University of Toronto
Department of Economics

Working Paper 408

Modelling Realized Covariances and Returns

By Xin Jin and John M Maheu

July 16, 2010

Modelling Realized Covariances and Returns*

Xin Jin† John M. Maheu‡

This draft June 2010

Abstract

This paper proposes new dynamic component models of realized covariance (RCOV) matrices based on recent work in time-varying Wishart distributions. The specifications are linked to returns for a joint multivariate model of returns and covariance dynamics that is both easy to estimate and forecast. Realized covariance matrices are constructed for 5 stocks using high-frequency intraday prices based on positive semi-definite realized kernel estimates. The models are compared based on a term-structure of density forecasts of returns for multiple forecast horizons. Relative to multivariate GARCH models that use only daily returns, the joint RCOV and return models provide significant improvements in density forecasts from forecast horizons of 1 day to 3 months ahead. Global minimum variance portfolio selection is improved for forecast horizons up to 3 weeks out.

key words: eigenvalues, dynamic conditional correlation, predictive likelihoods, MCMC.
JEL: C11, C32, C53, G17

*A previous version of this paper was titled "Modelling Realized Covariances". We are grateful for many helpful comments from Luc Bauwens, Christian Gourieroux, Tom McCurdy, Cathy Ning and seminar participants at Ryerson University and the Applied Financial Time Series conference HEC Montreal. Maheu thanks the Social Sciences and Humanities Research Council of Canada for financial support.

†Department of Economics, University of Toronto, Canada, reynold.jin@utoronto.ca

‡Department of Economics, University of Toronto, Canada, and RCEA Italy, jmaheu@chass.utoronto.ca

1 Introduction

This paper proposes new dynamic component models of realized covariance (RCOV) matrices based on recent work in time-varying Wishart distributions.[1] The specifications are linked to returns for a joint multivariate model of returns and covariance dynamics that is both easy to estimate and forecast. The models are compared based on an out-of-sample term structure of density forecasts of returns. While the existing literature has focused on the forecasting of RCOV, this paper demonstrates the benefits to forecasts of the return distribution from the joint modelling of RCOV and returns.

Multivariate volatility modelling is a key input into portfolio optimization, risk measurement and management. There has arose a voluminous literature on how to approach this problem. The two popular approaches based on return data are multivariate GARCH (MGARCH) and multivariate stochastic volatility (MSV). Bauwens et al. (2006) provide a recent survey of MGARCH modelling while Asai et al. (2006) review the MSV literature. Despite the important advances in this literature there remain significant challenges. In practise the covariance of returns is unknown and is either projected onto past data in the case of MGARCH or is assumed to be latent in the case of MSV. For MSV sophisticated simulation methods must be used to deal with the unobserved nature of the conditional covariances. However, if an accurate measure of the covariance matrix could be obtained many of these difficulties could be avoided.

Recently, a new paradigm has emerged in which the latent covariance of returns is replaced by an accurate estimate based upon intraperiod return data. The estimator is nonparametric in the sense that we can obtain an accurate measure of daily ex post covariation without knowing the underlying data generating process. Realized covariance (RCOV) matrices open the door to standard time series analysis. See Andersen et al. (2003), Barndorff-Nielsen and Shephard (2004b) and Bandi and Russell (2005b) for the theoretical foundations and Andersen et al. (2009) and McAleer and Medeiros (2008) for surveys of the literature.

The purpose of this paper is to propose new joint models for time-varying RCOV matrices and returns. Among the few models in the literature for RCOV matrices is the Wishart autoregressive model of Gourieroux, Jasiak, and Sufana (2009). The process is defined by the Laplace transform and naturally leads to method of moments estimation (see also Chiriac (2006)) while the transition density is a noncentered Wishart. In a different approach Bauer and Vorkink (2010) decompose the RCOV matrix by a log-transformation and then use various time-series approaches to model the elements. Chiriac and Voev (2010) use a 3-step procedure, by first decomposing the RCOV matrices into Cholesky factors and modelling them with a VARFIMA process before transforming them back.

This paper is related to this literature but differs in that our model builds on the MSV Wishart specifications of Asai and McAleer (2009) and Philipov and Glickman (2006). These models specify a standard Wishart transition density for the covariance of returns.[2] In contrast to modelling the Cholesky factor or log-transformation of RCOV, contemporaneous

[1]The Wishart distribution is a generalization of the univariate gamma distribution to nonnegative-definite matrices.

[2]An advantage to working with the Wishart distribution is that the pdf and simulation methods for random draws are readily available, while this is not the case for the noncentral Wishart distribution (Gauthier and Possamai 2009).

covariances between elements in the RCOV matrix are straightforward to interpret and model using a Wishart law of motion.

RCOV is estimated following Barndorff-Nielsen et al. (2008) to obtain an *observable* MSV model.[3] The empirical analysis of 5 stocks show the strong persistence of the daily time series of RCOV elements. We propose new MSV Wishart specifications with components to capture the persistence properties in realized covariances. A component is defined as a sample average of past RCOV matrices based on a particular window of data. Different windows of data give different components. Two types of time-varying Wishart model are considered. The first assumes the components affect the scale matrix in a multiplicative fashion while the second has the components enter additively. The additive specification performs the best in our analysis. Relative to the Wishart MSV models based only on returns the estimation is considerably simplified.

The models are estimated from a Bayesian perspective. We show how to estimate the length of data windows that enter into the components of the model. For each of the RCOV models the second component is associated with 2 weeks of data while the third component is associated with about 3 months of past data. The component models deliver a dramatic improvement in capturing the time series autocorrelations of the smallest and largest eigenvalues of the RCOV matrices.

Besides providing new tractable models for multivariate observable SV we also evaluate the models over a term structure of density forecasts of returns and a term structure of global minimum-variance portfolios.[4] It is important to consider density forecasts of returns since this is the quantity that in principle enters into all financial decisions such as risk measurement and management. In general the covariance of future returns is not a sufficient statistic for the density of returns.[5] Daily returns are common to both the MGARCH and RCOV models and provide a common metric to compare models that use high and low-frequency data. In contrast to the value-at-risk measures that focus on the tails of a distribution the cumulative log-predictive likelihoods measure the accuracy of the whole return distribution. A term structure of forecasts from 1 to 60 days ahead is considered in order to assess model forecast strength at many different horizons. Our results on density forecasts of returns are a new contribution to the literature and demonstrate the forecast gains that joint modelling of RCOV and returns provide compared to traditional MGARCH approaches.

The improvements from using high-frequency data are substantial. The RCOV component models have the best multiperiod density forecasts as well as the smallest portfolio variance. The best RCOV model provides significant improvements in density forecasts of returns over MGARCH models for up to 3 months ahead. The gains from using high frequency data for global minimum variance portfolio selection are important up to 3 weeks ahead.

[3]Estimation of RCOV this way has several benefits including imposing the positive definiteness and accounting for the bias that market microstructure and nonsynchronous trading can have.

[4]Maheu and McCurdy (2009) introduced the term structure of density forecasts for returns using joint models for returns and realized volatility for individual assets. We extend this to include multivariate assets and global minimum variance portfolios.

[5]For instance, the predictive density of returns in the models integrate out both parameter uncertainty and uncertainty regarding future RCOV values, making the density highly non-Gaussian.

In summary, we provide a new approach to modelling multivariate returns that consists of joint models of returns and RCOV matrices. We find that it is critical to include the components to obtain improved performance relative to MGARCH models. Compared to MSV models, our specifications provide the additional flexibility of SV at a considerably lower computational cost. This paper is organized as follows. In Section 2, we review the theory and the procedures of constructing the RCOV estimator and the data. In Section 3, two models for RCOV are introduced after briefly discussing two benchmark multivariate GARCH models of volatility based on daily returns. Section 4 explains the estimation procedure. Estimation results are reported in Section 5, followed by model comparison. Section 6 concludes. The Appendix contains details on posterior simulation.

2 Realized Covariance

2.1 RCOV Construction

Suppose the k-dimensional efficient log-price $Y(t)$, follows a continuous time diffusion process defined as follows:

$$Y(t) = \int_0^t a(u)\mathrm{d}u + \int_0^t \Phi(u)\mathrm{d}W(u), \tag{1}$$

where $a(t)$ is a vector of drift components, $\Phi(t)$ is the instantaneous volatility matrix, and $W(t)$ is a vector of standard independent Brownian motions.[6] The quantity of interest here is $\int_0^\tau \Phi(u)\Phi'(u)du$, known as the integrated covariance of $Y(t)$ over the interval $[0,\tau]$. It is a measure of the ex-post covariation of $Y(t)$. For simplicity, we normalize τ to be 1. Results from stochastic process theory (e.g. Protter (2004)) imply that the integrated covariance of $Y(t)$,

$$\int_0^1 \Phi(u)\Phi'(u)du, \tag{2}$$

is equal to its quadratic variation over the same interval,

$$[Y](1) \equiv \mathrm{plim}_{n\to\infty} \sum_{j=1}^{n} \{Y(t_j) - Y(t_{j-1})\}\{Y(t_j) - Y(t_{j-1})\}' \tag{3}$$

for any sequence of partitions $0 = t_0 < t_1 < \ldots < t_n = 1$ with $\sup_j\{t_{j+1} - t_j\} \to 0$ for $n \to \infty$.

An important motivation for our modelling approach is Theorem 2 from Andersen, Bollerslev, Diebold and Labys (2003). They show that the daily log-return follows,

$$Y(1) - Y(0)|\sigma\{a(v), \Phi(v)\}_{0\le v\le 1} \sim N\left(\int_0^1 a(u)\mathrm{d}u, \int_0^1 \Phi(u)\Phi'(u)du\right),$$

where $\sigma\{a(v), \Phi(v)\}_{0\le v\le 1}$ denotes the sigma-field generated by $\{a(v), \Phi(v)\}_{0\le v\le 1}$. In our empirical work we will assume the drift term is approximately 0 while the integrated covariance

[6] Jumps are not considered in this paper. How to model individual asset jumps and common jumps among several assets is an open question which we leave for future work.

can be replaced by an accurate estimate using high-frequency intraday data. We discuss the estimation of this next.

We are interested in obtaining an estimator of the quadratic variation of Y over a day, which is a measure of the ex post daily covariation. This estimator is referred to as realized covariance, or RCOV. We require RCOV to be positive definite. One way of constructing RCOV for a particular day t is

$$\widehat{RCOV}_t = \sum_{i=1}^{n_t} r_{i,t} r'_{i,t},$$

where n_t is the number of intraday log-returns for day t, $r_{i,t}$ is the i^{th} intraday return: $r_{i,t} = Y_{t,i} - Y_{t,i-1}$, $i = 1, 2, \ldots n_t$. In the absence of market microstructure noise, Barndorff-Nielsen and Shephard (2004b) shows that $\widehat{RCOV}_t$ is a consistent estimator of quadratic covariation as $n_t \to \infty$.

In the real world there is microstructure noise that affects the log price process, and intraday prices for different stocks are not observed at the same time, nor the same frequency. In other words, the price process is not synchronized, introducing another source of bias, known as the Epps effect. In the presence of microstructure noise, Bandi and Russell (2005b) show that the realized covariation estimator given above is not consistent. They propose a method for selecting the optimal sampling frequency as a trade-off between bias and efficiency. Instead of using all the intraday prices available, prices are sampled at a fixed frequency, say every 15 minutes. This method does not come without a cost. A majority of intraday information is discarded and a smaller sample size (the number of intraday prices to construct RCOV) usually means larger variation in the estimator.

We follow the procedure in Barndorff-Nielsen et al. (2008) (BNHLS) to construct RCOV using the high-frequency stock returns. BNHLS propose a multivariate realized kernel to estimate the ex-post covariation of log-prices. They show this new estimator is consistent, guaranteed to be positive semi-definite, can accommodate endogenous measurement noise and can also handle non-synchronous trading. To synchronize the data, they use the idea of *refresh time*. A kernel estimation approach is used to minimize the effect of the microstructure noise, and to ensure positive semi-definiteness. They choose the Parzen weight function for the kernel. We review these key ideas.

The econometrician observes the log price process $X = \left(X^{(1)}, X^{(2)}, \ldots, X^{(k)}\right)'$, which is generated by Y, but is contaminated with market microstructure noise. Prices arrive at different times and at different frequencies for different stocks over the unit interval, $t \in [0, 1]$.

Suppose the observation times for the i-th stock are written as $t_1^{(i)}, t_2^{(i)}, \ldots, i = 1, 2, \ldots, k$. Let $N_t^{(i)}$ count the number of distinct data points available for the i-th asset up to time t. The observed history of prices for the day is $X^{(i)}(t_j^{(i)})$, for $j = 1, 2, \ldots, N_1^{(i)}$, i.e, the j-th price update for asset i is $X^{(i)}(t_j^{(i)})$, it arrives at $t_j^{(i)}$. The steps to computing daily RCOV are the following.

1. Synchronizing the data.
   The first key step is to deal with the non-synchronous nature of the data. The idea of refresh time is used here. Define the first refresh time as $\tau_1 = \max\left(t_1^{(1)}, \ldots, t_1^{(k)}\right)$, and then subsequent refresh times as $\tau_{j+1} = \max\left(t^{(1)}_{N^{(1)}_{\tau_j}+1}, \ldots, t^{(k)}_{N^{(k)}_{\tau_j}+1}\right)$. τ_1 is the

time it has taken for all the assets to trade, i.e. all their posted prices have been updated at least once. τ_2 is the first time when all the prices are again updated, etc. From now on, we will base our analysis on this new conformed time clock $\{\tau_j\}$, and treat the entire k-dimensional vector of price updates as if it is observed at these refreshed times $\{\tau_j\}$. The number of observations of the synchronized price vector is $n+1$, which is no larger than the number of observations of the stock with the fewest price updates. Then, the synchronized high frequency return vector is defined as $x_j = X(\tau_j) - X(\tau_{j-1}), j = 1, 2, \ldots, n$, where n is the number of refresh return observations for the day.

2. Compute the positive semi-definite realized kernel. Having synchronized the high frequency vector returns $\{x_j\}, j = 1, 2, \ldots, n$, daily $RCOV_t$ is calculated as,

$$RCOV_t = \sum_{s=-n}^{n} f\left(\frac{s}{S+1}\right)\Gamma_s. \tag{4}$$

The selection of the bandwidth S is discussed in BNHLS while

$$f(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \leq x \leq 1/2 \\ 2(1-x)^3 & 1/2 \leq x \leq 1 \\ 0 & x > 1. \end{cases}$$

Γ_s is the s-th realized autocovariance:

$$\Gamma_s = \begin{cases} \sum_{j=|s|+1}^{n} x_j x'_{j-s}, & s \geq 0 \\ \\ \sum_{j=|s|+1}^{n} x_{j-s} x'_j, & s < 0. \end{cases}$$

We apply this multivariate realized kernel estimation to our high-frequency data, obtaining a series of daily $RCOV_t$ matrices, which will then be fitted by our proposed Wishart Model. The j-th diagonal element of $RCOV_t$ is called realized volatility[7] and is an ex post measure of the variance for asset j. Realized correlation between asset i and j is $RCOV_{t,ij}/\sqrt{RCOV_{t,ii}RCOV_{t,jj}}$ where $RCOV_{t,ij}$ is the element from the i-th row and j-th column.

2.2 Data

We use high-frequency stock prices for 5 assets, namely Standard and Poor's Depository Receipt (SPY), General Electric Co. (GE), Citigroup Inc.(C), Alcoa Inc. (AA) and Boeing Co. (BA). The sample period runs from 1998/12/04 – 2007/12/31 delivering 2281 days. We reserve the data back to 1998/01/02 (219 observations) as conditioning data for the components models. The data are obtained from the TAQ database. We use transaction prices and closely follow Barndorff-Nielsen et al. (2008) to construct daily RCOV matrices.

[7] Also called realized variance in the literature.

The data is cleaned as follows. First, trades before 9:30 AM or after 4:00 PM are removed as well as any trades with a zero price. We delete entries with a corrected trade condition, or an abnormal sale condition.[8] Finally, any trade that has a price increase (decrease) of more than 5% followed by a price decrease (increase) of more than 5% is removed. For multiple transactions that have the same time stamp the price is set to the median of the transaction prices. From this cleaned data we proceed to compute the refresh time and the realized kernel discussed in the previous section. The daily return r_t, is the continuously compounded return from the open and close prices and matches RCOV. Table 1 reports the average number of daily transaction for each stock. The average number of transactions based on the refresh time is much lower at 1835. This represents just under 5 transactions per minute. Based on this our sample is quite liquid.

Table 2 shows the sample covariance from daily returns along with the average RCOV. Figure 1 displays daily returns while the corresponding realized volatilities (RV) are in Figure 2.

3 Models

First we review two representative multivariate volatility models that use daily returns and will be compared to the RCOV models.

3.1 GARCH Models of Daily Returns

3.1.1 Vector-diagonal GARCH Model

Ding and Engle (2001) introduce the vector-diagonal GARCH (VD-GARCH-t) model to which we add Student-t innovations as follows

$$\begin{aligned} r_t|r_{1:t-1} &\sim t(0, H_t, \zeta) && (5) \\ H_t &= CC' + aa' \odot r_{t-1}r'_{t-1} + bb' \odot H_{t-1}, && (6) \end{aligned}$$

where r_t is a k-dimensional daily return series, and $r_{1:t-1} = \{r_1, \dots, r_{t-1}\}$. The parameters are C, a $k \times k$ lower trianglar matrix; a and b are $k \times 1$ vectors, and ζ is the degree of freedom in the Student-t density. $\odot$ denotes the Hadamard product of two matrices. In estimation, covariance targeting is achieved by replacing CC' with $\text{Cov}(r) \odot \left(\frac{\zeta-2}{\zeta}\iota\iota' - aa' - \frac{\zeta-2}{\zeta}bb'\right)$, where $\text{Cov}(r)$ is the sample covariance matrix estimated from daily returns, and ι is a $k \times 1$ vector of 1. This model assumes that the conditional covariance $h_{ij,t}$ is only a function of the past shock $r_{i,t-1}r_{j,t-1}$, and the past conditional covariance $h_{ij,t-1}$. The conditional covariance of returns is $\frac{\zeta}{\zeta-2}H_t$ assuming $\zeta > 2$.

[8]Specifically we remove a trade with CORR $\neq 0$, or a trade that has COND letter other than E or F in the TAQ database.

3.1.2 Dynamic Conditional Correlation Model

The second model is a dynamic conditional correlation (DCC-t) model of Engle (2002) with Student-t innovations,

$$
\begin{aligned}
r_t|r_{1:t-1} &\sim t(0, H_t, \zeta) && (7)\\
H_t &= D_t R_t D_t && (8)\\
D_t &= \text{diag}(\sigma_{i,t}) && (9)\\
\sigma^2_{i,t} &= \omega_i + \kappa_i r^2_{i,t-1} + \lambda_i \sigma^2_{i,t-1}, i = 1, \ldots, k && (10)\\
\epsilon_t &= \left(\frac{\zeta - 2}{\zeta}\right)^{1/2} D_t^{-1} r_t && (11)\\
Q_t &= \text{Corr}(r)(1 - \alpha - \beta) + \alpha \epsilon_{t-1}\epsilon'_{t-1} + \beta Q_{t-1} && (12)\\
R_t &= \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2}. && (13)
\end{aligned}
$$

Corr(r) is the sample correlation matrix estimated from daily returns. D_t, R_t, Q_t are all $k \times k$ matrices. The parameters are $\omega_1, \ldots, \omega_k$, $\kappa_1, \ldots, \kappa_k$, $\lambda_1, \ldots, \lambda_k, \alpha, \beta, \zeta$. In this specification, the unconditional mean of Q_t is equal to Corr(r), the sample correlation. This is called correlation targeting and in this way the number of parameters is greatly reduced from $\frac{k^2+5k}{2} + 2$ to $3k + 2$. Equation (10) governs the dynamics of the conditional variances of each individual return by a univariate GARCH process; equation (12) governs the dynamics of the time-varying conditional correlation of the whole return vector. Because Corr(r) is symmetric positive definite, and $\epsilon_t \epsilon'_t$ is symmetric positive semi-definite, the conditional correlation matrices are guaranteed to be symmetric positive definite.

3.2 Models of Realized Covariances

Compared to existing approaches which model factors of RCOV matrices (Cholesky factors, Chiriac and Voev (2010), principle components, Bauer and Vorkink (2010)) an advantage of the Wishart distribution is that it has support over symmetric positive definite matrices and allows for the joint modelling of all elements of a covariance matrix. Conditional moments between realized variances and covariances have closed form expressions.

Motivated by Philipov and Glickman (2006) and Asai and McAleer (2009), we propose to model the dynamics of RCOV by a time-varying Wishart distribution. This choice is similar to Gourieroux, Jasiak, and Sufana (2009) however they use a noncentered Wishart. We have also explored the inverse Wishart density as another distribution to govern the dynamics of realized covariances but found the Wishart provided superior performance.[9]

Two models are presented in which the scale matrix of the Wishart distribution follows a multiplicative structure and an additive structure. Both models feature components, which is important to providing gains against standard multivariate GARCH models, and accounting for persistence in RCOV elements.

The approach to modelling components is related to Andersen, Bollerslev and Diebold (2007), Corsi (2009), Maheu and McCurdy (2009) among others which uses the Heteroge-

[9]Note that the choice of the distribution governing the dynamics of Σ_t is unrelated to the Bayesian conjugate analysis that uses the Wishart as a conjugate prior for Σ_t^{-1} for Gaussian observations.

neous AutoRegressive (HAR) model of realized variance in the univariate case in order to capture long-memory like features of volatility parsimoniously.

3.2.1 A Multiplicative Component Wishart Model

Let $\Sigma_t \equiv RCOV_t$, then the Wishart-RCOV-M(K) model with $K \geq 1$ components is defined as,

$$\begin{aligned} \Sigma_t|\nu, S_{t-1} &\sim Wishart_k(\nu, S_{t-1}) \\ S_t &= \frac{1}{\nu}\left[\prod_{j=K}^{1} \Gamma_{t,\ell_j}^{\frac{d_j}{2}}\right] A \left[\prod_{j=1}^{K} \Gamma_{t,\ell_j}^{\frac{d_j}{2}}\right] \quad (14) \\ \Gamma_{t,\ell} &= \frac{1}{\ell}\sum_{i=0}^{\ell-1} \Sigma_{t-i} \quad (15) \\ 1 = \ell_1 < &\cdots < \ell_K. \quad (16) \end{aligned}$$

$Wishart_k(\nu, S_{t-1})$ denotes a Wishart distribution over positive definite matrices of dimension k with $\nu \geq k$ degrees of freedom and scale matrix S_{t-1}. A is a positive definite symmetric parameter matrix and d_j is a positive scalar.

The components enter as a sample average of past Σ_t raised to a different matrix power $d_k/2$.[10] The first component is assumed to be a function of only Σ_t, $\ell_1 = 1$. The component terms $\Gamma_{t,l}$ allow for more persistence in the location of Σ_t while the different values of d_j allow the effect to be dampened or amplified. In (14) the order of the product operator is important and differs in the two terms. The window width ℓ of each component $\Gamma_{t,l}$ could be preset or be estimated. We show how to estimate this in the next Section.

To discuss some of the features of this model consider the special case with $K = 1$ component, $S_t = \frac{1}{\nu}(\Sigma_t^{d_1/2})A(\Sigma_t^{d_1/2})$. By the properties of the Wishart distribution, the conditional expectation of Σ_t is:

$$E(\Sigma_t|\Sigma_{t-1}) = \nu S_{t-1} = (\Sigma_{t-1}^{d_1/2})A(\Sigma_{t-1}^{d_1/2}). \quad (17)$$

Conditional moments are straightforward to obtain and interpret. The conditional variance of element (i,j) is $\mathrm{Var}(\Sigma_{t,ij}|S_{t-1},\nu) = \frac{1}{\nu}[\tilde{S}_{t-1,ij}^2 + \tilde{S}_{t-1,ii}\tilde{S}_{t-1,jj}]$ where $\tilde{S}_{t-1,ij}$ is element (i,j) of (17). The conditional variance is increasing in $\tilde{S}_{t-1,ij}$, $\tilde{S}_{t-1,ii}$, and $\tilde{S}_{t-1,jj}$. The conditional covariance between elements has a similar form, $\mathrm{Cov}(\Sigma_{t,ij}, \Sigma_{t,km}|S_{t-1},\nu) = \frac{1}{\nu}[\tilde{S}_{t-1,ij}\tilde{S}_{t-1,km} + \tilde{S}_{t-1,ii}\tilde{S}_{t-1,jj}]$. The degree of freedom parameter ν determines how tight the density of Σ_t is centered around its conditional mean, with larger ν meaning the random matrices are more concentrated around νS_{t-1}. Thus, the modelling of the scale matrix and ν are the key factors in affecting the conditional moments of Σ_t.

[10] We also examined a *geometric* average version using the following specification: $\Gamma_{t,\ell}^d \equiv \Sigma_{t-\ell+1}^{\frac{d}{\ell}}\Sigma_{t-\ell+2}^{\frac{d}{\ell}}\cdots\Sigma_t^{\frac{d}{\ell}}$. We found this geometric average version, while it has similar performance in almost every aspect, is computationally more costly. We will hence focus our results on the sample average version.

The inverse of RCOV follows the inverse-Wishart distribution with the conditional expectation being:

$$E(\Sigma_t^{-1}|\Sigma_{t-1}) = (\nu - k - 1)^{-1} S_{t-1}^{-1} = \frac{\nu}{\nu - k - 1} (\Sigma_{t-1}^{-d_1/2}) A^{-1} (\Sigma_{t-1}^{-d_1/2}). \tag{18}$$

The elements of A determine how each element of Σ_t is related to elements of Σ_{t-1}. For example, if $W \sim Wishart_k(\nu, I_k)$ then

$$\Sigma_t = \frac{1}{\nu} (\Sigma_{t-1}^{d_1/2}) A^{1/2} W (A^{1/2})' (\Sigma_{t-1}^{d_1/2}). \tag{19}$$

The scalar parameter d_1 measures the overall influence of past RCOV on current RCOV. This parameter is closely related to the degree of persistence present in the RCOV series, with larger d_1 the stronger the persistence. Suppose A is the identity matrix and $d_1 = 1$, then by equation (17), $E(\Sigma_t|\Sigma_{t-1}) = \nu S_{t-1} = \Sigma_{t-1}$, which is a random walk in matrix form. If $d_1 = 0$, then $E(\Sigma_t|\Sigma_{t-1}) = A$, so the RCOV matrix follows an *i.i.d.* Wishart distribution over time.

By expanding to several components each with a different window lag length ℓ_j and parameter d_j, we obtain a richer model to capture the time series dependencies in realized covariances. The components are shown to be important in Section 5.

3.2.2 An Additive Component Wishart Model

A very similar model assumes the K components affect S_t in a additive fashion (Wishart-RCOV-A(K)),

$$\begin{aligned}
\Sigma_t|\nu, S_{t-1} &\sim Wishart_k(\nu, S_{t-1}) \\
\nu S_t &= B_0 + \left[\sum_{j=1}^{K} B_j \odot \Gamma_{t,\ell_j}\right] && (20) \\
\Gamma_{t,\ell} &= \frac{1}{\ell}\sum_{i=0}^{\ell-1} \Sigma_{t-i} && (21) \\
B_j &= b_j b_j', \qquad j = 1, \ldots, K && (22) \\
1 = \ell_1 < &\cdots < \ell_K. && (23)
\end{aligned}$$

Parameters are B_0, ν, $b_1, \ldots, b_k$, $\ell_2, \ldots, \ell_K$. B_0 is a $k \times k$ symmetric positive definite matrix, and b_j's are $k \times 1$ vectors making B_j rank 1. This specification ensures S_t is symmetric positive definite. Instead of estimating B_0, we implement RCOV targeting by setting $B_0 = (\iota\iota' - B_1 - B_2 - B_3) \odot \overline{\Sigma}_t$, where $\overline{\Sigma}_t$ is the sample mean of Σ_t. This ensures that the long-run mean of Σ_t is equal to $\overline{\Sigma}_t$. In estimation we impose the condition that every element of $(\iota\iota' - B_1 - B_2 - B_3)$ is positive.

This model only differs from the previous one in the form of the scale matrix. Like the VD-GARCH model, element (i, j) of the scale matrix is a function only of element (i, j) of lagged Σ_t. Many other parameterizations could be considered but this specification is reasonably parsimonious and performs well in the empirical work. All of the conditional moments discussed in the previous subsection hold for this model using the scale matrix S_{t-1} from (20).

3.2.3 Linking Returns and RCOV

Along with either the Wishart-RCOV-M or Wishart-RCOV-A specification for Σ_t the joint model with returns is closed with

$$r_t|\Sigma_t \sim N(0, \Sigma_t^{1/2}\Lambda(\Sigma_t^{1/2})'). \tag{24}$$

Λ is a symmetric positive definite matrix and allows the covariance of returns to deviate from the RCOV measure. Λ is estimated and to the extent that Λ is different than I_k it tells us something about the bias in the RCOV estimator. Therefore, Λ provides a simple way to adjust for estimation error in RCOV.

4 Model Estimation

We apply standard Bayesian estimation techniques to estimate the models using MCMC methods for posterior simulation. The posterior distribution is unknown for all the models considered, but a Markov Chain that has as its limiting distribution the posterior distribution of the parameters of interest can be sampled from using MCMC simulations. Features of the posterior density can then be estimated consistently based on the samples obtained from the posterior. For example, we can estimate the posterior mean of model parameters by the sample average of the MCMC draws. For more details on MCMC methods see Chib (2001).

To apply Bayesian inference, we need to first assign priors to the parameters. For the parameters common to the different RCOV models we assume $A^{-1} \sim Wishart_k(\gamma_0, Q_0)$, a Wishart distribution with $Q_0 = I_k$ and $\gamma_0 = k + 1$ set to reflect a proper but relatively uninformative prior. Each d_j follows a uniform prior, $d_j \sim U(-1, 1)$, and $\nu \sim \exp(\lambda_0)I_{\nu>k}$, an exponential distribution with support truncated to be greater than k. To make the prior flat, λ_0 is set to 100. The priors on the elements of b_j's and ν are all $N(0, 100)$, except the first element of each b_j is truncated to be positive for identification purposes. In the empirical work focus is given to $K = 3$ components as this was found to produce good results. The priors for ℓ_2 and ℓ_3 are uniform discrete with support $\{2, 3, \ldots, 200\}$, with the restriction that $\ell_2 < \ell_3$ for identification. We assume independence among the prior distributions of parameters.

If $\Sigma_{1:t-1} = \{\Sigma_1, \ldots, , \Sigma_{t-1}\}$, then the joint density of returns and realized covariances is decomposed as

$$p(r_t, \Sigma_t|\Lambda, \Theta, r_{1:t-1}, \Sigma_{1:t-1}) = p(r_t|\Lambda, \Sigma_t)p(\Sigma_t|, \Theta, \Sigma_{1:t-1}) \tag{25}$$

where Θ is the parameters in the RCOV specification. $p(r_t|\Lambda, \Sigma_t)$ has a density in (24) while $p(\Sigma_t|\Theta, \Sigma_{1:t-1})$ has the density from (14) for the Wishart-RCOV-M(K) model or (20) for the Wishart-RCOV-A(K) model. Equation (25) implies that estimation of Λ and Θ can be done separately.

Bayes' rule gives the posterior for Θ in the Wishart models as

$$p(\Theta|\Sigma_{1:T}) \propto \left[\prod_{t=1}^{T} p(\Sigma_t|\Theta, \Sigma_{1:t-1})\right] p(\Theta) \tag{26}$$

where $p(\Theta)$ is the prior discussed above. Conditional distributions used in posterior simulation are proportional to this density.

The Wishart-RCOV-M(3) model has parameters $\Theta = \{A, d, \ell, \nu\}$, with $d = \{d_1, d_2, d_3\}$ and $\ell = \{\ell_2, \ell_3\}$. MCMC sampling iterates making parameter draws from the following conditional distributions.

- $A^{-1}|\nu, d, \ell, \Sigma_{1:T}$
- $d_i|A, d_{-i}, \ell, \nu, \Sigma_{1:T}, \;\; i = 1, 2, 3$
- $\ell_i|\nu, d, \ell_{-i}, \Sigma_{1:T}, \;\; i = 2, 3$
- $\nu|A, d, \ell, \Sigma_{1:T}$

Sampling $A^{-1}|\nu, d, \ell, \Sigma_{1:T}$ is a typical Gibbs sampling step from a Wishart density while the remaining conditional densities are unknown and require a Metropolis-Hastings step. For these a random walk proposal is used. Similar steps are used to simulate from the posterior of the Wishart-RCOV-A(3) model.

Taking a draw from all of the conditional distributions constitutes one sweep of the sampler. After dropping an initial set of draws as burnin we collect M draws to obtain $\{\Theta^{(i)}\}_{i=1}^{M}$. Simulation consistent estimates of posterior moments can be obtained as sample averages of the draws. For instance, the posterior mean of Θ can be estimated as $M^{-1}\sum_{i=1}^{M}\Theta^{(i)}$.

Posterior simulation from $\Lambda|r_{1:T}, \Sigma_{1:T}$ is based on recognizing that $\tilde{r}_t = \Sigma_t^{-1/2} r_t \sim N(0, \Lambda)$. Setting the prior density of Λ^{-1} to Wishart results in a standard conjugate result for the multivariate normal model. This is done separately from the estimation for the RCOV models.

All of the details of the conditional distributions and proposal distributions along with details for the multivariate GARCH are collected in the Appendix.

5 Results

All data is used for estimation of the models. In each case, the first 1000 draws are discarded as burnin in posterior simulation and the next 5000 MCMC draws are used for inference. Before the parameter estimates are discussed we consider some features of the Wishart-RCOV(K)-M model and compare it to the DCC model.

To investigate the time-series persistence in RCOV, we plot the sample autocorrelation function (ACF) of both the largest and smallest eigenvalues of the RCOV series observed, and compare it to the posterior mean ACF obtained from the Wishart-RCOV-M(K) models.[11] The interpretation of the eigenvalues follows Gourieroux, Jasiak, and Sufana (2009). The largest (smallest) eigenvalue is equal to the maximum (minimum) risk from the portfolio with variance $\omega'\Sigma_t\omega$, given standardized portfolio weights $\omega'\omega = 1$. In Figure 3 and 4 both the largest and smallest eigenvalues show strong persistence and are different from zero even

[11]The posterior mean of the ACF of the eigenvalues is obtained as follows. A parameter draw is taken from the posterior density and used to simulate 2281 observations of Σ_t. From this the ACF is computed and saved. This is repeated many times and the average ACF is displayed.

400 lags out. The one component Wishart-RCOV-M(1) is completely unable to match the data.[12] The $K = 2$, and 3 component models provide significant improvements. The kinks in the ACF function from the simulated models occur at exactly the window width of the components. For instance, the Wishart-RCOV-M(3) has posterior mean of ℓ_2 equal to 9 and ℓ_3 equal to 64, and the change in the slope of the ACF can be seen at these locations. As we shall see in the remaining results and those in the next section, capturing the persistence in the RCOV matrices is critical to providing better in-sample and out-of-sample predictions.

Table 3 reports the in-sample mean square error (MSE) for the Wishart-RCOV-M(K) model against the observed RCOV series. For each model, the MSE is calculated in the following way: for each period t, we calculate the expected RCOV according to $E(\Sigma_t|\Sigma_{t-1}) = \nu S_{t-1}$, and subtract it from the observed RCOV, Σ_t. We then square each element of the difference matrix and sum them. As a comparison, the MSE for another 3 models: a naive model that sets the expected RCOV as last period's RCOV (i.e. $E(\Sigma_t|\Sigma_{t-1}) = \Sigma_{t-1}$); a model that sets the predicted RCOV equal to the sample covariance computed from daily returns; and a DCC model are included. The DCC model (Section 3.1.2) is estimated using daily returns to produce an in-sample estimate of the conditional covariance. The DCC model provides an improvement over the sample covariance estimate and also beats the naive RCOV model. However, each of the Wishart-RCOV-M(K) models provide significant improvements relative to any method that uses daily returns. The 3-component has the smallest MSE but this is only marginally smaller that the 2-component version.

To compare how the model tracks correlations, Figures 5 and 6 display *realized correlation* computed from the elements of the RCOV matrix along with the fitted estimates from the Wishart-RCOV-M(3) specification and the DCC model. For the realized covariance models, the fitted correlation is extracted from $E(\Sigma_t|\Sigma_{t-1})$. The first figure is the correlation between SPY and GE and the second is the correlation between GE and Citigroup. Both models track the realized correlation closely with the RCOV model displaying a clear advantage. In some episodes the DCC wanders away from the realized correlations. This is not surprising since the DCC can only infer correlations from noisy daily returns.

Based on this discussion the remainder of the paper will focus on $K = 3$ components in the Wishart-RCOV-M(K) and Wishart-RCOV-A(K) models. Estimates of the elements of A are found in Table 4 for Wishart-RCOV-M(3) while the remaining parameters are in Table 5. This table reports numerical standard errors based on the Newey-West estimator for the long-run variance with a lag length of 1000. The *inefficiency factors* are the ratio of the long-run variance estimate to the sample variance where the latter assumes an i.i.d. sample. The lower the value is, the closer the sampling is to i.i.d. This shows that the chain mixes well. The 95 percent density intervals are constructed using the 2.5th percentile and 97.5th percentile of the MCMC draws for the corresponding parameters.

The posterior mean of ℓ_2 and ℓ_3 are 9 and 64 respectively, and correspond to about 2 weeks and 3 months of data used in the second and third component. They are also accurately estimated as their density intervals indicate. The estimates of d_i show that the second component has relatively larger memory that the other two. Estimates of A show it to be very close to a diagonal matrix.

Table 6 reports parameter estimates for the Wishart-RCOV-A(3) parameterization. The

[12]Similar results are obtained for the ACF of individual elements of the RCOV matrices.

elements of the vectors b_1, b_2, and b_3 are reported along with the estimated lag lengths of the components. The posterior means for ℓ_2 and ℓ_3 are almost identical to the Wishart-RCOV-M(3) model as is the degree of freedom at 14. Once again the parameters are accurately estimated with tight density intervals.

In summary, the estimates provide strong evidence of components in RCOV with fairly long lag lengths in the third component. The in-sample one-step ahead forecasts of RCOV show improvements in using the RCOV models as compared to the multivariate GARCH. In addition, the RCOV models are better able to track realized correlations. The next section considers if these gains transfer to out-of-sample density forecasts of returns.

5.1 Density Forecasts

In this section, we compare the joint RCOV and returns models to the benchmark multivariate GARCH models, focusing on their out-of-sample performance. We compare each candidate's predictive density of returns through predictive likelihood, which is a popular approach in the literature (Maheu and McCurdy (2009), Amisano and Giacomini (2007), Bao, Lee, and Saltoglu (2007), Weigend and Shi (2000)).

It is important to consider density forecasts of returns since this is the quantity that in principle enters into all financial decisions such as portfolio choice and risk measurement.[13] Another reason for comparing models this way is that the daily returns are common to both the GARCH and the joint RCOV and return models and provides a common metric to compare models that use high and low frequency data. In contrast to the value-at-risk measures that focus on the tails of a distribution the predictive likelihoods test the accuracy of the whole distribution. Finally, a term structure of forecasts is considered in order to assess model forecast strength at many different horizons.

From a Bayesian perspective the predictive likelihoods are a key input into model comparison through predictive Bayes factors (Geweke (2005)). Following Maheu and McCurdy (2009) we evaluate a term structure of a model's density forecasts of returns. This is the cumulative log-predictive likelihood based on out-of-sample data for $h = 1, ..., H$ period ahead density forecasts of returns.

For a candidate model $\mathcal{A}$, we compute the following cumulative log-predictive likelihood:

$$\hat{p}_h^{\mathcal{A}} = \sum_{t=T_0-h}^{T-h} \log(p(r_{t+h}|I_t, \mathcal{A})), \tag{27}$$

for $h = 1, 2, \ldots, H$ and $T_0 < T$. For each h, $\hat{p}_h^{\mathcal{A}}$ measures the forecast performance based on the *same common set of returns*: $r_{T_0}, ..., r_T$. Therefore, $\hat{p}_1^{\mathcal{A}}$ is comparable with $\hat{p}_{10}^{\mathcal{A}}$ and allows us to measure the decline in forecast performance as we move from 1 day ahead forecasts to 10 day ahead forecasts using model $\mathcal{A}$. We are also interested in comparing $\hat{p}_h^{\mathcal{A}}$ for a fixed h with another specification $\mathcal{B}$, using its cumulative log-predictive likelihood $\hat{p}_h^{\mathcal{B}}$. Better models, in terms of more accurate predictive densities, will have larger (27).

For the RCOV models $I_t = \{r_{1:t}, \Sigma_{1:t}\}$ while for the MGARCH models $I_t = \{r_{1:t}\}$. The predictive likelihood $p(r_{t+h}|I_t, \mathcal{A})$, is the h-period ahead predictive density for model $\mathcal{A}$

[13]In general the covariance of returns is not a sufficient statistic for the future return distribution except with a Gaussian assumption.

evaluated at the realized return r_{t+h},

$$p(r_{t+h}|I_t, \mathcal{A}) = \int p(r_{r+h}|\theta, \Omega_{t+h}, \mathcal{A})p(\Omega_{t+h}|\theta, I_t, \mathcal{A})p(\theta|I_t, \mathcal{A})\mathrm{d}\theta\mathrm{d}\Omega_{t+h}. \tag{28}$$

Parameter uncertainty from θ and the future latent covariance of returns Ω_{t+h} are both integrated out and will in general result in a highly non-Gaussian density on the left hand side of (28). In the DCC-t and VD-GARCH-t models $\Omega_t \equiv H_t$ while for the Wishart-RCOV-M(3) and Wishart-RCOV-A(3) models $\Omega_t \equiv \Sigma_t$, while θ is the respective parameter vector. The integration is approximated as

$$\int p(r_{t+h}|\theta, \Omega_{t+h}, \mathcal{A})p(\Omega_{t+h}|\theta, I_t, \mathcal{A})p(\theta|I_t, \mathcal{A})\mathrm{d}\theta\mathrm{d}\Omega_{t+h} \approx \frac{1}{M}\sum_{i=1}^{M} p(r_{t+h}|\theta^{(i)}, \Omega_{t+h}^{(i)}, \mathcal{A}), \tag{29}$$

where $\Omega_{t+h}^{(i)} \sim p(\Omega_{t+h}|\theta^{(i)}, I_t, \mathcal{A})$, and $\theta^{(i)} \sim p(\theta|I_t, \mathcal{A})$. $\{\theta^{(i)}\}_{i=1}^{M}$ are the MCMC draws from the posterior distribution $p(\theta|I_t, \mathcal{A})$ for the model.

For the GARCH models, $p(r_{t+h}|\theta^{(i)}, \Omega_{t+h}^{(i)}, \mathcal{A})$ is the pdf of a multivariate Student-t density with mean 0, scale matrix $H_{t+h}^{(i)}$ and degree of freedom $\zeta^{(i)}$ evaluated at r_{t+h}. $H_{t+h}^{(i)}$ is simulated out from the last in-sample value $H_t^{(i)}$ which is computed using the GARCH recursion and the parameter draw $\theta^{(i)}$ from the posterior density given data $I_t = \{r_{1:t}\}$.

For the Wishart models, realized covariance dynamics are linked to returns through equation (24). $p(r_{t+h}|\theta^{(i)}, \Omega_{t+h}^{(i)}, \mathcal{A})$ is the pdf of a multivariate Normal density with mean 0 and covariance $(\Sigma_{t+h}^{(i)})^{1/2}\Lambda^{(i)}((\Sigma_{t+h}^{(i)})^{1/2})'$ evaluated at r_{t+h}. $\Sigma_{t+h}^{(i)}$ is simulated out using the Wishart dynamics of the particular model and conditional on $\theta^{(i)}, \Lambda^{(i)}$ from the posterior density, given data $I_t = \{r_{1:t}, \Sigma_{1:t}\}$.

Note that for each term $p(r_{t+h}|I_t, \mathcal{A})$ in the out-of-sample period we *re-estimate the model to obtain a new set of draws from the posterior to compute* (29). In other words the full set of models is recursively estimated for $t = T_0 - H, \ldots, T - 1$.

Given a model $\mathcal{A}$ with predictive likelihood $\hat{p}^{\mathcal{A}}$, and model $\mathcal{B}$ with predictive likelihood $\hat{p}^{\mathcal{B}}$, based on the common data $\{r_{T_0}, \ldots, r_T\}$, the predictive Bayes factor in favor of model $\mathcal{A}$ versus model $\mathcal{B}$ is $BF_{\mathcal{AB}} = \frac{\hat{p}^{\mathcal{A}}}{\hat{p}^{\mathcal{B}}}$. The Bayes factor is a relative ranking of the ability of the models to account for the data. A positive value means that model $\mathcal{A}$ is better able to account for the data compared to model $\mathcal{B}$. Kass and Raftery (1995) suggest interpreting the evidence for $\mathcal{A}$ as: not worth more than a bare mention if $0 \leq BF_{\mathcal{AB}} < 3$; positive if $3 \leq BF_{\mathcal{AB}} < 20$; strong if $20 \leq BF_{\mathcal{AB}} < 150$; and very strong if $BF_{\mathcal{AB}} \geq 150$.

Figure 7 plots $\hat{p}_h$ for all the models against $h = 1, 2, \ldots, H = 60$, giving each model a cumulative log-predictive likelihood term structure. Included are Wishart-RCOV-M(3), Wishart-RCOV-A(3) models using realized covariances computed from high frequency intraday data. The models based only on daily returns are a DCC with Gaussian innovations, the DCC-t and VD-GARCH-t model which both have Student-t innovations.

The out-of-sample data begins at T_0=2006/03/31 and ends at 2007/12/31 for a total of 441 observations. This is true for each model and each forecast horizon h. All specifications have a downward sloping term structure. Intuitively, forecasting further out is more difficult.

Generally, the best model across the full term structure of density forecasts is the Wishart-RCOV-A(3) model. The worst model is the DCC with normal innovations. Despite the fact that GARCH models with Normal innovations provide fat-tailed predictive densities for $h > 1$ the Student-t density is critically important for long horizon forecasts as the DCC-t model is uniformly better than the Gaussian DCC across all h.

The Wishart-RCOV-M(3) provide good short-term density forecasts (5 days ahead) but after this its performance decays relative the the MGARCH alternatives that do not employ RCOV data.

Are the differences in cumulative log-predictive likelihood values important? Figure 8 plots the log-predictive Bayes factors for the models that exploit intraday data versus the MGARCH models that do not. This figure is computed by subtracting the the log-predictive likelihood values in Figure 7. The top panel is the evidence supporting the RCOV models against the VD-GARCH-t model. The Wishart-RCOV-A(3) everywhere dominates except around 30 days ahead where the difference in forecast precision is very similar. The improvements for the Wishart-RCOV-A(3) are considerable. The log-Bayes factors in favor of this specification are 15.19 ($h = 1$), 12.07 ($h = 10$), 4.07 ($h = 30$), and 24.88 ($h = 60$). This means that for $h = 1$ the Wishart-RCOV-A(3) model is exp(15.10) times better at accounting for the out-of-sample data than the VD-GARCH-t model, while the RCOV model is exp(24.88) times better for $h = 60$. In other words, the RCOV model offers significant improvements in out-of-sample forecasts of the return distribution. The second plot in Figure 8 gives slightly stronger evidence for the RCOV model compared to the DCC-t model.

The evidence for the RCOV model is weaker using the Wishart-RCOV-M(3) specification. For short-run forecasts up to 15 days ahead it beats the MGARCH models but thereafter the evidence is mixed and often in support of the MGARCH.

Finally, the log-predictive Bayes factor found in Figure 9 compares the model of returns (24) with Λ estimated versus $\Lambda = I$. The latter says that RCOV is perfectly synonymous with the covariance of daily returns. The evidence across the term structure of forecasts is in favor of Λ being a free estimated parameter.

5.2 Economic Evaluation

In this section, we evaluate the out-of-sample performance of the models from a portfolio optimization perspective. We focus on the simple problem of finding the global minimum variance portfolio, so the issue of specifying the expected return is avoided. The h-period ahead global minimum variance portfolio (GMVP) is computed as the solution to

$$\min_{w_{t+h|t}} w'_{t+h|t}\Omega_{t+h|t}w_{t+h|t}$$
$$\text{s.t.} \quad w'_{t+h|t}\iota = 1.$$

$\Omega_{t+h|t}$ is the predictive mean of the covariance matrix at time $t+h$ given time t information for a particular model. Each model is re-estimated at each data point in the out-of-sample period. From the posterior draws the predictive mean of the covariance matrix at time $t+h$ is simulated along the lines of the previous subsection. $w_{t+h|t}$ is the portfolio weight, and ι is a vector with all the elements equal to 1. The optimal portfolio weight is

$$w_{t+h|t} = \frac{\Omega^{-1}_{t+h|t}\iota}{\iota'\Omega^{-1}_{t+h|t}\iota}. \tag{30}$$

It can be shown (Engle and Colacito (2006)) that if the portfolio weights, w_t, are constructed from the true conditional covariance, then the variance of a portfolio computed using the GMVP from any other model must be larger.

We evaluate model performance starting at 2006/03/31 to 2007/12/31 for a total of 441 observations for $h = 1, ..., H = 60$. The specifications considered are: Wishart-RCOV-A(3), Wishart-RCOV-M(3), DCC, DCC-t and the static sample covariance. As in the density forecasts the models are estimated using data up to and including time t and weights are computed from (30). For the MGARCH models $\Omega_{t+h|t} = E[\frac{\zeta}{\zeta-2}H_{t+h}|r_{1:t}]$ and for the RCOV models $\Omega_{t+h|t} = E[\Sigma_{t+h}^{1/2}\Lambda(\Sigma_{t+h}^{1/2})'|r_{1:t}, \Sigma_{1:t}]$. These terms are computed by simulation and have parameter uncertainty integrated out. Observation $t+1$ is added and the models are re-estimated and the new weights computed, etc.

We report the sample variances of the GMVPs across models in Figure 10. As in the density forecast exercise we use a common set of returns to evaluate the performance over different h. As a result, the upwards sloping portfolio variances indicates that time-series information is most useful for short term portfolio choice. All of the time-series models improve upon the sample covariance. The Wishart-RCOV-M(3) model provides the lowest portfolio variance across the term structure of forecasts. Both RCOV models provides gains compared to the DCC for about 15 days out after which the portfolio variance is similar.

6 Conclusion

This paper proposes to model the dynamics of realized covariance matrices (RCOV) and returns based on recent work in time-varying Wishart distributions. Realized covariance matrices are constructed for 5 stock returns using high-frequency intraday prices based on positive semi-definite realized kernel estimation introduced by Bardorff-Nielson et al. (2008). We explore the time-series properties of our RCOV models and propose component models to capture persistence. Out-of-sample performance of our models are compared to that of multivariate GARCH (MGARCH) models that only uses daily returns. The best RCOV models provide significant improvements over the MGARCH models in terms of density forecasts of returns for up to 3 months ahead. The gains from using high frequency data for global minimum variance portfolios is important up to 3 weeks ahead.

7 Appendix

7.1 Wishart-RCOV-M(K) Estimation

The parameters are $\{A, d, \ell, \nu\} = \Theta$.

Given the priors listed in Section 4 the conditional posterior distributions for the parameters are as follows.

$$
\begin{aligned}
p(A^{-1}|\nu, d, \ell, \Sigma_{t:T}) &\propto Wishart_k(A^{-1}|\gamma_0, Q_0) \times \prod_{t=1}^{T} Wishart_k(\Sigma_t|\nu, S_{t-1}) \\
&\propto \frac{|A^{-1}|^{\frac{\gamma_0-k-1}{2}}|Q_0^{-1}|^{\frac{\gamma_0}{2}}}{2^{\frac{\gamma_0 k}{2}}\prod_{j=1}^{k}\Gamma(\frac{\gamma_0+1-j}{2})} \exp\left(-\frac{1}{2}Tr(A^{-1}Q_0^{-1})\right) \\
&\times \prod_{t=1}^{T} \frac{|\Sigma_t|^{\frac{\nu-k-1}{2}}|S_{t-1}^{-1}|^{\frac{\nu}{2}}}{2^{\frac{\nu k}{2}}\prod_{j=1}^{k}\Gamma(\frac{\nu+1-j}{2})} \exp\left(-\frac{1}{2}Tr(\Sigma_t S_{t-1}^{-1})\right) \\
&\propto |A^{-1}|^{\frac{T\nu+\gamma_0-k-1}{2}} \exp\left(-\frac{1}{2}Tr[A^{-1}(Q_0^{-1} + \nu\sum_{t=1}^{T}\left[\prod_{j=1}^{K}\Gamma_{t-1,\ell_j}^{\frac{-d_j}{2}}\right]\Sigma_t\left[\prod_{j=K}^{1}\Gamma_{t-1,\ell_j}^{\frac{-d_j}{2}}\right])]\right) \\
&\propto Wishart_k(A^{-1}|\widetilde{\gamma}, \widetilde{Q})
\end{aligned} \tag{31}
$$

Where $\widetilde{Q}^{-1} = \nu\sum_{t=1}^{T}\left[\prod_{j=1}^{K}\Gamma_{t-1,\ell_j}^{\frac{-d_j}{2}}\right]\Sigma_t\left[\prod_{j=K}^{1}\Gamma_{t-1,\ell_j}^{\frac{-d_j}{2}}\right] + Q_0^{-1}$, $\widetilde{\gamma} = T\nu + \gamma_0$.

For $d_i, i = 1, \ldots, K$ we have,

$$
\begin{aligned}
p(d_i|A, d_{-i}, \ell, \nu, \Sigma_{1:T}) &\propto p(d_i) \times \prod_{t=1}^{T} Wishart_k(\Sigma_t|\nu, S_{t-1}) \\
&= p(d_i)\prod_{t=1}^{T} \frac{|\Sigma_t|^{\frac{\nu-k-1}{2}}|S_{t-1}^{-1}|^{\frac{\nu}{2}}}{2^{\frac{\nu k}{2}}\prod_{j=1}^{k}\Gamma(\frac{\nu+1-j}{2})} \exp\left(-\frac{1}{2}Tr(\Sigma_t S_{t-1}^{-1})\right) \\
&\propto p(d_i)\exp\left(-\frac{d_i\nu\phi_i}{2} - \frac{1}{2}Tr(\nu A^{-1}Q^{-1})\right),
\end{aligned} \tag{32}
$$

where $\phi_i = \sum_{t=1}^{T}\log|\Gamma_{t-1,\ell_i}|$, and $Q^{-1} = \sum_{t=1}^{T}\left[\prod_{j=1}^{K}\Gamma_{t-1,\ell_j}^{\frac{-d_j}{2}}\right]\Sigma_t\left[\prod_{j=K}^{1}\Gamma_{t-1,\ell_j}^{\frac{-d_j}{2}}\right]$. To sample from this density we do the following. If d_i is the previous value in the chain we propose $d_i^{'} = d_i + u$ where $u \sim N(0, \sigma^2)$ and accept $d_i^{'}$ with probability

$$
\min\left\{\frac{p(d_i^{'}|A, d_{-i}, \ell, \nu, \Sigma_{1:T})}{p(d_i|A, d_{-i}, \ell, \nu, \Sigma_{1:T})}, 1\right\}, \tag{33}
$$

and otherwise retain d_i. σ^2 is selected to achieve a rate of acceptance between 0.3-0.5.

For $\ell_i, i = 2, \ldots, K$ we have,

$$
\begin{aligned}
p(\ell_i|A, d, \ell_{-i}, \nu, \Sigma_{1:T}) &\propto p(\ell_i) \times \prod_{t=1}^{T} Wishart_k(\Sigma_t|\nu, S_{t-1}) \\
&\propto p(\ell_i)\exp\left(-\frac{d_i\nu\phi_i}{2} - \frac{1}{2}Tr(\nu A^{-1}Q^{-1})\right),
\end{aligned} \tag{34}
$$

where ϕ_i and Q^{-1} are defined the same way as in the previous case. To sample from the conditional posterior we use a simple random walk proposal. The proposal distribution is a Poisson random variable multiplied by a random variable that takes on values 1 and -1 with equal probability. The density of the proposal is

$$q(\ell) = \begin{cases} \frac{\lambda^{\ell} e^{-\lambda}}{2\ell!} & \ell = 1, 2, \cdots \\ e^{-\lambda} & \ell = 0 \\ \frac{\lambda^{-\ell} e^{-\lambda}}{2(-\ell)!} & \ell = -1, -2, \cdots \end{cases}$$

In the empirical work $\lambda = 2$. Given the value ℓ_i in the Markov chain, the new proposal $\ell_i' \sim q(\ell)$ is accepted with probability

$$\min\left\{\frac{p(\ell_i'|A, d, \ell_{-i}, \nu, \Sigma_{1:T})}{p(\ell_i|A, d, \ell_{-i}, \nu, \Sigma_{1:T})}, 1\right\}. \tag{35}$$

Finally, ν has the conditional posterior density

$$\begin{aligned} p(\nu|A, d, \ell, \Sigma_{1:T}) &\propto p(\nu) \times p(\Sigma_{t:T}|A, d, \nu) \\ &= p(\nu) \prod_{t=1}^{T} \frac{|\Sigma_t|^{\frac{\nu-k-1}{2}} |S_{t-1}^{-1}|^{\frac{\nu}{2}}}{2^{\frac{\nu k}{2}} \prod_{j=1}^{k} \Gamma(\frac{\nu+1-j}{2})} \exp\left(-\frac{1}{2} Tr(\Sigma_t S_{t-1}^{-1})\right) \\ &= p(\nu) \frac{\prod_{t=1}^{T} |\Sigma_t|^{\frac{\nu}{2}} \times \nu^{\frac{T\nu k}{2}} |A^{-1}|^{\frac{T\nu}{2}} \times \prod_{t=1}^{T} \prod_{i=1}^{K} |\Gamma_{t-1,\ell_i}|^{-\frac{d_i \nu}{2}}}{2^{\frac{T\nu k}{2}} (\prod_{j=1}^{k} \Gamma(\frac{\nu+1-j}{2}))^T} \\ &\quad \times \exp\left(-\frac{1}{2} Tr(\nu A^{-1} Q^{-1})\right) \\ &\propto \exp\left(-\lambda_0 \nu + \frac{T\nu}{2} \log|A|^{-1} + \frac{T\nu k}{2} \log\frac{\nu}{2} - T \sum_{j=1}^{k} \log\Gamma\left(\frac{\nu+1-j}{2}\right)\right) \\ &\quad \times \exp\left(\frac{\nu}{2} \sum_{t=1}^{T} \log|\Sigma_t| - \frac{\nu}{2} \sum_{i=1}^{K} d_i \phi_i - \frac{1}{2} Tr(\nu A^{-1} Q^{-1})\right) \end{aligned} \tag{36}$$

where Q^{-1} and ϕ_i are defined as in previous cases. This is a nonstandard distribution which we sample using a Metropolis-Hastings step with a random walk proposal analogous to the sampling in the previous step above.

7.2 Wishart-RCOV-A(K) Estimation

Parameters are $\Theta = \{\nu, b_1, \ldots, b_k, \ell_2, \ldots, \ell_K\}$. Given Θ the data density is the product of the Wishart densities,

$$p(\Sigma_{1:T}|\Theta) = \prod_{t=1}^{T} Wishart_k(\Sigma_t|\nu, S_{t-1}). \tag{37}$$

The joint posterior distribution of the parameters $p(\Theta|\Sigma_{1:T})$ then is the product of the data density and the individual priors for each parameter. We iteratively sample from the conditional posterior distribution of each parameter conditional on the other parameters by Metropolis-Hastings scheme. For each parameter, a random walk with normal proposal is applied, except for $\ell_i, i = 2, \ldots, K$, in which case the proposal distribution is a "symmetric Poisson" as in Wishart-RCOV-M(K). For a particular parameter $\theta_i \in \Theta$, the conditional posterior distribution is:

$$\begin{aligned} p(\theta_i|\theta_{-i}, \Sigma_{1:T}) &\propto p(\theta_i) \times \prod_{t=1}^{T} Wishart_k(\Sigma_t|\nu, S_{t-1}) \\ &= p(\theta_i) \times p(\Sigma_{1:T}|\Theta) \\ &= p(\theta_i) \prod_{t=1}^{T} \frac{|\Sigma_t|^{\frac{\nu-k-1}{2}} |S_{t-1}^{-1}|^{\frac{\nu}{2}}}{2^{\frac{\nu k}{2}} \prod_{j=1}^{k} \Gamma(\frac{\nu+1-j}{2})} \exp\left(-\frac{1}{2} Tr(\Sigma_t S_{t-1}^{-1})\right) \end{aligned} \quad (38)$$

7.3 Sampling from $\Lambda|r_{1:T}, \Sigma_{1:T}$

To estimate Λ, define $\tilde{r}_t = \Sigma_t^{-1/2} r_t$, then $\tilde{r}_t \sim N(0, \Lambda)$. We let the prior of Λ^{-1} be $Wishart_k(\gamma_1, Q_1)$, and we set $Q_1 = I$ and $\gamma_1 = k + 1$. The posterior distribution of Λ^{-1} is

$$\begin{aligned} p(\Lambda^{-1}|\tilde{r}_{1:T}) &\propto Wishart_k(\Lambda^{-1}|\gamma_1, Q_1) \times \prod_{t=1}^{T} N(\tilde{r}_t|0, \Lambda) \\ &\propto Wishart_k(\Lambda^{-1}|\widehat{\gamma}, \widehat{Q}) \end{aligned} \quad (39)$$

by the conjugacy of the Wishart prior of the precision matrix with respect to a multivariate Normal likelihood. Here $\widehat{\gamma} = T + \gamma_1$, and $\widehat{Q} = (\sum_{t=1}^{T} \tilde{r}_t \tilde{r}_t' + Q_1^{-1})^{-1}$

7.4 DCC-t Estimation

The parameters are $\{\omega_1, \ldots, \omega_k, \kappa_1, \ldots, \kappa_k, \lambda_1, \ldots, \lambda_k, \alpha, \beta, \zeta\} = \Psi$. All parameters are assigned an independent Normal prior with mean 0 and variance 100, with the following restrictions are imposed:

$$\omega_i > 0,\ \kappa_i \geq 0,\ \lambda_i \geq 0,\ \zeta > 2,\ \frac{\kappa_i \zeta}{\zeta - 2} + \lambda_i < 1,\ i = 1, ..., k,\ \alpha \geq 0,\ \beta \geq 0,\ \alpha + \beta < 1. \quad (40)$$

The joint prior $p(\Psi)$ is just the product of the individual priors. The likelihood function $p(r_{1:T}|\Psi)$ is:

$$p(r_{1:T}|\Psi) = \prod_{t=1}^{T} \frac{\Gamma[(\zeta + k)/2][1 + \frac{1}{\zeta} r_t' H_t^{-1} r_t]^{-(\zeta+k)/2}}{\Gamma(\zeta/2)(\zeta\pi)^{k/2}|H_t|^{1/2}}. \quad (41)$$

The posterior of the parameters $p(\Psi|r_{1:T})$ is:

$$p(\Psi|r_{1:T}) \propto p(\Psi) \prod_{t=1}^{T} \frac{\Gamma[(\zeta + k)/2][1 + \frac{1}{\zeta} r_t' H_t^{-1} r_t]^{-(\zeta+k)/2}}{\Gamma(\zeta/2)(\zeta\pi)^{k/2}|H_t|^{1/2}}. \quad (42)$$

For the special case of DCC with Normal innovations, the parameters are

$$\{\omega_1, \ldots, \omega_k, \kappa_1, \ldots, \kappa_k, \lambda_1, \ldots, \lambda_k, \alpha, \beta, \} = \Psi.$$

The restriction on the priors are similar:

$$\omega_i > 0,\ \kappa_i \geq 0,\ \lambda_i \geq 0,\ \kappa_i + \lambda_i < 1,\ i = 1, ..., k,\ \alpha \geq 0,\ \beta \geq 0,\ \alpha + \beta < 1. \tag{43}$$

The posterior of the parameters $p(\Psi|r_{1:T})$ is:

$$p(\Psi|r_{1:T}) \propto p(\Psi)(2\pi)^{\frac{Tk}{2}} \prod_{t=1}^{T} |D_t R_t D_t|^{-\frac{1}{2}} \times \exp\left(-\frac{1}{2}\sum_{t=1}^{T} r_t'(D_t R_t D_t)^{-1} r_t\right) \tag{44}$$

To sample from the joint posterior distribution $p(\Psi|r_{1:T})$, we do the following steps: We first adopt a single move sampler. For each iteration, the chain cycles through the conditional posterior densities of the parameters in a fixed order. For each parameter, a random walk with normal proposal is applied. After dropping an initial set of draws as burnin, we collect M draws and use them to calculate the sample covariance matrix of the joint posterior. Then, a block sampler is used to jointly sample the full posterior. The proposal density is a multivariate normal random walk with the covariance matrix set to the sample covariance, obtained from the draws of the single-move sampler, scaled by a scalar. When the model is recursively estimated as a new observation arrives the previous sample covariance is used as the next covariance in the multivariate normal random walk. This results in fast efficient sampling.

7.5 VD-GARCH-t Estimation

The parameters are $\{a_1, \ldots, a_k, b_1, \ldots, b_k, \zeta\} = \Psi$. All parameters are assigned an independent Normal prior with mean 0 and variance 100, with a_1 and b_1 restricted to be positive for identification purpose. The joint prior $p(\Psi)$ is just the product of the individual priors. The likelihood function $p(r_{1:T}|\Psi)$ is:

$$p(r_{1:T}|\Psi) = \prod_{t=1}^{T} \frac{\Gamma[(\zeta + k)/2][1 + \frac{1}{\zeta} r_t' H_t^{-1} r_t]^{-(\zeta+k)/2}}{\Gamma(\zeta/2)(\zeta\pi)^{k/2}|H_t|^{1/2}} \tag{45}$$

The posterior of the parameters $p(\Theta|r_{1:T})$ is:

$$p(\Psi|r_{1:T}) \propto p(\Psi) \prod_{t=1}^{T} \frac{\Gamma[(\zeta + k)/2][1 + \frac{1}{\zeta} r_t' H_t^{-1} r_t]^{-(\zeta+k)/2}}{\Gamma(\zeta/2)(\zeta\pi)^{k/2}|H_t|^{1/2}} \tag{46}$$

The sampling procedure is similar to that of the DCC.

Reference

1. Amisano, G., R. Giacomini (2007): "Comparing Density Forecasts via Weighted Likelihood Ratio Tests", *Journal of Business and Economic Statistics*, 25(2), 177-190.

2. Andersen, T., T. Bollerslev, F. X. Diebold (2007), "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility", *Review of Economics and Statistics*, 89, 701-720.

3. Andersen, T., T. Bollerslev, F. X. Diebold (2009), "Parametric and Nonparametric Volatility Measurement", in *Handbook of Financial Econometrics*, VOL 1, Eds Ait-Sahalia and Hansen, Elsevier.

4. Andersen, T., T. Bollerslev, F. X. Diebold, P. Labys (2003), "Modeling and Forecasting Realized Volatility", *Econometrica*, 71, 529-626.

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

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

7. Bandi, F. M., Russell, J. R. (2005b), "Realized covariation, realized beta and microstructure noise", Unpublished paper. Graduate School of Business, University of Chicago.

8. Bao, Y., T.-H. Lee, and B. Saltoglu (2007), "Comparing Density Forecast Models", *Journal of Forecasting*, 26(3), 203-225.

9. Barndorff-Nielsen, O. E., Hansen, P. H., Lunde, A., Shephard, N. (2008), "Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading", working paper, SSRN: http://ssrn.com/abstract=1154144

10. Barndorff-Nielsen, O. E., Shephard, N. (2004b), "Econometric analysis of realised covariation: High Frequency based covariance, regression and correlation in financial economics", *Econometrica*, 72, 885-925.

11. Bauer, G. H., Vorkink, K. (2010), "Forecasting Multivariate Realized Stock Market Volatility", forthcoming *Journal of Econometrics*

12. Bauwens, L., S. Laurent, J. Rombouts (2006), "Multivariate GARCH Models: a Survey", *Journal of Applied Econometrics*, 21, 79-109.

13. Chib S. (2001) "Markov Chain Monte Carlo Methods: Computation and Inference", in *Handbook of Econometrics*, Heckman and Leamer Eds., Elsevier Science.

14. Chiriac, R., Voev, V. (2010): "Modelling and Forecasting Multivariate Realized Volatility", forthcoming *Journal of Applied Econometrics*

15. Chiriac, R. (2006), "Estimating Realized Volatility Wishart Autoregressive Model", Working Paper, University of Konstanz.

16. Corsi, F. (2009), "A Simple Approximate Long-Memory Model of Realized Volatility",*Journal of Financial Econometrics*, Spring 7: 174 - 196.

17. Engle, R. F., Colacito, R. (2006), "Testing and Valuing Dynamic Correlations for Asset Allocation", *Journal of Business and Economic Statistics*, 24, 239-250.

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

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

20. Gauthier, P. and D. Possamai (2009) "Efficient Simulation of the Wishart model", http://ssrn.com/abstract=1474728

21. Geweke J. (2005) *Contemporary Bayesian Econometrics and Statistics*, Wiley.

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

23. Greene, W. H. (2002), *Econometric Analysis* (5th Edition), Prentice Hall.

24. Kass R. E. and A. E. Raftery (1995) "Bayes Factors", *Journal of the American Statistical Association*, 90(430), 773-795.

25. Maheu, J. M., T. H. McCurdy (2009), "Do high-frequency measures of volatility improve forecasts of return distributions?", forthcoming, *Journal of Econometrics.*

26. McAleer M. and M. Medeiros (2008) "Realized Volatility: A Review", *Econometric Reviews*, 27(1-3), pages 10-45.

27. Muirhead, R. J., (1982), *Aspects of Multivariate Statistical Theory*, Wiley, New York.

28. Philipov, A., M. E. Glickman (2006), "Multivariate Stochastic Volatility via Wishart Process", *Journal of Business and Economic Statistics*, 24(3), 313-328.

29. Protter, P. (2004), *Stochastic Integration and Differential Equations*, Springer-Verlag, New York.

30. Tsay, R. S., (2005), *Analysis of Financial Time Series*, Wiley.

31. Weigend, A. S., S. Shi (2000), "Predicting Daily Probability Distributions of S&P500 Returns", *Journal of Forecasting*, 19, 375-392.

Table 1: Average daily number of transactions and average daily refresh time (RT) observations per day

SPY	GE	C	AA	BA	RT
6985	7479	6121	3279	3745	1835

This table reports the average daily number of transactions (after data cleaning) for Standard and Poor's Depository Receipt (SPY), General Electric Co. (GE), Citigroup Inc.(C), Alcoa Inc. (AA) and Boeing Co. (BA). The total number of days is 2281. RT reports the average number of daily observations according to the refresh time.

Table 2: Summary statistics: Daily returns and RCOV

	Sample covariance from daily returns					Average of realized covariances				
	SPY	GE	C	AA	BA	SPY	GE	C	AA	BA
SPY	0.963	1.078	1.172	0.834	0.751	0.907	0.972	1.099	0.822	0.718
GE		2.410	1.500	1.062	0.931		2.327	1.250	0.897	0.796
C			2.826	1.014	0.931			3.176	0.982	0.835
AA				3.900	0.993				3.921	0.734
BA					2.933					2.910

This table reports the sample covariance from daily returns and the sample average of the realized covariances. The data are Standard and Poor's Depository Receipt (SPY), General Electric Co. (GE), Citigroup Inc.(C), Alcoa Inc. (AA) and Boeing Co. (BA). Total observations is 2281.

Table 3: Mean square error for different models

3-comp	2-comp	1-comp	$E(\Sigma_t \vert \Sigma_{t-1}) = \Sigma_{t-1}$	DCC	Sample covariance
70.21	70.80	75.12	102.03	85.20	118.52

Given a model's in-sample fitted value $\hat{\Sigma}_t$ and the data Σ_t this table reports $\sum_{t=1}^{T} \|\Sigma_t - \hat{\Sigma}_t\|^2$, where $\|\cdot\|$ denotes the Frobenius matrix norm.

Table 4: Posterior mean and standard deviation of lower triangular elements of A for Wishart-RCOV-M(3)

1.0480				
(0.0100)				
0.0129	1.0832			
(0.0064)	(0.0087)			
0.0176	0.0010	1.0951		
(0.0064)	(0.0062)	(0.0089)		
0.0058	0.0059	0.0045	1.1138	
(0.0061)	(0.0060)	(0.0061)	(0.0102)	
0.0029	0.0019	0.0018	0.0037	1.1082
(0.0060)	(0.0060)	(0.0061)	(0.0062)	(0.0093)

This table reports the posterior mean, and the posterior standard deviation in parentheses for the lower triangle of A.

Table 5: Estimation results for Wishart-RCOV-M(3)

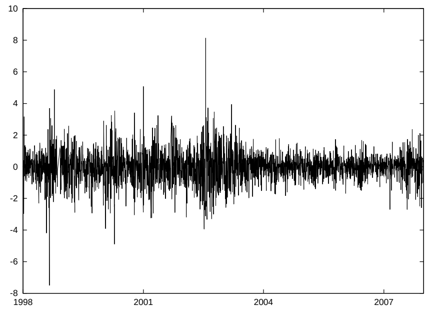
Parameter	Mean	NSE	0.95 DI	Ineff
d_1	0.2553	0.0004	$(0.2415, 0.2671)$	17.6101
d_2	0.4502	0.0006	$(0.4303, 0.4715)$	17.0676
d_3	0.2651	0.0006	$(0.2413, 0.2858)$	15.5695
ν	14.6679	0.0032	$(14.4736, 14.8603)$	5.3509
ℓ_2	9.0280	0.0219	$(8.0000, 10.0000)$	13.5203
ℓ_3	64.1822	0.0294	$(63.0000, 67.0000)$	3.0019

This table reports the posterior mean, its numerical standard error (NSE), a 0.95 density interval (DI) and the inefficiency factor for model parameters.

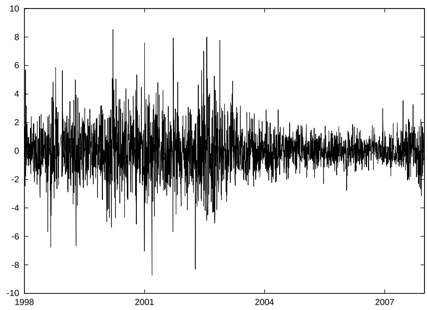
Table 6: Estimation results for Wishart-RCOV-A(3)

Parameter	Mean	NSE	0.95 DI	Ineff
b_{11}	0.5744	0.0012	$(0.5556, 0.5905)$	55.6555
b_{12}	0.5579	0.0019	$(0.5354, 0.5783)$	90.5498
b_{13}	0.5995	0.0017	$(0.5835, 0.6155)$	138.5520
b_{14}	0.4888	0.0005	$(0.4628, 0.5127)$	3.5098
b_{15}	0.5878	0.0010	$(0.5668, 0.6077)$	25.4376
b_{21}	0.6732	0.0010	$(0.6518, 0.6932)$	28.8173
b_{22}	0.6536	0.0023	$(0.6314, 0.6821)$	104.4400
b_{23}	0.6536	0.0013	$(0.6381, 0.6691)$	73.1365
b_{24}	0.6918	0.0005	$(0.6649, 0.7174)$	3.7541
b_{25}	0.5623	0.0017	$(0.5290, 0.5963)$	33.6903
b_{31}	0.4242	0.0010	$(0.3992, 0.4519)$	16.6831
b_{32}	0.4854	0.0024	$(0.4544, 0.5111)$	76.9035
b_{33}	0.4410	0.0019	$(0.4187, 0.4644)$	85.4987
b_{34}	0.4475	0.0010	$(0.4066, 0.4833)$	7.9212
b_{35}	0.5384	0.0013	$(0.5064, 0.5709)$	17.9957
ν	14.6666	0.0037	$(14.4875, 14.8439)$	4.6488
ℓ_2	8.9967	0.0031	$(9.0000, 9.0000)$	8.7925
ℓ_3	63.8190	0.0379	$(62.0000, 66.0000)$	3.3739

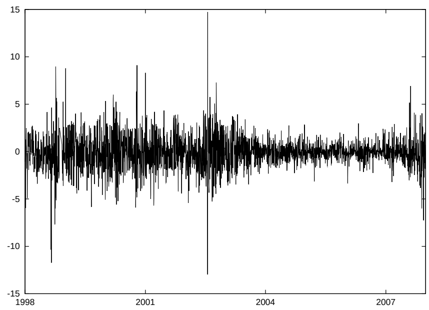
This table reports the posterior mean, its numerical standard error (NSE), a 0.95 density interval (DI) and the inefficiency factor for model parameters.



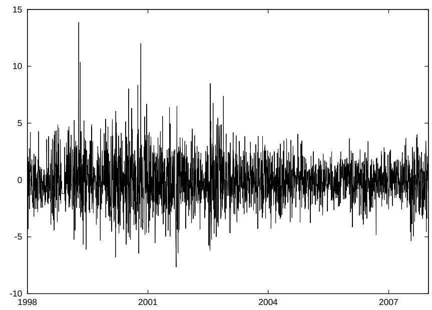
(a) SPYDER



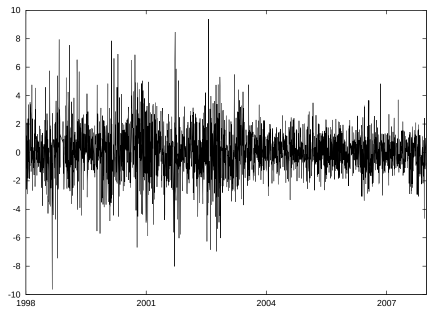
(b) GE



(c) Citigroup Inc.

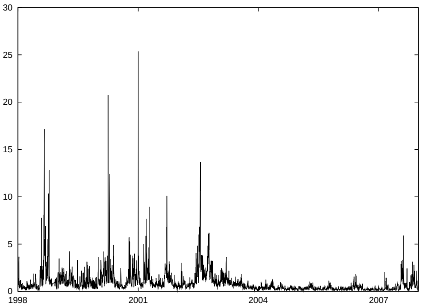


(d) Alcoa Inc.

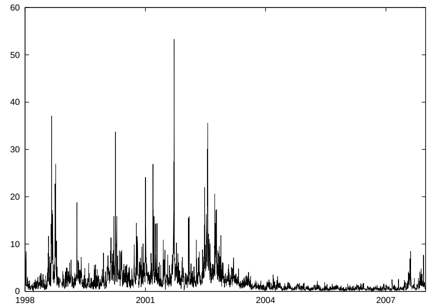


(e) Boeing Co.

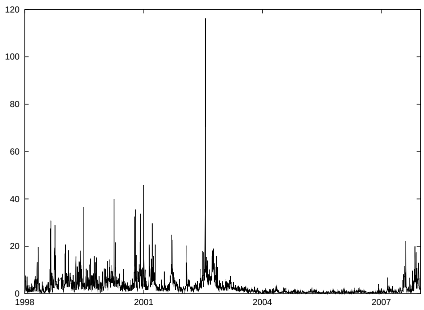
Figure 1: Daily returns



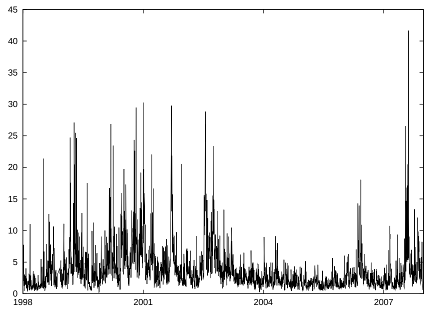
(a) SPYDER



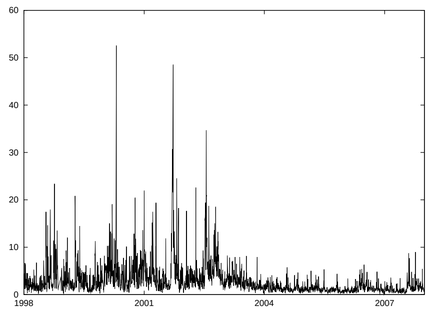
(b) GE



(c) Citigroup Inc.



(d) Alcoa Inc.



(e) Boeing Co.

Figure 2: RV for individual assets

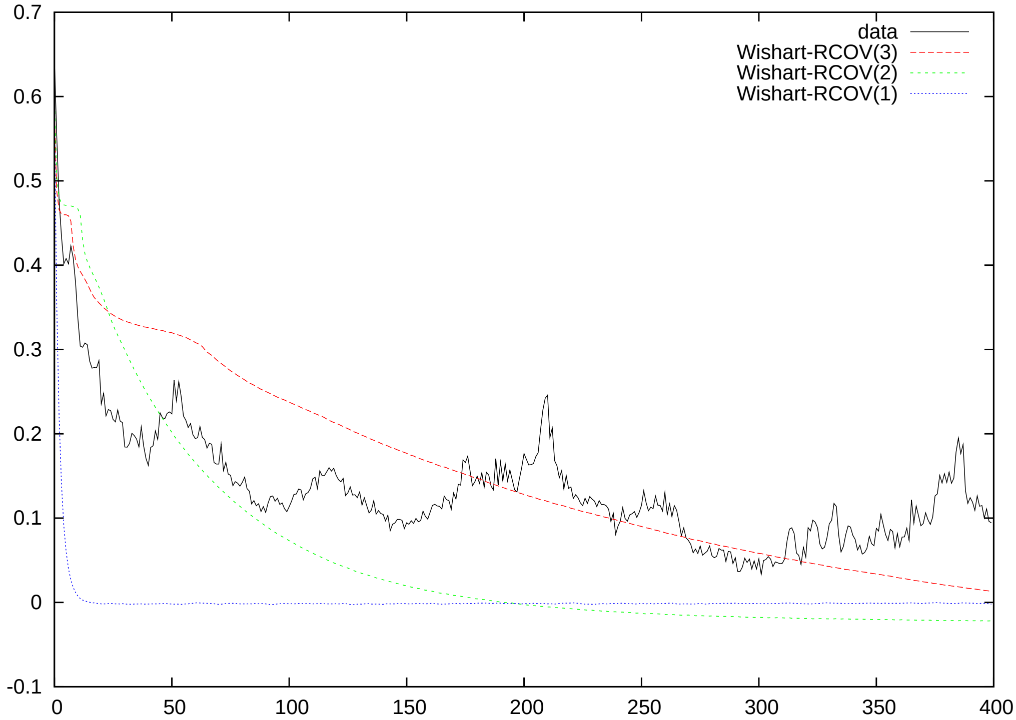


Figure 3: Sample autocorrelation functions of the largest eigenvalues of RCOVs

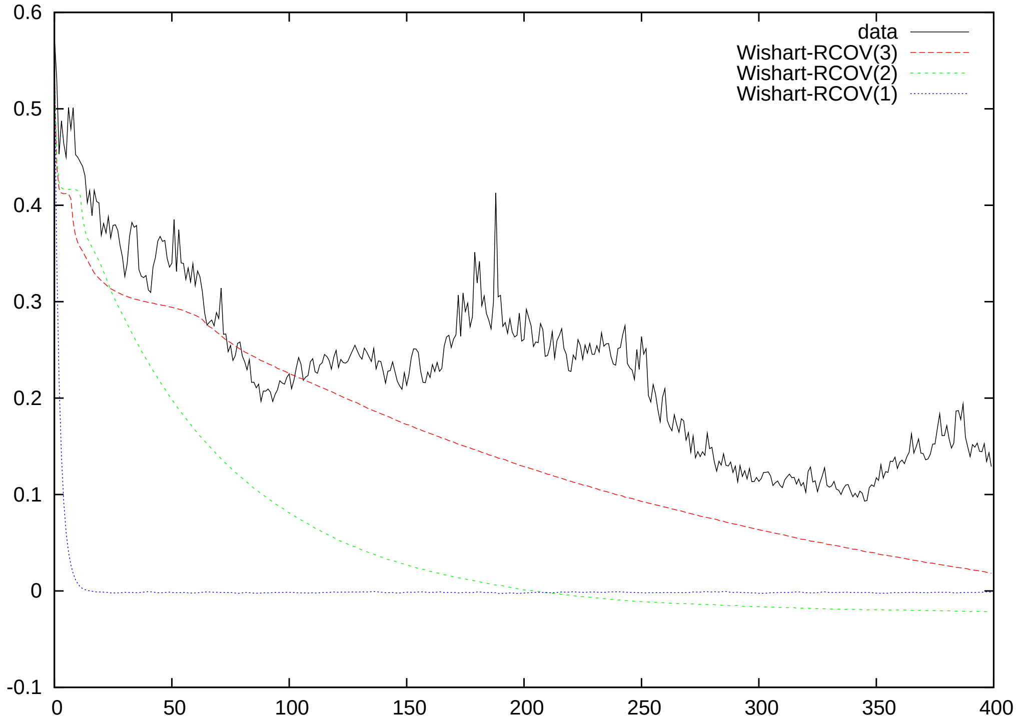


Figure 4: Sample autocorrelation functions of the smallest eigenvalues of RCOVs

Figure 5: Correlation between SPYDER and GE

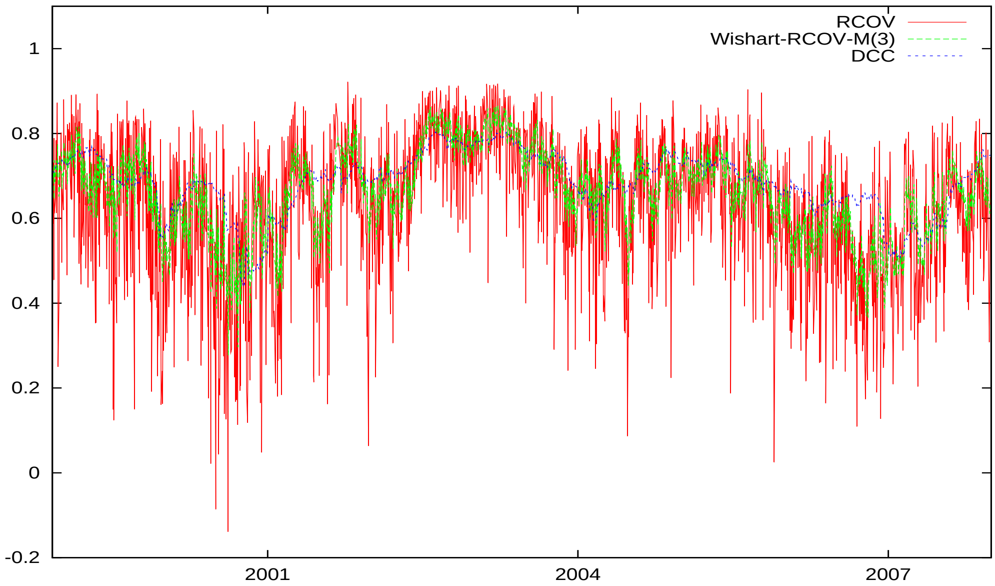
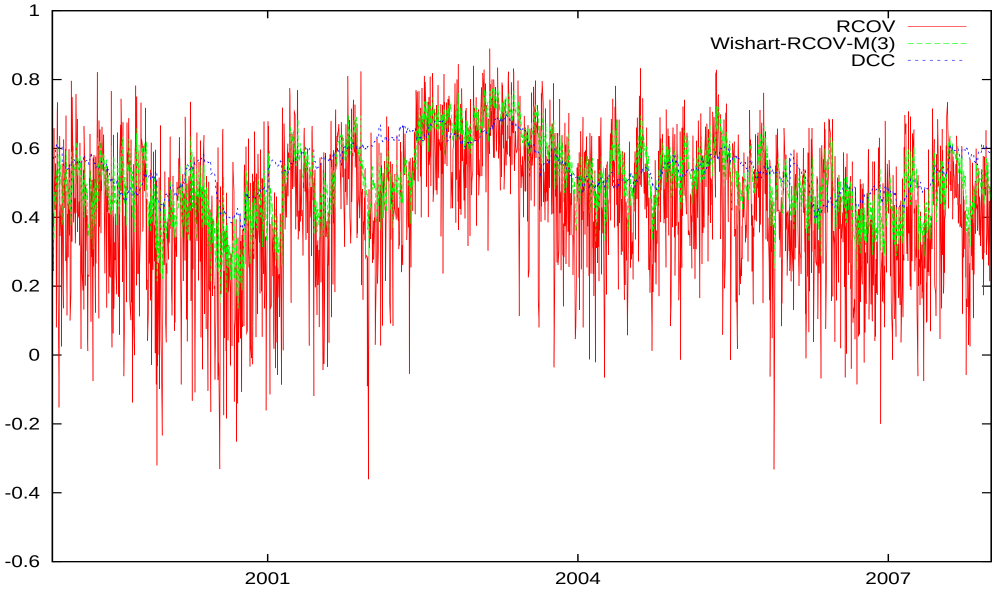


Figure 6: Correlation between GE and Citigroup Inc.



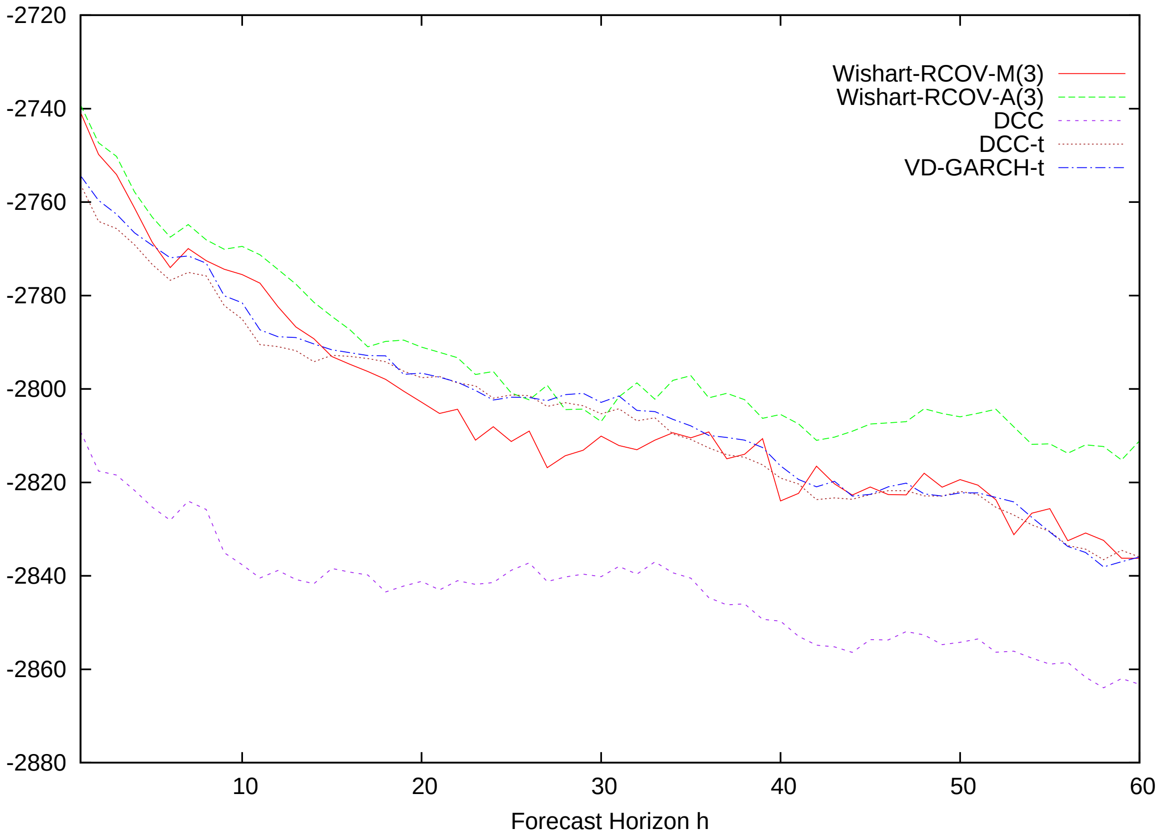


Figure 7: Term structure of cumulative log-predictive likelihoods for different models

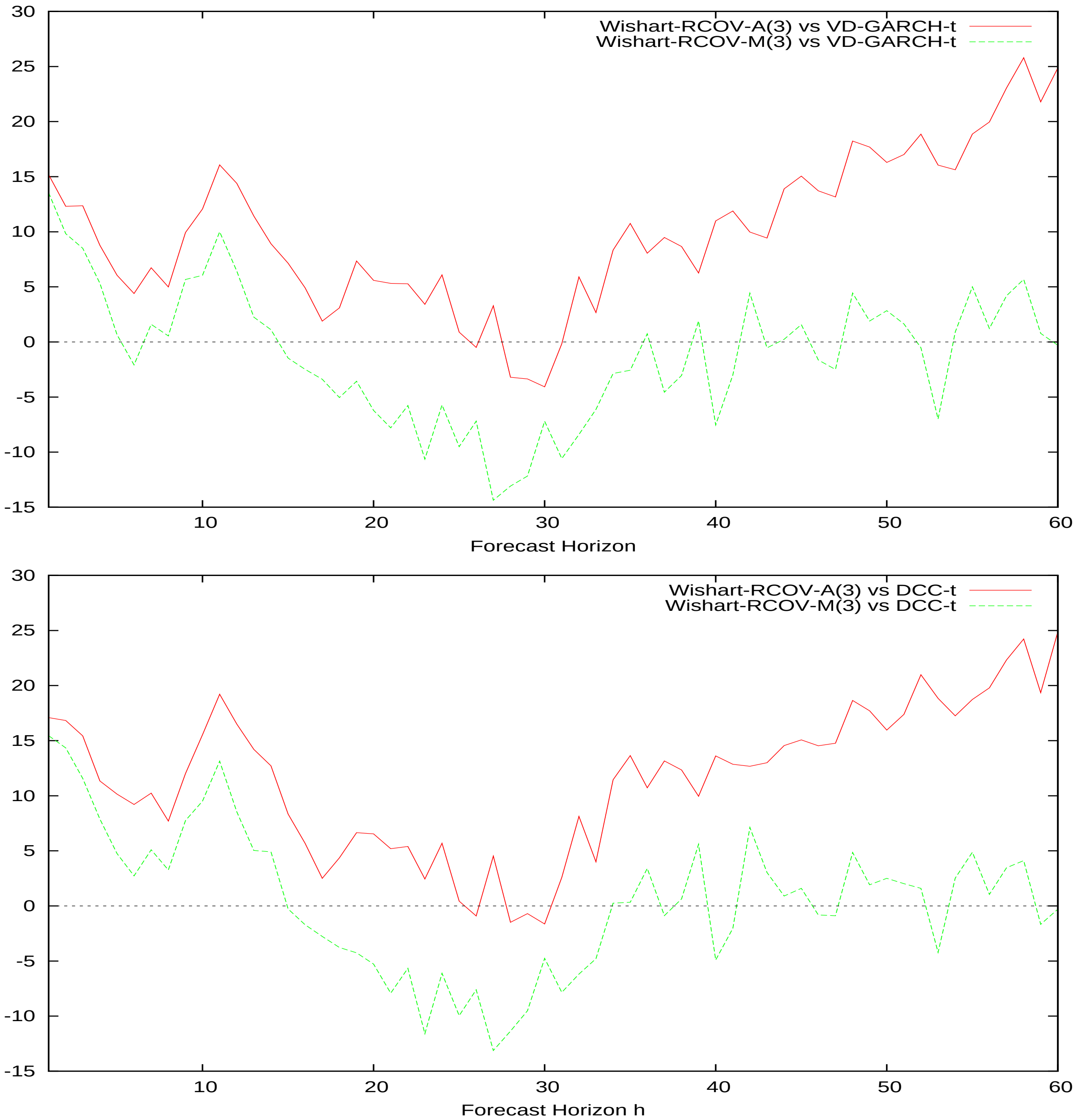


Figure 8: Log Predictive Bayes Factors

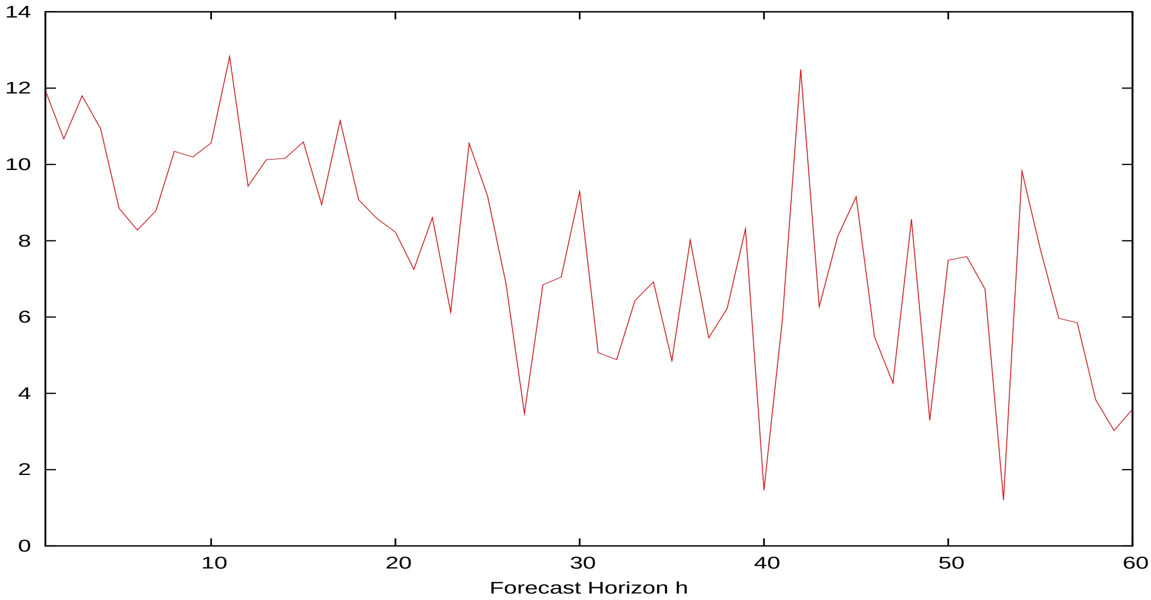


Figure 9: Log Predictive Bayes Factor: Λ estimated vs $\Lambda = I$

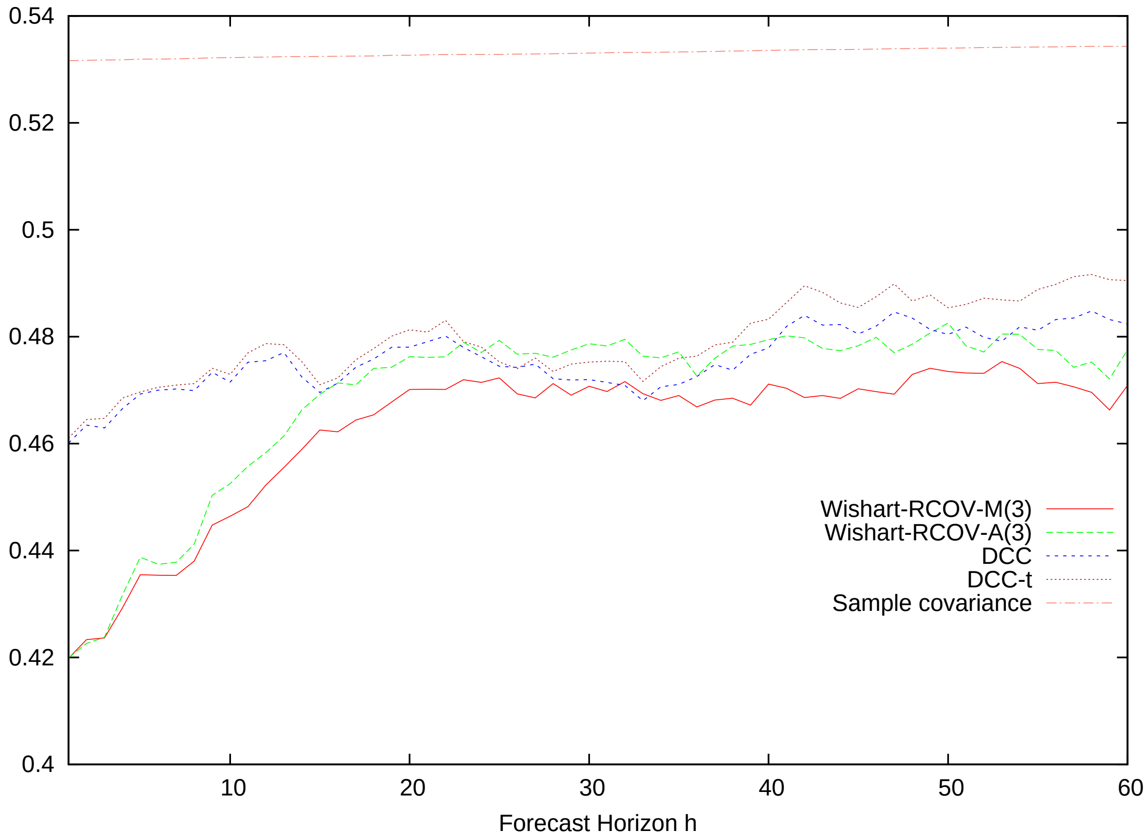


Figure 10: Sample variances of GMVPs across models against forecast horizon