Asia-Pacific Journal of Financial Studies

Asia-Pacific Journal of Financial Studies (2012) **41**, 103–124 doi:10.1111/j.2041-6156.2011.01066.x

Empirical Comparison of Alternative Implied Volatility Measures of the Forecasting Performance of Future Volatility*

Dong Woo Rhee
Samsung Asset Management

Suk Joon Byun
Business School, Korea Advanced Institute of Science and Technology

Sol Kim**
College of Business Administration, Hankuk University of Foreign Studies

Received 10 May 2011; Accepted 28 November 2011

Abstract

Implied volatility from the Black and Scholes (*Journal of Political Economy* 81, 1973, p. 637) model has been empirically analyzed for the forecasting performance of future volatility and is well known to be biased. Based on the belief that implied volatility from option prices can best estimate future volatility, this study identifies the best way to derive implied volatility to overcome the forecast bias associated with the Black–Scholes model. For this, the following three models are considered: Heston's model, which best addresses the problems associated with the Black–Scholes model for pricing and hedging options; Britten-Jones and Neuberger's model-free implied volatility (MFIV), which eliminates the model-oriented bias; and VKOSPI, the Korean version of the Chicago Board Options Exchange Market Volatility Index. This study conducts a comparative analysis of implied volatilities from the Black–Scholes model, Heston's model, the MFIV, and VKOSPI for their abilities to forecast future volatility. The results of the empirical analysis of the KOSPI 200 options market indicate that Heston's model can eliminate most of the bias associated with the Black–Scholes model, whereas the MFIV and VKOSPI do not show any improvement in terms of forecasting performance.

Keywords Options; Stochastic volatility; Volatility; VKOSPI; Forecasting

JEL Classification: G13, G14

*Acknowledgments: This work was supported by the Hankuk University of Foreign Studies Research Fund of 2012. We are grateful to the Editor and two anonymous referees for many constructive comments and suggestions.
**Corresponding author: Sol Kim, College of Business Administration, Hankuk University of Foreign Studies, 270 Imun-dong, Dongdaemun-Gu, Seoul, Korea. Tel: +82-2-2173-3124, Fax: +82-2-959-4645, email: solkim@hufs.ac.kr.

1. Introduction

Under the stochastic volatility assumption, Hull and White (1987) suggest that the Black and Scholes (1973) model to derive implied volatility from option prices is the best model for estimating future volatility. Based on this theoretical background, a number of empirical studies have examined whether implied volatility can efficiently forecast future realized volatility. After analyzing cross-sectional data on stock options, Latane and Rendleman (1976) suggest that stocks with high implied volatility are likely to exhibit high ex-post realized volatility. Later studies employing time-series analysis offer contrasting findings on the forecasting performance of implied volatility. Day and Lewis (1992), Lamoureux and Lastrapes (1993), Jorion (1995), and Fleming (1998) suggest that Black–Scholes implied volatility is a biased estimator of future realized volatility.

This raises the question of whether using the Black–Scholes model to derive implied volatility from option prices is the best option even when an option price is believed to reflect the market expectation of future volatility. As is inferred from the famous "volatility smile," the Black–Scholes model does not correctly explain the options market. To overcome this empirical deficiency, previous studies have developed a number of models that make various assumptions about the Black–Scholes model, including assumptions about stochastic volatility, stochastic interest rates, and jumps. Bakshi *et al.* (1997, 2000) show that the stochastic volatility model can best address the problems associated with the Black–Scholes model. Thus, the present study uses a stochastic volatility model to assess the ability of the implied volatility to forecast future volatility. Among various stochastic models, this study uses Heston's (1993) model, not only because it follows the continuous time stochastic volatility model that Kim and Kim (2005) demonstrate to be the best in terms of pricing and hedging efficiency but also because it takes into account the correlation between the volatility of and return on the underlying asset and provides closed-form solutions.

This study fills the gap in previous research by conducting an empirical analysis of the ability of the implied volatility from option prices to forecast future volatility. Poteshman (2000) was the first to demonstrate the ability of the implied volatility to forecast future volatility. Poteshman (2000) used a stochastic volatility model to evaluate the ability of the implied volatility to forecast 1-month-ahead future volatility and employed time-series data on the underlying asset to estimate some of the unobservable parameters of Heston's model by assuming that these structural parameters were constant throughout the sample period. By contrast, our method considers all the estimated parameters to be time-varying variables whose values are determined by option prices on the observation date. According to Bates (1991), time-varying parameters are more valuable than constant parameters not only because time-varying parameters can reflect the market sentiment when they are estimated but also because they can offer the future specification of complex dynamic models.

In addition to Black and Scholes (1973) and Heston (1993) implied volatility, this study considers the model-free implied volatility (MFIV) and VKOSPI as comparative implied volatility measures for assessing the ability of the implied volatility to forecast future volatility. Since Britten-Jones and Neuberger (2000) derived the MFIV, a number of studies have evaluated its ability to forecast future volatility. In particular, Jiang and Tian (2005) demonstrate the superiority of the MFIV over Black–Scholes implied volatility in forecasting performance and conclude that the MFIV performs better than Black–Scholes implied volatility because it eliminates model misspecification. In this regard, the MFIV may be the best estimator of future volatility even if it leads to another bias when implemented. In contrast, introduced in 1993, the Chicago Board Options Exchange Market Volatility Index (VIX) is a well-known measure of risk in stock index options markets. VIX is considered a special case of the model-free approach, and its ability to forecast future realized volatility has attracted considerable interest. In addition, some studies, such as Corrado and Miller (2005), Fleming *et al.* (1995) and Eom *et al.* (2008) suggest the superiority of VIX over historical volatility in forecasting future realized volatility. In this study, we investigate KOSPI 200 options data, and compare VKOSPI, the Korean version of VIX, with other implied volatilities.

Finally, the present study investigates the KOSPI 200 options market to examine the ability of the implied volatility to forecast future volatility. Many studies have examined the forecasting performances of future volatility on the KOSPI 200 options market. Chang (2001) reports that the exponentially weighted moving average or generalized autoregressive conditional heteroskedasticity (GARCH) method forecasts future volatility better than implied volatility. Kim and Park (2006) compare the forecasting performances of some implied volatilities on future realized volatility and show that VKOSPI performs the best among them. Eom *et al.* (2008) estimate a realized volatility through intraday data and show that VKOSPI is an unbiased estimator of future realized volatility. The KOSPI 200 options market, the largest stock index options market in the world, provides a unique opportunity for empirical research because of its huge liquidity. As Figlewski (1997) implies, it is important to choose an options market that is liquid because mispriced options in an illiquid market may lead to forecast bias for implied volatility measures of future realized volatility. According to the Futures Industry Association, KOSPI 200 options contracts have accounted for more than 20% of the world's trading volume since 2000. Therefore, it is unlikely that the noise from the options market distorted the present study's findings on the ability of the implied volatility to forecast future volatility because we used data on the KOSPI 200 options market for the period from January 2000 to June 2007.

According to the results, Heston's (1993) implied volatility dominated other implied volatility measures as well as historical volatility in forecasting future volatility. It even eliminated most of the bias associated with Black–Scholes implied volatility. The MFIV and the VKOSPI did not perform better than Black–Scholes implied volatility.

The rest of this paper is organized as follows. Section 2 presents the models and the derivation of Heston's model, the MFIV, and the VKOSPI, followed by the estimation of realized volatility. Section 3 discusses the KOSPI 200 options market and the data and presents a basic statistical analysis of realized volatility, historical volatility, Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV, and VKOSPI on a monthly basis for the final analysis. Section 4 illustrates the relationship between realized volatility and implied volatility. It discusses an econometric analysis model for determining the ability of the implied volatility to forecast future volatility and presents the forecasting results for future realized volatility. Section 5 concludes by summarizing the results and suggesting some interesting avenues for future research.

2. Volatility Model

2.1. Stochastic Volatility Model

Stochastic volatility models are largely divided into two types: continuous time stochastic models and discrete time stochastic models. Hull and White (1987), Johnson and Shanno (1987), Scott (1987), Wiggins (1987), Melino and Turnbull (1990), Stein and Stein (1991), and Heston (1993) belong to the former, whereas Duan (1995) and Heston and Nandi (2000) belong to the latter. Among several stochastic volatility models, this study uses Heston's model for the following reasons. First, it is a continuous time stochastic model, and continuous time stochastic models are known to perform better than discrete time stochastic models (Bakshi *et al.*, 1997, 2000; Kim and Kim, 2005). Second, Heston's (1993) model can provide closed-form solutions and take into account the correlation between the volatility of and return on the underlying asset.[1] The stochastic process of stock prices and their volatility in Heston's model is as follows:

$$\mathrm{d}S = \mu S \mathrm{d}t + \sqrt{v_t} S \mathrm{d}W_S\,, \tag{1}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_v\,, \tag{2}$$

where S is the stock price; μ is the return on the stock; W is the Wiener process; W_S and W_v have a correlation of ρ; v_t is an instantaneous variance at time t; κ is the speed parameter reverting to the long-term average, θ; and σ is the volatility of the variance. Using the Fourier transform under the assumption of the stochastic process described above, the option pricing model can be expressed as:

[1]Previous empirical studies reveal a negative correlation between the volatility of and the return on the underlying asset, a risk-neutral distribution with negative skewness, and a low strike price with high volatility (referred to as the ‘‘volatility sneer’’). This is consistent with the leverage effect (Black, 1976; Christie, 1982). This negative correlation may exist because falling stock prices can bring about a higher debt/equity ratio, which, in turn, can have a leverage effect on the firm, making the volatility of earnings per share higher and eventually amplifying stock price volatility.

$$C = SP_1 - Ke^{-r\tau}P_2, \tag{3}$$

$$P_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{e^{-i\phi \ln[K]} f_j(x, v, \tau; \phi)}{i\phi}\right] \mathrm{d}\phi \quad (j = 1, 2), \tag{4}$$

where C is the call option; K is the exercise price of the call option; r is the risk-free interest rate; τ is time to maturity; $\mathrm{Re}[\cdot]$ is the real number part of a complex number; i is an imaginary number, $\sqrt{-1}$;$f_j(x, v, \tau; \phi) = \exp(A(\tau; \phi) + B(\tau; \phi)v + i\phi x)$;$x = \ln(S)$; and $A(\cdot)$ and $B(\cdot)$ are functions of θ, κ, ρ, and σ.

Various methods can be used to estimate the parameters of Heston's (1993) model. First, the parameters can be estimated from the historical return on underlying assets using maximum likelihood estimation or efficient method of moments. Second, the parameters can be estimated by minimizing the sum of squared percentage errors of the difference between the model price and the actual market price. This method is suggested by Bakshi *et al.* (1997, 2000) and Bates (1991, 2000). Third, the parameters can be estimated using both the time-series data for the underlying asset and the cross-sectional data for the option prices, as suggested by Pan (2002). Previous studies considering the forecasting performance of the implied volatility of the Black and Scholes (1973) model have used the second method, the nonlinear least squares procedure. Therefore, we use the second method, which is presented in equation (5), for comparison with the results of previous research:

$$\min_{\sigma,\theta,\kappa,\rho,v_t} \sum_{i=1}^{N} \left[\frac{O_i^*(t, \tau; K) - O_i(t, \tau; K)}{O_i(t, \tau; K)}\right]^2 \quad (t = 1, \ldots, T), \tag{5}$$

where $O_i^*(t, \tau; K)$ is the model price of option i at time t; $O_i(t, \tau; K)$ is the market price of option i at time t; N is the number of options at time t; and T is the number of days in the sample.

The minimization method of the sum of squared errors between the model and market prices is sometimes used to estimate the parameters using option prices instead of equation (5). This method gives more weight on relatively expensive at-the-money (ATM) or near-the-money options, as a result, makes the worse fit for out-of-the-money (OTM) options. As Lee and Kwon (2001) show, OTM options are more predictive of future realized volatility than ATM options because they are more liquid. Equation (5), which minimizes the sum of squared percentage errors between the model and the market prices, is superior in that it places more weight on relatively cheap OTM options, which are more actively transacted by investors.

2.2. Model-free Implied Volatility

Britten-Jones and Neuberger (2000) show that the variance in asset returns between two finite dates in a risk-neutral world can be described as follows:

$$E_0^Q\left[\int_{\tau_1}^{\tau_2}\left(\frac{\mathrm{d}S_t}{S_t}\right)^2\right] = 2\int_0^{\infty} \frac{C(\tau_2, K) - C(\tau_1, K)}{K^2}\mathrm{d}K, \tag{6}$$

where E^Q indicates the expectation in a risk-neutral world; S_t denotes the stock price at time t; and $C(\tau, K)$ is the price of the call option whose time to maturity and exercise price are τ and K, respectively. We adopted MFIV to compare our results with those of Jiang and Tian (2005) to demonstrate the superiority of the MFIV over Black–Scholes implied volatility in forecasting performance.

Because equation (6) does not assume any stochastic process of the underlying asset, it is referred to as the model-free implied volatility. Some modifications are required to implement the MFIV. First, equation (6) does not consider the interim cash flow that the underlying asset yields. Second, it is assumed that the interest rate is zero. Third, because the forecast period for volatility starts from each observation date in this study, τ_1 has to be set at zero. Finally, strike prices in a real options market exist finitely. As a result, a discrete version is required for an empirical analysis. The following equation reflects these modifications:

$$E_0^Q\left[\sum_{i=0}^{n-1}\left(\frac{S_{t_{i+1}}-S_{t_i}}{S_{t_i}}\right)^2\right]=\left(u-\frac{1}{u}\right)\sum_{j=-m}^{m}\frac{C(\tau,K_je^{r\tau})-\mathrm{Max}[S_0-K_j,0]}{K_j}, \tag{7}$$

where $t_i = i\text{h}$, for $i = 0, 1, 2, \ldots, n$; $h = \tau/n$; $K_j = S_0u^j$, for $j = 0, \pm1, \pm2, \ldots, \pm m$; $u = (1 + k)^{1/m}$; S_0 is the current stock price; r is the risk-free rate; n and m are positive integers; and k is a constant that is greater than zero.

To implement the MFIV, we require the prices of $2m + 1$ call options. Because it is not realistic that call options corresponding to $2m + 1$ exercise prices are traded at the same time, artificial prices are needed. We apply the curve-fitting method to obtain artificial prices from real prices. Following Bates (1991), Campa *et al.* (1998), and Jiang and Tian (2005), this study uses the cubic spline for the curve fitting. Because of the strong nonlinear relationship between option prices and exercise prices, this study applies the curve-fitting method to implied volatility, not to option prices, which is consistent with Shimko (1993), Ait-Sahalia and Lo (1998), and Jiang and Tian (2005), among others. If either k or m is too small, then the approximation error can be significant. This study sets k and m as 0.2 and 20, respectively, because Jiang and Tian (2005) show that the approximation error is too small to determine when k and m are equal to or greater than 0.2 and 20, respectively.

2.3. VKOSPI

VKOSPI is a volatility index indicating the expected volatility of the underlying asset from KOSPI 200 index options. The deriving method of VKOSPI is the same as that of VIX, the volatility index on S&P 500 index options. Since April 2009 when VKOSPI was first disclosed,[2] VKOSPI has been widely used by investors as a barometer of market volatility sentiments.

[2]Instead of obtaining VKOSPI from the Korea Exchange, we generate VKOSPI because time-series analysis starts on a date when VKOSPI is not released.

VKOSPI, which indicates the expected risk-neutral value of realized volatility in the discrete version, can be derived as follows. First, we set the forward index level, which is used to calculate VKOSPI:

$$F = K_{\mathrm{ATM}} + e^{r\tau} \cdot (C_{\mathrm{ATM}} - P_{\mathrm{ATM}}),$$

where K_{ATM} is the exercise price at which the difference between the call option price and the put option price is the smallest; C_{ATM} and P_{ATM} are the call and put option prices, respectively, and correspond to K_{ATM}; r is the risk-free rate; and τ is the amount of time left to expiration of the option in years.

Second, the strike price just below the forward index level is defined as the strike price (K_0) of the ATM option and is used to calculate VKOSPI. When calculating VKOSPI, all available OTM calls and puts are used because OTM options take up most of the trading volume and become in the money (ITM) options as option prices increase. The option price at K_0 is the average price of calls and puts. Calls are selected at a price higher than K_0, and puts are selected at a price lower than K_0. Finally, VKOSPI is calculated using the following equation:

$$\mathrm{VKOSPI} = \sqrt{\frac{2}{\tau}\sum_{i=1}^{m}\frac{\Delta K_i}{K_i^2}e^{r\tau}Q(K_i) - \frac{1}{\tau}\left[\frac{F}{K_0} - 1\right]^2}, \tag{9}$$

where F is the forward index level; K_i is the ith exercise price of an OTM option that is a call option if K_i is greater than F and a put option when K_i is less than F; m is the number of exercise prices; ΔK_i is $(K_{i+1} - K_{i-1})/2$ if i is between 2 and $m - 1$; ΔK_i is $(K_2 - K_1)$ if i is 1; ΔK_i is $(K_m - K_{m-1})$ if i is m; K_0 is the highest exercise price below F; r is the risk-free rate; τ is the time left to the expiration of the option; and $Q(K_i)$ is the transacted option price with the exercise price, K_i.

Because VKOSPI is the 30-day volatility, the interpolation of the volatility of the nearest contract whose maturity is less than 30 days and the volatility of the second nearest contract whose maturity is more than 30 days is typically applied when the time to maturity for the option is not exactly 30 days. However, this interpolation is not necessary in this study because we examine only those options whose time to maturity is 30 days, which is consistent with the maturity targeted by VKOSPI.

2.4. Realized Volatility

In the case of the existing method for estimating volatility through the use of daily prices, the volatility estimation can be distorted because not all price changes are reflected during a certain timeframe. For example, the real volatility can be underestimated in the case in which the intra-day price fluctuation is great or the difference among closing prices is not extreme. By using high-frequency data, we can reduce the estimation error when estimating volatility. Poteshman (2000) shows that by using high-frequency data (e.g. 5-min intervals) instead of existing daily data, approximately half the forecast bias can be eliminated in estimating realized volatility. This is supported by Andersen *et al.* (2003) and Pong *et al.* (2003), who suggest that high-frequency volatility shows better forecasting performance than low-frequency volatility not only in the short term but also in the long term because it reflects more data. Andersen *et al.* (2002, 2001) demonstrate that as the sampling frequency of the underlying returns approaches infinity, volatility estimates are theoretically free from measurement error. Some studies that examine the KOSPI 200 market, such as Lee *et al.* (2005) and Yoo and Koh (2009), also affirm the superiority of using high-frequency return data in estimating the realized volatility. Based on the above arguments, this study estimates both realized volatility and historical volatility using 5-min interval data (which contain more information than daily data) to estimate future realized volatility. The volatility estimation is as follows:

$$\mathrm{Vol} = \sqrt{\frac{1}{\Delta}\frac{1}{L-1}\sum_{i=1}^{L-1}\left[\ln\left({}^{S_{i+1}}/_{S_i}\right)\right]^2}, \tag{10}$$

where Δ is the time interval between i and $i+1$ measured in years; L is the number of stock price observations; and S_i is the stock price at time i.

3. Data

We considered the KOSPI 200 options market to analyze the forecasting performance of various implied volatility measures for future volatility. The KOSPI 200 options market is appropriate for this analysis because it is the largest equity options market in the world in terms of trading volume. Thus, it allows for the efficient determination of option prices. The maturity date for KOSPI 200 options is the second Thursday of the options contract month, and options contract months are consecutive 3 months and 1 more month from March, June, September, and December. There are at least five exercise prices for each options contract month, which can increase as option prices fluctuate. KOSPI 200 options contracts are fully automated and are European options; that is, they can be exercised only at maturity. Because the liquidity of options is concentrated in the nearest contract and because VKOSPI, one of this study's volatility measures, indicates the volatility of 30 days, we selected those options whose time to maturity was 30 days.

We used OTM calls and puts to calculate Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV, and the VKOSPI. We excluded ITM options because of their very low trading volume. Thus, the reliability of transacted ITM options was not fully guaranteed. The data covered the period from 10 January 2000 to 11 June 2007, and we obtained the minute-by-minute transaction prices of KOSPI 200 options from the Korea Stock Exchange. Because there is no widely used benchmark 30-day rate in Korea, we used the 91-day certificate of deposit rate, which is the representative short-term rate in Korea, for the risk-free rate. To filter data, which is necessary for an empirical analysis, we applied the following principles. The last price of each options contract before 14.50 hours[3] on each trading day was used for the empirical analysis. We included only the last transacted option in the sample if the same option was traded several times during any time period. To mitigate the price discreteness effect in the option valuation, we excluded those options whose prices were lower than 0.02. Finally, the prices that did not meet the arbitrage restriction were excluded.

Poteshman (2000) compared the estimation result from daily prices with that from 5-min interval prices in estimating realized volatility. He found a reduction in the estimation error and suggested that the forecasting performance of implied volatility improves when realized volatility is estimated using high-frequency (e.g. 5-min interval) data. Thus, we estimated both realized volatility and historical volatility using 5-min data on the KOSPI 200 index, which were obtained from the Korea Stock Exchange. Because we estimated historical volatility using historical data that preceded the date for which the implied volatility was calculated and because we estimated realized volatility using ex-post data that followed the date for which the implied volatility was calculated, the data period for the KOSPI 200 index (10 January 1999 to 12 July 2007) was longer than that for KOSPI 200 options.

Table 1 shows the descriptive statistics for monthly realized volatility, historical volatility, Black–Scholes implied volatility, Heston's (1993) implied volatility, MFIV, and VKOSPI. In addition, a correlation matrix of these volatilities is shown in Table 2. Realized volatilities are estimated by five minutes interval KOSPI 200 index data. We estimated 30-day historical volatility using 5-min interval KOSPI 200 index data, including option prices between 14.50 hours on the observed date and 14.50 hours 30 days previously. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. We estimated implied volatility using prices of options with 30 days left to the maturity date, which met the filtering principles. We estimated monthly data for the period from January 2000 to June 2007 for realized volatility, historical volatility, Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV, and VKOSPI.

We considered monthly data because this sampling procedure results in no overlap at all in the data. Thus, there is no need to account for an overlapping data problem when making statistical inferences. Most previous studies use daily data and options with horizons of 1–3 months. As a result, there is overlap in the error term, which biases downward the usual OLS standard errors. Beginning with Canin-a and Figlewski (1993), researchers have typically used variants of the Hansen (1982) generalized method of moments (GMM) method to correct the standard

[3]There are simultaneous bids and offers from 14.50 hours in the Korean stock market. Therefore, it is appropriate to use both KOSPI 200 index data and KOSPI 200 options data before 14.50 hours.

Table 1 Descriptive statistics

Descriptive statistics for 90 monthly estimates of realized volatility (RV), 30-day historical volatility (30D HV), 60-day historical volatility (60D HV), 90-day historical volatility (90D HV), 180-day historical volatility (180D HV), 365-day historical volatility (365D HV), Black–Scholes implied volatility (BSV), Heston's (1993) implied volatility (SV), the MFIV (MFV), and the VKOSPI for KOSPI 200 options for the period from 10 January 2000 to 11 June 2007, are presented. We selected options with 30 days to maturity to calculate Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV, and the VKOSPI. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility by using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility using 5-min interval KOSPI 200 index data including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. The number of observations, mean, median, maximum, minimum, standard deviations, augmented Dickey–Fuller (ADF) test statistics for the 90 monthly estimates of realized volatility (RV), 30-day historical volatility (30D HV), 60-day historical volatility (60D HV), 90-day historical volatility (90D HV), 180-day historical volatility (180D HV), 365-day historical volatility (365D HV), Black–Scholes implied volatility (BSV), Heston's (1993) implied volatility (SV), the MFIV (MFV), and VKOSPI are presented.

Volatility	Observations	Mean	Median	Maximum	Minimum	Standard deviation	ADF
RV	90	0.2763	0.2499	0.6284	0.1109	0.1141	−5.2373
30D HV	90	0.2667	0.2457	0.5668	0.1015	0.1105	−4.9866
60D HV	90	0.2768	0.2525	0.5784	0.1139	0.1075	−4.2973
90D HV	90	0.2826	0.2640	0.5428	0.1211	0.1059	−4.8166
180D HV	90	0.2885	0.3006	0.4863	0.1336	0.1001	−3.5697
365D HV	90	0.2983	0.2889	0.4725	0.1647	0.0929	−3.1686
BSV	90	0.2943	0.2757	0.5601	0.1452	0.1022	−3.4830
SV	90	0.2794	0.2514	0.5585	0.1079	0.1022	−4.6021
MFV	90	0.2984	0.2873	0.5442	0.1440	0.0945	−3.6149
VKOSPI	90	0.3123	0.2861	0.5955	0.1585	0.1016	−3.7639

errors. Simulations in Canina and Figlewski (1993) and Jorion (1995) indicate that in the absence of error in the variables, the GMM corrections make it unlikely that the literature's conclusion that option based forecasts of future volatility are biased is caused by the faulty calculation of standard errors.

The average value of realized volatility or historical volatility derived from actual KOSPI 200 index data ranged from 0.2667 to 0.2983. Among implied volatility measures using option prices, the average values of Black–Scholes implied volatility and Heston's (1993) implied volatility were within the range of average values for realized volatility or historical volatility. In contrast, the average values of the MFIV and VKOSPI were 0.2984 and 0.3123, respectively, which were above the upper range. However, in terms of maximum/minimum values, all implied volatility values estimated using option prices were within the values of the realized volatility derived

Table 2 Correlation statistics

A correlation matrix of 90 monthly time-series datasets for the estimated realized volatility (RV), 30-day historical volatility (30D HV), Black–Scholes implied volatility (BSV), Heston's (1993) implied volatility (SV), the MFIV (MFV), and the VKOSPI for KOSPI 200 options for the period from 10 January 2000 to 11 June 2007, are presented. We selected options with 30 days to maturity to calculate Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV (MFV), and the VKOSPI. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility using 5-min interval KOSPI 200 index data, including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day.

	RV	BSV	SV	MFV	VKOSPI	30D HV
RV	1	0.761	0.836	0.756	0.758	0.726
BSV		1	0.786	0.963	0.926	0.895
SV			1	0.787	0.742	0.719
MFV				1	0.908	0.851
VKOSPI					1	0.823
30DHV						1

from the actual KOSPI 200 index. Because the KOSPI 200 options market is very liquid, price discrepancies tend to be very small and, thus, implied volatility from option prices tends to reflect actual volatility. Figure 1 shows the movements of all volatility measures considered in this study (realized volatility is indicated by the bold line). For simplicity, Figure 1 includes only 30-day historical volatility as the representative historical volatility. All volatility measures show similar movements, suggesting that realized volatility can be forecast using any volatility measure.

Table 1 shows the results of the unit root test (the right column). Except for 365-day historical volatility, all volatility measures were rejected at the 5% level of significance; 365-day volatility could not be rejected even at the 10% level.

4. Relationship between Implied and Realized Volatility

4.1. Econometric Analysis Model

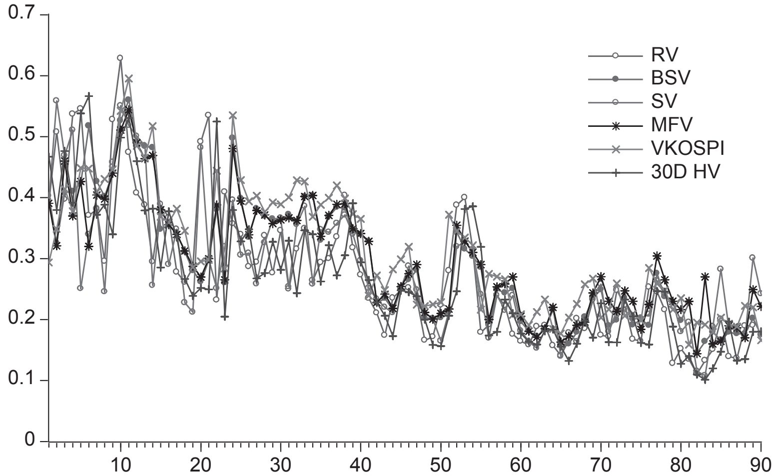
Most of the previous studies verifying the effectiveness of the implied volatility of options in forecasting future volatility use the following two regression equations:

$$\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \varepsilon(t), \tag{11}$$

$$\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \gamma \cdot \sigma_H(t) + \varepsilon(t), \tag{12}$$

where $\sigma_R(t)$ is the annualized ex-post realized volatility of the underlying asset at time t for the period between time t and the option expiration date; $\sigma_I(t)$ is

Figure 1 Volatility movements.



Movements of 90 monthly time-series datasets for the estimated realized volatility (RV), 30-day historical volatility (30D HV), Black–Scholes implied volatility (BSV), Heston's (1993) implied volatility (SV), the MFIV (MFV), and the VKOSPI for KOSPI 200 options for the period from 10 January 2000 to 11 June 2007, are shown. On the *x*-axis, 1 indicates 10 January 2000, and 90, 11 June 2007.

the annualized implied volatility derived from option prices at time t; $\sigma_H(t)$ is the annualized ex-post historical volatility of the underlying asset at time t for the period between past a certain point and time t; and $\epsilon(t)$ is the forecasting error that is uncorrelated with independent variables. Following previous studies, the present study uses these two equations.

If the value of β in equation (11) is significant and positive, then the implied volatility derived from option prices is a good indicator of future volatility. However, for implied volatility to be an unbiased estimator, the conditions $\alpha = 0$ and $\beta = 1$ have to hold simultaneously. Otherwise, the implied volatility may not contain sufficient information on future volatility.

Equation (12) tests the informational efficiency of the options market. If the value of γ is significantly different from zero, then implied volatility is an informationally inefficient estimator of future volatility. If implied volatility contains all information on past volatility, then it must be an unbiased estimator even with historical volatility. In other words, the conditions $\alpha = 0$, $\beta = 1$, and $\gamma = 0$ have to hold simultaneously. This means that option prices contain even the information contained in historical volatility.

4.2. Unbiasedness Tests

Table 3 shows the regression results, which indicate the forecasting performance of Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV, and the VKOSPI for future realized volatility. For implied volatility to be an

Table 3 Forecasting regression with implied volatility

The regression results for realized volatility when Black–Scholes implied volatility (BSV), Heston's (1993) implied volatility (SV), the MFIV (MFV), and the VKOSPI were each used as an independent variable are presented: $\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \varepsilon(t)$, where $\sigma_R(t)$ and $\sigma_I(t)$ are realized volatility and implied volatility, respectively, at time t. We used 90 monthly time-series datasets for each variable. The constant coefficient α and the slope coefficient β, with t-statistics values whose null hypotheses are $\alpha = 0$, and $\beta = 1$, respectively (in parentheses), are presented. There were 90 monthly time-series datasets for the period from 10 January 2000 to 11 June 2007, not only for realized volatility but also for each independent variable. We selected options with 30 days to maturity to calculate Black–Scholes implied volatility, Heston's (1993) implied volatility, the MFIV (MFV), and the VKOSPI. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility using 5-min interval KOSPI 200 index data. Adjusted R^2 indicates the portion of the total variation explained by a set of independent variables. We reported F-statistics to test the joint hypothesis that $\alpha = 0$ and $\beta = 1$ and p-values.

Volatility	α	β	Adjusted R^2	F-statistics	p-value
BSV	0.0309 (1.3107)	0.8260 (2.2952)	0.5749	6.0469	0.0035
SV	0.0125 (0.6547)	0.9444 (0.8692)	0.7114	0.4904	0.6141
MFV	0.0059 (0.2223)	0.9063 (1.0965)	0.5586	4.3559	0.0157
VKOSPI	0.0133 (0.5080)	0.8424 (1.9781)	0.5574	11.9189	0.0000

unbiased estimator of realized volatility, the conditions $\alpha = 0$ and $\beta = 1$ must hold simultaneously. Day and Lewis (1992) analyze the relationship between implied volatility and realized volatility using daily data on S&P 100 index options for the 1980s and suggest that forecasting ex-post realized volatility through implied volatility is biased and inefficient. In particular, Canina and Figlewski (1993) conclude that there is no correlation between implied volatility and realized volatility. In contrast, Jorion (1995) examines FX forward options on the Chicago Board Options Exchange Market, which were expected to show no substantial measurement error because of very brisk trading and the simultaneous closing of the underlying asset and options, and suggest that implied volatility is a better tool than historical volatility or GARCH for forecasting future realized volatility but that it is biased. In addition, Fleming (1998) analyzes the S&P 100 index options market and concludes that implied volatility has meaningful information on realized volatility but is biased. The present study, which analyzes the forecasting performance of the implied volatility of KOSPI 200 options, is consistent with previous studies in that the results indicate that Black–Scholes implied volatility is informative but biased.

Heston's (1993) implied volatility is superior to Black–Scholes implied volatility in forecasting future realized volatility. In addition, according to the F-statistics, the joint hypothesis that $\alpha = 0$ and $\beta = 1$ was not rejected, indicating that Heston's (1993) implied volatility is an unbiased estimator of future realized volatility. This result is consistent with the findings of Poteshman (2000), who suggest that Heston's (1993) implied volatility is better than Black–Scholes implied volatility at forecasting realized volatility. However, because our parameter estimation method for Heston's (1993) model is different from that of Poteshman (2000), the same result as Poteshman (2000) is not guaranteed. In addition to the fact that our parameter estimation method for Heston's model, which uses only option prices, is the same as that in Bakshi *et al.* (1997, 2000), it is also consistent with the intention of the present study to observe the bias of the model, which implied volatility is based on under the assumption that option prices are correct.

According to the F-statistics, the MFIV was the best estimator after Heston's (1993) implied volatility. That is, the t-statistic of MFIV on β is the second lowest. This means that MFIV is more informative about future volatility than Black–Scholes implied volatility or VKOSPI. However, as shown in Table 3, the MFIV is slightly better than or almost equal to Black–Scholes implied volatility in terms of its ability to forecast realized volatility, which is inconsistent with the findings of Jiang and Tian (2005), who suggest that the MFIV subsumes all information contained in Black–Scholes implied volatility.

We checked the ability of VKOSPI, a popular volatility index, to forecast future realized volatility. VKOSPI takes a non-parametric approach to derive the volatility from options. The results in Table 3 indicate that VKOSPI was as biased as Black–Scholes implied volatility in forecasting future volatility. Even if VKOSPI is less biased than Black–Scholes implied volatility in that both t-statistics of $\alpha = 0$ and $\beta = 1$, respectively, from VKOSPI are lower than those from Black–Scholes implied volatility, the R^2 result of Black–Scholes implied volatility is superior to that of VKOSPI. As shown in Table 3, Heston's (1993) implied volatility was the best in forecasting realized volatility.

Table 4 shows the regression results for the forecasting performance of historical volatility for future realized volatility. Among several time periods for estimating historical volatility, the 180-day period was the best for forecasting future volatility, suggesting that information contained in previous 6-month data is meaningful for forecasting future volatility. In particular, when we applied the historical volatility estimated using previous 6-month data, the forecasting performance of historical volatility was comparable to that of Black–Scholes implied volatility, the MFIV, and the VKOSPI. These results, which are inconsistent with the findings of previous studies suggesting the superiority of Black–Scholes implied volatility, the MFIV, and the VKOSPI over historical volatility in forecasting future volatility, may be attributed to our measurement of historical volatility using high-frequency data. However, our results do not contradict the findings of previous studies considering other markets because this study's results (Tables 5–8) suggest that implied volatility plays a more important role than historical volatility in forecasting future volatility when implied volatility and historical volatility are used simultaneously as independent variables.

Table 4 Forecasting regression with historical volatility

The regression results for realized volatility when 30-day (30D HV), 60-day (60D HV), 90-day (90D HV), 180-day (180D HV), and 365-day (365D HV) historical volatility measures were each used as an independent variable are presented: $\sigma_R(t) = \alpha + \beta \cdot \sigma_H(t) + \varepsilon(t)$, where $\sigma_R(t)$ and $\sigma_H(t)$ are realized volatility and historical volatility, respectively, at time t. We used 90 monthly time-series datasets for each variable. The constant coefficient α and the slope coefficient, β, with t-statistics values (in parentheses), are presented. There were 90 monthly time-series datasets for the period from 10 January 2000 to 11 June 2007, not only for realized volatility but also for each independent variable. We estimated realized volatility by using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility by using 5-min interval KOSPI 200 index data including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. Adjusted R^2 indicates the portion of the total variation explained by a set of independent variables. We reported F-statistics to test the joint hypothesis that $\alpha = 0$ and $\beta = 1$ and p-values.

Volatility	α	β	Adjusted R^2	F-statistic	p-value
30D HV	0.0757 (3.4558)	0.7523 (3.2636)	0.5249	5.9874	0.0037
60D HV	0.0596 (2.6161)	0.7829 (2.8252)	0.5389	3.9924	0.0219
90D HV	0.0555 (2.3133)	0.7814 (2.7468)	0.5203	4.0512	0.0208
180D HV	0.0289 (1.1787)	0.8576 (1.7697)	0.5614	2.7122	0.0720
365D HV	0.0030 (0.1012)	0.9170 (0.9469)	0.5523	4.1346	0.0193

Table 5 Variance inflation factor (VIF)

The VIF test, which examines the degree of multicollinearity between an implied volatility and a historical volatility, is presented: $\mathrm{VIF} = \frac{1}{1-R_j^2}$, where $R^2{}_j$ is the coefficient of determination of a regression of explanatory j on the other explanatory. We used 90 monthly time-series datasets for 30-day historical volatility (30D HV), 60-day historical volatility (60D HV), 90-day historical volatility (90D HV), 180-day historical volatility (180D HV), 365-day historical volatility (365D HV), Black–Scholes implied volatility (BSV), Heston's (1993) implied volatility (SV), the MFIV (MFV), and the VKOSPI for the period from 10 January 2000 to 11 June 2007. The left (right) number in each cell presents a VIF result when the coefficient of determination of a regression of an implied volatility (a historical volatility) on a historical volatility (an implied volatility) is applied.

Volatility	BSV	SV	MFV	VKOSPI
30D HV	3.66/4.96	1.60/1.99	2.22/3.56	2.02/3.09
60D HV	4.49/5.38	2.90/2.29	2.59/3.72	2.29/3.14
90D HV	4.51/5.10	1.92/2.00	2.76/3.73	2.45/3.16
180D HV	4.41/4.35	1.91/1.74	2.99/3.52	2.83/3.18
365D HV	3.28/2.67	1.87/1.40	2.54/2.47	2.62/2.42

4.3. Informational Efficiency Tests

Equation (12) is applied to test the informational efficiency of an implied volatility on a realized volatility. Because there are two independent variables in equation (12), the multicollinearity between two variables has to be investigated in advance. The variance inflation factor (VIF), which is presented in equation (13), is

Table 6 Forecasting regression with Black–Scholes implied volatility and historical volatility

The regression results for realized volatility when Black–Scholes implied volatility and historical volatility were used as independent variables are presented: $\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \gamma \cdot \sigma_H(t) + \varepsilon(t)$, where $\sigma_R(t)$, $\sigma_I(t)$, and $\sigma_H(t)$ are realized volatility, implied volatility, and historical volatility, respectively, at time t. The constant coefficient α and the slope coefficients β and γ, with t-statistics values whose null hypotheses are $\alpha = 0$, $\beta = 1$, and $\gamma = 0$, respectively (in parentheses), are presented. There were 90 monthly time-series datasets for the period from 10 January 2000 to 11 June 2007, not only for realized volatility but also for the independent variables. We selected options with 30 days to maturity to calculate Black–Scholes implied volatility. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility (30D HV) using 5-min interval KOSPI 200 index data including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. We used F-statistics to test the joint hypothesis that $\alpha = 0$, $\beta = 1$, and $\gamma = 0$.

HV	α	β	γ	Adjusted R^2	F-statistic	p-value
30D HV	0.0347 (1.5584)	0.6099 (1.9411)	0.2245 (1.0738)	0.5799	4.7598	0.0041
60D HV	0.0309 (1.3551)	0.6132 (1.5680)	0.2273 (0.9919)	0.5789	4.6734	0.0045
90D HV	0.0289 (1.2765)	0.6549 (1.2190)	0.1859 (0.7431)	0.5760	4.4539	0.0059
180D HV	0.0179 (0.8224)	0.5027 (2.2810)	0.3762 (1.7344)	0.5950	5.9879	0.0009
365D HV	−0.0013 (−0.0511)	0.5047 (2.3783)	0.4259 (1.8072)	0.6087	7.1915	0.0002

widely used for testing whether a multicollinearity problem exists in a multiple regression model:

$$\mathrm{VIF} = \frac{1}{1 - R_j^2}, \tag{13}$$

where R_j^2 is the coefficient of determination of a regression of explanator j on all the other explanators. Generally, VIF of 5 and above indicates a multicollinearity problem. Table 5 shows that multicollinearity between an implied volatility and a historical volatility does not lead to a statistical problem and implies that informational efficiency tests using equation (12) are effective.

Table 6 shows the regression results for the forecasting performance of Black–Scholes implied volatility and historical volatility for future realized volatility. To estimate historical volatility, we considered five time periods: 30 days, 60 days, 90 days, 180 days, and 365 days. We estimated each historical volatility using 5-min interval data for each time period. As shown in Table 6, the longer the estimation period for historical volatility, the better the historical volatility was at forecasting realized volatility. When we applied 30-day historical volatility, β was 0.6099, and γ was 0.2245; for 365-day historical volatility, β decreased to 0.5047, and γ increased to 0.4259. To calculate the standard error of each coefficient, we adopted the

Table 7 Forecasting regression with Heston's (1993) implied volatility and historical volatility

The regression results for realized volatility when Heston's (1993) implied volatility and historical volatility were used as independent variables are presented: $\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \gamma \cdot \sigma_H(t) + \varepsilon(t)$, where $\sigma_R(t)$, $\sigma_I(t)$, and $\sigma_H(t)$ are realized volatility, implied volatility, and historical volatility, respectively, at time t. The constant coefficient α and the slope coefficients β and γ, with t-statistic values whose null hypotheses are $\alpha = 0$, $\beta = 1$, and $\gamma = 0$, respectively (in parentheses), are presented. There were 90 monthly time-series datasets for the period from 10 January 2000 to 11 June 2007, not only for realized volatility but also for the independent variables. We selected options with 30 days to maturity to calculate Heston's (1993) implied volatility. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility (30D HV) using 5-min interval KOSPI 200 index data including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. We used F-statistics to test the joint hypothesis that $\alpha = 0$, $\beta = 1$, and $\gamma = 0$.

HV	α	β	γ	Adjusted R^2	DW	F-statistic	p-value
30D HV	0.0001 (0.0048)	0.7458 (1.7800)	0.2545 (1.6968)	0.7378	1.8573	3.6131	0.0164
60D HV	0.0009 (0.0553)	0.7598 (1.5478)	0.2283 (1.4321)	0.7274	1.6664	2.3858	0.0747
90D HV	−0.0035 (−0.2079)	0.7635 (1.6807)	0.2354 (1.5748)	0.7300	1.6232	2.6849	0.0516
180D HV	−0.0180 (−1.0785)	0.7053 (2.7604)	0.3372 (2.8729)	0.7507	1.5939	5.2887	0.0021
365D HV	−0.0330 (−1.5538)	0.7063 (3.2144)	0.3754 (3.0481)	0.7571	1.6044	6.1810	0.0003

Newey–West corrected standard error to correct the effects of correlation and heteroskedasticity in the error terms. As a result, the t-statistic reported in Table 6 is obtained from the value of a coefficient divided by the Newey–West corrected standard error. All t-statistics on coefficients in Table 6–9 are calculated in the same way. The fact that the coefficient of historical volatility, γ, was not statistically zero suggests that option prices are informationally inefficient in forecasting future volatility. It can be interpreted that a model deriving implied volatility from option prices is not appropriate.

Table 7 shows the regression results for the forecasting performance of Heston's (1993) implied volatility and historical volatility for future realized volatility. Compared with the results presented in Table 3, the t-statistic on Heston's (1993) implied volatility is much bigger. Because the null hypothesis of this t-statistic is $\beta = 1$, the bigger t-statistic on Heston's (1993) implied volatility in Table 5 implies that Heston's (1993) implied volatility explains the future realized volatility less when a historical volatility is added as an independent variable. This is because a historical volatility plays a partial role of what an implied volatility alone used to do. However, when we used 60- or 90-day historical volatility as historical volatility in the regression, the F-statistics did not reject the informational efficiency of Heston's (1993) implied volatility at the 95% confidence level. In other words, Heston's (1993) implied volatility reflected the information contained in past data,

Table 8 Forecasting regression with model-free implied volatility and historical volatility

The regression results for realized volatility when the MFIV and historical volatility were used as the independent variables are presented: $\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \gamma \cdot \sigma_H(t) + \varepsilon(t)$, where $\sigma_R(t)$, $\sigma_I(t)$, and $\sigma_H(t)$ are realized volatility, implied volatility, and historical volatility, respectively, at time t. The constant coefficient α and the slope coefficients β and γ, with t-statistic values whose null hypotheses are $\alpha = 0$, $\beta = 1$, and $\gamma = 0$, respectively (in parentheses), are presented. There were 90 monthly time-series datasets for the period from 10 January 2000 to 11 June 2007, not only for realized volatility but also for the independent variables. We selected options with 30 days to maturity to calculate the MFIV. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility (30D HV) using 5-min interval KOSPI 200 index data, including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. We used F-statistics to test the joint hypothesis that $\alpha = 0$, $\beta = 1$, and $\gamma = 0$.

HV	α	β	γ	Adjusted R^2	F-statistic	p-value
30D HV	0.0161 (0.6970)	0.5740 (2.6708)	0.3337 (1.9508)	0.5827	5.0812	0.0028
60D HV	0.0111 (0.4478)	0.5389 (2.2539)	0.3773 (2.0106)	0.5880	5.5151	0.0016
90D HV	0.0070 (0.2712)	0.5872 (1.7145)	0.3330 (1.5386)	0.5797	4.8382	0.0037
180D HV	−0.0036 (−0.1458)	0.4836 (2.6272)	0.4703 (2.4401)	0.6023	6.7444	0.0004
365D HV	−0.0260 (−0.9566)	0.5149 (2.5833)	0.4986 (2.4105)	0.6145	7.8638	0.0001

Table 9 Forecasting regression with VKOSPI and historical volatility

The regression results for realized volatility when VKOSPI and historical volatility were used as the independent variable are presented: $\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \gamma \cdot \sigma_H(t) + \varepsilon(t)$, where $\sigma_R(t)$, $\sigma_I(t)$, and $\sigma_H(t)$ are realized volatility, implied volatility, and historical volatility, respectively, at time t. The constant coefficient α and the slope coefficients β and γ, with t-statistic values whose null hypotheses are $\alpha = 0$, $\beta = 1$, and $\gamma = 0$, respectively (in parentheses), are presented. There were 90 monthly time-series datasets for the period from 10 January 2000 to 11 June 2007, not only for realized volatility but also for the independent variables. We selected options with 30 days to maturity to calculate VKOSPI. We followed four principles to filter the data for the empirical analysis. First, we selected the latest transacted options data before 14.50 hours for each day during the sample period. Second, if there is more than one transaction data for each day, only one data is used. Third, we excluded those options whose price was below 0.02. Fourth, we excluded those options that did not meet the arbitrage restriction. We estimated realized volatility by using 5-min interval KOSPI 200 index data. We estimated 30-day historical volatility (30D HV) by using 5-min interval KOSPI 200 index data including option prices between 14.50 hours on the observed date and 14.50 hours 30 days ago. If the sample day was a holiday, then we used 1 business day prior to that day. We estimated 60-, 90-, 180-, and 365-day historical volatility in the same way. We used F-statistics to test the joint hypothesis that $\alpha = 0$, $\beta = 1$, and $\gamma = 0$.

HV	α	β	γ	Adjusted R^2	DW	F-statistic	p-Value
30D HV	0.0178 (0.6763)	0.5245 (3.7444)	0.3551 (2.4189)	0.5912	2.0126	6.2134	0.0030
60D HV	0.0118 (0.4566)	0.4972 (2.8222)	0.3949 (2.3416)	0.5973	1.7401	6.9582	0.0016
90D HV	0.0085 (0.3224)	0.5343 (2.2060)	0.3573 (1.8336)	0.5877	1.6294	5.7985	0.0043
180D HV	−0.0019 (−0.0761)	0.4513 (2.7874)	0.4759 (2.3969)	0.6067	1.5310	8.1558	0.0006
365D HV	−0.0216 (−0.8060)	0.4763 (2.6069)	0.5002 (2.2240)	0.6133	1.5249	9.3124	0.0003

and, as a result, it was informationally efficient in forecasting future volatility. However, when we applied 180-day historical volatility (i.e. more information), Heston's (1993) implied volatility was no longer informationally efficient. The fact that the *p*-value (0.0021) was less than 1% indicates that the informational efficiency of Heston's (1993) implied volatility was rejected at the 99% confidence level. Consistent with the results in Table 6, historical volatility played a more important role as the estimation period increased. When we applied 30-day historical volatility, β was 0.7458, and γ was 0.2545; for 365-day historical volatility, β decreased to 0.7068, and γ increased to 0.3754.

Table 8 shows the regression results for the forecasting performance of the MFIV and historical volatility for future realized volatility. Unlike the results in Table 7, the *F*-statistic results presented in Table 8 indicate that MFIV was an informationally inefficient estimator for future volatility regardless of the length of data period for estimating historical volatility. Furthermore, the MFIV (as indicated by the results in Table 8) was not superior to Black–Scholes implied volatility (as indicated by the results in Table 6) in terms of informational efficiency. Even if the R^2 values in Table 8 are generally higher than those in Table 6, historical volatility accounted for most of these differences. The fact that the historical volatility in Table 8 contributed more than that in Table 6 indicates that the MFIV was informationally less efficient than Black–Scholes implied volatility.

Table 9 shows the regression results for the forecasting performance of VKOSPI and historical volatility for future realized volatility. Consistent with the results in Tables 6–8, the longer the estimation period for historical volatility, the better the historical volatility was at forecasting realized volatility. In particular, the coefficient for historical volatility was higher than that for VKOSPI when 180- or 365-day historical volatility was applied. This indicates that VKOSPI was inferior to Black–Scholes implied volatility in terms of informational efficiency in forecasting future volatility.

In sum, when we employed implied volatility and historical volatility as the independent variables for the regression for future volatility, Heston's (1993) implied volatility was the most informationally efficient estimator of future volatility. Although this efficiency was observed only for 60- or 90-day historical volatility, compared with Heston's (1993) implied volatility, other implied volatility measures did not show much informational efficiency. This provides strong evidence that Heston's (1993) implied volatility dominates other implied volatility measures in forecasting future volatility.

5. Conclusions

Previous studies have typically assessed the forecasting performance of Black–Scholes implied volatility for future realized volatility. It has been empirically demonstrated that Black–Scholes implied volatility is a biased estimator of future realized volatility, although it is generally better than historical volatility or GARCH in terms of forecasting performance. By assuming that option prices are unbiased, we looked for a better way to derive implied volatility from option prices.

We compared four implied volatility measures, including Black–Scholes implied volatility. First, we selected Heston's (1993) implied volatility because it best addresses the problems associated with the Black–Scholes model for pricing and hedging options, and therefore, we expected it to reduce the forecast bias associated with the Black–Scholes model. Second, we selected the MFIV as an alternative model for testing forecasting performance for future volatility because some recent papers have shown that the forecasting performance of the MFIV is superior to that of Black–Scholes implied volatility. Such papers have attributed the superiority of the MFIV to the fact that a test of the MFIV is a direct test of market efficiency, whereas a test of other implied volatility measures is a joint test of market efficiency and model stability. Finally, we adopted VKOSPI as an alternative measure because it has attracted considerable interest not only from practitioners but also from researchers. In addition, because the discrete version of VKOSPI is, by definition, the expected risk-neutral value of realized volatility, we expected it to be a good estimator of future realized volatility.

According to the results of this empirical study, among the four implied volatility measures, Heston's (1993) implied volatility was the best, and was a statistically unbiased estimator of future realized volatility. This is important in that option prices may not be biased for forecasting future volatility if an appropriate model is applied. Furthermore, Heston's (1993) implied volatility was an informationally efficient estimator of future volatility. This suggests that the flexible assumptions in Heston's model, such as non-zero market prices of volatility risk and non-zero correlation between innovations to the level and volatility of the underlying asset, can address the empirical deficiency of the Black–Scholes model.

However, the MFIV and VKOSPI did not reduce the bias associated with Black–Scholes implied volatility. Inconsistent with some recent papers suggesting the superiority of the MFIV over Black–Scholes implied volatility, the present study shows that the forecasting performance of the MFIV was as weak as that of Black–Scholes implied volatility. These results suggest that creating artificial option prices, which are required for implementing the MFIV, can lead to a bias even if the MFIV does not suffer from a model-specific bias. If the bias from creating artificial option prices exceeds the model-specific bias, then the forecasting performance of the MFIV may be inferior to that of any other model-oriented implied volatility measure. VKOSPI also failed to overcome Black–Scholes implied volatility in forecasting future volatility. This suggests that the MFIV and VKOSPI, which make the same assumption about the movement of the underlying asset, cannot outdo each other even if the procedure for deriving the implied volatility from option prices is different.

A jump factor can be significant in a short-term option market. Therefore, it could be interesting and valuable to check in the future whether the jump factor contributes to the future realized volatility forecast performance when it is included in Heston's (1993) model.

References

Ait-Sahalia, Y., and A. W. Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* 53, pp. 499–547.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, The distribution of stock return volatility, *Journal of Financial Econometrics* 61, pp. 43–76.

Andersen, T. G., T. Bollerslev, and F. X. Diebold, 2002, Parametric and nonparametric volatility measurement, Working Paper, prepared for the *Handbook of Financial Econometrics*.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2003, Modeling and forecasting realized volatility, Working Paper, National Bureau of Economic Research.

Bakshi, G., C. Cao, and Z. Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, pp. 2003–2049.

Bakshi, G., C. Cao, and Z. Chen, 2000, Pricing and hedging long-term options, *Journal of Econometrics* 94, pp. 277–318.

Bates, D., 1991, The crash of '87: Was it expected? The evidence from options market, *Journal of Finance* 46, pp. 1009–1044.

Bates, D., 2000, Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics* 94, pp. 181–238.

Black, F., 1976, Studies of stock price volatility changes, *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economic Statistic Section*, pp. 177–181.

Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, pp. 637–659.

Britten-Jones, M., and A. Neuberger, 2000, Option prices, implied price processes, and stochastic volatility, *Journal of Finance* 55, pp. 839–866.

Campa, J. M., K. P. Chang, and R. L. Reider, 1998, Implied exchange rate distributions: Evidence from OTC option markets, *Journal of International Money and Finance* 17, pp. 117–160.

Canina, L., and S. Figlewski, 1993, The informational content of implied volatility, *Review of Financial Studies* 6, pp. 659–681.

Chang, K. H., 2001, Volatility prediction and forecast performance in Korean stock index option markets, *Korean Journal of Futures and Options* 9, pp. 51–79.

Christie, A., 1982, The stochastic behavior of common for alternative stochastic processes, *Journal of Financial Economics* 10, pp. 407–432.

Corrado, C., and T. Miller, 2005, The forecast quality of CBOE implied volatility indexes, *Journal of Futures Markets* 25, pp. 339–373.

Day, T., and C. Lewis, 1992, The information content of option prices and a test of market efficiency, *Journal of Econometrics* 52, pp. 267–287.

Duan, J. C., 1995, The GARCH option pricing model, *Mathematical Finance* 5, pp. 13–32.

Eom, Y. H., H. J. Ji, and W. W. Jang, 2008, The predictability of volatility index in KOSPI 200 option market, *Journal of Money and Finance* 22, pp. 1–33.

Figlewski, S., 1997, Forecasting volatility, *Financial Market, Institutions and Instruments* 6, pp. 1–88.

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.

Fleming, J., B. Ostdiek, and R. Whaley, 1995, Predicting stock market volatility: A new measure, *Journal of Futures Markets* 15, pp. 265–302.

Hansen, L. P., 1982, Large sample properties of generalized method of moment estimators, *Econometrica* 50, pp. 1029–1054.

Heston, S. L., 1993, A closed-form solutions for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, pp. 327–343.

Heston, S. L., and S. Nandi, 2000, A closed-form GARCH option valuation model, *Review of Financial Studies* 13, pp. 585–625.

Hull, J., and A. White, 1987, The pricing of options with stochastic volatilities, *Journal of Finance* 42, pp. 281–300.

Jiang, G. J., and Y. S. Tian, 2005, The model-free implied volatility and its information content, *Review of Financial Studies* 18, pp. 1305–1342.

Johnson, H., and D. Shanno, 1987, Option pricing when the variance is changing, *Journal of Financial and Quantitative Analysis* 22, pp. 143–151.

Jorion, P., 1995, Predicting volatility in the foreign exchange market, *Journal of Finance* 50, pp. 507–528.

Kim, I. J., and S. Kim, 2005, Empirical comparison of alternative stochastic volatility option pricing models: Evidence from Korean KOSPI 200 index options market, *Pacific Basin Finance Journal* 12, pp. 117–142.

Kim, T., and J. Park, 2006, Measuring implied volatility for predicting realized volatility, *Korean Journal of Financial Engineering* 5, pp. 1–25.

Lamoureux, D., and W. Lastrapes, 1993, Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities, *Review of Financial Studies* 6, pp. 293–326.

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.

Lee, J. H., and S. S. Kwon, 2001, Predictive power of the implied volatility from the KOSPI 200 options market, *Korean Journal of Futures and Options* 9, pp. 25–50.

Lee, I., H. S. Oh, Y. R. Song, and Y. J. Yang, 2005, Realized volatility estimation and analysis of the KOSPI 200 index options prices, *Asia-Pacific Journal of Financial Studies* 34, pp. 181–208.

Melino, A., and S. M. Turnbull, 1990, Pricing foreign currency options with stochastic volatility, *Journal of Econometrics* 45, pp. 239–265.

Pan, J., 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, pp. 3–50.

Pong, S., M. Shackleton, S. Taylor, and X. Xu, 2003, Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models, Working Paper, Lancaster University, Lancashire, England.

Poteshman, A. M., 2000, Forecasting future volatility from option prices, Working Paper, University of Illinois at Urbana–Champaign.

Scott, L. O., 1987, Option pricing when the variance changes randomly: Theory, estimation, and an application, *Journal of Financial and Quantitative Analysis* 22, pp. 419–438.

Shimko, D., 1993, Bounds of probability, *Risk* 6, pp. 33–37.

Stein, E. M., and J. C. Stein, 1991, Stock price distribution with stochastic volatility: An analytic approach, *Review of Financial Studies* 4, pp. 727–752.

Wiggins, J. B., 1987, Option values under stochastic volatility: Theory and empirical estimates, *Journal of Financial Economics* 19, pp. 351–377.

Yoo, S., and J. Y. Koh, 2009, A research on enhancing forecasting power for the realized volatility of KOSPI 200, *Korean Journal of Futures and Options* 17, pp. 21–49.

