BAYESIAN ANALYSIS OF GENERAL ASYMMETRIC MULTIVARIATE GARCH MODELS AND NEWS IMPACT CURVES

Manabu Asai*

The BEKK model is a popular multivariate GARCH processes. The paper develops a new general asymmetric BEKK structure, which is based on recent empirical findings by semi-parametric news impact curves. For estimating the new model, a Markov chain Monte Carlo technique is used. Empirical results for trivariate asset returns from firms in the US indicate that the deviance information criterion favors the new model with a multivariate $t$ distribution, and that co-leverage effects exist among the three assets.

Key words and phrases: Asymmetry, Bayesian Markov chain Monte Carlo method, BEKK, leverage effect, multivariate GARCH.

1. Introduction

Multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models are frequently used in the analysis of dynamic covariance structures for multiple asset returns of financial time series. See survey papers such as Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009). One popular multivariate GARCH model is the BEKK model, which is named after Baba et al. (1987). See Engle and Kroner (1995), which is the final version of Baba et al. (1987). The BEKK model has the positive definite covariance process, and it is easy to verify its stationary conditions. To reduce the number of parameters, ‘diagonal BEKK’ and ‘scalar BEKK’ models are often used in empirical analysis.

For accommodating the leverage effects in a multivariate framework, Kroner and Ng (1998) developed the asymmetric BEKK (A-BEKK) model by extending the univariate asymmetric model of Glosten, Jagannathan, and Runkle (GJR) (1992), while Sentana (1995) suggested the multivariate quadratic GARCH model. However, recent empirical results from the semi-parametric news impact curve of Chen and Ghysels (2010) indicate that the shape of the asymmetric curve is more flexible than the GJR model and the quadratic GARCH model. Thus, there is room for improving the asymmetric specification.

Turning to estimation of the multivariate GARCH models, Vrontos et al. (2003) estimated several bivariate ARCH and GARCH models, and found that maximum likelihood estimates of the parameters were different from their Bayesian MCMC estimates. As the difference can be caused by the non-normality of the parameters, Vrontos et al. (2003) suggest careful interpretation of the ML

Received November 7, 2015. Revised January 9, 2016. Accepted January 10, 2016.

*Faculty of Economics, Soka University, 1-236 Tangi-machi, Hachioji, Tokyo 192-8577, Japan. Email: m-asai@soka.ac.jp
estimation. For this reason, the Bayesian Markov chain Monte Carlo (MCMC) technique is considered in the paper, so that non-normal posterior distributions are obtained.

In the univariate class, Geweke (1989) suggested the importance sampling technique for estimating ARCH models, while Bauwens and Lubrano (1998), Kim et al. (1998), Nakatsuma (2000), Vrontos et al. (2000), and Mitsui and Watanabe (2003) proposed several MCMC methods. Asai (2006) compared the above MCMC methods, and found that the best method is the so-called ‘tailored’ approach based on the acceptance-rejection Metropolis-Hastings (ARMH) algorithm, with respect to the mixing, efficiency and computational requirement. For the multivariate GARCH models, Vrontos et al. (2003) used the Metropolis-Hastings algorithm of Vrontos et al. (2000), while Ishihara et al. (2015) worked with the tailored ARMH algorithm. As discussed in Virbickaite et al. (2015), the recommended approach is the tailored ARMH algorithm as in the univariate case.

The first purpose of the paper is to suggest a new general asymmetric BEKK model by approximating the result of Chen and Ghysels (2010) in the multivariate framework, and to give the news impact curves. The new model includes the A-BEKK model of Kroner and Ng (1998) and the multivariate Quadratic GARCH model of Sentana (1995), as special cases. The second purpose of the paper is to explain the tailored ARMH algorithm for multivariate GARCH models, used in Ishihara et al. (2015), in detail. Note that Ishihara et al. (2015) estimated the A-BEKK model with the multivariate normal conditional distribution, while the current paper considers the new general asymmetric BEKK model with the multivariate standardized Student $t$ conditional distribution.

The remainder of the paper is organized as follows. Section 2 introduces the new general asymmetric BEKK model and its news impact curve. Section 3 explains the tailored ARMH algorithm for the Bayesian estimation in detail. Section 4 provides an empirical example for three stocks traded on the New York Stock Exchange. Section 4 compares six kinds of symmetric and asymmetric models with normal and non-normal conditional distributions, using the deviance information criterion of Spiegelhalter et al. (2002). Section 4 also examines the MCMC estimates of the parameters, and presents the news impact curves for describing the leverage and co-leverage effects. Finally, Section 5 gives some concluding remarks.

2. General asymmetric multivariate GARCH models

In this paper general asymmetric effects in a multivariate case are considered. Using a semi-parametric method with univariate realized volatility data, Chen and Ghysels (2010) observed that (1) negative shocks and large positive shocks increase future volatility via the current return; (2) the impact of a negative return is larger than that of a positive return; (3) small positive shocks decrease volatility. Such a phenomenon can be approximated by using the following lever-
age function

$$\lambda(x) = a(x - \gamma)^2 + cx^21(x < 0)$$

where $a > 0$, $c > 0$, $\gamma > 0$, and $1(x < 0)$ is the indicator function, which takes one if $x < 0$ and zero otherwise. Imposing the restriction $c = 0$ gives the leverage function of the quadratic GARCH model of Sentana (1995), while the leverage function reduces to that of the GJR model of Glosten et al. (1992) if $\gamma = 0$. This idea is used to consider multivariate leverage effects.

Let $y_t$ be an $m \times 1$ vector of financial asset returns. The new general asymmetric BEKK model with the $t$-distribution (GA-BEKK-$t$) is defined by

(2.1) $y_t = \mu_t + \varepsilon_t$,  $\varepsilon_t = H_t^{1/2}\xi_t$,

(2.2) $\xi_t = \zeta_t \sqrt{\frac{k - 2}{\kappa_t}}$,  $\zeta_t \sim N(0, I_m)$,  $\kappa_t \sim \chi^2(k)$,

(2.3) $H_t = W + A(\varepsilon_{t-1} - \gamma)(\varepsilon_{t-1} - \gamma)'A' + C\eta_{t-1}\eta_{t-1}'C' + BH_{t-1}B'$,

where $A$, $B$ and $C$ are $m$-dimensional square matrices, $\gamma$ is an $m$-vector of parameters, $W$ is an $m$-dimensional positive definite matrix, and $k$ is the parameter of the degree-of-freedom with $k > 2$. Also assume that $\zeta_t$ and $\kappa_t$ are mutually independent, $\eta_t = (\eta_{1t}, \ldots, \eta_{mt})'$ and $\eta_{it} = \varepsilon_{it}1(\varepsilon_{it} < 0)$. By the definition, $\xi_t$ follows the multivariate standardized $t$ distribution with the mean $0$ and covariance matrix $I_m$. Note that the GA-BEKK-$t$ model has the quadratic term, $A(\varepsilon_{t-1} - \gamma)(\varepsilon_{t-1} - \gamma)'A'$, and the threshold effect, $C\eta_{t-1}\eta_{t-1}'C'$. For the purpose of identification, the restrictions that $a_{11} \geq 0$, $b_{11} \geq 0$ and $c_{11} \geq 0$ are imposed. The model may be extended for the higher order specification as in Engle and Kroner (1995). It is assumed $\mu_t = 0$, but the assumption can be relaxed straightforwardly.

As argued in Laurent et al. (2012), the ‘variance targeting’ estimation is useful for modeling and forecasting conditional covariance matrices. Noting that $E[(\varepsilon_t - \gamma)(\varepsilon_t - \gamma)'] = E[\varepsilon_t\varepsilon_t'] + \gamma\gamma'$, it can be replaced by

(2.4) $W = \Omega - A(\Omega + \gamma\gamma')A' - B\Omega B' - C\text{NC}'$,

where $\Omega = E(H_t) = E(\varepsilon_t\varepsilon_t')$ and $N = E(\eta_t\eta_t')$. Again, assume that the right hand side of equation (2.4) is positive definite.

The new GA-BEKK-$t$ model includes several special cases: (i) When $\gamma = 0$, $C = O$ and $k \to \infty$, the model reduces to the BEKK-n model of Engle and Kroner (1995); (ii) When $\gamma = 0$ and $k \to \infty$, the resulting model is the asymmetric BEKK-n (A-BEKK-n) model of Kroner and Ng (1998), which is a multivariate extension of the GJR model of Glosten et al. (1992); (iii) The restrictions $C = O$ and $k \to \infty$ give the quadratic BEKK-n (Q-BEKK-n) model, which is an extension of the quadratic GARCH model of Sentana (1995); (iv) When $\gamma = 0$ and $C = O$, the model reduces to the BEKK-$t$ model; (v) When $\gamma = 0$, the A-BEKK-$t$ model is obtained; (vi) The restriction $C = O$ gives the Q-BEKK-$t$ model.
Engle and Kroner (1995) derived the stationary condition for the BEKK model. Applying the works of Sentana (1995), Ling and McAleer (2002), and Caporin and McAleer (2011) to the BEKK models, we obtain the sufficient condition for the covariance stationary for the GA-BEKK model which has a symmetric conditional distribution with finite fourth moments, such that all eigenvalues of

\[(A \otimes A) + (B \otimes B) + \frac{1}{2}(C \otimes C)\]

lie inside the unit circle.

The diagonal GARCH model of Ding and Engle (2001) can be extended by using the above general asymmetric effects. Based on the work of Cappiello et al. (2006), the general asymmetric diagonal GARCH (GAD-GARCH) model can be defined by equation (2.1) and

\[H_t = \Omega + A^* \circ \{ (\varepsilon_{t-1} - \gamma)(\varepsilon_{t-1} - \gamma)' - \Omega - \gamma \gamma' \} \]
\[+ C^* \circ (\eta_{t-1} \eta_{t-1}' - N) + B^* \circ (H_{t-1} - \Omega), \]

where \(A^*, B^*\) and \(C^*\) are positive (semi-)definite matrices.

For the purpose of reducing the number of parameters of the GA-BEKK specification, it is possible to consider the diagonal BEKK model which imposes restrictions such that \(A = \text{diag}(a), B = \text{diag}(b)\) and \(C = \text{diag}(c)\) in equation (2.3), where ‘\(\circ\)’ denotes the Hadamard (element-by-element) product, and \(a, b\) and \(c\) are \(m\)-vectors of parameters. By construction, the diagonal BEKK model is equivalent to assuming \(A^* = aa', B^* = bb'\) and \(C^* = cc'\) in equation (2.5).

Let \(\psi = (\theta, \Omega, N, k), \theta = (\text{vec}(A)', \text{vec}(B)', \text{vec}(C)', \gamma')'\) and \(Y_T = (y_1, \ldots, y_T)\). Then the joint probability density function of \(Y_T\) given \(\psi\) for the GA-BEKK-t model (2.1)–(2.4) is given by

\[f(Y_T | \psi) = \prod_{t=1}^{T} \frac{1}{[\pi(k-2)]^{-m/2}} \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} |H_t|^{-1/2} \]
\[\times \left[1 + \frac{1}{k-2} \varepsilon_t' H_t^{-1} \varepsilon_t\right]^{-(k+m)/2}, \]

where it is assumed that the initial conditional covariance matrix is equal to the unconditional one, i.e., \(H_1 = \Omega\).

Turning to the frequentist’s approach, Comte and Lieberman (2003) showed the consistency and asymptotic normality of the quasi-maximum likelihood estimator of the BEKK model under the existence of moments of order 8. However, there is no asymptotic theory for the variants of the BEKK model, indicating that the Bayesian approach is useful for the new GA-BEKK model at this moment. As a bridge to the gaps between the frequentist and Bayesian approaches, we should refer to the work of McAleer et al. (2008). Using an extension of
the univariate random coefficient autoregressive (RCA) process of Tsay (1987), McAleer et al. (2008) derived the BEKK model from a vector RCA process. Although McAleer et al. (2008) consider a general dynamic correlation model and derived the asymptotic theory, it is worth examining the appropriateness of the random coefficient for the new GA-BEKK model. We need to wait for further research.

In our empirical analysis, the GA-BEKK-\(t\) model is compared with its special cases including the BEKK-n, A-BEKK-n, GA-BEKK-n, BEKK-\(t\) and A-BEKK-\(t\) models.

3. Bayesian estimation

This section explains the Bayesian Markov chain Monte Carlo technique for estimating the new GA-BEKK-\(t\) models.

3.1. Prior distributions

For the prior distributions of \((\Omega, N)\), assume

\[
\Omega^{-1} \sim W(\nu_{\Omega,0}, Q_{\Omega,0}), \\
N^{-1} \sim W(\nu_{N,0}, Q_{N,0}),
\]

where \(W(p, S)\) denotes the Wishart distribution with the scale matrix \(S\) and the degrees-of-freedom parameter \(p\). The hyperparameters are specified as \(\nu_{\Omega,0} = m + 5\), \(Q_{\Omega,0} = \nu_{\Omega,0}I_m\), \(\nu_{N,0} = m + 5\), and \(Q_{N,0} = 0.5\nu_{N,0}I_m\). Regarding the parameter vector, \(\theta\), also assume the multivariate truncated normal distribution for the prior distribution as

\[
\theta \sim N(\delta_0, \Sigma_0) \times 1(\theta \text{ satisfies the stationary condition}),
\]

where \(\delta_0 = (\text{vec}(\sqrt{0.05}I_m)', \text{vec}(\sqrt{0.9}I_m)', \text{vec}(O)', \theta')\) and \(\Sigma_0 = 5I_m\) is specified, corresponding to the definition of \(\theta\). The truncated exponential distribution is used for the prior of \(k\),

\[
p(k) \propto \exp(-\lambda_0 k), \quad k > 4,
\]

with the hyperparameter, \(\lambda_0 = 5\).

3.2. ARMH algorithm

Here the acceptance-rejection Metropolis-Hastings (ARMH) algorithm is reviewed shortly. Consider sampling from the target density \(\pi(\psi \mid Y_T) \propto f(Y_T \mid \psi) \times p(\psi)\). Let \(q(\psi \mid Y_T)\) denote a candidate generating density. The simple acceptance-rejection method requires that there exists a constant \(c\) such that the condition

\[
D = \{\psi : f(Y_T \mid \psi)p(\psi) \leq cq(\psi \mid Y_T)\}
\]

holds for all \(\psi\) in the support \(\Psi\) of the target density. The ARMH algorithm is an MCMC sampling procedure, which allows that the condition is not satisfied for some \(\psi \in \Psi\). Let \(D^C\) be the complement of \(D\), and suppose that the current state of the chain is \(\psi\). Then the ARMH algorithm is defined as follows.
**A-R Step**: Generate a draw $\psi' \sim q(\psi \mid Y_T)$; accept $\psi'$ with probability

$$\alpha_{AR} = \min\left\{ 1, \frac{f(Y_T \mid \psi')p(\psi')}{cq(\psi' \mid Y_T)} \right\}.$$ 

Continue the process until a draw $\psi'$ has been accepted.

**M-H Step**: Given the current value $\psi$ and the candidate value $\psi'$:

1. if $\psi \in \mathcal{D}$, set $\alpha_{MH}(\psi, \psi' \mid Y_T) = 1$;
2. if $\psi \in \mathcal{D}^C$ and $\psi' \in \mathcal{D}$, set $\alpha_{MH}(\psi, \psi' \mid Y_T) = \frac{cq(\psi \mid Y_T)}{f(Y_T \mid \psi)p(\psi)}$;
3. if $\psi \in \mathcal{D}^C$ and $\psi' \in \mathcal{D}^C$, set $\alpha_{MH}(\psi, \psi' \mid Y_T) = \{1, \frac{f(Y_T \mid \psi')p(\psi')q(\psi \mid Y_T)}{f(Y_T \mid \psi)p(\psi)q(\psi' \mid Y_T)}\}$.

Return $\psi'$ with probability $\alpha_{MH}(\psi, \psi' \mid Y_T)$. Otherwise return $\psi$.

See also Chib (2001) for details of the ARMH algorithm.

### 3.3. MCMC method

For univariate GARCH models, Mitsui and Watanabe (2003) suggested to apply the ‘tailored’ ARMH algorithm, which is a natural extension of the tailored MH algorithm proposed by Chib and Greenberg (1994). Asai (2006) found that the approach is best for estimating univariate GARCH models, among the several MCMC methods, in regard to the mixing, efficiency and computational requirements of the MCMC.

For multivariate GARCH models, the direct application of the tailored ARMH algorithm can be inefficient for the following reason. As the works of Mitsui and Watanabe (2003) and Asai (2006) use the multivariate normal (or t) distributions for the candidate generating density of the MCMC algorithm, there is no guarantee that the generated samples of $\Omega$ and $N$ are positive definite. In the following, their works are extended by accommodating the generations of positive definite matrices.

Using the likelihood function (2.6) and the prior distributions, the posterior density function of $\psi$, is given by

$$\pi(\psi \mid Y_T) \propto f(Y_T \mid \psi) \times p(\Omega^{-1}) \times p(N^{-1}) \times p(\theta) \times p(k).$$

To obtain the posterior quantities of the parameters, this work implements the MCMC algorithm in five steps:

1. Initialize $\Omega, N, \theta, k$.
2. Generate $\Omega \mid N, \theta, k, Y_T$.
3. Generate $N \mid \Omega, \theta, k, Y_T$.
4. Generate $(\theta, k) \mid \Omega, N, Y_T$.
5. Go to Step 2.

For generating the positive definite matrices, $\Omega$ and $N$, the ARMH algorithm is used. The candidate generating density for $\Omega^{-1}$ is $W(\nu_{\Omega,1}, Q_{\Omega,1})$, where

$$\nu_{\Omega,1} = \nu_{\Omega,0} + T - m - 1, \quad Q_{\Omega,1} = \left(Q_{\Omega,0}^{-1} + \sum_{t=1}^{T} \varepsilon_t \varepsilon_t' \right)^{-1}.$$
It is motivated by an approximation of the unconditional distribution of $\varepsilon_t$ by $N(0, \Omega)$. Then the approximated posterior density function could be

$$p(\Omega^{-1}) \times |\Omega|^{-T/2} \exp \left\{ -\frac{1}{2} \text{tr} \left( \Omega^{-1} \sum_{t=1}^{T} \varepsilon_t \varepsilon_t' \right) \right\},$$

which gives the above candidate generating density. The candidate generating density for $N^{-1}$ is $W(\nu_{N,1}, Q_{N,1})$, where

$$\nu_{N,1} = \nu_{N,0} + T - m - 1, \quad Q_{N,1} = \left( Q_{N,0}^{-1} + \sum_{t=1}^{T} \eta_t \eta_t' \right)^{-1}.$$

In the $i$th step, generate $\Omega^{(i)}$ and $N^{(i)}$ from the conditional distributions, $\Omega^{-1} | N^{(i-1)}, \theta^{(i-1)}, Y_T$ and $N^{-1} | \Omega^{(i)}, \theta^{(i-1)}, Y_T$, via the ARMH algorithm, respectively.

For generating the remaining parameter vector, $\phi = (\theta', k)'$, the ARMH algorithm is used, as in Mitsui and Watanabe (2003) and Asai (2006). The candidate generating density is specified as the multivariate $t$ distribution with location given by the mode of $g(\phi) = \log f(Y_T | \Omega^{(i)}, N^{(i)}, \theta, k) + \log p(\theta) + \log p(k)$, and the dispersion given by the inverse of the Hessian evaluated at the mode. Specifically, the parameters of the proposal density are taken to be

$$m = \arg \max g(\phi),$$
$$V = \left\{ \frac{\partial^2 g(\phi)}{\partial \phi \partial \phi'} \right\}_{\phi=m}^{-1},$$

under the restriction that $\theta$ satisfies the stationary condition. In the $i$th step, generate $(\theta^{(i)}, k^{(i)})$ from the conditional distribution, $(\theta, k) | \Omega^{(i)}, N^{(i)}, Y_T$ via the ARMH algorithm.

As explained above, the approach of Mitsui and Watanabe (2003) and Asai (2006) use the multivariate truncated normal or $t$ distribution. Rather than generating $\Omega$ and $N$ from the multivariate truncated distribution with the restriction of positive definiteness, the new approach based on the Wishart distribution is more efficient because the approach can generate directly the positive definite matrices via the Wishart distribution.

For the implementation issues, the first 5,000 samples are discarded as the burn-in and the subsequent 10,000 samples are used for estimation. The convergence of the Markov chain is tested by the convergence diagnostic test of Geweke (1992), and the inverse of efficiency factor of Geweke (1992), which is called the ‘inefficiency factor’ is reported. The inefficiency factor is close to one when the Markov chain is close to being independent chain, i.e., it is most efficient.
4. Empirical analysis

4.1. Data and preliminary analysis

This section estimates and compares several asymmetric BEKK models with the heavy-tailed conditional distribution, including the BEKK-n, A-BEKK-n, GA-BEKK-n, BEKK-t, A-BEKK-t, and GA-BEKK-t models. Based on the estimation results, the news impact curves and cross leverage effects are also examined. Three stocks traded on the New York Stock Exchange are selected, namely: Bank of America (BAC), General Electric (GE), and International Business Machines (IBM). The close-to-close returns are calculated using the sample from September 13, 2007 to August 24, 2015, giving 2,000 observations.

Table 1 shows the descriptive statistics for all observations. Among the three series, the distributions of BAC and IBM are right-skewed, as in typical results for financial assets. On the other hand, the distribution of GE is rather symmetric.

As a preliminary analysis, the univariate GA-GARCH-\(t\) model is estimated. The model is given by

\[
y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t = \sqrt{h_t} \xi_t, \\
\xi_t = \zeta_t \sqrt{\frac{k - 2}{\kappa_t}}, \quad \zeta_t \sim N(0, 1), \quad \kappa_t \sim \chi^2(k), \\
h_t = \omega + \alpha(\varepsilon_{t-1} - \gamma)^2 + \lambda \eta_t^2 + \beta h_{t-1},
\]

where \(\eta_t\) is the negative part of \(\varepsilon_t\), and it is assumed that \(\mu_t = 0\) and \(\alpha + \beta + \lambda/2 < 1\). The GA-GARCH model reduces to the quadratic GARCH-\(t\) model when \(\lambda = 0\), while it becomes the GJR-\(t\) model when \(\gamma = 0\).

The tailored ARMH algorithm is used, as explained in subsection 3.2. With the log-likelihood function plus the log of the prior density function, its mode and negative inverse of the Hessian matrix at the mode are obtained for constructing the candidate generating function for the ARMH algorithm. The MCMC algorithm generates 15,000 samples, and discards the first 5,000 samples as the burn-in, using the subsequent 10,000 samples for estimation.

Table 2 shows the estimation results for the GA-GARCH-\(t\) model. The results of the convergence diagnostic test of Geweke (1992) indicate that the null hypothesis that the Markov chain converged to the stationary distribution for all cases can not be rejected. The values of the inefficiency factor implies that the sampling of \((\alpha, \beta, \lambda)\) is efficient compared to sampling \((\omega, k)\). The posterior means and 95% credible intervals for the degree-of-freedom parameter for the standardized \(t\) distribution suggests the fat-tailed conditional distribution. Also, there are overlaps of the 95% credible intervals, implying the adequateness of
Table 2. Estimation results for univariate GA-GARCH-t model.

<table>
<thead>
<tr>
<th>Param.</th>
<th>Mean</th>
<th>95% interval</th>
<th>CD</th>
<th>Inef.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) BAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>14.787</td>
<td>[10.774, 19.086]</td>
<td>0.9909</td>
<td>42404.</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1097</td>
<td>[0.0569, 0.1680]</td>
<td>0.9936</td>
<td>9.3914</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.8677</td>
<td>[0.8226, 0.9089]</td>
<td>0.9973</td>
<td>5.5661</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0272</td>
<td>[−0.0428, 0.1012]</td>
<td>0.8063</td>
<td>17.933</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.7940</td>
<td>[0.2715, 1.1512]</td>
<td>0.9908</td>
<td>492.19</td>
</tr>
<tr>
<td>$k$</td>
<td>4.9038</td>
<td>[4.2584, 5.7219]</td>
<td>0.9985</td>
<td>1444.4</td>
</tr>
<tr>
<td>(b) GE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>4.4821</td>
<td>[0.9929, 8.0366]</td>
<td>0.9954</td>
<td>32733.</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0174</td>
<td>[0.0011, 0.0400]</td>
<td>0.9989</td>
<td>1.1193</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.9058</td>
<td>[0.8723, 0.9362]</td>
<td>0.9998</td>
<td>2.6639</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.1341</td>
<td>[0.0752, 0.1988]</td>
<td>0.9990</td>
<td>9.8596</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>−0.1054</td>
<td>[−1.9610, 1.2826]</td>
<td>0.9288</td>
<td>7610.6</td>
</tr>
<tr>
<td>$k$</td>
<td>5.0486</td>
<td>[4.1513, 6.1249]</td>
<td>0.9990</td>
<td>2540.0</td>
</tr>
<tr>
<td>(c) IBM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>2.1445</td>
<td>[1.0435, 3.1885]</td>
<td>0.9982</td>
<td>29065.</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0772</td>
<td>[0.0179, 0.1405]</td>
<td>0.9943</td>
<td>9.7775</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.8641</td>
<td>[0.8120, 0.9124]</td>
<td>0.9996</td>
<td>6.4434</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0475</td>
<td>[−0.0493, 0.1493]</td>
<td>0.9886</td>
<td>24.852</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.6856</td>
<td>[0.1576, 1.1000]</td>
<td>0.9932</td>
<td>545.70</td>
</tr>
<tr>
<td>$k$</td>
<td>5.1287</td>
<td>[4.1797, 6.2885]</td>
<td>0.9988</td>
<td>2890.3</td>
</tr>
</tbody>
</table>

Note: ‘95% interval’ is the lower and upper quantiles of 95% credible interval. ‘CD’ is the $p$-value of the convergence diagnostic test by Geweke (1992), and ‘Inef’ is the value of the inefficiency factor.

specifying the multivariate models using the same degree-of-freedom parameter. The estimated 95% credible intervals of all parameters, except for $\lambda$, exclude zero for BAC and IBM, showing that the datasets prefer the quadratic GARCH model. On the other hand, the 95% credible intervals of $\gamma$ contain zero for GE, implying that the GJR model fits the dataset. Based on the results, the general asymmetric multivariate model is used for examining the news impact curves and cross-leverage effects.

The news impact curve (NIC) of the GA-GARCH model is given by

$$h_t = \begin{cases} 
\omega + \alpha (\varepsilon_{t-1} - \gamma)^2 & \varepsilon_{t-1} > 0 \\
\omega + \alpha (\varepsilon_{t-1} - \gamma)^2 + \lambda \varepsilon_{t-1}^2 & \varepsilon_{t-1} \leq 0 
\end{cases}$$

where $\omega = \omega + \beta \sigma^2$ and $\sigma^2$ is the unconditional variance. Figure 1 shows the NICs and their 95% credible intervals for the GAR-GARCH model, which are adjusted
The news impact curves are adjusted to annualized return.

Figure 1. News impact curves for GA-GARCH models.

to the annualized return series. Figure 1 indicates that the asymmetric patterns of the estimated NICs for BAC and IBM are similar, reflecting the estimation results of Table 2.
Table 3. Acceptance rates for ARMH candidates.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\Omega$</th>
<th>$\mathcal{N}$</th>
<th>$(\theta, k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEKK-n</td>
<td>1.000</td>
<td>1.000</td>
<td>—</td>
</tr>
<tr>
<td>A-BEKK-n</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>GA-BEKK-n</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>BEKK-$t$</td>
<td>1.000</td>
<td>1.000</td>
<td>—</td>
</tr>
<tr>
<td>A-BEKK-$t$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>GA-BEKK-$t$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: Regarding the BEKK-n, A-BEKK-n, and GA-BEKK-n models, the ML estimates are used for constructing candidate generating distribution for simplicity.

Table 4. DIC.

<table>
<thead>
<tr>
<th>Model</th>
<th>ranking</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEKK-n</td>
<td>6</td>
<td>71975.5</td>
</tr>
<tr>
<td>BEKK-$t$</td>
<td>5</td>
<td>71764.1</td>
</tr>
<tr>
<td>A-BEKK-n</td>
<td>4</td>
<td>71549.6</td>
</tr>
<tr>
<td>GA-BEKK-n</td>
<td>3</td>
<td>71530.5</td>
</tr>
<tr>
<td>A-BEKK-$t$</td>
<td>2</td>
<td>70908.1</td>
</tr>
<tr>
<td>GA-BEKK-$t$</td>
<td>1</td>
<td>70702.6</td>
</tr>
</tbody>
</table>

4.2. Results

The six kinds of BEKK models, including the BEKK-n, A-BEKK-n, GA-BEKK-n, BEKK-$t$, A-BEKK-$t$, and GA-BEKK-$t$ models, are estimated, using the trivariate returns of BAC, GE, and IBM.

Markov chains are constructed as explained in Section 3. As before, we generate 15,000 samples, and discard the first 5,000 samples as the burn-in, using the subsequent 10,000 samples for estimation. The ARMH algorithm is used for sampling from the conditional distributions of $\Omega \mid \cdot$, $\mathcal{N} \mid \cdot$, $(\theta, k) \mid \cdot$. In the $i$th iteration, it is required to obtain the mode of the log-likelihood function plus the log of the prior density function given $\Omega^{(i)}$ and $\mathcal{N}^{(i)}$, for constructing the candidate generating density for $(\theta, k)$. To save time when estimating the BEKK-n, A-BEKK-n, and GA-BEKK-n models, the maximum likelihood estimates and their covariance matrix estimates are used for constructing candidate generating density, in order to avoid maximization in each step. Table 3 gives the acceptance rate of the A-R step and M-H step in the ARMH algorithm, which are the estimates of the expected values $\alpha_{AR}$ and $\alpha_{MH}$. For generating $\Omega$ and $\mathcal{N}$, the Markov chain always accept the candidates. On the other hand, the acceptance rates for the tailored ARMH algorithm for $(\theta, k)$ are relatively low. Especially, the acceptance rate of the MH step is low. Table 3 also shows the importance of the maximization in each step for obtaining an appropriate candidate generating density.

The estimated models are compared by the deviance information criterion.
Figure 2. Posterior samples of $A$, $B$, and $C$ for the GA-BEKK-$t$ model.

(DIC) of Spiegelhalter et al. (2002), defined by

$$\text{DIC} = E_{\psi \mid Y_T} [D(\psi)] + p_d$$

$$p_D = E_{\psi \mid Y_T} [D(\psi)] - D(E_{\psi \mid Y_T} [\psi]),$$

$$D(\psi) = -2 \log f(Y_T \mid \psi) + C_y,$$

where $C_y$ is the constant term which depends only on the dataset $Y_T$. Using the posterior samples generated by the MCMC method, $E_{\psi \mid Y_T} [D(\psi)]$ is estimated by its sample analogue $\frac{1}{M} \sum_{i=1}^{R} D(\psi^{(i)})$. Also calculate $D(E_{\psi \mid Y_T} [\psi])$ by evaluating $D(\psi)$ at the average of the posterior samples. Table 4 presents the DICs for the six competing models. The symmetric BEKK-n and BEKK-$t$ models are the worst two among the six models. Table 4 indicates that the GA-BEKK-$t$ model has the smallest DIC, favoring the general asymmetric effects and the
standardized t distribution.

As the GA-BEKK-\( t \) model is selected, the samples of the MCMC technique for the model are examined further. Although the table of the convergence diagnostics of Geweke (1992) is omitted, tests indicate that all chains of the parameters converged to stationary distributions. Figure 2 presents the box plots of the posterior samples of \( A, B, \) and \( C \) for the GA-BEKK-\( t \) Model. Figure 2 implies that most of the posterior densities are skewed and have heavy tails, that is, they have non-normal distribution. Figure 3 also shows the box plots of the posterior samples of \( \Gamma \) and the kernel density of \( k \). Figure 3 indicate that the posterior densities of \( \Gamma \) have fat-tails, and that the posterior density of \( k \) is skewed and leptokurtic.

Figure 4 shows the NICs from the \( i \)th return to the \( j \)th conditional variance \((i, j = BAC, GE, IBM)\). The NICs are adjusted to annualized returns. The diagonal graphs in Fig. 4 are partial effects, removing the effects of other variables. From the off-diagonal graphs in Fig. 4, it is observed that co-leverage effects are not negligible.

5. Conclusions

The paper suggested the new GA-BEKK-\( t \) model, which is based on recent empirical findings. The paper also provided its news impact curves, and explained the MCMC method for estimating the new model. Empirical analysis using the three US firms indicates that the DIC favors the suggested general asymmetry and heavy-tailed distributions.

Although the paper has certain contributions, the paper can be extended in
several directions. First of all, the new model applied to the class of dynamic conditional correlations (DCC) of Engle (2002) can be considered, since the DCC is a popular approach for high dimensional volatility modeling. Secondly, factor specifications can be considered as in Engle et al. (1990), Vrontos et al. (2003b), and Lanne and Saikkonen (2007), since they are alternative approaches for reducing dimensions. Thirdly, the acceptance rate of the M-H step in the MCMC scheme may be improved by using the delayed rejection sampling of Mira (2001). As the delayed rejection sampling takes more CPU time compared to one sweep in the M-H algorithm, we need to use the approach carefully. Such tasks await future research.

Acknowledgements
The author is most grateful to Yoshi Baba, Hiroshi Kurata, Michael McAleer, and anonymous reviewers for very helpful comments and suggestions. The author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science, and
Australian Academy of Science.

REFERENCES


