International Journal of Business and Management Invention
ISSN (Online): 2319 – 8028, ISSN (Print): 2319 – 801X
www.ijbmi.org || Volume 5 Issue 10 || October. 2016 || PP—01-14

The Predictive Power of Intraday-Data Volatility Forecasting Models: A Case Study of Taiwan Stock Indices

Yu-Min Wang[1], Chia-Fei Lin[2], Ming-Kai Chen[3]
[1]*Department of Banking and Finance, National Chiayi University*
[2]*Department of Medical and Healthcare Business, Hsin Sheng College of Medical Care and Management*
[3]*Department of Business Administration, National Chiayi University*

ABSTRACT:*The purpose of this study was to compare the predictive power of various volatility forecasting models. Using intraday high-frequency data, this study investigated the influence of time frequency on the predictive power of a volatility forecasting model. The empirical results revealed that the realized volatility increased when the time frequency of forecasts reduced. The overall results showed that when the forecast range was 1 day, among various volatility forecasting models, the autoregressive moving average-generalized autoregressive conditional heteroskedasticity(1, 1) model presented the optimal forecasting performance and the implied volatility model presented the worst forecasting performance for all time frequencies.*
JEL classification:C14; C32; C53
Keywords: *High-Frequency Data, Long-Memory Model, Realized Volatility*

I. INTRODUCTION

Throughout the previous decades, financial derivatives have rapidly mushroomed and numerous disastrous events have frequently occurred in global financial markets. Financial instrument valuation and investment strategies are closely related to changes in asset return volatility; they mutually influence each other. Volatility represented bya percentage signifies underlying asset price fluctuations. Volatility refers to the amplitude of price fluctuations without considering the direction of price fluctuations; in other words, volatility means the severity of price fluctuations.

Numerous studies have emphasized how to enhance the accuracy of volatility forecasting. Particularly, when investors undertake investment activities such as determining portfolio strategies and asset allocation and conductingoptions valuation andrisk analysis, assessing investment risks and returns is extremely crucial. Previous studies on volatility mostly focused on historical volatility in investment markets. Another method for describing volatility is called realized volatility (RV), which applies high-frequency data to describe volatility fluctuations.

In addition to historical volatility and RV, implied volatility (IV) is a commonly used method for volatility forecasting. IV is the expectation value of volatility in a market. Regarding the calculation method of IV, the market prices of options are used as theoretical prices, and then the option valuation formula proposed by Black–Scholes (1973) is used along with observed market prices to derive IV. IV concurrently contains historical, current, and future information.

In the literature on volatility, some researchers have indicated that volatility in financial markets is not only related to time fluctuations, but also has a long-memory characteristic according to time series observation. The long-memory characteristic means that a time series is constantly influenced by previous series and therefore shows a slow hyperbolic decline curve. In other words, the interdependence of time-series observation values is influenced by previous series as time duration increases. Therefore, when a variable possesses a long-memory characteristic, the random external impact during each periodconstantly influences the variable for an extended period.

Generally, the impact of related information on economics and finance can be short-termor long-term and persistent. Granger and Joveux (1980) proposed the autoregressive integrated moving average model to explain the long-term impact of related information. Hosking (1981) indicated that the autocorrelation coefficients between series observation values decline more slowly than an autoregressive integrated moving average process. This phenomenon is called long-memory behavior.

This study compared various volatility forecasting models by comparing the predictive power of time series model, IV, and long-memory model for forecasting intraday data at various frequencies. The researchers of this study hoped to identify the methods and forecasting models suitable for evaluating the Taiwan stock market and to provide a reference for investors in trading markets.

II. LITERATURE REVIEW

This study compared various volatility forecasting models. Volatility measurement is a crucial research topic in financial economics. The investment strategies related to portfolio, asset pricing, or risk management are closely related to volatility measurement. Among numerous studies on return fluctuations, Engle (1982) proposed the autoregressive conditional heteroskedasticity(ARCH) model to explain volatility clustering. Bollersleve (1986) further developed the generalized autoregressive conditional heteroskedasticity (GARCH) model, which can not only capture the fluctuations of stock return volatility, but also rationalize the deferred structure of heteroskedasticity.

Hull and White (1987) indicated that when the information about the underlying asset market can effectively cause responses from the options market, IV should be the unbiased estimate of actual fluctuations in the future. Harvey and Whaley (1991) indicated that at-the-money options that expire in months clearly reflect the messages carried by the IV model. At-the-money options are susceptible to volatility fluctuations because theycover a maximal amount of information concerning volatility. Therefore, IV inferred from at-the-money options can reflect actual market situations.

In studies regarding stock market volatility forecasting, Christensen and Prabhala (1998) and Fleming (1998) have indicated that the predictive power of IV is significantly superior to that of historical volatility. Szakmary, Ors, Kim, and Davidson (2003) also found that the predictive power of IV is significantly superior to that of historical volatility. Andersen, Bollerslev, and Diebold (2001) indicated that at appropriate time frequencies, RV is the consistent, unbiased, and effective estimate of actual volatility and has the long-memory characteristic. Poon and Granger (2003) indicated thatthe strike prices of at-the-money options are frequently used to calculate IV because at-the-money options have high market liquidity and low measurement error.

Wong et al. (2009) assumed that the Taiwan stock market was an imperfect market, investigated the information content of RV and IV based on options, and empirically found that the IV inferred from Taiwan stock index options (TXO) prices containedthe most information regardingvolatility forecasting. Hung, Tzang, and Shyu (2009) compared the volatility forecasting efficiency of high and low volatility intervals in the Taiwan stock market with the volatility forecasting efficiency ofvolatility index (VIX). Hung, Tzang, and Shyu (2009) found that the forecast efficiencies of the two methods are similar and that for a small sample size, the volatility predictive power of VIX is significantly superior to that of the other method.

Mandelbrot and van Ness (1968) were the first to propose the concept of long memory. Long memory means that a time series is constantly influenced by previous series; therefore,such time series show a slow hyperbolic decline curve. Observing long memory requires performing data analysis for a certain period of time. By observing the long-memory characteristic, the persistence of the economic and financial impact can be clearly understood and the autocorrelation between time periods can be easily determined.

Brunetti and Gilbert (2000) found that a significant long-memory characteristic exists in markets. Baillie, Bollerslev, and Mikkelsen (1996) used a long-memory model to study changes in average daily gains in terms of German marks against U.S. dollars and found that the volatility predictive power of the long-memory model is superior to that of the traditional short-term time series model.

Choi and Zivot (2007) found that after multistructural changeswere adjusted, the sustainability of forward discount significantly reduced; however, the long-memory characteristic still existed. Choi and Zivot (2010) used RV to examine and estimate the long-memory model and found that volatility possessedthe long-memory characteristic. Ohanissian, Russell, and Tsay (2011) showed that the long-memory characteristic was caused by unstable structural changes or a slow regime-switching model.

Conventionally, most studies simulated daily data to investigatevolatility forecastingmodels. However, using intraday data for evaluations easily causes measurement errors. Because high-frequency trading has graduallybecome widespread, studies related to high-frequency intraday data have gradually received attention. Blain, Poon, and Taylor (2001) investigated the volatility of the S&P 100 Index and found that using high-frequency intraday data substantially enhanced the explanatory powerof actual volatility.

Zhou (1996) indicated that by using high-frequency data to estimate volatility, the volatility at various frequencies in any sample interval can be immediately estimated and a delay will not occur because of database updates. Drost and Nijman (1993) indicated that high- and low-frequency intraday data are highly correlated and that the results obtained by using high-frequency data for estimating low-frequency data are generally superior to the results obtained by directly using low-frequency data.

In a study by Giot (2005) regarding high-frequency intraday data, two types of data at two time frequencies (15 minutes and 30 minutes) were used and excellent estimation results were obtained. Guglielmo and Gil-Alana Luis (2010) used the intraday return data at various time frequencies and found that the data at low time frequencieswere correlated with the low orders of fractional difference and that when the time frequency for the selected data was 10 minutes, the value of the fractional difference parameter was less than 1, indicating mean reversion.Maheu and McCurdy (2011) used high-frequency data to measure volatility and compare the predictive power of time series models.

The aforementioned related studiesserved as references for this study. Previous studies rarely investigated volatility forecasting by using intraday data at various time frequencies. This study used intraday data forthe stock index as the research sample and categorized intraday data into various types of data at various time frequencies. In addition, regarding volatility forecasting models, this study referred to the volatility forecasting models proposed by other researchers and adopted the commonly used models as the representative models. To compare forecasting models, this study integrated time series models and compared those models with the IV model. Based on the characteristics of various forecasting models, the accuracy and suitability of such models were compared.

III. METHODS

A unit root test was performed on the selected sample data to determine whether the adopted time series data pattern met stationary conditions. Subsequently, the method for measuring RV was introduced and the baseline for comparing volatility predictive power was established. In addition, this study investigated whether time series models and the IV model for high-frequency data are useful for forecasting RV. Finally, the principle of error assessment was used to compare the predictive power of various volatility forecasting models.

1. Volatility measurement

(1) Realized volatility

Based on Andersen and Bollerslev (1998), this study defined RV as the sum of squared intraday returns, which is expressed as Equation (1).

$$RV_t = \sum_{n=1}^{N} R_{t,n}^2 \quad , \quad RSD_{t,1} = \sqrt{RV_{t,1}} \qquad (1),$$

where $R_{t,n} = \log(S_{t,n} / S_{t,n-1})$ denotes the intraday return in day *t*; *N* denotes the number of time intervals in a trading day,and $S_{t,n}$ denotes the stock index at time *n* in a trading day. When the frequency for intraday returns is set to be high (i.e., when *n* approaches infinity), RV will converge toward actual volatility. Regarding intraday data frequency selection for estimating RV, this study used intraday data at five frequencies (i.e., every 5, 10, 15, 30, and 60 minutes) and used the final quoted price as the baseline. Because the Taiwan Stock Exchange operates from 9 AM to 1:30 PM, using the data at the frequency of every 5 minutes as an example, *N* is equal to 54 during trading time (the same method was used to calculate *N* for other time frequencies).

The risk-free rate, the period of time before an expiry date, and the volatility ofthe Black–Scholes option valuation model used in this study are ona yearly basis. The volatility is annualized volatility. Nevertheless, RV is calculated only for particular forecast ranges. Therefore, based on the method developed by Pong et al. (2004), this study calculated RV on an annualized basis. Equation (2) shows how to calculate annualized RV ($RSD_{t,T}^A$).

$$RSD_{t,T}^A = 100 \times \sqrt{\frac{365}{T} RV_{t,T}} \quad (2)$$

According to Pong et al. (2004), this study indicated that the distribution of RV for high-frequency stock returns approximates a log-normal distribution. For RV measurement, this study calculated the logarithm of annualized RVby using Equation (3).

$$y_{t,T} = \log(RSD_{t,T}^A) \qquad (3)$$

(2) Implied volatility

Historical volatility means using time series data to estimate actual volatility in the future. If a drastic change suddenly occurs in markets, historical volatility cannot immediately and adequately forecast the change of actual volatility. However, because option contracts per se imply a large amount of market information, the IV obtained using the Black–Scholes option valuation model can solve the problem regarding market information asymmetry. In addition, because IV is calculated using option prices, IV also reflects investors' expectations concerning the volatility of underlying assets in the future.

According to the risk-free arbitrage principle, Black and Scholes (1973) indicated that options can be assessed using the risk-neutral valuation method. According to previous studies, the IV obtained fromat-the-money options represents the subjective volatility in markets; without assessing volatility risk, the IV obtained from at-the-money options should be, in theory, the unbiased estimate of the average volatility in the future. In addition, because at-the-money options have high market liquidity, the IV obtained from at-the-money options is a forecast representative, compared with the volatility obtained from in-the-money or out-of-the-money options. Therefore, this study investigated only the IV obtained from at-the-money options(call options). This study calculated the at-the-money options defined in this study according to the daily IV database established by Taiwan Economic Journal. This study adopted Newton's method to obtain the approximate solutionfor IV. Newton's method was adopted because the convergence rate of using itfor obtaining solutions is faster than that of using the dichotomy method.

2. Comparing the predictive power of various models

Based on Wash, David, Tsou, and Glenn (1998), this study used root mean squared error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) as the error indicators for forecasting models. This study used the three error indicators to assess the predictive values obtained using various volatility forecasting models and to determine the forecast efficiency of various models.

IV. EMPIRICAL ANALYSIS

This study explored volatility forecasting for the Taiwan stock market. This study used the intraday data collected every minute of every trading day regarding the Taiwan weighted stock index (TAIEX), electronic index, and financial and insurance stock index provided by the Taiwan Stock Exchange as the sample data. The sample data were collected from January 2, 2007 to December 28, 2012. In addition, this study used the time when the financial crisis of 2008 occurred as the cut-off point to divide the sample period into two sub periods. The two subperiods were from January 2, 2007 to December 31, 2008 and from January 5, 2009 to December 28, 2012. To understand how sample intraday data at various time frequencies changed, this study organized the sample data collected every minute of every day into the intraday data at various time frequencies (i.e., every 5, 10, 15, 30, and 60 minutes). According to Andersen and Bollerslev (1998), by using intraday data to measure volatility, volatility measurement is easily influenced by market-microstructure factors, thereby resulting in biased estimation if the time frequency for sample data is less than 5 minutes. Therefore, the time frequencies for high-frequency data used in this study were not less than 5 minutes.

The options data used in this study were the daily options trading data provided by the database established by Taiwan Economic Journal. The sample option contracts included TXO, electronic options (TEO), and financial options (TFO). Considering the characteristics of Taiwanese option contracts, because RV can be easily overestimated for in-the-money and out-the-money options, this study investigated only at-the-money options (call options).

Before constructing an empirical model, this study performed a unit root test on sample data to determine whether the sample data were stationary series. This study alsotabulated descriptive statistical results regarding RV for TAIEX, electronic index, and financial and insurance stock index. The empirical results indicated that for low time frequencies, the standard deviation of RV tended to be large, indicating a large change in RV for low time frequencies. Next, according to the descriptive statistical results, before the logarithm of RV was calculated, the distribution of RV apparently deviated from normal distribution; after the logarithm of RV was calculated, the problems regarding deviation from normal distribution and extreme values were substantially improved. In addition, the empirical results showed that the volatility in financial and insurance markets was susceptible to the impact of external events.

1. Time series models

(1) The empirical results for the GARCH (1, 1) model

After Engle (1982) proposed the ARCH model, Bollerslev (1986) proposed a generalized ARCH (GARCH) model. The GARCH model can adequately describe changes in stock returns and returns from other financial instruments, can capture the volatility clustering phenomenon in stock returns, and can solve the problems that the ARCH model cannot solve. Among numerous short-term time series models, the GARCH model is widely used. The performance of a complex model is not necessarily superior to that of a simple model. Therefore, this study used the GARCH model to forecast RV.

Table 1 presents the test results for TAIEX. The Lagrangian multiplier test (LM(1)) results and the Q^2 test results showed that no ARCH effect and no autocorrelation phenomenon existed forsubperiods or the entire sample period. Table 2 presents the test results for electronic index. The results of a heteroskedasticity test on residuals also showed that no ARCH effect and no autocorrelation existed for sub periods or the entire sample period. Therefore, the model was acceptable. Table 3 presents the test results for financial and insurance stock index. The results forfinancial and insurance stock index differed from the results for TAIEX and electronic index. According to the test results for financial and insurance stock index for the subperiod of 2007 to 2008, the results of the Q^2 test on the data at the frequency of 5 minutes performed at the 10^{th}5-minute interval showed that autocorrelation existed. Therefore, this study performed a Q^2 test at later 5-minute intervals and the results showed that the ARCH effect was effectively removed at the 12^{th}5-minute interval. In addition, other problems related to heteroskedasticity and autocorrelation in the model were also effectively eliminated. For the data regarding financial and insurance stock index at the time frequency of 5 minutes for the entire sample period, the results showed that heteroskedasticity existed. We postponed performing the ARCH-LM test until the 5^{th}5-minute interval and found that heteroskedasticity no longer existed.

In summary, although autocorrelation and heteroskedasticity existed in financial and insurance stock index, postponing the performanceof statistical tests effectively solved the problem. In addition to financial and insurance stock index, heteroskedasticity in TAIEX and electronic index was also effectively removed from the model, indicating the model was acceptable.

(2) The empirical results for the autoregressive moving average–generalized autoregressive conditional heteroskedasticity (ARMA–GARCH) model

Tables 4 to 6 present the test results for ARMA (p, q)–GARCH (1, 1). As shown in Table 4 about TAIEX, the results of performing the Q^2 test at the 10^{th} interval revealed that no autocorrelation existed, except that autocorrelation existed in the residuals for the data at the time frequency of 30 minutes for the 2009–2012sub period. However, postponing performing the Q^2 test until later intervals solved the problem regarding autocorrelation. In addition, the log-likelihood (LLH) value decreased as the time frequency increased. This study found that the values of $\beta_1 + \beta_2$ for various time frequencies for various periods mostly approximated 1, indicating that volatility was highly persistent.

Table 5 presents the test results for the electronic index. The results showed that no autocorrelation existed forsubperiods and the entire sample period and that most coefficients for autoregressive (AR) and moving average (MA) were significant. The results also indicated highly persistent volatility. Table 6 presents the empirical results for financial and insurance stock index. The Q^2 test results showed that no autocorrelation existed for various time frequencies and for various periods, except that autocorrelation existed in the data at the time frequency of 5 minutes for the subperiod of 2007–2008 and in the data at the time frequency of 60 minutes for the subperiod of 2009–2012 andfor the entire sample period. Regarding the persistency of volatility, it was relatively weaker for financial and insurance stock index was relatively weaker for the subperiod of 2007–2008 compared with TAIEX and electronic index.

Table 1. Estimated parameters for TAIEX based on the GARCH(1,1) model

2007–2008									
	c_0	c_1	β_0	β_1	β_2	LLH	Q(10)	Q^2(10)	LM(1)
5 min	3.1948 (0.00)***	0.6634 (0.00)***	0.2179 (0.00)***	0.2066 (0.02)**	-0.1577 (0.41)	-329.98	110.72 (0.00)***	3.5833 (0.89)	0.0033 (0.95)
10 min	3.2004 (0.00)***	0.6638 (0.00)***	0.2139 (0.00)***	0.2402 (0.02)**	-0.0945 (0.58)	-349.47	112.51 (0.00)***	4.3088 (0.83)	0.0037 (0.95)
15 min	3.1843 (0.00)***	0.5801 (0.00)***	0.2553 (0.04)**	0.0896 (0.22)	-0.1081 (0.82)	-358.31	124.46 (0.00)***	3.9859 (0.86)	0.0007 (0.98)
30 min	3.1163 (0.00)***	0.4744 (0.00)***	0.2265 (0.01)**	0.0646 (0.36)	0.1991 (0.53)	-409.58	158.94 (0.00)***	3.7702 (0.88)	0.0209 (0.88)
60 min	3.0907 (0.00)***	0.4131 (0.00)***	0.2338 (0.04)**	-0.0036 (0.94)	0.3571 (0.26)	-450.03	154.26 (0.00)***	8.5786 (0.38)	0.0132 (0.91)
2009-2012									
5 min	2.8787 (0.00)***	0.5189 (0.00)***	0.1319 (0.00)***	0.0897 (0.04)**	0.1858 (0.36)	-562.38	250.69 (0.00)***	10.4970 (0.23)	0.0788 (0.78)
10 min	2.8889 (0.00)***	0.4664 (0.00)***	0.1245 (0.09)*	0.0593 (0.11)	0.3144 (0.41)	-608.13	270.33 (0.00)***	9.2997 (0.32)	0.0231 (0.88)
15 min	2.8528 (0.00)***	0.4098 (0.00)***	0.1548 (0.01)**	0.0547 (0.10)	0.2429 (0.36)	-660.31	276.28 (0.00)***	9.2834 (0.32)	0.0027 (0.96)
30 min	2.7781 (0.00)***	0.3441 (0.00)***	0.0060 (0.15)	0.0286 (0.01)**	0.9487 (0.00)***	-746.64	246.27 (0.00)***	14.6615 (0.13)	0.2844 (0.59)
60 min	2.7271 (0.00)***	0.1949 (0.00)***	0.0173 (0.11)	0.0062 (0.58)	0.9479 (0.00)***	-932.11	289.54 (0.00)***	4.9728 (0.76)	0.0630 (0.80)
2007–2012									
5 min	2.9803 (0.00)***	0.6206 (0.00)***	0.1601 (0.00)***	0.1593 (0.00)***	0.0517 (0.75)	-912.27	344.75 (0.00)***	8.1509 (0.42)	0.0091 (0.92)
10 min	2.9902 (0.00)***	0.5845 (0.00)***	0.1714 (0.00)***	0.1362 (0.01)**	0.0892 (0.67)	-980.64	359.46 (0.00)***	10.4828 (0.23)	0.0012 (0.97)
15 min	2.9586 (0.00)***	0.5234 (0.00)***	0.1882 (0.00)***	0.0834 (0.03)**	0.1283 (0.65)	-1044.80	390.96 (0.00)***	7.8875 (0.44)	0.0003 (0.99)
30 min	2.8587 (0.00)***	0.3994 (0.00)***	0.0066 (0.09)*	0.0344 (0.00)***	0.9431 (0.00)***	-1175.23	445.09 (0.00)***	10.7815 (0.21)	0.0155 (0.90)
60 min	2.8307 (0.00)***	0.2961 (0.00)***	0.0109 (0.02)**	0.0212 (0.01)**	0.9512 (0.00)***	-1414.25	513.28 (0.00)***	3.1385 (0.93)	0.8235 (0.36)

Notes: 1. The values in parentheses are Pvalues.
2. The symbols *, **, and *** represent the 10%, 5%, 1% significance levels, respectively.

3. $y_{t,T} = c_0 + c_1 y_{t-1,T} + \varepsilon_t$, $h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$, $y_{t,T}$ isthelogarithmof RV, h_t is conditional heteroskedasticity.

Table 2. Estimated parameters for electronic index based on the GARCH(1,1) model

2007–2008									
	c_0	c_1	β_0	β_1	β_2	LLH	Q(10)	Q^2 (10)	LM(1)
5 min	3.2958 (0.00)***	0.6134 (0.00)***	0.2258 (0.00)***	0.1331 (0.05)**	-0.1986 (0.43)	-315.05	128.31 (0.00)***	3.7395 (0.88)	0.0802 (0.78)
10 min	3.3051 (0.00)***	0.5949 (0.00)***	0.2749 (0.00)***	0.1454 (0.06)*	-0.2822 (0.32)	-347.18	123.24 (0.00)***	5.4862 (0.70)	0.0399 (0.84)
15 min	3.2752 (0.00)***	0.5153 (0.00)***	0.4068 (0.00)***	0.0398 (0.37)	-0.6873 (0.09)*	-355.90	130.82 (0.00)***	4.6092 (0.80)	0.0046 (0.95)
30 min	3.1509 (0.00)***	0.3799 (0.00)***	0.0077 (0.21)	0.0280 (0.07)*	0.9487 (0.00)***	-410.76	171.20 (0.00)***	3.6723 (0.89)	0.1319 (0.71)
60 min	3.1820 (0.00)***	0.3772 (0.00)***	0.3155 (0.01)**	-0.0714 (0.09)*	0.1439 (0.70)	-433.68	159.09 (0.00)***	9.4954 (0.30)	0.0443 (0.83)
2009–2012									
5 min	2.9707 (0.00)***	0.4767 (0.00)***	0.1085 (0.00)***	0.1022 (0.03)**	0.2423 (0.21)	-513.39	260.94 (0.00)***	8.8062 (0.36)	0.0046 (0.95)
10 min	2.9701 (0.00)***	0.4281 (0.00)***	0.1033 (0.03)**	0.0748 (0.05)**	0.3604 (0.00)***	-565.50	266.96 (0.00)***	8.3816 (0.40)	0.0002 (0.99)
15 min	2.9303 (0.00)***	0.3680 (0.00)***	0.1338 (0.01)**	0.0514 (0.13)	0.2932 (0.24)	-622.34	274.89 (0.00)***	11.7662 (0.16)	0.0004 (0.98)
30 min	2.8443 (0.00)***	0.3021 (0.00)***	0.0057 (0.10)	0.0268 (0.01)**	0.9498 (0.00)***	-718.20	232.38 (0.00)***	15.1328 (0.12)	0.0586 (0.81)
60 min	2.7938 (0.00)***	0.1822 (0.00)***	0.0111 (0.05)**	0.0176 (0.08)*	0.9498 (0.00)***	-883.08	252.59 (0.00)***	11.0245 (0.20)	3.0404 (0.08)*
2007–2012									
5 min	3.0692 (0.00)***	0.5807 (0.00)***	0.1311 (0.00)***	0.1346 (0.00)***	0.1647 (0.00)***	-854.98	376.90 (0.00)***	7.8241 (0.45)	0.025 (0.87)
10 min	3.0293 (0.00)***	0.4739 (0.00)***	0.0049 (0.04)**	0.0338 (0.00)***	0.9429 (0.00)***	-934.87	445.21 (0.00)***	10.2980 (0.24)	1.9552 (0.16)
15 min	2.9942 (0.00)***	0.4318 (0.00)***	0.0071 (0.02)**	0.0336 (0.00)***	0.9353 (0.00)***	-1002.67	449.82 (0.00)***	5.3365 (0.72)	0.0217 (0.88)
30 min	2.9196 (0.00)***	0.3562 (0.00)***	0.0054 (0.03)**	0.0362 (0.00)***	0.9448 (0.00)***	-1149.58	432.11 (0.00)***	10.3064 (0.24)	0.0063 (0.94)
60 min	2.8993 (0.00)***	0.2897 (0.00)***	0.0081 (0.01)**	0.0273 (0.00)***	0.9508 (0.00)***	-1351.11	477.89 (0.00)***	11.0918 (0.20)	7.5331 (0.01)

Notes:1. The values in parentheses are Pvalues.
2. The symbols *, **, and *** represent the 10%, 5%, 1%significance levels, respectively.
3. $y_{t,T} = c_0 + c_1 y_{t-1,T} + \varepsilon_t$, $h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$, $y_{t,T}$ is the logarithm of RV, h_t is conditional heteroskedasticity.

Table 3. Estimated parameters for financial and insurance stock index based on the GARCH(1,1) model

2007–2008									
	c_0	c_1	β_0	β_1	β_2	LLH	Q(10)	Q^2 (10)	LM(1)
5 min	3.3326 (0.00)***	0.7358 (0.00)***	0.1670 (0.00)***	0.3973 (0.00)***	-0.0739 (0.65)	-324.77	89.90 (0.00)***	15.36 (0.05)**	0.0306 (0.86)
10 min	3.3251 (0.00)***	0.6444 (0.00)***	0.0506 (0.33)	0.1306 (0.02)**	0.6716 (0.01)**	-355.07	118.01 (0.00)***	6.1044 (0.64)	0.7329 (0.39)
15 min	3.3159 (0.00)***	0.6448 (0.00)***	0.1763 (0.03)**	0.1855 (0.01)**	0.2103 (0.50)	-389.96	110.38 (0.00)***	5.9116 (0.66)	1.51E-07 (0.99)
30 min	3.2371 (0.00)***	0.4639 (0.00)***	0.0609 (0.40)	0.1275 (0.02)**	0.7069 (0.00)***	-442.70	158.12 (0.00)***	7.5171 (0.48)	0.0477 (0.83)
60 min	3.2861 (0.00)***	0.4120 (0.00)***	0.7230 (0.00)***	-0.0488 (0.17)	-0.6146 (0.01)**	-494.78	186.04 (0.00)***	4.9109 (0.77)	0.0192 (0.89)
2009–2012									
5 min	3.1616 (0.00)***	0.6357 (0.00)***	0.0502 (0.65)	0.0949 (0.26)	0.5897 (0.45)	-490.56	193.26 (0.00)***	6.8496 (0.55)	1.1309 (0.29)
10 min	3.1458 (0.00)***	0.5541 (0.00)***	0.0037 (0.08)*	0.0414 (0.00)***	0.9395 (0.00)***	-581.84	207.38 (0.00)***	10.5262 (0.23)	0.0072 (0.93)
15 min	3.1186 (0.00)***	0.4821 (0.00)***	0.0059 (0.09)*	0.0326 (0.01)**	0.9414 (0.00)***	-668.87	246.33 (0.00)***	12.7756 (0.12)	0.2321 (0.63)
30 min	3.0708 (0.00)***	0.3838 (0.00)***	0.0051 (0.11)	0.0238 (0.01)**	0.9591 (0.00)***	-799.04	254.23 (0.00)***	10.0372 (0.26)	0.0065 (0.94)
60 min	3.0342	0.2618	0.7950	0.0281	-0.8869	-988.62	268.78	13.8479	0.7179

	(0.00)***	(0.00)***	(0.00)***	(0.08)*	(0.00)***		(0.00)***	(0.20)	(0.40)
2007–2012									
5 min	3.1849 (0.00)***	0.6380 (0.00)***	0.0265 (0.27)	0.1205 (0.03)**	0.7370 (0.00)***	-823.67	298.52 (0.00)***	12.463 (0.13)	5.1227 (0.02)**
10 min	3.1930 (0.00)***	0.5857 (0.00)***	0.0112 (0.25)	0.0731 (0.02)**	0.8759 (0.00)***	-946.93	333.53 (0.00)***	5.9002 (0.66)	0.9068 (0.34)
15 min	3.1672 (0.00)***	0.5189 (0.00)***	0.0134 (0.24)	0.0573 (0.02)**	0.8896 (0.00)***	-1068.55	391.27 (0.00)***	8.2164 (0.41)	1.5752 (0.21)
30 min	3.1161 (0.00)***	0.4136 (0.00)***	0.0213 (0.46)	0.0557 (0.15)	0.8779 (0.00)***	-1251.79	427.99 (0.00)***	9.6189 (0.29)	0.1186 (0.73)
60 min	3.0906 (0.00)***	0.3142 (0.00)***	0.0092 (0.13)	0.0146 (0.05)**	0.9645 (0.00)***	-1502.13	482.04 (0.00)***	15.04 (0.06)*	0.9980 (0.32)

Notes:1. The values in parentheses are Pvalues.

2. The symbols *, **, and *** represent the 10%, 5%, 1% significance levels, respectively.

3. $y_{t,T} = c_0 + c_1 y_{t-1,T} + \varepsilon_t$, $h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$, $y_{t,T}$ is the logarithm of RV, h_t is conditional heteroskedasticity.

Table 4. Estimated parameters for TAIEX based on the ARMA(p,q)-GARCH(1,1) model

2007–2008												
	c_0	c_1	c_2	b_1	b_2	β_0	β_1	β_2	AIC	LLH	Q(10)	$Q^2(10)$
5 min	3.2776 (0.00)***	0.0012 (0.94)	0.9749 (0.00)***	0.1532 (0.00)***	-0.8121 (0.00)***	0.0156 (0.52)	0.0194 (0.43)	0.8831 (0.00)***	1.0253	-245.25	10.005 (0.12)	6.2418 (0.40)
10 min	3.2536 (0.00)***	0.9843 (0.00)***		-0.8854 (0.00)***	0.0822 (0.08)*	0.0165 (0.42)	0.0188 (0.44)	0.8875 (0.00)***	1.1077	-267.1444	4.7823 (0.69)	8.2016 (0.32)
15 min	3.2766 (0.00)***	0.9872 (0.00)***		-0.8332 (0.00)***		0.0196 (0.51)	0.0020 (0.93)	0.8913 (0.00)***	1.1492	-278.4233	8.0341 (0.43)	4.5364 (0.81)
30 min	3.1387 (0.00)***	0.9832 (0.00)***		-0.9413 (0.00)***	0.1214 (0.01)**	0.0131 (0.23)	0.0153 (0.45)	0.9267 (0.00)***	1.3396	-324.5551	3.2830 (0.86)	6.6147 (0.47)
60 min	3.1921 (0.000)***	0.9858 (0.000)***		-0.8644 (0.000)***		0.1677 (0.40)	-0.0450 (0.19)	0.4377 (0.54)	1.5714	-382.9218	10.036 (0.26)	10.404 (0.24)
2009–2012												
5 min	2.8062 (0.00)***	0.9832 (0.00)***		-0.8365 (0.00)***		0.0518 (0.12)	0.0644 (0.05)**	0.5623 (0.03)**	0.8664	-426.3203	10.440 (0.24)	7.5205 (0.48)
10 min	2.8431 (0.00)***	0.9850 (0.00)***		-0.8428 (0.00)***		0.0009 (0.21)	0.0110 (0.03)**	0.9832 (0.00)***	0.9554	-470.7573	11.481 (0.18)	11.293 (0.19)
15 min	2.7568 (0.00)***	0.9842 (0.00)***		-0.9110 (0.00)***	0.049 (0.14)	0.0523 (0.10)	0.0581 (0.05)**	0.6391 (0.00)***	1.0875	-535.6673	10.304 (0.17)	6.258 (0.51)
30 min	2.7136 (0.00)***	0.9816 (0.00)***		-0.8642 (0.00)***		0.0676 (0.21)	0.0536 (0.08) *	0.6328 (0.02)**	1.3112	-648.9387	15.494 (0.05)**	6.6833 (0.57)
60 min	2.7058 (0.00)***	0.8568 (0.00)***	0.1202 (0.00)***	-0.8629 (0.00)***		0.6156 (0.00)***	0.015 (0.00)***	-1.0099 (0.00)***	1.6705	-825.7367	11.63 (0.11)	2.9538 (0.89)
2007–2012												
5 min	2.9574 (0.00)***	0.3479 (0.14)	0.6312 (0.01) **	-0.2314 (0.34)	-0.4818 (0.02) **	0.0041 (0.11)	0.0191 (0.01)**	0.9529 (0.00)***	0.9133	-673.7982	9.3434 (0.16)	7.2789 (0.30)
10 min	2.9697 (0.00)***	0.9867 (0.00)***		-0.8825 (0.00)***	0.0547 (0.04) **	0.0078 (0.05)**	0.0276 (0.01)**	0.9239 (0.00)***	1.0039	-724.9409	8.0589 (0.33)	6.7937 (0.45)
15 min	2.9171 (0.00)***	0.9880 (0.00)***		-0.8966 (0.00)***	0.0527 (0.05) **	0.0163 (0.16)	0.0279 (0.04)**	0.88 (0.00)***	1.1032	-817.0825	5.0207 (0.66)	5.4977 (0.60)
30 min	2.8614 (0.00)***	0.9841 (0.00)***		-0.9042 (0.00)***	0.0648 (0.00)***	0.0077 (0.05)**	0.0181 (0.05)**	0.9470 (0.00)***	1.3151	-975.3744	7.0009 (0.43)	11.499 (0.12)
60 min	2.8268 (0.00)***	0.9863 (0.00)***		-0.9533 (0.00)***	0.0798 (0.03)**	0.0127 (0.11)	-0.0030 (0.70)	0.9556 (0.00)***	1.6504	-1225.856	13.607 (0.06)*	4.344 (0.74)

Notes: 1 The symbols *, **, and *** represent the 10%, 5%, 1% significance levels, respectively.

2. The values in parentheses are Pvalues.

3. $y_{t,T} = c_0 + \sum_{i=1}^{p} c_i y_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{i,t}$, $h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \varepsilon^2{}_{t-1}$, $y_{t,T}$ is the logarithm of RV, h_t is conditional heteroskedasticity.

Table 5. Estimated parameters for electronic index based on the ARMA(p,q)-GARCH(1,1) model

	c_0	c_1	c_2	b_1	b_2	β_0	β_1	β_2	AIC	LLH	Q(10)	Q^2 (10)
2007–2008												
5 min	3.3675 (0.00)***	0.9861 (0.00)***		-0.9015 (0.00)***	0.0830 (0.10) *	0.0167 (0.53)	0.0161 (0.51)	0.8726 (0.00)***	0.9678	-230.0609	13.078 (0.07) *	5.9795 (0.54)
10 min	3.3539 (0.00)***	0.9840 (0.00)***		-0.8978 (0.00)***	0.0868 (0.08) *	0.0192 (0.50)	0.0190 (0.49)	0.8705 (0.00)***	1.1016	-265.6505	4.4603 (0.73)	5.7194 (0.57)
15 min	3.2524 (0.00)***	0.4935 (0.00)***				0.0260 (0.52)	0.0118 (0.66)	0.8842 (0.00)***	1.4560	-355.3604	139.90 (0.00)***	5.5575 (0.78)
30 min	3.2102 (0.00)***	0.9841 (0.00)***		-0.9499 (0.00)***	0.1082 (0.02)**	0.0063 (0.21)	0.0172 (0.28)	0.9571 (0.00)***	1.3804	-334.6509	1.6894 (0.98)	6.7702 (0.45)
60 min	3.2339 (0.00)***	0.9839 (0.00)***		-0.8585 (0.00)***		0.0155 (0.19)	0.0084 (0.65)	0.9344 (0.00)***	1.5179	-369.6755	8.7114 (0.37)	7.5988 (0.47)
2009–2012												
5 min	2.9224 (0.00)***	0.9815 (0.00)***		-0.8795 (0.00)***	0.0488 (0.11)	0.0024 (0.16)	0.0220 (0.01)**	0.9586 (0.00)***	0.7735	-378.9987	8.0783 (0.33)	11.409 (0.12)
10 min	2.9080 (0.00)***	0.9844 (0.00)***		-0.8848 (0.00)***	0.0322 (0.31)	0.0026 (0.15)	0.0262 (0.01)**	0.9559 (0.00)***	0.8908	-437.9150	10.122 (0.18)	7.9399 (0.34)
15 min	2.8347 (0.00)***	0.9866 (0.00)***		-0.8781 (0.00)***		0.0082 (0.20)	0.0343 (0.02)**	0.9153 (0.00)***	1.0266	-560.2655	12.756 (0.12)	10.754 (0.22)
30 min	2.8022 (0.00)***	0.9794 (0.00)***		-0.8721 (0.00)***		0.0834 (0.09)*	0.0835 (0.05)**	0.5157 (0.05) **	1.2704	-627.9241	12.583 (0.13)	8.4709 (0.39)
60 min	2.7557 (0.00)***	0.3315 (0.05) **	0.6323 (0.00)***	-0.3216 (0.08) *	-0.4780 (0.00)***	0.0248 (0.38)	0.0105 (0.51)	0.9023 (0.00)***	1.6090	-794.1109	3.8168 (0.70)	7.2206 (0.30)
2007–2012												
5 min	3.0432 (0.00)***	0.9870 (0.00)***		-0.8910 (0.00)***	0.0612 (0.02)**	0.0069 (0.08)*	0.03 (0.01) **	0.9190 (0.00)***	0.8323	-614.7586	12.616 (0.08)*	9.1677 (0.24)
10 min	3.0314 (0.00)***	0.9875 (0.00)***		-0.8929 (0.00)***	0.0525 (0.05)**	0.0049 (0.06)*	0.0315 (0.00)***	0.9369 (0.00)***	0.9585	-708.9772	5.3622 (0.62)	6.8963 (0.44)
15 min	2.9898 (0.00)***	0.9887 (0.00)***		-0.9036 (0.00)***	0.0486 (0.07) *	0.0113 (0.15)	0.0303 (0.03)**	0.9034 (0.00)***	1.0696	-791.9653	5.2323 (0.63)	11.158 (0.13)
30 min	2.9225 (0.00)***	0.9870 (0.00)***		-0.8682 (0.00)***		0.0053 (0.06)*	0.0227 (0.01)**	0.9533 (0.00)***	1.3066	-970.0308	10.573 (0.23)	17.327 (0.23)
60 min	2.8996 (0.00)***	0.4141 (0.04)**	0.5640 (0.00)***	-0.3807 (0.07) *	-0.4210 (0.03)**	0.0129 (0.20)	0.0138 (0.20)	0.9409 (0.00***	1.5800	-1171.476	2.0859 (0.91)	10.163 (0.12)

Notes: 1 The symbols *, **, and *** represent the 10%, 5%, 1% significance levels, respectively.

2. The values in parentheses are Pvalues.

3. $y_{t,T} = c_0 + \sum_{i=1}^{p} c_i y_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{i,t}$, $h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \varepsilon^2{}_{t-1}$, $y_{t,T}$ is the logarithm of RV, h_t is conditional heteroskedasticity.

Table 6. Estimated parameters for financial and insurance stock index based on the ARMA(p,q)-GARCH(1,1) model

	c_0	c_1	c_2	b_1	b_2	β_0	β_1	β_2	AIC	LLH	Q(10)	Q^2 (10)
2007–2008												
5 min	3.2692 (0.00)***	0.9854 (0.00)***		-0.7909 (0.00)***		0.1134 (0.00)***	0.2207 (0.00)***	0.1099 (0.58)	1.0361	-250.4278	9.8026 (0.28)	20.016 (0.01)**
10 min	3.3552 (0.00)***	0.9884 (0.00)***		-0.8078 (0.00)***		0.1596 (0.05) **	0.11 (0.05) **	0.01 (0.98)	1.1445	-227.2693	7.3979 (0.49)	6.5455 (0.59)
15 min	3.3489 (0.00)***	0.99 (0.00)***		-0.8269 (0.00)***		0.2205 (0.01) **	0.1206 (0.03) **	-0.1835 (0.63)	1.2760	-309.8039	5.6636 (0.69)	6.8267 (0.56)
30 min	3.3744 (0.00)***	0.9923 (0.00)***		-0.8672 (0.00)***		0.02021 (0.00)***	0.1886 (0.00)***	0.0138 (0.95)	1.4612	-355.6572	6.0640 (0.64)	9.2627 (0.32)
60 min	3.3794 (0.00)***	0.9868 (0.00)***		-0.9287 (0.00)***	0.0828 (0.05)**	0.0185 (0.02) **	-0.03 (0.04) **	0.973 (0.00)***	1.7137	-417.1517	5.6762 (0.58)	9.5338 (0.22)
2009–2012												
5 min	3.0914 (0.00)***	0.0875 (0.36)	0.8810 (0.00)***	0.1147 (0.23)	-0.7125 (0.00)***	0.0227 (0.04)**	0.0538 (0.02)**	0.7646 (0.00)***	0.7635	-372.6028	20.867 (0.00)***	6.1233 (0.41)
10 min	3.1300 (0.00)***	0.0653 (0.35)	0.9011 (0.00)***	0.1216 (0.08) *	-0.7434 (0.00)***	0.0007 (0.33)	0.0109 (0.03)**	(0.9848) (0.00)***	0.9541	-467.6013	22.421 (0.00)***	8.9979 (0.25)
15 min	3.0298 (0.00)***	0.0692 (0.50)	0.9071 (0.00)***	0.0716 (0.49)	-0.7815 (0.00)***	0.0075 (0.07) *	0.0378 (0.01)**	0.9209 (0.00)***	1.1197	-550.1887	24.153 (0.00)***	6.8643 (0.33)
30 min	2.8568 (0.00)***	0.9904 (0.00)***		-0.8860 (0.00)***		0.0835 (0.18)	0.0639 (0.05)**	0.5824 (0.04) **	1.3984	-691.8227	22.566 (0.00)***	8.4570 (0.39)
60 min	2.9418 (0.00)***	0.9060 (0.00)***	0.0790 (0.03)**	-0.8836 (0.00)***		0.0354 (0.09)*	0.0112 (0.39)	0.8857 (0.00)***	1.7938	-887.2105	20.015 (0.01)**	13.000 (0.11)
2007–2012												
5 min	3.1379 (0.00)***	0.9859 (0.00)***		-0.8015 (0.00)***		0.0225 (0.00)***	0.0822 (0.00)***	0.7567 (0.00)***	0.8512	-629.8248	18.553 (0.02) **	7.5459 (0.48)
10 min	3.1667 (0.00)***	0.9863 (0.00)***		-0.8126 (0.00)***		0.0114 (0.05) **	0.0388 (0.01)**	0.8913 (0.00)***	1.0168	-753.5788	20.078 (0.01) **	3.5504 (0.90)
15 min	3.1019 (0.00)***	0.9887 (0.00)***		-0.8437 (0.00)***		0.0182 (0.06) *	0.0470 (0.00)***	0.8568 (0.00)***	1.1698	-867.8648	19.791 (0.01) **	3.7170 (0.88)
30 min	2.9915 (0.00)***	0.9919 (0.00)***		-0.8775 (0.00)***		0.1624 (0.00)***	0.1113 (0.00)***	0.2179 (0.31)	1.4150	-1051.023	19.294 (0.01)**	9.2606 (0.32)
60 min	3.0636 (0.00)***	0.5009 (0.02)**	0.4813 (0.03)**	-0.4710 (0.04)**	-0.351 (0.08)*	0.0656 (0.54)	0.0143 (0.47)	0.7922 (0.08) *	1.7651	-1309.615	9.1825 (0.16)	13.882 (0.03)**

Notes: 1 The symbols *, **, and *** represent the 10%, 5%, 1% significance levels, respectively.

2. The values in parentheses are Pvalues.

3. $y_{t,T} = c_0 + \sum_{i=1}^{p} c_i y_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{i,t}$, $h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \varepsilon^2{}_{t-1}$, $y_{t,T}$ is the logarithm of RV, h_t is conditional heteroskedasticity.

(3) The empirical results for the autoregressive fractionally integrated moving average (ARFIMA) model

For long-memory models, the estimation of the fractional difference parameter (d) in the ARFIMA (p, d, q) model can influence our viewsregarding related data characteristics. Therefore, this study used the maximum likelihood estimation method proposed by Smith, Taylor, and Yadav (1997) to estimate the value of d.

Regarding the goodness of fit of AR (p) and MA (q) in the ARFIMA model, this study adopted Akaike's information criterion (AIC). A small AIC value indicated an excellent model fit. In addition, regarding the selection range for the goodness of fit of models, the parameters related to the RV of TAIEX, electronic index, and financial and insurance stock index used in this study were estimated within the interval range between ARFIMA (0, d, 0) and ARFIMA (2, d, 2). This study also used the value of the fractional difference parameter d to determine whether the model possessed the long-memory characteristic. If the value of the fractional difference parameter d was significant and satisfied conditions, then the series possessed the long-memory characteristic. Finally, based on the goodness of fit of models, the model with the maximum AIC value was selected from the interval range as the representative ARFIA (p, d, q) model.

Based on the aforementioned assessment principles, the empirical results for the ARFIMA models for TAIEX, electronic index, and financial and insurance stock index are presented as follows: First, regarding the fractional difference parameter d, Table 7 presents the empirical results for the ARFIMA model of TAIEX. As shown in Table 7, the values of the fractional difference parameter d for all periods were significant except for the time frequencies of 10 minutes and 60 minutes in Panel A of Table 7, indicating that RV possessed the long-memory characteristic. In addition, the empirical results for AR and MA were significant and lower time frequencies yielded a better fit. Table 8 presents the results of the parametric estimation for the ARFIMA model of electronic index. As shown in Table 8, the values of the fractional difference parameter d for all periods were significant except for the time frequencies of 10 minutes and 30 minutes for the subperiod of 2007 to 2008 shown in Panel A of Table 8. All periods contained the MA (1) term, probably implying error correction.

Table 9 presents the results of the parametric estimation for the ARFIMA model of the financial and insurance stock index. As shown in Table 9, the values of the fractional difference parameter d for all periods were significant except for the sub period of 2007 to 2008 shown in Panel A of Table 9, indicating that RV possessed the long-memory characteristic.

In summary, this study found that for all the ARFIMA models of TAIEX, electronic index, and financial and insurance indices, all the values of the fractional difference parameter d were significant except for the sub period of 2007 to 2008 shown in Panel A of Tables 7, 8, and 9, indicating that RV possessed the long-memory characteristic. Therefore, compared with the long-memory characteristic of the Taiwan stock and financial markets shown before the financial crisis of 2008, the long-memory characteristic of the Taiwan stock and financial markets shown after the financial crisis of 2008 was more apparent, indicatingthat the impact of pre- and postautocorrelation was persistent. In addition, the empirical results showed that the AIC value decreased as the time frequency for high-frequency data increased, indicating that a higher time frequency of data yielded better model fit.

Table 7. The parametric estimation for the ARFIMA model of TAIEX

Panel A: 2007–2008							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4869			-0.3579		-254.4143	1.0419
	(0.00)***			(0.00)***			
10 min	**0.1121**	0.6238	0.346	-0.6365	-0.1685	-269.4249	1.1146
	(0.66)	(0.04)**	(0.25)	(0.19)	(0.60)		
15 min	0.4845			-0.3665		-285.1715	1.1660
	(0.00)***			(0.00)***			
30 min	0.4832			-0.4066		-332.8868	1.3584
	(0.00)***			(0.00)***			
60 min	**0.0067**	0.9172	00653	-0.8458		-384.6853	1.5753
	(0.98)	(0.00)***	(0.62)	(0.00)***			
Panel B: 2009–2012							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4823			-0.3471		-435.2471	0.8794
	(0.00)***			(0.00)***			
10 min	0.4799			-0.3641		-483.40	0.9758
	(0.00)***			(0.00)***			
15 min	0.4732			-0.3791		-545.2915	1.0997
	(0.00)***			(0.00)***			
30 min	0.4599			-0.3730		-654.5644	1.3184
	(0.00)***			(0.00)***			
60 min	0.4670	-0.2935	0.1671	-0.1644	-0.2620	-839.0108	1.6937
	(0.00)***	(0.19)	(0.01)**	(0.48)	(0.01)**		
Panel C: 2007–2012							

	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4919			-0.3618		-688.9514	0.9270
	(0.00)***			(0.00)***			
10 min	0.4907			-0.3697		-756.8253	1.0178
	(0.00)***			(0.00)***			
15 min	-0.2581	1.0335	-0.0381	-0.6790		-822.4519	1.1083
	(0.08) *	(0.00)***	(0.50)	(0.00)***			
30 min	0.4859			-0.4063		-985.5009	1.3237
	(0.00)***			(0.00)***			
60 min	0.4855	-0.2199	0.1571	-0.2323	-0.2250	-1,227.7175	1.6518
	(0.00)***	(0.30)	(0.00)***	(0.27)	(0.03)**		

Notes:The symbols ***, **, and * represent the 1%, 5%, 10% significance levels, respectively.
The values in parentheses are Pvalues.

Table 8. The parametric estimation for the ARFIMA model of electronic index

Panel A: 2007–2008							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4870			-0.3726		-237.9797	0.9757
	(0.00)***			(0.00)***			
10 min	**0.1294**	0.9741		-1.0111	0.15	-267.6557	1.1035
	(0.74)	(0.00)***		(0.01)**	(0.48)		
15 min	0.4836			-0.3889		-286.8855	1.1729
	(0.00)***			(0.00)***			
30 min	**-0.0719**	0.9843		-0.8824	0.0816	-339.0252	1.3912
	(0.82)	(0.00)***		(0.00)***	(0.56)		
60 min	0.467			-0.3834		-376.0248	1.5324
	(0.00)***			(0.00)***			
Panel B: 2009–2012							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4817			-0.3684		-392.4193	0.7936
	(0.00)***			(0.00)***			
10 min	0.4750			-0.3720		-452.7362	0.9144
	(0.00)***			(0.00)***			
15 min	0.4836			-0.3889		-286.8855	1.1729
	(0.00)***			(0.00)***			
30 min	0.4658			-0.3832		-518.2624	1.0456
	(0.00)***			(0.00)***			
60 min	0.4586	-0.3282	0.1718	-0.1090	-0.2845	-801.4483	1.6185
	(0.00)***	(0.21)	(0.02)**	(0.68)	(0.01)**		
Panel C: 2007–2012							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4921			-0.3813		-630.3863	0.8487
	(0.00)***			(0.00)***			
10 min	0.4898			-0.3831		-725.1143	0.9754
	(0.00)***			(0.00)***			
15 min	0.4885			-0.4054		-804.6156	1.0818
	(0.00)***			(0.00)***			
30 min	-0.3183	1.512	-0.0552	-0.6659		-976.4616	1.3143
	(0.01) **	(0.00)***	(0.29)	(0.00)***			
60 min	0.4833	0.1807	0.0970	-0.6231		-1179.5327	1.5860
	(0.00)***	(0.02) **	(0.03)**	(0.00)***			

Notes:The symbols ***, **, and * represent the 1%, 5%, 10% significance levels, respectively.
The values in parentheses are Pvalues.

Table 9. The parametric estimation for the ARFIMA model of financial and insurance stock index

Panel A: 2007–2008							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	**0.2303**	0.974		-1.0737	0.1763	-261.4114	1.0783
	(0.44)	(0.00)***		(0.00)***	(0.00)***		
10 min	0.4858			-0.3181		-284.894	1.1649
	(0.00)***			(0.00)***			
15 min	0.4848			-0.3471		-318.474	1.3003
	(0.00)***			(0.00)***			
30 min	**-0.0545**	0.9887		-0.9026	0.0858	-362.0127	1.4839
	(0.86)	(0.00)***		(0.00)***	(0.57)		
60 min	0.4755			-0.402		-424.0001	1.7258
	(0.00)***			(0.00)***			
Panel B: 2009–2012							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC

5 min	0.4596	-0.2272				-378.9767	0.7667
	(0.00)***	(0.00)***					
10 min	0.4502	-0.2476				-475.2095	0.9594
	(0.00)***	(0.00)***					
15 min	0.4174	-0.2442				-559.8980	1.1289
	(0.00)***	(0.00)***					
30 min	0.4487			-0.3239		-694.8116	1.3990
	(0.00)***			(0.00)***			
60 min	0.4270			-0.3833		-892.1009	1.7940
	(0.00)***			(0.00)***			
Panel C: 2007–2012							
	d	AR(1)	AR(2)	MA(1)	MA(2)	LLH	AIC
5 min	0.4651	-0.2651				-651.7554	0.8773
	(0.00)***	(0.00)***					
10 min	0.4596	-0.2692				-761.7312	1.0244
	(0.00)***	(0.00)***					
15 min	0.4842			-0.3309		-877.7990	1.1797
	(0.00)***			(0.00)***			
30 min	0.4800			-0.3843		-1060.6982	1.4243
	(0.00)***			(0.00)***			
60 min	0.4647			-0.4135		-1315.3508	1.7650
	(0.00)***			(0.00)***			

Notes:The symbols ***, **, and * represent the 1%, 5%, 10% significance levels, respectively.
The values in parentheses are Pvalues.

2. Implied volatility

Regarding the calculation of IV, this study adopted the Black–Scholes option valuation model to analyze and annualize at-the-money options. Regarding data use, because of the characteristics of Taiwanese option contracts, this study used the option contracts (call options) in recent months to calculate IV. The sample option contracts were TXO, TEO, and TFO.

Table 10 presents the descriptive statistics for IV,which were relatively greater than those for TXO and TEO. The IV for TFO was also higher than those for TXO and TEO. Compared with the RV shown in a table10 presented previously, IV apparently approximated a normal distribution to a higher degree. Similar to the previously presented results, the means and standard deviations of financial and insurance stock index and TFO were the highest, indicating a large volatility.

Table 10. Descriptive statistics for IV

	TXO	TEO	TFO
Mean	22.03	23.12	28.38
Median	19.51	20.31	26.05
Standard deviation	10.65	10.47	11.83
Maximum value	74.46	82.93	86.23
Minimum value	3.41	4.27	5.28
Skewness	1.32	1.16	1.08
Kurtosis	5.64	5.02	4.58
JB value	866.88***	345.67***	445.32***

Note:The symbols ***, **, and * represent the 1%, 5%, 10% significance levels, respectively.

3. Comparison of predictive power of various models

This study used TAIEX, electronic index, and financial and insurance stock index as the research targets. In this study, the data pattern for stock index was intraday high-frequency data. The time frequencies used in this study included five time frequencies (i.e., 5, 10, 15, 30, and 60 minutes). This study compared the predictive power of various models for forecasting RV and the forecast range was 1 trading day. This study considered the data for the 250 trading days in 2012 as out-of-sample forecasting data. Table 11 shows the comparison results for the predictive power of various volatility forecasting models. The empirical results revealed that when the forecast range was 1 day, the time series model, ARMA (p, q)–GARCH (1, 1), yielded the minimum forecast error, indicating an excellent predictive power. As shown in Table 11, when the time frequency was lower than 30 minutes, the long-memory model, ARFIMA, was superior to the ARMA (p, q)–GARCH (1,1) model only inMAPE forecasts. Overall, the IV model presented the worst forecasting performance.

Regarding the predictive powerof volatility forecasting models, when the forecast range was 1 day, the predictive power of the short-memory model, ARMA (p, q)–GARCH (1, 1), surpassed those of other models. However, this study inferred that the capability of the long-memory model should improve as the forecast time frequency decreases (longer than 1-day interval). The overall assessment results showed that in all situations, the predictive power of the IV model was the worst,which may have been because of the sample data pattern adopted and the calculation formula used in this study or because of other related factors disregarded in this study. We suggest that further studies be cautious in these respects.

Table 11. Predictive power assessment

	GARCH(1,1)	ARMA-GARCH(1,1)	ARFIMA	IV
t = 5				
RMSE	0.4689	**0.3851**	0.5124	0.5044
MAE	0.3724	**0.3055**	0.4021	0.4245
MAPE	14.7803	**11.7415**	14.8150	16.9434
t = 10				
RMSE	0.4930	**0.4089**	0.5384	0.5230
MAE	0.3942	**0.3201**	0.4216	0.4424
MAPE	15.9351	**12.4571**	15.2870	17.8198
t = 15				
RMSE	0.5252	**0.4415**	0.5631	0.5466
MAE	0.4242	**0.3467**	0.4715	0.4578
MAPE	17.4950	**13.7506**	16.0550	18.8088
t = 30				
RMSE	0.5799	**0.5918**	0.6210	0.6503
MAE	0.4644	**0.4767**	0.5219	0.4912
MAPE	20.6685	**21.5714**	17.3010	21.8848
t = 60				
RMSE	0.6498	**0.5125**	0.6654	0.6502
MAE	0.5130	**0.4433**	0.5781	0.5098
MAPE	25.4509	**20.9927**	18.5430	24.6417
t = 5				
RMSE	0.4336	**0.3588**	0.4492	0.4558
MAE	0.3443	**0.2834**	0.3218	0.3651
MAPE	12.7388	**10.1560**	12.5440	13.3779
t = 10				
RMSE	0.4695	**0.3899**	0.4776	0.4789
MAE	0.3688	**0.3062**	0.3361	0.3841
MAPE	13.8312	**11.1078**	13.1310	14.2194
t = 15				
RMSE	0.4856	**0.4191**	0.5113	0.5017
MAE	0.3892	**0.3312**	0.4582	0.3989
MAPE	15.0867	**12.4071**	14.0110	15.2472
t = 30				
RMSE	0.5227	**0.4622**	0.5685	0.5404
MAE	0.4199	**0.3743**	0.4475	0.4369
MAPE	17.1816	**14.6893**	15.5230	17.6040
t = 60				
RMSE	0.6201	**0.5472**	0.6119	0.5983
MAE	0.4861	**0.4262**	0.5024	0.4633
MAPE	21.9274	**18.3180**	16.8690	20.1545
t = 5				
RMSE	0.4195	**0.3552**	0.5531	0.4660
MAE	0.3416	**0.2935**	0.4322	0.3905
MAPE	12.4182	**10.4451**	14.3310	14.3636
t = 10				
RMSE	0.4833	**0.4087**	0.6153	0.5203
MAE	0.4004	**0.3312**	0.5217	0.4331
MAPE	15.0480	**12.1160**	15.5512	16.4805
t = 15				
RMSE	0.5324	**0.4504**	0.6378	0.5491
MAE	0.4364	**0.3715**	0.5025	0.4571
MAPE	16.7646	**13.7775**	16.9821	17.6488
t = 30				
RMSE	0.6063	**0.5212**	0.7106	0.6132
MAE	0.4828	**0.4186**	0.5966	0.4960
MAPE	20.1127	**16.6177**	17.7601	20.5106
t = 60				
RMSE	0.6872	**0.5923**	0.7324	0.6557
MAE	0.5502	**0.4664**	0.6154	0.5284
MAPE	24.4161	**19.7786**	18.499	23.1640

Note: Numbers in bold are the minimum forecast errors.

V. CONCLUSION AND SUGGESTIONS

Volatility measurement and forecast accuracy are extremely crucial for determining the investment strategies related to portfolio management, asset pricing, or risk assessment. Volatility is mainly used to assess changes in underlying asset prices. This study involved forecasting the RV for TAIEX, electronic index, and financial and insurance stock index. Among the various volatility forecasting models applied to the data at various time frequencies, this study intended to identifythe forecasting models suitable for the Taiwan stock market.

Different from previous studies, this study used intraday high-frequency data to calculate RV. This study organized high-frequency data into five types of data at five time frequencies to investigate whether the predictive power of various models varied as time frequency varied. In addition, to avoid the errors caused by excess in-the-money or out-the-money options, the IV used in this study was derived using the concept of at-the-money options (call options). This study substituted higher, middle, and lower strike prices of at-the-money options into the Black–Scholes model to obtain IV. Then, the weighted average IV for the three strike prices was calculated. This weighted average IV was the IV defined in this study.

For the three stock indices used in this study, the volatility of financial and insurance stock index was more susceptible to market responses than those of the other two stock indices. Regarding the predictive power of time series models, the short-memory model, ARMA (p, q)–GARCH (1, 1) presented the optimal forecasting performance. Nevertheless, for lower time frequencies, the long-memory model, ARFIMA, presented more accurate forecasting performance than the short-memory model.

Overall, the sample data for stock index in this study were intraday high-frequency data and the forecast range was 1 day. Regarding predictive power, ARMA (p, q)–GARCH (1, 1) was the optimal forecast model and the IV model presented the worst forecasting performance. We suggest that for further studies, researchers can expand the prediction time range or employ other suitable estimation methods to obtain excellent forecast results. In addition, by including a regime-switching model for comparison, the market information implied by volatility under various regimes may be revealed through a regime-switching process.

REFERENCES

[1]. Andersen, T. G., Bollerslev, T., 1998, “Answering the skeptics: Yes, standard volatility models do provide accurate forecasts,”*International Economic Review 39 (4)*, 885-905.

[2]. Arek Ohanissian, Jeffrey R Russell and Ruey S Tsay, 2008, “True or Spurious Long Memory? A New Test,” *Journal of Business and Economics Statistics 26(2)*, 161-175.

[3]. Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen, 1996. “Fractionally integrated generalizedautore-gressive conditional heteroskedasticity,” *Journal of Econometrics74*, 3–30.

[4]. Black, F., Scholes, M. S., 1973, “The Pricing of Options and CorporateLiabilities,” *Journal of Political Economy81*, 637-659.

[5]. Bollerslev,T., 1986, “Generalized utoregressive conditional heteroskedasticity,” *Journal of Econometrics*31,307-327.

[6]. Brunetti C. and Gilbert C. L., 2002,“Bivariate FIGARCH and Fractional Cointegration,”*Journal of Empirical Finance* 7, 509-530.

[7]. Blair, B. J. Poon, S. H.and Taylor, S. J., 2001, “Forecasting S&P 100volatility: The incremental information content of implied volatilities and high frequency index returns,”*Journal of Econometrics*105,5-26.

[8]. Christensen B. J. and Prabhala N. R., 1998, “The Relation between Implied and Realized Volatility," *Journal of Financial Economics*50,125-150.

[9]. Choi, K., Zivot, E., 2007,“Long memory and structural changes in the forward discount: An empirical investigation,”*Journal of International Money and Finance*26, 342-363.

[10]. Choi, Kyongwook & Yu, Wei-Choun & Zivot, Eric, 2010,“Long memory versus structural breaks in modeling and forecasting realized volatility,”*Journal of International Money and Finance* 29(5), 857-875.

[11]. Drost, F.C.and NijmanT.E.,1993, “Temporal Aggregation of GARCHProcesses,” *Econometric* 61, 909-927.

[12]. Engle,R.F., 1982,“Autoregressive Conditional Heteroskedasticity with Estimates ofThe Variance of UK Inflation,”*Econometrica*50,987-1007.

[13]. Fleming, J., 1998,“The Quality of Market Volatility ForecastsImplied by S&P 100 Index Option Prices,”*Journal of EmpiricalFinance*5, 317-345.

[14]. Granger, C.W.J. and Joyeux R., 1980, “An Introduction To Long- MemoryTime Series Models And Fractional Differencing,”*The Journal of Time Series Analysis*1,15-29.

[15]. Giot, P., 2005,“Relationships between Implied Volatility Indexes and StockIndex Returns,”*Journal of Portfolio Management* 31,92-100.

[16]. Guglielmo M. C.and Gil-AlanaLuis A., 2010, “Long Memory and Fractional Integration in High Frequency Financial Time Series,”*Discussion Papers of DIW Berlin*, 1016.

[17]. Hull, J. and White, A.,1987, “The Pricing of Options on Assets with Stochastic Volatilities,”*Journal of Finance*,42, 281-300.

[18]. Harvey, C. and WhaleyR., 1991,“S&P 100 Index Option Volatility,” *TheJournal of Finance*46,1551-1561.

[19]. Hosking, J. R. M., 1981, “Fractional Differencing,”*Biometrika 68*, 165-176.

[20]. Hung, Chih-Hsing, Shyh-Weir Tzang and So-De Shyu, 2009, "Forecasting Efficiency of Implied Volatility and the Intraday High-Low Price Range in Taiwan Stock Market," *International Research Journal of Finance and Economics27*, 21-33.

[21]. Jorion, P., 1995,“Predicting Volatility in the Foreign Exchange Market,”*Journal of Finance*50,507-528.

[22]. Mandelbrot, B.B. and VanNessJ.W., 1968, “Fractional Brownianmotions, fractional noises and applications,”*SIAM Review*10, 422-437.

[23]. Maheu, J. M. and McCurdy, T. M., 2011,“Do high-frequency measures of volatility improve forecasts of return distributions?,”*Journal of Econometrics* 160(1), 69-76.

[24]. Pong, S., Shackleton, M. B., Taylor, S. J., Xu, X.,2004,“Forecasting currency volatility: Acomparison of implied volatilities and AR(FI)MA models,”*Journal of Banking and Finance28(10)*, 2541–2563.

[25]. Poon, S. H. and Granger C. W. J., 2003,“Forecasting Financial Market Volatility: A Review,”*Journal of Economic Literature*41(2),478-539.

[26]. Smith, J., Taylor, N. and Yadav, S., 1997,“Comparing the bias and misspecificationin ARFIMA models,”*Journal of Time Series*

Analysis18, 507-527.
[27]. Wash, David M., Tsou, Glenn Yu-Gen, 1998, "Forecasing index volatility: Sample interval and No-trading effects,"*Applied Financial Economics 8*, 477-485.
[28]. Wong, W. K., M. C. Chang, A. H. Tu,2009, "Are magnet effects caused by uninformed traders? Evidence from Taiwan Stock Exchange," *Pacific-Basin Finance Journal17*, 28-40.
[29]. Yang, C.W., Ming-Jeng Hwang, Bwo-Nung Huang, 2002. "An Analysis of Factors Affecting Price Volatility of the US oil Market," Energy Economics 24, 107-119.
[30]. Zhou, Bin, 1996, "High-Frequency Data and Volatility in Foreign-Exchange Rates,"*Journal of Business & Economic Statistics 14(1)*, 45-52.