A Generalized Discrete-Time Long Memory Volatility Model for Financial Stock Exchange

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

Abstract: We proposed a simple way to combine a few long memory models in financial market volatility modeling using daily, range and high frequency data. This model was able to fit the return, range of daily return or realized volatility under a parametric heavy-tailed distribution. Model was flexible to include additional volatility information as the contemporaneous variables. Empirical results found that the proposed model provides substantial improvement in the model fitting, specification and most importantly, a better out-of sample forecasting in the Malaysian stock market.

Keywords: realized volatility, fractionally integrated, autoregressive conditional heteroscedastic (ARCH).

INTRODUCTION

The time varying conditional volatility ARCH-GARCH models[1-2] are widely used in financial time series analysis such as risk management, portfolio analysis and derivative pricing. The presence of long persistence volatility in asset returns[3-4] has further improved the model specification in volatility modelling. Both the ARCH effect and long persistence volatility have enriched the definitions of efficiency market hypothesis[5]. A series of studies[6-7] investigate the market efficiency by analyzing the degree of Hurst's parameter[8] for different financial markets. With the evidence of long persistence property, some extension definitions of market efficiency hypothesis have been proposed such as fractal market hypothesis[9] and heterogeneous market hypothesis[10]. The most prominent application of volatility model is the measurement of value-at-risk(VaR) in risk and portfolio analysis. VaR measurement is directly related to the expected volatility over the relevant time horizon such as daily, weekly and etc. The commercial application of VaR by Morgan[11] with his risk management model, *RiskMetrics*[TM], has successfully applied VaR in portfolio investments. Hence, a volatility model with correct specification is very important in the market efficiency and risk analysis studies.

Besides the model specification issue, the appropriate data handling approach also contributes to the development of volatility modelling. The most common approach is based on the closing prices which daily returns are subsequently calculated. Due to its availability and simplicity, most of the empirical finance literatures using this daily return (squared returns/residual) as the measurement of latent volatility such as GARCH and stochastic volatility models. However, as indicated in[12], this estimation of volatility can be very noisy. Another alternative measurement of volatility is using the range, the logarithm difference between the highest and lowest prices. The literatures that focus on range-based volatility proxy include the earliest work by Parkinson[13] and Garman and Klass[14]. However, the latent volatility is also not directly measurable with the usage of daily information. Hence, it is hard to identify the outperform models in terms of model fitting performance and specification adequacy. The vast development of recent information and communication technology(ICT) has encouraged the usage of high frequency data as an observable proxy for latent volatility which facilitates a more accurate estimation as well as forecasting. Moreover, the availability of high frequency financial assets information such as stocks, stock indexes, foreign exchange, etc. has facilitated a more accurate estimation as well as forecasting. The study of observable(realized) volatility using high frequencies asset prices are discussed in[15]. The above study has motivated the usage realized volatility in both FX and equity markets due to the remarkably good out-of-sample forecasting.

In this paper, we have selected the Kuala Lumpur stock exchange(KLSE) index transaction prices during the recovery period 1st January 2003 to 15 January 2006 (745 and 266710 observations for low and high frequency data respectively). The Malaysian stock

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

market had experienced a massive slid due to the major factor of drastic depreciation of Malaysian Ringgit(RM) where the RM(RM2.50 in first half of 1997) was depreciated to the weakest rate recorded against USD as RM4.88 on 9th January, 1988. Malaysian government had implemented the one USD pegged to RM3.80 in the 1st September 1998 to stabilize the RM from currency speculation. After the implementation, the Malaysian economics show significant recovery. In our empirical study of recovery period, the Malaysia stock market was speculated by the RM-USD un-pegged regulation(implemented at year middle of 2005 where the RM was expected undervalued by approximately 6.5%), the merged of MESDAQ (http://www.klse.com.my) in KLSE besides the Main board and Second board previously started in year 2002, the fluctuating of petrol prices, etc. We intended to study the long persistence volatility, the presence of risk premium and how the market participants react with respect to good and bad events using the low and high frequency data.

In order to account the stylized empirical facts of the volatility, we proposed a combination of ARFIMAX, HAR and ARCH-type generalized models. The model can flexibly include the additional volatility information (such as weekly, monthly) as the contemporaneous variables. A battery of statistical tests has been employed to diagnose the model specifications. As a result, the proposed model shows substantial improvement in in-sample estimation compare to other models. For forecasting performance evaluations, the proposed model with the inclusion of additional contemporaneous variable is outperformed compare to other models.

MATERIALS AND METHODS

Data Source: In our empirical study, the KLSE index transaction prices are obtained from the DATA STREAM for 745 daily closing prices. On the other hand, the high frequency data1 is available from the data vendor which consists of 266710 data points. The availability of KLSE high frequency (minutely) data is not as popular as other mature market such as S&P 500, NIKKEI 250, CAC40, etc. where minutely data can be obtained up to 15 years start from 1991. The original dataset includes prices for every trade for 1-minute interval data. However, we further extract the 5 and 20-minute intervals in order to study the statistical behaviour of the emerging market. As a comparison to the mature market such as S&P500 etc., we like to study whether the most common 5-minute interval is

also suitable for the Malaysian emerging stock market. The interday return series, r_t is defined as the close-to-close prices on consecutive trading days. The percentage interday returns can be expressed as:

$$r_t = 100\left(\ln P_{t,close} - \ln P_{t-1,close}\right) \qquad (1)$$

The proposed HARFIMA-GARCH model: In this paper, we proposed a flexible volatility model which is able to form the ARCH-type, range-based and realized volatility model specifications. In addition, this model is capable to include the possible past realized volatility with different frequencies such as weekly or monthly information as the contemporaneous variables. In short, the proposed model integrates the ARFIMAX, HAR and FIGARCH models in general. The maximum likelihood estimation uses the iterative optimization algorithm to determine the second derivatives (Hessian matrix) under the standardized t-distribution with $\upsilon>2$ degree of freedom. The HARFIMA(f_k,p_I,d_I,q_I)-FIGARCH(p_{II},d_{II},q_{II}) can be expressed as follows:

$$(1-B)^{d_I}(1-\beta_I(B))(y_{I,t})$$
$$= c_o + c_1 r_{t-1}^+ + c_2 r_{t-1}^- + \sum_k\left(c_{h,k}\left(\frac{1}{n}\right)\sum_{i=1}^n y_{I,t-n,k}\right) + (1-\alpha_I(B))a_t,$$

$$a_t = y_{II,t}\varepsilon_t,$$
$$(1-B)^{d_{II}}(1-\beta_{II}(B))y_{II,t}^2 = \alpha_0 + \alpha_{II}(B)a_t^2 . \qquad (6)$$

where

- $y_{I,t}$ and $y_{II,t}$ represent either the return, range-based volatility, return's volatility or volatility of realized volatility series according to a specific model:

  -For ARCH-type model: the $y_{I,t}$ and $y_{II,t}$ represent the r_t and σ_t^2 respectively;

  -For range-based volatility model: the volatility estimator is adopted from the Parkinson[13] and Garman and Klass[14] approaches with the assumption of the expected return is equal to zero where the mean return is not statistically different from zero at 5% level under the *t-test*(t-statistic 1.7521);

  -For realized volatility model, Martens[15] suggested to use a scaled sum of squared intraday returns to represent the scaled realized volatility as follow:

$$RV_t = (1+c)\left(\sum_{a=1}^{\frac{\tau_M}{n}}\left(R_{t,a}^M\right)^2 + \sum_{b=1}^{\frac{\tau_A}{n}}\left(R_{t,b}^A\right)^2\right); \qquad (5)$$

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;

- r_{t-1}^{+} (r_{t-1}^{-}) indicates r_{t-1} when $r_{t-1}>0$($r_{t-1}<0$) and is zero otherwise. The present of r_{t-1}^{+} (r_{t-1}^{-}) suggests that the financial stock market volatility tends to rise in response to good (bad) news and responses reversely to bad (good) news;

- c_1 and c_2 denote the risk-return tradeoff. If the return-volatility poses a positive relationship, we assume that for a more volatile(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 favorable in saving;

- $\alpha_I(B)$, $\beta_I(B)$, $\alpha_{II}(B)$ and $\beta_{II}(B)$ denote the stationary finite polynomials in the backshift or lag operators with $B^k y_t = y_{t-k}$ where in general:

$$\alpha(B) = 1 - \alpha_1 B - ... - \alpha_p B^p;$$
$$\beta(B) = 1 - \beta_1 B - ... - \beta_q B^q; \tag{7}$$

- d_I and d_{II} represent the long memory parameters with the constraint vary to their respective models. The $(1-B)^d$ can be expressed as an infinite binomial distribution for non-integer powers:

$$(1-B)^d = \sum_{x=1}^{\infty} (-1)^x \binom{d}{x} B^x . \tag{8}$$

The HARFIMA(f_k,p_I,d_I,q_I)-FIGARCH(p_{II},d_{II},q_{II}) model includes other models as special cases:

- The ARFIMAX of Granger (1980) when $c_{h,k}$ =0 (*k*= weekly and monthly) with no conditional heteroscedastic volatility;

- The Heterogeneous AutoRegressive with time varying volatility model[16] when $d_I = c_1 = c_2 = 0$; The $\sum_{k}\left(c_{2,k}\left(\dfrac{1}{n}\right)\sum_{i=1}^{n}\ln RV_{t-n,k} \right)$ represents the contemporaneous realized volatility with the coefficient, $c_{h,k}$, which indicates the significant influences of possible different frequencies of past realized volatility with the k = lagged daily, weekly and monthly and its associates n =1, 5 and 20 respectively;

- The joint ARMA-GARCH[2] and ARMA-FIGARCH[3] respectively when $c_{h,k}$=0 (*k*= daily, weekly and monthly), $d_I = c_1 = c_2 = 0$; $0 < d_{II} < 0.5$ to ensure a stationary ARMA-FIGARCH. The shock term, a_t, follows a conditional time-varying variance and the ε_t~iid, with certain specific parametric distributions;

The ARFIMA-FIGARCH when $c_{h,k} = c_1 = c_2 = 0$. Davidson[17] argued that the d_I in ARFIMA is structural different from d_{II} where the persistence is increases when d_I approaches 0.5 compare to d_{II} approaches 0. The reverse behaviour may be due to the parameter acts directly on the squared errors but not on the conditional variance.

One-day-ahead Forecasting evaluation: The out-of-sample forecasting is evaluated by using some standard statistical measurements such as mean squared error(MSE), mean error(ME), mean absolute error(MAE) and Mincer-Zarnowitz Regression[18]. For Mincer-Zarnowitz regression, σ_t^2 is the proxy of the actual volatility (realized volatility) for time period t and $\hat{\sigma}_t^2$ is the forecasted conditional variance for time t. The simple linear regression model is illustrated as follow:

$$\sigma_t^2 = a + b\hat{\sigma}_t^2 + u_t \tag{13}$$

Conditioning upon the forecast, the forecast is unbiased and optimal only if a=0 and b=1. The determinant coefficient, R^2 indicates the power of predictability of the selected models.

RESULTS AND DISCUSSION

The results in Table 1 show that all the standard deviations for the realized volatility, $RV_{t,I}$ are slightly greater than $RV_{t,II}$ for all the selected frequencies which explained that with the presence of lunch-break and over-night effects, the realized volatility become more noise. Due to this, we stick to the $RV_{t,II}$ results for further analysis. The scaling value ($1+c$) is approximately 1.111 which is lower than the result reported by Marten[15] of 1.205 in S&P500 future index series.

Table 1: Break and Scaled Realized volatility

RV	$RV_{t,I}$ 1-min	$RV_{t,II}$ 1-min	$RV_{t,I}$ 5-min	$RV_{t,II}$ 5-min	$RV_{t,I}$ 20-min	$RV_{t,II}$ 20-min
Mean	0.5762	0.5952	0.3968	0.3874	0.3884	0.3835
Std. Dev.	0.3869	0.2945	0.3423	0.1999	0.4741	0.3925
Skewness	5.2064	2.3162	6.7664	2.1395	5.5368	4.5374
Kurtosis	46.2392	17.8137	66.1134	11.3014	45.1322	35.5703
(1+c)		1.113		1.111		1.111

Table 2: Range-based and scaled realized volatility estimators

	daily				Realized volatility						Range-based volatility			
	r_t^2	$RV_{t,1}$	$ln(RV_{t,1})$	$\sqrt{RV_{t,1}}$	$RV_{t,5}$	$ln(RV_{t,s})$	$\sqrt{RV_{t,5}}$	$RV_{t,20}$	$ln(RV_{t,20})$	$\sqrt{RV_{t,20}}$	$\sigma^2_{t,park}$	$\ln \sigma^2_{t,park}$	$\sigma^2_{t,GK}$	$\ln \sigma^2_{t,GK}$
Mean	0.4076	0.5952	-0.6315	0.7508	0.3874	-1.0605	0.6052	0.3835	-1.2669	0.5663	0.3120	-1.5905	0.2962	-1.6140
Std. Dev.	0.7781	0.2945	0.5007	0.1777	0.1999	0.4773	0.1456	0.3925	0.7090	0.2213	0.3963	0.8769	0.3848	0.8487
Skewness	4.3502	2.3162	-1.1662	0.6495	2.1395	-0.2507	0.9277	4.5374	0.2578	1.8261	4.6944	0.3829	5.9270	0.2827
Kurtosis	28.3303	17.8137	10.8317	5.6151	11.3014	4.9507	5.2998	35.5703	3.5384	9.4183	34.6527	3.0387	53.5148	3.3209
JB test	21699*	7286*	2019*	257*	2638*	122*	264*	49023*	16*	1649*	32883*	17.74*	81216*	12.75*

* denotes 1% level of significance

Table 3: Hurst's parameter estimations

	r_t^2	$\ln \sigma^2_{t,park}$	$\ln \sigma^2_{t,GK}$	$RV_{t,20}$	$ln(RV_{t,20})$	$\sqrt{RV_t,20}$
Variance –Time plot*	0.53	0.62	0.60	0.64	0.71	0.69
R/S**	0.59	0.73	0.76	0.70	0.74	0.74
average	*0.560*	*0.675*	*0.680*	*0.660*	*0.725*	*0.715*

*For the aggregated time series r(n) of a self-similar process, the variance obeys the following large sample property:

$$V[r^{(n)}] \sim \frac{V[r]}{n^\beta}, \text{ where } r_k^{(n)} = \frac{1}{n} \sum_{i=kn-(n-1)}^{kn} r_i \text{ and the self-similarity}$$

parameter H = 1- (β/2).

**The ratio of R/S is defined as:

$$\frac{R}{S} = \frac{\max_{1 \le j \le N}\left[\sum_{k=1}^{j}(r_k - M(L))\right] - \min_{1 \le j \le N}\left[\sum_{k=1}^{j}(r_k - M(L))\right]}{\sqrt{\frac{1}{N}\left[\sum_{k=1}^{N}(r_k - M(L))^2\right]}}$$

where M(L) is the sample mean over the time period L. The Hurst's parameter is determined by: R/S ~ (L/2)**H**,
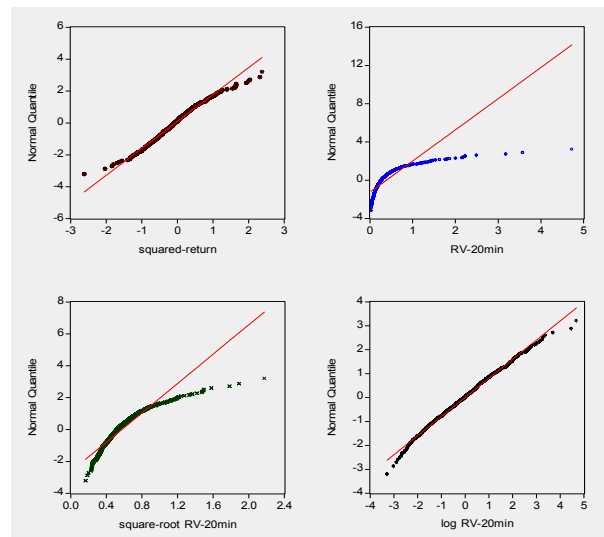

Fig. 1: Q-Q plots for volatility estimators

We have selected three properties of realized volatility namely the normality, the long persistence behavior and the risk premium to study in this session. Table 2 shows the squared-return (r_t^2) for daily volatility, range-based volatility estimators, $\sigma^2_{t,park}$ and $\sigma^2_{t,GK}$, and 1, 5 and 20-minute interval of scaled realized volatility in the form of variance (RV_t), logarithm ($lnRV_t$) and standard deviation ($\sqrt{RV_t}$). The series exhibit nearest to normality is chosen in the models estimations. At a glance, all the series show asymmetry, excess kurtosis and the Jarque-Bera normality tests are insignificant at 1% level. The $\sigma^2_{t,park}$, $\sigma^2_{t,GK}$ and logarithm 20-minute interval indicates the most closely to a Gaussian distribution with the kurtosis close to 3 (3.0387, 3.3209, 3.5384 respectively) and skewness nearly zero (0.3829, 0.2827,0.2578 accordingly). We further investigate the Q-Q plots for squared-returns and the 20-minute interval that are illustrated in Fig 1.

The $lnRV_t$ indicates approximate a linear line compare to return-squared which is skewed to the right. Our results are different from the work by Andersen and Bollerslev[12] which suggested 5-minute interval in FX. In our study, the dissimilar outcome might cause by the low trading in the emerging market. In addition, [19] shown that 25-minute and 15-minute returns provided the optimal sampling frequency in their studies.

Table 4: estimation results

Estimation	ARFIMAX	HARX-GARCH	HARFIMAX-GARCH	ARFIMAX-FIGARCH
			Realized volatility model	
c_0	-0.5721[c](0.000)	-0.2702 [c] (0.000)	-0.3763 [c] (0.000)	-0.5692 [c] (0.000)
d_I	0.2647 [c] (0.000)		0.1293 [a] (0.094)	0.2648 [c] (0.000)
c_1	0.0218(0.346)	0.0354(0.129)	0.0308(0.202)	0.0179(0.437)
c_2	-0.0661 [b] (0.013)	-0.0821 [c] (0.002)	-0.0768 [c] (0.004)	-0.0654 [b] (0.012)
α_I				
α_0		0.1253 [b] (0.014)	0.1290 [b] (0.011)	0.1120 [c] (0.009)
d_{II}		-0.1182(0.121)	-0.1216(0.100)	-0.0952(0.183)
α_{II}		0.3389 [a] (0.055)	0.3343 [a] (0.070)	0.3132(0.185)
β_{II}		0.1199(0.399)	0.1134(0.450)	0.1233(0.542)
$c_{h,day}$		0.2318 [c] (0.000)	0.1043(0.207)	
$c_{h,week}$		0.2492 [c] (0.000)	0.1703 [b] (0.011)	
$c_{h,month}$		0.0173(0.810)	0.0486(0.622)	
υ	10.4735 [b] (0.012)	13.5236 [b] (0.046)	12.0320 [b] (0.027)	10.9210 [b] (0.032)
$\mathcal{L}$	-111.37	-106.33	-104.78	-107.90
AIC	0.3508	0.3507	0.3491	0.3494
SIC	0.3913	0.4248	0.4299	0.4101
(1) Q-(12) on a_t	15.1091(0.235)			
Q-(120) on a_t	111.13(0.706)			
(2) Q-(12) on a_t^2	16.9897(0.149)			
Q-(120) on a_t^2	142.905(0.0754)			
(3) LM(12) ARCH test	1.4117(0.155)	0.9796(0.466)	0.9206(0.525)	0.9442(0.501)
(4) Q-(12) on $\tilde{a}_t$		21.3861(0.045)	17.7972(0.121)	16.6550(0.163)
(5) Q-(12) on $\tilde{a}_t^2$		12.1478(0.275)	11.3642(0.329)	11.6916(0.306)

(6)BDS test

	ARFIMAX p	ARFIMAX p_b	HARX-GARCH p	HARX-GARCH p_b	HARFIMAX-GARCH p	HARFIMAX-GARCH p_b	ARFIMAX-FIGARCH p	ARFIMAX-FIGARCH p_b
$m=2$								
0.5σ	0.023 [b]	0.042 [b]	0.241	0.296	0.239	0.290	0.121	0.168
1.00σ	0.025 [b]	0.030 [b]	0.111	0.116	0.186	0.178	0.078	0.082
1.50σ	0.013 [b]	0.026 [b]	0.087 [a]	0.102	0.098	0.104	0.047	0.04
$m=5$								
0.5σ	0.290	0.392	0.582	0.736	0.506	0.632	0.271	0.416
1.00σ	0.387	0.372	0.144	0.168	0.172	0.158	0.335	0.398
1.50σ	0.291	0.258	0.090	0.106	0.164	0.182	0.296	0.332

(7)Engle and Ng test

	ARFIMAX	HARX-GARCH	HARFIMAX-GARCH	ARFIMAX-FIGARCH
Sign-bias		1.9072(0.056) [a]	1.4825(0.138)	1.2410(0.214)
Negative size-bias		1.0753(0.282)	0.9785(0.327)	0.9555(0.339)
Positive size-bias		0.5464(0.584)	0.3757(0.707)	0.2815(0.778)

Notes: a, b and c denote 10%, 5% and 1% level of significance. The values in the parentheses represent the p-value. $\tilde{a}_t$ represents the standardized residual.

Table 5: Forecasting Evaluations

	Realized volatility ARFIMAX	HARX-GARCH	HARFIMAX-GARCH	ARFIMAX-FIGARCH	Range-based volatility ARFIMAX-FIGARCH	ARCH-type model AR-GARCH	ARFIMA-FIGARCH
MSE	0.0345	0.0368	0.0352	0.0343	0.5751	0.0073	0.0072
ME	-0.0515	0.0676	-0.0564	-0.0532	-0.1363	0.0071	-0.0017
MAE	0.1445	0.1465	0.1450	0.1436	0.6117	0.0656	0.0658
Theil	0.1328	0.1388	0.1346	0.1322	0.1821	0.2031	0.1967
Mincer-Zarnowitz							
a	-0.1690(0.482)	-0.1531(0.557)	-0.1478(0.563)	-0.1777(0.461)	-0.1202(0.531)	0.0105 (0.945)	0.0199 (0.873)
b	0.8243(0.026) [b]	0.8673(0.034) [b]	0.8640(0.029) [b]	0.8116(0.028) [b]	0.3052(0.002) [c]	0.9828 (0.214)	0.8950 (0.145)
R^2	0.0987	0.0896	0.0944	0.0957	0.1699	0.0321	0.0460

The observable volatility is represented by the realized volatility with 20-minute interval. Notes: a, b and c denote 10%, 5% and 1% level of significance. The values in the parentheses represent the p-value

Next, we exercise the variance-time plot and rescaled-range analysis to examine the Hurst's parameter. In Table 3, the $lnRV_t$ series shows the strongest persistency with the value 0.725, follows by the $\sqrt{RV_t}$ (0.715), RV_t (0.670), follows by the $ln\sigma_{t,GK}^2$ (0.680), $ln\sigma_{t,park}^2$ (0.675) and finally the squared-return (0.560). The long persistence behaviour encourages us to model the realized volatility using a fractionally integrated model.
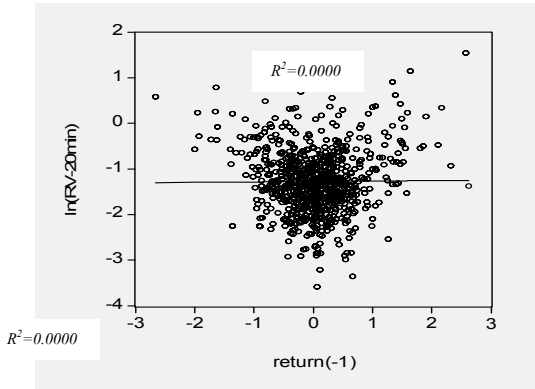
Fig. 2: News impact

Finally, the ARCH-Mean effect and news impact are examined by investigating the scatter plots with regression line for the realized volatility and current and lagged returns respectively. In Fig. 2, the scatters plots display an ordinary least squares(OLS) regression line with a constant and $lnRV_t$. From the plot of 20-minute realized volatility and current return, the R^2 shows insignificant linear relation between the $lnRV_t$ and current return which suggests the absence of ARCH-mean effect. However, for $lnRV_t$ and lagged return, the lagged positive and negative returns yield significant occurrences of high volatility as indicated in Fig. 2. This implied that the news impact may be significant in model estimation. The formal test for risk premium and relation between the news and realized volatility are carried out in the next session.

Maximum likelihood estimation results: Table 4 reports the estimation results of realized volatility, range-based and ARCH-type models with the student-t distributed innovation, ε_t. For all the realized volatility and range based models only the coefficient, c_2, is significantly different from zero which indicated that only the lagged negative returns yield great volatility.

(1), (2), (4) and (5): 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; (6): This table reports the descriptive statistics for all the residual and standardized residual series. All the series exhibit leptokurtic with kurtosis around 3.500 which is biased from normal distribution. Due to this, we calculated both the normality assumption and bootstrapped p-values for the BDS test statistics. Both the p-values show similar results and this is acceptable because our series do not differ a lot from the normal distribution. p and p_b denote the p-value for the BDS test statistic for assumption of asymptotic normal distribution and bootstrapped p-value respectively. (7): Engle and Ng[21] news impact test based on the regression $\tilde{a}_t^2 = a_1 + a_2 S_t^- + a_3 S_t^- a_{t-1} + a_4 S_t^- a_{t-1}^2 + \varepsilon_t$

This leverage effect implied that downward movement (shock) in the stock market is followed by a greater volatility than upward movement of the same magnitude. In addition, the negative value of c_2 implied that the market participants are mostly in the precaution condition or only observed the market movement instead of taking part in the stock market. In

another word, the market participants are preferred to save their money than taking the risk. This may be due to the bad experience or lack of confidence to the volatile market.

All the models fitted well in a heavy-tailed student-t distribution with degree of freedom (υ) approximately 12 except for range-based model. This stylized fact is commonly seen in almost all the financial asset pricing. The four long memory models indicated long memory parameters, ds with the 0.2674, 0.1293, 0.2648 (for the three former realized volatility models) and 0.2842 for the range-based model respectively. For the proposed HARFIMAX-GARCH model, the reduction of the long persistence behaviour is due to the inclusion of weekly volatility information as compare to the original HAR-GARCH with news impact model. This implied that the contemporaneous variables, for some extent but not fully explained the initial volatility model specification.

Overall, the highest log-likelihood value is -104.78 (HARFIMAX-GARCH) with the smallest AIC (0.3491) as compare to others. However, the BIC value for HARFIMAX-GARCH is suffered for greater BIC value 0.4299 as compare to others due to the additional parameters. For ARCH-type models, we have selected AR-GARCH and ARFIMA-FIGARCH models to compare the in-sample estimations as a whole. Both the models show almost similar results in log-likelihood function and estimation performance evaluations. The ARCH-type model indicated better in-sample estimation as compare to range-based model, however, inferior if compare to the realized volatility models.

Diagnostic results: In Table 4, the ARFIMAX's exhibits possibility of conditional heteroscedascity in their squared-residuals and LM-ARCH test and BDS test[20] as well. This implied that the most parsimony ARFIMAX model is not adequately specified with the existence of conditional heteroscedastic effect. On the other hand, the HARX-GARCH suffered from the misspecification in the standardized residual as well as the sign bias test under 5% and 10% significant levels respectively. Except the above mentioned models, Table 3 shown that all the volatility models are not significant at 1% level for the Ljung-Box test, serial correlation and ARCH effect test respectively. The further discussion of the BDS test is illustrated in the footnote of Table 4. On the other hand, the sign-bias and negative/positive tests for all the squared-standardized residual o the studied models show no evidence of unexplained non-linearity, sign or size bias in the negative/positive side at 1% significance level. As a result, the asymmetric models with news effect are adequately estimating the conditional standard deviations.

One-step-ahead forecasting evaluations: The in-sample estimations sometimes do not guarantee outperforms forecasting results. In order to verify this, we conducted a 50 one-step-ahead daily forecast for the

KLSE stock index. The one-day horizon forecast comparisons are based on MSE, ME, MAE and Mincer-Zarnowitz Regression. As indicated in Table 5, the ARCH-type models show smallest MSE, ME and MAE results. This followed by the realized-volatility models and finally the range-based models. However, the Mincer-Zarnowitz regression analysis indicates that the estimated parameter (b) for ARCH-type models is not significantly different from zero. This implied that the ARCH-type models could not be used to predict future volatility. On the other hand, the realized volatility and range-based models show strong rejection of the b parameter and the coefficients of determination, R^2, can be ranked based on the nearest to unity. The HARFIMAX-GARCH shows the highest b and R^2 (0.8690 and 0.0984) as compare to the other models. This implied that the out-of-sample forecasting for HARFIMAX-GARCH is superior as compare to the ran

CONCLUSION

In this paper, we constructed a generalized model which is able to account for time-varying heteroskedasticity, long-persistence and leverage effects in a specific case of Malaysian stock market volatility for low and high frequency data. Our empirical result found that the contemporaneous different time horizons realized volatilities have substantial reduced but not fully eliminated the volatility persistence which may cause by other unknown sources of volatility. In future work, the Value-at-risk (VaR) for long and short trading positions can be determined based on the estimated volatility model.

ACKNOWLEDGEMENTS

REFERENCES

1. Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50: 987-1007.
2. Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31: 307-327.
3. Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics, 66: 47-78.
4. Mandelbrot, B., 1997. Fractal and scaling in finance: Discontinuity, concentration, risk. New York: Springer.
5. Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25: 383-417.
6. Chin W.C., Abu Hassan S.M.N, Zaidi I., 2007. Asymmetry and long memory volatility: some empirical evidence using GARCH, Physica A, 337: 651-664.
7. Cajueiro, D.O., & Tabak, B.M., 2005. Ranking efficiency for emerging market II. Chaos, Solitons & Fractals, 23: 671-675.
8. Hurst, H.E., 1951. Long term storage capacity of reservoirs. Transaction of the American Society of Civil Engineers, 116: 770-799.
9. Peters, E.E., 1994. Fractal Market Analysis (A Wiley Finance Edition). New York: John Wiley & Sons.
10. Mullier, U., Dacorogna, M., Dav, R., Pictet, O., Olsen, R., & Ward, J., 1993. Fractals and intrinsic time-a challenge to econometricians. XXXIXth International AEA Conference on Real Time Econometrics, 14-15.
11. Morgan, J.P., 1996. RiskMetricsTM − Technical Document, New York.
12. Andersen, T.G., & Bollerslev, T., 1998. Answering the skeptics: Yes standard volatility models do provide accurate forecasts. International Economic Review, 39: 885-905.
13. Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return, Journal of Business, 53: 61-65.
14. Garman, M., & Klass, M., 1980. On the estimation of security price volatilities form historical data. Journal of Business, 53: 67-78.
15. Martens, M., 2002. Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. The Journal of Futures markets, 22: 497-518.
16. Corsi, F., Kretschemer, U., Mittnik, S., & Pigorsch, G., 2005. The volatility of realized volatility. Working Paper Center for Financial Studies, 2005/33.
17. Davidson, J., 2001. Moment and memory properties of linear conditional heteroscedasticity models, Manuscript, Cardiff University.
18. Mincer, J. and Zarnowitz, V., 1969. The evaluation of economic forecasts in J. Mincer (ed.), Economic Forecasts and Expectations, NBER, New York.
19. Giot, P., & Laurent, S., 2004. Modeling daily value-at-risk using realized volatility and ARCH type models. Journal of Empirical Finance, 11: 379-398.
20. Brock, W.A., Dechart, W.D., & Scheinkman, J.A., 1987. A Test for Independence based on the Correlation Dimension. Social Science Research Unit, University of Wisconsin: Mimeo.
21. Engle, R.F., & Ng, V.K., 1993. Measuring and testing the impact of news on volatility. Journal of Finance, 48: 1749-1778.