GARCH model, Glosten et al. [4] constructed a TGARCH model (also called GJR model) to introduce the leverage effect on volatility into the new model. GJR model (1,1) is
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
h_t = \omega + (\alpha + \beta N_{t-1}) \epsilon_{t-1}^2 + \theta h_{t-1} + \lambda RV_{t-1}.
\]

where \( N_{t-1} \) is the indicator variable of the negative \( \epsilon_{t-1} \)
\[
N_{t-1} = \begin{cases} 1, & \epsilon_{t-1} < 0 \\ 0, & \epsilon_{t-1} \geq 0. \end{cases}
\]

Similarly, using the method in Section 2.1, we introduce RV_{t-1} as an exogenous variable into the CJR(1,1) and construct the CJR-RV model:
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
h_t = \omega + (\alpha + \beta N_{t-1}) \epsilon_{t-1}^2 + \theta h_{t-1} + \lambda RV_{t-1}.
\]

We divide RV_{t-1} into \( C_{t-1} \) and \( J_{t-1} \), and we get the CJR-CJ model
\[
R_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} \cdot v_t,
\]
\[
h_t = \omega + (\alpha + \beta N_{t-1}) \epsilon_{t-1}^2 + \theta h_{t-1} + \lambda C_{t-1} + \gamma J_{t-1}.
\]
3. Empirical Study and Comparative Analysis of Models’ Predictive Power

3.1. Empirical Study

3.1.1. Samples and Their Statistics. For the empirical study, we take samples from the HUSHEN 300 index for Chinese stock market, and the data come from the Wind financial database. The time span of the samples covers from April 20, 2007, to April 20, 2012, including 1199 trading days. In the calculation of the realized volatility, the sampling frequency of intraday data has a great influence on the research results. On the one hand, too low sampling frequency cannot well capture the volatility information; on the other hand, too high sampling frequency will produce noise which will harm the results. Therefore, in accordance with previous research of other scholars, this paper uses the 5-minute high frequency sampling frequency will produce noise which will harm the results. Therefore, in accordance with previous research of other scholars, this paper uses the 5-minute high frequency data of HUSHEN 300 index. And we use the moving average interpolation method to make up for the missed data, which derives 58751 valid data, 49 pieces of transaction data for each day (including 1 overnight transaction data and 48 day transaction data). Variables needed in this paper are $R_t$ return, $RV_t$ realized volatility, continuous sample path variation $C_t$, and discontinuous jump variation $I_t$ and log realized volatility, log continuous sample path variation, and log discontinuous jump variation. We can see from Table 1 that $RV_t$ series do not follow the normal distribution and are leptokurtic. This implies that China’s stock market has a big volatility. In addition, ADF test shows that all the series reject the null hypothesis of unit root at the 99% confidence level; it can be considered that all series are stationary and thus can be further used in model analysis.

3.1.2. Model Parameter Estimation and Analysis. In this paper, maximum likelihood method is adopted to estimate the model in Section 1. Because the setting of the initial value has a great influence on the result in the estimation process, this paper adopts an approximate value from multiple fitting (also satisfying that the likelihood score be the maximum) as the initial parameter value. Tables 2, 3, and 4 list the estimates for GARCH and other eight models under the assumptions of the residuals following Gaussian distribution and $t$ distribution. Comparing the log likelihood and the AIC value for GARCH, EGARCH, and CJR, we can see that the goodness of fit for the asymmetric EGARCH model and the CJR models is better than that for the GARCH model, which indicates that the influence of favorable and of unfavorable news is asymmetric on the market volatility in China’s stock market. In addition, comparing the log likelihood and the AIC value

### Table 1: Descriptive statistics of each variable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Jarque-Bera</th>
<th>ADF-t statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_t$</td>
<td>-0.0152</td>
<td>2.1141</td>
<td>-0.2647</td>
<td>4.9345</td>
<td>200.95***</td>
<td>-32.901***</td>
</tr>
<tr>
<td>$RV_t$</td>
<td>4.3471</td>
<td>6.8083</td>
<td>5.9174</td>
<td>49.075</td>
<td>11305***</td>
<td>-10.710***</td>
</tr>
<tr>
<td>$C_t$</td>
<td>3.3412</td>
<td>4.3445</td>
<td>5.9456</td>
<td>65.248</td>
<td>20064***</td>
<td>-77.154***</td>
</tr>
<tr>
<td>$I_t$</td>
<td>1.0058</td>
<td>4.8285</td>
<td>9.5878</td>
<td>109.80</td>
<td>58817***</td>
<td>-16.145***</td>
</tr>
<tr>
<td>ln($RV_t$)</td>
<td>0.9600</td>
<td>0.9323</td>
<td>0.5180</td>
<td>55.961</td>
<td>59.611***</td>
<td>-4.9183***</td>
</tr>
<tr>
<td>ln($C_t$)</td>
<td>0.7650</td>
<td>0.8992</td>
<td>0.3525</td>
<td>26.138</td>
<td>26.138***</td>
<td>-5.3556***</td>
</tr>
<tr>
<td>ln($I_t + 1$)</td>
<td>0.2825</td>
<td>0.6209</td>
<td>3.0445</td>
<td>14.085</td>
<td>7990.6***</td>
<td>-16.266***</td>
</tr>
</tbody>
</table>

*** denotes significance at 1% significance level.

### Table 2: Estimation results for GARCH and its extended model.

<table>
<thead>
<tr>
<th></th>
<th>Residual following Gaussian distribution</th>
<th>Residual following $t$ distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GARCH</td>
<td>GARCH-RV</td>
</tr>
<tr>
<td>$\mu_{t-1}$</td>
<td>-0.0007</td>
<td>-0.0400</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.0324***</td>
<td>0.2759***</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0392***</td>
<td>-0.0468***</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.9520***</td>
<td>0.6835***</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.3254***</td>
<td>0.3865***</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.0678*</td>
<td>0.0906</td>
</tr>
<tr>
<td>DOF of $t$ distribution</td>
<td>5.9313***</td>
<td>6.7107***</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-2499.9</td>
<td>-2482.4</td>
</tr>
<tr>
<td>AIC</td>
<td>4.1802</td>
<td>4.1525</td>
</tr>
</tbody>
</table>

***, ***, and * denote significance at the 1%, 5%, and 10% significance level.
for those models, we can see that, with the assumption of a $t$ distribution for the residuals, the fitting performs better than with a Gaussian distribution assumption. This shows that the distribution of the return series is fat-tailed. Therefore, the assumption of a $t$ distribution for the residual error in the GARCH-CJ model and the CJR-CJ model is assumed to follow a Gaussian distribution, otherwise insignificant. Form this, we can know that, in China's stock market, the lagged continuous sample path variation contains relatively more information for predicting the current volatility, while the lagged discontinuous jump variation contains relatively less information for forecasting. In addition, regardless of whether the residual error follows a Gaussian distribution or a $t$ distribution, the AIC value for the GARCH-CJ-type model is lower than the GARCH-RV-type and the GARCH-type models, which fully demonstrates that the fitting of the GARCH-CJ-type model has a better fitting effect.

<table>
<thead>
<tr>
<th>Table 3: Estimation results of EGARCH and its extended model.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Residual following Gaussian distribution</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>EGARCH</td>
</tr>
<tr>
<td><strong>μ_{t-1}</strong></td>
</tr>
<tr>
<td><strong>ω</strong></td>
</tr>
<tr>
<td><strong>α</strong></td>
</tr>
<tr>
<td><strong>β</strong></td>
</tr>
<tr>
<td><strong>θ</strong></td>
</tr>
<tr>
<td><strong>λ</strong></td>
</tr>
<tr>
<td><strong>γ</strong></td>
</tr>
<tr>
<td>DOF of $t$ distribution</td>
</tr>
<tr>
<td>Log likelihood</td>
</tr>
<tr>
<td>AIC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4: Estimation results of EGJR and its extended model.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Residual following Gaussian distribution</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CJR</td>
</tr>
<tr>
<td><strong>μ_{t-1}</strong></td>
</tr>
<tr>
<td><strong>ω</strong></td>
</tr>
<tr>
<td><strong>α</strong></td>
</tr>
<tr>
<td><strong>β</strong></td>
</tr>
<tr>
<td><strong>θ</strong></td>
</tr>
<tr>
<td><strong>λ</strong></td>
</tr>
<tr>
<td><strong>γ</strong></td>
</tr>
<tr>
<td>DOF of $t$ distribution</td>
</tr>
<tr>
<td>Log likelihood</td>
</tr>
</tbody>
</table>

***, **, and * denote significance at the 1%, 5%, and 10% significance level.

3.2. The Comparison of Model Predictive Power

3.2.1. In-Sample Prediction. In order to confirm whether the GARCH-CJ-type model has more predictive power to future volatility than the GARCH-type model and the GARCH-RV-type model, this paper compares the predictive power...
of these three types of models using a loss function. We select Mean Absolute Error (MAE), Heteroskedastic adjusted Mean Absolute Error (HMAE), Root Mean Squared Error (RMSE), and Heteroskedastic adjusted Root Mean Squared Error (HRMSE) as 4 indexes to evaluate and analyze the performance of the volatility models. Generally, the smaller the four are, the stronger predictive power the corresponding model has to predict future volatility. The formulae for getting the values of MAE, HMAE, RMSE, and HRMSE are expressed in (21). Since volatility cannot be directly observed in the stock market, scholars ([6, 16, 17]) usually use the realized volatility (RV) as a substitute for the volatility in Day t. In this paper, RVt is also used as the substitute.

\[
\begin{align*}
\text{MAE} & = \frac{1}{n} \sum_{t=1}^{n} |\sigma_t^2 - \hat{\sigma}_t^2|, \\
\text{HMAE} & = \frac{1}{n} \sum_{t=1}^{n} \frac{|\sigma_t^2 - \hat{\sigma}_t^2|}{\sigma_t^2}, \\
\text{RMSE} & = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (\sigma_t^2 - \hat{\sigma}_t^2)^2}, \\
\text{HRMSE} & = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(\frac{\sigma_t^2 - \hat{\sigma}_t^2}{\sigma_t^2}\right)^2},
\end{align*}
\]

(21)

where \( n \) is the size of the predictive sample, and \( \sigma_t^2 \) is the real volatility; that is, \( RV_t; \hat{\sigma}_t^2 \) denotes the predicted volatility.

Table 5 lists the statistics of in-sample predictive power evaluation index values for the GARCH type model, the GARCH-RV type model, and the GARCH-CJ-type model when using lag 1 data to predict the current volatility. Comparing the value for each evaluation index, we can see that, except that the HRMSE value for the GARCH-RV type model is greater than that for the GARCH type model, the RMSE for the EGARCH-CJ-type model is larger than that for the EGARCH-RV-type model, all values for the GARCH-CJ-type model are smaller than that for the GARCH-RV-type model, and the value for the GARCH-RV type model is lesser than that for the GARCH-type model. Therefore, we can presume that in forecasting the in-sample volatility the GARCH-CJ-type model has a greater in-sample predictive power than the GARCH-RV-type model, and the GARCH-RV-type model has greater in-sample predictive power than the GARCH type model.

### 3.2.2 Out-of-Sample Prediction.

Compared with the in-sample predictive power of the model, we are more concerned about the out-of-sample predictive power, since it has more practical value. In order to effectively evaluate out-of-sample predictive power, we divide the sample (April 20, 2007–April 20, 2012) into two parts. The first part (April 20, 2007–November 20, 2011) totals 1099 samples to be used for model estimation; the second part (November 21, 2011–2012, April 20) totals 100 samples to be used for prediction. As in the in-sample part for model estimation, we still use the loss function to compare the effectiveness of the prediction performed by the models. The results are shown in Table 6.

Comparing the value for each evaluation index, we can see that, except that the RMSE value for the GARCH-RV type model is greater than that for the GARCH type model in the case where both the models’ residuals are assumed to follow a \( t \) distribution, the MAE value and the HRMSE value for the CJR-RV-type model are larger than that for the CJR type model in the case where both the models’ residuals are assumed to follow a \( t \) distribution, all values for the GARCH-CJ-type model are smaller than that for the GARCH-RV-type model, and the value for the GARCH-RV-type model is lesser than that for the GARCH type model. So we can presume that, in terms of the out-of-sample predictive power for volatility, GARCH-CJ-type model works better than the GARCH-RV-type model and, in turn, the latter is superior to the GARCH-type model.

Combining the discussion in Section 3.2.1 with that in Section 3.2.2, we can see that among the above three types of volatility models the GARCH-CJ-type model performs the best in predicting future volatility. Therefore, it makes sense to introduce the realized volatility (RV) into the GARCH-type model and decompose it into continuous sample path variation (C) and discontinuous jump variation (J) to enhance the model’s predictive power for volatility.
Table 6: Statistics of out-of-sample predictive power evaluation index.

<table>
<thead>
<tr>
<th></th>
<th>Residual following Gaussian distribution</th>
<th>Residual following $t$ distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>MAE = 1.1044</td>
<td>HRMSE = 1.2693</td>
</tr>
<tr>
<td>GARCH-RV</td>
<td>HMAE = 1.0219</td>
<td>RMSE = 1.3689</td>
</tr>
<tr>
<td>GARCH-CJ</td>
<td>RMSE = 1.474</td>
<td>HRMSE = 1.2693</td>
</tr>
<tr>
<td>EGARCH</td>
<td>MAE = 1.0533</td>
<td>HMAE = 0.9980</td>
</tr>
<tr>
<td>EGARCH-RV</td>
<td>HMAE = 0.9314</td>
<td>RMSE = 1.3743</td>
</tr>
<tr>
<td>EGARCH-CJ</td>
<td>RMSE = 1.3858</td>
<td>HRMSE = 1.2086</td>
</tr>
<tr>
<td>CJsR</td>
<td>MAE = 0.9690</td>
<td>HMAE = 0.9717</td>
</tr>
<tr>
<td>CJR-RV</td>
<td>HMAE = 0.8368</td>
<td>RMSE = 1.3529</td>
</tr>
<tr>
<td>CJR-CJ</td>
<td>RMSE = 1.3377</td>
<td>HRMSE = 1.0295</td>
</tr>
</tbody>
</table>

4. Conclusions

This paper constructs GARCH-CJ model on the basis of the GARCH-RV model to obtain a volatility model that can better measure and predict asset volatility. And, in order to test the validity of the model, an empirical study is carried out using the 5-minute high frequency data of HUSHEN 300 index in China (April 20, 2007, to April 20, 2012), we estimate the parameters of the GARCH-type model, the GARCH-RV-type model, and the GARCH-CJ-type model and evaluate all models’ predictive power for future market volatility using a loss function (MAE, HMAE, RMSE, and HRMSE).

From the results of the estimated parameters, we can see that favorable and unfavorable news have an asymmetric impact on the market volatility in China’s stock market, and the distribution of the market return series is leptokurtic. At the same time, through the empirical results, we can draw some conclusions as follows.

1. The past continuous sample path variation has more predictive power for future volatility, but the past discontinuous jump variation has less information to predict.

2. The GARCH-CJ-type model has a much better fitting of the future volatility than other two types of models (the GARCH-type model and GARCH-RV-type model).

3. According to the comparison of the predictive power of the three types of models, the GARCH-RV model performs the better in predicting the future volatility than the GARCH-type models, which is consistent with Koopman, Fuertes et al., and Lehnert et al. [6–8].

4. The proposed GARCH-CJ-type model in this paper has a better ability to predict the future volatility than the other two types of models, which means the application of GARCH-CJ model is more reasonable in measuring and predicting volatility in financial practices such as capital asset pricing, financial derivatives pricing, and risk measures.

Although GARCH-CJ model has a greater power to predict the market volatility, it is still necessary to further increase the accuracy of measuring and predicting the market volatility. Therefore, the GARCH-CJ-type model, further improvement in the fitting, and predictive accuracy of the volatility models will be our emphasis for further research.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The Research is sponsored by National Science Fund for Distinguished Young Scholars (70825006), Natural Science Foundation of China (71171076), and Scientific Research Foundation for the Returned Overseas Chinese Scholars (2011 [185]).

References


