ICAE

Instituto Complutense de Análisis Económico

Testing for Volatility Co-movement in Bivariate Stochastic Volatility Models

Jinghui Chen
Graduate School of International Social Sciences Yokohama National University

Masahito Kobayashi
Department of Economics Yokohama National University

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

Abstract

The paper considers the problem of volatility co-movement, namely as to whether two financial returns have perfectly correlated common volatility process, in the framework of multivariate stochastic volatility models and proposes a test which checks the volatility co-movement. The proposed test is a stochastic volatility version of the co-movement test proposed by Engle and Susmel (1993), who investigated whether international equity markets have volatility co-movement using the framework of the ARCH model.

In empirical analysis we found that volatility co-movement exists among closelylinked stock markets and that volatility co-movement of the exchange rate markets tends to be found when the overall volatility level is low, which is contrasting to the often-cited finding in the financial contagion literature that financial returns have co-movement in the level during the financial crisis.

Keywords Lagrange multiplier test; Volatility co-movement, Stock markets, Exchange rate Markets; Financial crisis

JEL Classification C12, C58, G01, G11

Working Paper nº 1710

February, 2017

UNIVERSIDAD
COMPLUTENSE
MADRID

ISSN: **2341-2356**

Testing for Volatility Co-movement in Bivariate Stochastic Volatility Models

Jinghui Chen

Graduate School of International Social Sciences
Yokohama National University

Masahito Kobayashi*

Department of Economics
Yokohama National University

Michael McAleer

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

Revised: February 2017

*corresponding author: Masahito Kobayashi, Department of Economics, Yokohama National University, Yokohama 240-8501, Japan; email: kobayashi-masahito-nz@ynu.ac.jp

Abstract

The paper considers the problem of volatility co-movement, namely as to whether two financial returns have perfectly correlated common volatility process, in the framework of multivariate stochastic volatility models and proposes a test which checks the volatility co-movement. The proposed test is a stochastic volatility version of the co-movement test proposed by Engle and Susmel (1993), who investigated whether international equity markets have volatility co-movement using the framework of the ARCH model.

In empirical analysis we found that volatility co-movement exists among closely-linked stock markets and that volatility co-movement of the exchange rate markets tends to be found when the overall volatility level is low, which is contrasting to the often-cited finding in the financial contagion literature that financial returns have co-movement in the level during the financial crisis.

Keywords: Lagrange multiplier test; Volatility co-movement, Stock markets, Exchange rate Markets; Financial crisis

JEL Classification: C12, C58, G01, G11

1 Introduction

This paper considers the problem as to whether financial returns have volatility co-movement using the framework of multivariate stochastic volatility models that were suggested by Harvey et al. (1994). We propose a stochastic volatility version of the ARCH test proposed by Engle and Kozicki (1993) and Engle and Susmel (1993), who investigated volatility co-movement, namely whether international equity markets have a common volatility process, using the multivariate ARCH model framework, and found groups of countries that had a similar time-varying volatility. Fleming et al. (1998) used the multivariate stochastic volatility model to estimate volatility linkages across stock, bond, and money markets, and found strong correlation between markets. Fleming et al. (1998) also tested perfectly correlated volatility processes, extending the model of Tauchen and Pitts. (1983). Their definition of volatility linkage is stronger than the mere presence of a common factor in volatility processes in that they have no idiosyncratic volatility factor. They also conducted a Wald-type test, using the GMM framework, and rejected the null hypothesis of perfectly correlated volatility and concluded that cross-market hedging is imperfect.

However, the use of the Wald and likelihood ratio tests in the classical hypothesis testing framework, which use the estimator of the volatility correlation parameter, is inappropriate for the null hypothesis of perfectly correlated volatility, as the asymptotic distribution of the Wald test statistics is different from the conventional chi-squared distribution, as shown, for example, in Chernoff (1954), since, as the correlation estimator cannot be greater than or equal to one in absolute value, the distribution of the estimator of the constrained parameter is asymmetric, and hence non-normal, when the true correlation coefficient is unity under the null hypothesis.

The paper proposes a new Lagrange multiplier (LM) test for volatility co-movement, namely the hypothesis that the volatility processes of bivariate series have a perfectly correlated common volatility factor. We use the framework of multivariate stochastic volatility model proposed by Harvey et al. (1994), where the log volatility follows vector autoregressive (VAR) process of order one with diagonal autoregressive coefficient matrix.

The Lagrange multiplier test principle is the only alternative for this problem in deriving the test statistics because it estimates only the null model and does not estimate the parameter on the boundary of the parameter space. Then the test statistic can follow the conventional chi-squared asymptotic distribution under the null hypothesis.

To the best of our knowledge, the Lagrange multiplier test statistic for the perfectly correlated volatility processes has not been proposed in the literature, except Chiba and Kobayashi (2013), who employed the unconventional assumption that the log of squared returns is normally distributed in deriving the test statistic. It is not without reason why the LM test has not been proposed; the conventional method to obtain a score function is unworkable in this problem, because the derivative of the transition density is intractable under the null hypothesis, as the transition disturbance has zero variance. We here derive the score function using the ingenious method devised by Chesher (1984), which is the main technical breakthrough in tackling this problem.

Our test can be regarded as a test for the number of stochastic volatility factors, in line with the definition of Harvey et al. (1994) and Cipollini and Kapetanios (2008), when the number of factors is one under the null hypothesis. Cipollini and Kapetanios (2008) used a linearized model for the log of squared returns, and used the principal component methodology of Stock and Watson (2002) in deciding the number of factors. Their method has the advantage in that it is applicable when the number of variables is large, though it is not a statistical test. The new test developed in the paper is the only existent statistical test for the hypothesis.

We employed the quadrature in evaluating the likelihood function proposed by Watanabe (1999) and Kitagawa (1987) and the method proposed by Hamilton (1989) in evaluating the score function.

The remainder of the paper is organized as follows. Section 2 presents the model, Section 3 develops the LM test statistic, Section 4 presents Monte Carlo experiments, and Section 5 illustrates two empirical analyses, and concluding remarks are given in Section 6. Appendices illustrate the pre-orthogonalization of data and the derivation of the score functions.

2 Model

We here consider the unconstrained bivariate stochastic volatility model. Under the alternative hypothesis, the observation vector $y_t = (y_{1t}, y_{2t})'$ is expressed as follows:

(Model under the Alternative Hypothesis):

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \exp\left(\frac{h_{1t}}{2}\right) & 0 \\ 0 & \exp(\frac{h_{2t}}{2}) \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, t = 1, \cdots, T, \tag{1}$$

$$\begin{pmatrix} h_{1t} \\ h_{2t} \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ 0 & \psi \end{pmatrix} \begin{pmatrix} h_{1,t-1} \\ h_{2,t-1} \end{pmatrix} + \begin{pmatrix} \sigma & 0 \\ \omega & \sqrt{\lambda} \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}, t = 1, \cdots, T, \tag{2}$$

$$\begin{pmatrix} h_{11} \\ h_{21} \end{pmatrix} = \begin{pmatrix} \sigma/\sqrt{(1-\rho^2)} & 0 \\ \omega/\sqrt{(1-\psi^2)} & \sqrt{\lambda/(1-\psi^2)} \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}, \tag{3}$$

$$(e_{1t}, e_{2t}, u_{1t}, u_{2t})' \sim N(0, \mathbf{I}_4),$$

where the log volatility $(h_{1t}, h_{2t})'$ follows a stationary bivariate autoregressive process of order one, and disturbance term is normally distributed contemporaneously and serially independent with zero mean unit variance.

The null hypothesis to be tested is

$$\textbf{(Null Hypothesis):} \quad \lambda = 0, \psi = \rho, \omega = \sigma. \tag{4}$$

Under the nullhypothesis the joint distribution of the state variable $(h_{1t}, h_{2t})'$ is degenerate, since the disturbance term of the transition equation (2) is $(\sigma u_{1t}, \omega u_{1t} + \sqrt{\lambda} u_{2t})'$; then the measurement equations have a single common volatility, which is expressed as

$$\textbf{(Null Model):} \quad \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \exp(h_{1t}/2) \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \tag{5}$$

$$h_{1t} = \rho h_{1,t-1} + \sigma e_{1t,}, \; h_{2t} = h_{1t}, \quad t = 1, \cdots, T. \tag{6}$$

The multivariate stochastic volatility model was originally suggested by Harvey et al. (1994) and was examined in detail in Danıelsson (1998) and Asai et al. (2006). The null model of our paper is the stochastic volatility factor model discussed in Harvey et al. (1994) and Cipollini and Kapetanios (2008) in the simple case when the number of factors is one.

A remark may be in order here on the assumption that the disturbance term of the measurement equation (1), namely $(e_{1t}, e_{2t})'$, is contemporaneously uncorrelated. We can justify the use of this simple assumption by showing in Appendices that the data can be transformed so as to satisfy the assumption of the model under the null hypothesis.

3 LM test statistic

3.1 Notation

We propose the LM test for the null hypothesis (4) for the observation series y_{1t} and y_{2t}. Define the unconstrained parameter vector as $\theta_1 = (c, \psi, \omega, \rho, \sigma)$, and the constrained parameter vector as $\theta_0 = (0, \psi, \omega, \psi, \omega)$.

First, we obtain the maximum likelihood estimator of the constrained parameter, θ_0, of the state space system (1) and (2). Denote $\mathbf{y}_1 = (y_{11}, y_{12}, \ldots, y_{1T})'$, $\mathbf{y}_2 = (y_{21}, y_{22}, \ldots, y_{2T})'$, $\mathbf{h}_1 = (h_{11}, h_{12}, \ldots, h_{1T})'$, and $\mathbf{h}_2 = (h_{21}, h_{22}, \ldots, h_{2T})'$. The likelihood is expressed as

$$f(\mathbf{y}_1, \mathbf{y}_2) = \int\int f(\mathbf{h}_1, \mathbf{h}_2, \mathbf{y}_1, \mathbf{y}_2) d\mathbf{h}_1 d\mathbf{h}_2 = \int\int f(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{h}_1, \mathbf{h}_2) f(\mathbf{h}_2 | \mathbf{h}_1) f(\mathbf{h}_1) d\mathbf{h}_1 d\mathbf{h}_2,$$

where the explicit form of $f(\mathbf{y}_1|\mathbf{h}_1), f(\mathbf{y}_2|\mathbf{y}_1, \mathbf{h}_1, \mathbf{h}_2), f(\mathbf{h}_2|\mathbf{h}_1)$, and $f(\mathbf{h}_1)$ are given in Appendices. We perform this integration numerically by the quadrature suggested by Kitagawa (1987) and Watanabe (1999) in evaluating the likelihood function.

Second, we derive the score function under the alternative hypothesis and evaluate it under the null hypothesis. Denote

$$\mathbf{y}_T = (y_{11}, y_{12}, \ldots, y_{1T},\ y_{21}, y_{22}, \ldots, y_{2T})' = (\mathbf{y}_1, \mathbf{y}_2)$$

and the score function as

$$\frac{\partial \log f(\mathbf{y}_t)}{\partial \theta_1} = \left(\frac{\partial \log f(\mathbf{y}_t)}{\partial \lambda}, \frac{\partial \log f(\mathbf{y}_t)}{\partial \psi}, \frac{\partial \log f(\mathbf{y}_t)}{\partial \omega}, \frac{\partial \log f(\mathbf{y}_t)}{\partial \rho}, \frac{\partial \log f(\mathbf{y}_t)}{\partial \sigma} \right).$$

Noting that $\log f(\mathbf{y}_t|\mathbf{y}_{t-1}) = \log f(\mathbf{y}_t) - \log f(\mathbf{y}_{t-1})$, conditional score function is expressed as

$$Q_t = \frac{\partial \log f(\mathbf{y}_t|\mathbf{y}_{t-1})}{\partial \theta_1'} = \frac{\partial \log f(\mathbf{y}_t)}{\partial \theta_1'} - \frac{\partial \log f(\mathbf{y}_{t-1})}{\partial \theta_1'}.$$

Then the estimated Fisher information matrix is

$$\textbf{(Fisher Information Matrix):}\ \mathcal{I}(\theta) = \frac{1}{T} \sum_{t=1}^{T} Q_t Q_t',$$

and the score vector is

$$\textbf{(Score Vector):}\ \mathcal{U}(\theta) = \frac{1}{T} \sum_{t=1}^{T} Q_t = \frac{1}{T} \frac{\partial \log f(\mathbf{y}_T)}{\partial \theta_1'}. \tag{7}$$

3.2 Test statistic

The LM test statistic is defined by

$$\textbf{(LM Test Statistic):}\quad LM = T \times \mathcal{U}'(\hat{\theta}_0)\mathcal{I}(\hat{\theta}_0)^{-1}\mathcal{U}(\hat{\theta}_0), \tag{8}$$

where $\hat{\theta}_0$ is the maximum likelihood estimator of θ_0 under the null hypothesis. Then, under the regularity condition that the parameters to be estimated lie in the interior of the parameter space and the estimated information matrix converges to a nonsingular matrix, and hence the estimators are normally distributed asymptotically, we have

$$\textbf{(Asymptotic Distribution):}\ LM \xrightarrow{L} \chi^2(3),$$

when T is sufficiently large, as shown by Davidson and MacKinnon (1993, p. 91), with three degrees of freedom of the asymptotic χ^2-distribution corresponding to the three restrictions of the null hypothesis (4).

The score function (7) cannot be evaluated by the conventional method; the derivative $\partial f(\mathbf{h}_2|\mathbf{h}_1)/\partial\lambda$ in

$$\frac{\partial f(\mathbf{y}_t)}{\partial\lambda} = \int f(\mathbf{y}_1, \mathbf{y}_2|\mathbf{h}_1, \mathbf{h}_2)\frac{\partial f(\mathbf{h}_2|\mathbf{h}_1)}{\partial\lambda} f(\mathbf{h}_1)d\mathbf{h}_1 d\mathbf{h}_2 \tag{9}$$

diverges as $\lambda \to 0$, because $f(\mathbf{h}_2|\mathbf{h}_1)$ is a normal density with variance λ and hence its derivative, as well as the score function, with respect λ under the null hypothesis cannot be evaluated directly. We have circumvented this difficulty using the ingenious method proposed in Chesher (1984). The algebraic details for the derivation of (10) are found Appendices.

We here give only the final formula of the score functions as follows:

(Score Functions)

$$\begin{aligned}
\frac{\partial \log f(\mathbf{y})}{\partial\lambda}|_{\mathbf{H}_0} &= \frac{1}{8} tr E_{\mathbf{h}_1|\mathbf{y}}\Big(-2\times \mathbf{V}_\psi \mathbf{Y}_2 \exp(-\mathbf{h}_1)\mathbf{1}_{1\times T} + \mathbf{1}_{T\times T}\mathbf{V}_\psi \\
&\quad + \mathbf{V}_\psi \mathbf{Y}_2 \exp(-\mathbf{h}_1)\exp(-\mathbf{h}_1')\mathbf{Y}_2 - 2\mathbf{V}_\psi \mathbf{Y}_2 \mathbf{H}_1^{-1}\Big), \\
\frac{\partial \log f(\mathbf{y})}{\partial\psi}|_{\mathbf{H}_0} &= \frac{1}{2}\mathbf{1}_{1\times T}\mathbf{V}_\psi^{1/2}\mathbf{Z}_\psi E_{\mathbf{h}_1|\mathbf{y}}[\mathbf{h}_1] - \frac{1}{2} tr\left[\mathbf{Y}_2 \mathbf{V}_\psi^{1/2}\mathbf{Z}_\psi E_{\mathbf{h}_1|\mathbf{y}}[\mathbf{h}_1\exp(-\mathbf{h}_1')]\right], \\
\frac{\partial \log f(\mathbf{y})}{\partial\omega}|_{\mathbf{H}_0} &= -\frac{1}{2\omega}\mathbf{1}_{1\times T}E_{\mathbf{h}_1|\mathbf{y}}[\mathbf{h}_1] + \frac{1}{2\omega}(\mathbf{y}_2\circ\mathbf{y}_2)' E_{\mathbf{h}_1|\mathbf{y}}[\exp(-\mathbf{h}_1)\circ\mathbf{h}_1], \\
\frac{\partial \log f(\mathbf{y})}{\partial\rho}|_{\mathbf{H}_0} &= -\frac{\partial \log f(\mathbf{y})}{\partial\psi}|_{\mathbf{H}_0} - \frac{\psi}{1-\psi^2} - \frac{1}{2}\omega^{-2} tr\left(\frac{\partial \mathbf{V}_\psi^{-1}}{\partial\psi}E_{\mathbf{h}_1|\mathbf{y}}\left(\mathbf{h}_1\mathbf{h}_1'\right)\right), \\
\frac{\partial \log f(\mathbf{y})}{\partial\sigma}|_{\mathbf{H}_0} &= -\frac{\partial \log f(\mathbf{y})}{\partial\omega}|_{\mathbf{H}_0} - \frac{t}{\omega} + \frac{1}{\omega^3} tr\left(\mathbf{V}_\psi^{-1}E_{\mathbf{h}_1|\mathbf{y}}\left(\mathbf{h}_1\mathbf{h}_1'\right)\right),
\end{aligned} \tag{10}$$

where $\circ$ denotes the operator of the element-by-element multiplication (the Hadamard product),

$$\exp(-\mathbf{h}_1) = (\exp(-h_{11}), \ldots, \exp(-h_{1T}))', \tag{11}$$

$$\mathbf{Y}_2 = diag(\mathbf{y}_2 \circ \mathbf{y}_2), \tag{12}$$

$$\mathbf{H}_1 = diag(\exp(h_{11}), \ldots, \exp(h_{1T})), \tag{13}$$

and $\mathbf{V}_\psi$ is the covariance matrix of $\mathbf{h}_2$, whose square root ${\mathbf{V}_\psi}^{1/2}$ is defined by the Cholesky decomposition.

The algebraic details of the derivation of (10) are found in Appendices.

4 Monte Carlo experiments

In order to confirm that the proposed statistic is asymptotically distributed as $\chi^2(3)$ under the null, and whether it has power to reject a false hypothesis, we conduct two Monte Carlo experiments. The number of iterations is 1000 for the experiment under the null model and 100 for the experiment under the alternative hypothesis. The number of iterations of the latter is not large, but this limitation is unavoidable because the convergence of the maximum likelihood estimation is slow and hence the calculation of the test statistics require considerable computational time when the data is generated from the alternative hypothesis.

It took two minutes on the average to calculate one iteration for the data for Tables 1 and 2 when the sample size is 500 with the parallel computing tool of MATLAB (8 threads) using a PC with Intel's Core I7 -3770K . The codes are available from the authors on request.

4.1 Size of Test and Null Distribution

First, we generate artificial samples drawn from the null hypothesis, calculate the test statistic, and obtain the empirical distribution of the test statistic under the null hypothesis. Second, we estimate the empirical power by the ratio of the test statistic that exceeds the theoretical critical value. We also obtain the empirical distribution of the test statistic under the null and alternative hypotheses by using kernel estimation and histogram to show that it follows the $\chi^2(3)$ distribution with sufficient precision.

The rejection rates for some critical values and sample sizes are shown in Table 1, when the data is generated under the null hypothesis, where $\theta_0 = (\lambda = 0, \psi, \omega, \rho = \psi, \sigma = \omega)$. Table

1 shows that, as the sample size T increases, the rejection rate approaches to the theoretical significance level. The empirical size of the test is sufficiently close to the theoretical value when the sample size is 500, which suggests that we should use data with at least 500 observations in practice. The empirical null distribution of the test statistic when $T = 500$ is shown in Figures 1-3.

Table 1: Rejection Rates of the Null under the Null Hypothesis

Parameter Values in H_0					Rejection rates					
λ	ψ	ω	ρ	σ	T=100		T=200		T=500	
					5%	1%	5%	1%	5%	1%
0	0.7	1	0.7	1	10.2%	4.3%	8.3%	2.7%	7.1%	1.3%
0	0.9	1	0.9	1	21.0%	9.5%	13.1%	4.8%	6.3%	1.7%
0	0.95	0.45	0.95	0.45	22.5%	10.2%	14.7%	5.6%	7.4%	1.6%

Note: The number of iterations is 1000.

Figure 1: Histogram of LM Test Statistic at ψ=0.7

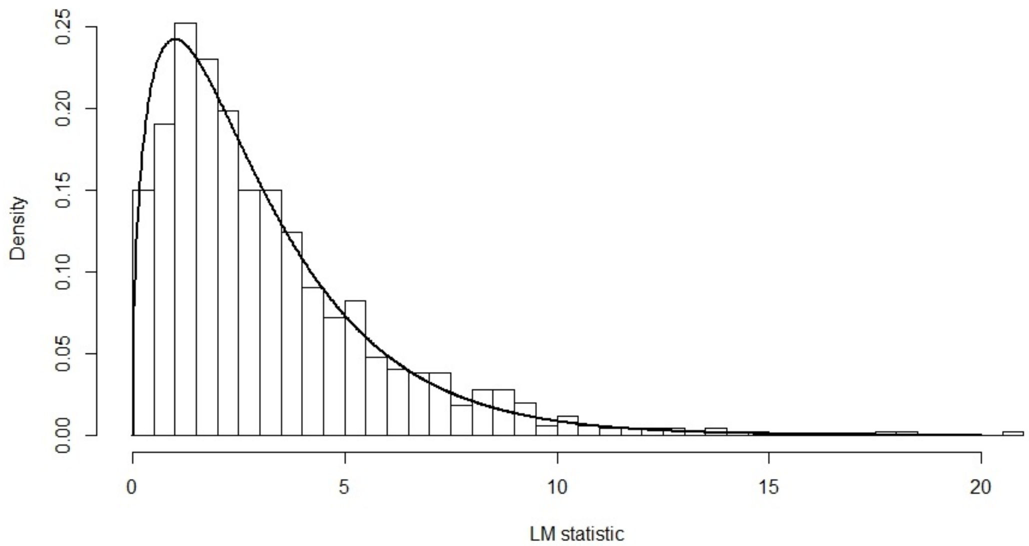


Figure 2: Histogram of LM Test Statistic at ψ=0.9

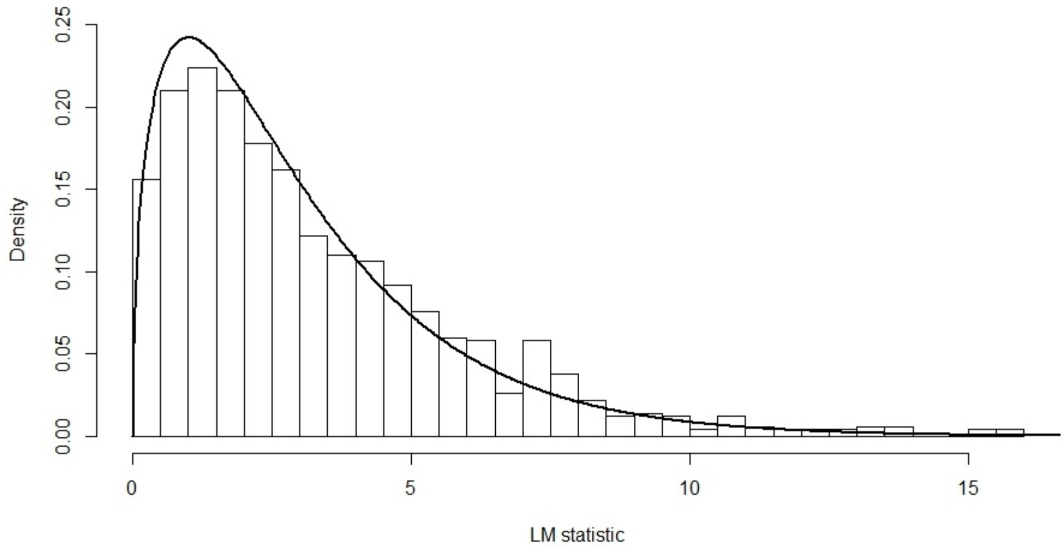
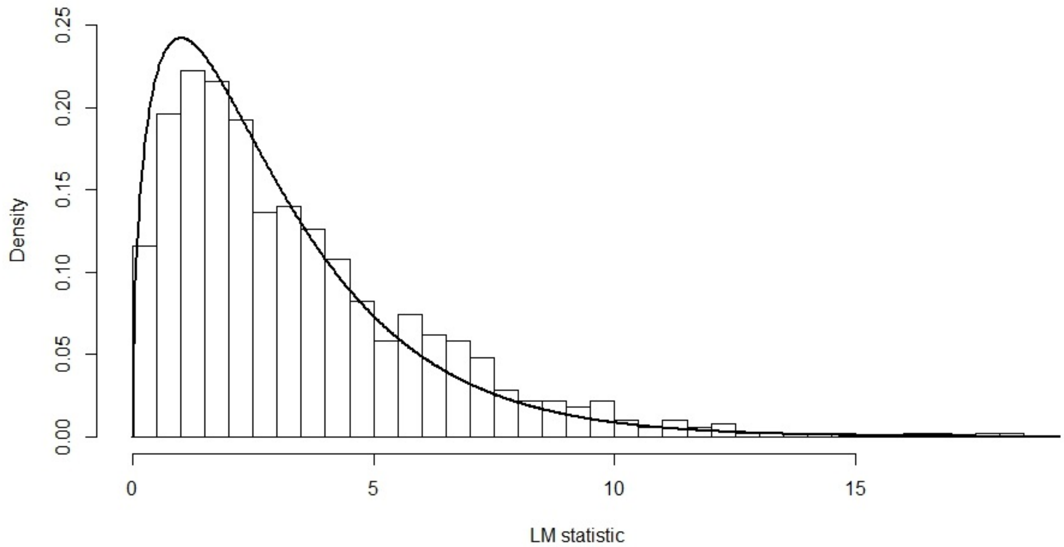


Figure 3: Histogram of LM Test Statistic at ψ=0.95



4.2 Power of test

We generate artificial data under the alternative hypothesis, and calculate the rejection rate to show that the proposed statistic has sufficient power. The Monte Carlo results are shown in

Table 2, where the parameter value deviates from that of the null hypothesis. For example, the value of λ is set at 0.32 and 0.45 in the first and second rows under the alternative hypothesis, whereas it should be $\lambda = 0$ under the null hypothesis.

Table 2: Rejection Rates of the Null under the Alternative Hypothesis

Parameter Values in H_1					Rejection Rates	
					T=500	
λ	ψ	ρ	ω	σ	5%	1%
0.32	0.7	0.7	0.32	0.32	28%	13%
0.45	0.7	0.7	0.32	0.32	72%	43%
0	0.5	0.7	0.32	0.32	14%	7%
0	0.9	0.7	0.32	0.32	86%	72%
0	0.7	0.7	0.25	0.32	24%	7%
0	0.7	0.7	0.19	0.32	51%	28%

Note: Null hypothesis is $\lambda = 0, \psi = \rho, \omega = \sigma$. The number of iterations is 100.

5 Empirical analysis

Using the proposed statistical test, we first examine the volatility co-movement between stock markets to find a group of countries with common volatility factor and show that our method can be applied to more than two markets. We next investigate the effect of overall volatility level on the co-movement of exchange rates by comparing the financial crisis period and low volatility period.

The value of the test statistic depends upon the the order of the variables in the pair and hence the empirical result is sometime inconsistent when the order is changed, because the pre-orthogonalization of data illustrated in Appendices is asymmetric with respect to the order of the variables. The null distribution of the test statistic is unaffected with respect to the order of variable when the the volatility co-movement exits under the null hypothesis by its construction; then, the the probability of type I error is correct after the pre-orthogonalization. Under the alternative hypothesis, however, the pre-orthogonalization can contaminate the joint distribution of the volatilities, and hence undermine the power of the test statistic.

We believe that we can solve this asymmetry problem in future by estimating the correlation parameter of the measurement equation in (1) simultaneously, by means of improvement in the accuracy and speed of the computation .

5.1 Stock markets

First, we checks whether there exists a group of stock markets that shows volatility co-movement consistently in different times. The data is the adjusted-close prices downloaded from Yahoo finance for the stock market indexes listed below:

We divided daily data from January, 2011 to December, 2014 into two periods to check the volatility co-movement in different periods. We excluded observations whenever at least one market is closed. The test is performed for the 28 pairs, and we have 56 values in Tables 4 and 5, since the value of the test statistic depends upon the order of the variables asymmetrically, on account of the data pre-orthogonalization process in Appendices.

We see that U.K., Singapore, and Australia can share the same volatility factor consistently even in different periods, where the null hypothesis is accepted, even if calculated in the different order, for every possible pairs in the group. China and Japan have volatility factor independent

Table 3: Stock Market Indexes

Stock Market	**Symbol**	**Country/Region**
Dow Jones Industrial Average	DOW	U.S.
FTSE Index	FTSE	U.K.
DAX Index	DAX	Germany
Shanghai Composite Index	SSCI	China
Nikkei 225 Stock Average Index	NIKKEI	Japan
Hang Seng Index	HSI	Hong Kong
Straits Times Index	STI	Singapore
All Ordinaries Index	AORD	Australia

Table 4: Test Statistic for Volatility Co-movement of Stock Markets for 2011-2012

	y_2							
y_1	U.S.	U.K.	Germany	China	Japan	Hong Kong	Singapore	Australia
U.S.		9.1*	9.19*	21.48**	24.22**	5.00	10.44*	4.07
U.K.	5.14		1.39	20.33**	25.33**	3.66	3.74	1.07
Germany	5.96	8.7*		17.84**	30.19**	2.52	6.22	6.82
China	21.07**	7.63	15.72**		18.07**	4.7	7.92*	8.46*
Japan	22.15**	8.88*	29.5**	15.23**		2.34	10.23*	5.22
Hong Kong	4.9	2.73	5.88	4.14	40.52**		2.87	4.69
Singapore	7.69	2.99	7.03	3.17	34.31**	18.12**		1.98
Australia	2.46	5.56	12.88**	5.19	69.35**	15.88**	2.68	

Note: * denotes significance at five percent, ** denotes significance at one percent.

mutually and of the other countries or regions in 2011 and 2012, where the null hypothesis is rejected with, at least, one of the two test statistic values. In 2013 and 2014, the number of groups that possibly share the same volatility factor is increased to three, namely

Table 5: Test Statistic for Volatility Co-movement of Stock Markets for 2013-2014

	y_2							
y_1	U.S.	U.K.	Germany	China	Japan	Hong Kong	Singapore	Australia
U.S.		0.39	8.05*	11.83**	15.43**	9.24*	4.97	5.21
U.K.	15.79**		9.9*	2.28	4.76	4.17	1.92	2.68
Germany	4.34	7.52		10.64*	21.47**	3.31	7.83*	10.22*
China	11.53**	2.14	10.72*		3.39	12.84**	3.79	6.11
Japan	15.05**	5.68	25.2**	2.58		9.92*	9.78*	9.88*
Hong Kong	7.88*	1.54	7.13	7.47	12.39**		19.45**	2.58
Singapore	11.42**	1.91	23.23**	7.08	7.13	8.32*		1.63
Australia	6.11	0.8	8.03*	2.31	5.71	3.99	4.13	

Note: * denotes significance at five percent, ** denotes significance at one percent.

Group 1 : U.K., China, Japan

Group 2 : U.K., Hong Kong, Australia

Group 3 : U.K., Singapore, Australia

Then U.K., Singapore and Australia share the same volatility factor consistently in the two period, probably because of their close economic ties.

We cannot suggest, at this stage, why the number of groups with possibly the same volatility factors increased. It is suspected that a determinant is the overall level of volatility and we will consider this hypothesis in the next subsection using the exchange rate data.

.

5.2 Exchange rate markets

We here investigate the volatility processes of the foreign exchange rates in the global financial crisis and the low volatility period.

First, we define two time periods representing the financial crisis and the low volatility period using the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) as an indicator of high volatility. Figures 4 and 5 show that volatility deviated drastically from the historical trend in the financial crisis. We choose Period 1: Oct/1/2008 - Oct/31/2008 as the global financial crisis, and Period 2: Oct/1/2012 - Oct/31/2012 as low volatility period.

Figure 4: VIX during the Global Financial Crisis (2008-2009)

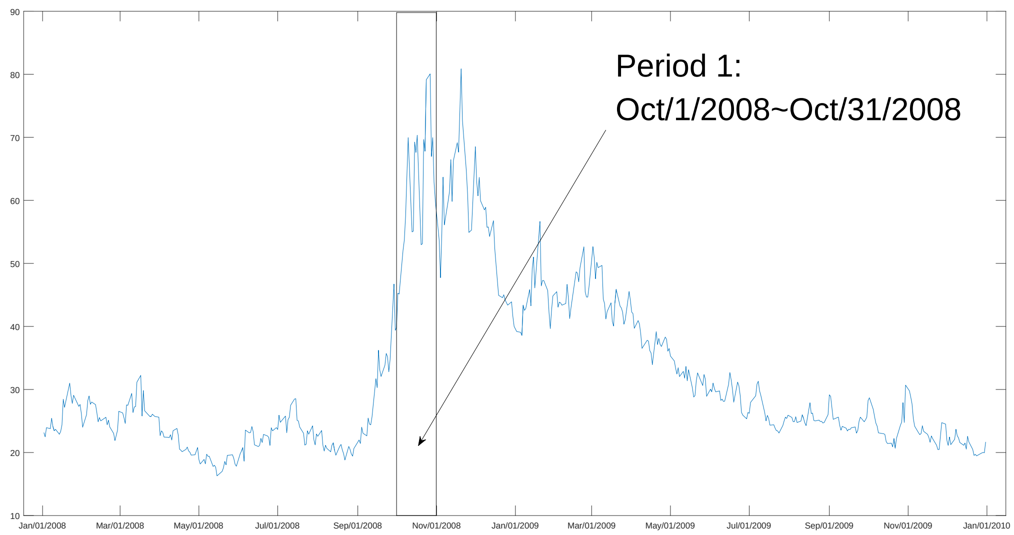
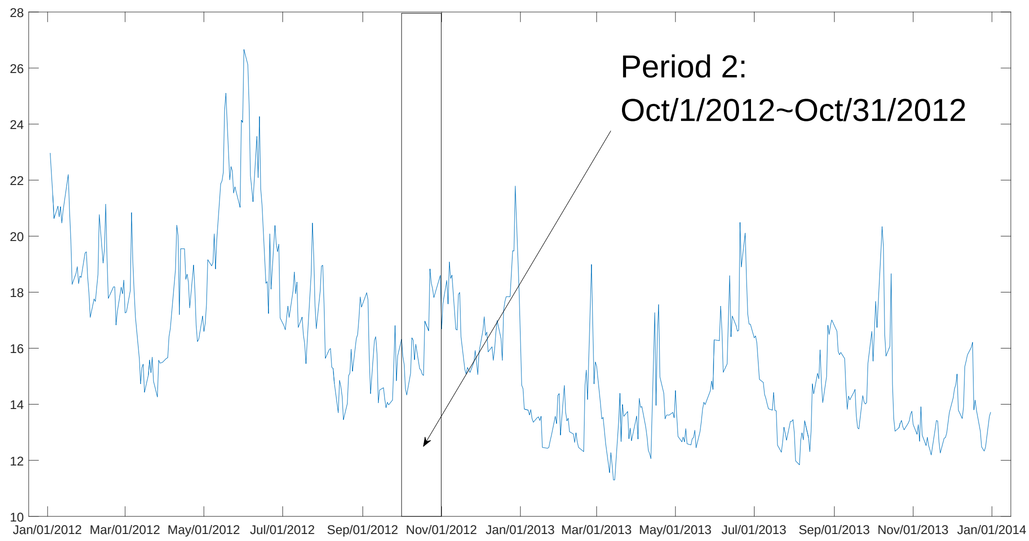


Figure 5: VIX during the Low Volatility Period (2012-2013)



Second, we analyze 6 major currency pairs, namely Euro (EUR), the United States Dollar (USD), Japanese Yen (JPY), British Pound (GBP), Australian Dollar (AUD), Swiss Franc (CHF), Canadian Dollar (CAD), using roughly 500 hourly observations for a month, and the results of the proposed test of volatility co-movement are shown in Tables 6 and 7.

Table 6: Test Statistic for Volatility Co-movement of Exchange Rates in High Volatility Period

	y_2					
y_1	EUR/USD	USD/JPY	GBP/USD	AUD/USD	USD/CHF	USD/CAD
EUR/USD		23.73**	12.06**	34.09**	10.43*	17.89**
USD/JPY	35.45**		17.28**	15.62**	28.23**	54.5**
GBP/USD	14.49**	14.34**		53.78**	18.74**	22.03**
AUD/USD	37.32**	33.16**	30.01**		25.47**	28.3**
USD/CHF	24.47**	33.96**	43.9**	39.49**		23.63**
USD/CAD	18.31**	37.44**	15.14**	19.02**	14.64**	

Note: * denotes significance at five percent, ** denotes significance at one percent. The exchange rate was downloaded from FXDD's historical database.

Table 7: Test Statistic for Volatility Co-movement of Exchange Rates in Low Volatility Period

	y_2					
y_1	EUR/USD	USD/JPY	GBP/USD	AUD/USD	USD/CHF	USD/CAD
EUR/USD		11.27*	5.17	7.56	6.61	25.69**
USD/JPY	15.88**		28.67**	12.97**	16.73**	14.86**
GBP/USD	2.15	23.74**		20.35**	5.39	8.38*
AUD/USD	28.07**	18.32**	26.51**		14.97**	22.86**
USD/CHF	3.22	9.53*	4.64	6.9		17.66**
USD/CAD	5.5	6.47	4.83	5.82	4.41	

Note: * denote significance at five percent, ** denotes significance at one percent. The exchange rates data was downloaded from FXDD's historical database.

During the financial crisis, when volatility is large, the null hypothesis of volatility co-movement was rejected in every case in Table 6. On the other hand, several currency pairs are suggested to share the same volatility factor during the low volatility period; the accepted rate of the null hypothesis is 43.3%, namely 13 pairs of 30 pairs, in Table 7. Then we can suggest that the volatility co-movement tends to be found during the low volatility period. This result is interesting and contrasting to the often-cited finding in the financial contagion literature that financial returns have co-movement in the level during the financial crisis, which is discussed critically by Forbes and Rigobon (2002) . It is suspected that, when the overall volatility level is low, the idiosyncratic volatility factor can be small, or often negligible, in comparison with the common volatility factor, whereas, when the overall volatility level is high in the financial crisis, the idiosyncratic volatility factor dominates the common volatility factor.

6 Conclusion

In this paper, we have proposed a Lagrange multiplier test statistic for the null hypothesis that the volatility processes of a bivariate series are perfectly correlated in the framework of the multivariate stochastic volatility model. The considered model is the simplest case of a multiple stochastic volatility model. The extension to a multiple factor model is left for further research, as it is a challenging problem computationally and theoretically. In order to improve the efficiency of the numerical calculation we are planning to use the particle filter method proposed by Kitagawa (1996) and the fast Gauss transform method proposed by Greengard and Strain (1991).

In the empirical analysis of stock markets, we found that the United Kingdom, Singapore and Australia share a common time-varying volatility factor consistently. It is also suspected that the common volatility factor in the global currency market was dominated by the idiosyncratic volatility factors during high volatility periods. A clear-cut conclusion cannot be obtained because of the asymmetry of the test statistic with respect to the order of the variables, which is left to the further research.

Appendices

A Derivation of likelihood and score functions

A.1 Likelihood

In order to express the transition equation (2) in matrix form, we express the log volatilities and disturbance terms used in (1) and (2) in vector form, as follows:

$$\mathbf{h}_1 = (h_{11}, \ldots, h_{1t})', \ \mathbf{h}_2 = (h_{21}, \ldots, h_{2t})',$$
$$\mathbf{u}_1 = (u_{11}, \ldots, u_{1T})', \ \mathbf{u}_2 = (u_{21}, \ldots, u_{2T})',$$
$$\mathbf{e}_1 = (e_{11}, \ldots, e_{1T})', \ \mathbf{e}_2 = (e_{21}, \ldots, e_{2T})'.$$

Then the transition equation (2) is

$$\mathbf{h}_1 = \mathbf{V}_\rho^{1/2}(\sigma \mathbf{u}_1), \ \mathbf{h}_2 = \mathbf{V}_\psi^{1/2}(\omega \mathbf{u}_1 + \sqrt{\lambda}\mathbf{u}_2) = \mathbf{V}_\psi^{1/2}(\mathbf{V}_\rho^{-1/2}\mathbf{h}_1 \omega/\sigma + \sqrt{\lambda}\mathbf{u}_2), \tag{14}$$

where $\mathbf{V}_\rho$ and $\mathbf{V}_\psi$ are the covariance matrices of the autoregressive processes of order one, $\mathbf{h}_1$ and $\mathbf{h}_2$, respectively, and $\mathbf{V}_\rho^{1/2}$ and $\mathbf{V}_\psi^{1/2}$ are defined by their Cholesky decomposition as follows:

$$\mathbf{V}_\rho = (\mathbf{V}_\rho^{1/2})(\mathbf{V}_\rho^{1/2})', \ \mathbf{V}_\psi = (\mathbf{V}_\psi^{1/2})(\mathbf{V}_\psi^{1/2})',$$

where

$$\mathbf{V}_\psi^{1/2} = \begin{pmatrix} 1/\sqrt{1-\psi^2} & 0 & \ldots & 0 & 0 \\ \psi/\sqrt{1-\psi^2} & 1 & \ldots & 0 & 0 \\ \psi^2/\sqrt{1-\psi^2} & \psi & \ldots & 0 & 0 \\ \vdots & & & & \\ \psi^{T-1}/\sqrt{1-\psi^2} & \psi^{T-2} & \ldots & \psi & 1 \end{pmatrix}, \ \mathbf{V}_\rho^{1/2} = \begin{pmatrix} 1/\sqrt{1-\rho^2} & 0 & \ldots & 0 & 0 \\ \rho/\sqrt{1-\rho^2} & 1 & \ldots & 0 & 0 \\ \rho^2/\sqrt{1-\rho^2} & \rho & \ldots & 0 & 0 \\ \vdots & & & & \\ \rho^{T-1}/\sqrt{1-\rho^2} & \rho^{T-2} & \ldots & \rho & 1 \end{pmatrix}, \tag{15}$$

Their inverses are decomposed as $\mathbf{V}_{\psi}{}^{-1} = (\mathbf{V}_{\psi}^{-\frac{1}{2}})'\mathbf{V}_{\psi}^{-\frac{1}{2}},\ \mathbf{V}_{\rho}{}^{-1} = (\mathbf{V}_{\rho}^{-\frac{1}{2}})'\mathbf{V}_{\rho}^{-\frac{1}{2}}$, where:

$$\mathbf{V}_{\psi}^{-\frac{1}{2}} = \begin{pmatrix} \sqrt{1-\psi^2} & 0 & \dots & 0 & 0 \\ -\psi & 1 & \dots & 0 & 0 \\ 0 & -\psi & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & -\psi & 1 \end{pmatrix}, \mathbf{V}_{\rho}^{-\frac{1}{2}} = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & \dots & 0 & 0 \\ -\rho & 1 & \dots & 0 & 0 \\ 0 & -\rho & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & -\rho & 1 \end{pmatrix}. \tag{16}$$

Then the density functions of the transition and measurement equations of the model is

$$f(\mathbf{h}_1) = \frac{1}{(2\pi)^{\frac{T}{2}}\sigma^T}\left|\mathbf{V}_{\rho}^{-1/2}\right| \exp\left(-\frac{1}{2}\sigma^{-2}\mathbf{h}_1'\mathbf{V}_{\rho}^{-1}\mathbf{h}_1\right), \tag{17}$$

$$f(\mathbf{h}_2|\mathbf{h}_1) = \frac{1}{(2\pi)^{\frac{T}{2}}(\sqrt{\lambda})^T}\left|\mathbf{V}_{\psi}^{-1/2}\right| \exp\left(-\frac{1}{2}\mathbf{u}_2'\mathbf{u}_2\right), \tag{18}$$

$$f(\mathbf{y}_1|\mathbf{h}_1) = \frac{1}{(2\pi)^{\frac{T}{2}}}\left|\mathbf{H}_1^{-1/2}\right| \exp\left(-\frac{1}{2}\mathbf{y}_1'\mathbf{H}_1^{-1}\mathbf{y}_1\right), \tag{19}$$

$$f(\mathbf{y}_2|\mathbf{h}_2) = \frac{1}{(2\pi)^{\frac{T}{2}}}\left|\mathbf{H}_2^{-1/2}\right| \exp\left(-\frac{1}{2}\mathbf{y}_2'\mathbf{H}_2^{-1}\mathbf{y}_2\right), \tag{20}$$

where

$$\mathbf{u}_2 = \left(\mathbf{V}_{\psi}^{-\frac{1}{2}}\mathbf{h}_2 - \mathbf{V}_{\rho}^{-\frac{1}{2}}\mathbf{h}_1\frac{\omega}{\sigma}\right)/\sqrt{\lambda}, \tag{21}$$

$$\mathbf{H}_1 = diag(\exp(h_{11}),\dots,\exp(h_{1T})),\ \mathbf{H}_2 = diag(\exp(h_{21}),\dots,\exp(h_{2T})). \tag{22}$$

Then, we can rewrite the likelihood function as

$$\textbf{(Likelihood):}\ f(\mathbf{y}_1,\mathbf{y}_2) = \int\int f(\mathbf{y}_2|\mathbf{h}_2)f(\mathbf{y}_1|\mathbf{h}_1)f(\mathbf{h}_2|\mathbf{h}_1)f(\mathbf{h}_1)d\mathbf{h}_2 d\mathbf{h}_1, \tag{23}$$

where

$$f(\mathbf{u}_2|\mathbf{h}_1) = \frac{1}{(2\pi)^{\frac{T}{2}}}\exp\left(-\frac{1}{2}\mathbf{u}_2'\mathbf{u}_2\right) \tag{24}$$

in terms of $\mathbf{u}_2$, instead of $\mathbf{h}_2$, by the variable transformation (21).

A.2 Score function with respect to λ

We obtain the score function with respect to λ as

$$\frac{\partial f(\mathbf{y})}{\partial\lambda} = \int\int \frac{\partial f(\mathbf{y}_2|\mathbf{u}_2,\mathbf{h}_1)}{\partial\lambda} f(\mathbf{y}_1|\mathbf{h}_1)f(\mathbf{u}_2|\mathbf{h}_1)f(\mathbf{h}_1)d\mathbf{u}_2 d\mathbf{h}_1, \tag{25}$$

because the variance parameter λ appears only in

$$f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2) = \frac{1}{(2\pi)^{\frac{T}{2}}}\left|\mathbf{H}_2^{-1/2}\right|\exp\left(-\frac{1}{2}\mathbf{y}_2'\mathbf{H}_2^{-1}\mathbf{y}_2\right)$$

through $\mathbf{h}_2$ in $\mathbf{H}_2 = diag(\exp(\mathbf{h}_2))$, since we have

$$\mathbf{h}_2 = \mathbf{V}_\psi^{-\frac{1}{2}}\left(\sqrt{\lambda}\mathbf{u}_2 + \mathbf{V}_\rho^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right) \tag{26}$$

from (14).

Then we obtain the derivative of $f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2,)$ with respect to λ as follows. First, noting (26), we define

$$f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2) = \mathbf{KF}, \tag{27}$$

where

$$\begin{aligned}
\mathbf{K} &= |\mathbf{H}_2|^{-1/2} = \exp\left(-\frac{1}{2}\mathbf{1}_{1\times T}\mathbf{h}_2\right) \\
&= \exp\left(-\frac{1}{2}\mathbf{1}_{1\times T}\mathbf{V}_\psi^{1/2}\left(\sqrt{\lambda}\mathbf{u}_2 + \mathbf{V}_\psi^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right)\right), \\
\mathbf{F} &= \exp\left(-\frac{1}{2}\mathbf{y}_2'\mathbf{H}_2^{-1}\mathbf{y}_2\right) \\
&= \exp\left(-\frac{1}{2}\big(\exp(-\mathbf{h}_2)\big)'(\mathbf{y}_2\circ\mathbf{y}_2)\right)
\end{aligned} \tag{28}$$

and, for notational convenience, we define

$$\exp(-\mathbf{h}_2) = (\exp(-h_{21}),\ldots,\exp(-h_{2T}))',\ \mathbf{y}_2\circ\mathbf{y}_2 = \left(y_{21}^2, y_{22}^2,\ldots,y_{2T}^2\right)'$$

and $\mathbf{h}_2$ denotes a function of $\mathbf{u}_2$ as the abbreviation of equation (26).

Then, defining

$$\mathbf{M}_1 = \frac{\partial\mathbf{K}}{\partial\lambda}\frac{1}{\sqrt{\lambda}},\ \mathbf{M}_2 = \frac{\partial\mathbf{F}}{\partial\lambda}\frac{1}{\sqrt{\lambda}}, \tag{29}$$

we have

$$\begin{aligned}
\mathbf{B} &= \lim_{\lambda\to 0}\frac{\partial f(\mathbf{y})}{\partial\lambda} \\
&= \lim_{\lambda\to 0}\int(other\ terms)\left(\mathbf{F}\frac{\partial\mathbf{K}}{\partial\lambda} + \mathbf{K}\frac{\partial\mathbf{F}}{\partial\lambda}\right)d\mathbf{u}_2 d\mathbf{h}_1 \\
&= \lim_{\lambda\to 0}\frac{\sqrt{\lambda}\int(other\ terms)(\mathbf{F}\ \mathbf{M}_1 + \mathbf{K}\ \mathbf{M}_2)d\mathbf{u}_2 d\mathbf{h}_1}{\lambda}
\end{aligned} \tag{30}$$

from (25). We then have that

$$\begin{aligned}\frac{\partial \mathbf{K}}{\partial \lambda} &= -\frac{1}{2}\mathbf{K}\ \frac{1}{2\sqrt{\lambda}}\mathbf{1}_{1\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2 = -\frac{1}{4\sqrt{\lambda}}\mathbf{K}\ \mathbf{1}_{1\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2,\\ \frac{\partial \mathbf{F}}{\partial \lambda} &= -\frac{1}{2}\mathbf{F}\ \frac{\partial}{\partial \lambda}\exp(-\mathbf{h}_2^{'})(\mathbf{y}_2\circ\mathbf{y}_2) = \frac{1}{4\sqrt{\lambda}}\mathbf{F}\ \mathbf{G},\\ \mathbf{G} &= \mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{H}_2^{-1}(\mathbf{y}_2\circ\mathbf{y}_2),\end{aligned} \tag{31}$$

since

$$\frac{\partial \mathbf{h}_2}{\partial \lambda} = \frac{1}{2\sqrt{\lambda}}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2, \frac{\partial \exp(-\mathbf{h}_2)}{\partial \lambda} = -\frac{1}{2\sqrt{\lambda}}\mathbf{H}_2^{-1}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2.$$

Note that the denominators of the derivatives (31) and (29) contain λ, which converges to zero, and hence is intractable by conventional method. We will use the ingenious method proposed by Chesher (1984) to solve this singularity. First, applying L'Hopital's rule with respect to λ , we obtain

$$\mathbf{B} \ = \tfrac{1}{2}\mathbf{B} + \lim_{\lambda\to 0}\sqrt{\lambda}\tfrac{\partial}{\partial\lambda}\int (other\, terms)(\mathbf{F}\ \mathbf{M}_1 + \mathbf{K}\ \mathbf{M}_2)d\mathbf{u}_2 d\mathbf{h}_1. \tag{32}$$

Comparing the both sides of equation (32), we have

$$\begin{aligned}\mathbf{B} &= 2\lim_{\lambda\to 0}\sqrt{\lambda}\int (other\, terms)\left(\frac{\partial \mathbf{F}}{\partial\lambda}\ \mathbf{M}_1 + \frac{\partial \mathbf{K}}{\partial\lambda}\ \mathbf{M}_2 + \mathbf{F}\frac{\partial \mathbf{M}_1}{\partial\lambda} + \mathbf{K}\frac{\partial \mathbf{M}_2}{\partial\lambda}\right)d\mathbf{u}_2 d\mathbf{h}_1\\ &= 2\lim_{\lambda\to 0}\sqrt{\lambda}\int (other\, terms)\left(2\mathbf{M}_1\ \mathbf{M}_2 + \mathbf{F}\frac{\partial \mathbf{M}_1}{\partial\lambda} + \mathbf{K}\frac{\partial \mathbf{M}_2}{\partial\lambda}\right)d\mathbf{u}_2 d\mathbf{h}_1.\end{aligned} \tag{33}$$

Defining $\mathbf{Y}_2 = diag(\mathbf{y}_2\circ\mathbf{y}_2)$, the terms in the integrand are

$$\begin{aligned}\mathbf{M}_1\mathbf{M}_2 &= -\frac{1}{4}\mathbf{K}\ \mathbf{1}_{1\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2 \times \frac{1}{4}\mathbf{F}\ \mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_2),\\ \frac{\partial \mathbf{M}_1}{\partial\lambda} &= -\frac{1}{4}\frac{\partial \mathbf{K}}{\partial\lambda}\mathbf{1}_{1\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2 = \frac{1}{16\sqrt{\lambda}}\mathbf{K}\ tr(\mathbf{1}_{T\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}),\\ \frac{\partial \mathbf{M}_2}{\partial\lambda} &= \frac{1}{4}\frac{\partial \mathbf{F}}{\partial\lambda}\mathbf{G} + \frac{1}{4}\mathbf{F}\frac{\partial \mathbf{G}}{\partial\lambda} = \frac{1}{16\sqrt{\lambda}}\mathbf{F}\mathbf{G}^2 + \frac{1}{4}\mathbf{F}\frac{\partial \mathbf{G}}{\partial\lambda},\end{aligned}$$

$$\frac{\partial \mathbf{G}}{\partial\lambda} = \mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\frac{\partial \exp(-\mathbf{h}_2)}{\partial\lambda} = -\frac{1}{2\sqrt{\lambda}}tr\left(\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\mathbf{H}_2^{-1}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\right),$$

$$\begin{aligned}\mathbf{G}^2 &= \mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_2)\exp(-\mathbf{h}_2^{'})\mathbf{Y}_2\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\\ &= tr\left(\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_2)\exp(-\mathbf{h}_2^{'})\mathbf{Y}_2\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\right).\end{aligned}$$

Then, we have

$$
\begin{aligned}
\mathbf{B} = \frac{1}{8}\lim_{\lambda\to 0}\int f(\mathbf{y}_1|\mathbf{h}_1)\frac{1}{(2\pi)^{\frac{T}{2}}}\mathbf{KF} \\
\Bigg[-2tr\left(\mathbf{1}_{1\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_2)\right) \\
+ tr\left(\mathbf{1}_{T\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\mathbf{V}_{\psi}^{1/2^{'}}\right) \\
+ tr\left(\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_2)\exp(-\mathbf{h}_2^{'})\mathbf{Y}_2\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\right) \\
- 2tr\left(\mathbf{V}_{\psi}^{1/2^{'}}\mathbf{Y}_2\mathbf{H}_2^{-1}\mathbf{V}_{\psi}^{1/2}\mathbf{u}_2\mathbf{u}_2^{'}\right)\Bigg] f(\mathbf{u}_2|\mathbf{h}_1)f(\mathbf{h}_1)d\mathbf{u}_2 d\mathbf{h}_1.
\end{aligned}
\tag{34}
$$

We can perform the integration with respect to $\mathbf{u}_2$ in (34) analytically. As $\mathbf{u}_2|\mathbf{h}_1$ follows the T-dimensitonal standard normal distribution, we have that

$$
\int \mathbf{u}_2\mathbf{u}_2^{'}f(\mathbf{u}_2|\mathbf{h}_1)d\mathbf{u}_2 = \mathbf{I}_T. \tag{35}
$$

Under the null hypothesis $\mathbf{h}_2 = \mathbf{h}_1$ and $\psi = \rho$, equation (34) is

$$
\begin{aligned}
\mathbf{B} &= \frac{1}{8}\int f(\mathbf{y}_1|\mathbf{h}_1)\frac{1}{(2\pi)^{\frac{T}{2}}}\mathbf{KF}\Big[-2tr\left(\mathbf{1}_{1\times T}\mathbf{V}_{\rho}^{1/2}\mathbf{V}_{\rho}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_1)\right) \\
&+ tr\left(\mathbf{1}_{T\times T}\mathbf{V}_{\rho}^{1/2}\mathbf{V}_{\rho}^{1/2^{'}}\right) \\
&+ tr\left(\mathbf{V}_{\rho}^{1/2^{'}}\mathbf{Y}_2\exp(-\mathbf{h}_1)\exp(-\mathbf{h}_1^{'})\mathbf{Y}_2\mathbf{V}_{\rho}^{1/2}\right) \\
&- 2tr\left(\mathbf{V}_{\rho}^{1/2^{'}}\mathbf{Y}_2\mathbf{H}_1^{-1}\mathbf{V}_{\rho}^{1/2}\right)\Big] f(\mathbf{h}_1)d\mathbf{h}_1.
\end{aligned}
\tag{36}
$$

Noting that $\mathbf{V}_{\rho} = \mathbf{V}_{\rho}^{1/2}\mathbf{V}_{\rho}^{1/2^{'}}$ and applying the cyclic property of the trace operator to simplify the equation (36), we have

$$
\mathbf{B} = \left.\frac{\partial f(\mathbf{y})}{\partial\lambda}\right|_{\mathbf{H}_0} = \int tr\mathbf{J}f(\mathbf{y},\mathbf{h}_1)d\mathbf{h}_1, \tag{37}
$$

where

$$
\begin{aligned}
\mathbf{J} = \frac{1}{8}\Big(&- 2\left(\mathbf{1}_{1\times T}\mathbf{V}_{\rho}\mathbf{Y}_2\exp(-\mathbf{h}_1)\right) + \mathbf{1}_{T\times T}\mathbf{V}_{\rho} \\
&+ \mathbf{V}_{\rho}\mathbf{Y}_2\exp(-\mathbf{h}_1)\exp(-\mathbf{h}_1^{'})\mathbf{Y}_2 - 2\mathbf{V}_{\rho}\mathbf{Y}_2\mathbf{H}_1^{-1}\Big).
\end{aligned}
\tag{38}
$$

Since we have

$$
\left.\frac{\partial \log f(\mathbf{y})}{\partial\lambda}\right|_{\mathbf{H}_0} = \lim_{\lambda\to 0}\frac{1}{f(\mathbf{y})}\frac{\partial f(\mathbf{y})}{\partial\lambda} = \int tr\mathbf{J}\frac{1}{f(\mathbf{y})}f(\mathbf{h}_1,\mathbf{y})d\mathbf{h}_1 = trE_{\mathbf{h}_1|\mathbf{y}}(\mathbf{J}) \tag{39}
$$

from $f(\mathbf{h}_1|\mathbf{y}) = f(\mathbf{h}_1,\mathbf{y})/f(\mathbf{y})$, we have only to evaluate $E_{\mathbf{h}_1|\mathbf{y}}\left[\exp(-\mathbf{h}_1)\right]$ and $E_{\mathbf{h}_1|\mathbf{y}}\left[\exp(-\mathbf{h}_1)\exp(-\mathbf{h}_1)'\right]$ to obtain the score function. These expected values have no analytic expressions so that they should be evaluated numerically.

A.3 Score function with respect to ψ

In the log-likelihood function, ψ appears only in $f(\mathbf{y}_1|\mathbf{h}_1,\mathbf{u}_2) = \mathbf{KF}$ in (27). The partial derivative of the likelihood with respect to ψ is

$$\begin{aligned}\frac{\partial f(\mathbf{y})}{\partial\psi} &= \int\left(\frac{\partial\mathbf{K}}{\partial\psi}\mathbf{K}^{-1}+\frac{\partial\mathbf{F}}{\partial\psi}\mathbf{F}^{-1}\right)f(\mathbf{y},\mathbf{u}_2,\mathbf{h}_1)d\mathbf{u}_2d\mathbf{h}_1\\ &= \int\left(\frac{\partial\mathbf{K}}{\partial\psi}\mathbf{K}^{-1}+\frac{\partial\mathbf{F}}{\partial\psi}\mathbf{F}^{-1}\right)f(\mathbf{y},\mathbf{h}_1)d\mathbf{h}_1,\end{aligned} \tag{40}$$

since, as will be seen later, $\mathbf{u}_2$ can be integrated out in $\left(\frac{\partial\mathbf{K}}{\partial\psi}\mathbf{K}^{-1}+\frac{\partial\mathbf{F}}{\partial\psi}\mathbf{F}^{-1}\right)$. Then we have

$$\frac{\partial\log f(\mathbf{y})}{\partial\psi}\Big|_{\mathbf{H}_0} = \int\left(\frac{\partial\mathbf{K}}{\partial\psi}\mathbf{K}^{-1}+\frac{\partial\mathbf{F}}{\partial\psi}\mathbf{F}^{-1}\right)f(\mathbf{h}_1|\mathbf{y})d\mathbf{h}_1 = E_{\mathbf{h}_1|\mathbf{y}}\left(\frac{\partial\mathbf{K}}{\partial\psi}\mathbf{K}^{-1}+\frac{\partial\mathbf{F}}{\partial\psi}\mathbf{F}^{-1}\right), \tag{41}$$

noting that

$$f(\mathbf{h}_1|\mathbf{y}) = f(\mathbf{h}_1,\mathbf{y})/f(\mathbf{y}).$$

First, using (28) and the formula

$$\frac{\partial\mathbf{V}_\psi^{1/2}}{\partial\psi} = -\mathbf{V}_\psi^{1/2}\mathbf{Z}_\psi\mathbf{V}_\psi^{1/2}, \tag{42}$$

where

$$\mathbf{Z}_\psi = \frac{\partial\mathbf{V}_\psi^{-1/2}}{\partial\psi},$$

we have that

$$\begin{aligned}\frac{\partial\mathbf{K}}{\partial\psi} &= -\frac{1}{2}\mathbf{K}\mathbf{1}_{1\times T}\frac{\partial\mathbf{V}_\psi^{1/2}}{\partial\psi}\left(\sqrt{\lambda}\mathbf{u}_2+\mathbf{V}_\rho^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right)\\ &= \frac{1}{2}\mathbf{K}\mathbf{1}_{1\times T}\mathbf{V}_\psi^{1/2}\mathbf{Z}_\psi\mathbf{V}_\psi^{1/2}\left(\sqrt{\lambda}\mathbf{u}_2+\mathbf{V}_\rho^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right),\\ \frac{\partial\mathbf{F}}{\partial\psi} &= -\frac{1}{2}\mathbf{F}\ \frac{\partial}{\partial\psi}\left[(\mathbf{y}_2\circ\mathbf{y}_2)'\exp(-\mathbf{h}_2)\right]\\ &= -\frac{1}{2}\mathbf{F}\ (\mathbf{y}_2\circ\mathbf{y}_2)'\mathbf{H}_2^{-1}\mathbf{V}_\psi^{1/2}\mathbf{Z}_\psi\mathbf{V}_\psi^{1/2}\left(\sqrt{\lambda}\mathbf{u}_2+\mathbf{V}_\rho^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right),\end{aligned} \tag{43}$$

as we have

$$\mathbf{h}_2 = \mathbf{V}_\psi^{1/2}\left(\sqrt{\lambda}\mathbf{u}_2+\mathbf{V}_\rho^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right), \tag{44}$$

and hence

$$\frac{\partial}{\partial\psi}\mathbf{h}_2 = \frac{\partial\mathbf{V}_\psi^{1/2}}{\partial\psi}\left(\sqrt{\lambda}\mathbf{u}_2+\mathbf{V}_\rho^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right). \tag{45}$$

Evaluating these terms under the null hypothesis $\lambda = 0$ and $\sigma = \omega$, we have

$$\frac{\partial \mathbf{K}}{\partial \psi}|_{\mathbf{H}_0} = \tfrac{1}{2}\mathbf{K}\mathbf{1}_{1\times T}\mathbf{V}_{\psi}^{1/2}\mathbf{Z}_{\psi}\mathbf{h}_1, \tag{46}$$

$$\frac{\partial \mathbf{F}}{\partial \psi}|_{\mathbf{H}_0} = -\tfrac{1}{2}\mathbf{F}\ tr\left[\mathbf{Y}_2\mathbf{V}_{\psi}^{1/2}\mathbf{Z}_{\psi}\mathbf{h}_1\exp(-\mathbf{h}_1^{'})\right], \tag{47}$$

using the identity

$$(\mathbf{y}_2 \circ \mathbf{y}_2)'\mathbf{H}_1^{-1} = \exp(-\mathbf{h}_1^{'})\mathbf{Y}_2.$$

Then, from (41), we have

$$\left.\frac{\partial \log f(\mathbf{y})}{\partial \psi}\right|_{\mathbf{H}_0} = \frac{1}{2}\mathbf{1}_{1\times T}\mathbf{V}_{\rho}^{1/2}\mathbf{Z}_{\rho}E_{\mathbf{h}_1|\mathbf{y}}[\mathbf{h}_1] - \frac{1}{2}tr\left[\mathbf{Y}_2\mathbf{V}_{\rho}^{1/2}\mathbf{Z}_{\rho}E_{\mathbf{h}_1|\mathbf{y}}[\mathbf{h}_1\exp(-\mathbf{h}_1^{'})]\right]. \tag{48}$$

Note that the matrix $\mathbf{Y}_2\mathbf{V}_{\rho}^{1/2}\mathbf{Z}_{\rho}$ is lower triangular, and we have only to calculate the upper triangular part of the matrix $E_{\mathbf{h}_1|\mathbf{y}}[\mathbf{h}_1\exp(-\mathbf{h}_1^{'})]$ in evaluating the score function (48).

A.4 Score function with respect to ω

First, note that, in the log-likelihood function, ω appears only in $f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2) = \mathbf{KF}$, through

$$\mathbf{h}_2 = \mathbf{V}_{\psi}^{1/2}\left(\sqrt{\lambda}\mathbf{u}_2 + \mathbf{V}_{\rho}^{-\frac{1}{2}}\frac{\omega}{\sigma}\mathbf{h}_1\right), \tag{49}$$

as shown in (27). Then, we have the formula

$$\frac{\partial \log f(\mathbf{y})}{\partial \omega} = E_{\mathbf{u}_2,\mathbf{h}_1|\mathbf{y}}\left(\frac{\partial \mathbf{K}}{\partial \omega}\mathbf{K}^{-1} + \frac{\partial \mathbf{F}}{\partial \omega}\mathbf{F}^{-1}\right), \tag{50}$$

using

$$\frac{\partial f(\mathbf{y})}{\partial \omega} = \int (\mathbf{y}_1|\mathbf{h}_1)\frac{\partial f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2)}{\partial \omega}f(\mathbf{h}_1)f(\mathbf{u}_2|\mathbf{h}_1)d\mathbf{u}_2\mathbf{h}_1$$
$$= \int \left(\frac{\partial \mathbf{K}}{\partial \omega}\mathbf{K}^{-1} + \frac{\partial \mathbf{F}}{\partial \omega}\mathbf{F}^{-1}\right)f(\mathbf{y},\mathbf{h}_1)d\mathbf{h}_1, \tag{51}$$

as we have

$$\frac{\partial f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2)}{\partial \omega} = \frac{\partial \mathbf{K}}{\partial \omega}\mathbf{F} + \mathbf{K}\frac{\partial \mathbf{F}}{\partial \omega} = \left(\frac{\partial \mathbf{K}}{\partial \omega}\mathbf{K}^{-1} + \frac{\partial \mathbf{F}}{\partial \omega}\mathbf{F}^{-1}\right)f(\mathbf{y}_2|\mathbf{h}_1,\mathbf{u}_2). \tag{52}$$

From (28) their partial derivatives of $\mathbf{K}$ and $\mathbf{F}$ are

$$\begin{aligned}\left.\frac{\partial \mathbf{K}}{\partial \omega}\right|_{\mathbf{H}_0} &= -\frac{1}{2}\mathbf{K}\mathbf{1}_{1\times T}\frac{1}{\sigma}\mathbf{h}_1,\\ \left.\frac{\partial \mathbf{F}}{\partial \omega}\right|_{\mathbf{H}_0} &= -\frac{1}{2}\mathbf{F}\ \frac{\partial}{\partial \omega}\left[(\mathbf{y}_2\circ\mathbf{y}_2)'\exp(-\mathbf{h}_2)\right]\Big|_{\mathbf{H}_0} = \frac{1}{2}\mathbf{F}\ tr\left[(\mathbf{y}_2\circ\mathbf{y}_2)'\mathbf{H}_2\frac{1}{\sigma}\mathbf{h}_1\right]\Big|_{\mathbf{H}_0},\end{aligned} \tag{53}$$

noting that, under the null hypothesis, we have $\rho = \psi, \mathbf{h}_1 = \mathbf{h}_2$, and $\mathbf{V}_\rho = \mathbf{V}_\psi$ and

$$\frac{\partial}{\partial \omega} \exp(-\mathbf{h}_2) = -\mathbf{H}_2^{-1} \omega / \sigma.$$

Then we have

$$\frac{\partial \log f(\mathbf{y})}{\partial \omega}|_{\mathbf{H}_0} = -\frac{1}{2\sigma} \mathbf{1}_{1\times T} E_{\mathbf{h}_1|\mathbf{y}} [\mathbf{h}_1] + \frac{1}{2\sigma} tr \left[(\mathbf{y}_2 \circ \mathbf{y}_2)' E_{\mathbf{h}_1|\mathbf{y}} [\exp(-\mathbf{h}_1) \circ \mathbf{h}_1] \right]. \quad (54)$$

A.5 Score function with respect to ρ

In the log-likelihood function, ρ appears only in $f(y_1|\mathbf{h}_1, \mathbf{u}_2) = \mathbf{KF}$ and $f(\mathbf{h}_1)$ in (17) and (27). Then we have the derivative using the formula

$$\frac{\partial \log f(\mathbf{y})}{\partial \rho}\Big|_{\mathbf{H}_0} = E_{\mathbf{h}_1|\mathbf{y}} \left(\frac{\partial \mathbf{K}}{\partial \rho} \mathbf{K}^{-1} + \frac{\partial \mathbf{F}}{\partial \rho} \mathbf{F}^{-1} + \frac{\partial f(\mathbf{h}_1)}{\partial \rho} f(\mathbf{h}_1)^{-1} \right) \quad (55)$$

analogously to that of (41). Noting (28) and (17) and defining

$$\mathbf{Z}_\rho = \frac{\partial \mathbf{V}_\rho^{-1/2}}{\partial \rho}, \quad (56)$$

the derivatives of $\mathbf{K}$ and $\mathbf{F}$ are

$$\frac{\partial \mathbf{K}}{\partial \rho} = -\frac{1}{2} \mathbf{K} \mathbf{1}_{1\times T} \mathbf{V}_\psi^{1/2} \mathbf{Z}_\rho \frac{\omega}{\sigma} \mathbf{h}_1, \quad (57)$$

$$\frac{\partial \mathbf{F}}{\partial \rho} = \frac{1}{2} \mathbf{F} \ (\mathbf{y}_2 \circ \mathbf{y}_2)' \mathbf{H}_1^{-1} \mathbf{V}_\psi^{1/2} \mathbf{Z}_\rho \frac{\omega}{\sigma} \mathbf{h}_1, \quad (58)$$

$$\frac{\partial f(\ \mathbf{h}_1)}{\partial \rho} = f(\mathbf{h}_1) \left[-\frac{\rho}{1-\rho^2} - \frac{1}{2} \sigma^{-2} tr \left(\frac{\partial \mathbf{V}_\rho^{-1}}{\partial \rho} \mathbf{h}_1 \mathbf{h}_1^{'} \right) \right]. \quad (59)$$

We have used $(\partial/\partial\rho) \left| \mathbf{V}_\rho^{1/2} \right| = 1/\sqrt{1-\rho^2}$ in deriving the first term of equation (59). Noting that $\exp(-\mathbf{h}_1{}')\mathbf{Y}_2 = (\mathbf{y}_2 \circ \mathbf{y}_2)' \mathbf{H}_1^{-1}$ under the null hypothesis, we have

$$\frac{\partial \mathbf{K}}{\partial \rho}|_{\mathbf{H}_0} = -\frac{1}{2} \mathbf{K} \left[\mathbf{1}_{1\times T} \mathbf{V}_\rho^{1/2} \mathbf{Z}_\rho \frac{\omega}{\sigma} \mathbf{h}_1 \right] = -\frac{\partial \mathbf{K}}{\partial \psi}|_{\mathbf{H}_0}, \quad (60)$$

$$\frac{\partial \mathbf{F}}{\partial \rho}|_{\mathbf{H}_0} = -\frac{1}{2} \mathbf{F} \ tr \left[\mathbf{Y}_2 \mathbf{V}_\rho^{1/2} \mathbf{Z}_\rho \frac{\omega}{\sigma} \mathbf{h}_1 \exp(-\mathbf{h}_1^{'}) \right] = -\frac{\partial \mathbf{F}}{\partial \psi}|_{\mathbf{H}_0}. \quad (61)$$

The score function with respect to ρ is

$$\frac{\partial \log f(\mathbf{y})}{\partial \rho}|_{\mathbf{H}_0} = -\frac{\partial \log f(\mathbf{y})}{\partial \psi}|_{\mathbf{H}_0} - \frac{\rho}{1-\rho^2} - \frac{1}{2} \sigma^{-2} tr \left(\frac{\partial \mathbf{V}_\rho^{-1}}{\partial \rho} E_{\mathbf{h}_1|\mathbf{y}} \left(\mathbf{h}_1 \mathbf{h}_1^{'} \right) \right). \quad (62)$$

A.6 Score Function with respect to σ

In the likelihood, σ appears only in $\mathbf{K}, \mathbf{F}$ in (28) and $f(\mathbf{h}_1)$. Then we can derive the score function with respect to σ using the formula analogous to that of ρ given in (55), with ρ replaced by σ. We can easily show from (28) that, under the null hypothesis $\omega = \sigma$, the derivatives of $\mathbf{K}$ and $\mathbf{F}$ with respect to σ are equal to the negative of the derivations with respect to ω , namely

$$
\begin{aligned}
\frac{\partial \mathbf{K}}{\partial \sigma}|_{\mathbf{H}_0} &= -\frac{\partial \mathbf{K}}{\partial \omega}|_{\mathbf{H}_0}, &(63)\\
\frac{\partial \mathbf{F}}{\partial \sigma}|_{\mathbf{H}_0} &= -\frac{\partial \mathbf{F}}{\partial \omega}|_{\mathbf{H}_0}, &(64)
\end{aligned}
$$

so that no additional calculations are necessary. From (17), the derivative of $f(\mathbf{h}_1)$ is

$$\frac{\partial f(\mathbf{h}_1)}{\partial \sigma} = f(\mathbf{h}_1)\left[-\frac{t}{\sigma} + \frac{1}{\sigma^3} tr\left(\mathbf{V}_\rho^{-1}\mathbf{h}_1\mathbf{h}_1^{'}\right)\right]. \tag{65}$$

Using the formula

$$\frac{\partial \log f(\mathbf{y})}{\partial \sigma}\Big|_{\mathbf{H}_0} = E_{\mathbf{h}_1|\mathbf{y}}\left(\frac{\partial \mathbf{K}}{\partial \sigma}\mathbf{K}^{-1} + \frac{\partial \mathbf{F}}{\partial \sigma}\mathbf{F}^{-1} + \frac{\partial f(\mathbf{h}_1)}{\partial \sigma} f(\mathbf{h}_1)^{-1}\right), \tag{66}$$

whose derivation is analogous to that of (55), and comparing the formua (50), we have

$$\frac{\partial \log f(\mathbf{y})}{\partial \sigma}|_{\mathbf{H}_0} = -\frac{\partial \log f(\mathbf{y})}{\partial \omega}|_{\mathbf{H}_0} - \frac{t}{\sigma} + \frac{1}{\sigma^3} tr\left(\mathbf{V}_\rho^{-1} E_{\mathbf{h}_1|\mathbf{y}}\left(\mathbf{h}_1\mathbf{h}_1^{'}\right)\right). \tag{67}$$

B Pre-orthogonalization of data

B.1 Purpose

Before estimating the model using actual data the observed return variables should be orthogonalized so that the error terms (e_{1t}, e_{2t}) in the measurement equation (5) are distributed contemporaneously independently with unit variance according to the assumption in (5), since the actual financial returns are contemporaneously correlated.

We cannot estimate the model under the assumption that $(e_{1t}, e_{2t})'$ have non-zero correlation and non-unit variances, because the increased number of the parameters to be estimated increases the computational time of the maximum likelihood estimation considerably. This problem is especially serious when the volatility series has high autocorrelation. We believe that this difficulty can be removed in future by improved algorithm. At present, however, this assumption is unavoidable to perform Monte Carlo experiments reported in Section 4 with sufficient number of iterations

We here show that, if the null hypothesis is true, namely $h_{1t} \equiv h_{2t}$, we can orthogonalize the observed series so as to satisfy the assumption of uncorrelatedness in (5).

This pre-orthogonalization is not without cost. The most serious demerit is that the result of the test depends upon the order of variables; we have a different value of the test statistic by exchanging the order of the variables, because the second variable is redefined by the Cholesky decomposition.

This asymmetry could be removed by including the correlation parameter in the measurement equation in (5) explicitly and then by estimating it simultaneously by the maximum likelihood method. However, we cannot use this method because the computational time is prohibitively large if the correlation parameter is included, so that we are obliged to drop the correlation parameter from (5) and to orthogonalize data before executing the test in the empirical analysis in Section 6.

B.2 Algebraic details

We assume that under the null hypothesis $h_{1t} = h_{2t}$ for any t and that the actual data, say $(\tilde{y}_{1t}, \tilde{y}_{2t})$, is written as

$$\textbf{(Unorthogonalized Model):} \quad \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix} = \exp\left(\frac{h_{1t}}{2}\right)\mathbf{A}\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_3 & \alpha_2 \end{pmatrix}, \tag{68}$$

in practice, namely when the disturbance term of the measurement equation has non-zero correlation and non-unit variance, instead of (5). This assumption is justifiable because the proposed Lagrange multiplier test statistic uses only the estimation of the null model.

We here estimate $\mathbf{A}^{-1}$, which is the desired orthogonalization matrix. First, note that the product moment of $(\tilde{y}_{1t}, \tilde{y}_{2t})$ is

$$\Lambda \equiv E\begin{pmatrix} \tilde{y}_{1t}^2 & \tilde{y}_{1t}\tilde{y}_{2t} \\ \tilde{y}_{1t}\tilde{y}_{2t} & \tilde{y}_{2t}^2 \end{pmatrix} = E(\exp(h_{1t}))\mathbf{A}\mathbf{A}'. \tag{69}$$

We can estimate Λ consistently using the sample moment of $(\tilde{y}_{1t}^2, \tilde{y}_{1t}\tilde{y}_{2t}, \tilde{y}_{2t}^2)$. Then we have only to estimate $E(\exp(h_{1t}))$ in order to obtain $\mathbf{A}$ using the formula

$$\mathbf{A} = (E(\exp(h_{1t})))^{-1/2}\Lambda^{1/2}, \ \mathbf{A}^{-1} = (E(\exp(h_{1t})))^{1/2}\Lambda^{-1/2}, \tag{70}$$

where $\Lambda^{1/2}$ denotes the Cholesky decomposition of Λ defined in (69).

Defining

$$\begin{pmatrix} \ddot{y}_{1t} \\ \ddot{y}_{2t} \end{pmatrix} \equiv \Lambda^{-1/2} \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix}, \tag{71}$$

we have that

$$\begin{pmatrix} \ddot{y}_{1t} \\ \ddot{y}_{2t} \end{pmatrix} = \mathbf{A}^{-1}(E(\exp(h_{1t})))^{-1/2} \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix} = (E(\exp(h_{1t})))^{-1/2} \exp(h_{1t}/2) \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix},$$

and hence

$$\log \ddot{y}_{1t}^2 = -\log(E(\exp(h_{1t}))) + h_{1t} + \log e_{1t}^2.$$

Then, since we have $E(\log(e_{1t}^2)) = -1.27$ as shown by Harvey et al. (1994) and $E(h_{1t}) = 0$ from the stationarity of h_{1t}, we have that

$$\frac{1}{T} \sum \log \ddot{y}_{1t}^2 \approx E[\log \ddot{y}_{1t}^2] = -\log(E(\exp(h_{1t}))) - 1.27$$

and hence we can estimate $E(\exp(h_{1t}))$ in (70) by

$$E(\exp(h_{1t})) \approx \exp\left(-(1/T) \sum \log \ddot{y}_{1t}^2 + 1.27\right).$$

Then we can have the orthogonalized data in (5) by

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \hat{\mathbf{A}}^{-1} \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix},$$

where

$$\hat{\mathbf{A}}^{-1} = \exp\left[-\left((1/T) \sum \log \ddot{y}_{1t}^2 + 1.27\right)/2\right] \hat{\Lambda}^{-\frac{1}{2}}, \tag{72}$$

$$\hat{\Lambda} = \frac{1}{T} \begin{pmatrix} \sum \tilde{y}_{1t}^2 & \sum \tilde{y}_{1t}\tilde{y}_{2t} \\ \sum \tilde{y}_{1t}\tilde{y}_{2t} & \sum \tilde{y}_{2t}^2 \end{pmatrix}. \tag{73}$$

Acknowledgement

The authors are grateful to anonymous reviewers for very helpful comments and suggestions. This research is partially supported by the Grants-in-aid for Scientific Research of Japan Society for the Promotion of Science (C) 15K03394. The third author acknowledges the Australian Research Council and the National Science Council, Ministry of Science and Technology (MOST), Taiwan.

References

Asai, M., McAleer, M. and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews*, 25(2-3):145–175.

Chernoff, H. (1954). On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics*, 25(3):573–578.

Chesher, A. (1984). Testing for neglected heterogeneity. *Econometrica*, 52(4):865–872.

Chiba, M. and Kobayashi, M. (2013). Testing for a single-factor stochastic volatility in bivariate series. *Journal of Risk of Financial Management*, 6(1):31–618.

Cipollini, A. and Kapetanios, G. (2008). A stochastic variance factor model for large datasets and an application to S&P data. *Economics Letters*, 100(1):130–134.

Danıelsson, J. (1998). Multivariate stochastic volatility models: estimation and a comparison with vgarch models. *Journal of Empirical Finance*, 5(2):155–173.

Davidson, R. and MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press.

Engle, R. F. and Kozicki, S. (1993). Testing for common features. *Journal of Business & Economic Statistics*, 11(4):369–380.

Engle, R. F. and Susmel, R. (1993). Common volatility in international equity markets. *Journal of Business & Economic Statistics*, 11(2):167–176.

Fleming, J., Kirby, C. and Ostdiek, B. (1998). Information and volatility linkages in the stock, bond, and money markets. *Journal of Financial Economics*, 49(1):111–137.

Forbes, K. J. and Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *The Journal of Finance*, 57(5):2223–2261.

Greengard, L. and Strain, J. (1991). The fast Gauss transform. *SIAM Journal on Scientific and Statistical Computing*, 12(1):79-94.

Hamlton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.

Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series [with discussion] *Journal of the American Statistical Association*, 82(1):1032–1063.

Kitagawa, G. (1995). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models *Journal of Computational and Graphical Statistics*, 5(1):1–125.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179.

Tauchen, G. E. and Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485–505.

Watanabe, T. (1999). A non-linear filtering approach to stochastic volatility models with an application to daily stock returns. *Journal of Applied Econometrics*, 14(2):101–121.