Volatility in Malaysian Stock Market: An Empirical Study Using Fractionally Integrated Approach

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

Abstract: This study explores the fractionally integrated (FI) time series analysis in Malaysian stock market. Four proxies of latent volatility, namely the absolute return, squared return and range-based (Parkinson and Garman and Klass) volatilities are selected for the empirical studies. In addition, the well-known FI autoregressive conditional variance (ARCH) type model is also taken into account for comparison purposes. Our empirical results evidence the proxy of absolute return and ARCH-type volatility model provides better performances in both the estimation and forecasting evaluations.

Key words: Range-based volatility, long-range dependence, econometrics, financial time series

INTRODUCTION

Recent literatures of long-range dependence[1,5] indicate that the fractionally integrated models are widely used in financial time series analysis such as risk management, portfolio analysis and derivative pricing. The long-range dependence processes commonly capture by the fractional integrated operator $(1-B)^d$ in ARFIMA model. Granger[6], Granger and Joyeux[4] and Hosking[7] employ this operator as a filter to transform a general time series into a constant plus a white noise process. Later, Baillie[1] introduces the fractionally integrated autoregressive conditional heteroscedasticity (FIGARCH) model which relates to financial volatility dynamics.

The phenomenon of long-range dependence asset returns has further improved the model specification in volatility modelling which previously concentrate on short memory basis. Moreover, these underlying predictability components have led to new implications[5,8,9] in efficiency market hypothesis (EMH) which stated that the future returns are unpredictable by using information on past returns. Another important application derives from the return's volatility is the measurement of value-at-risk (VaR) in risk management[10,11]. The VaR implementation is also recommended by international institutions such as the Bank For International Settlements, the American Federal Reserve Bank and the Securities and Exchange Commission for any derivatives market participants. The Kuala Lumpur stock exchange (KLSE), which is our focus, has received great attentions from researchers and investors as the empirical case studies country and potential investment alternatives in South East Asia. Cajueiro and Tabak[12] studied the long memory volatility of KLSE from 1992 to 2002 and found that the Hurst's parameter with the value of 0.628. Cheong et al.[3] investigated the asymmetry and long-memory volatility behaviour of the KLSE daily data over a period of 1991-2005 with four sub-periods. They fitted the asymmetry long memory GARCH models across the periods and the results shown the mixture of symmetry and asymmetry GARCH modelling. As a whole, the above literatures are mainly concentrated on the daily closing price indices. For further analysis, we have selected four volatility proxies and directly implement them in the ARFIMA models. On the other hand, we also consider the fractionally integrated GARCH by Baillie[1,13] for comparison in model specification evaluations. A battery of statistical tests has been employed to diagnose the model specifications. As a result, the model with absolute returns shows superior in in-sample estimation and forecasting evaluations as compare to other models.

MATERIALS AND METHODS

Data Source: The index transaction prices begin from 1st January 2000 until 30 November 2005 (1445 observations) in our empirical study. During this period, the Malaysia stock market was speculated by the RM-USD un-pegged regulation (implemented at year 2006 where the RM was expected undervalued by approximately 6.5%), the merged of MESDAQ in KLSE beside the Main board and Second board previously started in year 2002, the fluctuating of petrol prices, etc. We intended to study the stock market volatility and the reactions of market participants respected to good and bad events.

ARFIMA proxies volatility models: Absolute return and squared returns are the two volatility proxies commonly used in empirical financial time series analysis. Due to its availability and simplicity, most empirical finance literatures using this daily return (squared returns/residual) as the measure of latent volatility. Ding *et al.*[2] claim that the absolute return exhibits consistently higher long memory behaviour than squared returns in S and P 500. Ding and Granger[14] further examines world wide stock markets and foreign exchange and find similar results.

Another alternative measurement of volatility is using the range, the logarithm difference between the highest and lowest prices. Most of the free distributed financial time series, such as newspaper and website, are provided with the information of closing, opening, high and low stock prices, currencies, interest rate, etc. We focus on the range-based volatility proxies include the earliest work by Parkinson[15] and Garman and Klass[16] with the assumption of expected return equal to zero. The mean return is not statistically different from zero at 5% level under the t-test (t-statistic 1.7521). The above studies have proven that the Parkinson[15] approach is five times efficient than the classical average volatility estimator. Both the volatility estimators are define as:

$$\sigma^2_{t,park} = \frac{1}{4\ln 2}\left(H_t - L_t\right)^2 \tag{1}$$

$$\sigma^2_{t,GK} = 0.511\left(H_t - L_t\right)^2 \\ -0.019\left(C_t(H_t - L_t) - 2H_tL_t\right) - 0.383C_t^2 \tag{2}$$

where the definitions are follow Yang and Zhang[17] with H_t, L_t and C_t represent the normalized high, low and closing prices respectively.

Our volatility model appears to be fractionally integrated and correlated with the lagged return (risk premium), negative returns (leverage effect) and heavy-tailed distributed respectively. For risk premium analysis, if the return-volatility poses a positive relationship, we assume that for a more volatility (riskier) securities, the rational market participants require a greater risk premium. On the other hand, if the relationship is negative, it implies that the market participants are more favourable in saving. Whereas the asymmetric effect (leverage effect) implies that impact of bad news increases volatility is relatively deeper than good news. The model can be estimated by using the

maximum likelihood estimation in the following ARFIMAX model:

$$(1-B)^d.1-\beta_l(B)) \\ = c_o + c_1 r^-_{t-1} + c_2 r^-_{t-1} + (1+\alpha_l(B))a_t, \tag{3}$$

where r^-_{t-1} indicates r_{t-1} when $r_{t-1}<0$ and is zero otherwise. While coefficient, c_1 indicates the presence of risk premium. The conditional volatility component is applicable only if the estimated volatility performs the presence of further volatility components[18]. The shock term, a_t, follows a conditional time-varying variance and the $\varepsilon_t \sim$ iid, N $(0,\sigma^2)$ or t-dist (v).

Fractionally Integrated GARCH: The conditional mean equations of KLSE stock returns are an AR (1) model of (r_t) as below:

$$r_t = \theta_0 + \theta_1 r_{t-1} + a_t. \tag{4}$$

The infrequent trading that often occurs in emerging market can be adjusted[19] by a first order autoregressive model to avoid the spurious correlation in the conditional volatility models. For conditional variance specifications, the a_t is serially uncorrelated, but dependent to its lagged values or the conditional variance components as follow:

$$a_t = \sigma_t\varepsilon_t$$

where

$$\varepsilon_t \sim \text{iid, t-dist (v)} \tag{5}$$

β (B) and α (B) capture the shot-run of volatility while the fractional difference parameter d models the long run characteristics of volatility with the intraday information is as follow:

$$\sigma^2_t = \frac{\alpha_0}{1-\beta(B)} + \left\{1 - \frac{\varphi(B)(1-B)^d}{1-\beta(B)}\right\}a^2_t \tag{6}$$

with $0 \le d \le 1$. When $0 < d < 0.5$, the term $(1-B)^d$ has an infinite binomial distribution for non-integer powers. Finally, Tse[20] includes the new impact component and introduces the FIAPARCH (p,d,q):

$$\sigma^\delta_t = \frac{\alpha_0}{1-\beta(B)} + \left\{1 - \frac{\varphi(B)(1-B)^d}{1-\beta(B)}\right\}(|a_t| - \gamma a_t)^\delta \tag{7}$$

Table 1: Descriptive statistics of volatility proxies

|	$|r_t|$	$\log|r_t|$	r^2_t	$\log r^2_t$	$\sigma^2_{t,park}$	$\log \sigma^2_{t,park}$	$\sigma^2_{t,GK}$	$\log \sigma^2_{t,GK}$
Mean	0.6977	-0.8586	0.9784	-1.7173	0.6045	-1.0866	0.6643	-1.0333
Maximum	6.3422	1.8472	40.2235	3.6945	18.6944	2.9282	13.6433	2.6132
Minimum	0.0014	-6.5693	0.0000	-13.1386	0.0146	-4.2243	0.0177	-4.0344
Std. Dev.	0.7014	1.1693	2.6642	2.3387	0.9973	1.0328	1.0788	1.0700
Skewness	2.7549	-1.1783	8.3041	-1.1783	7.6903	0.2772	5.4592	0.3006
Kurtosis	15.4275	5.6188	96.5083	5.6188	103.4999	3.0075	45.5174	2.9084
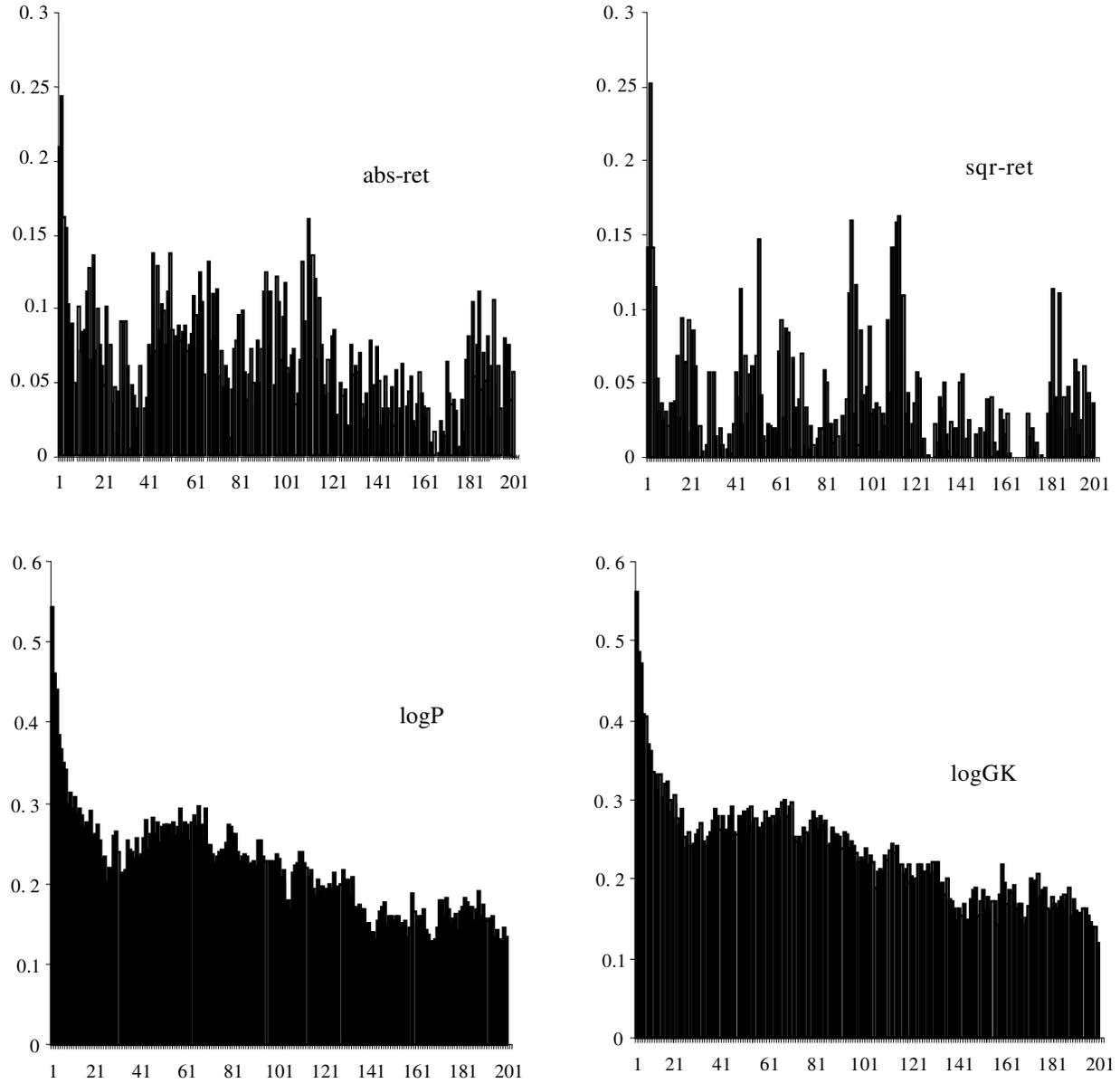
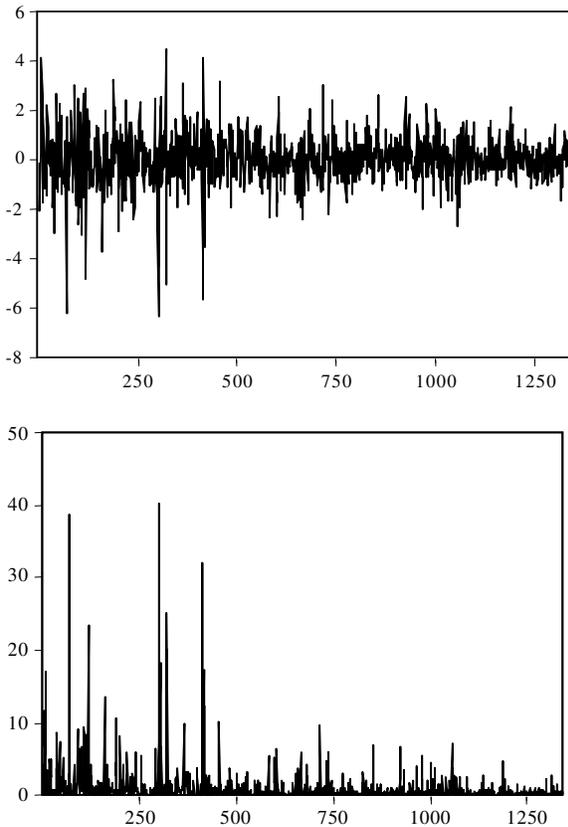Fig. 1: Sample autocorrelation function

Fig. 2: return and squared-return

where the δ and γ couple the possible varying power of standard deviation and leverage effect respectively.

The forecasting evaluations are base on mean squared error (MSE), mean error (ME), mean absolute error (MAE) and Mincer-Zarnowitz Regression. For Mincer-Zarnowitz regression, σ_t^2 is the proxy of the (realized volatility) for time period t and $\hat{\sigma}_t^2$ is the forecasted conditional variance for time t. The realized volatility is measures according to Chin *et al.*[3] approach.

RESULTS AND DISCUSSION

Descriptive statistics: Table 1 reports the descriptive statistics for all the volatility proxies. For $|r_t|$ and r_t^2 indicate serious excess kurtosis with the values 15.42 and 96.51 compare to a standard normal distribution with kurtosis 3. A parametric t-distribution is in favor to normally distribution assumption in the model specifications. On the other hand, both the range-based volatility estimators (after natural logarithm) show similar statistical property with normal distribution with unity standard deviation and kurtosis of 3 respectively. Therefore, this preliminary statistics suggest

estimations under normality assumption for both the range-based volatilities.

Long-range dependences: In Figure 1, the long-range dependences are observed in all the volatility proxies where the sample autocorrelation function indicates significant spikes even after long lags. All the figures also obviously show that range-based volatilities are far more long persistence as compare to transformed-returns volatility. This property can be further examined in the model estimations.

Estimation and diagnostic: Table 2 reports the estimation results of ARFIMAX and ARCH-type with the assumption of Gaussian or student-t distributions. Firstly, we examine the ARFIMAX (0,d,1) models and find that all the coefficient (except r_t^2) if the return-volatility poses a positive relationship, we assume that for a more volatility (riskier) securities, the rational market participants require a greater risk premium, are significantly different from zero at 5% level of significant. These findings suggest that the Malaysian stock asset returns display a tendency to be negatively correlated with the returns volatility. Or in short, the leverage effect indicates that stock market volatility tends to rise in response to bad news and responses reverse to good news. For positive return-volatility relationship, in fig. 2 we conclude that for a more volatility (riskier) securities, the rational market participants require a greater risk premium in their investment for all the proxies except r_t^2. The $|r_t|$ and r_t^2 are significant imply that both the proxies are excess kurtosis and exhibit heave-tailed compare to normally distributed range-based volatilities. For parameter d_{ARFIMA}, all the values are fall in the range of 0<d<0.5 which indicates the presence of long persistence volatility. The strength of long persistence is strongest in $\log \sigma_{t,GK}^2$ (0.4182), follows by $\log \sigma_{t,park}^2$ (0.4003), $|r_t|$ (0.083) and lowest in r_t^2 (0.0452) respectively. It is also worth notes that only $|r_t|$ exhibits the ARCH-effect in the standard residual of ARFIMA.

For ARCH-type models, the AR (1) t-distributed asymmetry power GARCH estimations are reported in Table 2. The δs are close to 2 suggest the conditional variance for the KLCI. On the other hand, the γs indicate similar results in ARFIMA models with leverage effect. The fractional parameter d_{GARCH} is statistically significant with the value 0.2388 also indicates identical inference of long-range dependence volatility. In Table 2, the $|r_t|$ ARFIMA (0,d,1)-GARCH (1,1) provides a slightly statistical improvement over other models based on Akaike and Schwarz information criterion (AIC and BIC). In Table 2, the Ljung-Box

Table 2: maximum likelihood estimation and model diagnostic

Estimation	ARFIMA					ARCH Type	
	$\lvert r_t \rvert$	r_t^2	$\log \sigma_{t,park}^2$	$\log \sigma_{t,GK}^2$		AR(1)-APARCH	AR(1)-FIAPARCH
c_0	0.0177(0.029)	0.0396(0.390)	-0.6752(0.009)	-0.7316(0.002)	c_0	-0.0055(0.824)	-0.0042(0.864)
$c_{1,\,retlag}$	0.9288(0.000)	-0.0722(0.330)	0.1076(0.045)	0.2162(0.000)	α_I	0.1670(0.000)	0.1630(0.000)
$c_{2,retneglag}$	-1.8906(0.000)	0.1072(0.442)	-0.2876(0.002)	-0.4901(0.000)			
α_I	-0.9135(0.000)	0.0465(0.291)	0.1636(0.004)	-0.1911 (0.000)	α_0	0.0639(0.000)	0.1567(0.107)
d_{ARFIMA}	0.0632(0.083)	0.0452(0.000)	0.4003(0.000)	0.4182(0.000)	α_{II}	0.1594(0.000)	-0.3621(0.438)
					β_{II}	0.7981(0.000)	-0.2857(0.572)
α_0	0.0692(0.001)	5.0900(0.324)	0.2662(0.580)	0.3233(0.582)	γ	0.2124(0.015)	0.1951(0.032)
α_{II}	0.1884(0.000)	6.5479(0.329)	0.0222(0.482)	0.0214(0.475)	δ	1.4875(0.000)	1.9454(0.000)
β_{II}	0.6949(0.000)	0.5420(0.000)	0.5968(0.398)	0.4470(0.647)	d_{GARCH}	0.2388(0.000)	
υ	2.9801(0.000)	2.0086(0.000)		51.4487(0.395)	υ	5.7000(0.000)	6.0147(0.000)
L	-1053	-1523	-1660	-1568		-1686	-1678
AIC	1.5858	2.2876	2.4910	2.3532		2.5276	2.5161
SIC	1.6208	2.3226	2.5220	2.3842		2.5586	2.5510
Diagnostic							
(1) Q-(12) on a_t	7.2009(0.782)	1.7258(0.999)	12.0312(0.361)	16.7476(0.115)		17.5587(0.092)	16.2388(0.132)
(2) Q-(12) on $\tilde{a}_t^{\,2}$						8.3195(0.597)	4.1171(0.941)
(3) ARCH test						0.6837(0.768)	0.4073(0.961)
Forecasting evaluation							
MSE	0.0844	0.1805	0.4235	0.391		0.03607	0.0219
ME	-0.0195	0.0694	-0.1306	-0.0974		-0.1198	-0.0335
MAE	0.2109	0.1812	0.5177	0.5053		0.1573	0.1111
Theil	0.3654	0.7190	0.1619	0.1583		0.2539	0.2244
MZ							
a	-0.0231(0.876)	0.1750(0.836)	-0.3430(0.663)	-0.6797(0.617)		-0.0079(0.932)	0.0065(0.948)
b	1.1121(0.044)	0.0081(0.008)	0.9228(0.427)	1.6124(0.479)		0.7243(0.004)	0.8743(0.009)
R^2	0.2023	0.2263	0.0070	0.0156		0.2129	0.2195

NOTES: a_t and $\tilde{a}_t$ represent the residual and standardized residual respectively, (1) and (2): Ljung Box Serial Correlation Test (Q-statistics) on $\tilde{a}_t$ and $\tilde{a}_t^2$: Null hypothesis – No serial correlation, (3) LM ARCH test: Null hypothesis - No ARCH effect, (4)Forecasting evaluation, Mincer-Zarnowitz (MZ) regression, $\sigma_t^2 = a + b\hat{\sigma}_t^2 + u_t$, where σ_t^2 : realized volatility with 20-minute interval and $\hat{\sigma}_t^2$: forecasted conditional variance for time t. The values in the parentheses represent the p-value.

statistics indicate that all serial correlations in the error terms are not significant at 1% significant level. For ARCH-type models, the asymmetry FIAPARCH shows better diagnostic results compare to APARCH model in the Ljung-Box standardized residual correlation test. For squared standardized residuals, both the tests exhibit no significant serial correlations and ARCH effect at 1% level of significance. Finally, Table 2 presents 100 one-step-ahead daily volatility forecasts for the KLSE stock index. The fractionally integrated ARCH-type models show smallest MSE and MAE while the $\lvert r_t \rvert$-ARFIMA indicates relative lesser in term of magnitude for ME. In the regression analysis, the proxies $\lvert r_t \rvert$, r_t^2 and FIAPACH show the highest R^2 value (0.2023 to 0.2263) compare to the range-based volatility models with insignificant bs. However, the $\lvert r_t \rvert$ exhibit superior in the regressor coefficient, b, (1.1121) which are nearest to unity compare to other models.

As a result, in terms of model fitting and specifications, we conclude that the $\lvert r_t \rvert$, ARFIMA (0,d,1)-GARCH model is out-performance compare to all the conditional and range-based volatility models.

CONCLUSION

This study investigates the fractionally integrated behaviour of KLSE's volatility over the year 2000 to year 2005. During this recovery period, the stock market exhibits the presence of long memory volatility, presence of risk premium, significant relation between news (lagged return) and volatility which implied that bad news have the higher predictive power for upcoming volatility and finally the existence of heavy-tailed property in selected FI volatility models. As a conclusion, the empirical results show that the inclusion of time-varying volatility in the absolute return

ARFIMA model contributes a better in-sample estimation as well as one-day-ahead forecasting. For further research, we are interested to use the estimated volatility in the Value-at-risk application for long and short trading analyses.

ACKNOWLEDGEMENTS

REFERENCES

1. Baillie, R.T., T. Bollerslev and H.O. Mikkelsen, 1996. Fractionally integrated autoregressive conditional heteroscedasticity. J. Econometrics, 74: 3-30.
2. Ding, Z., C.W.J. Granger and R.F. Engle, 1993. A Long Memory Property of Stock Market Returns and a New Model. J. Empirical Finance, 1: 83-106.
3. Cheong C.W., S.M.N. Abu Hassan and I. Zaidi, 2007. Asymmetry and long memory volatility: some empirical evidence using GARCH. Physica A: Statistical Mechanics and its Applications, 373: 651-664.
4. Granger, C.W.J. and R. Joyeux, 1980. An introduction to long memory time series models and fractional differencing. J. Time Series Analysis, 1: 15-29.
5. Mandelbrot, B., 1997. Fractal and scaling in finance: Discontinuity, concentration, risk. New York, Springer.
6. Granger, C., 1980. Long memory relationships and the aggregation of dynamic models. J. Econometrics, 14: 227-238.
7. Hosking, J., 1981. Fractional differencing. Biometrika, 68: 165-176.
8. Peters, E.E., 1994. Fractal Market Analysis. A Wiley Finance Edition, John Wiley and Sons, New York.
9. Mullier, U., M. Dacorogna, R. Dav, O. Pictet, R. Olsen and J. Ward, 1993. Fractals and intrinsic time-a challenge to econometricians. XXXIXth International AEA Conference on Real Time Econometrics, pp: 14-15.
10. Jorion, P., 2000. Value-at-risk: The New Benchmark for Manging Financial Risk. New York: McGraw-Hill.
11. Morgan, J.P., 1996. RiskMetricsTM-Technical Document, New York.
12. Cajueiro, D.O. and B.M. Tabak, 2004. The Hurst's exponent over time: testing the assertion that emerging markets are becoming more efficient. Physica A, 336: 521-537.
13. Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. J. Econometrics, 66: 47-78.
14. Ding, Z. and C.W.J. Granger, 1996. Modelling volatility persistence of speculative returns: a new approach. J. Econometrics, 73: 185-215.
15. Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Business, 53: 61-65.
16. Garman, M. and M. Klass, 1980. On the estimation of security price volatilities form historical data. J. Business, 53: 67-78.
17. Yang, D. and Q. Zhang, 2002. Drift independent volatility estimation based on high, low, open and close prices. J. Business, 73: 477-491.
18. Corsi, F., U. Kretschemer, S. Mittnik and G. Pigorsch, 2005. The volatility of realized volatility. Working paper Center for Financial Studies, 2005/33.
19. Tse, Y.K., 1998. The Conditional Heteroscedasticity of the Yen-Dollar Exchange Rate. J. Appl. Econometrics, 193: 49-55.
20. Miller, M.H., J. Muthuswamy and R.E. Whaley, 1994. Mean reversion of Standard and Poor 500 index basis changes: Arbitrage-induced or statistical illusion. J. Finance, 49: 479-513.