Information Technology and Quantitative Management (ITQM 2016)

An Option Pricing Model using High Frequency Data

Saebom Jeon\textsuperscript{a}, Won Chang\textsuperscript{b}, Yousung Park\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a} Korea University, Anam 5-1, Seoul 136-701, Republic of Korea
\textsuperscript{b} University of Chicago, 5734 S. University Avenue, Chicago 60637, USA

Abstract

We propose a European call option evaluation framework accommodating GARCH-M model and its extension to handle irregularly spaced high-frequency data. The framework takes Bayesian approach to derive the predictive distribution for option prices and their volatilities. These predictive distributions vary as time approaches to the expiry data and provide credibility intervals to evaluate the option market. In empirical study, we illustrate the application using KOSPI200 and its options. Our approach results well-suited in the simulation based option pricing and explains a behavior of option prices close to the expiry.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of the Organizing Committee of ITQM 2016

Keywords: European Call Option; Volatility; Bayesian; GARCH-M model; MCMC algorithm.

1. Introduction

Today we’ve entered the era of ‘Big Data’, so-called the explosion of information, and this trend has deep impact on science, engineering and business. The flood of information enables decision makings to become more and more data-driven, and rigorous data-science to become more increasingly important. In business and finance, various types of Big Data are almost continuously generated daily or at a finer time stamped transaction-by-transaction or tick-by-tick, and corporations utilize such a Big Data to maximize profits by marketing services and to minimize losses by risk management.

High-frequency financial data, observed at a very short-term, have been widely used to study real-time market dynamics, volatility weighting, and strategic behaviour of market participants. However, utilizing high-frequency data also poses substantial modelling challenges, which can be viewed as an example of big data problem. In particular, since practical financial asset trading occurs irregularly in time and hence generates irregularly spaced time series, existing option pricing approaches that assume regularly spaced asset return series are not readily applicable. To mitigate this challenge, we adopt a framework for irregularly spaced time

* Corresponding author. Tel.: +82-2-3290-2238; fax: +82-2-924-9895.
E-mail address: yspark@korea.ac.kr.
series data observed at ultra-high frequency (UHF) referred by [1]. This approach allows us to better utilize information from tick-by-tick asset return series by eliminating the need to aggregate them into regularly-spaced time series.

Volatility is one of the most common risk measure using high-frequency time series. Despite of its imperfections, it is a key factor in option price evaluation and a main interest in risk management. It can be commonly either estimated through the past process of underlying asset price, or implied through a derivative pricing model. Although the implied volatility is popular and common in option pricing, it is derived by unrealistic normal assumption and constant volatility regardless of time. In this respect, historical volatility through generalized autoregressive conditional heteroscedasticity (GARCH) process has been proposed and extended by several researchers [2-8].

However, in real market, practical trading option price generally has shown a gap, sometimes a great difference, from the expectation, and the uncertainty should be considered in the option valuation. For these requirements, Bayesian approach is useful in that predictive distribution enables one not only to forecast option price but also to diagnosis the anomalies of market. A Bayesian GARCH process has been proposed by [9-10], but their suggestions give improper results when distribution of option price is nonsymmetric. Therefore, we propose Bayesian UHF GARCH in mean process, to handle the irregularly spaced high frequency data properly and to reflect the mean process of option pricing model regardless of the distribution assumption.

The remainder of the paper is organized as follows. The Section 2 proposes a Bayesian UHF GARCH-M option pricing model. The Section 3 illustrates empirical results of KOSPI200 call option. The last section concludes the paper with final remarks.

2. A Bayesian UHF GARCH-M Option Pricing Model

Geometric Brownian motion in discrete time assumes a distribution of asset return as \( y_{t_i} = \mu \Delta t_i + \varepsilon_{t_i} e^{t_i} \sim N(0, \Delta t_i \sigma^2) \), for \( i = 1, \ldots, T \) and \( \Delta t_i = t_i - t_{i-1} \). It can be extended as GARCH-M model with \( \sigma_{t_i} \) and \( r + \lambda \sigma_{t_i} \) instead of \( \sigma \) and \( \mu \), in order to consider risk premium (\( \lambda \)) and conditional heteroscedasticity (\( \sigma_{t_i} \)).

\[
y_{t_i} = r \Delta t_i + \lambda \Delta t_i \sigma_{t_i} + \varepsilon_{t_i} e^{t_i} \sim N(0, \Delta t_i) \quad \text{under P-measure},
\]

\[
\sigma_{t_i}^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \varepsilon_{t_{i-j}}^2 + \sum_{k=1}^{q} \delta_k \sigma_{t_{i-k}}^2 \quad \text{under P-measure},
\]

The likelihood function of \( y = [y_{t_1}, \ldots, y_{t_T}] \) becomes \( L(y_{t_1}, \ldots, y_{t_T}|\lambda, \alpha, \delta) = \prod_{t=1}^{T} \phi \left( \frac{y_{t_i}}{\sqrt{\sigma_{t_i}^2}} \right) \), where \( \phi(\cdot) \) is the p.d.f of \( N(0,1) \). To satisfy stationary conditions of GARCH model, we assume non-informative normal prior for \( \lambda \) and improper priors for \( \alpha, \delta \) using \( I(\alpha, \delta); \alpha_j > 0, \delta_k > 0 \) and \( \sum_{j=1}^{p} \alpha_j + \sum_{k=1}^{q} \delta_k < 1 \) for \( 1 \leq j \leq p, 1 \leq k \leq q \). Then, the joint prior becomes \( p(\lambda, \alpha, \delta) \propto N(\beta_j, \phi_{j}) \times p(\alpha, \delta) \times I(\alpha, \delta) \), where \( N(\beta_j, \phi_{j}) \) is normal distribution with mean \( \theta_j \) and variance \( \phi_{j} \), and the indication function \( I(\alpha, \delta) \) is 1 when stationary condition is satisfied, otherwise 0. These priors produce proper posterior distribution of \( \lambda, \alpha, \delta \) for given \( y \).

\[
f(\lambda, \alpha, \delta|y) \propto l(y|\lambda, \alpha, \delta) \times p(\lambda, \alpha, \delta),
\]

when \( \sigma_{t_i} < \infty \) [9]. It is not possible to sample the marginal distribution directly from the posterior distribution in a closed form, we implement Metropolis algorithm to draw random samples of parameters and make an inference, by proceeding between sampling from its conditional distribution given current values of all other values and updating value of each parameter.
To produce proper option price using the parameter estimates and samples of risk neutral return and volatility of UHF GARCH-M model, we extend the idea of Duan [4] in Q-measure. Through MCMC algorithm, we obtain $L$ sets of random samples of UHF GARCH-M parameter estimates $(\lambda^{(1)}, \alpha^{(1)}, \delta^{(1)}), \cdots, (\lambda^{(L)}, \alpha^{(L)}, \delta^{(L)})$ and for each $l_{th}$ sample $(\lambda^{(l)}, \alpha^{(l)}, \delta^{(l)})$, we generate $M$ sets of $(y^{(l,m)}_{ti}, \sigma^{2(l,m)}_{ti})$ at time $t_i$. For $m = 1, \cdots, M$ and $l = 1, \cdots, L$.

\[ y^{(l,m)}_{ti} = r \Delta t_i + \varepsilon^{(l,m)}_{ti}, \varepsilon^{(l,m)}_{ti} \sim N(0, \Delta t_i \sigma^{(l,m)}_{ti}) \text{ under Q-measure,} \]

\[ \sigma^{2(l,m)}_{ti} = \alpha^{(l)}_0 + \sum_{j=1}^{p} \alpha^{(l)}_j \varepsilon^{(l,m)}_{t_{i-j}} - \lambda^{(l)}(\Delta t_i \sigma^{(l,m)}_{t_{i-j}})^2 + \sum_{k=1}^{q} \delta^{(l)}_k \sigma^{2(l,m)}_{t_{i-k}} \text{ under Q-measure,} \]

After burning out MCMC samples, we can obtain random sample of $\hat{C}^{T^*(l)}_{t_i}$, via $p^{(l,m)}_{t_{T+t}}$, where $p^{(l,m)}_{t_{T+t}} = \max(S^{(l,m)}_{t_{T+t}} - K, 0), S^{(l,m)}_{t_{T+t}} = S_T \prod_{t_{i}=T+1}^{t_{T+t}} y^{(l,m)}_{i}$. One notable thing is that data generating process under risk neutral is GARCH without risk premium, while its original process is GARCH-M model. This process includes [4] as a special case when $\lambda \to 0$ and $\Delta t \to 1$.

Finally, we can produce the credibility interval and distribution of call option prices using $L$ random samples, and estimate the proper option price by posterior mean of $\hat{C}^{T^*(l)}_{t_i}$,

\[ C^{pos}_{t_i} = \frac{1}{L} \sum_{l=1}^{L} \hat{C}^{T^*(l)}_{t_i}, \]

where $\hat{C}^{T^*(l)}_{t_i} = e^{-r \Delta t \tau} \frac{1}{M} \sum_{m=1}^{M} p^{(l,m)}_{t_{T+t}}$.

3. Empirical Study

We illustrate our Bayesian UHF GARCH-M model to KOSPI200 and KOSPI200 call option whose strikes lie between 140 and 182.5 and option prices above 0.01, between 9:30 and 15:00 from March 5, 2009 to May 14, 2009. To implement MCMC algorithm, we set the values of tuning constants $c_\lambda, c_\alpha, c_\delta$, or acceptance ratios, to 0.5, and also specify the variance-covariance matrix of proposal distribution as $\sigma^2, \Sigma^2, \Sigma^2 [11]-[12]$. Through the MCMC algorithm with $L = 10000$ iterations and $M = 100$ forecasts, we estimate GARCH(1,1) model, proven to be superior to GARCH(1,2) and GARCH(2,2) by bayes factor.

A detailed feature of volatility can give implications for risk management, and we explore the distribution of the volatility from Bayesian UHF GARCH-M model. Figure 1 shows asymmetric distribution of volatility, consistent with earlier studies [13]. We can also observe that the distribution of volatility is more concentrated to the mean as time approaches to the expiry date, the less uncertainty in time to maturity.

![Fig. 1. Posterior Distribution of Volatility](image)
Table 1 also reports a similar finding to the Fig. 1 that the volatility decreases towards the maturity, regardless of volatility types, including typical historical volatility - a sample standard deviation of asset return series - and implied volatility by Black-Scholes model [14]. We also observe that implied volatility is always larger than historical and Bayesian volatilities, as many literatures investigated its upward biasedness [15].

Table 1. Volatility comparison for duration.

<table>
<thead>
<tr>
<th>Volatility</th>
<th>D-31</th>
<th>D-21</th>
<th>D-11</th>
<th>H-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical volatility</td>
<td>0.2814</td>
<td>0.2775</td>
<td>0.2774</td>
<td>0.2703</td>
</tr>
<tr>
<td>Bayesian UHF GARCH-M volatility</td>
<td>0.2819</td>
<td>0.2761</td>
<td>0.2796</td>
<td>0.2727</td>
</tr>
<tr>
<td>Implied volatility</td>
<td>0.3376</td>
<td>0.3288</td>
<td>0.3232</td>
<td>0.3244</td>
</tr>
</tbody>
</table>

Fig. 2 displays call option prices from the market (circle), estimated option prices from the model (dot) and 95% credibility intervals. We can observe that the standard deviation of the option prices decreases as time approaches to the expiry and the option prices from the market are mostly within the credibility intervals, implying options are fairly priced in the market.

4. Conclusion

This paper proposes Bayesian UHF GARCH-M model as a European option valuation method with high-frequency data. Such a Bayesian approach enables market participants to examine detailed volatility features
and evaluate the option market’s behavior, by taking into account the uncertainty in the predictive distribution of option price. The empirical result shows that the market participants trade options rationally within the credibility intervals. However, the estimated option price would not be always fairly priced in the market, in that the standard deviation of the option prices decreases and the credibility interval becomes narrower as the strike gets higher. Therefore, we would extend the current model setting and its application, to examine in several view by considering option evaluations according to the moneyness and time to maturity.

Acknowledgements

This work was supported by the Korea University Grant and National Research Foundation of Korea Grant funded by the Korean Government (NRF-2013S1A5B8A01054750).

References