2. FORECASTING VOLATILITY IN FINANCIAL MARKETS USING A BIVARIATE STOCHASTIC VOLATILITY MODEL WITH SURPRISING INFORMATION

Beum-Jo PARK[1]

Abstract

Most asset returns exhibit high volatility and its persistence. Heuristically, this paper focuses on the role of surprising information in high volatility processes and indicates that dismissing surprising information may lead to considerable loss in forecast accuracy. In response, this paper considers the corresponding extension of the modified MDH to surprising information, and proposes a bivariate stochastic volatility model incorporating surprising information in the volatility equations (BSV-SI), which is also designed to capture the dynamics of returns and trading volume. Using the South Korea stock index and trading volume series, it turns out that performance of the one-step-ahead forecasts of the BSV-SI model is apparently superior to those of other competitive models.

Keyword: Volatility forecasting, Bivariate stochastic volatility model with surprising information, Modified mixture of distribution hypothesis, Realized volatility models, Markov Chain Monte Carlo (MCMC)

JEL Classification: G17, G12, C53

I. Introduction

Return volatility, considered as variation in the activities of a financial market, is a prevailing property. Its source has retained attention by academics and market participants, and recently has received much greater attention due to global financial turbulence that started in 2007 following the subprime mortgage crisis in the United States. One can reasonably infer that return volatility may be induced by trades related to the arrival of information in markets (as early influential work on this issue, see Clark, 1973; Copeland, 1976; Admati and Pfleiderer, 1988, among others). Much of the extant studies relating to information arrival focus on a single type of information

[1] Department of Economics, Dankook University, 44-1 Jukjeon-dong Yongin-si, Gyeonggi-do 448-701, South Korea, E-mail: bjpark@dankook.ac.kr.

such as public information. Andersen (1996), however, indicates that there are two or more types of information arrival each of which has a different effect on trading volume and volatility persistence. In particular, Park (2010) has lately advocated decomposing information into surprising and general information to account more accurately for this aspect of the relationship between volatility and trading volume[2]. Employing the Mixture of Distribution Hypothesis (MDH), Park (2010) theoretically demonstrates that, in contrast with general information, surprising information could give rise to a weak or even negative correlation between volatility and trading volume. His empirical findings with high frequency data also substantially support the different role of surprising information in a foreign exchange market.

In the spirit of Andersen (1996) and Park (2010), this paper highlights the effect of surprising information on volatility processes and forecasting performance. If surprising information is neglected, volatility models might be misspecified and lead to poor out-of-sample forecasts because exceptionally high volatility and its persistence caused by surprising information can give rise to a substantial overshooting in forecasting volatility. Therefore, this study is valuable in that none of previous studies focus on the crucial role of surprising information in predicting future volatility.

This paper also contributes to feasible extension of the modified MDH by Andersen, (1996), and corroborates the Park's (2010) idea of the negative effect of surprising information on the relationship between volatility and trading volume. Furthermore, to detect the unobservable surprising information, this paper suggests a reliable method based on the extension of Khmaladze's approach to quantile regression setting (Koenker and Xiao, 2002) in place of the method that uses a quantile regression of realized volatility (Andersen *et al*., 2003) on trading volumes[3]. The method based on Khmaladze's approach is derived from the intuition that when surprising information arrives in markets, the trading volume variable may alter the location and shape of the conditional distribution of realized volatility at any upper quantile because the influence of surprising information on the relation between return volatility and trading volume may be different from that of general information.

Forecasting volatility has received a great deal of concern from researchers as well as financial market participants because it provides a decisive key to risk management, the pricing of derivative securities, and portfolio optimization. Although the generalized autoregressive conditional heteroskedasticity (GARCH) models (Bollerslev, 1986) have been extensively used to model and forecast time-varying volatility, they have an obvious disadvantage in that current variance in the models is a deterministic function of past information. Realistically, however, volatility also depends on unconsidered factors and should still contain some random components. A more heuristic approach

[2] According to Park (2010), surprising information is closely related to announcement surprise (or news) that is widely recognized in existing literature (e.g. Andersen et al., 2003). More precisely, however, it shall be defined as information that is unanticipated and leads to large market changes in terms of prices and volatility. Examples include news of the 9/11 terror attack on the U.S., a North Korean nuclear test, the breathtaking collapse of investment bank Bear Stearns (3/19/08), and the Federal Open Market Committee's decision, contrary to market expectations, to decrease the target for the federal funds rates.

[3] As researched by Lanne (2007), separate modeling for the detection of surprising information in this study could help improving forecast accuracy.

to modeling time-varying volatility is a stochastic volatility (SV) model (Taylor, 1986 and Harvey *et al*., 1994) in which the variance contains an unobserved component following a particular stochastic process. As expected, empirical results demonstrating the superior forecasting performance of several SV models are reported by Heynen and Kat (1994), Alizadeh *et al*. (2002), Sadorsky (2005), among others.

On the other hand, an important information source for volatility prediction is found in trading volume which is viewed as a proxy for the rate of information flow to the stock market and thus contemporaneously correlated with price changes (see e.g., Lamoureux and Lastrapes, 1990; Andersen, 1996; Liesenfeld, 2001). Nonetheless, as in the study by Lamoureux and Lastrapes (1990), entering contemporaneous volume into the volatility equation of an univariate model suffers from a possible simultaneity bias, induced by treating trading volume as an exogenous variable.

Overall, both surprising information and the sharing of information indicated by trading volume play important roles. This paper therefore proposes a bivariate stochastic volatility model incorporating surprising information in the volatility equations (BSV-SI)[4], which captures the dynamics of returns and trading volume, and gives rise to improvement in the performance of measuring and forecasting volatility. This model is intuitively attractive but seems to be empirically challenging because it has no closed form, and hence cannot be estimated directly by classical parameter estimation like maximum likelihood estimation. Fortunately, several estimation methods are now available due to recent increasing computational power (see Ghysels *et al*., 1996, for reviews of the related literature). In particular, Monte Carlo integration using Markov Chain (MCMC) (Jacquier *et al*., 1994) provides an efficient estimation method for the model.

In this paper we also investigate whether the forecasting performance of the BSV-SI model provides significant improvement in general forecast evaluation criteria over several SV models (a simple stochastic autoregressive volatility model, a SV with jumps model and a volume-augmented SV model) and historical volatility models with realized volatility (an autoregressive model for the realized volatility (AR-RV) and a heterogeneous autoregressive model for the realized volatility (HAR-RV) (Müller et al., 1997; Corsi, 2004)[5]). The empirical investigation is for the South Korean stock index (KOSPI) and trading volume series for 1071 trading days in the period January 2, 2004 to April 30, 2008. Five-minute frequency data is used to compute realized

[4] If some jumps are caused by surprising information, the BSV-SI model seems to be closely linked to a stochastic volatility with jumps model. In contrast to the stochastic volatility with jumps model, however, it does not require unrealistic assumptions of jump specification (e.g., Eraker et al., 2003; Maheu and McCurdy 2004) and can consider the dynamic role of volumes as a proxy for the rate of information flow in volatility. More importantly, surprising information in the model reflects not only jump components but also volatility persistence.

[5] Although no volatility-forecasting model always dominates over other alternatives, recent literature (Andersen et al., 2003, 2004 among others) shows that even simple models of realized volatility tend to be superior to the standard volatility-forecasting models in out-of-sample forecasting. Further, a HAR-RV model is known to efficiently account for long-memory effect and to have substantially good performance in volatility forecasting. In response, here, the BSV-SI model is compared with the AR-RV and HAR-RV models, which are regarded as some of the competitive ones.

volatility. The performance of out-of-sample forecasting of the models was evaluated by calculating one-step-ahead volatility forecasts over the 5-, 10-, 20- and 50-day horizons.

The remainder of the paper is organized as follows. The next section briefly describes the theoretical framework with surprising information. The third section presents the BSV-SI model and suggests a method based on Khmaladze's approach to detect unobservable surprising information. Using high frequency and daily data of KOSPI index and trading volume, the fourth section provides an empirical illustration that the out-of-sample volatility forecasts of the BSV-SI model are apparently superior to those of other competitive SV models. The final section summarizes the paper, and offers some concluding remarks.

2. The theoretical framework with surprising information

2.1 The effect of surprising information on the volatility-volume relationship

In this section, in order to derive a bivariate mixture model for surprising information, we briefly extend the MDH that is modified by Andersen (1996) based on the theoretical framework of Glosten and Milgrom (1985). For convenience, we adopt the basic setting of Andersen's (1996) sequential trade model. It is assumed that each private information arrival induces a price discovery phase followed by an equilibrium phase. In his sequential trade model, therefore, each private information arrival is likely to generate trading by the informed traders, leading to price change, but noise trading is assumed to have no influence on price change because noise trades carry no information. On the basis of the market microstructure theory, however, this assumption seems to be too restrictive to derive the exact relationship between volatility and trading volume so that it is relaxed in this paper.

- *The Return Process:* For the large number of information arrivals, and under weak regularity conditions, Andersen (1996) postulates that daily returns are conditionally normal with mean zero and variances that reflect the intensity of information arrivals, K_t:

$$r_t \mid K_t \sim N(0, \sigma^2 K_t) \tag{2.1}$$

where: r_t is the continuously compounded return, σ^2 is a variance of random return.

This specification represents that return volatility is driven by the temporal property of the information flow. By contrast, Park (2010) argues more reasonably that return volatility depends on not only the intensity of information arrivals but the type of information: general and surprising information. It is clear that surprising information flow tends to result in higher return volatility than general information flow. Convincingly, the validity of the additional randomness associated with surprising information is supported by empirical features in financial markets such as fat tails and excess kurtosis in the distribution of returns.

- *The Trading Volume Process:* According to Andersen (1996), information arrival induces trades so that price and trading volume are jointly determined by the information flow. In period *t,* the trading volume ($\mathcal{V}_t$) consists of informed and noise components:

$$\mathcal{V}_t = \mathcal{V}_t^{\mathcal{I}} + \mathcal{V}_t^{\mathcal{N}} \tag{2.2}$$

where: each trading component is dictated by a Poisson arrival process with an arrival intensity. That is, noise component ($\mathcal{V}_t^{\mathcal{N}}$) is governed by a stochastic process with a constant mean intensity of m_0 per day, and informed component ($\mathcal{V}_t^{\mathcal{I}}$) by a stochastic process with an inconstant mean intensity of $m_1 K_t$. Combining the expressions for the informed and noise components, daily volume has the following Poisson distribution:

$$\mathcal{V}_t \mid K_t \sim c \cdot Po(\boldsymbol{m}_0 + \boldsymbol{m}_1 K_t) \tag{2.3}$$

where *c* is a scaling parameter.

- *Surprising Information and the Volatility-Volume Relationship:* To look into the effect of surprising information on the volatility-volume relationship, we can consider the covariance between squared returns and trading volume in the system of joint return and trading volume distribution. In Andersen's (1996) framework, the covariance is given by

$$\mathrm{Cov}(R_t^2, \mathcal{V}_t) = \mathrm{Cov}(R_t^2, \mathcal{V}_t^{\mathcal{I}}) + \mathrm{Cov}(R_t^2, \mathcal{V}_t^{\mathcal{N}}) \tag{2.4}$$

where: $\mathrm{Cov}(R_t^2, \mathcal{V}_t^{\mathcal{I}})$ is equal to $c\,\boldsymbol{m}_1 \mathrm{Var}(K_t)$, and $\mathrm{Cov}(R_t^2, \mathcal{V}_t^{\mathcal{N}})$ is assumed to be zero. That is, Andersen (1996) does not allow for the effect of noise trading on return volatility. According to the market microstructure theory, however, this assumption is quite unrealistic. Some studies (see, e.g., Kyle 1985; Easley *et al.*, 1996) state that noise trading is inversely related to return volatility due to a positive effect of noise trading on the depth of the market. On the other hand, surprising information is likely to result in a reduction in the level of disagreement between noise traders about the price change, which is caused by noise traders' nearly unanimous reaction to the surprising information. Thus, surprising information, even associated with high volatility, may reduce noise trading. In this context, this paper suggests an intuition that surprising information, leading to negative $\mathrm{Cov}(R_t^2, \mathcal{V}_t^{\mathcal{N}})$, may cause the relationship between volatility and trading volume to be ambiguous or negative, which is consistent with Park's (2010) theoretical result. We next turn to detect surprising information empirically. Following this detection, we can indirectly verify that the effect of surprising information on the volatility-volume relationship contrasts with that of general information.

2.2 A method for detecting surprising information

Even if surprising information is not observable directly, this paper suggests a reliable criterion that distinguishes surprising information from general information. This criterion is based on the idea that, as explained earlier, in contrast to general information, the surprising information arrival leads to high volatility but little change or decrease in trading volume. Hence, in a quantile regression of realized volatility on trading volumes, the coefficient of the relationship between volatility and trading volume is likely to be statistically less significant or be transferred from positive to negative at higher quantiles of volatility distribution, which are more closely associated with surprising information.

Within this context, we propose a method for detecting surprising information that is based on Khmaladze's (1981) approach extended to quantile regression setting (Koenker and Xiao, 2002), which can be applied to test location-scale shift hypothesis. If arrivals of surprising information make a different impact on the relationship between volatility and trading volume, it might affect both the location (or scale) and the shape of the conditional distribution of volatility given trading volume at high quantiles. That is, as the quantile of the distribution is closer to one, entailing arrivals of more surprising information, the trading volume variable substantially alters the shape of the conditional distribution of volatility and hence the location-scale shift hypothesis might be rejected with a much wider range of quantile intervals. Inspired by this intuition, we propose the following process for detecting days with surprising information.

1. Calculate realized volatility (RV) at each day.

   When general conditions are satisfied, RV ($\tilde{\sigma}_t^2$) is obtained by summing intraday squared returns over many small intervals within the day (*m*):

$$\tilde{\sigma}_t^2(m) = \sum_{k=0}^{m-1} r_{t+k/m}^2 \tag{2.5}$$

   where: $r_{t+k/m} = \mathcal{P}_{t+k/m} - \mathcal{P}_{t+(k-1)/m}$. Under weak regularity conditions, RV converges in probability to the quadratic variation of the diffusion process over the day as m goes to infinity (e.g., Andersen et al., 2001, 2003):

$$\bar{\sigma}_t^2 = p\lim_{m\to\infty} \tilde{\sigma}_t^2(m) \tag{2.6}$$

   where: $\bar{\sigma}_t^2$ is the integral of the instantaneous variances over the day t.

2. Produce time series of filtered realized volatility (FRV) and detrended log-trading volume ($\mathcal{V}$).

   To filter out the autocorrelation and a day-of-the-week effect, the following regression is estimated.

$$\mathrm{RV}_t = c + \sum_{i=1}^{L} a_i \mathrm{RV}_{t-i} + \sum_{j=1}^{4} b_j D_{jt} + \varepsilon_t \tag{2.7}$$

   where: c is a constant, D_{jt}'s are day-of-the-week dummies used to capture differences in mean volatility, and RV_{t-i}'s are lagged RVs. Thus, the absolute values of the residuals are considered as FRV. On the other hand, in order to obtain $\mathcal{V}$, the following regression is estimated.

$$\mathrm{LV}_t = c + a_1 \boldsymbol{t} + a_2 \boldsymbol{t}^2 + \sum_{j=1}^{4} b_j D_{jt} + \varepsilon_t \tag{2.8}$$

where: LV_t are log-trading volumes, c is a constant, t is a linear trend, t^2 is a quadratic trend, and D_{jt} 's are day-of-the-week dummies.

3. Consider the linear location-scale regression model:

$$\mathrm{y}_t = x_t\alpha + (x_t\lambda)\varepsilon_t \tag{2.9}$$

where: $\mathrm{y}_t = \mathrm{FRV}_t$, $x_t = V_t$, α and λ are parameters, and ε_t are assumed independent and identically-distributed random varibles from some distribution F_0.

4. Using the test statistic proposed by Koenker and Xiao (2002), implement the test of the location-scale shift hypothesis with the form:

$$R\beta(\tau) = r(\tau) \tag{2.10}$$

where: τ is some index set $\tau \in [0,1]$, $\beta(\tau) = \alpha + \lambda F_0^{-1}(\tau)$, R is a q × p matrix (q ≤ p), and $r \in \mathbb{R}^q$. Under the null hypothesis we have the process:

$$\hat{v}_T(\tau) = \sqrt{T}\varphi_0(\tau)(R\Omega R')^{-1/2}(R\hat{\beta}(\tau) - \hat{r}(\tau)) \tag{2.11}$$

where: $\varphi_0(\tau) = f_0(F_0^{-1}(\tau))$, $\Omega = H_0^{-1}J_0H_0^{-1}$ with $J_0 = \lim T^{-1}\sum x_t x_t'$ and $H_0 = \lim T^{-1}\sum x_t x_t'/(\lambda' x_t)$. Then, employing Khmaladze's (1981) martingale transformation, we have $\tilde{v}_T(\tau) \equiv Q_g \hat{v}_T(\tau)$ and Kolmogorov-Smirnov-type test statistic:

$$K_T = \sup_{\tau \in \mathcal{T}} \|\tilde{v}_T(\tau)\| \tag{2.12}$$

Using l_1 norm, Koenker and Xiao (2002) report asymptotic critical values of the test statistic K_T in an electronic appendix. Note that arrivals of surprising information guarantee the existence of quantiles at which the null hypothesis is rejected. In reverse, any rejection indirectly verifies that the effect of surprising information on the volatility-volume relationship differs from general information.

5. As the range of quantile intervals, $\mathcal{T} \in [a, 1-a]$, widens from $a = 0.25$ to $a = 0.05$[6], repeat the test of the location-scale shift hypothesis with a diminishment of 0.01 in a and find a lowest $\tilde{a}$ at which the hypothesis is rejected due to the impact of surprising information. Then, at quantiles higher than the quantile $1-\tilde{a}$, we can expect the arrival of surprising information in the market. Consequently, if on given a day, t, the daily realized volatility is greater than the value of the $(1-\tilde{a})^{\text{th}}$ conditional quantile function of FRV, we can decide that surprising information arrived on t.

[6] In Koenker and Xiao (2002), asymptotic critical values of the test statistic are not reported below $a = 0.05$ and according to author's experience, the range of quantile intervals $\mathcal{T} \in [a, 1-a]$ from $a = 0.25$ to $a = 0.05$ is wide enough to find $\tilde{a}$.

3. The BSV-SI model and estimation methodology

3.1 Model specification

Stochastic volatility models have flexibility due to allowance for error processes for the conditional variance as well as the conditional mean. With regard to forecasting performance, the stochastic volatility models can be extended in two ways: 1) incorporating surprising information and 2) specifying bivariate case for returns and trading volume based on the theoretical framework in Section 2. Setting y_t be a vector of returns and detrended log-trading volume, $(\mathrm{r}_t, V_t)'$, $\mathrm{t} = 1, \cdots, \mathrm{T}$, therefore, the BSV-SI model is defined as:

$$
\begin{aligned}
&\mathrm{y}_t = \mathrm{H}_t^{1/2}\varepsilon_t \\
&\mathrm{H}_t = \mathrm{diag}\{\exp(h_{rt}), \exp(h_{Vt})\} \qquad \begin{pmatrix}\varepsilon_t \\ \eta_t\end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\Sigma_\varepsilon & 0 \\ 0 & \Sigma_\eta\end{pmatrix}\right) \\
&h_t - \Psi_t = \phi \circ (h_{t-1} - \Psi_{t-1}) + \eta_t \\
&\Psi_t = \mu + \alpha \circ \mathrm{x}_{1t} + \beta \circ \mathrm{x}_{2t}
\end{aligned}
\tag{3.1}
$$

where: $h_t = (h_{rt}, h_{Vt})'$ is the logarithm of the latent volatilities and $\phi = (\phi_r, \phi_V)'$, $\mu = (\mu_r, \mu_V)'$, $\alpha = (\alpha_r, \alpha_V)'$, $\beta = (\beta_r, \beta_V)'$ are parameter vectors. In the transition equation, x_{1t} and x_{2t} are explanatory variables, defined as $\mathrm{x}_{1t} = (\mathrm{D}^{\mathrm{S}}_{\mathrm{t-1}}, \mathrm{D}^{\mathrm{S}}_{\mathrm{t-1}})'$, $\mathrm{x}_{2t} = (V_{t-1}, |\,\mathrm{r}_{\mathrm{t-1}}\,|)'$ where $\mathrm{D}^{\mathrm{S}}_{\mathrm{t-1}}$ is a dummy variable for surprising information. The operator $\circ$ denotes the Hadamard product. The persistent parameter is assumed to satisfy $|\,\phi\,| < 1$, implying that h_t is stable, and $\varepsilon_t = (\varepsilon_{rt}, \varepsilon_{Vt})'$, $\eta_t = (\eta_{rt}, \eta_{Vt})'$ are vectors of error terms. Corresponding to the MDH, it is assumed that Σ_ε is an identity matrix, but the off-diagonal elements of Σ_η are not all equal to zero:

$$
\Sigma_\eta = \begin{pmatrix} \sigma^2_{\eta\mathrm{r}} & \sigma^2_{\eta\mathrm{rv}} \\ \sigma^2_{\eta\mathrm{rv}} & \sigma^2_{\eta\mathrm{v}} \end{pmatrix}
\tag{3.2}
$$

3.2 Estimation using a MCMC methodology

The BSV-SI model cannot be easily estimated by standard maximum likelihood methods because both the parameter vector and the unobserved volatility vector have to be estimated simultaneously and further the observation equation is nonlinear in the state variable. Indeed, from the BSV-SI model, the log volatility vector is defined as $h_t = (h_1, \cdots, h_T)$ and the parameter vector is defined as $\Xi = (\mu_r, \mu_V, \phi_r, \phi_V, .\ \alpha_r, \alpha_V, \beta_r, \beta_V, \sigma^2_{\eta\mathrm{r}}, \sigma^2_{\eta\mathrm{v}}, \sigma^2_{\eta\mathrm{rv}})$. Therefore, the likelihood function of the model is the conditional density of data y:

$$L(\Xi) = f(y \mid \Xi) = \int f(y,h \mid \Xi)dh \propto \int f(y \mid h,\Xi) f(h \mid \Xi)dh \qquad (3.3)$$

Denoting $Y_{t-1} = (y_1, \cdots, y_{t-1})$, the density of the data can be represented as a mixture over the log volatility vector using law of total probability:

$$f(y \mid \Xi) = \prod_{t=1}^{T} f(y_t \mid Y_{t-1}, \Xi) \propto \prod_{t=1}^{T} \int f(y_t \mid h_t, Y_{t-1}, \Xi) f(h_t \mid Y_{t-1}, \Xi)\, dh_t \qquad (3.4)$$

where: this likelihood function is intractable because the density function $f(h_t \mid Y_{t-1}, \Xi)$ has no closed form so that $y_t \mid Y_{t-1}$ cannot be analytically expressed. To resolve this problem, the MCMC methodology, proposed by Shephard (1993), has been extensively used. According to the results of simulation for estimating stochastic models, the MCMC methodology is superior to the alternatives such as a quasi-maximum likelihood method or GMM method in terms of sampling properties (Jacquier *et al.*, 1994).

The procedure of the Bayesian MCMC estimation algorithms for the BSV-SI model is straightforward. At first, it constructs a Markov chain in which limiting invariant distribution is the target distribution, $f(\Xi, h \mid y)$. Simulating the Markov chain a large number of times and recording its values produces a sample of draws from the target distribution. However, the likelihood function for the BSV-SI model cannot be obtained and this precludes the direct construction of the posterior distribution function $f(\Xi \mid y)$. The problem requires that the parameter vector Ξ is augmented by the log volatility vector h to form a large parameter vector (Ξ, h) and the joint distribution of (Ξ, h) conditional upon y, i.e., $f(\Xi, h \mid y)$, is substituted for the posterior distribution function $f(\Xi \mid y)$. Consequently, this allows the MCMC methodology to sample the density without computing the likelihood function and to estimate both the parameter vector Ξ and the log volatility vector h[7].

4. Forecasting return volatility of KOSPI index

4.1 The Data

The data employed in this empirical application are daily and high-frequency KOSPI and trading volumes (the number of trading stocks) in the stock market of South Korea, covering the period between January 2, 2004 and April 30, 2008[8]. After missing values are dropped, the sample contains a total of 1070 observation days.

[7] For more discussion along this line, see Chib et al. (2002), Broto and Ruiz (2004), Asai et al. (2006).

[8] The span of the sampling period has high volatility due to the ‘U.S. subprime mortgage crisis,’ which has caused panic in global financial markets, and other several pieces of big news. For representative news, we can take the following: “China shock” (April 29, 2004), triggered by Chinese premier Wen Jiabao’s comments on cooling down the overheating Chinese economy and dealing a sharp blow to the Korean financial markets, “Announcement that North Korean implemented nuclear test” (October 9, 2006), “News about credit fears in global

To avoid problems arising from the non-stationary behavior usually observed in stock prices, we take the natural logarithmic differences between two successive trading days. The differences of log prices are defined as returns: $r_t = (\mathcal{P}_t - \mathcal{P}_{t-1}) \cdot 100$. This paper also calculates absolute daily return residuals ($|\hat{\varepsilon}_t|$) which have been widely used to estimate daily volatility. They are obtained from estimating the following regression model:

$$r_t = c + \sum_{i=1}^{2} a_i r_{t-i} + \sum_{j=1}^{4} b_j D_{jt} + \varepsilon_t \qquad (4.1)$$

where c is a constant, D_{jt} values' are day-of-the-week dummies used to capture differences in mean returns, and r_{t-i} values' are lagged returns. Based on Akaike's information criterion (AIC), the two lag length is chosen to control for serial dependence in returns[9]. Since existing literature (e.g., Andersen *et al.*, 2001; Bandi and Russell, 2008) indicate that five-minute observations are close to optimal sampling intervals, the high-frequency data used here to measure realized volatility are five-minute observations of spot markets, which are provided by the Korea Exchange (KRX, http://sm.krx.co.kr)[10]. The use of a five-minute frequency, corresponding to 60 intraday observations (*m*=60), means that the total data used in this study were obtained from 64,200 observations. Using the method in section 2.2, realized volatility estimates are calculated by summing squares of MA(1)-filtered intraday returns[11]: $\mathrm{RV}_t \equiv \sum_{j=0}^{59} r_{t+j/60}^2$.

4.2 Preliminary statistics and findings

In table 1, a wide range of preliminary statistics for returns (r_t), trading volume ($\mathcal{V}$), detrended log-trading volume (V), absolute return residuals ($|\hat{\varepsilon}_t|$), and realized volatility (RV) are reported. These include the following distributional parameters and test statistics: minimum, maximum, quartiles, mean, variance, skewness, kurtosis, Ljung-Box Q statistic for the null hypothesis of no serial correlation, Jarque-Bera (JB) statistic for the null hypothesis of normality, and augmented Dickey-Fuller statistic

financial markets triggered from the fact that France's biggest-listed bank, BNP Paribas, had frozen $2.2 billion worth of funds hit by subprime mortgage crisis" (Aug. 10, 2007) .

9 Unlike what has been found in some financial markets, there is little evidence of a weekday effect in returns because none of the coefficients for weekday dummy variables is significant. Thus, the day-of-the-week effect for returns is not taken into account. The estimation results are not reported in this paper, but can be obtained from the author upon request.

10 The Korea Exchange was established by consolidating three Korean domestic exchanges: Korea stock exchange, KOSDAQ market, and Korea future exchange.

11 Realized volatility suffers from a bias problem resulting from market microstructure noise, causing autocorrelation in the intraday returns (Hansen and Lunde, 2006). To remove this autocorrelation and resolve the problem, filtering techniques such as a moving average (MA) filter have been used by Andersen et al. (2001), Maheu and McCurdy (2002), and others.

(ADF) for the null hypothesis of unit root. The graphs of the series are presented in Figure 1. The first panel of this figure shows the movements of returns, which supports the view that the return series is highly dynamic. The return series also tends to be clustered together over time. For the series, the coefficients of skewness and kurtosis support the view of non-normality and, as expected, the Jarque-Bera (JB) statistic safely rejects normality.

The second and third panels of Figure 1 show the turbulent movements of absolute return residuals $|\hat{\varepsilon}_t|$ and RV, respectively. For the series, we observe two extremely high volatility periods which occurred during the second quarter of 2004 (influence of China shock) and after the summer of 2007 (influence of U.S. subprime mortgage crisis). Their movements seem to be similar but realized volatility might have much less variation than absolute return residuals except when measured at several extreme values. Hence, the use of such a noisy volatility estimator will undermine the inference regarding accuracy of volatility forecasts.

Table 1

Summary statistics for r_t, $|\hat{\varepsilon}_t|$, RV, $\mathbf{v}$ and $\mathcal{V}$

Statistics	r_t	$\mathbf{v}$	$\mathcal{V}$	RV	$\lvert\hat{\varepsilon}_t\rvert$
Minimum	-7.4078	1.2348e+08	-0.8296	0.0820	0.0008
Maximum	5.3769	9.3472e+08	1.0254	10.1767	7.4299
First Quartile	-0.5530	2.5874e+08	-0.2265	0.4910	0.3072
Third Quartile	0.8462	4.1523e+08	0.2214	1.2144	1.3376
Median	0.1687	3.3000e+08	-0.0146	0.7433	0.7146
Mean	0.0744	3.4366e+08	0.0000	1.0144	0.9510
Variance	1.6922	1.3389e+16	0.1031	0.8938	0.7750
Skewness(Sk=0)	-0.6577	0.7170	0.0950	3.6466	1.8899
Kurtosis(Ku=0)	2.4910	0.6460	-0.4581	19.8852	5.7525
Q(10)	13.6 (0.1898)	6048.9 (0.0000)	6462.2 (0.0000)	2164.6 (0.0000)	243.6 (0.0000)
Q(50)	56.8 (0.2347)	15340.1 (0.0000)	18184.5 (0.0000)	3912.3 (0.0000)	468.0 (0.0000)
JB	355.96 (0.0000)	110.99 (0.0000)	10.85 (0.0043)	2120.06 (0.0000)	2116.59 (0.0000)
ADF	-30.946 (0.0000)	-4.6690 (0.0001)	-4.9026 (0.0000)	-7.5278 (0.0000)	-7.5217 (0.0000)

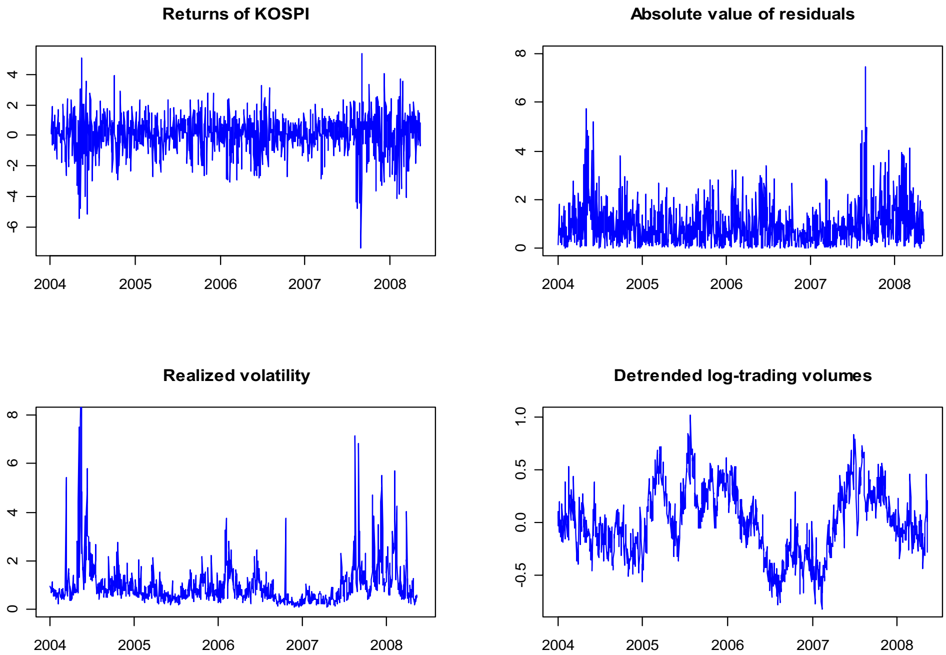
Note: Skewness and Kurtosis are coefficients of skewness and kurtosis, respectively. Q(M) is the Ljung-Box Q statistics at lag M for series, JB is the Jarque-Bera normality test and ADF is the augmented Dickey-Fuller statistic including an intercept term. P-values are in parentheses.

The coefficients of skewness and kurtosis for RV and $\mathbf{v}$ support the view that each distribution is not normal, coinciding with the other empirical findings. In particular, the high kurtosis of RV is attributed to large outliers that should be strongly associated with surprising information flows. Jarque-Bera (JB) statistics for RV and $\mathbf{v}$ also reject normality at the conventional 5-percent level. Another important property of the series is revealed by the joint test of serial correlation. The Ljung-Box Q statistics are computed up to the tenth and fiftieth lags. Under the null hypothesis with no serial correlation, such statistics have an asymptotic chi-square distribution with ten and fifty degrees of freedom. The Ljung-Box Q statistics indicate a high serial correlation in RV and $\mathcal{V}$ series, which reveals a long memory feature. The fourth panel of Figure 1 displays the detrended log-trading volume series ($\mathcal{V}$)[12], which will be used in our empirical study. As expected, for detrended log-trading volume, non-normality is reduced considerably and more correlation structure in the series appears in comparison with the raw trading volume series.

In order to test for the stationarity of the series we use the augmented Dickey-Fuller (1979) tests of the null of unit root against the stationary series. The results, provided in Table 1, indicate that all series are clearly stationary. Especially, the stationarity nature for the series of RV and $\mathcal{V}$ allows us to apply traditional economic models to empirical works.

Figure 1

KOSPI returns, absolute value of residuals, RV, and DTLV



[12] It is necessary to filter out the trend as previous studies have done (e.g., Jorion, 1996; Park, 2007). The estimation results using Eq. (2.7) are not reported in this paper, but can be obtained from the author upon request.

4.3. Detecting unobservable surprising information

Using the method described in the section 2.2, we can detect days with arrival of surprising information. On the basis of the linear location-scale regression model Eq. (2.9), we implement the test of the location-scale shift hypothesis at the range of quantile intervals ($\mathcal{T} \in [a, 1-a]$) with a diminishment from $a = 0.25$ to $a = 0.05$ by 0.01 and find a lowest $\tilde{a}$ at which the hypothesis is rejected due to the impact of surprising information.

Table 2 reports the estimates of Kolmogorov-Smirnov-type test statistic with l_1 norm. According to asymptotic critical values for the test statistic provided by Koenker and Xiao (2002), the 5-percent critical values for $a = 0.1$ and $a = 0.05$ are 2.102 and 2.140, respectively. Therefore, $a = 0.07$ is a lowest $\tilde{a}$ at which the hypothesis is rejected and consequently, if on given a day t, the filtered realized volatility (FRV) is greater than the value of $Q_{FRV_t}(0.93 \mid \mathcal{V}_t)$, we can decide that surprising information arrived on t. For instance, the days of “China shock” (April 29, 2004), “North Korea nuclear test” (October 9, 2006), and “News about credit fears in global financial markets” (Aug. 10, 2007) are detected as days having surprising information.

This empirical result supports the argument that, as the quantile of the distribution is closer to one, entailing arrivals of more surprising information, the trading volume variable substantially alters the shape of the conditional distribution of volatility and hence the location-scale shift hypothesis is rejected with a much wider range of quantile intervals. Therefore, it corroborates the Park’s (2010) idea of the negative effect of surprising information on the relationship between volatility and trading volume.

Table 2

Test of the location-scale shift hypothesis

$\mathcal{T} \in [a, 1-a]$	Test statistic	$\mathcal{T} \in [a, 1-a]$	Test statistic
[0.25, 0.75]	0.1461	[0.10, 0.90]	1.5277
[0.20, 0.80]	0.4986	[0.09, 0.91]	1.5796
[0.15, 0.85]	0.1753	[0.08, 0.92]	1.9882
[0.14, 0.86]	0.8622	[0.07, 0.93]	2.1981*
[0.13, 0.87]	1.2168	[0.06, 0.94]	2.6449*
[0.12, 0.88]	1.1561	[0.05, 0.95]	3.1881*
[0.11, 0.89]	1.3426		

Note: Asterisks indicate parameters statistically significant at the 5-percent level.

4.4 Forecasting results and the relative performance of the BSV-SI model

Given our emphasis on surprising information, we now turn to estimation of the BSV-SI model using KOSPI returns and trading volumes, and compare its forecasting performance with SV models and historical volatility models with realized volatility that have been commonly specified in the literature:

- SV Model (simple stochastic volatility model):
$$r_t = \exp(h_t/2)\varepsilon_t,\ \ h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t.$$
- SV-J Model (stochastic volatility model including a jump variable whose coefficient is changed over time): $r_t = \exp(h_t/2)\varepsilon_t,\ h_t = \mu + \phi(h_{t-1} - \mu) + \lambda_t q_t + \eta_t$, where q_t is a Bernoulli random variable that takes 1 with probability κ and 0 with probability $1-\kappa$ and the size of jumps is denoted by $\lambda_t \sim \mathrm{N}(m, v)$.
- SV-V Model (stochastic volatility model including a lagged volume variable): $r_t = \exp(h_t/2)\varepsilon_t, h_t = \mu + \beta V_{t-1} + \phi(h_{t-1} - (\mu + \beta V_{t-2})) + \eta_t$, where V_{t-1} is a lagged variable for detrended log-trading volume.
- AR-RV Model (autoregressive model for the realized volatility): $\mathrm{RV}_t = b_0 + \sum_{i=1}^{17} b_i \mathrm{RV}_{t-i} + e_t$, where the AR(17) process is constructed to consider the autocorrelation in the square root of realized volatility based on AIC.
- HAR-RV Model (heterogeneous autoregressive model for the realized volatility): $\mathrm{RV}_t = b_0 + b_D \mathrm{RV}_{t-1} + b_W \mathrm{RV}_{t-6,t-1} + b_M \mathrm{RV}_{t-23,t-1} + e_t$, where $\mathrm{RV}_{t,t+k}$ is the multiperiod realized volatility that is normalized sums of the one-period realized volatilities (Corsi, 2004). That is, it is denoted by $\mathrm{RV}_{t,t+k} = k^{-1}(\mathrm{RV}_{t+1} + \mathrm{RV}_{t+2} + \cdots + \mathrm{RV}_{t+k})$ and $k = 1, 2, \cdots$ so that $\mathrm{RV}_{t,t+1} \equiv \mathrm{RV}_{t+1}$.

For the posterior computation in the stochastic volatility models, WinBUGS is used because it is well known that WinBUGS provides an efficient implementation of the MCMC algorithm. The posterior quantities are computed from 10,000 draws of the MCMC algorithm, collected after an initial burn-in period of 1,000 iterations[13]. To estimate the stochastic volatility models via the MCMC algorithm, we should specify suitable prior distributions on the unknown parameter vector Ξ and assume each parameter to be independent. For instance, a prior density function of $\Xi = (\mu, \phi, \sigma_\eta^2)$ for SV Model is $f(\Xi) = f(\mu)f(\phi)f(\sigma_\eta^2)$. In this empirical study, prior distributions on Ξ basically follow from the specification of Chib et al. (2002) or Yu and Meyer (2006). Thus, the prior distributions for parameters $\mu_i, \phi_i, \sigma_{\eta i}^2,\ \mathrm{i} = r, V$ are the following: $\mu_i \sim \mathrm{N}(0,100)$; $\phi_i^* \sim \mathrm{Beta}(20,1.5)$ where $\phi_i^* = (\phi_i + 1)/2$; $\sigma_{\eta i}^2 \sim \mathrm{IG}(2.5, 0.025)$. Furthermore, we assume $\alpha_i \sim \mathrm{N}(0,1)$; $\beta_i \sim \mathrm{N}(0,1)$; $\lambda_t \sim \mathrm{N}(m, v)$, where $m \sim \mathrm{N}(0,100)$, $v \sim \mathrm{IG}(2.5, 0.025)$.

[13] Although it is known that 5,000 draws are large enough to get accurate estimation using an efficient Metropolis algorithm (e.g., Chib et al., 2002), this paper takes 10,000 draws for optimal implementation in terms of speed and accuracy.

In Table 3 the estimation results of stochastic volatility models are reported and in Table 4 those of historical volatility models with realized volatility are reported. The first 1020 days of the data set are used for estimating the volatility models and the last 50 days for out-of-sample forecasting. It is noted that we use only one set of parameter estimates and not re-estimate the parameters within the out-of-sample period. The reason for not using rolling estimates is the computational burden of the BSV-SI model.

Table 3

Estimation results of BSV-SI and SV-type models

Parameter	SV-J Model			SV Model		
	Mean	SD	95% CI	Mean	SD	95% CI
μ	0.1899	0.1662	(-0.1429 0.5147)	0. 2482	0.1580	(-0.062 0.5374)
ϕ	0.9571	0.0196	(0.9036 0.9855)	0.9648	0.0206	(0.9111 0.9879)
κ	0.0017	0.0012	(0.0002 0.0049)			
m	-14.720	0.1610	(-14.980 -			
σ_η	0.1971	0.0480	(0.1355 0.3199)	0.1833	0.0486	(0.1676 0.2999)
RMSE	0.9264			1.0162		
MAE	0.6387			0.6984		

Parameter	BSV-SI Model			SV-V Model		
	Mean	SD	95% CI	Mean	SD	95% CI
μ or μ_r	-0.1610	0.1764	(-0.5494 0.1693)	0.0703	0.0650	(-0.0637 0.1753)
μ_V	-0.6570	0.3494	(-0.1323 -			
ϕ or ϕ_r	0.9698	0.0704	(0.7802 0.9981)	0.9810	0.0095	(0.9519 0.9941)
ϕ_V	0.6568	0.1677	(0.2454 0.8308)			
α_r	0.7245	0.2448	(0.4004 1.2700)			
α_V	0.3640	0.2043	(0.0058 0.6965)			
β or β_r	-0.6297	0.2897	(-1.0950 -	-1.2210	0.2646	(-1.6660 -0.7915)
β_V	-0.0346	0.0369	(-0.0952 0.0304)			
σ_η or $\sigma_{\eta r}$	0.0742	0.0024	(0.0723 0.0801)	0.1634	0.0320	(0.1288 0.2653)
$\sigma_{\eta v}$	0.0620	0.0019	(0.0587 0.0670)			
ρ	0.0212	0.0005	(0.0204 0.0225)			
RMSE	0.8202			1.2034		
MAE	0.5307			0.7883		

Note: SD stands for standard deviation, '95% CI' denotes 95% confidence interval, and

$\rho = \sigma^2_{\eta rv} / (\sigma_{\eta r}\sigma_{\eta v})$.

Table 3 contains the mean, standard deviation, 95% confidence interval of the posterior distribution for parameters, and two general loss functions (Root-Mean-Squared Errors: RMSE and Mean-Absolute Errors: MAE) to measure the in-sample forecasting accuracy of the models. The loss functions are presented in the following equations:

$$\text{RMSE} = \sqrt{T^{-1}\sum_{t=1}^{T}\left(\sigma_{t+1}^{2} - \hat{\sigma}_{t+1}^{2}\right)^{2}} \quad \text{MAE} = T^{-1}\sum_{t=1}^{T}\mid \sigma_{t+1}^{2} - \hat{\sigma}_{t+1}^{2} \mid$$

where: σ_{t+1}^{2} is the actual volatility and $\hat{\sigma}_{t+1}^{2}$ is the predicted volatility obtained from volatility models. It is argued earlier that the actual volatility is not observable.

Thus, we use realized volatility (2.5) as a proxy for the actual volatility[14] so that the forecast error here is obtained from the difference between realized volatility and predicted volatility.

Table 4

Estimation results of realized volatility models

	AR-RV Model			HAR-RV Model	
Parameter	Estimate	S.E.	Parameter	Estimate	S.E.
b_0	0.1323	0.0428	b_0	0.1314	0.0510
b_1	0.4715	0.0638	b_D	0.4179	0.1020
b_2	0.1359	0.0640	b_W	0.2954	0.1115
b_3	0.0317	0.0507	b_M	0.1625	0.0947
⋮	⋮	⋮			
b_{16}	0.0418	0.0331			
b_{17}	0.1045	0.0403			
RMSE	0.6724		RMSE	0.6878	
MAE	0.3957		MAE	0.4004	

Note: S.E. is standard errors that are calculated by Newey-West heteroskedasticity consistent covariance matrix estimator.

The estimation results of SV-type models show that the parameter of volatility persistence ϕ is highly credible and its values in all models are close to, but less than unity. This implies high persistence in the volatility process which may be useful for purpose of forecasting conditional volatility (Akgiray, 1989). Turning to the results of the HAR-RV model, the estimates for b_D, b_W, and b_M also support the existence of highly persistent volatility dependence. Interestingly, the high persistence in volatility is not substantially diminished by inclusion of the lagged trading-volume variable, jump variable, or surprising information dummy variable. Whereas this peculiarity is not anticipated, we can attribute it to recent long-term turbulence due to global financial crisis. On the other hand, the estimate of ϕ_V is 0.6568 and it reveals that the volatility persistence in the volume process is weaker compared to the return process. This

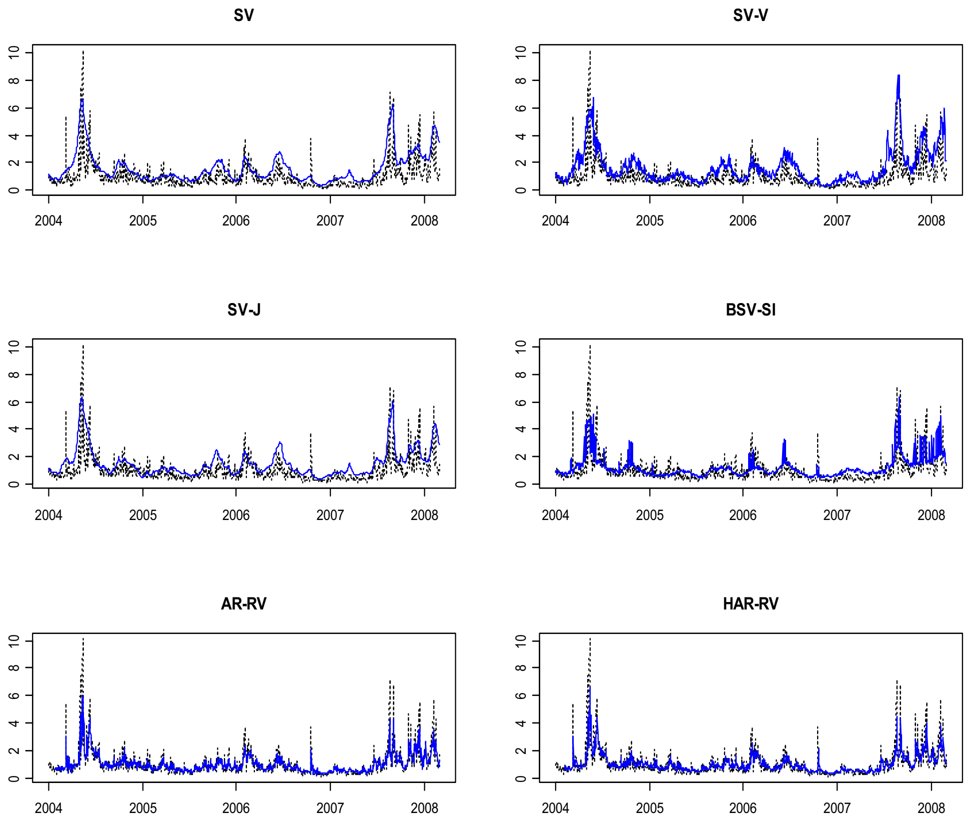
[14] It is known that the integrated volatility is a natural measure of return volatility. Despite the non-observability of integrated volatility, the theory of quadratic variation guarantees that the integrated volatility is converged uniformly in probability by realized volatility as m goes to infinity (e.g., Andersen et al., 2001, 2003).

feature is also found by Liesenfeld (2001). Another empirical finding reported in this paper indicates that volatility in both daily stock returns and trading volumes is significantly amplified by surprising information and the impact of surprising information on volatility is more prominent for returns rather than trading volumes, i.e. $\alpha_r = 0.7245$ $\alpha_V = 0.3640$.

A nontrivial observation from the plots in Figure 2 is that all SV-type models except the BSV-SI Model significantly over-predict volatility. This can be inferred from the statistical finding that daily squared return residuals are much more noisy estimates than realized volatility. Furthermore, since realized volatility is used to proxy the actual volatility in this study, it is natural to find that SV-type models are inferior to realized volatility models in in-sample forecasts. Based on the RMSE and MAE, the AR-RV model and HAR-RV model are the best and second best fitting models out of the six models, respectively. The difference between the RMSE for the AR-RV model and its nearest apparent competitor, the BSV-SI Model, is about 22 percent.

Figure 2

The in-sample forecasts (line = forecasts, dots = RV)



Next, we measure the accuracy of the models in forecasting the one-step-ahead (one-day-ahead) conditional volatility via two general loss functions: RMSE, MAE and two more elaborate loss functions: Heteroskedasticity-Adjusted RMSE (HRMSE) and Heteroskedasticity-Adjusted MAE (HMAE), which take into account the heteroskedasticity environment, were considered by West *et al*. (1993), Bollerslev and Ghysels (1996) and others. The heteroskedasticity-adjusted loss functions that may produce reliable results in the highly non-linear environment are defined by the following equations:

$$\mathrm{HRMSE} = \sqrt{T^{-1}\sum_{t=1}^{T}\left(\frac{\sigma_{t+1}^2}{\hat{\sigma}_{t+1}^2}-1\right)^2} \quad \mathrm{HMAE} = T^{-1}\sum_{t=1}^{T}\left|\frac{\sigma_{t+1}^2}{\hat{\sigma}_{t+1}^2}-1\right|$$

where: T is the total number of observations.

Table 5

Out-of-sample forecasting criteria (evaluated against $\sigma_{t+1}^2 = \mathrm{RV}_{t+1}$)

		BSV-SI	SV	SV-J	SV-V	AR-RV	HAR-RV
D=5	RMSE	0.3481 ①	0.5963 ④	0.4398 ②	1.0742 ⑥	0.6272 ⑤	0.5636 ③
	MAE	0.3010 ①	0.5295 ⑤	0.3627 ②	0.8731 ⑥	0.4377 ③	0.4933 ④
	HRMSE	0.4046 ③	0.4543 ⑤	0.3855 ①	0.5387 ⑥	0.3995 ②	0.4361 ④
	HMAE	0.3592 ③	0.4040 ⑤	0.3252 ②	0.4822 ⑥	0.3236 ①	0.3989 ④
D =10	RMSE	0.3306 ①	0.6392 ④	0.4308 ②	1.1586 ⑥	0.5074 ⑤	0.4466 ③
	MAE	0.2839 ①	0.5842 ⑤	0.3661 ②	1.0489 ⑥	0.3876 ④	0.3771 ③
	HRMSE	0.3781 ①	0.4883 ⑤	0.4002 ③	0.6021 ⑥	0.4077 ④	0.3978 ②
	HMAE	0.3292 ①	0.4469 ⑤	0.3436 ②	0.5673 ⑥	0.3613 ③	0.3631 ④
D=20	RMSE	0.7229 ①	0.8853 ⑤	0.9432 ④	1.1283 ⑥	0.8011 ②	0.8193 ③
	MAE	0.5149 ①	0.7043 ⑤	0.6461 ④	0.9509 ⑥	0.5899 ②	0.6012 ③
	HRMSE	0.5192 ①	0.6954 ③	1.1402 ⑥	0.6037 ②	0.7814 ⑤	0.7672 ④
	HMAE	0.4338 ①	0.5485 ③	0.7205 ⑥	0.5266 ②	0.5670 ④	0.5712 ⑤
D=50	RMSE	0.5469 ①	0.7463 ⑤	0.6378 ④	1.0326 ⑥	0.5871 ②	0.5974 ③
	MAE	0.4112 ④	0.6342 ⑤	0.4071 ③	0.9104 ⑥	0.4053 ①	0.4093 ②
	HRMSE	0.4660 ①	0.6040 ③	0.7835 ⑥	0.6042 ④	0.6128 ⑤	0.5737 ②
	HMAE	0.4016 ①	0.5154 ⑤	0.4892 ④	0.5529 ⑥	0.4398 ③	0.4215 ②

Note: Numbers inside circles are model rankings for forecasting performance in terms of each loss function. D is the number of days beyond the sample.

For out-of-sample volatility forecasting, we consider four periods: two short terms (D = 5, 10), medium term (D = 20) and long term (D = 50). The forecast summary statistics from Tables 5 exhibit some features. Most interestingly, in contrast to the results of in-sample forecasting, the realized volatility models do not dominate over the SV-type models in every category of loss functions. Between the AR-RV and HAR-RV models, there is no clear distinction in accuracy. Although there is no uniformly most accurate model, the BSV-SI Model tends to outperform all other models, whatever the forecast period. It appears that its superiority of forecasting is prominent in short forecasting horizon. For example, in 5-day forecasting period the RMSE for the BSV-SI model is 20.85 percent less than that for the second best model (SV-J model), and 38.24 percent less than that for third best model (HAR-RV model). As anticipated, since other models are likely to be sensitive to long-term turbulence in volatility triggered by surprising information, they tend to over-predict volatility and their forecasting performance are uniformly poor.

For completeness, we also report the model rankings according to each criterion. In terms of the model rankings, the BSV-SI Model is remarkably superior to other models. In particular, the RMSE ranks the BSV-SI Model first across all forecasting periods. Besides the BSV-SI Model, the SV-J model provide more accurate forecasts when the forecasting period is shorter, whereas the realized volatility models do so as forecasting period is longer. For instance, the HRMSE ranks the SV-J model first in 5-day horizon and the MAE ranks the AR-RV model first in 50-day horizon. Overall, it is worthwhile to note that although the empirical results only apply to the Korean financial market, they confirm the important role of surprising information in volatility forecasting.

5. Discussion and Conclusions

Forecasting the volatilities of asset returns has been a central theme in the recent literature of financial economics. Yet, no research has focused on the role of surprising information in forecasting volatility. By definition, surprising information causes exceptionally high volatility and persistence. Therefore, this paper argues that standard models that fail to account for surprising information may over-predict volatility and suffer from poor forecasting performance. In addition, it is well known that trading volume as a proxy for the rate of information flow to the stock market is a matter for volatility forecasting. Therefore, on the basis of the extension of the modified MDH this paper has proposed a bivariate stochastic volatility model with surprising information (BSV-SI), which is also designed to capture the dynamics of returns and trading volume. This paper has presented a reliable method for detecting days with surprising information using Khmaladze's approach extended to quantile regression setting. Furthermore, from a practical perspective, the paper has demonstrated that the MCMC provides an efficient method of estimating the model.

To evaluate predictability of the BSV-SI model as compared with other volatility models, we have used five-minute frequency data on KOSPI and daily trading volume data in the stock market of South Korea, covering the period from January 2, 2004 to April 30, 2008. Empirical results show that although the AR-RV and HAR-RV models dominate SV type models in in-sample forecasting, this domination does not extend to out-of-sample forecasting. Despite the inexistence of a uniformly most accurate model, it is obvious that the BSV-SI model does provide more accurate volatility forecasts than realized volatility models and other standard SV models in the empirical study with KOSPI data. The accuracy gains are fairly robust with respect to loss functions and forecasting periods. As such, these empirical results clearly demonstrate the important role that surprising information and trading volumes play in volatility forecasting.

In this research stream there are some of important next steps: First, we need to consider asymmetry in volatility that is well-known phenomenon in financial markets. Including this asymmetry into the BSV-SI model may improve the performance in forecasting future volatility significantly. Second, it is necessary to develop a more sophisticated test statistic for detecting surprising information or extension of Koenker and Xiao's (2002) approach, which does not require the assumption of identical and independent distribution of error terms in quantile regression. Third, it is important to investigate whether the considerable accuracy gain of the BSV-SI model may be obtained when other data samples are employed from a variety of markets (e.g. foreign exchange market, option market, or forward market). Thus, we leave further work along these lines for future research.

Acknowledgements

I would also like to express my appreciation to J.E. Song, Y.Y. Kim, C.H. Hwang and seminar participants at the Financial Economic Meeting for their valuable comments. The present research was conducted by the research fund of Dankook University in 2009.

References

Admati, A.R. and Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. *Review of Financial Studies,* 1, pp.3-40.

Akgiray, V. 1989. Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of Business*, 62, pp.55-80.

Alizadeh, S. Brandt, M. W. and Diebold, F. X., 2002. Range-based estimation of stochastic volatility models. *Journal of Finance*, 57, pp.1047-1091.

Andersen, T.G. 1996. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance*, 51, pp.169-204.

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

Andersen, T.G. Bollerslev, T. Diebold, F. X. and Labys, P., 2003. Modeling and forecasting realized volatility. *Econometrica*, 71, pp.579-625.

Andersen, T.G. Bollerslev, T. and Meddahi, N., 2004. Analytic evaluation of volatility forecasts. *International Economic Review*, 45:1079–1110.

Asai, M., McAeer, M. and Yu, J., 2006. Multivariate stochastic volatility: A review. *Econometric Reviews*, 25, pp.145-175.

Bandi, F.M. and Russell, J.R., 2008. Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75, pp.339-369.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics.* 31, pp.307-327.

Bollerslev, T. and Ghysels, E., 1996. Periodic autoregressive conditional heteroskedasticity. *Journal of Business and Economic Statistics*, 14(2), pp.139-151

Broto, C. and Ruiz, E., 2004. Estimation methods for stochastic volatility models: a survey. *Journal of Economic Surveys*, 18, pp.613–649.

Chib, S. Nardari, F. and Shephard, N., 2002. Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, 108, pp.281-316.

Clark, P.K. 1973. A subordinated stochastic process model with finite variance for speculative process. *Econometrica* , 41, pp.135-155.

Corsi, F., 2004. A simple long memory model of realized volatility. Manuscript, University of Lugano.

Copeland, T.E., 1976. A model of asset trading under the assumption of sequential information arrival. *Journal of Finance,* 31, pp.1149-1168.

Dickey, D. and Fuller, W.A., 1979. Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, pp.427-431.

Easley, D. Kiefer, N.M., O'Hara, M. and Paperman, J. B., 1996. Liquidity, information, and infrequently traded stocks. *Journal of Finance*, 51(4), pp.1405-1436.

Eraker, B. Johannes, M. S. and Polson, N. G., 2003. The impact of jumps in volatility and returns. *Journal of Finance*, 58, pp.1269-1300.

Ghysels, E. Harvey, A. and Renault, E., 1996. Stochastic volatility”. In: Rao, C.R., Maddala, G.S. (Eds.), Statistical Methods in Finance, North-Holland, Amsterdam, pp. 119-191.

Glosten, L.R. and Milgrom, P.R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), pp.71-100.

Harvey, A.C. Ruiz, E. and Shepard, N., 1994. Multivariate stochastic variance models. *Review of Economic Studies*, 61, pp.247-264.

Hansen, P.R. and Lunde, A., 2006. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24(2), pp.127–161.

Heynen, R.C. and Kat, H.M., 1994. Volatility prediction: A comparison of stochastic volatility, GARCH(1,1) and EGARCH(1,1) models. *Journal of Derivatives*, 2, pp.50-65.

Jacquier, E. Polson, N.G. and Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics*, 12(4), pp.371-417.

Jorion, P., 1996. Risk and turnover in the foreign exchange market”. In: Frankel, Galli and Giovannini (Eds.) *The Microstructure of foreign exchange markets*, University of Chicago Press, Chicago, pp. 19-36.

Khmaladze, E.V., 1981. Martingale approach in the theory of goodness-of-fit tests. *Theory of Probability and its Applications*, 26, pp.240-257.

Koenker, R. and Xiao, Z. 2002. Inference on the quantile regression process. *Econometrica*, 70, pp.1583-1612.

Kyle, A.S., 1985. Continuous auctions and insider trading. *Econometrica*, 53, pp.1315-1336.

Lamoureux, C.G. and Lastrapes, W.D., 1990. Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, 45(1), pp.221-229.

Lanne, M., 2007. Forecasting realized exchange rate volatility by decomposition. *International Journal of Forecasting*, 23, pp.307-320.

Liesenfeld, R., 2001. A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics*, 104, pp.141-178.

Maheu, J. and McCurdy, T., 2002. Nonlinear features of FX realized volatility. *Review of Economics and Statistics*, 84(4), pp.668–681.

Maheu, J. and McCurdy, T., 2004. News arrival, jump dynamics, and volatility components for individual stock returns. *Journal of Finance*, 59, pp.755-793.

Müller, U.A. Dacorogna, M.M. Davé, R.D. Olsen, R.B. Puctet, O.V. and von Weizsäcker, J., 1997. Volatilities of different time resolutions - Analyzing the dynamics of market components. *Journal of Empirical Finance*, 4, pp.213–239.

Park, B.J., 2007. Trading volume, volatility, and GARCH effects in the Korean won-U.S. dollar exchange market: Some evidence from conditional quantile estimation. *Japanese Economic Review*, 58(3), pp.382-399.

Park, B.J., 2010. Surprising information, the MDH, and the relationship between volatility and trading volume. *Journal of Financial Markets*, 13, pp.344-366.

Park, B.J., 2011. “Asymmetric herding as a source of asymmetric return volatility. *Journal of Banking and Finance*, doi:10.1016/j.jbankfin.2011.02.025.

Sadorsky, P., 2005. Stochastic volatility forecasting and risk management. *Applied Financial Economics*, 15, pp.121-135.

Shephard, N., 1993. Fitting non-linear time series models, with application to stochastic variance models. *Journal of Applied Econometrics*, 8, pp.135–152.

Tauchen, G.E. and Pitts, M., 1983. The price variability-volume relationship on speculative markets. *Econometrica*, 5, pp.485-550.

Taylor, S.J., 1986. *Modelling Financial Time Series*, Chichester: John Wiley.

West, K. Edison, H. and Cho, D., 1993. A utility-based comparison of some models of exchange rate volatility. *Journal of International Economics*, 35, pp.23-45.

Yu, J. and Meyer, R., 2006. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometrics Reviews*, 25, pp.361-384.