American Journal of Applied Sciences 5 (2): 146-151, 2008
ISSN 1546-9239

The Gaussianity Evaluations of Malaysian Stock Return Volatility

Chin Wen Cheong
Faculty of Information Technology
Multimedia University, 63100 Cyberjaya, Selangor, Malaysia

Abstract: We study the distribution of standardized returns by using various frequencies data. The empirical standardized returns are obtained by using the unobserved and observable daily volatility. Our empirical results evidence the realized-standardized returns follow nearest to a Gaussian distribution. On the other hand, the standardized returns using daily closing and range-based data are able to reduce but not fully eliminate the excess kurtosis condition compare to the realized standardized returns.

Keywords: financial time series, realized volatility, discrete time-domain modelling

INTRODUCTION

The financial time varying volatility is closely related to the risk management applications such as risk diversification, portfolio analysis and derivative pricing prediction. One of the famous risk analysis applications is the immediate determination of value-at-risk (VaR) from the estimated volatility. The commercial application of *RiskMetrics*TM[1] is successfully applied VaR in portfolio investments. Since the estimated volatility hinges crucially on its associated conditional return distribution, thus, an insight understanding of the conditional return distribution is important in order to provide useful information in risk management analysis. In addition, the identified underlying distributions of conditional returns are important in the parametric discrete-time ARCH, stochastic volatility and continuous-time diffusion processes modelling under a particular underlying distribution assumption.

In this paper, we have selected the Malaysian stock exchange and as an emerging market, the nature of the stock market is characterized by low liquidity, infrequent trading, low quality of information and rapid changes in regulatory framework. The daily closing returns are adjusted for infrequent trading behaviour which might cause spurious correlation. We later focus on the distinction of standardized return from fractional integrated autoregressive moving average (ARFIMA) model, long-range dependence autoregressive conditional hetoroscedasticity (ARCH) models, range-based volatility and realized volatility in term of their Gaussianity. In this study, the standardized return is defined as σ-standardized return where σ is either the estimated or realized volatility standard deviation. In realized standardized returns, the 10- and 20-minute interval shows the nearest Gaussianity compare to others. The Gaussianity of these return series are analysed by their moments, empirical cumulative distribution function (CDF) plot, quantile-quantile (Q-Q) plots, Jacque-Bera test and a series of empirical distribution tests. The overall results indicate the realized standardized returns are nearly Gaussian distributed compare to other approaches modelling where improvement of reducing but not eliminate the leptokurtic.

DATA AND METHODOLOGY

Data Source: The Kuala Lumpur Composite Index (KLCI) transaction prices cover the recovery period 1st 2003 to 2006. This price index is weighted by market capitalization with the base year 1977 of 100 listed companies. After the bad experience of Asian Financial crisis, Malaysia implements the selective capital control in 1st September 1998 where RM is pegs at 3.80 to the USD. This action stabilizes the RM where non-residents from Malaysia and abroad are restrict to trade the RM. In addition, Securities Commission and the KLSE implements recovery strategic such as strengthening market intermediaries, improving market transparency and improving liquidity in corporate sectors. In this recovery period, the Malaysia stock market is speculated by the RM-USD

Corresponding Author: Chin Wen Cheong, Faculty of Information Technology, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia. Tel: +603-83125249, Fax: +603-83125264

un-pegged regulation(implemented at year middle of 2005 where the RM was expected undervalued by approximately 6.5%), the merged of MESDAQ in KLSE besides the Main board and Second board previously started in year 2002, the fluctuating of petrol prices, etc. We intend to study the reaction of market participants and market volatility to the good and bad news.

Adjusted return: Emerging markets are often related to the infrequently traded shares activities. This phenomena occurs when stocks market do not trade at every consecutive interval. We adopt the method proposed by Miller *et al* [2] as follows:

$$r_t = a_0 + a_1 r_{t-1} + \varepsilon_t \tag{1}$$

$$r_{t(adj)} = \frac{\varepsilon_t}{(1-a_1)}\,. \tag{2}$$

The model assumes that the non-trading adjustment required to adjusted returns is constant throughout the periods in most of the high traded markets. Let assumes that the return series after adjustment are decomposed as $r_{t(adj)} = \sigma_t \varepsilon_t$ where ε_t is independently and identically distributed with mean zero and unit variance. Therefore, the σ-standardized return can be rearranged become,

$$\varepsilon_t = \frac{r_t}{\sigma_t}\,. \tag{3}$$

Based on this equation, if given the σ_t, the distribution and structure of ε_t can be determined straightforward. However, the conditional standard deviation is not directly observable and has to be estimated from ARCH-type models, range-based volatility or realized volatility respectively.

Unobservable estimated ARCH-types volatility:
Component GARCH: Ding and Granger[3] and Engle and Lee[4] decomposed volatility into two components with one component captures the short-run innovation impact and the other captures the long-run impact of an innovation as follows:

$$\begin{aligned} \sigma_t^2 &= \sigma_{t,q}^2 + \sigma_{t,s}^2 \\ \sigma_{t,q}^2 &= \omega + \gamma_{1q}\sigma_{t-1,q}^2 + \gamma_{2q}(a_{t-1}^2 - \sigma_{t-1}^2) \\ \sigma_{t,s}^2 &= \gamma_{1s}\sigma_{t-1,s}^2 + \gamma_{2s}(a_{t-1}^2 - \sigma_{t-1}^2) \end{aligned} \tag{4}$$

Fractionally Integrated GARCH: The conditional variance of FIGARCH(p,d,q) introduced by Baillie et.al[5] can be expressed as:

$$\sigma_t^2 = \frac{\alpha_0}{1-\beta(B)} + \left\{1 - \frac{\varphi(B)(1-B)^d}{1-\beta(B)}\right\} a_t^2\,, \tag{5}$$

with $0 \le d \le 1$. if d=0, the model will become $\beta(B)\sigma_t^2 = \alpha(B)a_t^2$, which is a GARCH model. If d=1, the model $\beta(B)(1-B)\sigma_t^2 = \alpha(B)a_t^2$, will follow a IGARCH model. And when d is $0 < d < 0.5$, the term $(1\text{-}B)^d$ has an infinite binomial distribution for non-integer powers.

ARFIMA-FIGARCH: The ARFIMA($p_1,d1,q_1$)-FIGARCH($p_2,d2,q_2$) model is able to capture the possibility of long-range dependence in both the conditional return and its volatility as follows:

$(1\text{-}B)^{d1}(1\text{-}\beta_1(B))r_t = (1\text{-}\alpha_1(B))a_t$, $\quad a_t = \sigma_t\varepsilon_t$,

$$\sigma_t^2 = \frac{\alpha_0}{1-\beta_2(B)} + \left\{1 - \frac{\alpha_2(B)\ (1-B)^{d2}}{1-\beta_2(B)}\right\} a_t\,. \tag{6}$$

The shock term, a_t, follows a conditional time-varying variance and the $\varepsilon_t \sim iid,\ N(0,\sigma^2)$. Davidson[6] and Caporin[7] argued that the $d1$ in ARFIMA is structural different from $d2$ where the persistence is increase when $d1$ approaches 0.5 compare to $d2$ approaches 0. The reverse behaviour may be due to the parameter acts directly on the squared errors but not on the conditional variance.

Unobservable range-based volatility: We have adopted the Parkinson[8] and Garman and Klass[9] approaches with the assumption of the expected return is equal to zero. The mean return is not statistically different from zero at 5% level under the *t-test*(t-statistic 1.7521). In addition, the selected range-based volatility model is without the inclusion of bid-ask information. This is due to the limitation of data source. The Garman and Klass is the extension of Parkinson with the inclusion of opening and closing price. Both the volatility estimators can be expressed as:

$$\sigma_{t,park}^2 = \frac{1}{4\ln 2}(H_t - L_t)^2\,; \tag{7}$$

$$\sigma_{t,GK}^2 = 0.511(H_t - L_t)^2 - 0.019\big(C_t(H_t - L_t) - 2H_tL_t\big) - 0.383C_t^2 \tag{8}$$

where the definitions are followed the Yang and Zhang [10] with H_t, L_t and C_t are the normalized high, low and closing prices respectively.

Observable realized volatility: In stock market, the intraday returns are obtained by summing the trading hours with the absence of overnight trading. However, we are able to observe the close-to-open return for the overnight period. The stock market encounters a short break in the afternoon and an overnight non-trading period. Under these conditions, we expect relative larger changes in the stock index price during the closing period compared to the n-minute returns observed during trading hours. Therefore the overnight and afternoon break will provide a distorting effect on the volatility estimation. A better alternative by using only intraday returns are proposed by Martens[11] and Hansen and Lunde[12]. Similarly, they suggested to use a scaled sum of squared intraday returns as follow:

$$RV_t = (1+c)\left(\sum_{a=1}^{\frac{T_M}{n}} \left(R_{t,a}^{M}\right)^2 + \sum_{b=1}^{\frac{T_A}{n}} \left(R_{t,b}^{A}\right)^2\right); \qquad (9)$$

where $c = \dfrac{\sigma_{oc}^2 + \sigma_{co}^2}{\sigma_{oc}^2}$. The OC and CO represent close-to-open and open-to-close respectively.

RESULTS AND DISCUSSION

Descriptive statistics: Table 1, the unadjusted return exhibits excess kurtosis and the Box-Ljung Q(12) statistic indicates the presence of serially correlation. The highly significant value for the first-order serial correlation is caused by the infrequent trading of emerging market. After the correction of thin trading effect, the adjustments appear to have eliminated the apparent serial correlation of the return series where the Q(12) shows insignificant serially correlation at 1% significant level.

Comparison of standardized returns: The distributions of each of the standardized-returns series are examined relative to a standard Gaussian distribution. We look into their moments, the empirical cumulative distribution function (CDF) plots, the Q-Q plots and empirical distribution test respectively.

Geometrically, the preliminary Gaussianity analyses are illustrated by the CDF plots and Q-Q plots respectively in Figure 1 and Figure 2. Most of the ARCH-standardized returns move closely to a simulated Gaussian distribution. On the other hand, especially the range-based standardized returns series show higher and lower cumulative probabilities in the early and end tails respectively relative to the simulated Gaussian CDF. In addition, the range-based standardized returns series also indicate relative wider range in the volatility attribute axis.

The Q-Q plot for adjusted return exhibit s-shaped pattern with symmetric and heavy tailed at both the ends. A few points fall on the end of the line indicate that the extreme values in the returns series. For ARCH,

Table 1: Return and adjusted return

	return	adjusted return
Mean	0.0493	-0.00002
Std. Dev.	0.6474	0.7572
Skewness	0.3064 ^{c} (3.410)	0.2525 ^{c}(2.810)
Kurtosis	4.5780 ^{c}(8.780)	4.3414 ^{c}(7.464)
Jarque-Bera	88.5970 ^{c}(0.0000)	63.5211 ^{c}(0.0000)
Normality test		
D	0.0499 ^{c}(0.0003)	0.01470 ^{c}(0.0057)
A^2	0.3888 ^{c}(0.0000)	0.2544 ^{c}(0.0011)
W^2	2.6389 ^{c}(0.0000)	1.7459 ^{c}(0.0002)
Autocorrelation		
lag 1	0.155	-0.005
lag 2	0.056	0.032
lag 3	0.020	0.005
lag 4	0.046	0.044
lag 5	0.005	-0.006
Q(12)	26.165	5.462
p-value	0.010^b	0.941

Note: the indications are the same in notes Table 2.

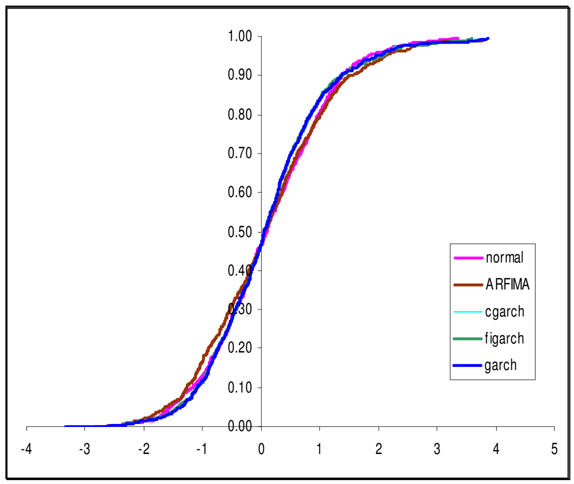


Fig. 1: CDF plots

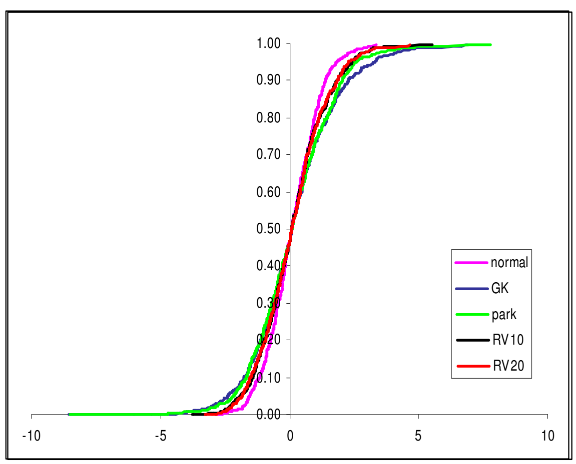


Fig. 2: CDF plots

Table 2: σ-standardized adjusted return Comparison

	σ-standardized adjusted return							
	σ(GARCH)	σ(CGARCH)	σ(FIGARCH)	σ(ARFIMA)	σ(park)	σ(GK)	σ(RV-10)	σ(RV-20)
Mean	-0.0040	-0.0062	-0.0076	-0.0110	-0.0007	0.0243	-0.0085	-0.0062
Std. Dev.	1.0005	0.9980	1.0057	1.1040	1.7945	1.8246	1.2718	1.2548
Skewness	0.3256^{c}	0.2837^{c}	0.2813^{c}	0.2746^{c}	0.2991^{c}	0.5250^{c}	0.1961^{b}	0.2442^{c}
	(3.623)	(3.157)	(3.131)	(3.052)	(3.3233)	(5.8333)	(2.182)	(2.717)
Kurtosis	4.0479^{c}	3.9448^{c}	3.9077^{c}	3.1758	4.5508^{c}	5.0530^{c}	3.2931	3.0602
	(5.831)	(5.257)	(5.050)	(0.977)	(8.6155)	(28.0722)	(1.631)	(0.335)
Jarque-Bera	46.9936^{c}	37.5039^{c}	35.2117^{c}	10.2710^{c}	84.2730^{c}	162.1903^{c}	7.4106^{b}	7.4863^{b}
	(0.0000)	(0.0000)	(0.0000)	(0.0059)	(0.0000)	(0.0000)	(0.0246)	(0.0237)
EDT								
	0.0147^{c}	0.0399^{c}	0.0410^{c}	0.0220	0.0781^{c}	0.0863^{c}	0.0315^{a}	0.0254
D	(0.0057)	$(0.0099)^{c}$	(0.0072)	(>0.100)	$(0.0000)^{c}$	$(0.0000)^{c}$	(0.0953)	(0.1135)
	0.2544^{c}	0.2402^{c}	0.2527^{c}	0.6403^{a}	5.7606^{c}	6.5152^{c}	0.0865	0.0604
A^2	(0.0011)	(0.0017)	(0.0012)	(0.0946)	(0.0000)	(0.0000)	(0.1703)	(0.3729)
	1.7459^{c}	1.6838^{c}	1.8156^{c}	0.0798	1.0203^{c}	1.1223^{c}	0.5535	0.4873
W^2	(0.0002)	(0.0003)	(0.0001)	(0.2081)	(0.0000)	(0.0000)	(0.1531)	(0.2235)
ACF								
lag 1	0.012	0.006	0.007	0.004	-0.024	-0.026	-0.058	-0.089
lag 2	0.031	0.029	0.038	0.037	0.037	0.033	0.036	0.029
lag 3	0.003	0.003	0.002	-0.001	0.032	0.035	0.004	0.009
lag 4	0.058	0.058	0.054	0.039	0.052	0.036	0.011	-0.003
lag 5	-0.030	-0.031	-0.029	-0.026	-0.049	-0.049	-0.02	-0.011
Q(12)	7.123	6.943	7.331	7.0826	-0.003	0.019	10.971	15.090
p-value	0.849	0.861	0.835	0.852	0.461	0.405	0.531	0.237

Notes: (1) *t-test* for Gaussian skewness and kurtosis.

The standard error for Gaussian skewness and kurtosis are $\sqrt{6/T}=0.090$ and $\sqrt{24/T}=0.180$. The parentheses indicate the *t*-statistics.

The null hypothesis indicates $\tilde{S}=0$ and $\tilde{K}=3$ respectively.

(2) Empirical distribution test(EDT): The test statistics are Kolmogorov-Smirnov(D), Anderson-Darling(A^2) and Cramer-von Mises(W^2) respectively. The parentheses indicate the p-values.

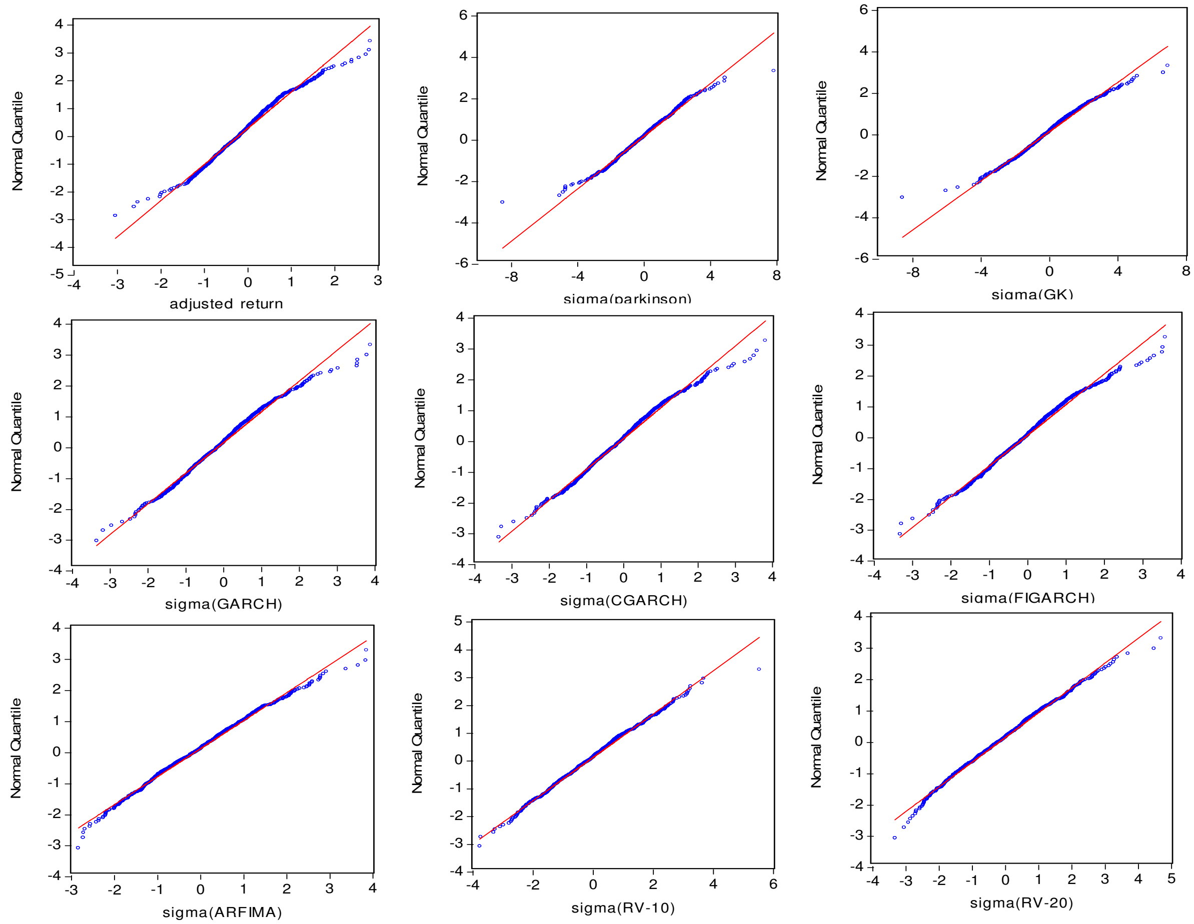
H_0: The return series follows the Gaussian distribution.

H_1: The return series does not follow the Gaussian distribution.

a, b and c denote 10%, 5% and 1% level of significance.

range-based and realized -standardized returns series, a nearly linear line indicating that Gaussian distribution provides a better approximation to the series. However, a series of statistical tests have to be implemented to examine the Gaussianity of the standardized returns series.

In Table 2, the standardized returns by long-range dependence ARCH-type estimations show a better approximation to Gaussian distribution with nearly unity standard deviation and smaller skewness compare to the adjusted and unadjusted returns series. The complete ARCH-type estimations are provided upon request. However, the ARCH-standardized returns still exhibit leptokurtic even though overall they show smaller excess kurtosis. On the other hand, the range-based-standardized return show similar results compare to ARCH. However, the range-based approach exhibits stronger violation from Gaussian distribution which indicated by the smaller p-value compare to ARCH approach. Finally, the realized-standardized returns show closest to a Gaussian distribution with decreased positive skewness and kurtosis approximate 3 that are 3.293 and 3.060 for 10 and 20-minute interval respectively. The *t*-statistics for gaussianity skewness and kurtosis tests in Table 2 confirm that only the realized standardized returns with 10-minute interval are insignificant at 5% significant level. Finally, formal tests by Jacque-Bera and empirical distribution tests are implemented to further analyze the presence of Gaussianity of the return series. From Table 2, only the realized standardized returns with 10-minute and 20-minute interval are normally distributed at 5% significant level. These findings are inline with the work by Andersen et.al.[13] and Andersen et.al.[14] who reported that the stock realized stock returns are approximately normally distributed.



Note: The symmetric red line indicates a normal-normal QQ-plot

Fig. 3: Quantile-quantile plots for the overall standardized returns series

CONCLUSION

In this paper, our empirical results evidence that the realized-standardized return series with 10- and 20-minute interval show remarkably near to a Gaussian distribution compare to ARCH and range-based standardized return series. The result suggests that for the underlying Gaussianity distribution assumption in volatility modelling, the realized volatility approach provide a better theoretical modelling framework. Our findings may offer some statistical implications in the distribution of returns series for any further theoretical modelling and prediction of others market financial time series.

ACKNOWLEDGEMENTS

The author would like to gratefully acknowledge the financial support from Multimedia University and the Research Design and Development and Innovation Grant in publishing this research.

REFERENCES

1. Morgan, J.P., 1996. RiskMetricsTM – Technical Document, New York.
2. Miller M.H., Muthuswamy, J. and Whaley, R.E., 1994. Mean reversion of Standard and Poor 500 index basis changes: Arbitrage-induced or statistical illusion? Journal of Finance, 49:479-513.
3. Ding Z., and C.W.J. Granger, 1996. Modelling volatility persistence of speculative returns: a new approach. Journal of Econometrics 73: 185-215.
4. Engle R.F., and G.G.J. Lee, 1999. A long-run and sort-run component model of stock return volatility, in cointegration, causality and forecasting, ed. By Engle RF, and White H, Oxford University Press.

5. Baillie R. T., Bollerslev T., and Mikkelsen H.O., 1996. Fractionally integrated autoregressive conditional heteroscedasticity, Journal of Econometrics 74**:** 3-30.
6. Davidson, J., 2001. Moment and memory properties of linear conditional heteroscedasticity models, Manuscript, Cardiff University.
7. Caporin, 2001. Estimation and identification of FIGARCH, Universita ca' Foscari di Venezia, Venezia, I.
8. Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return, Journal of Business, 53: 61-65.
9. Garman M., and Klass M., 1980. On the estimation of security price volatilities form historical data. Journal of Business 53: 67-78.
10. Yang, D., & Zhang, Q., 2002. Drift independent volatility estimation based on high, low, open and close prices. Journal of Business, 73: 477-491
11. Martens, M., 2002. Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. The Journal of Futures markets, 22: 497-518.
12. Hansen, P.R., & Lunde, A., 2004. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? Journal of Applied Econometrics, 20: 873-889.
13. Andersen, T.G., & Bollerslev, T., 1998. Answering the skeptics: Yes standard volatility models do provide accurate forecasts. International Economic Review, 39: 885-905.
14. Andersen, T.G., Bollerslev, T., Diebold, F.X., & Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica, 71: 529-626.