Finance Research Letters 28 (2019) 438–443

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

A new variant of RealGARCH for volatility modeling

Haibin Xie*[a], Nan Qi[a], Shouyang Wang[b]

[a] University of International Business and Economics, Huixin East Road 10, Chaoyang District, Beijing 100029, China
[b] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Zhongguancun East Road, Haidian District, Beijing 100190, China

ARTICLE INFO

Keywords:
CARR
GARCH@CARR
RealGARCH
Volatility

JEL classification:
C5
C22

ABSTRACT

A new variant of RealGARCH model is proposed for volatility modeling and forecasting. The main difference between this variant and the existed variants is that we use a multiplicative error model (MEM) structure to the measurement equation. Empirical studies are performed on several stock indices to evaluate our model specification, and the results turn out to be promising.

1. Introduction

With the accessibility of high-frequency financial data, a number of new informative volatility measures have been proposed, including realized variance, bipower variation, the realized kernel, price range, realized range, and many related quantities (see Parkinson, 1980; Andersen and Bollerslev, 1998; Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2002, 2004; Barndorff-Nielsen et al., 2008; Martens and van Dijk, 2007; and references therein). Any of these measures is far more informative for modeling and forecasting future volatility than is the squared return.

Different methods have been proposed to incorporate more informative volatility measures into return volatility estimation. A natural idea is to estimate a GARCH model that includes a volatility measure in the GARCH equation, known as a GARCH-X model (Engle, 2002; Forsberg and Bollerslev, 2002). A problem with the GARCH-X framework is that it pays no effort to explain the variation in the volatility measures, so these GARCH-X models fails when volatility beyond a single period is required.

The recent proposed RealGARCH (Hansen et al., 2012) provides an efficient framework for volatility modeling. The main and the most important feature of RealGARCH is that it is a joint modeling of return and realized volatility measure, which thus makes possible the multi-step-ahead volatility forecasts.

The main contribution in this paper is that we propose a new variant of RealGARCH model. Our variant differs from the existed variants in that we specify a Multiplicative Error Model structure, MEM (Engle, 2002) to the measurement equation. Compared with the existed variants, our variant has fewer parameters to estimate. Sharma and Vipul (2016) perform a comprehensive empirical study on 16 stock indices and find that better volatility forecasts can be obtained if simpler time series models with realized measures are used. Therefore our RealGARCH variant with MEM structure is an advantage.

Taking price range as a realized volatility measure, we investigate the empirical performance of our RealGARCH variant on several stock indices, and the results turn out to be promising. The structure of this paper is organized as follows. Section 2 proposes the new RealGARCH variant together with some discussions. Section 3 present the empirical results for both in-sample and out-of-

* Corresponding author.
E-mail addresses: hbxie@amss.ac.cn (H. Xie), clytzepi@126.com (N. Qi), sywang@amss.ac.cn (S. Wang).

https://doi.org/10.1016/j.frl.2018.06.015
Received 27 April 2018; Received in revised form 14 June 2018; Accepted 23 June 2018
Available online 18 July 2018

sample estimations. Section 4 concludes.

2. The model

In this section we first present the general RealGARCH framework, and then introduce a new variant of RealGARCH. To simplify the presentation in the following passage we will assume that $E(r_t|\mathscr{F}_{t-1}) = 0$, where $\{r_t\}$ a return series. A more general specifications for conditional mean, such as a constant or the GARCH-in-mean by Engle et al. (1987), is accommodated by reinterpreting r_t as the return less its conditional mean.

2.1. RealGARCH framework

The general structure of RealGARCH(p, q) model is given by

$$r_t = \sqrt{h_t} z_t, \tag{1}$$

$$h_t = \upsilon(h_{t-1}, \ldots, h_{t-p}, \ldots, x_{t-1}, \ldots, x_{t-q}), \tag{2}$$

$$x_t = m(h_t, z_t, u_t), \tag{3}$$

where x_t is a kind of realized volatility measure and $z_t \sim$ i.i.d. (0,1), with z_t and u_t being mutually independent. The first two equations are referred to as *return equation* and *GARCH equation*, and the last one is called *measurement equation*. Different specifications to functions $\upsilon(.)$ and $m(.)$ give different RealGARCH variants.

2.2. RealGARCH with a MEM to measurement equation

The new RealGARCH(p,q) variant proposed in this paper is presented as follows

$$r_t = \rho \lambda_t z_t, \tag{4}$$

$$ln(\lambda_t) = \omega + \sum_{i=1}^{p} \alpha_i ln(\lambda_{t-i}) + \sum_{j=1}^{q} \beta_j ln(x_{t-j}) + \tau(z_{t-1}), \tag{5}$$

$$x_t = \lambda_t u_t, \tag{6}$$

where $u_t \sim$ i.i.d. $D(1, \theta)$ with unit mean,[1] $\tau(z_{t-1})$ is called the *leverage function*, and λ_t is the conditional mean of the realized volatility. Scale factor ρ is used to make sure $\rho\lambda_t$ is an unbiased volatility estimation.

The most significant difference between our model specification and those RealGARCH variants proposed in Hansen et al. (2012) is that we specify a MEM structure to the measurement equation. The MEM structure simplifies the measurement equation structure, and thus makes our model has fewer parameters to estimate.

It is clear from our model specification that it is the conditional mean of the realized volatility that determines the dynamics of the asset volatility. Since the conditional mean of the realized volatility is specified to follow a conditional autoregressive structure, we thus call our variant as a GARCH model with Conditional AutoRegressive Realized volatility structure, or GARCH@CARR for short.

2.3. Model estimation

The joint log-likelihood function of the our model (conditionally on $\mathscr{F}_{t-1}$) is presented as

$$logL(\{r_t, x_t\}_{t=1}^{n}; \Theta) = \sum_{t=1}^{n} logH(r_t, x_t|\mathscr{F}_{t-1}),$$

where Θ is the set of the parameters to be estimated, $H(.)$ is a joint density for (r_t, x_t). With the assumption that z_t is independent of u_t, we can factorize the joint conditional density by

$$H(r_t, x_t|\mathscr{F}_{t-1}) = h_r(r_t|\mathscr{F}_{t-1}) h_x(x_t|\mathscr{F}_{t-1}),$$

where $h_.(.)$ is the marginal density. Thus the log-likelihood function can be rewritten as

$$logL(\{r_t, x_t\}_{t=1}^{n}; \Theta) = \sum_{t=1}^{n} logh_1(r_t|\mathscr{F}_{t-1}) + \sum_{t=1}^{n} logh_r(x_t|\mathscr{F}_{t-1}),$$

which is a sum of two partial log-likelihoods,

[1] The density function of u_t should be positively valued given that the realized volatility measure x_t is positive.

$$l(r) \;:=\; \sum_{t=1}^{n} log h_r(r_t|\mathscr{F}_{t-1}),$$

and

$$l(x) \;:=\; \sum_{t=1}^{n} log h_x(x_t|\mathscr{F}_{t-1}).$$

These two log-likelihoods are informative and quite useful: the partial log-likelihood, $l(r)$ can be used to compare with that of a return model, and the partial log-likelihood, $l(x)$ can be used to compare with that of a realized volatility measure model.

3. Empirical examples

High-low price range is of great interest to practitioners and proves to be a more informative volatility estimator than the return-based one (Parkinson, 1980). Using price range as a realized volatility measure, we present our GARCH@CARR model as follows

$$r_t \;=\; \rho\lambda_t z_t,\; z_t \sim N(0, 1), \tag{7}$$

$$ln(\lambda_t) \;=\; \omega + \alpha_1 ln(\lambda_{t-1}) + \beta_1 ln(R_{t-1}) + \pi_1 z_{t-1}, \tag{8}$$

$$R_t \;=\; \lambda_t u_t,\; u_t \sim LN(-\sigma^2/2, \sigma), \tag{9}$$

where R_t is the price range defined as the difference between log high price and log low price,[2] $LN(.)$ means log normal distribution,[3] parameter π_1 is used to capture the leverage effect. Eqs. (8) and (9) are also known as the CARR model (Chou, 2005; Chiang et al., 2016).

3.1. Data description

We investigate the empirical performance of our model specification on several stock indices, including S&P500 Index, Dow Jones Industrial Average (DJIA), Hang Seng Index (HSI) of Hong Kong and DAX Index of Germany. We collect the daily data for period from January 3rd, 1995 to December 30th, 2016. For each day four pieces of price information, the high, low, opening and closing prices are reported. All the data sets are downloaded from the website "www.finance.yahoo.com".

Table 1 presents the summary statistics for the range, the log range, the return, and the squared return of each index. Consistent with the well documented fact, asset return is found to be negatively skewed with high kurtosis. The log range reports the smallest Jarque–Bera statistics, which is consistent with Alizadeh et al. (2002) that log range approximates normal distribution. It is more interesting to compare the Q statistics for the range with that for the squared return. The Q statistics show that there is a significant difference between the range and the squared return, which means the volatility information in the range is different from that in the squared return.

3.2. In-sample estimation

For in-sample volatility estimation, we use both GARCH and EGARCH as the benchmarks. We only consider models of order (1, 1) because it has been well documented that order of (1, 1) is sufficient for capturing the persistence in volatility.

Table 2 presents the maximum likelihood estimation results. First, the point estimates show significant leverage effect since π_1s are reported to be negative and statistically significant. Next, we compare the partial log likelihood statistics, $l(r)$ in the GARCH@CARR with the those in the GARCH and EGARCH. The log likelihood statistic shows that the GARCH@CARR model, though does not maximize the log likelihood function of the return equation, still produces a better empirical fit than both the GARCH and EGARCH models. This result is consistent with the finding of Li and Hong (2011) since range is a more informative volatility measure. Scale factor ρ shows deviation from 0.6267, which is consistent with our expectation since the assumptions in Feller (1951) are hardly satisfied in reality.

For in-sample realized volatility measure estimation, we use the CARR model of Chiang et al. (2016) as a benchmark, and the results are also presented in Table 2. It interesting to note that the log likelihood statistics, l(R)s reported by the GARCH@CARR model are almost equal to those reported by the CARR model, which means that the GARCH@CARR model almost has the same ability with the CARR in explaining the range variation.

3.3. Out-of-sample forecast

We further evaluate our model through out-of-sample forecasting performance. A recursive (expanding) window forecasting

[2] Under the assumption that the logarithmic price p_s follows a Brownian motion with zero drift and diffusion σ, Feller (1951) have shown price range is related to the diffusion σ by $\lambda_t = E[R_t] = \sqrt{\frac{8}{\pi}}\sigma_t \approx 0.6267\sigma_t$.

[3] Log normal assumption is used because Alizadeh et al. (2002) showed that log range estimator is approximately Gaussian.

Table 1
Summary statistics of the range, log range, the return and the squared return.

	S&P500				DJIA			
	R_t	$ln(R_t)$	r_t	r_t^2	R_t	$ln(R_t)$	r_t	r_t^2
Mean	0.014	−4.488	2.86E−04	1.43E−04	0.013	−4.484	2.97E−04	1.27E−04
Std	9.75E−03	0.595	0.012	4.54E−04	9.38E−03	0.570	0.011	3.99E−04
Min	2.01E−03	−6.210	−0.095	0.000	1.67E−03	−6.397	−0.082	0.000
Max	0.109	−2.216	0.110	0.012	0.122	−2.107	0.105	0.011
Skewness	3.127	0.243	−0.250	12.374	3.345	0.263	−0.167	12.763
Kurtosis	20.303	3.111	11.138	231.280	23.358	3.301	10.875	254.615
Jarque-Bera	7.81E04***	57.35***	1.53E04***	1.22E07***	1.06E05***	84.56***	1.43E04***	1.48E07***
Q(12)	20487***	18057***	74.877***	5201.1***	19358***	17146***	68.302***	4566.6***
	DAX				HSI			
Mean	0.018	−4.222	3.05E−04	2.24E−04	0.015	−4.349	1.90E−04	2.68E−04
Std	0.012	0.601	0.015	5.53E−04	0.010	0.550	0.016	9.28E−04
Min	1.44E−03	−6.545	−0.089	0.000	2.65E−03	−5.935	−0.147	0.000
Max	0.148	−1.912	0.108	0.012	0.176	−1.735	0.172	0.030
Skewness	2.381	0.128	−0.105	8.119	3.429	0.390	0.078	15.253
Kurtosis	12.800	2.980	7.112	110.801	28.287	3.204	12.988	345.791
Jarque-Bera	2.76E04***	15.33***	3.94E03***	2.76E06***	1.55E05***	147.28***	2.26E05***	2.68E07***
Q(12)	22972***	22730***	25.807**	3414.8***	15919***	16575***	24.583**	2937***

Notes: We use Q as the Ljung-Box Q statistics. Symbols *, **, *** mean respectively significance at 10%, 5%, and 1%.

Table 2
In-sample estimation results.

Panel A:	S&P500				DJIA			
Model	GARCH	EGARCH	GARCH@CARR	CARR	GARCH	EGARCH	GARCH@CARR	CARR
Panel A1: Point estimates								
ω	1.840E − 06 (1.963E−07)	−0.221 (0.02)	−0.097 (0.007)	−0.101 (0.009)	1.822E − 06 (1.883E−07)	−0.227 (0.017)	−0.094 (0.008)	−0.099 (0.010)
α_1	0.888 (0.007)	0.976 (0.002)	0.853 (0.000)	0.843 (0.000)	0.883 (0.006)	0.975 (0.002)	0.845 (0.000)	0.838 (0.000)
β_1	0.098 (0.006)	0.129 (0.009)	0.123 (0.002)	0.132 (0.002)	0.103 (0.006)	0.114 (0.010)	0.132 (0.002)	0.137 (0.002)
π_1		−0.139 (0.006)	−0.077 (0.003)	−0.099 (0.004)		−0.127 (0.006)	−0.070 (0.000)	−0.092 (0.004)
ρ			0.766 (0.009)				0.736 (0.008)	
σ			0.404 (0.004)	0.404 (0.004)			0.395 (0.004)	0.395 (0.004)
Panel A2: Auxiliary statistics								
ψ	0.986	0.976	0.976	0.975	0.986	0.975	0.976	0.975
l(r)	17799	17933	18030		18036	18151	18249	
l(R)			22029	22030			22121	22122
Panel B	DAX				HSI			
Panel B1: Point estimates								
ω	2.080E − 07 (2.026E−06)	−0.184 (0.018)	−0.078 (0.008)	−0.059 (0.000)	1.894E − 06 (3.022E−07)	−0.128 (0.015)	−0.037 (0.007)	−0.044 (0.008)
α_1	0.906 (0.005)	0.979 (0.002)	0.848 (0.007)	0.848 (0.008)	0.922 (0.005)	0.985 (0.002)	0.870 (0.000)	0.860 (0.000)
β_1	0.086 (0.005)	0.152 (0.009)	0.131 (0.006)	0.136 (0.007)	0.071 (0.005)	0.140 (0.009)	0.119 (0.002)	0.127 (0.002)
π_1		−0.093 (0.005)	−0.053 (0.003)	−0.064 (0.004)		−0.060 (0.005)	−0.024 (0.002)	−0.019 (0.003)
ρ			0.777 (0.009)				0.941 (0.011)	
σ			0.390 (0.004)	0.390 (0.004)			0.398 (0.004)	0.398 (0.004)
Panel B2: Auxiliary statistics								
ψ	0.992	0.979	0.978	0.983	0.993	0.984	0.989	0.987
l(r)	16456	16533	16606		15702	15755	15799	
l(R)			20864	20867			20899	20904

Notes: The return series has been demeaned by a constant. Numbers in the parenthesis are the standard errors. Parameter ψ measures the persistence.

Table 3
MAE and RMSE statistics for both volatility and high-low price range.

Panel A	GARCH	EGARCH	GARCH@CARR	GARCH	EGARCH	GARCH@CARR
Panel A1	S&P500			DJIA		
MAE	3.6E−03	3.1E−03	**2.7E−03**	3.2E−03	2.9E−03	**2.5E−03**
RMSE	5.0E−03	4.1E−03	**3.8E−03**	4.7E−03	4.2E−03	**3.9E−03**
Panel A2:	DAX			HSI		
MAE	3.9E−03	3.6E−03	**3.2E−03**	6.1E−03	6.0E−03	**4.9E−03**
RMSE	5.1E−03	4.6E−03	**4.3E−03**	7.6E−03	7.1E−03	**5.7E−03**
Panel B		CARR	GARCH@CARR		CARR	GARCH@CARR
Panel B1	S&P500			DJIA		
MAE		4.4E−03	4.4E−03		4.4E−03	4.4E−03
RMSE		6.8E−03	6.8E−03		6.7E−03	6.7E−03
Panel B2:	DAX			HSI		
MAE		5.4E−03	5.3E−03		4.8E−03	4.9E−03
RMSE		7.7E−03	7.6E−03		7.7E−03	7.7E−03

Notes: Panel A presents the MAE and RMSE for return volatility; Panel B presents the MAE and RMSE for high-low price range. The MAE and RMSE are defined as
$MAE = \frac{1}{N}\sum_{t=1}^{N}|Obs_t - F_t(M_i)|$, $RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(Obs_t - F_t(M_i))^2}$,
where Obs_t are data observations and $F_t(M_i)$ are forecasts reported by Model i.

Table 4
DM Statistics.

	EGARCH	GARCH@CARR	EGARCH	GARCH@CARR
Panel A:	S&P500		DJIA	
GARCH	4.609***	9.441***	4.683***	8.599***
EGARCH		5.903***		6.458***
Panel B:	DAX		HSI	
GARCH	4.550***	3.663***	2.099**	8.584***
EGARCH		6.581***		5.967***

Notes: Each table entry represents a t-statistic of the estimate of $\mu_{i,\,j}$ in $\epsilon_{i,t}^2 - \epsilon_{j,t}^2 = \mu_{i,j} + \eta_t$, where $\epsilon_{i,\,t}$ is the forecast error of model i in period t. A positive t-statistic indicates that model j is preferred to model i, whereas a negative t-statistic indicates the opposite. ***, ** and * mean respectively statistical significance at the level of 1%, 5% and 10%.

procedure is used and the last 10-year data observations are used for evaluation. As volatility is unobservable, we use the realized volatility as the true volatility proxy.[4]

Table 3 presents the MAE and RMSE statistics for return volatility in Panel A. Unanimously, both MAE and RMSE show the dominance of EGARCH over GARCH and the dominance of GARCH@CARR over both GARCH and EGARCH. We also calculate the MAE and RMSE for range forecasts and the results are given in Panel B. The results report competitive performance of GARCH@CARR for range forecasting relative to the CARR model.

To see if there is significant difference between two competing models for out-of-sample forecasting, we employ the DM statistic (Diebold and Mariano, 1995). Let the forecasting error of model i be

$$\epsilon_{i,t} = Obs_t - FV_t(M_i), \tag{10}$$

where Obs_t and $FV_t(M_i)$ are respectively the data observation and the forecasted value reported by model i. We test the superiority of model j over model i with a t-test of $\mu_{i,\,j}$ coefficient in

$$\epsilon_{i,t}^2 - \epsilon_{j,t}^2 = \mu_{i,j} + \eta_t, \tag{11}$$

where a positive estimate of $\mu_{i,\,j}$ indicates support for model j.

The DM statistics are reported in Table 4. Consistent with the MAE and RMSE, the DM statistics show clear dominance of the GARCH@CARR model over both the EGARCH and GARCH with great significance[5].

[4] Realized volatility is computed with 5-min high frequency data. The realized volatility is downloaded from https://realized.oxford-man.ox.ac.uk/dataHeber, Gerd, Asger Lunde, Neil Shephard and Kevin Sheppard (2009) "Oxford-Man Institute's realized library", Oxford -Man Institute, University of Oxford.

[5] We also calculate the DM statistic for range forecast. The results show no clear dominance of either GARCH@CARR over CARR or CARR over GARCH@CARR.

4. Conclusions

The RealGARCH model provides an efficient and flexible framework for a joint modeling of return and realized volatility measure. Within the RealGARCH framework, this paper presented a new variant of RealGARCH model by specifying a MEM structure to the measurement equation. The MEM structure simplifies the measurement equation dynamics and thus has fewer parameters to estimate compared with the existed variants of RealGARCH model.

Empirical studies are performed on several stock indices and the results show that our RealGARCH model specification is also promising. Of course, we haven't compared our model specification with the existed RealGARCH variants in this paper, which is of great interest and will be our future research efforts.

Acknowledgement

This research was supported by National Natural Science Foundation of China under Grant No.71401033 and the Program for Young Excellent Talents, UIBE under Grant No. 15YQ08.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.frl.2018.06.015.

References

Alizadeh, S., Brandt, M., Diebold, F., 2002. Range-based estimation of stochastic volatility models. J. Finance 57 (3), 1047–1091.
Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int. Econ. Rev. 39 (4), 885–905.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.
Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76 (6), 1481–1536.
Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. B 64 (2), 253–280.
Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. J. Financial Econom. 2 (1), 1–48.
Chiang, M.H., Chou, R.Y., Wang, L.M., 2016. Outlier detection in the lognormal logarithmic conditional autoregressive range model. Oxford Bull. Econ. Stat. 78 (1), 126–144.
Chou, R., 2005. Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. J. Money Credit Bank. 37 (3), 561–582.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13 (3), 253–263.
Engle, R., 2002. New frontiers of ARCH models. J. Appl. Econom. 17 (5), 425–446.
Engle, R., Lilien, D.M., Robins, R.P., 1987. Estimating time varying risk premia in the term structure: the ARCH-m model. Econometrica 55 (2), 391–407.
Feller, W., 1951. The asymptotic distribution of the range of sums of independent random variables. Ann. Math. Stat. 22 (3), 427–432.
Forsberg, L., Bollerslev, T., 2002. Bridging the gap between the distribution of realized volatility and ARCH modeling: the GARCH-NIG model. J. Appl. Econom. 17 (5), 535–548.
Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: a joint model for returns and realized measures of volatility. J. Appl. Econom. 27 (6), 877–906.
Li, H.Q., Hong, Y.M., 2011. Financial volatility forecasting with range-based autoregressive volatility model. Financ. Res. Lett. 8 (2), 69–76.
Martens, M., van Dijk, D., 2007. Measuring volatility with the realized range. J. Econom. 138, 181–207.
Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53 (1), 61–65.
Sharma, P., Vipul, 2016. Forecasting stock market volatility using realized GARCH model: international evidence. Q. Rev. Econ. Finance 59, 222–230.