Forecasting Future Volatility from Option Prices
Under the Stochastic Volatility Model

Suk Joon Byun, Sol Kim, and Dong Woo Rhee*

Abstract

The implied volatility from Black and Scholes (1973) model has been empirically tested for the forecasting performance of future volatility and commonly shown to be biased. Based on the belief that the implied volatility from option prices is the best estimate of future volatility, this study tries to find out a better model, which can derive the implied volatility from option prices, to overcome the forecasting bias from Black and Scholes (1973) model. Heston (1993)'s model which improves on the problems of Black and Scholes (1973) model the most for pricing and hedging options is one candidate, and VIX which is the expected risk neutral value of realized volatility under the discrete version is the other. This study conducts a comparative analysis on the implied volatility from Black and Scholes (1973) model, that from Heston (1993)'s model, and VIX for the forecasting performance of future volatility. From the empirical analysis on KOSPI200 option market, it is found that Heston (1993)'s implied volatility eliminates the bias mostly which Black and Scholes (1973) implied volatility has. VIX, on the other hand, does not show any improvement for the forecasting performance.

Keywords: Options, Stochastic Volatility, VIX, Forecasting, Regression
JEL classification: G13, G14

Suk Joon Byun is an assistant professor of the Business School at the Korea Advanced Institute of Science and Technology

Sol Kim is an assistant professor of College of Business Administration at Hankuk University of Foreign Studies.

Dong Woo Rhee is a doctorial candidate of the Business School at the Korea Advanced Institute of Science and Technology.

* Corresponding author, Dong Woo Rhee, Korea Advanced Institute of Science & Technology (KAIST) Business School , 87 Hoegiro Dongdaemoon-gu Seoul 130-722 KOREA.

Phone: 82-2-958-3968, Fax: 82-2-958-3618, E-mail: dwrhee@business.kaist.ac.kr

1 Introduction

Under the stochastic volatility assumption, Hull and White (1987) showed that Black and Scholes (1973) (henceforth BS) implied volatility from option price is considered the best estimate of future volatility. Upon such theoretical background, there have been numerous empirical studies on whether implied volatility can efficiently forecast the future realized volatility. The early study by Latane and Rendleman(1976) showed that stocks with high implied volatility had high ex-post realized volatility through the analysis on stock option cross-section data. Time series analysis conducted in the later studies offer a contrary view on the forecasting performance of the implied volatility. Day and Lewis (1992), Lamoureux and Lastrapes (1993), Jorion (1995), and Fleming (1998) commonly show that BS implied volatility is the biased estimator of the future realized volatility.

A question can arise whether deriving the implied volatility from option prices using BS model is correct even if an option price is believed to reflect the market expectation of the future volatility. As inferred from the famous phenomenon, volatility smile, it is generally accepted that BS model fails to explain the option market correctly. To overcome this empirical deficiency, many models which make the assumptions of BS model flexible have developed. There are alternative models based on the assumptions of stochastic volatility, stochastic interest rate, and jumps. Bakshi, Cao, and Chen (1997, 2000) show that the stochastic volatility model improves on the problems of BS model the most. Thus, this study tests the forecasting performance of future volatility through a stochastic volatility model. Among various stochastic models, we choose Heston (1993)'s model not only because it follows the continuous time stochastic volatility model which, Kim and Kim (2005) proves, is the best in pricing and hedging efficiency but also because it takes into account the correlation between the volatility and the return of the underlying asset and provides the closed form solution.

In the empirical analysis on the forecasting performance of future volatility, we fill the gap that is not resolved in previous researches. First, this is the first study which shows the forecasting performance of future volatility through Heston (1993)'s model whose unobservable parameters are estimated by only option prices. Even if there is an existing study, Poteshman (2000), which uses the stochastic volatility model in forecasting performance of future volatility, Poteshman (2000) uses the past time-series data of the underlying asset in estimating some of the unobservable parameters of Heston (1993)'s model. Furthermore, Poteshman (2000) assumes that these structural parameters are constant during the whole sample period. On the contrary, all parameters estimated by our method are considered as time varying variables whose values are determined by option prices at observation date. According to Bates (1991), time varying parameters are more valuable than constant parameters not only because the time varying parameters can reflect the market sentiment when those are estimated but also because this time varying parameters can offer the future specification of complex dynamic models.

Second, this study investigates KOSPI200 options market for the forecasting performance of future volatility. KOSPI200 options market, the biggest stock index options market in the world, is a quite adequate one for an empirical research because of its huge liquidity. As Figlewski (1997) implies, to choose an option market which is as liquid as possible is very important, otherwise the mispriced options from an illiquid market might lead to the forecasting bias of the implied volatility on future realized volatility. According to Futures Industry Association, KOSPI200 option contracts occupy, on average, more than a fifth of world wide trading volumes since 2000. Therefore, it is less likely happened in this study that the noise from a friction market distorts the result on forecasting performance of future volatility because the data employed in this analysis is on KOSPI200 options market during the period from January 2000 to June 2007.

Finally, this study also includes VIX as one of the comparative implied volatilities in forecasting performance of future volatility. First published in 1993, VIX has been considered a popular measure of risk in stock index option market. Because VIX is defined as the expected risk neutral value of realized volatility under the discrete version, its performance on future realized volatility has drawn interest. Even if some studies, such as Corrado and Miller (2005) and Fleming, Ostdiek, and Whaley (1995), show the superiority of VIX over historical volatility in forecasting performance of future realized volatility, there has been no comparison between VIX and any other implied volatility. This study initially shows the comparison results on forecasting performance of VIX against other implied volatilities on future volatility.

The main finding from this study is that the implied volatility from the stochastic volatility model is superior to other implied volatilities as well as the historical volatility in forecasting performance of future volatility. It even eliminates the bias mostly which BS implied volatility has. VIX, on the other hand, does not show any improvement over BS implied volatility in this empirical analysis.

The remainder of this study is organized in the following way. Section 2 exhibits the models which are used in this study. The process of deriving the Heston (1993)'s model and VIX is presented, followed by the method of estimating realized volatility. Section 3 explains the KOSPI200 option market and the data used in this study. It presents the basic statistical analyses of realized volatility, historical volatility, BS implied volatility, Heston (1993)'s implied volatility and VIX on monthly basis for the final analysis. Section 4 is to show the relationship between realized and implied volatilities. First, econometric analysis model to determine the forecasting ability of realized volatility is explained. Next, the forecasting results of future realized volatility, which this study ultimately aims to show, by using not only implied volatilities but also historical volatilities are presented through a variety of econometric analysis techniques. Section 5 is the conclusion which summaries this study and presents the future research areas.

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. Whereas 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, Duan(1995) and Heston and Nandi(2000) belong to the latter. Among several stochastic volatility models, this study uses Heston (1993)'s model in the following reasons. First, it belongs to continuous time stochastic models which are proven to be better than discrete time stochastic models by Bakshi, Cao, and Chen (1997, 2000) and Kim and Kim (2005). Next, Heston (1993)'s model can provide the closed form solution and take into account the correlation between the volatility and the return of the underlying asset.[1] Stochastic process of stock price and volatility which is assumed in Heston (1993)'s model is as below.

$$dS = \mu S dt + \sqrt{v_t} S dW_S \tag{1}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}\, dW_v \tag{2}$$

where, S is a stock value, μ is the return on the stock, W is a 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 volatility. Using the Fourier transform under the assumption of the stochastic process described above, the option pricing model follows below.

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

[1] The existing empirical studies showed the negative correlation between the volatility and the return of the underlying asset, risk neutral distribution with negative skewness and the low strike price which has large volatility, called volatility sneer. This is consistent with the leverage effect documented by Black (1976) and Christie (1982). The negative correlation phenomenon can be explained that falling stock prices will bring about relatively higher debt equity ratio, which in turn will have a leverage effect on the enterprise, which makes the volatility of the earnings per share greater, which eventually has the effect of amplifying the stock price volatility.

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

where, C is a call option, K is the exercise price of the call option, r is risk-free interest rate, τ is time to maturity, Re[·]is the real number part of a complex number, i is the 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(·) and B(·) are functions of θ, κ, ρ, and σ.

What to do next is to estimate structural parameters which are not observable. This study uses the same estimation methods as the standard practices which were applied by Bakshi, Cao, and Chen (1997, 2000) and Bates (1991, 2000). If there is a closed form solution in the option pricing model, then the starting point of estimating the parameters in the pricing formula is non-linear least squares procedure. This is a method of estimating the parameters by minimizing the sum of squared percentage errors of the difference between the model price and the actual price. Even if the parameters can be estimated from the historical returns of asset, this method is limited in that the historical data only reflect the past. A major advantage of estimating the parameters from the option prices is that the method can take advantage of forward looking information contained in option prices. In the case of Heston (1993)'s model, we estimate the parameters by minimizing the sum of squared percentage errors of the difference between the model price and the actual price as in the following equation.

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

$O_i^*(t, \tau; K)$ is the model price of option i at time t, and $O_i(t, \tau; K)$ 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.

2.2 VIX

VIX means volatility index which represents the expected volatility of the returns from equity index options for thirty days which is valued as of today. Since 1993 when CBOE (Chicago Board Options Exchange) began public disclosure of VIX, it has been widely used by investors not only as a barometer of market volatility sentiment but also as the underlying asset of volatility products. At September 22 2003, new VIX was released to make up some distortions which previous one had. Compared with previous VIX which is derived from only at the money (thereafter ATM) option prices, new VIX is calculated through all available out of the money (thereafter OTM) option prices. Since more information from more options is reflected, new VIX shows a more robust result. Therefore, new VIX is adopted in this study.

VIX, which is the expected risk neutral value of realized volatility under the discrete version, can be derived in the following way. First, the forward index level which is to be used in the calculation of VIX has to be set.

$$F = K_{ATM} + e^{r\tau} \cdot (C_{ATM} - P_{ATM}) \tag{6}$$

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

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 in the calculation of VIX. When calculating VIX, all available OTM calls and puts are used because OTM options take up most of the trading volume, and become in the money (thereafter ITM) option as the value of the option increases. The option price at K_0 is the average of price of call and put, and call is selected at a higher price than K_0, and put is selected at a price lower than K_0. Finally, VIX is calculated using the equation below.

$$VIX = \sqrt{\frac{2}{\tau}\sum_{i=1}^{m}\frac{\Delta K_i}{K_i^2}e^{r\tau}Q(K_i) - \frac{1}{\tau}\left[\frac{F}{K_0} - 1\right]^2} \tag{7}$$

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

Because VIX is the 30days 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 applied, when the time to maturity of the option is not exactly 30 days. However, interpolation is unnecessary in this study because it only examines option whose time to maturity is 30 days, which is consistent with the maturity VIX is targeting.

2.3 Realized Volatility

In the case of the existing estimation method of volatility using daily price, the volatility estimation can be distorted because not all price changes are reflected during a certain time frame. For example, the real volatility can be underestimated in the case in which the intra-day price fluctuation is great or the difference among the closing price is not extreme. By using high frequency data, one can reduce the estimation error when estimating volatility. Poteshman (2000) showed that about one-half of the forecasting bias can be eliminated in estimating realized volatility using the high frequency data of 5 minute intervals instead of the existing daily data. This result is supported by studies by Anderson, Bollerslew, Diebold, and Labys (2003) and Pong, Shackleton, Taylor, and Shu (2003). According to the studies, high-frequency volatility has greater forecasting strengths compared to low frequency volatility not only in the short-term, but also in the long term since it contains more data. Anderson, Bollerslev, Deiebold, and Ebens (2001) and Anderson, Bollerslev, and Diebold(2002) present that, as the sampling frequency of the underlying returns approaches infinitely, volatility estimates are, in theory, free from measurement error. Upon the arguments above, this study estimates both realized volatility and historical volatility by using the five minutes interval data which contains more information than daily data to estimate the future realized volatility. The volatility estimation is described as below.

$$Vol = \sqrt{\frac{1}{\Delta}\frac{1}{L-1}\sum_{i=1}^{L-1}\left[\ln\left(S_{i+1}/S_i\right)\right]^2} \tag{8}$$

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

3 Data

KOSPI200 option market in Korea is analyzed for testing forecasting performance for future volatility of the implied volatilities. KOSPI200 option market is appropriate for this analysis because it is the biggest equity option market in the world in terms of trading volume, as a result, option prices can be efficiently determined. The maturity date of KOSPI200 option is the second Thursday of the option contract month, and option contract months are consecutive three months and one more month from March, June, September, and December. There are at least five exercise prices per each option contract month, which can be increased as option prices move. KOSPI200 option contract is fully automated and is European option which can be exercised only at maturity. Because the liquidity of the option is concentrated on the nearest contract and one of comparative volatilities in this study, VIX, means the volatility which is expired in 30 days, this study selected options whose maturities are left 30 days every month. OTM calls and puts are used for calculating BS implied volatility, Heston (1993)'s implied volatility, and VIX. ITM options are excluded because their trading volume is very small, as a result, the reliability of the transacted price is not fully satisfied. Data period is from January 10th 2000 to June 11th 2007 and minute by minute transaction prices of KOSPI200 option on the day in which there are 30 days until maturity is obtained from Korea Stock Exchange. Because there is no significant benchmark rate for 30 days in Korea, 91days CD(certificate deposit) rate, which is the representative short rate in Korea, is used for a risk free rate. To filter data which is necessary for empirical analysis, the following principals are applied. The last trading price prior to PM2:50[2] of each option contract on every sample day is applied to the empirical analysis. The last transacted option is only included in the sample if the same options are traded several times during any time window. To mitigate the price discreteness effect in the option valuation, the options whose prices are lower than 0.02 are excluded. Finally, the prices which do not meet arbitrage restriction are not included.

Anderson, Bollerslev, and Diebold (2002) mentioned that the estimation error can be reduced through using high frequency data for estimating volatility because existing method for estimating volatility, which is based on daily price, can distort it because all price movement information during the estimation period can not be reflected. Poteshman (2000) compared the estimation result from using daily prices with that from using five minutes interval prices in estimating realized volatility. In his study, estimation error is reduced, as a result, forecasting performance of implied volatility is improved when realized volatility is estimated from five-minutes high frequency data. Therefore, this study estimates both realized volatility and historical volatility through five minutes high frequency data of KOSPI200 index. Five minutes interval KOSPI200 index data is obtained from Korea Stock Exchange. Because historical volatility is estimated from the historical data which is preceding to the date when implied volatility is calculated, and realized volatility is estimated from ex-post data which is following the date when implied volatility is calculated, data period for KOSPI200 index, which is from January 10th 1999 to July 12th 2007, is longer than that for KOSPI200 option.

Table 1 is the descriptive statistics for monthly realized volatility, historical volatility, BS implied volatility, Heston (1993)'s implied volatility, and VIX. In addition, correlation matrix of these volatilities is shown in Table 2. Realized volatilities are estimated by five minutes interval

---

[2] There are simultaneous bids and offers from PM2:50 in the Korean stock market. Therefore, it is appropriate to use both KOSPI200 index and KOSPI200 option data prior to PM2:50.

KOSPI200 index data from the market open price of the next day of 30 days left to option maturity to the market close price of the option expiration date. 30days historical volatilities are estimated by five minutes interval KOSPI200 index data from 2:50PM price of the option price observed date to 2:50PM price of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60days, 90days, 180days, and 365days historical volatilities are estimated as the same way. Implied volatilities are estimated by option prices, which meet the filtering principals, of 30 days left to option maturity. From January 2000 to June 2007, 90 monthly data are estimated for realized volatility, historical volatility, BS implied volatility, Heston (1993)'s implied volatility, and VIX respectively. Average numbers of realized volatility or historical volatility which is derived from actual KOSPI200 index data are ranged from 0.2667 to 0.2983. Among implied volatilities derived from option prices, the average numbers of BS implied volatilities and Heston (1993)'s implied volatilities are within the range of average values from realized volatility or historical volatility. On the other hand, the average value of VIX, 0.3123, is higher than the upper range. However, considering Max, Min values, all implied volatilities estimated by option prices are within the volatilities derived from actual KOSPI200 index values. Because KOSPI200 option market in Korea is very liquid, as a result, the price discrepancy is tiny, the implied volatilities from option prices are not remote from the actual volatilities. Figure 1 shows the movements of all volatilities presented in this study, especially, realized volatility is depicted as a bold line. To avoid the complexity of this figure, only 30days historical volatility is included as a representative of historical volatilities. As shown in Figure 1, movements of all volatilities are similar, so we can guess that realized volatility can be forecasted by any volatility predictor in a sense.

Unit root test result is present in the most right column of Table 1. All volatilities but 365days historical volatility are rejected unit root under the 5% significance level. Even 365days volatility can be rejected unit root when the significance level is increased to 10%. Therefore, all volatilities which are analyzed in the empirical test can be used as levels not differences in their time series data.

4 The relationship between implied and realized volatility

4.1 Econometric analysis model

Most studies on verification of forecasting effectiveness of future volatility using the implied volatility of options use the two regression equations below. This study also uses the two equations for consistency in comparison with existing studies.

$$\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \varepsilon(t) \tag{9}$$

$$\sigma_R(t) = \alpha + \beta \cdot \sigma_I(t) + \gamma \cdot \sigma_H(t) + \varepsilon(t) \tag{10}$$

where $\sigma_R(t)$ is the annualized ex-post realized volatility of the underlying asset at time t during the period between time t and the option expiration date, $\sigma_I(t)$ is 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 during the period between past certain point and time t, and $\varepsilon(t)$ is the forecasting error which is uncorrelated with the independent variables.

If the value of β is statistically significantly positive in Equation (9), it means that the implied volatility derived from option prices is informative about future volatility. However, in order for the implied volatility to be an unbiased estimator, the joint condition, $\alpha=0$, $\beta=1$, has to be held simultaneously. Otherwise, the implied volatility might contain the information of future volatility insufficiently. Equation (10) tests the informational efficiency in the option market. If the value of γ is significantly different from zero, it means that the implied volatility is informationally inefficient estimator of future volatility. If the implied volatility contains all information about past volatility, the implied volatility has to be an unbiased estimator even if historical volatility is added. In other words, the condition, $\alpha=0$, $\beta=1$, $\gamma=0$, has to be held at the same time. It means that option prices contain even the information which historical volatility has.

4.2 Unbiasedness Tests

Table 3 is the regression result which shows the forecasting performances of BS implied volatility, Heston (1993)'s implied volatility, and VIX for the future realized volatility. In order to accept that the implied volatility is the unbiased estimator of the realized volatility, the condition, $\alpha=0$, $\beta=1$, has to be satisfied at the same time. Day and Lewis(1992) showed that by analyzing the relationship between implied volatility and realized volatility based on daily data of US S&P100 index option of 1980s, the effectiveness of forecasting ex-post realized volatility through implied volatility was biased and inefficient. Especially, Canina and Figlewski(1993) conclude that there is no correlation between implied volatility and realized volatility. On the other hand, Jorion(1995) stated that the results of a study on FX forward options of CBOE which is expected not to have a great measurement error due to very brisk trading and the simultaneous closing of the underlying asset and option show that implied volatility was a better tool in forecasting the future realized volatility than historical volatility or GARCH but it is biased. Also Fleming(1998), after analysis on the S&P100 index option market, concludes that implied volatility has meaningful information about realized volatility but it was also biased. This study, which analyzed the forecasting performance of the implied volatility in KOSPI 200 option market, is consistent with the previous researches in that BS implied volatility is informative but biased.

On the other hand, Heston (1993)'s implied volatility is superior to BS implied volatility in terms of the forecasting performance of future realized volatility. Further to the superiority to BS implied volatility, Heston (1993)'s implied volatility was proven to be an unbiased estimator of the future realized volatility through the F statistics result that joint hypothesis, α=0, β=1, is not rejected. This result is consistent with that of Poteshman(2000) which shows that Heston (1993)'s implied volatility forecasts realized volatility better than BS implied volatility. However, since our parameter estimation method for Heston(1993)'s model is different from that of Poteshman(2000), the same result as Poteshman(2000) is not guaranteed. Besides the fact that our parameters estimation of the Heston (1993)'s model using only the option prices is the same as the recent studies by Bakshi, Cao, and Chen (1997, 2000), it is also in line with the intentions of this study to observe the bias of the model which the implied volatility is based on under the assumption that option prices are correct.

On the other hand, this study adopted VIX, which is a popular volatility index in a market, to check the forecasting performance of the future realized volatility. VIX follows more non-parametric approach to derive the volatility from option prices than BS implied volatility though the assumption that the underlying asset follows the log normal process is the same. Previous researches showed that both BS implied volatility and VIX are superior to historical volatility in terms of the forecasting performance of the future realized volatility. According to the empirical result of this study, VIX is the biased estimator of the future volatility as BS implied volatility is. Even if VIX is less biased than BS implied volatility in that both t statistics of α=0 and β=1 respectively from VIX are lower than those from BS implied volatility, R^2 result of BS implied volatility is superior to that of VIX. Therefore, this study concludes the mixed result for the comparison of the forecasting performance between BS implied volatility and VIX. In the end, the result of Table 3 indicates that Heston (1993)'s implied volatility is the best for the forecasting performance of the realized volatility.

Table 4 is the regression result which shows the forecasting performances of historical volatilities for the future realized volatility. As time period for estimating historical volatility increases, the forecasting performance is shown to be better. This fact implies that the information in the past data within one year is meaningful for the forecast on future volatility. Especially, when 180days or 365days historical volatility is applied, the forecasting performance of historical volatility is not worse than that of BS implied volatility or VIX. These results look inconsistent with those from previous studies which mostly showed the superiority of BS implied volatility or VIX against historical volatility in the forecasting performance of future volatility. One of the reasons which lead to improved results of historical volatility can be attributed to the volatility measurement through high frequency data. However, our results do not deny the previous findings from other markets because implied volatilities are shown to play more important role than historical volatilities in the forecast of future volatility when an implied volatility and a historical volatility are used at the same time as independent variables. These results are presented from Table 5 through Table 7.

4.3 Informational Efficiency Tests

Table 5 is the regression result of the future realized volatility on both BS implied volatility and historical volatility. For estimating historical volatility, 5 time periods of data, which are 30 days, 60 days, 90 days, 180 days, and 365 days, are applied. Each historical data is estimated from the 5 minutes interval data of each time period. Table 5 presents that as the estimation period of historical volatility is longer, historical volatility tends to be more significant for forecasting the realized volatility. Whereas β is 0.6099 and γ is 0.2245 when 30 days historical volatility is applied, β decreases to 0.5047 and γ increases to 0.4259 when 365 days historical volatility is applied. The fact[3] that the coefficient of historical volatility, γ, is not statistically zero means that option prices are informationally inefficient for forecasting future volatility. Table 6 is the regression result of the future realized volatility on both Heston (1993)'s implied volatility and historical volatility. The fact shown in Table 6 is that the addition of new independent variable, historical volatility, is also significant for the forecasting performance of the future volatility. Even if the F statistics in Table 3 does not reject the unbiasedness of Heston (1993)'s implied volatility, the F statistics in Table 6 rejects the informational efficiency of Heston (1993)'s implied volatility. In other words, Heston (1993)'s implied volatility is informationally inefficient for forecasting future volatility in spite of the fact that it is the unbiased estimator of future volatility. Like the result presented in Table 5, historical volatility plays more important role as the estimation period of historical volatility is longer. Whereas β is 0.7458 and γ is 0.2545 when 30 days historical volatility is applied, β decreases to 0.7068 and γ increases to 0.3754 when 365 days historical volatility is applied. Table 7 is the regression result of the future realized volatility on both VIX and historical volatility. Like the results presented in Table 5 and Table 6, the longer period historical volatility is estimated, the more significantly the historical volatility influences on forecasting future volatility. Especially, the coefficient of historical volatility is bigger than that of VIX when 180 days historical volatility or 365 days historical volatility is applied. This fact can be interpreted that VIX is inferior to BS implied volatility in terms of the informational efficiency for the forecast of future volatility. To sum up, there are mainly two findings when an implied volatility and a historical volatility are applied as independent variables for the regression of future volatility. First, implied volatility from option

[3] However, it can be interpreted that the model which derives the implied volatility from option prices is not appropriate.

prices is informationally inefficient no matter what implied volatility is applied. In other words, option prices can not contain full information which past data of underlying asset has. Next, the importance of historical volatility is bigger as the estimation period of historical volatility is longer. This phenomenon is consistent with that shown in Table 4.

4.4 Relative Strength Tests

Table 8 is the regression result of future realized volatility when two independent variables are selected from BS implied volatility, Heston (1993)'s implied volatility, and VIX. In this result, the superiority of Heston (1993)'s implied volatility for the forecasting performance of future realized volatility is proven directly, compared with BS implied volatility and VIX. When Heston (1993)'s implied volatility and BS implied volatility are used at the same time, the coefficient of Heston (1993)'s implied volatility is 0.7342, which is much greater than that of BS implied volatility, 0.2706. On the other hand, when Heston (1993)'s implied volatility and VIX are used in the same regression, the coefficient of Heston (1993)'s implied volatility, 0.7147, is much larger than that of VIX, 0.3142. One interesting result is observed when the results of Table 6 and Table 8 are compared. Whereas the R^2 in Table 6 is more than 0.75 when Heston (1993)'s implied volatility and a historical volatility whose estimation period is either 180 days or 365 days are applied, the R^2 in Table 8 is less than 0.75 when Heston (1993)'s implied volatility and one of the other two implied volatility are applied. This comparison implies that what Heston (1993)'s implied volatility is short of can be complemented more by historical volatility than other implied volatility. It can be inferred from the correlation matrix of volatilities presented in Table 2.

On the other hand, the superiority can not be determined between BS implied volatility and VIX because the coefficients of both variables are 0.4559 and 0.4174, respectively when these two are applied as independent variables. This is supported by the F statistics for testing the hypothesis that $\beta=\gamma$. The p-value from F statistics when two independent variables are BS implied volatility and VIX is 0.8418, which strongly support that BS implied volatility and VIX can't beat each other in the forecasting contribution for future volatility.

5 Conclusion

Previously, most efforts to check the forecasting performance of future realized volatility were concentrated on BS implied volatility. It is an empirically proven common result that BS implied volatility is a biased estimator for the future realized volatility even if it is generally better than historical volatility or GARCH in terms of forecasting performance. Under the assumption that option prices are unbiased, we wanted to find out a better model to derive implied volatility from option prices. For the alternatives, Heston (1993)'s model was selected because it improved on the problems of BS model the most for pricing and hedging options, so was expected to reduce the forecasting bias BS model has as well. On the other hand, this study selected VIX as a comparative model for testing forecasting performance of future volatility because VIX is currently paid attention to not only from practical aspect but also from academic area. In addition, because VIX is, by definition, the expected risk neutral value of realized volatility under the discrete version, it is expected to be a good forecaster of future realized volatility.

This study tried first to compare the forecasting performance of VIX with that of BS implied volatility and Heston (1993)'s implied volatility. This is an interesting trial because no body touches this comparison before regardless of the fact that the forecasting performance of VIX was proven to be superior to that of historical volatility in some previous researches such as Corrado and Miller (2005) and Fleming, Ostdiek, and Whaley (1995).

The main finding in this empirical study is that the Heston (1993)'s implied volatility is the best estimator of future realized volatility among three volatilities analyzed in this study. In particular, Heston (1993)'s implied volatility was shown to be a statistically unbiased estimator of future volatility. This result is very significant because option prices may not be biased for forecasting future volatility if an appropriate model is applied. However, VIX was not shown to improve the bias which BS implied volatility has. This result can be inferred that two volatilities with the same assumption of the underlying asset movement can't beat each other significantly even if the deriving procedure from option prices to volatility is different. In this aspect, to make the assumptions flexible in Heston (1993)'s model, such as non-zero market price of volatility risk and non-zero correlation between innovations to the level and volatility of the underlying asset, is the meaningful effort to improve the empirical deficiency which BS model has. However, the fact that Heston (1993)'s implied volatility was shown to be informatively inefficient when historical volatility was added as another independent variable is what we have to study further in the future.

References

Andersen TG, Bollerslev T, Diebold FX (2002) Parametric and Nonparametric Volatility Measurement. Working paper, prepared for Handbook of Financial Econometrics

Andersen TG, Bollerslev T, Diebold FX, Ebens H (2001) The Distribution of Stock Return Volatility. J Financial Econometrics 61:43-76

Andersen TG, Bollerslev T, Diebold FX, Labys P (2003) Modeling and Forecasting Realized Volatility. Working Paper, National Bureau of Economic Research

Bakshi G, Cao C, Chen Z (1997) Empirical performance of alternative option pricing models. J Finance 52:2003-2049

Bakshi G, Cao C, Chen Z (2000) Pricing and hedging long-term options. J Econometrics 94:277-318

Bates D (1991) The Crash of ’87: Was It Expected? The Evidence from Options Market. J Finance 46:1009-1044

Bates D (2000) Post-’87 Crash Fears in the S&P 500 Futures Option Market. J Econometrics 94:181-238

Black F (1976) Studies of Stock Price Volatility Changes. Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economic Statistic Section:177-181

Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Political Economy 81:637-659

Canina L, Figlewski S (1993) The informational content of implied volatility. Rev Financial Studies 6:659-681

Christensen BJ, Prabhala NR (1998) The relation between implied and realized volatility. J Financial Economics 50:125-150

Christie A (1982) The Stochastic Behavior of Common for Alternative Stochastic Processes. J Financial Economics 10:407-432

Corrado C, Miller T (2005) The forecast quality of CBOE implied volatility indexes. J Futures Markets 25:339-373

Day T, Lewis C (1992) The information content of option prices and a test of market efficiency. J Econometrics 52:267-287

Duan JC (1995) The GARCH Option Pricing Model. Math Finance 5:13-32

Figlewski S (1997) Forecasting volatility. Financial Market Institutions & Instruments 6:1-88

Fleming J (1998) The quality of market volatility forecasts implied by S&P 100 index option prices. J Empir Finance 5:317-345

Fleming J, Ostdiek B, Whaley R (1995) Predicting stock market volatility: A new measure. J Futures Markets 15:265-302

Hansen LP (1982) Large sample properties of generalized method of moment estimators. Econometrica 50:1029-1054

Heston SL (1993) A Closed-Form Solutions for Options with Stochastic Volatility with Applications to Bond and Currency Options. Rev Financial Studies 6:327-343

Heston SL, Nandi S (2000) A Closed-Form GARCH Option Valuation Model. Rev of Financial Studies 13:585-625

Hull J, White A (1987) The Pricing of Options with Stochastic Volatilities. J Finance 42:281-300

Johnson H, Shanno D (1987) Option Pricing when the Variance is Changing. J Financial and Quantitative Analysis 22:143-151

Jorion P (1995) Predicting volatility in the foreign exchange market. J Finance 50:507-528

Kim IJ, Kim S (2005) Empirical Comparison of Alternative Stochastic Volatility Option Pricing Models: Evidence from Korean KOSPI 200 Index Options Market. Pac Basin Finance J 12:117-142

Lamoureux D, Lastrapes W (1993) Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. Rev Financial Studies 6:293-326

Latane HA, Rendleman RJ (1976) Standard deviations of stock price ratios implied in option prices. J Finance 31:369-381

Melino A, Turnbull SM (1990) Pricing Foreign Currency Options with Stochastic Volatility. J Econometrics 45:239-265

Pong S, Shackleton M, Taylor S, Xu X (2003) Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models. Working paper

Poteshman AM (2000) Forecasting Future Volatility from Option Prices. Working paper, University of Illinois at Urbana-Champaign

Scott LO (1987) Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. J Financial and Quantitative Analysis 22:419-438

Stein EM, Stein JC (1991) Stock Price Distribution with Stochastic Volatility: An Analytic Approach. Rev Financial Studies 4:727-752

Wiggins JB (1987) Option Values under Stochastic Volatility: Theory and Empirical Estimates. J Financial Economics 19:351-377

Table 1 Descriptive Statistics. Descriptive Statistics for 90 monthly estimated realized volatility(RV), 30days historical volatility(30D HV), 60days historical volatility(60D HV), 90days historical volatility(90D HV), 180days historical volatility(180D HV), 365days historical volatility(365D HV), BS implied volatility(BSV), Heston (1993)'s implied volatility(SV), and VIX on KOSPI200 option for the period from January 10 2000 through June 11 2007 are presented. Options with 30 days to maturity are selected for calculating BS implied volatility, Heston (1993)'s implied volatility, and VIX. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. 30days historical volatilities are estimated by 5 minutes interval KOSPI200 index data from 2:50PM data of the option price observed date to 2:50PM data of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60day, 90day, 180day, and 365day historical volatilities are estimated as the same way. Mean, Media, Max, Min, Standard deviation, Augmented Dickey-Fuller test statistics for 90 monthly estimated realized volatility(RV), 30days historical volatility(30D HV), 60days historical volatility(60D HV), 90days historical volatility(90D HV), 180days historical volatility(180D HV), 365days historical volatility(365D HV), BS implied volatility(BSV), Heston (1993)'s implied volatility(SV), and VIX are presented.

Volatility	Obs.	Mean	Median	Max	Min	S.D.	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
VIX	90	0.3123	0.2861	0.5955	0.1585	0.1016	-3.7639

Table 2 Correlation Statistics. Correlation matrix of 90 monthly time series data for estimated realized volatility (RV), 30days historical volatility (30D HV), BS implied volatility (BSV), Heston (1993)'s implied volatility (SV), and VIX on KOSPI200 option for the period from January 10 2000 through June 11 2007 are shown. Options with 30 days to maturity are selected for calculating BS implied volatility, Heston (1993)'s implied volatility, and VIX. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. 30days historical volatilities are estimated by 5 minutes interval KOSPI200 index data from 2:50PM data of the option price observed date to 2:50PM data of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60day, 90day, 180day, and 365day historical volatilities are estimated as the same way.

	RV	BSV	SV	VIX	30DHV	60DHV	90DHV	180DHV	365DHV
RV	1	0.768	0.845	0.750	0.728	0.738	0.725	0.753	0.747
BSV		1	0.792	0.923	0.896	0.905	0.900	0.888	0.839
SV			1	0.735	0.722	0.768	0.741	0.723	0.698
VIX				1	0.823	0.826	0.828	0.834	0.800
30DHV					1	0.941	0.896	0.848	0.791
60DHV						1	0.966	0.909	0.853
90DHV							1	0.945	0.891
180DHV								1	0.955
365DHV									1

Table 3 Forecasting Regression with Implied Volatilities. The regression results for the realized volatility when BS implied volatility (BSV), Heston (1993)'s implied volatility (SV), and VIX are applied respectively 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 the realized volatility and the implied volatility at time t, respectively.

90 monthly time series data for each variable are used. Constant coefficient, α, and Slope coefficient, β, with t statistics for $\alpha=0$, $1-\beta=0$ (in parentheses) are presented. There are 90 monthly time series data starting from January 10 2000 through June 11 2007 not only for the realized volatility but also for each independent variable. Options with 30 days to maturity are selected for calculating BS implied volatility, Heston (1993)'s implied volatility, and VIX. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. Adj.R^2 is statistics which shows how much of total variation is explained by a set of independent variables. DW stands for Durbin-Watson statistics which tests the autocorrelation of the error terms. The F statistics is for testing the joint hypothesis that $\alpha=0$, $\beta=1$. The p-value is a measure of how likely the sample results are, assuming the joint hypothesis is true.

Volatility	α	β	Adj. R^2	DW	F-stat	p-value
BSV	0.0309 (1.3107)	0.8260 (2.2952)	0.5749	1.7268	5.2681	0.0242
SV	0.0125 (0.6547)	0.9444 (0.8692)	0.7114	1.5851	0.7555	0.3871
VIX	0.0133 (0.5080)	0.8424 (1.9781)	0.5574	1.5925	3.9128	0.0511

Table 4 Forecasting Regression with Historical Volatilities. The regression results for the realized volatility when 30day(30D HV), 60day(60D HV), 90day(90D HV), 180day(180D HV), and 365day(365D HV) historical volatilities are applied respectively 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 the realized volatility and the historical volatility at time t, respectively.

90 monthly time series data for each variable are used. Constant coefficient, α, and Slope coefficient, β, with t statistics for $\alpha=0$, $1-\beta=0$ (in parentheses) are presented. There are 90 monthly time series data starting from January 10 2000 through June 11 2007 not only for the realized volatility but also for each independent variable. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. 30days historical volatilities are estimated by 5 minutes interval KOSPI200 index data from 2:50PM data of the option price observed date to 2:50PM data of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60day, 90day, 180day, and 365day historical volatilities are estimated as the same way. Adj.R^2 is statistics which shows how much of total variation is explained by a set of independent variables. DW stands for Durbin-Watson statistics which tests the autocorrelation of the error terms. The F statistics is for testing the joint hypothesis that $\alpha=0$, $\beta=1$. The p-value is a measure of how likely the sample results are, assuming the joint hypothesis is true.

Volatility	α	β	Adj. R^2	DW	F-stat	p-value
30D HV	0.0757 (3.4558)	0.7523 (3.2636)	0.5249	2.0734	10.6510	0.0018
60D HV	0.0596 (2.6161)	0.7829 (2.8252)	0.5389	1.5892	7.9819	0.0059
90D HV	0.0555 (2.3133)	0.7814 (2.7468)	0.5203	1.3970	7.5451	0.0073
180D HV	0.0289 (1.1787)	0.8576 (1.7697)	0.5614	1.2559	3.1318	0.0803
365D HV	0.0030 (0.1012)	0.9170 (0.9469)	0.5523	1.1576	0.8966	0.3463

Table 5 Forecasting Regression with BS Implied and Historical Volatilities. The regression results for the realized volatility when BS implied volatility is applied as independent variables with a historical volatility as the second 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 the realized volatility, the implied volatility, and the historical volatility at time t, respectively.

Constant coefficient, α, and Slope coefficients, β and γ, with t statistics for α=0, 1-β=0, γ=0 (in parentheses) are presented. There are 90 monthly time series data starting from January 10 2000 through June 11 2007 not only for the realized volatility but also for independent variables. Options with 30 days to maturity are selected for calculating BS implied volatility. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. 30days(30D HV) historical volatilities are estimated by 5 minutes interval KOSPI200 index data from 2:50PM data of the option price observed date to 2:50PM data of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60days(60D HV), 90days(90D HV), 180days(180D HV), and 365days(365D HV) historical volatilities are estimated as the same way. The F statistics is for testing the joint hypothesis that α=0, β=1, and γ=0.

HV	α	β	γ	Adj. R^2	DW	F-stat	p-value
30D HV	0.0347 (1.4736)	0.6099 (2.3078)	0.2245 (1.4281)	0.5799	1.9229	3.6856	0.0292
60D HV	0.0309 (1.3175)	0.6132 (2.2084)	0.2273 (1.3461)	0.5789	1.7346	3.5650	0.0326
90D HV	0.0289 (1.2251)	0.6549 (2.0115)	0.1859 (1.1112)	0.5760	1.6870	3.2587	0.0433
180D HV	0.0179 (0.7546)	0.5027 (3.1260)	0.3762 (2.2955)	0.5950	1.5974	5.3996	0.0062
365D HV	-0.0013 (-0.0507)	0.5047 (3.7412)	0.4259 (2.9043)	0.6087	1.6197	7.0794	0.0014

Table 6 Forecasting Regression with Heston's implied and Historical Volatilities. The regression results for the realized volatility when Heston (1993)'s implied volatility is applied as independent variables with historical volatility as the second 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 the realized volatility, the implied volatility, and the historical volatility at time t, respectively.

Constant coefficient, α, and Slope coefficients, β and γ, with t statistics for α=0, 1-β=0, γ=0 (in parentheses) are presented. There are 90 monthly time series data starting from January 10 2000 through June 11 2007 not only for the realized volatility but also for independent variables. Options with 30 days to maturity are selected for calculating Heston (1993)'s implied volatility. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. 30days(30D HV) historical volatilities are estimated by 5 minutes interval KOSPI200 index data from 2:50PM data of the option price observed date to 2:50PM data of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60days(60D HV), 90days(90D HV), 180days(180D HV), and 365days(365D HV) historical volatilities are estimated as the same way. The F statistics is for testing the joint hypothesis that α=0, β=1, and γ=0.

HV	α	β	γ	Adj. R^2	DW	F-stat	p-value
30D HV	0.0001 (0.0036)	0.7458 (2.8859)	0.2545 (3.1241)	0.7378	1.8573	5.2956	0.0068
60D HV	0.0009 (0.0448)	0.7598 (2.4731)	0.2283 (2.4737)	0.7274	1.6664	3.4595	0.0359
90D HV	-0.0035 (-0.1822)	0.7635 (2.5653)	0.2354 (2.6469)	0.7300	1.6232	3.9069	0.0238
180D HV	-0.0180 (-0.9288)	0.7053 (3.4222)	0.3372 (3.8380)	0.7507	1.5939	7.8026	0.0008
365D HV	-0.0330 (-1.6036)	0.7063 (3.5866)	0.3754 (4.1686)	0.7571	1.6044	9.1376	0.0003

Table 7 Forecasting Regression with VIX and Historical Volatilities. The regression results for the realized volatility when VIX is applied as independent variables with historical volatility as the second 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 the realized volatility, the implied volatility, and the historical volatility at time t, respectively.

Constant coefficient, α, and Slope coefficients, β and γ, with t statistics for $\alpha=0$, $1-\beta=0$, $\gamma=0$ (in parentheses) are presented. There are 90 monthly time series data starting from January 10 2000 through June 11 2007 not only for the realized volatility but also for independent variables. Options with 30 days to maturity are selected for calculating VIX. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. 30days(30D HV) historical volatilities are estimated by 5 minutes interval KOSPI200 index data from 2:50PM data of the option price observed date to 2:50PM data of 30 days ago. If that day is a holiday, one business day prior to that day is applied. 60days(60D HV), 90days(90D HV), 180days(180D HV), and 365days(365D HV) historical volatilities are estimated as the same way. The F statistics is for testing the joint hypothesis that $\alpha=0$, $\beta=1$, and $\gamma=0$.

HV	α	β	γ	Adj. R^2	DW	F-stat	p-value
30D HV	0.0178 (0.7083)	0.5245 (3.5245)	0.3551 (2.8619)	0.5912	2.0126	6.2134	0.0030
60D HV	0.0118 (0.4718)	0.4972 (3.7303)	0.3949 (3.1010)	0.5973	1.7401	6.9582	0.0016
90D HV	0.0085 (0.3376)	0.5343 (3.4013)	0.3573 (2.7197)	0.5877	1.6294	5.7985	0.0043
180D HV	-0.0019 (-0.0756)	0.4513 (4.0357)	0.4759 (3.4508)	0.6067	1.5310	8.1558	0.0006
365D HV	-0.0216 (-0.8247)	0.4763 (4.2164)	0.5002 (3.6830)	0.6133	1.5249	9.3124	0.0003

Table 8 Forecasting Regression with Two Implied Volatilities. The regression results for the realized volatility when two from BS implied volatility (BSV), Heston (1993)'s implied volatility (SV), and VIX are applied as independent variables are presented.

$$\sigma_R(t) = \alpha + \beta \cdot \sigma_{I_1}(t) + \gamma \cdot \sigma_{I_2}(t) + \varepsilon(t)$$

where $\sigma_R(t)$, $\sigma_{I1}(t)$, and $\sigma_{I2}(t)$ are the realized volatility, the 1st implied volatility, and the 2nd implied volatility at time t, respectively.

Constant coefficient, α, and Slope coefficient, β, with t statistics for α=0, 1-β=0 (in parentheses) are presented. There are 90 monthly time series data starting from January 10 2000 through June 11 2007 not only for the realized volatility but also for independent variables. Options with 30 days to maturity are selected for calculating BS implied volatility, Heston (1993)'s implied volatility, and VIX. There are 4 principals to filter data for empirical analysis. First, the latest transacted option data prior to 2:50 PM for each day during the sample period are selected. Second, if there are more than one transaction data for each day, only one data is used. Third, the option whose price is below 0.02 is excluded. Fourth, the option which doesn't meet the arbitrage restriction is not included. Realized volatilities are estimated by 5 minutes interval KOSPI200 index data from the market open price of the next day of the option price observed date to the market close price of the option expiration date. Adj.R^2 is statistics which shows how much of total variation is explained by a set of independent variables. DW stands for Durbin-Watson statistics which tests the autocorrelation of the error terms. The F statistics is for testing the hypothesis that β=γ. The p-value is a measure of how likely the sample results are, assuming the joint hypothesis is true.

Two Volatilities	C	BS	SV	VIX	Adj. R^2	DW	F-stat	p-value
BS+SV	-0.0084 (-0.4256)	0.2706 (2.7670)	0.7342 (7.5036)		0.7319	1.6449	0.7269	0.4693
BS+VIX	0.0118 (0.4626)	0.4559 (2.2124)		0.4174 (2.0136)	0.5764	1.6228	0.2001	0.8418
SV+VIX	-0.0215 (-1.0580)		0.7147 (8.0575)	0.3142 (3.5229)	0.7449	1.6652	1.2912	0.2001

Fig. 1 Movements of volatilities. Movements of 90 monthly time series data for estimated realized volatility (RV), 30days historical volatility (30D HV), BS implied volatility (BSV), Heston (1993)'s implied volatility (SV), and VIX on KOSPI200 option for the period from January 10 2000 through June 11 2007 are shown. On X-axis, 1 is correspondent to January 10 2000 and 90 is June 11 2007.