KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE

OF

ECONOMIC RESEARCH

KYOTO UNIVERSITY

KYOTO, JAPAN

A Fractionally Integrated Wishart Stochastic Volatility Model[*]

Manabu Asai

Faculty of Economics
Soka University, Japan
and
Wharton School
University of Pennsylvania

Michael McAleer

Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Institute of Economic Research
Kyoto University, Japan
and
Department of Quantitative Economics
Complutense University of Madrid, Spain

January 2013

Abstract

There has recently been growing interest in modeling and estimating alternative continuous time multivariate stochastic volatility models. We propose a continuous time fractionally integrated Wishart stochastic volatility (FIWSV) process. We derive the conditional Laplace transform of the FIWSV model in order to obtain a closed form expression of moments. We conduct a two-step procedure, namely estimating the parameter of fractional integration via log-periodogram regression in the first step, and estimating the remaining parameters via the generalized method of moments in the second step. Monte Carlo results for the procedure shows reasonable performances in finite samples. The empirical results for the bivariate data of the S&P 500 and FTSE 100 indexes show that the data favor the new FIWSV processes rather than one-factor and two-factor models of Wishart autoregressive processes for the covariance structure.

1 Introduction

In the framework of discrete time models, the empirical literature has found evidence of slowly decaying autocorrelations in the volatility of financial time series. This literature is based on models suggested by Ding, Granger and Engle (1993), Baillie, Bollerslev and Mikkelsen (1996), and Bollerslev and Mikkelsen (1996) for the ARCH family, Breidt, Crato and de Lima (1998) and Harvey (1993) for stochastic volatility models, and Andersen et al. (2001), Asai, McAleer and Medeiros (2012), Koopman, Jungbacker and Hol (2005), and Pong et al. (2004) for models of realized volatility. With respect to the continuous time framework, Comte and Renault (1998) suggested a long memory stochastic volatility model. Recently, Bollerslev, Sizova and Tauchen (2012) developed a volatility equilibrium model.

For multivariate volatility models, a wide range of multivariate GARCH and stochastic volatility (SV) models has been developed, analyzed, and applied extensively in recent years to characterize the volatility that is inherent in financial time series data. While Bauwens, Laurent and Rombouts (2006) provided a survey of multivariate GARCH models, Asai, McAleer and Yu (2006) and Chib, Omori and Asai (2009) deal with multivariate SV models. Recently, there has been growing interest in continuous time multivariate SV processes. With respect to multivariate derivative pricing models, Gourieroux (2006) and Gourieroux and Sufana (2010) extended Wishart Autoregressive (WAR) diffusion processes, as originally considered by Bru (1991). Recently, Muhle-Karbe, Pfaffel and Stelzer (2012) developed a multivariate Ornstein-Uhlenbeck-type SV model based on a Lévy process.

The purpose of the paper is to develop a new continuous time fractionally integrated multivariate SV model, along the lines of Bru (1991) and Gourieroux (2006), combined with a continuous time fractionally integrated process, as suggested by Comte and Renault (1996, 1998).

3

The remainder of the paper is organized as follows. Section 2 proposes the continuous time fractionally integrated Wishart processes, and derives the conditional Laplace transform. Section 3 considers three kinds of Wishart stochastic volatility (WSV) models, namely one-factor and two-factor WSV models and the fractionally integrated WSV (FIWSV) model. Section 4 suggests a two-step procedure for estimating the new FIWSV models, and reports the results of Monte Carlo experiments. Section 6 provides empirical results for the bivariate returns vector of the S&P 500 and FTSE 100 indexes. Section 7 gives some conclusions.

In the following, for any symmetric matrix A, $A^{1/2}$ is defined by the spectral decomposition of A, so that $A^{1/2}A^{1/2} = A$. For any square matrix A, the matrix-exponential operator is defined by $\text{Exp}(A) = \sum_{i=0}^{\infty}(1/i!)A^i$, with $A^0 = I$.

2 Fractionally Integrated Wishart Autoregressive Model

We develop new continuous time multivariate stochastic volatility (MSV) models, based on the Wishart Autoregressive (WAR) process, as suggested by Bru (1991) and analyzed by Gourieroux (2006). The WAR(ν,Φ,Θ) process is defined by

$$dW(t) = \left(\nu\Theta\Theta' + W(t)\Phi' + \Phi W(t)\right)dt + W(t)^{1/2}dB(t)\Theta' + \Theta dB(t)W(t)^{1/2}, \tag{1}$$

where the $B(t)$ are m dimensional and symmetric matrix-variate standard Brownian motions, ν is a scalar such that $\nu > m$, Φ is an $m \times m$ matrix, and Θ is an m dimensional lower triangular matrix. The Wishart process can be driven by the sum of the cross-products of the vector Ornstein-Uhlenbeck (OU) process, as

$$\begin{aligned} W(t) &= \sum_{i=1}^{\nu} Y_i(t)Y_i(t)', \\ dY_i(t) &= \Phi Y_i(t)dt + \Theta d\tilde{B}_i(t), \end{aligned} \tag{2}$$

4

where $\tilde{B}_i(t)$ are m dimensional vector of mutually independent standard Brownian motions. However, the process also exists for non-integer degrees of freedom ν, like the χ^2 distribution and the Wishart distribution.

We employ the fractional Brownian motion, following the work of Comte and Renault (1996, 1998). Comte and Renault (1996) suggested continuous time fractional ARMA processes, while Comte and Renault (1998) developed continuous time SV models. Based on an $m \times m$ standard Brownian motion, define the fractional Brownian motion of order α as

$$B_\alpha(t) = \int_0^t \frac{(t-s)}{\Gamma(\alpha+1)} dB(s), \quad 0 \le \alpha < \frac{1}{2}, \tag{3}$$

with $B_\alpha(0) = O$, where $\Gamma(x)$ is the gamma function. This 'one-sided' definition corresponds to the seminal paper by Mandelbrot and van Ness (1968). For $\alpha = 0$, $B_\alpha(t)$ reduces to the standard Brownian motion $B(t)$. We exclude the case that $-1/2 < \alpha < 0$, as long memory occurs on the range that $0 < \alpha < 1/2$. As noted in Comte and Renault (1996), $H = \alpha + \frac{1}{2}$ is the Hurst parameter. We may define a two-sided range for an increment of a fractional Brownian motion process, but we follow Comte and Renault (1996).

The fractional Brownian motion process has several features. First, it has zero mean and a covariance matrix given by

$$\text{cov}\left(\text{vech}\{B_\alpha(t)\}, \text{vech}\{B_\alpha(s)\}\right) = \frac{1}{2}\left(|t|^{2\alpha+1} + |s|^{2\alpha+1} - |t-s|^{2\alpha+1}\right) I_{m(m+1)/2}.$$

Second, the fractional Brownian motion process has a property of H self-similarity, which means the probability distributions of $B_\alpha(\beta t)$ and $|\beta|^H B_\alpha(t)$ are identical. Third, the conditional variance of increment $B_\alpha(t+h) - B_\alpha(t)$ is given by

$$V_t\left(\text{vech}\{B_\alpha(t+h) - B_\alpha(t)\}\right) = \frac{h^{2\alpha+1}}{(2\alpha+1)\{\Gamma(\alpha+1)\}^2} I_{m(m+1)/2}.$$

5

See Comte and Renault (1996) for further details. Hence, the conditional variance of $dB_\alpha(t)$ is proportional to $(dt)^{2\alpha+1}$.

We can consider two kinds of extensions of the WAR process in order to develop a fractionally integrated WAR (FIWAR) process. One extension is to replace $B(t)$ with $B_\alpha(t)$ in equation (1), while another is to use a vector fractional Brownian motion, $\tilde{B}_{\alpha,i}(t)$, instead of $\tilde{B}_i(t)$ in equation (2). The former approach produces the FIWAR1 model, given by

$$dW(t) = \left(\nu\Theta\Theta' + W(t)\Phi' + \Phi W(t)\right) dt + W(t)^{1/2}dB_\alpha(t)\Theta' + \Theta dB_\alpha(t)W(t)^{1/2}, \qquad (4)$$

with $W(0) = W_0$. The latter approach yields the FIWAR2 model, defined by

$$W(t) = \sum_{i=1}^{\nu} Y_i(t)Y_i(t)',$$
$$dY_i(t) = \Phi Y_i(t)dt + \Theta d\tilde{B}_{\alpha,i}(t), \qquad (5)$$

giving a slightly different process:

$$dW(t) = \frac{\nu}{(2\alpha+1)\{\Gamma(\alpha+1)\}^2}\Theta\Theta'(dt)^{2\alpha+1} + \left(W(t)\Phi' + \Phi W(t)\right) dt$$
$$+ W(t)^{1/2}dB_\alpha(t)\Theta' + \Theta dB_\alpha(t)W(t)^{1/2}, \quad W(0) = W_0. \qquad (6)$$

The FIWAR1 process (4) and FIWAR2 process (6) reduce to the WAR model when $\alpha = 0$. The difference between the two FIWAR models is characterized by $(dt)^{2\alpha+1}$. If $\alpha > 0$, the first term on the right-hand side of $dW(t)$ in equation (6) vanishes as a consequence of application of the Ito formula. See Dai and Heyde (1996) for further details of the Ito formula based on fractional Brownian motion process.

In the following, we derive the conditional and unconditional Laplace transforms (moment generating function) of the FIWAR processes. As explained in Duffie, Pan and Singleton (2000), the risk neutral Laplace transform is the basis for derivative pricing because it can be used to obtain explicit or quasi-explicit prices for various derivatives. Applying the approach of Gourieroux

and Sufana (2010), we can show that $W(t)$ is an affine process. In order to derive the conditional Laplace transform, we use the matrix Riccati linearization technique suggested by Fonseca, Grasselli and Tebaldi (2008), instead of the approach of Gourieroux and Sufana (2010).

Before we derive conditional Laplace transforms for the two FIWAR processes, we introduce the conditional Laplace transform of the basic WAR process, which is examined by Bru (1991) and Gourieroux (2006). The conditional Laplace transform of $W(t + h)$, given $W(t)$, is

$$\Psi_{\Xi,t}(h) \equiv E\left[\exp\left(\mathrm{tr}\{\Xi W(t + h)\}\right)|W(t)\right]$$

$$= \frac{\exp(\mathrm{tr}\{(I_m - 2\Xi\Omega(h))^{-1})\Xi M(h)W(t)M(h)'\})}{[\det(I_m - 2\Xi\Omega(h))]^{\nu/2}}, \tag{7}$$

where Ξ is a symmetric matrix,

$$M(h) = \mathrm{Exp}(Ah), \quad \Omega(h) = \int_0^h \mathrm{Exp}(As)\Theta\Theta'[\mathrm{Exp}(As)]'ds. \tag{8}$$

Matrices $M(h)$ and $\Omega(h)$ can be interpreted using the analogy of the OU process in continuous time and the Gaussian VAR(1) in discrete time. For the OU process in equation (1), the distribution of $Y_i(t + h)$ conditional on $Y_i(t)$ is given by $N(M(h)Y_i(t), \Omega(h))$, which constitutes the Gaussian VAR(1) in discrete time. It is possible to obtain the solution of $\Omega(h)$. The conditional distribution of $W(t + h)$ is a noncentral Wishart distribution.

Proposition 1 *(i) For the WAR process defined by (1), the conditional Laplace transform is given by equations (7) and (8). If Φ^{-1} exists, we have*

$$\mathrm{vec}(\Omega(h)) = [(I_m \otimes \Phi) + (\Phi \otimes I_m)]^{-1} \mathrm{vec}(\mathrm{Exp}(\Phi h)\Theta\Theta'\mathrm{Exp}(\Phi'h) - \Theta\Theta').$$

(ii) The WAR process (1) is stationary if and only if all the eigenvalues of Φ are negative. Under this condition, the WAR process has the invariant distribution, and follows the Wishart distribution, $W(\Omega, \nu)$, where

$$\mathrm{vec}(\Omega) = -[(I_m \otimes \Phi) + (\Phi \otimes I_m)]^{-1} \mathrm{vec}(\Theta\Theta').$$

7

Proof. See Appendix A.1.

Remark 1. By the property of the Wishart distribution, $W(t)$ is positive definite if and only if (a) $\nu > m$ and (b) Ω is positive definite. The second condition is guaranteed by the two conditions; (b1) all the eigenvalues of Φ are negative, and (b2) the diagonal elements of C are non-zero.

Before we turn to the two FIWAR processes, we give the conditional Laplace transform of the vector OU-type process with fractional Brownian motion.

Lemma 1 *For the vector OU-type process with fractional Brownian motion defined by equation (5), the conditional Laplace transform of $Y_i(t+h)$, given $Y_i(t)$, is*

$$\psi_{\Xi,t}^\alpha(h) \equiv E\left[\exp\left(\gamma' Y(t+h)\right)|Y(t)\right] = \exp\left(b^\dagger(h)' Y(t) + c^\dagger(h)\right)$$

where

$$b^\dagger(h) = \text{Exp}(\Phi' h)\gamma, \qquad c^\dagger(h) = \frac{1}{2}\text{tr}\left\{\Omega_\alpha(h)\gamma\gamma'\right\}, \tag{9}$$

with

$$\Omega_\alpha(h) = c_\alpha^{-1}\int_0^h s^{2\alpha}\text{Exp}(\Phi s)\Theta\Theta'\text{Exp}(\Phi' s)ds,$$

where $c_\alpha = (2\alpha+1)^2\{\Gamma(\alpha+1)\}^2$. If Φ^{-1} exists and $0 < \alpha < 1/2$, we have

$$\text{vec}(\Omega_\alpha(h)) = c_\alpha^{-1}h^{2\alpha}\left[I_{m^2} - (I_m \otimes \Phi) - (\Phi \otimes I_m)\right]^{-1}\text{vec}(\Omega(h) - \text{Exp}(\Phi h)\Theta\Theta'\text{Exp}(\Phi' h)), \tag{10}$$

where $\Omega(h)$ is defined by Proposition 1.

Proof. See Appendix A.2.

Remark 2. The distribution of $Y_i(t+h)$ conditional on $Y_i(t)$ is given by $N(M(h)Y_i(t), \Omega_\alpha(h))$.

The following proposition shows the conditional Laplace transform regarding the FIWAR2 process.

Proposition 2 *For the FIWAR2 process (5), the conditional Laplace transform of $W(t + h)$, given $W(t)$, is*

$$\Psi^\alpha_{2,\Xi,t}(h) \equiv E\left[\exp\left(\mathrm{tr}\{\Xi W(t+h)\}\right) | W(t)\right]$$

$$= \frac{\exp(\mathrm{tr}\{(I_m - 2\Xi\Omega_\alpha(h))^{-1}\Xi M(h)W(t)M(h)'\})}{[\det(I_m - 2\Xi\Omega_\alpha(h))]^{\nu/2}}, \quad (11)$$

where Ξ is a symmetric matrix, $M(h) = \mathrm{Exp}(Ah)$, and $\Omega_\alpha(h)$ is defined by Lemma 1.

Proof. With respect to the case that ν is an integer, it is straightforward from Lemma 1. As in the case of the WAR process, the conditional distribution of $W(t + h)$ is a noncentral Wishart distribution. It is known that this distribution exists also for noninteger degrees of freedom ν. ■

Now we move to the conditional Laplace transform for the FIWAR1 process.

Proposition 3 *For the FIWAR1 process (4), the conditional Laplace transform of $W(t + h)$, given $W(t)$, is*

$$\Psi^\alpha_{1,\Xi,t}(h) \equiv E\left[\exp\left(\mathrm{tr}\{\Xi W(t+h)\}\right) | W(t)\right]$$

$$= \exp(c(h) + \mathrm{tr}\{(I_m - 2\Xi\Omega_\alpha(h))^{-1}\Xi M(h)W(t)M(h)'\}), \quad (12)$$

where Ξ is a symmetric matrix,

$$c(h) = \int_0^h \nu\,\mathrm{tr}\left\{[I_m - 2\Xi\Omega_\alpha(s)]^{-1}\Xi\,\mathrm{Exp}(\Phi s)\Theta\Theta'\mathrm{Exp}(\Phi's)\right\} ds, \quad (13)$$

and $M(h)$ and $\Omega_\alpha(h)$ are as before.

Proof. See Appendix A.3.

Remark 3. We can apply the approach shown in Appendix A.3 to obtain the conditional Laplace transform of equation (6). The result is the same as Proposition 2.

Owing to the integral of $c(h)$, it is not straightforward to use the Laplace transform of FIWAR1. For this reason, we concentrate on the FIWAR2 model defined by equation (5) for the remainder of the paper.

3   Fractionally Integrated Wishart Stochastic Volatility

Let p_t be an m dimensional vector of financial log-prices evolving in continuous time. The continuous time multivariate stochastic volatility (MSV) model is defined by

$$dp(t) = \mu(p, t)dt + \Sigma(t)^{1/2}d\tilde{B}_0(t), \tag{14}$$

where $\tilde{B}_0(t)$ are m dimensional vector of mutually independent standard Brownian motions, and $\Sigma(t)^{1/2}$ ($m \times m$) is a function of diffusion processes. Here, $\Sigma(t)$ is the instantaneous covariance matrix. For modeling $\Sigma(t)$, we consider three kinds of specifications based on the WAR and FIWAR processes, as follows:

(i) Wishart Stochastic Volatility (WSV) Model

$$\Sigma(t) = CV(t)C', \text{ where } V(t) \sim \text{WAR}(\nu, \Phi, I_m);$$

(ii) Two Factor WSV (2WSV) Model

$$\Sigma(t) = C[V_1(t) + V_2(t)]C', \text{ where } V_k(t) \sim \text{WAR}(\nu_k, \Phi_k, I_m) \quad (k = 1, 2);$$

(iii) Fractionally Integrated WSV (FIWSV) Model

$$\Sigma(t) = CV(t)C', \text{ where } V(t) \sim \text{FIWAR2}(\nu, \Phi, I_m, \delta),$$

where C is an m dimensional lower triangular matrix. For purposes of identification, we restrict the $(1, 1)$ element of C to be positive. As our analysis focuses on modeling covariance processes rather than correlation processes (see McAleer et al. (2008) and Caporin and McAleer (2012)), we assume $\mu(p, t) = \mathbf{0}$.

The first model is based on the WAR process, and $\Sigma(t)$ is equivalent to $W(t)$ in equation (1). The second model is motivated by recent empirical analyses on univariate two-factor SV models,

including the papers of Bollerslev and Zhou (2002) and Chernov et al. (2003). By introducing a second factor, the model can describe a longer persistence process than the simple one-factor model. The third model works with the FIWAR2 model examined in the previous section.

Now we turn to the conditional Laplace transform of the log-price process $p(t)$. Regarding the WSV and 2WSV processes, we obtain the conditional Laplace transform as special cases of Proposition 1 of Asai and McAleer (2012). For the general K factor WSV model, we have

$$\tilde{\Psi}_{\gamma,t}(h) \equiv E_t \left[\exp \left(\gamma' p(t+h) \right) \right] = \exp \left[\sum_{k=1}^{K} \operatorname{tr} \left(A_k(h) V_k(t) \right) + b(h)' p(t) + c(h) \right], \qquad (15)$$

where

$$A_k(h) = [\tilde{N}_{22,k}(h)]^{-1} \tilde{N}_{21,k}(h), \quad (k = 1, \ldots, K),$$

$$b(h) = \gamma,$$

$$c(h) = -\frac{1}{2} \sum_{k=1}^{K} \nu_k \left[\log \det(\tilde{N}_{22,k}(h)) + h \operatorname{tr}(\Phi_k') \right],$$

with

$$\left(\begin{array}{cc} \tilde{N}_{11,k}(h) & \tilde{N}_{12,k}(h) \\ \tilde{N}_{21,k}(h) & \tilde{N}_{22,k}(h) \end{array} \right) = \operatorname{Exp} h \left(\begin{array}{cc} \Phi_k & -2I_m \\ \frac{1}{2} C \gamma \gamma' C' & -\Phi_k' \end{array} \right),$$

if $[\tilde{N}_{22,k}(h)]^{-1}$ exists.

The next proposition gives the conditional Laplace transform of the log-price process $p(t)$ of the FIWSV process.

Proposition 4 *For the FIWSV model with equation (6), $\Theta = I_m$, and equation (14), the conditional Laplace transform of $p(t+h)$, given $(p(t), V(t))$, is*

$$\tilde{\Psi}_{\gamma,t}^{\alpha}(h) \equiv E_t \left[\exp \left(\gamma' p(t+h) \right) \right] = \exp \left[\operatorname{tr} \left(\tilde{A}(h) V(t) \right) + \tilde{b}(h)' p(t) + \tilde{c}(h) \right], \qquad (16)$$

11

where the symmetric matrices $\tilde{A}(h)$, vector $\tilde{b}(h)$ and scalar $\tilde{c}(h)$ satisfy the system of Riccati equations:

$$\frac{d}{dh}\tilde{A}(h) = \Phi'\tilde{A}(h) + \tilde{A}(h)\Phi + 2c_\alpha^{-1}h^{2\alpha}[\tilde{A}(h)]^2 + \tilde{\Gamma},$$

$$\frac{d}{dh}\tilde{b}(h) = 0,$$

$$\frac{d}{dh}\tilde{c}(h) = \nu c_\alpha^{-1}h^{2\alpha}\mathrm{tr}\left(\tilde{A}(h)\right),$$

with $\tilde{\Gamma} = \frac{1}{2}\Theta\gamma\gamma'\Theta'$, c_α defined by Lemma 1, and the initial conditions that $\tilde{A}(0) = O$, $\tilde{b}(0) = \gamma$ and $\tilde{c}(0) = 0$. The solutions are given by

$$\tilde{A}(h) = [\tilde{N}_{22}(h)]^{-1}\tilde{N}_{21}(h),$$

$$\tilde{b}(h) = \gamma,$$

$$\tilde{c}(h) = -\frac{1}{2}\nu\left[\log\det(\tilde{N}_{22}(h)) + h\mathrm{tr}(\Phi')\right],$$

where

$$\begin{pmatrix} \tilde{N}_{11}(h) & \tilde{N}_{12}(h) \\ \tilde{N}_{21}(h) & \tilde{N}_{22}(h) \end{pmatrix} = \mathrm{Exp}\,h\begin{pmatrix} \Phi_k & -2c_\alpha^{-1}h^{2\alpha}I_m \\ \tilde{\Gamma} & -\Phi'_k \end{pmatrix},$$

if $[\tilde{N}_{22}(h)]^{-1}$ exists.

Proof. See Appendix A.4.

Define the vector of asset returns as $y(t + h) = p(t + h) - p(t)$. Noting that the fractional Brownian motion process (3) starts from zero, we can obtain the moment generating function of $y(t + h)$, conditional on $V(0)$, as follows:

$$E\left[\exp(\gamma'y(t + h))|V(0)\right] = E_0 E_t\left[\exp(\gamma'(p(t + h) - p(t)))\right]$$

$$= E_0\exp(-\gamma'p(t))\tilde{\Psi}^\alpha_{\gamma,t}(h)$$

$$= E_0\exp\left(\tilde{c}(h) + \mathrm{tr}\left\{\tilde{A}(h)V(t)\right\}\right)$$

$$= \exp\left(\tilde{c}(h)\right)\Psi^\alpha_{2,\tilde{A}(h),0}(t), \tag{17}$$

12

where $\Psi_{2,\Xi,t}^{\alpha}(h)$ is defined by equation (11). This result constitutes the basics of the estimation procedure shown in the next section.

4   Estimation

For estimating the FIWSV models, we use the following two-step method: (i) Estimate the order of the fractional Brownian motion, α, via the Geweke and Porter-Hudak (1993) (GPH) log-periodogram regression estimator; (ii) Estimate the remaining parameters via a generalized method of moments (GMM), based on the moment generating function of $y(t + h)$.

Robinson (1995) gives the formal proof for the asymptotic normality of the GPH estimator and its multivariate extension. Deo and Hurvich (2001) present Monte Carlo results for estimating the degree of fractional integration of ARFIMA plus noise models, using the GPH estimator. Deo and Hurvich (2001) investigated an appropriate order of the length of periodgrams, and found that it is $o(T^{0.3})$. For the data of $\{\text{vech}(y(t)y(t)')\}_{t=1}^{T}$, we estimate the single parameter α by imposing linear restrictions on the multivariate log-periodogram regression.

In the second step, we conduct GMM using equation (17). Following Gourieroux and Sufana (2010), we work with the moments of $\exp(u_2'y(t + 2h) + u_1'y(t + h))$ for $y(t)$. For the case of $n = 2$, we consider 17 moment conditions of (u_2', u_1'): $(0, 0, 1, 0)$, $(0, 0, 0, 1)$, $(0, 0, 2, 0)$, $(0, 0, 0, 2)$, $(0, 0, 1, 1)$, $(1, 0, 1, 0)$, $(0, 1, 0, 1)$, $(1, 0, 0, 1)$, $(0, 1, 1, 0)$, $(1, 0, 2, 0)$, $(0, 1, 0, 2)$, $(1, 0, 0, 2)$, $(0, 1, 2, 0)$, $(2, 0, 2, 0)$, $(0, 2, 0, 2)$, $(2, 0, 0, 2)$, $(0, 2, 2, 0)$. For the parameter vector θ, we define the 17×1 vector as

$$f_t(\theta) = \left\{ \exp(u_2'y(t + 2h) + u_1'y(t + h))) - E[\exp(u_2'y(t + 2h) + u_1'y(t + h)))|V_0] \right\}.$$

By construction, $E_t[f_t(\theta)] = \mathbf{0}$, and we can define the corresponding GMM estimator by $\hat{\theta}_T = \text{argmin} g_T(\theta)'W g_T(\theta)$, where $g_T(\theta)$ is the sample mean of the moment conditions, $g_T(\theta) = T^{-1}\sum_{t=1}^{T-2} f_t(\theta)$,

13

and W denotes the inverse of the asymptotic covariance matrix of $g_T(\theta)$ (Hansen, 1982). Under standard regularity conditions, the minimized value of the objective function multiplied by the sample size follows the chi-square distribution asymptotically, which enables us to conduct an omnibus test of the overidentifying restrictions. Moreover, we have $\sqrt{T}(\hat{\theta} - \theta_0) \to^d N(\mathbf{0}, (G'WG)^{-1})$, where θ_0 is the true value of θ and $G = \frac{dg_T(\theta_0)}{d\theta'}$.

At this point, we need to examine estimators of W^{-1}. As we are dealing with long memory processes, the conventional heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimators, such as those of Newey and West (1987) and Andrews (1991), are inconsistent, leading to asymptotically invalid tests and inconsistent interval estimates. In order to cope with this problem, Robinson (2005) developed a consistent estimator in the presence of long memory. Also, we use this approach for estimating $\Sigma(0)$ in $V(0) = C^{-1}\Sigma(0)C'^{-1}$ for the initial value $V(0)$, as we can use the estimator of Robinson (2005) for the special case $d = 0$.

We present the results of a Monte Carlo study to investigate the finite sample performance of the GMM estimation for $n = 2$. We generate R simulated time series of $\{y_t, x_t\}_{t=1}^{T}$ for the FIWSV model (equation (5), $\Theta = I_m$, and equation (14)) and for some given "true" parameter vector. Subsequently, we treat the parameter vector as unknown and estimate it for each series using GMM. We compute the sample mean, standard deviation, and root mean squared error (RMSE), and compare it with the true parameter values.

The two sets of true parameter values for generating Monte Carlo samples are given in the first column of Tables 1(a)(b), which are based on the empirical analysis in Section 4. More precisely, we use $\lceil \hat{\nu} \rceil$ for the true value of ν, where $\lceil x \rceil$ is the integer from rounding up a scalar value x, in order to generate samples from equation (5). Furthermore, we consider $\hat{\alpha} + 0.1$ for the second DGP in order to examine the effects of the long memory parameter. The results given in Table 1

are for sample size $T = 2500$, with the number of iterations set to $R = 2000$.

Table 1(a) shows the downward bias for the GPH estimator of α, as $T = 2500$ is a relatively small sample size for estimating the long memory parameter. With respect to the second step, the GMM estimator has an upward bias in Θ and ν, which is mainly caused by the bias in α. The estimators of the diagonal elements of Φ have a downward bias, while those of the off-diagonal elements have an upward bias. Table 1(a) shows that most of the values of the standard deviations are close to those of the RMSE, indicating that the biases in finite samples are negligible. Table 1(b) deals with the case of longer persistence, that is, $\hat{\alpha} + 0.1$. Compared with Table 1(a), the biases, standard deviations and RMSE become larger due to the long range persistence, as expected.

5 Empirical Analysis

This section presents the estimates of the alternative continuous time WSV models using bivariate data for Standard & Poor's 500 Stock Index (S&P) and the Financial Times Stock Exchange 100 Index (FTSE). The sample period for both series is Jan/2/2001 to Jan/28/2011, giving $T = 2534$ observations. Returns, y_{it}, are defined as $\{\log P_{it} - \log P_{i,t-1}\}$, where P_{it} is the closing price on day t for stock i.

We estimate one-factor and two-factor WSV models, including the simple WSV, 2WSV and FIWSV models given in Section 3. As a diagnostic statistic, we use the conventional GMM J-test of Hansen (1982). Under standard regularity conditions, the minimized value of the objective function multiplied by the sample size is asymptotically chi-square distributed, which allows for an omnibus test of the overidentifying restrictions.

The first column of Table 2 presents the GMM estimates of the WSV model. The estimate of Φ

15

is close to the zero matrix, implying that $\text{Exp}(\hat{\Phi})$ is close to the identity matrix (see the first column of Table 3). Intuitively, we may consider the process $V(t)$ for constructing covariance matrices as the sum of outer products of VAR processes, $Y_i(t+h) \sim N(\text{Exp}(\Phi)Y_i(t), I_n)$, as described in the vector OU process (2). In this sense, the $Y_i(t)$ which constitute $V(t)$ may be considered as a process integrated of order one. Hence, we need to consider even longer persistence in the covariance structure. As the estimates of Θ and ν are insignificant, this implies the possibility that the covariance matrix estimator is inconsistent for the presence of longer memory. The J-test of overidentifying restrictions strongly rejects the WSV model.

Table 2 also gives the GMM estimates of the 2WSV model, which produces longer persistence in the covariance matrix process than does the WSV model. With respect to the first factor, the estimate of ν is 4.20 and significant. Although the estimate of Φ is close to the zero matrix, the estimates of Φ_{11}, Φ_{21} and Φ_{22} are significant. For the second factor, the estimate of ν_2 is 7.82. The sum of ν and ν_2 is 12, which is close to the estimates for the one-factor model. The estimates of ν_2 and Φ_2 are insignificant.

The second column of Table 3 shows the persistence of these two factors, with $\text{Exp}(\hat{\Phi}) \simeq 0.94 I_2$ and $\text{Exp}(\hat{\Phi}_2) \simeq O$. Returning to Table 2, the J-test of overidentifying restrictions rejects the 2WSV model. Noting that the J-test statistic is the minimized value of the objective function multiplied by the sample size, the objective function for the 2WSV model is improved compared with the case of the WSV model. Note that it is not possible to impose the restriction on Φ_2 of $\text{Exp}(\Phi_2) = O$ in order to reduce the number of parameters, as the restriction indicates that the diagonal elements of Φ_2 are $-\infty$.

Now we turn to the estimates of the FIWSV process shown in Table 2. By construction, the FIWSV model has a longer memory in the covariance structure than the two-factor WAR process.

16

The GPH estimate of the memory parameter α is 0.335 and significant. As the model with $\alpha = 0$ corresponds to the simple WSV model, this shows that the data prefer the FIWSV model. The estimate of ν is 2.51 and significant. The estimates of the elements of Φ are insignificant, corresponding to $\text{Exp}(\hat{\Phi}) \simeq O$, which is shown in the third column of Table 3. Intuitively, $Y_i(t)$, which constitutes $V(t)$ in equation (5), can be considered as a process with fractional Gaussian noise, which is defined by the difference between a fractional Brownian motion process with its one-period lag. Again we are unable to restrict $\Phi = \text{Log}(O)$ in order to reduce the number of parameters. The J-test of overidentifying restrictions does not reject the FIWSV model. Therefore, the data prefer the FIWSV model, as compared with the WSV and 2WSV models.

6 Conclusion

In this paper, we suggested the continuous time fractionally integrated Wishart stochastic volatility (FIWSV) process, which is an extension of the Wishart autoregressive process of Bru (1991) and Gourieroux (2006). We also derived the closed-form expression of the conditional Laplace transform (moment generating function) of the FIWSV model. We proposed a two-step procedure for estimating the FIWSV model. The first step is to estimate the parameter of fractional integration via the GPH estimator. In the second step, we estimate the remaining parameter using the GMM technique, with Robinson's (2005) asymptotic covariance matrix estimator for the long memory process. We reported the results for Monte Carlo experiments. The empirical results for the bivariate data of the S&P 500 and FTSE 100 indexes showed the usefulness of the new FIWSV specification as compared with the one-factor and two-factor WSV models.

Appendix

A.1 Proof of Proposition 1

In order to prove Proposition 1, we need to show some properties of the matrix-exponential operator. Before we proceed, we should note that for any $m \times m$ matrix A, the spectral decomposition is given by $A = ULU^H$, where U is the unitary matrix, and L is the diagonal matrix of eigenvalues such that $L = \text{diag}\{\lambda_1, \ldots, \lambda_m\}$.

We work with the following two lemmas.

Lemma 2 *For any square matrix A $(m \times m)$, we have the decomposition,*

$$\text{Exp}(A) = U \begin{pmatrix} \exp(\lambda_1) & & O \\ & \ddots & \\ O & & \exp(\lambda_m) \end{pmatrix} U^H.$$

Proof. See Bellman (1970) and Chiu et al. (1996).

Lemma 3 *For any square matrix A $(m \times m)$,*

$$\frac{d}{ds} \text{Exp}(As) = A \, \text{Exp}(As).$$

Proof. It is straightforward from the definition, as we have

$$\frac{d}{ds} \text{Exp}(As) = \frac{d}{ds} \left\{ I + \lim_{n \to \infty} \sum_{i=1}^{n} \frac{1}{i!} A^i s^i \right\} = \lim_{n \to \infty} \frac{d}{ds} \sum_{i=1}^{n} \frac{1}{i!} A^i s^i$$

$$= \lim_{n \to \infty} A \sum_{i=1}^{n} \frac{1}{(i-1)!} A^{i-1} s^{i-1} = A \, \text{Exp}(As). \quad \blacksquare$$

Under the stationary condition, the transition at horizon h tends to the invariant distribution of the process in the long run $(h \to \infty)$.

Proof of Proposition 1. By Lemma 1, we have $M(\infty) = O$ if and only if the stationary condition is satisfied.

Denote $g(h) = \text{Exp}(\Phi h)\Theta\Theta'\text{Exp}(\Phi'h)$. Then, $\Omega(h) = \int_0^h G(s)ds$. Noting that $\frac{dg(s)}{ds} = \Phi g(s) + g(s)\Phi'$ by Lemma 2, we have

$$\text{vec}(\Omega(h)) = [(I_m \otimes \Phi) + (\Phi \otimes I_m)]^{-1}\text{vec}(\text{Exp}(\Phi h)\Theta\Theta'\text{Exp}(\Phi'h) - \Theta\Theta'),$$

if Φ^{-1} exists. Hence,

$$\text{vec}(\Omega(\infty)) = -[(I_m \otimes \Phi) + (\Phi \otimes I_m)]^{-1}\text{vec}(\Theta\Theta'),$$

if and only if the stationary condition is satisfied.

By setting $h \to \infty$ in equations (7) and (8), we can show that the invariant distribution of $W(t)$ exists, if and only if the stationary condition is satisfied. The invariant distribution which is characterized by the moment generating function, $\Psi_{\Xi,t}(\infty) = [\det(I_m - 2\Xi\Omega(\infty))]^{-\nu/2}$, is the Wishart distribution. ∎

A.2 Proof of Lemma 1

As $Y_i(t)$ is an affine process, the conditional Laplace transform is exponential affine such that $\exp\left(b^\dagger(h)'Y_i(t) + c^\dagger(h)\right)$. By applying the Feynman-Kac argument, we obtain

$$\begin{aligned}
0 = &-\frac{d}{dh}b^\dagger(h)'Y_i - \frac{d}{dh}c^\dagger(h) \\
&+ b^\dagger(h)'\Phi Y_i + \frac{1}{2}\text{tr}\left\{c_\alpha^{-1}h^{2\alpha}\Theta\Theta'b^\dagger(h)b^\dagger(h)'\right\},
\end{aligned}$$

with boundary conditions $b(0) = \gamma$ and $c(0) = 0$. Note that $c_\alpha^{-1}h^{2\alpha}I_m$ is the first derivative of $E[(\tilde{B}_\alpha(t+h) - \tilde{B}_\alpha(t))(\tilde{B}_\alpha(t+h) - \tilde{B}_\alpha(t))']$. By identifying the coefficients of Y_i, we have the ordinary differential equations for $b^\dagger(h)$ as

$$\begin{aligned}
\frac{d}{dh}b^\dagger(h) &= \Phi b^\dagger(h) \\
\frac{d}{dh}c^\dagger(h) &= \frac{1}{2c_\alpha}\text{tr}\left\{h^{2\alpha}\Theta\Theta'b^\dagger(h)b^\dagger(h)'\right\}.
\end{aligned}$$

Thus, we have the solutions in (9).

Using $g(h)$ in Appendix A.1, we have

$$\Omega_\alpha(h) = c_\alpha^{-1} \int_0^h s^{2\alpha} g(s) ds,$$

where c_α is defined by Proposition 2. If $0 < \alpha < 1/2$, we obtain

$$
\begin{aligned}
c_\alpha \Omega_\alpha(h) &= [s^{2\alpha} \Omega(s)]_0^h - \int_0^h (2\alpha) s^{2\alpha-1} g(s) ds \\
&= h^{2\alpha} \Omega(h) - [s^{2\alpha} g(s)]_0^h + \int_0^h s^{2\alpha} \frac{dg(s)}{ds} ds \\
&= h^{2\alpha} \{\Omega(h) - g(h)\} + c_\alpha \Phi \Omega_\alpha(h) + c_\alpha \Omega_\alpha(h) \Phi'.
\end{aligned}
$$

The last equality is based on the fact that $\frac{dg(s)}{ds} = \Phi g(s) + g(s) \Phi'$. If Φ^{-1} exists, we have equation

(10). ∎

A.3 Proof of Proposition 3

By applying the same approach of Gourieroux and Sufana (2010), it is straightforward to show that

the FIWAR process is an affine process. Hence, the conditional Laplace transform is exponential

affine such that $\exp(A(h)W(t) + c(h))$. By applying the Feynman-Kac argument, we obtain

$$
\begin{aligned}
0 = -\mathrm{tr} &\left[\frac{d}{dh} A(h) W \right] - \frac{d}{dh} c(h) \\
&+ \mathrm{tr} \left\{ (\nu \Theta \Theta' + W \Phi' + \Phi W) A(h) + 2 c_\alpha^{-1} h^{2\alpha} W A(h) \Theta \Theta' A(h) \right\}
\end{aligned}
$$

with boundary conditions $A(0) = O$ and $c(0) = 0$. By identifying the coefficients of W, we have

the ordinary differential equations for $A(h)$ and $c(h)$ given by

$$
\begin{aligned}
\frac{d}{dh} A(h) &= \Phi A(h) + \Phi A(h) + 2 c_\alpha^{-1} h^{2\alpha} A(h) \Theta \Theta' A(h) \\
\frac{d}{dh} c(h) &= \mathrm{tr} \left\{ \nu \Theta \Theta' A(h) \right\}.
\end{aligned}
$$

As usual, the function $c(h)$ can be obtained by direct integration.

Following Fonseca, Grasselli and Tebaldi (2008), we linearize the Matrix Riccati equation for $B(h)$ by doubling the dimension of the problem. Consider a decomposition which satisfies

$$A(h) = [F(h)]^{-1} G(h).$$

Then we have

$$\frac{d}{dh}[F(h)A(h)] - \frac{d}{dh}[F(h)]A(h) = F(h)\frac{d}{dh}[A(h)].$$

Substituting the Matrix Riccati equation for $A(h)$ into the above equation, we obtain

$$\frac{d}{dh}G(h) - \frac{d}{dh}[F(h)]A(h) = G(h)\Phi + \left[F(h)\Phi' + 2c_\alpha^{-1}h^{2\alpha}G(h)\Theta\Theta'\right]A_k(h).$$

Hence, we can construct a system of $(2m)$ linear equations given by

$$\frac{d}{dh}G(h) = G(h)\Phi_k,$$

$$\frac{d}{dh}F(h) = -F(h)\Phi' - 2c_\alpha^{-1}h^{2\alpha}G(h)\Theta\Theta',$$

with initial conditions $G(0) = \Xi$ and $F_k(0) = I_m$. The above equations can be written as follows:

$$\frac{d}{dh}\left(G(h)\ \ F(h)\right) = \left(G(h)\ \ F(h)\right)\begin{pmatrix} \Phi & -2c_\alpha^{-1}h^{2\alpha}\Theta\Theta' \\ O & -\Phi' \end{pmatrix}.$$

Hence, we obtain the solution by matrix-exponentiation:

$$\left(G(h)\ \ F(h)\right) = \left(\Xi\ \ I_m\right)N(h), \qquad N(h) = \mathrm{Exp}\,h\begin{pmatrix} \Phi & -2c_\alpha^{-1}h^{2\alpha}\Theta\Theta' \\ O & -\Phi' \end{pmatrix}.$$

Denoting the partitioned matrices of $N(h)$ as $N_{ij}(h)$ $i,j = 1,2$, we have $N_{11}(h) = \mathrm{Exp}(\Phi h)$, $N_{22}(h) = \mathrm{Exp}(-\Phi'h)$ and $N_{21}(h) = O$. With tedious but cumbersome matrix calculus, we have $N_{12}(h) = -2\Omega_\alpha(h)\mathrm{Exp}(-\Phi'h)$, if Φ^{-1} exists. Since we have

$$\left(G(h)\ \ F(h)\right) = \left(\Xi N_{11}(h) + N_{21}(h)\ \ \Xi N_{21}(h) + N_{22}(h)\right),$$

21

we obtain the solution of $A(h)$ as

$$A(h) = \text{Exp}(\Phi' h)(I_m - 2\Xi\Omega_\alpha(h))^{-1})\Xi\text{Exp}(\Phi h).$$

Using the solution of $A(h)$ and $c(h)$, we have equation (12). ∎

A.3 Proof of Proposition 4

We can use the approach, which is analogous to Proposition 3, for obtaining the sytem of Riccati equations and solutions of $\tilde{A}(h)$ and $\tilde{b}(h)$. Hence, we explain the way to derive the solution of $\tilde{c}(h)$.

By linearizing the Matrix Riccati equation for $\tilde{A}(h)$, we have $\tilde{F}(h)$ which satisfies

$$[\tilde{F}(h)]^{-1}\frac{d}{dh}\tilde{F}(h) = -\Phi' - 2c_\alpha^{-1}h^{2\alpha}\tilde{A}(h).$$

We can remove $\tilde{A}(h)$ from the ODE of $\tilde{c}(h)$, in order to have

$$\frac{d}{dh}\tilde{c}(h) = -\frac{1}{2}\nu\text{tr}\left([\tilde{F}(h)]^{-1}\frac{d}{dh}\tilde{F}(h) + \Phi'\right).$$

Note that $2c_\alpha^{-1}h^{2\alpha}$ was cancelled out by the substitution. On the other hand, we also have $\tilde{F}(h) = \tilde{N}_{22}(h)$ defined by Proposition 4, to derive the solution of $\tilde{A}(h)$. Now we can integrate the last equation and obtain the solution of $\tilde{c}(h)$ given by Proposition 4.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001), "The Distribution of Realized Exchange Rate Volatility", *Journal of the American Statistical Association*, **96**, 42–55.

Asai, M. and M. McAleer (2012), "Leverage and Feedback Effects on Multifactor Wishart Stochastic Volatility for Option Pricing", Unpublished paper, Faculty of Economics, Soka University.

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

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

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica* **59**, 817–858.

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

Bauwens, L., S. Laurent, and J.V.K. Rombouts (2006), "Multivariate GARCH: A Survey", *Journal of Applied Econometrics*, **21**, 79–109

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

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 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.

Bru, M.-F. (1991), "Wishart Processes", *Journal of Theoretical Probability*, **4**, 725–751.

Caporin, M. and M. McAleer (2012), "Do We Really Need Both BEKK and DCC? A Tale of Two Multivariate GARCH Models", *Journal of Economic Surveys*, **26**, 736–751.

Chernov, M., A.R. Gallant, E. Ghysels, and G. Tauchen (2003), "Alternative Models for Stock Price Dynamics", *Journal of Econometrics*, **116**, 225–257.

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*, Springer-Verlag, New York, 365–400.

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

Comte, F. and E. Renault (1996), "Long Memory Continuous Time Models", *Journal of Econometrics*, **73**, 101–149.

Comte, F. and E. Renault (1998), "Long Memory in Continuous-Time Stochastic Volatility Models", *Mathematical Finance*, **8**, 291–323.

Dai, W. and C.C. Heyde (1996), "It's Formula with Respect to Fractional Brownian Motion and Its Application", *Journal of Applied Mathematics and Stochastic Analysis*, **9**, 439–448.

Deo, R.S. and C.M. Hurvich (2001), "On the Log Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models", *Econometric Theory*, **17**, 686–710.

Ding, Z., C.W.J. Granger, and R. Engle (1993), "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance*, **1**, 83–106.

Duffie, D., J. Pan, and K. Singleton (2000), "Transform Analysis and Asset Pricing for Affine Jump Diffusions", *Econometrica*, **68**, 1343–1376.

Fonseca, J., M. Grasselli, and C. Tebaldi (2008), "A Multifactor Volatility Heston Model", *Quantitative Finance*, **8**, 591–604.

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.

Gourieroux, C. (2006), "Continuous Time Wishart Process for Stochastic Risk", *Econometric Reviews*, **25**, 177–217.

Gourieroux, C. and R. Sufana (2010), "Derivative Pricing with Wishart Multivariate Stochastic Volatility", *Journal of Business & Economic Statistics*, **28**, 438–451.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, **50**, 1029–1054.

Harvey, A.C. (1993), "Long Memory in Stochastic Volatility", Unpublished Paper, London School of Economics.

Koopman, S.J., B. Jungbacker, and E. Hol (2005), "Forecasting Daily Variability of the S&P 100 Stock Index Using Historical Realized and Implied Volatility Measurements", *Journal of Empirical Finance*, **12**, 445–475.

McAleer, M., F. Chan, S. Hoti and O. Lieberman (2008), "Generalized Autoregressive Conditional Correlation", *Econometric Theory*, **24**, 1554–1583.

Mandelbrot, M.M and J.W. van Ness (1968), "Fractional Brownian Motions, Fractional Noises and Applications", *SIAM Review*, **10**, 422-437.

Muhle-Karbe, J., O. Pfaffel, and R. Stelzer (2012), "Option Pricing in Multivariate Stochastic Volatility Models of OU Type", *SIAM Journal on Financial Mathematics*, **3**, 66–94.

Newey, W.K., and K.D. West (1987), "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica* **55**, 703–708.

Pong S., M.B. Shackelton, S.J. Taylor, and X. Xu (2004), "Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models", *Journal of Banking and Finance*, **28**, 2541–2563.

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

Robinson, P.M. (2005), "Robust Covariance Matrix Estimation: HAC Estimates with Long Memory/Antipersistence Correction", *Econometric Theory*, **21**, 171–180.

Table 1: Monte Carlo Results for FIWSV Model

(a) DGP1

Para.	True	Bias	Std.Dev.	RMSE
α	0.3354	-0.0892	0.1853	0.2057
Θ_{11}	0.0470	0.0121	0.2143	0.2146
Θ_{12}	-0.0155	0.0120	0.3243	0.3245
Θ_{22}	-0.0765	0.0515	0.4930	0.4956
ν_1	3.0000	0.2035	3.0808	3.0875
Φ_{11}	-49.367	-0.2430	2.5288	2.5405
Φ_{21}	-9.8194	0.4994	2.2700	2.3243
Φ_{12}	-9.8870	0.3420	2.3569	2.3816
Φ_{22}	-42.784	-0.7457	2.9187	3.0125

(b) DGP2

Para.	True	Bias	Std.Dev.	RMSE
α	0.4354	-0.1489	0.2180	0.2581
Θ_{11}	0.0470	0.0201	0.2324	0.2333
Θ_{12}	-0.0155	0.0322	0.1693	0.1723
Θ_{22}	-0.0765	0.0575	0.3315	0.3534
ν_1	3.0000	0.2183	3.1611	3.1723
Φ_{11}	-49.367	-0.2270	2.1101	2.1223
Φ_{21}	-9.8194	0.5105	2.3513	2.4061
Φ_{12}	-9.8870	0.4675	2.1337	2.1843
Φ_{22}	-42.784	-0.8073	2.9107	2.9954

Note: We estimate the long memory parameter α by the GPH method in the first step, and obtain the GMM estimates based on $\hat{\alpha}$ in the second step.

Table 2: Estimates for WSV Processes

Parameters	WSV		2WSV		FIWSV	
GPH Estimate						
α					0.3354	(0.0237)
GMM Estimates						
Θ_{11}	0.0009	(0.0012)	0.0022	(2.1251×10^{-5})	0.0470	(0.0154)
Θ_{21}	0.0001	(0.0011)	0.0004	(5.8339×10^{-5})	-0.0155	(0.0126)
Θ_{22}	0.0006	(0.0008)	0.0021	(2.0196×10^{-5})	-0.0765	(0.0062)
ν	12.937	(36.371)	4.1960	(0.4509)	2.5142	(0.4688)
Φ_{11}	-0.0393	(0.0986)	-0.0711	(0.0014)	-49.367	(31.823)
Φ_{21}	0.0160	(0.1427)	0.0065	(0.0038)	-9.8194	(19.208)
Φ_{12}	-0.0140	(0.0877)	-0.0273	(0.0029)	-9.8870	(23.268)
Φ_{22}	-0.0164	(0.0447)	-0.0639	(0.0014)	-42.784	(6.8394)
ν_2			7.8179	(6.9510)		
$\Phi_{2,11}$			-12.637	(19.700)		
$\Phi_{2,21}$			1.7311	(9.9463)		
$\Phi_{2,12}$			2.5824	(9.8638)		
$\Phi_{2,22}$			-3.8836	(2.2194)		
J-test	70.663		15.727		16.144	
d.o.f.	9		4		9	
p-Value	[0.0000]		[0.0034]		[0.0639]	

Note: Standard errors are in parentheses. 'J-test' is the GMM test of overidentifying restrictions. The inverse of the weighting matrix is calculated using the asymptotic covariance matrix estimator of Robinson (2005) for the moment conditions of FIWSV, and the estimator of Newey and West (1987) with a lag length of 20 for the WSV model and 200 for 2WSV.

Table 3: Estimates for Persistence Parameters

$$V(t) = \sum_{i=1}^{\nu} Y_i(t)Y_i(t)', \qquad dY_i(t) = \Phi Y_i(t)dt + \Theta d\tilde{B}_{\alpha,i}(t)$$

$$Y_i(t+1)|Y_i(t) \sim N(\text{Exp}(\Phi)Y_i(t), \Omega_\alpha(h))$$

Parameters	WSV	2WSV	FIWSV
α	0	0	0.3354
$\text{Exp}(\Phi)$			
(1,1)	0.9614	0.9366	0.0000
(2,1)	0.0156	0.0092	0.0000
(1,2)	-0.0136	-0.0196	0.0000
(2,2)	0.9836	0.9363	0.0000
$\text{Exp}(\Phi_2)$			
(1,1)		0.0019	
(2,1)		0.0067	
(1,2)		0.0099	
(2,2)		0.0351	

Note: Although the estimates of ν are not integers, this expression gives an intuitive interpretation of the WSV models.