Stochastic Covariance Models*

Manabu Asai† and Mike K.P. So‡

†Faculty of Economics, Soka University
‡The Hong Kong University of Science and Technology

30 August 2010

Abstract

A new class of stochastic covariance models based on Wishart distribution is proposed. Three categories of dynamic correlation models are introduced depending on how the time-varying covariance matrix is formulated and whether or not it is a latent variable. A stochastic covariance filter is also developed for filtering and predicting covariances. Extensions of the basic models enable the study of the long memory properties of dynamic correlations, threshold correlation effects and portfolio analysis. Suitable parameterization in the stochastic covariance models and the stochastic covariance filter facilitate efficient calculation of the likelihood function in high-dimensional problems, no matter whether the covariance matrix is observable or latent. Monte Carlo experiments investigating finite sample properties of the maximum likelihood estimator are conducted. Two empirical examples are presented. One deals with the realized covariance of high frequency exchange rate data, while the other examines daily stock returns.

Keywords: Covariance prediction; Dynamic correlation; Nonlinear time series; Long memory; Stochastic covariance filter; Threshold models.

---

*The authors wish to acknowledge the helpful comments and suggestions of Yoshi Baba, Tim Bollerslev, Yasuhiro Omori, Wolfgang Polasek, George Tauchen and Phillip Yu. The first author acknowledges the financial support of the Japan Society for the Promotion of Science and the Australian Academy of Science. The research of Mike So is supported by HKUST grant SBI08/09.BM06. Address for correspondence: Faculty of Economics, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan. Email address: m-asai@soka.ac.jp.

1 Introduction

Covariance matrix estimation has been an active research topic, partly because covariance matrices appear naturally in many stochastic models as a statistical measure of the interdependence of variables. Inference about covariance can be complicated when its dimension is high, implying a large number of unknown parameters which may make the positive definite condition of the covariance matrix difficult to satisfy in the estimation process. To alleviate the high dimensionality problem, there are three major classes of methods for covariance estimation. The first one is based on spectral decomposition using eigenvalues and orthogonal eigenvectors (Daniel and Kass, 1999). Another method follows from the Cholesky decomposition of covariance matrices. Daniels and Pourahmadi (2002) modeled covariance matrices and studied the relationship between a prior they proposed and the inverse Wishart prior. Smith and Kohn (2002) proposed parsimonious specifications based on Cholesky decomposition. Also falling into this category of methods is that of Pourahmadi (2000, 2007). The third method is to analyze the covariance matrix via variances and covariances. Barnard, McCulloch and Meng (2000) called this the separation strategy. Using this idea, Liechty, Liechty and Muller (2004) proposed priors for Bayesian correlation modeling. Other methods include the matrix logarithmic transformation prior formulation of Leonard and Hsu (1992) and the use of an inverse covariance matrix prior in Wong, Carter and Kohn (2003) and Pitt, Chan and Kohn (2006) for covariance selection. Rather than assuming homogenous covariances as in the above-mentioned procedures, this study develops a modeling framework for time-varying covariance matrices.

The study of time-dependent covariance or multivariate stochastic volatility (MSV) models has been an active area of research in the past decade, primarily because of the need to capture the dependence of high-dimensional financial returns. The first MSV model introduced by Harvey, Ruiz and Shephard (1994) was considered to be an alternative to the multivariate GARCH models proposed by Bollerslev (1990). Both types of models assume constant correlations. With respect to the multivariate GARCH, Engle and Kroner (1995) and Ding and Engle (2001) developed new models which guarantee the positive definiteness of the covariance matrices, allowing conditional correlations to be time-varying. Later, Engle (2002) and Tse and Tsui (2002) proposed the dynamic conditional correlation (DCC) models which specify the process of conditional correlations directly. The DCC specification reduces the number of parameters dramatically. See McAleer (2005) for the survey of multivariate GARCH models.

Turning back to the MSV model, there have been two streams of research. One has worked with the factor model, while the other is based on the Wishart distribution (see Wishart, 1928). Among several factor models, Pitt and Shephard (1999a) introduced mean factors which have stochastic variances, in addition to the original specification of Harvey, Ruiz and Shephard (1994). One of the byproducts of this specification is to enable the correlation matrix to vary over time. Chib, Nardari and Shephard (2006) developed a further generalization, allowing jumps and heavy tails in the conditional distributions. Regarding the Wishart distribution, several scholars employ it in order to guarantee the positive definiteness of the covariance matrix. Philipov and Glickman (2006a,b) proposed a high-dimensional MSV model in which the covariance matrices are driven by Wishart random processes. Philipov and Glickman (2006a) specified that the log of the determinant of the covariance matrix follows an AR(1) process, while Philipov and Glickman (2006b) extended it to a factor MSV model for the purpose of reducing the number of the parameters. Asai and McAleer (2009) suggested a dynamic correlation model, which is an extension of the dynamic conditional correlation (DCC) model of Engle (2002). Gourieroux (2006) and Gourieroux et al. (2009) proposed the Wishart autoregressive (WAR) multivariate process of stochastic positive definite matrices. Intuitively, the WAR model is based on the sum of the outer-products of the vector autoregressive processes. Among these MSV models, Asai and McAleer (2009) assumed a process with dynamic correlations, while all the other have considered dynamic covariance which produces dynamic correlation as a byproduct. Asai, McAleer and Yu (2006) and Chib, Omori and Asai (2009) give surveys of the various MSV models.

The purpose of the paper is to introduce a new class of stochastic covariance models using Wishart distribution. Our specification clarifies the relationship between the stochastic and conditional covariance matrices. There are four ways we contribute to the covariance modeling literature. First, we propose three categories of dynamic correlation models depending on how we formulate the time-varying covariance matrix and whether it is a latent variable. Second, we develop a stochastic covariance filter, a matrix analog of the Kalman filter, for filtering and prediction of covariances. Third, extensions of the basic models enable us to study long memory properties in the dynamic correlations, threshold correlation effects and to do portfolio analysis. Finally, suitable parameterization in our dynamic correlation models and the stochastic covariance filter facilitate efficient calculation of the likelihood function in high-dimensional problems, no matter whether the covariance matrix is observable or latent. The organization of the paper is as follows. Section 2 proposes the new multivariate process with the stochastic covariance, and also introduces the stochastic covariance filter. Section 3 suggests the three categories of dynamic correlation models. Section 4 extends the models to accommodate the long memory, asymmetry and portfolio analysis. Section 5 discusses likelihood inference procedures for parameter estimation, and Section 6 conducts Monte Carlo experiments to study finite-sample properties of estimators. Section 7 shows two empirical examples; one is for the bivariate exchange rates using high frequency data, while the other deals with the trivariate daily returns for stock indices.

Section 8 gives conclusions.

2 Multivariate Processes with Stochastic Covariance

2.1 A new stochastic covariance model

Let y_t be an m-dimensional random vector with the stochastic covariance matrix C_t $(m \times m)$. Let $\Im_t$ be the information set up to t. Assume that y_t is observable, while C_t can be either observable or unobservable. We define the conditional mean and covariance matrix as

$$\begin{aligned} \mu_t &\equiv E[y_t|\Im_{t-1}], \\ C_{t|t-1} &\equiv E[C_t|\Im_{t-1}]. \end{aligned} \tag{1}$$

In order to avoid the problem of the simultaneous equation bias, we excluded the exogenous variables for period t from the conditional covariance matrix. On the other hand, the conditional mean can be extended to include such exogenous variables[1]. It is also possible to extend μ_t to depend on C_t, as in the stochastic volatility in mean model suggested by Koopman and Uspensky (2002). Now, we introduce the new stochastic covariance model as

$$y_t = \mu_t + C_t^{1/2} z_t, \quad z_t \sim N(\mathbf{0}, I_m), \tag{2}$$

$$C_t = C_{t|t-1}^{1/2} \left\{(\nu - m - 1)E_t\right\} C_{t|t-1}^{1/2}, \quad E_t^{-1} \sim W(I_m, \nu), \tag{3}$$

where $W(A, p)$ denotes the Wishart distribution with the scale matrix A and the degrees-of-freedom parameter p. The new model has an alternative form as,

$$\begin{aligned} y_t|\Im_{t-1}, C_t &\sim N(\mu_t, C_t), \\ C_t|\Im_{t-1} &\sim IW\left((\nu - m - 1)C_{t|t-1}, \nu\right), \end{aligned} \tag{4}$$

where $IW(B, p)$ denotes the inverse Wishart distribution with the scale matrix B and the degrees-of-freedom parameter p. By the properties of the inverse Wishart distribution, $E[C_t|\Im_{t-1}] =$

[1]Ren and Polasek (2000) considered the VAR-GARCH in mean model. Although we do not deal with such models here, our new specification includes conditional covariance models with feedback effects in the mean as a special case.

$(\nu - m - 1)^{-1} \times (\nu - m - 1)C_{t|t-1} = C_{t|t-1}$, which matches the definition in (1). Our stochastic covariance model is developed by defining suitable structures in the conditional mean of C_t given previous information. We specify the dynamic structure of both the time-varying variances and the correlations explicitly in (4). This is unlike the classical methods of Harvey, Ruiz and Shephard (1994) and Chib, Nardari and Shephard (2006) which implicitly captured stochastic correlation by innovation correlations or by defining correlations using the variances of latent factors and factor loadings. The model in (4) does not belong to the class of Wishart process as discussed by Philipov and Glickman (2006a) or Gourieroux, Jasiak and Sufana (2009), as $C_{t|t-1}$ is not necessarily a function of $C_1, ..., C_t$. One novelty of this new model is that $C_{t|t-1}$ is formulated by observations up to time $t - 1$. This enables us to derive useful statistical properties based on the model in (4), namely (i) the connection between y_t and the conditional covariance, (ii) multi-step-ahead forecasts of C_t and the variance of y_t, and (iii) the filtered estimate of C_t.

First of all, define $\theta = (\theta_1', \theta_2')'$ as the parameter vector, where θ_2 includes ν. Denote the density functions of y_t and C_t as $f(y_t|\Im_{t-1}, C_t, \theta_1)$ and $f(C_t|\Im_{t-1}, \theta_2)$, respectively. Integrating out C_t, we have the distribution of $y_t|\Im_{t-1}$ as

$$\begin{aligned} f(y_t|\Im_{t-1}, \theta) &= \int f(y_t|\Im_{t-1}, C_t, \theta_1) f(C_t|\Im_{t-1}, \theta_2) dC_t \\ &= (\pi k)^{-\frac{m}{2}} \frac{\Gamma(\frac{k+m}{2})}{\Gamma(\frac{k}{2})} |H_t|^{-\frac{1}{2}} \left[1 + \frac{(y_t - \mu_t)' H_t^{-1} (y_t - \mu_t)}{k} \right]^{-\frac{k+m}{2}}, \qquad (5) \end{aligned}$$

where $k = \nu - m + 1$ and $H_t = k^{-1}(\nu - m - 1)C_{t|t-1}$. This defines $y_t|\Im_{t-1} \sim MT(\mu_t, H_t, k)$, a multivariate t distribution with the location parameter μ_t, scale parameter H_t and k degrees of freedom. Since each component of y_t follows a univariate conditional t distribution with k degrees of freedom, it has a finite $4th$ moment only when $k > 4$ or $\nu > m + 3$. The higher the dimension of m, the more fat-tailed is the conditional distribution of y_t induced by C_t. The conditional mean and variance of y_t are

$$E(y_t|\Im_{t-1}) = \mu_t,$$

$$Var(y_t|\Im_{t-1}) = \frac{k}{k-2}H_t = C_{t|t-1}.$$

Hence, the specification in (4) makes the conditional mean of stochastic covariance, C_t, to be the conditional covariance matrix of y_t.

Secondly, we consider multi-step-ahead prediction given $\Im_t$, i.e., $E[C_{t+h}|\Im_t]$ and $Var(y_{t+h}|\Im_t)$ for $h \geq 1$. By definition, $C_{t+1|t}$ is the one-step-ahead forecast of C_{t+1} given $\Im_t$. Hence, we turn to the prediction of C_{t+h} $(h > 1)$ given $\Im_t$. We can obtain $f(y_{t+1}, ..., y_{t+h}|\Im_t)$ as a product of multivariate t distributions:

$$f(y_{t+1}, ..., y_{t+h}|\Im_t) = \prod_{i=1}^{h} f(y_{t+i}|\Im_{t+i-1}) = \prod_{i=1}^{h} MT(\mu_{t+i}, H_{t+i}, k).$$

For multiple-step prediction, we can then use

$$E[C_{t+h}|\Im_t] = E[E[C_{t+h}|\Im_{t+h-1}]|\Im_t]] = E[C_{t+h|t+h-1}|\Im_t], \; h \geq 1.$$

From this, we can estimate $E[C_{t+h}|\Im_t]$ (for $h > 1$) by simulating $y_{t+1}, ..., y_{t+h-1}$ from the product of the multivariate t distributions and approximate it by the Monte Carlo average

$$E[C_{t+h}|\Im_t] \approx \frac{1}{N}\sum_{j=1}^{N} C^{(j)}_{t+h|t+h-1}, \tag{6}$$

where $C^{(j)}_{t+h|t+h-1}$ is computed by the jth draw from $f(y_{t+1}, ..., y_{t+h-1}|\Im_t)$. Similarly, we can estimate $Var(\mu_{t+h}|\Im_t)$ by

$$Var(\mu_{t+h}|\Im_t) \approx \frac{1}{N}\sum_{j=1}^{N} \mu^{(j)}_{t+h}\, \mu^{(j)}_{t+h}{}' - \left(\frac{1}{N}\sum_{j=1}^{N} \mu^{(j)}_{t+h}\right)\left(\frac{1}{N}\sum_{j=1}^{N} \mu^{(j)}_{t+h}\right)', \tag{7}$$

where $\mu^{(j)}_{t+h}$ is determined by the jth draw from $f(y_{t+1}, ..., y_{t+h-1}|\Im_t)$. Adding the Monte Carlo estimators in (6) and (7) gives an estimator of the h-step-ahead predictive variance of y_{t+h}, i.e.

$Var(y_{t+h}|\Im_t)$, by using the result

$$Var(y_{t+h}|\Im_t) = E[C_{t+h}|\Im_t] + Var(\mu_{t+h}|\Im_t). \tag{8}$$

To prove (8), note that $E[y_{t+h}|\Im_t] = E[E[y_{t+h}|\Im_{t+h-1}, C_{t+h}]|\Im_t] = E[\mu_{t+h}|\Im_t]$, and

$$\begin{aligned}
E[y_{t+h}\, y_{t+h}{}'\, |\Im_t] &= E[E[y_{t+h}\, y_{t+h}{}'\, |\Im_{t+h-1}, C_{t+h}]|\Im_t] \\
&= E[Var(y_{t+h}|\Im_{t+h-1}, C_{t+h}) + E[y_{t+h}|\Im_{t+h-1}, C_{t+h}]\; E[y_{t+h}|\Im_{t+h-1}, C_{t+h}]'] \\
&= E[C_{t+h} + \mu_{t+h}\, \mu_{t+h}{}'\, |\Im_t].
\end{aligned}$$

Equation (8) is especially important when we are interested in the multiple-step forecasting of the variance of y_t.

Thirdly, we also suggest an approach for updating C_t analogous to the updating step in Kalman filter. While the prediction of C_t can be accomplished using the conditional distribution of $C_t|\Im_{t-1}$, we can also filter C_t based on the information up to time t. This can be done by noting that

$$\begin{aligned}
f(C_t|\Im_t) &= f(C_t|y_t, \Im_{t-1}) \\
&\propto f(y_t|C_t, \Im_{t-1}) f(C_t|\Im_{t-1}) \\
&\propto |C_t|^{-\frac{(\nu+1)+m+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\Phi_t C_t^{-1}\right)\right\},
\end{aligned}$$

where $\Phi_t = (\nu - m - 1)C_{t|t-1} + (y_t - \mu_t)(y_t - \mu_t)'$. Then we have

$$C_t|\Im_t \sim IW(\Phi_t, \nu + 1), \text{ or } C_t^{-1}|\Im_t \sim W(\Phi_t^{-1}, \nu + 1), \tag{9}$$

giving

$$E[C_t|\Im_t] = (\nu - m)^{-1}\Phi_t = (\nu - m)^{-1}[(\nu - m - 1)C_{t|t-1} + (y_t - \mu_t)(y_t - \mu_t)']. \tag{10}$$

Such updating is especially important when C_t is unobservable.

2.2 A stochastic covariance filter

Classical signal extraction approaches help predict and filter latent state variables under state space models. The usual technique involved is the Kalman filter (Harvey, 1989; 1993) which proceeds recursively to the calculation of predicted and filtered state values for statistical inference of the unobservable signal. In this study, we develop a stochastic covariance filter (SCF), which can be regarded as a matrix analog of the Kalman filter. Instead of having an unknown state variable, in this section $\{C_t\}$ is assumed to be a latent process so that $\Im_t$ only includes the history of $y_1, ..., y_t$. Following the notation common in the signal extraction literature, denote the filtered value of C_t, i.e. $E[C_t|\Im_t]$, by $\hat{C}_t$. The prediction and filtering equations for the proposed SCF are then

Prediction

$$C_{t|t-1} = g_{t-1}(C_{t-1|t-2}, \hat{C}_{t-1}), \tag{11}$$

Updating

$$\hat{C}_{t-1} = \frac{\nu - m - 1}{\nu - m} C_{t-1|t-2} + \frac{1}{\nu - m}(y_{t-1} - \mu_{t-1})(y_{t-1} - \mu_{t-1})', \tag{12}$$

where g_{t-1} is a function of $C_{t-1|t-2}$ and $\hat{C}_{t-1}$ based on the information $\Im_{t-1}$. While the prediction equation takes a very general form in terms of the function g_t, the updating equation follows from (10). From the updating formula in (12), the parameter ν and the dimension m determine the contribution of the information up to time $t-2$, summarized in $C_{t-1|t-2}$, and y_{t-1}, to form the updated value $\hat{C}_{t-1}$, which in turn governs the conditional variance of y_t given $\Im_{t-1}$. A greater contribution is from y_t, the information at time t, if the data is very fat-tailed or $\nu - m$ is small. So the mixture Wishart representation in (4) induces weights in the construction of $\hat{C}_t$, and iteratively produces $C_{t|t-1}$ using our SCF. The time series evolution of $C_{t|t-1}$ in (11) will be used to define a new dynamic correlation model in the next section. The choice of the function $g(\cdot)$ is flexible, but a special form is discussed in Section 4 which can help capture any long-range dependence in stochastic covariance.

3 Dynamic Correlations

Returning to the stochastic covariance model in (4), consider three categories, depending on how we formulate C_t and whether C_t is a latent variable.

1. The conditional covariance $C_{t|t-1}$ includes past information of C_t, which is observable.

2. The conditional covariance $C_{t|t-1}$ includes past information of C_t, which is a latent variable.

3. The conditional covariance $C_{t|t-1}$ excludes past information of C_t. In this case, past values of C_t are replaced by their filtered estimates, $\hat{C}_{t-s}(s \geq 1)$.

In the first category, we can define a time series model with realized covariance matrices, while the second category deals with a class of multivariate stochastic volatility models. The final category introduces a class of multivariate GARCH (MGARCH) models based on our stochastic covariance filter.

3.1 A new modeling framework

In this section, we address specifications for the conditional covariance, $C_{t|t-1}$, in detail. In the literature on multivariate GARCH models, there are two approaches to specifying $C_{t|t-1}$. The first is to model $C_{t|t-1}$ itself, as in Engle and Kroner (1995) and Ding and Engle (2001). The other is modeling the conditional correlations, $P_{t|t-1} = \Delta_{t|t-1}^{-1/2} C_{t|t-1} \Delta_{t|t-1}^{-1/2}$, rather than conditional covariances, as Engle (2002) and Tse and Tsui (2002) have done, where $\Delta_{t|t-1} = \mathrm{diag}(C_{t|t-1})$ contains the conditional variances of each component of y_t given $\Im_{t-1}$, and $\mathrm{diag}(A)$ is the diagonal part of a matrix A. The latter approach allows the conditional correlation to vary parsimoniously. Recently, Asai and McAleer (2009) have suggested a dynamic stochastic correlation model which allows time-varying correlation in MSV models. In this paper, we develop a new parsimonious specification for $P_{t|t-1}$ in multivariate covariance models to handle all three categories of problems.

Define the stochastic correlation matrix by $P_t = \Delta_t^{-1/2} C_t \Delta_t^{-1/2}$, where $\Delta_t = diag(C_t)$ is the stochastic variance matrix. The corresponding updated estimates of P_t and Δ_t obtained from $\hat{C}_t$ in the updating step are denoted by $\hat{P}_t = \hat{\Delta}_t^{-1/2} \hat{C}_t \hat{\Delta}_t^{-1/2}$ and $\hat{\Delta}_t = diag(\hat{C}_t)$, respectively. Now, we propose a new dynamic correlation model, as follows.

$$C_{t|t-1} = \Delta_{t|t-1}^{1/2} P_{t|t-1} \Delta_{t|t-1}^{1/2}, \tag{13}$$

with

$$\begin{aligned} P_{t|t-1} &= S \circ (\iota\iota' - A - B) + A \circ P_{t-1} + B \circ P_{t-1|t-2}, \\ \text{vecd}(\Delta_{t|t-1}) &= \kappa + \gamma \circ \text{vecd}(\Delta_{t-1}) + \delta \circ \text{vecd}(\Delta_{t-1|t-2}), \end{aligned} \tag{14}$$

where ι is the vector of ones, $\circ$ is the Hadamard product (element-by-element multiplication), S is the unconditional correlation matrix, A and B are (possibly) positive semi-definite matrices of parameters, and κ, γ and δ are $m \times 1$ vectors of parameters. Also, the operator vecd(A) creates a vector from the diagonal elements of the square matrix A. An alternative formulation of (14) is to replace P_{t-1} by $\hat{P}_{t-1}$ and Δ_{t-1} by $\hat{\Delta}_{t-1}$. This will be discussed in a later section. For convenience, we call the model defined by (4), (13) and (14) the dynamic correlation, or 'DC' model.

In the DC model, the conditional variance and conditional correlations are specified separately. The vector of conditional variance, vecd$(\Delta_{t|t-1})$ has a structure similar to that of a GARCH model. Instead of the squared residuals in GARCH models, we work with the realized variances, Δ_{t-1}. With respect to the conditional correlation, $P_{t|t-1}$, note that its specification requires no standardization, as S, P_{t-1} and $P_{t-1|t-2}$, like the specification of Tse and Tsui (2002), are correlation matrices. The dynamic conditional correlation (DCC) model of Engle (2002) instead employed an additional process of Q_t to produce the correlation matrix by defining $P_{t|t-1} = [\mathrm{diag}(Q_t)]^{-1/2}\, Q_t\, [\mathrm{diag}(Q_t)]^{-1/2}$. That approach accommodates $z_{t-1}z'_{t-1}$, where $z_t = \Delta_{t|t-1}^{-1/2}(y_t - \mu_t)$ rather than the P_{t-1} in (14). Using the approach taken by Ding and Engle (2001), we can show that $P_{t|t-1}$ is positive definite if $(\iota\iota' - A - B)$, A and B are positive definite. Note that P_{t-1} is always positive definite, unlike the outer product of the standardized vector in the original DCC model, and the positive definiteness of P_{t-1} supports $P_{t|t-1}$ to be positive-definite. Each element of $\mathrm{vecd}(\Delta_{t|t-1})$ follows a GARCH-type process, so $\kappa_i > 0$, $\gamma_i > 0$, $\delta_i > 0$ and $\gamma_i + \delta_i < 1$ for all $i = 1, 2, \ldots, m$. As in Engle (2002), a parsimonious specification for P_t is given by

$$P_{t|t-1} = S(1 - \alpha - \beta) + \alpha P_{t-1} + \beta P_{t-1|t-2}. \tag{15}$$

We will use this specification in our empirical analysis. In this specification, we may consider $(1-\alpha-\beta)$, α and β as the weights for the three kinds of correlation matrices, i.e. the unconditional, the realized and the conditional.

Although the DC model starts from the specification of the process of the correlation matrices, general multivariate covariance models yield time-varying correlations as well. For example, we can use a BEKK-type model (named after Baba, Engle, Kraft and Kroner and introduced in Engle and Kroner, 1995). The first order form is given by $C_{t|t-1} = \underline{C}'\underline{C} + AC_{t-1}A' + BC_{t-1|t-2}B'$, where A and B are $m \times m$ matrices of parameters, and $\underline{C}$ is the upper triangular matrix. Another possibility is a variant of the diagonal GARCH model (Bollerslev, Engle and Wooldridge, 1988) given by $C_{t|t-1} = \underline{\Omega} \circ (\iota\iota' - A - B) + A \circ C_{t-1} + B \circ C_{t-1|t-2}$, where $\underline{\Omega}$ is the positive definite matrix of parameters, and A and B are (possibly) positive semi-definite matrices. These specifications also introduce dynamic patterns in the correlation matrices. In the same manner, it is possible to develop various kinds of stochastic covariance models by assuming $C_{t|t-1} = g_{t-1}(C_{t-1|t-2}, C_{t-1})$

as in the SCF. For instance, we may also use the matrix exponential model suggested by Chiu et al. (1996) and its variants including that of Kawakatsu (2006). Before moving to the next topic, we should emphasis that the SC model (14) is not in the class of multivariate GARCH models because the right hand side involving P_{t-1} and Δ_{t-1} is not deterministic given y_1,..., y_{t-1}. On the other hand, the flexibility of using $\hat{P}_{t-1}$ and $\hat{\Delta}_{t-1}$ generated from the updating step in place of P_{t-1} and Δ_{t-1} in (14) enables the modeling of stochastic covariance using observation-driven models like the multivariate GARCH. This will be explored in Section 3.3.

3.2 Time series models with realized covariance

The three categories of scenarios can be used to examine the hierarchical structures of the DC models. The first category assumes that the conditional covariance $C_{t|t-1}$ depends on past information of C_t, which is observable. In this case, the realized covariance (RC) matrix can be considered as the observed value of C_t or the observed value with some measurement error[2]. With a high-frequency time series (data collected with a less-than-one-day horizon), daily realized covariance can be obtained. Studies show that this realized covariance can help in various financial applications (Bandi, Russell and Zhu, 2008) and thus there is theoretical and practical interest in modeling realized covariance. Much of the prior work on this topic, including that of Andersen, Bollerslev, Diebold and Labys (2003), Corsi (2004) and McAleer and Medeiros (2008) focused on univariate modeling, i.e. developing time series models of realized variance. Some exceptions are Voev (2007) and Gourieroux, Jasiak and Sufana (2009). In this paper, we work on multivariate modeling where y_t can be a financial return on day t which has the property $var(y_t|\Im_{t-1}, C_t)$ equals C_t, the realized covariance on day t. We call this DC model as the 'DC-RC' when re-

[2]For the realized volatility, Bandorff-Nielsen and Shephard (2002) called such measurement error as 'realized volatility error'. The error is also caused by the microstructure noise common in ultra high frequency data. See also Andersen et al. (2003) and Asai, McAleer and Medeiros (2008).

alized covariance is available for C_t. We model the conditional mean of C_t, i.e. $C_{t|t-1}$ using (14). With the hierarchical structure in (4), the predictive distribution $f(y_t, C_t|\Im_{t-1})$ is written as $f(y_t|C_t, \Im_{t-1})f(C_t|\Im_{t-1})$, where the first component is multivariate normal and the second part is inverse Wishart, from which the likelihood function is formulated as

$$\prod_t f(y_t, C_t|\Im_{t-1}) = \prod_t N(\mu_t, C_t) \times IW((\nu - m - 1)C_{t|t-1}, \nu). \tag{16}$$

Since in this case the information set $\Im_t$ includes past history of both y_t and C_t, our stochastic covariance model can help incorporate both the information of y_t and C_t in forecasting future realized covariance. For example, we can have

$$P_{t|t-1} = S \circ (\iota\iota' - A - B - D) + A \circ P_{t-1} + B \circ P_{t-1|t-2} + D \circ R_{t-1},$$

where D is positive definite and R_{t-1} is any correlation matrix constructed from y_1,..., y_{t-1}. The information of y_t enters into the distribution of C_t through R_t, but the likelihood decomposition in (16) is still valid. For the case of unobservable C_t, (4), (13) and (14) define a new class of MSV models, which is referred to as the 'DC-MSV' model.

3.3 A MGARCH model based on stochastic covariance filter

With respect to the Category 3, we propose to replace C_t by its filtered estimate $\hat{C}_t = E[C_t|\Im_t]$ in equation (10), which is the weighted average of $C_{t|t-1}$ and $(y_t - \mu_t)(y_t - \mu_t)'$ with weights $(\nu - m - 1)/(\nu - m)$ and $1/(\nu - m)$, respectively. We then have $\hat{\Delta}_t = \text{diag}(\hat{C}_t)$ and $\hat{P}_t = \hat{\Delta}_t^{-1/2}\hat{C}_t\hat{\Delta}_t^{-1/2}$. Equation (4) gives a time series model in hierarchical form with the stochastic variance C_t following a conditional Wishart distribution. Integrating out C_t to get the multivariate t predictive distribution in (5) and with the prediction equation in (11), we can define a new

MGARCH model as

$$y_t|\Im_{t-1} \sim MT(\mu_t, \frac{k-2}{k} C_{t|t-1}), \qquad C_{t|t-1} = \Delta_{t|t-1}^{1/2} P_{t|t-1} \Delta_{t|t-1}^{1/2},$$

$$P_{t|t-1} = S \circ (\iota\iota' - A - B) + A \circ \hat{P}_{t-1} + B \circ P_{t-1|t-2}, \tag{17}$$

$$\text{vecd}(\Delta_{t|t-1}) = \kappa + \gamma \circ \text{vecd}(\hat{\Delta}_{t-1}) + \delta \circ \text{vecd}(\Delta_{t-1|t-2}),$$

where $k = \nu - m + 1$. Intrinsically, the normal-Wishart mixture in (4) implies a conditional multivariate t distribution for y_t given $\Im_{t-1}$ with degrees of freedom $k = \nu - m - 1$. $C_{t|t-1}$ is equivalent to $var(y_t|\Im_{t-1})$ as shown in section 2. Empirical data analyses in the literature show that the fat-tailed characteristics of the t distribution are more suitable than a multivariate normal distribution. The degrees of freedom also reveals higher conditional kurtosis in larger m if ν is fixed. Equation (17) lies in the framework of our SCF, where g_{t-1} in this case is a nonlinear function of $C_{t-1|t-2}$ and $\hat{C}_{t-1}$.

A novelty of our MGARCH approach is to be able to filter out C_t using the SCF, whereas in traditional MGARCH models, no such step is possible because the conditional variance, $var(y_t|\Im_{t-1})$, is 'assumed' rather than constructed by an underlying stochastic covariance process, which is the source of the changing variance of y_t. Even though filtering can be done under MSV models, for examples in Kitagawa (1996) and Pitt and Shephard (1999b), the filtering can be very time consuming and usually requires Monte Carlo simulations. In practice, $\hat{C}_t = E[C_t|\Im_t]$ is the best estimator of C_t in the mean square sense given the information up to time t. So from the signal extraction point of view, it is more accurate statistically to use $\hat{C}_t$, instead of $C_{t|t-1}$ to analyze the unknown C_t. With respect to financial applications (where y_t stands for a return at time t), it makes more sense to use $\hat{C}_t$ than $C_{t|t-1}$ for portfolio analysis and risk calculations at time t. By construction, $P_{t|t-1}$ and $\Delta_{t|t-1}$ are determined from the past information of y_t, and they do not depend on the past values of C_t. Hence, $C_{t|t-1}$ depends only on the past information of y_t. In this sense, the above model accommodating the filtered estimates can be interpreted as a new specification of the DCC model. Furthermore, $y_t|\Im_{t-1}$ follows a multivariate t distribution and so we call the new DCC model as the 'DCCt model with Filtered Covariance' (DCCt-FC).

Let us examine some properties of the DCCt-FC model to show its similarities with and differences from Engle's (2002) DCC models. We may write the (i,i) element of Δ_t in the DCCt-FC model as

$$\Delta_{ii,t} = \kappa_i + \gamma_i \hat{D}_{ii,t-1} + \delta_i \Delta_{ii,t-1},$$

where $\hat{D}_{ii,t} = (\nu - m)^{-1}[(\nu - m - 1)\Delta_{ii,t} + (y_{it} - \mu_{it})^2]$. This leads to

$$\Delta_{ii,t} = \kappa_i + \gamma_i^* (y_{it} - \mu_{it})^2 + \delta_i^* \Delta_{ii,t-1},$$

where $\gamma_i^* = \gamma_i(\nu - m)^{-1}$ and $\delta_i^* = \gamma_i + \delta_i - \gamma_i^*$. This is the same as the GARCH specification. Therefore, there is no difference between the DCCt-FC and DCC models in terms of their conditional variances. On the other hand, such a correspondence is unavailable with respect to the correlation process, because of the differences between $\hat{P}_{t-1}$ in the DCCt-FC and $z_{t-1}z_{t-1}'$ in the DCC. There is therefore a crucial difference between the DCCt-FC and DCC models in terms of the correlation process.

4 Extensions

In this section, we propose several extensions of the new DC model. For convenience, we will refer these variants using the abbreviations, DC-RC, DC-MSV and DCCt-FC.

4.1 A long memory model

We extend the DC model in the previous section to handle any long memory properties observed in the realized volatility, and possibly also present in realized correlations. Baillie et al. (1996) and Bollerslev and Mikkelsen (1996) have previously suggested long-memory conditional variance models, and Andersen et al. (2003) proposed a vector autoregressive approach (VAR) for the realized variances. To reduce the number of parameters, assume that the degrees of long-run dependency are common to all the variances (See Andersen et al., 2003). In this paper, we develop a DC model which also captures any long memory property in the time-varying correlations. Assume also a common long-run dependency in the correlations, though we may relax these assumptions easily. A new multivariate long memory time series model in C_t is introduced by separating it into P_t and Δ_t for modeling

$$\begin{aligned}(1-\phi L)(1-L)^{d_c}\mathrm{vech}(P_t - S) &= (1-\beta L)e_t,\\ (I_m-\lambda L)(1-L)^{d_v}\mathrm{vecd}(\Delta_t) &= \kappa + (I_m - \delta L)u_t,\end{aligned} \tag{18}$$

where L is the lag operator, $e_t = \mathrm{vech}(P_t - P_{t|t-1})$ and $u_t = \mathrm{vecd}(\Delta_t - \Delta_{t|t-1})$ play the role of innovations, I_m is the $m \times m$ identity matrix, κ is a $m \times 1$ vector, ϕ and β are scalar parameters, λ and δ are the $m \times m$ diagonal parameter matrices $\mathrm{diag}\{\lambda_1, ..., \lambda_m\}$ and $\mathrm{diag}\{\delta_1, ..., \delta_m\}$ respectively, and d_c and d_v are the common degrees of the long-run dependencies for variances and correlations, respectively. Rewriting (18), we propose the Fractionally Integrated DC (FIDC) model, which is given by equation (13) and

$$\begin{aligned}\mathrm{vech}(P_{t|t-1}) =& \mathrm{vech}(S) + \beta \mathrm{vech}(P_{t-1|t-2} - S)\\ &+ \left\{1 - \beta L - (1-\phi L)(1-L)^{d_c}\right\}\mathrm{vech}(P_t - S),\\ \mathrm{vecd}(\Delta_{t|t-1}) =& \kappa + \delta \circ \mathrm{vecd}(\Delta_{t-1|t-2})\\ &+ \left\{I_m - \delta L - (I_m - \lambda L)(1-L)^{d_v}\right\}\mathrm{vecd}(\Delta_t).\end{aligned} \tag{19}$$

The assumptions about the common degree of long-range dependency governed by d_c and d_v can be relaxed easily. In equation (19) we encounter the lag operator, which takes the form $(1-aL)(1-L)^d$. It may be convenient to show its alternative representation given by Hosking (1981), $\sum_{k=0}^{\infty} b_k L^k$, where

$$b_k = \left(1 - a - \frac{1+d}{k}\right)\frac{\Gamma(k-d-1)}{\Gamma(-d)\Gamma(k)}.$$

As a special case, setting $d_c = d_v = 0$ in (19) leads to the DC model in (15) with the volatility equation given in (14). In this case, the relationships between the parameters are given by $\phi = \alpha+\beta$ and $\lambda_i = \gamma_i + \delta_i$ $(i = 1, 2, \ldots, m)$.

When C_t is taken as the realized covariance, the FIDC belongs to the first category of models proposed in Section 3. In this case, we name (19) as the FIDC-RC model. Andersen et al. (2003) constructed a fractionally integrated VAR model for the log of the realized variances with the common parameter d_v. Compared with that model, our model also includes the process for the realized correlations, which is new in the literature. On the other hand, when C_t is a latent variable, equation (19) gives a new specification for the long memory MSV model (see So and Kwok, 2009, for an example). This case produces the FIDC-MSV model with the volatilities and correlations following different long memory processes. Focusing on the volatility part, the structure of each element of Δ_t is similar to the fractionally integrated GARCH (FIGARCH) model suggested by Baillie et al. (1996). While the FIGARCH is based on the squared residuals, this new model works with the realized variance, which is the element of vecd(Δ_t). We can modify the FIDC model to specify a new multivariate long memory GARCH model by replacing P_t and Δ_t with the updated values $\hat{P}_t$ and $\hat{\Delta}_t$, respectively. Hence, by using our SCF, we can construct a 'FIDC with Filtered Covariance' (FIDC-FC) model, where $P_{t|t-1}$ and $\Delta_{t|t-1}$ in (19) can be calculated recursively using $\hat{P}_t$ and $\hat{\Delta}_t$ to substitute for P_t and Δ_t.

For the estimation of the FIDC models, we fix all the pre-sample values of Δ_t and P_t for $t = 0, -1, -2, \cdots$ at the vector of the unconditional variance and the unconditional correlation matrix, as suggested by Baillie et al. (1996) and Bollerslev and Mikkelsen (1996). They also truncated the infinite expansion of $(1 - aL)(1 - L)^d$ at $k = 1000$.

4.2 A threshold model

Researchers have observed that market variability reacts differently to the rise and fall of financial markets. This kind of asymmetry is likely to exist in correlations as well. Extending Tse and Tsui's (2002) varying-correlation multivariate GARCH model to have threshold nonlinearity (Tong, 1980; Tong, 1983; Tong, 1990, p.116) in the first and second conditional moments, Kwan, Li and Ng (2007) proposed a threshold model which incorporates (i) a threshold AR model for the conditional mean, (ii) a threshold GARCH model for the conditional variance, and (iii) a threshold model for the conditional correlation.

Let n be the delay parameter and r_{t-n} be a real-valued threshold variable. Now, we purpose the Threshold DC (TDC) model, given by equations (4), (13) with

$$\mu_t = \begin{cases} \tau_0^{(1)} + \sum_{k=1}^{p_1} \tau_k^{(1)} \circ y_{t-k}, & r_{t-n} \leq l, \\ \tau_0^{(2)} + \sum_{k=1}^{p_2} \tau_k^{(2)} \circ y_{t-k}, & r_{t-n} > l, \end{cases} \tag{20}$$

for the conditional mean, and

$$\mathrm{vecd}(\Delta_{t|t-1}) = \begin{cases} \kappa^{(1)} + \gamma^{(1)} \circ \mathrm{vecd}(\Delta_{t-1}) + \delta^{(1)} \circ \mathrm{vecd}(\Delta_{t-1|t-2}), & r_{t-n} \leq l, \\ \kappa^{(2)} + \gamma^{(2)} \circ \mathrm{vecd}(\Delta_{t-1}) + \delta^{(2)} \circ \mathrm{vecd}(\Delta_{t-1|t-2}), & r_{t-n} > l, \end{cases} \tag{21}$$

with

$$P_{t|t-1} = \begin{cases} S^{(1)}(1-\alpha^{(1)}-\beta^{(1)}) + \alpha^{(1)} P_{t-1} + \beta^{(1)} P_{t-1|t-2}, & r_{t-n} \leq l, \\ S^{(2)}(1-\alpha^{(2)}-\beta^{(2)}) + \alpha^{(2)} P_{t-1} + \beta^{(2)} P_{t-1|t-2}, & r_{t-n} > l, \end{cases} \tag{22}$$

for the conditional covariance, $C_{t|t-1}$. Depending on the regime i ($i = 1, 2$), $\phi_k^{(i)}$, $\kappa^{(i)}$, $\gamma^{(i)}$ and $\delta^{(i)}$ are m-dimensional vectors of parameters, $S^{(i)}$ is the unconditional correlation matrix, and $\alpha^{(i)}$ and $\beta^{(i)}$ are scalar parameters. Although this is a two-regime specification, it is possible to extend it to multiple-regimes as in Kwan, Li and Ng (2007) and So and Yip (2009). Following the categorization in Section 3, we also have three versions of such TDC models. Version one is defined when the realized covariance is available and is taken as C_t. In this case, it is known as the TDC-RC model and its likelihood function has the form given in (16). The second version is defined when C_t is a latent variable, where a new multivariate stochastic volatility model is formed. The third category is obtained when P_t and Δ_t in (21) and (22) are replaced by their filtered estimates, $\hat{P}_t$ and $\hat{\Delta}_t$ generated from the SCF. The likelihood function is then obtained as a product of multivariate t densities. Based on our notational convention, it is called the TDCCt-FC model. The setup in (21) and (22) differs from the formulation of Kwan, Li and Ng (2007) in that instead of using the squared returns and the sample correlation matrices of recent observations, $\hat{\Delta}_{t-1}$ and $\hat{P}_{t-1}$ are used instead. We believe that our version is more reasonable because the filtered estimates, $\hat{\Delta}_{t-1}$ and $\hat{P}_{t-1}$, are generated from the the best estimator, $\hat{C}_{t-1}$, of C_{t-1}.

Kwan, Li and Ng (2007) worked with the method of Tsay (1989) to estimate the delay parameter d, while they suggested a technique for estimating the threshold value l. Instead of their approach, So, Li and Lam (2002) and So and Choi (2009) employed the convenience of setting $d = 1$ and $l = 0$. Furthermore, So and Choi (2009) specified the threshold variable as the first element of y_t, i.e., $r_t = y_{1t}$. Knowing the estimation techniques for l and d, we may then employ ML/QML methods for estimating the TDC-RC and TDCCt-FC models, as will be discussed in the next section.

4.3 A portfolio index model

Let y_t be the vector of returns of financial assets following the model in (2) and (3). The return of the portfolio consisting of m assets is denoted by

$$y_{pt} = w_t' y_t = w_t' \mu_t + w_t' C_t^{1/2} z_t, \tag{23}$$

where $z_t \sim N(0, I_m)$ and $w_t = (w_{1t}, \ldots, w_{mt})'$ are the portfolio weights, such that $\sum_{i=1}^{m} w_{it} = 1$. The weight vector w_t is predetermined. For simplicity, assume that $w_t = w$ (constant) in this section, though it is straightforward to consider time-varying weights, and we will do so in the empirical analysis. Note that the stochastic variance of the portfolio return is given by $c_{pt} = Var(y_{pt}|\Im_{t-1}, C_t) = w'C_t w$. Deriving from (2) and (3), we have

$$y_{pt}|C_t, \Im_{t-1} \sim N(\mu_{pt}, c_{pt}),$$

where $\mu_{pt} = E(y_{pt}|\Im_{t-1}) = w'\mu_t$ is the conditional mean of the portfolio return, assuming $E(y_{pt}|C_t, \Im_{t-1}) = E(y_{pt}|\Im_{t-1})$ as in the original setting. According to (5), we also have the conditional mean of c_{pt} and the conditional variance of y_{pt} given $\Im_{t-1}$ as

$$\omega_{pt} = Var(y_{pt}|\Im_{t-1}) = w'E(C_t|\Im_{t-1})w = w'C_{t|t-1}w.$$

When C_t is available, for example when C_t is the realized covariance at time t, we can state that y_t is normally distributed with mean μ_{pt} and variance c_{pt}, conditional on the past information and the stochastic variance c_{pt}. Instead of specifying $C_{t|t-1}$ as before, we introduce an approach which models the conditional variance of the portfolio, ω_{pt}, as follows

$$\omega_{pt} = \kappa_p + w'(\Gamma_p \circ C_{t-1})w + \delta_p \omega_{p,t-1}, \tag{24}$$

where κ_p and δ_p are scalar parameters, and Γ_p is the $m \times m$ matrix of parameters. This is obtained using an idea similar to that suggested by Asai and McAleer (2008). Equation (24) models ω_{pt} explicitly using the information of C_{t-1}, and it is called the 'Portfolio Index DC' model. Since C_t is available as the realized covariance, ω_{pt} can be calculated by the recursion in (24) and the likelihood function of the portfolio returns and covariance, y_{pt} and c_{pt}, follows (16).

By the property of the multivariate t distribution in equation (5), it is straightforward to show that conditional on $\Im_{t-1}$,

$$y_{pt} = \mu_{pt} + \omega_{pt}^{1/2} \eta_{pt},$$

where η_{pt} follows the standardized t distribution with k degrees of freedom. When C_t is unavailable, the conditional distribution of y_{pt} given $\Im_{t-1}$ is the t distribution with the conditional mean and variance given by μ_{pt} and ω_{pt}, respectively. In this case, we may use $\frac{1}{m}\sum_{i=1}^{m}(y_{t-1-i} - \mu_{t-1-i})(y_{t-1-i} - \mu_{t-1-i})'$ instead of C_{t-1} in (24). When $m = 1$, our formulation reduces to the Portfolio Index GARCH specification of Asai and McAleer (2008).

5 Estimation

In this section, we show parameter estimation methods for the three categories in Section 3. Define $\theta = (\theta_1, \theta_2)'$ to be the vector of unknown parameters, where θ_1 is the parameter vector for the mean part, μ_t, and θ_2 is the covariance parameter characterizing the evolution of the dynamic correlation and variance in (14). Consider the first category where $C_{t|t-1}$ includes past information of an observable C_t. A likely example is when C_t represents the realized covariance at time t as in Section 3.2, giving us the DC-RC model. In this case, the information set $\Im_t$ contains the history of both y_t and C_t. Following (16), the likelihood function can be written as $L(\theta) = \prod_{t=1}^{T} f(y_t, C_t|\Im_{t-1}, \theta)$ $= L_1(\theta_1) \times L_2(\theta_2)$, where $L_1(\theta_1) = \prod_{t=1}^{T} f(y_t|\Im_{t-1}, C_t; \theta_1)$, $L_2(\theta_2) = \prod_{t=1}^{T} f(C_t|\Im_{t-1}; \theta_2)$,

$$\ln L_1(\theta_1) = \sum_{t=1}^{T}\left[-\frac{m}{2}\ln(2\pi) - \frac{1}{2}\ln|C_t| - \frac{1}{2}(y_t - \mu_t)'C_t^{-1}(y_t - \mu_t)\right],$$

$$\ln L_2(\theta_2) = \sum_{t=1}^{T}\left[-\frac{\nu m}{2}\ln 2 + \frac{\nu}{2}\ln|S_t| - \ln\Gamma_m\left(\frac{\nu}{2}\right) - \frac{\nu + m + 1}{2}\ln|C_t| - \frac{1}{2}\mathrm{tr}\left(S_t C_t^{-1}\right)\right],$$

$\Gamma_m(\nu/2) = \pi^{m(m-1)/4}\prod_{j=1}^{m}\Gamma[(\nu+1-j)/2]$, $S_t = (\nu - m - 1)C_{t|t-1}$ and $C_{t|t-1}$ is obtained from (14) as a function of θ_2. The log-likelihood function can be maximized over the parameters of the model to obtain the maximum likelihood (ML) estimator. Without assuming a Wishart distribution, it can be interpreted as the quasi-maximum likelihood (QML) estimator. The decomposition of $L(\theta)$ into $L_1(\theta_1)$ and $L_2(\theta_2)$ means that we can estimate θ_1 and θ_2 separately, thus making the estimation computationally feasible, especially in high dimension situations. If we further generalize (14) to be a function of μ_t as well, we will have $L(\theta) = \prod_{t=1}^{T} f(y_t, C_t|\Im_{t-1}, \theta) = L_1(\theta_1) \times L_2(\theta_1, \theta_2)$ because the density $f(C_t|\Im_{t-1})$ will also depend on θ_1. Since the first part of the likelihood is still a function of θ_1 only, computing the ML estimates can also be done effectively using the method of Song et al. (2005). As far as forecasting the covariance is concerned, Category 1 uses more information (both y_t and C_t) for modeling $C_{t|t-1}$ compared to the other two categories. This is one advantage of Category 1 that may produce a better one-step-ahead forecast for future C_t given $\Im_{t-1}$. This time, we do not need to use the estimate of C_t given $\Im_t$ in (12) because C_t is available.

For Category 2, when $C_{t|t-1}$ is a function of an unobservable C_t, we may classify it as the MSV models. In this case, the likelihood function is given by

$$L(\theta) = \prod_{t=1}^{T} f(y_t|\Im_{t-1};\theta) = \int \cdots \int \prod_{t=1}^{T} f(y_t|\Im_{t-1}, C_t; \theta_1) f(C_t|\Im_{t-1};\theta_2) dC_1 \cdots dC_T,$$

and is hard to obtain an analytical solution because of the high-dimensional integration. We can employ computationally intensive methods such as the Bayesian Markov chain Monte Carlo (MCMC) technique of Chib, Nardari and Shephard (2006), the Monte Carlo likelihood methods of Durbin and Koopman (1997), or the efficient importance sampling approach of Liesenfeld and Richard (2003). All such methods suggest estimates of C_t and forecasts of C_{T+k} for $k \geq 1$.

For Category 3, where $C_{t|t-1}$ is specified according to (17) in the DCCt-FC model, we can work with the conditional density of y_t in (5) to get the likelihood function $L(\theta) = \prod_{t=1}^{T} f(y_t|\Im_{t-1}, \theta)$, where

$$\begin{aligned} \ln L(\theta) = \sum_{t=1}^{T} \Bigg[& -\frac{m}{2} \ln(\pi(\nu + m - 1) + \ln \Gamma\left(\frac{\nu+1}{2}\right) - \ln \Gamma\left(\frac{\nu - m + 1}{2}\right) - \frac{1}{2} \ln |C_{t|t-1}| \\ & -\frac{\nu+1}{2} \ln \left(1 + \frac{(y_t - \mu_t)' C_{t|t-1}^{-1} (y_t - \mu_t)}{\nu - m - 1}\right) \Bigg]. \end{aligned} \tag{25}$$

In other words, we may conduct ML estimation via the multivariate t distribution The ML estimator becomes the quasi-ML (QML) estimator, if z_t is non-Gaussian and/or E_t^{-1} follows a non-Wishart distribution. The calculation of (25) is analogous to the prediction error decomposition using Kalamn filter (Harvey, 1989; 1993). Like using Kalman filter in state space models, we first obtain the filtered estimate $\hat{C}_{t-1}$ and $C_{t|t-1}$ sequentially for $t = 1, ..., T$ using the SCF in Section 2, and then determine the likelihood as a product of the multivariate t densities. Evaluating the likelihood function requires only one pass of our SCF. This modeling approach inherently allows the conditional density, $f(y_t|\Im_{t-1})$ to be fat-tailed, which matches well with empirical stylized facts in financial markets. By construction, $C_{T+1|T}$ forecasts C_{T+1} based on $\Im_T$. Risk analysis with financial time series is one application. In such risk analyses, Value-at-Risk (VaR) is also available. Taking account of the tail behavior of y_t, we should work with the multivariate t distribution rather than the multivariate normal distribution.

6 Monte Carlo Experiments

Let us now investigate the finite sample properties of the ML estimators through simulations for four DC models, namely, the DC-RC and FIDC-RC models for C_t and the DCCt-FC and TDCCt-FC models for y_t. The former two models are in Category 1, while the latter two are in Category 3.

6.1 DC-RC

With respect to the DC-RC model specified in (3), (13) and (14) with the correlation equation of $P_{t|t-1}$ simplified to (15), consider three data-generating processes (DGP). Specify the parameters of the dynamic correlations in (15) as

$$(\alpha, \beta, s_{12}, \nu) = \begin{cases} (0.10, 0.8, -0.3, 5) & \text{for DGP1,} \\ (0.10, 0.8, 0.3, 10) & \text{for DGP2,} \\ (0.05, 0.9, -0.3, 5) & \text{for DGP3,} \end{cases}$$

and set the parameters of the conditional variances, $\Delta_{t|t-1}$, in (14) as $(\kappa_1, \gamma_1, \delta_1) = (0.05, 0.15, 0.8)$ and $(\kappa_2, \gamma_2, \delta_2) = (0.02, 0.1, 0.88)$. Here, s_{12} is the $(1, 2)th$ element of S. With respect to the dynamic correlations, the degree of persistence, $\alpha + \beta$, is 0.9 for DGP1 and 2, while that is

0.95 for DGP3. Hence, the last case shows higher persistence in the correlation dynamic. The parameter of the unconditional correlation, s_{12}, is -0.3 for DGP1 and DGP3, while that is 0.3 for DGP2. The Wishart distribution has 5 degees of freedom for DGP1 and DGP3, but 10 for DGP2. Turning to the conditional variance, the second component indicates higher persistence in the variance process than the first component does. These two sets are common in the three DGPs. We consider a sample size of $T = 500$ with 1000 replications.

Table 1 shows the sample means, standard deviations and root mean squared errors (RMSE) of the ML estimators. Small biases and standard deviations are observed in almost all parameters. The RMSEs and the corresponding standard deviations are close, indicating that the biases are negligible. The only noticeable exception is the small downward bias in β. Such a bias is expected to disappear with a larger sample size, as shown by the Monte Carlo experiments with univariate and multivariate GRACH models (e.g., Lumsdaine, 1995; and Bollerslev and Wooldridge, 1992).

6.2 FIDC-RC

We now turn to the long memory model, the FIDC-RC ((3), (13) and (19)). Consider three DGPs and specify the parameters for the dynamic correlations as

$$(\phi, \beta, s_{12}, d_c) = \begin{cases} (0.00, 0.40, -0.30, 0.45) & \text{for DGP4 and DGP6,} \\ (0.95, 0.80, -0.30, 0.20) & \text{for DGP5,} \end{cases}$$

and parameters for the variances of

$$\{(\kappa_1, \lambda_1, \delta_1), (\kappa_2, \lambda_2, \delta_2), d_v\} = \begin{cases} \{(0.01, 0.0, 0.2), (0.002, 0.0, 0.4), 0.45\} & \text{for DGP4 and DGP5,} \\ \{(0.01, 0.7, 0.5), (0.002, 0.9, 0.8), 0.20\} & \text{for DGP6.} \end{cases}$$

Set $\nu = 6$ for the three DGPs. All the dynamic variances and correlations are stationary long-memory processes. The pair of long-range dependencies, (d_c, d_v), controls the long memory characteristics of the DGPs. With respect to DGP5, the variance process has a longer memory than the correlation process ($d_v = 0.45$ versus $d_c = 0.2$). On the other hand, the variance process of

DGP6 has a shorter memory than the correlation process ($d_v = 0.2$ versus $d_c = 0.45$). For DGP4, the long-range dependencies for the variance and correlation processes are the same. We consider a sample size of $T = 1500$ with 1000 replications.

Table 2 presents the sample means, standard deviations, and root mean squared errors of the ML estimators for the FIDC-RC model. As in the DC-RC simulations, all biases are within one standard deviation of the estimators. Most standard deviations are within 0.1, implying that the standard errors of most estimates are within one decimal place. The long memory parameters, d_v and d_c, are well estimated, especially when the long memory effect is strong (or the long memory parameters are large). For the sample size of $T = 1500$, the results show that the biases are quite close to zero and small biases are observed in s_{12} and δ_1. With respect to δ_1, the bias becomes smaller when δ_1 is large.

6.3 DCCt-FC

The next model to consider is based on the filtered estimates, i.e. the DCCt-FC ((3) and (17)). Let us specify the parameters for the conditional correlations as

$$(\alpha, \beta, s_{12}, \nu) = \begin{cases} (0.4, 0.5, -0.3, 5) & \text{for DGP7,} \\ (0.6, 0.3, 0.3, 5) & \text{for DGP8,} \\ (0.6, 0.3, -0.3, 7) & \text{for DGP9,} \end{cases}$$

in the dynamic correlations, and set the parameters for the conditional variances as $(\kappa_1, \gamma_1, \delta_1) = (0.05, 0.45, 0.5)$ and $(\kappa_2, \gamma_2, \delta_2) = (0.02, 0.3, 0.68)$. Again, s_{12} is the $(1, 2)$ element of S. With respect to the dynamic correlations, the degree of persistence is 0.9 for all three DGPs. The parameter of the unconditional correlation, S, is -0.3 for DGP7 and DGP9, but 0.3 for DGP8. The Wishart distribution has 5 degrees of freedom for DGP7 and DGP8, but 7 for DGP9. The conditional variance resembles a GARCH process. As discussed in Section 3, we have

$$\Delta_{11,t} = \kappa_1 + \gamma_1^*(y_{1t} - \mu_{1t})^2 + \delta_1^*\Delta_{11,t-1},$$

$$\Delta_{22,t} = \kappa_2 + \gamma_2^*(y_{2t} - \mu_{2t})^2 + \delta_2^*\Delta_{22,t-1},$$

where $\gamma_i^* = \gamma_i(\nu-2)^{-1}$ and $\delta_i^* = \gamma_i + \delta_i - \gamma_i^*$ for the bivariate case. For the conditional variance, $(\gamma_1^*, \delta_1^*) = (0.15, 0.8)$ and $(\gamma_2^*, \delta_2^*) = (0.1, 0.88)$ for $\nu = 5$, and $(\gamma_1^*, \delta_1^*) = (0.09, 0.86)$ and $(\gamma_2^*, \delta_2^*) = (0.06, 0.92)$ for $\nu = 7$. The parameter settings imply that the second component, $\Delta_{22,t}$, has higher persistence ($\gamma_2^* + \delta_2^* = 0.98$) in the variance process than the first component does ($\gamma_1^* + \delta_1^* = 0.95$). Furthermore, the contributions of γ_i^* are relatively small for the case of $\nu = 7$. We consider a sample size of $T = 1000$ with 1000 replications.

Table 3 reports the sample means, standard deviations, and root mean squared errors of the ML estimators for the DCCt-FC model. For DGP7 and DGP8, in which $\nu = 5$, there is no major difference between the two cases in terms of the biases and root mean squared errors. In both cases, all biases are negligibly small and all RMSEs are small relative to the true values. With DGP9, the biases are predominantly larger than with DGP7 or DGP8, especially in κ_1, κ_2 and s_{12}. It seems that when γ_i^* is small, the bias in γ_2 increases. All biases in the three DGPs tend to zero when the sample size is increased to $T = 2000$ (results are available on request).

6.4 TDCCt

Regarding the TDCCt model ((4), (13) and (20)-(22)), we specify the parameters for the experiment as

Regime	Var.	$\tau_{0i}^{(j)}$	$\tau_{1i}^{(j)}$	$\kappa_i^{(j)}$	$\gamma_i^{(j)}$	$\delta_i^{(j)}$	$s_{12}^{(j)}$	$\alpha^{(j)}$	$\beta^{(j)}$	k
$j = 1$	y_{1t}	0.0	-0.1	0.05	0.05	0.90	-0.50	0.15	0.80	5.00
	y_{2t}	0.0	0.0	0.02	0.10	0.88				
$j = 2$	y_{1t}	0.0	0.1	0.02	0.20	0.78	-0.20	0.05	0.85	
	y_{2t}	0.0	-0.1	0.05	0.15	0.80				

where y_{it} denotes the ith component of y_t. The threshold variable r_t is selected to be y_{1t}. The number of regimes is fixed at two, the delay $n = 1$ and the threshold value of $l = 0$. In other words, whether or not $y_{it-1} > 0$ determines the regime. The persistence of the conditional variance for y_{it} is $\gamma_i^{(j)} + \delta_i^{(j)} = 0.95$ or 0.98, depending on the regime j. Regarding the persistence of the conditional correlation process $(\alpha^{(j)} + \beta^{(j)})$, regime 1 ($j = 1$) has higher persistence than regime 2 does. The degrees-of-freedom parameter for the Wishart distribution is ν=6, indicating that the conditional distribution of y_t is the multivariate t distribution with the degrees-of-freedom parameter k=5 $(\nu - m + 1)$. As before, the sample size is $T = 500$, and we conduct 1000 replications.

Table 4 shows the sample means, standard deviations, and root mean squared errors of the ML estimators for the TDCCt model. As the sample size is not very large, the results shows a small downward bias in $\gamma_i^{(j)}$, $\delta_i^{(j)}$ and $\beta^{(j)}$. The biases in $\gamma_i^{(j)}$ and $\delta_i^{(j)}$ make the biases in $\kappa_i^{(j)}$ bigger (see Lumsdaine (1995)). Even for the relatively small sample size of $T = 500$, the standard deviations of the estimators are mostly less than 0.1, indicating precise estimation in terms of having small standard errors. Exceptions are $\kappa_2^{(1)}$, $s_{12}^{(1)}$, $s_{12}^{(2)}$ and k, where the biases are still acceptable. The statistical precision should improve with a larger sample size like $T = 1000$, corresponding to roughly four years of daily observations in a financial market. Note that the number of parameters is almost double compared to the DCCt specification. Hence, we need to be careful about the sample size when reliable statistical tests are required in empirical analysis.

7 Empirical Examples

7.1 A bivariate FIDC model for realized covariance of exchange rates returns

In this subsection, an empirical analysis is conducted applying the bivariate DC-RC and FIDC-RC models to data on the spot exchange rates for the U.S. dollar, the Japanese yen, and the Singapore dollar. The raw data consists of all interbank JPY/USD and SGD/USD bid/ask quotes from

August 1, 2001 through November 30, 2007, obtained from Olsen & Associates. The return series are constructed as follows. We first calculate fifteen-minute prices from the linearly interpolated logarithmic average of the bid and ask quotes. We then determine the fifteen-minute returns as the first difference of the logarithmic prices. We exclude all Saturday and Sunday returns to avoid modeling weekend effects. We then obtain the realized covariance matrices by the method of Hayashi and Yoshida (2005)[3]. As a result, we have a series of realized covariances covering $T = 1644$ days.

As a preliminary analysis, consider the case of $m = 1$ with two univariate models for the realized variance, c_t. One is a GARCH-type model based on the realized variance (RV-GARCH), which is the univariate version of (3) and (14),

$$c_t = (\nu - 2)e_t c_{t|t-1}, \quad e_t^{-1} \sim \chi^2(\nu),$$
$$c_{t|t-1} = \kappa + \gamma c_{t-1} + \delta c_{t-1|t-2}.$$

The other is a RV-FIGARCH model, which is the univariate version of (3) and (19),

$$c_t = (\nu - 2)e_t c_{t|t-1}, \quad e_t^{-1} \sim \chi^2(\nu),$$
$$c_{t|t-1} = \kappa + \delta c_{t-1|t-2} + (1 - \delta L - (1 - \lambda L)(1 - L)^d)c_t,$$

where L is the lag operator as before.

Table 5 shows the ML estimates of the RV-GARCH and RV-FIGARCH models for the JPY/USD and SGD/USD rates. All estimates are significant at the five percent level. With respect to the RV-GARCH model, the estimates of ν are 8.42 and 8.13 for JPY/USD and SGD/USD,

[3]For more frequent time series such as one-minute or five-minute data, we need to remove the microstructure noise by the methods proposed by Bandi and Russel (2005), Zhang (2005), Shephard (2006), Voev and Lunde (2007), or Griffin and Oomen (2007). These are multivariate extensions of the univariate approach. For instance, Barndorff-Nielsen et al. (2008) showed how to use realized kernels to carry out efficient inference on the ex post variation of asset prices, while Voev and Lunde (2007) formulated its multivariate extension.

respectively. Like a typical result from a GARCH model, the estimates of $\gamma + \delta$, which are about 0.95, are larger than 0.9 and quite close to one. Models with fractional integration can thus be a promising alternative to the RV-GARCH model. For the RV-FIGARCH model fitting, the estimates of d are 0.2978 and 0.1661 for JPY/USD and SGD/USD, respectively. The results indicate that the two processes of the realized variance exhibit clear long memory effect and are stationary. Incorporating the long memory property does not seem to affect the implied distribution of c_t, because the estimates of ν are close to those of the RV-GARCH. For the both series, AIC and BIC select the RV-FIGARCH model. We consider the likelihood ratio test for the null of RV-GARCH against the RV-FIGARCH. The test statistic rejects the null at the five percent significance level, implying the we prefer RV-FIGARCH to RV-GARCH with both series.

Table 6 gives the ML estimates for the DC-RC model in (3), (13), (14) and (15), i.e.

$$\begin{aligned}
C_t &= C_{t|t-1}^{1/2} \{(\nu - m - 1)E_t\} C_{t|t-1}^{1/2}, \quad E_t^{-1} \sim W(I_m, \nu),\ C_{t|t-1} = \Delta_{t|t-1}^{1/2} P_{t|t-1} \Delta_{t|t-1}^{1/2}, \\
P_{t|t-1} &= S(1 - \alpha - \beta) + \alpha P_{t-1} + \beta P_{t-1|t-2}, \\
\mathrm{vecd}(\Delta_{t|t-1}) &= \kappa + \gamma \circ \mathrm{vecd}(\Delta_{t-1}) + \delta \circ \mathrm{vecd}(\Delta_{t-1|t-2}).
\end{aligned}$$

All estimates are significant at the five percent level. In particular, the estimates for κ_i, γ_i and δ_i are close to the values given in Table 5. After accounting for the dynamic correlation, the estimate of the degrees-of-freedom parameter, ν, becomes 11.71, which is higher than the two values in Table 5. The estimate of s_{12} is 0.1802, implying that the unconditional correlation is positive and significant. The estimate of $\alpha + \beta$ is 0.98, revealing that the realized correlation process is more persistent than the two realized variance processes. Given the high persistence in both realized variances and realized correlation, we explore if the dynamic process of C_t is more suitably described by a fractionally integrated process. Table 7 presents the ML estimates of the

FIDC-RC model in (3), (13) and (19), i.e.

$$C_t = C_{t|t-1}^{1/2} \{(\nu - m - 1)E_t\} C_{t|t-1}^{1/2}, \quad E_t^{-1} \sim W(I_m, \nu),\ C_{t|t-1} = \Delta_{t|t-1}^{1/2} P_{t|t-1} \Delta_{t|t-1}^{1/2},$$

$$\mathrm{vech}(P_{t|t-1}) = \mathrm{vech}(S) + \beta \mathrm{vech}(P_{t-1|t-2} - S) + \left\{1 - \beta L - (1 - \phi L)(1 - L)^{d_c}\right\} \mathrm{vech}(P_t - S),$$

$$\mathrm{vecd}(\Delta_{t|t-1}) = \kappa + \delta \circ \mathrm{vecd}(\Delta_{t-1|t-2}) + \left\{I_m - \delta L - (I_m - \lambda L)(1 - L)^{d_v}\right\} \mathrm{vecd}(\Delta_t).$$

Compared with the case of $m = 1$ in separate modeling of the two realized variance processes in Table 5, both λ_i and δ_i are smaller than their univariate counterparts, whereas d_v (0.3699) is larger than both long-range dependence parameters for the JPY/USD (0.2978) and SGD/USD (0.1661) series. After capturing the dynamic correlation property, the long memory characteristic of the realized variance is more evident. Regarding the conditional variance, the estimate of the parameter for the common long-run dependence is 0.3699, implying that the processes are fractionally integrated. All the dynamic correlation parameters except ϕ are significant, indicating that the realized correlation follows an ARFIMA(0,d_c,1). The estimate of d_c is 0.4372, which is larger than d_c. This empirical finding is consistent with Table 6 that the realized correlation process is more persistent than the realized variance process. As the estimates of d_c and d_v are smaller than 0.5, the covariance structure is stationary. The estimate of ν is almost identical to the one in Table 6. In terms of information criteria, both AIC and BIC select the FIDC-RC model. In addition, the likelihood ratio test for the null of DC-RC against the FIDC-RC rejects the null hypothesis at the five percent level.

By virtue of the fact that $E[C_t|\Im_{t-1}] = C_{t|t-1}$, the conditional covariance matrix, $C_{t|t-1}$, can be used for one-step-ahead prediction of C_t. Based on the parameter estimates obtained from the first $T = 1300$ observations, predicted values can be calculated for the last 300 observations, i.e. $C_{T+1|T}$,..., $C_{T+300|T+299}$. Figure 1 shows the forecasts of the conditional correlation based on the FIDC-RC model together with the realized correlations. The FIDC-RC model captures the long term trend well by the nature of fractional integration, but it may take time for the model to accommodate the effects of big and sudden changes.

Consider now a portfolio of the two exchange rate returns. From the realized covariance matrices $C_t = \{c_{ij,t}\}$, the realized portfolio variance is given by

$$\sigma_{p,t}^2 = w_1^2 c_{11,t} + (1 - w_1)^2 c_{22,t} + 2w_1(1 - w_1)c_{12,t},$$

where w_1 is the weight for JPY/USD and $(1 - w_1)$ is for SGD/USD. The minimum variance portfolio is given by

$$w_{1t}^{\text{opt}} = \frac{c_{22,t} - c_{12,t}}{c_{11,t} + c_{22,t} - 2c_{12,t}},$$

while the one-step-ahead forecast of w_{1t}^{opt} is given by

$$w_{1t}^{f} = \frac{\omega_{22,t} - \omega_{12,t}}{\omega_{11,t} + \omega_{22,t} - 2\omega_{12,t}},$$

which is calculated using $C_{t|t-1} = \{\omega_{ij,t}\}$ in place of C_t. Figure 2 shows the forecasts and the realized value of w_{1t}^{opt}, demonstrating that the realized values fluctuate strongly but the forecasts capture the trend. To compare the performance of the predicted weights, we calculate the ratio of the portfolio risk, defined by

$$\sqrt{\frac{\sigma_{p,t}^2(w_{1t}^f)}{\sigma_{p,t}^2(w_{1t}^{\text{opt}})}},$$

where $\sigma_{p,t}^2(w_{1t}^f)$ and $\sigma_{p,t}^2(w_{1t}^{\text{opt}})$ are the portfolio variance based on the weights w_{1t}^f and w_{1t}^{opt}, respectively. Figure 3 shows the ratio of the portfolio risk. It is always greater than one, as the risk based on the forecast is higher than that of the optimal value. The average of the ratio is 1.040, so the difference is very minor. The ratio exceeds 1.1 twenty-six times among 300 forecasts, while it exceeds 1.2 only eight times. The percentage difference between $\sqrt{\sigma_{p,t}^2(w_{1t}^f)}$ and $\sqrt{\sigma_{p,t}^2(w_{1t}^{\text{opt}})}$ is within 10% for most of the predicted values. Hence, the FIDC-RC model produces minimum variance portfolios whose volatility approximates the true minimum portfolio risk very well.

7.2 A trivariate threshold DCCt-FC model of stock returns

This subsection presents an empirical analysis with the trivariate DCCt-FC and TDCCt-FC models. The datasets consist of daily closing values of the Nikkei 225, Hang Seng, and Singapore Straits Times (SST) stock indices. The return series are calculated as $y_{it} = 100 \times (\ln P_{it} - \ln P_{i,t-1})$, where P_{it} is the closing price of the ith index. The sample period is July 9, 2001 through August 28, 2009, giving $T = 2000$ observations. Table 8 presents the ML estimates for the DCCt-FC model in (17) with μ_t given by $\tau_0 + \tau_1 y_{t-1}$. Denote the ith component of τ_0 and τ_1 by τ_{0i} and τ_{1i}, respectively. For the mean equation, all the constant terms, τ_{0i}, are positive, while all the AR(1) coefficients, τ_{1i}, are negative. The constant term for the Hang Seng and SST are significant, while the AR(1) coefficient for the SST is significant. The magnitude of τ_{1i}'s are smaller than 0.1, implying weakly negative correlations. With respect to the conditional variance, the persistence of the conditional variance, $\Delta_{t|t-1}$, is estimated to be $\gamma_i + \delta_i > 0.99$. Regarding the conditional correlations, the persistence estimate $\alpha + \beta$ is 0.9724. The estimate of the degrees-of-freedom parameter, ν, is 8.89. All of these results are typical for dynamic correlation model fitting.

Table 9 shows the ML estimates for the TDCCt-FC model in Section 4.2. The lag-one Nikkei return is chosen as the threshold variable. In other words, regime 1 is defined by the negative sign of the yesterday's Nikkei return, and regime 2 is by the positive sign. While all $\tau_{1i}^{(1)}$ for regime 1 are negative, some $\tau_{1i}^{(2)}$ for regime 2 are positive. This implies evidence of asymmetry in the mean equation. Other evidence of the need for a threshold DC model lies in the estimates of $\gamma_i^{(2)} + \delta_i^{(2)}$ which are close to 0.93, whereas those of $\gamma_i^{(1)} + \delta_i^{(1)}$ exceed one. However, the averages of $\gamma_i^{(1)} + \delta_i^{(1)}$ and $\gamma_i^{(2)} + \delta_i^{(2)}$ are 0.99, which is similar to the estimates of $\gamma_i + \delta_i$ of the DCCt-FC model in Table 8. For the dynamic correlation, the estimate of $\alpha^{(1)} + \beta^{(1)}$ exceeds one, whereas that of $\alpha^{(2)} + \beta^{(2)}$ is 0.9216. The average is about 0.96, and similar to the estimate of $\alpha + \beta$ with the

DCCt-FC model. The estimate of the degrees-of-freedom ν is 9.12, which is consistent with the corresponding estimate using the DCCt-FC model. In addition to the above empirical evidence supporting the use of the TDCCt-FC over the DCCt-FC, the LR test for the null hypothesis of the DCCt-FC model also favors the TDCCt-FC model. The AIC selects the TDCCt-FC model, while the BIC chooses the DCCt-FC.

To further compare the DCCt-FC and the TDCCt-FC models, we examine the Value-at-Risk (VaR) thresholds of the h-step-ahead predictions ($h = 1, 5, 10, 20$) for the equally weighted portfolio with $w_t = (1/3, 1/3, 1/3)'$. Fixing the sample size at T=1800, we re-estimate the two models and obtain estimates of the conditional variance, $Var(w_{T+h}{}' y_{T+h}|\Im_T) = w_{T+h}{}' Var(y_{T+h}|\Im_T)w_{T+h}$, for the last 200 days. We determine $C_{T+1|T}$ as $Var(y_{T+1}|\Im_T)$ for one-step-ahead prediction. Based on equation (8), we produce estimates of $Var(y_{T+h}|\Im_T)$, $h = 5, 10, 20$, by obtaining h-step-ahead forecasts of $E(C_{T+h}|\Im_T)$ and $Var(\mu_{t+h}|\Im_t)$ with the Monte Carlo methods described in equations (6) and (7), respectively. We calculate the VaR thresholds, accommodating (i) a standardized t distribution and (ii) a filtered historical simulation (FHS) approach. The former is derived by our model specification, while the latter is an effective method for predicting VaR thresholds (see Kuester et al. (2006) for some studies discussing the FHS approach). In short, the FHS approach estimates the empirical distribution of the standardized returns and then obtains the 100pth empirical percentiles as the 100p% VaR thresholds. In our analysis, we adopt a rolling-sample method. Each time, we estimate the models with 1800 observations and estimate the VaR for various combinations h and p. The models are re-estimated when the 1800 observations in the estimation window are updated.

To assess the estimated VaR thresholds, we conduct the unconditional coverage and independence tests developed by Christoffersen (1998). Consider the “hit sequence” of VaR exceedences, which takes one if the return is smaller than (or exceeds) the VaR threshold, and takes zero if the VaR is not exceeded. According to Christoffersen (1998), if we can predict the hit sequence, then that information may help to construct a better model. For a good model, the hit sequence should therefore be unpredictable, and it should follow an independent Bernoulli distribution. The unconditional coverage test is the likelihood test for examining the null hypothesis that the Bernoulli probability is p. The likelihood ratio test of independence is constructed against a first-order Markov alternative. Both tests have an asymptotic $\chi^2(1)$ distribution. Table 10 presents the proportion of VaR exceedences and test results for the DCCt-FC and TDCCt-FC models with $h = 1$, 5, 10 and 20 with the FHS approach and the standardized t distribution. We compute each proportion of exceedence using the sample size of $200 - h + 1$ for h-step-ahead prediction. For the DCCt-FC model, the percentage exceedences are far from the nominal values of $p = 1\%$ or 5%, leading to rejection in the unconditional coverage tests in many cases. With respect to the TDCCt-FC model, the tests do not reject the null hypothesis for the 5% and 1% VaR thresholds, indicating that the estimated VaR thresholds are satisfactory for both the standardized t and the FHS approaches. The FHS approach does not seem to produce any benefit over the implied standardized t distribution from our DCCt-FC and TDCCt-FC models with respect to VaR prediction.

8 Conclusion

In this paper, we introduce a new class of stochastic covariance models accommodating the dynamic correlation. We classify the models into three categories; (i) the conditional covariance $C_{t|t-1}$ includes past information of C_t, which is observable; (ii) the conditional covariance $C_{t|t-1}$ includes past information of C_t, which is a latent variable; (iii) the conditional covariance $C_{t|t-1}$ excludes past information of C_t, where past values of C_t are replaced by their filtered estimates,

$\hat{C}_{t-s}(s \geq 1)$. We also develop a stochastic covariance filter for filtering and prediction of covariances. We suggest three kinds of extensions of the new models, namely, the fractionally integrated model for correlations, the threshold model, and the portfolio index model. Our methodology facilitates modeling stochastic covariance no matter whether the covariance matrix is observable or latent. Empirical examples with FIDC-RC and TDCC-FC show that the new specifications perform well in predicting the minimum variance and in estimating portfolio VaR. Further research on how to enrich these stochastic variance models for incorporating more time series properties of the dynamic correlation is fruitful.

References

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

Asai, M. and M. McAleer (2008), "A Portfolio Index GARCH Model", *International Journal of Forecasting*, **24**, 449–461.

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

Asai, M., M. McAleer and M. Medeiros (2008), "Modeling and Forecasting Daily Volatility with Noisy Realized Volatility Measures", Unpublished paper, Soka University.

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

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

Barnard, J., R. McCulloch and X.L. Meng (2000), "Modeling Covariance Matrices in Terms of Standard Deviations and Correlations, With Applications to Shrinkage", *Statistica Sinica*, **10**, 1281–1311.

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

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

Bandi, F.M. and J.R. Russell (2005), "Realized Covariation, Realized Beta and Microstructure Noise", Unpublished Paper, Graduate School of Business, University of Chicago.

Bandi, F.M., J.R. Russell and Y. Zhu (2008), "Using High-Frequency Data in Dynamic Portfolio Choice", *Econometric Reviews*, **27**, 163-198.

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

Bollerslev, T. and J.M. Wooldridge (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances", *Econometric Reviews*, **11**, 143–172.

Chib, S., F. Nardari and N. Shephard (2006), "Analysis of High Dimensional Multivariate Stochastic Volatility Models", *Journal of Econometrics*, **134**, 341-371.

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

Corsi, F. (2004), "A Simple Long Memory Model of Realized Volatility". Working Paper, University of Southern Switzerland.

Daniels, M.J. and R.E. Kass (1999), "Nonconjugate Bayesian Estimation of Covariance Matrices and Its Use in Hierarchical Models", *Journal of the American Statistical Association*, **94**, 1254–1263.

Daniels, M.J. and M. Pourahmadi (2002), "Bayesian Analysis of Covariance Matrices and Dynamic Models for Longitudinal Data", *Biometrika*, **89**, 553–566.

Ding, Z. and R.F. Engle (2001), "Large Scale Conditional Covariance Matrix Modeling, estimation and Testing", *Academia Economic Papers*, **1**, 83–106.

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

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

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

Engle, R.F. and K. Sheppard (2001), "Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH," NBER Working Papers 8554, National Bureau of Economic Research, Inc.

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

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

Griffin, J.E. and R.C.A. Oomen (2006), "Covariance Measurement in the Presence of Non-synchronous Trading Market Microstructure Noise", Unpublished Paper, Department of Statistics, University of Warwick.

Harvey, A.C. (1989), *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A.C. (1993), *Time Series Models*. Cambridge: MIT Press.

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

Hayashi, T. and N. Yoshida (2005), "On Covariance Estimation of Non-Synchronously Observed Diffusion Processes", *Bernoulli*, **11**, 359–379.

Hosking, J.R.M. (1981), "Fractional Differencing", *Biometrika*, **68**, 165–176.

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

Kitagawa, G. (1996), "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Model", *Journal of Computational and Graphical Statistics*, **5**, 1-25.

Koopman, S.J. and E.H. Uspensky (2002), "The Stochastic Volatility in Mean Model: Empirical Evidence from International Stock Markets", *Journal of Applied Econometrics*, **17**, 667–689.

Kuester, K., S. Mittnik and M.S. Paolella (2006), "Value-at-Risk Prediction: A Comparison of Alternative Strategies", *Journal of Financial Econometrics*, **4**, 53–89.

Kwan, W., W.K. Li and K. Ng (2007), "A Multivariate Threshold Varying Conditional Correlations Model", forthcoming in *Econometric Reviews*.

Leonard, T. and J.S.J. Hsu (1992), "Bayesian Inference for a Covariance Matrix", *The Annals of Statistics*, **20**, 1669-1696.

Liechty, J.C., M.W. Liechty and P. Müller (2004), "Bayesian Correlation Estimation", *Biometrika*, **91**, 1-14.

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

Lumsdaine, R.L., (1995), "Finite-Sample Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models: A Monte Carlo Investigation", *Journal of Business & Economics Statistics*, **13**, 1–10.

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

McAleer, M. and M. C. Medeiros (2008), " A Multiple Regime Smooth Transition Heterogeneous Autoregressive Model for Long Memory and Asymmetries", *Journal of Econometrics*, **147**, 104-119.

Philipov, A., and M.E. Glickman (2006a), "Multivariate Stochastic Volatility via Wishart Processes", *Journal of Business & Economic Statistics*, **24**, 313–328.

Philipov, A., and M.E. Glickman (2006b), "Factor Multivariate Stochastic Volatility via Wishart Processes", *Econometric Reviews*, **25**, 311–334.

Pitt, M., D. Chan and R. Kohn (2006), "Efficient Bayesian Inference for Gaussian Copula Regression Models", *Biometrika*, **93**, 537-554.

Pitt, M. and N. Shephard (1999a), "Time Varying Covariances: A Factor Stochastic Volatility Approach", In: Bernardo, J. M., Berger, J. O., David, A. P., Smith, A. F. M., eds. *Bayesian Statistics* **6**, Oxford University Press, pp. 547–70.

Pitt, M. K. and N. Shephard (1999b), “Filtering via simulation: auxiliary particle filter. *Journal of the American Statistical Association*, **94**, 590-599.

Pourahmadi, M. (2000), “Maximum Likelihood Estimation of Generalised Linear Models for Multivariate Normal Covariance Matrix”, *Biometrika*, **87**, 425-435.

Pourahmadi, M. (2007), “Cholesky Decompositions and Estimation of a Covariance Matrix: Orthogonality of Variance-Correlation Parameters”, *Biometrika*, **94**, 1006-1013.

Ren, L. and W. Polasek (2000), “A Multivariate GARCH Model for Exchange Rates in The Us, Germany and Japan”, Computing in Economics and Finance 2000, No.223, Society for Computational Economics.

Shephard, N. (2006), “Measuring Realized Covariance”, Unpublished Paper, Department of Economics, University of Oxford.

Smith, M. and R. Kohn (2002), “Parsimonious Covariance Matrix Estimation for Longitudinal Data”, *Journal of the American Statistical Association*, **97**, 1141–1153.

So, M.K.P. and C.Y. Choi (2008), “A Multivariate Threshold Stochastic Volatility Model”. *Mathematics and Computers in Simulation*, **79**, 306-317.

Tse, Y.K. and A.K.C. Tsui (2002), “A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations”, *Journal of Business & Economic Statistics*, **20**, 351–361.

Tong, H. (1983). *Threshold Models in Non-linear Time Series Analysis* (vol. 21 of Lecture Notes in Statistics). New York: Springer-Verlag.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach.* Oxford: Oxford University Press.

Tong, H. and K.S. Lim (1980). "Threshold Autoregression, Limit Cycles and Cyclical Data (With Discussion)". *Journal of the Royal Statistical Society, Series B*, **42**, 245-292.

Voev, V. (2007), "Dynamic Modelling of Large Dimensional Covariance Matrices", Working Paper, University of Konstanz.

Voev, V. and A. Lunde (2007), "Integrated Covariance Estimation Using High-Frequency Data in the Presence of Noise", *Journal of Financial Econometrics*, **5**, 68–104.

Wishart, J., (1928), "The Generalized Product Moment Distribution in Samples from A Multinomial Population", *Biometrika*, **20**, 32–52, and 424.

Wong, F., C.K. Carter and R. Kohn (2003), "Efficient Estimation of Covariance Selection Models", *Biometrika*, **90**, 809-830.

Zhang, L. (2005), "Estimating Covariation: Epps Effect and Microstructure Noise", Unpublished Paper, Department of Finance, University of Illinois, Chicago.

Table 1: Finite Sample Performance of the ML Estimator for the DCC-RC model with $T = 500$

Parameters	DGP1		DGP2		DGP3	
	True value		True value		True value	
κ_1	0.05	0.0524	0.05	0.0570	0.05	0.0526
		(0.0140)		(0.0216)		(0.0140)
		[0.0142]		[0.0227]		[0.0142]
γ_1	0.15	0.1488	0.15	0.1480	0.15	0.1489
		(0.0232)		(0.0247)		(0.0232)
		[0.0232]		[0.0248]		[0.0232]
δ_1	0.8	0.7982	0.8	0.7945	0.8	0.7980
		(0.0305)		(0.0374)		(0.0306)
		[0.0306]		[0.0378]		[0.0306]
κ_2	0.02	0.0223	0.02	0.0273	0.02	0.0223
		(0.0090)		(0.0163)		(0.0090)
		[0.0092]		[0.0178]		[0.0093]
γ_2	0.1	0.0996	0.1	0.0995	0.1	0.0997
		(0.0167)		(0.0189)		(0.0168)
		[0.0167]		[0.0189]		[0.0168]
δ_2	0.88	0.8775	0.88	0.8729	0.88	0.8773
		(0.0217)		(0.0282)		(0.0218)
		[0.0218]		[0.0291]		[0.0220]
α	0.1	0.1003	0.1	0.1000	0.05	0.0527
		(0.0320)		(0.0306)		(0.0256)
		[0.0320]		[0.0306]		[0.0258]
β	0.8	0.7643	0.8	0.7645	0.9	0.8375
		(0.1268)		(0.1154)		(0.1686)
		[0.1317]		[0.1207]		[0.1798]
s_{12}	-0.3	-0.3006	0.3	0.2997	-0.3	-0.3012
		(0.0383)		(0.0265)		(0.0422)
		[0.0383]		[0.0265]		[0.0422]
ν	5	5.0201	10	10.056	5	5.0190
		(0.1360)		(0.3312)		(0.1362)
		[0.1375]		[0.3360]		[0.1375]

Note: Standard deviations are in parentheses and root mean squared errors are in brackets.

Table 2: Finite Sample Performance of the ML Estimator for the FIDC-RC model with $T = 1500$

Parameters	DGP4		DGP5		DGP6	
	True value		True value		True value	
κ_1	0.01	0.0100	0.01	0.0099	0.01	0.0106
		(0.0058)		(0.0058)		(0.0024)
		[0.0058]		[0.0058]		[0.0025]
κ_2	0.002	0.0021	0.002	0.0021	0.002	0.0022
		(0.0012)		(0.0010)		(0.0007)
		[0.0012]		[0.0011]		[0.0008]
λ_1	0.0	-0.0145	0.0	-0.0157	0.7	0.6952
		(0.1273)		(0.1221)		(0.0554)
		[0.1281]		[0.1231]		[0.0556]
λ_2	0.0	-0.0025	0.0	-0.0017	0.9	0.8946
		(0.0564)		(0.0541)		(0.0291)
		[0.0565]		[0.0541]		[0.0296]
δ_1	0.2	0.1790	0.2	0.1767	0.5	0.4898
		(0.1422)		(0.1363)		(0.0627)
		[0.1437]		[0.1382]		[0.0635]
δ_2	0.4	0.3871	0.4	0.3861	0.8	0.7906
		(0.0709)		(0.0681)		(0.0453)
		[0.0720]		[0.0695]		[0.0462]
d_v	0.45	0.4400	0.45	0.4382	0.2	0.1918
		(0.0358)		(0.0351)		(0.0367)
		[0.0371]		[0.0371]		[0.0376]
ϕ	0.0	-0.0116	0.95	0.9498	0.0	-0.0107
		(0.0695)		(0.0139)		(0.0692)
		[0.0704]		[0.0139]		[0.0700]
β	0.4	0.3839	0.8	0.7820	0.4	0.3858
		(0.1019)		(0.0787)		(0.1087)
		[0.1032]		[0.0807]		[0.1147]
s_{12}	-0.3	-0.1882	-0.3	-0.2535	-0.3	-0.1840
		(0.1364)		(0.1598)		(0.1392)
		[0.1780]		[0.1665]		[0.1812]
d_c	0.45	0.4441	0.2	0.1788	0.45	0.4452
		(0.0554)		(0.0931)		(0.0560)
		[0.0558]		[0.0955]		[0.0563]
ν	6	6.0579	6	6.0630	6	6.0363
		(0.1092)		(0.1075)		(0.1087)
		[0.1236]		[0.1246]		[0.1147]

Note: Standard deviations are in parentheses and root mean squared errors are in brackets.

Table 3: Finite Sample Performance of the ML Estimator for the DCCt-FC model with $T = 1000$

Parameters	DGP7		DGP8		DGP9	
	True value		True value		True value	
κ_1	0.05	0.0560	0.05	0.0565	0.05	0.0654
		(0.0229)		(0.0241)		(0.0434)
		[0.0237]		[0.0250]		[0.0460]
γ_1	0.45	0.4581	0.45	0.4531	0.45	0.4665
		(0.1048)		(0.1002)		(0.1209)
		[0.1052]		[0.1003]		[0.1221]
δ_1	0.5	0.4847	0.5	0.4897	0.5	0.4870
		(0.1107)		(0.1062)		(0.1368)
		[0.1117]		[0.1067]		[0.1375]
κ_2	0.02	0.0240	0.02	0.0245	0.02	0.0304
		(0.0130)		(0.0146)		(0.0274)
		[0.0136]		[0.0153]		[0.0293]
γ_2	0.3	0.3001	0.3	0.2996	0.3	0.3552
		(0.0713)		(0.0716)		(0.0806)
		[0.0713]		[0.0716]		[0.0812]
δ_2	0.68	0.6732	0.68	0.6739	0.68	0.6672
		(0.0758)		(0.0754)		(0.0885)
		[0.0761]		[0.0756]		[0.0894]
α	0.4	0.4118	0.6	0.6083	0.6	0.6224
		(0.1149)		(0.1122)		(0.1205)
		[0.1155]		[0.1125]		[0.1226]
β	0.5	0.4642	0.3	0.2792	0.3	0.2732
		(0.1556)		(0.1316)		(0.1417)
		[0.1597]		[0.1332]		[0.1442]
s_{12}	-0.3	-0.2909	0.3	0.2963	-0.3	-0.2772
		(0.0587)		(0.0613)		(0.0691)
		[0.0593]		[0.0614]		[0.0727]
ν	5	5.1020	5	5.1147	7	7.1730
		(0.3750)		(0.3717)		(1.8663)
		[0.3886]		[0.3890]		[1.9387]

Note: Standard deviations are in parentheses and root mean squared errors are in brackets.

Table 4: Finite Sample Performance of ML Estimator for the Bivariate Threshold DCCt Model for $T = 500$

Parameters	Regime 1 ($j=1$)				Regime 2 ($j=2$)			
			y_{1t}	y_{2t}			y_{1t}	y_{2t}
	True value				True value			
$\tau_{0i}^{(j)}$	0.0	0.0	-0.0010	0.0012	0.0	0.0	-0.0023	0.0001
			(0.0578)	(0.0475)			(0.0537)	(0.0436)
			[0.0578]	[0.0475]			[0.0537]	[0.0436]
$\tau_{1i}^{(j)}$	-0.1	0.0	-0.1007	-0.0010	0.1	-0.1	0.0974	-0.0999
			(0.0800)	(0.0598)			(0.0841)	(0.0582)
			[0.0800]	[0.0598]			[0.0841]	[0.0582]
$\kappa_i^{(j)}$	0.05	0.02	0.0567	0.0462	0.02	0.05	0.0376	0.0580
			(0.0421)	(0.2013)			(0.0922)	(0.0512)
			[0.0427]	[0.2030]			[0.0939]	[0.0518]
$\gamma_i^{(j)}$	0.05	0.10	0.0434	0.0932	0.20	0.15	0.1890	0.1458
			(0.0322)	(0.0627)			(0.0660)	(0.0666)
			[0.0329]	[0.0631]			[0.0669]	[0.0667]
$\delta_i^{(j)}$	0.90	0.88	0.8913	0.8615	0.78	0.80	0.7697	0.7939
			(0.0787)	(0.0885)			(0.0893)	(0.0877)
			[0.0792]	[0.0904]			[0.0899]	[0.0879]
$s_{12}^{(j)}$	-0.50		-0.4442		-0.20		-0.1800	
			(0.5472)				(0.5358)	
			[0.5501]				[0.5361]	
$\alpha^{(j)}$	0.15		0.1500		0.05		0.0475	
			(0.0508)				(0.0291)	
			[0.0508]				[0.0292]	
$\beta^{(j)}$	0.80		0.7723		0.85		0.8446	
			(0.1048)				(0.0946)	
			[0.1084]				[0.0947]	
k	5.00				5.5008			
					(1.0799)			
					[1.1903]			

Note: Standard deviations are in parentheses and root mean squared errors are in brackets.

Table 5: ML Estimates for the RV-GARCH and RV-FIGARCH Models

Model	RV-GARCH		RV-FIGARCH	
Parameters	JPY/USD	SGD/USD	JPY/USD	SGD/USD
κ	0.0176	0.0075	0.0012	0.0019
	(0.0049)	(0.0017)	(0.0006)	(0.0007)
γ	0.2375	0.1317		
	(0.0301)	(0.0205)		
λ			0.9719	0.9457
			(0.0126)	(0.0161)
δ	0.7192	0.8072	0.9457	0.9117
	(0.0391)	(0.0320)	(0.0190)	(0.0243)
ν	8.4203	8.1304	8.7348	8.1818
	(0.2787)	(0.2736)	(0.2701)	(0.2737)
d			0.2978	0.1661
			(0.0318)	(0.031296)
LogLike	809.199	2511.83	835.432	2524.55
AIC	-1610.40	-5015.67	-1660.86	-5039.09
BIC	-1588.78	-4994.05	-1633.84	-5012.07

Note: Standard errors are in parentheses.

Table 6: ML Estimates for the Bivariate DC-RC Model

Parameter	JPY/USD	SGD/USD	Parameter	Corr.	Parameter	Common
κ_i	0.0178	0.0073	α	0.1305	ν	11.708
	(0.0039)	(0.0013)		(0.0307)		(0.2183)
γ_i	0.2296	0.1305	β	0.8522		
	(0.0234)	(0.0162)		(0.0365)	LogLike	6623.02
δ_i	0.7262	0.8052	s_{12}	0.1802	AIC	-13226.0
	(0.0306)	(0.0257)		(0.0646)	BIC	-13172.0

Note: Standard errors are in parentheses.

Table 7: ML Estimates for the Bivariate FIDC-RC Model

Parameter	JPY/USD	SGD/USD	Parameter	Corr.	Parameter	Common
κ_i	0.0108	0.0021	ϕ	0.2510	ν	11.793
	(0.0050)	(0.0006)		(0.1559)		(0.2062)
λ_i	0.5418	0.6296	β	0.5156		
	(0.1680)	(0.0550)		(0.1891)		
δ_i	0.5805	0.7404	s_{12}	0.1656		
	(0.1680)	(0.0442)		(0.0759)	LogLike	6660.69
d_v	0.3699		d_c	0.4372	AIC	-13297.4
	(0.0326)			(0.0780)	BIC	-13232.5

Note: Standard errors are in parentheses.

Table 8: ML Estimates for the Trivariate DCC*t*-FC Model

Parameters	Nikkei	Hang Seng	SST	Parameters	Others
τ_{0i}	0.0221	0.0554	0.0523	α	0.2075
	(0.0261)	(0.0232)	(0.0199)		(0.0461)
τ_{1i}	-0.0271	-0.0296	-0.0760	β	0.7649
	(0.0185)	(0.0167)	(0.0169)		(0.0537)
κ_i	0.0121	0.0115	0.0098	ν	8.8863
	(0.0049)	(0.0038)	(0.0030)		(0.5695)
γ_i	0.3118	0.2092	0.2289		
	(0.0540)	(0.0408)	(0.0408)		
δ_i	0.6863	0.7832	0.7616		
	(0.0537)	(0.0379)	(0.0417)		
s_{2i}	0.5883				
	(0.0318)			LogLike	-8734.9
s_{3i}	0.5854	0.6747		AIC	17511.8
	(0.0326)	(0.0267)		BIC	17629.4

Note: Standard errors are in parentheses.

Table 9: ML Estimates for the Trivariate Threshold DCCt-FC Model

	State 1 ($j=1$)			State 2 ($j=2$)		
Parameters	Nikkei	Hang Seng	SST	Nikkei	Hang Seng	SST
$\tau_{0i}^{(j)}$	0.0103	0.0485	0.0563	-0.1005	0.0302	0.0427
	(0.0502)	(0.0357)	(0.0305)	(0.0472)	(0.0340)	(0.0288)
$\tau_{1i}^{(j)}$	-0.0818	-0.0792	-0.0954	0.0911	0.0160	-0.0549
	(0.0372)	(0.0263)	(0.0261)	(0.0378)	(0.0241)	(0.0242)
$\kappa_i^{(j)}$	5.6616×10^{-11}	3.2100×10^{-11}	1.4252×10^{-12}	0.0293	0.0265	0.0172
	(0.0428)	(0.0304)	(0.0232)	(0.0324)	(0.0225)	(0.0172)
$\gamma_i^{(j)}$	0.4162	0.2993	0.2873	0.2336	0.1510	0.1413
	(0.0833)	(0.0651)	(0.0565)	(0.0745)	(0.0428)	(0.0491)
$\delta_i^{(j)}$	0.6515	0.7513	0.7686	0.6960	0.7852	0.7910
	(0.0856)	(0.0735)	(0.0654)	(0.0685)	(0.0476)	(0.0523)
$s_{2i}^{(j)}$	-0.9976			0.4332		
	(1.9917)			(0.0758)		
$s_{3i}^{(j)}$	-0.6463	-0.6309		0.4629	0.5437	
	(1.6570)	(1.7285)		(0.0765)	(0.0666)	
$\alpha^{(j)}$		0.2241			0.2081	
		(0.0535)			(0.0770)	
$\beta^{(j)}$		0.7838			0.7135	
		(0.0549)			(0.1034)	
ν			9.1160			
			(0.6141)			
LogLike			-8709.4			
AIC			17500.8			
BIC			17730.4			
LR Test			51.02 [0.0001]			

Note: Standard errors are in parentheses. LR test is the likelihood ratio test for the null hypothesis of the DCCt-FC model. P-values are in the square brackets.

Table 10: Tests for Value-at-Risk Thresholds

(a) Filtering Historical Simulation

Model	DCCt-FC			Threshold DCCt-FC		
Test Stat.	% Failure	UC	IND	% Failure	UC	IND
1-step-ahead						
VaR 5%	0.0200	0.0275*	0.6854	0.0350	0.3047	0.4749
VaR 1%	0.0050	0.4315	0.9199	0.0050	0.4315	0.9199
5-step-ahead						
VaR 5%	0.0205	0.0329*	0.6815	0.0513	0.9348	0.2970
VaR 1%	0.0051	0.4507	0.9189	0.0051	0.4507	0.9189
10-step-ahead						
VaR 5%	0.0158	0.0120*	0.7557	0.0316	0.2127	0.5305
VaR 1%	0.0000	NA	NA	0.0053	0.4706	0.9179
20-step-ahead						
VaR 5%	0.0111	0.0040*	0.8316	0.0222	0.0558	0.6689
VaR 1%	0.0000	NA	NA	0.0056	0.5130	0.9156

(b) Standardized t distribution

Model	DCCt-FC			Threshold DCCt-FC		
Test Stat.	% Exceedence	UC	IND	% Exceedence	UC	IND
1-step-ahead						
VaR 5%	0.0250	0.0737	0.6117	0.0550	0.7493	0.2564
VaR 1%	0.0050	0.4315	0.9200	0.0050	0.4315	0.9199
5-step-ahead						
VaR 5%	0.0308	0.1860	0.5360	0.0667	0.3085	0.1716
VaR 1%	0.0051	0.4507	0.9189	0.0103	0.9714	0.8383
10-step-ahead						
VaR 5%	0.0263	0.1008	0.6022	0.0579	0.6258	0.2435
VaR 1%	0.0000	NA	NA	0.0053	0.4706	0.9179
20-step-ahead						
VaR 5%	0.0167	0.0178*	0.7491	0.0278	0.1367	0.5919
VaR 1%	0.0056	0.5130	0.9156	0.0056	0.5130	0.9156

Note: '% Exceedence' is the proportion of days when the return exceeds the VaR threshold. UC and IND are the likelihood ratio tests for unconditional coverage and independence (against a first-order Markov alternative) developed by Christoffersen (1998). The two columns under 'UC' and 'IND' present P-values of the UC and IND tests with '*' indicating the significance at the five percent level.

Figure 1: Plot of the estimates of the conditional correlations, $C_{T+1|T},...,C_{T+300|T+299}$ (the blue one), and the realized correlations (the red one).

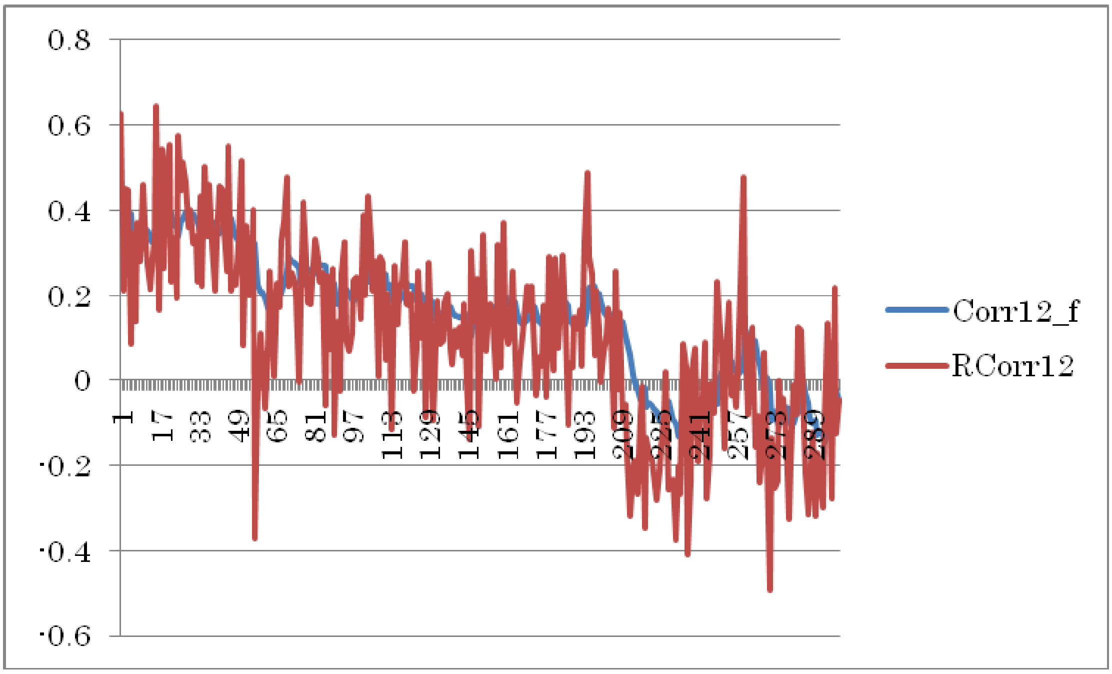

Figure 2: Plot of the forecasts of ω_{1t}^{f} (blue); and the realized value of ω_{1t}^{opt} (red).

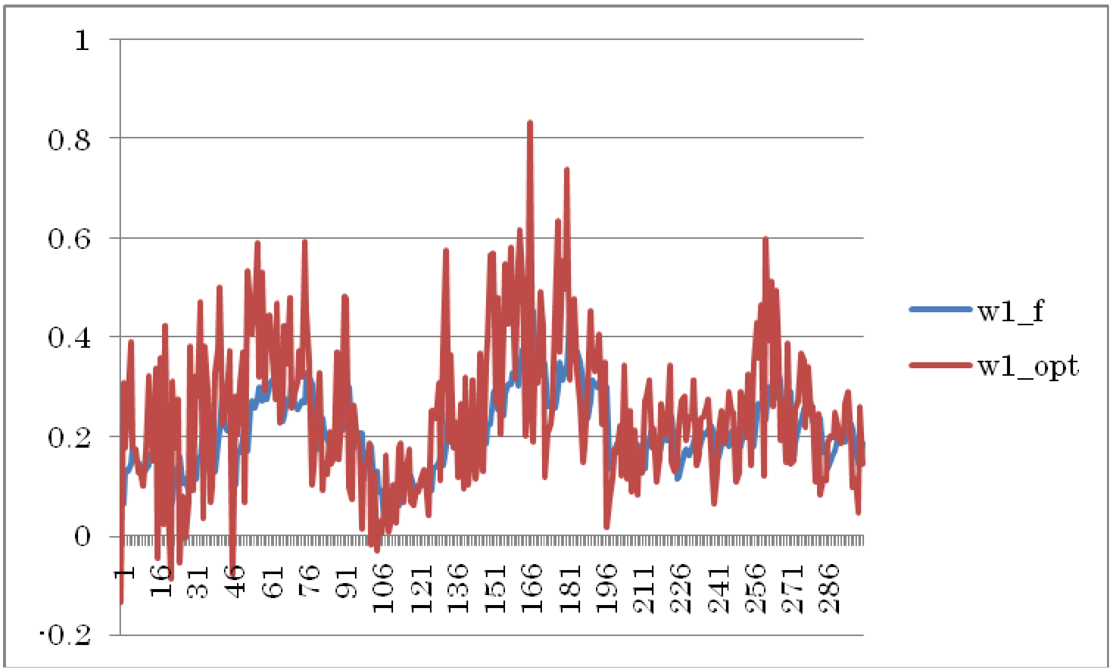

Figure 3: The ratio of the portfolio risks.