SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

SMU
SINGAPORE MANAGEMENT
UNIVERSITY

Multivariate Stochastic Volatility

Manabu Asai, Michael McAleer, Jun Yu
December 2006

Paper No. 33-2006

Multivariate Stochastic Volatility*

Manabu Asai
Faculty of Economics
Tokyo Metropolitan University

Michael McAleer
School of Economics and Commerce
University of Western Australia

Jun Yu
School of Economics and Social Science
Singapore Management University

January 2005

* The authors wish to acknowledge helpful discussions with Yoshi Baba, Jiti Gao, Suhejla Hoti, Essie Maasoumi, Peter Phillips, Peter Robinson and Ruey Tsay. The first author appreciates the financial support of the Grant-in-Aid for the 21st Century COE program "Microstructure and Mechanism Design in Financial Markets" from the Ministry of Education, Culture, Sports, Science and Technology of Japan, the second author is grateful for the financial support of an Australian Research Council Discovery Grant, and the third author gratefully acknowledges financial support from the Wharton-SMU Research Centre at SMU.

Abstract

The literature on multivariate stochastic volatility (MSV) models has developed significantly over the last few years. This paper reviews the substantial literature on specification, estimation and evaluation of MSV models. A wide range of MSV models is presented according to various categories, namely (i) asymmetric models; (ii) factor models; (iii) time-varying correlation models; and (iv) alternative MSV specifications, including models based on the matrix exponential transformation, Cholesky decomposition, Wishart autoregressive process, and the empirical range. Alternative methods of estimation, including quasi-maximum likelihood, simulated maximum likelihood, Monte Carlo likelihood, and Markov chain Monte Carlo methods, are discussed and compared. Various methods of diagnostic checking and model comparison are also examined.

Keywords and phrases: multivariate stochastic volatility, asymmetry, leverage, thresholds, factor models, time-varying correlations, transformations, estimation, diagnostic checking, model comparison.

1. Introduction

A wide range of multivariate GARCH and stochastic volatility (SV) models has been developed, analysed and applied extensively in recent years to characterize the volatility that is inherent in financial time series data. Bauwens et al. (2004) provided a recent survey of multivariate GARCH**,** or conditional volatility, models. The GARCH literature has expanded considerably since the univariate ARCH process was developed by Engle (1982). The univariate SV model was proposed by, among others, Taylor (1982, 1986), and the univariate SV literature was surveyed in Ghysels et al. (1996).

Although there have already been many practical applications of multivariate GARCH models, the theoretical literature on multivariate stochastic volatility (MSV) models has developed significantly over the last few years. Some of the more important existing univariate and multivariate GARCH and SV models have been analysed in McAleer (2005). However, a comprehensive review of the important aspects of existing MSV models in the literature does not yet seem to exist. Owing to the development of a wide variety of MSV models in recent years, this paper reviews the substantial literature on the specification, estimation and evaluation of MSV models.

There are both economic and econometric reasons why multivariate volatility models are important. The knowledge of correlation structures is vital in many financial applications, such as optimal portfolio risk management and asset allocation, so that multivariate volatility models are useful for making financial decisions. Moreover, as financial volatility moves together across different assets and markets, modelling volatility in a multivariate framework can lead to greater statistical efficiency.

The remainder of the paper is organized as follows. Section 2 presents a range of MSV models according to various categories, including asymmetric models, factor models, time-varying correlation models, and several alternative specifications, including the matrix exponential transformation, Cholesky decomposition, Wishart autoregressive models, and an empirical range-based model. Section 3 compares and discusses alternative methods of estimation, including the quasi-maximum likelihood, simulated maximum likelihood**,** Monte Carlo likelihood, and Markov Chain Monte Carlo techniques. Various methods of diagnostic checking and model comparison are examined in Section 4. Some concluding comments are given in Section 5.

2. MSV Models

This section reviews several variants of MSV models according to four categories, as follows: (i) asymmetric models; (ii) factor models; (iii) time-varying correlation models; and (iv) alternative MSV specifications, including models based on the matrix exponential transformation, Cholesky decomposition, Wishart autoregressive process, and empirical range.

In what follows, $y = (y_1, \ldots, y_m)'$ denotes a vector of returns for m financial assets. For expositional purposes, it is assumed that the conditional mean vector of y is zero, although this can easily be relaxed. Moreover, exp(.) and log(.) denote the element-by-element exponential and logarithmic operators, respectively, and $diag\{x\} = diag\{x_1, \ldots, x_m\}$ denotes the $m \times m$ diagonal matrix with diagonal elements given by $x = (x_1, \ldots, x_m)'$.

2.1 Basic Models

The first MSV model proposed in the literature is due to Harvey et al. (1994), as follows:

$$\begin{aligned} y_t &= D_t \varepsilon_t, \\ D_t &= diag\left\{e^{h_{1t}/2}, \ldots e^{h_{mt}/2}\right\} = diag\left\{\exp\left(0.5 h_t\right)\right\}, \end{aligned} \tag{1}$$

$$h_{t+1} = \mu + \phi \circ h_t + \eta_t, \tag{2}$$

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} P_\varepsilon & O \\ O & \Sigma_\eta \end{pmatrix}\right], \tag{3}$$

where $h_t = (h_{1t}, \ldots, h_{mt})'$ is an $m \times 1$ vector of unobserved log-volatility, μ and ϕ are $m \times 1$ parameter vectors, the operator $\circ$ denotes the Hadamard (or element-by-element) product, $\Sigma_\eta = \left\{\sigma_{\eta,ij}\right\}$ is a positive-definite covariance matrix, and $P_\varepsilon = \{\rho_{ij}\}$ is the correlation matrix, that is, P_ε is a positive definite matrix with

$\rho_{ii} = 1$ and $|\rho_{ij}| < 1$ for any $i \neq j$, $i, j = 1, \ldots, m$. Harvey et al. (1994) also considered the multivariate t distribution for ε_t as this specification permits greater kurtosis as compared with the Gaussian assumption.

The model of Harvey et al. (1994) can easily be extended to a VARMA structure for h_t, as follows:

$$\Phi(L)h_{t+1} = \mu + \Theta(L)\eta_t,$$

where

$$\Phi(L) = I_m - \sum_{i=1}^{p} \Phi_i L^i,$$

$$\Theta(L) = I_m - \sum_{j=1}^{q} \Theta_j L^j,$$

and L is the lag operator.

Assuming that the off-diagonal elements of Σ_η are all equal to zero, the model corresponds to the constant conditional correlation (CCC) model proposed by Bollerslev et al. (1988) and Bollerslev (1990) in the framework of multivariate GARCH processes. In the CCC model, each conditional variance is specified as a univariate GARCH model, that is, with no spillovers from other assets, while each conditional covariance is given as a constant correlation coefficient times the product of the corresponding conditional standard deviations. If the off-diagonal elements of Σ_η are not equal to zero, then the elements of h_t are not independent.

Before introducing various MSV models, we present the long-memory MSV model analyzed by Anderson et al. (2003). By using a common degree of fractional integration, d, Anderson et al. (2003) specified the multivariate long-memory model as follows:

$$\Phi(L)(1-L)^d h_{t+1} = \mu + \eta_t . \qquad (4)$$

As their analysis depends on the realized value of volatility, h_t, estimation of this model is relatively straightforward, as follows: (i) estimate the common d using a multivariate extension of the GPH (Geweke and Porter-Hudak (1983)) estimator, as developed by Robinson (1995); and (ii) estimate the model by applying OLS to each equation separately.

2.2 Asymmetric Models

It has long been recognized that the volatility of many financial assets responds differently to bad news and good news. This is especially true for stock returns. In particular, while bad news tends to increase the future volatility, good news of the same size will increase the future volatility by a smaller amount, or may even decrease the future volatility. A popular explanation for this asymmetry is the leverage effect, as first proposed by Black (1976) (see also Christie (1982))**,** which predicts that volatility tends to decrease in response to good news but increase in response to bad news. Other forms of asymmetry, such as the asymmetric V-shape, have to be explained by reasons other than the leverage effect. Alternative reasons for the volatility asymmetry that has been suggested in the literature include the volatility feedback effect (Campbell and Hentschel (1992)).

The volatility asymmetry has been examined extensively in the context of univariate SV models. The news impact function (NIF) of Engle and Ng (1993) is a powerful tool for analysing the volatility asymmetry for GARCH-type models. The idea of the NIF is to examine the relationship between conditional volatility in period t+1 (defined by σ_{t+1}^2) and the standardized shock to returns in period t (defined by ε_t) in isolation. Yu (2004b) generalized the NIF for the relationship between $E(\ln \sigma_{t+1}^2 \mid \varepsilon_t)$ and ε_t in isolation, so that the NIF is also applicable to SV models.

It is now possible to review the various asymmetric MSV models according to the different shapes of the NIF.

2.2.1 Leverage Effect

Yu (2004a) defined the leverage effect to be a negative relationship between $E(\ln \sigma_{t+1}^2 \mid \varepsilon_t)$ and ε_t, holding everything else constant. According to this definition, the leverage effect must lead to a decreasing NIF.

A univariate discrete time SV model with the leverage effect was first proposed by Harvey and Shephard (1996), although Wiggins (1987) and Chesney and Scott (1989), among others, considered a continuous time model itself and discretized it for purposes of estimation. The model of Harvey and Shephard (1996) may be regarded as the Euler approximation to the continuous time SV model that is used widely in the option price literature (see, for example, Hull and White (1987), who generalized the Black-Scholes option pricing formula to analyse SV and leverage). These papers assume the negative correlation between the innovations. Yu (2004b) showed that, after rewriting the model in a Gaussian non-linear state space form with uncorrelated measurement and transition equation errors, the NIF is a straight line which slopes downwards.

An alternative discrete time SV model with "leverage effect" was proposed by Jacquier et al. (2004), which differs from the specification in Harvey and Shephard (1996) in how the correlation of two error processes is modelled. Yu (2004a) argued that it is difficult to interpret the leverage effect in the latter specification, whereas the interpretation of the leverage effect is straightforward in the former.

Danielsson (1998) and Chan et al. (2003) considered a multivariate extension of the model of Jacquier et al. (2004). The model is given by equations (1) and (2), together with

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} P_\varepsilon & L \\ L & \Sigma_\eta \end{pmatrix} \right], \qquad (5)$$

$$L = \{\lambda_{ij}\sigma_{\eta,ij}\},$$

where the parameter λ_{ij} captures asymmetry.[1] In the empirical analysis, Danielsson (1998) did not estimate the multivariate DL model because the data used in his analysis did not suggest any asymmetry in the estimated univariate models. Chan et al. (2003) employed the Bayesian Markov Chain Monte Carlo (MCMC) procedure to estimate the DL model. However, the argument of Yu (2004a) regarding the leverage effect also applies to the model in (5). Thus, the interpretation of the leverage effect in (5) is unclear and, even if $\lambda_{ii} < 0$, there is no guarantee that there will, in fact, be a leverage effect.

Asai and McAleer (2004b) considered a multivariate extension of the model of Harvey and Shephard (1996). The model is given by equations (1) and (2), together with

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} P_\varepsilon & L \\ L & \Sigma_\eta \end{pmatrix} \right],$$
$$L = diag\{\lambda_1 \sigma_{\eta,11}, \ldots, \lambda_m \sigma_{\eta,mm}\}, \tag{6}$$

where the parameter λ_i, $i = 1,\ldots,m$, is expected to be negative. Asai and McAleer (2004b) developed an estimation method for this MSV model based on the Monte Carlo likelihood (MCL) technique proposed in Durbin and Koopman (1997).

2.2.2 Threshold Effect

In the GARCH literature, Glosten et al. (1993) proposed modelling the asymmetric responses in conditional volatility using thresholds. In the univariate SV literature, So et al. (2002) proposed a threshold SV model, in which the constant term and the autoregressive parameter in the SV equation changed according to the sign of the previous return.

Although the multivariate threshold SV model has not yet been developed in the literature, it is straightforward to introduce a multivariate threshold SV model with

[1] The case $\lambda_{ii} < 0$ implies a leverage effect for asset i as it implies that bad news for asset i at period t tends to increase the volatility of asset j at period $t+1$.

mean equation given in (1), together with the volatility equation, as follows:

$$h_{t+1} = \mu(s_t) + \phi(s_t) \circ h_t + \eta_t, \tag{7}$$

where

$$\mu(s_t) = \left(\mu_1(s_{1t}), \ldots, \mu_m(s_{mt})\right)',$$

$$\phi(s_t) = \left(\phi_1(s_{1t}), \ldots, \phi_m(s_{mt})\right)',$$

and s_t is a state vector, with elements given by

$$s_{it} = \begin{cases} 0, & \text{if } y_{it} < 0, \\ 1, & \text{otherwise.} \end{cases} \tag{8}$$

It is straightforward to show that the NIF of a threshold SV model can have a flexible shape.

2.2.3 General Asymmetric Effect

Within the univariate framework, Danielsson (1994) suggested an alternative asymmetric specification to the leverage SV model, which is similar in spirit to an extension of the univariate exponential GARCH (EGARCH) model of Nelson (1991). The EGARCH model incorporated the absolute value function to capture the sign and magnitude of the previous normalized returns shocks to accommodate asymmetric behaviour.

In the model suggested by Danielsson (1994), which was termed the asymmetric leverage (AL) model in Asai and McAleer (2004a), two additional terms, $|y_t|$ and y_t, were included in the volatility equation, while the correlation between the two error terms was assumed to be zero. Asai and McAleer (2004a) noted that the absolute values of the previous realized returns were included because it was not computationally straightforward to incorporate the previous normalized shocks in the framework of SV models. Yu (2004b) proposed an alternative specification to Danielsson (1994) and

extended the volatility equation of the standard leverage SV model by including an additional term, namely $|y_t|$. Finally, Asai and McAleer (2004a) proposed a more general model which nests both the model of Yu (2004b) and the AL model as special cases. It is straightforward to show that the NIF of all three models can have a flexible shape.

In the multivariate context, Asai and McAleer (2004b) suggested an extension of the AL SV model of Danielsson (1994) as equation (1), together with

$$h_{t+1} = \mu + \gamma_1 \circ y_t + \gamma_2 \circ |y_t| + \phi \circ h_t + \eta_t, \tag{9}$$

where γ_1 and γ_2 are $p \times 1$ parameter vectors. Asai and McAleer (2004b) developed an estimation method for this MSV model, based on the Monte Carlo likelihood (MCL) technique.

2.3 Factor Models

In an attempt to reduce the dimensionality of the parameter space, various factor MSV models have been proposed. The factor MSV model was originally proposed by Harvey et al. (1994), and extended by Shephard (1996) and Jacquier et al. (1999). This model has several attractive features, including parsimony of the parameter space, and the ability to capture the common features in asset returns and volatilities. Alternative dynamic factor models are the latent factor ARCH model of Diebold and Nerlove (1989), and the factor ARCH representation used in Engle et al. (1990).

2.3.1 Additive Factor Models

The additive factor MSV model was first introduced by Harvey et al (1994), and subsequently extended in Jacquier et al (1995, 1999), Shephard (1996), Pitt and Shephard (1999a), and Aguilar and West (2000). The basic idea is borrowed from the factor multivariate ARCH models, where the returns are decomposed into two additive components. The first component has a smaller number of factors which captures the information relevant to the pricing of all assets, while the other one is idiosyncratic noise which captures the asset specific information (for further details, see Diebold and Nerlove (1989)).

Denote the $K \times 1$ vector of factors as f_t $(K < m)$, and D is an $m \times K$ dimensional matrix of factor loadings. The additive K factor MSV model presented by Jacquier et al (1995) can be written as

$$
\begin{aligned}
y_t &= Df_t + e_t, \\
f_{it} &= \exp(h_{it}/2)\varepsilon_{it}, \\
h_{i,t+1} &= \mu_i + \phi_i h_{it} + \eta_{it}, \quad i = 1, \ldots, K,
\end{aligned}
\tag{10}
$$

where e_t, ε_{it} and η_{it} are assumed to be mutually independent. In order to guarantee the identification of D and f_t uniquely, the restrictions $D_{ij} = 0$ and $D_{ii} = 1$ for $i = 1, \ldots, m$ and $j < i$ are usually adopted (see Aguilar and West (2000)).

Model (10) was extended in Pitt and Shephard (1999a) by allowing each element in e_t to evolve according to a univariate SV model. Chib et al. (2005) further extended the model by allowing for jumps and for idiosyncratic errors which follow the student t SV process.

Jacquier et al. (1999), Pitt and Shephard (1999a) and Aguilar and West (2000) proposed estimation methods based on single-move MCMC algorithms. Chib et al (2005) argued that the single-move algorithms can be simulation-inefficient and suggested a more efficient multi-move MCMC algorithm. Based on the Efficient Importance Sampling (EIS) method proposed by Richard and Zhang (2004), Liesenfeld and Richard (2004) developed an alternative multi-move MCMC algorithm. Liesenfeld and Richard (2003) showed that the EIS method can be used to approximate the likelihood function, so that it can facilitate a simulated ML approach.

Yu and Meyer (2004) showed that additive factor models accommodate both time varying volatility and time varying correlations. In the context of the bivariate one-factor SV model given by:

$$y_t = Df_t + e_t,\ \ e_t \sim N(0, diag\{\sigma_{e1}^2, \sigma_{e2}^2\})$$

$$f_t = \exp(h_t / 2)\varepsilon_t,\ \ \varepsilon_t \sim N(0,1),$$

$$h_{t+1} = \mu + \phi h_t + \eta_t,\ \ \eta_t \sim N(0,1),$$

Yu and Meyer (2004) derived the conditional correlation coefficient between y_{1t} and y_{2t} as

$$\frac{d}{\sqrt{(1+\sigma_{e1}^2 \exp(-h_t))(d^2 + \sigma_{e2}^2 \exp(-h_t))}},$$

where $(1, d)' = D$. It is clear from the above expression that the correlation depends on the volatility of the factor.

Philipov and Glickman (2004a) proposed a high-dimensional additive factor MSV model in which the factor covariance matrices are driven by Wishart random processes, as follows:

$$y_t = Df_t + e_t,\ \ e_t \sim N(0, \Omega),$$

$$f_t \mid V_t \sim N(0, V_t),$$

$$V_t^{-1} \mid V_{t-1}^{-1}, A, v,\ \ d \sim Wish(v, S_{t-1}),$$

$$S_{t-1} = \frac{1}{v} A^{1/2} (V_{t-1}^{-1})^d (A^{1/2})',$$

where V_t^{-1} is a matrix of factor volatility, A is a symmetric positive definite matrix, d is a scalar persistence parameter, *Wish* is the Wishart distribution, v is the degrees of freedom parameter of the Wishart distribution. A Bayesian MCMC algorithm is developed to estimate the model.

2.3.1 Multiplicative Factor Models

The multiplicative factor MSV model, also known as the stochastic discount factor model, was considered in Quintana and West (1987). The one-factor model from this class decomposes the returns into two multiplicative components, a scalar common factor and a vector of idiosyncratic noise, as follows:

$$\begin{aligned} y_t &= \exp(h_t / 2)\varepsilon_t, \\ h_{t+1} &= \mu + \phi(h_t - \mu) + \eta_t. \end{aligned} \tag{11}$$

The first element in Σ_ε is assumed to be one for purposes of identification. Compared with the basic MSV model, this model has a smaller number of parameters, which makes it more convenient computationally. Unlike the additive factor MSV model, however, the correlations are now invariant with respect to time. Moreover, the correlation in log-volatilities is always equal to one.

Ray and Tsay (2000) extended the one-factor model to a *k*-factor model, in which long range dependence is accommodated in the factor volatility:

$$y_t = \exp(h_t' v_t / 2)\varepsilon_t, \ \ \varepsilon_t \sim N(0, P_\varepsilon),$$

$$(1-L)^d h_t = \mu + \eta_t,$$

where v_t is an $(m \times k)$ matrix of rank *k*, with $k < m$.

2.4 Time-Varying Correlation Models

The assumption of constant correlations in the correlation matrix P_ε in equation (3) can be relaxed by considering the time-varying correlation matrix, $P_{\varepsilon t} = \{\rho_{ij,t}\}$, where $\rho_{ii,t} = 1$ and $\rho_{ij,t} = \rho_{ji,t}$.

Following the suggestion made by Tsay (2002) and Christodoulakis and Satchell (2002) in the bivariate GARCH framework, Yu and Meyer (2004) proposed that the Fisher transformation of $\rho_{12,t}$ could be modelled in a bivariate SV framework, as follows:

$$\begin{aligned} \rho_{12,t} &= \left(\exp(v_t)-1\right)\big/\left(\exp(v_t)+1\right), \\ v_{t+1} &= \mu_v + \varphi(v_t - \mu_v) + u_t, \quad u_t \sim N(0,\sigma_u^2). \end{aligned} \tag{12}$$

The first equality in (12) guarantees that $|\rho_{12,t}|<1$. Yu and Meyer (2004) estimated the bivariate model in equations (1), (2) and (12) using the Bayesian MCMC method. The obvious drawback with this specification is the difficulty in generalizing it to a higher-dimension.

In order to develop an MSV model which accommodates time-varying correlation, Asai and McAleer (2004c) and Yu and Meyer (2004) suggested two alternative SV extensions of the dynamic conditional correlation (DCC) model of Engle (2002) (see the VCC model of Tse and Tsui (2002) for a related development). Suppose that y_t, conditional on Σ_t, has a multivariate normal distribution, $N(0,\Sigma_t)$, where the covariance matrix is given by

$$\Sigma_t = D_t \Gamma_t D_t . \tag{13}$$

In equation (13), the time-varying correlation matrix is given by Γ_t, while the diagonal matrix D_t is defined by equations (1) and (2). For the DCC model, Γ_t is specified as follows:

$$\Gamma_t = Q_t^{*-1} Q_t Q_t^{*-1},$$

where $Q_t^* = \left(diag\left\{\text{vecd}\left(Q_t\right)\right\}\right)^{1/2}$, by using some positive definite matrix Q_t.

Asai and McAleer (2004c) extended the DCC model by specifying Q_t as follows:

$$\begin{aligned} Q_{t+1} &= (1-\psi)\bar{Q} + \psi Q_t + \Xi_t, \\ \Xi_t &\sim W_m(\nu, \Lambda), \end{aligned} \tag{14}$$

where $W_m(\nu, \Lambda)$ denotes a Wishart distribution. This dynamic correlation MSV model guarantees the positive definiteness of Γ_t under the assumption that $\bar{Q}$ is positive definite and $|\psi| < 1$. The latter condition also implies that the time-varying correlations are mean reverting. In the special case where $\nu = 1$, Ξ_t can be expressed as the cross-product of a multivariate normal variate with mean zero and covariance matrix given by Λ.

Yu and Meyer (2004) proposed an alternative MSV extension of DCC by specifying Q_t as follows:

$$\begin{aligned} Q_{t+1} &= S + B \circ (Q_t - S) + A \circ (e_t e_t' - S) \\ &= S \circ (\iota\iota' - A - B) + B \circ Q_t + A \circ e_t e_t', \end{aligned} \tag{15}$$

where $e_t \sim N(0, I_m)$. According to Ding and Engle (2001) and Engle (2002), if A, B and $(\iota\iota' - A - B)$ are positive semi-definite, then Q_t will also be positive semi-definite. Moreover, if any one of the matrices is positive definite, then Q_t will also be positive definite.

2.5 Alternative Specifications

This sub-section introduces four alternative MSV models based on the matrix exponential transformation, the Cholesky decomposition, the Wishart autoregressive process, and the observed range**,** respectively.

2.5.1 Matrix Exponential Transformation

Chiu et al. (1996) proposed a general framework for the logarithmic covariance matrix based on the matrix exponential transformation, which is well known in the mathematics literature (see, for example, Bellman (1970)). In this sub-section, we denote Exp(.) as the matrix exponential operation to distinguish it from the standard exponential operation. For any $m \times m$ matrix A, the matrix exponential transformation is defined by the power series expansion:

$$\text{Exp}(A) = \sum_{s=0}^{\infty} (1/s!) A^S, \tag{16}$$

where A^0 reduces to the $m \times m$ identity matrix and A^s denotes the standard matrix multiplication of A s times. Thus, in general, the elements of $\text{Exp}(A)$ do not typically exponentiate the elements of A.

The properties of the matrix exponential and matrix logarithm are summarized in Chiu et al. (1996). For any real symmetric matrix A, we note the singular value decomposition $A = TDT'$, where the columns of the $m \times m$ orthonormal matrix T denote the appropriate eigenvectors of A, and D is an $m \times m$ diagonal matrix, with elements equal to the eigenvalues of A. Therefore, $\text{Exp}(A) = T\,\text{Exp}(D)T'$, where $\text{Exp}(D)$ is an $m \times m$ diagonal matrix, with diagonal elements equal to the exponential of the corresponding eigenvalues of A. If it is assumed that $\Sigma_{\varepsilon t} = \text{Exp}(A_t)$ for any symmetric matrix A_t, then Σ_t is positive definite.

Similarly, the matrix logarithmic transformation, $\text{Log}(B)$, for any $m \times m$ positive definite matrix, B, is defined by using the spectral decomposition of B.

Using the matrix exponential operator, we propose the following model:

$$\begin{aligned} y_t &\sim N(0, \Sigma_t), \\ \Sigma_t &= \text{Exp}(A_t), \end{aligned} \tag{17}$$

where $\alpha_t = \text{vech}(A_t)$ is a vector autoregressive process, as follows:

$$\alpha_{t+1} = \mu + \Upsilon x_t + \phi \circ \alpha_t + \eta_t,$$
$$\eta_t \sim N(0, \Sigma_\eta), \tag{18}$$

with $x_t = \left(y_t', |y_t'|\right)'$, $n \times 1$ parameter vectors μ and ϕ, where $n = 0.5m(m+1)$, $n \times n$ covariance matrix Σ_η, and an $n \times 2m$ matrix of parameters Υ. A limitation of this specification is that it is not straightforward to interpret the relationship between the elements of Σ_t and A_t.

2.5.2 Cholesky Decomposition

One of the most serious difficulties in modelling multivariate volatility is to ensure that the covariance matrix is positive semi-definiteness (see, for example, Engle and Kroner (1995)). Although the conventional approach is to impose suitable parametric restrictions, Tsay (2002) advocated an alternative approach, which uses the Cholesky decomposition. For a symmetric, positive-definite matrix Σ_t, the Cholesky decomposition factors the matrix Σ_t uniquely in the form $\Sigma_t = L_t G_t L_t^{'}$, where L_t is a lower triangular matrix with unit diagonal elements, and G_t is a diagonal matrix with positive elements.

The MSV model of Tsay (2002) is given as follows:

$$y_t \mid \Sigma_t \sim N(0, \Sigma_t),$$

$$\Sigma_t = L_t G_t L_t^{'},$$

$$L_t = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ q_{21,t} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ q_{m1,t} & q_{m2,t} & \cdots & 1 \end{bmatrix},$$

$$G_t = diag\{g_{11,t}, \ldots, g_{mm,t}\} = diag\{\exp(h_{1t}), \ldots, \exp(h_{mt})\},$$

$$h_{it+1} = \mu_i + \phi_i h_{it} + \eta_{it}, \ i = 1, \ldots, m,$$

$$q_{ij,t+1} = \alpha_{ij} + \beta_{ij} q_{ij,t} + u_{ij,t}, \ i > j.$$

The elements in G_t are always positive due to the exponential transformation. Consequently, the Cholesky decomposition guarantees the positive semi-definiteness of Σ_t. It can be seen that the elements in L_t and G_t are assumed to follow an AR(1) process. Moreover, it is straightforward to derive the relationship between the variances and correlations, on the one hand, and the variables in L_t and G_t, on the other, as follows:

$$\sigma_{ii,t} = \sum_{k=1}^{i} q_{ik,t}^2 g_{kk,t}, \ i = 1, \ldots, m,$$

$$\sigma_{ij,t} = \sum_{k=1}^{j} q_{ik,t} q_{jk,t} g_{kk,t}, \ i > j, \ i = 2, \ldots, m,$$

$$\rho_{ij,t} = \frac{\sigma_{ij,t}}{\sqrt{\sigma_{ii,t}\sigma_{jj,t}}} = \frac{\sum_{k=1}^{j} q_{ik,t} q_{jk,t} g_{kk,t}}{\sqrt{\sum_{k=1}^{i} q_{ik,t}^2 g_{kk,t} \sum_{k=1}^{j} q_{jk,t}^2 g_{kk,t}}}.$$

It is clear from these expressions that the dynamics in $g_{ii,t}$ and $q_{ij,t}$ are the

driving forces underlying the time-varying volatility and the time-varying correlation. However, the dynamics underlying volatility are not determined separately from those associated with the correlations, as both are dependent on their corresponding AR(1) processes. This restriction is, at least in spirit, similar to that associated with factor MSV models.

2.5.3 Wishart Autoregressive Models

Gourieroux et al. (2004) proposed the Wishart autoregressive (WAR) multivariate process of stochastic positive semi-definite matrices to develop an altogether different type of dynamic MSV model. Let Σ_t denote a time-varying covariance matrix of y_t. Gourieroux et al. (2004) defined the WAR(p) process, as follows:

$$\Sigma_t = \sum_{k=1}^{K} x_{kt} x_{kt}^{'} \tag{19}$$

where $K > m-1$ and each x_{kt} follows the VAR(p) model, given by:

$$x_{kt} = \sum_{i=1}^{p} A_i x_{k,t-i} + \varepsilon_{kt}, \quad \varepsilon_{kt} \sim N(0, \Sigma).$$

By using the realized value of volatility, Gourieroux et al. (2004) estimated the parameters of the WAR(1) process using a two-step procedure based on nonlinear least squares.

Philipov and Glickman (2004b) suggested an alternative model based on Wishart processes, as follows:

$$\begin{aligned} &y_t \mid \Sigma_t \sim N(0, \Sigma_t), \\ &\Sigma_t^{-1} \mid \nu, S_{t-1} \sim W_m(\nu, S_{t-1}), \end{aligned} \tag{20}$$

where ν and S_t are the degrees of freedom and the time-dependent scale parameter of the Wishart distribution, respectively. With a time-invariant covariance structure, the above model may be considered as a traditional Normal-Wishart representation of the

behavior of multivariate returns. However, Philipov and Glickman (2004b) introduced time variation in the scale parameter, as follows:

$$S_t = \frac{1}{\nu}\left(A^{1/2}\right)\left(\Sigma_t^{-1}\right)^d \left(A^{1/2}\right),$$

where A is a positive definite symmetric parameter matrix that is decomposed through a Cholesky decomposition as $A = \left(A^{1/2}\right)\left(A^{1/2}\right)'$, and d is a scalar parameter. The quadratic expression ensures that the covariance matrices are symmetric positive definite. Philipov and Glickman (2004b) estimated the parameters of the above model using the Bayesian MCMC technique.

2.5.4 Range-Based Model

Tims and Mahieu (2003) proposed a range-based MSV model. As the range can be used as a measure of volatility, which is observed (or realized) when the high and low prices are recorded, Tims and Mahieu (2003) suggested a multivariate model for volatility directly, as follows:

$$\log(range_t) = D' f_t + \varepsilon_t, \ \ \varepsilon_t \sim N(0, \Sigma),$$

$$f_{t+1} = \Phi f_t + \eta_t, \ \ \eta_t \sim N(0, \Sigma_\eta).$$

As the volatility is not latent in this model, efficient estimation of the parameters is achieved through the use of the Kalman filter. It is not known, however, how to use this model for purposes of asset pricing.

3. Estimation

As SV models typically do not have a closed-form expression for the likelihood function, the estimation of the parameters for the wide range of univariate and multivariate SV models has attracted significant attention in the literature.

An important concern for the choice of a particular estimation method lies in its

efficiency. In addition to efficiency, other important issues related to estimation include: (1) estimation of the latent volatility; (2) determination of the optimal filtering, smoothing and forecasting methods; (3) computational efficiency; (4) applicability for flexible modelling. Broto and Ruiz (2004) provided a recent survey regarding the numerous estimation techniques for SV models, with an emphasis on univariate SV models and methods. Some of these techniques have also been applied to the estimation of MSV models.

3.1 Quasi-Maximum Likelihood

In order to estimate the parameters of the model (1)-(3), Harvey et al. (1994) proposed a Quasi-Maximum Likelihood (QML) method based on the property that the transformed vector $y_t^* = (\ln y_{1t}^2, \ldots, \ln y_{mt}^2)'$ has a state space form with the measurement equation given by:

$$\begin{aligned} y_t^* &= h_t + \xi_t, \\ \xi_t &= \ln \varepsilon_t^2 = (\ln \varepsilon_{1t}^2, \ldots, \ln \varepsilon_{mt}^2)' \end{aligned} \tag{21}$$

and the transition equation (2). The measurement equation errors, ξ_t, are non-normal, with mean vector $E(\xi_t) = -1.2793\iota$, where ι is an $m \times 1$ vector of unit elements. Harvey et al. (1994) showed that the covariance matrix of ξ_t, denoted Σ_ξ, is given by $\Sigma_\xi = (\pi^2/2)\{\rho_{ij}^*\}$, where $\rho_{ii}^* = 1$ and

$$\rho_{ij}^* = \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{(n-1)!}{(1/2)_n n} \rho_{ij}^{2n}, \tag{22}$$

where $(x)_n = x(x+1)\cdots(x+n-1)$ and ρ_{ij} is defined by (3). Treating ξ_t as a Gaussian error term, QML estimates may be obtained by applying the Kalman filter to equations (2) and (21). Taking account of the non-normality in ξ_t, the asymptotic standard errors can be obtained by using the results established in Dunsmuir (1979).

If ρ_{ij}^{*} can be estimated, then it is also possible to estimate the absolute value of ρ_{ij}, and the cross-correlations between the different values of ε_{it}. Estimation of the signs of ρ_{ij} may be obtained by returning to the untransformed observations, and noting that the sign of each of the pairs, $\varepsilon_{it}\varepsilon_{jt}$ $(i, j = 1, \ldots, m)$, will be the same as the corresponding pairs of observed values, $y_{it}y_{jt}$. Therefore, the sign of ρ_{ij} is estimated as positive if more than one-half of the pairs, $y_{it}y_{jt}$, is positive.

One of the main features of this transformation is that ξ_t and η_t are uncorrelated even if the original ε_t and η_t are correlated (see Harvey et al. (1994)). Since the leverage effects assume a negative correlation between ε_t and η_t, as in equation (5), the transformation may ignore the information regarding the leverage effects. In the univariate case, Harvey and Shephard (1996) recovered it in the state space form, given the signs of the observed values. As for the multivariate case, Asai and McAleer (2004b) derived the state space form for the leverage effects in the model (1), (2) and (5), based on pairs of the signs of y_{it} and y_{jt}. This representation enables use of the QML method based on the Kalman filter. However, Asai and McAleer (2004b) adopted the Monte Carlo likelihood method for purposes of efficient estimation.

The main advantages of the QML method are that it is computationally convenient, and also straightforward for purposes of filtering, smoothing, and forecasting. Unfortunately, the available (though limited) Monte Carlo experiments in the context of the basic univariate SV model suggest that the QML method is generally less efficient than the Bayesian MCMC technique and the likelihood approach based on Monte Carlo simulation (for further details, see Jacquier et al. (1994) and the discussions contained therein). It is natural to believe that inefficiency remains for the QML method relative to the Bayesian MCMC technique and the likelihood approach in the multivariate context,

although no Monte Carlo evidence is available to date.

3.2 Simulated Maximum Likelihood

One Simulated Maximum Likelihood (SML) method is the Accelerated Gaussian Importance Sampling (AGIS) approach, as developed in Danielsson and Richard (1993). The AGIS approach is designed to estimate dynamic latent variable models, whereby Monte Carlo methods are used to integrate the latent variables out of the joint density of the latent and observable variables to obtain the marginal densities of the observable variables. Danielsson's comments on Jacquier et al. (1994) show that the finite sample property of this SML estimator is close to that of the Bayesian MCMC method. As for MSV models, Danielsson (1998) applied the AGIS approach to estimate the parameters of the MSV model in (1)-(3). It seems difficult to extend the AGIS approach to accommodate more flexible SV models, as the method is specifically designed for models with a latent Gaussian process.

While the AGIS technique has limited applicability, the Efficient Importance Sampling (EIS) procedure proposed by Richard and Zhang (2004), and applied by Liesenfeld and Richard (2003, 2004), is applicable to models with more flexible classes of distributions for the latent variables. As in the case of AGIS, EIS is a Monte Carlo technique for the evaluation of high-dimensional integrals. The EIS relies on a sequence of simple low-dimensional least squares regressions to obtain a very accurate global approximation of the integrand. This approximation leads to a Monte Carlo sampler, which produces highly accurate Monte Carlo estimates of the likelihood. In order to estimate the parameters of the additive factor model (10) with one factor, Liesenfeld and Richard (2003) proposed an SML approach by approximating the likelihood function based on EIS, while Liesenfeld and Richard (2004) developed a multi-move MCMC algorithm by sampling the latent variables based on EIS.

Let λ_t denote a q-dimensional vector of latent variables, and $f(Y,\Lambda;\theta)$ be the joint density of $Y=\{y_t\}_{t=1}^{T}$ and $\Lambda=\{\lambda_t\}_{t=1}^{T}$. The likelihood function associated with the observable variables, Y, is given by the $(T\times q)$-dimensional integral $L(\theta;Y)=\int f(Y,\Lambda;\theta)d\Lambda$. The likelihood function can be factorized as:

$$L(\theta;Y)=\int\prod_{t=1}^{T}f(y_t,\lambda_t\mid Y_{t-1},\Lambda_{t-1},\theta)d\Lambda=\int\prod_{t=1}^{T}g(y_t\mid\lambda_t,Y_{t-1},\theta)\,p(\lambda_t\mid Y_{t-1},\Lambda_{t-1},\theta)d\Lambda,$$

where $Y_t=\{y_s\}_{s=1}^{t}$ and $\Lambda_t=\{\lambda_s\}_{s=1}^{t}$. It should be noted that the second equality implies that y_t is independent of Λ_{t-1}, given (λ_t,Y_{t-1}), which is a standard assumption in the analysis of SV models. Although it is assumed, for notational convenience, that the initial values $\{y_0,y_{-1},\ldots\}$ and $\{\lambda_0,\lambda_{-1},\ldots\}$ are known constants, this condition can be relaxed. A Monte Carlo estimate of $L(\theta;Y)$ based on the above factorization is given by:

$$\hat{L}(\theta;Y)=\frac{1}{N}\sum_{i=1}^{N}\left\{\prod_{t=1}^{T}g(y_t\mid\tilde{\lambda}_t^{(i)}(\theta),Y_{t-1},\theta)\right\},$$

where the $\left\{\tilde{\lambda}_t^{(i)}(\theta)\right\}_{t=1}^{T}$ are samples drawn from the conditional density $p(\lambda_t\mid Y_{t-1},\tilde{\Lambda}_{t-1}^{(i)},\theta)$. This Monte Carlo estimate is inefficient in the sense that $\tilde{\lambda}_t^{(i)}(\theta)$ has no relation to the actual value λ_t as it does not use any information about y_t to generate $\tilde{\lambda}_t^{(i)}(\theta)$ from $p(\lambda_t\mid Y_{t-1},\tilde{\Lambda}_{t-1}^{(i)},\theta)$.

In order to cope with this problem, the AGIS proposed by Danielsson and Richard (1993) and Danielsson (1994), and the EIS suggested by Richard and Zhang (2004), consider an auxiliary sampler, $m(\lambda_t\mid\Lambda_{t-1},a_t)$. These methods enable the factorization given above to be rewritten as:

$$L(\theta;Y)=\int\prod_{t=1}^{T}\left[\frac{f(y_t,\lambda_t\mid Y_{t-1},\Lambda_{t-1},\theta)}{m(\lambda_t\mid\Lambda_{t-1},a_t)}\right]\prod_{t=1}^{T}m(\lambda_t\mid\Lambda_{t-1},a_t)d\Lambda,$$

which yields an importance sampling Monte Carlo estimate, as follows:

$$\tilde{L}(\theta;Y)=\frac{1}{N}\sum_{i=1}^{N}\left[\prod_{t=1}^{T}\frac{f(y_t,\tilde{\lambda}_t^{(i)}(a_t)\mid Y_{t-1},\tilde{\Lambda}_{t-1}^{(i)}(a_{t-1}),\theta)}{m(\tilde{\lambda}_t^{(i)}(a_t)\mid\tilde{\Lambda}_{t-1}^{(i)}(a_{t-1}),a_t)}\right],\qquad(23)$$

where the $\left\{\tilde{\lambda}_t^{(i)}(a_t)\right\}_{t=1}^{T}$ are samples drawn from importance density $m(\lambda_t \mid \tilde{\Lambda}_{t-1}^{(i)}, a_t)$.

The EIS method denotes an approximation of the density $f(y_t, \lambda_t \mid Y_{t-1}, \Lambda_{t-1}, \theta)$ as $\kappa(\Lambda_t; a_t)$, and constructs the auxiliary density $m(\lambda_t \mid \Lambda_{t-1}, a_t)$, as follows:

$$m(\lambda_t \mid \Lambda_{t-1}, a_t) = \frac{\kappa(\Lambda_t; a_t)}{\chi(\Lambda_{t-1}; a_t)},$$

where $\chi(\Lambda_{t-1}; a_t) = \int \kappa(\Lambda_t; a_t) d\lambda_t$. It should be noted that matching $f(y_t, \lambda_t \mid Y_{t-1}, \Lambda_{t-1}, \theta)$ with $\kappa(\Lambda_t; a_t)$ may leave $\chi(\Lambda_{t-1}; a_t)$ unexplained. As $\chi(\Lambda_{t-1}; a_t)$ does not depend on λ_t, it can be transferred back to the period $t-1$ minimization sub-problem. Taken together, EIS requires solving a simple back-recursive sequence of low-dimensional least squares problems of the form:

$$\hat{a}_t = \arg \min_{a_t} \sum_{i=1}^{N} \left\{ \log\left[f\left(y_t, \tilde{\lambda}_t^{(i)}(\theta) \mid Y_{t-1}, \Lambda_{t-1}^{(i)}(\theta), \theta\right) \cdot \chi\left(\tilde{\Lambda}_t^{(i)}(\theta); \hat{a}_{t+1}\right)\right] - c_t - \log \kappa\left(\tilde{\Lambda}_t^{(i)}(\theta); \hat{a}_{t+1}\right) \right\}$$

for $t: T \to 1$, with $\chi(\Lambda_T; a_{T+1}) \equiv 1$. The c_t are unknown constants to be estimated jointly with the unknown a_t. In order to obtain highly efficient importance samplers, a small number of iterations of the EIS algorithm is required. When such iterations converge to fixed values of the auxiliary parameters, $\hat{a}_t$, this would be expected to produce optimal importance samplers. Finally, the EIS estimate of the likelihood function for a given value of θ is obtained by substituting $\{\hat{a}_t\}_{t=1}^{T}$ for $\{a_t\}_{t=1}^{T}$ in equation (23).

Moreover, the EIS method can be used to compute the filtered estimates of the latent variables. Let $h(\lambda_t)$ denote a function such as, for example, $\exp(\lambda_t)$, which represents the conditional return variance. Then the sequence of conditional expectations of $h(\lambda_t)$, given Y_{t-1}, the past observations of the returns, provides a sequence of filtered estimates of $h(\lambda_t)$. For the SV model, these expectations take the

form of a ratio of integrals, as follows:

$$E\left[h(\lambda_t)\,|\,Y_{t-1}\right]=\frac{\int h(\lambda_t)\,p(\lambda_t\,|\,\Lambda_{t-1},Y_{t-1},\theta)f(Y_{t-1},\Lambda_{t-1};\theta)d\Lambda_t}{\int f(Y_{t-1},\Lambda_{t-1};\theta)d\Lambda_{t-1}},$$

in which both the numerator and denominator can be estimated by the EIS algorithm.

In the application of the Bayesian MCMC method, Liesenfeld and Richard (2004) proposed using a combination of the EIS-sampler with Tierney's (1994) Acceptance-Rejection Metropolis-Hastings (AR-MH) algorithm to simulate $\Lambda\,|\,Y,\theta$. The basis of such a procedure is the fact that the EIS density for Λ provides a very close approximation to $f(\Lambda\,|\,Y,\theta)$.

3.3 Monte Carlo Likelihood

The Monte Carlo likelihood (MCL) approach for non-Gaussian models is based on importance sampling techniques, so that the method may be classified as an SML method. The MCL method can approximate the likelihood function to an arbitrary degree of accuracy by decomposing it into a Gaussian part, which is constructed by the Kalman filter, and a remainder function, whose expectation is evaluated through simulation.

Durbin and Koopman (1997) demonstrated that the log-likelihood function of state space models with non-Gaussian measurement disturbances could be expressed simply as

$$\ln L(y\,|\,\theta)=\ln L_G(y\,|\,\theta)+\ln E_G\left[\frac{p_\xi(\xi\,|\,\theta)}{p_G(\xi\,|\,y,\theta)}\right], \tag{24}$$

where $y=(y_1,\ldots y_T)'$, $y_t=(y_{1t},\ldots y_{mt})'$, and $\xi=(\xi_1,\ldots,\xi_T)'$ and $\ln L_G(y\,|\,\theta)$ are the vectors of measurement disturbances and the log-likelihood function of the approximating Gaussian model, respectively, $p_\xi(\xi\,|\,\theta)$ is the true density function, $p_G(\xi\,|\,y,\theta)$ is the Gaussian density of the measurement disturbances of the

approximating model, and E_G refers to the expectation with respect to the 'so-called' importance density $p_G(\xi \mid y, \theta)$ associated with the approximating model. Equation (24) shows that the non-Gaussian log-likelihood function can be expressed as the log-likelihood function of the Gaussian approximating model plus a correction for the departures from the Gaussian assumptions relative to the true model.

A key feature of the MCL method is that only the minor part of the likelihood function requires simulations, unlike other SML methods. Therefore, the method is computationally efficient in the sense that it needs only a small number of simulations to achieve the desirable accuracy for empirical analysis.

The MCL estimates of the parameters, θ, are obtained by numerical optimization of the unbiased estimate of equation (24). The log-likelihood function of the approximating model, $\ln L_G(y \mid \theta)$, can be used to obtain the starting values. Sandmann and Koopman (1998) is the first paper to have used this MCL approach in the SV literature. Asai and McAleer (2004b) developed the MCL method for asymmetric MSV models. As noted in Asai (2004), this MCL method is also able to accommodate the additive factor MSV model.

3.4 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) methods are used widely in the SV literature, following the development in Jacquier et al. (1994), which has been greatly refined and simplified by Shephard and Pitt (1997) and Kim et al. (1998). The idea behind MCMC methods is to produce variates from a given multivariate density (the posterior density in Bayesian applications) by repeatedly sampling a Markov chain whose invariant distribution is the target density of interest. The MCMC method focuses on the density $\pi(\theta, h \mid y)$ instead of the usual posterior density, $\pi(\theta \mid y)$, since the latter requires computation of the likelihood function $f(y \mid \theta) = \int f(y \mid h, \theta) f(h \mid \theta) dh$. The MCMC procedure only requires alternating back and forth between drawing from $f(h \mid \theta, y)$ and $f(\theta \mid h, y)$. This process of alternating between conditional distributions produces a cyclic chain.

As for the property of sample variates from an MCMC algorithm, they are a high-dimensional sample from the target density of interest. These draws can be used as the basis for drawing inferences by appealing to suitable ergodic theorems for Markov chains. For example, posterior moments and marginal densities can be estimated (or simulated consistently) by averaging the relevant function of interest over the sampled variates. The posterior mean of θ is estimated simply as the sample mean of the simulated θ values. These estimates can be made arbitrarily accurate by increasing the simulation sample size.

One particularly important technical advantage of the Bayesian MCMC method over classical inferential techniques is that MCMC does not need to use numerical optimization. This advantage becomes especially important when the number of parameters to be estimated is large, as in the application of MSV models to the analysis of financial data.

Jacquier et al. (1999), Pitt and Shephard (1999a), and Aguilar and West (2000) have applied the MCMC procedure to estimate additive factor MSV models, while Yu and Meyer (2004) compared various MSV models. Moreover, Yu and Meyer (2004) employed the purpose-built Bayesian software package called BUGS (Bayesian Analysis Using the Gibbs Sampler). Each of these MCMC algorithms is based on a single-move algorithm.

The main drawback with the single-move algorithm for MSV models lies in its slow convergence. This is not surprising since the components of the latent volatility process are highly persistent (Kim et al. (1998)). In order to improve the simulation efficiency, Chib et al. (2005) developed a new MCMC algorithm which greatly improves simulation efficiency for a factor MSV model augmented with jumps. Liesenfeld and Richard (2004) proposed an alternative multi-move MCMC method based on EIS, which can be used to estimate SV models by maximum likelihood, as well as simulation smoothing.

Bos and Shephard (2004) modelled the Gaussian errors in the standard Gaussian, linear state space model as an SV process, and showed that conventional MCMC algorithms for this class of models are ineffective. Rather than sampling the unobserved

variance series directly, Bos and Shephard (2004) sampled in the space of the disturbances, which decreased the correlation in the sampler and increased the quality of the Markov chain. Using the reparameterized MCMC sampler, they showed how to estimate an unobserved factor model.

Smith and Pitts (2005) used a bivariate SV model to measure the effects of intervention in stabilization policy. Missing observations were accommodated in the model and a data-based Wishart prior for the precision matrix of the errors in the transition equation were suggested. A threshold model for the transition equation was estimated by MCMC jointly with the bivariate SV model.

4. Diagnostic Checking and Model Comparison

Pitt and Shephard (1999a) conducted diagnostic checking which is applicable to the MSV models. Although their method is based on the particle filter algorithm (Pitt and Shephard (1999b)), other simulation filtering techniques, such as the EIS filter of Liesenfeld and Richard (2003) and the reprojection technique of Gallant and Tauchen (1998), may also be applicable. By using these filtering methods, we can obtain samples from the prediction density, $f(h_{t+1} \mid Y_t; \theta)$, where $Y_t = (y_1, \ldots, y_t)'$. Pitt and Shephard (1999a) focus on four quantities for assessing overall model fit, outliers and observations which have substantial influence on the fitted model.

The first quantity is the log-likelihood for $t+1$, $l_{t+1} = \log f(y_{t+1} \mid Y_t; \theta)$. As we have the following:

$$f(y_{t+1} \mid Y_t; \theta) = \int f(y_{t+1} \mid h_{t+1}; \theta) dF(h_{t+1} \mid Y_t; \theta),$$

Monte Carlo integration may be used as

$$\hat{f}(y_{t+1} \mid Y_t; \theta) = \frac{1}{M} \sum_{i=1}^{M} f(y_{t+1} \mid h_{t+1}^i; \theta),$$

where $h_{t+1}^i \sim f(h_{t+1} \mid Y_t; \theta)$. It is possible to evaluate the log-likelihood at the ML (MCL, SML) estimates or at the posterior means.

The second quantity is the normalized log-likelihood, l_t^n. Pitt and Shephard (1999a) used samples from z^j ($j = 1, \ldots, S$), where $z^j \sim f(y_{t+1} \mid Y_t; \theta)$, to obtain samples l_{t+1}^i using the above method. Denote the sample mean and standard deviation of the samples of log-likelihood as μ_{t+1}^l and σ_{t+1}^l, respectively. The normalized log-likelihood at $t+1$ may be computed as $l_{t+1}^n = \left(l_{t+1} - \mu_{t+1}^l\right) \Big/ \sigma_{t+1}^l$. If the model and parameters are correct, then this statistic should have mean zero and variance one. Large negative values indicate that an observation is less likely than would be expected from the model.

The third quantity is the uniform residual, $u_{t+1} = F(l_{t+1} \mid Y_t; \theta)$, which may be estimated as

$$\hat{u}_{t+1} = \hat{F}(l_{t+1}) = \left(1/S\right) \sum_{j=1}^{S} I(l_{t+1}^j < l_{t+1}),$$

where the l_{t+1}^j are constructed as above. Assuming that the parameter vector θ is known, under the null hypothesis that the model is correct, it follows that $\hat{u}_{t+1} \sim UID(0,1)$.

Finally, the fourth quantity is the distance measure, d_t, which may be computed as follows:

$$\Sigma_{t+1} = V(y_{t+1} \mid Y_t; \theta) \approx \left(1/M\right) \sum_{i=1}^{M} V(y_{t+1} \mid h_{t+1}^i; \theta),$$

where $h_{t+1}^{i} \sim f(h_{t+1} \mid Y_t; \theta)$. If the conditional distribution of y_t is multivariate normal, then the quantity $d_t = y_t' \Sigma_t^{-1} y_t$ is independently distributed as χ_m^2 under the null hypothesis that the parameters and model are correct. Therefore, we may use $\sum_{t=1}^{T} d_t \sim \chi_{mT}^2$ as a test statistic.

When the MCMC procedure is used, it may require checking convergence of Markov chains and prior sensitivities. The former can be assessed by correlograms, and the latter by using alternative priors (for further details, see Kim et al. (1998), Chib (2001) and Chib et al. (2005)).

Turning to model selection, we may use the likelihood ratio test for the nested models, and Akaike information criterion (AIC) or Bayesian information criterion (BIC) for the non-nested models, in the context of the likelihood-based methods, such as SML and MCL. In the Bayesian framework, model comparison can be conducted via the posterior odds ratio or Bayes factor. For both values, the marginal likelihood needs to be calculated, for which estimation is based on the procedure proposed by Chib (1995) and its various extensions.

The AIC is inappropriate for the MCMC method because, when MCMC is used to estimate the SV models, as mentioned above, the parameter space is augmented. For example, in the basic univariate SV model, we include the T latent volatilities in the parameter space, with T being the sample size. As these volatilities are dependent, they cannot be counted as T additional free parameters. Consequently, AIC is not applicable for comparing SV models. Recently, Berg et al. (2004) showed that model selection of alternative univariate SV models can be performed easily using the deviance information criterion (DIC) proposed by Spiegelhalter et al. (2002), while Yu and Meyer (2004) compared alternative MSV models using DIC.

5. Concluding Remarks

As the literature on multivariate stochastic volatility (MSV) models has developed significantly over the last few years, this paper reviewed the substantial literature on specification, estimation and evaluation of MSV models. A wide range of MSV models was presented according to various categories, namely (i) asymmetric models; (ii) factor models; (iii) time-varying correlation models; and (iv) alternative MSV specifications, including models based on the matrix exponential transformation, Cholesky ecomposition, Wishart autoregressive process, and the empirical range. Alternative methods of estimation, including quasi-maximum likelihood, simulated maximum likelihood, Monte Carlo likelihood, and Markov chain Monte Carlo methods, were discussed and compared. Various methods of diagnostic checking and model comparison were also examined.

Relative to the extensive theoretical and empirical multivariate GARCH literature, the MSV literature is still in its infancy. The majority of existing research in the MSV literature deals with specifications and/or estimation techniques, which are often illustrated by fitting a particular symmetric or asymmetric MSV model to financial returns series. Few papers have directly addressed important economic issues using MSV models. To our knowledge, Nardari and Scruggs (2003) and Han (2002) are two exceptions. Nardari and Scrugg used MSV models to address the restrictions in the APT theory while Han examined the economic values of MSV models. Clearly, further applications of MSV models are needed.

Most of the MSV models discussed in this paper have been estimated using at most 3 or 4 return series. Chib et al. (2005) is the first paper in the literature where genuinely high-dimensional MSV models have been estimated. Chan et al. (2003) and Nadari and Scruggs (2003) also estimated high-dimensional MSV models. With the development of superior estimation techniques and the availability of greater computing power, the literature on specification, estimation and evaluation of high-dimensional MSV models will be broadened appreciably.

References

Aguilar, O., and M. West (2000), "Bayesian Dynamic Factor Models and Portfolio Allocation", *Journal of Business and Economic Statistics*, **18**, 338-357.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003), "Modelling and Forecasting Realized Volatility", *Econometrica*, **71**, 579-625.

Asai, M. (2004), "A Comparative Analysis of Multi-Factor Stochastic Volatility Models with Heavy-Tailed Models and Higher-Order Autoregressive Models", unpublished paper, Faculty of Economics, Tokyo Metropolitan University.

Asai, M., and M. McAleer (2004a), "Dynamic Asymmetric Leverage in Stochastic Volatility Models", unpublished paper, Faculty of Economics, Tokyo Metropolitan University.

Asai, M., and M. McAleer (2004b), "Asymmetric Multivariate Stochastic Volatility", unpublished paper, Faculty of Economics, Tokyo Metropolitan University.

Asai, M., and M. McAleer (2004c), "Dynamic Correlations in Stochastic Volatility Models", unpublished paper, Faculty of Economics, Tokyo Metropolitan University.

Bauwens, L., S. Laurent and J.V.K. Rombouts (2004), "Multivariate GARCH: A Survey", to appear in *Journal of Applied Econometrics*.

Bellman, R. (1970), *Introduction to Matrix Analysis*, McGraw-Hill, New York.

Berg, A., R. Meyer and J. Yu (2004), "Deviance Information Criterion for Comparing Stochastic Volatility Models", *Journal of Business and Economic Statistics*, **22**, 107-120.

Black, F. (1976), "Studies of Stock Market Volatility Changes", *1976 Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp.177-181.

Bollerslev, T. (1990), "Modelling the Coherence in Short-run Nominal Exchange Rates:

A Multivariate Generalized ARCH Approach", *Review of Economics and Statistics*, 72, 498-505.

Bollerslev, T., R.F. Engle and J. Wooldridge (1988), "A Capital Asset Pricing Model with Time Varying Covariance", *Journal of Political Economy*, **96**, 116–131.

Bos, C.S and N. Shephard (2004), "Inference for Adaptive Time Series Models: Stochastic Volatility and Conditionally Gaussian State Space Form", unpublished paper, Tinbergen Institute and Vrije Universiteit Amsterdam.

Broto, C., and E. Ruiz (2004), "Estimation Methods for Stochastic Volatility Models: A Survey", *Journal of Economic Surveys*, **18**, 613-649.

Campbell, J.Y., and L. Hentschel (1992), "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns", *Journal of Financial Economics*, **31**, 281-318.

Chan, D., R. Kohn and C. Kirby (2003), "Multivariate Stochastic Volatility with Leverage", unpublished paper, School of Economics, University of New South Wales.

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**, 267-284.

Chib, S. (1995), "Marginal Likelihood from the Gibbs Output", *Journal of the American Statistical Association*, **90**, 1313-1321.

Chib, S. (2001), "Markov Chain Monte Carlo Methods: Computation and Inference", in J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, Volume 5, North-Holland , pp. 3569–3649.

Chib, S., F. Nardari and N. Shephard (2005), "Analysis of High Dimensional Multivariate Stochastic Volatility Models", to appear in *Journal of Econometrics*.

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.

Christie, A.A. (1982), "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects", *Journal of Financial Economics*, **10**, 407-432.

Christodoulakis, G., and S.E. Satchell (2002), "Correlated ARCH (CorrARCH): Modelling the time-varying conditional correlation between fianacial asset returns", *European Journal of Operational Research*, **139**, 351-370.

Danielsson, J. (1994), "Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood", *Journal of Econometrics*, **64**, 375-400.

Danielsson, J. (1998), "Multivariate Stochastic Volatility Models: Estimation and a Comparison with VGARCH Models", *Journal of Empirical Finance*, **5**, 155–173.

Danielsson, J., and J.-F. Richard (1993), "Quadratic Acceleration for Simulated Maximum Likelihood Evaluation", *Journal of Applied Econometrics*, **8**, 153–173.

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 –22.

Ding, Z., and R. Engle (2001), "Large Scale Conditional Covariance Modelling, 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.

Dunsmuir, W. (1979), "A Central Limit Theorem for Parameter Estimation in Stationary Vector Time Series and Its Application to Models for a Signal Observed with Noise", *Annals of Statistics*, 7, 490-506.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica,* **50**, 987-1007.

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

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

Engle, R., and V. Ng (1993), "Measuring and Testing the Impact of News in Volatility", *Journal of Finance*, **43**, 1749-1778.

Engle, R.F., V. Ng, and M. Rothschild (1990), "Asset Pricing with a Factor-ARCH Covariance Structure," *Journal of Econometrics*, **45**, 213–237.

Gallant, A.R., and G.E. Tauchen (1998), "Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions", *Journal of the American Statistical Association*, **93**, 10- 24.

Geweke, J., and S. Porter-Hudak (1983), "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis*, **4**, 221-238.

Ghysels, E., A.C. Harvey and E. Renault (1996), "Stochastic Volatility", in C.R. Rao and G.S. Maddala (eds.), *Statistical Models in Finance*, North-Holland, Amsterdam, pp. 119-191.

Glosten, L., R. Jagannathan and D. Runkle (1993), "On the Relation Between the Expected Value and Volatility of Nominal Excess Returns on Stocks", *Journal of Finance*, **46**, 1779-1801.

Gourieroux, C., J. Jasiak and R. Sufana (2004), "The Wishart Autoregressive Process of Multivariate Stochastic Volatility", unpublished paper, CRES and CEPREMAP.

Han, Y., (2002), "The Economics Value of Volatility Modelling: Asset Allocation with a High Dimensional Dynamic Latent Factor Multivariate Stochastic Volatility Model", unpublished paper, Olin School of Business, Washington University in St. Louis.

Harvey, A.C., E. Ruiz and N. Shephard (1994), "Multivariate Stochastic Variance Models", *Review of Economic Studies*, **61**, 247-264.

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.

Hull, J., and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatility", *Journal of Finance*, **42**, 281-300.

Jacquier, E., N.G. Polson and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models (with discussion)", *Journal of Business and Economic Statistics*, **12**, 371– 389.

Jacquier, E., N.G. Polson and P.E. Rossi (1995), "Models and Priors for Multivariate Stochastic Volatility", CIRANO Working Paper No. 95s-18, Montreal.

Jacquier, E., N.G. Polson and P.E. Rossi (1999), "Stochastic Volatility: Univariate and Multivariate Extensions", CIRANO Working paper 99s-26, Montreal.

Jacquier, E., N.G. Polson and P.E. Rossi (2004), "Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors", *Journal of Econometrics*, **122**, 185-212.

Kim, S., N. Shephard and S. Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models", *Review of Economic Studies*, **65**, 361-393.

Liesenfeld, R., and J.-F. Richard (2003), "Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics", *Journal of Empirical Finance*, **10**, 505-531.

Liesenfeld, R., and J.-F. Richard (2004), "Classical and Bayesian Analysis of Univariate and Multivariate Stochastic Volatility Models", unpublished paper, Department of Economics, Christian-Albrecht-Universitat, Germany.

McAleer, M. (2005), "Automated Inference and Learning in Modeling Financial Volatility", *Econometric Theory*, 21, 232-261.

Nardari, F., and J. Scruggs (2003), "Analysis of Linear Factor Models with Multivariate Stochastic Volatility for Stock and Bond Returns", unpublished paper, Department of

Finance, Arizona State University.

Nelson, D.B. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach", *Econometrica*, 59, 347-370.

Philipov, A., and M.E. Glickman (2004a), "Factor Multivariate Stochastic Volatility Via Wishart Processes", unpublished paper, Department of Health Services, Boston University.

Philipov, A., and M.E. Glickman (2004b), "Multivariate Stochastic Volatility Via Wishart Processes", unpublished paper, Department of Health Services, Boston University.

Pitt, M., and N. Shephard (1999a), "Time Varying Covariances: A Factor Stochastic Volatility Approach", in J.M. Bernardo, J.O. Berger, A.P. David and A.F.M. Smith (eds.), *Bayesian Statistics 6*, Oxford University Press, pp. 547-570.

Pitt, M., and N. Shephard (1999b), "Filtering via Simulation: Auxiliary Particle Filter", *Journal of the American Statistical Association*, **94**, 590–599.

Quintana, J.M., and M. West (1987), "An Analysis of International Exchange Rates Using Multivariate DLMs", *The Statistician*, **36**, 275-281.

Ray, B.K., and R.S. Tsay (2000), "Long-range Dependence in Daily Stock Volatilities", *Journal of Business and Economic Statistics*, **18**, 254–262.

Richard, J.-F., and W. Zhang (2004), "Efficient High-Dimensional Importance Sampling", unpublished paper, Department of Economics, University of Pittsburg.

Robinson, P.M. (1995), "Log-Periodogram Regression with Time Series with Long-Range Dependence", *Annals of Statistics*, **23**, 1048-1072.

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. (1996), "Statistical Aspects of ARCH and Stochastic Volatility", in D.R.

Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (eds.), *Time Series Models in Econometrics, Finance and Other Fields*. Chapman & Hall, London, pp. 1-67.

Shephard, N., and M.K. Pitt (1997), "Likelihood Analysis of Non-Gaussian Measurement Time Series", *Biometrika*, **84**, 653-667.

Smith, M. and A. Pitts (2005), "Foreign Exchange Intervention by the Bank of Japan: Bayesian Analysis Using a Bivariate Stochastic Volatility Model", unpublished paper, Econometrics and Business Statistics, University of Sydney.

So, M.K.P., W.K. Li and K. Lam (2002), "A Threshold Stochastic Volatility Model", *Journal of Forecasting*, **21**, 473-500.

Spiegelhalter, D.J., N.G. Best, B.P. Carlin and A. van der Linde (2002), "Bayesian Measures of Model Complexity and Fit (with discussion)", *Journal of the Royal Statistical Society, Series B***, 64**, 583-639.

Taylor, S.J. (1982), "Financial Returns Modelled by the Product of Two Stochastic Processes — A Study of the Daily Sugar Prices 1961-75", in O.D. Anderson (ed.) *Time Series Analysis: Theory and Practice*, Vol. 1, North-Holland, pp.203-226.

Taylor, S.J. (1986), *Modelling Financial Time Series*, Wiley, Chichester.

Tierney, L. (1994), "Markov Chains for Exploring Posterior Distributions", *Annals of Statistics*, **22**, 1701-1762.

Tims, B. and R. Mahieu (2003), "A Range-Based Multivariate Model for Exchange Rate Volatility", unpublished paper, Department of Financial Management, Erasmus University Rotterdam.

Tsay, R.S. (2002), *Analysis of Financial Time Series: Financial Econometrics*, Wiley, New York.

Tse, Y.K., and A.K.C. Tsui (2002), "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations", *Journal of Business and Economic Statistics*, **20**, 351-362.

Wiggins, J.B. (1987), "Option Values Under Stochastic Volatility: Theory and Empirical Estimates", *Journal of Financial Economics*, **19**, 351-372.

Yu, J. (2004a), "On Leverage in a Stochastic Volatility Model", to appear in *Journal of Econometrics*.

Yu, J. (2004b), "Asymmetric Response of Volatility: Evidence from Stochastic Volatility Models and Realized Volatility", School of Economics and Social Sciences, Singapore Management University, Working Paper 24-2004.

Yu, J., and R. Meyer (2004), "Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison", School of Economics and Social Sciences, Singapore Management University, Working Paper 23-2004.