Comparison of Several Volatility Forecasting Models

Zheng Hao	A0197899R	Literature review, data analysis, R code for GARCH models
Xiao Chao	A0197872J	Literature review, R code for EW and RW
Zheng Pin	A0197997U	Literature review, R code for MA, model evaluation
Liu Yonghao	A0197875A	Loss function analysis, slides compilation
Shan Changhan	A0197908J	R code for MSE&QLIKE and visualization, report writing

1 Introduction

Volatility forecasting of financial assets has important implications for option pricing, portfolio selection, risk-management and volatility trading strategies. This article compares the forecasting ability of Realized GARCH model with that of the standard GARCH models using only the daily returns, and the other time series models based on the realized measures of volatility.

2 Methology

2.1 Realized measures

We use three realized measures of volatility.

First, subsampled realized variance measure (RVS). The RV measure for day t is

$$RV_t = \sum_{j=1}^{M} (r_{t,j})^2,$$

where $r_{t,j}$ is the jth intraday return of the day and M is the total number of intraday returns for the day. RVS is calculated using 5-min returns with 1-min subsampling.

Second, subsampled realized bipower variance (BVS). The BV measure for day t is

$$BV_t = \frac{\pi}{2} \sum_{j=2}^{M} |r_{t,j}||r_{t,j-1}|,$$

and BVS is calculated using 5-min returns with 1-min subsampling.

Third, realized kernel estimator (RK):

$RK_t = \sum_{h=-H}^{H} \kappa\left(\frac{h}{H+1}\right)\gamma_h$, where $\gamma_h = \sum_{j=|h|+1}^{M} r_{t,j} r_{t,j-|h|}$ and $\kappa(x)$ is a kernel weight function.

2.2 Forecasting models

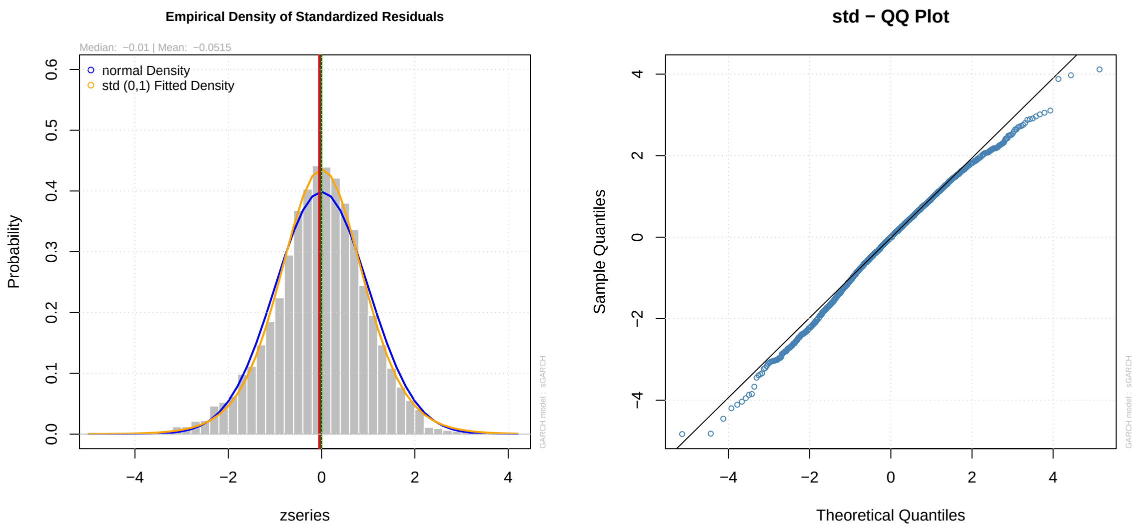
We use an AR(1) specification for modeling the conditional mean of the GARCH models:

AR(1): $r_t = c + \phi_1 r_{t-1} + \varepsilon_t$.

The conditional variance equations for the various GARCH models with Student's t-distribution are specified as

$$\begin{aligned}
& \varepsilon_t = \sigma_t z_t, \quad z_t \sim i.i.d.t(d), \\
\text{GARCH(1,1):}\ & \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \\
\text{EGARCH(1,1):}\ & \log\sigma_t^2 = \omega + \beta\log\sigma_{t-1}^2 + \tau_1 z_{t-1} + \tau_2(|z_{t-1}| - E|z_{t-1}|), \\
\text{Realized GARCH(1,1):}\ & \log\sigma_t^2 = \omega + \beta\log\sigma_{t-1}^2 + \gamma\log x_{t-1}, \\
& \log x_t = \xi + \varphi\log\sigma_t^2 + \delta(z_t) + u_t, \\
& \delta(z_t) = \delta_1 z_t + \delta_2(z_t^2 - 1), \\
& \text{where } u_t \sim i.i.d.\, N(0, \sigma_u^2) \text{ and } x_t \text{ is realized volatility.}
\end{aligned}$$

We use a rolling window of the most recent 2000 daily observations for the estimation of GARCH models. Next, we describe the forecasting models based on the realized



(a) Empirical Density of Standardized Residuals (b) QQ Plot

Figure 1: Student's t-distribution

measures. As the realized measures estimate the variance for the open-to-close period, we scale them to obtain the estimate of close-to-close variance. The close-to-close variance for day t is estimated as $\sigma_t^2 = \eta x_t$, where η is the scaling factor and $x_t = RVS_t, BVS_t$ or RK_t. The scaling factor η is calculated as

$$\eta = \frac{T^{-1}\sum_{t=1}^{T}(r_t - \mu_{cc})^2}{T^{-1}\sum_{t=1}^{T}(r_{oc,t} - \mu_{oc})^2},$$

where T is the total number of days in the sample period, $r_{oc,t}$ is the open-to-close log return for day t, $\mu_{cc} = T^{-1}\sum_{t=1}^{T} r_t$, and $\mu_{cc} = T^{-1}\sum_{t=1}^{T} r_{cc,t}$. The random walk model (RW), the moving average model (MA), and the exponentially weighted moving average model (EW), are specified as

$$\text{RW: } \hat{\sigma}^2_{t+1} = \sigma^2_t, \text{ where } \sigma^2_t \text{ is realized volatility,}$$
$$\text{EW: } \hat{\sigma}^2_{t+1} = \lambda\sigma^2_t + (1-\lambda)\hat{\sigma}^2_t, \text{ where } \lambda = 0.4,$$
$$\text{MA: } \hat{\sigma}^2_{t+1} = p^{-1}\sum_{i=1}^{p}\sigma^2_{t+1-i}, \text{ where } p = 5.$$

2.3 Forecast evaluation

We select two approaches for evaluating the accuracy of volatility forecasts, by implementing mean squared error (MSE) and quasi-likelihood (QLIKE) loss functions. MSE is a loss criterion that penalizes the forecasting errors in a symmetrical manner, while QLIKE is an asymmetric loss function that penalizes the under-prediction more heavily than the over-prediction, which is more suitable for the applications like risk management and VaR forecasting.

$$\text{MSE} = E(L_{1,k,t}), \quad \text{where } L_{1,k,t} = (\sigma^2_t - \hat{\sigma}^2_t)^2.$$
$$\text{QLIKE} = E(L_{2,k,t}), \quad \text{where } L_{2,k,t} = (\log(\hat{\sigma}^2_t) + \sigma^2_t\hat{\sigma}^{-2}_t).$$

Here, $L_{1,k,t}$ and $L_{2,k,t}$ are the losses for the forecasting model k, with the MSE and QLIKE loss functions, respectively.

We use the Diebold-Mariano test for comparing the predictive accuracy of forecasting models i and j. The test statistic is based on the loss differencial $d_{w,t} = L_{w,i,t} - L_{w,j,t}$. The null hypothesis of equal predictive accuracy is $H_0 : E(d_{w,t}) = 0$, the Diebold-Mariano test statistic is $DM = \bar{d}/\sqrt{Var(\bar{d})}$, where $\bar{d} = N^{-1}\sum_{j=1}^{N} d_{w,t+j}$.

3 Empirical Results

We use the daily return and realized variance for 12 stock indices across the world (Table 1). The sample period extends from 1 January 2000 to 20 September 2019. For each index, we generate N variance forecasts on a rolling basis, where $N = T - 2000$ and T is the total number of daily observations.

#	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}
Ticker	FTSE	N225	GDAXI	DJI	FCHI	KS11	AEX	SSMI	IBEX	NSEI	MXX	STOXX50E
Index	FTSE 100	Nikkei 225	DAX	DJIA	CAC 40	KOSPI	AEX	SMI	IBEX 35	S&P CNX Nifty	IPC	Euro STOXX 50
Country	United Kingdom	Japan	Germany	United States	Canada	South Korea	Netherlands	Switzerland	Spain	India	Mexico	Eurozone

Table 1: 12 stock indices

Table 2 provides models' performance based on the MSE and QLIKE criterion for

12 indexes, lower value indicating better performance. Under both criteria, Realized GARCH performs best in GARCH models, while original GARCH performs worst. As for the other models, exponentially weighted moving average model performs best and the random walk model performs worst.

Table 3 reports the statistics for the Diebold-Mariano tests, where the EGARCH model

	Loss function:MSE (10^{-8})												Loss function:QLIKE											
Model	I$_1$	I$_2$	I$_3$	I$_4$	I$_5$	I$_6$	I$_7$	I$_8$	I$_9$	I$_{10}$	I$_{11}$	I$_{12}$	I$_1$	I$_2$	I$_3$	I$_4$	I$_5$	I$_6$	I$_7$	I$_8$	I$_9$	I$_{10}$	I$_{11}$	I$_{12}$
GARCH	8.0	33.5	7.4	5.0	6.3	5.9	5.3	6.3	12.8	21.6	3.7	13.8	-8.39	-7.94	-8.00	-8.64	-7.91	-8.39	-8.14	-8.42	-7.64	-8.14	-8.45	-7.94
EGARCH	7.2	32.7	8.0	3.9	6.3	7.0	5.5	5.4	11.0	17.9	3.1	14.2	-8.39	-7.92	-8.00	-8.66	-7.90	-8.34	-8.12	-8.42	-7.64	-8.14	-8.46	-7.93
RGARCH_rvs	7.3	33.0	6.6	4.5	5.8	5.3	5.7	5.7	10.3	22.0	3.6	13.2	-8.43	-7.98	-8.04	-8.69	-7.95	-8.43	-8.17	-8.47	-7.68	-8.18	-8.46	-7.95
RGARCH_bvs	7.2	31.7	7.0	4.4	6.0	5.4	5.9	6.1	10.2	21.2	3.3	13.5	-8.43	-7.98	-8.04	-8.70	-7.95	-8.43	-8.17	-8.46	-7.68	-8.18	-8.47	-7.94
RGARCH_rk	7.4	34.3	6.9	4.8	5.9	6.0	5.9	5.7	10.8	22.3	3.4	13.0	-8.43	-7.96	-8.04	-8.68	-7.95	-8.42	-8.16	-8.46	-7.68	-8.17	-8.48	-7.97
RW_rvs	11.1	30.7	8.0	7.0	7.7	6.8	6.3	7.0	12.1	35.9	5.4	19.2	-8.32	-7.91	-7.98	-8.59	-7.88	-8.37	-8.12	-8.46	-7.63	-8.10	-8.33	-5.45
RW_bvs	8.7	30.2	6.6	4.5	6.8	6.8	6.4	6.7	10.8	18.7	3.9	14.1	-8.21	-7.84	-7.95	-8.50	-7.83	-8.33	-8.08	-8.44	-7.61	-8.05	-8.23	-0.37
RW_rk	12.0	47.1	7.7	5.3	7.1	6.3	7.3	8.9	16.0	34.2	5.1	20.3	-8.29	-7.60	-7.90	-8.48	-7.70	-8.07	-7.92	-8.39	-7.52	-8.03	-8.26	-3.82
MA_rvs	8.0	34.0	7.9	4.8	6.7	6.9	5.8	5.9	11.8	26.0	3.9	14.8	-8.41	-7.99	-8.04	-8.68	-7.94	-8.41	-8.16	-8.46	-7.66	-8.16	-8.45	-7.98
MA_bvs	7.7	34.0	7.4	4.1	6.4	6.9	5.8	5.8	11.4	21.7	3.8	13.9	-8.31	-7.94	-8.02	-8.59	-7.93	-8.39	-8.15	-8.45	-7.65	-8.15	-8.37	-7.94
MA_rk	8.1	37.5	8.1	4.3	6.5	6.0	5.9	6.3	12.7	25.1	3.7	14.5	-8.41	-7.97	-8.03	-8.68	-7.92	-8.38	-8.14	-8.45	-7.66	-8.15	-8.47	-7.96
EW_rvs	7.7	31.9	6.7	4.5	6.0	5.9	5.3	5.6	10.8	24.8	3.7	14.1	-8.42	-8.01	-8.05	-8.70	-7.95	-8.42	-8.17	-8.47	-7.68	-8.18	-8.47	-7.99
EW_bvs	7.5	31.8	6.2	3.8	5.8	5.9	5.4	5.6	10.5	19.8	3.7	13.1	-8.33	-7.96	-8.03	-8.62	-7.94	-8.40	-8.16	-8.46	-7.66	-8.16	-8.39	-7.96
EW_rk	7.7	34.2	6.6	4.0	5.7	5.1	5.4	6.1	11.8	24.4	3.5	13.7	-8.42	-7.98	-8.04	-8.70	-7.94	-8.39	-8.16	-8.46	-7.67	-8.17	-8.48	-7.97

Table 2: MSE&QLIKE criteria

is used as the benchmark. A positive value of the statistic indicates that the benchmark model performs better than the competing model. All the statistics that are significant at the 5% level are marked with an asterisk. It is obvious that RGARCH models performs best while original GARCH model performs worst under QLIKE criterion.

	Loss function: MSE (benchmark: EGARCH)				Loss function:QLIKE (benchmark: EGARCH)			
Index	GARCH	RGARCH_rvs	RGARCH_bvs	RGARCH_rk	GARCH	RGARCH_rvs	RGARCH_bvs	RGARCH_rk
FTSE	0.6	1.3	1.3	1.2	1.1	0.9	0.9	0.9
N225	-0.5	0.5	1.5	-1.5	1.2	0.4	0.5	0.3
GDAXI	1.3	1.4	1.4	1.4	1.4	-8.7*	-8.4*	-8.2*
DJI	-1.2	-1.2	-0.6	-1.2	5.9*	-5.4*	-7.4*	-4.8*
FCHI	1.0	1.0	-0.3	1.5	0.6	-7.4*	-7.0*	-6.1*
KS11	1.1	1.1	1.1	1.0	-2.6	-7.3*	-6.5*	-6.5*
AEX	1.0	-1.0	-1.5	-2.1	0.8	-6.7*	-6.3*	-5.6*
SSMI	0.0	1.0	0.8	0.9	1.1	0.7	0.8	0.8
IBEX	-1.0	1.0	1.0	1.0	1.4	-5.3*	-5.2*	-5.9*
NSEI	-1.1	-1.3	-1.4	-1.3	0.9	-4.6*	-4.2*	-4.0*
MXX	-1.0	-1.4	-1.3	-1.4	-0.3	1.1	-0.5	-1.4
STOXX50E	1.9	1.6	1.7	1.3	1.2	0.6	0.2	0.8

Table 3: DM test statistics

Table 4 provides a comparison of the Bayesian Information Criterion for GARCH models, indicating that Realized GARCH models rank lowest.

Model	FTSE	N225	GDAXI	DJI	FCHI	KS11	AEX	SSMI	IBEX	NSEI	MXX	STOXX50E	Mean
GARCH	2	2	2	2	2	2	2	2	2	2	2	2	2
EGARCH	1	1	1	1	1	1	1	1	1	1	1	1	1
RGARCH_rvs	4	3	4	4	4	3	3	3	4	3	4	4	4
RGARCH_bvs	3	4	3	3	3	4	4	4	3	4	3	3	3
RGARCH_rk	5	5	5	5	5	5	5	5	5	5	5	5	5

Table 4: BIC rank

4 Conclusion

The Realized GARCH model provides a unique framework for the joint modeling of conditional variance and realized measures of volatility. This article attempts to bridge the gap by comparing the predictive ability of the Realized GARCH with that of the GARCH and EGARCH models based on daily returns, and the EW, RW and MA models based on realized measures of volatility.

With extra information of realized volatility, Reazlied GARCH model provides better forecast than GARCH and EGARCH models, while the EW model performs best. We propose two reasons for the superior forecasting performance of the EW model. First, even the most basic Realized GARCH model requires an estimation of nine parameters, whereas the EW model requires the estimation of a single parameter. The estimation of a large number of parameters can often lead to considerable estimation errors that may make the model ill-suited for forecasting applications, and also leads to high BIC. Second, the EW model is quicker to respond to the changes in the variance process as compared to the GARCH estimate.

Reference

Prateek Sharma et al. Forecasting stock market volatility using realized garchmodel: International evidence. The Quarterly Review of Economics and Finance, 59:222–230, 2016.

Peter Reinhard Hansen, Zhuo Huang, Howard Howan Shek. Realized garch: a joint model for returns and realized measures of volatility. Journal of Applied Econometrics, 27(6):877–906, 2012.

Spyros Makridakis, A Andersen, Robert Carbone, Robert Fildes, Michele Hibon, Rudolf Lewandowski, Joseph Newton, Emanuel Parzen, and Robert Winkler. Theaccuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of forecasting, 1(2):111–153, 1982.