Estimation of Stochastic Volatility and Option Prices

Suk Joon Byun, Jung-Soon Hyun, Woon Jun Sung*

April 6, 2020

ABSTRACT

This study suggests methods to estimate a stochastic volatility model using both daily high and low prices and close prices of underlying asset. To estimate parameters and initial volatility of volatility process, the likelihood-based inferences of Markov chain Monte Carlo (MCMC) are conducted. Simulation studies reveal that i) the model with high/low prices as well as close prices is superior to the traditional model using close prices only in both estimation and option prices, and ii) the leverage effect (the negative correlation effect of Black (1976)) is not be a crucial factor in pricing options when high and low prices are considered in estimation. This finding suggests that high and low prices substitute for the effect of correlation, known difficult to be estimated, in pricing options.

JEL Classification: G10; G17; C11; C13; C15

Keywords: Stochastic Volatility, Leverage Effect, High and Low prices, MCMC, Bayes rule

---

* Corresponding author: Sung: KAIST Business School, 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Korea (email: assmir@business.kaist.ac.kr)

Byun: KAIST Business School, 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Korea (email: sjbyun@business.kaist.ac.kr); Hyun: KAIST Business School, 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Korea (email: jshyun@ kaist.ac.kr);.

1. Introduction

Stochastic movement of volatility has been well documented market-wide, and measuring volatility is an important issue for valuation of derivatives. Pricing options, Heston (1993) suggests a continuous-time square-root stochastic volatility model, in which innovations to volatility are partially correlated with innovations to the price of stock. He demonstrates that out-of-the-money (OTM) put option prices depend on the negative correlation. However, correlation is hard to be estimated and its estimates widely vary time to time. In this paper, we suggest a couple of estimators mitigating the correlation effect in option pricing. The estimators use daily high and low prices of underlying asset (H/L prices hereafter) in addition to daily close prices. Daily H/L prices reflect intraday movement of underlying asset on estimation for initial volatility as well as parameters of volatility process.

Over-the-counter (OTC) derivatives have in general longer maturity than exchange traded options do. Pricing OTC options requires volatility term structure implied in plain vanilla options, but longer-term maturity options are not frequently traded in exchanges. Black and Scholes (1973) or Heston models are popular in option pricing since both models provides analytic solutions. Unlike Heston model, Jacquier et al. (2004) and Yu (2005) propose non-affine lognormal autoregressive conditional volatility models. Although their models are less popular than Heston model, they fit well movement of stocks empirically. Because of such reason, we choose a lognormal autoregressive volatility process of Yu (2005).

While most studies including Jacquier et al. (2004) and Yu (2005) use daily or weekly close price data for the estimation, Gallant et al. (1999) use close-to-close returns along with the range data. Range is defined by the difference between the daily H/L log prices. Alizadeh et al. (2002) also show that range-based volatility proxies are more efficient than volatilities estimated on the basis of absolute daily returns and robust to market microstructure noise. Recently, high frequency data are available, so that estimations can be easily performed on the continuous time basis. High frequency data, however, are expensive but daily open, close, high and low prices are cost-free and available with lengthy time-series.

Both bid-ask bounce and asynchronous trading problem are fairly mild as Brandt and Diebold (2006) point out. Horst et al. (2012) also estimates a stochastic volatility model using full opening, high, low and closing prices and demonstrate that the use of full information can improve estimation of volatility. Those studies focus on estimation, but we do option prices. In this paper, we examine whether range or H/L prices improve option prices and test which estimators are best.

All models using H/L prices assume, to the best of our knowledge, independent innovations between prices and volatility, i.e., zero correlation. On the other way, studies on stochastic volatility models with correlation do not consider H/L prices. It may be because there is no explicit joint likelihood of H/L prices and close prices when prices and volatilities are correlated. Nonetheless, a negative correlation of prices and volatilities is not only empirically observed in stock market but theoretically meaningful since the negative correlation is well known as the leverage effect as in Black (1976). The leverage effect refers to the increase of volatility when stock price decreases. This leverage effect is a stylized fact observed in stock market. It tends to be more profound in index rather than in individual stocks

In this context, this paper suggests two stochastic volatility models allowing both correlation and additional information on H/L prices. The first one uses daily close (close-to-close prices) returns and ranges, and the other adds to close returns high returns (open-to-high) and low returns (open-to-low) separately. Since range is identical to high return minus low return, range cannot detect asymmetric effect of rise or fall of prices. That is why we distinguish two models. Our models are closely related to those of Alizadeh et al. (2002), Brandt and Jones (2005) and Horst et al. (2012), but are extended to include the leverage effect.

Initial volatility needs to be estimated since it is unobservable. Following Eraker et al. (2003), we apply the likelihood-based Markov chain Monte Carlo (MCMC) filtering method to estimate both parameters and initial volatility[1]. Posterior distributions are calculated by decomposition of the entire likelihood into two parts applying the Bayes rule: the joint likelihood of close, high, and low returns and the joint likelihood of daily returns and volatilities. This is an approximation for the true likelihood.

Conducting simulation studies, we test performance of estimations and option values. There are several findings. First, the parameters estimated by our model are closer to the true values than those estimated by the baseline model that uses daily close returns only. All parameters of volatility process except correlation are estimated significantly across various scenarios. Second, as for estimated volatilities, the root mean squared errors (RMSEs) of the baseline model compared to the true volatilities increase 1.54 – 1.76 times of those for our model on average. When correlation is smaller than -0.4, the suggested model estimates volatilities significantly better than the previous models such as Brandt and Jones (2005) and Horst et al. (2012) which assume zero correlation. However, range and H/L price models show an almost identical estimation result to initial volatility. Thirds, the close return-based baseline model performs in option pricing much worse than range- or H/L-based model. This observation is more apparent in low persistency than in high persistency. Fourth, the negative correlation between return and volatility does not have so large impact on put option prices as Black (1976) claims. Adjusting the long-term mean of volatility, two models with/without correlation shows minor difference in RMSEs performance, even though correlation is significantly estimated as nonzero.

The rest of this paper is organized as follows. In Section 2, we introduce the stochastic volatility models and explain how the information on high/low prices is incorporated into the model. In Section 3 we present the estimation method. In Section 4, we demonstrate estimation result of models through simulation. In Section 5, we compare performance of models for option prices. Finally, Section 6

---

[1] There are alternative filtering methods such as particle filtering (Horst et al. (2015)) and Kalman filter. Kalman filter is more efficient in computation time than other two methods but assumes Gaussian state space for an unobservable variable. See Johannes et al. (2003) for details for MCMC filtering method

summarizes the contents of this paper and suggests directions for further research.

2. Stochastic Volatility Models

2.1. Baseline model

Among the various stochastic volatility models in continuous-time economy, affine or log-volatility models are widely used. While people prefer affine models to log-volatility models from the perspective of pricing options, log-volatility models are often used for examining statistical behavior of an underlying asset. We also choose a log-volatility process in a continuous time economy. Let S_t be an asset's price and σ_t volatility of log-return at time t. We assume that the log-returns of stock prices evolve below:

$$\begin{aligned} d\ln S_t &= \mu dt + \sigma_t dW_t^1 \\ d\ln \sigma_t &= \kappa(\theta - \ln \sigma_t)dt + \upsilon dW_t^2 \end{aligned} \tag{1}$$

where W_t^1 and W_t^2 are Wiener processes correlated with $< dW_t^1, dW_t^2 >= \rho dt$ for a constant ρ and μ is the instantaneous expected rate of log return. If $\rho = 0$, it coincides with the models proposed by Alizadeh et al. (2002) and Horst et al. (2012). The log-volatility process of eq. (1) is a latent variable and is assumed that it follows a mean-reverting Ornstein–Uhlenbeck process. The parameter κ is a mean-reverting factor, θ is a long-term mean of log-volatility, and υ is the volatility of stock returns' volatility.

For empirical purpose, following Yu (2005), we assume a basic stochastic volatility model discretized as follows: For $t = 1,2,\cdots,T$, trading days

$$\begin{aligned} y_t &= \sigma_{t-1}\varepsilon_t^y \\ \ln \sigma_t &= \alpha + \delta \ln \sigma_{t-1} + \upsilon \varepsilon_t^\sigma \end{aligned} \tag{2}$$

where y_t is a log return of the asset, and $\varepsilon_t^y, \varepsilon_t^\sigma$ have a bivariate standard normal distribution with the correlation ρ. Here y_t becomes a daily log return and α equals $\kappa\theta$, and δ corresponds to $1-\kappa$ which represents the autocorrelation of volatilities. If δ is positively large, autocorrelation is strong so that the persistence of volatility is high. We denote this discretized model of eq. (2) as SV. It is worthy to note that the return at t is conditioned on the volatility estimated at $t-1$. We ignore the drift term μ in eq. (2) because short-term data like daily prices are used, so that we focus on parameters of volatility and volatility itself. Horst et al. (2012) also support the driftless model by estimating the drift of weekly log price near 0.

2.2. The model using ranges and high/low prices

In order to incorporate intraday data into the model, the continuous model in eq. (1) cannot be discretized as eq. (2). Instead we keep a continuous framework as follows: with the same notation as in eq. (2) and for $t = 1,2,\cdots,T$, trading days,

$$\begin{aligned} y_s &= \sigma_t \varepsilon_s^y, \qquad \text{for } t-1 < s \leq t, \\ \ln \sigma_t &= \alpha + \delta \ln \sigma_{t-1} + \upsilon \varepsilon_t^\sigma. \end{aligned} \tag{3}$$

Note that unlike the basic model, returns continuously move with constant conditional volatility σ_t over intraday $(t-1, t]$ but volatility is approximately discretized. Since volatility is constant during a day, the correlation measuring leverage effect is not taken into account over intraday but updated daily. Two discretized models in eq. (2) and eq. (3) are slightly different in the conditional volatility affecting the return. Jacquire et al. (2004) and Brant and Jones (2005) adopt the same volatility as in eq. (3) but Yu (2005) argues that the contemporaneous dependence between the two disturbances of eq. (2) makes y_t a martingale difference whereas y_t in eq. (3) is not.[2]. Thus the model of eq. (3) is not consistent with the efficient market hypothesis. The reason that we mix discretization of volatility

[2] Instead of the log-volatility process, Yu (2005) assumes the log-variance process. In the process, the conditional volatility of a return is just a square root of updated variance, so the difference of the two processes is little.

with continuous movement of returns is to incorporate both information on extreme values and the leverage effect simultaneously. This will be explained in more detail in the next section.

3. Estimation Methodology

In this section we combine the model given by eq. (3) with intraday data through two different rules. In the first rule, we exploit range information with daily returns, precisely open-to-close returns. Since range is defined by the value of subtraction of the open-to-low log return from the open-to-high log return, it does not distinguish the rise and fall over a day as long as the high and low returns have same absolute value. The model of eq. (3) with returns and ranges is denoted by RR. This RR estimator is a correlated proxy corresponding to the model of Brandt and Jones (2005) that assumes zero correlation. The other estimator uses H/L returns separately. It is denoted by RHL. The RHL is a correlated counterpart corresponding to the CHLO model of Horst et al. (2012). They demonstrate information on levels of extreme values improves the estimation of volatility, although zero correlation is assumed. We test whether the same phenomena are observed in a correlated case. Using separate returns rather than ranges seems more useful but has both pros and cons. Ranges use only high and low prices that occur mainly during consecutive trades, so that these prices may be considered as the values from a theoretical continuous-time series. On the contrary, the RHL exploits both the H/L and open prices but open prices may be easily influenced by market microstructure due to trading mechanisms of stock markets.[3]

3.1 The RHL model

We first explain the RHL estimator since the density of range can be derived from the joint density of high and low returns. To exploit information on high and low prices during trading hours, we define

[3] Amihud and Mendelson (1987) show that open-to-open returns exhibit greater dispersion and non-normality by the trading mechanism. There are studies that opening or closing prices do not represent appropriate stock values but are affected by market structures (e.g., Harris (1989), Stoll and Whaley (1990)), and those prices may be noisy.

daily high and low returns as $H_t = \max_{t-1<s\le t} y_s$ and $L_t = \min_{t-1<s\le t} y_s$, respectively, i.e. daily open-to-high and open-to-low returns. The joint density of high and returns conditioned on return and the volatility can all be derived from the results of Feller (1951), Freedman (1971), and Klebaner (2005). Since the return y_t is a driftless Brownian Motion with a constant volatility σ_t over the interval $(t-1, t]$, the joint density of H_t and L_t is

$$\begin{aligned} & p(H_t \in db,\ L_t \in da|\ y_t = y, \sigma_t) \\ & = \frac{1}{\sigma_t^2}\frac{1}{\phi\left(\frac{y}{\sigma_t}\right)}\sum_{n=-\infty}^{\infty}\left[4n^2\left(\frac{(2n(b-a)-y)^2}{\sigma_t^2}-1\right)\phi\left(\frac{2n(b-a)-y}{\sigma_t}\right)\right. \\ & \qquad \left. -4n(n-1)\left(\frac{(2n(b-a)+y-2b)^2}{\sigma_t^2}-1\right)\phi\left(\frac{2n(b-a)+y-2b}{\sigma_t}\right)\right] \end{aligned} \tag{4}$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$, a standard normal density. The density of eq. (4) involves a calculation of an infinite sum but the sum quickly converges. (see Choi and Roh (2013)).

In order to consider the joint likelihood for extreme and daily returns, we decompose the likelihood into two terms using Bayes rule:

$$\begin{aligned} P(y, H, L|\Theta, \sigma) &= P(H, L|y, \Theta, \sigma)P(y|\Theta, \sigma) \\ &\approx P\big(H, L|y, \Theta, \sigma, (-\rho)\big)P(y|\Theta, \sigma) \end{aligned} \tag{5}$$

where y, H and L are vectors of time series data of open-to-close, open-to-high, and open-to-low returns. Θ represents a vector of model parameters $\{\alpha, \delta, \upsilon\}$ and σ is a vector of a time series of volatility. The first term, $P(H, L|y, \Theta, \sigma)$, is the joint density of high and low returns conditioned on returns, volatility, and parameters. Due to this term, information on the high and low prices can be incorporated when the model parameters are estimated. The second term, $P(y|\Theta, \sigma)$, is the likelihood of open-to-close returns given volatility and parameters. In eq. (3), the model assumes the daily correlation between the return process and the volatility process, so the leverage effect can be measured from the second term as in the case of general stochastic volatility models. The distribution, $P(y|\Theta,\sigma)$, can be calculated exactly from return data, but in this model the joint likelihood $P(H,L|y,\Theta,\sigma)$ is approximated by $P\big(H,L|y,\Theta,\sigma,(-\rho)\big)$ which denotes the likelihood for the case of zero correlation. This approximation is inevitable from the model since volatility is assumed constant over intraday and is updated at the end of a day. If innovations are uncorrelated, eq. (5) holds exactly.

Since we measure the leverage effect only from the $P(y|\Theta,\sigma)$ term, correlation may be underestimated. Nonetheless we may argue that the contemporaneous dependence between the return process and the volatility process of eq. (3) gauges the leverage effect appropriately. Suppose volatility increases from σ_{t-1} to σ_t, i.e., $\sigma_{t-1} < \sigma_t$. Then the return y_t is more likely to decrease but the increase in σ_t can cause a large gap between the maximum and the minimum returns, which means that returns and ranges are negatively correlated. For the case of the inter-temporal dependence as in eq. (2), the gap $H_t - L_t$ tends to have a small value because the likelihood of returns and extreme returns are affected by the lower value of σ_{t-1} instead of σ_t.[4]

3.2 The RR model

Instead of using the level of extreme values we also use ranges to exploit the symmetric information of high and low prices level. By a change of variables, the density of range conditioned on constant volatility and an absolute value of return can be derived from eq. (5) as follows:

$$p((H_t - L_t) \in dR\, |\sigma_t, |y_t| = |y|)$$
$$= \frac{1}{\sigma_t^2 \phi\left(\frac{|y|}{\sigma_t}\right)} \sum_{n=-\infty}^{\infty} \left[4n^2 \left(\frac{(2nR - |y|)^2}{\sigma_t^2} - 1\right) \phi\left(\frac{2nR - |y|}{\sigma_t}\right)(R - |y|)\right. \tag{6}$$

[4] We tested the inter-temporal dependence model, but leverage effect was considerably underestimated when high and low prices are used.

$$-2n(n-1)(2(n-1)R+|y|)\phi\left(\frac{2(n-1)R+|y|}{\sigma_t}\right)$$

$$\left.+2n(n-1)(2nR-|y|)\phi\left(\frac{2nR-|y|}{\sigma_t}\right)\right].$$

Again, Bayes rule yields the joint density of ranges and returns below:

$$\begin{aligned} P(y,R|\Theta,\sigma) &= P(R|y,\Theta,\sigma)P(y|\Theta,\sigma) \\ &\approx P\big(R|y,\Theta,\sigma,(-\rho)\big)P(y|\Theta,\sigma) \end{aligned} \tag{7}$$

where R is a vector of a time series of ranges and other notations are the same with those of eq. (6). From the first term, $P(R|y,\Theta,\sigma)$, we exploit the information on ranges by using the density in equation (6). Correlation is estimated through the second term, $P(y|\Theta,\sigma)$. As in the case of RHL estimator, the likelihood $P(R|y,\Theta,\sigma)$ is approximated by $P\big(R|y,\Theta,\sigma,(-\rho)\big)$ because piecewise constant volatility is assumed.

3.3 MCMC Method

The Markov chain Monte Carlo (MCMC) method is an exact likelihood-based inference and highly efficient, making it widely used to estimate model parameters and unobservable variables such as volatility (see Eraker et al. (2003), Eraker (2004)). In addition to estimating latent variables, MCMC provides parameters estimation risk and is a useful tool for estimation under complex distribution. Since our estimation involves an unobservable variable, volatility and its posterior distribution is not known, MCMC is an appropriate methodology. Moreover, Jacquier et al. (1994) insist MCMC outperforms GMM and QMLE in the estimation of stochastic volatility models.

The MCMC draws samples from each conditional posterior distribution which is factored into likelihood and a prior by Bayes rule. In the case of the stochastic volatility model without considering extreme values, the likelihood of $P(y|\Theta,\sigma)$ is well known, so that sampling is routine. This is no longer true for the RR and RHL models, specifically when return and volatility are correlated. Bayes rule makes the posteriors factored below:

$$RHL: P\left(\Theta_i|\Theta_{(-i)}, y, H, L, \sigma\right) \propto P(y, H, L|\Theta, \sigma)P(\sigma|\Theta)P(\Theta_i)$$
$$RR: P\left(\Theta_i|\Theta_{(-i)}, y, R, \sigma\right) \propto P(y, R|\Theta, \sigma)P(\sigma|\Theta)P(\Theta_i) \tag{8}$$

where Θ_i represents each parameter of the volatility process and $\Theta_{(-i)}$ indicates the parameter set except the parameter Θ_i. The likelihoods, $P(y, H, L|\Theta, \sigma)$ and $P(y, R|\Theta, \sigma)$, are calculated from the approximated likelihoods of eq. (5) and eq. (7).

The approximated likelihoods are advantageous in the aspect of estimation. Since the joint densities given in eq. (4) and eq. (6) are not dependent on any parameters except for volatility, Gibbs sampler can be used for sampling the parameters α, κ, and υ. That is, these parameters are sampled directly from known distributions whose conjugate priors are given by

$$(\alpha, \kappa) \sim BVN\left(0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \text{ and } \upsilon^2 \sim IG(2.5, 0.1)$$

where BVN and IG refer to a bivariate normal and an inverse gamma distribution, respectively. In order to sample the parameter ρ, we use an independent Metropolis-Hastings algorithm. The proposal density we use in the Metropolis-Hastings algorithm is $U(-1,1)$ where U represents a uniform distribution.

As for volatility estimation, it is well-known that the conditional posterior of volatility is not recognizable, so the Gibbs sampler is no longer applicable. In its place, a random walk Metropolis-Hastings algorithm is commonly used. We sample volatilities using the same algorithm for the RR and RHL estimators. The conditional posteriors are calculated as follows

$$
\begin{aligned}
RHL&: P(\sigma_t|\Theta,\sigma_{t-1},\sigma_{t+1},y_t,H_t,L_t,y_{t+1},H_{t+1},L_{t+1}) \\
&\propto P(y_{t+1},H_{t+1},L_{t+1}|\Theta,\sigma_t,\sigma_{t+1})P(y_t,H_t,L_t|\Theta,\sigma_t,\sigma_{t-1}) \\
&\times P(\sigma_{t+1}|\Theta,\sigma_t)P(\sigma_t|\Theta,\sigma_{t-1}) \\
RR&: P(\sigma_t|\Theta,\sigma_{t-1},\sigma_{t+1},y_t,R_t,y_{t+1},R_{t+1}) \\
&\propto P(y_{t+1},R_{t+1}|\Theta,\sigma_t,\sigma_{t+1})P(y_t,R_t|\Theta,\sigma_t,\sigma_{t-1}) \\
&\times P(\sigma_{t+1}|\Theta,\sigma_t)P(\sigma_t|\Theta,\sigma_{t-1}).
\end{aligned} \tag{9}
$$

In the MCMC algorithm, the sampling of parameters and volatility is iterated from the eq. (8) and eq. (9). Before the Markov chain converges, initial sampling values are discarded. We use the first 10,000 samplings as a burn-in period and 90,000 samplings for estimation after the burn-in period. Through trace plots, we check the convergence of the algorithm in the next simulation and empirical analysis sections.

4. Simulation Analysis

The RR and the RHL are the first estimators considering both the leverage effect and information on extreme prices, to the best of our knowledge. Through the simulation, we examine performance on two aspects. In one aspect, we see how close the estimated parameters are to the realized parameters. Another aspect is to contrast the RMSEs between volatilities estimated and volatilities calculated from simulation. We compare the RR and the HRL with the basic SV model and existing estimators, the $\mathrm{RR}(-\rho)$ and the $\mathrm{RHL}(-\rho)$ which are the counter parts with no correlation to the RR and the RHL, respectively. In other words, the $\mathrm{RR}(-\rho)$ and the $(\mathrm{RHL}(-\rho))$ are the same estimators as the RR (RHL) except imposition of zero correlation.

First, we test the case that each simulation path has 500 lengths which mean 500 trading days, about 2 years. Each path is generated from the SV model. Next we extend the total number of trading days to 1,000 days. For both cases, we generate 500 paths. Each trading day consists of 1,000 sub-periods to produce intraday returns whose maximum and minimum returns correspond to the high and the low returns, respectively. In order to compare against the results of Brandt and Jones (2005), we use the same parameter values and test three levels of persistency of volatility: low, medium, and high. The persistency is determined by the parameter, $\delta(= 1 - \kappa)$, which is the autocorrelation coefficient of volatility process. The higher the mean-reversion coefficient is, the lower autocorrelation is, which makes persistency low. Because the leverage effect is taken into account in this paper, we also test four cases of correlation, 0.0, -0.2, -0.4 and -0.6.

Table 1 provides interesting features of estimation results for total simulation length T=500. The mean of estimates is reported and the values in parentheses are the root mean square errors (RMSEs). Regardless of correlations and persistency, the parameters of the volatility process, α, δ, and υ, are better estimated in estimators relating intraday data than in SV model both in terms of the means and the RMSEs even though paths are generated by the SV model. Specifically, the parameters measuring persistency and variation of volatility are much more improved. This result is because additional information on high/low prices gets rid of noises embedded in returns and so lead to accurate inference of volatility dynamics. One special feature of Table 1 is that the SV model overestimates the volatility parameter of volatility process υ considerably and the extent of overestimation also severely grows as persistency increases. As Brandt and Jones (2005) point out, a highly persistent volatility process generates less volatile daily volatility and hinders the inference of υ. In other words, intraday data makes estimation more accurate than returns only do when market is calm. As for the estimate of α, it seems to be sensitive in all models. All estimates are statistically significant in all estimators but marginal in SV model. This happens since α heavily depends on the persistency of parameter δ. Nonetheless, all models have analogous long-term means of volatility when these values are calculated from the estimated parameters. The average level of volatility is well measured without high/low prices. This result seems natural since the intraday data movement has a small effect on the average value of volatilities. In the comparison between the RR and the RHL, it seems there is no difference as we find no evidence that there is an asymmetric impact of high/low price levels in the simulation. Horst et al. (2012) also document similar evidence.

Panel A to D show the results according to the correlation levels 0.0, -0.2, -0.4 and -0.6. We confirm that the absolute value of estimates of ρ increases in the SV, RR, RHL models as the absolute value of the true ρ increases. Panel A shows, however, that the estimated correlations are close to 0 but insignificant in all three models. While all three estimators are able to estimate various magnitudes of the leverage effect suitably, the estimates of ρ are slightly underestimated when compared to the true values across all the estimators. Although no distinction between the RR and the RHL appear, the mean values of estimates of ρ in the SV model are closer to the true values in the low and medium persistence cases. This may occur because the simulated paths are not continuous but discrete or due to the approximation of joint densities given by eq. (5) and eq. (7). For high persistency cases, regardless of how large the correlation is, the mean values of estimated ρ from the RR and the RHL are closer to the true values than those of the SV. In the high persistence case of the SV model, since the parameters defining the volatility dynamics are poorly estimated, the estimation of correlation is also affected. As for RMSEs of ρ, the RR and the RHL outperform the SV model in all cases.

Panel A to D also show the parameter estimates by the $\mathrm{RR}(-\rho)$ and the $\mathrm{RHL}(-\rho)$ i.e., no correlation assumed. In Panel B, since correlation is low, the parameter estimates by the $\mathrm{RR}(-\rho)$ and the $\mathrm{RHL}(-\rho)$ are very close to those estimated by the RR and the RHL in low and medium persistency. For high persistency, estimates of $\mathrm{RR}(-\rho)$ are not significant while estimates of $\mathrm{RHL}(-\rho)$ are significant. From Panel C and D, we find that the parameter estimates are closer to the true value in the RR/RHL than in the $\mathrm{RR}(-\rho)/\mathrm{RHL}(-\rho)$ and the RMSEs are smaller. However, the difference is not as much as we originally expected. This may be due to the limitations of the RR and the RHL estimator, i.e., volatility is constant over intraday but is daily correlated with returns.

(Table 1)

Table 2 present the results of estimations, where sample paths are extended to 1,000 days. The results show similar patterns of parameter estimation as seen in Table 1 but provide better estimates in terms of both the means and the RMSEs. However, the degree of improvement is more noticeable in the SV model than in other models. The same phenomenon is observed in Alizadel et al. (2002). This can be interpreted as the effect of information. While the RR and the RHL estimator exploit the additional information on high/low prices more effectively in a short period, it appears the improvements of estimates are not large enough given a longer period.[5] This is anticipatable in light of Parkinson (1980) which shows theoretically that in order to obtain same amount of standard deviation of continuous random walk using daily returns, the number of observations for returns is needed about 2.5 times more than that for ranges. In examination of correlation parameter, mean values of SV are a little bit closer to the true values except for the high persistency cases, but the RMSEs of the RR and the RHL models are mostly smaller than those of SV.

(Table 2)

In addition to inferring parameters, estimating volatility is also important in stochastic volatility models. Table 3 reports ratios of the RMSEs between models. Each simulated volatility series with 500 lengths is produced from the simulation in Table 1. Using the simulated volatility series and the estimated volatility series, root mean square error is calculated per path. The table shows the means and the values at the 5% and 95% percentiles (values in parenthesis). For example, "SV/RR" denotes value of the RMSE from SV divided by the RMSE from RR. Other symbols have similar meaning. The closer to 1 the ratio, the more similar the models are in sense of volatility estimates. We find the ratios SV/RR and SV/RHL moderately increase as persistency increases but the RR/RHL is almost identical. It is worthy to note RMSEs of SV are at least 1.5 times larger than those of RR/RHL, which implies that

[5] We also tested 2000 lengths of the sample period. There is not much difference with the results of T=1000. When the sample period increases, the benefit of using high/low prices still appears.

volatilities are estimated more accurately when intraday data are added rather than when only returns are used. Since the SV model shows poor results in the case of high persistency, this also affects volatility estimates. Within the same persistency level, the ratios remain stable across correlations. These finding suggests volatility estimates are not heavily dependent upon correlation but rather persistency which is dependent upon information about fluctuation embedded in maximum and minimum values. The $\mathrm{RR}(-\rho)/\mathrm{RR}$ reports the comparison of our RR estimator with the $\mathrm{RR}(-\rho)$, a proxy used by Brandt and Jones (2005). When correlation is low $\rho = -0.2$, the two estimators are statistically identical with 95% confidence levels. When the correlation is absolutely larger than -0.4, the RMSEs of RR are $4 - 10\%$ smaller on average than those of $\mathrm{RR}(-\rho)$ and the difference is statistically significant. As we can see, the similar results hold for RHL and $\mathrm{RHL}(-\rho)$.[6]

(Table 3)

5. Option prices

Given parameters and volatilities estimated in simulation, we evaluate European put options with two different maturities: one month and one year. We assume that the current asset price is 100, zero risk-free rate, and no risk premium for volatility. Since there is no closed-form solution for our models, we simulate 1,000 paths with 20-time steps for a 1-month option and 500 paths with 50-time steps for a 1-year option, respectively. Since RR and RHL show no difference in parameters as well as volatility, options are calculated in $\mathrm{RHL}(-\rho)$ and RHL when correlations are -0.4 and -0.6.

Table 4 reports RMSEs of prices for European put OTM options with strike price 90, ATM with 100, and ITM with 110 maturing in 1 month. Panel A presents the results when correlation is -0.4. ATM

[6] We perform the same simulation with 1,000 days. Unlike enhancement of parameter estimates for SV model shown in Table2, SV/RR and SV/RHL ratios do not improve much. Hence the effect of incorporating H/L prices appears more precisely in estimation of volatilities than of parameters.

options' error is the most in terms of the absolute error but ITM is the least in relative error since ITM put is the most expensive. As for comparison among models, SV performs poorly compared to the other two model in both absolute and relative error. RMSEs of the absolute error for RHLs decrease to around two thirds of those for SV across all scenarios regardless of strike prices. In relative error terms, RHLs dominate SV for OTM puts, which errors of RHLs are one half of those of SV. However, performance drastically diminishes for ATM puts and shows no difference for ITM puts. This is due to the level of option prices. ITM put prices are 10 times larger than OTM puts. This result implies that RHLs performs better than SV does for OTM puts. One interesting point of Table 4 is that option prices show minor difference between $\mathrm{RHL}(-\rho)$ and RHL. RHL improves $\mathrm{RHL}(-\rho)$ within 2% at most. We recall that estimates of correlation are significant in Table 1 although they are underestimated compared to the true value. Instead, $\mathrm{RHL}(-\rho)$ consistently estimates the higher long-term mean of volatility than RHL to adjust impact of correlation.

Panel B presents the results when correlation is -0.6. Both absolute and relative errors have similar patterns as in Panel A overall. Magnitudes of RMSEs are also close in both panels. These results clearly provide evidence that information of H/L prices helps estimating stochastic volatility and calculating option values as a result.

(Table 4)

Table 5 presents a parallel result of Table 4 when option's maturity is 1 year. Its prices are around 6 units higher on average than 1-month option's price due to the time value. Compared to Table 4, we find a couple of distinctions. First, performance of RHLs to SV is less dominant for the options with longer maturity. Prices of option with long maturity tend to be affected more by time value rather than volatility itself. Second, relative errors over OTM, ATM and ITM are flatter than those of short maturity options. In Table 4, relative errors between SV and RHLs for ITM options are negligible, 1% to 2%, but they are slightly higher 3% to 5% in Table 5. Finally, we do not find noticeable difference between $\mathrm{RHL}(-\rho)$ and RHL for puts with long maturity. The leverage effect seems to be weakening as option's maturity becomes longer.

This fact suggests that correlation of stochastic volatility model can be a minor factor in an option price whose maturity is long although it could be important in static analysis, all else being equal, as in Heston (1993). If we assume that correlation is zero, not only the exact likelihood for close and H/L returns is known, but the analytic formula for European options is also given. Zero correlation between returns and volatility makes empirical estimation simple and allows OTC products to be priced easily. The longer the maturity of OTC derivatives is, the more negligible the correlation effect is.

(Table 5)

6. Conclusions

Volatility changes over time. It is also well known that volatility rises when market returns fall. Many researches model this financial leverage effect as the negative correlation between market returns and changes of volatility. In this context, we have suggested stochastic volatility models that incorporate both the leverage effect and information on high /low prices.

To compare performance of models, we executed the MCMC simulation study. With changes in the correlation, the models using the additional information on ranges or high/low prices give better estimates of the true parameters than the baseline stochastic volatility model does. The same was also true when estimating volatility. As the correlation level increased, our proposed model performed better than the proxy models of Alizadeh et al. (2002) and Horst et al. (2012) and it was statistically significant. Given parameters and volatility estimated, we test performance of each model in option pricing. We confirmed that information on daily high/low prices helps put option prices improved regardless of moneyness. We did not find the notable effect of correlation in put option prices, which indicates that zero correlation between return and volatility may be acceptable in estimation of stochastic volatility model for the purpose of pricing option. Empirical analysis with option data leaves for future work.

REFERENCES

Alizadeh, Sassan, Michael W Brandt, and Francis X Diebold, 2002, Range-based estimation of stochastic volatility models, *The Journal of Finance* 57, 1047-1091.

Amihud, Yakov, and Haim Mendelson, 1987, Trading mechanisms and stock returns: An empirical investigation, *The Journal of Finance* 42, 533-553.

Black, F, 1976, Studies of stock price volatility changes, *In: Proceedings of the 1976 Meetings of the American Statistical Association* 171-181.

Brandt, Michael W, and Christopher S Jones, 2005, Bayesian range-based estimation of stochastic volatility models, *Finance Research Letters* 2, 201-209.

Choi, ByoungSeon, and JeongHo Roh, 2013, On the trivariate joint distribution of brownian motion and its maximum and minimum, *Statistics & Probability Letters 83, 1046-1053*.

Eraker, Bjørn, 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *The Journal of Finance* 59, 1367-1404.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in volatility and returns, *The Journal of Finance* 58, 1269-1300.

Feller, William, 1951, The asymptotic distribution of the range of sums of independent random variables, *The Annals of Mathematical Statistics* 427-432.

Gallant, A Ronald, Chien-Te Hsu, and George Tauchen, 1999, Using daily range data to calibrate volatility diffusions and extract the forward integrated variance, *Review of Economics and Statistics* 81, 617-631.

Harris, Lawrence, 1989, A day-end transaction price anomaly, *Journal of Financial and Quantitative Analysis* 24, 29-45.

Horst, Enrique Ter, Abel Rodriguez, Henryk Gzyl, and German Molina, 2012, Stochastic volatility models including open, close, high and low prices, *Quantitative Finance* 12, 199-212.

Jacquier, Eric, Nicholas G Polson, and Peter E Rossi, 1994, Bayesian analysis of stochastic volatility models, *Journal of Business and Economic Statistics* 12, 413-417.

Jacquier, Eric, Nicholas G Polson, and Peter E Rossi, 2004, Bayesian analysis of stochastic volatility models with fat-tails and correlated errors, *Journal of Econometrics* 122, 185-212.

Johannes, Michael, and Nick Polson, 2003, Mcmc methods for continuous-time financial econometrics, *Available at SSRN 480461*.

Stoll, Hans R, and Robert E Whaley, 1990, Stock market structure and volatility, *Review of Financial Studies* 3, 37-71.

Yu, Jun, 2005, On leverage in a stochastic volatility model, *Journal of Econometrics* 127, 165-178.

Table 1

Parameter Estimation Results of Simulation with T=500

This table shows the result of parameter estimation of simulated 500 paths. Each path is produced from a discretized stochastic volatility model:

$$y_t = \sigma_{t-1}\varepsilon_t^y$$
$$\ln\sigma_t = \alpha + \delta\ln\sigma_{t-1} + \upsilon\varepsilon_t^\sigma$$

where ε_t^y and ε_t^σ have a bivariate standard normal distribution with correlation ρ. The return and volatility series of each path has 500 lengths, so a trading day t has a value from 1 to 500. During each time step from $t-1$ to t, there are 1,000 sub-periods. We set three different levels of persistence: "Low", "Medium", and "High", by changing the values of volatility parameters. We vary the level of correlation: the case of no correlation (Panel A), -0.2 (Panel B), -0.4 (Panel C), and -0.6 (Panel D). Parameter estimation is conducted through three different models. "SV" denotes the basic stochastic volatility model. "RR" denotes the stochastic volatility model incorporating ranges. "RHL" denotes the stochastic volatility model incorporating high and low prices. These values are means of estimates and the values in parenthesis are root mean square errors. "RR$(-\rho)$" $\big(\text{RHL}(-\rho)\big)$ is a same estimator as RR (RHL) except imposition of zero correlation.

Panel A: Parameter Estimation with ρ=0

	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	0.000	-0.184	0.950	0.130	0.000	-0.074	0.980	0.083	0.000
SV	-0.521	0.859	0.202	-0.004	-0.362	0.902	0.174	-0.002	-0.278	0.927	0.147	-0.009
	(0.246)	(0.066)	(0.032)	(0.140)	(0.255)	(0.068)	(0.049)	(0.142)	(0.285)	(0.074)	(0.067)	(0.160)
RR (-ρ)	-0.382	0.897	0.173	Null	-0.236	0.937	0.138	Null	-0.150	0.960	0.106	Null
	(0.092)	(0.025)	(0.015)	Null	(0.090)	(0.024)	(0.013)	Null	(0.104)	(0.027)	(0.024)	Null
RHL (-ρ)	-0.383	0.897	0.174	Null	-0.236	0.937	0.138	Null	-0.150	0.960	0.107	Null
	(0.093)	(0.025)	(0.015)	Null	(0.091)	(0.024)	(0.013)	Null	(0.105)	(0.027)	(0.025)	Null
RR	-0.384	0.896	0.174	0.005	-0.237	0.936	0.139	0.001	-0.151	0.960	0.107	0.002
	(0.093)	(0.025)	(0.015)	(0.079)	(0.091)	(0.024)	(0.014)	(0.085)	(0.106)	(0.027)	(0.026)	(0.101)
RHL	-0.385	0.896	0.175	0.006	-0.238	0.936	0.139	0.002	-0.152	0.960	0.108	0.002
	(0.093)	(0.025)	(0.015)	(0.078)	(0.092)	(0.024)	(0.014)	(0.084)	(0.106)	(0.027)	(0.026)	(0.100)

Panel B: Parameter Estimation with ρ=-0.2

	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	-0.200	-0.184	0.950	0.130	-0.200	-0.074	0.980	0.083	-0.200
SV	-0.549	0.851	0.202	-0.165	-0.364	0.902	0.171	-0.145	-0.276	0.928	0.146	-0.125
	(0.274)	(0.074)	(0.033)	(0.138)	(0.242)	(0.064)	(0.045)	(0.158)	(0.273)	(0.071)	(0.065)	(0.179)
RR (-ρ)	-0.393	0.894	0.175	Null	-0.243	0.935	0.138	Null	-0.238	0.936	0.138	Null
	(0.092)	(0.024)	(0.015)	Null	(0.097)	(0.025)	(0.013)	Null	(0.092)	(0.024)	(0.013)	Null
RHL (-ρ)	-0.394	0.894	0.175	Null	-0.244	0.935	0.138	Null	-0.154	0.960	0.106	Null
	(0.093)	(0.024)	(0.015)	Null	(0.098)	(0.025)	(0.013)	Null	(0.105)	(0.027)	(0.024)	Null
RR	-0.390	0.895	0.175	-0.141	-0.242	0.935	0.139	-0.138	-0.152	0.960	0.106	-0.134
	(0.090)	(0.024)	(0.015)	(0.094)	(0.096)	(0.025)	(0.013)	(0.108)	(0.104)	(0.027)	(0.024)	(0.122)
RHL	-0.391	0.895	0.175	-0.140	-0.242	0.935	0.139	-0.137	-0.153	0.960	0.106	-0.133
	(0.091)	(0.024)	(0.015)	(0.094)	(0.097)	(0.025)	(0.013)	(0.108)	(0.105)	(0.027)	(0.024)	(0.123)

Panel C: Parameter Estimation with ρ=-0.4

	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	-0.400	-0.184	0.950	0.130	-0.400	-0.074	0.980	0.083	-0.400
SV	-0.526	0.857	0.201	-0.309	-0.363	0.902	0.172	-0.306	-0.283	0.925	0.147	-0.265
	(0.235)	(0.064)	(0.031)	(0.155)	(0.241)	(0.065)	(0.046)	(0.169)	(0.289)	(0.075)	(0.066)	(0.205)
RR (-ρ)	-0.386	0.896	0.175	Null	-0.238	0.936	0.138	Null	-0.144	0.962	0.105	Null
	(0.089)	(0.024)	(0.014)	Null	(0.092)	(0.024)	(0.013)	Null	(0.094)	(0.024)	(0.023)	Null
RHL (-ρ)	-0.386	0.896	0.176	Null	-0.238	0.936	0.138	Null	-0.153	0.960	0.106	Null
	(0.089)	(0.024)	(0.014)	Null	(0.092)	(0.024)	(0.013)	Null	(0.104)	(0.027)	(0.024)	Null
RR	-0.367	0.900	0.173	-0.286	-0.225	0.939	0.136	-0.294	-0.144	0.962	0.105	-0.286
	(0.079)	(0.021)	(0.015)	(0.135)	(0.078)	(0.021)	(0.012)	(0.136)	(0.094)	(0.024)	(0.023)	(0.152)
RHL	-0.368	0.900	0.173	-0.284	-0.226	0.939	0.136	-0.292	-0.144	0.962	0.106	-0.284
	(0.079)	(0.021)	(0.015)	(0.136)	(0.078)	(0.021)	(0.012)	(0.137)	(0.094)	(0.024)	(0.024)	(0.154)

Panel D: Parameter Estimation with ρ=-0.6												
	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	-0.600	-0.184	0.950	0.130	-0.600	-0.074	0.980	0.083	-0.600
SV	-0.509	0.862	0.197	-0.477	-0.358	0.904	0.169	-0.458	-0.269	0.929	0.145	-0.394
	(0.209)	(0.057)	(0.026)	(0.168)	(0.225)	(0.060)	(0.043)	(0.192)	(0.262)	(0.068)	(0.064)	(0.254)
RR (-ρ)	-0.382	0.897	0.174	Null	-0.245	0.934	0.138	Null	-0.135	0.964	0.104	Null
	(0.089)	(0.024)	(0.015)	Null	(0.097)	(0.025)	(0.013)	Null	(0.084)	(0.022)	(0.022)	Null
RHL (-ρ)	-0.383	0.897	0.175	Null	-0.245	0.934	0.138	Null	-0.156	0.959	0.106	Null
	(0.090)	(0.024)	(0.015)	Null	(0.097)	(0.025)	(0.013)	Null	(0.109)	(0.028)	(0.024)	Null
RR	-0.340	0.907	0.167	-0.454	-0.215	0.941	0.133	-0.458	-0.135	0.964	0.104	-0.446
	(0.074)	(0.020)	(0.018)	(0.162)	(0.066)	(0.018)	(0.009)	(0.162)	(0.084)	(0.022)	(0.022)	(0.180)
RHL	-0.340	0.907	0.167	-0.453	-0.215	0.941	0.133	-0.457	-0.135	0.964	0.104	-0.444
	(0.074)	(0.020)	(0.018)	(0.164)	(0.066)	(0.018)	(0.009)	(0.164)	(0.084)	(0.022)	(0.022)	(0.181)

Table 2

Parameter Estimation Results of Simulation with T=1,000

This table shows the result of parameter estimation of simulated 500 paths. Each path is produced from a discretized stochastic volatility model:

$$y_t = \sigma_{t-1}\varepsilon_t^y$$
$$\ln\sigma_t = \alpha + \delta\ln\sigma_{t-1} + \upsilon\varepsilon_t^\sigma$$

where ε_t^y and ε_t^σ have a bivariate standard normal distribution with correlation ρ. The return and volatility series of each path has 1,000 lengths, so a trading day t has a value from 1 to 1,000. During each time step from $t-1$ to t, there are 1,000 sub-periods. We set three different levels of persistence: "Low", "Medium", and "High", by changing the values of volatility parameters. We vary the level of correlation: the case of no correlation (Panel A), -0.2 (Panel B), -0.4 (Panel C), and -0.6 (Panel D). Parameter estimation is conducted through three different models. "SV" denotes the basic stochastic volatility model. "RR" denotes the stochastic volatility model incorporating ranges. "RHL" denotes the stochastic volatility model incorporating high and low prices. These values are means of estimates and the values in parenthesis are root mean square errors.

Panel A: Parameter Estimation with ρ=-0.2

	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	-0.200	-0.184	0.950	0.130	-0.200	-0.074	0.980	0.083	-0.200
SV	-0.442	0.880	0.188	-0.159	-0.274	0.926	0.153	-0.151	-0.171	0.954	0.122	-0.140
	(0.139)	(0.037)	(0.022)	(0.107)	(0.129)	(0.034)	(0.028)	(0.116)	(0.126)	(0.033)	(0.040)	(0.134)
RR	-0.364	0.902	0.172	-0.144	-0.211	0.943	0.133	-0.142	-0.114	0.970	0.097	-0.143
	(0.061)	(0.016)	(0.014)	(0.077)	(0.056)	(0.015)	(0.008)	(0.086)	(0.055)	(0.014)	(0.015)	(0.094)
RHL	-0.365	0.902	0.173	-0.143	-0.212	0.943	0.133	-0.141	-0.114	0.970	0.097	-0.142
	(0.061)	(0.016)	(0.014)	(0.077)	(0.056)	(0.015)	(0.008)	(0.086)	(0.056)	(0.014)	(0.015)	(0.094)

Panel B: Parameter Estimation with ρ=-0.4												
	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	-0.400	-0.184	0.950	0.130	-0.400	-0.074	0.980	0.083	-0.400
SV	-0.429	0.883	0.186	-0.330	-0.269	0.927	0.153	-0.327	-0.168	0.955	0.123	-0.307
	(0.120)	(0.033)	(0.019)	(0.116)	(0.115)	(0.031)	(0.028)	(0.123)	(0.114)	(0.030)	(0.041)	(0.149)
RR	-0.341	0.907	0.169	-0.297	-0.198	0.946	0.131	-0.305	-0.108	0.971	0.096	-0.309
	(0.062)	(0.017)	(0.016)	(0.116)	(0.043)	(0.011)	(0.008)	(0.111)	(0.047)	(0.012)	(0.014)	(0.114)
RHL	-0.342	0.907	0.169	-0.296	-0.199	0.946	0.131	-0.303	-0.108	0.971	0.096	-0.307
	(0.061)	(0.017)	(0.016)	(0.117)	(0.043)	(0.012)	(0.008)	(0.112)	(0.047)	(0.012)	(0.014)	(0.116)

Panel C: Parameter Estimation with ρ=-0.6												
	Low persistence				Medium persistence				High persistence			
	α	δ	ν	ρ	α	δ	ν	ρ	α	δ	ν	ρ
True	-0.368	0.900	0.182	-0.600	-0.184	0.950	0.130	-0.600	-0.074	0.980	0.083	-0.600
SV	-0.433	0.882	0.184	-0.502	-0.272	0.926	0.151	-0.493	-0.167	0.955	0.122	-0.454
	(0.111)	(0.030)	(0.018)	(0.131)	(0.111)	(0.030)	(0.025)	(0.141)	(0.111)	(0.029)	(0.040)	(0.179)
RR	-0.322	0.912	0.164	-0.465	-0.189	0.948	0.127	-0.473	-0.101	0.972	0.094	-0.478
	(0.065)	(0.017)	(0.020)	(0.144)	(0.036)	(0.010)	(0.007)	(0.139)	(0.040)	(0.011)	(0.012)	(0.139)
RHL	-0.322	0.911	0.164	-0.463	-0.189	0.948	0.127	-0.471	-0.102	0.972	0.094	-0.476
	(0.065)	(0.017)	(0.020)	(0.146)	(0.036)	(0.010)	(0.007)	(0.140)	(0.041)	(0.011)	(0.012)	(0.141)

Table 3

Ratio of Root Mean Square Error for Volatility with T=500

This table reports means of the ratio of root mean square errors for volatility. Each simulated volatility series with 500 lengths is produced from the simulation in Table 1. Throughout the parameter estimation in Table 1, latent volatilities are estimated from the three SV, RR, and RHL models. Using the simulated volatility series and the estimated volatility series, root mean square error is calculated per path. “SV/RR” denotes the value of the root mean square error from SV over that from RR. “SV/ RHL” denotes the value of the root mean square error from SV over that from RHL. “RR/ RHL” denotes the value of the root mean square error from RR over that from RHL. From the 500 simulated paths, this table shows means and 5% and 95% percentiles (values in parenthesis) of the ratios. Varying the levels of persistence and correlation, the ratio of root mean square error is calculated. “RR (-ρ) / RR” denotes the value of the root mean square error from RR over that from RR (-ρ). “RHL (-ρ) / RHL” denotes the value of the root mean square error from RR over that from RR (-ρ).

Panel A: Low Persistence

Model	ρ = 0	ρ = -0.2	ρ = -0.4	ρ = -0.6
SV / RR	1.55 (1.36, 1.74)	1.55 (1.37, 1.78)	1.55 (1.36, 1.78)	1.54 (1.33, 1.77)
SV / RHL	1.55 (1.36, 1.75)	1.55 (1.36, 1.78)	1.55 (1.36, 1.77)	1.53 (1.32, 1.76)
RR / RHL	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
RR (-ρ) / RR	1.00 (1.00, 1.01)	1.01 (1.00, 1.02)	1.04 (1.02, 1.06)	1.10 (1.07, 1.14)
RHL (-ρ) / RHL	1.00 (1.00, 1.01)	1.01 (1.00, 1.02)	1.04 (1.02, 1.06)	1.10 (1.07, 1.14)

Panel B: Medium Persistence

Model	ρ = 0	ρ = -0.2	ρ = -0.4	ρ = -0.6
SV / RR	1.61 (1.37, 1.92)	1.63 (1.38, 1.92)	1.62 (1.37, 1.91)	1.64 (1.38, 1.94)
SV / RHL	1.61 (1.37, 1.92)	1.63 (1.38, 1.92)	1.62 (1.37, 1.90)	1.64 (1.37, 1.95)
RR / RHL	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
RR (-ρ) / RR	1.00 (1.00, 1.01)	1.01 (1.00, 1.02)	1.04 (1.02, 1.06)	1.11 (1.07, 1.15)
RHL (-ρ) / RHL	1.00 (1.00, 1.01)	1.01 (1.00, 1.02)	1.04 (1.02, 1.06)	1.11 (1.07, 1.15)

Panel C: High Persistence

Model	ρ = 0	ρ = -0.2	ρ = -0.4	ρ = -0.6
SV / RR	1.69 (1.37, 2.09)	1.72 (1.39, 2.12)	1.72 (1.38, 2.09)	1.76 (1.4, 2.21)
SV / RHL	1.68 (1.36, 2.09)	1.71 (1.39, 2.12)	1.72 (1.38, 2.09)	1.76 (1.40, 2.20)
RR / RHL	1.00 (0.98, 1.01)	1.00 (0.99, 1.01)	1.00 (0.98, 1.01)	1.00 (0.98, 1.01)
RR (-ρ) / RR	1.00 (0.99, 1.02)	1.01 (0.99, 1.02)	1.04 (1.01, 1.07)	1.10 (1.06, 1.15)
RHL (-ρ) / RHL	1.00 (0.99, 1.02)	1.01 (0.99, 1.02)	1.04 (1.01, 1.07)	1.10 (1.06, 1.15)

Table 4

RMSEs of European put option prices when option's maturity is 1 month

This table reports root mean square errors for European put option prices with the parameters given Table 1. It is assumed that a risk-free rate is 0%, current price of stock is 100, and risk premium for variance is 0. Maturity of each option is 1 month and time steps are assumed by 20. Option prices are averaged over 1,000 simulations. K represents strike price. Values in parenthesis are 5 and 95 percentiles. "SV" denotes the basic stochastic volatility model. "RHL" denotes the stochastic volatility model incorporating high and low prices. "RHL($-\rho$)" is the same model as RHL except imposition of zero correlation.

Panel A: Correlation = -0.4

Absolute error									
	OTM put option with K=90			ATM put option with K=100			ITM put option with K=110		
Model	Low persistency	Medium	High	Low	Medium	High	Low	Medium	High
SV	0.76(0.03,1.58)	0.68(0.03,1.45)	0.55(0.02,1.17)	1.05(0.06,2.07)	0.91(0.05,1.83)	0.75(0.04,1.51)	0.75(0.03,1.58)	0.66(0.02,1.40)	0.53(0.01,1.14)
RHL (-ρ)	0.52(0.02,1.12)	0.45(0.02,0.96)	0.34(0.01,0.73)	0.70(0.04,1.42)	0.58(0.03,1.19)	0.45(0.02,0.92)	0.51(0.01,1.09)	0.42(0.01,0.91)	0.32(0.00,0.70)
RHL	0.50(0.02,1.07)	0.43(0.02,0.92)	0.32(0.01,0.70)	0.67(0.03,1.37)	0.56(0.03,1.15)	0.43(0.02,0.89)	0.49(0.01,1.05)	0.41(0.01,0.89)	0.31(0.00,0.68)
Relative error									
SV	1.28(0.04,2.55)	0.91(0.03,1.77)	0.74(0.03,1.45)	0.26(0.01,0.52)	0.22(0.01,0.43)	0.18(0.01,0.35)	0.06(0.00,0.13)	0.05(0.00,0.11)	0.04(0.00,0.09)
RHL (-ρ)	0.55(0.03,1.01)	0.42(0.02,0.75)	0.33(0.02,0.63)	0.15(0.01,0.29)	0.12(0.01,0.24)	0.10(0.01,0.19)	0.04(0.00,0.09)	0.03(0.00,0.07)	0.03(0.00,0.06)
RHL	0.53(0.01,1.07)	0.40(0.02,0.73)	0.32(0.02,0.61)	0.14(0.01,0.28)	0.12(0.01,0.23)	0.10(0.01,0.19)	0.04(0.00,0.08)	0.03(0.00,0.07)	0.03(0.00,0.06)

Panel B: Correlation = -0.6

Absolute error									
	OTM put option with K=90			ATM put option with K=100			ITM put option with K=110		
Model	Low persistency	Medium	High	Low	Medium	High	Low	Medium	High
SV	0.70(0.03,1.47)	0.63(0.03,1.34)	0.52(0.02,1.11)	0.97(0.06,1.92)	0.85(0.05,1.72)	0.71(0.04,1.44)	0.69(0.02,1.44)	0.61(0.01,1.30)	0.50(0.01,1.08)
RHL (-ρ)	0.52(0.02,1.11)	0.44(0.02,0.94)	0.33(0.01,0.72)	0.69(0.04,1.40)	0.57(0.03,1.17)	0.44(0.02,0.91)	0.49(0.01,1.06)	0.41(0.01,0.89)	0.31(0.01,0.68)
RHL	0.47(0.02,1.00)	0.39(0.01,0.84)	0.30(0.01,0.64)	0.67(0.03,1.27)	0.52(0.03,1.06)	0.40(0.02,0.83)	0.45(0.01,0.97)	0.37(0.01,0.81)	0.29(0.00,0.62)
Relative error									
SV	1.15(0.03,2.27)	0.85(0.03,1.66)	0.72(0.03,1.41)	0.24(0.01,0.48)	0.21(0.01,0.41)	0.17(0.01,0.34)	0.06(0.00,0.12)	0.05(0.00,0.11)	0.04(0.00,0.09)
RHL (-ρ)	0.56(0.03,1.00)	0.42(0.02,0.75)	0.34(0.02,0.63)	0.15(0.01,0.29)	0.13(0.01,0.24)	0.10(0.01,0.19)	0.04(0.00,0.09)	0.03(0.00,0.07)	0.03(0.00,0.06)
RHL	0.50(0.02,0.91)	0.37(0.02,0.68)	0.30(0.02,0.58)	0.14(0.01,0.26)	0.11(0.01,0.22)	0.09(0.01,0.18)	0.04(0.00,0.08)	0.03(0.00,0.07)	0.02(0.00,0.05)

Table 5

RMSEs of European put option prices when option's maturity is 1 month

This table reports root mean square errors for European put option prices with the parameters given Table 1. It is assumed that a risk-free rate is 0%, current price of stock is 100, and risk premium for variance is 0. Maturity of each option is 1 year and time steps are assumed by 50. Option prices are averaged over 500 simulations. K represents strike price. Values in parenthesis are 5 and 95 percentiles. "SV" denotes the basic stochastic volatility model. "RHL" denotes the stochastic volatility model incorporating high and low prices. "RHL($-\rho$)" is the same model as RHL except imposition of zero correlation.

Panel A: Correlation = -0.4

Absolute error									
	OTM put option with K=90			ATM put option with K=100			ITM put option with K=110		
Model	Low persistency	Medium	High	Low	Medium	High	Low	Medium	High
SV	2.07(0.12,4.14)	1.93(0.11,3.91)	1.69(0.09,3.48)	2.20(0.13,4.37)	2.04(0.12,4.11)	1.79(0.10,3.66)	2.07(0.12,4.13)	1.93(0.11,3.92)	1.70(0.09,3.50)
RHL (-ρ)	1.37(0.08,2.76)	1.23(0.07,2.51)	1.03(0.05,2.11)	1.43(0.08,2.87)	1.28(0.07,2.57)	1.07(0.06,2.17)	1.35(0.08,2.72)	1.19(0.07,2.42)	0.99(0.05,2.03)
RHL	1.30(0.07,2.64)	1.18(0.06,2.10)	1.00(0.05,2.06)	1.38(0.08,2.76)	1.24(0.07,2.51)	1.06(0.06,2.15)	1.30(0.07,2.62)	1.17(0.06,2.39)	1.00(0.05,2.05)
Relative error									
SV	0.39(0.02,0.72)	0.31(0.02,0.57)	0.26(0.01,0.47)	0.20(0.01,0.39)	0.17(0.01,0.33)	0.15(0.01,0.28)	0.12(0.01,0.23)	0.10(0.01,0.20)	0.09(0.01,0.18)
RHL (-ρ)	0.21(0.01,0.39)	0.17(0.01,0.33)	0.15(0.01,0.28)	0.13(0.01,0.24)	0.10(0.01,0.20)	0.09(0.01,0.17)	0.08(0.00,0.15)	0.06(0.00,0.13)	0.05(0.00,0.10)
RHL	0.20(0.01,0.38)	0.16(0.02,0.31)	0.14(0.01,0.27)	0.12(0.01,0.23)	0.10(0.01,0.20)	0.09(0.01,0.17)	0.07(0.00,0.14)	0.06(0.00,0.12)	0.05(0.00,0.10)

Panel B: Correlation = -0.6

Absolute error									
	OTM put option with K=90			ATM put option with K=100			ITM put option with K=110		
Model	Low persistency	Medium	High	Low	Medium	High	Low	Medium	High
SV	1.90(0.11,3.80)	1.80(0.10,3.64)	1.58(0.08,3.24)	2.03(0.12,4.02)	1.90(0.11,3.83)	1.68(0.09,3.41)	1.90(0.11,3.79)	0.85(0.03,1.81)	0.70(0.01,1.49)
RHL (-ρ)	1.35(0.08,2.71)	1.21(0.07,2.46)	1.03(0.05,2.09)	1.41(0.08,2.82)	1.25(0.07,2.51)	1.05(0.06,2.13)	1.34(0.08,2.68)	0.59(0.01,1.25)	0.44(0.01,0.94)
RHL	1.20(0.07,2.42)	1.08(0.06,2.19)	0.92(0.05,1.89)	1.27(0.07,2.54)	1.13(0.07,2.28)	0.98(0.05,1.98)	1.20(0.07,2.40)	0.53(0.01,1.13)	0.40(0.01,0.86)
Relative error									
SV	0.36(0.02,0.66)	0.29(0.01,0.53)	0.25(0.01,0.46)	0.19(0.01,0.37)	0.16(0.01,0.32)	0.14(0.01,0.28)	0.11(0.01,0.21)	0.10(0.01,0.19)	0.09(0.01,0.17)
RHL (-ρ)	0.21(0.01,0.40)	0.17(0.01,0.33)	0.15(0.01,0.29)	0.13(0.01,0.25)	0.11(0.01,0.20)	0.09(0.01,0.17)	0.08(0.00,0.15)	0.06(0.00,0.13)	0.05(0.00,0.10)
RHL	0.19(0.01,1.36)	0.15(0.01,0.29)	0.13(0.01,0.26)	0.11(0.01,0.22)	0.10(0.01,0.18)	0.08(0.01,0.16)	0.07(0.00,0.13)	0.06(0.00,0.13)	0.05(0.00,0.10)