

Modeling Stock Returns Volatility of the Nairobi Securities Exchange Index and Other Indices

Kalovwe, Sebastian Kaweto^{1*} and Mwaniki, Ivivi Joseph²

¹ Faculty of Science, Department of Mathematics and Computer Science, the Catholic University of Eastern Africa, Nairobi, Kenya

² School of Mathematics, University of Nairobi, Nairobi, Kenya
Email: skaweto@gmail.com

Abstract. This paper seeks to model daily, weekly and monthly stock indices returns using GARCH (1,1) model which is expected to reproduce most of the stylized facts of financial time series data which, in most cases, are found in different types of market. In addition, the distributional behavior of returns as the data changes from daily through to monthly returns is investigated by performing the JB and K-S tests. The results indicate evidence of volatility clustering, leverage effects, Gaussianity and leptokurtic distribution in the stock returns. A key observation is that the monthly returns of the three indices follow a Gaussian distribution (i.e. as the data changes from daily through to monthly returns it follows a normal distribution).

Keywords: GARCH, Gaussianity, Stock returns, volatility, heteroscedasticity.

1 Introduction

Statistical properties of stock prices and market indexes have been studied using data from various markets and instruments over the past couple of years. A set of these properties have been observed to be common across many instruments, markets and time periods and have been classified as 'stylized facts'. For instance, Cont [9] examined stylized statistical properties of asset returns, which are common to a wide set of financial assets, such as heavy tails, leptokurtic distribution, volatility clustering, absence of autocorrelations, leverage effects, aggregation Gaussianity, etc. The Autoregressive Conditional Heteroskedasticity (ARCH) model with normal innovations first introduced by Engle [11] captured some of stylized characteristics of financial assets. To improve the fit of the GARCH and EGARCH models into international equity markets, Fernandez and Steel [13] and Harris et al. [14] used the Skewed Generalized Student's t-distribution to capture the skewness and leverage effects of daily returns. Bollerslev [6], Baillie and Bollerslev [1], and Beine et al. [2] have used the Student's t-distribution to embrace the thick tails property of high frequency financial time series data. Further, studies of Mandelbrot [15], Fama [12], Black [4], Christie [8], Dowd [10] and Poon [16] indicate that, financial time series data is characterized by volatility clustering, leptokurtosis, leverage effects, have fat tails and a greater peak at the mean than the normal distribution. In economic time series analysis, it is quite common to develop models and theories under the assumption of Gaussianity and apply them to real data. The research findings of Mandelbrot [15] indicate that the empirical distributions of daily stock returns differ significantly from the traditional Gaussian model.

The main focus of this paper is to model the indices returns using GARCH(1,1) model which attempts to capture most of the stylized features of return series. Moreover, we investigate the distribution of returns by using data of various frequencies. All the parameters are estimated from historical data for FTSE100, S&P500 and NSE20 indices from Jan 3, 2000 to Dec 31, 2012.

This paper is organized as follows. Section 2 provides an overview of the ARCH and GARCH models and the data used in this study. Section 3 provides the descriptive statistics and the general discussion of the study findings. Finally, section 4 concludes the paper.

2 Methodology

2.1 ARCH Model

An ARCH model is a stochastic process with autoregressive conditional heteroscedasticity. The autoregressive property describes a feedback mechanism that incorporates past observations into the present while conditionality implies a dependence on the observations of the immediate past and heteroscedasticity means time-varying variance (volatility). The model was first introduced by Engle [11] when modeling the United Kingdom inflation. ARCH models are simple models capable to describe a stochastic process which is locally non-stationary but asymptotically stationary. If the stochastic process exhibits a time dependent variance, i.e. volatility, then the ARCH models are particularly useful and therefore have been applied to many different areas of economics such as interest rates, stock returns, foreign exchange rates, etc. In an ARCH process, the variance at a time t depends on some past values and it is characterized by a certain number of parameters. An ARCH(p) model assumes that

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_p a_{t-p}^2 \quad (1)$$

where ε_t is a sequence of *i.i.d.* random variables with mean 0 and variance 1, and $\alpha_i \geq 0$ for $i > 0$.

2.2 GARCH Model

GARCH-models which refer to generalized autoregressive conditional heteroskedasticity models have been widely used in financial and economic modeling and analysis. These models are characterized by their ability to capture volatility clustering, and they are widely used to account for non-uniform variance in time-series data.

The GARCH model is based on the assumption that forecasts of variance changing in time depend on the lagged variance of capital assets. An unexpected increase or fall in the returns of an asset at time t will generate an increase in the variability expected in the period to come. These models were proposed by Bollerslev [5] as a useful extension of ARCH model. For a log return series, r_t , we assume that the mean equation of the process can be adequately described by an ARMA model. Let $a_t = r_t - \mu_t$ be the mean-corrected log return. Then a_t follows a GARCH(p, q) model if

$$a_t = \alpha_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (2)$$

where q is the degree of GARCH; p is the degree of the ARCH process, ε_t is a sequence of *i.i.d.* random variables with mean 0 and variance 1, $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$ and $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_j) < 1$.

We shall note that $\alpha_i = 0$ for $i > p$, and $\beta_j = 0$ for $j > q$. The constraint $\alpha_i + \beta_j$ implies that the unconditional variance of a_t is finite, whereas its conditional variance a_t^2 evolves over time. Equation (2) reduces to a pure ARCH (p) model if $q = 0$. The basic and most widespread model is GARCH(1,1), which can be expressed as

$$a_t = \alpha_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (3)$$

According to Brook and Burke [7], the GARCH(1,1) is sufficient to capture all the volatility clustering that is present in a data. This study therefore applies this model to investigate the volatility clustering and other stylized facts of the time series data in question.

2.3 Jarque-Bera Test

In statistics, the Jarque-Bera test is a goodness-of-fit test of whether sample data have the skewness and kurtosis matching a normal distribution. The test statistic JB is defined as

$$JB = \frac{n}{6} \left[s^2 + \frac{1}{4} (k - 3)^2 \right] \quad (4)$$

where n is the number of observations (or degrees of freedom in general), s is the sample skewness and k is the sample kurtosis. If the data come from a normal distribution, the JB statistic asymptotically has a chi-squared distribution with two degrees of freedom, so the statistic can be used to test the hypothesis that the data are from a normal distribution. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being zero. Samples from a normal distribution have an expected skewness of zero and an excess kurtosis of zero (which is the same as kurtosis of 3). As the definition of JB shows, any deviation from this increases the JB statistic.

2.4 The Kolmogorov-Smirnov Test

The K-S test is a nonparametric test of the equality of continuous, one dimensional probability distributions that can be used to compare a sample with a reference probability distribution (one-sample K-S test), or to compare two samples (two-sample K-S test). The K-S statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution, or between the empirical distribution functions of the two samples. The K-S statistic for a given cumulative distribution function $F(x)$ is

$$D_n = \sup_x |F_n(x) - F(x)| \quad (5)$$

where \sup_x is the supremum of the set of distances, $F_n(x)$ is the empirical distribution function of the sample for all x . By the Glivenko-Cantelli theorem, if the sample comes from distribution $F(x)$, then D_n converges to zero almost surely in the limit when n goes to infinity.

2.5 Empirical Data

In our analysis we focus on the daily, weekly and monthly closing indices as reported in Nairobi Securities Exchange for NSE20 share index, FTSE100 index in London and S&P500 index in New-York Stock Exchange. Daily, weekly and monthly log-returns, r_t , of FTSE100, S&P500 and NSE20 share indexes are computed from Jan 3, 2000 to Dec 31, 2012. Let p_t denote the successive closing price observation at time t , the continuously compounded (or log) return is defined as

$$r_t = \ln p_t - \ln p_{t-1} \quad (6)$$

3 Results and Discussion

3.1 Descriptive Statistics

Table 1 summarizes the basic statistical properties of the data. The mean returns are positive except for the FTSE100 index and the S&P500 daily observations but close to zero. The returns appear to be somewhat asymmetric as reflected by skewness estimates: there are more observations in the left-hand tail than in the right-hand tail. All the series returns have heavy tails and show strong departure from normality (skewness and kurtosis coefficients are all statistically different from those of the normal distribution which are 0 and 3 respectively). The kurtosis and the value of Jarque-Bera (JB) statistic decrease as the time scale increases, however, the JB statistics clearly rejects the null hypothesis of normality.

3.2 Empirical Findings and Discussion

We estimate GARCH(1,1) type model assuming conditional normality. The results of estimating the daily, weekly and monthly returns of the three market indices are presented in tables 3. The results indicate that the estimated coefficients of the model (α and β) meet the requirement that $\alpha + \beta < 1$ and $\alpha_0 > 0, \alpha_1 > 0, \beta_1 \geq 0$, which is a crucial condition for a mean reverting process. This implies that the fitted GARCH(1,1) is weakly stationary and that conditional volatilities are mean reverting for all the time series and frequencies. Also in the variance equation the first three coefficients ω (constant), ARCH

term (ω_1) and GARCH term (β_1) are highly significant for the three indices returns in daily, weekly and monthly observations with the exception that omega (ω) is not significant for monthly observations of the three market indices returns. The significance of α and β indicates that lagged conditional variance and squared disturbance has an impact on conditional variance, this means that news about volatility from previous periods has an explanatory power on current volatility. The GARCH coefficient β_1 is found to be around 0.9 and thus it's obvious that large values of σ_{t-1}^2 will be followed by large value of σ_t^2 , and small values of σ_{t-1}^2 will be followed by small values of σ_t^2 . The coefficient α_1 measures the degree to which volatility shock that occurs now feeds through into next period's volatility.

The value of β_1 is high in most of the return series data from the three indices, hence it can be inferred that the shocks to conditional variance dies after a long time. In this case we note that volatility is persistent. We also note that the sum of α_1 and β_1 is close to the one which implies that a shock at a time t will persist for a long time in the future. α_1 is less than β_1 . It can be inferred that the volatility of the stock market index is affected by past volatility more than by related news from the previous period.

To study the characteristics of the distribution of the returns of the three markets indices, we plotted the empirical probability density functions and compared it with the normal distribution (see figure 1). The plots indicate that as the data changes from daily through to monthly, their distribution looks more and more like a normal distribution. In particular, the shape of the distribution is not the same at different time scales. These results are also confirmed by the results of the Kolmogorov-Smirnov test shown in table 2 where the null hypothesis that the returns follow normal distribution was rejected for the daily and weekly returns but was accepted for the monthly returns of the three markets.

Table 1. Descriptive statistics for daily, weekly and monthly stock returns

Market	Obs	nobs	Min	Max	Mean	Std Dev	skewness	Ex. Kurt	J-B test
FTSE 100 Index	D	3390	-0.093	0.094	-4.8e-5	0.013	-0.143	5.992	5092
	W	678	-0.236	0.126	-9.7e-5	0.026	-1.099	11.50	3898
	M	155	-0.140	0.083	-3.9e-4	0.042	-0.694	0.618	15.5
S&P 500 Index	D	3268	-0.095	0.110	-6.0e-6	0.014	-0.158	7.321	7323
	W	677	-0.201	0.114	2.5e-5	0.027	-0.768	6.206	1162
	M	155	-0.186	0.102	1.45e-4	0.047	-0.668	1.061	19.8
NSE20 Index	D	3257	-0.052	0.070	1.80e-4	0.009	0.464	7.571	7908
	W	678	-0.128	0.155	8.64e-4	0.027	0.512	5.160	788
	M	155	-0.263	0.178	3.67e-3	0.062	-0.272	2.192	34.9

D=daily, W=weekly and M=monthly observations

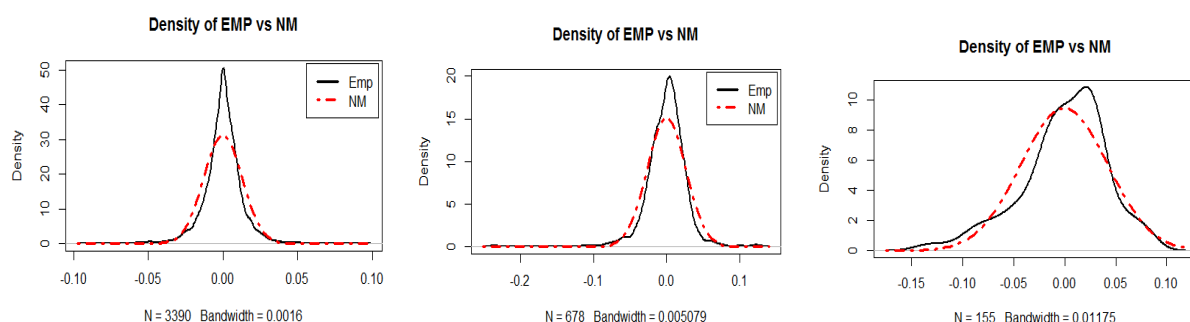


Figure 1. Empirical vs. normal densities for FTSE100 index returns, daily (left), weekly (centre) and monthly (right)

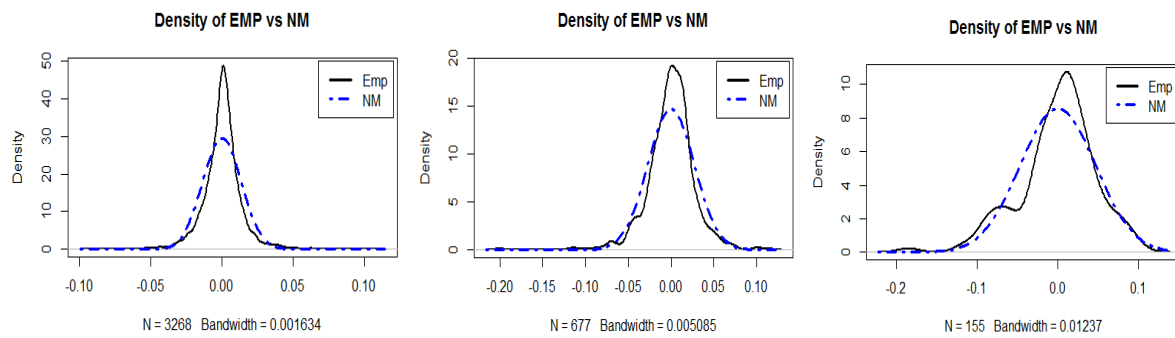


Figure 2. Empirical vs. normal densities for S&P500 index returns, daily (left), weekly (centre) and monthly (right)

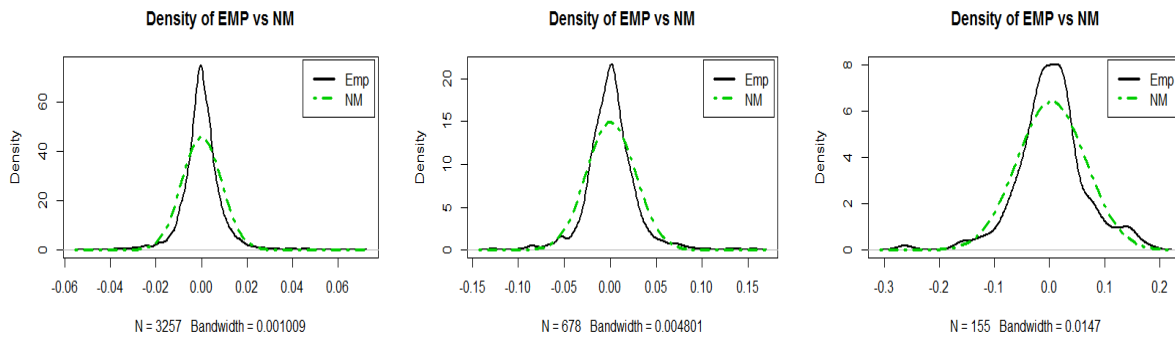


Figure 3. Empirical vs. normal densities for NSE20 index returns, daily (left), weekly (centre) and monthly (right)

Table 2. Kolmogorov-Smirnov test for density of returns

INDEX	Observations	Normal distribution		NIG distribution	
		D- statistic	p-value	D- statistic	p-value
FTSE100	Daily	0.0776	< 2.2e-16	0.0381	0.000105
	Weekly	0.0648	0.006775	0.0437	0.1497
	Monthly	0.0777	0.3066	0.1064	0.05999
S&P500	Daily	0.0797	< 2.2e-16	0.0471	1.035e-06
	Weekly	0.071	0.002155	0.0392	0.2486
	Monthly	0.0968	0.1095	0.0912	0.152
NSE20	Daily	0.0882	< 2.2e-16	0.0465	1.546e-06
	Weekly	0.0777	0.0005621	0.0397	0.2344
	Monthly	0.0892	0.1699	0.0435	0.9308

Table 3. GARCH(1,1) for returns

Parameters	FTSE100 index returns			S&P500 index returns			NSE20 index returns		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly
μ	3.600e-04 0.0128*	9.723e-04 0.1847	3.748e-03 0.1847	6.166e-05 0.697	2.540e-04 0.7485	1.452e-03 0.6059	-2.262e-5 0.849	4.63e-04 0.5766	0.0018 0.7053
ω	1.238e-06 1.05e-4***	3.322e-05 0.0052 **	1.103e-04 0.2091	1.511e-06 2.3e-6***	4.119e-05 0.0037**	6.080e-05 0.3129	4.039e-06 1.3e-7***	1.12e-04 3.0e-4***	0.0005 0.1008
α	1.056e-01 <2e-16***	2.561e-01 5.6e-06***	2.199e-01 0.0114*	8.568e-02 <2e-16***	2.109e-01 5.3e-7***	2.409e-01 0.0046**	2.138e-01 <2e-16***	4.5e-01 1.1e-1***	0.2675 0.0213*
β	8.889e-01 <2e-16***	7.278e-01 <2e-16 ***	7.309e-01 8.9e-16***	9.049e-01 <2e-16***	7.412e-01 <2e-16***	7.463e-01 <2e-16***	7.385e-01 <2e-16***	4.44e-01 1.1e-1***	0.6109 5.8e-6***
$\alpha+\beta$	0.9945	0.9839	0.9508	0.9906	0.9521	0.9872	0.9523	0.894	0.8785
Logl	10692.9	1597.2	281.6	10172.5	1577.1	271.2	11459.2	1599.5	220.9
nl	3.154	2.356	1.817	3.113	2.330	1.750	3.518	2.359	1.425
JB	94.71 0	499.0 0	16.03 0.0003	244 0	80.90 0	7.641 0.0219	2618.6 0	82.44 0	2.046 0.3595
Q(10)	13.18 0.2139	10.38 0.4081	3.366 0.9714	18.20 0.0516	7.426 0.6847	5.362 0.8657	502.4 0	76.28 2.7e-12	30.55 0.0007
Q(15)	17.52 0.2886	20.40 0.1571	7.195 0.9520	27.98 0.0218	11.560 7121	6.441 0.9714	523.9 0	88.03 2.3e-12	35.15 0.0023
LM Test	17.54 0.1306	17.61 0.1282	8.322 0.7595	28.36 0.0049	14.26 0.2846	6.869 0.8661	13.77 0.3154	6.244 0.9033	16.64 0.1638
AIC	-6.306	-4.700	-3.582	-6.223	-4.647	-3.448	-7.034	-4.707	-2.798
BIC	-6.299	-4.673	-3.503	-6.216	-4.621	-3.370	-7.027	-4.680	-2.720

Note: *, ** and*** indicate significant at 5%, 1% and 0.1% respectively.
 The estimated GARCH parameters are all highly significant.
 Q (10) is Ljung-Box test at lag 10 and Q (15) is Ljung-Box test at lag 15
 nl is normalized
 Logl is log likelihood
 AIC is the Akaike Information Criterion;
 BIC is the Bayesian Information Criterion;
 JB is Jarque-Bera test

4 Conclusion

In this paper, our empirical results provide evidence that the behavior of the three stock indices returns has the common characteristics of many daily, weekly and monthly financial time series. The data exhibits considerable level of excess kurtosis, which can be related to the time-dependence in conditional variance and also the distribution of all return series is relatively asymmetric. As a consequence of these two characteristics, all the return series data shows a significant departure from normality and existence of conditional heteroscedasticity. All the three series returns exhibit asymmetric behavior in the conditional variance, related by many authors to leverage effects. The skewness and leptokurtosis observed in the original series of returns can be partly explained by the GARCH model. The study, thus, finds strong evidence of volatility clustering, leverage effects and leptokurtic distribution for the indices returns. The results of K-S test provide evidence that as the data changes from daily through to monthly returns, the distribution is close to that of normal.

Acknowledgement

Comments from the editor and anonymous referees are highly acknowledged.

References

1. Baillie, R. and Bollerslev, T. (1989). Common stochastic trends in a system of exchange rates, *Journal of Monetary Economics*, 44, 167-81.
2. Beine, et.al (2002). Accounting for conditional leptokurtosis and closing days effects in FIGARCH models of daily exchange rates, *Applied Financial Economics*, 12, 589-601.
3. Bekaert, G and Harvey, C.R.(1997).“Emerging Market Volatility,” *Journal of Financial Economics*, 43: 29-77.
4. Black, F. (1976). “Studies of Stock Market Volatility Changes”, *Proceedings of the American Association, Business and Economic Statistics Section*, pp. 177-181
5. Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of econometrics*, 31,307-327.
6. Bollerslev, T. (1987). A conditionally Heteroskedastic time series model for speculative prices and rates of return, *review of economics & statistics*, 69,542-7.
7. Brook, C. and Burke, S.P. (2003). “Information Criteria for GARCH Model Selection:
8. Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage, and nterest rate effect. *Journal of financial economics*, 10,407-432.
9. Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223-236.
10. Dowd, K. (2005). *Measuring Market Risk*. John Wiley & Sons Ltd, England.
11. Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with estimates of the Variance of U.K inflation, *Econometrica*, 50,987-1008.
12. Fama, E. (1965). “The Behavior of Stock Market Prices”, *Journal of Business*, 38 (1), 34-105.
13. Fernandez, C. and Steel, M. (1998). On Bayesian modeling of fat tails and skewness, *Journal of the America Statistical Association*, 93,359-71.
14. Harris, et.al (2004). Skewness in the conditional distribution of daily equity returns, *Journal of Applied Financial Economics*, 14,195-202.
15. Mandelbrot. (1963).The variation of speculative prices. *J.R. Stat. Soc. B47*, 165-202.
16. Poon, S.H. (2005). *A practical Guide to forecasting financial market volatility*. Wiley Finance, England.
17. Taylor, S. (1994). “Modeling Stochastic Volatility: A Review and Comparative Study”, *Mathematical Finance*, 4, 183-204.