

# “The predictive power of volatility models: evidence from the ETF market”

## AUTHORS

Chang-Wen Duan  
Jung-Chu Lin

## ARTICLE INFO

Chang-Wen Duan and Jung-Chu Lin (2014). The predictive power of volatility models: evidence from the ETF market. *Investment Management and Financial Innovations*, 11(2)

## RELEASED ON

Tuesday, 17 June 2014

## JOURNAL

"Investment Management and Financial Innovations"

## FOUNDER

LLC “Consulting Publishing Company “Business Perspectives”



NUMBER OF REFERENCES

0



NUMBER OF FIGURES

0



NUMBER OF TABLES

0

© The author(s) 2025. This publication is an open access article.

Chang-Wen Duan (Taiwan), Jung-Chu Lin (Taiwan)

## The predictive power of volatility models: evidence from the ETF market

### Abstract

This study uses exchange-traded fund (ETF) data to investigate the ability of the time-series volatility model, the implied volatility model, and the intraday return volatility model to forecast return volatility. Among various ETFs, we adopt NASDAQ 100 Index Tracking Stock (QQQ) as the sample because it has corresponding volatility index (VIX) issued which is necessary. The results show that all volatility models applied in this study can reliably forecast volatility. The Giotto-Jagannathan-Runkle GARCH model is superior to the GARCH model, implying that the return volatility of QQQ is asymmetric. Among the added incremental information, QQQ Volatility Index (QQV) of the American Stock Exchange has better ability in forecasting the return volatility of QQQ, followed by the NASDAQ Volatility Index (VIX) of the Chicago Board Options Exchange, and then by the intraday return volatility. The probable reason is that the turnover of QQQ options is higher than that of the NASDAQ 100 Index Options (NDX) and causes QQQ to contain substantially more information than VIX and to predict volatility better. We also find the predictive power of the time-series GARCH model is weaker than that of the volatility model with QQQ embedded as incremental information. Since QQQ, as an ETF, has diversified its non-systematic risks, the GARCH model using non-systematic risk information to predict volatility is inevitably worse than that using implied volatility. Identical results are achieved when examining out-of-sample forecasting performance.

**Keywords:** volatility model, implied volatility, volatility index, incremental information.

**JEL Classification:** G14, G17.

### Introduction

Developing a feasible volatility model to assist with describing and predicting the volatility of returns on financial assets has long been a focus of research. This is because volatility underpins the risks, pricing, and allocation of assets. Among the various incremental informational variables embedded in volatility models, it is widely believed that if the options market is informationally efficient, implied volatility extracted from options prices is the optimal predictor of future volatility (implied volatility hypothesis). However, the empirical evidence on whether options prices or historical data in time-series models contains much more information about future volatility is mixed.

Lamoureux and Lastrapes (1993) track 10 individual stock options to test several volatility models. Their results reject the orthogonality restriction that the forecast from time-series models should not have predictive power on top of implied volatility and are thus inconsistent with the implied volatility hypothesis. Such contradiction motivates a lot of subsequent studies to explore the performance of implied volatility using different data and more complete methods.

According to Rubenstein (1994), fundamental structural change occurred in options markets after the US market crash of October 1987. Options prices contain much more valuable information than other asset prices since then. Relevant studies, e.g. Day and Lewis (1992), Christensen and Prabhala

(1998), Fleming (1998), and Mayhew and Stivers (2003), among others, find that options with higher trading volume provide more information on future volatility. To improve the predictive ability of time-series volatility models, studies also include high-frequency data in the models. A key finding is that if the research samples consist of individual stocks, indices, or options with high trading volume or high frequency data, the volatility models using these samples gain informational superiority over those models that use only historical return data (Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold and Ebens, 2001; Mayhew and Stivers, 2003).

Past studies focus their attention on individual stocks or composite indices while testing the implied volatility hypothesis. Exchange-traded fund (ETF), as a highly innovative and convenient instrument for spot index trading, has become the highlight of market transactions. Yet few studies target the ETF market to investigate the predictability of volatilities extracted from various models. To address the literature gap, this paper aims to explore the performance of various volatility models using ETF data. Among various ETFs, we adopt an ETF that has corresponding options issued as sample in order that the implied volatility can be calculated. Moreover, two types of options correspond to the ETF: one is based on the underlying index the ETF track, and the other is based on the ETF itself. This study thus chooses an ETF that has both types of options issued in order to explore implied volatility of which type can better predict the return volatility. Finally, since ETF, as a

fund tracking some index, has diversified its non-systematic risks, we are interested in investigating whether time-series volatility models which include only non-systematic risk information, e.g. the generalized autoregressive conditional heteroscedasticity (GARCH) model, can provide all relevant information on future volatility of ETFs.

American Stock Exchange (AMEX) issued a volatility index (VIX) based on NASDAQ 100 Index Tracking Stock (QQQ) options on January 27, 2001. This VIX, known as the AMEX QQQ Volatility Index (QQV), is thus an index for the implied volatility of QQQ options. On the other hand, the underlying index that the QQQ tracks, i.e. the NASDAQ 100 Index, has corresponding NASDAQ 100 Index Options (NDX) issued. Chicago Board Options Exchange (CBOE) also compiles a volatility index, known as the CBOE NASDAQ Volatility Index (VXN), to represent the implied volatility of NDX. Therefore, the tracker fund, QQQ, conform to the sample selection standard of this study that an ETF has two types of options issued at the same time. Moreover, according to the evidence of past literature, studies using actively traded commodities as samples can achieve consistent and meaningful results. The ETF QQQ, on this point, has extremely high trading volumes compared to any other stock, e.g. on September 14, 2007, the trading volume of QQQ on the NASDAQ<sup>1</sup> reached US\$99,801,165, ranking the second highest. We thus adopt QQQ as the sample to investigate the predictive power of volatility models and to explore implied volatility extracted from which type of options can better predict the return volatility. The results contribute to the thorough understanding of the predictive power of various volatility indices.

We begin the analysis by comparing the ability of two time series models, the GARCH and Glosten-Jagannathan-Runkle GARCH (GJR-GARCH)<sup>2</sup> models, to explain return volatility. Next, we introduce different volatility indices into models and compare their incremental information effects. Further, we observe whether 5-minute intraday returns provide better information for the volatility of the underlying asset returns than the lags of daily return volatility. The results show that all volatility models applied in this paper have forecasting ability. The GJR-GARCH model is superior to the GARCH model, which implies that the return volatility of QQQ is asymmetric. In terms of incremental information from volatility indices, the

model embedding QQV is better than that incorporating VXN. Finally, we find that since ETF has diversified its non-systematic risks, the time series model using the lag of the error term to predict ETF volatility is inevitably worse than that using implied volatility.

This article proceeds as follows. Section 1 reviews the literature. Sections 2 and 3 describe the research models and methodology, respectively. Section 4 presents the empirical results. The last section presents the conclusions.

## 1. Literature review

The literature regarding the volatility of stock returns usually assumes stock returns stochastic and normally distributed and assumes the variance of stock returns constant. The assumption of constant variance, however, is called into question by many researchers. Fama (1965) considers the distribution of stock prices leptokurtic and fat-tailed and the changes in prices of stock not independent. If a larger volatility appears in a particular period, another larger volatility will follow in a subsequent period, known as volatility clustering. Therefore, the return volatility should not be a constant. Morgan (1976) finds the variance of stock returns varying over time and demonstrates the heteroscedasticity of return volatility on stock time series.

To consider the heteroscedasticity of volatility, Engle (1982) develops the autoregressive conditional heteroscedasticity (ARCH) model. The model defines the distribution of conditional error terms as a normal distribution and lets the conditional variance have a linear relationship with the square of past error terms. Engle (1982) finds that the ARCH model not only improves the predictive performance of the ordinary least square method, but also acquires a more accurate forecast of variance. Bollerslev (1986) adds conditional variance to the ARCH model, extending the ARCH model to the GARCH model. The GARCH model makes the lag structure of the conditional variance more flexible and reasonable. Considering that the return volatility varies over time, Domowitz and Hakkio (1985) use the ARCH model to fit the time-varying variance and tested the existence of time-varying risk premium. Such a model is the so-called ARCH-Mean model. Bollerslev, Engle and Wooldridge (1988) extend the ARCH-Mean model to the GARCH-Mean model.

Black (1976), Christie (1982), and Schwert (1990) indicate that the distribution of return volatilities was asymmetric. They demonstrate that negative return shocks have larger effects on return volatilities than positive return shocks. They also point out that the asymmetry of volatility could be

<sup>1</sup> AMEX was merged with NASDAQ in 1998 and all its three earliest ETFs were moved to NASDAQ to trade.

<sup>2</sup> Glosten, Jagannathan and Runkle (1993).

explained by the financial leverage effect. In other words, negative return shocks make stock prices decline, making the debt/equity ratio rise and resulting in increased future return volatility on stocks. On the contrary, a positive return shock will decrease the volatility. Therefore, in order to describe precisely the asymmetry of return volatility, asymmetric models based on the GARCH model are subsequently proposed, e.g. asymmetric GARCH (AGARCH) of Engle (1990), exponential GARCH (EGARCH) of Nelson (1991), threshold GARCH (TGARCH) of Zakoian (1994), and GJR model of Glosten, Jagannathan and Runkle (1993).

Some studies have demonstrated the implied volatility hypothesis to be true. There is a school of opinion attributes the result to the fact that while estimating time-series volatility models, these studies use only closing price data of stocks in their computation. For instance, Latane and Rendleman (1976) and Schmalensee and Trippi (1978) find that the power of implied volatility is superior to that of historical volatility on forecasting realized volatility. Chiras and Manaster (1978) show that if dividend yields are incorporated, the predictive power of implied volatility is no longer significantly better than that of historical volatility.

The evidence is mixed regarding the implied volatility hypothesis. Day and Lewis (1992) use data of options on S&P 100 index futures to compare the predictive powers of implied volatility, historical volatility, GARCH model, and EGARCH model. They demonstrate that the time series models might provide substantially more information than the implied volatility model. Lamoureux and Lastrapes (1993) exploit data of 10 stock options on CBOE and obtain a conclusion identical to that of Day and Lewis (1992). Canina and Figlewski (1993) adopt a data sample drawn from the set of weekly settlement prices of all call options on the S&P 100 Index from 1983 to 1987. They find implied volatility to be a biased and inefficient estimator and incapable of gathering the information the historical volatility contains. Becker, Clements and White (2007) and Becker and Clements (2008) also indicate that historical data subsume important information that is not incorporated in option prices, suggesting that implied volatility has poor performance on volatility forecasting.

However, the implied volatility from the index option has been widely studied and totally different results are obtained. Jorion (1995) reports that implied volatility is superior to historical return volatility in terms of both predictive power and the extent of information content using the data of options on foreign currency futures. Christensen and Prabhala (1998) adopt the same S&P 100 Index

options<sup>1</sup> as those of Canina and Figlewski (1993) as their sample and acquire exactly the opposite results. They find that not only the predictive power of implied volatility is superior to that of historical volatility, but also the implied volatility incorporates substantially more information on future volatility. Christensen and Hansen (2002) further include both in-the-money and out-of-the-money options on the S&P 100 Index to construct a trade weighted average of implied volatilities. They also incorporate the data of the put option. Their results are identical to those of Christensen and Prabhala (1998). In recent studies, the empirical evidence from Becker, Clements and McClelland (2009) and Frijns, Tallau and Tourani-Rad (2010) documents that the implied volatilities from index options can capture most of the relevant information in the historical data.

Some studies attempt to uncover the predictive information from intraday data to forecast return volatility. Andersen and Bollerslev (1998) use tick data to compare the predictive powers of ARCH and stochastic volatility models over volatility. They find that both models provide superior volatility forecasts and the use of the high-frequency data contributes to the accuracy of volatility measurements. Andersen, Bollerslev, Diebold and Labys (2001) believe that the five-minute horizon of intraday data is short enough to make the estimated realized volatilities free from measurement error. Moreover, Andersen, Bollerslev, Diebold and Ebens (2001) focus on intraday data for 30 stocks in the Dow Jones Industrial Average to observe the distributions of realized volatilities. They find that the distributions of realized volatilities are highly right-skewed, implying asymmetry of return volatilities. Blair, Poon and Taylor (2001) choose the S&P 100 Index as the underlying to evaluate the predictive power of return volatilities. They incorporate implied volatilities and 5-minute high frequency data to compare their information content in the GJR model. They find that high frequency return data are able to enhance the adaptive and predictive abilities of models, but that the implied VIX of the S&P 100 Index options provides the most accurate forecasts.

Engle and Ng (1993) using the Japanese stock return to compare the EGARCH and GJR models find that the GJR is the best at parsimoniously capturing the asymmetric effect. Mayhew and Stivers (2003) expand their sample to cover 50 individual stocks, which are divided into three sub-samples by trading volume. They employ GARCH and GJR time series models, implied volatility model, and high-frequency intraday return data to compare the forecasting power of incremental information over

---

<sup>1</sup> At-the-money options.

return volatility. Their results indicate that the implied volatility reliably outperforms the GARCH model for both high and low trading volume stocks. In addition, the implied volatility of index options subsumes reliable incremental volatility information and the GARCH model can explain the returns of individual stocks without corresponding options issued. For those stocks that have insufficient volatility information content because of the lower liquidity of the corresponding options, the implied volatility of index options also provides superior incremental information about future firm-level volatility. Liu and Hung (2010) also investigate the performance of various volatility forecasts for the S&P100 stock index series. They compare the symmetric GARCH model with three different types of distributions against GJR-GARCH and EGARCH models. Their empirical results indicate that the GJR-GARCH model achieves the most accurate volatility forecasts.

## 2. Empirical models

The study uses the time-series volatility model, the implied volatility model, and the high-frequency intraday return volatility model to examine the predictive power of various volatilities. Among time-series models, we adopt GARCH, a widely accepted property of volatility, as the basis of the model. According to the empirical evidence of Baillier and Bollerslev (1989), Bollerslev (1987), Engle and Bollerslev (1986), and Hsieh (1989), the existence of volatility clustering in speculative returns is ubiquitous. Many of these studies find that the simple GARCH (1, 1) model provides a decent first approximation of the observed temporal dependencies in daily data. Thus, we adopt the GARCH (1, 1) model as our time-series volatility model.

The model for return dynamics is set and estimated as follows:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \tag{1}$$

$$\varepsilon_t \sim N(0, h_t) \tag{2}$$

$$h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^-) \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{ETF,t-1}^2 + \beta_4 IV_{index,t-1}^2 + \frac{\beta_5 R_{\Sigma 5min,t-1}^2}{1 - \gamma_2 L}, \tag{3}$$

where  $R_t = \ln(p_t/p_{t-1})$ ,  $p$  is the spot price of underlying asset,  $\alpha$ ,  $\beta$  and  $\gamma$  are parameters of the model,  $\varepsilon$  is the error term,  $h$  is the function of generalized conditional return heteroscedasticity;  $D_{t-1}^-$  is a dummy variable that takes a value of 1 if

$\varepsilon_{t-1} < 0$ , otherwise a value of zero;  $L$  indicates the lag operator,  $IV$  is implied volatility,  $R_{\Sigma 5min}^2$  is the sum of squared 5-minute intraday returns of the ETF and a proxy for realized volatility. By restricting  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$  and  $\beta_5$  in equation (3) to a value of zero, this specification nests the GARCH (1, 1) model:

$$h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L}. \tag{4}$$

Allowing for the asymmetric phenomenon of financial asset returns that GARCH (1, 1) is unable to describe, we restrict  $\beta_3$ ,  $\beta_4$  and  $\beta_5$  in equation (3) to a value of zero and thus turn this equation into the standard GJR-GARCH (1, 1) model:

$$h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^-) \varepsilon_{t-1}^2}{1 - \gamma_1 L}. \tag{5}$$

This equation is divided into two subsets by the positive/negative of return errors. We can gauge, through the significance of  $\beta_2$ , whether the explanatory ability of each error term of the two subsets is significantly different from each other. The significance of  $\beta_2$  indicates the asymmetry of volatilities.

For purposes of discussing the predictive effects of implied volatility on stock-return volatility, we restrict  $\beta_1$ ,  $\beta_2$ ,  $\beta_4$  and  $\beta_5$  in equation (3) to a value of zero and thus turn the conditional heteroscedasticity model into a single-factor volatility model that considers only one factor, the implied volatility of the underlying:

$$h_t = \beta_0 + \beta_3 IV_{ETF,t-1}^3, \tag{6}$$

where  $IV_{ETF}$  is the implied variance from options on the ETF; the statistical significance of  $\beta_3$  reveals that the implied volatility has sufficient information content to predict return volatility. Restricting  $\beta_2$ ,  $\beta_5$  and  $\gamma_2$  in equation (3) to a value of zero, we get:

$$h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{ETF,t-1}^3 + \beta_4 IV_{index,t-1}^2 \tag{7}$$

which is an equation for modeling GARCH, but incorporating two other variations to compare the effects of incremental information:  $IV_{ETF}$  is the implied volatility from options on the ETF, and  $IV_{index}$  is the implied volatility from options on the index. Mayhew and Stivers (2003) indicate that comparing the values of  $\beta_3$  and  $\beta_4$  can help assess which has better predictive power. That is to say, the significance of either coefficient implies that the respective variation has sufficient information content for volatility forecasting.

Andersen and Bollerslev (1998) believe that the incorporation of high frequency intraday data in a volatility model can enhance the model's ability to explain return volatility. Therefore, we subsume both the intraday returns and implied volatilities to find out which has the more powerful incremental information. In accordance with Blair et al. (2001), we choose a 5-minute frequency to sieve returns from everyday trading data for the period 08:30 to 15:00 CST<sup>1</sup>, then square and sum these 5-minute returns to proxy for intraday return volatility; furthermore, in consideration of the overnight effect, the trading data after 15:00 of the previous day are also included to compute the intraday volatility of each day. This volatility model is as follows:

$$h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{ETF,t-1}^3 + \frac{\beta_5 + R_{\Sigma 5min,t-1}^2}{1 - \gamma_2 L}, \quad (8)$$

where  $R_{\Sigma 5min}^2$  is the intraday volatility based on 5-minute returns. If  $\beta_5$  is significant, then the information that intraday data contain has forecasting power.

Mayhew and Stivers (2003) also apply the above model, but allow for lagged daily return shocks ( $\varepsilon_{t-1}^2$ ) and lagged 5-minute sum of squared returns ( $R_{\Sigma 5min,t-1}^2$ ) to have a different lagged decay structure. Thus, this specification enables one to compare volatility information from daily return shocks versus the intraday return volatility of different lags. According to evidence from Mayhew and Stivers (2003), return-shocks and intraday return volatility of older lags add essentially no explanatory power. Therefore, we restrict  $\gamma_2$  to a value of zero and discuss only the power of intraday return volatility of lag 1 to explain daily return volatility while verifying the GARCH+V5 model, a GARCH (1, 1) model with intraday return volatility as its incremental information; that is,  $t - 2$  and older return-shock and intraday return volatility are not considered. Such an empirical method does not model the GJR asymmetry (Mayhew and Stivers, 2003).

### 3. Methodology

AMEX and CBOE began issuing QQQ options on March 3, 1999 and February 27, 2001, respectively. AMEX started to issue QQV on January 31, 2001. To ease the shocks that the issuances of QQQ options and QQV have created in the market, the data period begins at March 30, 2001, two months after the launch of QQV, and ends at June 30, 2003. A total of 27 trading months of the data period provide 563 daily observations. Daily data are used and high-frequency intraday price data of QQQ are

gathered to compute intraday 5-minute return volatility. In addition, due to the measurement errors of implied volatility based on option pricing theory and the smile effect exhibited by the implied volatility, we use two VIXs, QQV of QQQ options and VXN of NASDAQ 100 Index options, compiled with the Whaley (1993) method and by CBOE, to avoid the measurement errors. Options data are from the Prophet Company in the United States and intraday data are from the Tickdata Company.

Whaley (1993) uses the implied volatilities of eight near-the-money options to calculate an implied volatility index. There are four calls and four puts in his sample and pairs of nearest-the-money exercise prices are chosen to calculate a weighted average of eight implied volatilities. As such, the implied volatility of an at-the-money option with a constant 22 trading days to expiry could be constructed. We also have an out-of-sample prediction period for comparing the forecasting ability of the empirical models. Since volatility index, however, is based on the at-the-money options with 22 trading days to expiry, here we allow an out-of-sample prediction period to have 22 trading days to expiry also to facilitate the comparison of the predictive performance.

The numerical method of Berndt, Hall, Hall, and Hausman (1974) is applied to estimate the model parameters. The log-likelihood function value (Log-L) of Bollerslev and Wooldridge (1992) is used in order to compare the fit ability of the models. As regards the issue of comparing the incremental information of various volatilities, the likelihood ratio (LR) tests are employed to execute tests. Finally, three test statistics, mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE), are computed to examine the predictive power of the models while comparing the out-of-sample forecast performance.

### 4. Empirical results

If a time series is non-stationary, the execution of a regression will cause spurious regression. We use an augmented Dickey-Fuller (ADF) test to test the stationarity of the data series. In addition, we perform normality, stability, autocorrelation, and heteroscedasticity tests on variations.

Table 1 presents the descriptive statistics for QQQ. Observing the skewness and kurtosis coefficients, we find that the distribution of QQQ price returns is both right-skewed and leptokurtic. The Jarque-Bera test also indicates that the distribution is not normal. The results evidence the fat-tailed characteristic of QQQ price returns and thus using ARCH or GARCH to describe the heteroscedasticity of price returns appears correct.

<sup>1</sup> US Central Standard Time.

Table 1. Basic statistics of QQQ returns

Statistic variables	Statistic values
Sample size	584
Mean	0.081
Standard deviation	2.713
Maximum	10.163
Minimum	-8.888
Skewness coefficient	0.119
Kurtosis coefficient	3.767
Jarque-Bera test value	13.903**

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

The results of the stability<sup>1</sup>, heteroscedasticity, and autocorrelation tests are reported in Table 2. The result of the unit root test shows that both daily price returns and intraday price returns reject the existence of unit root, indicating their stationary. The *Q* statistic test of Ljung and Box (1978) shows that the *Q* statistic is statistically significant in lag 6, 12, 18, 24, and 30 of the square of residual terms, indicating the existence of heteroscedasticity in residual variance. Furthermore, the results of the ARCH-LM test (Engle, 1982) show the significance of the LM statistic and thus demonstrate the heteroscedasticity of price return residuals. Observing the autocorrelation of residuals, while setting the number of lag term 40, the correlation coefficients of lag 2, 11, 13, 29, and 33 are statistically significant, demonstrating the autocorrelation of return residuals.

Table 2. Statistical test of QQQ

ADF's unit root test				
Intercept and trend included			Only intercept included	
Data type	Lag	Test statistic	Lag	Test statistic
Price return	1	-18.627***	1	-18.605***
Intraday return	4	-11.515***	4	-11.455***

Table 3. GARCH, GJR-GARCH and the conditional variance models with incremental information QQQ

$$\text{Model: } R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, h_t), \quad h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^-) \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{ETF,t-1}^2 + \beta_4 IV_{index,t-1}^2 + \frac{\beta_5 R_{\Sigma 5min,t-1}^2}{1 - \gamma_2 L},$$

where  $R_t = \ln(p_t/p_{t-1})$ ,  $p$  is the spot price of an underlying asset,  $\alpha$ ,  $\beta$  and  $\gamma$  are parameters of the model,  $\varepsilon$  is the error term,  $h$  is the function of generalized conditional return heteroscedasticity;  $D_{t-1}^-$  is a dummy variable that takes on a value of 1 if  $\varepsilon_{t-1} < 0$ , otherwise it takes on a value of 0;  $L$  indicates the lag operator;  $IV_{ETF}$  is QQQ, an implied volatility for QQQ options;  $IV_{index}$  is VXN, an implied volatility for NASDAQ 100 Index options;  $R_{\Sigma 5min}^2$  is the sum of squared 5-minute intraday returns and proxy for intraday volatility.

Coefficient	Model				
	GARCH	GJR-GARCH	QQV	GARCH + QQQ	GJR-GARCH + QQQ
$\beta_0 \times 10^{-4}$	0.538 (0.97)	0.346 (-1.08)	-4.384*** (-6.12)	-2.852*** (-1.78)	-0.234* (-1.22)
$\beta_1$	0.048*** (2.82)	-0.017* (-1.48)		-0.022 (-0.66)	-0.015 (-0.93)
$\beta_2$		0.084*** (3.49)			0.097** (2.89)
$\beta_3$			0.056*** (11.19)	0.034* (1.62)	0.003 (1.32)

<sup>1</sup> We use the Akaike's Information Criterion (AIC) rule, recommended by Engle and Yoo (1987), to choose the best lag term.

Ljung and Box Q test for heteroscedasticity		Autocorrelation test	
Q(lag)	Statistic	Lag	Correlation
Q(6)	77.329***	2	-0.095305**
Q(12)	142.203***	11	-0.081555**
Q(18)	174.801***	13	0.089302**
Q(24)	196.592***	29	0.098872**
Q(30)	237.402***	33	-0.086684**
ARCH-LM test			
T · F statistic	52.60911***		

Note: ADF unit root test uses AIC rule to choose the best lag term. ARCH-LM is the ARCH (6) statistic; \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 3 presents the empirical results of the GARCH, GJR-GARCH and QQQ volatility models. The coefficient  $\beta_2$  of GJR-GARCH is statistically significant, revealing the asymmetry of QQQ return volatility. Comparing the GARCH, GJR-GARCH, and QQQ models, the Log-L value of the QQQ model is the highest, showing that this model has the best predictive power over other return volatility models. Further, we use LR statistics to compare the incremental information effects of GARCH and GJR-GARCH models with QQQ added and find that both LR values of the two models are statistically significant, demonstrating that adding QQQ into the models can enhance the ability to forecast future return volatility. Therefore, QQQ has incremental information for predicting the return volatility of QQQ. Comparing the Log-L values of these two models, the GJR-GARCH model with QQQ embedded provides better predictive power than the GARCH model with QQQ embedded. Nevertheless, the Log-L values of these two models are about the same. The coefficient of QQQ,  $\beta_3$ , is statistically significant only when it is embedded in a GARCH model.

Table 3 (cont.). GARCH, GJR-GARCH and the conditional variance models with incremental information QQV

Coefficient	Model				
	GARCH	GJR-GARCH	QQV	GARCH + QQV	GJR-GARCH + QQV
$\beta_4$					
$\beta_5$					
$\gamma_1$	0.943*** (44.52)	968*** (72.25)		0.449* (1.18)	0.918*** (20.51)
$\gamma_2$					
Log-L	1831.3	1835.9	1842.1	1848.6	1849.5
LR				17.25***	13.55***

Note: Values in parentheses are  $t$  statistics calculated with Bollerslov-Wooldrige robust standard errors; \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

QQQ is an ETF tracking NASDAQ-100 Index. Table 4 shows a very high correlation between QQV and VXN. The correlation coefficient of the two VIX is as high as 0.9847. Figure 1 displays the movements of the QQV, VXN, and QQQ prices. Also, evident from this figure is the high correlation between QQV and VXN. In light of Mayhew and Stivers (2003), the implied volatility subsumes reliable information content about the return volatility of the underlying. Therefore, we further observe whether VXN, a VIX based on the NASDAQ

100 Index, also subsumes information about the price return volatility of QQQ and compare VXN with QQV.

Table 4. Correlation coefficients matrix of QQQ price returns, QQV and VXN

	QQQ	QQV	VXN
QQQ	1	0.0800	0.0650
QQV	0.0800	1	0.9847
VXN	0.0650	0.9847	1

Note: The data period starts from 2001/03/31 to 2003/06/30.

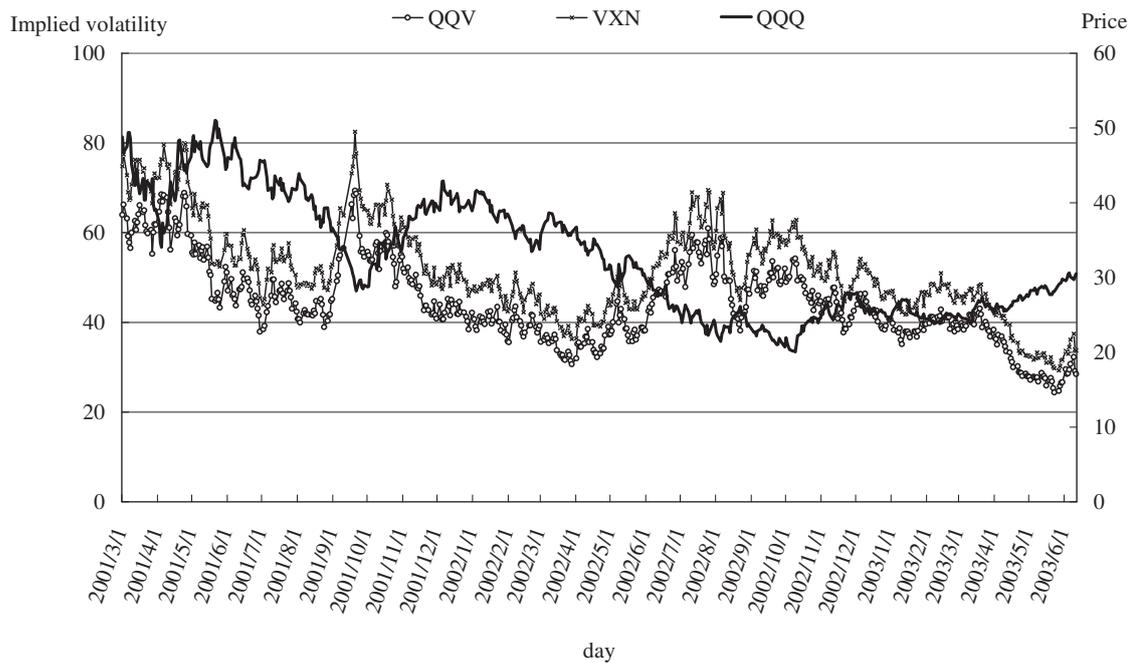


Fig. 1. The dynamics of QQQ prices, QQV and VXN

Table 5 presents the results of executing volatility models with VXN. All the coefficient  $\beta_4$ s are statistically significant, indicating that the VIX of the NASDAQ-100 Index has predictive power over QQQ return volatility. This outcome is identical to that of Mayhew and Stivers (2003). Compared to Table 3, the Log-L value of the VXN model is 1840.93, higher than that of the GARCH (1, 1) model, indicating that the power to predict return volatility in the VXN model is better than that of GARCH. Comparing the power to predict return

volatility of the four models in Table 5, the GARCH model with both QQV and VXN has the strongest power and the most incremental information. If only comparing the extent to which two variations, QQV and VXN, affect return volatility, we find that the coefficients in model QQV + VXN are all statistically significant,  $\beta_3$  equals 0.031,  $\beta_4$  equals 0.023, and  $\beta_3 > \beta_4$ , indicating that QQV's marginal explanatory power for QQQ return volatility is higher than that of VXN. This outcome is identical to that in model GARCH + QQV + VXN.

Table 5. GARCH and the conditional variance model with incremental information QQV and VXN

$$\text{Model: } R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, h_t), \quad h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^-) \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{ETF,t-1}^2 + \beta_4 IV_{index,t-1}^2 + \frac{\beta_5 R_{\Sigma 5min,t-1}^2}{1 - \gamma_2 L},$$

where  $R_t = \ln(p_t/p_{t-1})$ ,  $p$  is the spot price of an underlying asset,  $\alpha$ ,  $\beta$  and  $\gamma$  are parameters of the model,  $\varepsilon$  is the error term,  $h$  is the function of generalized conditional return heteroscedasticity;  $D_{t-1}^-$  is a dummy variable that takes on a value of 1 if  $\varepsilon_{t-1} < 0$ , otherwise it takes on a value of 0;  $L$  indicates the lag operator;  $IV_{ETF}$  is QQV, an implied volatility for QQQ options;  $IV_{index}$  is VXN, an implied volatility for NASDAQ 100 index options;  $R_{\Sigma 5min}^2$  is the sum of squared 5-minute intraday returns and proxy for intraday volatility.

Coefficients	Models			
	VXN	GARCH + VXN	QQV + VXN	GARCH + QQQ + VXN
$\beta_0 \times 10^{-4}$	15.293*** (9.61)	-7.062** (-2.20)	-5.947*** (-5.09)	-8.715*** (-6.22)
$\beta_1$		-0.389 (-1.27)		-0.049** (-1.67)
$\beta_2$				
$\beta_3$			0.031*** (2.83)	0.048*** (4.33)
$\beta_4$	0.015*** (11.61)	0.067** (2.18)	0.023* (1.58)	0.029*** (2.96)
$\beta_5$				
$\gamma_1$		0.232 (0.38)		0.269 (0.92)
$\gamma_2$				
Log-L	1840.9	1849.9	1848.8	1851.1
LR		18.51***	10.83	19.72***

Note: Values in parentheses are  $t$  statistics calculated with Bollerslov-Wooldridge robust standard errors; \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

Finally, we incorporate intraday 5-minute return volatility data into the models to observe whether intraday information enhances the modeling and forecasting of the QQQ return volatility. Table 6 shows that the coefficient  $\beta_5$  of the intraday 5-minute return volatility (V5) model is statistically significant, revealing that intraday return volatility can predict the volatility of QQQ returns. Compared to that of the GARCH (1, 1) model in Table 3, the Log-L value of the V5 model is 1838.4 and is larger than that of the GARCH (1, 1) model, which uses daily data, indicating that using high-frequency data could indeed reduce noise<sup>1</sup> from price returns and enhance the ability to forecast the volatility of returns. Compared to the QQV model in Table 3 and the VXN model in Table 5, the Log-L value of the V5 model is the smallest, showing that the V5 model does not provide better information content than the QQV or VXN models. The model that has the highest Log-L value is the QQV model.

Furthermore, we add the intraday return volatility variable information to the GARCH model. The coefficient  $\beta_5$  of GARCH + V5 in Table 6 is statistically significant – the coefficient  $\beta_1$  of the residual term is smaller than that of GARCH, and  $\gamma_1$  drops from 0.943 in GARCH to 0.645 in GARCH + V5, indicating that the model with extra volatility information on QQQ intraday returns has a better

ability to forecast return volatility than GARCH. In terms of the incremental information effects of intraday returns, the LR statistic of the GARCH + V5 model in Table 6 is 9.16 and is statistically significant, showing that intraday return volatility can indeed provide more information, increasing the ability of GARCH (1, 1) to forecast volatility. This outcome is identical to those of Blair et al. (2001), and Mayhew and Stivers (2003).

Table 6. GARCH, ARCH and the conditional variance model with intraday incremental information

$$\text{Model: } R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, h_t),$$

$$h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^-) \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{ETF,t-1}^2 + \beta_4 IV_{index,t-1}^2 + \frac{\beta_5 R_{\Sigma 5min,t-1}^2}{1 - \gamma_2 L},$$

where  $R_t = \ln(p_t/p_{t-1})$ ,  $p$  is the spot price of an underlying asset,  $\alpha$ ,  $\beta$  and  $\gamma$  are parameters of the model,  $\varepsilon$  is the error term,  $h$  is the function of generalized conditional return heteroscedasticity;  $D_{t-1}^-$  is a dummy variable that takes on a value of 1 if  $\varepsilon_{t-1} < 0$ , otherwise it takes on a value of 0;  $L$  indicates the lag operator;  $IV_{ETF}$  is QQV, an implied volatility for QQQ options;  $IV_{index}$  is VXN, an implied volatility for NASDAQ 100 index options;  $R_{\Sigma 5min}^2$  is the sum of squared 5-minute intraday returns and proxy for intraday volatility.

Coefficients	Models		
	V5	GARCH + V5	ARCH + QQQ + V5
$\beta_0 \times 10^{-4}$	0.053 (0.93)	0.059** (1.19)	-4.967*** (-6.90)

<sup>1</sup> Referring to Andersen and Bollerslev (1998).

Table 6 (cont.). GARCH, ARCH and the conditional variance model with intraday incremental information

Coefficients	Models		
	V5	GARCH + V5	ARCH + QQV + V5
$\beta_1$		0.01*** (2.36)	0.015** (1.70)
$\beta_2$			
$\beta_3$			0.06*** (11.01)
$\beta_4$			
$\beta_5$	0.044** (1.32)	0.055** (1.48)	0.029** (2.06)
$\gamma_1$		0.645*** (38.31)	
$\gamma_2$	0.946*** (46.01)		
Log-L	1838.4	1840.5	1850.1
LR		9.16***	

Note: This paper discusses only the predictive power of lag 1 return residuals and intraday returns (V5) over volatility of QQQ returns, and hence we restrict  $\gamma_2$  to a value of zero. Values in parentheses are  $t$ -statistics calculated with Bollerslov-Wooldridge robust standard errors; \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

Since Andersen and Bollerslev (1998) use intraday data to forecast future volatility and find that the ARCH model performs better, we incorporate QQV and V5 into the ARCH model and find that all coefficients are statistically significant.  $\beta_3$  is greater than  $\beta_5$ , showing again that the ability of QQV to forecast the return volatility of QQQ is superior to that of intraday return volatility.

Table 7 presents the comparison of the performance of various volatility models in out-of-sample forecasts. The results indicate that the error is smallest for the QQV model, demonstrating that using QQV to forecast QQQ return volatility has a smaller prediction error.

Table 7. Predictive errors

Performance index	Models				
	GARCH	GJR-GARCH	QQV	VXN	V5
RMSE	0.563	0.587	0.418	0.441	0.452
MAE	0.474	0.498	0.366	0.392	0.399
MAPE	43.667	44.705	26.951	27.975	30.700

## Conclusion

This paper investigates the ability of various return volatility models to forecast future return volatility of QQQ, an ETF with diversification advantages. We compare the time series volatility model, the implied volatility model, and the intraday return

volatility model, in an effort to determine which has the best predictive power for the volatility in the ETF market. The evidence shows that all empirical volatility models considered in this study have predictive power to forecast volatility, but employing only a GARCH model to forecast volatility cannot subsume all information content. Previous studies demonstrate the asymmetry of return volatility. Thus, when describing the time-varying process of return volatility, it is better to use models that consider this property. The empirical result that the GJR-GARCH model is superior to the GARCH model supports this viewpoint.

Moreover, incremental information incorporated into the models all enhances the ability to forecast return volatility. QQV has the best power to predict the volatility of QQQ returns, VXN is in the second place, and intraday return volatility has the lowest predictive power. Since the trading volume of QQQ is far more than that of NDX, such results are identical to the conclusion of prior studies that the implied volatility of option with higher liquidity would have better predictive power or more information over return volatility than option with lower liquidity.

Since ETFs are index funds that diversify almost all non-systematic risks, using lagged error terms as proxy variables that represent non-systematic risks to predict return volatility of ETFs would not subsume enough information content. The empirical results demonstrate this point of view. Incorporating other incremental information into a GARCH model could increase its ability to forecast future return volatility, indicating that there exist incremental information effects within a GARCH model. QQV is the most valuable among all such incremental information.

One of the major shortcomings of this study is the usage of not so current data set. As a result of using old data set, the findings may be not so robust. Consequently, given that the authors of this study have limitations to more recent data and to enhance the robustness of the findings, a few directions for future research are recommended. First of all, since the current study analyzes old data set, future research can be set up to extend or investigate the topic of this study by using a more recent data set or data sets belonging to other countries. In addition, future research can look into incorporating other econometric techniques such as EGARCH or other asymmetric GARCH models (or regime-switching models) in order to provide more robust findings.

## References

1. Andersen, T.G. and T. Bollerslev (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts, *International Economic Review*, 49, pp. 885-905.
2. Andersen, T.G., T. Bollerslev, F. Diebold and H. Ebens (2001). The Distribution of Realized Stock Return Volatility, *Journal of Financial Economics*, 61, pp. 43-76.

3. Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001). The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association*, 96, pp. 42-55.
4. Baillier, T. and T. Bollerslev (1989). The Message in Daily Exchange Rates: A Conditional Variance Tale, *Journal of Business and Economic Statistics*, 7, pp. 297-305.
5. Becker, R. and A.E. Clements (2008). Are Combination Forecasts of S&P 500 Volatility Statistically Superior? *International Journal of Forecasting*, 24, pp. 122-133.
6. Becker, R., A.E. Clements and A. McClelland (2009). The Jump Component of S&P 500 Volatility and the VIX Index, *Journal of Banking & Finance*, 33, pp. 1033-1038.
7. Becker, R., A.E. Clements and S.I. White (2007). Does Implied Volatility Provide any Information beyond that Captured in Model-based Volatility Forecasts? *Journal of Banking & Finance*, 31, pp. 2535-2549.
8. Berndt, E., B. Hall, R. Hall and J. Hausman (1974). Estimation and Inference in Nonlinear Structural Models, *Annals of Economic and Social Measurement*, 3, pp. 653-665.
9. Black, F. (1976). Studies in Stock Price Volatility Changes, in Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association, pp. 177-181.
10. Blair, B., S. Poon and S. Taylor (2001). Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High-frequency Index Returns, *Journal of Econometrics*, 105, pp. 5-26.
11. Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, pp. 1307-1327.
12. Bollerslev, T. (1987). A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics*, 69, pp. 542-547.
13. Bollerslev, T., R.F. Engle and J.M. Wooldridge (1988). A Capital Asset Pricing Model with Time-Varying Covariances, *Journal of Political Economy*, 96, pp. 116-131.
14. Bollerslev, T. and J. Wooldridge (1992). Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances, *Econometric Reviews*, 11, pp. 143-172.
15. Canina, L. and S. Figlewski (1993). The Informational Content of Implied Volatility, *Review of Financial Studies*, 6, pp. 659-681.
16. Chiras, D.P. and S. Manaster (1978). The Information Content of Option Prices and a Test of Market Efficiency, *Journal of Financial Economics*, 6, pp. 213-234.
17. Christensen, B.J. and N.R. Prabhala (1998). The Relation between Implied and Realized Volatility, *Journal of Financial Economics*, 50, pp. 125-150.
18. Christensen, B.J. and C.S. Hansen (2002). New Evidence on the Implied-realized Volatility Relation, *European Journal of Finance*, 8, pp. 187-205.
19. Christie, A.A. (1982). The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects, *Journal of Financial Economics*, 10, pp. 407-432.
20. Day, T. and C. Lewis (1992). Stock Market Volatility and the Information Content of Stock Index Options, *Journal of Econometrics*, 52, pp. 267-287.
21. Domowitz, I. and C.S. Hakkio (1985). Conditional Variance and the Risk Premium in the Foreign Exchange Market, *Journal of International Economics*, 19, pp. 47-66.
22. Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50, pp. 987-1008.
23. Engle, R.F. (1990). Discussion: Stock Market Volatility and the Crash of '87, *Review of Financial Studies*, 3, pp. 103-106.
24. Engle, R.F. and T. Bollerslev (1986). Modeling the Persistence of Conditional Variances, *Econometric Reviews*, 5, pp. 81-87.
25. Engle, R.F. and V.M. Ng (1993). Measuring and testing the impact of news on volatility, *Journal of Finance*, 48, pp. 1749-1788.
26. Engle, R.F. and B.S. Yoo (1987). Forecasting and Testing in Co-integrated Systems, *Journal of Econometrics*, 35, pp. 143-159.
27. Fama, E. (1965). The Behavior of Stock Market Prices, *Journal of Business*, 38, pp. 34-105.
28. Fleming, J. (1998). The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices, *Journal of Empirical Finance*, 5, pp. 317-345.
29. Frijns, B., Tallau, C. and A. Tourani-Rad (2010). Australian Implied Volatility Index, *Finsia Journal of Applied Finance*, 1, pp. 31-35.
30. Glosten, L., R. Jagannathan and D. Runkle (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48, pp. 1779-1801.
31. Hsieh, D.A. (1989). Modeling Heteroskedasticity in Daily Foreign Exchange Rates, *Journal of Business and Economic Statistics*, 7, pp. 307-317.
32. Jorion, P. (1995). Predicting Volatility in the Foreign Exchange Market, *Journal of Finance*, 50, pp. 507-528.
33. Lamoureux, C.G. and W.D. Lastrapes (1993). Forecasting Stock-return Variances: Toward an Understanding of Stochastic Implied Volatilities, *Review of Financial Studies*, 6, pp. 293-326.
34. Latane, H.A. and R.J. Rendleman (1976). Standard Deviations of Stock Price Ratios Implied in Option Prices, *Journal of Finance*, 31, pp. 369-381.

35. Liu, H.C. and J.C. Hung (2010). Forecasting S&P100 Stock Index Volatility: The Role of Volatility Asymmetric and Distribution Assumption in GARCH Models, *Expert Systems with Applications*, 37, pp. 4928-4934.
36. Ljung, G.M. and G.E.P. Box (1978). On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 65, pp. 297-303.
37. Mayhew, S. and S. Stivers (2003). Stock Return Dynamics, Option Volume, and the Information Content of Implied Volatility, *Journal of Futures Markets*, 23, pp. 615-646.
38. Morgan, I.G. (1976). Stock Prices and Heteroscedasticity, *Journal of Business*, 49, pp. 496-508.
39. Nelson, D.B. (1991). Conditional Heteroscedasticity in Asset Returns: A New Approach, *Econometrica*, 59, pp. 347-370.
40. Rubenstein, M. (1994). Implied Binomial Trees, *Journal of Finance*, 49, pp. 771-818.
41. Schmalensee, R. and R. Trippi (1978). Common Stock Volatility Expectations Implied by Option Premia, *Journal of Finance*, 33, pp. 129-147.
42. Schwert, G.W. (1990). Stock Volatility and the Crash of '87, *Review of Financial Studies*, 3, pp. 77-102.
43. Whaley, R.E. (1993). Derivatives on Market Volatility: Hedging Tools Long Overdue, *Journal of Derivatives*, 1, pp. 71-84.
44. Zakoian, J.M. (1994). Threshold Heteroskedastic Models, *Journal of Economic Dynamics and Control*, 18, pp. 931-955.