Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Realized stochastic volatility with leverage and long memory

Shinichiro Shirota [a], Takayuki Hizu [b], Yasuhiro Omori [c,*]

[a] *Graduate school of Economics, The University of Tokyo, Japan*

[b] *Mitsubishi UFJ Trust and Banking, Japan*

[c] *Faculty of Economics, The University of Tokyo, Japan*

ARTICLE INFO

Article history:
Received 4 November 2012
Received in revised form 26 August 2013
Accepted 28 August 2013
Available online xxxx

Keywords:
ARFIMA
Leverage effect
Long memory
Markov Chain Monte Carlo
Mixture sampler
Realized volatility
Realized stochastic volatility model
State space model

ABSTRACT

The daily return and the realized volatility are simultaneously modeled in the stochastic volatility model with leverage and long memory. The dependent variable in the stochastic volatility model is the logarithm of the squared return, and its error distribution is approximated by a mixture of normals. In addition, the logarithm of the realized volatility is incorporated into the measurement equation, assuming that the latent log volatility follows an Autoregressive Fractionally Integrated Moving Average (ARFIMA) process to describe its long memory property. The efficient Bayesian estimation method using Markov chain Monte Carlo method (MCMC) was proposed and implemented in the state space representation. Model comparisons are performed based on the marginal likelihood, and the volatility forecasting performances are investigated using S&P500 stock index returns.

1. Introduction

The realized volatility is defined as the sum of the squared intraday returns over a specified time interval such as a day (e.g., Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2001)). This measure would provide a consistent estimator of the latent volatility under the ideal market assumption. The theory of the realized volatility is discussed in Barndorff-Nielsen and Shephard (2002) and Meddahi (2002), and there have been extensive studies on its time series structure and performance in volatility prediction (e.g., Andersen et al. (2003, 2007, 2004), Koopman et al. (2005) and Maheu and McCurdy (2007)).

In the real market, however, two major problems arise in measuring the daily realized volatility using high frequency return data: (1) the presence of non-trading hours and (2) market microstructure noise in transaction prices. The first problem arises because the stock market is usually open for only part of the day. For example, the Tokyo Stock Exchange (TSE) is open for 4.5 h a day and there is a lunch break. If we calculate the realized volatility as the sum of the squared intraday returns when the market is open, we may underestimate the latent one-day volatility. To avoid this underestimation, Hansen and Lunde (2005) proposed a scale realized volatility that adjusts the realized volatility by the ratio of the variance of the daily return to the mean of the realized volatility.

Market microstructure noise has various causes, including bid-ask spread and variation in trade sizes (see O'Hara (1995) and Hasbrouck (2007) for details) and can cause the realized volatility to be a biased estimator of the latent volatility. As the

* Corresponding author. Tel.: +81 3 5841 5516; fax: +81 3 5841 5521.
E-mail addresses: 6614272319@mail.ecc.u-tokyo.ac.jp (S. Shirota), omori@ja2.so-net.ne.jp, omori@e.u-tokyo.ac.jp (Y. Omori).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.08.013

sample time interval approaches zero, the bias owing to microstructure noise is expected to increase significantly. At-Sahalia et al. (2005) and Bandi and Russell (2008) propose a procedure to determine the optimal sampling interval, and Zhang et al. (2005) propose a bias adjusting method by assigning different weights to the realized volatilities calculated using different time intervals. In addition, Barndorff-Nielsen et al. (2008) derive the Realized Kernel (RK) as a consistent estimator of the latent volatility using high frequency data with noise.

Whereas, the intraday returns are heavily contaminated by microstructure noise, the daily returns are less subject to the noise. The daily returns could, therefore, provide additional information to eliminate the bias owing to microstructure noise and non-trading hours simultaneously. Takahashi et al. (2009) propose an extension of the stochastic volatility (SV) model to include such simultaneous modeling of the daily returns and realized volatility known as the Realized Stochastic Volatility (RSV) model. Hansen et al. (2012) implement a similar simultaneous modeling approach within the GARCH framework, called the Realized GARCH model, and demonstrate the superior, performance of the proposed model compared to GARCH (using daily returns only). Maheu and McCurdy (2011) consider the simultaneous modeling of S&P500 and IBM data and show that this approach outperforms the conventional EGARCH model.

Two important properties of the stochastic volatility and realized volatility have been discussed in previous empirical studies: (i) the leverage effect and (ii) long memory. The leverage effect refers to the correlation between the return at time t and the logarithm of the volatility at time $t+1$ and has been well established in empirical studies of stock returns (see, e.g., the survey by Shephard (2005)). To account for leverage effects, Melino and Turnbull (1990), for example, use the GMM (generalized methods of moments), and Harvey and Shephard (1996) use the QML (quasi-maximum likelihood method) with the Kalman filter for their estimation. Bayesian estimations have been described in various studies (e.g., Jacquier et al. (2004), Omori et al. (2007) and Omori and Watanabe (2008)). Takahashi et al. (2009) further propose a Bayesian estimation method for the RSV model with leverage where they use a single realized measure, while multiple realized measures are used in Venter and de Jongh (forthcoming) and Koopman and Scharth (2013). Superposition model, in which the logarithm of the volatility is a sum of latent factor processes, is proposed to describe the long-range dependence of the volatility in Dobrev and Szerszen (2010) with jumps in latent processes, and in Koopman and Scharth (2013) with a correlation between returns and measurement errors.

The long memory property of the realized volatility has also been investigated in many empirical studies using the high frequency data (e.g., Andersen et al. (2001)) and Raggi and Bordignon (2012) modeled the realized volatility with long memory and Markov switching dynamics using a Bayesian estimation method for the state space model. The SV model with long memory is discussed in Breidt et al. (1998) using the frequency domain approach (spectral likelihood estimator) and in So (2002) using a Bayesian approach with the state space model (So, 1999). Ruiz and Veiga (2008) investigate the statistical property of the stochastic volatility model with leverage and long memory (but without using the realized volatility), and compare with those of FIEGARCH models. Further, the autocorrelation function of powered absolute returns and their cross-correlations with original returns are derived in Pérez et al. (2009).

This paper extends the RSV model by incorporating both the leverage effects in the SV model and the long memory property of the realized volatility, and proposes a highly efficient Bayesian estimation method with a Markov Chain Monte Carlo (MCMC) implementation. Instead of the block sampler used in Takahashi et al. (2009), we employ the mixture sampler, a highly efficient Bayesian estimation method proposed by Kim et al. (1998) and Omori et al. (2007). In these methods, we take the logarithm of the squared asset return as a dependent variable to obtain linear measurement equations and approximate the error distribution by a mixture of normal distributions. In addition to the transformed stochastic volatility model, we assume an Autoregressive Fractionally Integrated Moving Average (ARFIMA) process for the logarithm of the log volatility to describe the long memory property of the realized volatility.

The paper is organized as follows. In Section 2, we introduce our model and its motivation. Section 3 describes the Bayesian estimation procedure based on the state space representation and Markov chain Monte Carlo methods. We illustrate our proposed method through numerical examples using simulated data in Section 4. In Section 5, we present our empirical studies using S&P500 realized volatility and realized kernels, perform model comparisons based on the marginal likelihood, and investigate the volatility forecast performances. We conclude in Section 6.

2. Realized stochastic volatility with leverage and long memory

2.1. Realized stochastic volatility with leverage

The simple stochastic volatility model with leverage is given by

$$y_{1t} = \exp(h_t/2)\epsilon_t, \quad t = 1, 2, \ldots, n, \tag{1}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \ldots, n, \tag{2}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \text{ i.i.d. } N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{pmatrix} \right), \qquad |\phi| < 1, \tag{3}$$

$$h_1 \sim (\mu, \sigma^2/(1-\phi^2)), \tag{4}$$

Please cite this article in press as: Shirota, S., et al., Realized stochastic volatility with leverage and long memory. Computational Statistics and Data Analysis (2013), http://dx.doi.org/10.1016/j.csda.2013.08.013