Instituto Complutense de Análisis Económico

Realized Matrix-Exponential Stochastic Volatility with Asymmetry, Long Memory and Spillovers

Manabu Asai
Faculty of Economics
Soka University, Japan

Chia-Lin Chang
Department of Applied Economics
Department of Finance
National Chung Hsing University, Taiwan

Michael McAleer
Department of Quantitative Finance
National Tsing Hua University, Taiwan and
Econometric Institute Erasmus School of Economics
Erasmus University Rotterdam, The Netherlands and
Department of Quantitative Economics Complutense University of Madrid, Spain
And Institute of Advanced Sciences Yokohama National University, Japan

Abstract

The paper develops a novel realized matrix-exponential stochastic volatility model of multivariate returns and realized covariances that incorporates asymmetry and long memory (hereafter the RMESV-ALM model). The matrix exponential transformation guarantees the positivedefiniteness of the dynamic covariance matrix. The contribution of the paper ties in with Robert Basmann's seminal work in terms of the estimation of highly non-linear model specifications ("Causality tests and observationally equivalent representations of econometric models", Journal of Econometrics, 1988, 39(1-2), 69–104), especially for developing tests for leverage and spillover effects in the covariance dynamics. Efficient importance sampling is used to maximize the likelihood function of RMESV-ALM, and the finite sample properties of the quasi-maximum likelihood estimator of the parameters are analysed. Using high frequency data for three US financial assets, the new model is estimated and evaluated. The forecasting performance of the new model is compared with a novel dynamic realized matrix-exponential conditional covariance model. The volatility and co-volatility spillovers are examined via the news impact curves and the impulse response functions from returns to volatility and co-volatility.

Keywords Matrix-exponential transformation, Realized stochastic covariances, Realized conditional covariances, Asymmetry, Long memory, Spillovers, Dynamic covariance matrix, Finite sample properties, Forecasting performance.

JEL Classification C22, C32, C58, G32.

Working Paper nº 1615
September, 2016

UNIVERSIDAD
COMPLUTENSE
MADRID

ISSN: **2341-2356**
WEB DE LA COLECCIÓN: http://www.ucm.es/fundamentos-analisis-economico2/documentos-de-trabajo-del-icaeWorking papers are in draft form and are distributed for discussion. It may not be reproduced without permission of the author/s.

Realized Matrix-Exponential Stochastic Volatility with Asymmetry, Long Memory and Spillovers*

Manabu Asai

Faculty of Economics
Soka University, Japan

Chia-Lin Chang

Department of Applied Economics
Department of Finance
National Chung Hsing University, Taiwan

Michael McAleer

Department of Quantitative Finance
National Tsing Hua University, Taiwan
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam, The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid, Spain
and
Institute of Advanced Sciences
Yokohama National University, Japan

September 2016

*The authors are most grateful to Yoshi Baba and Karen Lewis for very helpful comments and suggestions. The first 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. The second and third authors are most grateful for the financial support of the Australian Research Council, National Science Council, Ministry of Science and Technology (MOST), Taiwan, and the Japan Society for the Promotion of Science. Address for correspondence: Faculty of Economics, Soka University, 1-236 Tangi-machi, Hachioji, Tokyo 192-8577, Japan. Email address: m-asai@soka.ac.jp.

Abstract

The paper develops a novel realized matrix-exponential stochastic volatility model of multivariate returns and realized covariances that incorporates asymmetry and long memory (hereafter the RMESV-ALM model). The matrix exponential transformation guarantees the positive-definiteness of the dynamic covariance matrix. The contribution of the paper ties in with Robert Basmann's seminal work in terms of the estimation of highly non-linear model specifications ("Causality tests and observationally equivalent representations of econometric models", Journal of Econometrics, 1988, 39(1-2), 69–104), especially for developing tests for leverage and spillover effects in the covariance dynamics. Efficient importance sampling is used to maximize the likelihood function of RMESV-ALM, and the finite sample properties of the quasi-maximum likelihood estimator of the parameters are analysed. Using high frequency data for three US financial assets, the new model is estimated and evaluated. The forecasting performance of the new model is compared with a novel dynamic realized matrix-exponential conditional covariance model. The volatility and co-volatility spillovers are examined via the news impact curves and the impulse response functions from returns to volatility and co-volatility.

Keywords: Matrix-exponential transformation, Realized stochastic covariances, Realized conditional covariances, Asymmetry, Long memory, Spillovers, Dynamic covariance matrix, Finite sample properties, Forecasting performance.

JEL classifications: C22, C32, C58, G32.

1 Introduction

Recent empirical analyses for estimating and forecasting volatility emphasizes realized measures such as the realized kernel of Barndorff-Nielsen et al. (2008). Even though we can obtain a consistent estimator of true volatility, there are non-negligible differences referred to as the 'realized volatility error' (see Barndorff-Nielsen and Shephard (2002)). For removing the estimation bias caused by the realized volatility error in estimating stochastic volatility (SV) models, Barndorff-Nielsen and Shephard (2002), Bollerslev and Zhou (2002), Takahashi, Omori and Watanabe (2009), and Asai, McAleer and Medeiros (2012a,b) showed that it is useful to use an ad hoc approach that accommodates an error term with constant variance.

As the information of returns and realized volatility measures are available simultaneously, Engle and Gallo (2006), Shephard and Sheppard (2010), and Hansen, Huang and Shek (2012), among others, extended the class of generalized autoregressive conditional heteroskedasiticy (GARCH) models using information such as the range, squared returns, and realized measure of volatility. Hansen, Huang and Shek (2012) developed the 'realized GARCH' model, which is based on the traditional returns equation and an additional equation of a realized measure. From this viewpoint, we may call the specification of Takahashi, Omori and Watanabe (2009) as the 'realized SV' model, since they use daily returns and a realized volatility measure simultaneously. Recently, Hansen and Huang (2016) extend the work of Hansen, Huang and Shek (2012) to develop realized exponential GARCH (EGARCH) models (see Martinet and McAleer (2016) and McAleer and Hafner (2014) for theoretical problems associated with EGARCH models).

In the univariate case, it is popular to specify that the log-volatility follows an autoregressive and moving average (ARMA) process. By considering a model of log-volatility rather than volatility itself, the specified model has no need to impose additional restrictions, apart from stationary and invertibility. In the multivariate framework, there are several approaches to guarantee the positive definiteness of time-varying covariance matrices, including the BEKK model of Engle and Kroner (1995) and the dynamic conditional correlation model of Engle (2002). In multivariate SV models with or without realized covariance, several specifications including the Cholesky decomposition models of Chiriac and Voev (2011) and Loddo, Ni and Sun (2011), the matrix-exponential models of Bauer and Vorkink (2011) and Ishihara, Omori, and Asai (2016), the Wishart autoregressive model of Gourieroux et al. (2009), and the dynamic correlation model of Asai and McAleer (2009a), among others, guarantee the covariance matrix to be positive definite. Among them, the matrix-exponential transformation enables us to have the advantages of specifying log-volatility and positive definiteness simultaneously.

The seminal work of Chiu, Leonard, and Tsui (1996) suggested using the matrix-exponential transformation to guarantee the positive definiteness of the covariance matrix. While Kawakatsu (2006) developed matrix-exponential GARCH models, Asai, McAleer and Yu (2006) and Ishihara, Omori and Asai (2016) proposed matrix-exponential SV models. For modeling realized covariances, Bauer and Vorkink (2011) applied the matrix-exponential transformation, and Asai and McAleer (2015) accommodate a factor structure, asymmetric effects and long-range dependence in order to develop a matrix-exponential factor MSV model.

The first purpose of the paper is to develop realized matrix-exponential SV models with asymmetry and long memory (RMESV-ALM), by extending the specifications of Bauer and Vorkink (2011) and Asai and McAleer (2015). The new model includes the realized SV model with asymmetric effect, developed by Takahashi, Omori and Watanabe (2009), and the matrix-exponential SV model of Ishihara, Omori and Asai (2016), as special cases. For the RMESV-ALM model, we accommodate the RV error as a realized measure of covariances, instead of a direct specification of realized covariance. As the specification assumes that the covariance matrix of a return vector is latent, we use the Monte Carlo likelihood approach of Durbin and Koopman (1997) for estimating the new model. For this purpose, we use the simulation smoother for long memory with an additive noise process developed by Asai and So (2016), which is an extension of simulation smoothers of de Jong and Shephard (1995) and So (1999).

Using the new RMESV-ALM model, we examine the leverage and spillover effects from a return to the own and other future volatilities, respectively. We also compare the forecasting performance with a novel realized matrix-exponential GARCH model, which has not previously been estimated.

The remainder of the paper is organized as follows. Section 2 develops the new RMESV-ALM model, and derives a representation for the asymmetric effects. Section 3 explains the MCL approach of Durbin and Koopman (1997) and the simulation smoother of Asai and So (2016). Section 3 also discusses the semi-parametric estimation of long memory parameters in a vector stochastic process, as suggested by Shimotsu (2007). Section 4 provides an empirical example for three stocks traded on the New York Stock Exchange. Section 5 gives some concluding remarks.

2 Realized Matrix-Exponential SV Model with Asymmetry and Long Memory

2.1 Realized Matrix-Exponential SV Model

In order to model dynamic covariances, we consider the matrix-exponential transformation that guarantees the positive definiteness of the covariance matrix. Chiu, Leonard, and Tsui (1996) proposed the idea of specifying the time-varying covariance matrix, Kawakatsu (2006) considered the matrix-exponential GARCH model, while Asai, McAleer and Yu (2006) and Ishihara, Omori, and Asai (2016) developed the matrix-exponential SV model. Compared with the matrix-exponential GARCH model, the matrix-exponential SV model has flexibility in the error term of the volatility equation. Note that Kawakatsu (2006) uses the unstandardized residuals, so that the univariate EGARCH model is not a special case of its purported multivariate counterpart, but we may develop alternative specifications based on standardized residuals, as in Nelson (1991) and Asai and McAleer (2015). In the specification developed below, we consider the model such that the standardized residuals affect future volatility.

For any square matrix A, the matrix-exponential transformation is defined by $\mathrm{Exp}(A) = \sum_{i=0}^{\infty}(1/i!)A^i$, with $A^0 = I$. The same result is obtained by using the spectral decomposition, as we have $\mathrm{Exp}(A)$ by replacing the eigenvalues of A by their respective exponential transformations. Note that $\mathrm{Exp}(A)$ is positive definite, whenever A is symmetric. In the same manner, $\mathrm{Log}(B)$ is defined by its spectral decomposition of a positive definite matrix, B, with replacement of the logarithmic transformation of the eigenvalues. We also denote $\log(x)$ $(\exp(x))$ for the vector x as the element-by-element logarithmic (exponential) transformation of x.

Let $\boldsymbol{y}_t$ and $\boldsymbol{X}_t$ be an $m \times 1$ vector of financial asset returns and an $m \times m$ matrix created by the matrix-logarithmic transformation of a realized covolatility matrix measure, respectively. We consider the realized matrix-exponential SV (RMESV) model, as follows:

$$\boldsymbol{y}_t = \Omega_t^{1/2} \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim i.i.d.N(0, I_m), \tag{1}$$

$$\Omega_t = \boldsymbol{S} \mathrm{Exp}(\boldsymbol{H}_t) \boldsymbol{S}, \tag{2}$$

$$\boldsymbol{X}_t = K + \boldsymbol{H}_t + U_t, \qquad \boldsymbol{u}_t = \mathrm{vech}(U_t) \sim i.i.d.N(0, \Sigma_u), \tag{3}$$

where $\boldsymbol{H}_t = \{h_{ij,t}\}$ is an $m \times m$ symmetric matrix of unobserved components, with $\boldsymbol{H}_t = O$, and $\boldsymbol{S} = \{s_{ij}\}$ is an $m \times m$ positive definite matrix, with $s_{ii} > 0$ $(i = 1, \ldots, m)$, and $K = \{\kappa_{ij}\}$ is an $m \times m$ symmetric matrix of parameters. For the error terms, $U_t = \{u_{ij,t}\}$ is an $m \times m$ symmetric matrix of normal variates, and it is assumed that $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{mt})'$ and U_s are mutually independent for any t, s. The matrix-exponential transformation of $\boldsymbol{H}_t$ guarantees the positive definiteness of Ω_t, which is the stochastic covariance matrix of $\boldsymbol{y}_t$. The RMESV model in equations (1)-(3) is a multivariate extension of Takahashi, Omori and Watanabe (2009). We examine alternative representations of the RMESV model in this section.

There is strong evidence of long-range dependence in the volatility models of financial returns. For long-range dependence in financial volatility, Baillie, Bollerslev and Mikkelsen (1996) developed the fractionally-integrated GARCH model, while Bollerslev and Mikkelsen (1996) suggested the fractionally-integrated EGARCH model (see Martinet and McAleer (2016) and McAleer and Hafner (2014) for caveats regarding EGARCH). In addition to GARCH specifications, Breidt, Crato and de Lima (1998), Harvey (1998), Pérez and Ruiz (2001), So (2002) and So and Kwok

(2006) examined long-memory stochastic volatility models. For the unobserved components, we assume that each element of $\boldsymbol{H}_t$ follows an autoregressive fractionally-integrated moving average (ARFIMA) process. In matrix form, we can write:

$$\Phi(B) \circ D(B) \circ \boldsymbol{H}_{t+1} = \Theta(B) \circ V_t, \tag{4}$$

with

$$\Phi(B) = \iota_m \iota_m' - \Phi_1 B - \cdots - \Phi_p B^p,$$

$$\Theta(B) = \iota_m \iota_m' + \Theta_1 B + \cdots + \Theta_q B^q,$$
$$D(B) = \begin{pmatrix} (1-B)^{d_{11}} & \cdots & (1-B)^{d_{1m}} \\ \vdots & \ddots & \vdots \\ (1-B)^{d_{m1}} & \cdots & (1-B)^{d_{mm}} \end{pmatrix},$$

where ι_m is a vector of ones, $\Phi_l = \{\phi_{ij,l}\}$ and $\Theta_l = \{\theta_{ij,l}\}$ are $m \times m$ symmetric matrices of parameters, and $V_t = \{v_{ii,t}\}$ is an $m \times m$ symmetric matrix of error terms, as defined below. As the ARFIMA(1,d,0) model is typically used in the literature, we follow the simple specification in our empirical analysis.

2.2 Asymmetric Effects

For incorporating asymmetric effects in SV models, Wiggins (1987), Chesney and Scott (1989), Harvey and Shephard (1996), and Asai and McAleer (2006) incorporate a negative correlation between the disturbances of return and future log-volatility. In the specification, a negative return leads to an increase in future volatility, which is called the leverage effect. We will consider not only leverage effects, but also the asymmetric effects for other volatilities. For this purpose, we consider the specification:

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{u}_t \\ \boldsymbol{v}_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_m & O & \Lambda' \\ O & \Sigma_u & O \\ \Lambda & O & \Sigma_v \end{bmatrix} \right), \tag{5}$$

where $\boldsymbol{v}_t = \text{vech}(V_t)$, and Λ is an $m^* \times m$ matrix of parameters, with $m^* = m(m+1)/2$. In order to guarantee the positive definiteness of the covariance matrix of $(\boldsymbol{\varepsilon}_t', \boldsymbol{u}_t', \boldsymbol{v}_t')'$, we need to assume that $\Sigma_v - \Lambda\Lambda'$ is positive definite.

In considering the negative correlation between ε_{it} and $v_{ii,t}$, we use the following notation. Denote the duplication and elimination matrices as D_m ($m^2 \times m^*$) and L_m ($m^* \times m^2$), respectively. For any $m \times m$ symmetric matrix A, the duplication matrix is defined such that $\text{vec}(A) = D_m \text{vech}(A)$, while the elimination matrix is defined such that $\text{vech}(A) = L_m \text{vec}(A)$. Defining an m vector of diagonal elements of V_t as $\boldsymbol{z}_t = (v_{11,t}, v_{22,t}, \ldots, v_{mm,t})'$, we obtain $\boldsymbol{z}_t = \underline{S}_m D_m \boldsymbol{v}_t$, where $\underline{S}_m = \{\underline{s}_{ij}\}$ is the $m \times m^2$ selection matrix, with $\underline{s}_{ij} = 1$ if $j = m(i-1)+i$ $(i = 1, 2, \ldots, m)$, and zero otherwise. As $D_m \boldsymbol{v_t} = \text{vec}(V_t)$, $\underline{S}_m$ selects the diagonal elements of V_t via $\underline{S}_m \text{vec}(V_t)$, and we can specify:

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{z}_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_m & P\underline{\Sigma}_z^{1/2} \\ \underline{\Sigma}_z^{1/2} P & \Sigma_z \end{bmatrix} \right),$$

where $\Sigma_z = \underline{S}_m D_m \Sigma_v D_m' \underline{S}_m'$, $\underline{\Sigma}_z$ is the diagonal matrix, with the diagonal elements of Σ_z, $P = \text{diag}(\rho_1, \rho_2, \ldots, \rho_m)$. Note that $\underline{\Sigma}_z^{1/2} P$ is a diagonal matrix. In the specification, we obtain:

$$\text{corr}(\varepsilon_{it}, v_{ii,t}) = \rho_i, \quad \text{corr}(\varepsilon_{it}, v_{ij,t}) = 0 \ (j \neq i),$$

where ρ_i is in the range $(-1, 0)$ for purposes of the leverage effect from the ith return to its own one-step-ahead volatility.

Returning to the specification of Λ, it is straightforward to consider $\Lambda = L_m \underline{S}'_m \underline{\Sigma}_z^{1/2} P$ for the simple leverage effects, as it only gives the negative correlation between ε_{it} and $v_{ii,t}$. By extending the specification, we obtain an equivalent representation of Λ based on leverage effects, as:

$$\Lambda = Q \underline{\Sigma}_z^{1/2} P, \tag{6}$$

where $Q = \{q_{ij}\}$ is an $m^* \times m$ matrix, with $q_{ij} = 1$ if the (i,j) element of $L_m \underline{S}'_m$ is one. Noting that the numbers of parameters in Q and P are $m(m^* - 1)$ and m, respectively, it is obvious that there is no restriction on Λ, except for $-1 < \rho_i < 0$. This specification reduces to the model with simple leverage effects, when $Q = L_m \underline{S}'_m$. We call equation (6) the 'concentrated leverage representation'.

For understanding the structure of the concentrated leverage representation, we consider the distribution of $\boldsymbol{v}_t$ conditional on $\boldsymbol{\varepsilon}_t$, which is given as $\boldsymbol{v}_t | \boldsymbol{\varepsilon}_t \sim N(\Lambda \boldsymbol{\varepsilon}_t, \Sigma_v - \Lambda\Lambda')$. For $m = 2$, the conditional mean vector is given by:

$$\Lambda \boldsymbol{\varepsilon}_t = \begin{pmatrix} 1 & q_{12} \\ q_{21} & q_{22} \\ q_{31} & 1 \end{pmatrix} \begin{pmatrix} \sigma_{v,11} & 0 \\ 0 & \sigma_{v,33} \end{pmatrix} \begin{pmatrix} \rho_1 \varepsilon_{1t} \\ \rho_2 \varepsilon_{2t} \end{pmatrix}.$$

In general, $h_{ij,t+1}$ is affected by the linear combination of $\{\rho_1 \varepsilon_{1t}, \ldots, \rho_m \varepsilon_{mt}\}$. In other words, we can describe the cross-leverage effects from the ith return to $h_{ii,t+1}$, $h_{ij,t+1}$ and $h_{jj,t+1}$ ($i \neq j$, $i, j = 1, \ldots, m$). We will refer to the REMSV model with asymmetry and long memory, equations (1)–(6), as 'REMSV-ALM'.

In order to show how the shocks in returns at time t affect the volatilities at time $t+1$, we describe the news impact curve, following Engle and Ng (1993). Similar ideas for stochastic volatility models are discussed in Yu (2005) and Asai and McAleer (2009b). Let $\boldsymbol{H}_t = O$ and

$V_t = O$ for their past values, and consider the case where:

$$\Omega_{t+1} = \boldsymbol{S}\mathrm{Exp}(\boldsymbol{H}_{t+1})\boldsymbol{S}, \quad \boldsymbol{h}_{t+1} = \Lambda \boldsymbol{S}^{-1}\boldsymbol{y}_t, \tag{7}$$

where $\boldsymbol{h}_t = \mathrm{vech}(\boldsymbol{H}_t)$, and Λ is defined in equation (6). We examine the news impact curves on the standard deviations and correlations of $\boldsymbol{y}_{t+1}$ from $\boldsymbol{y}_t$, using equation (7).

2.3 Observationally Equivalent Representation and Tests for Leverage and Spillover Effects

Before considering the causality from $\boldsymbol{y}_t$ to Ω_{t+1}, we return to the important concept of the 'observationally equivalent representation' of Basmann (1988) for causality analysis. Basmann (1988) uses a structural vector autoregression (SVAR) model, which has its reduced form derived by multiplying a matrix for normalization. Basmann (1988) shows that a causal relationship will change, depending on the choice of the matrix, which is an example of the problems associated with observationally equivalent representations. For this problem, estimation of the SVAR model requires imposing restrictions, as discussed in Waggoner and Zha (2003).

For the RMESV model, the observationally equivalent representation of (2) is given by:

$$\Omega_t = \mathrm{Exp}(\boldsymbol{A}_t), \quad E(\boldsymbol{A}_t) = M, \tag{8}$$

where $\boldsymbol{A}_t$ is an unobservable process of the $m \times m$ symmetric matrix, and M is an $m \times m$ symmetric matrix of parameters. Compared with $\boldsymbol{S}$ for the specification (2), it has the same number of parameters. By a property of the matrix-logarithmic transformation, we can obtain $\mathrm{Log}(\Omega_t) = \mathrm{Log}\left(\boldsymbol{S}\mathrm{Exp}(\boldsymbol{H}_t)\boldsymbol{S}\right) \neq 2\mathrm{Log}(\boldsymbol{S}) + \boldsymbol{H}_t$, in general (see Chiu, Leonard, and Tsui (1996)). The result indicates that $\boldsymbol{A}_t \neq 2\mathrm{Log}(\boldsymbol{S}) + \boldsymbol{H}_t$ and $\boldsymbol{H}_t \neq \boldsymbol{A}_t - \boldsymbol{M}$, but these two specifications have the same Ω_t. This is the reason why we can avoid specifying $\boldsymbol{S}$ as a diagonal matrix for the structure of the RMESV model, (1)-(3). In other words, a parsimonious specification of S causes a (hidden) restriction on the specification in (8).

As the matrix-exponential transformation is highly nonlinear, it is important to check such an observationally equivalent representation before estimating the models and testing for non-causality. One of the merits in specification (2) is that we can identify two sources of spillover effects separately: (i) $\boldsymbol{S}$ as the constant component in the dynamic covariance structure of Ω_t; and (ii) via $\boldsymbol{H}_t$, which represents the dynamic component.

Based on the concentrated leverage representation (6) under Ω_t in equation (2), we consider three kinds of tests. The first is a test for the leverage effect:

$$\mathrm{H}_0 : \rho_i = 0 \quad \text{vs} \quad \mathrm{H}_1 : \rho_i < 0,$$

for the ith asset return ($i = 1, \dots, m$). The second is a test for constant spillover effects:

$$\mathrm{H}_0 : s_{ij} = 0 \quad (i \neq j) \quad \text{vs} \quad \mathrm{H}_1 : s_{ij} \neq 0,$$

for the off-diagonal elements of $\boldsymbol{S}$. The third one is a test for dynamic spillover effects:

$$\mathrm{H}_0 : Q = L_m \underline{S}_m \quad \text{vs} \quad \mathrm{H}_1 : Q \neq L_m \underline{S}_m,$$

as specified in equation (6). We can test the first constraint via the t test, while we use the likelihood ratio test for the remaining two tests. For the likelihood ratio test, the test statistics are expected to follow the asymptotic $\chi^2\left(m(m-1)/2\right)$ and $\chi^2\left(m(m^*-1)\right)$ distribution, respectively, under the null hypothesis of no spillover effects.

We provide a detailed discussion of each test, and consider first the leverage effect. Consider the case $\rho_i < 0$ under $Q = L_m \underline{S}_m$ and a diagonal $\boldsymbol{S}$. By the definition of the matrix-exponential transformation, a negative shock in the ith element of $\boldsymbol{y}_t$ in equation (7) produces a diagonal matrix, Ω_{t+1}, with $\omega_{ii,t+1} > 1$ and $\omega_{jj,t+1} = 1$ $(j \neq i,\ j = 1, \dots, m)$. Hence, the effect is restricted to the relationship between the ith return and the associated ith volatility. The negative correlation between ε_{it} and $\Omega_{ii,t}$ remains under the general Q and $\boldsymbol{S}$, if $\rho_i < 0$.

For constant spillover effects, the restriction is straightforward under $\boldsymbol{H}_t = O$, that is, $\Omega_t = \boldsymbol{S}^2$. For dynamic spillover effects, we start from the volatility matrix for news impacts in (7). For no dynamic spillover effects under $Q = L_m \underline{S}_m$, such that at least one of the ρ_i is not equal to zero, any shock in $\boldsymbol{y}_t$ in equation (7) yields diagonal matrices, $\boldsymbol{H}_{t+1}$ and Ω_{t+1}. By the diagonal structure, no dynamic spillover effects on $\boldsymbol{H}_{t+1}$ will correspond to those on Ω_{t+1}. When there is a spillover effect in h_{t+1} via $Q \neq L_m \underline{S}_m$, Ω_{t+1} is no longer diagonal. As a result, the test of dynamic spillover effects works for $\boldsymbol{H}_t = O$ and $V_t = O$.

Although $\boldsymbol{H}_t \neq O$ and $V_t \neq O$ in the RMESV model, we can approximate $\mathrm{Exp}(\boldsymbol{H}_t)$ as:

$$\mathrm{Exp}(\boldsymbol{H}_t) \simeq I_m + \boldsymbol{H}_t,$$

by a first-order Taylor series expansion. The approximation is accurate when $\boldsymbol{H}_t \simeq O$, and we can apply the above discussion for the RMESV model to test dynamic spillover effects.

3 Monte Carlo Maximum Likelihood Estimation

This section develops the Monte Carlo likelihood (MCL) method to estimate the new REMSV-ALM model. We will explain below the general framework of the MCL approach proposed by

Durbin and Koopman (1997), and construct the approximating densities, as required by the MCL approach.

3.1 Likelihood Evaluation via Importance Sampling

For the MCL method, the likelihood function can be approximated arbitrarily by decomposing it into a Gaussian part, which is constructed by the Kalman filter, and a remainder function, for which the expectation is evaluated through simulation.

Let $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T)$, $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_T)$, and $\boldsymbol{H} = (\boldsymbol{H}_1, \ldots, \boldsymbol{H}_T)$. With the vector of unknown parameters, $\boldsymbol{\psi}$, we can express the density of $(\boldsymbol{Y}, \boldsymbol{X})$ as:

$$p(\boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi}) = \int p(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{H}; \boldsymbol{\psi}) d\boldsymbol{H} = \int p(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi}) p(\boldsymbol{H}; \boldsymbol{\psi}) d\boldsymbol{H}. \tag{9}$$

Durbin and Koopman (1997) considered the likelihood of the approximating Gaussian model as:

$$L_q(\boldsymbol{\psi}) = q(\boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi}) = \frac{q(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi}) p(\boldsymbol{H}; \boldsymbol{\psi})}{q(\boldsymbol{H} | \boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi})}. \tag{10}$$

Note that the MCL method uses the same density of $(\boldsymbol{H}; \boldsymbol{\psi})$ as the true model to construct the approximating Gaussian model. Substituting $p(\boldsymbol{H}; \boldsymbol{\psi})$ from the above equation into (9) gives:

$$L(\boldsymbol{\psi}) = \int L_g(\boldsymbol{\psi}) \frac{p(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi})}{q(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi})} q(\boldsymbol{H} | \boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi}) d\boldsymbol{H} = L_q(\boldsymbol{\psi}) E_q \left[\frac{p(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi})}{q(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi})} \right],$$

where E_q denotes the expectation with respect to $q(\boldsymbol{H} | \boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi})$. The advantage of the approach of Durbin and Koopman (1997) is that it requires simulation only to estimate departures of the likelihood from the Gaussian likelihood, rather than the likelihood itself. Durbin and Koopman (1997) suggested that $q(\boldsymbol{H} | \boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi})$ be used as the importance density for the simulations.

The log-likelihood function is given by:

$$\log L(\boldsymbol{\psi}) = \log L_q(\boldsymbol{\psi}) + \log E_q \left[\frac{p(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi})}{q(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}; \boldsymbol{\psi})} \right], \tag{11}$$

and its consistent estimator is given by:

$$\log L(\hat{\boldsymbol{\psi}}) = \log L_q(\boldsymbol{\psi}) + \log \bar{w} + \frac{\bar{w}^2}{2N s_w^2}, \tag{12}$$

where N is the number of simulations:

$$\bar{w} = \frac{1}{N} \sum_{i=1}^{N} w_i, \qquad s_w^2 = \frac{1}{N-1} \sum_{i=1}^{N} (w_i - \bar{w})^2, \qquad w_i = \frac{p(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}^{(i)}; \boldsymbol{\psi})}{q(\boldsymbol{Y}, \boldsymbol{X} | \boldsymbol{H}^{(i)}; \boldsymbol{\psi})},$$

and $\boldsymbol{H}^{(i)}$ denotes a draw from the importance density $q(\boldsymbol{H} | \boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi})$ (for further details, see Durbin and Koopman (1997, 2001)).

3.2 Constructing the Candidate Density

This subsection develops a candidate density for the importance sampling of the likelihood. Defining $\boldsymbol{x}_t = \text{vech}(\boldsymbol{X}_t)$ and $\boldsymbol{w}_t = (\boldsymbol{y}_t, \boldsymbol{x}_t)$, we can write the observation equation as:

$$\boldsymbol{w}_t = \begin{pmatrix} \boldsymbol{y}_t \\ \boldsymbol{x}_t \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\kappa} + \boldsymbol{h}_t \end{pmatrix} + \begin{pmatrix} \Omega_t^{1/2} & O \\ O & I_{m^*} \end{pmatrix} \begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{u}_t \end{pmatrix},$$

where $\boldsymbol{\kappa} = \text{vech}(K)$ and $\boldsymbol{h}_t = \text{vech}(\boldsymbol{H}_t)$. Conditional on $\boldsymbol{h}_t$, the contribution of the true log-likelihood at t is given by:

$$\begin{aligned}
p_t &= p_{1t} + p_{2t}, \\
p_{1t} &= -\frac{m}{2} \log(2\pi) - \frac{1}{2} \log |\Omega_t| - \frac{1}{2} \boldsymbol{y}_t' \Omega_t^{-1} \boldsymbol{y}_t, \\
p_{2t} &= -\frac{m^*}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (\boldsymbol{x}_t - \boldsymbol{\kappa} - \boldsymbol{h}_t)' \Sigma^{-1} (\boldsymbol{x}_t - \boldsymbol{\kappa} - \boldsymbol{h}_t).
\end{aligned}$$

We linearize the observation equation for the part of $\boldsymbol{y}_t$.

For the structure of the approximating model, consider:

$$\tilde{\boldsymbol{w}}_t = \tilde{Z}\boldsymbol{h}_t + \begin{pmatrix} \tilde{\boldsymbol{\varepsilon}}_t \\ \boldsymbol{u}_t \end{pmatrix}, \qquad \begin{pmatrix} \tilde{\boldsymbol{\varepsilon}}_t \\ \boldsymbol{u}_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} C_t & O \\ O & \Sigma_u \end{bmatrix} \right), \tag{13}$$

where:

$$\tilde{\boldsymbol{w}}_t = \begin{pmatrix} \tilde{\boldsymbol{y}}_t \\ \boldsymbol{x}_t - \boldsymbol{\kappa} \end{pmatrix}, \qquad \tilde{Z} = \begin{pmatrix} Z \\ I_{m^*} \end{pmatrix},$$

with $Z = \underline{S}_m D_m$, which selects the diagonal elements of $\boldsymbol{H}_t$ via $Z\boldsymbol{h}_t$, C_t is a time-varying m-dimensional positive-definite matrix, $\tilde{\boldsymbol{y}}_t$ will be defined below, using $\boldsymbol{y}_t$ and $\tilde{\boldsymbol{h}}_t$, where $\tilde{\boldsymbol{h}}_t$ is a trial value of $\boldsymbol{h}_t$. Conditional on $\boldsymbol{h}_t$, the contribution of the approximating log-likelihood at t is given by:

$$q_t = q_{1t} + p_{2t},$$
$$q_{1t} = -\frac{m}{2}\log(2\pi) - \frac{1}{2}\log|C_t| - \frac{1}{2}(\tilde{\boldsymbol{y}}_t - Z\boldsymbol{h}_t)' C_t^{-1} (\tilde{\boldsymbol{y}}_t - Z\boldsymbol{h}_t).$$

Define $\boldsymbol{\xi}_t = Z\boldsymbol{h}_t$, and consider an approximation of the first derivative of p_{1t} around $\tilde{\boldsymbol{\xi}}_t = Z\tilde{\boldsymbol{h}}_t$, to obtain:

$$\frac{\partial p_{1t}}{\partial \boldsymbol{\xi}_t} \approx \dot{p}_{1t} + \ddot{p}_{1t}(\boldsymbol{\xi}_t - \tilde{\boldsymbol{\xi}}_t), \tag{14}$$

where

$$\dot{\boldsymbol{p}}_{1t} = \left.\frac{\partial p_{1t}}{\partial \boldsymbol{\xi}_t}\right|_{h_t = \tilde{h}_t} = \frac{1}{2}\left[\nabla\Omega_t|_{h_t=\tilde{h}_t} \right]' \left(\tilde{\Omega}_t^{-1} \otimes \tilde{\Omega}_t^{-1} \right) \mathrm{vec}(\boldsymbol{y}_t\boldsymbol{y}_t' - \tilde{\Omega}_t),$$
$$\ddot{\boldsymbol{p}}_{1t} = E\left(\left. \frac{\partial^2 p_{1t}}{\partial \boldsymbol{\xi}_t \partial \boldsymbol{\xi}_t'} \right| \boldsymbol{h}_t = \tilde{\boldsymbol{h}}_t \right) = -\frac{1}{2}\left[\nabla\Omega_t|_{h_t=\tilde{h}_t} \right]' \left(\tilde{\Omega}_t^{-1} \otimes \tilde{\Omega}_t^{-1} \right) \left[\nabla\Omega_t|_{h_t=\tilde{h}_t} \right],$$

with $\nabla\Omega_t = \partial \mathrm{vec}(\Omega_t)/\partial \boldsymbol{\xi}_t'$, and $\tilde{\Omega}_t$ and $\tilde{\boldsymbol{H}}_t$ are Ω_t and $\boldsymbol{H}_t$ evaluated at $\boldsymbol{h}_t = \tilde{\boldsymbol{h}}_t$, respectively.

Noting that $\partial \mathrm{vec}(X^i)/\partial \mathrm{vec}(X)' = \sum_{j=0}^{i-1}(X')^{i-1-j} \otimes X^j$ for any $m \times m$ matrix X, we obtain:

$$\nabla\Omega_t = \frac{\partial \mathrm{vec}(\Omega_t)}{\partial \mathrm{vec}(\boldsymbol{H}_t)'}\underline{S}_m' = (\boldsymbol{S} \otimes \boldsymbol{S}) \sum_{i=0}^{\infty} \frac{1}{i!} \frac{\partial \mathrm{vec}(\boldsymbol{H}_t^i)}{\partial \mathrm{vec}(\boldsymbol{H}_t)'}\underline{S}_m' = (\boldsymbol{S} \otimes \boldsymbol{S}) \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{j=0}^{i-1} \left(\boldsymbol{H}_t^{i-j-1} \otimes \boldsymbol{H}_t^j\right) \underline{S}_m'.$$

For the approximating density, we specify:

$$\tilde{\boldsymbol{y}}_t = \tilde{\boldsymbol{\xi}}_t + \ddot{\boldsymbol{p}}_{1t}^{-1}\dot{\boldsymbol{p}}_{1t}, \qquad C_t = \ddot{\boldsymbol{p}}_t^{-1}, \tag{15}$$

which, together with equation (14), gives $\partial p_{1t}/\partial \boldsymbol{\xi}_t \approx \partial q_{1t}/\partial \boldsymbol{\xi}_t$.

By the transformation (15), we lose the information on the sign of the elements of $\boldsymbol{y}_t$, but we can recover it by the approach of Harvey and Shephard (1996) and Asai and McAleer (2006). Define $\boldsymbol{s}_t = (s_{1t}, \dots, s_{mt})'$, where s_{it} takes one (minus one) if y_{it} is positive (negative or equal to zero). Conditional on the signs of each element of $\boldsymbol{y}_t$, we can write equation (4) as:

$$\boldsymbol{\phi}(B) \circ \boldsymbol{d}(B) \circ (\boldsymbol{h}_{t+1} - \boldsymbol{\mu}) = \boldsymbol{\theta}(B) \circ \tilde{\boldsymbol{v}}_t, \tag{16}$$

where $\boldsymbol{\phi}(B) = \mathrm{vech}(\Phi(B))$, $\boldsymbol{d}(B) = \mathrm{vech}(D(B))$, $\boldsymbol{\theta}(B) = \mathrm{vech}(\Theta(B))$, $\boldsymbol{\mu} = \mathrm{vech}(M)$, and

$$\tilde{\boldsymbol{v}}_t \sim N(\boldsymbol{\mu}_{vt}, \Sigma_{vt}), \qquad \boldsymbol{\mu}_{vt} = \sqrt{\frac{2}{\pi}}\Lambda \boldsymbol{s}_t, \qquad \Sigma_{vt} = \Sigma_v - \Lambda\Lambda' + \left(1 - \frac{2}{\pi}\right)\Lambda \boldsymbol{s}_t \boldsymbol{s}_t' \Lambda'.$$

Noting that $\boldsymbol{v}_t | \boldsymbol{\varepsilon}_t \sim N(\Lambda\boldsymbol{\varepsilon}_t, \Sigma_v - \Lambda\Lambda')$, and using the properties of the half normal distribution, we can derive the above results (see Asai and McAleer (2006) for further details).

Corresponding to the asymmetric effects defined in equation (5), we consider the covariance matrix of $\tilde{\boldsymbol{\varepsilon}}_t$ and $\tilde{\boldsymbol{v}}_t$ conditional on $\boldsymbol{s}_t$. Noting that:

$$\begin{aligned}\tilde{\boldsymbol{\varepsilon}}_t = &\frac{1}{2}C_t^{1/2}\left[\nabla\Omega_t|_{h_t=\tilde{h}_t}\right]\left(\tilde{\Omega}_t^{-1} \otimes \tilde{\Omega}_t^{-1}\right)(\Omega_t \otimes \Omega_t)\,\mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' - I_m) \\ &- \frac{1}{2}C_t^{1/2}\left[\nabla\Omega_t|_{h_t=\tilde{h}_t}\right]'\mathrm{vec}(\tilde{\Omega}_t - \Omega_t) + C_t^{-1/2}(\tilde{\xi}_t - \xi_t),\end{aligned}$$

we need to consider the covariance of $\mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' - I_m)$ and $\tilde{\boldsymbol{v}}_t$ conditional on $\boldsymbol{s}_t$. As $\boldsymbol{v}_t|\boldsymbol{\varepsilon}_t \sim N(\Lambda\boldsymbol{\varepsilon}_t, \Sigma_v - \Lambda\Lambda')$, we can derive the covariance matrix, by using the first three moments of the half normal distribution:

$$E|\varepsilon_{it}| = \sqrt{\frac{2}{\pi}}, \quad E|\varepsilon_{it}|^2 = 1, \quad E|\varepsilon_{it}|^3 = 2\sqrt{\frac{2}{\pi}}, \quad (i = 1, \dots, m).$$

See Elandt (1961) for the moments of the half normal distribution. Then we have:

$$E[\boldsymbol{v}_t \mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' - I_m)'|\boldsymbol{s}_t] = \Lambda \left[G^+ \circ \left(\boldsymbol{s}_t \mathrm{vec}(\boldsymbol{s}_t\boldsymbol{s}_t')'\right) - \sqrt{\frac{2}{\pi}}\boldsymbol{s}_t \mathrm{vec}(I_m)' \right],$$

where $G^+ = E[\boldsymbol{\varepsilon}_t^+ \mathrm{vec}\left((\boldsymbol{\varepsilon}_t^+)(\boldsymbol{\varepsilon}_t^+)\right)']$ and $\boldsymbol{\varepsilon}_t^+ = (|\varepsilon_{1t}|, \dots, |\varepsilon_{mt}|)'$. Neglecting the differences in $\tilde{\Omega}_t$ and Ω_t, we can approximate the covariance of $\tilde{\boldsymbol{v}}_t$ and $\tilde{\boldsymbol{\varepsilon}}_t$ as:

$$\Lambda_t = \frac{1}{2}\Lambda \left[G^+ \circ \left(\boldsymbol{s}_t \mathrm{vec}(\boldsymbol{s}_t\boldsymbol{s}_t')'\right) - \sqrt{\frac{2}{\pi}}\boldsymbol{s}_t \mathrm{vec}(I_m)' \right] \left[\nabla\Omega_t|_{h_t = \tilde{h}_t} \right] C_t^{1/2}. \tag{17}$$

As a result, we can obtain $(\tilde{\boldsymbol{\varepsilon}}_t', \boldsymbol{u}_t', \tilde{\boldsymbol{v}}_t')' \sim N(\boldsymbol{\mu}_t, \Sigma_t)$, where:

$$\boldsymbol{\mu}_t = \begin{pmatrix} 0 \\ 0 \\ \boldsymbol{\mu}_{vt} \end{pmatrix}, \qquad \Sigma_t = \begin{pmatrix} C_t & O & \Lambda_t' \\ O & \Sigma_u & O \\ \Lambda_t & O & \Sigma_{vt} \end{pmatrix}.$$

We use the candidate density in equations (13), (15)–(17) for the importance sampling of the likelihood, as explained in the previous subsection. We may improve the approximation by considering $\partial^2 p_{1t}/\partial\boldsymbol{\xi}_t\boldsymbol{\xi}_t'$ rather than its expected value in (15), and by accommodating the effects of the differences in $\tilde{\Omega}_t$ and Ω_t in (17). Rather than considering further approximations, we will place a priority on computational convenience.

3.3 Implementation Issues

For obtaining a consistent estimate of the log-likelihood function as (12), we draw samples $\boldsymbol{H}^{(i)}$ $(i = 1, \dots, N)$ from the approximating density, $q(\boldsymbol{H}|\boldsymbol{Y}, \boldsymbol{X}; \boldsymbol{\psi})$, which is defined in equations

(13), (15)–(17). For this purpose, we use the simulation smoother of Asai and So (2016) for long memory processes with additive noise, which is an extension of de Jong and Shephard (1995) and So (1999).

The simulation smoother of Asai and So (2016) is based on the Choleski decomposition of the $(m+m^*)T \times (m+m^*)T$ matrix, $\Sigma_w = LML'$, in which the (i,j)th block is given by the sample covariance matrix of $\boldsymbol{w}_i$ and $\boldsymbol{w}_j$, where L is a block lower triangular matrix, with the (i,j)th block given by L_{ij}, and M is a block diagonal matrix, with element M_i $(i,j = 1,2,\dots,T)$. Similarly, define the Choleski decomposition of the $m^*T \times m^*T$ matrix $\Gamma = \underline{L}\,\underline{M}\,\underline{L}'$, where Γ is the covariance matrix of $(\boldsymbol{h}_1', \dots, \boldsymbol{h}_T')'$, and $\boldsymbol{h}_t$ is generated by equation (4). Given the specification, we have $\Gamma^{1/2} = \underline{L}\,\underline{M}^{1/2}$. We decompose $\Sigma_t^{1/2}$ into two matrices, and denote the first $(m+m^*) \times (2m+m^*)$ matrix and the second $m \times (2m+m^*)$ of $\Sigma_t^{1/2}$ matrix as A_t and B_t, respectively. For $\boldsymbol{h}_t$ generated by equation (16), we can express the covariance matrix between $\boldsymbol{h}_i$ and $\boldsymbol{h}_j$, conditional on the sign of $\boldsymbol{y}_{t-1}$, as:

$$\mathrm{Cov}(\boldsymbol{h}_i, \boldsymbol{h}_j) = \sum_{k=1}^{\min(i,j)} \underline{L}_{i,k} \underline{M}_k^{1/2} B_{k-1} B_{k-1}' \underline{M}_k^{1/2} \underline{L}_{j,k}'.$$

In order to implement the simulation smoother of Asai and So (2016), we first obtain the Choleski quantities, $L_{t,j}$, and the prediction errors, v_t, from the Choleski decomposition of Σ_w for $t = 1, \dots, n$ and $j = 1, \dots, t$, by the following recursive equations:

$$\begin{aligned}
a_{t+1|t} &= \sum_{j=1}^{t} \Psi_{t,j} e_{t+1-j}, \qquad e_t = \boldsymbol{w}_t - \tilde{\boldsymbol{w}}_{t|t-1}, \qquad \tilde{w}_{t|t-1} = \tilde{Z}(\boldsymbol{\mu} + \boldsymbol{\mu}_{vt}) + \tilde{Z} a_{t|t-1}, \\
F_{t+1} &= \tilde{Z} P_{t+1|t} \tilde{Z}', \qquad\qquad (18) \\
P_{t+1|t} &= \sum_{k=1}^{t+1} \underline{L}_{t+1,k} \underline{M}_k^{1/2} B_{k-1} B_{k-1}' \underline{M}_k^{1/2} \underline{L}_{t+1,k}' - \sum_{k=1}^{t} \Psi_{t,t-k-1} F_k \Psi_{t,t-k-1}',
\end{aligned}$$

where

$$\Psi_{t,j} = \left[\left(\sum_{k=1}^{t-j+1} \underline{L}_{t+1,k} \underline{M}_k^{1/2} B_{k-1} B_{k-1}' \underline{M}_k^{1/2} \underline{L}_{t+1,k}' - \sum_{k=0}^{j-1} \Psi_{t,t-k} F_{k+1} \Psi_{j,j-k}' \right) \tilde{Z}' \right.$$
$$\left. + \underline{L}_{t+1,t-j+2} \underline{M}_{t-j+2}^{1/2} B_{t-j+1} A_{t-j+1}' \right] F_{t+1-j}^{-1}.$$

Then we calculate $N_{t,s}$, $\Upsilon_{t,s}$ and Ξ_t, and draw ξ_t from $N(0, \Xi_t)$, where:

$$N_{t,t} = A_t', \quad N_{t,s} = -A_t' L_{s,t}' - \sum_{j=0}^{s-2} N_{t,j+1} L_{s,j+1}' \quad (s = t+1, \ldots, n), \tag{19}$$

$$\Upsilon_{t,s} = -N_{t,s} F_s^{-1} A_s B_s' - \sum_{k=s+1}^{n} \left\{ N_{t,k} F_k^{-1} N_{s,k}' B_s' + \Upsilon_{t,k} \Xi_k^{-1} \Upsilon_{s,k}' B_s' \right\}, \tag{20}$$

$$\Xi_t = B_t \left(I - A_t' F_t^{-1} A_t - \sum_{k=t+1}^{n} \left\{ N_{t,k} F_k^{-1} N_{t,k}' + \Upsilon_{t,k} \Xi_k^{-1} \Upsilon_{t,k}' \right\} \right) B_t'. \tag{21}$$

We evaluate $\Upsilon_{t,s}$ and Ξ_t by iterating between equations (20) and (21) according to the sequence:

$$\Xi_n; \quad \Upsilon_{n-1,n}, \Xi_{n-1}; \quad \Upsilon_{n-2,n}, \Upsilon_{n-2,n-1}, \Xi_{n-2}; \quad \Upsilon_{n-3,n}, \Upsilon_{n-3,n-1}, \Upsilon_{n-3,n-2}, \Xi_{n-3}; \quad \ldots.$$

Finally, set:

$$v_t = \xi_t + B_t \left[A_t' F_t^{-1} e_t + \sum_{s=t+1}^{n} \left\{ N_{t,s} F_s^{-1} e_s + \Upsilon_{t,s} \Xi_s^{-1} \xi_s \right\} \right],$$

and generate $\{\boldsymbol{h}_t\}$ via $(\boldsymbol{h}_1', \ldots, \boldsymbol{h}_T')' = \Gamma^{1/2} (\boldsymbol{v}_0', \ldots, \boldsymbol{v}_{T-1}')'$.

In order to construct the approximating density, we need to choose trial values, $\{\tilde{\boldsymbol{h}}_t\}$, for equation (15). Following Durbin and Koopman (1997, 2001), we apply the recursive scheme:

(i) Start at, for example, $\tilde{\boldsymbol{h}}_t = \mathbf{0}$, or the previous value of $\tilde{\boldsymbol{h}}_t$.

(ii) Solve $\{C_t\}$, and obtain $\{\tilde{y}_t\}$.

(iii) Use the above smoother without simulations to extract a new $\tilde{\boldsymbol{h}}_t$.

(iv) Return to (ii) until $\tilde{\boldsymbol{h}}_t$ converges.

Generally, we need a low (such as 5-15) number of iterations. Note that we can obtain smoothed estimates of $\{\tilde{\boldsymbol{h}}_t\}$ by setting $\xi_t = 0$ and $\Upsilon_{t,s} = O$ in the above simulation smoother. After obtaining the appropriate values of $\{\tilde{\boldsymbol{h}}_t\}$, we can draw $\{\boldsymbol{H}^{(i)}\}$ $(i = 1, \ldots, N)$. We set $N = 200$ in equation (12) to obtain a consistent estimate of the log-likelihood function.

3.4 Estimating and Forecasting Ω_t

For estimating Ω_t, we extend the approach of Sandmann and Koopman (1998). Consider the following:

$$\tilde{\Omega}_t = \bar{J}_T \hat{\boldsymbol{S}} \mathrm{Exp}(\tilde{\boldsymbol{H}}_t) \hat{\boldsymbol{S}} \bar{J}_T$$

where $\tilde{\boldsymbol{H}}_t$, is the smoothed estimate, obtained by the algorithm above, and a diagonal matrix $\bar{J}_T$ is defined by the square roots of the diagonal elements of $\bar{\Omega}_T$, given by:

$$\bar{\Omega}_T = \frac{1}{T} \sum_{t=1}^{T} \left[\hat{\boldsymbol{S}} \mathrm{Exp}\left(\tilde{H}_t \right) \hat{\boldsymbol{S}} \right]^{-1/2} \boldsymbol{y}_t \boldsymbol{y}_t' \left[\hat{\boldsymbol{S}} \mathrm{Exp}\left(\tilde{H}_t \right) \hat{\boldsymbol{S}} \right]^{-1/2}.$$

If $m = 1$, the approach reduces to that of Sandmann and Koopman (1998).

Applying the filtering technique in equation (18), we obtain the k-step-ahead forecast and its error covariance as:

$$\boldsymbol{h}_{T+s|T} = \boldsymbol{\mu}_{vT} + \sum_{j=s}^{T+s-1} \Psi_{T+s-1,j} e_{T+s-j},$$

$$P_{T+s|T} = P_{T+s|T+s-1} - \sum_{j=s}^{T+s-1} \Psi_{T+s-1,j} F_{T+j} \Psi'_{T+s-1,j}.$$

See Asai and So (2016) for further details. Then we can obtain the forecasts of Ω_t as $\Omega_{T+s|T} = \bar{J}_T \hat{\boldsymbol{S}} \mathrm{Exp}(\boldsymbol{H}_{T+s|T}) \hat{\boldsymbol{S}} \bar{J}_T$.

3.5 Two-Step Estimation

For efficiency of the estimators, it is often preferred to estimate the long memory parameter separately using the semiparametric approach, such as local Whittle (LW) estimation. For this purpose, we may use the following two-step method: (i) estimate d_{ij} using $\boldsymbol{X}$ via multivariate Gaussian semiparametric estimation, as suggested by Shimotsu (2007), which is a multivariate extension of the local Whittle (LW) estimator of Shimotsu and Phillips (2006); and (ii) estimate the remaining parameters via the MCL approach, as explained above. We apply two-step estimation for estimating the RMESV-ALM model.

4 Empirical Examples

4.1 Data

We estimate the RMESV-ALM model using daily returns and realized covariance matrices for three major stocks traded on the New York Stock Exchange, namely: Bank of America (BAC), General Electric (GE), and International Business Machines (IBM). Based on the vector of returns for the $m = 3$ stocks computed for 1-min intervals of the trading day at t between 9:30a.m. and 4:00 p.m., we calculated the daily multivariate realized kernel (RK) estimates of Barndorff-Nielsen et al. (2009). Note that the multivariate RK estimator gives a consistent estimator of the integrated covariance matrix, and is robust to microstructure noise and non-synchronous trading.

We also calculate the corresponding open-close returns for the three assets. Denote the vector of returns, and the covariance matrix estimator as $\boldsymbol{y}_t$ and $\tilde{\boldsymbol{X}}_t$, respectively. By definition, $\boldsymbol{X}_t = \mathrm{Log}\left(\tilde{\boldsymbol{X}}_t\right)$. The sample period starts at October 14, 2010, and ends on October 4, 2012, giving

500 observations.

Table 1 presents the descriptive statistics of the returns, volatilities and covolatilities. The empirical distribution of the returns is highly leptokurtic, and is heavily skewed to the left, except for GE. The stock price of GE grows rapidly from January 2009, causing the right-skewness of the empirical distribution of the return series. Regarding volatilities and co-volatilities, they are skewed to the right, with strong evidence of heavy tails in all the series.

4.2 Benchmark Model and Preliminary Results

As a benchmark model, we use a realized BEKK model with asymmetry and long memory. Instead of considering a multivariate extension of the fractionally-integrated GARCH model of Baillie, Bollerslev, and Mikkelsen (1996), we use the heterogeneousness asymmetric BEKK (HABEKK) model of Asai and McAleer (2016). The HABEKK model captures long-range dependence in the volatility matrix, as in the heterogeneous autoregressive (HAR) model of Corsi (2009) and heterogeneous ARCH model of Müller et al. (1997).

For the HABEKK model, consider the mean returns for the past h days as:

$$\left(\boldsymbol{y}_{t-1}\right)_h = h^{-1}(\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{y}_{t-h}),$$

so that we can obtain the weekly ($h = 5$) and monthly ($h = 22$) mean returns of the past $\boldsymbol{y}_t$ as $\left(\boldsymbol{y}_{t-1}\right)_5$ and $\left(\boldsymbol{y}_{t-1}\right)_{22}$, respectively. Define a negative part of the ith element of $\left(\boldsymbol{y}_{t-1}\right)_h$ as $\left(\boldsymbol{y}_{i,t-1}\right)_h^- = \left(\boldsymbol{y}_{i,t-1}\right)_h \times \mathbf{1}\left(\left(\boldsymbol{y}_{i,t-1}\right)_h < 0\right)$, where $\mathbf{1}(x < 0)$ is one if $x < 0$, and is zero otherwise.

Then we can obtain the realized HABEKK model as:

$$\begin{aligned}
&\boldsymbol{y}_t = \Omega_t^{1/2}\boldsymbol{y}_t, \quad \boldsymbol{y}_t \sim i.i.d.N(0, I_m),\\
&\tilde{\boldsymbol{X}}_t = \tilde{K} + \Omega_t + \tilde{U}_t, \quad \tilde{\boldsymbol{u}}_t = \mathrm{vech}(\tilde{U}_t) \sim i.i.d.N(0, \Sigma_{\tilde{u}}),\\
&\Omega_t = \boldsymbol{W} + \boldsymbol{A}_d \boldsymbol{y}_{t-1}\boldsymbol{y}_{t-1}' \boldsymbol{A}_d' + \boldsymbol{A}_w \left(\boldsymbol{y}_{t-1}\right)_5 \left(\boldsymbol{y}_{t-1}\right)_5' \boldsymbol{A}_w' + \boldsymbol{A}_m \left(\boldsymbol{y}_{t-1}\right)_{22} \left(\boldsymbol{y}_{t-1}\right)_{22}' \boldsymbol{A}_m'\\
&\qquad + \boldsymbol{C}_d \boldsymbol{y}_{t-1}^{-} \boldsymbol{y}_{t-1}^{-\prime} \boldsymbol{C}_d' + \boldsymbol{C}_w \left(\boldsymbol{y}_{t-1}\right)_5^{-} \left(\boldsymbol{y}_{t-1}\right)_5^{-\prime} \boldsymbol{C}_w' + \boldsymbol{C}_m \left(\boldsymbol{y}_{t-1}\right)_{22}^{-} \left(\boldsymbol{y}_{t-1}\right)_{22}^{-\prime} \boldsymbol{C}_m'\\
&\qquad + \boldsymbol{B}\Omega_{t-1}\boldsymbol{B}',
\end{aligned}$$

where $\tilde{K}$ is an $m \times m$ square matrix, W is a positive definite matrix, and $\boldsymbol{A}_i$, $\boldsymbol{C}_i$ $(i = d, w, m)$, and B are $m \times m$ matrices of parameters, with $a_{11,i} > 0$, $c_{11,i} > 0$ and $b_{11} > 0$. The (1,1) elements $a_{11,i} > 0$, $c_{11,i} > 0$ and $b_{11} > 0$ are required for identifiability. In order to reduce the number of parameters, and to ensure regularity, invertibility and asymptotic properties of the QML estimator, $\boldsymbol{A}_i$, $\boldsymbol{C}_i$, and B are assumed to be diagonal matrices (see Asai and McAleer (2016) for further details).

Table 2 shows the ML estimates of the realized HABEKK model. The estimates of $a_{ii,d}$ and b_{ii} $(i = 1, 2, 3)$ are significant at the five percent level, as in the literature of diagonal BEKK specifications. The estimates of $c_{ii,d}$ $(i = 1, 2, 3)$ are significant, indicating that a negative shock in return increases future volatility. For the parameters of the weekly effects, $a_{ii,w}$ and $c_{ii,w}$ are significant, except for the case $i = 2$. The estimates of $a_{ii,m}$ and $c_{ii,m}$ for the monthly effects are insignificant in all cases. For univariate models, Müller et al. (1997) recommend removing past values of conditional volatility to obtain significant heterogeneous-time effects for the squared residuals. In our case, the estimates of $a_{ii,m}$ and $c_{ii,m}$ are significant by setting $\boldsymbol{B} = O$, but the

Akaike information criterion (AIC) and the Bayesian information criterion (BIC) prefer the model with $\boldsymbol{B} \neq O$.

4.3 Estimation Results

Table 3 shows the estimates of the RMESV-ALM model, based on the first $T = 250$ observations. Compared with the realized HABEKK model, the RMESV-ALM model has smaller AIC and BIC. The LW estimates of d_{ij} are significant at the five percent level, indicating that all the elements of $\boldsymbol{X}_t$ and $\boldsymbol{H}_t$ follow long memory processes. The estimates of d_{ij} are significant, and are located between 0.26 to 0.41.

All the estimates of ϕ_{ij} and κ_{ij} are positive and significant. For the parameters of the concentrated leverage specification (6), the estimates of ρ_i are negative and significant, indicating the existence of a leverage effect. For the constant part of the dynamic covariances, the estimates of s_{ij} are significant, except for s_{32}. For the dynamic spillover effects, q_{61}, q_{62} and q_{i3} $(i = 1, \ldots, 6)$ are significant.

Table 3 also presents the results of the likelihood ratio tests for constant and dynamic spillover effects, with the empirical results indicating the existence of the dynamic and constant spillover effects.

We compare the out-of-sample forecasting performance of the RMESV-ALM model with that of realized HABEKK, using the last $T_2 = 250$ observations. We use $\hat{\Omega}_{T_1+h} = \mathrm{Exp}(\boldsymbol{X}_{T_1+h})$ as a proxy for the unobservable Ω_{T_1+h} ($T_1 = 250$, $h = 1, \ldots, T_2$), and define the forecast error matrix as:

$$E_{T_1+h} = \hat{\Omega}_{T_1+h} - \tilde{\Omega}_{T_1+h|T_1+h-1},$$

where $\tilde{\Omega}_{T_1+h|T_1+h-1}$ is the one-step-ahead forecast of Ω_{T_1+h}. Following Chiriac and Voev (2011), we compare the out-of-sample forecast root-mean-squared errors (RMSE) of the two models based on the Frobenius norm of the forecast error, which is defined by:

$$FN = \frac{1}{T_2}\sum_{h=1}^{T_2} ||E_{T_1+h}|| = \frac{1}{T_2}\sum_{h=1}^{T_2}\left[\sum_{i,j} e^2_{ij,T_1+h}\right]^{1/2}. \tag{22}$$

We also calculate a measure based on the Wishart distribution:

$$WM = \frac{1}{T_2}\sum_{h=1}^{T_2}\left[\text{tr}\left(\hat{\Omega}^{-1}_{T_1+h}\tilde{\Omega}_{T_1+h|T_1+h-1}\right)\right]^{1/2}. \tag{23}$$

Table 4 shows the results of the forecasting performance. Compared with the realized HABEKK specification, the RMESV-ALM model gives smaller FN and WM values, indicating a better performance for the latter model.

We examine the news impact curves (NICs) of the standard deviations and correlations of $\boldsymbol{y}_{t+1}$ by a shock in $\boldsymbol{y}_t$, using Exp($\boldsymbol{H}_{t+1}$) in equation (7). Figure 1 shows the NICs from the ith return on the one-step-ahead standard deviation of j (i, j =BAC,GE,IBM). The horizontal axis indicates the values of annualized returns, while the vertical axis gives the values of annualized one-step-ahead standard deviations. We adjusted the curves so that they take zero at the origin. As implied by the estimates of ρ_i, the NICs from the return to its own future volatility indicate a negative relation between return and future volatility for BAC, GE, and IBM. For the spillover effects, the effects from the GE returns are greater than those of the GE one-step-ahead standard deviations. The spillover effects between BAC and GE are negligible.

Figure 2 also shows the news impact from a return to the one-step-ahead correlation coefficients. Figure 2 indicates that the impacts on the dynamic correlations change, depending on the signs and magnitudes of the returns. The range of the change is 0.2, at most. It should be noted that changes in a return can affect the correlation coefficient between the other two stocks. For example, a positive return of IBM decreases the correlation between BAC and GE for the period.

Figure 3 illustrates the impulse response functions (IRFs) from returns to future volatilities. We use -10 percent of the annual return as the impulse at time zero in order to draw responses at time h ($h = 1, \ldots, 200$). A shock on the ith return ($i =$BAC,GE,IBM) to its own future standard deviation decreases slowly toward its corresponding mean, $\{\boldsymbol{S}^2\}_{ii}$, as a result of the long range dependence in $\boldsymbol{H}_t$.

Figure 3 also indicates the spillover effects on the IRFs. For example, a negative impulse in the return of BAC produces an undershooting (overshooting) response in the future standard deviations of GE (IBM).

5 Concluding Remarks

The paper developed a novel realized matrix-exponential stochastic volatility model of multivariate returns and realized covariances that incorporated asymmetry and long memory (hereafter the RMESV-ALM model), as well as new tests for volatility and co-volatility spillovers. The matrix-exponential transformation guarantees the positive definiteness of the dynamic covariance matrix. Efficient importance sampling was used to maximize the likelihood function of RMESV-ALM.

Using high frequency data for three major US financial assets, the RMESV-ALM model was estimated, and compared with the novel realized HABEKK model via in-sample model-fitness and out-of-sample forecasting performance. The empirical results suggested the RMESV-ALM specification to be superior. The news impact curve and impulse response functions were also estimated, and spillover effects were found from returns to the remaining volatilities and correlation dynamics.

As shown in the paper, it is useful to consider the matrix-exponential specification using returns and realized volatility measures simultaneously. The theoretical developments may be extended by developing a general asymmetric function which accommodates both spillover effects and the quadratic form of the standardized residuals of returns, based on the alternative SV models developed in Asai and McAleer (2011). The Bayesian Markov Chain Monte Carlo (MCMC) technique may also be developed for estimating the RMESV-ALM model by extending the work of Ishihara, Omori and Asai (2016).

References

Asai, M. and M. McAleer (2006), "Asymmetric Multivariate Stochastic Volatility", *Econometric Reviews*, **25**, 453–473.

Asai, M. and M. McAleer (2009a), "The Structure of Dynamic Correlations in Multivariate Stochastic Volatility Models", *Journal of Econometrics*, **150**, 182–192.

Asai, M. and M. McAleer (2009b), "Multivariate Stochastic Volatility, Leverage and News Impact Surfaces", *Econometrics Journal*, **12**, 292–309.

Asai, M. and M. McAleer (2011), "Alternative Asymmetric Stochastic Volatility Models" *Econometric Reviews*, **30**, 548–564.

Asai, M. and M. McAleer (2015), "Forecasting Co-Volatilities via Factor Models with Asymmetry and Long Memory in Realized Covariance", *Journal of Econometrics*, **189**, 251–262.

Asai, M. and M. McAleer (2016), "Asymptotic Theory for Extended Asymmetric Multivariate GARCH Processes", Tinbergen Institute DP 16-071/III, Amsterdam and Rotterdam, The Netherlands.

Asai, M., M. McAleer and M. Medeiros (2012a), "Estimation and Forecasting with Noisy Realized Volatility", *Computational Statistics & Data Analysis*, **56**, 217–230.

Asai, M., M. McAleer and M. Medeiros (2012b), "Asymmetry and Long Memory in Volatility Modeling", *Journal of Financial Econometrics*, **10**, 495–512.

Asai, M., M. McAleer, and J. Yu (2006), "Multivariate Stochastic Volatility: A Review", *Econometric Reviews*, **25**, 145–175.

Asai, M., and M.K.P. So (2016), "A Simulation Smoother for Long Memory Time Series with Additive Noise", Unpublished Paper, Soka University.

Baillie R.T., T. Bollerslev and H.O. Mikkelsen (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, **74**, 3–30.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde, and N. Shephard (2008), "Designing Realised Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise", *Econometrica*, **76**, 1481–1536.

Barndorff-Nielsen, O.E., and N. Shephard (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models", *Journal of the Royal Statistical Society, Series B*, **64**, 253–280.

Basmann, R. (1988), "Causality Tests and Observationally Equivalent Representations of Econometric Models", *Journal of Econometrics*, **39**, 69–104.

Bauer, G.H. and K. Vorkink (2011), "Forecasting Multivariate Realized Stock Market Volatility", *Journal of Econometrics*, **160**, 93–101.

Bollerslev, T. and H.O. Mikkelsen (1996), "Modeling and Pricing Long-Memory in Stock Market Volatility", *Journal of Econometrics*, **73**, 151–184.

Bollerslev, T., and H. Zhou (2002), "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility", *Journal of Econometrics*, **109**, 33–65.

Breidt, F.J., N. Crato and P. de Lima (1998), "The Detection and Estimation of Long Memory", *Journal of Econometrics*, **83**, 325–348.

Chesney, M., and L.O. Scott (1989), "Pricing European Currency Options: A Comparison of The Modified Black-Scholes Model and A Random Variance model", *Journal of Financial and Quantitative Analysis*, **24**, 267–84.

Chiriac, R., and V. Voev (2011), "Modelling and Forecasting Multivariate Realized Volatility", *Journal of Applied Econometrics*, **26**, 922–947.

Chiu, T.Y.M., T. Leonard and K.W. Tsui (1996), "The Matrix-Logarithmic Covariance Model", *Journal of the American Statistical Association*, **91**, 198–210.

Corsi, F. (2009), "A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics*, **7**, 174–196.

Durbin, J., and S.J. Koopman (1997), "Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models", *Biometrika*, **84**, 669-684.

Durbin, J., and S.J. Koopman (2001), *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford.

Elandt, R. C. (1961), "The Folded Normal Distribution: Two Methods of Estimating Parameters from Moments", *Technometrics*, **3**, 551–562.

Engle, R.F. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models", *Journal of Business & Economic Statistics*, **20**, 339–350.

Engle, R. F., and G. Gallo (2006), "A Multiple Indicators Model for Volatility Using Intra-Daily Data", *Journal of Econometrics*, **131**, 3–27.

Engle, R.F. and K.F. Kroner (1995), "Multivariate Simultaneous Generalized ARCH", *Econometric Theory*, **11**, 122–150.

Gourieroux, C., J. Jasiak, and R. Sufana (2009), "The Wishart Autoregressive Process of Multivariate Stochastic Volatility", *Journal of Econometrics*, **150**, 167–181.

Harvey, A. (1998), "Long Memory in Stochastic Volatility", In: Knight, J. and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, Oxford: Butterworth-Haineman, 307–320.

Hansen, P.R., Z. Huang, and H.H. Shek (2012), "Realized GARCH: A Complete Model of Returns and Realized Measures of Volatility", *Journal of Applied Econometrics*, **27**, 877–906.

Hansen, P.R., and Z. Huang (2016), "Exponential GARCH Modeling with Realized Measures of Volatility", *Journal of Business & Economic Statistics*, **34**, 269–287.

Harvey, A.C., and N. Shephard (1996), "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns", *Journal of Business and Economic Statistics*, **14**, 429–434.

Ishihara, T., Y. Omori and M. Asai (2016), "Matrix Exponential Stochastic Volatility with Cross Leverage", *Computational Statistics & Data Analysis*, **100**, 331–350.

de Jong, P., and N. Shephard (1995), "The simulation smoother for time series models", *Biometrika*, **82**, 339–50.

Kawakatsu, H. (2006), "Matrix Exponential GARCH", *Journal of Econometrics*, **134**, 95–128.

Loddo, A., S. Ni, and D. Sun (2011), "Selection of Multivariate Stochastic Volatility Models via Bayesian Stochastic Search", *Journal of Business & Economic Statistics*, **29**, 342–55.

Martinet, G. and M. McAleer (2016), "On The Invertibility of EGARCH(p,q)", to appear in *Econometric Reviews*.

McAleer, M. and C. Hafner (2014), "A One Line Derivation of EGARCH", *Econometrics*, **2**, 92–97.

Müller, U., M. Dacorogna, R. Dav, R. Olsen, O. Pictet, and J. von Weizsacker (1997), "Volatilities of Different Time Resolutions: Analysing The Dynamics of Market Components", *Journal of Empirical Finance*, **4**, 213–239.

Pérez, A. and E. Ruiz (2001), "Finite Sample Properties of a QML Estimator of Stochastic Volatility Models with Long Memory", *Economics Letters*, **70**, 157–164.

Sandmann, G., and S. J. Koopman (1998), "Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood", *Journal of Econometrics*, **87**, 271–301.

Shephard, N, and K. Sheppard (2010), "Realising The Future: Forecasting with High Frequency Based Volatility (HEAVY) Models", *Journal of Applied Econometrics*, **25**, 197–231.

Shimotsu, K. (2007), "Gaussian Semiparametric Estimation of Multivariate Fractionally Integrated Processes", *Journal of Econometrics*, **137**, 277–310.

Shimotsu, K. and P.C.B. Phillips (2006), "Local Whittle Estimation of Fractional Integration and Some of Its Variants", *Journal of Econometrics*, **130**, 209–233.

So, M.K.P. (1999), "Time Series with Additive Noise", *Biometrika*, **86**, 474–82.

So, M.K.P. (2002), "Bayesian Analysis of Long Memory Stochastic Volatility Models", *Sankhyā*, **24**, 1–10.

So, M.K.P. and S.W.Y. Kwok (2006), "A Multivariate Long Memory Stochastic Volatility Model", *Physica A*, **362**, 450–464.

Takahashi, M., Y. Omori, and T. Watanabe (2009), "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously", *Computational Statistics & Data Analysis*, **53**, 2404–2426.

Waggoner, D.F. and T. Zha (2003), "Likelihood Preserving Normalization in Multiple Equation Models", *Journal of Econometrics*, **114**, 329–347.

Wiggins, J.B. (1987), "Option Values under Stochastic Volatility: Theory and Empirical Estimates", *Journal of Financial Economics*, **19**, 351–72.

Table 1: Descriptive Statistics for Returns, Realized Volatilities and Co-Volatilities

Stock	Mean	Min	Max	Std.Dev.	Skew.	Kurt.
Returns						
BAC	−0.2105	−17.737	12.493	2.8161	−0.7993	8.2415
GE	0.1334	−13.334	14.458	2.7201	0.2109	8.3415
IBM	−0.3329	−25.378	20.080	4.1100	−0.3400	10.916
Volatilities						
BAC	10.550	1.1550	403.48	22.278	10.177	147.72
GE	8.3140	0.4397	334.09	19.300	7.9431	103.64
IBM	16.733	0.7791	688.46	41.977	7.8817	93.775
Co-volatilities						
BAC-GE	3.3653	−0.1901	98.200	6.7915	6.0948	61.548
BAC-IBM	2.2917	−24.138	81.676	6.9318	5.1632	43.130
GE-IBM	2.8881	−9.4994	84.010	8.3987	4.4224	28.214

Note: The number of observations for each series is 500.

Table 2: Estimates of the Realized Diagonal HABEKK Model

Parameters	$i=1$		$i=2$		$i=3$	
$a_{ii,d}$	0.3314	(0.1208)	0.3805	(0.0609)	0.5150	(0.1007)
$a_{ii,w}$	0.3385	(0.1252)	0.3279	(0.1664)	0.3913	(0.1043)
$a_{ii,m}$	0.3308	(0.4948)	0.3295	(0.5686)	0.3568	(0.2681)
$c_{ii,d}$	0.3031	(0.0844)	0.3625	(0.0637)	0.4558	(0.1378)
$c_{ii,w}$	0.3274	(0.1357)	0.3414	(0.1920)	0.3975	(0.1220)
$c_{ii,m}$	0.3314	(0.5150)	0.3289	(1.8452)	0.3547	(0.2793)
b_{ii}	0.3499	(0.1597)	0.3551	(0.0623)	0.4567	(0.0505)
LogLike	-12331.1					
AIC	24758.4					
BIC	25192.4					

Note: Standard errors are given in parentheses. We have omitted the estimates of Σ_u and W to save space.

Table 3: Estimates of the RMESV-ALM Model

(i,j)	d_{ij}	ϕ_{ij}	κ_{ij}	s_{ij}
(1,1)	0.2638	0.7353	1.7205	1.9292
	(0.0239)	(0.0867)	(0.0274)	(0.0553)
(2,1)	0.3967	0.4800	0.3630	0.3975
	(0.0179)	(0.4107)	(0.0052)	(0.0342)
(3,1)	0.2833	0.6717	0.0890	0.2053
	(0.0080)	(0.1102)	(0.0050)	(0.0703)
(2,2)	0.3805	0.4828	1.1192	1.5386
	(0.0217)	(0.1488)	(0.0362)	(0.0412)
(3,2)	0.4081	0.4996	0.1394	0.0186
	(0.0223)	(0.2397)	(0.0064)	(0.0425)
(3,3)	0.2967	0.5302	1.9762	2.5650
	(0.0225)	(0.0886)	(0.0340)	(0.0863)
(i,j)	q_{i1}	q_{i2}	q_{i3}	ρ_i
(1,j)	1	-0.0421	-0.6415	-0.1165
		(0.1055)	(0.0431)	(0.0209)
(2,j)	0.1101	-0.0367	0.1344	-0.1105
	(0.0673)	(0.0366)	(0.0101)	(0.0167)
(3,j)	-0.0650	0.0762	0.2948	-0.2090
	(0.0903)	(0.0462)	(0.0221)	(0.0063)
(4,j)	0.1966	1	0.4078	
	(0.3395)		(0.0498)	
(5,j)	0.3870	0.1765	0.4078	
	(0.3034)	(0.1324)	(0.0439)	
(6,j)	0.4078	-0.5243	1	
	(0.1125)	(0.0557)		
LogLike	-2865.7			
AIC	5863.4			
BIC	6643.3			
LR_{cs}	218.38	[0.0000]		
LR_{ds}	70.343	[0.0000]		

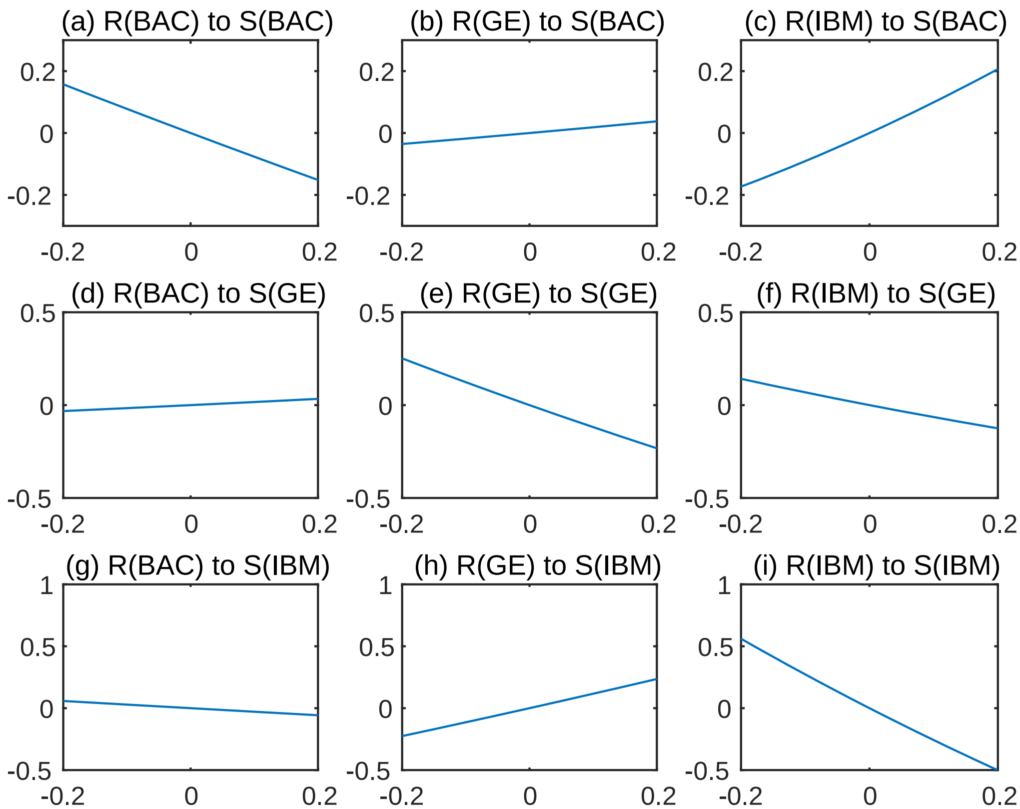
Note: The entries show the MCL estimates, except for the long memory parameter, d_{ij}, which are the LW estimates. Standard errors are given in parentheses. LR_{cs} and LR_{ds} denote the likelihood ratio tests for constant and dynamic spillover effects, respectively. P-values are given in brackets. We have omitted the estimates of Σ_u and Σ_v to save space.

Table 4: Results for Out-of-Sample Forecasting Performance

Model	FN	WM
MER-GARCH	4.7256	3.2324
MERSV-ALM	2.4580	2.5751

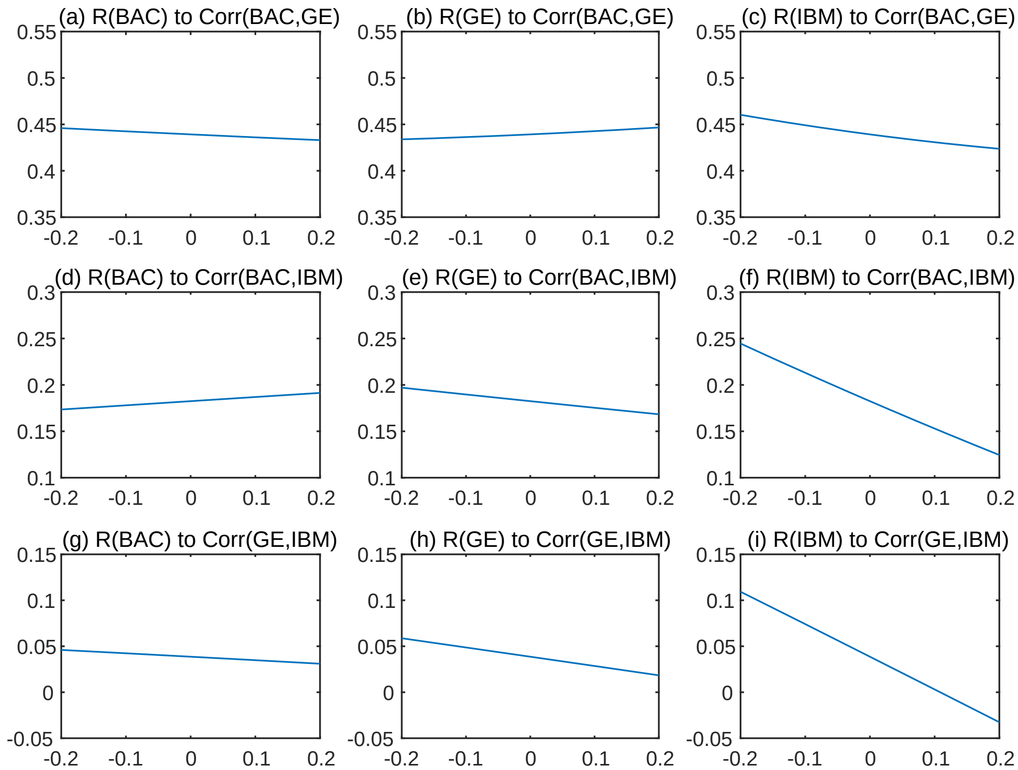
Note: FN denotes the RMSE based on the Frobenius norm of the forecasting error (22), while WM is the Wishart-type measure defined by equation (23).

Figure 1: News Impact Curves for the RMESV-ALM Model



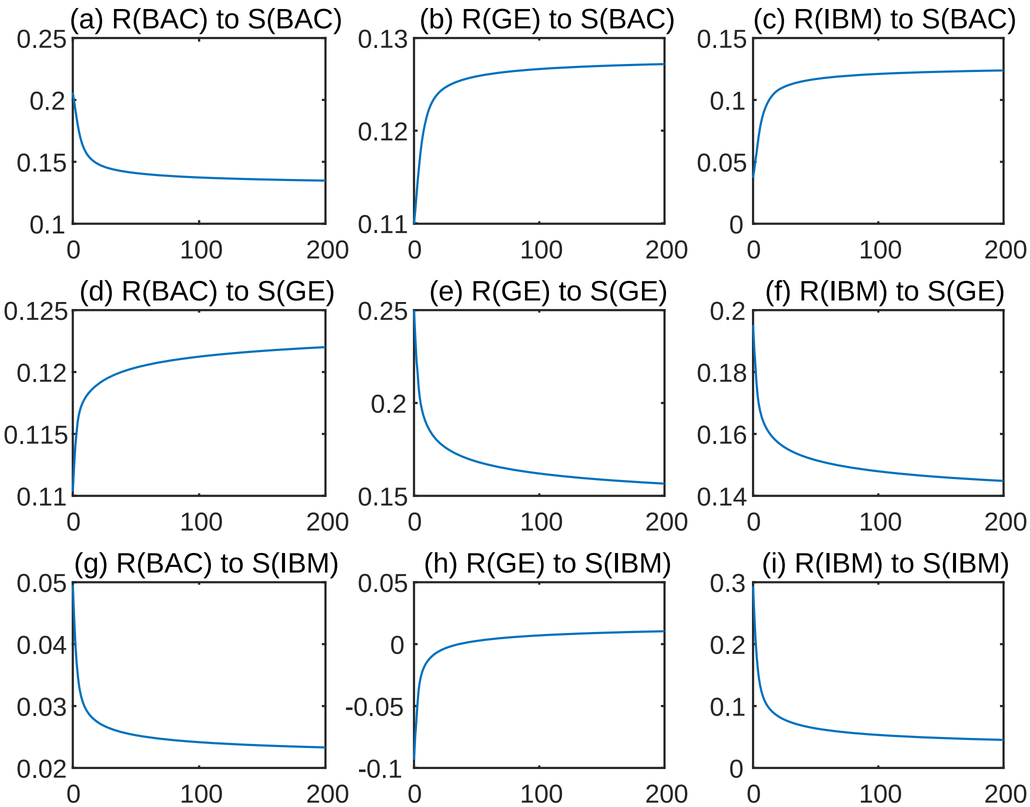
Note: R(i) and S(i) ($i =$ BAC,GE,IBM) indicate the i-th return and its one-step-ahead standard deviation, respectively. The horizontal axis indicates the value of the annualized return, while the vertical axis gives the value of the annualized standard deviation. We adjusted the curves so that they are zero at the origin.

Figure 2: News Impacts on Correlation Dynamics for the RMESV-ALM Model



Note: R(i) and Corr(i,j) (i,j = BAC,GE,IBM) indicate the return and one-step-ahead correlation coefficient between the i-th and j-th returns, respectively. The horizontal axis indicates the value of the annualized return, while the vertical axis gives the value of the correlation coefficient.

Figure 3: Impulse Response Functions from Negative Returns to Future Volatility for the RMESV-ALM Model



Note: R(i) and S(i) (i = BAC,GE,IBM) indicate the i-th return and its future standard deviation, respectively. We use -10 percent of annual return as the impulse at time zero in order to draw responses at time h ($h = 1, \ldots, 100$). The horizontal axis represents time h, while the vertical axis gives the value of the annualized standard deviation.