A multivariate semi-parametric portfolio risk optimization and forecasting framework

Storti, Giuseppe and Wang, Chao

University of Salerno, University of Sydney

August 2022

A multivariate semi-parametric portfolio risk optimization and forecasting framework

Giuseppe Storti[1], Chao Wang[2]

[1] Department of Economics and Statistics, University of Salerno

[2] Discipline of Business Analytics, The University of Sydney

Abstract

A new multivariate semi-parametric risk forecasting framework is proposed, to enable the portfolio Value-at-Risk (VaR) and Expected Shortfall (ES) optimization and forecasting. The proposed framework accounts for the dependence structure among asset returns, without assuming their distribution. A simulation study is conducted to evaluate the finite sample properties of the employed estimator for the proposed model. An empirically motivated portfolio optimization method, that can be utilized to optimize the portfolio VaR and ES, is developed. A forecasting study on 2.5% level evaluates the performance of the model in risk forecasting and portfolio optimization, based on the components of the Dow Jones index for the out-of-sample period from December 2016 to September 2021. Comparing to the standard models in the literature, the empirical results are favourable for the proposed model class, in particular the effectiveness of the proposed framework in portfolio risk optimization is demonstrated.

Keywords: semi-parametric; Value-at-Risk; Expected Shortfall; multivariate; portfolio optimization.

1 Introduction

Value-at-Risk (VaR) and Expected Shortfall (ES) play a central role in the risk management systems of banks and other financial institutions. For more than two decades, VaR has been the official risk measure adopted worldwide by financial intermediaries operating in the global financial system. Nevertheless, VaR has some important theoretical limits. First, VaR cannot measure the expected loss for extreme (violating) returns. In addition, it can be shown that VaR is not always a *coherent* risk measure, due to failure to match the *subadditivity* property. For these reasons, the Basle Committee on Banking Supervision proposed in May 2012 to replace VaR with the ES (Artzner 1997; Artzner et al. 1999). ES is defined as the expectation of the return conditional on it having exceeded the VaR. Differently from VaR, ES is a coherent measure and "measures the riskiness of a position by considering both the size and the likelihood of losses above a certain confidence level" (Basel Committee on Banking Supervision, 2013). Thus, in recent years ES has been increasingly employed for tail risk measurement.

Comparing to VaR, there is much less existing work on modeling ES, which is partly due to the non-elicitability of ES alone. However, the recent work in Fissler and Ziegel (2016) has shown that the pair (VaR,ES) is jointly elicitable. They develop a family of joint loss, or *scoring*, functions that are strictly consistent for the true VaR and ES, i.e., they are uniquely minimized by the true VaR and ES series. This result has important implications for the estimation of conditional VaR and ES as well as for ranking risk forecasts from alternative competing models. Taylor (2019) proposes a joint VaR and ES modelling approach (named as ES-CAViaR in this paper) based on the minimization of the negative of an Asymmetric Laplace (AL) log-likelihood function that can be derived as a special case of the Fissler and Ziegel (2016) class of loss functions, under specific choices of the functions involved. Furthermore, Patton et al. (2019) propose new dynamic models for VaR and ES, through adapting the generalized autoregressive score (GAS) framework (Creal et al., 2013) and utilizing a 0-degree homogeneous loss function falling in the Fissler and Ziegel (2016) class (FZ0).

The works mentioned above focus on univariate time series and do not take into account the correlation among assets in financial markets. Several quantile-based methods,

see for example the works in Baur (2013), Bernardi et al. (2015), White et al. (2015), have been developed to estimate VaR in a multivariate setting and model the tail interdependence between assets, while the ES component is not specified in these frameworks.

In this paper we contribute to the literature on portfolio risk forecasting by proposing a class of semi-parametric marginalized multivariate GARCH (MGARCH) models that generate portfolio VaR and ES forecasts jointly, taking multivariate information on the returns of portfolio constituents as input. Our main reference model is given by a marginalized version of the dynamic conditional correlations (DCC) model (Engle, 2002) whose parameters are estimated by minimization of a negative AL log-likelihood as in Taylor (2019). In the remainder, for short we will refer to this model as the DCC-AL model. The proposed framework can be easily generalized to other parameterizations from multivariate GARCH literature such as the corrected DCC (cDCC) of Aielli (2013) or dynamic covariance models such as the BEKK model of Engle and Kroner (1995). Our preference for the DCC parameterization is motivated by the flexibility of this specification and its widespread diffusion among practitioners [1]. At the same time, the AL scoring function can be replaced in a straightforward manner by any other strictly consistent loss for the pair (VaR, ES) such as the FZ0 used by Patton et al. (2019).

In a risk forecasting perspective, the DCC-AL model extends the univariate approach of Taylor (2019) to a multivariate framework, with the objective of obtaining joint estimates of both VaR and ES for portfolios of financial assets, accounting for their cross-sectional correlation structure. Due to its semi-parametric nature, the proposed framework does not require the formulation of any assumption on the conditional distribution of portfolio returns, since the required VaR and ES factors are estimated minimizing a strictly consistent loss function. However, compared to the univariate semi-parametric approaches in risk forecasting, due to its underlying multivariate GARCH structure, the DCC-AL model can be used for a wider range of applications such as portfolio optimization and hedging.

[1] We also derived a marginalized version of the cDCC model, cDCC-AL for short. However, consistently with previous findings, the cDCC-AL and DCC-AL models turned out to give very similar estimates, motivating our decision to keep the mathematically simpler DCC-AL as the main reference model. Results for the cDCC-AL model are available upon request.

The DCC-AL model develops a multivariate portfolio risk optimization framework based on the minimization of a risk-targeted strictly consistent loss function. It has significant differences to the existing semi-parametric multivariate risk forecasting frameworks. First, Merlo et al. (2021) generalize the AL distribution to a multivariate setting for portfolio risk optimization, while their approach is limited to low portfolio dimensions, i.e., three market indices are studied in their paper. Second, the DCC model could be estimated semi-parametrically by using the Gaussian Quasi Maximum Likelihood (DCC-QML) method. Then the portfolio VaR and ES can be produced by the filtered historical simulation (HS) approach by calculating the sample quantiles and tail averages of the standardized returns (Francq and Zakoïan, 2015). The DCC-QML with HS is a two-step process, while the DCC-AL can jointly estimate and forecast portfolio VaR and ES in one step, which is more convenient from a practical perspective. Further, the DCC-QML approach optimizes a Gaussian based quasi-likelihood, while the DCC-AL optimizes a risk-targeted strictly consistent loss function for VaR and ES, e.g., the AL based joint loss function.

In an empirical application to a panel of 28 assets included in the Dow Jones index, we find that, when forecasting risk for an equally weighted portfolio, DCC-AL models are competitive with state-of-the-art univariate semi-parametric approaches and perform better than parametric DCC models. When focusing on the out-of-sample hedging performance of the model, our results show that the DCC-AL model clearly outperforms a benchmark equally weighted allocation strategy. We also find evidence that in a high volatility period, at the outbreak of the COVID-19 crisis in 2020, compared to the DCC-QML, the DCC-AL model is characterized by a better out-of-sample minimum variance hedging performance. In addition, our empirical results show that the DCC-AL model can be effectively used as a modelling platform for the generation of minimum risk (VaR, ES) portfolios without requiring any parametric assumption on the conditional distribution of returns, including its ellipticity.

The paper is structured as follows. Section 2 reviews the semi-parametric univariate risk forecasting approaches and presents the proposed framework and its technical details. A two stage model estimation procedure is shown in Section 3. Section 4 presents the simulation study. An empirically motivated portfolio risk optimization procedure is

developed in Section 5. The empirical results of the proposed model in risk forecasting and portfolio optimization are discussed in Section 6. Section 7 concludes the paper.

2 Statistical framework

2.1 Description of the environment

First, let $\mathbf{r}_t = (r_{t,1}, \ldots, r_{t,n})'$ be a vector of returns on the n portfolio assets at time t and generated by the process:

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

where $\mathbf{z}_t \overset{iid}{\sim} D_1(\mathbf{0}, I_n)$, D_1 is a multivariate distribution with zero mean and identity covariance matrix, $\boldsymbol{\mu}_t = E(\mathbf{r}_t|\mathcal{I}_{t-1})$ is the conditional mean vector of returns with $\mathcal{I}_{t-1}$ as the information available at time $t-1$, and $H_t^{1/2}$ is a positive definite matrix such that $H_t^{1/2}(H_t^{1/2})' = H_t$, with $H_t = var(\mathbf{r}_t|\mathcal{I}_{t-1})$ being the conditional covariance matrix of returns.

Pre-multiplying both members of Equation (1) by the (transposed) vector of portfolio weights $\mathbf{w} = (w_1, \ldots, w_n)'$, we obtain the following (univariate) portfolio returns:

$$r_t(\mathbf{w}) = \mathbf{w}'\mathbf{r}_t = \mathbf{w}'\boldsymbol{\mu}_t + \mathbf{w}'H_t^{1/2}\mathbf{z}_t.$$

It is easy to infer that these can be equivalently represented as:

$$r_t(\mathbf{w}) = \mathbf{w}'\boldsymbol{\mu}_t + e_t\sqrt{\mathbf{w}'H_t\mathbf{w}}, \tag{2}$$

where e_t is a scalar continuous error term such that $e_t \overset{iid}{\sim} (0, 1)$. In order to simplify the presentation, we here assume that $\mathbf{z}_t$ has a spherical distribution, thus the distribution of e_t will be the same as the marginal distribution of the components of $\mathbf{z}_t$. Assuming sphericity of the distribution of $\mathbf{z}_t$ is equivalent to assume ellipticity of the conditional distribution of $r_t(\mathbf{w})$ (see e.g. Francq and Zakoïan, 2020). However, as it will be discussed later, this condition is not strictly required in our setting but is here invoked with the only purpose of facilitating the illustration of the proposed approach. Further, we assume that the Cumulative Distribution Function (CDF) of e_t is strictly increasing on the real line $\Re$.

Assuming that asset returns follow the process in (2) and given the set of weights $\mathbf{w}$, the α-level portfolio conditional VaR and ES are given by:

$$Q_{t,\alpha} = \mathbf{w}^{'}\boldsymbol{\mu}_t + q_\alpha\sqrt{\mathbf{w}^{'}H_t\mathbf{w}}, \qquad \mathrm{ES}_{t,\alpha} = \mathbf{w}^{'}\boldsymbol{\mu}_t + c_\alpha\sqrt{\mathbf{w}^{'}H_t\mathbf{w}}, \tag{3}$$

where $\mathbf{w}^{'}\boldsymbol{\mu}_t$ and $\mathbf{w}^{'}H_t\mathbf{w}$ are the conditional mean and variance of portfolio returns $r_t(\mathbf{w})$, respectively. $0 < \alpha << 1$ denotes the target level for the estimation of VaR and ES. In the empirical application we focus on $\alpha = 2.5\%$. $q_\alpha = F^{-1}_{e,t-1}(\alpha)$ and $c_\alpha = E(e_t|\mathcal{I}_{t-1}, e_t \leq q_\alpha)$, where $F_{e,t-1}(.)$ represents the cumulative distribution function (CDF) of e_t conditional on $\mathcal{I}_{t-1}$. Therefore, q_α and c_α are the portfolio VaR and ES factors calculated based on the definitions of VaR and ES.

In order to simplify notation, in the remainder, unless differently specified, the following notational conventions are adopted: $Q_{t,\alpha} \equiv Q_t$, $\mathrm{ES}_{t,\alpha} \equiv \mathrm{ES}_t$, $q_\alpha \equiv q$ and $c_\alpha \equiv c$, $r_t \equiv r_t(\mathbf{w})$.

2.2 Univariate approaches to portfolio risk forecasting

The literature on semi-parametric forecasting of portfolio risk has so far mostly been limited to univariate approaches. In this section, we present a selective review of the most relevant contributions to the research on this topic.

Focusing on VaR forecasting, Koenker and Machado (1999) note that the usual quantile regression estimator is equivalent to a maximum likelihood estimator based on the AL density with a mode at the quantile. The parameters in the model for Q_t can then be estimated maximizing a quasi-likelihood based on:

$$p(r_t|\mathcal{I}_{t-1}) = \frac{\alpha(1-\alpha)}{\sigma}\exp\left(-\frac{(r_t - Q_t)(\alpha - I(r_t \leq Q_t))}{\sigma}\right),$$

for $t = 1, \ldots, N$ and σ is a scale parameter.

Taylor (2019) extends this result to incorporate the associated ES quantity into the likelihood expression, noting a link between ES_t and a dynamic σ_t, resulting in the conditional density function:

$$p(r_t|\mathcal{I}_{t-1}) = \frac{(\alpha-1)}{\mathrm{ES}_t}\exp\left(\frac{(r_t - Q_t)(\alpha - I(r_t \leq Q_t))}{\alpha\mathrm{ES}_t}\right). \tag{4}$$

This allows a likelihood function to be built and maximised, given model expressions for (Q_t, ES_t). Taylor (2019) notes that the negative logarithm of the resulting likelihood function is strictly consistent for (Q_t, ES_t) considered jointly, e.g., it fits into the class of jointly consistent scoring functions for VaR and ES developed by Fissler and Ziegel (2016). Along the same lines, Patton et al. (2019) investigate a class of semi-parametric models, including some observation driven models, whose parameters can be estimated minimizing a 0-degree homogeneous loss function (FZ0) still included in the same class. Gerlach and Wang (2020) extend the framework in Taylor (2019) by incorporating realized measures as exogenous variables, showing improved VaR and ES forecast accuracy.

As mentioned above, all these papers focus on univariate semi-parametric modelling approaches that, when applied to portfolio returns, do not explicitly assess the impact of cross-sectional correlations among assets. Although this issue has been extensively analyzed in the literature on parametric MGARCH models (Bauwens et al., 2006), to the extent of our knowledge, no attempts have been made to address it in a multivariate semi-parametric risk-targeted framework (except the framework given by Merlo et al. (2021) as discussed in Section 1). In order to fill this gap, in the next section we propose a novel modelling strategy based on the use of what we call a semi-parametric marginalized MGARCH model.

2.3 Semi-parametric marginalized MGARCH models

In this section, we present the proposed semi-parametric marginalized MGARCH framework. In order to explicitly highlight the link between portfolio risk and H_t, Equations (3) can be reparameterized as:

$$Q_t = \mathbf{w}^{'}\boldsymbol{\mu}_t + \sqrt{\mathbf{w}' H_t^{(q)} \mathbf{w}}, \qquad \text{ES}_t = \mathbf{w}^{'}\boldsymbol{\mu}_t + \sqrt{\mathbf{w}' H_t^{(c)} \mathbf{w}}, \tag{5}$$

where we define $H_t^{(q)} = H_t q^2$ and $H_t^{(c)} = H_t c^2$. In the remainder, we will conventionally use the superscript (q) to denote variables or quantities related to "q"uantile (VaR), while (c) will be used in variables or quantities related to "c"onditional tail expectation (ES). Inspired by Taylor (2019), the ratio between c^2 and q^2 is then modelled as

$$\frac{c^2}{q^2} = \frac{H_t^{(c)}}{H_t^{(q)}} = 1 + \exp(\gamma_0). \tag{6}$$

By construction, this formulation guarantees that the VaR and ES do not cross. It is evident that the VaR and ES dynamics are driven by those of H_t in expression (2). These can be modelled using a wide range of specifications from the MGARCH literature. Without any loss of generality, due to its flexibility and widespread diffusion among practitioners, our proposed modelling approach builds on the DCC model of Engle (2002) and shares the same volatility and correlation dynamics. Under the DCC specification, the conditional variance and covariance matrix H_t is decomposed as:

$$H_t = D_t P_t D_t, \tag{7}$$

where D_t is a $(n \times n)$ diagonal matrix such that its i-th diagonal element is $D_{t,ii} = h_{t,ii}$, with $h^2_{t,ii} = var(r_{t,i}|\mathcal{I}_{t-1})$. Since our approach is developed in a fully semi-parametric framework, the specification for $h^2_{t,ii}$ is indirectly recovered assuming an ES-CAViaR type model for the individual asset VaRs and ESs. Among the several diverse specifications that have been proposed in the literature (Engle and Manganelli, 2004; Taylor, 2019), for presentation purposes, without any loss of generality, we focus on the ES-CAViaR model with the Indirect GARCH (IG) specification for VaR and multiplicative VaR to ES relationship:

$$\begin{aligned} Q_{t,i} &= -\sqrt{\omega_i^{(q)} + \alpha_i^{(q)} r^2_{t-1,i} + \beta_i Q^2_{t-1,i}}, \\ ES^2_{t,i} &= (1 + \exp(\gamma_{0,i}))Q^2_{t,i}, \qquad i = 1, \ldots, n, \end{aligned} \tag{8}$$

where, by a variance targeting argument, the intercept $\omega_i^{(q)}$ is parameterized as:

$$\omega_i^{(q)} = (q_i^2(1 - \beta_i) - \alpha_i^{(q)})\, var(r_{t,i}) \tag{9}$$

so to make explicit the dependence on q_i, that is the VaR factor for the i-th asset. This is easily derived considering that the VaR model implies

$$Q^2_{t,i} = \omega_i^{(q)} + \alpha_i^{(q)} r^2_{t-1,i} + \beta_i Q^2_{t-1,i}. \tag{10}$$

Noting that, with $\mu_t = 0$, $Q_{t,i} = q_i h_{t,ii}$, the expression in (9) is then obtained applying targeting to the recursion in (10).

The matrix P_t in Equation (7) is the conditional correlation matrix of the returns vector $\mathbf{r}_t$. Multiplying both sides of (7) by q^2 and further introducing $P_t q^2 = P_t^{(q)}$, $H_t^{(q)}$

can be parameterized as follows:

$$H_t^{(q)} = q^2 D_t P_t D_t = D_t P_t^{(q)} D_t, \tag{11}$$

which makes explicit the dependence of $H_t^{(q)}$ on individual assets volatilities (D_t), conditional correlations (P_t) and tail properties of the error component (q). The matrix $P_t^{(q)}$ is given by the conditional correlation matrix P_t multiplied by q^2. Hence, it will have all its diagonal elements equal to q^2.

Based on Equations (6) and (11), the proposed framework can be shown as:

$$\begin{aligned} H_t^{(q)} &= D_t P_t^{(q)} D_t, \\ H_t^{(c)} &= (1 + \exp(\gamma_0)) H_t^{(q)}, \end{aligned} \tag{12}$$

We refer to our proposed model (12) as a marginalized semi-parametric DCC model. We call it *marginalized* since, given a portfolio composition $\mathbf{w}$, it is fitted to the univariate portfolio returns $r_t(\mathbf{w})$ rather than to the vector process $\mathbf{r}_t$, see discussion related to Equations (1), (2) and (3) for details. The *semi-parametric* nature of the model derives from the fact that, as it will be later discussed in Section 3, its coefficients are fitted minimizing a jointly consistent loss for Q_t and ES_t without assuming the parametric family of the conditional distribution of returns. In particular, we use the negative of the AL log-likelihood in Equation (4) as the loss function. For this reason we denote the proposed semi-parametric DCC model (12) as *DCC-AL* model. The proposed framework can be further explored with a time varying $H_t^{(c)}$ to $H_t^{(q)}$ relationship. However, the results in Taylor (2019) and Gerlach and Wang (2020) show that a constant multiplicative factor between VaR and ES is capable of producing very competitive risk forecasts in univariate risk forecasting models. Therefore, to limit the focus of the paper we employ the constant multiplicative factor $(1 + \exp(\gamma_0))$.

To model P_t in Equation (11) and ensure unit correlations on its main diagonal, according to the DCC model we use:

$$P_t = S_t^{-1/2} R_t S_t^{-1/2}, \tag{13}$$

where $S_t^{1/2}$ is a diagonal matrix containing the diagonal entries of $R_t^{1/2}$, i.e., $S_{t,ii}^{1/2} = R_{t,ii}^{1/2}$. The dynamics of R_t can be parsimoniously modelled as:

$$R_t = \Omega + a\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{'} + bR_{t-1}, \tag{14}$$

where $\boldsymbol{\epsilon}_t$ is a $(n \times 1)$ vector whose i-th element is given by $\epsilon_{t,i} = r_{t,i}/h_{t,ii}$; a and b are non-negative coefficients satisfying the stationarity condition $(a+b) < 1$. When Ω and R_0 are positive definite and symmetric (PDS), the condition $(a+b) < 1$ is sufficient to ensure that R_t is PDS, for any time point t. As further explained in Section 3, the maximum number of simultaneously estimated coefficients can be further reduced by applying correlation targeting (Engle, 2002) in Equation (14):

$$R_t = (1 - a - b)\hat{\Sigma}_{\boldsymbol{\epsilon}} + a\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}^{'} + bR_{t-1}, \tag{15}$$

where

$$\hat{\Sigma}_{\boldsymbol{\epsilon}} = T^{-1}\sum_{t=1}^{T}\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^{'}. \tag{16}$$

Given the assumptions on a and b, no identifiability issue arises and only four coefficients need to be estimated in the DCC-AL model: q, a, b and γ_0. The parameter estimate q (VaR factor) allows us to directly estimate the quantile of the error distribution in the model for portfolio returns. Furthermore, in the Monte Carlo simulations to be shown in Section 4, we can compare the estimates obtained for this coefficient with its theoretical value. Since we have defined $c^2/q^2 = 1 + \exp(\gamma_0)$, after estimating parameters q and γ_0, the value of the ES factor c can be immediately recovered.

Remark 1. It is worth noting that the proposed semi-parametric framework can be easily adapted to consider other multivariate GARCH specifications for H_t. For example, this could be estimated by the BEKK model of Engle and Kroner (1995) that we here consider in its scalar version:

$$H_t = \Omega + a\mathbf{u}_{t-1}\mathbf{u}_{t-1}^{'} + bH_{t-1}, \tag{17}$$

with $\mathbf{u}_t = \mathbf{r}_t - \boldsymbol{\mu}_t$; $a, b \geq 0$ and $a + b < 1$; $\Omega = CC'$ where C is upper triangular. To avoid the estimation of the intercept matrix Ω with $(n+1) \times n/2$ elements, from standard MGARCH literature we can apply covariance targeting to Equation (17) :

$$H_t = \widehat{\Sigma}(1 - a - b) + a\mathbf{u}_{t-1}\mathbf{u}_{t-1}^{'} + bH_{t-1}, \tag{18}$$

where $\widehat{\Sigma} = \widehat{var}(\mathbf{u}_t)$, with $\widehat{var}(.)$ denoting the empirical variance and covariance matrix operator. Multiplying both sides of Equation (18) by q^2, we get:

$$H_t q^2 = \hat{\Sigma}(q^2 - aq^2 - bq^2) + aq^2\mathbf{u}_{t-1}\mathbf{u}_{t-1}^{'} + bq^2 H_{t-1}.$$

Letting $aq^2 = a^{(q)}$, $H_t^{(q)} = H_t q^2$ and $H_t^{(c)} = H_t c^2$, the marginalized semi-parametric BEKK model is then obtained as:

$$H_t^{(q)} = \Sigma(q^2 - a^{(q)} - bq^2) + a^{(q)}\mathbf{u}_{t-1}\mathbf{u}_{t-1}^{'} + bH_{t-1}^{(q)}, \tag{19}$$
$$H_t^{(c)} = (1 + \exp(\gamma_0))H_t^{(q)}.$$

In the remainder, given the DCC type models' greater flexibility and appeal for practitioners, we will exclusively focus on evaluating the performance of the DCC-AL model in risk forecasting and optimization. However, in general, the proposed framework could be easily extended by modelling the dynamics of H_t through a wide range of parameterizations selected from the MGARCH literature. For example, when dealing with large dimensional systems, an interesting addition could be to consider the Dynamic Equicorrelated Conditional Correlation (DECO) model by Engle and Kelly (2012).

3 Estimation

The estimation of the DCC-AL model in (14) can performed by means of a two step procedure in the spirit of Engle (2002). To simplify the presentation, we here set $\mu_{t,i} = \mu_i = 0$ that, in practical applications, is equivalent to work with the demeaned data. The DCC-AL estimation procedure can be then outlined as follows.

Step 1 is dedicated to the estimation of the individual assets volatilities. In order to gain robustness against heavy tailed distributions, the $h_{t,ii}$ are indirectly obtained via the estimation of n separate ES-CAViaR Indirect GARCH (ES-CAViaR-IG, Taylor 2019) models employing a multiplicative VaR to ES factor, as defined in Equation (9)[2].

Namely, for each asset, the coefficients of the individual risk models

$$\boldsymbol{\xi}_i = (\alpha_i^{(q)}, \beta_i, q_i, \gamma_{0,i})$$

[2]Alternatively, the $h_{t,ii}$ could be obtained by Gaussian QML estimation of GARCH(1,1) models. We prefer not to take this route since the Gaussian QML estimator can be characterized by substantial efficiency losses in the presence of heavy tailed errors (Hall and Yao, 2003).

are separately estimated minimizing an AL loss:

$$\widehat{\boldsymbol{\xi}}_i = \arg\min_{\boldsymbol{\xi}_i} \quad \ell_1(r_i; \boldsymbol{\xi}_i) = -\sum_{t=1}^{T} \left(\log \frac{(\alpha - 1)}{\mathrm{ES}_{t,i}} + \frac{(r_{t,i} - Q_{t,i})(\alpha - I(r_{t,i} \leq Q_{t,i}))}{\alpha \mathrm{ES}_{t,i}} \right),$$

for $i = 1, \ldots, n$. The set of estimated 1-st stage coefficients is denoted as: $\widehat{\boldsymbol{\xi}} = (\widehat{\boldsymbol{\xi}}_1, \ldots, \widehat{\boldsymbol{\xi}}_n)$.

Step 2 is dedicated to the estimation, conditional on $\widehat{\boldsymbol{\xi}}$, of the coefficients controlling correlation dynamics $(\mathrm{vech}(\Omega), a, b)$ and the tail properties of the conditional distribution of portfolio returns (q, γ_0). Here, as usual, the notation $\mathrm{vech}(\Omega)$ denotes the column-stacking operator applied to the upper portion of the symmetric matrix Ω. Again, these coefficients can be jointly estimated minimizing the AL loss function:

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \qquad \ell_2(\mathbf{r}; \boldsymbol{\theta} | \widehat{\boldsymbol{\xi}}) = -\sum_{t=1}^{T} \left(\log \frac{(\alpha - 1)}{\mathrm{ES}_t} + \frac{(r_t - Q_t)(\alpha - I(r_t \leq Q_t))}{\alpha \mathrm{ES}_t} \right), \quad (20)$$

where Q_t and ES_t are the portfolio VaR and ES as defined in Equation (5), $\boldsymbol{\theta} = (\mathrm{vech}(\Omega), a, b, q, \gamma_0)$.

Namely, the estimated volatilities from Step 1 are used to compute the estimated standardized asset returns:

$$\widehat{\epsilon}_t = \widehat{D}_t^{-1} \mathbf{r}_t,$$

where $\widehat{D}_{t,ii} = \widehat{h}_{t,ii} = \widehat{Q}_{t,i} / \hat{q}_i$, with the hat denoting estimated quantities; $\widehat{\epsilon}_t$ is then plugged into Equation (14) to obtain $\widehat{R}_t$ that can in turn be used to obtain $\widehat{P}_t$ using Equation (13); $\widehat{P}_t$ and $\widehat{D}_t$ are used to produce the $\widehat{H}_t^{(q)}$ and $\widehat{H}_t^{(c)}$ with (12). Finally, portfolio VaR and ES, $\widehat{Q}_t$ and $\widehat{\mathrm{ES}}_t$, can be calculated with (5).

Remark 2. It is worth noting that direct estimation of $\mathrm{vech}(\Omega)$ would imply optimizing the AL loss wrt $n(n+1)/2$ additional parameters, that can be hardly feasible for even moderately large cross-sectional dimensions. For example, in our empirical analysis, where we work with a portfolio of dimension $n = 28$, we would have to estimate 406 distinct Ω coefficients. To overcome this issue, correlation targeting can be applied. When correlation targeting is used and R_t is modelled as in Equation (15), the estimation procedure is modified to incorporate an intermediate step in which $var(\boldsymbol{\epsilon}_t)$ is estimated by the sample variance and covariance matrix of $\widehat{\boldsymbol{\epsilon}}_t = \widehat{D}_t^{-1} \mathbf{r}_t$ ($\widehat{\Sigma}_{\boldsymbol{\epsilon}}$). Therefore, the loss in Equation (20)

is then optimized only wrt (a, b, q, γ_0).

In step 1, we use the Matlab 2021b "fmincon" optimization routine to estimate the ES-CAViaR-IG models. In step 2, the Matlab "MultiStart" facility is further employed for "fmincon" to improve the robustness of the estimation results. We use 5 separate sets of starting points, generating 5 local solutions, then the optimum set among these is finally chosen as the parameter estimates. The detail of "MultiStart" is explained in the the Matlab documentation `https://au.mathworks.com/help/gads/multistart.html`. Using a standard laptop with i7 1.8GHz CPU and 16GB RAM, the estimation of the DCC-AL model with a dataset of 3000 sample size and $n = 28$ assets takes approximately 5 minutes.

4 Simulation Study

In this section a simulation study is conducted to assess the statistical properties of the two-step AL loss based estimation procedure discussed in the previous section. The aim of this study is firstly to assess the bias and efficiency of the estimators of the DCC-AL parameters (a, b, q, γ_0). Furthermore, we also assess the the risk forecasting performances of the estimated DCC-AL model in an equally weighted portfolio, through evaluating the one-step-ahead 2.5% level portfolio VaR and ES forecast accuracy as compared to the "true" simulated values.

Consistent with the empirical findings arising from our empirical application, we have considered as Data Generating Process (DGP) a parametric DCC model of dimension n=28 with the true values of the correlation dynamic parameters in Equation (15) given by $a = 0.12$ and $b = 0.78$:

$$Q_t = (1 - 0.12 - 0.78)\Sigma_{\boldsymbol{\epsilon}} + 0.12\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}^{'} + 0.78Q_{t-1}, \tag{21}$$

where $\Sigma_{\boldsymbol{\epsilon}}$ is a $n \times n$ matrix with unit diagonal and off-diagonal elements equal to 0.5.

In the return equation $\mathbf{r}_t = \boldsymbol{\mu}_t + H_t^{1/2}\mathbf{z}_t$, the conditional mean vector $\boldsymbol{\mu}_t$ of returns is chosen to be zero. The DGP univariate volatilities are assumed to follow the GARCH(1,1)

model:

$$h_{t,i}^2 = 0.1 + 0.1r_{t-1,i}^2 + 0.8h_{t-1,i}^2, \qquad i = 1, \ldots, n.$$

The parametric distribution of returns is assumed to be $\mathbf{z}_t \overset{iid}{\sim} D_1(\mathbf{0}, I_n)$, where the following choices of D_1 have been considered:

- Standardized multivariate Normal distribution: $\mathbf{z}_t \overset{iid}{\sim} \mathcal{N}_n(\mathbf{0}, I_n)$.
- Standardized multivariate Student's t distribution: $\mathbf{z}_t \overset{iid}{\sim} t_n(\mathbf{0}, I_n; \nu)$, where the degrees of freedom parameter ν has been set equal to 10.
- A multivariate non-spherical distribution with Student's t marginals (*nst*: $\mathbf{z}_t \overset{iid}{\sim} nst_n(\mathbf{0}, I_n; \boldsymbol{\nu})$; the density of this distribution is obtained taking the product of n independent univariate standardized $t_1(0, 1; \nu_i)$ densities; $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ is the n-dimensional vector of marginal degrees of freedom parameters. When $\nu_i = \nu$ $(\forall i)$, the marginal densities of the product are the same as those of the multivariate t, although the joint density is different (Bauwens and Laurent, 2005). In our simulations, the values of ν_i have been uniformly drawn over the interval [5,15].

While the first two distributions belong to the spherical family, the same does not hold for the third one (*nst*).

The DCC-AL model is then fitted to the time series of equally weighted ($\mathbf{w} = 1/n$) portfolio returns generated from the chosen DGP, for three different sample sizes $T \in \{2000, 3000, 5000\}$. Overall, matching 3 different distributional assumptions and sample sizes, we have 9 simulation settings and, under each of these, 250 return series have been generated.

Conditional on the chosen error distribution and on the values of the correlation parameters a and b, the true values of the other two parameters q and γ_0 in the DCC-AL model can be then calculated as follows.

In the spherical case, multivariate Normal and t, letting $F(.)$ be the CDF of $z_{t,i}$, $q = F^{-1}(\alpha)$ and $c = E(z_{t,i}|z_{t,i} \leq q)$. From (6) it is then easy to obtain:

$$\gamma_0 = \log\left(\frac{c^2}{q^2} - 1\right).$$

Therefore, taking the multivariate Normal case as an example, in the DCC-AL the true value of q^2 is $\Phi^{-1}(\alpha)^2 = (-1.96)^2 = 3.8415$. The true value for c^2 is $\left(\frac{\phi(\Phi^{-1}(\alpha))}{\alpha}\right)^2 =$

$(-2.3378)^2$, where Φ is the standard Normal CDF and ϕ is standard Normal Probability Density Function (PDF). Further, we have $(1+\exp(\gamma_0)) = \frac{c^2}{q^2} = \frac{(-2.3378)^2}{(-1.9600)^2} = 1.4227$, thus $\gamma_0 = \log(1.4227-1) = -0.8610$. These true values of a, b, γ_0 and q^2 for the multivariate Normal distribution are shown in the *True* rows in Table 1. When $\mathbf{z}_t \overset{iid}{\sim} t_n(\mathbf{0}, I_n; \nu = 10)$, the true values of q, c and γ_0 are obtained through a similar approach, replacing the analytical formulas for q and c with the equivalent expressions for an univariate t distribution.

In the *nst* case, the true values of q, c and γ_0 need to be calculated via simulations since this distribution is not a member of the spherical family. It also follows that, differently from what observed for the multivariate Normal and t cases, the values of these coefficients will be dependent on the chosen portfolio allocation. So they could change when moving from equal weighting to a different allocation scheme.

More specifically, q and c can be calculated by the following procedure.

1. For a given set of randomly generated degrees of freedom values $\boldsymbol{\nu}$ and portfolio allocation $\mathbf{w}$, a return series with size $T_{sim} \times n$ ($T_{sim} = 10^5$) is simulated from the specific DCC process taken as DGP: $\mathbf{r}_j^*$, $j = 1, \ldots, T_{sim}$.

2. Given the simulated returns $\mathbf{r}_j^*$, conditional covariance matrix H_j^* and portfolio allocation $\mathbf{w}$, the time series of simulated standardized portfolio returns is computed as:
$$z_j^* = \frac{r_j^*}{h_j^*},$$
for $j = 1, \ldots, T_{sim}$ and where
$$r_j^* = \mathbf{w}'\mathbf{r}_\mathbf{j}^*, \qquad h_j^* = \sqrt{\mathbf{w}' H_j^* \mathbf{w}}.$$

3. The empirical quantile and conditional tail average of the z_j^* series are then used as the simulated "true" values for q and c. These values can be used to calculate the true values for γ_0 through inverting Equation (6).

For ease of reference, for each DGP, the true parameter values for (a, b, q, γ_0) are included in the *True* rows in Table 1.

The DCC-AL model is then fitted to each simulated dataset, using the estimation procedure as outlined in Section 3. The simulation results for the DCC-AL model are shown in Table 1 where, for ease of presentation, we have chosen not to report parameter estimates of the fitted first stage ES-CAViaR-IG models.

The rows labeled as *True* report, for each DGP, the values of the parameter true values used for simulation. The empirical averages of the 250 parameter estimates, for various return distributions and sample sizes, are shown in the *Mean* rows. The Root Mean Squared Error (RMSE) values between the parameter estimates and the true values are shown in the *RMSE* rows in Table 1.

To further evaluate the accuracy of the portfolio VaR and ES forecasts from the estimated DCC-AL model, for all 250 simulated datasets, we compare the one-step-ahead $\widehat{Q}_{T+1}$ and $\widehat{\mathrm{ES}}_{T+1}$ forecasts based on the *estimated* parameters with their counterparts based on the *true* DGP coefficients. For $\widehat{Q}_{T+1}$ and $\widehat{\mathrm{ES}}_{T+1}$, the *True* rows then report the averages of the 250 risk forecasts based on the *true* parameters, for different return distributions and sample sizes.

First, the results provide support to the use of the two-step AL based estimation method and show that it is able to produce relatively accurate parameter estimates. Overall, the estimated bias ($|Mean - True|$) is reasonably low. When the distribution is fixed, the absolute bias decreases as the time series length T increases. For all parameters, the values of the RMSE also monotonically decrease as T increases, suggesting consistency of the estimation procedure. When the sample size is fixed, as expected, the RMSE values are clearly larger when the errors follow a multivariate Student's t, comparing to multivariate Normal distribution.

The estimation results are still relatively accurate for the non-spherical *nst* distribution that, for correlation parameters, returns RMSE values slightly higher than those obtained in the Normal case. For parameters q and γ_0, the simulated RMSE is in line with the Normal case, for γ_0, and even lower for q.

These results suggest that the proposed semi-parametric estimation procedure is able not only to keep track of the volatility and correlation dynamics but also of the distributional properties of portfolio returns, through the estimation of q and γ_0 (implicitly c).

Finally, reminding that the main motivation for the DCC-AL model is the generation of accurate portfolio risk forecasts, the last two columns of Table 1 provide a very important benchmark for assessing the properties of the proposed estimation procedure. Comparing the $\alpha = 2.5\%$ *estimated* and *true* risk forecasts, for both VaR and ES, it can be noted that these two series are on average very close even for the shortest sample size $T = 2000$. The RMSE values are also remarkably low: with $T = 2000$ they do not exceed 0.0811 for VaR and 0.1034 for ES, and their values monotonically decrease as T increases across three return distributions. This last set of results confirms the ability of the proposed two-stage estimation procedure to accurately reproduce the portfolio risk dynamics.

5 A portfolio risk optimization procedure using the proposed semi-parametric framework

A further appeal of the proposed DCC-AL models is that they can be used to optimize portfolio risk and produce minimum VaR and ES portfolios, that is evidently not possible when using univariate semi-parametric approaches.

Portfolio optimization (PO) targeted on the minimization of VaR/ES has by far received much less attention than the traditional mean-variance optimization. Under the assumption of normally distributed returns, Rockafellar and Uryasev (2000) show that the computation of the mean-VaR, mean-ES and mean-variance frontiers result in equivalent optimization problems. Namely, they show that the efficient frontiers obtained from mean-VaR and mean-ES optimization are subsets of the mean-variance efficient frontier. It is easy to show that a parallel result holds for the family of elliptical distributions.

Assuming that the conditional distribution of returns on portfolio components is elliptical, through rewriting Equation (5) we can get the portfolio VaR and ES analytical solution as:

$$Q_t = \mathbf{w}'\boldsymbol{\mu}_t + q\sqrt{\mathbf{w}'H_t\mathbf{w}}, \qquad \mathrm{ES}_t = \mathbf{w}'\boldsymbol{\mu}_t + c\sqrt{\mathbf{w}'H_t\mathbf{w}}, \tag{22}$$

where VaR and ES factors q and c depend on the level α but do not depend on the weight vector $\mathbf{w}$ and, hence, on portfolio configuration. Therefore, under the ellipticity

Table 1: DCC-AL model parameter estimates and VaR and ES forecasting results with simulated datasets.

DGP		a	b	γ_0	q^2	$\widehat{Q}_{T+1}$	$\widehat{\mathrm{ES}}_{T+1}$
$\mathcal{N}$, $T = 2000$	True	0.1200	0.7800	-0.8610	3.8415	-1.3704	-1.6346
	Mean	0.1360	0.7150	-0.9131	3.9308	-1.3849	-1.6420
	RMSE	0.0751	0.1995	0.1639	0.2057	0.0606	0.0762
$\mathcal{N}$, $T = 3000$	True	0.1200	0.7800	-0.8610	3.8415	-1.3565	-1.6180
	Mean	0.1310	0.7319	-0.8919	3.9079	-1.3669	-1.6244
	RMSE	0.0615	0.1736	0.1200	0.1613	0.0521	0.0611
$\mathcal{N}$, $T = 5000$	True	0.1200	0.7800	-0.8610	3.8415	-1.3665	-1.6299
	Mean	0.1265	0.7528	-0.8830	3.9018	-1.3760	-1.6367
	RMSE	0.0466	0.1118	0.1029	0.1363	0.0392	0.0454
t, $T = 2000$	True	0.1200	0.7800	-0.5097	3.9717	-1.4353	-1.8159
	Mean	0.1591	0.6719	-0.5844	4.1136	-1.4479	-1.8108
	RMSE	0.1023	0.2419	0.1894	0.2967	0.0811	0.1034
t, $T = 3000$	True	0.1200	0.7800	-0.5097	3.9717	-1.3825	-1.7491
	Mean	0.1497	0.6894	-0.5658	4.0686	-1.3990	-1.7543
	RMSE	0.0884	0.2248	0.1527	0.2167	0.0692	0.0850
t, $T = 5000$	True	0.1200	0.7800	-0.5097	3.9717	-1.4155	-1.7909
	Mean	0.1303	0.7395	-0.5459	4.0619	-1.4304	-1.7995
	RMSE	0.0617	0.1564	0.1172	0.1613	0.0534	0.0703
nst, $T = 2000$	True	0.1200	0.7800	-0.8311	3.8772	-1.3570	-1.6259
	Mean	0.1440	0.7030	-0.8893	3.9012	-1.3538	-1.6096
	RMSE	0.0894	0.2245	0.1656	0.1849	0.0639	0.0762
nst, $T = 3000$	True	0.1200	0.7800	-0.8311	3.8772	-1.3806	-1.6542
	Mean	0.1356	0.7223	-0.8657	3.8978	-1.3799	-1.6461
	RMSE	0.0729	0.1926	0.1246	0.1539	0.0576	0.0706
nst, $T = 5000$	True	0.1200	0.7800	-0.8311	3.8772	-1.4001	-1.6775
	Mean	0.1277	0.7486	-0.8569	3.9091	-1.4057	-1.6788
	RMSE	0.0496	0.1195	0.1029	0.1251	0.0383	0.0465

assumption, the same VaR and ES factors could be used for computing risk for different portfolio configurations. On the other hand, if this assumption is removed, q and c could change as a function of $\mathbf{w}$. In practical applications this would imply that the risk model should be re-estimated each time the portfolio configuration changes.

In financial applications, deviations from the ellipticity assumption often occur due to the presence of specific features such as skewness. A notable example of non-elliptical distributions is given by the multivariate skew-t distribution that is derived by Bauwens and Laurent (2005) through the "contamination" of a symmetric multivariate t distribution via the method proposed by Fernández and Steel (1998).

In the parametric world, if the ellipticity assumption is removed, the availability of an analytical formula for portfolio VaR and ES forecasting is not guaranteed and, in most cases, simulation methods should be used. Using a semi-parametric approach allows to obtain VaR and ES forecasts through an analytical formula where the VaR and ES factors are not constant but their values depend on $\mathbf{w}$. This complicates the implementation of minimum risk PO procedures since the risk model should be in principle re-estimated at each PO iteration, making the numerical solution of the problem computationally challenging.

In this section, to overcome these difficulties we propose a simple and empirically motivated minimum risk PO procedure that is here presented with reference to the DCC-AL model. However, it is worth remarking that the procedure is not specific to this model but can be immediately extended to other semi-parametric specifications such as DCC models fitted by both Gaussian QML and Composite QML. The proposed PO procedure can be illustrated as follows.

1. Estimate the DCC-AL model with variance targeting by minimizing the AL joint loss, to produce parameter estimates: $\hat{a}$, $\hat{b}$, $\hat{q}$ and $\hat{\gamma}_0$. We use equal weights to start the process.

2. Relying on the estimated DCC-AL model, use the estimated $\widehat{H}_t^{(q)}$ and $\hat{q}$ to calculate the conditional covariance matrix $\widehat{H}_t = \frac{\widehat{H}_t^{(q)}}{\hat{q}^2}$.

3. Repeat the following process for N_{sim} times (we use $N_{sim} = 1000$):

   (a) Randomly generate a vector of portfolio weights $\mathbf{w}$ of size $n \times 1$, where

   $$w_i = \frac{u_i}{\sum_{i=1}^{n} u_i},$$

   with $u_i \overset{iid}{\sim} U(0,1)$, and $U(0,1)$ represents a uniform distribution on the interval [0,1].

(b) Compute the portfolio conditional variance: $\mathbf{w}'\widehat{H}_t\mathbf{w}$, $t = 1, \ldots, T$.

(c) Compute the series of standardized portfolio returns $\hat{z}_t = \frac{\mathbf{w}'(\mathbf{r}_t - \hat{\boldsymbol{\mu}})}{\sqrt{\mathbf{w}'\widehat{H}_t\mathbf{w}}}$, $t = 1, \ldots, T$, where $\hat{\boldsymbol{\mu}}$, of size $n \times 1$, is the vector of estimated mean return levels for n assets and used to demean the data.

(d) For the generated $\hat{z}_t$ series, compute sample quantile $\tilde{q}$ and sample conditional tail average $\tilde{c}$.

After N_{sim} iterations, we obtain the following entities:

- a matrix of simulated portfolio weights: $\mathbf{w}$, of size $N_{sim} \times n$,
- a vector of simulated VaR factors: $\tilde{\mathbf{q}}$, of size $N_{sim} \times 1$,
- a vector of simulated ES factors: $\tilde{\mathbf{c}}$, of size $N_{sim} \times 1$,

4. Run two separate ordinary least squares (OLS) regressions using $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{c}}$ as dependent variables, respectively, and $\mathbf{w}$ as the matrix of regressors[3], to get the regression coefficients $\widehat{\boldsymbol{\beta}}_q$ and $\widehat{\boldsymbol{\beta}}_c$, respectively.

5. For a given $\mathbf{w}$, the portfolio VaR and ES can be calculated as:

$$\widehat{Q}_t = \mathbf{w}'\hat{\boldsymbol{\mu}}_t + \mathbf{w}'\widehat{\boldsymbol{\beta}}_q\sqrt{\mathbf{w}'\widehat{H}_t\mathbf{w}}, \qquad \widehat{\mathrm{ES}}_t = \mathbf{w}'\hat{\boldsymbol{\mu}}_t + \mathbf{w}'\widehat{\boldsymbol{\beta}}_c\sqrt{\mathbf{w}'\widehat{H}_t\mathbf{w}}, \tag{23}$$

here the OLS predictions $\mathbf{w}'\widehat{\boldsymbol{\beta}}_q$ and $\mathbf{w}'\widehat{\boldsymbol{\beta}}_c$ can be treated as predictions for the underlying VaR and ES factors (q and c).

6. Given a target portfolio return μ_0, minimum risk portfolios are then calculated minimizing $\widehat{Q}_t$ or $\widehat{\mathrm{ES}}_t$ with respect to $\mathbf{w}$ under the set of constraints[4]:

$$\mathbf{w}'\hat{\boldsymbol{\mu}}_t \geq \mu_0; \qquad 0 \leq w_i \leq 1; \qquad \sum_{i=1}^{n} w_i = 1.$$

In the paper we focus on the left tail of the return distribution, thus the portfolio VaR and ES have negative values. Therefore, the minimization objectives are based on the absolute values of portfolio VaR and ES. In order to simplify the illustration of the procedure, we have ruled out short selling practices by imposing positive portfolio

[3]The constant term is omitted to avoid multicollinearity due to the fact that $\sum_{i=1}^{n} w_i = 1$

[4]Minimization has been performed using the Matlab "fmincon" optimization routine in Matlab 2021b.

weights. However, it is worth noting that this doesn't imply any loss of generality since the framework can be readily extended to consider the case of negative weights.

Steps 1-4 are aimed at predicting the VaR and ES factors as a function of the portfolio composition $\mathbf{w}$, without assuming ellipticity and having to re-estimate the DCC-AL model at each PO iteration. A simple *naive* alternative would be to replace, in the $\widehat{Q}_t$ and $\widehat{\mathrm{ES}}_t$ calculations as in Equation (23), the OLS predictions $\mathbf{w}'\widehat{\boldsymbol{\beta}}_q$ and $\mathbf{w}'\widehat{\boldsymbol{\beta}}_c$ by the noisy simulated VaR and ES factors $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{c}}$ respectively. Under this respect, the regression in step 4 can then be seen as a linear filter whose aim is to attenuate the noise in the simulated $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{c}}$ series (as can be seen in the Figure 6 to be shown in Section 6.3). Alternative choices of the OLS filter could in principle be considered, such as a Neural Network trained on the simulated factors. The motivations for our preference for the OLS filter are twofold. First, the use of the OLS is computationally convenient. Second, in our empirical study, such as Figure 7 to be shown in Section 6.3, we found the OLS filter is able to approximate the underlying (unobserved) VaR and ES factors (q and c) with good accuracy.

6 Empirical Study

The performance of the proposed models in forecasting risk and optimizing minimum risk portfolios has been assessed via an empirical study on a multivariate time series of US stock returns.

After providing a short description of the data and describing the forecasting design in Section 6.1, in Section 6.2 we focus on assessing the risk forecasting performance of the proposed models for an equally weighted portfolio. This choice does not imply any loss of generality since our investigation could be easily replicated under alternative portfolio configurations. Our preference for the equally weighted scheme is motivated by the robust performance of this simple allocation rule that, in many instances, has been found to be competitive with more sophisticated benchmarks (DeMiguel et al., 2007).

In addition, following the discussion in Section 5, in Section 6.3 we also evaluate the effectiveness of proposed framework in generating minimum risk portfolios.

6.1 Equally weighted portfolio risk forecasting

Daily closing price data are collected for 28 out of the 30 components of the Dow Jones index, for the period 4 January 2005 to 28 September 2021. Only assets providing full coverage of the period of interest have been considered for the analysis.

For day t, denoting the daily closing prices as C_t, the daily log return can be calculated as:

$$r_{t,i} = \log(C_{t,i}) - \log(C_{t-1,i}),$$

where $r_{t,i}$ is the return for i-th asset in the portfolio on day t, $i = 1, \ldots, n$, and $t = 1, \ldots, T$. The total sample size T is 4213. Therefore, at time t we have $\mathbf{r}_t = (r_{t,1}, \ldots, r_{t,n})$ which is a $(n \times 1)$ vector of returns on the $n = 28$ portfolio assets. For the equally weighted portfolio considered in our analysis, the portfolio return for day t is then given by:

$$r_t(\mathbf{w} = 1/n) = \frac{\sum_{i=1}^{n} r_{t,i}}{n},$$

The computed time series of equally weighted portfolio returns is shown in Figure 1. As can be seen in the plot, two major high volatility periods are clearly visible. The outbreak of coronavirus disease (COVID-19) has caused a highly volatile period in 2020 and, less recently, the 2008 Global Financial Crisis has also greatly impacted the financial market.

A rolling window scheme, with fixed in-sample size $T_{\text{in}} = 3000$ and daily re-estimation, is then implemented to generate $T_{\text{out}} = 1213$ out-of-sample one-step-ahead forecasts of VaR and ES at 2.5% level. Therefore, the in-sample period is from 4 January 2005 to 1 December 2016, and the out-of-sample period covers the time range from 2 December 2016 to 28 September 2021.

Several existing univariate semi-parametric models are selected as benchmarks for comparison. First we consider the recently proposed ES-CAViaR model (Taylor, 2019) with Symmetric Absolute Value (SAV) and Indirect GARCH (IG) specifications for the quantile regression equation. Regarding the ES specification, to facilitate the performance comparison with the DCC-AL, we choose the model with the constant multiplicative ES to VaR factor $(1 + \exp(\gamma_0))$, which is the factor that we use in the proposed DCC-AL model.

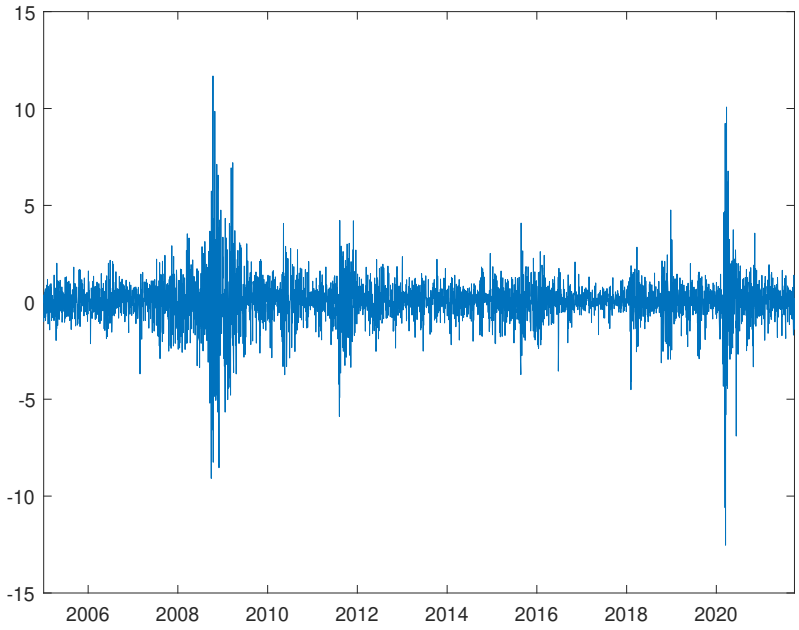

Figure 1: Equally weighted portfolio returns for 28 assets in the Dow Jones index.

In addition, the semi-parametric Conditional Autoregressive Expectile framework (CARE, Taylor 2008), with SAV and IG specifications, is also included.

Furthermore, the risk forecasting performances of the proposed framework are compared to those yielded by the conventional DCC models fitted through different estimation approaches. First, we consider a standard DCC model fitted by a two stage procedure (Engle, 2002) combining maximization of Gaussian likelihoods in the first (volatility) and second (correlation) stage of the estimation procedure with the application of correlation targeting. At the VaR and ES forecasting stage, depending on the assumptions formulated on the error distribution, two different series of forecasts are generated labeled as DCC-N, when a multivariate Normal distribution is assumed, and DCC-QML, when a semi-parametric approach is taken. Namely, for the DCC-N approach, the theoretical quantile and conditional quantile average based on the Normal distribution are used for VaR and ES calculation. For the QML, differently, the semi-parametric filtered historical simulation approach is used to calculate the VaR and ES. The error quantiles $\hat{q}$ and tail expectations $\hat{c}$ are then estimated by computing the relevant sample quantiles and tail averages of standardized returns (r_t divided by its volatility). Finally, level-α VaR and ES

forecasts are obtained by multiplying $\hat{q}$ and $\hat{c}$, respectively, by the portfolio conditional standard deviation forecast from the fitted DCC model.

When applied to even moderately large datasets, such as the one that is here considered, the original approach to the estimation of DCC parameters described in Engle (2002) has been found to be prone to return biased estimates of the correlation dynamic parameters. This motivates our choice to consider, as a further benchmark, the Composite Likelihood (clik) approach developed in Pakel et al. (2021). Along the same lines discussed above, the estimated volatility and correlation parameters are then used to generate two different sets of VaR and ES forecasts labeled as DCC-clik-N and DCC-clik-QML, respectively.

Finally, the picture is completed by considering a parametric DCC model fitted by ML using multivariate Student's t likelihoods in the estimation of volatility and correlation parameters (DCC-t). VaR and ES forecasts are generated considering theoretical quantiles and tail expectations for a standardized Student's t distribution.

For all the DCC benchmarks, in order to guarantee a fair comparison with the DCC-AL model, the fitted univariate volatility specifications are given by GARCH(1,1) models.

In Figure (2), it is interesting to note that the estimated correlation parameters in the DCC-AL model vary over the forecasting period, reacting to changes in the underlying market volatility level. In particular, the estimated value of a is characterized by a positive trend originating at the outbreak of the pandemic COVID-19 crisis. An opposite behaviour is observed for b.

Table 2 compares the DCC-AL estimates of correlation coefficients with those obtained for the benchmark DCC specifications considered, reporting the average estimates of correlation parameters a and b across all forecasting steps. Here, consistent with previous findings in the literature, standard DCC estimators tend to oversmooth (very close to 1 estimates for b) correlations, while this is not the case for DCC-AL models. The DCC-clik estimates stay in between. It is here worth noting that the DCC-clik and DCC-AL models are based on different first stage volatility estimators: QML-GARCH for DCC-clik, and ES-CAViaR-IG as in Equation (8) for DCC-AL. In addition, Figure 3 shows the cross-sectional averages of the estimated conditional correlations (calculated based on P_t in Equation (13)) across the forecasting period. As can be seen, the dynamics of the

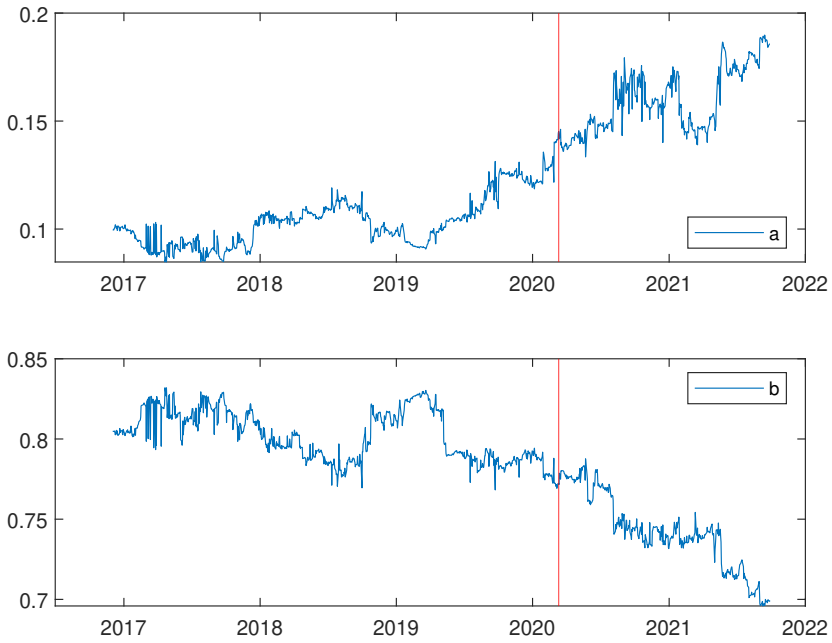

Figure 2: Plots of estimated correlation parameters a and b of the DCC-AL model. The vertical lines denote the outbreak of the COVID-19 emergency (as officially declared by the World Health Organization on 30 January 2020). Whole forecasting sample (1213 days).

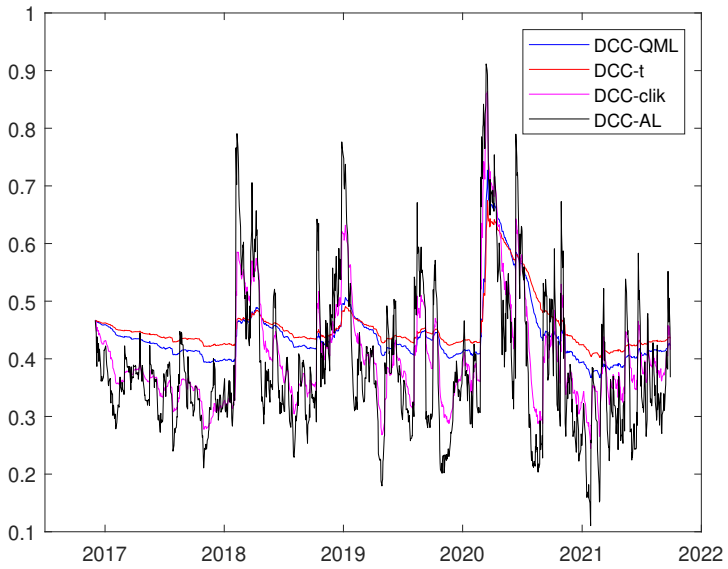

Figure 3: Cross-sectional averages of the estimated conditional correlation for DCC-QML, DCC-t, DCC-clik and DCC-AL across the whole forecasting sample (1213 days).

estimated conditional correlations between DCC-AL and DCC-clik are relatively close to each other.

Table 2: Average estimates (across time) of correlation dynamic parameters (a and b) from different models and estimation methods. Whole forecasting sample (1213 days).

	DCC-AL	DCC-QML	DCC-t	DCC-clik
a	0.1217	0.0040	0.0023	0.0308
b	0.7827	0.9788	0.9816	0.9198

The average of estimated first stage univariate modelling coefficients, across all assets and forecasting steps, is reported in Table 3. GARCH-t and QML-GARCH, used in DCC-t and DCC-QML respectively, appear to be less reactive (smaller α estimates) to past shocks compared to the ES-CAViaR-IG, used in the DCC-AL model.

Table 3: Average estimates (across times and assets) of the first stage univariate modelling parameters (α and β) from different models and estimation methods across the full forecasting sample (1213 days). For ES-CAVIAR-IG, we report $\alpha = \alpha^{(q)}/q^2$.

	GARCH-t	QML-GARCH	ES-CAViaR-IG
α	0.0902	0.0852	0.1369
β	0.8702	0.8944	0.8059

6.2 Evaluation of VaR and ES forecasts

Assuming equally weighted portfolio returns $r_t(\mathbf{w} = 1/n)$, one-step-ahead forecasts of daily VaR and ES from the proposed DCC-AL model are produced by using Equation (5). In this section, these forecasts are compared with the ones from the competing models presented in Section 6.1.

The standard quantile loss function is employed to compare the models for VaR forecast accuracy: the most accurate VaR forecasts should minimize the quantile loss function, given as:

$$\sum_{t=T_{\text{in}}+1}^{T_{\text{in}}+T_{\text{out}}} (r_t - \widehat{Q}_t)(\alpha - I(r_t \leq \widehat{Q}_t)) \ , \tag{24}$$

where T_{in} is the in-sample size, T_{out} is the out-of-sample size and $\widehat{Q}_{T_{\text{in}}+1}, \ldots, \widehat{Q}_{T_{\text{in}}+T_{\text{out}}}$ is a series of VaR forecasts at level α for observations $r_{T_{\text{in}}+1}, \ldots, r_{T_{\text{in}}+T_{\text{out}}}$.

Moving to the assessment of joint (VaR, ES) forecasts, as discussed in Section 2.2, Taylor (2019) shows that the negative logarithm of the likelihood function built from Equation (4) is strictly consistent for Q_t and ES_t considered jointly, and fits into the class of strictly consistent joint loss functions for VaR and ES developed by Fissler and Ziegel (2016). This loss function is also called the AL log-score in Taylor (2019) and is defined as:

$$S_t(r_t, \widehat{Q}_t, \widehat{\text{ES}}_t) = -\log\left(\frac{\alpha - 1}{\widehat{\text{ES}}_t}\right) - \frac{(r_t - \widehat{Q}_t)(\alpha - I(r_t \leq \widehat{Q}_t))}{\alpha \widehat{\text{ES}}_t}. \tag{25}$$

In our analysis, we use the joint loss $S = \sum_{t=T_{\text{in}}+1}^{T_{\text{in}}+T_{\text{out}}} S_t$ to formally and jointly assess and compare the VaR and ES forecasts from all models.

First, Figure 4 visualizes the 2.5% portfolio ES forecasts from the DCC-QML, DCC-

clik-QML and DCC-AL models, using equally weighted Dow Jones returns. In general, we can see that the ES forecasts produced from the DCC-AL model are comparable to the ones from the DCC-QML and DCC-clik-QML. In order to more clearly visualise the difference in the ES forecasts from the three models, Figure 5 shows the zoomed in results of the Figure 4. Inspecting the ES forecasts at the start of 2020 period, when the financial market is greatly impacted by the outbreak of COVID-19, we can see that the three models have distinctive behaviours. Comparing to the DCC-QML model, the DCC-AL model is more reactive to the return shocks.This is because of the larger a and α estimates as discussed in Tables 2 and 3. The ES forecasts from the DCC-clik-QML stay in between the ones from DCC-QML and DCC-AL, which is also consistent with the findings from Table2.

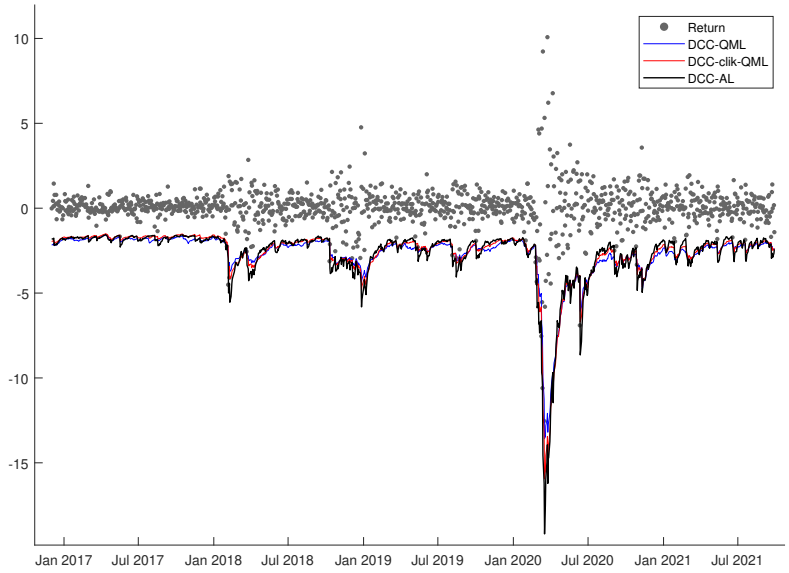

Figure 4: 2.5% ES forecasts from the DCC-QML, DCC-clik-QML and DCC-AL models, using equally weighted Dow Jones returns.

The quantile and joint loss results from the proposed DCC-AL model and other competing models are shown in Table 4. Overall, the results show that the proposed DCC-AL model generates competitive loss results comparing to other models, which lends evidence on the validity of the proposed semi-parametric framework. Furthermore, the DCC-AL model generates smaller loss values than all other parametric or semi-parametric

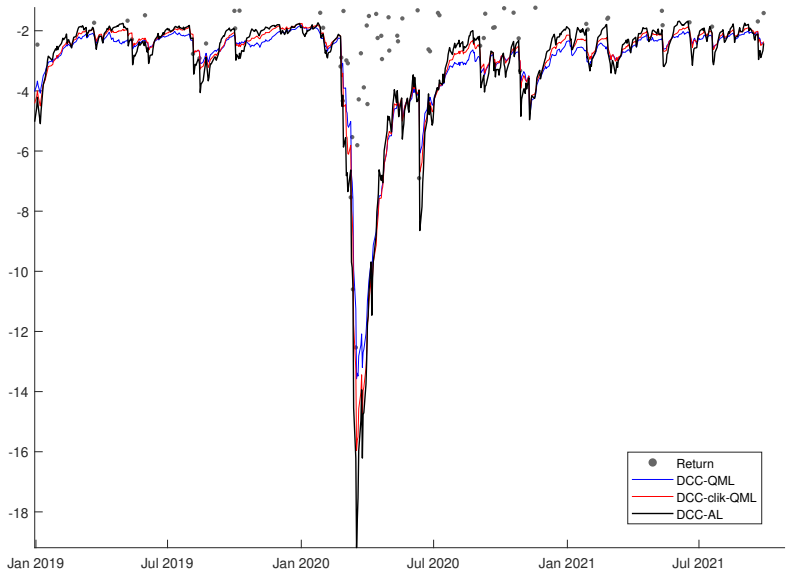

Figure 5: Zoomed in 2.5% ES forecasts from the DCC-QML, DCC-clik-QML and DCC-AL models, using equally weighted Dow Jones returns.

DCC models that have been taken as benchmarks.

The Model Confidence Set (MCS) of Hansen et al. (2011) is employed to statistically compare the quantile loss (Equation (24)) and joint loss (Equation (25)) values yielded by the different models. A MCS is a set of models, constructed such that it contains the best model with a given level of confidence, selected as either 75% or 90% in our paper. Two methods, R and SQ (see Hansen et al., 2011, for details), are employed to calculate the MCS test statistic. The MCS results, using both quantile and joint loss functions, are shown in Table 5. As can be seen, the DCC-AL is included in MCS for both methods across both 90% and 75% tests. However, the DCC-QML, DCC-N and DCC-clik-N models are excluded from the 90% MCS with the SQ method and joint loss. On the 75% level with the SQ method and joint loss, all the DCC models, except the proposed DCC-AL, are excluded from the MCS.

To further backtest the VaR forecasts, we have employed several quantile accuracy and independence tests, including: the unconditional coverage (UC) test (Kupiec et al., 1995); the conditional coverage (CC) test (Christoffersen, 1998); the dynamic quantile (DQ) test (Engle and Manganelli, 2004); and the quantile regression based VaR (VQR)

Table 4: The quantile loss (Equation (24)) and joint loss function (Equation (25)) values for all competing models across the full forecasting sample.

Model	Quantile loss	Joint loss
CARE-SAV	90.4	2,405.4
CARE-IG	89.1	2,392.1
ES-CAViaR-SAV	89.4	2,398.1
ES-CAViaR-IG	87.3	2,369.5
DCC-QML	98.7	2,492.6
DCC-N	101.2	2,566.8
DCC-clik-QML	95.2	2,453.9
DCC-clik-N	97.7	2,528.9
DCC-t	101.5	2,555.3
DCC-AL	89.1	2,386.8

test (Gaglianone et al., 2011). Table 6 presents the UC, CC, DQ (lag 1 and lag 4), and VQR backtests' p-values for the 10 competing models. The "Total" columns show the total number of rejections at the 5% significance level. As can be seen, the DCC-AL is the only framework which receives 0 rejections. The DCC-N model receives 5 rejections, while the DCC-t model gets rejected 3 times. This demonstrates the importance of the return distribution selection in parametric MGARCH models. The DCC-AL model, which is semi-parametric, shows advantage from this perspective. Comparing the DCC-AL model to the semi-parametric DCC models with QML, the DCC-AL still has better performance considering these backtests. These results again lend evidence on the validity and effectiveness of the DCC-AL in forecasting VaR.

Lastly, Bayer and Dimitriadis (2022) propose three versions of ES backtests named as Auxiliary, Strict and Intercept ES regression (ESR) backtests. These ESR backtests (two-sided), labeled as V1, V2, and V3, respectively, are also employed to backtest the ES forecasts from the 10 competing models.

For the implementation, we use the R package developed by the authors, which can be found at: `https://search.r-project.org/CRAN/refmans/esback/html/esr_`

Table 5: The 90% and 75% MCS results based on the quantile loss and joint loss function values, using both the R and SQ methods.

Model	R - quantile	SQ - quantile	R - joint	SQ - joint
90% MCS				
CARE-SAV	1	1	1	1
CARE-IG	1	1	1	1
ES-CAViaR-SAV	1	1	1	1
ES-CAViaR-IG	1	1	1	1
DCC-QML	1	1	1	0
DCC-N	1	1	1	0
DCC-clik-QML	1	1	1	1
DCC-clik-N	1	1	1	0
DCC-t	1	1	1	1
DCC-AL	1	1	1	1
75% MCS				
CARE-SAV	1	1	1	1
CARE-IG	1	1	1	1
ES-CAViaR-SAV	1	1	1	1
ES-CAViaR-IG	1	1	1	1
DCC-QML	1	1	1	0
DCC-N	1	1	1	0
DCC-clik-QML	1	0	1	0
DCC-clik-N	1	0	1	0
DCC-t	1	1	1	0
DCC-AL	1	1	1	1

Note:"0" in red represents models that are not in MCS.

`backtest.html`. The link also includes the details of the three versions of the backtests. As in Table 7, the three backtests return quite consistent results, with the DCC-AL model being not rejected. The models that get rejected at the 5% level by all three backtests are DCC-N, DCC-clik-N and DCC-t. Meanwhile, the semi-parametric DCC-QML and DCC-clik-QML do not get rejected.

Table 6: Summary of 2.5% VaR forecasts UC, CC, DQ1, DQ4, and VQR backtests' p-values and the total number of rejections at the 5% significance level.

Model	UC	CC	DQ1	DQ4	VQR	Total
CARE-SAV	76%	16%	9%	3%	46%	1
CARE-IG	95%	46%	41%	6%	0%	1
ES-CAViaR-SAV	90%	14%	8%	2%	68%	1
ES-CAViaR-IG	67%	83%	82%	1%	84%	1
DCC-QML	76%	3%	0%	0%	0%	4
DCC-N	4%	2%	0%	0%	0%	5
DCC-clik-QML	51%	18%	6%	0%	20%	1
DCC-clik-N	9%	10%	4%	0%	5%	2
DCC-t	13%	4%	0%	0%	11%	3
DCC-AL	67%	83%	81%	8%	42%	0

Table 7: Summary of three versions (V1, V2, V3) of ES backtests' p-values and the total number of rejections at the 5% significance level.

Model	V1	V2	V3	Total
CARE-SAV	83%	83%	70%	0
CARE-IG	95%	95%	76%	0
ES-CAViaR-Mult-SAV	92%	92%	78%	0
ES-CAViaR-IG	96%	96%	96%	0
DCC-QML	20%	20%	24%	0
DCC-N	0%	0%	1%	3
DCC-clik-QML	35%	35%	21%	0
DCC-clik-N	0%	0%	1%	3
DCC-t	2%	2%	1%	3
DCC-AL	87%	87%	71%	0

6.3 Hedging performance

In this section we evaluate the hedging performance of the DCC-AL model in producing minimum variance, VaR and ES portfolios and compare results with those obtained by standard DCC models taken as benchmarks. Minimum risk (VaR and ES) portfolios are computed using the PO procedure illustrated in Section 5 with the main difference that, since our interest is in hedging, at this stage we do not impose any constraint on the mean return target μ_0 of the optimized portfolio. On the other hand, minimum variance portfolios are, as usual, computed by numerically minimizing portfolio volatility with respect to the weights vector $\mathbf{w}$.

First, we present some empirical results to further demonstrate how the PO procedure proposed in Section 5 works and lend evidence on its validity. Figure 6 visualizes the simulated VaR and ES factors $\tilde{q}$ and $\tilde{c}$ (of size $N_{sim} \times 1$) generated from the in-sample data used for the first forecasting step. We can see that the simulated $\tilde{q}$ and $\tilde{c}$ series are centered around -2.1 and -2.7 respectively, with certain level of variation (noise) as expected. These two series are then used as the target variables for the OLS Step 4 of the PO procedure. As discussed, the OLS step aims to filter out the noise in the $\tilde{q}$ and $\tilde{c}$ series and potentially improve the approximation accuracy to the underlying VaR and ES factors.

With Steps 5 and 6 of the PO procedure, we can produce the optimized minimum VaR and ES portfolios over the out-of-sample period. Figure 7 presents the out-of-sample OLS predicted VaR and ES factors ($\mathbf{w}'\widehat{\boldsymbol{\beta}}_q$ and $\mathbf{w}'\widehat{\boldsymbol{\beta}}_c$ as in Equation (23), obtained through the OLS regression embedded in Step 4 of the procedure. In addition, relying on the out-of-sample optimized minimum VaR and ES portfolio weights, at each forecasting step we also re-estimate the DCC-AL model on the optimized portfolio returns, thus obtaining a new set of estimates for the underlying VaR and ES factors that we denote as $\hat{q}$ and $\hat{c}$, respectively. Here, we remind that $\hat{q}$ is directly estimated while $\hat{c}$ can be recovered from estimated DCC-AL parameters $\hat{q}$ and $\hat{\gamma}_0$ using Equation (6). Being based on the minimization of a strictly consistent loss, under correct specification of the underlying risk dynamics, $\hat{q}$ and $\hat{c}$ provide consistent estimates for the underlying q and c. Hence, a comparison of these estimates with the OLS based predictions at the optimum can be

used to check the accuracy of the latter approximation for the risk factors of the optimized portfolios. It is evident that the OLS predicted VaR factors and estimated $\hat{q}$ series are very close. Similar observations hold for ES. Therefore, such results demonstrate the validity of the PO procedure which can produce VaR and ES factors that can approximate the underlying q and c with good accuracy. In this way, the PO procedure, while still not assuming ellipticity, avoids the need of re-estimating the DCC-AL model at each portfolio configuration. Now we present additional empirical results, via comparing the DCC-AL and standard DCC models, to further demonstrate the effectiveness in PO of the DCC-AL model and of the proposed PO procedure.

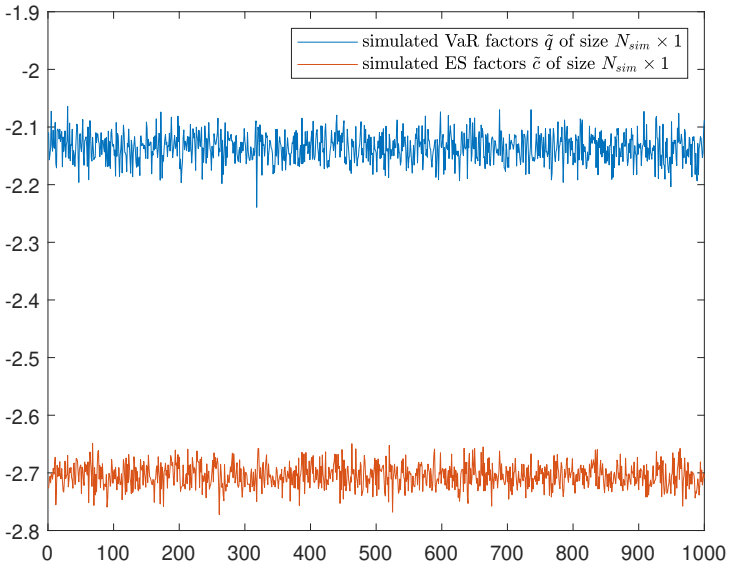

Figure 6: The simulated VaR and ES factors $\tilde{q}$ and $\tilde{c}$ of size $N_{sim} \times 1$ generated from the in-sample data used for the first forecasting step.

A summary of the PO optimization results is shown in Table 8. For all the competing models, including DCC-QML, DCC-t, DCC-clik-QML and the proposed DCC-AL, the minimum variance portfolios are firstly calculated. It is worth noting that, for DCC-t and DCC-N, due to the spherical nature of the chosen error distribution, minimum variance and minimum risk hedged portfolios are virtually identical. Minimal differences in the optimal portfolio allocation could only occur due to the (here negligible) conditional mean term appearing in the VaR and ES measures. Furthermore, DCC-(clik)-QML and DCC-

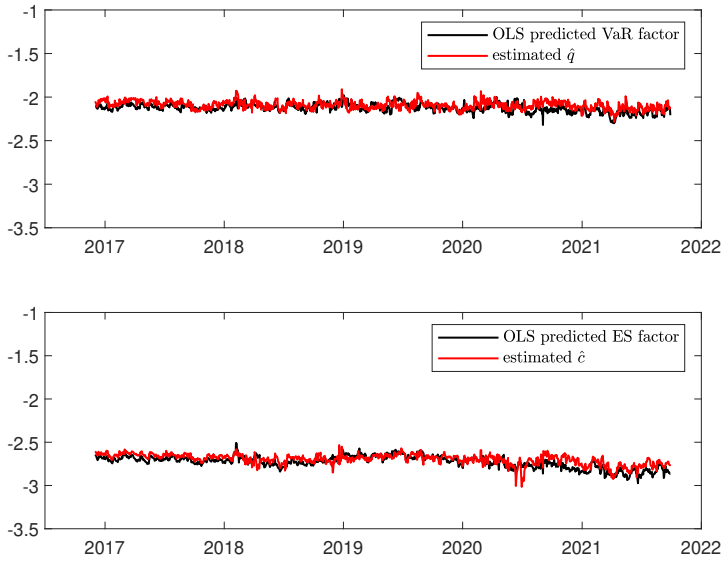

Figure 7: The top plot shows the OLS predicted VaR factor $\mathbf{w}'\widehat{\boldsymbol{\beta}}_q$ and the estimated $\hat{q}$ in DCC-AL for the optimized minimum VaR portfolio. The bottom plot shows the OLS predicted ES factor $\mathbf{w}'\widehat{\boldsymbol{\beta}}_c$ and the estimated $\hat{c}$ in DCC-AL for the optimized minimum ES portfolio. Whole forecasting sample (1213 days).

(clik)-N by construction return identical minimum variance allocations, being based on the same estimated conditional variance and covariance matrix.

In Tables 8 and 9, *DCC-AL min. VaR* and *DCC-AL min. ES* represent the optimized portfolios built from DCC-AL using the proposed PO procedure as shown in Section 5. *DCC-(clik)-QML min. VaR* and *DCC-(clik)-QML min. ES* are the optimized portfolios built from the DCC-(clik)-QML model using the PO procedure for minimum risk portfolios described in Section 5. The portfolio allocation exercise is based on the same rolling window scheme described in Section 6.2. At each time point t, models are estimated on a rolling window of 3000 observations extending up to time t and the predicted $\widehat{H}_{t+1}$ is used to optimize portfolio allocation for time $t+1$. We rule out short-selling operations by restricting portfolio weights to take values in [0,1].

The columns in Table 8 report the out-of-sample empirical variances, VaR and ES computed over the hedged portfolios for different models. Lastly, an equally weighted (EW) portfolio is also included as a benchmark. Compared to the benchmark EW case,

all the estimated DCC models yield substantial reductions in terms of out-of-sample empirical variance, VaR and ES. When focusing on the whole out-of-sample period, the hedging performance of the DCC-AL model is quite competitive comparing to the other DCC models, which also lends evidence on the validity of the DCC-AL and the proposed portfolio risk optimization process in Section 5. On the other hand, more relevant differences appear in a short term perspective when focusing on specific critical events such as the COVID-19 outbreak. Table 9 replicates the analysis in Table 8 focusing on a forecasting subsample including 100 consecutive days, from 7 February 2020 to 1 July 2020, in a neighbourhood of the COVID-19 outbreak. Compared to the empirical findings arising for the whole forecasting period, it shows that DCC-AL and DCC-clik-QML models are able to hedge ES risk more effectively than DCC-QML and DCC-t. For example, the empirical $\widehat{ES}(r_t^{(p)}, \alpha)$ produced by the DCC-AL and DCC-clik-QML minimum ES portfolios are (in absolute values) smaller than that of the DCC-QML and DCC-t. Comparing the DCC-AL and DCC-clik-QML models, we find that the two models return close performances. The empirical variance $\widehat{var}(r_t^{(p)})$ of the DCC-AL minimum variance portfolio is slightly smaller than that of the DCC-clik-QML minimum variance portfolio. Similarly, for the minimum VaR portfolios produced by the DCC-AL and DCC-clik-QML, the empirical VaR from DCC-AL is, in absolute value, also slightly smaller. The empirical ES values from DCC-AL and DCC-clik-QML minimum VaR portfolios are identical.

Summarizing the empirical evidence, what we have found is that in the long-term, when referring to the whole forecasting period, the selected performance indicators, based on unconditional variance, VaR and ES, do not reveal any striking differences between minimum variance and minimum risk portfolios. This suggests that the conditional distribution of returns is, for most days, likely to be close to some distribution falling within the elliptical family.

Nevertheless, in the short term, the selected optimal portfolio allocations can differ quite substantially across different strategies: that is minimum variance, VaR and ES. Evidence in this direction is indirectly provided by the analysis performed on the 100 days "COVID-19" subsample. More specifically, this intuition is confirmed by Figure 8, that reports the boxplots of the DCC-AL model optimized weights for all the assets included in our portfolio and for all allocation strategies for the 100 days "COVID-19" subsample, and

even more clearly by Figure 9, that reports differences between asset allocations yielded by minimum variance and minimum risk portfolios. Even if the median differences tend to stay close to zero, the tails of the boxplots reveal that there are more than a few time periods in which the weights assigned to specific portfolio assets can differ quite a lot across different allocation strategies.

Table 8: Descriptive statistics of out-of-sample returns $r_t^{(p)}$ for different portfolios with $\alpha = 2.5\%$. $\widehat{var}(r_t^{(p)})$: empirical variance; $\widehat{VaR}(r_t^{(p)}, \alpha)$: empirical VaR; $\widehat{ES}(r_t^{(p)}, \alpha)$: empirical ES. Whole forecasting sample (1213 days).

	$\widehat{var}(r_t^{(p)})$	$\widehat{VaR}(r_t^{(p)}, \alpha)$	$\widehat{ES}(r_t^{(p)}, \alpha)$
EW	1.5233	- 2.4682	- 4.1951
DCC-QML min. variance	1.0342	- 2.0135	- 3.6360
DCC-t min. variance	1.0310	- 1.9407	- 3.5127
DCC-QML min. VaR	1.0001	- 1.9722	- 3.5255
DCC-QML min. ES	1.0262	- 1.9953	- 3.5928
DCC-clik-QML min. variance	0.9788	- 1.9517	-3.2625
DCC-clik-QML min. VaR	0.9844	- 1.9992	- 3.3384
DCC-clik-QML min. ES	0.9818	- 1.9755	- 3.3407
DCC-AL min. variance	1.0037	- 2.0884	- 3.5945
DCC-AL min. VaR	0.9841	- 2.1251	- 3.5503
DCC-AL min. ES	0.9840	- 2.1077	- 3.5534

7 Conclusions

In this paper, we propose an innovative tail risk forecasting framework in a multivariate and semi-parametric setting. Through introducing variance targeting, the proposed framework is capable of modelling high dimensional return series parsimoniously and efficiently. In addition, a statistical procedure is designed to employ the proposed framework for optimising the portfolio VaR and ES. Compared to the state-of-the-art univariate semi-parametric models, the proposed framework delivers competitive risk forecasting

Table 9: Descriptive statistics of out-of-sample returns $r_t^{(p)}$ for different portfolios with $\alpha = 2.5\%$. $\widehat{var}(r_t^{(p)})$: empirical variance; $\widehat{VaR}(r_t^{(p)}, \alpha)$: empirical VaR; $\widehat{ES}(r_t^{(p)}, \alpha)$: empirical ES. Period: forecasting steps 801-900 (100 days) corresponding to the interval 7 February 2020 to 1 July 2020.

	$\widehat{var}(r_t^{(p)})$	$\widehat{VaR}(r_t^{(p)}, \alpha)$	$\widehat{ES}(r_t^{(p)}, \alpha)$
EW	11.2036	- 7.5400	- 11.5660
DCC-QML min. variance	6.6620	- 6.3077	- 9.3131
DCC-t min. variance	6.5782	- 6.4684	- 9.7652
DCC-QML min. VaR	6.3404	- 5.4028	- 9.1153
DCC-QML min. ES	6.5718	- 5.9916	- 9.3512
DCC-clik-QML min. variance	6.0174	- 5.9946	-7.0825
DCC-clik-QML min. VaR	6.0992	- 6.8327	- 7.2388
DCC-clik-QML min. ES	6.0821	- 6.7658	- 7.1063
DCC-AL min. variance	6.0097	- 5.9419	- 7.0825
DCC-AL min. VaR	5.8914	- 6.3003	- 7.0356
DCC-AL min. ES	5.8973	- 6.2114	- 7.1063

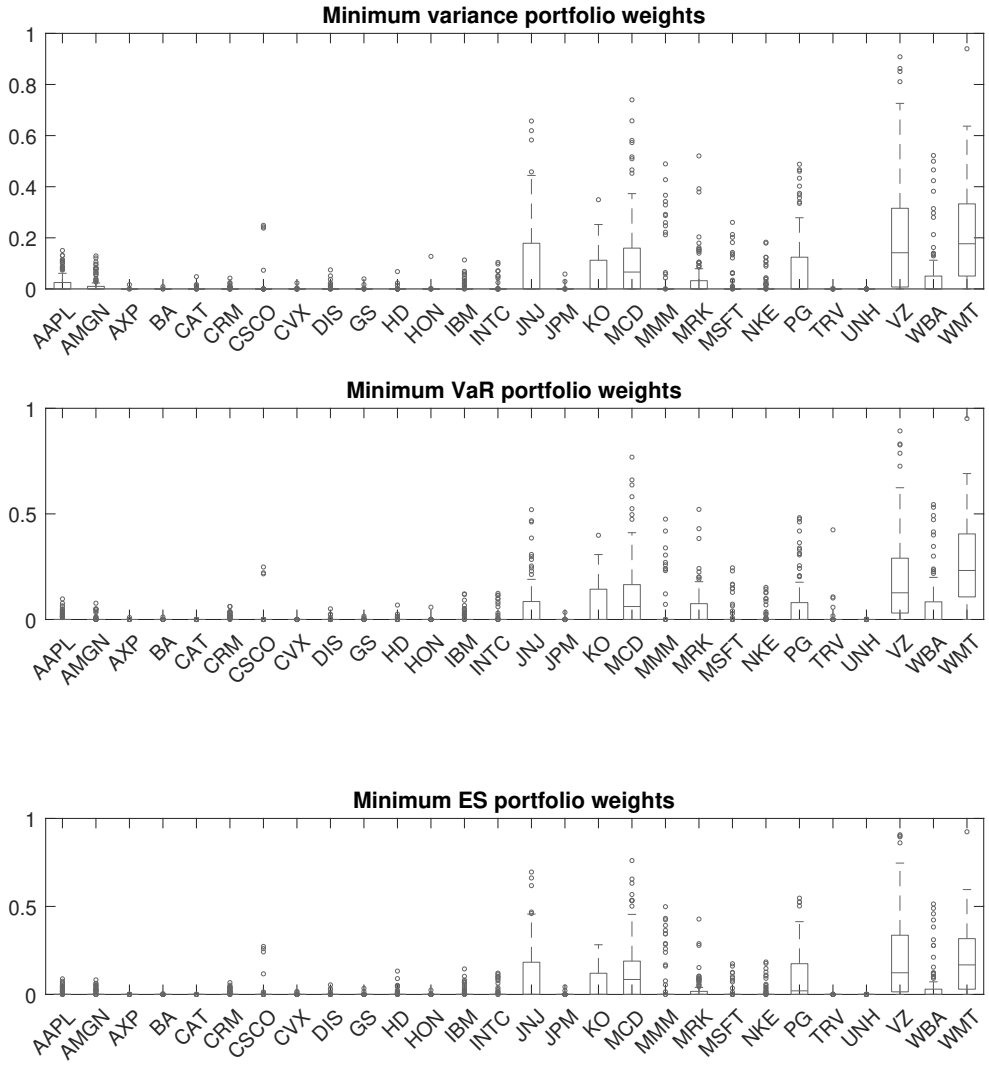

Figure 8: Boxplots of DCC-AL model minimum variance, VaR and ES portfolio weights. The x-axis shows the 28 components of the Dow Jones index. Period: forecasting steps 801-900 (100 days) corresponding to the interval 7 February 2020 to 1 July 2020.

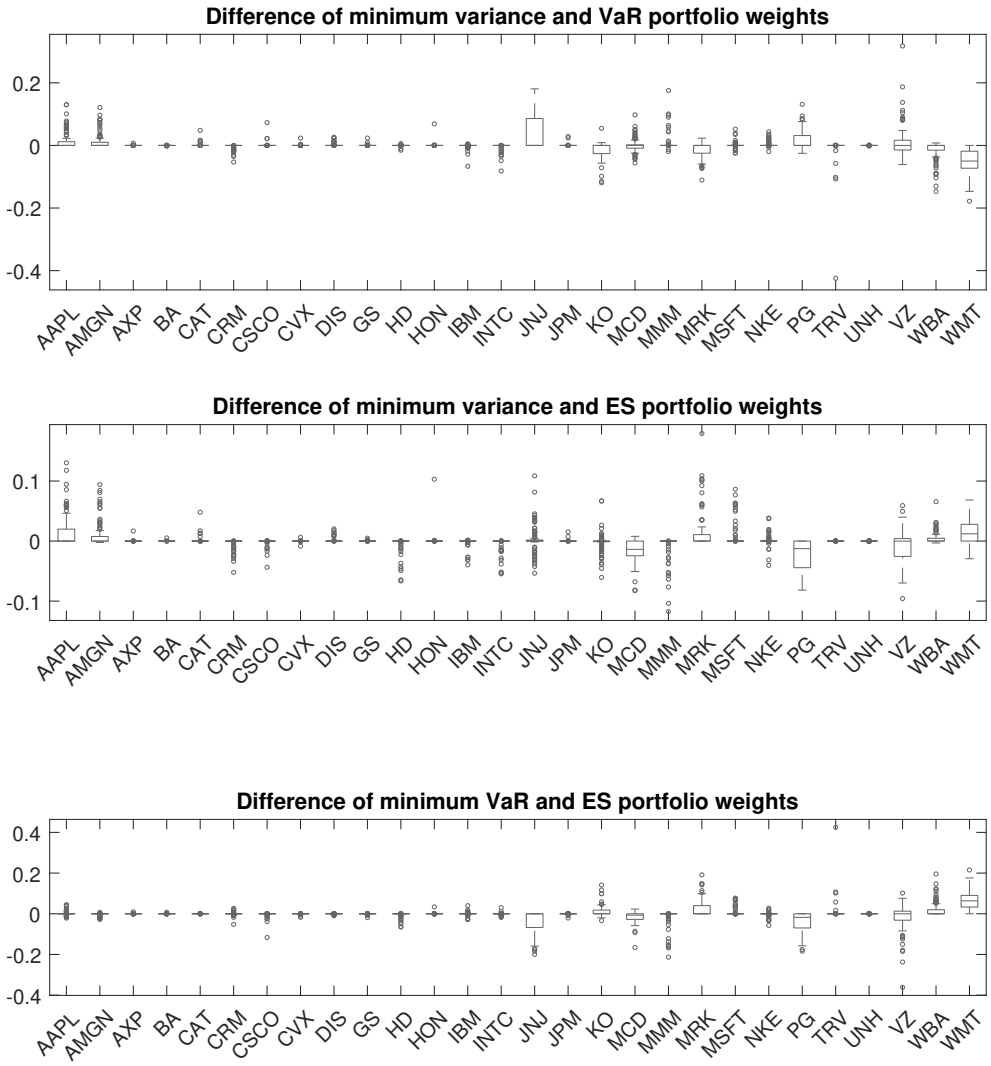

Figure 9: Boxplots of differences of DCC-AL model minimum variance and VaR, minimum variance and ES, and minimum VaR and ES portfolio weights. The x-axis shows the 28 components of the Dow Jones index. Period: forecasting steps 801-900 (100 days) corresponding to the interval 7 February 2020 to 1 July 2020..

performances and, in addition, it can be used for a wider range of applications such as portfolio optimisation and investigation of risk spillovers. Compared to existing DCC estimation approaches, our empirical application shows that the DCC-AL is able to deliver more accurate risk forecasts and hedging performances in line with those of DCC-clik-QML estimators (Pakel et al., 2021).

From a different viewpoint, it is worth noting that our framework also delivers a semi-parametric approach to the generation of robust estimates of DCC coefficients. Under this respect, in our empirical application we find that semi-parametric estimates of dynamic correlation parameters obtained from the DCC-AL model, as the DCC-clik estimator, are potentially not affected by the well known severe downward bias towards zero that, in even moderately large dimensions, typically characterizes estimates obtained through Gaussian QML (Pakel et al., 2021). A comparison of the statistical properties of our estimator with those of QML and clik-QML would be of great interest. However, this goes beyond the scope of this paper and is currently left for future research.

In addition to this, our projects for future research are focused on two main additional issues. First, we are extending the DCC-AL model via including the high frequency realized measures, to further improve its performance. A second issue of interest is related to a refinement of the minimum risk PO procedure discussed in Section 5. In particular, the OLS regression in Step 4 of the procedure could be replaced with a more flexible Neural Network regression.

References

Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics 31*(3), 282–299.

Artzner, P. (1997). Thinking coherently. *Risk*, 68–71.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Basel Committee on Banking Supervision (2013). *Fundamental review of the trading book: A revised market risk framework*. Bank for International Settlements.

Baur, D. G. (2013). The structure and degree of dependence: A quantile regression approach. *Journal of Banking & Finance 37*(3), 786–798.

Bauwens, L. and S. Laurent (2005). A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics 23*(3), 346–354.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics 21*(1), 79–109.

Bayer, S. and T. Dimitriadis (2022). Regression-based expected shortfall backtesting. *Journal of Financial Econometrics 20*(3), 437–471.

Bernardi, M., G. Gayraud, and L. Petrella (2015). Bayesian tail risk interdependence using quantile regression. *Bayesian Analysis 10*(3), 553–603.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, 841–862.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

DeMiguel, V., L. Garlappi, and R. Uppal (2007, 12). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies 22*(5), 1915–1953.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics 30*(2), 212–228.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized arch. *Econometric theory 11*(1), 122–150.

Engle, R. F. and S. Manganelli (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics 22*(4), 367–381.

Fernández, C. and M. F. Steel (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association 93*(441), 359–371.

Fissler, T. and J. F. Ziegel (2016). Higher order elicitability and Osband's principle. *The Annals of Statistics 44*(4), 1680–1707.

Francq, C. and J.-M. Zakoïan (2015). Risk-parameter estimation in volatility models. *Journal of Econometrics 184*(1), 158–173.

Francq, C. and J.-M. Zakoïan (2020). Virtual Historical Simulation for estimating the conditional VaR of large portfolios. *Journal of Econometrics 217*(2), 356–380.

Gaglianone, W. P., L. R. Lima, O. Linton, and D. R. Smith (2011). Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics 29*(1), 150–160.

Gerlach, R. and C. Wang (2020). Semi-parametric dynamic asymmetric laplace models for tail risk forecasting, incorporating realized measures. *International Journal of Forecasting 36*(2), 489–506.

Hall, P. and Q. Yao (2003). Inference in arch and garch models with heavy–tailed errors. *Econometrica 71*(1), 285–317.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Koenker, R. and J. A. F. Machado (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association 94*(448), 1296–1310.

Kupiec, P. H. et al. (1995). *Techniques for verifying the accuracy of risk measurement models*, Volume 95. Division of Research and Statistics, Division of Monetary Affairs, Federal ....

Merlo, L., L. Petrella, and V. Raponi (2021). Forecasting var and es using a joint quantile regression and its implications in portfolio allocation. *Journal of Banking & Finance 133*, 106248.

Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle (2021). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics 39*(3), 652–668.

Patton, A. J., J. F. Ziegel, and R. Chen (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of Econometrics 211*(2), 388 – 413.

Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*, 21–41.

Taylor, J. W. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics 6*(2), 231–252.

Taylor, J. W. (2019). Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric laplace distribution. *Journal of Business & Economic Statistics 37*(1), 121–133.

White, H., T.-H. Kim, and S. Manganelli (2015). Var for var: Measuring tail dependence using multivariate regression quantiles. *Journal of Econometrics 187*(1), 169–188.