Leccadito *et al. Financial Innovation* (2024) 10:116
https://doi.org/10.1186/s40854-024-00642-2

Financial Innovation

RESEARCH

Open Access

# A novel robust method for estimating the covariance matrix of financial returns with applications to risk management

Arturo Leccadito[1,2*], Alessandro Staino[1] and Pietro Toscano[3]

*Correspondence:
arturo.leccadito@unical.it

[1] Department of Economics, Statistics and Finance, University of Calabria, 87036 Ponte Bucci, Rende, CS, Italy
[2] LFIN/LIDAM, UCLouvain, Voie du Roman Pays 34, 1348 Louvain la neuve, Belgium
[3] Fidelity Investments, 245 Summer Street, Boston, MA 02210, USA

**Abstract**

This study introduces the dynamic Gerber model (DGC) and evaluates its performance in the prediction of Value at Risk (VaR) and Expected Shortfall (ES) compared to alternative parametric, non-parametric and semi-parametric methods for estimating the covariance matrix of returns. Based on ES backtests, the DGC method produces, overall, accurate ES forecasts. Furthermore, we use the Model Confidence Set procedure to identify the superior set of models (SSM). For all the portfolios and VaR/ES confidence levels we consider, the DGC is found to belong to the SSM.

**Keywords:** Value at risk, Expected shortfall, Gerber statistic, Model confidence set, Superior set of models

**JEL Classifcation:** C51, C52, C58, G15

## Introduction

In a globalized economy, turbulence in financial markets has become more frequent. Recent events like the COVID-19 pandemic, the surge in inflation, and the Russian-Ukrainian conflict have all triggered considerable losses for investors worldwide. Providing methods that accurately measure financial market risk is therefore an increasingly crucial task. Over the past two decades, Value at Risk (VaR)—a measure related to the quantile of the conditional portfolio return distribution—has become the standard measure of market risk. While VaR is a measure of risk easy to understand for laymen, it has a number of shortcomings. First, it provides no information about the magnitude of the losses that exceed the quantile (see for instance Christoffersen 2012, Ch. 2). Second, VaR is not a coherent risk measure (see Artzner et al. 1999, for the properties of coherent risk measures). In particular, VaR is not subadditive, implying that the portfolio VaR could be larger than the sum of the VaRs of its components (see for instance Dhaene et al. 2006). VaR has been used in the context of credit scoring by Xu et al. (2024), who introduce the worst-case conditional value-at-risk metric to measure the loss incurred from employing a classification model in credit scoring under the deterioration of cost parameters. Unlike VaR, Expected Shortfall (ES)—defined as the conditional expectation of exceedances beyond the VaR—is a coherent risk measure (Acerbi and Tasche

2002). Several methods to estimate risk measures are available (see Nieto and Ruiz 2016, forasurveyofestimationmethodsconcerningVaR). A well-known non-parametric method is historical simulation (HS), which estimates VaR and ES using the empirical counterparts under the assumption of i.i.d. portfolio returns. Parametric methods typically involve estimating GARCH-type models; a few examples include Brooks and Persand (2003), Chu et al. (2017), and Long et al. (2020). Common alternatives are models based on extreme value theory, recent examples of which are represented by Bekiros et al. (2019) and Echaust and Just (2020). A semi-parametric approach to VaR/ES estimation combining a parametric model and HS is the filtered historical simulation (FHS) method of Barone-Adesi et al. (2002). The FHS method exploits the idea of using the empirical quantile of random draws obtained with replacement from the standardized residuals of a parametric model. VaR is then obtained by rescaling this quantile using the predicted volatility from the parametric model. When the focus is only on forecasting VaR, a popular approach is to directly model the conditional quantile using quantile regression, as in the conditional autoregressive VaR model of Engle and Manganelli (2004). Alternative semi-parametric methods that jointly estimate VaR and ES models include Patton et al. (2019) and Taylor (2019). Recent works are concerned with the VaR/ES risk measures in the context of financial (James et al. 2023), insurance (Fan et al. 2023), or commodity markets (Vancsura et al. 2023).

As in Lopez and Walter (2000) and Skintzi and Xanthopoulos-Sisinis (2007), who discuss the importance of covariance matrix forecasting for risk management, in this study we consider alternative methods to estimate the covariance matrix of returns. In particular, beside HS, we compare—in terms of accuracy in forecasting VaR or jointly VaR and ES—a number of different methods (parametric, non-parametric, and semi-parametric) for obtaining the correlation and hence the covariance matrix. In addition to methods employing the standard Pearson correlation, we rely on methods based on a static or dynamic version of the robust correlation proposed by Gerber et al. (2022). The measure is an extension of Kendall's Tau robust measure of pairwise movements of two series of returns. In particular, it is built based on the proportions of co-movements in the series of interest (i.e., on how many times the series simultaneously pierce some pre-specified thresholds). We contribute to the literature by introducing a dynamic version of the Gerber correlation matrix that we call the dynamic Gerber model (DGC). Like dynamic conditional correlation (DCC) models, the DGC model relies in the first stage of the estimation on univariate GARCH models for the volatility of each asset return. In the second stage, based on the marginally standardized residuals, a dynamic Gerber matrix is established. The evolution of such a matrix depends on six parameters only, even when the number of assets in the portfolio, $k$, is large. Like DCC models, but contrary to BEKK models, our DGC model cannot be used to directly capture volatility spillovers between different markets (like for instance in Kondoz et al. 2019). However, our model can easily be estimated when $k$ is large, whereas BEKK models become problematic as the number of parameters grows with the square of $k$. Yet another way of obtaining a time-varying correlation matrix would be one based on Markov switching models[1]. Again, unless parsimonious specifications are considered that rarely are effective in practical situations, such an approach could suffer from a curse of dimensionality. An alternative study that considers dynamic dependence via threshold exceedances is the one of Gong and Huser (2022). To have time-varying copula models[2], the authors estimate a family of parameters for each time point and replace the likelihood function by a family of weighted local likelihoods to be maximized. However, in their approach, the choice of the bandwidth associated with the weight functions (kernel) is crucial. Small bandwidths lead to parameter estimates that are very variable, while large bandwidths lead to smooth estimates with low variability over time. Our approach, instead, does not require specifying a bandwidth. Furthermore, the approach presented in Gong and Huser (2022) is bivariate, whereas our methodology is valid for a generic number of assets.

A study concerning robust forecasting of conditional correlations is that of Boudt et al. (2013), where an extension of the DCC model to reduce the biases in the volatility and correlation dynamics caused by large one-off events is presented. The authors apply their model to forecast the covariance matrix of the daily EUR/USD and Yen/USD return series, and they find that their model guarantees more precise out-of-sample covariance forecasts than the DCC model. In the same line of research, Jarjour and Chan (2020) introduce the concept of angular correlation for estimating the instantaneous correlation matrix and then generalize the DCC model to the dynamic conditional angular correlation (DCAC) model. They illustrate the better performance of the DCAC model compared to the DCC model in portfolio construction. A further study of robust forecasting of conditional correlations is conducted by Fiszeder et al. (2023), who suggest a new version of the DCC model based on daily opening, high, low, and closing prices. Using two different datasets, five exchange-traded funds and five currencies, they show that their model significantly outperforms the DCC model when forecasting conditional covariance matrices. Our paper can be placed in this context of empirical research, where more robust estimates of the conditional correlations than those obtained using the DCC model are proposed and tested.

In our empirical application, we consider three different portfolios consisting of developed equities (S &P 500 index), emerging equities (MXEF index), bonds (LBUSTRUU index), and gold. For different probability levels, we derive out-of-sample VaR and ES for the three portfolios. We use a recently proposed procedure to backtest ES and the Model Confidence Set (MCS) procedure to identify the superior set of models (SSM). For both the risk measures considered in the study, we find that models based on the DGC approach are part of the SSM across portfolios and confidence levels. A practical implication of our study is that the proposed DGC approach should be taken into consideration by risk managers, investors, and regulators when evaluating the risk of asset portfolios. The DGC model is indeed never eliminated from the SSM and, in the case of joint VaR and ES forecasting, is the only model not to be rejected by the MCS procedure. A model that offers a more robust estimate of conditional correlations represents

[1] Examples of application of Markov switching models are Athari et al. (2022) and Pelletier (2006).

[2] Copulas with time-varying parameters have been studied also by Patton (2006), Oh and Patton (2018), and Cortese (2019).

a valuable tool. Indeed, as some empirical studies show (see, e.g., Jiang et al. 2017 and Athari and Hung, 2022), adverse market conditions usually make assets more interdependent, and a model that accurately describes these dynamics, like the DGC model, could ensure more reliable risk estimates. Potentially, our method can also be employed in the context of bankruptcy prediction (Kou et al. 2021) in the case of loan portfolios as an alternative to Pearson correlation-based methods (Düllmann et al. 2007, 2010). Indeed, our method could be applied to the time series of Moody's KMV asset values as the correlation of unobservable asset returns is a key component for the measurement of portfolio risk. The remainder of the paper is organized as follows. Section "Methodology" presents the methods used in the estimations of VaR and ES, introduces the DGC model, and describes the MCS procedure. Section "Empirical analysis" presents the data used in this study and the results of the empirical analysis. Section "Conclusions" concludes.

## Methodology

In this section, we first present the static version of the robust measure of correlation introduced by Gerber et al. (2022). Next, we discuss DCC models and introduce the novel DGC method. Finally, we explain how the risk measures of interest are derived and how to evaluate their predictions via the MCS procedure.

### The Gerber statistic

We denote by $r_{i,t}$ and $r_{j,t}$ the returns of asset $i$ and asset $j$ at time $t$. The Gerber statistic, introduced by Gerber et al. (2022), is a robust measure of pairwise movements of the two series of returns defined as

$$g(i,j) = \frac{n_{ij}^c - n_{ij}^d}{n_{ij}^c + n_{ij}^d}, \tag{1}$$

where

$$\begin{aligned} n_{ij}^c &= \sum_{t=1}^{T} I\left(r_{i,t} \geq Q_i\right) I\left(r_{j,t} \geq Q_j\right) + \sum_{t=1}^{T} I\left(r_{i,t} \leq -Q_i\right) I\left(r_{j,t} \leq -Q_j\right) = n_{ij}^{UU} + n_{ij}^{DD} \\ n_{ij}^d &= \sum_{t=1}^{T} I\left(r_{i,t} \geq Q_i\right) I\left(r_{j,t} \leq -Q_j\right) + \sum_{t=1}^{T} I\left(r_{i,t} \leq -Q_i\right) I\left(r_{j,t} \geq Q_j\right) = n_{ij}^{UD} + n_{ij}^{DU}. \end{aligned}$$

Here, $T$ is the number of observations, $Q_i$ and $Q_j$ are thresholds, and $I(A)$ denotes the indicator function for the event $A$. Hence, $n_{ij}^c$ denotes the number of concordant pairs (i.e., the number of times both returns pierce their thresholds while moving in the same direction). Indeed, $n_{ij}^c$ is equal to the sum of $n_{ij}^{UU}$—the number of pairs for which both returns are larger than their threshold—and $n_{ij}^{DD}$, the number of pairs for which both returns are smaller than their threshold times minus one. Conversely, $n_{ij}^d = n_{ij}^{UD} + n_{ij}^{DU}$ represents the number of discordant pairs in the sample, that is, the number of times both returns pierce their thresholds while moving in the opposite direction.

When several pairs of returns are involved, using (1) to construct a correlation matrix may lead to covariance matrices that are not positive semidefinite. Therefore, Gerber et al. (2022) define $n_{ij}^{NN} = \sum_{t=1}^{T} I\left(|r_{i,t}| \leq Q_i\right) I\left(|r_{j,t}| \leq Q_j\right)$ and propose replacing (1) with

$$g(i,j) = \frac{n_{ij}^{c} - n_{ij}^{d}}{T - n_{ij}^{NN}}, \tag{2}$$

which instead yields positive semidefinite covariance matrices.

We now consider the case of $k$ different assets and denote by $\boldsymbol{U}$ and $\boldsymbol{D}$ the two $T \times k$ matrices with generic element $u_{t,j} = I\left(r_{j,t} \geq Q_j\right)$ and $d_{t,j} = I\left(r_{j,t} \leq -Q_j\right)$, respectively, $t = 1, \ldots, T$ and $j = 1, \ldots, k$. The $k \times k$ matrix $\boldsymbol{G}$ with element in position $(i, j)$ given by the Gerber correlation (1) is then given by

$$\boldsymbol{G} = \left(\boldsymbol{U}'\boldsymbol{U} + \boldsymbol{D}'\boldsymbol{D} - \boldsymbol{U}'\boldsymbol{D} - \boldsymbol{D}'\boldsymbol{U}\right) \oslash \left(\boldsymbol{U}'\boldsymbol{U} + \boldsymbol{D}'\boldsymbol{D} + \boldsymbol{U}'\boldsymbol{D} + \boldsymbol{D}'\boldsymbol{U}\right), \tag{3}$$

where $\oslash$ means elementwise division.

With the further definition of the $T \times k$ matrix $\boldsymbol{N}$ with generic element $n_{t,j} = I\left(|r_{j,t}| \leq Q_j\right)$, $t = 1, \ldots, T$ and $j = 1, \ldots, k$, we can express the $k \times k$ matrix $\boldsymbol{G}$ with generic element $g(i, j)$ of equation (2) as

$$\boldsymbol{G} = (\boldsymbol{\mathcal{C}} - \boldsymbol{\mathcal{D}}) \oslash (\mathbf{1} - \boldsymbol{\mathcal{N}}), \tag{4}$$

where $\boldsymbol{\mathcal{C}} = \left(\boldsymbol{U}'\boldsymbol{U} + \boldsymbol{D}'\boldsymbol{D}\right)/T$, $\boldsymbol{\mathcal{D}} = \left(\boldsymbol{U}'\boldsymbol{D} + \boldsymbol{D}'\boldsymbol{U}\right)/T$, $\boldsymbol{\mathcal{N}} = \boldsymbol{N}'\boldsymbol{N}/T$, and $\mathbf{1}$ denotes the $k \times k$ matrix with all entries 1.

### DCC models

DCC models assume that the multivariate time series of returns $\boldsymbol{r}_t = (r_{1,t}, \ldots, r_{k,t})'$ is described by

$$\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{z}_t, \tag{5}$$

where $\{\boldsymbol{z}_t\}_t$ is a sequence of independent and identically distributed random vectors such that $\mathbb{E}(\boldsymbol{z}_t) = 0$ and $\text{cov}(\boldsymbol{z}_t) = \boldsymbol{I}_\text{k}$, with $\boldsymbol{\Sigma}_t^{1/2}$ denoting the positive-definite square-root matrix of the conditional covariance matrix of the returns, $\boldsymbol{\Sigma}_t$.

In DCC models (Engle 2002), the $\boldsymbol{\Sigma}_t$ matrix, whose generic element is $\sigma_{ij,t}$, is decomposed as

$$\boldsymbol{\Sigma}_t = \text{diag}\,(\sigma_{11,t}^{1/2}, \ldots, \sigma_{kk,t}^{1/2}) \boldsymbol{R}_t \,\text{diag}\,(\sigma_{11,t}^{1/2}, \ldots, \sigma_{kk,t}^{1/2}), \tag{6}$$

where $\boldsymbol{R}_t$ is the positive definite conditional correlation matrix. In this way, if $\rho_{ij,t}$ is the element of position $(i, j)$ of the correlation matrix $\boldsymbol{R}_t$, then the corresponding element of $\boldsymbol{\Sigma}_t$ is found to be $\rho_{ij,t}\sqrt{\sigma_{ii,t}\sigma_{jj,t}}$.

Let $\hat{u}_{i,t} = r_{i,t} - \mu_{i,t}$ be the time $t$ residual from the mean equation of asset $i$. We denote by $\hat{\boldsymbol{\eta}}_t = (\hat{\eta}_{1,t}, \ldots, \hat{\eta}_{k,t})'$ the marginally standardized innovation vector:

$$\hat{\eta}_{i,t} = \frac{\hat{u}_{i,t}}{\sqrt{\sigma_{ii,t}}} \quad i = 1, \ldots, k.$$

In this way, $\boldsymbol{R}_t$ is the covariance matrix of $\boldsymbol{\eta}_t$. Engle (2002) proposes modelling the correlation matrix as

$$\begin{aligned} \boldsymbol{S}_t &= (1-a-b)\bar{\boldsymbol{S}} + a\hat{\boldsymbol{\eta}}_{t-1}\hat{\boldsymbol{\eta}}'_{t-1} + b\boldsymbol{S}_{t-1} \\ \boldsymbol{R}_t &= \boldsymbol{J}_t\boldsymbol{S}_t\boldsymbol{J}_t, \end{aligned} \tag{7}$$

where $\bar{\boldsymbol{S}}$ is the unconditional covariance matrix of $\hat{\boldsymbol{\eta}}_t$, and $\boldsymbol{J}_t = \text{diag}\,(s_{11,t}^{-1/2}, \ldots, s_{kk,t}^{-1/2})$, with $s_{ii,t}$ denoting the element of position $(i, i)$ of $\boldsymbol{S}_t$. In the first stage of the DCC model estimation, $k$ univariate GARCH models are independently estimated for each of the return series. In the second stage, the marginally standardized innovation vectors are derived and the parameters $a$ and $b$ in (7) are estimated.

### A dynamic Gerber model

We propose a dynamic Gerber correlation (DGC) model based on a two-stage estimation like DCC models. In the first stage, the volatility of each asset is independently modelled via a GARCH model. Hence, as in DCC models, in the first stage $k$ univariate GARCH models are independently estimated. In the second stage, based on the the marginally standardized innovation vectors, a dynamic is given to the Gerber correlation matrix (see Eq. (4)) rather than the Pearson correlation matrix. To be more precise, we identify the thresholds $Q_i^{\eta}$ for each of the $k$ time-series of marginally standardized innovation $\hat{\eta}_i$ $(i = 1, \ldots, k)$ and consider the following dynamics:

$$\begin{aligned} \boldsymbol{\mathcal{C}}_t &= (1-a_C-b_C)\overline{\boldsymbol{\mathcal{C}}} + a_C\hat{\boldsymbol{\mathcal{I}}}_{C,t-1} + b_C\boldsymbol{\mathcal{C}}_{t-1} \\ \boldsymbol{\mathcal{D}}_t &= (1-a_D-b_D)\overline{\boldsymbol{\mathcal{D}}} + a_D\hat{\boldsymbol{\mathcal{I}}}_{D,t-1} + b_D\boldsymbol{\mathcal{D}}_{t-1} \\ \boldsymbol{\mathcal{N}}_t &= (1-a_N-b_N)\overline{\boldsymbol{\mathcal{N}}} + a_N\hat{\boldsymbol{\mathcal{I}}}_{N,t-1} + b_N\boldsymbol{\mathcal{N}}_{t-1} \\ \boldsymbol{G}_t &= (\boldsymbol{\mathcal{C}}_t - \boldsymbol{\mathcal{D}}_t) \oslash (\boldsymbol{1} - \boldsymbol{\mathcal{N}}_t). \end{aligned} \tag{8}$$

The $\hat{\boldsymbol{\mathcal{I}}}$ matrices appearing in (8) are obtained as follows:

$$\begin{aligned} \hat{\boldsymbol{\mathcal{I}}}_{C,t} &= \hat{\boldsymbol{u}}_t\hat{\boldsymbol{u}}'_t + \hat{\boldsymbol{d}}_t\hat{\boldsymbol{d}}'_t \\ \hat{\boldsymbol{\mathcal{I}}}_{D,t} &= \hat{\boldsymbol{u}}_t\hat{\boldsymbol{d}}'_t + \hat{\boldsymbol{d}}_t\hat{\boldsymbol{u}}'_t \\ \hat{\boldsymbol{\mathcal{I}}}_{N,t} &= \hat{\boldsymbol{n}}_t\hat{\boldsymbol{n}}'_t, \end{aligned}$$

where the element of position $i$ of the $k \times 1$ vectors $\hat{\boldsymbol{u}}_t$, $\hat{\boldsymbol{d}}_t$, and $\hat{\boldsymbol{n}}_t$ is $I\left(\hat{\eta}_{i,t} \geq Q_i^{\eta}\right)$, $I\left(\hat{\eta}_{i,t} \leq -Q_i^{\eta}\right)$, and $I\left(|\hat{\eta}_{i,t}| \leq Q_i^{\eta}\right)$, respectively. The matrices $\overline{\boldsymbol{\mathcal{C}}}$, $\overline{\boldsymbol{\mathcal{D}}}$, and $\overline{\boldsymbol{\mathcal{N}}}$ are obtained as the unconditional expectations of $\hat{\boldsymbol{\mathcal{I}}}_C$, $\hat{\boldsymbol{\mathcal{I}}}_D$, and $\hat{\boldsymbol{\mathcal{I}}}_N$, respectively. Contrary to the approach of Algieri et al. (2021), who make the Gerber correlation dynamic by assuming two parameters for each possible pair of assets, the proposed model employs only six parameters in total. To estimate the parameters appearing in (8), we assume the multivariate skew-Student distribution used in the context of DCC models by Bauwens and Laurent (2005). The model again consists of equations (5)–(6) with the difference that i) the innovations follow the standardized multivariate skew-Student distribution for which the density function is given in Appendix A and ii) $\boldsymbol{R}_t$ is replaced by $\boldsymbol{G}_t$ obtained[3] from (8). With the considered density, each marginal has a specific asymmetry

[3] As in DCC models, we also consider the normalization given by the second equation of Eq. (7).

coefficient, related to the parameters in the vector $\boldsymbol{\xi}$. Kurtosis (i.e., a measure of the thickness of the tails of the return distribution), is instead captured by the parameter $\nu$, which is assumed to be common to all the $k$ assets, as in Bauwens and Laurent (2005). Under the assumption of a non-normal distribution for the innovations, the decomposition proposed by Engle (2002) is no longer possible. Nevertheless, to follow the spirit of the DCC model, as in Bauwens and Laurent (2005), we estimate the $k$ univariate GARCH-type model by quasi maximum likelihood, or QML, (to estimate the vector $\boldsymbol{\eta}$), and then estimate the parameters in the Gerber correlation part together with the vector $\boldsymbol{\xi}$ and $\nu$ related to the skew-Student density. When estimating the parameters for the dynamic correlation, we impose the constraints $\max\left\{|b_x|, |a_x|, |a_x + b_x|\right\} < 1$, $x \in \{C, D, N\}$ that imply the stationarity of the three processes appearing in (8) (see Douc et al. 2013).

### Risk measures

In this study, we make predictions for the portfolio VaR and ES assuming different forecasts for the covariance matrix (and hence for portfolio volatility). Indeed, our main aim is to assess the impact of competing methods for estimating the covariance matrix. The VaR measure is defined implicitly as

$$\mathrm{P}\left(\mathrm{r_{p,T+1}} \leq -\mathrm{VaR_{T+1|T}}(\tau)|\mathcal{F}_\mathrm{T}\right) = \tau,$$

where $r_{p,T+1}$ is the portfolio return at time $T+1$ and $\mathcal{F}_T$ is the information available up to time $T$. Expected shortfall is instead defined as

$$\mathrm{ES_{T+1|T}}(\tau) = -\frac{1}{\tau}\int_{-\infty}^{-\mathrm{VaR_{T+1|T}}(\tau)} \mathrm{yf(y)dy},$$

where $f(y)$ is the predicted density for $r_{p,T+1}$ conditional on $\mathcal{F}_T$. Given $\boldsymbol{\omega}$, the $k \times 1$ vector of portfolio weights, the forecasts for $\tau$-VaR and $\tau$-ES in the normal case are given by

$$\mathrm{VaR_{T+1|T}}(\tau) = -\boldsymbol{\omega}'\boldsymbol{\mu}_{T+1|T} - \sqrt{\boldsymbol{\omega}'\boldsymbol{\Sigma}_{T+1|T}\boldsymbol{\omega}}\Phi^{-1}(\tau) \tag{9}$$

$$\mathrm{ES_{T+1|T}}(\tau) = -\boldsymbol{\omega}'\boldsymbol{\mu}_{T+1|T} + \sqrt{\boldsymbol{\omega}'\boldsymbol{\Sigma}_{T+1|T}\boldsymbol{\omega}}\frac{\phi(\Phi^{-1}(\tau))}{\tau}, \tag{10}$$

where $\boldsymbol{\mu}_{T+1|T}$ and $\boldsymbol{\Sigma}_{T+1|T}$ are the predictions we make based on the information up to time $T$ for the mean vector and covariance matrix at time $T+1$, respectively. In the case of models involving the multivariate skew-Student distribution, no analytic formula is available for VaR and ES. We therefore rely on Monte Carlo simulations that entail simulating portfolio returns for time $T+1$ as $r_{p,j} = \boldsymbol{\omega}'\left(\boldsymbol{\mu}_{T+1|T} + \boldsymbol{\Sigma}^{1/2}_{T+1|T}\boldsymbol{z}_j\right)$, where $\boldsymbol{z}_j$ is the $j$-th simulation from the standardized skew-Student distribution and $j = 1, \ldots, J$. Using this approach, the $\tau$-VaR for time $T+1$ is estimated as minus one times the $\tau$ empirical quantile of $r_{p,j}$ over the $J$ simulations. $\mathrm{ES_{T+1|T}}(\tau)$ can be estimated as minus one times the average of simulations that are smaller than the $\tau$ empirical quantile. In our empirical application, we set $J = 100,000$.

**Evaluating VaR and ES predictions**

In this study, we employ the MCS procedure of Hansen et al. (2011) to classify the models based on their out-of-sample performance. The procedure is based on an optimality criterion such that the resulting superior set of models $M^*$ will contain the best model with a given confidence level $1-\alpha$.

It uses the idea of sequential testing, for which the generic set $M^0$, containing $m_0$ competing models, gets reduced in the number of elements by an elimination rule if the Equal Predictive Ability (EPA) null hypothesis is rejected. The procedure is iterated until the EPA hypothesis is not rejected for all the models left in the set, constituting the optimal model confidence set $M^*_{1-\alpha}$.

We use a loss function to compare forecasts from different models. In particular, the smaller is the value of the loss function for a given model, the more accurate are the predictions from the model. We denote by $l_{i,t}$ the loss associated with model $i$ at time $t$. To evaluate VaR forecasts, we use the following loss function (see for instance González-Rivera et al. 2004):

$$l_{i,t}\left(r_{p,t}, \mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)\right) = \rho_\tau\left(r_{p,t} + \mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)\right), \tag{11}$$

where $\mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)$ is the predicted $\tau$-VaR at time $t$ based on model $i$, $r_{p,t}$ is the realized portfolio return, and $\rho_\tau(u) = u(\tau - I(u<0))$.

ES is not elicitable[4] on its own, but it is jointly elicitable together with VaR using a suitable scoring function. Hence, we jointly assess VaR and ES forecasts considering the following functional form proposed by Fissler et al. (2015):

$$\begin{aligned} & l_{i,t}(r_{p,t}, \mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau), \mathrm{ES}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)) \\ & \quad = \rho_\tau\left(r_{p,t} + \mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)\right) - \tau r_{p,t} - \frac{\mathrm{ES}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)}{1+\exp\left(-\mathrm{ES}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)\right)} \times \\ & \left(\mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau) - \mathrm{ES}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau) - \mathrm{I}(\mathrm{r_{p,t}} \le -\mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)) \times \frac{\mathrm{VaR}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau) + \mathrm{r_{p,t}}}{\tau}\right) \\ & \quad + \log\left(\frac{2}{1+\exp\left(-\mathrm{ES}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)\right)}\right), \end{aligned} \tag{12}$$

where $\mathrm{ES}^{\mathrm{i}}_{\mathrm{t|t-1}}(\tau)$ is the prediction model $i$ makes for $\tau$-ES at time $t$. The remaining details of the MCS procedure are given in Appendix B.

## Empirical analysis

The dataset we use consists of weekly data for the indices described in Table 1. The data spans the period from January 22, 1999 to January 7, 2022 (1199 observations) and has been downloaded from Bloomberg. The motivation for choosing these four indices is that investors may build highly diversified portfolios by purchasing passive funds mimicking them.

[4] A measure is said to be elicitable if there exists at least one scoring function such that the correct forecast of the measure is the unique minimizer of the expectation of the scoring function.

**Table 1** Data description

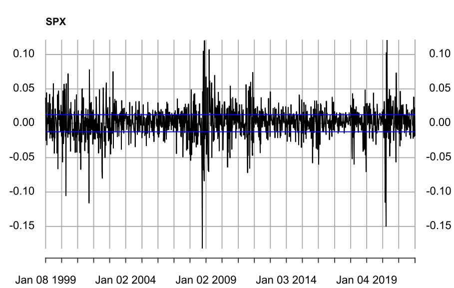
| Description | Bloomberg ticker |
|---|---|
| S &P 500 index | SPX |
| MSCI Emerging Markets index | MXEF |
| Bloomberg Barclays U.S. Aggregate Bond index | LBUSTRUU |
| Gold | XAU |

**Table 2** Summary statistics for the returns of the four indices

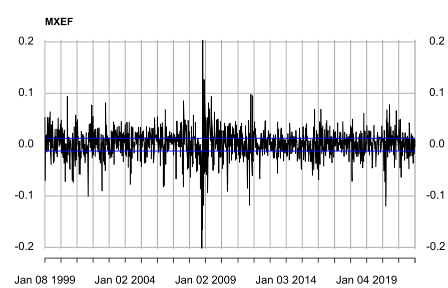
| | Min | 1st Qu. | Median | Mean | 3rd Qu. | Max | Volatility |
|---|---|---|---|---|---|---|---|
| SPX | −0.1820 | −0.0103 | 0.0014 | 0.0014 | 0.0143 | 0.1210 | 0.0248 |
| MXEF | −0.2020 | −0.0142 | 0.0016 | 0.0016 | 0.0178 | 0.2037 | 0.0296 |
| LBUSTRUU | −0.0317 | −0.0018 | 0.0009 | 0.0009 | 0.0042 | 0.0265 | 0.0051 |
| XAU | −0.3179 | −0.0276 | 0.0018 | 0.0018 | 0.0329 | 0.2359 | 0.0507 |

The return of index *j* is calculated as $r_{j,t} = P_{j,t}/P_{j,t-1} - 1$, where $P_{j,t}$ is the index price at the end of week *t*

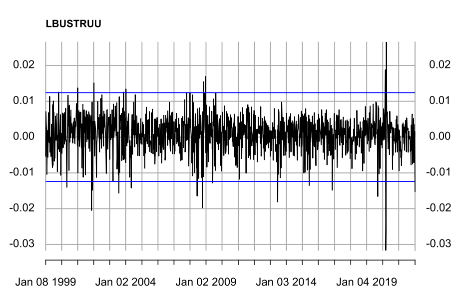
For index *j*, we compute returns as $r_{j,t} = P_{j,t}/P_{j,t-1} - 1$, where $P_{j,t}$ is the index price at the end of week *t*. Table 2 presents descriptive statistics for the simple returns of the four indices, and Fig. 1 plots their time series. In Fig. 2 we report the Pearson and Gerber correlation matrices calculated on the four time-series of returns. In the case of Gerber correlations, the thresholds are set to half the unconditional standard deviations.



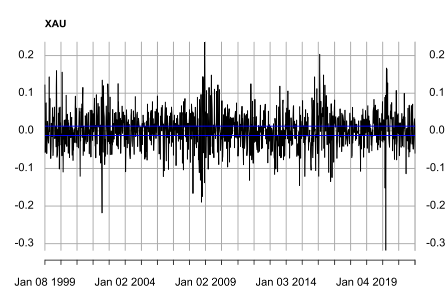

(a) SPX




(b) MXEF

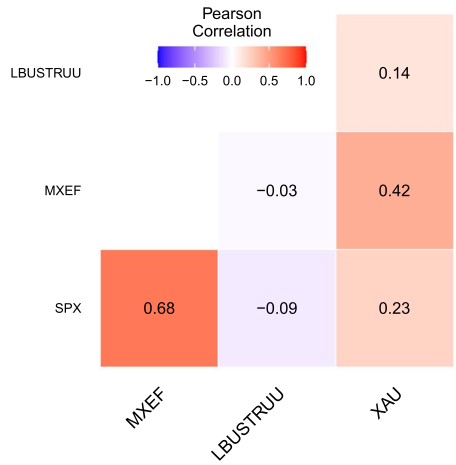



(c) LBUSTRUU

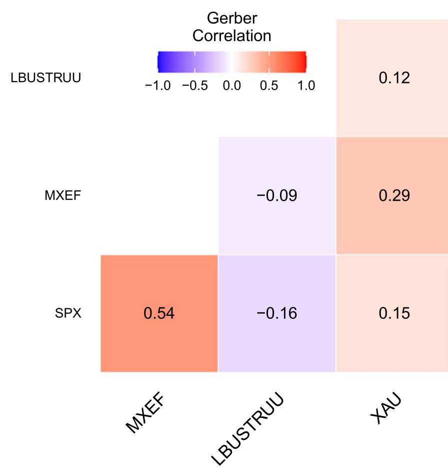



(d) XAU

**Fig. 1** Time-series plots for the returns of the four indices. The return of index *j* is calculated as $r_{j,t} = P_{j,t}/P_{j,t-1} - 1$, where $P_{j,t}$ is the index price at the end of week *t*. The horizontal blue lines represent the thresholds used in the calculation of Gerber correlations




(a) Pearson




(b) Gerber

**Fig. 2** Correlation plots for the returns of the four indices. The thresholds for the Gerber statistic are $Q_j = 0.5 \times \sigma_j$, $j = 1, \ldots, 4$, where $\sigma_j$ is the standard deviation of the returns of index $j$

## In sample analysis

In this section, we use the whole sample to estimate the DGC model. For the first stage, we use an ARMA(1,1)-GARCH(1,1) model for each of the return series. The thresholds for the Gerber statistic are assumed to be half the unconditional volatility. The estimated parameters are reported in Table 3. It is interesting to notice that the persistence[5], measured by the sum of the parameters $a$ and $b$, is lower for the equation related to the dynamics of the $\mathcal{N}$ matrix and higher for the dynamics of the $\mathcal{C}$ and $\mathcal{D}$ matrices. The estimated parameters for the part of the model involving the skew-Student distribution point to a negative skewness ($\hat{\xi}_j < 1$ for $j = 1, \ldots, 4$) and excess kurtosis due to the relatively small estimate for the degrees of freedom parameter ($\hat{\nu} = 6.91$). Fig. 3

[5] The persistence is related to the speed at which the process reverts back to its long run mean. A high persistence, i.e. $a + b$ close to 1, implies that shocks that push the process away from its long run mean will persist for a long time.

**Table 3** Univariate GARCH Models (Panel A), skew-Student (Panel B), and DGC Model Estimates (Panel C)

| | SPX | MXEF | LBUSTRUU | XAU |
|---|---|---|---|---|
| Panel A: Univariate GARCH models | | | | |
| Const. (mean eq.) | 0.0025*** | 0.0021** | 0.0008*** | 0.0002 |
| | (0.0005) | (0.0009) | (0.0001) | (0.0010) |
| ar1 | −0.7533*** | 0.7130*** | −0.3248 | 0.9167*** |
| | (0.1338) | (0.2387) | (0.2619) | (0.0360) |
| ma1 | 0.6826*** | −0.6639*** | 0.2237 | −0.9387*** |
| | (0.1500) | (0.2553) | (0.2686) | (0.0327) |
| const. (variance eq.) | 0.0000*** | 0.0001*** | 0.0000*** | 0.0002*** |
| | (0.0000) | (0.0000) | (0.0000) | (0.0001) |
| ARCH | 0.2635*** | 0.1310*** | 0.1103*** | 0.0981*** |
| | (0.0355) | (0.0234) | (0.0107) | (0.0210) |
| GARCH | 0.6996*** | 0.7907*** | 0.7861*** | 0.8251*** |
| | (0.0341) | (0.0350) | (0.0168) | (0.0416) |
| Panel B: Skew-Student | | | | |
| $\xi$ | 0.8086*** | 0.8518*** | 0.8214*** | 0.9089*** |
| | (0.0328) | (0.0348) | (0.0324) | (0.0363) |
| $\nu$ | 6.9184*** | | | |
| | (0.5453) | | | |
| Panel C: DGC Model | | | | |
| | | $a_C$ | 0.0235*** | |
| | | | (0.0006) | |
| | | $b_C$ | 0.9775*** | |
| | | | (0.0067) | |
| | | $a_D$ | 0.0546*** | |
| | | | (0.0097) | |
| | | $b_D$ | 0.9234*** | |
| | | | (0.0346) | |
| | | $a_N$ | 0.0120*** | |
| | | | (0.0185) | |
| | | $b_N$ | 0.7835*** | |
| | | | (0.0080) | |

We fit an ARMA(1,1)-GARCH(1,1) model for each of the return series. The thresholds for the Gerber statistic are assumed to be half the unconditional volatility. The table reports the parameters estimated via QML (standard errors in parenthesis). '***', '** ' and '*' denote significance at the 1%, 5% and 10% level, respectively

plots the Gerber correlations obtained from the estimated DGC model via the recursions (8) using the estimated parameters of Table 3. Some interesting features can be gleaned from the figure. For instance, it is possible to observe an increase in the correlation between equities and gold after 2019 (see panels c and d). Furthermore, the correlation between bonds and gold became negative in the aftermath of the financial crisis of 2008 and became again positive around 2015.

### Out-of-sample analysis

Here, the focus is on estimating VaR and ES for the three portfolios of Table 4. The first one is the equally-weighted portfolio. The remaining two portfolios both invest 10% in gold. However, they differ in the equity and bond allocation. Indeed, portfolio 2 invests

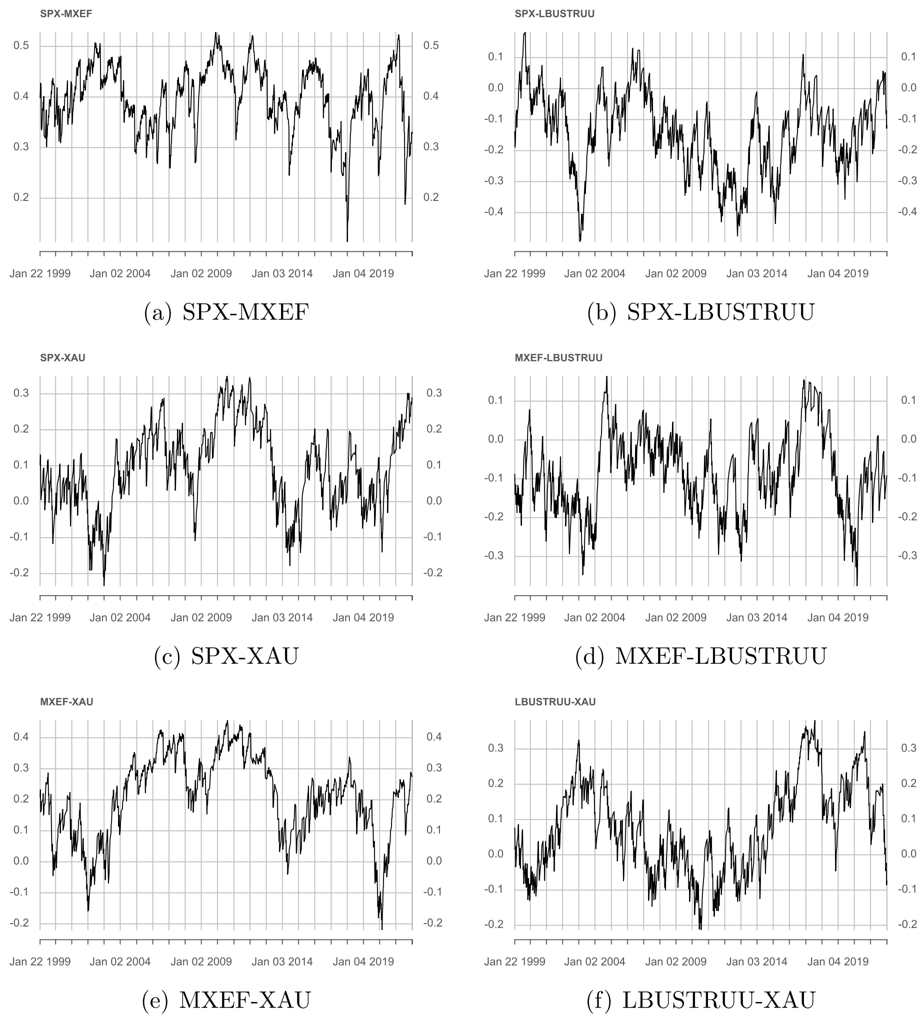



**Fig. 3** Time-Series Plots of the Gerber correlations based on the estimated DGC model

approximately 60% of the allocation (excluding gold) in equities. On the contrary, portfolio 3 invests approximately 60% of the allocation (excluding gold) in bonds. The rationale for choosing the last two sets of weights is that many investors hold a 60/40 or 40/60 portfolio with the addition of a relatively small share of gold[6].

We assess how well the proposed DGC model predicts VaR and ES relative to a number of alternative methods. We consider non-parametric, parametric, and semi-parametric methods. The first benchmark is a non-parametric method, historical simulation (HS). The method considers the last $T$ portfolio returns and estimates $\tau$-VaR as the negative of their $\tau$-quantile. Instead, $\tau$-ES is estimated under HS as minus one times the mean of portfolio returns that are less than the $\tau$-quantile. The next two methods we consider, labelled PearsonHist and GerberHist, are for the 'static' models relying on the Pearson and Gerber correlation matrices, respectively.

[6] For example, a recent article about the 60/40 portfolio and the '10% Golden rule' can be found at https://www.forbes.com/sites/greatspeculations/2023/01/10/is-the-6040-portfolio-a-thing-of-the-past-not-so-fast/

**Table 4** Weights of the three portfolios

| Asset | SPX | MXEF | LBUSTRUU | XAU |
|---|---|---|---|---|
| Portfolio 1 | 0.25 | 0.25 | 0.25 | 0.25 |
| Portfolio 2 | 0.475 | 0.075 | 0.35 | 0.1 |
| Portfolio 3 | 0.3 | 0.05 | 0.55 | 0.1 |

In particular, they use the last $T$ observations of the returns on the $k$ assets to estimate the (Pearson or Gerber) correlation matrix, $\boldsymbol{\Omega}_{T+1|T}$, and estimate the covariance matrix as $\boldsymbol{\Sigma}_{T+1|T} = \boldsymbol{\Lambda}_{T+1|T}\boldsymbol{\Omega}_{T+1|T}\boldsymbol{\Lambda}_{T+1|T}$, where $\boldsymbol{\Lambda}_{T+1|T}$ is the matrix with the sample volatilities in the main diagonal and zero elsewhere. In the case of the GerberHist method, Eq. (2) is used and the parameters of the multivariate skew-Student are estimated. With this method, VaR and ES are derived using Monte Carlo simulations as described at the end of section "Risk measures". However, for the PearsonHist method, VaR and ES are estimated using (9)–(10), with $\boldsymbol{\mu}_{T+1|T}$ equal to the vector of sample means for the returns on the $k$ assets. The next alternative method we consider is the DCC model (see Sect. 2.2). In this case, we use ARMA(1,1)-GARCH(1,1) for the univariate specifications. VaR and ES are based on the mean prediction from each univariate ARMA(1,1) model and on the predicted covariance matrix $\boldsymbol{\Sigma}_{T+1|T} = \boldsymbol{\Lambda}_{T+1|T}\boldsymbol{R}_{T+1|T}\boldsymbol{\Lambda}_{T+1|T}$, where $\boldsymbol{\Lambda}_{T+1|T}$ is the diagonal matrix consisting of the predicted volatilities from each univariate GARCH(1,1) model. For the proposed DGC model, the covariance matrix we use in the estimation of VaR and ES is instead $\boldsymbol{\Sigma}_{T+1|T} = \boldsymbol{\Lambda}_{T+1|T}\boldsymbol{G}_{T+1|T}\boldsymbol{\Lambda}_{T+1|T}$. Note that in the DCC and DGC models, we use the same univariate specification—namely, the ARMA(1,1)-GARCH(1,1) model—for each of the return series. Furthermore, to have a meaningful comparison of the DGC and DCC models, for the latter we assume that the innovations follow the multivariate skew-Student distribution as well. Hence, similarly for DCC models, the risk measures of interest are derived using Monte Carlo simulations. Finally, we consider the semi-parametric method filtered historical simulation (FHS) together with the DCC or DGC model. We implement the method as follows. We first estimate an ARMA(1,1)-GARCH(1,1) model on the last $T$ portfolio returns. Denote by $\{z_t\}_{t=1,\ldots,T}$ the standardized portfolio returns, that is, $z_t = y_t/\sigma_t$, where $y_t$ is the portfolio return at time $t$, and $\sigma_t$ the time $t$ volatility from the estimated GARCH(1,1) model. VaR and ES are then estimated by modifying (9)–(10) in the following way: i) $\Phi^{-1}(\tau)$ is replaced by the sample $\tau$-quantile of the $z$ series, $q_z(\tau)$, and ii) $\frac{\phi(\Phi^{-1}(\tau))}{\tau}$ is replaced by minus one times the mean of the $z_t$ that are smaller than $q_z(\tau)$.

Before presenting the results of the out-of-sample analysis, we first report the results of a Monte Carlo experiment we ran to evaluate the considered methods. To generate scenarios for the assets of interest, we use the moment-matching method of Høyland et al. (2003). We chose an agnostic approach to identify the data-generating process for the simulations to have a fair comparison of the alternative methods for deriving the covariance matrix. For each simulation, we generate 501 observations for the returns of the four assets and use the first 500 observations[7] to estimate all the models and predict

[7] The window coincides with the one we move in the out-of-sample analysis based on the real-world returns.

**Table 5** Average of the loss functions (11)–(12) over the 1000 Monte Carlo simulations. The minimum loss is in bold

| | 1% VaR | 5% VaR | 10% VaR | 1% VaR - 1% ES | 5% VaR - 5% ES | 10% VaR - 10% ES |
|---|---|---|---|---|---|---|
| Panel A: Portfolio 1 | | | | | | |
| HS | 0.7699 | 2.5558 | 3.9685 | 39.1269 | 28.0308 | 23.7402 |
| PearsonHist | 0.8293 | 2.5635 | 4.0199 | 27.3049 | 25.3705 | 21.6851 |
| GerberHist | 0.7704 | 2.5604 | 3.9598 | 37.2304 | 25.6947 | 22.1826 |
| DGC | 0.8002 | 2.5653 | **3.9464** | **24.9296** | 24.0601 | **20.8163** |
| DGC FHS | **0.7260** | **2.5548** | 3.9621 | 38.5319 | 27.7227 | 23.5431 |
| DCC | 0.7676 | 2.5615 | 3.9884 | 37.1663 | **23.4299** | 23.0511 |
| DCC FHS | 0.7541 | 2.5667 | 3.9759 | 39.0725 | 28.2222 | 24.0479 |
| Panel B: Portfolio 2 | | | | | | |
| HS | 0.6161 | 1.8858 | 3.0129 | 32.4051 | 21.7228 | 17.8902 |
| PearsonHist | 0.6113 | 1.9277 | 3.0384 | 26.5444 | **17.4298** | 16.1318 |
| GerberHist | 0.5805 | 1.9273 | 3.0050 | 28.6044 | 19.5020 | 16.7387 |
| DGC | 0.5828 | 1.9290 | **3.0011** | **24.9169** | 18.7227 | **16.0257** |
| DGC FHS | **0.5183** | 1.8857 | 3.0069 | 31.8438 | 21.5732 | 17.8413 |
| DCC | 0.5778 | 1.9212 | 3.0241 | 28.3112 | 19.9969 | 17.3912 |
| DCC FHS | 0.6090 | **1.8842** | 3.0072 | 31.9622 | 21.9165 | 18.2532 |
| Panel C: Portfolio 3 | | | | | | |
| HS | 0.4303 | 1.3579 | 2.1738 | 22.0491 | 15.2718 | 12.7932 |
| PearsonHist | 0.4318 | 1.3761 | 2.1937 | 20.0108 | 12.6227 | 11.6485 |
| GerberHist | 0.4141 | 1.3745 | **2.1730** | 20.3838 | 14.0323 | 12.0815 |
| DGC | **0.4092** | 1.3736 | 2.1754 | 19.6704 | **12.3838** | **11.5167** |
| DGC FHS | 0.4340 | **1.3525** | 2.1736 | **19.6430** | 15.1511 | 12.7428 |
| DCC | 0.4157 | 1.3768 | 2.1868 | 20.2607 | 14.4648 | 12.6236 |
| DCC FHS | 0.4353 | 1.3560 | 2.1742 | 21.9015 | 15.5490 | 13.1683 |

VaR and ES for the three portfolios described in Table 4. We consider three different values of $\tau$—1%, 5%, and 10%—when predicting VaR and ES. The last observation is used to derive the simulated portfolio return and hence to compute the losses (11)–(12) together with the predicted risk measures. The number of simulations in our experiment is 1,000. The results of the Monte Carlo experiment are given in Table 5. From the table, we can see that in the considered experiments the proposed DGC and DGC FHS methods are often the ones for which the two losses are minimized.

In the out-of-sample analysis based on real-world (rather than simulated) returns, we move a window of length $T = 500$ to estimate the parameters (for parametric models) or to derive the empirical VaR or ES under the HS method. Consequently, the first prediction we make is for August 22, 2008. As in the Monte Carlo experiments, $\tau$ is equal to 1%, 5%, or 10%.

## Results

In this section, we first present the results of the ES-backtesting procedure[8] based on the results of Khalaf et al. (2021) (see Appendix C). We opt for this recently proposed procedure because it allows backtesting ES only based on a sequence of violations for an appropriately chosen sequence of VaRs. Therefore, contrary to the Du and Escanciano (2017) tests, it does not require knowing the entire conditional cumulative density function, which is difficult to derive, for instance, in the case of FHS methods. We consider conditional tests based on lags corresponding to one week, one month, and two months, that is, $m = 1$, $m = 5$, and $m = 10$. The results of the tests are reported for all the alternative models we consider in Table 6. Overall, models based on the dynamic Gerber correlation seem to have the best performance. Indeed, they perform well for portfolio 3 (Panel C) for all the three values of $\tau$. For the first two portfolios (Panels A and B), instead, the null of "accurate" ES predictions (see Eq. (24)) is rejected for some values of $m$ when $\tau = 1\%$ but not for the two remaining values of $\tau$. Despite the fact that it is based on a distribution that captures skewness and excess kurtosis in the returns—like all the alternative models—in this particular case, the DGC model fails to accurately capture the tail expectation for small values of $\tau$ if portfolios consist mainly of highly volatile assets such as equities (such as portfolio 1 and even more so portfolio 2).

As a robustness check, we run the backtests for two sub-samples: the first one covers a relatively tranquil period (January 2017–November 2019), while the second one covers a more turbulent period (February 2020–January 2022). Regarding the results for the two sub-samples, reported in Tables 7 and 8, two observations are in order. First, the proposed methods are not rejected taken as a whole; second, the results for the full sample, not surprisingly, seem to be driven by the more turbulent period that includes the COVID-19 pandemic and the beginning of the Russian-Ukrainian conflict.

Before presenting the results of the MCS procedure, we run a number of Diebold-Mariano (DM) tests (Diebold and Mariano 2002) to assess the accuracy of the proposed models in predicting VaR alone and VaR and ES jointly. We focus on the three portfolios of Table 4 and test the null hypothesis $H_0 : \mathbb{E}(d_{ij}) = 0$, where $d_{ij}$ is the loss differential between model $i$ and model $j$. We assume that model $i$ is one of the proposed models (i.e., DGC or DGC FHS) and model $j$ is one of the competing models (i.e., HS, PearsonHist, GerberHist, DCC, or DCC FHS). The results of the DM tests for the forecasting of VaR are given in Table 9, whereas the results for the case of joint VaR/ES forecasting can be found in Table 10. It is evident from both tables that often (and especially in the case of joint forecasting of the two risk measures) the null hypothesis is rejected and the test statistic is negative. When this is the case, the proposed methods (DGC or DGC FHS) are more accurate in predicting the two risk measures than the competing methods. In the few cases in which the test statistic is positive, we instead observe a lack of stars in the table, meaning that the null is not rejected and hence the proposed methods are as accurate as the alternative methods.

[8] We do not report the results of VaR backtesting procedures (see for instance Christoffersen 1998) since the statistical test we consider is based on cumulative violations associated with a sequence of quantiles in the left tail. Indeed some of the tests of Khalaf et al. (2021) are obtained as combinations of the $p$ values from VaR backtesting procedures. However, the results of VaR backtesting procedures are available on request from the authors.

**Table 6** *p* values (×100) for the backtesting procedure of Khalaf et al. (2021), whole sample

| | $\tau = 1\%$ | | | $\tau = 5\%$ | | | $\tau = 10\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $m = 1$ | $m = 5$ | $m = 10$ | $m = 1$ | $m = 5$ | $m = 10$ | $m = 1$ | $m = 5$ | $m = 10$ |
| Panel A: Portfolio 1 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| GerberHist | 0.0020 | 0.0000 | 0.0000 | 0.0000 | 0.0040 | 0.0000 | 0.0001 | 0.0000 | 0.0100 |
| DGC | 6.0025 | 0.0001 | 0.0106 | 7.1795 | 14.0262 | 21.9860 | 2.6068 | 26.2252 | 58.9214 |
| DGC FHS | 7.4067 | 4.7297 | 3.1148 | 8.0244 | 7.7298 | 7.1632 | 1.3875 | 20.3608 | 46.1753 |
| DCC | 0.0028 | 0.0000 | 0.0020 | 0.1015 | 1.7016 | 12.9372 | 0.6747 | 12.6533 | 34.7828 |
| DCC FHS | 0.0000 | 0.0000 | 0.0000 | 0.0374 | 0.3705 | 4.3314 | 0.3061 | 6.4895 | 22.7918 |
| Panel B: Portfolio 2 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| GerberHist | 2.2204 | 1.2204 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| DGC | 1.5374 | 0.9000 | 6.0233 | 15.4523 | 12.1693 | 13.3083 | 9.4346 | 6.2292 | 16.0223 |
| DGC FHS | 8.3597 | 0.0000 | 0.0000 | 0.0019 | 8.0403 | 0.7399 | 2.6583 | 5.7975 | 18.3711 |
| DCC | 0.0000 | 0.0000 | 0.0000 | 0.0590 | 0.7849 | 6.4119 | 1.7306 | 4.7454 | 15.5384 |
| DCC FHS | 0.0000 | 0.0000 | 0.0000 | 0.0162 | 0.0803 | 1.1697 | 0.3522 | 1.7616 | 8.0228 |
| Panel C: Portfolio 3 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0006 | 0.0000 | 0.0000 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0019 | 0.0000 | 0.0000 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.0027 | 0.0100 |
| DGC | 7.5072 | 10.0000 | 9.0013 | 13.8995 | 11.3663 | 41.5989 | 19.0026 | 36.6683 | 70.7112 |
| DGC FHS | 8.0002 | 3.1419 | 11.3863 | 2.8305 | 1.7527 | 8.7166 | 23.1811 | 32.4184 | 76.6040 |
| DCC | 5.0045 | 1.2786 | 3.4637 | 4.7322 | 3.5413 | 14.6516 | 42.6397 | 45.8370 | 85.8272 |
| DCC FHS | 1.0040 | 1.7231 | 6.7201 | 1.1881 | 0.5072 | 3.4448 | 30.1073 | 37.3137 | 78.0058 |

For each portfolio, for $m \in \{1, 5, 10\}$, and for $\tau \in \{1, 5, 10\}/100$, the table reports the *p* values (multiplied by 100) for the test statistic $C_m(\tau)$, Eq. (26). *K* is set to 100 and the *p* values are obtained with the Monte Carlo method based on 50,000 simulations

Next, we perform the MCS procedure using $\alpha = 5\%$ and 5,000 bootstrap replications to derive the *p* values associated with the test statistics involved. We report the SSMs for the three portfolios we consider in Tables 11, 12, and 13 for the case of VaR predictions and in Tables 14, 15, and 16 for the case of joint VaR/ES predictions. Each table reports, for each model *i* belonging to the identified SSM, the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (11) when we forecast VaR or the average of (12) when we jointly forecast VaR and ES.

In the tables involving VaR predictions alone, there are just a few cases where the SSM does not include all the considered models. When that happens, just one model is excluded. For portfolio 1, we have that all the models belong to the SSM when the VaR confidence level, $\tau$, is equal to 1% or 5%, and that only model PearsonHist gets excluded when $\tau = 10\%$. For portfolio 2, all the models belong to the SSM when $\tau = 1\%$, only the HS model gets excluded when $\tau = 5\%$, and only the PearsonHist model gets excluded when $\tau = 10\%$. Finally, for portfolio 3, all the models belong to the SSM for all three levels of $\tau$ we consider. Hence, the results regarding VaR predictions do not evidence much difference between static and dynamic models in terms of inclusion in the SSM, although the few excluded ones are static. It is

**Table 7** *p* values (×100) for the backtesting procedure of Khalaf et al. (2021), January 2017–November 2019

| | $\tau = 1\%$ | | | $\tau = 5\%$ | | | $\tau = 10\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $m = 1$ | $m = 5$ | $m = 10$ | $m = 1$ | $m = 5$ | $m = 10$ | $m = 1$ | $m = 5$ | $m = 10$ |
| Panel A: Portfolio 1 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 0.1455 | 2.8975 | 5.6362 | 0.0082 | 0.1756 | 0.0253 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 20.2966 | 89.2576 | 29.1040 | 0.0610 | 2.3095 | 0.1113 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 1.7266 | 33.5739 | 9.7694 | 0.1265 | 3.5479 | 0.4600 |
| GDC | 7.2121 | 17.7198 | 10.2747 | 13.5794 | 11.1435 | 7.9907 | 12.5412 | 21.4258 | 18.0071 |
| GDC FHS | 24.8057 | 31.4193 | 13.8628 | 14.6578 | 13.5873 | 11.6964 | 10.2616 | 12.7191 | 8.2929 |
| DCC | 0.0045 | 0.0000 | 0.0000 | 0.2443 | 8.7058 | 8.9560 | 0.8453 | 13.3445 | 10.3673 |
| DCC FHS | 0.0040 | 0.0000 | 0.0000 | 0.5764 | 15.6800 | 12.8061 | 0.8507 | 12.2145 | 7.5399 |
| Panel B: Portfolio 2 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 1.1194 | 23.3195 | 6.2130 | 11.8774 | 23.8972 | 9.7138 |
| PearsonHist | 0.0001 | 0.0000 | 0.0000 | 1.7168 | 22.5589 | 9.0363 | 7.6930 | 36.6896 | 16.9050 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 2.3827 | 24.2020 | 8.6890 | 15.8267 | 38.6520 | 14.5047 |
| GDC | 5.3416 | 28.3986 | 10.4706 | 24.4839 | 23.0942 | 14.3320 | 16.5033 | 26.2443 | 28.0296 |
| GDC FHS | 9.7831 | 36.0637 | 0.3027 | 12.3551 | 32.0767 | 32.3979 | 33.2941 | 28.6718 | 22.5274 |
| DCC | 3.6401 | 3.3933 | 5.4757 | 13.6402 | 35.3717 | 40.3591 | 35.3736 | 27.4298 | 22.4661 |
| DCC FHS | 1.8576 | 0.0829 | 0.4928 | 12.7322 | 30.3934 | 39.9827 | 29.2314 | 28.6863 | 25.1951 |
| Panel C: Portfolio 3 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 9.7026 | 37.3556 | 13.5201 | 58.5694 | 5.6765 | 2.4581 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 16.3205 | 42.0567 | 13.0008 | 55.0385 | 8.6600 | 2.9224 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 18.7918 | 32.5652 | 8.0423 | 62.3442 | 6.3332 | 3.0806 |
| GDC | 24.7931 | 10.8623 | 7.1023 | 16.8481 | 13.3563 | 6.2865 | 14.2379 | 8.0766 | 10.4443 |
| GDC FHS | 48.8139 | 47.5196 | 5.2555 | 26.1578 | 20.0726 | 8.7811 | 6.8517 | 17.2849 | 15.8078 |
| DCC | 6.6921 | 0.2232 | 0.0270 | 22.1855 | 24.7335 | 14.1008 | 17.3643 | 28.5591 | 32.2569 |
| DCC FHS | 4.9044 | 1.2805 | 2.1786 | 28.9667 | 27.3396 | 13.2530 | 24.2171 | 31.3911 | 35.8729 |

For each portfolio, for $m \in \{1, 5, 10\}$, and for $\tau \in \{1, 5, 10\}/100$, the table reports the *p* values (multiplied by 100) for the test statistic $C_m(\tau)$, Eq. (26). *K* is set to 100 and the *p* values are obtained with the Monte Carlo method based on 50,000 simulations

worth highlighting that the newly proposed models, DGC and DGC FHS, can produce VaR predictions that allow them to belong to the SSM. From the second and fifth columns in each panel of Tables 11–13, we see that dynamic models usually rank high among the considered models. Based on the statistic $\max_j T_{ij}$, the models DCC, DCC FHS, DGC, and DGC FHS are always in the first four positions. Based instead on the statistics $T_{i\cdot}$, there are more cases where some of the static models are ranked in the first four positions, but, in any case, dynamic models are always in the first two positions. However, at least one of the newly proposed models, DGC and DGC FHS, usually ranks high with the statistic $T_{i\cdot}$. To justify these results, we observe that the out-of-sample analysis uses a period—August 22, 2008 to January 7, 2022—characterized by episodes of turmoil in the markets. Indeed, the period includes the final part of the global financial crisis of 2007–2008, the European sovereign debt crisis of 2009 through the mid-to-late 2010s, 2011 Black Monday, and the stock market crash of 2020 caused by the COVID-19 pandemic. Dynamic models rank high in Tables 11–13 because they describe financial time series in the presence of turmoil better than static models do. However, the rolling-window approach applied to make VaR predictions allows static models to stay in the SSM.

**Table 8** *p* values (×100) for the backtesting procedure of Khalaf et al. (2021), February 2020–January 2022

| | $\tau = 1\%$ | | | $\tau = 5\%$ | | | $\tau = 10\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $m = 1$ | $m = 5$ | $m = 10$ | $m = 1$ | $m = 5$ | $m = 10$ | $m = 1$ | $m = 5$ | $m = 10$ |
| Panel A: Portfolio 1 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 1.1977 | 0.0754 | 1.8350 | 28.6672 | 50.4950 | 82.5708 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 0.7232 | 0.0457 | 1.2576 | 15.8907 | 24.5196 | 61.1537 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 2.6894 | 0.8391 | 9.6165 | 34.1538 | 67.2814 | 89.8548 |
| GDC | 0.1721 | 1.7720 | 1.2747 | 59.4891 | 43.8343 | 73.9660 | 49.7254 | 84.1690 | 92.3429 |
| GDC FHS | 12.4806 | 31.4193 | 13.8628 | 62.1201 | 14.0269 | 49.6200 | 45.2885 | 66.8318 | 87.4482 |
| DCC | 0.0045 | 0.0000 | 0.0000 | 61.3328 | 18.2047 | 49.2745 | 40.6929 | 63.1000 | 84.5219 |
| DCC FHS | 0.0040 | 0.0000 | 0.0000 | 62.3677 | 7.2243 | 35.9006 | 40.0455 | 59.5790 | 85.1275 |
| Panel B: Portfolio 2 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 26.2429 | 50.0545 | 85.2947 | 66.6324 | 90.1387 | 97.5782 |
| PearsonHist | 0.0001 | 0.0000 | 0.0000 | 28.4325 | 55.6232 | 86.9214 | 71.9227 | 95.4357 | 97.7074 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 26.3508 | 54.3791 | 87.8084 | 68.2522 | 92.8547 | 97.5494 |
| GDC | 15.3416 | 2.8399 | 2.4706 | 98.4346 | 79.7314 | 90.2162 | 68.1556 | 88.8441 | 97.3513 |
| GDC FHS | 15.9783 | 36.0637 | 30.2659 | 78.0868 | 81.1346 | 91.1787 | 89.2281 | 96.6572 | 98.2757 |
| DCC | 3.6401 | 3.3933 | 5.4757 | 57.9466 | 87.7371 | 94.3259 | 81.6790 | 98.3507 | 97.7324 |
| DCC FHS | 1.8576 | 0.0829 | 0.4928 | 58.1527 | 77.9000 | 89.1498 | 77.8662 | 98.2108 | 97.5602 |
| Panel C: Portfolio 3 | | | | | | | | | |
| HS | 0.0000 | 0.0000 | 0.0000 | 39.5732 | 47.3020 | 80.8825 | 24.6207 | 35.4951 | 62.5620 |
| PearsonHist | 0.0000 | 0.0000 | 0.0000 | 55.8345 | 74.1486 | 91.8199 | 74.4330 | 89.4889 | 96.9218 |
| GerberHist | 0.0000 | 0.0000 | 0.0000 | 44.6282 | 57.8897 | 85.8300 | 42.8926 | 61.0831 | 86.4168 |
| GDC | 10.6723 | 32.4039 | 17.1023 | 44.6146 | 43.1850 | 70.9851 | 23.1106 | 22.5168 | 37.9144 |
| GDC FHS | 14.8934 | 10.4756 | 12.3221 | 49.3767 | 56.6540 | 82.0182 | 30.1941 | 48.2064 | 74.2310 |
| DCC | 6.6921 | 0.2232 | 0.0270 | 75.6561 | 73.2815 | 87.7229 | 63.0551 | 73.7918 | 90.7133 |
| DCC FHS | 4.9044 | 1.2805 | 2.1786 | 84.6642 | 72.5471 | 88.1409 | 66.1312 | 76.0137 | 90.3140 |

For each portfolio, for $m \in \{1, 5, 10\}$, and for $\tau \in \{1, 5, 10\}/100$, the table reports the *p* values (multiplied by 100) for the test statistic $C_m(\tau)$, Eq. (26). *K* is set to 100 and the *p* values are obtained with the Monte Carlo method based on 50,000 simulations

As far as the joint prediction of VaR and ES is concerned, from Tables 14–16 we see that all the SSM consist only of the newly proposed DGC model. The results are consistent across portfolios and confidence levels of the two risk measures we forecast. This means that, for the three portfolios consisting of the four assets considered in our study and for the period of investigation, the DGC model produces joint VaR/ES predictions that are more accurate than the competing models.

## Conclusions

Methods that produce accurate forecasts of risk measures like VaR and ES are essential in an environment where market turmoil and substantial losses for investors are becoming increasingly frequent. This study introduces a new method, DGC, to predict the two risk measures based on the dynamic version of the robust correlation proposed by Gerber et al. (2022) that extends Kendall's Tau. As in DCC models, the

**Table 9** Test statistics for the Diebold-Mariano testing procedures, $\tau$-VaR forecasting

| Model *j* | Portfolio 1 | | | Portfolio 2 | | | Portfolio 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\tau = 1\%$ | $\tau = 5\%$ | $\tau = 10\%$ | $\tau = 1\%$ | $\tau = 5\%$ | $\tau = 10\%$ | $\tau = 1\%$ | $\tau = 5\%$ | $\tau = 10\%$ |
| Panel A: Model *i* DGC Model | | | | | | | | | |
| HS | − 0.9658 | − 1.6677* | − 1.1785 | − 1.2933 | − 2.0165** | − 1.7740* | − 1.0237 | − 1.7315* | − 1.2048 |
| Pearson-Hist | − 1.1679 | − 1.3498 | − 2.0140** | − 1.3826 | − 1.7239* | − 2.4308** | − 1.0963 | − 1.6086 | − 1.8235* |
| Gerber-Hist | − 1.1743 | − 1.2581 | − 1.3949 | − 1.3419 | − 1.6773* | − 1.9701** | − 1.0437 | − 1.4613 | − 1.2531 |
| DCC | 1.7133 | 0.5443 | 0.2518 | 1.5017 | 1.5500 | 0.6556 | 1.5830 | 0.9800 | 0.6346 |
| DCC FHS | 0.8929 | 0.1543 | − 0.1117 | 0.5383 | 1.0075 | 0.1149 | 0.8264 | 0.4573 | 0.2659 |
| Panel B: Model *i* DGC FHS Model | | | | | | | | | |
| HS | − 1.1586 | − 1.7355* | − 1.2173 | − 1.3569 | − 2.1184** | − 1.6040 | − 1.1540 | − 1.7161* | − 1.1023 |
| Pearson-Hist | − 1.1640 | − 1.4443 | − 2.1491** | − 1.3823 | − 1.8315* | − 2.3495** | − 1.1640 | − 1.6150 | − 1.7716* |
| Gerber-Hist | − 1.2163 | − 1.3174 | − 1.4395 | − 1.3716 | − 1.7661* | − 1.8111* | − 1.1393 | − 1.4499 | − 1.1627 |
| DCC | 1.2040 | 0.7576 | 0.0181 | 0.9357 | 1.9561* | 1.3835 | 1.3413 | 1.4659 | 1.1195 |
| DCC FHS | 1.4538 | − 0.1292 | − 0.6799 | 0.7873 | 1.0003 | 0.4812 | 1.1271 | 0.5378 | 0.5872 |

The null hypothesis is $H_0 : \mathbb{E}(d_{ij}) = 0$, where $d_{ij}$ is the loss differential between model *i* (DGC in Panel A and DGC FHS in Panel B) and model *j* (one of the alternative model). '***', '** ' and '*' denote significance at the 1%, 5% and 10% level, respectively

**Table 10** Test statistics for the Diebold-Mariano testing procedures, joint $\tau$-VaR and $\tau$-ES forecasting

| Model *j* | Portfolio 1 | | | Portfolio 2 | | | Portfolio 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\tau = 1\%$ | $\tau = 5\%$ | $\tau = 10\%$ | $\tau = 1\%$ | $\tau = 5\%$ | $\tau = 10\%$ | $\tau = 1\%$ | $\tau = 5\%$ | $\tau = 10\%$ |
| Panel A: Model *i* DGC Model | | | | | | | | | |
| HS | − 10.3140*** | − 5.7359*** | − 5.0365*** | − 6.1134*** | − 4.0855*** | − 3.4918*** | − 7.4212*** | − 4.4894*** | − 3.6717*** |
| Pearson-Hist | − 0.0931 | − 2.2036** | − 3.3620*** | 0.3686 | − 1.2839 | − 2.1531** | 0.1628 | − 1.5048 | − 2.3574** |
| Gerber-Hist | − 4.5518*** | − 3.0447*** | − 3.1253*** | − 2.3297** | − 2.1089** | − 2.2157** | − 2.6893*** | − 2.1460** | − 2.1539** |
| DCC | − 14.1124*** | − 15.6085*** | − 16.6673*** | − 19.1266*** | − 37.1434*** | − 32.3286*** | − 17.5401*** | − 23.5166*** | − 24.7270*** |
| DCC FHS | − 11.3959*** | − 12.9127*** | − 13.9270*** | − 13.6566*** | − 18.2288*** | − 21.6271*** | − 13.7780*** | − 14.6926*** | − 18.2965*** |
| Panel B: Model *i* DGC FHS Model | | | | | | | | | |
| HS | − 4.2635*** | − 1.6281 | − 1.2469 | − 3.2288*** | − 1.4229 | − 1.0029 | − 3.4493*** | − 1.4325 | − 0.9623 |
| Pearson-Hist | − 4.2472*** | 1.4752 | 0.1399 | − 3.1267*** | 1.0866 | 0.1207 | − 3.8656*** | 1.0681 | 0.1103 |
| Gerber-Hist | − 0.7340 | − 0.7031 | − 0.3586 | 0.5153 | 0.3409 | 0.1125 | 1.2544 | 0.4958 | 0.3006 |
| DCC | − 0.6037 | − 0.7510 | − 0.5277 | − 1.0536 | − 0.9846 | − 0.6778 | − 1.0607 | − 0.6848 | − 2.1082 |
| DCC FHS | − 3.5248*** | − 8.7083*** | − 9.9503*** | − 5.7700*** | − 22.3854*** | − 23.2669*** | − 5.7832*** | − 14.8289*** | − 19.8898*** |

The null hypothesis is $H_0 : \mathbb{E}(d_{ij}) = 0$, where $d_{ij}$ is the loss differential between model *i* (DGC in Panel A and DGC FHS in Panel B) and model *j* (one of the alternative model). '***', '** ' and '*' denote significance at the 1%, 5% and 10% level, respectively

**Table 11** Superior Set of Models, $\tau$-VaR forecasting, Portfolio 1

| **Model *i*** | $\text{Rank}_{T_R}$ | $\max_j T_{ij}$ | ***p*-value**$_{T_R}$ | $\text{Rank}_{T_{max}}$ | $T_{i\cdot}$ | ***p*-value**$_{T_{max}}$ | **Loss**$\times 10^3$ |
|---|---|---|---|---|---|---|---|
| Panel A: $\tau = 1\%$ | | | | | | | |
| HS | 7 | 1.5925 | 0.2172 | 7 | 1.8422 | 0.2372 | 1.1390 |
| PearsonHist | 5 | 1.2833 | 0.3644 | 4 | 1.5532 | 0.4540 | 1.1923 |
| GerberHist | 6 | 1.4991 | 0.2648 | 5 | 1.7188 | 0.3320 | 1.1804 |
| DGC | 4 | − 0.7569 | 1.0000 | 6 | 1.7227 | 0.3284 | 1.0454 |
| DGC FHS | 3 | − 1.0995 | 1.0000 | 3 | 1.4253 | 1.0000 | 1.0123 |
| DCC | 1 | − 2.0116 | 1.0000 | 1 | − 0.0246 | 1.0000 | 0.9771 |
| DCC FHS | 2 | − 1.2487 | 1.0000 | 2 | 0.0246 | 1.0000 | 0.9781 |
| Eliminated | - | | | | | | |
| Panel B: $\tau = 5\%$ | | | | | | | |
| HS | 7 | 2.0214 | 0.0614 | 7 | 3.2729 | 0.0072 | 2.9443 |
| PearsonHist | 6 | 1.4878 | 0.2362 | 6 | 1.6091 | 0.4874 | 2.8978 |
| GerberHist | 5 | 1.2910 | 0.3384 | 5 | 1.4917 | 0.5702 | 2.8845 |
| DGC | 4 | − 1.2983 | 1.0000 | 2 | 0.5544 | 0.9920 | 2.7260 |
| DGC FHS | 2 | − 1.6567 | 1.0000 | 3 | 0.8290 | 0.9920 | 2.7135 |
| DCC | 1 | − 1.7133 | 1.0000 | 1 | − 0.5544 | 1.0000 | 2.7004 |
| DCC FHS | 3 | − 1.3419 | 1.0000 | 4 | 1.1286 | 0.9920 | 2.7170 |
| Eliminated | – | | | | | | |
| Panel C: $\tau = 10\%$ | | | | | | | |
| HS | 5 | 1.1508 | 0.3942 | 4 | 1.2622 | 0.6414 | 4.2724 |
| GerberHist | 6 | 1.4914 | 0.2232 | 5 | 2.1128 | 0.1670 | 4.3014 |
| DGC | 3 | − 0.9106 | 1.0000 | 3 | 0.4063 | 0.9982 | 4.1307 |
| DGC FHS | 1 | − 1.4091 | 1.0000 | 2 | 0.0191 | 1.0000 | 4.1152 |
| DCC | 2 | − 1.3844 | 1.0000 | 1 | − 0.0191 | 1.0000 | 4.1146 |
| DCC FHS | 4 | − 0.7969 | 1.0000 | 6 | 2.1687 | 0.1500 | 4.1385 |
| Eliminated | PearsonHist | | | | | | |

For each model *i* in the SSM, the table reports the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (11) multiplied by $10^3$

proposed model is based—in a first stage of the estimation process—on univariate GARCH models. In an efficient way, the parameters in the recursions for the dynamic robust correlation matrix are estimated in the second stage. The distribution we use for estimating the DGC model captures asymmetries and fat tails in the returns, hence making a realistic assumption—especially in the presence of market turbulence. In an out-of-sample exercise, we tested the performance of the proposed DGC method in accurately forecasting only VaR or VaR and ES jointly for portfolios consisting of four assets. For three different diversified portfolios realistically selected by many investors, we first backtested ES for the alternative models under scrutiny. With the exception of the portfolios consisting mainly of equities for the case of the probability level 1%, we did not reject the null of accurate ES predictions from models based on the DGC method. Finally, for VaR and ES corresponding to different probability levels, we derived the superior set of models using the Model Confidence Set procedure.

**Table 12** Superior Set of Models, $\tau$-VaR forecasting, Portfolio 2

| **Model *i*** | $\text{Rank}_{T_R}$ | $\max_j T_{ij}$ | ***p*-value**$_{T_R}$ | $\text{Rank}_{T_{max}}$ | $T_{i\cdot}$ | ***p*-value**$_{T_{max}}$ | **Loss**$\times 10^3$ |
|---|---|---|---|---|---|---|---|
| Panel A: $\tau = 1\%$ | | | | | | | |
| HS | 5 | 1.4314 | 0.2310 | 4 | 1.5447 | 0.4220 | 0.8991 |
| PearsonHist | 7 | 1.5113 | 0.1926 | 7 | 1.5856 | 0.3930 | 0.9449 |
| GerberHist | 6 | 1.5041 | 0.1958 | 6 | 1.5768 | 0.4008 | 0.9223 |
| DGC | 4 | − 1.2871 | 1.0000 | 5 | 1.5472 | 0.4208 | 0.7287 |
| DGC FHS | 2 | − 1.4614 | 1.0000 | 3 | 0.9934 | 1.0000 | 0.7127 |
| DCC | 1 | − 1.6586 | 1.0000 | 1 | − 0.2573 | 1.0000 | 0.6913 |
| DCC FHS | 3 | − 1.3490 | 1.0000 | 2 | 0.2573 | 1.0000 | 0.6988 |
| Eliminated | - | | | | | | |
| Panel B: $\tau = 5\%$ | | | | | | | |
| PearsonHist | 6 | 2.0783 | 0.0510 | 6 | 2.2083 | 0.1058 | 2.2252 |
| GerberHist | 5 | 1.9637 | 0.0670 | 5 | 2.1243 | 0.1280 | 2.2225 |
| DGC | 4 | − 1.0085 | 1.0000 | 3 | 1.6544 | 0.3562 | 2.0358 |
| DGC FHS | 2 | − 1.8756 | 1.0000 | 4 | 2.0737 | 0.1424 | 2.0113 |
| DCC | 1 | − 2.3524 | 1.0000 | 1 | − 0.5351 | 1.0000 | 1.9620 |
| DCC FHS | 3 | − 1.7342 | 1.0000 | 2 | 0.5351 | 0.9800 | 1.9716 |
| Eliminated | HS | | | | | | |
| Panel C: $\tau = 10\%$ | | | | | | | |
| HS | 5 | 1.8186 | 0.1174 | 4 | 1.9949 | 0.2072 | 3.2746 |
| GerberHist | 6 | 2.1443 | 0.0534 | 5 | 2.2231 | 0.1212 | 3.2936 |
| DGC | 4 | − 1.4551 | 1.0000 | 2 | 0.7378 | 0.9458 | 3.0855 |
| DGC FHS | 3 | − 1.5227 | 1.0000 | 3 | 1.3510 | 0.9458 | 3.0951 |
| DCC | 1 | − 2.1778 | 1.0000 | 1 | − 0.7378 | 1.0000 | 3.0521 |
| DCC FHS | 2 | − 1.6679 | 1.0000 | 6 | 2.7889 | 0.0212 | 3.0792 |
| Eliminated | PearsonHist | | | | | | |

For each model *i* in the SSM, the table reports the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (11) multiplied by $10^3$

We showed that for all the portfolios and VaR/ES confidence levels we consider, the DGC model is part of the superior set of models. Because the DGC method is based on realistic assumptions about financial returns and, compared to existing methods, makes accurate predictions of market risk, it represents a valid tool for investors and financial regulators, both of whom would be concerned about the impact of losses in a turbulent market environment. Additionally, the results of our empirical application provide insight useful for investors who wish to diversify their portfolios across different asset classes. A possible limitation of the proposed methodology is that it implies the same thickness of tails for each marginal density. Another limitation is that the skewness parameters are assumed to be static. Future research may delve into these issues. Indeed, it would be interesting to extend the model by including the skewness parameters evolving according to a Generalized Autoregressive Score model. Another interesting development of the paper could be the use of the proposed method for portfolio selection, which we also leave for future research.

**Table 13** Superior Set of Models, $\tau$-VaR forecasting, Portfolio 3

| **Model *i*** | $\text{Rank}_{T_R}$ | $\max_j T_{ij}$ | ***p*-value**$_{T_R}$ | $\text{Rank}_{T_{max}}$ | $T_{i\cdot}$ | ***p*-value**$_{T_{max}}$ | **Loss**$\times 10^3$ |
|---|---|---|---|---|---|---|---|
| Panel A: $\tau = 1\%$ | | | | | | | |
| HS | 7 | 1.3003 | 0.2850 | 6 | 1.4065 | 0.5264 | 0.6545 |
| PearsonHist | 6 | 1.2969 | 0.2864 | 5 | 1.3851 | 0.5432 | 0.6843 |
| GerberHist | 5 | 1.2574 | 0.3008 | 4 | 1.3644 | 0.5564 | 0.6726 |
| DGC | 4 | − 0.8948 | 1.0000 | 7 | 1.6621 | 0.3462 | 0.5705 |
| DGC FHS | 3 | − 1.2403 | 1.0000 | 3 | 1.3557 | 1.0000 | 0.5581 |
| DCC | 1 | − 1.4987 | 1.0000 | 1 | − 0.0075 | 1.0000 | 0.5393 |
| DCC FHS | 2 | − 1.2563 | 1.0000 | 2 | 0.0075 | 1.0000 | 0.5394 |
| Eliminated | – | | | | | | |
| Panel B: $\tau = 5\%$ | | | | | | | |
| HS | 7 | 2.0066 | 0.0648 | 7 | 3.7049 | 0.0024 | 1.6212 |
| PearsonHist | 6 | 1.8402 | 0.0972 | 5 | 1.9213 | 0.2602 | 1.6054 |
| GerberHist | 5 | 1.5346 | 0.1826 | 4 | 1.7486 | 0.3608 | 1.5959 |
| DGC | 3 | − 1.5101 | 1.0000 | 2 | 1.1025 | 0.7938 | 1.4858 |
| DGC FHS | 2 | − 1.7199 | 1.0000 | 3 | 1.6392 | 0.7938 | 1.4805 |
| DCC | 1 | − 1.9628 | 1.0000 | 1 | − 1.1025 | 1.0000 | 1.4505 |
| DCC FHS | 4 | − 1.4624 | 1.0000 | 6 | 1.9387 | 0.2602 | 1.4662 |
| Eliminated | – | | | | | | |
| Panel C: $\tau = 10\%$ | | | | | | | |
| HS | 5 | 1.3767 | 0.3162 | 4 | 1.5547 | 0.4724 | 2.3411 |
| GerberHist | 6 | 1.4919 | 0.2626 | 5 | 1.6432 | 0.4142 | 2.3446 |
| DGC | 3 | − 0.8517 | 1.0000 | 2 | 0.7753 | 0.9374 | 2.2585 |
| DGC FHS | 4 | − 0.7836 | 1.0000 | 3 | 1.2702 | 0.9374 | 2.2652 |
| DCC | 1 | − 1.7186 | 1.0000 | 1 | − 0.7753 | 1.0000 | 2.2333 |
| DCC FHS | 2 | − 1.2547 | 1.0000 | 6 | 2.5088 | 0.0558 | 2.2471 |
| Eliminated | PearsonHist | | | | | | |

For each model *i* in the SSM, the table reports the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (11) multiplied by $10^3$

**Table 14** Superior Set of Models, joint $\tau$-VaR and $\tau$-ES forecasting, Portfolio 1

| **Model *i*** | $\text{Rank}_{T_R}$ | $\max_j T_{ij}$ | ***p*-value**$_{T_R}$ | $\text{Rank}_{T_{max}}$ | $T_{i\cdot}$ | ***p*-value**$_{T_{max}}$ | **Loss**$\times 10^3$ |
|---|---|---|---|---|---|---|---|
| Panel A: $\tau = 1\%$ | | | | | | | |
| DGC | 1 | − 0.1343 | 1.0000 | 1 | − 0.1343 | 1.0000 | 21.1566 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |
| Panel B: $\tau = 5\%$ | | | | | | | |
| DGC | 1 | -5.4361 | 1 | 1 | -5.4361 | 1.0000 | 19.2650 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |
| Panel C: $\tau = 10\%$ | | | | | | | |
| DGC | 1 | -5.5584 | 1 | 1 | -5.5584 | 1.0000 | 18.0268 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |

For each model *i* in the SSM, the table reports the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (12) multiplied by $10^3$

**Table 15** Superior Set of Models, joint $\tau$-VaR and $\tau$-ES forecasting, Portfolio 2

| **Panel A:** $\tau = 1\%$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Model *i*** | $\text{Rank}_{T_R}$ | $\max_j T_{ij}$ | ***p*-value**$_{T_R}$ | $\text{Rank}_{T_{max}}$ | $T_{i\cdot}$ | ***p*-value**$_{T_{max}}$ | **Loss**$\times 10^3$ |
| DGC | 1 | − 10.9627 | 1.0000 | 1 | − 10.9627 | 1.0000 | 15.3850 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |
| Panel B: $\tau = 5\%$ | | | | | | | |
| DGC | 1 | − 10.9627 | 1.0000 | 1 | − 10.9627 | 1.0000 | 12.3850 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |
| Panel C: $\tau = 10\%$ | | | | | | | |
| DGC | 1 | − 5.0087 | 1.0000 | 1 | − 2.3449 | 1.0000 | 14.6779 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |

For each model *i* in the SSM, the table reports the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (12) multiplied by $10^3$

**Table 16** Superior Set of Models, joint $\tau$-VaR and $\tau$-ES forecasting, Portfolio 3

| **Model *i*** | $\text{Rank}_{T_R}$ | $\max_j T_{ij}$ | ***p*-value**$_{T_R}$ | $\text{Rank}_{T_{max}}$ | $T_{i\cdot}$ | ***p*-value**$_{T_{max}}$ | **Loss**$\times 10^3$ |
|---|---|---|---|---|---|---|---|
| Panel A: $\tau = 1\%$ | | | | | | | |
| DGC | 1 | − 15.2774 | 1 | 1 | − 15.2774 | 1 | 12.3018 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |
| Panel B: $\tau = 5\%$ | | | | | | | |
| DGC | 1 | − 3.5558 | 1.0000 | 1 | − 3.5558 | 1.0000 | 10.8580 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |
| Panel C: $\tau = 10\%$ | | | | | | | |
| DGC | 1 | − 8.2694 | 1.0000 | 1 | − 3.3863 | 1.0000 | 10.5988 |
| Eliminated | HS, | PearsonHist, | GerberHist, | | | | |
| | DGC FHS, | DCC, | DCC FHS | | | | |

For each model *i* in the SSM, the table reports the statistics $\max_j T_{ij}$ and $T_{i\cdot}$, the *p* values associated with the test statistics $T_R$ and $T_{max}$, the ranking of the models in the SSM based on $T_R$ and $T_{max}$, and the average of the loss function (12) multiplied by $10^3$

## Appendices

### Appendix A: The standardized skew-student distribution

The standardized skew-Student density (Bauwens and Laurent 2005) is given by

$$f(\boldsymbol{z}|\boldsymbol{\xi},\nu) = \left(\frac{2}{\sqrt{\pi(\nu-2)}}\right)^k \left(\prod_{i=1}^k \frac{\xi_i s_i}{1+\xi_i^2}\right) \frac{\Gamma((\nu+k)/2)}{\Gamma(\nu/2)} \left(1+\frac{\boldsymbol{z}^{*'}\boldsymbol{z}^*}{\nu-2}\right), \tag{13}$$

where the $i$-th component of the vector $\boldsymbol{z}^*$ is $z_i^* = (s_i z_i + m_i)\xi_i^{I_i}$, with

$$
\begin{aligned}
m_i &= \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \left(\xi_i - \xi_i^{-1}\right) \\
s_i &= \left(\xi_i^2 + \xi_i^{-2} - 1\right) - m_i^2,
\end{aligned}
$$

and $I_i = 2 \times I(z_i < -m_i/s_i) - 1$.

## Appendix B: The MCS procedure

The relative performance between models $i$ and $j$ is obtained via the differential

$$
d_{ij,t} = l_{i,t} - l_{j,t} \qquad \forall i,j \in M_0 \quad t = 1, \ldots, n, \tag{14}
$$

and the simple average loss of model $i$ relative to the other models $j \in M$ at time $t$ as

$$
d_{i\cdot,t} = (m-1)^{-1} \sum_{j \in M \setminus i} d_{ij,t}. \tag{15}
$$

For the elimination of inferior elements within the set $M_0$, two alternative sets of hypotheses are available to test the EPA:

$$
\begin{cases} H_0 : \mathbb{E}(d_{ij}) = 0 \quad \forall i,j = 1, \ldots, m, \qquad \text{against} \\ H_1 : \mathbb{E}(d_{ij}) \neq 0 \end{cases} \tag{16}
$$

or

$$
\begin{cases} H_0 : \mathbb{E}(d_{i\cdot}) = 0 \quad \forall i = 1, \ldots, m, \qquad \text{against} \\ H_1 : \mathbb{E}(d_{i\cdot}) \neq 0 \end{cases} \tag{17}
$$

Two statistics are then constructed to test the above hypotheses:

$$
T_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}}; \qquad T_{i\cdot} = \frac{\bar{d}_{i\cdot}}{\sqrt{\widehat{\text{var}}(\bar{d}_{i\cdot})}}, \tag{18}
$$

where $\bar{d}_{ij}$ constitutes the relative average loss between models $i$ and $j$ and $\bar{d}_{i\cdot}$ represents the average loss of model $i$ relative to the average loss across the models belonging to the set $M$:

$$
\bar{d}_{ij} = n^{-1} \sum_{t=1}^{n} d_{ij,t}; \qquad \bar{d}_{i\cdot} = (m-1)^{-1} \sum_{j \in M \setminus i} \bar{d}_{ij}. \tag{19}
$$

The standard errors in the denominators of (18) are constructed by block bootstrap.

The two hypotheses from (16) and (17) are mapped into two test statistics:

$$
T_R = \max_{i,j \in M} T_{ij}; \qquad T_{max} = \max_{i \in M} T_{i\cdot}. \tag{20}
$$

Because their distributions under the null are not known, they are also simulated by the bootstrap method. When the null hypothesis is rejected, the following elimination rules establish which of the models can be discarded:

$$e_R = \arg\max_i \left\{ \sup_{j\in M} T_{ij} \right\}; \qquad e_{max} = \arg\max_i \{T_{i\cdot}\}. \tag{21}$$

Bernardi and Catania (2016) summarized the algorithm for the procedure as follows:

1. Set $M = M_0$.
2. Compute the test statistics under the null EPA hypothesis. If it is not rejected, set $M^*_{1-\alpha} = M$ and terminate the algorithm. If it is rejected, use the elimination rule to determine the worst model.
3. Discard the model and repeat step 2.

The elimination rule defines a sequence of sets $M = M_0 \supset M_1 \cdots \supset M_m$, where $M_i = (e_{M_i}, \ldots, e_{M_m})$, each of which has a $p$ value associated with an EPA test. Let $P_{H_0,M_i}$ be the $p$ value associated with the null hypothesis $H_{0,M_i}$. The MCS $p$ value for model $e_{M_j} \supset M$ is defined as $\hat{p}_{e_{M_j}} = \max_{i\leq j} P_{H_0,M_i}$.

## Appendix C: Backtesting ES

In this appendix we provide the details of the procedure we implement to backtest ES.

Du and Escanciano (2017) proposed a procedure to backtest $\tau$-ES based on the so-called cumulative violations (CV) process. The time-$t$ value of this process is given by

$$H_t(\tau) = \frac{1}{\tau}[\tau - u_t]I(u_t \leq \tau), \tag{22}$$

where $u_t = G(r_{p,t}|\mathcal{F}_{t-1})$ is the Probability Integral (PIT) transform (Rosenblatt 1952) and $G(\cdot|\mathcal{F}_{t-1})$ is the conditional cumulative distribution function of the portfolio return $r_{p,t}$.

To avoid estimating the portfolio distribution function, Khalaf et al. (2021) consider $K$ equally spaced VaR levels with the larger one coinciding with $\tau$, i.e. $\tau_j = (K - j + 1)\frac{\tau}{K}$ for $j = 1, \ldots, K$. Their ES backtesting procedures are based on the sum of VaR violations

$$N_t^K(\tau) = \sum_{j=1}^{K} I\left(r_{p,t} \leq -\mathrm{VaR}_{t|t-1}\left(\mathrm{j}\frac{\tau}{\mathrm{K}}\right)\right), \tag{23}$$

since they show that, when returns are absolutely continuous, $\frac{N_t^K(\tau)}{K} \xrightarrow{D} H_t(\tau)$. Khalaf et al. (2021) hence consider the null

$$H_0 : \begin{cases} N_t^K(\tau) = j \text{ with probability } \theta_j \text{ for } j = 0, \ldots, K, \\ N_t^K(\tau) \perp\!\!\!\perp N_{t-h}^K(\tau), \forall h \neq 0, \end{cases} \tag{24}$$

where $\perp\!\!\!\perp$ denotes independence, and

$$\theta_0 = 1 - \tau, \qquad \theta_j = \frac{\tau}{K}, \quad j = 1, \ldots, K. \tag{25}$$

Note that $\mathbb{E}\left[N_t^K(\tau)\right] = \frac{K+1}{2}\tau$ under (24)–(25). As a consequence, in this paper we use a conditional backtest based on the idea that under the null, $\left\{N_t^K(\tau) - \frac{K+1}{2}\tau\right\}_t$ is a martingale difference sequence.

To this end, we use, for a sample of length $n$ for $N_t^K(\tau)$, the Box-Pierce test statistic

$$C_m(\tau) = n\sum_{i=1}^{m}\rho_i^2(\tau), \tag{26}$$

where

$$\gamma_i(\tau) = \frac{1}{n-i}\sum_{t=i+1}^{n}\left(N_t^K(\tau) - \frac{K+1}{2}\tau\right)\left(N_{t-i}^K(\tau) - \frac{K+1}{2}\tau\right),$$

and $\rho_i(\tau) = \frac{\gamma_i(\tau)}{\gamma_0(\tau)}$. The $p$ values associated with the test statistic (26) can be obtained using the Monte Carlo test technique (Dufour 2006) since, under the null (24)–(25), it is easy to simulate (23) and hence the test statistic (26), see Khalaf et al. (2021) for further details.

**Abbreviations**

| | |
|---|---|
| DGC | Dynamic Gerber correlation |
| ARMA | Autoregressive moving average |
| CV | Cumulative violations |
| DCC | Dynamic conditional correlation |
| DM | Diebold-Mariano |
| ES | Expected shortfall |
| FHS | Filtered historical simulation |
| GARCH | Generalized Autoregressive Conditional Heteroskedasticity |
| HS | Historical simulation |
| MCS | Model confidence set |
| QML | Quasi maximum likelihood |
| SSM | Superior set of models |
| VaR | Value at Risk |

**Acknowledgements**
We would like to thank the editor and four anonymous reviewers for their insightful suggestions. We thank Geoffrey Groesbeck for carefully reviewing the first draft of the manuscript. The views expressed here are those of the authors alone and do not necessarily reflect those of their employers. This article is intended to stimulate further research and is not a recommendation for adopting the proposed method.

**Author Contributions**
AL: methodology, formal analysis, software, validation, formal analysis, investigation, data curation, writing.original draft, writing.review and editing. AS: methodology, formal analysis, investigation, data curation, visualization, writing.original draft, writing.review and editing. PT: conceptualization, methodology, writing.original draft, writing.review and editing. All authors read and approved the final manuscript.

**Funding:**
None

**Availability of data and material:**
The dataset analysed during the current study is available from Bloomberg but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Bloomberg.

## Declarations

**Competing interests**
The author declare that they have no competing interests.

Received: 7 November 2022 Accepted: 15 February 2024
Published online: 02 August 2024

**References**

Acerbi C, Tasche D (2002) On the coherence of expected shortfall. J Bank Finance 26(7):1487–1503
Algieri B, Leccadito A, Toscano P (2021) A time-varying Gerber statistic: application of a novel correlation metric to commodity price co-movements. Forecasting 3(2):339–354
Artzner P, Delbaen F, Eber JM, Heath D (1999) Coherent measures of risk. Math Financ 9(3):203–228
Athari SA, Hung NT (2022) Time-frequency return co-movement among asset classes around the COVID-19 outbreak: portfolio implications. J Econ Finance 46(4):736–756. https://doi.org/10.1007/s12197-022-09594-8
Athari SA, Kirikkaleli D, Adebayo TS (2022) World pandemic uncertainty and German stock market: evidence from Markov regime-switching and Fourier based approaches. Quality Quantity 57(2):1923–1936. https://doi.org/10.1007/s11135-022-01435-4
Barone-Adesi G, Giannopoulos K, Vosper L (2002) Backtesting derivative portfolios with filtered historical simulation (FHS). Eur Financ Manag 8(1):31–58
Bauwens L, Laurent S (2005) A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. J Bus Econ Stat 23(3):346–354. https://doi.org/10.1198/073500104000000523
Bekiros S, Loukeris N, Eleftheriadis I, Avdoulas C (2019) Tail-related risk measurement and forecasting in equity markets. Comput Econ 53(2):783–816
Bernardi M, Catania L (2016) Comparison of Value-at-Risk models using the MCS approach. Comput Stat 31(2):579–608
Boudt K, Daníelsson J, Laurent S (2013) Robust forecasting of dynamic conditional correlation GARCH models. Int J Forecast 29(2):244–257. https://doi.org/10.1016/j.ijforecast.2012.06.003
Brooks C, Persand G (2003) Volatility forecasting for risk management. J Forecast 22(1):1–22
Christoffersen PF (1998) Evaluating interval forecasts. Int Econ Rev 39(4):841–862
Christoffersen P (2012) Elements of financial risk management, 2nd edn. Elsevier Science
Chu J, Chan S, Nadarajah S, Osterrieder J (2017) GARCH modelling of cryptocurrencies. J Risk Financ Manag. https://doi.org/10.3390/jrfm10040017
Cortese FP (2019) Tail dependence in financial markets: a dynamic copula approach. Risks. https://doi.org/10.3390/risks7040116
Dhaene J, Vanduffel S, Goovaerts MJ, Kaas R, Tang Q, Vyncke D (2006) Risk measures and comonotonicity: a review. Stoch Model 22(4):573–606
Diebold FX, Mariano RS (2002) Comparing predictive accuracy. J Bus Econ Stat 20(1):134–144
Douc R, Doukhan P, Moulines E (2013) Ergodicity of observation-driven time series models and consistency of the maximum likelihood estimator. Stochastic Process Appl 123(7):2620–2647
Du Z, Escanciano JC (2017) Backtesting expected shortfall: accounting for tail risk. Manage Sci 63(4):940–958
Dufour JM (2006) Monte Carlo tests with nuisance parameters: a general approach to finite-sample inference and nonstandard asymptotics. J Econom 133(2):443–477
Düllmann K, Küll J, Kunisch M (2010) Estimating asset correlations from stock prices or default rates–which method is superior? J Econ Dyn Control 34(11):2341–2357. https://doi.org/10.1016/j.jedc.2010.06.003
Düllmann K, Scheicher M, Schmieder C (2007) Asset correlations and credit portfolio risk: an empirical analysis. Discussion paper no. 2007,13, Bundesbank. https://doi.org/10.2139/ssrn.2793999
Echaust K, Just M (2020) Value at risk estimation using the GARCH-EVT approach with optimal tail selection. Mathematics. https://doi.org/10.3390/math8010114
Engle R (2002) Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J Bus Econ Stat 20(3):339–350
Engle RF, Manganelli S (2004) CAViaR: conditional autoregressive Value-at-Risk by regression quantiles. J Bus Econ Stat 22(4):367–381
Fan Q, Tan KS, Zhang J (2023) Empirical tail risk management with model-based annealing random search. Insur: Math Econ 110:106–124. https://doi.org/10.1016/j.insmatheco.2023.02.005
Fissler T, Ziegel JF, Gneiting T (2015) Expected shortfall is jointly elicitable with value at risk - implications for backtesting. Risk Management. arXiv:1507.00244
Fiszeder P, Fałdziński M, Molnár P (2023) Modeling and forecasting dynamic conditional correlations with opening, high, low, and closing prices. J Empir Financ 70:308–321. https://doi.org/10.1016/j.jempfin.2022.12.007
Gerber S, Markowitz H, Ernst PA, Miao Y, Javid B, Sargen P (2022) The Gerber statistic: a robust co-movement measure for portfolio optimization. J Portf Manag 48(8):87–102
Gong Y, Huser R (2022) Asymmetric tail dependence modeling, with application to cryptocurrency market data. Ann Appl Stat 16(3):1822–1847. https://doi.org/10.1214/21-AOAS1568
González-Rivera G, Lee TH, Mishra S (2004) Forecasting volatility: a reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. Int J Forecast 20(4):629–645
Hansen PR, Lunde A, Nason JM (2011) The model confidence set. Econometrica 79(2):453–497
Høyland K, Kaut M, Wallace SW (2003) A heuristic for moment-matching scenario generation. Comput Optim Appl 24(2):169–185
James R, Leung H, Leung JWY, Prokhorov A (2023) Forecasting tail risk measures for financial time series: an extreme value approach with covariates. J Empir Financ 71:29–50. https://doi.org/10.1016/j.jempfin.2023.01.002
Jarjour R, Chan KS (2020) Dynamic conditional angular correlation. J Econom 216(1):137–150. https://doi.org/10.1016/j.jeconom.2020.01.010, annals Issue in honor of George Tiao: Statistical Learning for Dependent Data
Jiang Y, Yu M, Hashmi SM (2017) The financial crisis and co-movement of global stock markets–a case of six major economies. Sustainability. https://doi.org/10.3390/su9020260
Khalaf L, Leccadito A, Urga G (2021) Multilevel and tail risk management. J Financ Econom. https://doi.org/10.1093/jjfinec/nbaa044
Kondoz M, Bora I, Kirikkaleli D, Athari SA (2019) Testing the volatility spillover between crude oil price and the U.S. stock market returns. Management Science Letters p 1221–1230, https://doi.org/10.5267/j.msl.2019.4.019

Kou G, Xu Y, Peng Y, Shen F, Chen Y, Chang K, Kou S (2021) Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. Decis Support Syst 140(113):429. https://doi.org/10.1016/j.dss.2020.113429

Long HV, Jebreen HB, Dassios I, Baleanu D (2020) On the statistical GARCH model for managing the risk by employing a fat-tailed distribution in finance. Symmetry. https://doi.org/10.3390/sym12101698

Lopez JA, Walter C (2000) Evaluating covariance matrix forecasts in a Value-at-Risk framework. Working Paper Series 2000-21, Federal Reserve Bank of San Francisco

Nieto MR, Ruiz E (2016) Frontiers in VaR forecasting and backtesting. Int J Forecast 32(2):475–501

Oh DH, Patton AJ (2018) Time-varying systemic risk: evidence from a dynamic copula model of CDS spreads. J Bus Econ Stat 36(2):181–195. https://doi.org/10.1080/07350015.2016.1177535

Patton AJ (2006) Modelling asymmetric exchange rate dependence. Int Econ Rev 47(2):527–556. https://doi.org/10.1111/j.1468-2354.2006.00387.x

Patton AJ, Ziegel JF, Chen R (2019) Dynamic semiparametric models for expected shortfall (and Value-at-Risk). J Econom 211(2):388–413

Pelletier D (2006) Regime switching for dynamic correlations. J Econom 131(1):445–473. https://doi.org/10.1016/j.jeconom.2005.01.013

Rosenblatt M (1952) Remarks on a multivariate transformation. Ann Math Stat 23(3):470–472

Skintzi VD, Xanthopoulos-Sisinis S (2007) Evaluation of correlation forecasting models for risk management. J Forecast 26(7):497–526

Taylor JW (2019) Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution. J Bus Econ Stat 37(1):121–133

Vancsura L, Tatay T, Bareith T (2023) Evaluating the effectiveness of modern forecasting models in predicting commodity futures prices in volatile economic times. Risks. https://doi.org/10.3390/risks11020027

Xu Y, Kou G, Peng Y, Ding K, Ergu D, Alotaibi FS (2024) Profit- and risk-driven credit scoring under parameter uncertainty: a multiobjective approach. Omega 125(103):004. https://doi.org/10.1016/j.omega.2023.103004

## Publisher's Note