Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 386721, 8 pages
http://dx.doi.org/10.1155/2014/386721

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

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:

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t} \cdot v_t, \qquad h_t = \omega + \alpha\varepsilon_{t-1}^2 + \theta h_{t-1}, \tag{1}$$

where R_t is the return, μ_{t-1} denotes the conditional mean of R_t based on all available information, h_t is the volatility, ν_t is the white noise disturbance, and ω, α, and θ 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

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t} \cdot v_t, \qquad h_t = \omega + \alpha\varepsilon_{t-1}^2 + \theta h_{t-1} + \lambda \mathrm{RV}_{t-1}, \tag{2}$$

where λ is also a parameter to be estimated as ω, α, θ, 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

$$\mathrm{RV}_t = \sum_{i=1}^{N} r_{t,j}^2 + r_{t,n}^2 = \sum_{j=1}^{M} r_{t,j}^2, \quad M = N + 1, \tag{3}$$

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, $r_{t,1} = 100(\ln P_{t,1} - \ln P_{t,o})$, $P_{t,1}$ is the first closing price at Day t, $P_{t,o}$ is the opening price at Day t, $r_{t,2}$ is the second return after opening at Day t, and $r_{t,2} = 100(\ln P_{t,2} - \ln P_{t,1})$; ..., so $r_{t,N}$ expresses the Nth 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, $r_{t,n} = r_{t,M} = 100(\ln P_{t,o} - \ln P_{t-1,c})$, and $P_{t-1,c}$ 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 (RBV); that is,

$$\mathrm{RBV}_t^{[r,s]} = \left(\frac{h}{M}\right)^{1-(r+s)/2} \sum_{j=1}^{M-1} \left|r_{j,t}\right|^r \left|r_{j+1,t}\right|^s, \tag{4}$$

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 RBV_t is equivalent to a consistent estimator for discontinuous jump variation J_t

$$\mathrm{RV}_t - \mathrm{RBV}_t \xrightarrow{M \to \infty} J_t. \tag{5}$$

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:

$$J_t = \max\left[\mathrm{RV}_t - \mathrm{RBV}_t, 0\right]. \tag{6}$$

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:

$$Z_t = \frac{\left(\mathrm{RV}_t - \mathrm{RBV}_t\right)\mathrm{RV}_t^{-1}}{\sqrt{\left((\pi/2)^2 + \pi - 5\right)(1/M)\max\left(1, \mathrm{RTQ}_t/\mathrm{RBV}_t^2\right)}} \longrightarrow N(0,1), \tag{7}$$

where

$$\mathrm{RTQ}_t = M\mu_{4/3}^{-3}\left(\frac{M}{M-4}\right)\sum_{j=4}^{M}\left|r_{t,j-4}\right|^{4/3}\left|r_{t,j-2}\right|^{4/3}\left|r_{t,j}\right|^{4/3},$$
$$\left(\mu_{4/3} = E\left(\left|Z_T\right|^{4/3}\right) = 2^{2/3}\Gamma\left(\frac{7}{6}\right)\Gamma\left(\frac{1}{2}\right)^{-1}\right). \tag{8}$$

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_t as the robust estimator for J_t is biased, and this paper adopts MedRV_t proposed by Andersen et al. [15] as a robust estimator instead. MedRV_t can be expressed as follows:

$$\mathrm{MedRV}_t = \frac{\pi}{6-4\sqrt{3}+\pi}\left(\frac{M}{M-2}\right) \times \sum_{i=2}^{M-1}\mathrm{Med}\left(\left|r_{t,i-1}\right|\left|r_{t,i}\right|\left|r_{t,i+1}\right|\right)^2. \tag{9}$$

Accordingly, $\mathrm{RTQ}_{1,t}$, the statistic for Z_t in (6), is replaced by MedRTQ_t, which is expressed as follows:

$$\mathrm{MedRTQ}_t = \frac{3\pi M}{9\pi+72+-52\sqrt{3}}\left(\frac{M}{M-2}\right) \times \sum_{i=2}^{M-1}\mathrm{Med}\left(\left|r_{t,i-1}\right|,\left|r_{t,i}\right|,\left|r_{t,i+1}\right|\right)^4. \tag{10}$$

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:

$$J_t = I\left(Z_t > \phi_\alpha\right)\left(\mathrm{RV}_t - \mathrm{MedRV}_t\right). \tag{11}$$

Accordingly, the continuous sample path variation estimator is

$$C_t = I\left(Z_t \le \phi_\alpha\right)\mathrm{RV}_t + I\left(Z_t > \phi_\alpha\right)\mathrm{MedRV}_t. \tag{12}$$

In actual calculation, we need to select a suitable confidence level α. Drawing on previous research, we choose 0.99 as the confidence level α 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

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}\cdot v_t,$$
$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \theta h_{t-1} + \lambda C_{t-1} + \gamma J_{t-1}. \tag{13}$$

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(1,1) can be presented as

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}v_t,$$
$$\ln\left(h_t\right) = \omega + \alpha\left|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right| + \beta\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \theta\ln\left(h_{t-1}\right). \tag{14}$$

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

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}v_t,$$
$$\ln\left(h_t\right) = \omega + \alpha\left|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right| + \beta\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \theta\ln\left(h_{t-1}\right) + \lambda\ln\left(\mathrm{RV}_{t-1}\right). \tag{15}$$

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

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}v_t,$$
$$\ln\left(h_t\right) = \omega + \alpha\left|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right| + \beta\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \theta\ln\left(h_{t-1}\right) + \lambda\ln\left(C_{t-1}\right) + \gamma\ln\left(J_{t-1}+1\right). \tag{16}$$

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

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}v_t,$$
$$h_t = \omega + \left(\alpha + \beta N_{t-1}\right)\varepsilon_{t-1}^2 + \theta h_{t-1}, \tag{17}$$

where N_{t-1} is the indicator variable of the negative ε_{t-1}

$$N_{t-1} = \begin{cases} 1, & \varepsilon_{t-1} < 0 \\ 0, & \varepsilon_{t-1} \ge 0. \end{cases} \tag{18}$$

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:

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}\cdot v_t,$$
$$h_t = \omega + \left(\alpha + \beta N_{t-1}\right)\varepsilon_{t-1}^2 + \theta h_{t-1} + \lambda\mathrm{RV}_{t-1}. \tag{19}$$

We divide RV_{t-1} into C_{t-1} and J_{t-1}, and we get the CJR-CJ model

$$R_t = \mu_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}\cdot v_t,$$
$$h_t = \omega + \left(\alpha + \beta N_{t-1}\right)\varepsilon_{t-1}^2 + \theta h_{t-1} + \lambda C_{t-1} + \gamma J_{t-1}. \tag{20}$$

Table 1: Descriptive statistics of each variable.

	Mean	Std. dev.	Skewness	Kurtosis	Jarque-Bera	ADF-t statistic
R_t	−0.0152	2.1141	−0.2647	4.9345	200.95***	−32.901***
RV_t	4.3471	6.8083	5.9174	49.075	113055***	−10.710***
C_t	3.3412	4.3445	5.9456	65.248	200641***	−7.7154***
J_t	1.0058	4.8285	9.5878	109.80	588177***	−16.145***
$\ln(\mathrm{RV}_t)$	0.9600	0.9323	0.5180	3.2163	55.961***	−4.9183***
$\ln(C_t)$	0.7650	0.8992	0.3525	2.8382	26.138***	−5.3556***
$\ln(J_t + 1)$	0.2825	0.6209	3.0445	14.085	7990.6***	−16.266***

***denotes significance at 1% significance level.

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

	Residual following Gaussian distribution			Residual following t distribution		
	GARCH	GARCH-RV	GARCH-CJ	GARCH	GARCH-RV	GARCH-CJ
μ_{t-1}	−0.0007	−0.0400	−0.0457	0.0438	0.0091	−0.0003
ω	0.0324***	0.2759***	0.2170***	0.0132	0.2808***	0.1982**
α	0.0392***	−0.0468***	−0.0558***	0.0368***	−0.0667***	−0.0620***
θ	0.9520***	0.6835***	0.7055***	0.9607***	0.6537***	0.6750***
λ		0.3254***	0.3865***		0.3903***	0.4508***
γ			0.0678*			0.0906
DOF of t distribution				5.9313***	6.7107***	7.3077***
Log likelihood	−2499.9	−2482.4	−2469.3	−2474.6	−2463.2	−2453.0
AIC	4.1802	4.1525	4.1324	4.1396	4.1222	4.1068

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

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 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 J_t. All are processed on Matlab 7.0 or Excel 2003.

Table 1 is the descriptive statistics of return R_t, realized volatility RV_t, continuous sample path variation C_t, and discontinuous jump variation J_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 3: Estimation results of EGARCH and its extended model.

	Residual following Gaussian distribution			Residual following t distribution		
	EGARCH	EGARCH-RV	EGARCH-CJ	EGARCH	EGARCH-RV	EGARCH-CJ
μ_{t-1}	−0.0361	0.0002	−0.0089	0.0259	0.0374	0.0279
ω	-0.0598^{***}	0.2982^{***}	0.3384^{***}	-0.0502^{***}	0.3079^{***}	0.3769^{***}
α	0.1055^{***}	-0.1682^{***}	-0.1686^{***}	0.0730^{***}	-0.1951^{***}	-0.2131^{***}
β	-0.0374^{***}	-0.1070^{***}	-0.1074^{***}	-0.0320^{***}	-0.1093^{***}	-0.1142^{***}
θ	0.9849^{***}	0.6568^{***}	0.6513^{***}	0.9954^{***}	0.6544^{***}	0.6253^{***}
λ		0.2780^{***}	0.2849^{***}		0.2928^{***}	0.3247^{***}
γ			0.0524			0.0498
DOF of t distribution				6.2963^{***}	7.3426^{***}	7.1035^{***}
Log likelihood	−2497.0	−2465.8	−2462.9	−2472.3	−2450.1	−2446.5
AIC	4.1769	4.1265	4.1234	4.1374	4.1020	4.0976

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

Table 4: Estimation results of EGJR and its extended model.

	Residual following Gaussian distribution			Residual following t distribution		
	CJR	CJR-RV	CJR-CJ	CJR	CJR-RV	CJR-CJ
μ_{t-1}	−0.0235	0.0123	−0.0106	0.0361	0.0433	0.0168
ω	0.0505^{***}	0.1510^{***}	0.1904^{***}	0.0152	0.1470^{***}	0.1844^{***}
α	0.0258^{***}	-0.0581^{***}	-0.0864^{***}	0.0150	-0.0608^{***}	-0.0960^{***}
β	0.0371^{***}	0.0982^{***}	0.0752^{***}	0.0342^{**}	0.0952^{***}	0.0703^{**}
θ	0.9420^{***}	0.8437^{***}	0.7588^{***}	0.9629^{***}	0.8253^{***}	0.7112^{***}
λ		0.1309^{***}	0.3132^{***}		0.1595^{***}	0.4042^{***}
γ			0.0472^{*}			0.0635
DOF of t distribution				6.2568^{***}	7.4200^{***}	7.6072^{***}
Log likelihood	−2497.6	−2474.6	−2463.8	−2471.8	−2459.5	−2449.6
AIC	4.1780	4.1413	4.1249	4.1366	4.1176	4.1028

***, **, 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 of the GARCH model is more reasonable.

From the analysis of Tables 2–4, the coefficients of RV_{t-1} or $\ln(\mathrm{RV}_{t-1})$ of newly added exogenous variable, λ, in the volatility equation of the GARCH-type models are significantly positive at 1% significance level, which shows that market volatility in China's stock exhibits pronounced persistence and the last period volatility may serve as an indicator for the current period volatility. In addition, comparing the AIC values for the GARCH-RV-type model and the GARCH-type model, we can see that the fitting for the GARCH-RV model works better, which is consistent with Koopman et al. [6]. When it comes to the estimation results for this paper's newly built GARCH-CJ model, the coefficients for C_{t-1} (λ) are significantly positive at the 1% significance level, and the coefficients for J_{t-1} (γ) are significant only when 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.

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

Table 5: Statistics of in-sample predictive power evaluation index.

	Residual following Gaussian distribution				Residual following t distribution			
	MAE	HMAE	RMSE	HRMSE	MAE	HMAE	RMSE	HRMSE
GARCH	2.7463	0.9407	6.1528	1.2807	2.8447	0.9643	6.1921	1.3061
GARCH-RV	2.7329	0.9099	5.8897	1.4140	2.8234	0.9428	5.9979	1.4902
GARCH-CJ	2.5880	0.8296	5.7739	1.1701	2.6306	0.8407	5.7581	1.1966
EGARCH	2.7051	0.9258	6.1524	1.2666	2.8544	0.9582	6.2123	1.3075
EGARCH-RV	2.5862	0.8383	5.9059	1.1755	2.6549	0.8600	5.9339	1.2134
EGARCH-CJ	2.5143	0.8156	6.0180	1.1164	2.6222	0.8445	6.0682	1.1671
CJR	2.7088	0.9249	6.1899	1.2645	2.8155	0.9466	6.2264	1.2905
CJR-RV	2.5851	0.8657	5.8593	1.2567	2.6234	0.8756	5.8409	1.2858
CJR-CJ	2.5550	0.8145	5.7933	1.1479	2.6015	0.8257	5.7778	1.1748

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_t) as a substitute for the volatility in Day t. In this paper, RV_t is also used as the substitute

$$\begin{aligned}
\mathrm{MAE} &= \frac{1}{n}\sum_{t=1}^{n}\left|\sigma_t^2 - \widehat{\sigma}_t^2\right|, \\
\mathrm{HMAE} &= \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\sigma_t^2 - \widehat{\sigma}_t^2}{\sigma_t^2}\right|, \\
\mathrm{RMSE} &= \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\sigma_t^2 - \widehat{\sigma}_t^2\right)^2}, \\
\mathrm{HRMSE} &= \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left[\frac{\sigma_t^2 - \widehat{\sigma}_t^2}{\sigma_t^2}\right]^2},
\end{aligned} \tag{21}$$

where n is the size of the predictive sample, and σ_t^2 is the real volatility; that is, RV_t; $\widehat{\sigma}_t^2$ 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.

	Residual following Gaussian distribution				Residual following t distribution			
	MAE	HMAE	RMSE	HRMSE	MAE	HMAE	RMSE	HRMSE
GARCH	1.1044	1.0219	1.4174	1.2949	1.0789	0.9980	1.3689	1.2693
GARCH-RV	1.0533	0.9314	1.3858	1.1903	1.0719	0.9418	1.3743	1.2086
GARCH-CJ	0.9690	0.8368	1.3377	1.0546	0.9717	0.8211	1.3529	1.0295
EGARCH	1.2632	1.1594	1.4958	1.4401	1.1376	1.0481	1.4418	1.3155
EGARCH-RV	1.0578	0.9539	1.4287	1.2637	1.0541	0.9387	1.4341	1.2548
EGARCH-CJ	1.0130	0.8897	1.4067	1.1672	1.0212	0.8741	1.3940	1.1505
CJR	1.1685	1.0884	1.4006	1.3629	0.9992	0.9218	1.3649	1.1777
CJR-RV	1.0270	0.9330	1.3272	1.2248	1.0131	0.9099	1.3327	1.1913
CJR-CJ	0.9629	0.8521	1.3085	1.0939	0.9490	0.8155	1.3171	1.0385

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

[1] R. E. Engle, "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kindom inflation," *Econometrica*, no. 50, pp. 987–1007, 1982.

[2] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[3] D. B. Nelson, "Conditional heteroscedasticity in asset returns: a new approach," *Econometrica*, vol. 59, no. 2, pp. 347–370, 1991.

[4] L. Glosten, R. Jagannathan, and D. Runkle, "On the Relation between expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, vol. 48, no. 5, pp. 1779–1801, 1993.

[5] T. G. Andersen and T. Bollerslev, "Answering the skeptics: yes, standard volatility models do provide accurate forecasts," *International Economic Review*, vol. 39, no. 4, pp. 885–905, 1998.

[6] S. J. Koopman, B. Jungbacker, and E. Hol, "Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements," *Journal of Empirical Finance*, vol. 12, no. 3, pp. 445–475, 2005.

[7] A. Fuertes, M. Izzeldin, and E. Kalotychou, "On forecasting daily stock volatility: the role of intraday information and market conditions," *International Journal of Forecasting*, vol. 25, no. 2, pp. 259–281, 2009.

[8] B. Frijns, T. Lehnert, and R. C. J. Zwinkels, "Modeling structural changes in the volatility process," *Journal of Empirical Finance*, vol. 18, no. 3, pp. 522–532, 2011.

[9] T. G. Andersen, T. Bollerslev, and F. X. Diebold, "Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility," *Review of Economics and Statistics*, vol. 89, no. 4, pp. 701–720, 2007.

[10] C. Huang, X. Gong, X. Chen, and F. Wen, "Measuring and forecasting volatility in Chinese stock market using HAR-CJ-M model," *Abstract and Applied Analysis*, vol. 2013, Article ID 143194, 13 pages, 2013.

[11] M. Martens, "Measuring and forecasting S&P 500 index-futures volatility using high-frequency data," *Journal of Futures Markets*, vol. 22, no. 6, pp. 497–518, 2002.

[12] F. Wen, Z. Li, C. Xie, and S. David, "Study on the fractal and chaotic features of the Shanghai composite index," *Fractals-Complex Geometry Patterns and Scaling in Nature and Society*, vol. 20, no. 2, pp. 133–140, 2012.

[13] O. E. Barndorff-Nielsen and N. Shephard, "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, vol. 4, no. 1, pp. 1–30, 2006.

[14] O. E. Barndorff-Nielsen and N. Shephard, "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, vol. 4, no. 1, pp. 1–30, 2006.

[15] T. G. Andersen, D. Dobrev, and E. Schaumburg, "Jump-robust volatility estimation using nearest neighbor truncation," *Journal of Econometrics*, vol. 169, no. 1, pp. 75–93, 2012.

[16] F. Corsi, "A simple approximate long-memory model of realized volatility," *Journal of Financial Econometrics*, vol. 7, no. 2, pp. 174–196, 2009.

[17] Y. Wei, "Volatility forecasting models for CSI300 index futures," *Journal of Management Sciences in China*, no. 2, pp. 66–76, 2010 (Chinese).