

**DEPARTMENT OF ECONOMICS AND FINANCE  
COLLEGE OF BUSINESS AND ECONOMICS  
UNIVERSITY OF CANTERBURY  
CHRISTCHURCH, NEW ZEALAND**

**Forecasting Co-Volatilities via Factor Models with  
Asymmetry and Long Memory in Realized Covariance**

**Manabu Asai  
Michael McAleer**

***WORKING PAPER***

**No. 10/2014**

**Department of Economics and Finance  
College of Business and Economics  
University of Canterbury  
Private Bag 4800, Christchurch  
New Zealand**

## WORKING PAPER No. 10/2014

### Forecasting Co-Volatilities via Factor Models with Asymmetry and Long Memory in Realized Covariance

Manabu Asai<sup>1\*</sup>  
Michael McAleer<sup>2</sup>

March 17, 2014

**Abstract:** Modelling covariance structures is known to suffer from the curse of dimensionality. In order to avoid this problem for forecasting, the authors propose a new factor multivariate stochastic volatility (fMSV) model for realized covariance measures that accommodates asymmetry and long memory. Using the basic structure of the fMSV model, the authors extend the dynamic correlation MSV model, the conditional/stochastic Wishart autoregressive models, the matrix-exponential MSV model, and the Cholesky MSV model. Empirical results for 7 financial asset returns for US stock returns indicate that the new fMSV models outperform existing dynamic conditional correlation models for forecasting future covariances. Among the new fMSV models, the Cholesky MSV model with long memory and asymmetry shows stable and better forecasting performance for one-day, five-day and ten-day horizons in the periods before, during and after the global financial crisis.

**Keywords:** Dimension reduction; Factor Model; Multivariate Stochastic Volatility; Leverage Effects; Long Memory; Realized Volatility.

**JEL Classifications:** C32, C53, C58, G17

**Acknowledgements:** The authors are most grateful to Yoshi Baba 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 author is most grateful for the financial support of the Australian Research Council, National Science Council, Taiwan, and the Japan Society for the Promotion of Science.

<sup>1</sup> Faculty of Economics, Soka University, Japan

<sup>2</sup> Department of Quantitative Finance, National Tsing Hua University, Taiwan; Econometric Institute, Erasmus School of Economics, Erasmus University, Rotterdam and Tinbergen Institute, The Netherlands; and Department of Quantitative Economics, Complutense University of Madrid, Spain.

\*Corresponding Author: Manabu Asai, Faculty of Economics, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan. Email address: m-asai@soka.ac.jp.

# 1 Introduction

Modeling and forecasting covariance matrices of financial asset returns has been attracting interests especially for obtaining optimal portfolio of assets. Two major issues are to guarantee the positive definiteness of covariance matrices, and to reduce the number of parameters which grows with the order of squares of the number of assets. There are three classes in the fields, namely the multivariate autoregressive conditional heteroskedasticity (ARCH) class, the multivariate stochastic volatility (MSV) models, and the models of realized covariance matrix.

Regarding the multivariate ARCH family, the dynamic conditional correlation (DCC) model of Engle (2002) is most popular, and its variants are still developed in the literature. Engle (2002) suggested a structure for time-varying correlation matrices, working with a univariate ARCH class model for each conditional variance. As the estimation of the conditional variance model is conducted separately, it concentrates on the estimation of the dynamic correlations. By its structure, the DCC model reduces the number of parameters. Also with simple restrictions, Engle (2002) showed that it can guarantee the positive definiteness of the correlation matrix. Apart from the DCC model, it is worth noting the diagonal ARCH model of Ding and Engle (2001), as it enables us to estimate the model based on the estimates of unconditional covariance matrix, which can be obtained by the average of the outer-products of mean-subtracted return vectors. The approach is known as the ‘variance targeting’, and it largely decreases the number of parameters for estimating in a single step. See the survey papers of McAleer (2005), Bauwens, Laurent, and Rombouts (2006) and Silvennoinen and Teräsvirta (2009) for the models of univariate and multivariate ARCH family.

In the class of the MSV models, volatilities (variances) of stock return vector are assumed to be unobservable. Harvey, Ruiz and Shephard (1990) suggested basic MSV models with the

constant correlation matrix. As the volatilities are latent variables, it is hard to obtain the likelihood function analytically, and hence there are several techniques for estimating SV models, including the Markov chain Monte Carlo method for Jacquire, Polson and Rossi (1994) and Chib, Nardari and Shephard (2006), the efficient method of moments of Gallant and Tauchen (1996), and the Monte Carlo likelihood method proposed by Durbin and Koopman (1997), Sandmann and Koopman (1998) and Asai and McAleer (2006). See the survey papers for Asai, McAleer and Yu (2006) and Chib, Omori and Asai (2009) for alternative MSV models.

Recent developments on estimating realized co-volatility (covariance) measure of asset returns enable us to obtain consistent estimators for the unobservable volatility and co-volatilities, which can be used for estimating univariate/multivariate conditional/stochastic volatility models. In univariate SV case, Barndorff-Nielsen and Shephard (2002) and Bollerslev and Zhou (2002) suggested to use realized volatilities for estimating SV models, as the realized volatility can be considered as the true volatility with the microstructure noise. See Gouriéroux Jasiak, and Sufana (2009), Chiriac and Voev (2011), Bauer and Vorkink (2011), Tao et al. (2011) and Golosnoy, Gribisch, and Liesenfeld (2012) for such applications in multivariate context.

As in Engle (2002) and Asai and McAleer (2009a), working with dynamic correlations is useful for simplifying structures of multivariate volatility models. An alternative approach for approximating the dynamic covariance structures is to use factors as in the papers of Diebold and Nerlove (1989), Harvey, Ruiz, and Shephard (1994), Vrontos, Dellaportas and Politis (2003), Chib, Nardari, and Shephard (2006) and Philipov and Glickman (2006a) and Lanne and Saikkonen (2007).

Long memory and asymmetric effects are important features for analysis on financial time series. Regarding the long memory, Ray and Tsay (2000) introduced long range dependence into

the volatility factor model by Harvey, Ruiz, and Shephard (1994), although these papers assume the constant correlation matrix. Without any factor specifications, Chiriac and Voev (2011) suggested a long memory model, using the Cholesky decomposition of realized covariance matrices. Our factor models will involve long memory in time-varying covariances/correlations. It is worth noting that the heterogeneous autoregressive (HAR) model of Corsi (2009) and its variants are often used for approximating longer memory processes than the first order autoregressive term. See Corsi and Renò (2010), Martens, van Dijk, and de Pooter (2009), Bollerslev, Sizova, and Tauchen (2011), Chiriac and Voev (2011) and Golosnoy, Gribisch, and Liesenfeld (2012). We will also consider such HAR terms.

Turning to asymmetric effects, there are several papers that developed the approaches to incorporate negative correlation between stock returns and future volatility in the multivariate volatility models. Kroner and Ng (1998) and Kawakatsu (2006) suggested asymmetric multivariate GARCH models, while Cappiello, Engle and Sheppard (2006), Asai (2013) and Asai and So (2014) proposed asymmetric dynamic correlation models. For MSV models, Chan, Kohn, and Kirby (2006) and Asai and McAleer (2006, 2009b), Chib, Omori and Asai (2009), So and Choi (2009), Ishihara and Omori (2012), Ishihara, Omori and Asai (2013) incorporated asymmetric effects. However, there have been still no previous works on the multivariate volatility models which accommodate all of factor specifications, dynamic correlations, and asymmetric effects. The current paper will fulfill the gap.

In this paper, we will consider alternative factor MSV (fMSV) models with long memory and asymmetric effects, using realized volatilities and co-volatilities. Starting from a conventional factor model, we derive an alternative representation which decomposes the covariance matrix into the matrices of common factors and idiosyncratic errors orthogonally. Based on the decomposition,

we consider several specifications of the factor covariance matrix based on the dynamic correlation model of Asai and McAleer (2009a), the conditional autoregressive Wishart model of Golosnoy, Gribisch, and Liesenfeld (2012), the Wishart autoregressive model of Gouriéroux, Jasiak and Sufana (2009), the matrix-exponential MSV model of Ishihara, Omori and Asai (2013), and the Cholesky MSV model of Chiriac and Voev (2011). We should note that Chiriac and Voev (2011) and Golosnoy, Gribisch, and Liesenfeld (2012) used the HAR term, and Chiriac and Voev (2011) considered a long memory model for their specification. In addition to these works, we extend all models by incorporating asymmetric effects and long memory (or HAR terms).

The remainder of the paper is organized as follows. Section 2 derives a new representation of the fMSV model, and discusses the specifications of asymmetric effects and long memory. Section 2 also shows the estimation procedure to obtain consistent estimators of common factors and their covariance matrices, based on the alternative form. Section 3 suggests alternative specifications for the fMSV models, and Section 4 gives an empirical example for seven stocks traded at the New York Stock Exchange.

In the following, for any positive definite matrix  $X$  and real number  $\alpha$ , we define  $X^\alpha$  by the spectral decomposition. For example,  $X^{1/2}X^{1/2} = X$ .

## 2 Asymmetry and Long Memory for Factor MSV Models

### 2.1 Factor Multivariate SV Models

Let  $\mathbf{y}_t$  be  $m \times 1$  vector of asset returns at time  $t$ . Consider the factor model

$$\mathbf{y}_t = \mathbf{B}\mathbf{f}_t + \mathbf{u}_t, \tag{1}$$

where  $\mathbf{f}_t$  is a  $k \times 1$  vector of common factors,  $\mathbf{B}$  is the  $m \times k$  matrix of factor loading, and  $\mathbf{u}_t$  is the vector of idiosyncratic errors, which are uncorrelated with  $\mathbf{f}_t$ . We can write  $E(\mathbf{u}_t|\mathbf{f}_t, \mathfrak{S}_{t-1}) = \mathbf{0}$

and  $Var(\mathbf{u}_t|\mathbf{f}_t, \mathfrak{F}_{t-1}) = \boldsymbol{\Sigma}_u$ , where  $\mathfrak{F}_t$  is the information set up to time  $t$ . We assume that the vector of common factors has the multivariate SV as

$$\mathbf{f}_t = (\boldsymbol{\Omega}_t^\dagger)^{1/2} \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim IID_k(\mathbf{0}, I_k), \quad (2)$$

where  $\boldsymbol{\Omega}_t^\dagger$  is the  $k \times k$  positive definite matrix. Hence,  $\boldsymbol{\Omega}_t^\dagger = Var(\mathbf{f}_t|\boldsymbol{\Omega}_t^\dagger)$ . Conditional on  $\boldsymbol{\Omega}_t^\dagger$ , the  $m \times m$  covariance matrix of  $\mathbf{y}_t$  is given by

$$\boldsymbol{\Sigma}_t = \mathbf{B}\boldsymbol{\Omega}_t^\dagger\mathbf{B}' + \boldsymbol{\Sigma}_u. \quad (3)$$

We will call the model (1)-(3) as the ‘factor MSV’ (fMSV) model.

For various fMSV models, Pitt and Shephard (1999), Liesenfeld and Richard (2003), and Chib, Nardari and Shephard (2006) among others restricted the factor loading matrix  $\mathbf{B}$  such that  $b_{ij} = 0$  for  $(i < j, i \leq k)$  and  $b_{ii} = 1$  ( $i \leq k$ ) for identification purpose. We consider an alternative specification.

There is no loss in generality in assuming  $\mathbf{B}$  in (1) satisfying the condition  $\mathbf{B}'\mathbf{B} = I_k$ . Then there exists an  $m \times (m - k)$  matrix  $\mathbf{A}$  for which  $\mathbf{A}'\mathbf{B} = \mathbf{O}$  and  $\mathbf{A}'\mathbf{A} = I_{m-k}$ . In other words,  $\mathbf{S} = [\mathbf{B} \ \mathbf{A}]$  is an  $m \times m$  orthogonal matrix, and  $\mathbf{S}'\mathbf{S} = \mathbf{S}\mathbf{S}' = I_m$ . Then we have an alternative representation of the fMSV model as

$$\mathbf{y}_t = \mathbf{B}\boldsymbol{\varphi}_t + \mathbf{A}\boldsymbol{\zeta}_t, \quad (4)$$

$$\boldsymbol{\Sigma}_t = \mathbf{B}\boldsymbol{\Omega}_t\mathbf{B}' + \mathbf{A}\boldsymbol{\Sigma}_{\zeta\zeta}\mathbf{A}', \quad (5)$$

where  $\boldsymbol{\varphi}_t = \mathbf{f}_t + \mathbf{B}'\mathbf{u}_t$ ,  $\boldsymbol{\zeta}_t = \mathbf{A}'\mathbf{u}_t$ ,  $\boldsymbol{\Omega}_t = \boldsymbol{\Omega}_t^\dagger + \mathbf{B}'\boldsymbol{\Sigma}_u\mathbf{B}$  and  $\boldsymbol{\Sigma}_{\zeta\zeta} = \mathbf{A}'\boldsymbol{\Sigma}_u\mathbf{A}$ . We should note that

$$E \left[ \begin{pmatrix} \boldsymbol{\varphi}_t \\ \boldsymbol{\zeta}_t \end{pmatrix} \middle| \boldsymbol{\Omega}_t \right] = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad Var \left[ \begin{pmatrix} \boldsymbol{\varphi}_t \\ \boldsymbol{\zeta}_t \end{pmatrix} \middle| \boldsymbol{\Omega}_t \right] = \begin{pmatrix} \boldsymbol{\Omega}_t & \boldsymbol{\Sigma}_{\varphi\zeta} \\ \boldsymbol{\Sigma}_{\zeta\varphi} & \boldsymbol{\Sigma}_{\zeta\zeta} \end{pmatrix},$$

where  $\boldsymbol{\Sigma}_{\varphi\zeta} = \mathbf{B}'\boldsymbol{\Sigma}_u\mathbf{A}$ . Since  $\mathbf{A}'\mathbf{B} = \mathbf{O}$ , the equations (4) and (5) can be considered as the ‘orthogonal decomposition’ of the fMSV model (1)-(3).

There are two merits in using the orthogonal decomposition. The first one is that we can define leverage effect from return vector to future volatilities and co-volatilities by considering a relationship between  $\varphi_t$  and future values of  $\Omega_t$ . We will consider the asymmetric effects in the next subsection. The second merit is that we can obtain consistent estimators of  $\varphi_t$  and  $\Omega_t$  by the approach which will be explained later. For these reasons, we will consider modeling, estimation and forecasting of  $\{\Omega_t\}$  rather than  $\{\Omega_t^\dagger\}$  for the purpose of forecasting future values of  $\{\Sigma_t\}$ .

## 2.2 Asymmetric Effects on Factor Covariance

We specify asymmetric effects from the current returns,  $\mathbf{y}_t$ , to the one-step-ahead volatilities and co-volatilities,  $\Sigma_{t+1}$ , via the covariance of the common factors,  $\Omega_{t+1}$ . For this purpose, we work with  $k$ -vector,

$$\mathbf{z}_t = (\mathbf{B}'\Sigma_t\mathbf{B})^{-1/2}(\mathbf{B}'\mathbf{y}_t) = \Omega_t^{-1/2}\varphi_t = (\Omega_t^\dagger + \mathbf{B}'\Sigma_u\mathbf{B})^{-1/2}(\mathbf{f}_t + \mathbf{B}'\mathbf{u}_t). \quad (6)$$

By definition, we have  $E(\mathbf{z}_t|\Omega_t) = \mathbf{0}$  and  $\text{Var}(\mathbf{z}_t|\Omega_t) = I_k$ , and thus  $E(\mathbf{z}_t|\Omega_t) = \mathbf{0}$  and  $V(\mathbf{z}_t|\Omega_t) = I_k$ .

The two popular approaches for introducing asymmetry for volatility dynamics are employing threshold function, as in the GJR model of Glosten, Jagannathan and Runkle (1992), and working with an absolute value function, as in the specification of Nelson (1991). Kroner and Ng (1998) developed a multivariate extension of the GJR model, while Kawakatsu (2006) suggested an extension for the EGARCH model. Unlike these approaches, we consider the second order Hermit polynomial for the error term. Hansen, Huang and Shek (2012) suggested this approach for a univariate model. In this case, we have

$$\lambda_{it} = \lambda_{1i}z_{it} + \lambda_{2i}(z_{it}^2 - E(z_{it}^2)). \quad (7)$$

Noting that  $E(z_{it}^2) = 1$ , we also have  $E(\lambda_{it}) = 0$ ,  $\text{Var}(\lambda_{it}) = \lambda_{1i}^2 + 2\lambda_{2i}^2$  and  $E[z_{it}\lambda_{it}] = \lambda_{1i}$ , as

we assume that each element of  $z_t$  has zero mean and unite variance. Hence, if  $\lambda_{1i}$  is negative, the model describes the leverage effect. On the other hand,  $\lambda_{2i}$  is expected to be positive, as in the conventional GARCH families. We will use the vector specification,  $\boldsymbol{\lambda}_t = (\lambda_{1t}, \dots, \lambda_{kt})'$ , especially when we consider asymmetric effects on the diagonal elements of  $\boldsymbol{\Omega}_{t+1}$ .

Now we suggest the following multivariate extension of the above approach:

$$\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}_1 \circ (\mathbf{z}_t \mathbf{l}'_k + \iota_k \mathbf{z}'_t) + \boldsymbol{\Lambda}_2 \circ (\mathbf{z}_t \mathbf{z}'_t - E(\mathbf{z}_t \mathbf{z}'_t)). \quad (8)$$

Noting that  $E(\mathbf{z}_t \mathbf{z}'_t) = I_k$ , it is straightforward to show that  $E(\boldsymbol{\Lambda}_t) = O$  and that each element has finite variance. For the leverage effects,  $\lambda_{ii,1} < 0$  and  $\lambda_{ii,2} > 0$ . For the purpose of considering a positive semi-definite matrix of the multivariate asymmetric effects, we modify (8) as

$$\boldsymbol{\Lambda}_t^\dagger = \boldsymbol{\Lambda}^\dagger \circ (\mathbf{z}_t - \boldsymbol{\gamma})(\mathbf{z}_t - \boldsymbol{\gamma})' \quad (9)$$

where  $\boldsymbol{\gamma}$  is  $k$ -vector of parameters with  $\gamma_i > 0$ , and  $\boldsymbol{\Lambda}^\dagger$  is a  $k$  dimensional positive (semi-)definite matrix.

We may accommodate the asymmetric effects to the fMSV models, by introducing the error vector,  $\boldsymbol{\lambda}_t$ , the error matrix,  $\boldsymbol{\Lambda}_t$  and/or the positive semi-definite matrix,  $\boldsymbol{\Lambda}_t^\dagger$ . We will develop the new fMSV models in Section 3.

We can examine news impacts from an asset return to its future volatility, following the works of Engle and Ng (1993), Yu (2005), Asai and McAleer (2009b), Caporin and McAleer (2011) and Chen and Ghysels (2010). Engle and Ng (1993) developed the news impact curve (NIC), which is a useful tool for measuring the effects of news on the conditional variances. They showed, graphically, the asymmetric reactions of the conditional variances to positive and negative shocks of equal magnitude for the GJR model of Glosten et al. (1992) and the EGARCH model of Nelson (1991). Regarding multivariate conditional volatility models, Caporin and McAleer (2011)

developed news impact surfaces specifically for their dynamic asymmetric multivariate GARCH. In the framework of univariate SV models, Yu (2005) developed the news impact function for evaluating the effects of news on the log-volatilities as an adaptation of the NIC. Asai and McAleer (2009b) extended it to apply volatilities and log-volatilities for MSV with leverage models. When realized volatility is available, we can estimate the NIC using the MIDAS (mixed data sampling) regressions, as suggested by Chen and Ghysels (2010).

Since we accommodate factors which reduce the number of parameters, we need to check whether or not the fMSV model with asymmetric effects can approximate true NICs. We may define the news impact surface for the fMSV model as

$$\begin{aligned}\Sigma_t^{NIS}(\mathbf{y}) &= E(\Sigma_t | \mathbf{y}_{t-1} = \mathbf{y}, \Omega_{t-1} = \bar{\Omega}, \Omega_{t-2} = \bar{\Omega}, \dots) \\ &= \mathbf{B} \{ E(\Omega_t | \mathbf{y}_{t-1} = \mathbf{y}, \Omega_{t-1} = \bar{\Omega}, \Omega_{t-2} = \bar{\Omega}, \dots) \} \mathbf{B}' + \mathbf{A} \Sigma_{\zeta \zeta} \mathbf{A}'\end{aligned}$$

where  $\mathbf{y}$  is an  $m$ -vector and  $\bar{\Omega} = E(\Omega_t)$ . By definition, we can examine news impact from one asset return not only to its own future volatility but also to future volatilities of remaining assets. We will give illustrative examples in empirical analysis.

### 2.3 Long Memory for Covariance structure

For long-range dependences 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. In addition to GARCH specifications, Breidt et al. (1998), Harvey (1998), Pérez and Ruiz (2001), So (2002) and So and Kwok (2006) studied a long-memory stochastic volatility model.

Let  $L$  be the lag operator. For any  $k$ -dimensional symmetric matrix process  $\mathbf{X}_t$ , we define a

$k$ -dimensional matrix  $D(L)$  as

$$D(L) = \begin{pmatrix} (1-L)^{d_{11}} & (1-L)^{d_{12}} & \dots & (1-L)^{d_{1k}} \\ (1-L)^{d_{21}} & (1-L)^{d_{22}} & \dots & (1-L)^{d_{2k}} \\ \vdots & \vdots & \ddots & \vdots \\ (1-L)^{d_{k1}} & (1-L)^{d_{k2}} & \dots & (1-L)^{d_{kk}} \end{pmatrix}, \quad (10)$$

where  $d_{ij} = d_{ji}$  and  $0 \leq d_{ij} \leq 1$ , in order to describe the fractional difference of  $\mathbf{X}_t$  as  $D(L) \circ \mathbf{X}_t$ .

We do not consider negative values of  $d_{ij}$ , as the empirical evidences of the fractional integration are found in the positive region in financial time series. For any process of  $k$ -vector  $\mathbf{x}_t$ , we also define a  $k$ -dimensional diagonal matrix  $D_k(L)$  as

$$D_k(L) = \text{diag}((1-L)^{d_1^v}, \dots, (1-L)^{d_k^v}), \quad (11)$$

where  $0 \leq d_i^v \leq 1$ , so that  $D_k(L)\mathbf{x}_t$  is the fractional difference of  $\mathbf{x}_t$ . We can separately estimate the order of the fractional difference,  $d_{i,j}$  and/or  $d_i^v$ , via the multivariate Gaussian semiparametric estimation suggested by Shimotsu (2007), which is considered as a multivariate extension of the local Whittle (LW) estimator of Shimotsu and Phillips (2006).

Recent papers including Andersen, Bollerslev and Meddahi (2011), Bollerslev, Sizova and Tauchen (2011), Bollerslev and Todorov (2011), Chiriac and Voev (2011) work with the heterogeneous autoregressive (HAR) model of Corsi (2009) and its variants in order to describe longer-range dependence than that of AR(1) models. Following Corsi (2009), define  $h$  days average of past values of  $\mathbf{X}_t$  and  $\mathbf{x}_t$  as

$$\mathbf{X}_{t-1,t-h} = \frac{1}{h}(\mathbf{X}_{t-1} + \dots + \mathbf{X}_{t-h}), \quad \mathbf{x}_{t-1,t-h} = \frac{1}{h}(\mathbf{x}_{t-1} + \dots + \mathbf{x}_{t-h}), \quad (12)$$

respectively. In addition to daily effect,  $\mathbf{X}_{t-1}$ , we can consider weekly and monthly effects as  $\mathbf{X}_{t-1,t-5}$  and  $\mathbf{X}_{t-1,t-22}$ , respectively.

## 2.4 Estimation of Factor Covariance

Consider a consistent estimator of  $\Sigma_t$ , which is a realized covariance matrix, and we denote it as  $\hat{\Sigma}_t$ . By using the sequence of realized covariance matrix, Tao et al. (2011) developed a technique to estimate the factor covariance matrix,  $\Omega_t$ . We will shortly explain the approach. Define

$$\bar{\Sigma} = \frac{1}{T} \sum_{t=1}^T \Sigma_t, \quad C = \frac{1}{T} \sum_{t=1}^T (\Sigma_t - \bar{\Sigma})^2.$$

Recalling that  $A'B = O$ , we obtain  $A'\Sigma_t = A'\Sigma_u$  from equation (3). Hence, we obtain that  $A'\Sigma_t = A'\bar{\Sigma}$  and

$$A'CA = \frac{1}{T} \sum_{t=1}^T (A'\Sigma_t - A'\bar{\Sigma}) (\Sigma_t A - \bar{\Sigma} A) = O$$

From this result, we can consider  $A$  as the  $m - k$  orthonormal eigenvectors of  $C$ , corresponding to the  $(m - k)$ -fold eigenvalue 0. The remaining  $k$  orthonormal eigenvectors of  $C$ , corresponding to the  $k$  nonzero eigenvalues, can be taken as the columns of the factor matrix  $B$ .

Regarding the consistent estimator,  $\hat{\Sigma}_t$ , let

$$\bar{\hat{\Sigma}} = \frac{1}{T} \sum_{t=1}^T \hat{\Sigma}_t, \quad \bar{C} = \frac{1}{T} \sum_{t=1}^T (\hat{\Sigma}_t - \bar{\hat{\Sigma}})^2. \quad (13)$$

Then we obtain the estimator  $\hat{B}$  using the  $k$  orthonormal eigenvectors of  $\hat{C}$ , corresponding to the  $k$  largest eigenvalues, as its columns. Consequently, we obtain the estimated factor and its estimated covariance matrix as

$$\hat{\varphi}_t = \hat{B}' \hat{y}_t, \quad \hat{\Omega}_t = \hat{B}' \hat{\Sigma}_t \hat{B}, \quad t = 1, \dots, T,$$

respectively. We can also estimate  $\Sigma_{\zeta\zeta}$  by  $\hat{\Sigma}_{\zeta\zeta} = \hat{A}' \bar{\hat{\Sigma}} \hat{A}$ . Tao et al. (2011) showed the consistency of  $\hat{B}$  and  $\hat{\Omega}_t$  for fixed  $k$ , under several kinds of realized covariance matrices developed by Barndorff-Nielsen et al. (2008, 2011), Christensen, Kinnebrock, and Podolskij (2010), Griffin and

Oomen (2011), Hautsch, Kyj, and Oomen (2012), Wong and Zou (2010) and Zhang (2011). It is straightforward to show the consistency of  $\hat{\varphi}_t$  and  $\hat{\Sigma}_{\zeta\zeta}$  by applying the framework of Tao et al. (2011).

Once we obtained the estimates of the common factors,  $\hat{\varphi}_t$ , and covariance matrix of common factors,  $\hat{\Omega}_t$ , we can estimate alternative models for the covariance matrix of the common factors.

### 3 Specifications of Factor Covariance

We will extend five kinds of models for the covariance matrix, in order to accommodating asymmetry and long memory on the factor covariance matrix,  $\Omega_t$ .

#### 3.1 Dynamic Correlation fMSV Models

As in Engle (2002), we start from the decomposition of the covariance matrix,

$$\Omega_t = \mathbf{V}_t^{1/2} \mathbf{P}_t \mathbf{V}_t^{1/2}, \quad (14)$$

where  $\mathbf{V}_t$  is the diagonal matrix of volatilities of the vector of common factors, defined by  $\mathbf{V}_t = \text{diag}(v_{1t}, \dots, v_{kt})$ , and  $\mathbf{P}_t$  is its correlation matrix. By definition, we can write  $\mathbf{V}_t = (I_k \circ \Omega_t)$ , where ‘ $\circ$ ’ denotes the Hadamard (element-by-element) product. As in Engle (2002), we consider a positive definite matrix,  $\mathbf{Q}_t$ , in order to describe the correlation dynamics with  $\mathbf{P}_t = \mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2}$ , where  $\mathbf{Q}_t^* = (I_k \circ \mathbf{Q}_t)$ . Following most of the works on MSV models, we consider the vector of log-volatilities,  $\mathbf{h}_t = (h_{1t}, \dots, h_{kt})' = (\log v_{1t}, \dots, \log v_{kt})'$ , rather than volatilities themselves.

We use the dynamic correlation MSV (DC MSV) model of Asai and McAleer (2009a), in order

to specify the process of  $(\mathbf{Q}_t, \mathbf{h}_t)$  as

$$\mathbf{Q}_t^{-1} \sim W(\mathbf{R}_t, \delta), \quad \mathbf{R}_t = \frac{1}{\delta} (\mathbf{Q}_{t-1}^{-1})^{\frac{\alpha}{2}} \bar{\mathbf{R}} (\mathbf{Q}_{t-1}^{-1})^{\frac{\alpha}{2}}, \quad (15)$$

$$\mathbf{h}_t = \boldsymbol{\mu} + \boldsymbol{\Phi}_1 \{\mathbf{h}_{t-1} - \boldsymbol{\mu}\} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \quad (16)$$

where  $\delta$  ( $\delta > k - 1$ ) and  $\alpha$  ( $|\alpha| < 1$ ) are scalars,  $\bar{\mathbf{R}}$  and  $\boldsymbol{\Sigma}_\eta$  are  $k$  dimensional positive definite matrices,  $\boldsymbol{\mu}$  is  $k$ -vector of parameters,  $\boldsymbol{\Phi}_1$  is  $k$  dimensional diagonal matrix defined by  $\boldsymbol{\Phi} = \text{diag}(\phi_1, \dots, \phi_k)$ , and  $W(S, p)$  denotes the Wishart distribution with the scale matrix  $S$  and the degrees-of-freedom parameter  $p$ . Here, we define  $(\mathbf{P}_t^{-1})^{\alpha/2}$  by the spectral decomposition. We impose restrictions that the diagonal elements of  $\bar{\mathbf{R}}$  are ones, and that  $\mathbf{Q}_t = \bar{\mathbf{V}} \mathbf{R}_t \bar{\mathbf{V}}$ , where  $\bar{\mathbf{V}} = E(\mathbf{V}_t)$ .

Now we introduce the following general model for  $\mathbf{Q}_t$ , accommodating asymmetry and long memory.

$$\begin{aligned} \mathbf{Q}_t^{-1} &\sim W(\mathbf{R}_t, \delta), \\ \mathbf{R}_t &= \frac{1}{\delta} \left( \mathbf{P}_{t-1, t-22}^{-1} \right)^{\frac{\alpha_m}{2}} \left( \mathbf{P}_{t-1, t-5}^{-1} \right)^{\frac{\alpha_w}{2}} \left( \mathbf{P}_{t-1}^{-1} \right)^{\frac{\alpha_d}{2}} \\ &\quad \times [w \bar{\mathbf{R}} + (1-w) \boldsymbol{\Lambda}_{t-1}^\dagger] \left( \mathbf{P}_{t-1}^{-1} \right)^{\frac{\alpha_d}{2}} \left( \mathbf{P}_{t-1, t-5}^{-1} \right)^{\frac{\alpha_w}{2}} \left( \mathbf{P}_{t-1, t-22}^{-1} \right)^{\frac{\alpha_m}{2}}, \end{aligned} \quad (17)$$

where  $\alpha_d$ ,  $\alpha_w$  and  $\alpha_m$  are scalar parameters,  $w$  ( $0 < w \leq 1$ ) is a scalar weight parameter,  $\mathbf{P}_{t-1, t-5}$  and  $\mathbf{P}_{t-1, t-22}$  are weekly and monthly averages defined by (12), and  $\boldsymbol{\Lambda}_t$  is the positive semi-definite matrix defined by equation (9) such that the  $(i, j)$ th elements of  $\boldsymbol{\Lambda}^\dagger$  is  $\frac{1}{(1+\gamma_i)(1+\gamma_j)}$ . While  $\mathbf{P}_{t-1, t-5}$  and  $\mathbf{P}_{t-1, t-22}$  produce longer memory than the single  $\mathbf{P}_{t-1}$ ,  $\boldsymbol{\Lambda}_t$  yields asymmetric effects on the correlation dynamics.

Regarding the vector of log-volatilities,  $\mathbf{h}_t$ , a general model is given by

$$\boldsymbol{\Phi}(L) D_k(L) \{\mathbf{h}_t - \boldsymbol{\mu}\} = \boldsymbol{\lambda}_{t-1} + \boldsymbol{\eta}_t, \quad (18)$$

where  $D_k(L)$  is the fractional difference operator defined by equation (11),  $\boldsymbol{\lambda}_t$  is vector of asymmetric function defined by equation (7). We set  $\boldsymbol{\Phi}(L) = I_k - \boldsymbol{\Phi}_1 L$ . The competing heterogeneous VAR specification is given by

$$\mathbf{h}_t = \boldsymbol{\Phi}_d \{\mathbf{h}_{t-1} - \boldsymbol{\mu}\} + \boldsymbol{\Phi}_w \{\mathbf{h}_{t-1,t-5} - \boldsymbol{\mu}\} + \boldsymbol{\Phi}_m \{\mathbf{h}_{t-1,t-22} - \boldsymbol{\mu}\} + \boldsymbol{\lambda}_{t-1} + \boldsymbol{\eta}_t, \quad (19)$$

where  $\boldsymbol{\Phi}_d$ ,  $\boldsymbol{\Phi}_w$  and  $\boldsymbol{\Phi}_m$  are diagonal matrices.

It is convenient to have acronyms for the various fMSV-DC models: (i) **fMSV-DC** means the basic model defined by equations (14)-(16); (ii) **fMSV-DC-cAvA** stands for the fMSV-DC model with asymmetric effects on correlation and volatility dynamics, defined by equations (14), (17) and (18) such that  $\alpha_w = \alpha_m = 0$  and  $D_k(L) = I_k$ ; (iii) **fMSV-DC-cHvL** represents the fMSV-DC model with heterogeneous-time effects on correlation dynamics and with long-memory in volatility, defined by equations (14), (17) and (18) such that  $w = 1$ ,  $\boldsymbol{\Lambda}_t^\dagger = O$  and  $\boldsymbol{\lambda}_t = \mathbf{O}$ ; (iv) **fMSV-DC-cHvH** is the fMSV-DC model with heterogeneous-time effects on correlation and volatility dynamics, defined by equations (14), (17) and (19) such that  $w = 1$ ,  $\boldsymbol{\Lambda}_t^\dagger = O$  and  $\boldsymbol{\lambda}_t = \mathbf{O}$ ; (v) **fMSV-DC-cAHvAL** means the fMSV-DC model with asymmetric and heterogeneous-time effects on correlation dynamics and with asymmetric effects and long-memory in volatility dynamics, defined by equations (14), (17) and (18); (vi) **fMSV-DC-cAHvAH** stands for the fMSV-DC model with asymmetric and heterogeneous-time effects on correlation and volatility dynamics, defined by equations (14), (17) and (19).

We can estimate the fMSV-DC model (14), (17) and (18) (or (19)) using the likelihood function based on the Wishart distribution, when  $\boldsymbol{\Omega}_t$  is available. As noted in previous section, we can estimate long memory parameter  $d_i^v$  separately. In order to reduce the number of parameters in estimation, we may estimate the diagonal elements of  $\bar{\mathbf{V}}$  by the diagonal elements of the average of  $\hat{\boldsymbol{\Omega}}_t$ .

### 3.2 Wishart Disturbance Models for fMSV

Philipov and Glickman (2006a,b), Golosnoy, Gribisch and Liesenfeld (2012) and Asai and So (2013) assume that the covariance matrix follows a Wishart distribution (or an inverse Wishart distribution) with parameters that depend on the past covariance matrix. For the case of Wishart distribution, we can introduce a disturbance matrix for the covariance matrix of  $\varphi_t$  as

$$\mathbf{\Omega}_t = \mathbf{H}_t^{1/2} \mathbf{\Xi}_t \mathbf{H}_t^{1/2}, \quad \mathbf{\Xi}_t \sim W((1/\nu)I_k, \nu), \quad (20)$$

where  $\mathbf{H}_t$  is a  $k$  dimensional positive definite matrix, determined by the information up to  $t-1$ . By the property of the Wishart distribution,  $E(\mathbf{\Xi}_t) = I_k$  and thus  $E(\mathbf{\Omega}_t | \mathfrak{S}_{t-1}, \mathfrak{S}_{t-1}^\varphi) = \mathbf{H}_t$ , where  $\mathfrak{S}_{t-1}^\varphi$  consists of  $\{(\varphi_{t-1}, \mathbf{\Omega}_{t-1}), (\varphi_{t-2}, \mathbf{\Omega}_{t-2}), \dots\}$ . By definition, we also have  $\mathbf{\Omega}_t | \mathbf{H}_t \sim W((1/\nu)\mathbf{H}_t, \nu)$ .

Analogous to the BEKK model of Engle and Kroner (1995) and the diagonal MGARCH model of Ding and Engle (2001), we consider the following two specifications

$$\mathbf{H}_t = \mathbf{W} + \mathbf{G}' \mathbf{\Omega}_{t-1} \mathbf{G} + \mathbf{K}' \mathbf{H}_{t-1} \mathbf{K}, \quad (21)$$

$$\mathbf{H}_t = (\iota \iota' - \mathbf{G}^* - \mathbf{K}^*) \circ \mathbf{\Omega}^* + \mathbf{G}^* \circ \mathbf{\Omega}_{t-1} + \mathbf{K}^* \circ \mathbf{H}_{t-1}, \quad (22)$$

where  $\mathbf{W}$  and  $\mathbf{\Omega}^*$  are  $k$  dimensional positive definite matrices of parameters,  $\mathbf{G}$  and  $\mathbf{K}$  are  $k \times k$  matrices,  $\iota$  is a  $k$ -vector of ones, and  $\mathbf{F}^*$  and  $\mathbf{G}^*$  are  $k$  dimensional positive (semi-)definite matrices. By putting  $E(\mathbf{\Omega}_t) = E(\mathbf{H}_t)$  in equation (22), we have  $\mathbf{\Omega}^* = E(\mathbf{\Omega}_t)$ , which is the unconditional expectation of  $\mathbf{\Omega}_t$ . The equations (21) and (22) have the same parsimonious specification, which is obtained by setting  $\mathbf{G} = \text{diag}(\mathbf{g})$ ,  $\mathbf{K} = \text{diag}(\mathbf{k})$  and  $\mathbf{W} = (\iota \iota' - \mathbf{g}\mathbf{g}' - \mathbf{k}\mathbf{k}') \circ \mathbf{\Omega}^*$  in equation (21), namely,

$$\mathbf{H}_t = (\iota \iota' - \mathbf{g}\mathbf{g}' - \mathbf{k}\mathbf{k}') \circ \mathbf{\Omega}^* + (\mathbf{g}\mathbf{g}') \circ \mathbf{\Omega}_{t-1} + (\mathbf{k}\mathbf{k}') \circ \mathbf{H}_{t-1}, \quad (23)$$

where  $\mathbf{g}$  and  $\mathbf{k}$  are  $k$ -vector of parameters.

We will work with the Wishart disturbance model (20) and (23) in our analysis. We will call the model as the ‘conditional autoregressive Wishart fMSV’ (fMSV-CAW) model, following Golosnoy, Gribisch and Liesenfeld (2012).

The general specification for the fMSV-CAW model is given by equation (20) and

$$\begin{aligned} \mathbf{H}_t = & (\boldsymbol{\nu}' - \mathbf{g}_d \mathbf{g}'_d - \mathbf{g}_w \mathbf{g}'_w - \mathbf{g}_m \mathbf{g}'_m - \mathbf{k} \mathbf{k}') \circ \boldsymbol{\Omega}_t^* + (\mathbf{k} \mathbf{k}') \circ \mathbf{H}_{t-1} \\ & + (\mathbf{g}_d \mathbf{g}'_d) \circ \boldsymbol{\Omega}_{t-1} + (\mathbf{g}_w \mathbf{g}'_w) \circ \boldsymbol{\Omega}_{t-1,t-5} + (\mathbf{g}_m \mathbf{g}'_m) \circ \boldsymbol{\Omega}_{t-1,t-22} + \boldsymbol{\Lambda}_{t-1}^\dagger - \bar{\boldsymbol{\Lambda}}^\dagger, \end{aligned} \quad (24)$$

where  $\boldsymbol{\Omega}_{t-1,t-5}$  and  $\boldsymbol{\Omega}_{t-1,t-22}$  are weekly and monthly averages defined by (12),  $\boldsymbol{\Lambda}_t^\dagger$  is asymmetric matrix function defined by (9),  $\boldsymbol{\Omega}_{t-1,t-5}$ ,  $\bar{\boldsymbol{\Lambda}}^\dagger = E(\boldsymbol{\Lambda}_t^\dagger) = \boldsymbol{\Lambda}^\dagger \circ (I_k + \gamma \gamma')$  and  $\boldsymbol{\Lambda}^\dagger = \boldsymbol{\lambda}^\dagger \boldsymbol{\lambda}^{\dagger'}$  with  $k$ -vector  $\boldsymbol{\lambda}^\dagger$ .

The acronyms for the various fMSV-CAW models is as follows: (i) **fMSV-CAW** means the basic fMSV-CAW model, defined by equations (20) and (23); (ii) **fMSV-CAW-A** means the fMSV-CAW model with asymmetric effects, defined by equations (20) and (24) such that  $\mathbf{g}_w = \mathbf{g}_m = \mathbf{0}$ ; (iii) **fMSV-CAW-H** means the fMSV-CAW model with heterogeneous-time effects, defined by equations (20) and (24) such that  $\boldsymbol{\Lambda}_t^\dagger = \bar{\boldsymbol{\Lambda}}^\dagger = O$  and  $\mathbf{k} = \mathbf{0}$ ; (iv) **fMSV-CAW-AH** means the fMSV-CAW model with asymmetric and heterogeneous-time effects, defined by equations (20) and (24) with  $\mathbf{k} = \mathbf{0}$ .

When  $\boldsymbol{\Omega}_t$  is available, we can estimate these models by minimizing the log-likelihood function based on the Wishart distribution.

### 3.3 Wishart Autoregressive fMSV Models

Gourieroux, Jasiak and Sufana (2009) employed non-central Wishart distribution to define autoregressive property of a process of covariance matrices. They defined the model by the conditional

moment generating function, which is

$$E[\exp(\text{tr}\{\Upsilon\Omega_t\})|\Omega_{t-1}] = \frac{\exp(\text{tr}\{\mathbf{G}\Upsilon(I_k - 2\Omega^*\Upsilon)^{-1}\mathbf{G}'\Omega_{t-1}\})}{[\det(I_k - 2\Omega^*\Upsilon)^{\nu/2}],} \quad (25)$$

where  $\mathbf{G}$ ,  $\mathbf{W}$  and  $\nu$  are the same as in equations (20) and (21), and  $\Upsilon$  is a  $k \times k$  symmetric matrix which satisfies  $\|\mathbf{W}^{1/2}\Upsilon\mathbf{W}^{1/2}\| < 1$ . By Proposition 2 of Gouriéroux, Jasiak and Sufana (2009), we can write

$$\Omega_t = \nu\mathbf{W} + \mathbf{G}'\Omega_{t-1}\mathbf{G} + \mathbf{N}_t, \quad (26)$$

where  $\mathbf{N}_t$  is a heteroskedastic error term with conditional mean zero. The conditional mean of  $\Omega_t$  in (26) is similar to the specification in (21), but the approach of introducing error term is different.

As argued in Laurent, Rombouts, and Violante (2012) variance targeting estimation is useful when forecasting conditional covariance matrices. In this case, we may replace  $\nu\mathbf{W}$  with  $\Omega^* - \mathbf{G}'\Omega^*\mathbf{G}$ , where  $\Omega^* = E(\Omega_t)$ .

The general specification of the fMSV-WAR model is given by

$$\begin{aligned} \Omega_t = & (\omega' - \mathbf{g}_d\mathbf{g}'_d - \mathbf{g}_w\mathbf{g}'_w - \mathbf{g}_m\mathbf{g}'_m - \mathbf{k}\mathbf{k}') \circ \Omega^* + \Lambda_{t-1}^\dagger - \bar{\Lambda}^\dagger \\ & + (\mathbf{g}_d\mathbf{g}'_d) \circ \Omega_{t-1} + (\mathbf{g}_w\mathbf{g}'_w) \circ \Omega_{t-1,t-5} + (\mathbf{g}_m\mathbf{g}'_m) \circ \Omega_{t-1,t-22} + N_t, \end{aligned} \quad (27)$$

where parameters are the same as fMSV-CAW models. Note that the fMSV-WAR models has no past conditional covariance matrix unlike the fMSV-CAW, by construction. The acronyms for the various fMSV-WAR models is as follows: (i) **fMSV-WAR** means the basic fMSV-WAR model, defined by equation (26); (ii) **fMSV-WAR-A** means the fMSV-WAR model with asymmetric effects, defined by equation (27) such that  $\mathbf{g}_w = \mathbf{g}_m = \mathbf{0}$ ; (iii) **fMSV-WAR-H** means the fMSV-WAR model with heterogeneous-time effects, defined by equation (27) such that  $\Lambda_t^\dagger = \bar{\Lambda}^\dagger = O$ ;

(iv) **fMSV-WAR-AH** means the fMSV-WAR model with asymmetric and heterogeneous-time effects, defined by equation (27).

Rather than conducting ML estimation via non-central Wishart distribution, Gouriéroux, Jasiak and Sufana (2009) suggest to estimate via method of moments with first two moments. We will follow the latter approach.

### 3.4 Matrix-Exponential fMSV Models

Chiu, Leonard and Tsui (1996) suggested to use matrix-exponential transformation in order to guarantee the positive definiteness of covariance matrix, and Kawakatsu (2006) developed matrix-exponential GARCH models. Recently, Asai, McAleer and Yu (2006) and Ishihara, Omori and Asai (2012) proposed matrix-exponential SV models, while Bauer and Vorkink (2011) applied the matrix-exponential transformation for modeling realized covariances.

For any square matrix  $X$ , the matrix-exponential transformation is defined by  $\text{Exp}(X) = \sum_{i=0}^{\infty} (1/i!)X^i$  with  $X^0 = I$ . The same result is obtained by working with the spectral decomposition, as we have  $\text{Exp}(X)$  by replacing the eigenvalues by their exponential transformation. Note that  $\text{Exp}(X)$  is positive definite, whenever  $X$  is symmetric. In the same manner,  $\text{Log}(Y)$  is defined by its spectral decomposition of a positive definite matrix,  $Y$ , with replacement of the logarithmic transformation of eigenvalues.

Following Ishihara, Omori and Asai (2012), we consider the following matrix-exponential fMSV (ME-fMSV) model.

$$\mathbf{\Omega}_t = \text{Exp}(\mathbf{\Theta}_t), \tag{28}$$

$$\mathbf{\Theta}_t = \mathbf{M} + \mathbf{\Psi}_1 \circ \{\mathbf{\Theta}_{t-1} - \mathbf{M}\} + \mathbf{E}_t, \tag{29}$$

where  $\boldsymbol{\xi}_t = \text{vech}(\mathbf{E}_t) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\xi)$ ,  $\mathbf{\Theta}_t$  are  $k \times k$  symmetric matrices,  $\mathbf{M}$  and  $\mathbf{\Psi}_1$  is the  $k$  dimensional

symmetric matrix of parameters,  $\mathbf{E}_t$  are  $k$  dimensional symmetric matrix of disturbance, and  $\Sigma_\xi$  is the  $k(k+1)/2$  dimensional positive definite matrix.

The matrix-exponential fMSV model with asymmetry and long-memory is given by equation (28) and

$$\Psi(L) \circ D(L) \circ \{\Theta_t - M\} = \Lambda_{t-1} + \mathbf{E}_t, \quad (30)$$

where  $\Psi(L) = \iota' - \Psi_1 L$ , the asymmetric error matrix  $\Lambda_t$  is defined by (8), and the matrix of fractional difference operator  $D(L)$  is defined by (11). Instead of long-memory, we can incorporate heterogeneous-time effect into the fMSV model as,

$$\Theta_t = M + \Psi_d \circ \{\Theta_{t-1} - M\} + \Psi_w \circ \{\Theta_{t-1,t-5} - M\} + \Psi_m \circ \{\Theta_{t-1,t-22} - M\} + \Lambda_{t-1} + \mathbf{E}_t, \quad (31)$$

where  $\Psi_d$ ,  $\Psi_w$  and  $\Psi_m$  is the  $k$  dimensional symmetric matrix of parameters. The acronyms for the various fMSV-EXP models is as follows: (i) **fMSV-EXP** means the basic fMSV-EXP model, defined by equation (29); (ii) **fMSV-EXP-A** means the fMSV-EXP model with asymmetric effects, defined by equation (30) such that  $D(L) = \iota'$ ; (iii) **fMSV-EXP-L** means the fMSV-EXP model with long memory, defined by equation (30) such that  $\Lambda_1 = \bar{\Lambda}_2 = O$ ; (iv) **fMSV-EXP-AL** means the fMSV-EXP model with asymmetry and long memory, defined by equation (30); (v) **fMSV-EXP-H** means the fMSV-EXP model with heterogeneous-time effects, defined by equation (31) such that  $\Lambda_1 = \bar{\Lambda}_2 = O$ ; (vi) **fMSV-EXP-AH** means the fMSV-EXP model with asymmetric and heterogeneous-time effects, defined by equation (31).

When  $\Omega_t$  is available, we can calculate  $\Theta_t = \text{Log}(\Omega_t)$  in order to conduct the OLS estimation for each equation. As noted in previous section, we can estimate long memory parameter  $d_{ij}$  separately.

### 3.5 Cholesky fMSV Models

Chiriac and Voev (2011) applied the Cholesky decomposition for modeling realized covariances. For the covariance matrix of  $\varphi_t$ , we consider the Cholesky decomposition of  $\mathbf{\Omega}_t$  as

$$\mathbf{\Omega}_t = \mathbf{L}_t \mathbf{L}'_t, \quad (32)$$

where  $\mathbf{L}_t$  are lower triangular matrices. Define  $\mathbf{a}_t = \text{vech}(\mathbf{L}_t)$ . Then we can consider VAR model for  $\mathbf{a}_t$  as,

$$\mathbf{a}_t = \boldsymbol{\mu}_c + \mathbf{\Phi}_c \{\mathbf{a}_{t-1} - \boldsymbol{\mu}_c\} + \boldsymbol{\xi}_t \quad (33)$$

where  $\boldsymbol{\mu}_c$  is the  $k(k+1)/2$  vector of parameters,  $\mathbf{\Phi}_c$  is the  $k(k+1)/2$  dimensional diagonal matrix, and  $\boldsymbol{\xi}_t$  is the  $k(k+1)/2$  vector of error terms.

Analogous to the fMSV-EXP models, we can develop **fMSV-CH**, **fMSV-CH-A**, **fMSV-CH-L**, **fMSV-CH-AL**, **fMSV-CH-H**, and **fMSV-CH-AH** models. When  $\mathbf{\Omega}_t$  is available, we can calculate  $\mathbf{a}_t$  in order to conduct the OLS estimation for each equation. As noted in previous section, we can estimate long memory parameter  $d_{ij}$  separately.

## 4 Empirical Analysis

### 4.1 Data and Preliminary Analysis

We examine forecasting performances of five classes of fMSV models using daily realized covariance matrices for seven stocks traded at the New York Stock Exchange: Alcoa Inc. (AA), American Express (AXP), Bank of America (BAC), du Pont de Nemours and Company (DD), General Electric (GE), International Business Machines (IBM), and Coca-Cola Company (KO). Based on the vector of returns for the  $m = 7$  stocks computed for 1-min interval of trading day at  $t$  between 9:30a.m. and 4:00 p.m., we calculated daily realized volatilities and co-volatilities by

the averaging approach of Wang and Zou (2010) using 5-min averages with adjustments. Note that the method of Wang and Zou (2010) is robust to microstructure noise, and we can use the estimator for estimating factor covariances as proposed by Tao et al. (2011). As suggested by Bickel and Levina (2008a,b), Johnstone and Lu (2009) and Wang and Zou (2010), we impose some sparsity structure on the estimator in order to guarantee the consistency. For this purpose these authors work with threshold which means retaining its elements whose absolute values exceed a given value and replacing others by zero. For convenience, we find the threshold for co-volatilities on each day  $t$ , so that the estimated covariance matrix is positive definite. We also calculated the corresponding open-close returns for seven assets. The vector of returns and its covariance matrix estimator are denoted by  $\mathbf{y}_t$  and  $\hat{\Sigma}_t$ , respectively. The sample period starts at February 20, 1997, and ends on October 4, 2012, covering 3900 observations. Table 1 presents the descriptive statistics of the returns, volatilities and co-volatilities. The empirical distribution of the returns is heavily skewed to the left and is highly leptokurtic. Regarding volatilities and co-volatilities, they are skewed to the right except for co-volatilities for (AXP-KO) and (BAC-DD), with evidences of heavy-tails for all series. It should be noted that extremely large values for volatilities are observed during the IT bubble in the US.

We will use latter 1500 observations for forecasting analysis later in this section. Figure 1 shows the estimated volatilities for the period of forecasting. Figure 1 indicates that volatilities are high in the period of turbulence caused by the GFC. Figure 2 presents some of estimated correlation dynamics. We chose ten series which have stronger correlations than remaining eleven series. Because of the threshold explained above, correlation coefficients often take zero when they are close to zero. An interesting feature is that the correlation dynamics of (AA-BAC), (AXP-BAC), (GE-BAC), (IBM-BAC) fluctuate around zero just after the bankruptcy of Lehman

Brothers, then they show stronger correlation on average. One year after the bankruptcy of Lehman Brothers, these four series shows weak correlations around zero. On the contrary, other 6 correlation processes show strong connection for the whole period for forecasting.

As a preliminary analysis with former  $T = 2400$  observations, Table 2 gives the contribution rates of eigenvalues of  $\hat{C}$ , defined by equation (13). As reported in Sugiyama and Tong (1976) and Sheena (2013), the estimator of contribution rates are suffered from bias in finite sample, and hence Table 2 also gives the modified estimator of Sheena (2013). In this case, there are no major differences in these two estimators of contribution rates. The cumulative percentage of the eigenvalues up to the first 4 components is 89 percent, while it is 95% for the 5 components. For this reason, we will employ 4 and 5 factors in our empirical analysis.

We also estimated the news impact surface based on the estimates from the fMSV-CAW-A model, using the former periods from 1997 to 2006. Figure 3 gives NICs the NIC from  $y_{i,t}$  to  $\sigma_{j,t+1}^{1/2}$  for  $i, j = AA, AXP, BAC$ , after removing the effects from other assets. For convenience, the news impact is scaled by the mean of realized volatility for each series. While the NIC from each asset to its own future volatility shows asymmetric effects, the news impacts for other volatilities are negligible. Regarding BAC, the positive relation between current return and future volatility reflects the rapid growth for the years from 2000 to 2006.

## 4.2 Benchmark Models

For the benchmark models, we employ two kinds of dynamic conditional correlation (DCC) models; one is the asymmetric DCC (ADCC) model suggested by Cappiello, Engle and Sheppard (2006), while the other is the fractionally integrated matrix-exponential DCC (FIEDCC) model of Asai and So (2014). While ADCC model is popular and it captures the asymmetric effects in dynamic correlations, the FIEDCC model accommodates the long-memory and asymmetry in the dynamic

correlation process.

We consider the conventional specification of dynamic correlation model as

$$\begin{aligned} \mathbf{y}_t &= \mathbf{m}_t + \mathbf{u}_t, & E(\mathbf{u}_t | \mathfrak{S}_{t-1}) &= \mathbf{0}, & Var(\mathbf{u}_t | \mathfrak{S}_{t-1}) &= \mathbf{\Sigma}_t = \mathbf{V}_t^{1/2} \mathbf{P}_t \mathbf{V}_t^{1/2}, \\ \mathbf{V}_t &= \text{diag}(\mathbf{v}_t), & \mathbf{P}_t &= \mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2}, & \mathbf{Q}_t^* &= (I_k \circ \mathbf{Q}_t), \end{aligned} \quad (34)$$

where  $\mathbf{m}_t = E(\mathbf{y}_t | \mathfrak{S}_{t-1})$  are the conditional mean of  $\mathbf{y}_t$ ,  $\mathbf{v}_t = (v_{1t}, \dots, v_{mt})'$  are the vector of conditional volatilities, and  $\mathbf{Q}_t$  are  $m$  dimensional positive definite matrices which determine the process of conditional correlations. We also use the standardization defined by  $\mathbf{z}_t = \mathbf{V}_t^{-1/2}(\mathbf{y}_t - \mathbf{m}_t)$ , which produces  $E(\mathbf{z}_t \mathbf{z}_t' | \mathfrak{S}_{t-1}) = \mathbf{P}_t$ . We may simply assume that  $\mathbf{m}_t = \mathbf{0}$ .

Our ADCC model is given by equation (34) and

$$\begin{aligned} \mathbf{Q}_t &= (\bar{\mathbf{P}} - a\bar{\mathbf{P}} - b\bar{\mathbf{P}} - g\bar{\mathbf{N}}) + a(\mathbf{z}_{t-1} \mathbf{z}_{t-1}') + g(\mathbf{z}_{t-1}^- \mathbf{z}_{t-1}^-') + b\mathbf{Q}_{t-1}, \\ \mathbf{v}_t &= \boldsymbol{\omega} + \boldsymbol{\delta} \circ \mathbf{u}_{t-1} \circ \mathbf{u}_{t-1} + \boldsymbol{\eta} \circ \mathbf{v}_{t-1} + \boldsymbol{\xi} \circ \mathbf{u}_{t-1}^- \circ \mathbf{u}_{t-1}^-, \end{aligned} \quad (35)$$

where  $\boldsymbol{\omega}$ ,  $\boldsymbol{\delta}$ ,  $\boldsymbol{\eta}$  and  $\boldsymbol{\xi}$  are vectors of parameters,  $a$ ,  $b$  and  $g$  are scalar parameters,  $\mathbf{z}_t^- = (z_{1t}^-, \dots, z_{mt}^-)'$ ,  $z_{it}^- = z_{it} I(z_{it} < 0)$ ,  $I(\cdot)$  is an indicator function which takes the value of 1 if the argument is true and 0 otherwise,  $\bar{\mathbf{P}} = E(\mathbf{z}_t \mathbf{z}_t')$  and  $\bar{\mathbf{N}} = E(\mathbf{z}_t^- \mathbf{z}_t^-')$ . We work with the GJR model for the volatility process due to its status as one of popular asymmetric models in the ARCH family. As proposed by Cappiello, Engle and Sheppard (2006), we replace  $\bar{\mathbf{P}}$  and  $\bar{\mathbf{N}}$  by the sample covariance matrices.

The second model is the FIEDCC model given by equation (34) and

$$\begin{aligned} (1 - \underline{\alpha}L)(1 - L)^{d^c} (\text{Log} \mathbf{Q}_t - \mathbf{M}) &= \underline{\lambda}_1 (\mathbf{z}_{t-1} \iota_m' + \iota_m \mathbf{z}_{t-1}') + \underline{\lambda}_2 (\mathbf{z}_{t-1} \mathbf{z}_{t-1}' - \bar{\mathbf{P}}), \\ (1 - \underline{\phi}_i L)(1 - L)^{d_i^v} (\log v_{it} - \underline{\kappa}_i) &= \lambda_{1i} z_{i,t-1} + \lambda_{2i} (z_{i,t-1}^2 - 1), \end{aligned} \quad (36)$$

where  $d^c$  and  $d_i^v$  are scalar parameters of fractional difference,  $\underline{\alpha}$  and  $\underline{\phi}_i$  are AR(1) coefficients,  $\lambda_{1i}$ ,  $\lambda_{2i}$ ,  $\underline{\lambda}_1$  and  $\underline{\lambda}_2$  are scalar parameters for error terms which accommodate asymmetric effects,  $\mathbf{M} = E(\text{Log} \mathbf{Q}_t)$ ,  $\underline{\kappa}_i = E(\log v_{it})$ . As suggested by Asai and So (2014), we specify  $\mathbf{M} = \text{Log}(\bar{\mathbf{P}})$  and replace  $\bar{\mathbf{P}}$  by the sample covariance matrix.

We employ these two DCC models as the benchmarks for the empirical analysis.

### 4.3 Results

We set  $T = 2400$  for estimating alternative forecasting models, focusing on predict horizons of  $h = \{1, 5, 10\}$  days. We denote the  $h$ -step ahead forecast of  $\Sigma_t$  as  $\hat{\Sigma}_{t+h}^f$ . We re-estimate every model fixing sample size as  $T = 2400$ , and obtain new forecasts based on updated parameters. In our forecasting analysis, we use three out-of-sample windows, taking account for the effects of the Global Financial Crisis (GFC). The first window covers the period from September 8, 2006 to September 12, 2008. The second period starts from the bankruptcy of Lehman Brothers, i.e. September 15, 2008, and ends on September 24, 2010. The third window covers the period from September 27, 2010 to October 4, 2012. We will call these three windows as periods ‘before GFC’ (B-GFC), ‘during GFC’ (D-GFC) and ‘after GFC’ (A-GFC), respectively. The size of each window is 500.

We use  $\hat{\Sigma}_{t+h}$  as a proxy for the unobservable  $\Sigma_{t+h}$ , and define forecast error matrix as

$$E_{t+h} = \hat{\Sigma}_{t+h} - \hat{\Sigma}_{t+h}^f,$$

where  $\hat{\Sigma}_{t+h}^f$  is the  $h$ -step-ahead forecast of  $\Sigma_t$ . Following Chiriac and Voev (2011) and Golosnoy, Gribisch, and Liesenfeld (2012), we compare the seven models’ out-of-sample forecast root-mean-squared error (RMSE) based on the Frobenius norm of the forecast error, which is defined by

$$FN_h = \frac{1}{T_h} \sum_t \|E_{t+h}\| = \frac{1}{T_h} \sum_t \left[ \sum_{i,j} e_{ij,t+h}^2 \right]^{1/2}, \quad (37)$$

where  $T_h$  is the number of forecast periods. Especially,  $T_1 = 500$ ,  $T_5 = 496$  and  $T_{10} = 491$ .

Table 3 shows the results for forecasting performances. For 5- and 10-step-ahead forecasts, the statistic is standardized by adjusting by the number of steps to make the results comparable. Among the benchmark DCC models, the FIEDCC model always performs better than the ADCC model. However, all the factor MSV models performs better than the FIEDCC model.

Before examining the results of out-of-sample forecasts, we compare the results of in-sample forecasts for  $T = 2400$  observations, presented by Table 3. For the 4 factor MSV models, the DC-cAH-vAL, DC-cAH-vAH, EXP, EX-A, EXP-H and EXP-AH have smaller RMSE than the other models. Regarding the 5 factor models, the DC-cAH-vAL, DC-cAH-vAH, EXP-L, EX-AL, EXP-H and EXP-AH perform better than the other models. The WAR type and CAW type models give relatively larger RMSEs.

Regarding the out-of-sample forecasts, Table 3 implies three general results. First of all, the results obviously depend on the time periods, ‘before’, ‘during’ and ‘after’ the GFC, rather than the length of forecasting horizons. Compare to the RMSEs before the GFC, those for the period during the GFC is about 1.2 times, while those for the period after GFC is less than the half. The sizes of volatilities in Figure 1 support the results. Secondly, Table 3 indicate that there are no major differences between 4 and 5 factor models, showing the success of approximating via factors. Thirdly, accommodating asymmetric effects generally improves the RMSE of 1-day-ahead forecasts. We may improve the results by introducing asymmetric effects for heterogeneous time horizons, as in Chen and Ghysels (2010) and Asai (2013).

The best forecasting model depends on the time period and forecasting horizons. However, the CH-L and CH-AL models have relatively small RMSEs for all cases, supporting the results of Chiriac and Voev (2011). Although the simple WAR and WAR-A shows the best performance during the GFC, the CH-L and CH-AL models competitive results. In addition to these models, the DC-cAH-vAL, EXP-L and EXP-AL models often give the competitive results. We should note the similarity of the structure of CAW-type and WAR-type models, and differences of the forecasting performances brought by the differences of estimation procedures. Minimizing the difference of first moment in WAR-type models has an advantage on the above  $FN_h$ .

## 5 Conclusion

In this paper, we developed an approach for modeling asymmetry and long memory on the factor MSV (fMSV) model, using its alternative representation. Based on the approach, we extended DC, CAW, WAR, matrix-exponential MSV and Cholesky MSV model for specifying the covariance matrix of common factors. In forecasting application with seven stocks traded at the New York Stock Exchange, we show that the new fMSV models performs than the asymmetric DCC and fractionally-integrated matrix-exponential DCC models. Among the new fMSV models, the Cholesky fMSV model with long memory and asymmetry is the best model, since it has smallest RMSEs (or competitive results), and since it is robust to forecasting horizons and the periods before, during and after the GFC.

Our new factor approach for asymmetric effects and long memory opens to many interesting research directions. We may include not only daily asymmetric effects but also weekly and monthly effects, as in Chen and Ghysels (2010) and Asai (2013). We neglected the jumps and the impact of macroeconomic variables on future volatility and co-volatility. For these issues, we need to wait further researches on co-jumps and parsimonious specifications for macroeconomic variables.

## References

- Andersen, T.G., T. Bollerslev and N. Meddahi (2011), “Realized Volatility Forecasting and Market Microstructure Noise”, *Journal of Econometrics*, **160**, 220–234.
- Asai, M. (2013), “Heterogeneous Asymmetric Dynamic Conditional Correlation Model with Stock Return and Range”, *Journal of Forecasting*, **32**, 469–480.
- 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., M. McAleer, and J. Yu (2006), “Multivariate Stochastic Volatility: A Review”, *Econometric Reviews*, **25**, 145–175.
- Asai, M. and M.K.P. So (2013), “Stochastic Covariance Models”, *Journal of Japan Statistical Society*, **43**, 127–162.
- Asai, M. and M.K.P. So (2014), “Long Memory and Asymmetry for Matrix-Exponential Dynamic Correlation Processes”, Working Paper, Faculty of Economics, 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., P.R. Hansen, A. Lunde, and N. Shephard (2011), “Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of The Covariation of Equity Prices with Noise and Non-Synchronous Trading”, *Journal of Econometrics*, **162**, 149–169.
- 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.
- Bauer, G.H. and K. Vorkink (2011), “Forecasting Multivariate Realized Stock Market Volatility”, *Journal of Econometrics*, **160**, 93–101.
- Bauwens, L., S. Laurent, and J.V.K. Rombouts (2006), “Multivariate GARCH Models: A Survey”, *Journal of Applied Econometrics*, **21**, 79–109.
- Bickel, P. J., and E. Levina (2008a), “Regularized Estimation of Large Covariance Matrices”, *Annals of Statistics*, **36**, 199–277.
- Bickel, P. J., and E. Levina (2008b), “Covariance Regularization by Thresholding”, *Annals of Statistics*, **36**, 2577–2604.
- Bollerslev, T. and H.O. Mikkelsen (1996), “Modeling and Pricing Long-Memory in Stock Market Volatility”, *Journal of Econometrics*, **73**, 151–184.
- Bollerslev, T., N. Sizova, and G. Tauchen (2011), “Volatility in Equilibrium: Asymmetries and Dynamic Dependencies”, *Review of Finance*, **16**, 31–80.

- Bollerslev, T. and V. Todorov (2011), “Tails, Fears, and Risk Premia”, *Journal of Finance*, **66**, 2165–2211.
- 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.
- Caporin, M., and M. McAleer (2011), “Threshold, News Impact Surfaces and Dynamic Asymmetric Multivariate GARCH”, *Statistica Neerlandica*, **65**, 125–163.
- Cappiello, L., R.F. Engle and K. Sheppard (2006), “Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns”, *Journal of Financial Econometrics*, **4**, 537–572.
- Chen, X. and E. Ghysels (2010), “News - Good or Bad - and Its Impact on Volatility Predictions over Multiple Horizons”, *Review of Financial Studies*, **24**, 46–81.
- Chiriac, R., and V. Voev (2011), “Modelling and Forecasting Multivariate Realized Volatility”, *Journal of Applied Econometrics*, **26**, 922–947.
- Chib, S., F. Nardari, and N. Shephard (2006), “Analysis of High Dimensional Multivariate Stochastic Volatility Models”, *Journal of Econometrics*, **134**, 341–371.
- Chib, S., Y. Omori, and M. Asai (2009), “Multivariate Stochastic Volatility”, In T. G. Andersen, R.A. Davis, J.P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp.365–400, New York: Springer-Verlag.
- 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.
- Christensen, K., S. Kinnebrock and M. Podolskij (2010), “Pre-Averaging Estimators of the Ex-Post Covariance Matrix in Noisy Diffusion Models with Non-Synchronous Data”, *Journal of Econometrics*, **159**, 116–133.
- Corsi, F. (2009), “A Simple Approximate Long-Memory Model of Realized Volatility”, *Journal of Financial Econometrics*, **7**, 174–196.
- Corsi, F., and R. Renò (2010), “HAR Volatility Modelling with Heterogeneous Leverage and Jumps”, Unpublished Paper, Università di Siena.
- Diebold, F.X., and M. Nerlove (1989), “The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model”, *Journal of Applied Econometrics*, **4**, 1–21.
- Ding, Z. and Engle, R. F. (2001), “Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing”, *Academia Economic Papers*, **29**, 157–184.
- Durbin, J. and S.J. Koopman (1997), “Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models”, *Biometrika*, **84**, 669–684.
- 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 K.F. Kroner (1995), “Multivariate Simultaneous Generalized ARCH”, *Econometric Theory*, **11**, 122–150.
- Engle, R.F., and V. Ng (1993), “Measuring and Testing the Impact of News on Volatility”, *Journal of Finance*, **48**, 1749–1778.

- Gallant, A.R., and G. Tauchen (1996), “Which Moments to Match?”, *Econometric Theory*, **12**, 657–681.
- Glosten, L., R. Jagannathan, and D. Runkle (1992), “On the Relation between the Expected Value and Volatility of Nominal Excess Returns on Stocks”, *Journal of Finance*, **46**, 1779–1801.
- Golosnoy, V., B. Gribisch, and R. Liesenfeld (2012), “The Conditional Autoregressive Wishart Model for Multivariate Stock Market Volatility”, *Journal of Econometrics*, **167**, 211–223.
- Gourieroux, C., J. Jasiak, and R. Sufana (2009), “The Wishart Autoregressive Process of Multivariate Stochastic Volatility”, *Journal of Econometrics*, **150**, 167–181.
- Griffin, J.E., and R.C.A. Oomen (2011), “Covariance Measurement in the Presence of Non-Synchronous Trading and Market Microstructure Noise”, *Journal of Econometrics*, **160**, 58–68.
- Hautsch, N., R.M. Kyj and R.C.A. Oomen (2012), “A Blocking and Regularization Approach to High Dimensional Realized Covariance Estimation”, *Journal of Applied Econometrics*, **27**, 625–645.
- 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.
- 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.
- Harvey, A.C., E. Ruiz, and N. Shephard (1994), “Multivariate Stochastic Variance Models”, *Review of Economic Studies*, **61**, 247–264.
- Ishihara, T., Y. Omori and M. Asai (2012), “Matrix Exponential Stochastic Volatility with Cross Leverage”, Working paper, Faculty of Economics, University of Tokyo.
- Jacquier, E., N.G. Polson, and P.E. Rossi (1994), “Bayesian Analysis of Stochastic Volatility Models” (with discussion), *Journal of Business & Economic Statistics*, **12**, 371–389.
- Johnstone, I.M., and A.Y. Lu (2009), “On Consistency and Sparsity for Principal Component Analysis in High Dimensions” (with discussions), *Journal of the American Statistical Association*, **104**, 682–703.
- Kawakatsu, H. (2006), “Matrix Exponential GARCH”, *Journal of Econometrics*, **134**, 95–128.
- Kroner, K. and V. Ng (1998), “Modeling Asymmetric Comovements of Asset Returns”, *Review Financial Studies*, **11**, 817–844.
- Lanne, M., and P. Saikkonen (2007), “A Multivariate Generalized Orthogonal Factor GARCH Model”, *Journal of Business & Economic Statistics*, **25**, 61–75.
- Laurent, S., J.V.K. Rombouts and F. Violante (2012), “On the Forecasting Accuracy of Multivariate GARCH Models”, *Journal of Applied Econometrics*, **27**, 934–955.
- Liesenfeld, R., and J.-F. Richard (2003), “Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics”, *Journal of Empirical Finance*, **10**, 505–531.
- Martens, M., D. van Dijk, and M. de Pooter (2009), “Forecasting S&P 500 Volatility: Long Memory, Level Shifts, Leverage Effects, Day-of-the-Week Seasonality, and Macroeconomic Announcements”, *International Journal of Forecasting*, **25**, 282–303.
- Nelson, D.B. (1991), “Conditional Heteroskedasticity in Asset Returns: A New Approach”, *Econometrica*, **59**, 347–370.
- 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.

- Philipov, A. and M.E. Glickman (2006a), “Factor Multivariate Stochastic Volatility via Wishart Processes”, *Econometric Reviews*, **25**, 311–334.
- Philipov, A. and M.E. Glickman (2006b), “Multivariate Stochastic Volatility via Wishart Processes”, *Journal of Business & Economic Statistics*, **24**, 313–328.
- Pitt, M.K. and N. Shephard (1999), “Time Varying Covariances: A Factor Stochastic Volatility Approach”, In J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith (Eds.), *Bayesian Statistics*, Volume 6, pp.547–570, Oxford: Oxford University Press.
- Ray, B.K., and R.S. Tsay (2000), “Long-Range Dependence in Daily Stock Volatilities”, *Journal of Business & Economic Statistics*, **18**, 254–262.
- Sandmann, G. and S.J. Koopman (1998), “Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood”, *Journal of Econometrics*, **87**, 271–301.
- Sheena (2013), “Modified Estimators of the Contribution Rates of Population Eigenvalues”, *Journal of Multivariate Analysis*, **115**, 301–316.
- 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.
- Silvennoinen, A., and T. Teräsvirta (2009), “Multivariate GARCH Models”, In T. G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch (eds.), *siHandbook of Financial Time Series*, 201–229, New York: Springer.
- 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.
- Sugiyama, T and H. Tong (1976), “On A Statistic Useful for Dimensionality Reduction of Linear Stochastic Systems”, *Communications in Statistics: Theory and Methods*, **5**, 711–721.
- Tao, M., Y. Wang, Q. Yao and J. Zou (2011), “Large Volatility Matrix Inference via Combining Low-Frequency and High-Frequency Approaches”, *Journal of the American Statistical Association*, **106**, 1025–1040
- Vrontos, I.D., P. Dellaportas, and D.N. Politis (2003), “A Full-Factor Multivariate GARCH Model”, **6**, 311–333.
- Wang, Y., J. and Zou (2010), “Vast Volatility Matrix Estimation for High-Frequency Financial Data”, *Annals of Statistics*, **38**, 943–978.
- Yu, J. (2005), “On Leverage in a Stochastic Volatility Model”, *Journal of Econometrics*, **127**, 165–178.
- Zhang, L. (2011), “Estimating Covariation: Epps Effect, Microstructure Noise”, *Journal of Econometrics*, **160**, 33–47.

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

Stock	Mean	Min	Max	Std.Dev.	Skew.	Kurt.
Returns						
AA	-0.1290	-70.939	13.737	2.8079	-8.0234	200.61
AXP	0.0270	-107.15	14.458	2.8477	-13.455	511.57
BAC	-0.0871	-67.659	20.080	2.9418	-3.3396	82.829
DD	-0.0201	-67.922	10.052	2.1055	-8.5250	277.95
GE	0.0145	-71.076	14.125	2.2669	-14.702	453.93
IBM	-0.0376	-66.247	11.748	2.0435	-8.6308	283.87
KO	0.0288	-11.158	9.9392	1.4375	-0.0529	8.6970
Volatilities						
AA	6.1297	0.0561	3797.4	74.397	41.868	1920.3
AXP	6.4466	0.0610	9624.3	153.43	62.408	3912.4
BAC	5.7171	0.0813	2847.6	47.653	54.040	3206.1
DD	3.7965	0.0118	3101.2	52.793	53.705	3056.3
GE	3.1839	0.0612	2713.5	46.938	51.792	2882.3
IBM	3.4917	0.0857	2680.2	43.241	60.159	3720.5
KO	1.7498	0.0174	605.39	9.9454	56.786	3438.9
Co-volatilities						
AA-AXP	1.4739	-59.937	137.88	4.3539	11.621	299.80
AA-BAC	1.0984	-42.190	84.610	3.7854	7.6459	112.70
AA-DD	0.4821	-30.114	26.276	1.2324	0.5513	159.21
AA-GE	0.8989	-3.6298	76.053	2.2981	14.243	350.98
AA-IBM	1.3141	-2.8196	125.46	3.5966	15.833	435.35
AA-KO	0.3413	-11.475	16.942	0.9482	4.9720	82.635
AXP-BAC	1.6783	-14.746	81.759	5.1727	7.6358	81.302
AXP-DD	0.5802	-27.327	59.436	1.7817	11.081	356.76
AXP-GE	1.0623	-11.736	44.476	2.1931	7.5522	99.438
AXP-IBM	1.5056	-34.963	99.649	3.6069	9.8920	193.75
AXP-KO	0.4554	-134.08	19.622	2.5052	-38.759	2114.3
BAC-DD	0.5001	-65.812	19.828	1.7622	-10.627	532.2
BAC-GE	0.8557	-15.578	50.629	2.1967	7.9632	117.02
BAC-IBM	1.3033	-7.1569	96.732	4.4039	10.500	157.94
BAC-KO	0.4435	-7.9434	15.860	1.1695	5.4442	57.057
DD-GE	0.5024	-5.2749	107.63	2.0553	37.012	1878.1
DD-IBM	0.569	-7.0968	21.231	1.3087	5.5390	58.218
DD-KO	0.4330	-7.0070	18.929	1.0510	6.1137	74.406
GE-IBM	1.0140	-9.3817	45.273	1.9967	8.4410	128.99
GE-KO	0.4258	-8.6030	16.380	1.0004	5.6546	62.400
IBM-KO	0.5188	-24.579	18.491	1.2088	2.7633	96.704

Note: The number of observations for each series is 3900.

Table 2: Estimated Contribution Rates of Eigenvalues of  $\mathcal{C}$

Component	1	2	3	4	5	6	7
Estimates	0.6182	0.1440	0.0737	0.0578	0.0550	0.0486	0.0027
Modified Estimates	0.6172	0.1438	0.0737	0.0578	0.0550	0.0486	0.0027

Note: We use the approach of Sheena (2013) to calculate modified estimators of contribution rates.

Table 3: RMSE based on the Frobenius Norm of the Forecasting Error

Model	In-sample	1 day-ahead forecasts			5 day-ahead forecasts			10 day-ahead forecasts		
		B-GFC	D-GFC	A-GFC	B-GFC	D-GFC	A-GFC	B-GFC	D-GFC	A-GFC
Benchmark Models										
ADCC	4.8929	4.3007	7.6422	1.9062	4.6859	8.1614	1.9238	5.5587	8.8414	1.9448
FIEDCC	3.2042	4.1386	7.0064	1.8731	4.1002	7.0117	1.8130	4.1183	7.0344	1.8407
4 factor MSV models										
DC	2.6840	3.7689	4.6917	1.6633	3.8142	4.6945	1.6663	3.8375	5.3801	1.7341
DC-cAvA	2.3775	3.7639	4.6666	1.6883	3.8098	4.7293	1.7175	3.8167	5.3996	1.7179
DC-cHvL	2.3811	3.7885	5.0152	1.6520	3.7927	5.0165	1.6701	3.8534	5.3890	1.6851
DC-cAHvAL	2.1983	3.7669	4.8222	1.6270	3.7877	4.9981	1.6526	3.8530	5.0022	1.6714
DC-cHvH	2.2003	3.8116	5.5722	1.7002	3.8145	5.5874	1.7005	3.8798	5.6707	1.7211
DC-cAHvAH	2.1992	3.8125	5.4812	1.6989	3.8127	5.5289	1.7014	3.8769	5.7067	1.7192
CAW	2.9872	3.8961	5.2071	1.6806	4.1996	5.7201	2.3647	4.3954	5.8522	1.8367
CAW-A	2.8897	3.8953	5.1027	1.6814	4.7276	5.3257	2.2683	4.2358	5.8350	1.8182
CAW-H	2.9988	3.9266	6.3134	1.7590	4.4863	5.8106	2.6220	4.7860	5.9327	1.9341
CAW-AH	2.9929	3.9322	5.4074	1.7491	4.7581	5.7743	2.9160	4.8078	5.9797	1.9687
WAR	2.5504	3.9842	4.5689	1.6137	4.0532	4.6287	1.6348	4.0757	4.7149	1.6592
WAR-A	2.5502	3.9846	4.5642	1.6135	4.0532	4.6287	1.6348	4.0757	4.7149	1.6592
WAR-H	2.6141	3.9816	4.7387	1.6233	4.0110	4.7422	1.6461	4.0415	4.7971	1.6721
WAR-AH	2.5872	3.9923	4.6689	1.6190	4.0128	4.7147	1.6402	4.0386	4.7813	1.6651
EXP	2.1941	3.7901	5.0084	1.6602	3.7902	5.0091	1.6729	3.8561	5.1876	1.6766
EXP-A	2.1934	3.7695	4.8280	1.6283	3.7898	4.9905	1.6614	3.8561	4.9973	1.6783
EXP-L	2.3703	3.7524	4.6113	1.6209	3.7849	4.6815	1.6374	3.8170	5.0710	1.6813
EXP-AL	2.3619	3.7517	4.6164	1.6178	3.7837	4.6749	1.6306	3.8153	5.0612	1.6793
EXP-H	2.1948	3.7988	5.3729	1.6896	3.8031	5.3900	1.6897	3.8681	5.4781	1.7092
EXP-AH	2.1939	3.8003	5.2919	1.6882	3.8020	5.3356	1.6909	3.8666	5.4945	1.7073
CH	2.2600	3.7792	5.1872	1.6528	3.7798	5.2019	1.6671	3.8444	5.3421	1.6762
CH-A	2.2600	3.7572	5.0136	1.6175	3.7797	5.1831	1.6539	3.8445	5.1935	1.6685
CH-L	2.2509	3.7346	4.6633	1.5987	3.7357	4.6931	1.6073	3.8031	5.1612	1.6676
CH-AL	2.2511	3.7342	4.6603	1.6008	3.7375	4.6811	1.6110	3.8027	5.1620	1.6693
CH-H	2.2188	3.7745	5.5917	1.6848	3.7824	5.6133	1.6898	3.8465	5.6146	1.6988
CH-AH	2.2187	3.7763	5.5588	1.6851	3.7819	5.5610	1.6910	3.8459	5.6117	1.6993

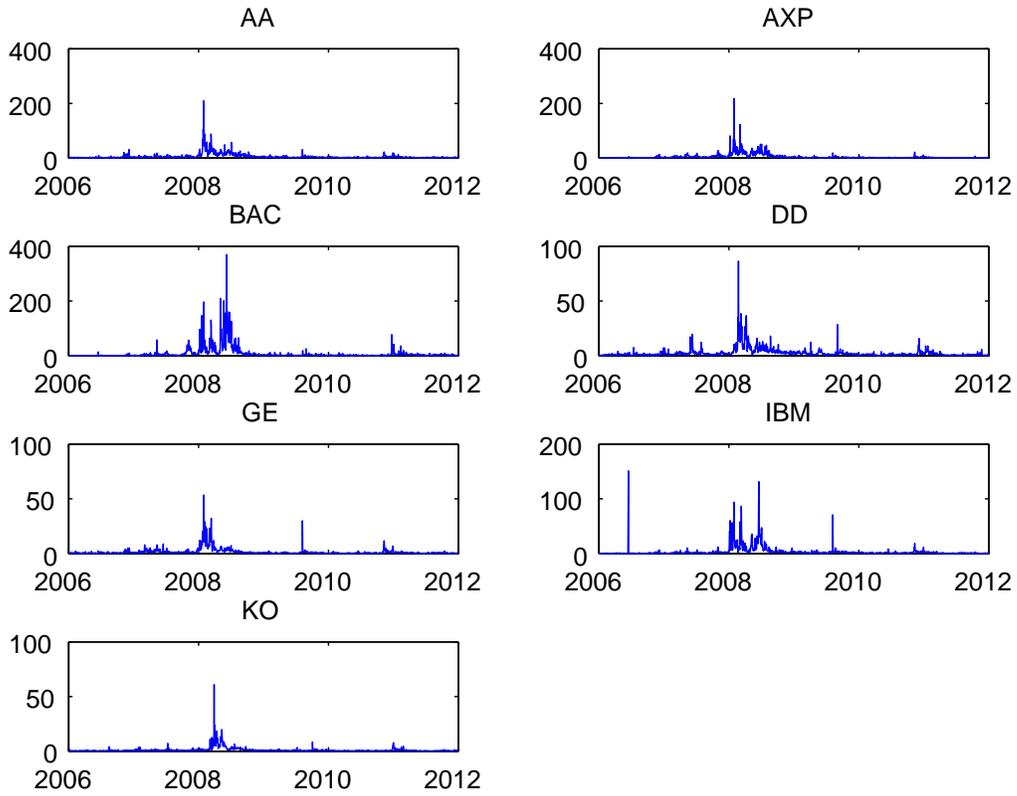
Note: 'B-GFC', 'D-GFC' and 'A-GFC' stand for 'Before GFC', 'During GFC' and 'After GFC'. For 5- and 10-step-ahead forecasts, the statistic is standardized by adjusting by the number of steps to make the results comparable.

Table 3: (Cont.) RMSE based on the Frobenius Norm of the Forecasting Error

Model	In-sample	1 day-ahead forecasts			5 day-ahead forecasts			10 day-ahead forecasts		
		B-GFC	D-GFC	A-GFC	B-GFC	D-GFC	A-GFC	B-GFC	D-GFC	A-GFC
5 factor MSV models										
DC	2.7771	3.8026	4.7625	1.7066	3.8402	4.7732	1.7164	3.8825	5.5601	1.7618
DC-cAhA	2.3503	3.8421	5.6042	1.7177	3.8645	4.8898	1.7494	4.0052	4.8928	1.7658
DC-cHvL	2.3612	3.8072	5.0369	1.6962	3.8083	5.0616	1.7182	3.8873	5.3640	1.7195
DC-cAHvAL	2.1398	3.7692	4.7689	1.6571	3.7977	4.9782	1.6874	3.8727	4.9836	1.7101
DC-cHvH	2.1373	3.8318	5.8305	1.7365	3.8467	5.8734	1.7450	3.9266	5.8824	1.7699
DC-cAHvAH	2.1302	3.8261	5.7057	1.7301	3.8319	5.7299	1.7310	3.9031	5.8989	1.7552
CAW	2.9490	3.8642	5.0261	1.9058	4.8888	5.7741	2.3074	4.1089	6.6824	1.8398
CAW-A	2.8732	3.8570	5.1365	1.9800	4.9455	5.8053	2.3311	4.1978	6.8520	1.8746
CAW-H	2.9997	4.0525	6.0349	1.7682	4.4054	5.9182	2.1278	4.4738	6.9780	1.9182
CAW-AH	2.9477	3.9112	5.8848	1.6662	5.2221	5.9633	2.7807	4.4332	6.6572	1.9439
WAR	2.5489	3.9447	4.5536	1.6042	3.9610	4.6183	1.6248	4.0326	4.7043	1.6488
WAR-A	2.5485	3.9440	4.5532	1.6040	3.9610	4.6183	1.6248	4.0326	4.7043	1.6488
WAR-H	2.6169	3.9431	4.7547	1.6163	3.9516	4.7762	1.6381	4.0106	4.7993	1.6632
WAR-AH	2.6185	3.9436	4.7580	1.6172	3.9519	4.7818	1.6387	4.0111	4.8013	1.6638
EXP	2.3227	3.7363	4.6218	1.6317	3.7829	4.6975	1.6472	3.7990	5.1123	1.6933
EXP-A	2.3120	3.7340	4.6220	1.6288	3.7813	4.6895	1.6419	3.7946	5.0959	1.6916
EXP-L	2.1247	3.7925	4.9899	1.6714	3.7959	5.0278	1.6881	3.8592	5.2902	1.6901
EXP-AL	2.1238	3.7695	4.8031	1.6402	3.7928	4.9795	1.6727	3.8600	5.0198	1.6900
EXP-H	2.1183	3.8017	5.4224	1.7080	3.8046	5.4385	1.7083	3.8683	5.5351	1.7266
EXP-AH	2.1174	3.8036	5.3638	1.7077	3.8038	5.3999	1.7099	3.8675	5.5578	1.7262
CH	2.2674	3.7916	5.1831	1.6827	3.7972	5.2093	1.6963	3.8572	5.4507	1.6999
CH-A	2.2672	3.7645	4.9832	1.6470	3.7916	5.1817	1.6837	3.8575	5.2085	1.6977
CH-L	2.1957	3.6927	4.6856	1.6130	3.7383	4.7124	1.6223	3.7594	5.2709	1.6831
CH-AL	2.1957	3.6921	4.6854	1.6109	3.7369	4.7165	1.6188	3.7583	5.2683	1.6824
CH-H	2.1438	3.7827	5.6585	1.7020	3.7829	5.6718	1.7056	3.8438	5.6985	1.7164
CH-AH	2.1439	3.7820	5.6414	1.7026	3.7844	5.6429	1.7061	3.8430	5.7004	1.7172

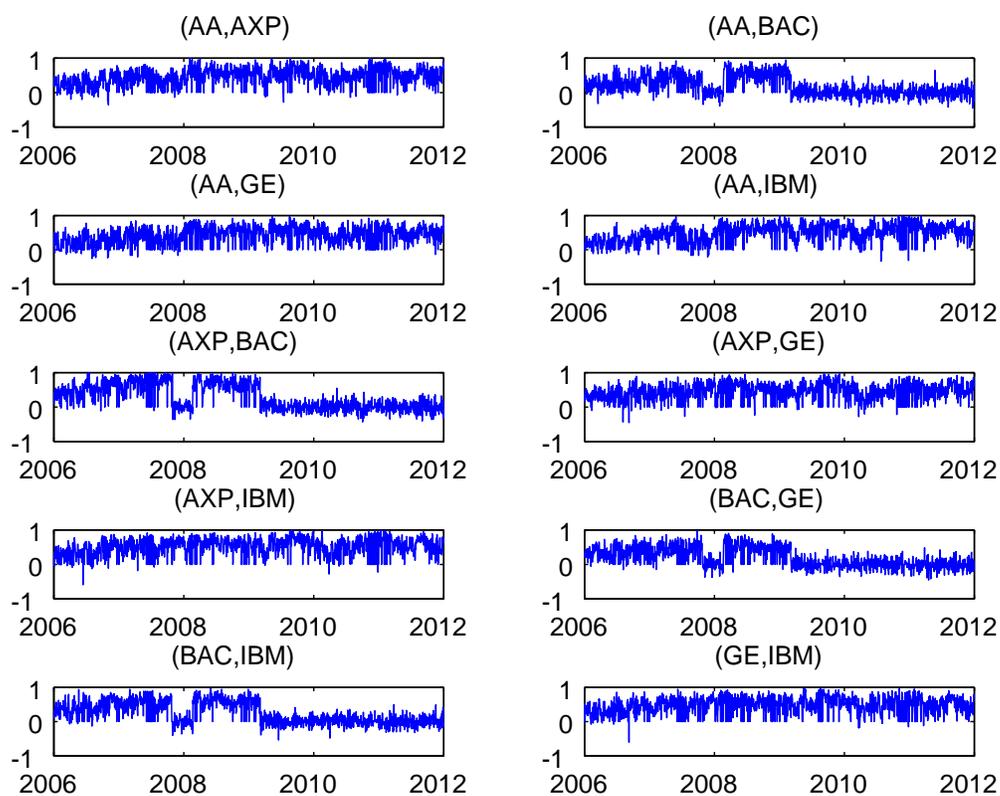
Note: ‘B-GFC’, ‘D-GFC’ and ‘A-GFC’ stand for ‘Before GFC’, ‘During GFC’ and ‘After GFC’. For 5- and 10-step-ahead forecasts, the statistic is standardized by adjusting by the number of steps to make the results comparable.

Figure 1: Time Series Plots of Realized Volatilities



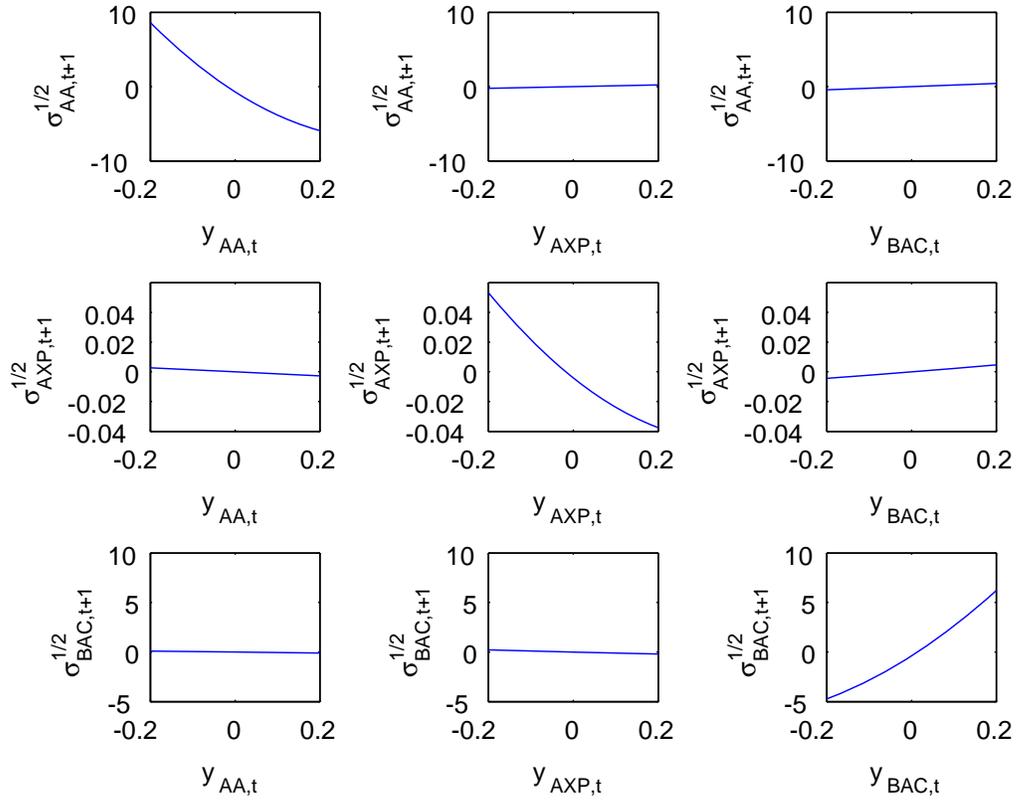
Note: The periods covers September 8, 2006 to October 4, 2012, giving 1,500 observations.

Figure 2: Time Series Plots of Realized Correlations



Note: The periods covers September 8, 2006 to October 4, 2012, giving 1,500 observations.

Figure 3: News Impact Curves from Returns to Future Volatilities



Note: Figure shows the news impact curves from  $y_{i,t}$  to  $\sigma_{j,t+1}^{1/2}$  ( $i, j = AA, AXP, BAC$ ).