Journal of Commerce & Accounting Research
9 (3) 2020, 41-50
http://publishingindia.com/jcar/

Submitted: 16 April, 2020
Accepted: 29 May, 2020

RETURN INNOVATION DISTRIBUTION IN BEST-FIT GARCH MODELS FOR HIGH-FREQUENCY DATA

Yelamanchili Rama Krishna*

Abstract *This paper offers empirical evidence to extant literature on choice of return innovation distribution in GARCH family models. Statistical software packages have Gaussian distribution as default option and offer other distribution options to choose as per properties of data series. Most common choices are Gaussian, t-distribution, and GED. The choice of appropriate return innovation distribution for best-fit GARCH model is still an open inquiry. We posit that choice of return innovation distribution is influenced by frequency of data, time horizon, and symmetric or asymmetric models. This paper estimates four symmetric and asymmetric GARCH models with three return innovation distributions using high frequency and long horizon stock market return series. Daily log returns of NIFTY50 index over the period January 1, 1996, to December 31, 2019, a total of 5,971 observations are analyzed. Results suggest significant evidence in support of volatility clustering, fat-tailedness, mean reversion, volatility persistence, leverage effect, and long memory in return series. Results recommend student's t-distribution as ideal return innovation process for both symmetric and asymmetric GARCH models. The basic models have serial correlations, but no remaining GARCH effects. After multiple trials, it is observed that the ARMA (1,1)-GARCH (1,1) models have better log likelihood values, minimum information criterion values, no serial correlations, and no remaining GARCH effects. Finally, among the three asymmetric models, ARMA-EGARCH model has superior information criterion values. We recommend ARMA-GARCH (1,1) as best-fit symmetric model and ARMA-EGARCH (1,1) as best-fit asymmetric model for high-frequency data spread over long horizon.*

Keywords: *Volatility Modeling, Distribution Density, GARCH Models, High-Frequency Data, Long Horizon*

INTRODUCTION

Over the past few decades, the autoregressive conditional heteroscedasticity (ARCH) models have become the most popular models of conditional variance of financial time series. The ARCH framework was firstly proposed by Engle (1982) and was further extended to the GARCH model (Bollerslev, 1986), EGARCH (Nelson, 1991), GJR-GARCH (Glosten et al., 1993), APARCH (Ding et al., 1993), and many other models popularly known as GARCH family models. The popularity of GARCH models is particularly due to their ability to capture characteristics of financial time series such as volatility clustering, fat tail behavior, long memory, mean reversion behavior, time varying heteroscedasticity, and leverage effect. Because of high dynamic nature of financial time series data, researchers are continuously adding more and more models to existing GARCH family models. Return innovation distribution plays an important role in estimation process of GARCH family models. While estimating GARCH models, research applies different innovation distributions depending on the frequency of data, time horizon, and type of estimation model. Statistical software packages have Gaussian distribution as a default option and offer other distribution options to choose as per properties of data series. Most common choices are Gaussian, t-distribution, and GED. The choice of appropriate return innovation distribution for best-fit GARCH models is still an open inquiry. This paper offers empirical evidence to extant literature on choice of return innovation distribution in GARCH family models. This paper estimates symmetric and asymmetric GARCH models with three return innovation distributions using high frequency and long horizon stock market daily return series and suggests appropriate return innovation distribution for best-fit GARCH symmetric and asymmetric models.

The structure of this paper is as follows. Section 2 discusses extant literature. Section 3 explains data and methodology. Section 4 presents estimation results. Finally, Section 5 draws conclusions from empirical tests.

LITERATURE REVIEW

Different return innovation distributions are applied by researchers to model financial assets returns volatility.

* Associate Professor, Department of Finance, ICFAI Business School (IBS), ICFAI University, Hyderabad, Telangana, India. Email: yrk@ibsindia.org

Gaussian distribution is used to measure volatility in symmetric and asymmetric GARCH models to measure various characteristics of financial asset returns (Dedi & Yavas, 2017; Ghangare, 2016; John & Amudha, 2019; Rizwan et al., 2018; Sharma and Vipul, 2016). On the other side, student's t-distribution is used to model volatility of stock markets and to model leverage effect and forecasting ability of GARCH models (Aziz & Iqbal, 2016; Olbrys & Majewska, 2017; and Roni Bhowmik et al., 2017). In other context, GED distribution is used with symmetric and asymmetric GARCH models (Vasudevan & Vetrivel, 2016). In extant literature, there are only few studies that compared volatility estimates resulted from two different distributions. For example, Marie-Eliette and Bing (2018) compared results from Gaussian distribution and t-distribution and reported that t-distribution is appropriate for higher Kurtosis GARCH models. Some papers applied three different distributions to model stock market volatility. Sasikanta Tripathy and Abdul Rahman (2013) reported t-distribution as appropriate to model stock market returns volatility. Similarly, Hemanth and Basavaraj (2016) applied three distributions and reported that GARCH (1,1) model with GED distribution outperformed all other models.

Notwithstanding many distinctions, the use of Gaussian distribution in GARCH models is most common (Hansen & Lunde, 2006). However, substantial evidence suggests that financial time series is rarely Gaussian but typically leptokurtic and exhibits heavy-tail behavior. Theoretically, GARCH model can accommodate for fat-tail through its specification (Bollerslev & Wooldridge, 1992). In practice, however, there is still excess kurtosis left in the standardized residuals in most cases. To solve this problem, a common solution is to employ a fat-tailed distribution such as the student's t-distribution or Generalized Error Distribution (GED) (Chkili, Aloui & Nguyen, 2012). In this context, we posit that choice of return innovation distribution depends on frequency of data, time horizon, and volatility model.

DATA AND METHODOLOGY

The data we analyze in this paper is the National Stock Exchange (NSE), India's broad market index, - NIFTY50 (Hereafter, NIFTY) daily closing price index returns. In year 1996, the NSE introduced NIFTY as a broad market index and currently it is computed based on free-float methodology. NIFTY contains most active and highly liquid 50 stocks listed in NSE and represents about 66.8% of the free-float market capitalization of the stocks listed on NSE. NIFTY is used for different purposes such as benchmarking fund portfolios, index-based derivatives, and index funds. The data is from www.nseindia.com. Over the study period, there are altogether 5,971 daily observations from Jan. 1, 1996, to Dec. 31, 2019. During this period, NIFTY moved from meager 913.11 points to 12247.10 points. We calculate absolute return [$r_t = 100*(P_t - P_{t-1})/P_{t-1}$], continuously compounding return or log return [$r_t = 100*\ln(P_t/P_{t-1})$], and squared return ($r^2_t$), where, P_t is closing value of the index at time t.

Symmetric and Asymmetric Models

GARCH Model (Bollerslev, 1986)

$$\sigma^2_t = \alpha_0 + \alpha_1\varepsilon^2_{t-1} + \beta_1\sigma^2_{t-1} \tag{1}$$

The standard GARCH model assumes that positive and negative error terms have symmetric effect on the volatility. This is because in the GARCH model only squared residuals ε^2_{t-1} enter the conditional variance equation, the signs of the residuals or shocks have no effect on conditional volatility. So, GARCH is unable to express the leverage effect. The drawbacks of the GARCH model are that it cannot explain the negative correlation between the fluctuations in stock returns. It assumes that the conditional variance is a function of lagged squared residuals. So, the symbol does not affect the residual volatility, that is, positive and negative changes are symmetric to conditional variance. Next, GARCH model assumes all coefficients are greater than zero, which also makes the model hard to apply.

In order to measure return volatility, symmetric basic GARCH model was extended. Among the many asymmetric models, we choose conventional asymmetric models to model asymmetric properties of returns volatility. In this paper, we apply three asymmetric GARCH models, that is, the EGARCH model, the GJR-GARCH model, and the APARCH models.

Exponential GARCH (EGARCH) Model (Nelson, 1991)

$$\log(\sigma^2_t) = \alpha_0 + \alpha_1\varepsilon_{t-1} + \lambda_1(|\varepsilon_{t-1}| - E|\varepsilon_{t-1}|) + \beta_1\log(\sigma^2_{t-1}) \tag{2}$$

GJR-GARCH Model (Glosten, Jaganathan and Runkle, 1993)

$$\sigma^2_t = \alpha_0 + \alpha_1\varepsilon^2_{t-1} + \lambda_1 d_{t-1}\varepsilon^2_{t-1} + \beta_1\sigma^2_{t-1} \tag{3}$$

APARCH model (Ding et al., 1993)

$$\sigma^\delta_t = \alpha_0 + \alpha_1(|\varepsilon_{t-1}| - \lambda\varepsilon_{t-1})^\delta + \beta_1\sigma^\delta_{t-1} \tag{4}$$

where, parameter δ (δ < 0) plays the role of a Box–Cox transformation of the conditional standard deviation σ_t, while λ reflects the leverage effect.

Different return innovation distributions are used in GARCH models. Most common and conventional return innovation distributions are Gaussian, t-distribution, and GED. The application of simple Gaussian distribution is most common. However, due to fat-tail behavior in financial asset return series, other return innovation distributions are considered. The t-distribution density curves are symmetric and bell-shaped such as the normal distribution and have their peak at 0. However, the spread is more than that of the standard normal distribution. The degrees of freedom is larger, the t-density is closer to normal density. Next, Nelson (1991) proposed to use the GED to capture the fat tails usually observed in the distribution of financial time series. The v (degrees of freedom) is a positive parameter governing the thickness of the tail behavior of the distribution. When v = 2 the probability density distribution (pdf) reduces to the standard normal pdf; when $v < 2$, the density has thicker tails than the normal density; when $v > 2$, the density has thinner tails than the normal density. When the tail thickness parameter $v = 1$, the pdf of GED reduces to the pdf of double exponential distribution.

RESULTS

In this section, we present results of this paper. First we describe few stylized facts of data series through plots. We then present preliminary analysis. Finally, we discuss results of symmetric and asymmetric GARCH models under different return innovation distributions. Fig. 1 plots the daily close prices, daily relative returns, daily log returns, and squared returns of NIFTY index. It also plots auto correlation function (ACF) and partial auto correlation function (PACF) plots of relative returns, log returns, and squared returns. There is no clear visible pattern of behavior in the relative returns and log returns series. In close price plot, we can see the movement as an upward trend, which shows the volatility clustering – low values of volatility followed by low values and high values of volatility followed by high values. This behavior is confirmed in ACF plots of relative and log returns series. The log returns show no evidence of serial correlation, but the squared returns are positively auto correlated. Also, the decay rates of the sample auto correlations of squared returns appear much slower, suggesting possible long memory behavior.

Table 1 gives some standard summary statistics along with results of unit root tests for stationarity and Jarque-Bera test for normality. The highest single day log return is 16.33 percent and lowest is -13.05 percent. The average daily log return is close to zero that is 0.04 percent with a standard deviation of 1.48. The distribution of daily log return series is clearly non-normal with negative Skewness (-0.16) and pronounced excess kurtosis (10.83). The excess kurtosis value shows the daily log return series have the fat-tail characteristic. The Jarque–Bera for the log return series is 15,280.83 and statistically significant at one percent level. The higher Jarque–Bera statistics indicates the non-normality of the return series. The Augmented Dickey Fuller test (ADF) and Dickey Fuller Generalized Least Squares (DF-GLS) test results indicate that the log return series is stationary at level. With Ljung-Box Q-statistics, we check for serial correlations in return series at 5, 10, and 20 lags representing a week, fortnight, and monthly trading days, respectively. The Q-statistics for all the three lags is statistically significant indicating serial correlations. This result suggests GARCH models are appropriate to model conditional variance of log return series.

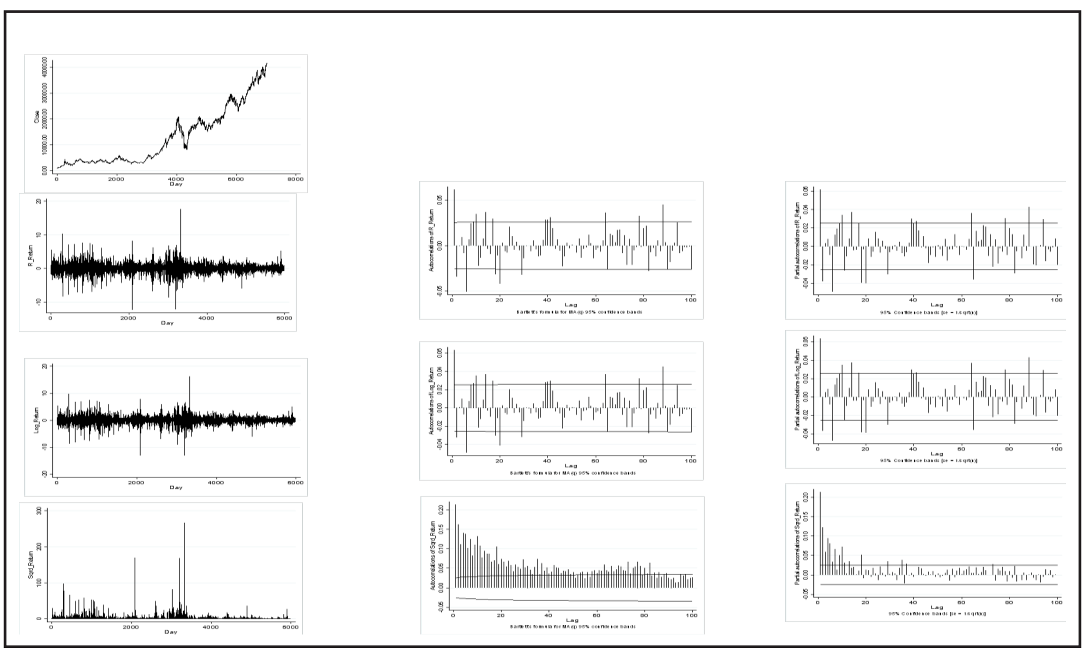

Fig. 1: Time Series, ACF, and PACF Plots of Absolute Value, Relative Return, Log Return, and Squared Return

Table 1: Summary Statistics of Daily Close Price, Relative Return, Log Return, and Squared Return

Descriptive	Daily Close	Relative Return	Log Return	Squared Return
Minimum	788.15	-12.23774	-13.05386	0.00
Maximum	12271.8	17.74407	16.33432	266.8098
Mean	4441.296	0.0544938	0.0434659	2.205982
Std. Dev	3298.882	1.484572	1.484743	6.905526
Skewness	0.6407455	0.0575066	-0.1645892	17.15456
Kurtosis	2.242334	11.30887	10.83019	501.8806
JB (p-value)	551.39 (0.00)	17179.19 (0.00)	15280.83 (0.00)	62212599 (0.00)
Q (5) (Chi- Prob)		30.55 (0.00)	31.51 (0.00)	746.53 (0.00)
Q (10) (Chi- Prob)		61.27 (0.00)	62.12 (0.00)	1127.43 (0.00)
Q (20) (Chi- Prob)		96.91 (0.00)	97.08 (0.00)	1563.01 (0.00)
ADF		-72.63 (0.00)	-72.48 (0.00)	-62.12 (0.00)
DF-GLS		-48.92 (0.00)	-48.67 (0.00)	-41.43 (0.00)
Observations	5,971	5,971	5,971	5,971

We first estimate basic symmetric model (GARCH 1,1) and asymmetric model (EGARCH (1,1), GJR-GARCH (1,1) and APARCH (1,1)) with different return innovation distributions (Gaussian, t, and GED) and check model efficiency with two information criterions (AIC and BIC) and log likelihood value. Furthermore, using Ljung-Box Q-statistics for 5, 10, and 20 lags, we diagnose standardized residuals and squared standardized residuals for no remaining serial correlations and no remaining GARCH effects. When we analyze basic models with diagnostic test results, we notice remaining serial correlations in standardized residuals of log returns. In this context, we experiment with multiple models such as GARCH-M, AR (1), MA (1), and ARMA (1,1) and few other specifications of GARCH (*p, q*). Finally, we find that ARMA (1,1)–GARCH (1,1) models are of white noise process with no remaining serial correlations and no remaining GARCH effects. We present results of model specifications in two tables. Table 2 presents estimated parameters of basic models and Table 3 presents estimated parameters of enhanced ARMA (1,1)-GARCH (1,1) models.

Model Estimation Results

We first check whether estimated parameters of the models meet their specifications and assumptions. The GARCH, GJR-GARCH, and APARCH models assume that $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $(\alpha_1 + \beta_1) < 1$. However, the EGARCH model uses logged conditional variance to relax the positiveness constraint of model coefficients. Results present in Tables 2 and 3 clearly evince that model coefficients for all the models are in accordance with assumptions and are statistically significant at one percent significance level. We then look at log likelihood values and information criterion values of each model under three different return innovation distributions. Models with student's t-distribution have larger log likelihood and minimum Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC) than other two return innovation distributions. This specifies that the models under t-distribution are better fitted.

To understand volatility clustering, we look at β_1 and find that in all models this coefficient is close to 0.9. Given the value of high β_1, it is obvious that large values of previous day's conditional variance will be followed by large values of σ^2_t, and small values of σ^2_{t-1} will be followed by small values of σ^2_t. This indicates that the daily log return series of NIFTY has well known behavior of volatility clustering. In addition, NIFTY daily log return series also exhibits fat-tailed behavior, which is depicted in the shape of the distribution. By applying the estimated parameters in GARCH model, we can estimate the unconditional variance of ε_t as α_0 / (1- α_1- β_1). The estimated unconditional variance in our basic GARCH model is 0.0831 / (1 – 0.0941 – 0.8984) = 11.08 and unconditional volatility is 3.33. And the estimated unconditional variance in our ARMA-GARCH model is 0.0822 / (1 – 0.0959 – 0.8963) = 10.54 and unconditional volatility is 3.24. We see volatility clustering, fat-tailed behavior, and excessive volatility from time to time in NIFTY returns. However, the excessive volatility will eventually settle down to a long-run level. This process is defined as mean reversion behavior. With the help of GARCH parameters, we measure the magnitude of mean reversion. The magnitude of $\alpha_1 + \beta_1$ controls the speed of mean reversion. The half-life of volatility shocks, defined as ln (0.5) / ln($\alpha_1 + \beta_1$), measures the average time it takes for $| \varepsilon^2 - \sigma^2|$ to decrease by one half. The closer $\alpha_1 + \beta_1$ is to one the longer is the half-life of a volatility shock. For our basic GARCH model, the half-life of volatility is -0.6931 / -0.0075 = 92 days. The half-life volatility in ARMA-GARCH model is 88 days.

Table 2: Parameter Estimates of Symmetric and Asymmetric GARCH (1,1) Models

Parameter	GARCH			GJR-GARCH		
	N	**t**	**GED**	**N**	**t**	**GED**
α_0	0.0814***	0.0831***	0.0842***	0.0577***	0.0671***	0.0688***
	(0.0134)	(0.0134)	(0.0130)	(0.0139)	(0.0134)	(0.0131)
α_1	0.1050***	0.0941***	0.0968***	0.1505***	0.1507***	0.1488***
	(0.0050)	(0.0082)	(0.0079)	(0.0074)	(0.0125)	(0.0117)
β_1	0.8907***	0.8984***	0.8967***	0.8867***	0.8888***	0.8888***
	(0.0047)	(0.0080)	(0.0076)	(0.0050)	(0.0083)	(0.0080)
Λ	-	-	-	-0.0882***	-0.1016***	-0.0953***
				(0.0078)	(0.0132)	(0.0124)
N	-	6.87	1.38	-	7.02	1.39
		(0.5107)	(0.0265)		(0.5199)	(0.0260)
Log(L)	-9893.45	-9739.23	-9764.59	-9860.85	-9707.25	-9736.60
Diagnostic Checking						
Standardized Residuals						
Q (5) p value in parentheses	57.36 (0.00)	56.59 (0.00)	56.84 (0.00)	57.48 (0.00)	56.25 (0.00)	56.45 (0.00)
Q (10) p value in parentheses	75.72 (0.00)	75.18 (0.00)	75.33 (0.00)	76.68 (0.00)	76.00 (0.00)	75.94 (0.00)
Q (20) p value in parentheses	87.75 (0.00)	87.25 (0.00)	87.36 (0.00)	88.59 (0.00)	88.08 (0.00)	87.96 (0.00)
Sqrd Standardized Residuals						
Q (5) p value in parentheses	3.035(0.65)	6.42(0.27)	5.42 (0.37)	0.92 (0.97)	1.35 (0.93)	1.33 (0.95)
Q (10) p value in parentheses	7.95 (0.63)	10.36 (0.41)	9.57 (0.48)	5.56 (0.85)	5.82 (0.83)	5.61 (0.85)
Q (20) p value in parentheses	20.90 (0.40)	22.05 (0.34)	21.64 (0.36)	17.88 (0.59)	17.35 (0.63)	17.40 (0.63)
Information Criteria						
AIC	19794.89 (df 4)	19488.46 (df 5)	19539.17 (df 5)	19731.70 (df 5)	19426.49 (df 6)	19485.20 (df 6)
BIC	19821.67 (df 4)	19521.93 (df 5)	19572.64 (df 5)	19765.17 (df 5)	19466.66 (df 6)	19525.37 (df 6)
OBS	5,971	5,971	5,971	5,971	5,971	5,971

Parameter	EGARCH			APARCH		
	N	**t**	**GED**	**N**	**t**	**GED**
α_0	0.0509***	0.0629***	0.0644***	0.0544***	0.0645***	0.0663***
	(0.0135)	(0.0133)	(0.0130)	(0.014)	(0.0134)	(0.0131)
α_1	0.2062***	0.1957***	0.1991***	0.1096***	0.1041***	0.1055***
	(0.0078)	(0.0145)	(0.0135)	(0.0048)	(0.0091)	(0.0082)
β_1	0.9802***	0.9792***	0.9798***	0.8950***	0.8975***	0.8969***
	(0.0018)	(0.0033)	(0.0031)	(0.0046)	(0.0079)	(0.0074)
λ	-0.0678***	-0.0746***	-0.0716***	-0.2797***	-0.3514***	-0.3222***
	(0.0049)	(0.0083)	(0.0079)	(0.0270)	(0.0484)	(0.0452)
ν	-	7.05	1.4	-	7.02	1.39
		(0.5201)	(0.0259)		(0.5194)	(0.0261)
δ				1.4418***	1.3215***	1.3674***
				(0.0705)	(0.1406)	(0.1228)

Log(L)	-9851.79	-9697.77	-9728.11	-9852.67	-9699.27	-9729.45
Diagnostic Checking						
Standardized Residuals						
Q (5) p value in parentheses	55.93 (0.00)	54.50 (0.00)	54.67 (0.00)	56.75 (0.00)	55.21 (0.00)	55.49 (0.00)
Q (10) p value in parentheses	75.35 (0.00)	74.22 (0.00)	74.24 (0.00)	76.05 (0.00)	74.97 (0.00)	75.02 (0.00)
Q (20) p value in parentheses	88.85 (0.00)	87.92 (0.00)	87.89 (0.00)	88.63 (0.00)	87.99 (0.00)	87.89 (0.00)
Sqrd Standardized Residuals						
Q (5) p value in parentheses	3.43 (0.63)	5.14 (0.40)	4.53 (0.48)	2.04 (0.84)	4.26 (0.51)	3.35 (0.64)
Q (10) p value in parentheses	8.65 (0.56)	10.91 (0.36)	10.02 (0.44)	6.82 (0.74)	9.66 (0.47)	8.40 (0.59)
Q (20) p value in parentheses	33.01 (0.03)	34.93 (0.02)	34.33 (0.02)	24.73 (0.21)	29.24 (0.08)	27.27 (0.13)
Information Criteria						
AIC	19713.57 (df 5)	19407.54 (df 6)	19468.22 (df 6)	19717.35 (df 6)	19412.55 (df 7)	19472.90 (df 7)
BIC	19747.05 (df 5)	19447.71 (df 6)	19508.38 (df 6)	19757.52 (df 6)	19459.41 (df 7)	19519.76 (df 7)
OBS	5,971	5,971	5,971	5,971	5,971	5,971

Table 3: Parameter Estimates of Symmetric and Asymmetric ARMA-GARCH (1,1) Models

Parameter	ARMA-GARCH			ARMA-GJR-GARCH		
	N	**t**	**GED**	**N**	**t**	**GED**
ar_1	-.1475	-0.2263	-0.1475	-0.0989	-0.1939	-0.1008
	(0.1525)	(0.1513)	(0.1622)	(0.1431)	(0.1440)	(0.1544)
ma_1	.2280	0.3061	0.224	0.1868	0.2793	0.1827
	(0.1511)	(0.1479)	(0.1601)	(0.1421)	(0.1409)	(0.1528)
α_0	0.0824***	0.0822***	0.0822***	0.0543***	0.0622***	0.0629***
	(0.0145)	(0.0143)	(0.0139)	(0.0151)	(0.0143)	(0.0140)
α_1	0.10607***	0.0959***	0.0983***	0.1566***	0.1584***	0.1558***
	(0.0051)	(0.0083)	(0.0008)	(0.0078)	(0.0132)	(0.0123)
β_1	0.8894***	0.8963***	0.8949***	0.8847***	0.8858***	0.8862***
	(0.0047)	(0.0082)	(0.0077)	(0.0050)	(0.0085)	(0.0081)
λ	-	-	-	-0.0963***	-0.1109***	-0.1038***
				(0.0084)	(0.0141)	(0.0131)
ν	-	6.85	1.38	-	7.02	1.39
		(0.5065)	(0.0265)		(0.5171)	(0.0261)
Log(L)	-9876.45	-9720.09	-9747.11	-9841.29	-9685.72	-9717.04
Diagnostic Checking						
Standardized Residuals						
Q (5) p value in parentheses	8.58 (0.13)	8.48 (0.13)	9.38 (0.09)	7.63 (0.18)	7.57 (0.18)	8.55 (0.13)
Q (10) p value in parentheses	26.78 (0.00)	26.88 (0.00)	27.70 (0.00)	26.65 (0.00)	27.13 (0.00)	27.86 (0.00)
Q (20) p value in parentheses	39.76 (0.00)	39.88 (0.00)	40..65 (0.00)	39.63 (0.00)	40.18 (0.00)	40.88 (0.00)
Sqrd Standardized Residuals						
Q (5) p value in parentheses	3.60 (0.61)	6.28 (0.28)	5.40 (0.37)	1.03 (0.96)	1.36 (0.93)	1.20 (0.94)
Q (10) p value in parentheses	7.85 (0.65)	9.90 (0.45)	9.24 (0.51)	5.18 (0.88)	5.30 (0.87)	5.19 (0.88)
Q (20) p value in parentheses	21.22 (0.38)	22.03 (0.34)	21.72 (0.35)	18.61 (0.55)	18.11 (0.58)	18.17 (0.58)
Information Criteria						

AIC	19765.29 (df 6)	19454.19 (df 7)	19508.22 (df 7)	19696.59 (df 7)	19387.45 (df 8)	19450.08 (df 8)
BIC	19805.46 (df 6)	19501.05 (df 7)	19555.09 (df 7)	19743.45 (df 7)	19441.01 (df 8)	19503.64 (df 8)
OBS	5,971	5,971	5,971	5,971	5,971	5,971
Wald chi2(2)	36.00 (0.00)	46.25 (0.00)	37.99 (0.00)	41.79 (0.00)	49.59 (0.00)	40.87 (0.00)

Parameter	**ARMA-GARCH**			**ARMA-GJR-GARCH**		
	N	**t**	**GED**	**N**	**t**	**GED**
ar_1	0.1165	-0.183	-0.0286	-0.0153	-0.1866	-0.0664
	(0.1384)	(0.0134)	(0.1554)	(0.1446)	(0.1449)	(0.1559)
ma_1	-0.0281	0.2668	0.1099	0.1033	0.271	0.1477
	(0.1394)	(0.0131)	(0.1547)	(0.1442)	(0.1420)	(0.1546)
α_0	0.0477***	0.0587***	0.0580***	0.0509***	0.0596***	0.0601***
	(0.0150)	(0.0160)	(0.0140)	(0.0152)	(0.0143)	(0.0141)
α_1	0.2083***	0.1991***	0.2022***	0.1109***	0.1062***	0.1075***
	(0.0081)	(0.0161)	(0.0136)	(0.0050)	(0.0092)	(0.0083)
β_1	0.9794***	0.9783***	0.9789***	0.8930***	0.8947***	0.8944***
	(0.0019)	(0.0037)	(0.0032)	(0.0047)	(0.0080)	(0.0075)
λ	-0.0754***	-0.0814***	-0.0786***	-0.3053***	-0.3772***	-0.3472***
	(0.0054)	(0.009)	(0.0084)	(0.0288)	(0.0511)	(0.0476)
ν	-	7.05	1.39	-	7.02	1.39
		(0.5763)	(0.0262)		(0.5165)	(0.0263)
δ				1.4314***	1.3243***	1.3648***
				(0.0727)	(0.1410)	(0.1247)
Log(L)	-9831.44	-9676.62	-9708.52	-9833.37	-9678.05	-9710.15
Diagnostic Checking						
Standardized Residuals						
Q (5) p value in parentheses	9.83 (0.08)	7.09 (0.21)	8.90 (0.11)	8.27 (0.14)	7.40 (0.19)	8.81 (0.11)
Q (10) p value in parentheses	28.69 (0.00)	26.34 (0.00)	27.96 (0.00)	27.29 (0.00)	26.87 (0.00)	28.06 (0.00)
Q (20) p value in parentheses	43.73 (0.00)	41.10 (0.00)	42.83 (0.00)	41.15 (0.00)	40.92 (0.00)	42.06 (0.00)
Sqrd Standardized Residuals						
Q (5) p value in parentheses	3.56 (0.61)	4.64 (0.46)	4.24 (0.51)	2.25 (0.81)	3.98 (0.55)	3.28 (0.65)
Q (10) p value in parentheses	8.76 (0.55)	10.27 (0.41)	9.66 (0.47)	6.91 (0.73)	9.24 (0.51)	8.22 (0.61)
Q (20) p value in parentheses	34.11 (0.02)	35.19 (0.02)	34.90 (0.02)	26.32 (0.15)	30.14 (0.07)	28.53 (0.10)
Information Criteria						
AIC	19676.87 (df 7)	19369.23 (df 8)	19433.05 (df 8)	19682.74 (df 8)	19374.10 (df 9)	19438.31 (df 9)
BIC	19723.74 (df 7)	19422.79 (df 8)	19486.60 (df 8)	19736.30 (df 8)	19434.35 (df 9)	19498.56 (df 9)
OBS	5,971	5,971	5,971	5,971	5,971	5,971
Wald chi2(2)	45.64 (0.00)	41.10 (0.00)	39.59 (0.00)	41.01 (0.00)	48.79 (0.00)	39.74 (0.00)

Asymmetric Leverage Effect

In asymmetric GARCH models, λ_1 measures the asymmetry in volatility. These models respond asymmetrically to positive and negative lagged values of standardized residuals. Since negative shocks / bad news tend to have a larger impact on volatility, the value of λ_1 would be expected to be negative. In all our models, λ_1 denotes the coefficient of leverage effects. The EGARCH, GJR-GARCH, and APARCH models are capable of modeling leverage effects. Tables 2 and 3 give estimation results of leverage effects of three asymmetric basic and enhanced models, respectively. We can clearly see the impact of leverage effects in these models. All of the asymmetric models show statistically significant leverage effect at one percent significance level. For EGARCH, GJR-GARCH, and APARCH models under t-distribution, the leverage effect values are -0.0746, -0.1016, and -0.3514, respectively, and for ARMA-EGARCH, ARMA-GJR-GARCH, and ARMA-APARCH models, the leverage effect values are -0.0754, -0.1109, and -0.3772, respectively. These results indicate that NIFTY log returns have significant leverage effect and bad news have a larger impact on volatility. All the parameters in the equations for ln(h_t) in EGARCH and ARMA-EGARCH with student's t-distribution are highly significant. This result indicates that let h_{t-1} = 1, a one unit decline in ε_{t-1} will increase the log of conditional volatility by 0.2703 units (0.1957 + 0.0746) in EGARCH model, and 0.2805 units (0.1991 + 0.0814) in ARMA-EGARCH model. However, a one unit increase in ε_{t-1} is estimated to induce a smaller effect in the log of the conditional variance by 0.1211 units (0.1957 - 0.0746) and 0.1177 units (0.1991 - 0.0814). The implication of these results is that bad news has a large effect on the conditional volatility and good news has small effect in volatility.

Diagnostic Checks for Model Adequacy

To check model adequacy, we form standardized residuals and squared standardized residuals of each symmetric and asymmetric models. We use these estimated residuals to obtain Ljung-Box Q-statistics for 5, 10, and 20 lags. Standardized residuals are analyzed to observe for no remaining serial correlations and squared standardized residuals are analyzed to observe for no remaining GARCH effects. Table 2 provides diagnostic checks statistics for basic GARCH models and Table 3 provides the same for ARMA-GARCH models. In basic GARCH models, we do not find any remaining GARCH effects. The Q-statistics for 5, 10, and 20 lags are statistically insignificant at any conventional significance levels. At the same time in basic GARCH models, we find remaining serial correlations. The Q-statistics are statistically significant indicating that the standardized residuals are not white noise process and there are remaining serial correlations. We improved our basic model by including ARMA (1,1) coefficients. In the ARMA-GARCH models, the ARMA coefficients are insignificant. We once again find that the Q-statistics of squared standardized residuals at 5, 10, and 20 lags are statistically insignificant, thus indicating the absence of remaining GARCH effects. One improvement we notice in these enhanced models is that there are no remaining serial correlations in Q(5). However, there are still some remaining serial correlations in Q(10) and Q(20). With such a large volume of return data, we can expect traces of serial correlations in Q(10) and Q(20) lags. Results suggest that we get best-fit models when ARMA (1,1) coefficients are included in modeling GARCH models with high-frequency data spread over long horizon.

Best-Fit Model

We use two information criterions and log likelihood values to choose best-fit models. Undoubtedly, the ARMA-GARCH model under student's t-distribution explains key stylized facts of NIFTY return series. This model portrays volatility clustering, fat-tail behavior, long memory, and mean reversion behavior of asset returns. The leverage effect is pertinently model by asymmetric models. However, ARMA-EGARCH model with student's t-distribution has better log likelihood value and minimum information criterion values. Among the asymmetric models, the ARMA-EGARCH model appropriately captures the leverage effect. For this model, the AIC = 19369.23 and BIC = 19422.79. These information criterion values are least among information criterion values of all the models considered in this paper. This paper finds that the ARMA-GARCH model is appropriate to capture the stylized facts of NIFTY return series and the ARMA-EGARCH model is appropriate to capture leverage effect.

CONCLUSION

Using high frequency, long horizon, and daily stock market returns, this paper empirically examined choice of return innovation distribution in symmetric and asymmetric GARCH models. As presented and discussed in introduction and literature review sections of this paper, the choice of return innovation distribution is dissimilar. We obtained NSE NIFTY50 data from January 1, 1996 to December 31, 2019. This is the first study that covers such a long horizon and high-frequency daily returns of NIFTY. We applied Gaussian distribution, t-distribution, and GED distribution to estimate multiple GARCH, EGARCH, GJR-GARCH, and APARCH models using 5,971 daily observations. Again, to the best of our knowledge, no other research paper estimated these many models of NIFTY index under multiple return innovation distributions. We observed that the NIFTY return series is leptokurtic. The log return series is negatively skewed and there is presence of excess kurtosis. This behavior is confirmed with J-B test of normality. We also observed that the log return series is stationary at level. NIFTY index close price series had upward movement. The return series and log return series are stochastic. From estimated models, we noticed volatility clustering, fat-tails, mean reversion, long memory, volatility persistence, and leverage effect. All the models are statistically significant at any conventional significance level, making over decision to choose appropriate return innovation distribution difficult. We relied on log likelihood value, AIC, and BIC to choose appropriate innovation distribution. These three values are in favor of student's t-distribution among each GARCH model. Based on our results, we recommend student's t-distribution as return innovation distribution for long horizon, high-frequency stock market returns. We extended our research further and performed diagnostic tests on standardized residuals and squared standardized residuals to examine presence of serial correlations and GARCH effects. We estimated Ljung-Box Q-statistics for 5, 10, and 20 lags representing week, fortnight, and month, respectively. In basic GARCH (1,1) models, we observed no remaining GARCH effect. However, there are remaining serial correlations. We experimented with different model specification and finally we enhanced our basic models with ARMA (1,1)-GARCH (1,1) models. In the ARMA (1,1)-GARCH (1,1) models, there are no remaining serial correlations and remaining GARCH effects. In the enhanced models also, model-selection criterion values are in favor of student's t-distribution. Among the asymmetric GARCH models, ARMA-EGARCH (1,1) had better model-selection criterion values and appropriately modeled leverage effect. As per the ARMA-GARCH (1,1) specifications, the unconditional volatility of NIFTY log returns is 3.24 and the mean reversion period is 88 days. The ARMA-EGARCH (1,1) model predicts for every unit change in unconditional volatility bad news will have 2.805 units impact indicating significant leverage effect. Finally, the v values in all the models indicated that NIFTY log return series has thicker tail behavior than normal distribution. We conclude that student's t-distribution is appropriate distribution density for high-frequency NIFTY log return series. In addition, we recommend ARMA-GARCH (1,1) model and ARMA-EGARCH (1,1) with t-distribution as best-fit symmetric and asymmetric models to estimate stylized facts of NIFTY log return series. In this paper, we do not consider low-frequency – monthly return series. Future studies may consider low-frequency data and other GARCH family models along with the models estimated in this paper.

REFERENCES

Ahmed, R. R., Vveinhardt, J., Štreimikienė, D., Ghauri, S. P., & Ashraf, M. (2018). Stock returns, volatility and mean reversion in emerging and developed financial markets. *Technological and Economic Development of Economy, 24*(3), 1149-1177.

Aziz, Z., & Iqbal, J. (2016). Comparing volatility forecasts of univariate and multivariate GARCH models: Evidence from the Asian stock markets. *Journal of International and Global Economic Studies, 9*(2), 67-78.

Bhowmik, R., Chao, W. U., Kumar, J. R., & Shouyang, W. (2017). A study on the volatility of the Bangladesh stock market - Based on GARCH type models. *Journal of Systems Science and Information, 5*(3), 193-215.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews, 11*(2), 143-172.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Economics, 31*(3), 307-327.

Chkili, W., Aloui, C., & Nguyen, D. K. (2012). Asymmetric effects and long memory in dynamic volatility relationships between stock returns and exchange rates. *Journal of International Financial Markets, Institutions and Money, 22*(4), 738-757.

Dedi, L., & Yavas, B. F. (2017). Equity returns and volatilities before and after the 2007-08 financial crisis. *Zagreb International Review of Economics & Business*, *20*(1), 65-79.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory of stock returns and a new model. *Journal of Empirical Finance, 1*, 83-106.

Dury, M. E., & Xiao, B. (2018). Forecasting the volatility of the Chinese gold market by ARCH family models and extension to stable models, hal-01709321.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. Inflation. *Econometrica, 50*(4), 987-1008.

Ghangare, A. S. (2016). A comparative analysis of volatility models of share prices of State Bank of India using ARCH GARCH family models. *International Journal of Pure and Applied Mathematics, 119*(16), 1337-1352.

Glosten, L. R., Jagannathan, R., & Runkle, D. E (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance, 48*(5), 1779-1801.

Hansen, P. R., & Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics, 131*(1/2), 97-121.

Hemanth, K. P., & Basavaraj, P. S. (2016). Volatility forecasting - A performance measure of GARCH techniques with different distribution models. *International Journal of Soft Computing, Mathematics and Control*, *5*(2/3), 1-13.

John, J. K., & Amudha, R. (2019). Asymmetric volatility modeling and leverage effect of Nifty stocks. *International Journal of Innovative Technology and Exploring Engineering, 8*(6), 280-289.

Nelson, (1991). Conditional heteroscedasticity in asset returns: A new approach, *Econometrica, 59*(2), 347-370.

Olbrys, J., & Majewska, E. (2017). Asymmetry effects in volatility on the major European stock markets: The EGARCH based approach. *Quantitative Finance and Economics, 1*(4), 411-427.

Sharma, P., & Vipul. (2016). Forecasting stock market volatility using realized GARCH model: International evidence. *The Quarterly Review of Economics and Finance, 59*, 222-230.

Tripathy, S., & Rahman, A. (2013). Forecasting daily stock volatility using GARCH Model: A comparison between BSE and NSE. *The IUP Journal of Applied Finance, 19*(4), 71-83.

Vasudevan, R. D., & Vetrivel, S. C (2016). Forecasting stock market volatility using GARCH models: Evidence from the Indian stock market. *Asian Journal of Research in Social Sciences and Humanities, 6*(8), 1565-1574.