



A Comparison of the Forecasting Performances of Multivariate Volatility Models

Vincenzo Candila

ISSN: 1971-3029

Dipartimento di Scienze Economiche e Statistiche
Università Degli Studi di Salerno
Via Ponte Don Melillo – 84084; Fisciano (SA) – Italy

Tel +39-089-96.21.55
Fax +39-089-96.20.49
E-mail dises@unisa.it
Web www.dises.unisa.it

A COMPARISON OF THE FORECASTING PERFORMANCES OF MULTIVARIATE VOLATILITY MODELS

Vincenzo Candila*

Abstract. The consistent ranking of multivariate volatility models by means of statistical loss function is a challenging research field, because it concerns the quality of the proxy chosen to replace the unobserved volatility, the set of competing models to be ranked and the kind of loss function. The existent works only consider the ranking of multivariate GARCH (MGARCH) models, based on daily frequency of the returns. Less is known about the behaviour of the models that directly use the realized covariance (RCOV), the proxy that generally provides a consistent estimate of the unobserved volatility. The aim of this paper is to evaluate which model has the best forecast volatility accuracy, from a statistical and economic point of view. For the first point, we empirically rank a set of MGARCH and RCOV models by means of four consistent statistical loss functions. For the second point, we evaluate if these rankings are coherent with those resulting from the use of an economic loss function. The evaluation of the volatility models through the economic loss function is usually done by looking at the Value at Risk (VaR) measures and its violations. A violation occurs every time the portfolio losses exceed the VaR. To assess the performances of the volatility models from an economic point of view, different tests regarding the violations have been proposed. In this work, the unconditional and conditional tests are considered. The analysis is based on a Monte Carlo experiment that samples from a trivariate continuous-time stochastic process a vector of observation each five minutes per two years.

Keywords: Volatility, Multivariate GARCH, Loss function.

JEL classifications: C10, C32, C52, C53, G10.

1. Introduction

In the literature there is a substantial debate on the quantification of the risk related to holding financial instruments, like for instance assets listed on a Stock Exchange. The specification of the risk is a challenging task for several reasons, among which is the lack of a commonly accepted definition. Moreover, there are many different risk measures, each of which has to own precise characteristics. But it is undoubted that the greatest problem relies on the latent nature of the variable “risk” that we do not observe. Even though we had a risk measure, how could be sure that our estimate is close to the real but unobserved latent variable? This task is arduous because it involves three aspects: first, the measure of the risk as much as possible close to the latent variable; second, the forecast of the risk, deriving from a model; third, a function that assesses the forecast accuracy, comparing the former to the latter quantity. This paper fits between the last two points, considering some specifications of multivariate models and comparing them with the proxy of the unobserved latent variable by means of statistical and economic loss functions.

*DISES, Via Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy, vcandila@unisa.it

The risk of holding assets listed on a Stock Exchange is identified with the volatility, that is the unobserved variable of interest. With respect to the previous first point, a volatility proxy is represented by the co-variation of the assets, that indeed is not constant over time (Mandelbrot (1963)). Therefore, the more an asset varies over time, the more it is said to be volatile. A widely accepted volatility proxy is represented by the realized volatility, ascribable to Andersen and Bollerslev (1998). Barndorff-Nielsen and Shephard (2004) provide the theoretical foundations of this approach, while in McAleer and Medeiros (2008) can be found a survey of the recent literature. Acting in a multivariate context, the proxy of the volatility is the ex-post realized covariance matrix as defined in Andersen et al. (2003). The realized volatility technique measures ex-post the variability of the assets by means of cumulative squared intraday returns¹. In this work, we use the realized covariance as volatility proxy, as defined in the work of Andersen et al. (2003).

With respect to the second point, that is a set of forecasting models that compete each other for the estimation of volatility, it is undoubted that a large literature has arisen during the last two decades. Among all the approaches, we focus on the multivariate GARCH models and some specifications that identify the volatility as a combination of the realized covariance. The first approach, of which a review can be found in Bauwens et al. (2003), estimates the (expected) volatility as a function of past returns and other observable variables. For the second family of volatility forecasts, we consider two specifications: the Rolling Covariance model (Fleming et al. (2003); Bandi et al. (2008); de Pooter et al. (2008)) and the Conditionally Autoregressive Wishart (CAW) model, introduced by Golosnoy et al. (2012).

Finally, with respect to the third point cited above, in the literature there are two approaches to evaluate the forecast accuracy: by using a statistical or an economic loss function. Usually these functions are used separately, in the sense that the results of the evaluation are not compared with each other. This paper aims to compare both the approaches. Firstly, some consistent statistical loss function are used in order to assess the distance between the volatility proxy and the forecast. A loss function is said to be consistent or robust² if the ranking of any two volatility forecasts is the same of the ranking that would be obtained if the true volatility had been observable. In the univariate field, Hansen and Lunde (2006) provide the sufficient condition for a loss function to be consistent and Patton (2006). For the multivariate framework, Laurent et al. (2013) define necessary and sufficient conditions for a loss function to be consistent. Taking as inspiration this last work, we evaluate the forecast performances of a set of multivariate GARCH models, of the Rolling Covariance and CAW models. The evaluation is realized by using four consistent loss functions, two symmetric and two asymmetric. The asymmetric loss function computes penalizes differently the over and the under predictions.

However the assessment of the forecast accuracy can be also done through a different approach, as stated above, based on an economic point of view. We are motivated to check if the models that are closer to the volatility proxy (as resulting by a consistent statistical loss function) are also the ones that exhibit better economic performances. This is done using the Value at Risk method, that represents the potential losses that a portfolio exhibits over a defined period for a given confidence interval. To (indirectly) evaluate the performances of the set of forecasting models we look at the VaR violations, the number of days on which the portfolio losses exceed the VaR. For a comprehensive overview of Value at Risk and

¹Nelson (1990) setted the continuous-time stochastic volatility framework.

²The adjective robust has not to be confused with its usual meaning: in estimation theory, the term 'robust' is related to the estimators that are insensitive to the presence of outliers in the data while in this context the term robust refers to the loss functions that are insensitive to the noise in the volatility proxy. Hence, a 'robust' (or consistent) loss function could be not robust to the presence of outliers in the data.

its measures, see Duffie and Pan (1997) and Jorion (2007). Even though the VaR has been criticized for its statistical properties (Artzner et al. (1999)), it still have an important role in Basel III (see EBA Guidelines on Stressed Value at Risk, 2012).

The aim of this paper is to evaluate which model has the best forecast volatility accuracy, from a statistical and economic point of view. Moreover, we are interested in investigating the forecast accuracy of the Rolling Covariance and CAW models, specifications that directly use the realized covariance with respect to the forecast accuracy of the MGARCH models.

A Monte Carlo experiment is carried out for the evaluation of the forecasting models. In particular, the data generating process is a trivariate continuous-time stochastic process, where we assume that each instantaneous variance is the GARCH(1,1) diffusion as proposed by Andersen and Bollerslev (1998).

The main results of the paper are as follows: first, we do not find a clear correspondence between the ranking produced by the statistical loss function and by the economic loss function. For a risk manager, decisions only based on the VaR violations could lead to use models that are far from the true (but unobserved) volatility. Second, the scalar CAW and Rolling Covariance assure better forecast accuracy than the standard multivariate GARCH models. Moreover, the importance of the covariances in the evaluation of the forecast accuracy seems to be low. The ranking given by the loss function that uses only the variances is the same of the ranking given by the loss function that uses both variances and covariances.

The remainder of this work is organised as follows. Section 2 describes the statistical and economic methods to compare the volatility models. Section 3 presents an overview of the volatility models used in this work. The setting of the Monte Carlo experiment is in Section 4 and the answers to our questions are in Section 5. The conclusions and some suggestions for future research are in Section 6.

2. Ranking of volatility models

2.1 Statistical loss functions

A loss function L is a function that reports the distance between the actual and forecasted value for a variable. For the univariate case, it assumes the form:

$$L(x, \hat{x}), \quad (1)$$

where x is the actual value of the variable of interest and \hat{x} is the forecasted value. If we have several estimates of \hat{x} indexed by m , such that \dot{X} is the set of estimates and $\hat{x}_m \in \dot{X}$, it is possible to rank these estimates on the basis of the loss function: obviously, the smaller the loss function is, the better that estimate is. If the loss function has the following three features, it is said to be *well defined*.

Assumption 1. $L(\cdot, \cdot)$ is continuous in \dot{X} and it is minimized at \hat{x}^* which represents the optimal forecast.

Assumption 2. $L(\cdot, \cdot)$ is such that the optimal forecast \hat{x}^* equals the true value, formally:

$$\hat{x}^* = \underset{\hat{x} \in \dot{X}}{\operatorname{argmin}} L(x, \hat{x}) \iff \hat{x}^* = x. \quad (2)$$

Assumption 3. The loss function gives zero loss then $\hat{x}^* = x$.

In addition to the previous assumptions, a loss function yields an increasing penalty if the distance between x and \hat{x} increases.

Unfortunately, in the framework we are in, we do not have the true value for the volatility. This is because the volatility is a latent variable. In an multivariate time series context, the (univariate) loss function presented above becomes:

$$L(\Sigma_t, H_t), \quad (3)$$

where Σ_t is the true but unobservable covariance matrix and H_t is its estimate. The consistency of ranking between any two models m and l , whose estimated conditional covariance matrix are $H_{m,t}$ and $H_{l,t}$, respectively, with $m \neq l$, holds if

$$E(L(\Sigma_t, H_{l,t})) \geq E(L(\Sigma_t, H_{m,t})) \iff E(L(\hat{\Sigma}_t, H_{l,t})) \geq E(L(\hat{\Sigma}_t, H_{m,t})), \quad (4)$$

where $L(\cdot, \cdot)$ is a well defined loss function. The definition (4) has a fundamental importance: even though we do not observe the true conditional covariance matrix, we can consistently order any estimate of it. The conditions, given by Laurent et al. (2013), that make a loss function consistent are: (i) replacing Σ_t with any conditionally unbiased proxy; (ii) the well defined $L(\cdot, \cdot)$ is twice continuously differentiable with respect to $\hat{\sigma}_t$ and h_t , indicating the element of matrices $\hat{\Sigma}_t$ and H_t , respectively; (iii) the second derivative $\frac{\partial^2 L(\Sigma_t, H_t)}{\partial \sigma_{l,t} \partial \sigma_{m,t}}$ is finite and independent of $H_t, \forall l, m$.

However, not all the most common loss functions are consistent. Laurent et al. (2013) define a family of consistent loss functions based on the (observed) forecast error $\hat{\Sigma}_t - H_t$, that assumes the following quadratic form:

$$L(\hat{\Sigma}_t, H_t) = \text{vech}(\hat{\Sigma}_t - H_t)' \Lambda \text{vech}(\hat{\Sigma}_t - H_t), \quad (5)$$

where the $\text{vech}(\cdot)$ is the operator that stacks the lower triangular portion of a matrix into a vector and Λ is a matrix that assigns the weights to each element of forecast error matrix $\hat{\Sigma}_t - H_t$. In this work, we use four specifications of Λ , as summarized in Table 1, in order to have two symmetric and two asymmetric loss functions.

- insert Table 1 about here -

The Euclidean distance is the matrix version of the Mean Squared Error distance. The Λ matrix is a diagonal matrix of 1s, such that each forecast error term is first squared and then summed, $\forall t$. The squared weighted Euclidean distance considers only the variances: the Λ matrix is a diagonal matrix of 0 and 1, such that only the forecast error for the variances is computed. The *over prediction* version of the Mahalanobis distance penalizes the negative forecast errors, that are present when the forecasted value is larger than the correspondent value of the volatility proxy. If this happens, the diagonal Λ matrix is such that the negative terms of the forecast error $\hat{\Sigma}_t - H_t$ count twice. The last loss function we use in this work is the opposite of the Mahalanobis distance just presented: the under prediction version penalizes the cases in which the term of the forecast error is positive, meaning that the forecasted value has been under predicted. If this happens, the diagonal Λ matrix is such that each positive term of the forecast error matrix counts twice, $\forall t$.

2.2 Economic loss functions

The economic loss function provides an indirect evaluation of the risk, because we do not directly assess the distance between a volatility proxy and the volatility as obtained by a

forecasting model. In fact we first compute some risk measures and then we verify if these risk measures are coherent with our assumptions. Moreover, we move from multivariate to univariate context, with the referring to *portfolio* return and variance. The indirect assessment is done in two steps: first, we calculate a (daily) ex ante risk measure; second, we evaluate this risk measure against an ex post realized portfolio loss. To assure that the comparability of the models is only based on the $H_{m,t}$ matrix, we give the same weights to each asset, such that the resulting portfolio variance changes over the models only for the conditional covariance matrix. Let $\sigma_t^2(p, m) = wH_{m,t}w'$ be the portfolio variance at time t for the model m , and w the (fixed) $1 \times k$ vector of weights. Let $r_t(p) = wr_t$ be the portfolio return at time t . The ex ante risk measure we use is the Value at Risk (VaR), calculated by means of the mean-variance approach. Other risk measures normally used in this context are the Expected Shortfall and the distributional forecast, for instance. For a survey of the Value at Risk against the other traditional risk measures, see Kaplanski and Kroll (2002).

The VaR_t , representing the Value at Risk constructed on day $t - 1$ for the day t , for long trading position, is obtained as:

$$VaR_t = r_t(p) + f_\alpha \sigma_t(p), \quad (6)$$

where f_α is the left quantile at $\alpha\%$ of the distribution f . The sense of α is that

$$Pr(r_t(p) < VaR_t) = \alpha. \quad (7)$$

In other words, α represents the probability that the portfolio loss on day t exceeds VaR_t . Observing the series of VaR_t , for $t = 1, \dots, T$ and portfolio daily returns, we can define the sequence of hit function $\{I_t\}_{t=1}^T$ as the number of VaR violations occurring in a given time period, where the hit function at time t is obtained as follows:

$$I_t = \begin{cases} 1 & \text{if } r_t(p) < VaR_t \\ 0 & \text{if } r_t(p) > VaR_t \end{cases} \quad (8)$$

The assessment of a model through the VaR and the hit function can be done in many methods, among which there are the Time Until First Failure (TUFF) test, the unconditional coverage test and the independence test. The TUFF test, proposed by Kupiec (1995), reports the first day in which a VaR violation occurs. The unconditional coverage test, always proposed by Kupiec, tests if the empirical frequency of violations is statistically equal to the prefixed α . The independence test, due to Christoffersen (1998), checks if the VaR violations are clustered in time or are independently distributed over time. In this work we only consider the last two.

The null hypothesis of the unconditional coverage test is $H_0 : E[I_t] \equiv \pi = \alpha$, where π stands for the unconditional probability of a violation of a model. Assuming the independence of I_t for each t , the likelihood of the hit sequence will be given by the productory of a Bernoulli random variable, that is:

$$L(\pi) = \prod_{t=1}^T (1 - \pi)^{1-I_t} \pi^{I_t} = (1 - \pi)^{T_0} \pi^{T_1},$$

where T_0 and T_1 the number of 0s and 1s, that is the number of non-violations and violations of the VaR in the sample, respectively. Let $\hat{\pi} = T_1/T$ the observed number of violations in the sample. If we insert $\hat{\pi}$ into the likelihood, we have:

$$L(\hat{\pi}) = (1 - T_1/T)^{T_0} (T_1/T)^{T_1}.$$

Under the null hypothesis that $\pi = \alpha$, we have the following likelihood:

$$L(\alpha) = \prod_{t=1}^T (1 - \alpha)^{1 - I_{t+1}} \alpha^{I_{t+1}} = (1 - \alpha)^{T_0} \alpha^{T_1}.$$

Finally, we can check the null by using the likelihood test

$$LR_{un} = -2 \log [L(\alpha) / L(\hat{\pi})] \sim \chi_1^2. \quad (9)$$

Also the independence test is based on a likelihood ratio test. If the hit sequence is dependent over time such that it can be defined as a first-order Markov sequence with the following transition probability matrix:

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix},$$

where π_{01} is the probability that, given today being a non-violation, tomorrow a violation occurs, meaning that $I_t = 0$ for today and $I_{t+1} = 1$ for tomorrow. Moreover, π_{11} is the probability of a violation tomorrow given today being a violation ($I_t = I_{t+1} = 1$). Conversely, the probability of a non-violation following a non-violation is denoted as $(1 - \pi_{01})$ and the probability of a non-violation following a violation as $(1 - \pi_{11})$. Let the likelihood function of the first-order Markov process be

$$L(\Pi_1) = (1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}},$$

where T_{ij} is the number of observations with a j following an i . Now, π_{01} and π_{11} can be estimated by taking the first derivatives of $L(\Pi_1)$. It is easy to demonstrate that:

$$\hat{\pi}_{01} = \frac{T_{01}}{T_{00} + T_{01}} \quad \text{and} \quad \hat{\pi}_{11} = \frac{T_{11}}{T_{10} + T_{11}}.$$

Because the probability has to sum to one, we have $\hat{\pi}_{00} = 1 - \hat{\pi}_{01}$ and $\hat{\pi}_{10} = 1 - \hat{\pi}_{11}$. At this point, we can formulate the null hypothesis of the independence test: $\pi_{01} = \pi_{11} = \pi$, based on the observation that under independence a violation tomorrow should not depend on today value of the hit function. The null hypothesis is tested using a likelihood ratio that assumes the form:

$$LR_{ind} = -2 \log \left[L(\hat{\pi}) / L(\hat{\Pi}_1) \right] \sim \chi_1^2,$$

where $L(\hat{\pi})$ has already been defined for the unconditional coverage test.

In the empirical part, we will use both the unconditional and the independence test in a jointly test as proposed by Christoffersen. In fact, this test, denoted as conditional coverage (*cc*), verifies the empirical rate of failures and the independence of the violations jointly. The conditional coverage test is given by following likelihood ratio test:

$$LR_{cc} = -2 \log [L(p) / L(\hat{\Pi})] \sim \chi_2^2,$$

that equals to test the hypothesis $\pi_{01} = \pi_{11} = \alpha$. Moreover, we have that $LR_{cc} = LR_{uc} + LR_{ind}$.

It remains to underline that the VaR forecasts are highly sensitive to distribution of the (portfolio) returns, as evidenced by Giot and Laurent (2002). To take into account the distribution of the portfolio returns, possibly different from the Gaussianity, the VaR forecasts used in this work are based on two different assumptions: the assumption of Normal distribution and the assumption of skewed and leptokurtic distribution of the returns. If the

distribution of the returns is Normal, then f_α in (6) is replaced by the left α quantile of the standard Normal distribution. But, if the returns exhibit skewness and severe kurtosis, as it has well documented in literature (for details, see Cont (2001)), then the VaR forecasts based on the Normal distribution assumption could be misleading. In order to explicitly consider the skewness and kurtosis of the returns, two approaches giving different values to f_α are used: the modified VaR approach and the skewed Student's t distribution. The first is due to Favre and Galeano (2002) that proposed an alternative version of the VaR quantile, through the use of a Cornish Fisher expansion. This modified version quantifies the α quantile as follows:

$$z_{cf} = z_c + [(z_c^2 - 1)S]/6 + [(z_c^3 - 3z_c)K]/24 - [(2z_c^3 - 5z_c)S^2]/36, \quad (10)$$

where z_c is the α quantile of the Normal distribution, S and K are the skewness and the excess kurtosis of the daily returns, respectively. Note that if $S = K = 0$, as in the case of Normal distribution, then $z_{cf} = z_c$, and the modified VaR collapses to the standard VaR.

The other approach considers the standardized skewed Student's t distribution (in short, SKST). Following Bauwens and Laurent (2002), assuming that the daily return r_t are such that $r_t = \sigma_t z_t$, where z_t is an i.i.d. process with $E(z_t) = 0$ and $Var(z_t) = 1$ and σ_t^2 represent the conditional variance, the excess of kurtosis and the skewness can be accommodated by using the (standardized) skewed- t distribution for z_t , formally $z_t \sim SKST(0, 1, \xi, v)$. The log transformation of ξ measures the skewness: if $\log(\xi) > 0$ the distribution is skew to the right and vice versa. The parameter v represents the degrees of freedom. This formalization is a generalization of the Student's t distribution: if $\xi = 1$, then $SKST$ collapses to a standard Student's t distribution. Assuming the conditional variance σ_t^2 can be modeled as an univariate GARCH (for details on GARCH models, see below), such that $\sigma_t^2 = \omega + ar_t^2 + b\sigma_{t-1}^2$, in the empirical part of this work we derive the quantile at $\alpha\%$ of the $SKST$ distribution, after having obtained the unknown parameters ξ and v , by maximizing the following log likelihood:

$$l_t(\theta) = \log\left(\frac{2}{\xi + \frac{1}{\xi}}\right) + \log\Gamma\left(\frac{v+1}{2}\right) - 0.5\pi(v-2) - \log\Gamma\left(\frac{v}{2}\right) + \log\frac{s}{\sigma_t} - 0.5(1+v)\log\left[1 + \frac{(sz_t + m)^2 \xi^{-2J_t}}{v-2}\right], \quad (11)$$

where $\theta = (\omega, a, b, \xi, v)$, $z_t = r_t/\sigma_t$,

$$m = \frac{\Gamma\left(\frac{v-1}{2}\right)(v-2)^{0.5}}{\pi^{0.5}\Gamma\left(\frac{v}{2}\right)}\left(\xi - \frac{1}{\xi}\right),$$

$$s = \left[\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2\right]^{0.5},$$

and

$$J_t = \begin{cases} 1 & \text{if } z_t \geq -\frac{m}{s} \\ -1 & \text{if } z_t < -\frac{m}{s} \end{cases}.$$

3. The models for the volatility

In this section we introduce the theoretical framework for the volatility proxy and the competing models for the forecasting of volatility.

3.1 Volatility proxy

Let $\{P_t^*, t \geq 0\}$ the continuous latent price process of a generic liquid asset be determined by the stochastic differential equation

$$d \log(P_t) = \sigma_t dW_t, \quad (12)$$

where W_t is a standard Brownian motion and σ_t is the spot volatility, that is continuous and predictable. Moreover, for simplicity we avoid to insert a drift term, we assume that σ_t and W_t are uncorrelated and the time unit interval is the day. For a survey about the stochastic differential equation see Protter (1992). We are interested in the estimate of the *integrated* volatility using the information up to day t . Thus, the one-period ahead integrated volatility is

$$IV_{t+1} = \int_t^{t+1} \sigma_\tau^2 d\tau. \quad (13)$$

Unfortunately, IV_{t+1} is not directly observable but the realized volatility, denoted as RV_{t+1} , represents its consistent estimate, as showed by Andersen and Bollerslev (1998), among others. In other words, RV_{t+1} is the volatility proxy, in the univariate context. The realized volatility is computed summing the squared intraday returns at a prefix frequency. Formally:

$$RV_{t+1} = \sum_{d=1}^D (r_{t,d})^2, \quad (14)$$

where D stands for the intraday period and $(r_{t,d}^D)$ for the observed intraday return, that is:

$$r_{t,d} = \log(P_{t,d}) - \log(P_{t,d-1}). \quad (15)$$

Hence, we need to sum squared intraday returns drawn at a very small intervals to correctly identify (13). Nevertheless, there are important implications in the choosing a small or a high frequency. From one hand, if the frequency D represents a day (so we have a very small frequency) and if the price at the end is the same of the price at the beginning of the trading day, then the corresponding realized volatility would be zero, even though the prices had experienced huge variations during the time interval. From the other hand, if the frequency is very high, some problems may occur. In fact, it has been documented that increasing to infinity the sampling frequency does not take to the real volatility but to a noise estimation of it, due to presence of micro-structure noise (Russell and Bandi (2004)).

The extension to the multivariate is the following. Let \mathbf{P}_t be the $k \times 1$ observed price vector for k assets, at time t . The equation (12) is replaced by:

$$d \log(\mathbf{P}_t) = \boldsymbol{\sigma}_t d\mathbf{W}_t, \quad (16)$$

where $\boldsymbol{\sigma}_t$ is the multivariate version of the spot volatility and \mathbf{W}_t is the k -dimensional standard Brownian motion. The estimation of $\boldsymbol{\sigma}_t$ is given by the $k \times k$ realized covariance $RCov_{t+1}$, whose principal diagonal represents the realized variance for each asset and whose i, j^{th} extra-diagonal element represents the realized covariance between the i^{th} and

j^{th} asset, with $i \neq j$. Formally, the realized covariance, based on observed prices and consequent returns, is:

$$RCov_{t+1} = \sum_{d=1}^D \mathbf{r}_{t,d} \mathbf{r}'_{t,d}, \quad (17)$$

where $\mathbf{r}_{t,d}$ denotes the $(k \times 1)$ vector of the differences between the log vector price at d^{th} intraday period and the log vector price at $(d - 1)^{th}$ intraday period for the day t . For ease of notation, hereafter the bold characterization for the multivariate processes will be replaced by the standard notation. Thus, $r_{t,d}$ indicates now the $(k \times 1)$ vector of log price differences.

3.2 Competing models for the volatility forecasting

We start with the models that belong to the multivariate GARCH family. Let r_t be the $k \times 1$ vector of daily log returns at time t . Moreover, let I^{t-1} be the information set at time $t - 1$. We assume that $E(r_t | I^{t-1}) = 0$. The volatility forecast for the model m is indicated with $H_{m,t}$, namely the conditional covariance matrix at time t . We use a standard parametrization, that has been extensively used in the multivariate GARCH literature, that is:

$$r_t = H_{m,t}^{0.5} z_t, \quad (18)$$

where z_t is the multivariate normal distributed innovation vector, such that $z_t \sim NID(0, I_k)$. We consider 9 specifications for $H_{m,t}$ that are frequently used in practice, that are: scalar, diagonal and full BEKK (the scalar BEKK model used in this work is the version with covariance targeting (Engle and Mezrich (1996)). For details on the BEKK model, see Engle and Kroner (1995)). Moreover, the DCC (Engle (2002)) and GOGARCH (van der Weide (2002)) models are considered. The univariate GARCH specifications for the conditional variance used in DCC and GOGARCH models are: GARCH (Bollerslev (1986)), GJR (Glosten et al. (1993)) and IGARCH (Engle and Bollerslev (1986)). These models have been chosen because most of them had been used in the work of Laurent et al. (2013). Table 2 provides the functional form for H_t for each model. The log-likelihood used to estimate all the parameters of the models and hence to have the estimates of H_t is:

$$l_T(\theta) = -\frac{1}{2} \sum_{t=1}^T \log |H_{m,t}| - \frac{1}{2} \sum_{t=1}^T (r_t)' H_{m,t}^{-1} (r_t), \quad (19)$$

where θ is the vector of parameters to estimate and T is the size of the estimation sample.

As highlighted before, among the set of competing models for the evaluation of the forecast accuracy we consider also some specifications that combine the realized covariance. These models are the Rolling Covariance (Fleming (2001)) and the Conditionally Autoregressive Wishart (CAW, Golosnoy et al. (2012)).

The Rolling Covariance derives H_t as

$$H_t = \exp(-\alpha) H_{t-1} + \alpha \exp(-\alpha) RCov_{t-1}, \quad (20)$$

where α is the parameter to estimate, also named decay parameter and $RCov_{t-1}$ is the realized covariance at time $t - 1$. The decay parameter is obtained by maximizing the log likelihood of (19).

The CAW specification assumes that the realized covariance $RCov_t$ follows a central Wishart distribution, given the past recorded in the information set I^{t-1} :

$$RCov_t | I^{t-1} \sim \mathcal{W}_k(\nu, S_t / \nu), \quad (21)$$

where $\nu > k - 1$ is the scalar degree of freedom and S_t/ν is the $k \times k$ symmetric positive definite matrix. It can be shown (see Anderson (1984)) that $E(RCov_t|I^{t-1}) = H_t$, the conditional covariance matrix. We opt for two specifications of the CAW model, based on the BEKK updating structure. The first is the scalar CAW with covariance targeting, that is:

$$H_t = (1 - a^2 - b^2) * \overline{RCov} + a^2 * RCov_t + b^2 H_{t-1}, \quad (22)$$

where a and b are scalars to estimate and \overline{RCov} is the sample average of the realized covariances, that is: $\overline{RCov} = \sum_{t=1}^T RCov_t$.

The second specification of the CAW model used in this work is the diagonal CAW with covariance targeting, that is:

$$H_t = \overline{RCov} - \overline{ARCov}A' - \overline{BRCov}B' + ARCov_tA' + BH_{t-1}B', \quad (23)$$

where A and B are two diagonal matrices to estimate. The estimation of (22) and (23) is done by maximizing the following log-likelihood, provided by Golosnoy et al. (2012):

$$l_T(\theta_w) = \sum_{t=2}^T \left\{ -\frac{\nu n}{2} \ln(2) - \frac{k(k-1)}{4} \ln(\pi) - \sum_{i=1}^k \ln \Gamma\left(\frac{\nu+1-i}{2}\right) - \frac{\nu}{2} \ln \left| \frac{RCov_t}{\nu} \right| + \frac{\nu-k-1}{2} \ln |RCov_t| - \frac{1}{2} \text{tr}(\nu H_t^{-1} RCov_t) \right\}, \quad (24)$$

where $\Gamma(\cdot)$ denotes the Gamma function, and $w = 1, 2$ is the suffix indicating the CAW models presented above. Hence, the unknown parameters are: $\theta_1 = (\nu, a, b)$ for the scalar CAW and $\theta_2 = (\nu, A_{11}, \dots, A_{kk}, B_{11}, \dots, B_{kk})$ for the diagonal CAW.

- insert Table 2 about here -

4. Monte Carlo experiment

In this section we illustrate the setting of the Monte Carlo experiment through which we investigate the behaviour of the competing models with reference to the forecast accuracy, from a statistical and economic point of view. For sake of simplicity, we consider a portfolio only composed by three assets. Let p_t denote the vector of the prices at time t . We assume that p_t is driven by the following stochastic differential equation:

$$dp(t) = \Theta(t)dW(t), \quad (25)$$

with $\Sigma(t) = \Theta(t)\Theta(t)'$, where

$$\Sigma(t) = \begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) & \Sigma_{13}(t) \\ \Sigma_{21}(t) & \Sigma_{22}(t) & \Sigma_{23}(t) \\ \Sigma_{31}(t) & \Sigma_{32}(t) & \Sigma_{33}(t) \end{pmatrix} = \begin{pmatrix} \sigma_1^2(t) & \sigma_{12}(t) & \sigma_{13}(t) \\ \sigma_{21}(t) & \sigma_2^2(t) & \sigma_{23}(t) \\ \sigma_{31}(t) & \sigma_{32}(t) & \sigma_3^2(t) \end{pmatrix},$$

and $\sigma_{ij}(t) = \sigma_i(t)\sigma_j(t)\rho_{ij}(t)$, with $i, j = 1, 2, 3$ and $i \neq j$. Moreover, $W(t)$ is a 3×1 vector of standard Brownian motions. The model for $\sigma_i^2(t)$ is the GARCH(1,1) diffusion studied in Andersen and Bollerslev (1998): $d\sigma_i^2(t) = (\omega - \theta * \sigma_i^2(t))dt + \lambda * \sigma_i^2(t)db_i(t)$, where $b_i(t)$ is a standard Brownian motion independent of $W(t)$. Following Donovan

et al. (2010), we set $\omega = 0.636$, $\theta = 0.035$ and $\lambda = 0.236$ for each i , that differs from each other for the initial point. Then, we set the instantaneous correlation $\rho_{ij} = (e^{2x(t)} - 1)/(e^{2x(t)} + 1)$, where x follows the GARCH diffusion: $dx(t) = (0.0192 - 0.03 * x(t))dt + 0.018 * x(t) * db_1(t)$, where b_1 is a standard Brownian motion. To make different the three instantaneous correlations, we divide each of them by a random number sampled from an Uniform distribution. The solution for all the stochastic differential equations has been obtained by an Euler discretization method, based on a equally spaced time increments. If these time increments are small, then we can approximate $t_i - t_{i-1}$ with dt . We have chosen a time increments of 0.0001. For details on the stochastic differential equation solutions and Euler discretization scheme, see Iacus (2008). In this framework, we simulate 500 times the trivariate vector p_t for two years. From the continuous time process (25) we drawn 288 observations per day, i.e. one observation each 5 minutes, for a total of 210,528 simulated observations (for each replicate). The forecasting sample, that is the period used to evaluate the forecast accuracy of the models, is of one year.

In this context, we have simulated three high-frequency prices following a data generating process with GARCH diffusion as variances and with changing instantaneous correlation. In our idea each price should approximate the behaviour of a stock. For this reason we sample only the observations between the 9.30 and 16.00, representing the standard trading day³. Then, we use these high-frequency data to obtain the volatility proxy ($RCov_t$) for different level of price aggregation and for the estimation of the competing models presented above. All the simulations have been carried out using R 2.15.3.

5. Results

In this section we present the answers to our research questions, that is: is the ranking of the models the same if we use a statistical or an economic loss functions? Do the MGARCH models have a worse forecast accuracy than that of the rolling Covariance and CAW models? The answers to the questions are given by using the simulated data presented above, needed to have 500 replicates of the conditional covariance matrix, denoted as $H_{m,t}$ for the $m - th$ model (summarized in Table 2, with $m = 1, \dots, 14$) and the 500 replicates of the volatility proxy, that is the realized covariance, for $t = 1, \dots, 730$. The forecasting sample is of one year, in a pure in-sample perspective: all conditional covariance matrices are computed for the whole period and the last year has been used for the statistical and economic loss evaluation.

Let us start with the rankings produced by the statistical loss functions. The evaluation is performed with respect to the deterioration of the volatility proxy, obtained sampling the intraday returns at lower frequencies. It is well known that, in absence of micro-structure frictions, the higher the frequency is, the better the proxy is (Russell and Bandi (2004)). For instance, using the data at 30 minutes will produce a realized covariance less noisy than using the data at 300 minutes. Having simulated the data by ourselves, we do not take in consideration the problem of the micro-structure frictions.

The results of the symmetric loss functions illustrated in Section 2 are summarized in Table 3. Here the loss functions are the Euclidean distance and the squared weighted Euclidean distance. The former considers the distance between each element of the volatility proxy matrix and $H_{m,t}$ while the latter considers only the diagonal elements, excluding the covariance entries. We do not find any significant difference between the two loss functions: the ranking of the models is almost the same. This is coherent with the literature that states the larger importance of variances with respect to the covariances. For the Euclidean

³The trading day is the time span that a particular Stock Exchange is open.

distance loss function, it results that when the volatility proxy is computed with the 5 minutes sampling frequency, the scalar CAW ranks first about 54% of the times and the Rolling Covariance at 5 minutes⁴ 41%. When the quality of the proxy deteriorates, first the Rolling Covariance at 5 minutes and then at 30 minutes emerge.

The same ranking (with different frequencies) can be observed for the squared weighted Euclidean distance. For instance, the scalar CAW ranks first about 46% of the times, a sensible lower percentage with respect to the previous loss function.

- insert Table 3 about here -

The frequencies at which each model is ranked first when the asymmetric loss function are used are presented in Table 4. In the upper part of the table the results for the penalizing over predictions Mahalanobis distance are reported, while in lower part there are those for the penalizing under predictions loss function. For the former loss function, the Rolling Covariance at 30 minutes ranks first when the sampling frequency drops from 10 minutes to 390 minutes (i.e. the daily frequency). Only when the volatility proxy is computed with 5 minutes frequency, the Rolling Covariance at 15 minutes ranks first about 50% of the times. We can state that the Rolling Covariance model seldom produces forecasts larger than the volatility proxy. This is an important result in a portfolio management optical. Not surprisingly, when the under prediction version of the Mahalanobis distance is used, the ranking of the model has different patterns. For a good quality of the realized covariance, the best model is the scalar CAW: when the realized covariance is obtained by using 5 minutes frequency, this model ranks first about 32% of the cases. Then, a multivariate GARCH model, for the first time, emerges: the IGOGARCH, that is the closest to the volatility proxy 25% of the times, for the realized covariance at 20 minutes. Finally, as seen also for the other loss functions, the Rolling Covariance at 30 minutes, when the quality of the proxy is low, always ranks first.

To sum up, when the statistical loss functions are used, the results award the models that directly use the realized covariance, as expected, excluding the too-parametrized diagonal CAW, which rarely ranks first. We can state that the statistical loss functions reward the model with less parameters to estimate, given that the Rolling Covariance and the scalar CAW have only one unknown parameter. Moreover, the frequently used multivariate GARCH models never rank first, except for the IGOGARCH model when used in some circumstances. After having reported the results of the statistical loss function, we can move to the economic part of the analysis in order to check if the same ranking is obtained.

- insert Table 4 about here -

The Christoffersen test is the method used to indirectly rank the set of forecasting models. The sample period used for such evaluation is of one year, as done for the statistical evaluation. In the spirit of Bauwens and Laurent (2002), we assign 3 different vectors of constant weights to the daily returns and daily conditional covariance matrix $H_{m,t}$, in order to obtain the portfolio mean and variance:

$$w_1 = (1/3, 1/3, 1/3), w_2 = (0.5, 0.2, 0.3) \text{ and } w_3 = (1.4, -0.2, -0.2).$$

⁴The frequency 5 minutes means here that the Rolling Covariance uses the 5 minutes sampling frequency to compute the conditional covariance matrix.

The Christoffersen test for all the models, with the one day ahead VaR for $\alpha = 0.05$ and the quantile for the long position obtained by means of the normal distribution, the Cornish Fisher expansion and the skewed Student's t distribution is reported in Table 5.

- insert Table 5 about here -

Looking at the table, the smaller the frequency is, the better that model is, because each frequency indicates how many times the test has been rejected over the 500 replicates. First of all, we note the differences between the frequencies when the VaR is computed with the normal distribution or the Cornish Fisher expansion and the frequencies when the VaR is computed with the skewed Student's t distribution. These latter are smaller even though we do not work with real financial data that suffer from the kurtosis excess and skewness. Secondly, the rejections of the Christoffersen test are not only few but are also similar in the number among the models: when using the scalar CAW, the vector weights w_1 and the normal distribution, these rejections are about 6% (smallest value) against about 10% of the worse model, that is the DCC. Using the weights w_1 , the best model is the GJR-GOGARCH, for all the specifications of the quantile. Instead, using the weights w_3 , the best models are the diagonal CAW and Rolling Covariance. We can state that there is not too much correspondence between the ranking of the statistical loss functions and that of the economic loss function. In fact, the GJR-GOGARCH model and the diagonal CAW never rank first when the statistical loss functions are used. Instead now these models yield to best economic performances. The issue is worth further consideration. Our idea is to look at the mean of the VaR violations among the models, where for *mean* we refer to the average number of VaR violations for all the replicates. Intuitively, for each replicate, we should have a value close to 0.05. The aim of taking the mean among all the replications is to approximate the whole behaviour of the model, independently of the each single replicate. The mean of the VaR violations for all the replicates is reported in Table 6.

- insert Table 6 about here -

The highest mean of VaR violations is of the GOGARCH model, for all the univariate specifications. Excluding the GOGARCH model, all the means are closer to the expected values 0.05. Combining this information with those furnished in Table 5, we can state that when the economic loss function is used alone there is no model that clearly emerges as the "best" model. This is because neither changing the weights nor the distributional function for the returns, we find the same ranking of the statistical function. In this sense, the economic function used here yields results that diverge from those of the statistical loss function. An issue that opens new questions.

6. Conclusion

Thanks to the work of Hansen and Lunde (2006) before, and that of Patton (2006) afterwards for the univariate framework and that of Laurent et al. (2013) for the multivariate, a new research field is developed: the consistent ranking of volatility models. If in Economics a prediction of a quantity, like for instance the GDP, can be evaluated ex-post once the quantity is observed, for the measure of the risk this is not possible. In fact, the risk is a latent variable that cannot be observed and only a proxy can be used. Normally, the

co-variation of the assets is used to measure the risk. Generally speaking, we refer to the co-variation with the term volatility. The cited works demonstrate that even if a proxy is used, the ranking of the forecasting models is the same of the ranking that would be obtained if the true volatility had been observable, under the constraint of the consistency of the loss function and of the conditional unbiasedness of the volatility proxy. A common choice of the volatility proxy that assures the conditional unbiasedness is the realized covariance. After these works, some questions arise. First question is about the differences between the ranking of a set of competing models using a statistical and economic loss function. Second related question relies on the forecast accuracy of the multivariate GARCH models compared to that of the models that directly use the realized covariance. This paper aims to give empirical answer to these questions by using a Monte Carlo experiment replicating 500 times a trivariate continuous-time stochastic process, from which we sampled one observation each 5 minutes for 2 years. Hence, we rank fourteen models using a forecasting sample of one year from a statistical and economic point of view. The forecasts are only made from an in-sample perspective. The statistical loss functions taken in consideration in this work are: (i) the matrix version of the Mean Squared Error function, named Euclidean distance; (ii) the squared weighted Euclidean distance, that considers only the variances for the computing of the distances; (iii) the Mahalanobis distance in the version that penalizes the over predictions; (iv) the Mahalanobis distance in the version that penalizes the under predictions. The economic loss function uses first the Value at risk methodology and then focuses on the violation of the one day ahead VaR computed with a 95% confidence level. In particular we look at the results of the Christoffersen test, that jointly tests if the number of violations are coherent with the expected number of violations and if the violations are not clustered in time (i.e. independence hypothesis). Given the VaR is sensible to the underlying assumption on the distribution of returns, we use three methods to calculate the VaR: the standard method that considers the normal distribution of the returns, the Corner Fisher expansion and the skewed Student's t distribution.

Our main findings are: first, we do not find a clear correspondence between the rankings resulting from the statistical and economic loss functions. A portfolio manager that had only used economic criteria for his decision would have preferred models not exactly closer to the volatility proxy. The statistical based ranking, when the volatility proxy is good, in the sense that it is based on high frequency, awards the scalar CAW and the rolling Covariance models, for the symmetric and asymmetric loss functions, respectively. When the quality of the proxy deteriorates, meaning that the realized covariance is computed using low frequencies, up to the use of daily returns, the statistical approach always ranks first the rolling Covariance model. Instead the economic loss function methodology used here is sensible to the choice of the distribution and of the weights such that there is any model that clearly emerges. Second, looking at the statistical loss function, we note that the MGARCH models yield worse forecast accuracy than that of the realized covariance based models, if these models have few parameters to estimate. Moreover, the impact of the covariances for the rankings seems to be irrelevant, given the ranking based only on variances is the same of the ranking based on both variances and covariances. Finally, if a portfolio manager is interested in studying the models such that the over predictions are rare, the Rolling Covariance at 30 minutes should be considered.

For future research, many questions remain open. First of all, it might be interesting to investigate the results to our questions by using real data, considering different forecasting methodologies, like for instance a rolling window to compute pure out-of-sample forecasts. Second, some other economic loss functions could be included into the analysis. A natural choice could be using the Dynamic Quantile (DQ) test, proposed by Engle and Manganelli (2004), that verifies if the probability of getting a VaR violation at time

$t + 1$ is independent of any variable observed at time t . An observed variable could be the contemporaneous or lagged VaR estimate of the model that uses the realized covariance at different sampling frequencies, instead of the conditional covariance matrix resulting from a multivariate GARCH model, in order to have a direct link between the statistical loss function and the economic loss function. Moreover, what could be further explored are the reasons of the different ranking between statistical and loss functions.

Appendix

Table 1: Specifications of loss functions used in the work

	Name	form of Λ	symmetry
L_E	Euclidean distance	$\Lambda = I_k$	symmetric
L_{SE}	Squared weighted Euclidean distance	$\lambda_{i,i} > 0$ and $\lambda_{i,j} = 0$	symmetric
L_{M-O}	Mahalanobis distance over prediction vers.	$\lambda_{i,i} = \lambda_{i,j} = 1 + I_{\{ov\}}$, where $I_{\{ov\}} = \begin{cases} 0 & \text{if } \hat{\sigma}_t - h_t \geq 0 \\ 1 & \text{if } \hat{\sigma}_t - h_t < 0 \end{cases}$	asymmetric
L_{M-U}	Mahalanobis distance under prediction vers.	$\lambda_{i,i} = \lambda_{i,j} = 1 + I_{\{ov\}}$, where $I_{\{ov\}} = \begin{cases} 0 & \text{if } \hat{\sigma}_t - h_t \leq 0 \\ 1 & \text{if } \hat{\sigma}_t - h_t > 0 \end{cases}$	asymmetric

Notes: λ_{ii} indicates the element of Λ referring to the variance element ii of the forecast error matrix; λ_{ij} to the covariance element ij , with $i, j = 1, \dots, k$. L_E equally weights the variance and covariance elements. L_{SE} weights only the variance elements. L_{M-O} penalizes the over predictions, such that if there is an over prediction at time t , the loss function counts twice that forecast error. L_{M-U} penalizes the under predictions, such that if there is an under prediction at time t , the loss function counts twice that forecast error.

Table 2: Forecasting models functional forms

Model	Multivariate GARCH models	Parameters
sBEKK(1,1)	$H_t = (1 - A - B) * \bar{H} + Ar_{t-1}r'_{t-1}A + BH_{t-1}B$	2
dBEKK(1,1)	$H_t = CC' + Ar_{t-1}r'_{t-1}A + BH_{t-1}B$	$k(k-1) + 2k$
BEKK(1,1)	$H_t = CC' + Ar_{t-1}r'_{t-1}A' + BH_{t-1}B'$	$5(5k+1)/2$
DCC(1,1)	$H_t = D_t R_t D_t$ $R_t = \text{diag}(q_{11,t}^{-1/2}, \dots, q_{kk,t}^{-1/2}) Q_t \text{diag}(q_{11,t}^{-1/2}, \dots, q_{kk,t}^{-1/2})$ $D_t = \text{diag}(h_{11,t}^{1/2}, \dots, h_{kk,t}^{1/2})$ $u_t = D_t^{-1} \epsilon_t$ $Q_t = (1 - \alpha - \beta) \bar{Q} + \alpha u_t u_t' + \beta Q_{t-1}$	$2 + k + \text{univ}$
GOG(1,1)	$V^{-1/2} \epsilon_t = L f_t$ $H_t = V^{1/2} L Z_t L V^{1/2}$ $Z_t = \text{diag}(\sigma_{f_{1t}}^2, \dots, \sigma_{f_{kt}}^2)$ $L = P \Lambda^{1/2} U, U = \prod_{i < j} R_{i,j}(\delta_{i,j}), -\pi \leq \delta_{i,j} \leq \pi$	$k(k-1)/2 + \text{univ}$
Univariate GARCH models in D_t and Z_t ($l = 1, \dots, k$)		
GARCH(1,1)	$h_{l,t} = \omega_l + \alpha_l \epsilon_{l,t-1}^2 + \beta_l h_{l,t-1}$	$3k$
GJR(1,1)	$h_{l,t} = \omega_l + \alpha_l \epsilon_{l,t-1}^2 + \gamma_l S_{l,t-1}^- \epsilon_{l,t-1}^2 + \beta_l h_{l,t-1}$ $S_{l,t}^- = 1 \text{ if } \epsilon_{l,t} < 0; S_{l,t}^- = 0 \text{ if } \epsilon_{l,t} \geq 0$	$4k$
IGARCH(1,1)	$h_{l,t} = \omega_l + \alpha_l \epsilon_{l,t-1}^2 + \beta_l h_{l,t-1}$ $\alpha_l + \beta_l = 1, \forall l$	$3k$
Model	Realized Covariance based models	Parameters
Rolling Cov.	$H_t = \exp(-\alpha) H_{t-1} + \alpha \exp(-\alpha) RCov_{t-1}$	1
sCAW	$H_t = (1 - a^2 - b^2) * \overline{RCov} + a^2 RCov_t + b^2 H_{t-1}$	2
dCAW	$H_t = \overline{RCov} - \overline{ARCOv} A' - \overline{BRCov} B' +$ $\overline{ARCOv} A' + B H_{t-1} B'$	$2k$

Notes: sBEKK: scalar BEKK; dBEKK: diagonal BEKK; GOG: GOGARCH; Rolling Cov.: Rolling Covariance; sCAW: scalar CAW; dCAW: diagonal CAW.

Table 3: Frequencies at which each model exhibits the smallest loss. *Symmetric loss functions*

	sBEKK	dBEKK	BEKK	DCC	gjrDCC	iDCC	GOG	gjrGOG	iGOG	RC_{5m}	RC_{15m}	RC_{30m}	sCAW	dCAW
Euclidean distance														
5 min	0.000	0.000	0.000	0.002	0.000	0.004	0.002	0.000	0.004	0.410	0.000	0.000	0.540	0.038
10 min	0.000	0.000	0.000	0.004	0.000	0.000	0.006	0.000	0.000	0.492	0.048	0.000	0.426	0.024
15 min	0.000	0.002	0.000	0.012	0.002	0.000	0.006	0.002	0.006	0.422	0.236	0.016	0.284	0.012
20 min	0.000	0.002	0.002	0.008	0.002	0.002	0.008	0.002	0.010	0.436	0.066	0.000	0.422	0.040
30 min	0.002	0.000	0.000	0.036	0.012	0.000	0.010	0.002	0.000	0.114	0.352	0.378	0.096	0.004
40 min	0.006	0.006	0.002	0.038	0.018	0.002	0.006	0.008	0.000	0.118	0.326	0.356	0.112	0.002
50 min	0.012	0.002	0.002	0.044	0.018	0.000	0.002	0.006	0.000	0.048	0.160	0.638	0.078	0.000
60 min	0.004	0.006	0.002	0.050	0.024	0.000	0.014	0.026	0.000	0.122	0.300	0.340	0.114	0.004
120 min	0.010	0.016	0.020	0.054	0.052	0.004	0.012	0.060	0.002	0.110	0.226	0.306	0.118	0.014
180 min	0.084	0.018	0.000	0.034	0.052	0.000	0.018	0.038	0.020	0.016	0.068	0.606	0.052	0.002
240 min	0.086	0.018	0.000	0.038	0.056	0.000	0.018	0.040	0.020	0.014	0.064	0.600	0.048	0.002
300 min	0.092	0.018	0.000	0.042	0.056	0.000	0.022	0.042	0.024	0.014	0.066	0.578	0.048	0.002
360 min	0.086	0.018	0.000	0.038	0.056	0.000	0.018	0.040	0.020	0.014	0.064	0.600	0.048	0.002
Daily	0.076	0.014	0.000	0.036	0.038	0.000	0.070	0.088	0.062	0.014	0.058	0.502	0.046	0.000
Squared weighted Euclidean distance														
5 min	0.000	0.000	0.000	0.016	0.002	0.002	0.010	0.000	0.012	0.448	0.008	0.004	0.460	0.046
10 min	0.000	0.004	0.002	0.016	0.000	0.000	0.012	0.000	0.006	0.552	0.054	0.004	0.322	0.032
15 min	0.000	0.006	0.000	0.014	0.006	0.004	0.012	0.002	0.014	0.444	0.242	0.028	0.218	0.014
20 min	0.000	0.006	0.002	0.016	0.006	0.008	0.016	0.000	0.018	0.468	0.070	0.008	0.346	0.040
30 min	0.010	0.002	0.000	0.056	0.012	0.002	0.012	0.002	0.000	0.088	0.312	0.430	0.078	0.002
40 min	0.014	0.004	0.000	0.044	0.016	0.004	0.014	0.010	0.000	0.102	0.310	0.394	0.092	0.002
50 min	0.024	0.006	0.002	0.046	0.026	0.002	0.004	0.008	0.000	0.030	0.134	0.654	0.068	0.000
60 min	0.008	0.008	0.008	0.048	0.028	0.002	0.008	0.018	0.004	0.106	0.308	0.376	0.084	0.000
120 min	0.016	0.022	0.020	0.050	0.048	0.006	0.006	0.050	0.006	0.106	0.236	0.330	0.100	0.012
180 min	0.070	0.012	0.000	0.034	0.038	0.002	0.028	0.072	0.032	0.016	0.060	0.596	0.042	0.002
240 min	0.066	0.014	0.002	0.034	0.052	0.002	0.034	0.060	0.030	0.016	0.066	0.586	0.040	0.002
300 min	0.078	0.014	0.000	0.034	0.044	0.002	0.034	0.076	0.034	0.016	0.062	0.570	0.038	0.002
360 min	0.066	0.014	0.002	0.034	0.052	0.002	0.034	0.060	0.030	0.016	0.066	0.586	0.040	0.002
Daily	0.066	0.012	0.002	0.036	0.036	0.002	0.070	0.100	0.076	0.016	0.050	0.506	0.032	0.000

Notes: sBEKK: scalar BEKK; dBEKK: diagonal BEKK; DCC, gjrDCC and iDCC: DCC with GARCH, GJR-GARCH and IGARCH univariate variances; GOG, gjrGOG and iGOG: GOGARCH with GARCH, GJR-GARCH and IGARCH univariate variances; RC_{tm} : Rolling Covariance with realized covariance at t minutes; sCAW: scalar CAW; dCAW: diagonal CAW.

Table 4: Frequencies at which each model exhibits the smallest loss. *Asymmetric loss functions*

	sBEKK	dBEKK	BEKK	DCC	gjrDCC	iDCC	GOG	gjrGOG	iGOG	RC_{5m}	RC_{15m}	RC_{30m}	sCAW	dCAW
Mahalanobis distance, over prediction version														
5 min	0.002	0.002	0.002	0.016	0.002	0.002	0.018	0.000	0.000	0.160	0.498	0.186	0.112	0.004
10 min	0.006	0.006	0.000	0.036	0.008	0.000	0.010	0.000	0.002	0.044	0.138	0.668	0.082	0.004
15 min	0.022	0.006	0.000	0.044	0.012	0.000	0.004	0.006	0.000	0.034	0.094	0.708	0.074	0.002
20 min	0.026	0.010	0.000	0.044	0.010	0.000	0.010	0.002	0.000	0.030	0.104	0.694	0.074	0.002
30 min	0.056	0.018	0.000	0.054	0.026	0.000	0.000	0.006	0.000	0.022	0.086	0.678	0.060	0.002
40 min	0.062	0.020	0.000	0.052	0.030	0.000	0.006	0.008	0.000	0.018	0.082	0.664	0.064	0.002
50 min	0.078	0.018	0.000	0.048	0.036	0.000	0.006	0.008	0.002	0.018	0.082	0.652	0.060	0.000
60 min	0.074	0.018	0.000	0.052	0.040	0.000	0.006	0.008	0.000	0.018	0.080	0.658	0.056	0.002
120 min	0.080	0.020	0.000	0.046	0.060	0.000	0.008	0.008	0.006	0.018	0.074	0.638	0.054	0.000
180 min	0.066	0.016	0.000	0.036	0.036	0.000	0.076	0.094	0.078	0.014	0.060	0.482	0.046	0.000
240 min	0.070	0.016	0.000	0.034	0.038	0.000	0.080	0.094	0.078	0.014	0.060	0.478	0.042	0.000
300 min	0.072	0.016	0.000	0.038	0.036	0.000	0.068	0.102	0.080	0.014	0.060	0.474	0.044	0.000
360 min	0.070	0.016	0.000	0.034	0.038	0.000	0.080	0.094	0.078	0.014	0.060	0.478	0.042	0.000
Daily	0.074	0.014	0.000	0.038	0.034	0.000	0.064	0.108	0.080	0.014	0.062	0.474	0.042	0.000
Mahalanobis distance, under prediction version														
5 min	0.012	0.038	0.008	0.064	0.010	0.058	0.066	0.004	0.162	0.102	0.004	0.012	0.326	0.142
10 min	0.012	0.052	0.014	0.052	0.014	0.074	0.060	0.010	0.204	0.076	0.006	0.008	0.270	0.156
15 min	0.018	0.072	0.014	0.058	0.022	0.072	0.072	0.004	0.188	0.088	0.006	0.010	0.246	0.138
20 min	0.038	0.090	0.026	0.038	0.022	0.096	0.044	0.016	0.250	0.062	0.008	0.008	0.164	0.146
30 min	0.014	0.082	0.016	0.058	0.022	0.056	0.066	0.018	0.170	0.114	0.006	0.008	0.264	0.114
40 min	0.028	0.104	0.038	0.040	0.032	0.072	0.038	0.024	0.220	0.072	0.008	0.010	0.208	0.118
50 min	0.010	0.072	0.032	0.036	0.042	0.052	0.062	0.024	0.166	0.130	0.006	0.014	0.262	0.096
60 min	0.050	0.132	0.068	0.016	0.026	0.068	0.040	0.046	0.220	0.048	0.004	0.012	0.156	0.122
120 min	0.090	0.144	0.154	0.008	0.026	0.062	0.012	0.086	0.198	0.020	0.002	0.002	0.082	0.118
180 min	0.026	0.014	0.000	0.048	0.042	0.000	0.002	0.018	0.000	0.020	0.110	0.656	0.060	0.004
240 min	0.022	0.026	0.068	0.050	0.108	0.002	0.008	0.052	0.000	0.048	0.120	0.434	0.058	0.008
300 min	0.032	0.024	0.034	0.040	0.094	0.000	0.002	0.048	0.000	0.038	0.106	0.534	0.042	0.010
360 min	0.022	0.026	0.068	0.050	0.108	0.002	0.008	0.052	0.000	0.048	0.120	0.434	0.058	0.008
Daily	0.064	0.018	0.002	0.044	0.048	0.000	0.010	0.024	0.006	0.016	0.076	0.642	0.054	0.000

Notes: sBEKK: scalar BEKK; dBEKK: diagonal BEKK; DCC, gjrDCC and iDCC: DCC with GARCH, GJR-GARCH and IGARCH univariate variances; GOG, gjrGOG and iGOG: GOGARCH with GARCH, GJR-GARCH and IGARCH univariate variances; RC_{tm} : Rolling Covariance with realized covariance at t minutes; sCAW: scalar CAW; dCAW: diagonal CAW.

Table 5: Frequencies at which the Christoffersen Test is rejected

Weights/ Model	(1/3, 1/3, 1/3)			(.5, .2, .3)			(.14, -.2, -.2)		
	VaR	Mod. VaR	Skew VaR	VaR	Mod. VaR	Skew VaR	VaR	Mod. VaR	Skew VaR
sBEKK	0.082	0.090	0.058	0.082	0.092	0.048	0.072	0.090	0.050
dBEKK	0.096	0.086	0.048	0.094	0.088	0.048	0.094	0.104	0.058
BEKK	0.084	0.096	0.056	0.070	0.100	0.052	0.104	0.098	0.066
DCC	0.108	0.102	0.066	0.088	0.100	0.064	0.072	0.086	0.070
gjrDCC	0.100	0.112	0.066	0.090	0.100	0.046	0.070	0.082	0.068
iDCC	0.090	0.096	0.058	0.088	0.096	0.060	0.078	0.086	0.086
GOG	0.072	0.076	0.042	0.070	0.068	0.042	0.056	0.078	0.078
gjrGOG	0.074	0.084	0.042	0.062	0.062	0.032	0.070	0.092	0.056
iGOG	0.068	0.070	0.040	0.074	0.070	0.042	0.060	0.072	0.037
RC_{5m}	0.082	0.096	0.064	0.090	0.068	0.050	0.062	0.082	0.036
RC_{15m}	0.086	0.090	0.048	0.088	0.090	0.048	0.074	0.088	0.046
RC_{30m}	0.076	0.078	0.044	0.092	0.080	0.046	0.062	0.074	0.034
sCAW	0.084	0.092	0.062	0.082	0.076	0.046	0.072	0.084	0.044
dCAW	0.062	0.080	0.062	0.088	0.070	0.058	0.050	0.056	0.042

Notes: sBEKK: scalar BEKK; dBEKK: diagonal BEKK; DCC, gjrDCC and iDCC: DCC with GARCH, GJR-GARCH and IGARCH univariate variances; GOG, gjrGOG and iGOG: GOGARCH with GARCH, GJR-GARCH and IGARCH univariate variances; RC_{tm} : Rolling Covariance with realized covariance at t minutes; sCAW: scalar CAW; dCAW: diagonal CAW. Columns *VaR* indicate the frequency at which the null hypothesis of the Christoffersen test is rejected, considering the 500 replicates, when the Normal quantile at 5% is used, Columns *Mod. VaR* when the Cornish Fisher expansion is used and Column *Skew VaR* when the skewed Student's t quantile is used.

Table 6: VaR violations mean for $\alpha = 0.05$

Weights/ Model	(1/3, 1/3, 1/3)			(.5, .2, .3)			(.1, .4, -.2, -.2)		
	VaR	Mod. VaR	Skew VaR	VaR	Mod. VaR	Skew VaR	VaR	Mod. VaR	Skew VaR
sBEKK	0.050	0.051	0.047	0.051	0.051	0.047	0.050	0.050	0.046
dBEKK	0.049	0.049	0.046	0.049	0.049	0.046	0.049	0.049	0.046
BEKK	0.048	0.048	0.048	0.049	0.049	0.046	0.048	0.049	0.045
DCC	0.051	0.051	0.048	0.051	0.051	0.047	0.051	0.051	0.047
GJR_DCC	0.050	0.050	0.047	0.050	0.050	0.047	0.051	0.051	0.047
iDCC	0.049	0.049	0.046	0.049	0.049	0.046	0.049	0.049	0.045
GOG	0.079	0.079	0.076	0.077	0.077	0.074	0.072	0.072	0.068
gjrGOG	0.079	0.079	0.076	0.076	0.076	0.073	0.070	0.070	0.067
iGOG	0.079	0.079	0.076	0.075	0.076	0.073	0.068	0.068	0.064
RC_{5min}	0.052	0.052	0.048	0.052	0.052	0.048	0.052	0.052	0.048
RC_{15min}	0.053	0.053	0.050	0.053	0.053	0.050	0.053	0.053	0.049
RC_{30min}	0.055	0.055	0.052	0.055	0.055	0.052	0.055	0.055	0.051
sCAW	0.051	0.051	0.048	0.051	0.051	0.048	0.051	0.051	0.047
dCAW	0.048	0.048	0.048	0.048	0.048	0.045	0.043	0.044	0.040

Notes: sBEKK: scalar BEKK; dBEKK: diagonal BEKK; DCC, gjrDCC and iDCC: DCC with GARCH, GJR-GARCH and IGARCH univariate variances; GOG, gjrGOG and iGOG: GOGARCH with GARCH, GJR-GARCH and IGARCH univariate variances; RC_{tm} : Rolling Covariance with realized covariance at t minutes; sCAW: scalar CAW; dCAW: diagonal CAW. All the columns indicate the mean of all the replications of the VaR violations, for different portfolios and different quantile f_α specifications.

References

- Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39(4), 885–905.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 71(2), 579–625.
- Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis* (second ed.). New York, NY: Wiley.
- Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance* 9(3), 203–228.
- Bandi, F., J. Russell, and Y. Zhu (2008). Using high-frequency data in dynamic portfolio choice. *Econometric Reviews* 27(1-3), 163–198.
- Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* 72(3), 885–925.
- Bauwens, L. and S. Laurent (2002). A new class of multivariate skew densities, with application to garch models. CORE Discussion Papers 2002020, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).
- Bauwens, L., S. Laurent, and J. Roumbouts (2003). Multivariate garch models: a survey. CORE Discussion Papers 2003031, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review* 39(4), 841–62.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* 1(2), 223–236.
- de Pooter, M., M. Martens, and D. van Dijk (2008). Predicting the daily covariance matrix for s&p 100 stocks using intraday data - but which frequency to use? *Econometric Reviews* 27(1-3), 199–229.
- Dovonon, P., S. Goncalves, and N. Meddahi (2010). Bootstrapping realized multivariate volatility measures. MPRA Paper 40123, University Library of Munich, Germany.
- Duffie, D. and J. Pan (1997). An overview of value at risk. Technical report, Stanford University.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20(3), 339–50.
- Engle, R. and J. Mezrich (1996). Garch for groups. *Risk* 9, 36–40.
- Engle, R. F. and T. Bollerslev (1986). Modelling the persistence of conditional variances. *Econometric Reviews* 5(1), 1–50.

- Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized arch. *Econometric Theory* 11(01), 122–150.
- Engle, R. F. and S. Manganelli (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics* 22, 367–381.
- Favre, L. and J. Galeano (2002). Mean-modified value-at-risk optimization with hedge funds. *Journal of Alternative Investments* 5, 21–25.
- Fleming, J. (2001). The economic value of volatility timing. *Journal of Finance* 56(1), 329–352.
- Fleming, J., C. Kirby, and B. Ostdiek (2003). The economic value of volatility timing using “realized” volatility. *Journal of Financial Economics* 67(3), 473–509.
- Giot, P. and S. Laurent (2002). Modelling daily value-at-risk using realized volatility and arch type models. *Computing in Economics and Finance 2002* 52, Society for Computational Economics.
- Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Staff Report 157, Federal Reserve Bank of Minneapolis.
- Golosnoy, V., B. Gribisch, and R. Liesenfeld (2012). The conditional autoregressive wishart model for multivariate stock market volatility. *Journal of Econometrics* 167(1), 211–223.
- Hansen, P. R. and A. Lunde (2006). Consistent ranking of volatility models. *Journal of Econometrics* 131(1-2), 97–121.
- Iacus, S. M. (2008). *Simulation and Inference for Stochastic Differential Equations: With R Examples*. New York: Springer. ISBN 978-0-387-75838-1.
- Jorion, P. (2007). *Value at risk* (3. ed. ed.). New York, NY: McGraw-Hill.
- Kaplanski, G. and Y. Kroll (2002). Var risk measures versus traditional risk measures: An analysis and survey. *Journal of Risk* 4(3), 1–27.
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. Finance and Economics Discussion Series 95-24, Board of Governors of the Federal Reserve System (U.S.).
- Laurent, S., J. V. Rombouts, and F. Violante (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics* 173(1), 1–10.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business* 36, 394.
- McAleer, M. and M. Medeiros (2008). Realized volatility: A review. *Econometric Reviews* 27(1-3), 10–45.
- Nelson, D. B. (1990). Arch models as diffusion approximations. *Journal of Econometrics* 45(1-2), 7–38.
- Patton, A. (2006). Volatility forecast comparison using imperfect volatility proxies. Research Paper Series 175, Quantitative Finance Research Centre, University of Technology, Sydney.

- Protter, P. (1992). *Stochastic Integration and Differential Equation* (Second ed.). Berlin, Heidelberg: Springer-Verlag.
- Russell, J. R. and F. M. Bandi (2004). Microstructure noise, realized volatility, and optimal sampling. *Econometric Society 2004 Latin American Meetings 220*, Econometric Society.
- van der Weide, R. (2002). Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics* 17(5), 549–564.