Physica A 401 (2014) 330–343

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Semi-nonparametric VaR forecasts for hedge funds during the recent crisis

Esther B. Del Brio [a], Andrés Mora-Valencia [b], Javier Perote [c,*]

[a] Faculty of Economics and Business, Department of Business, University of Salamanca, Spain
[b] School of Economics and Finance, Department of Finance, EAFIT University, Colombia
[c] Faculty of Economics and Business, Department of Economics, University of Salamanca, Spain

HIGHLIGHTS

- Traditional VaR measures fail to capture risk in highly volatile scenarios.
- Hedge funds demand accurate techniques for risk management.
- Semi-nonparametric methods and EVT capture risk accurately when volatility is high.
- Gram–Charlier copula incorporates non-linear dependences in portfolio risk assessment.
- Two-step estimation of short Gram–Charlier series simplifies the VaR computation.

ARTICLE INFO

Article history:
Received 12 July 2013
Received in revised form 29 October 2013
Available online 24 January 2014

Keywords:
Hedge funds
Value-at-risk
Backtesting
Extreme value theory
Gram–Charlier series
Copulas

ABSTRACT

The need to provide accurate value-at-risk (VaR) forecasting measures has triggered an important literature in econophysics. Although these accurate VaR models and methodologies are particularly demanded for hedge fund managers, there exist few articles specifically devoted to implement new techniques in hedge fund returns VaR forecasting. This article advances in these issues by comparing the performance of risk measures based on parametric distributions (the normal, Student's t and skewed-t), semi-nonparametric (SNP) methodologies based on Gram–Charlier (GC) series and the extreme value theory (EVT) approach. Our results show that normal-, Student's t- and Skewed t- based methodologies fail to forecast hedge fund VaR, whilst SNP and EVT approaches accurately success on it. We extend these results to the multivariate framework by providing an explicit formula for the GC copula and its density that encompasses the Gaussian copula and accounts for non-linear dependences. We show that the VaR obtained by the meta GC accurately captures portfolio risk and outperforms regulatory VaR estimates obtained through the meta Gaussian and Student's t distributions.

1. Introduction

The literature on hedge funds management has undergone a major impulse in the last decades (see Agarwal and Naik [1] for an extensive review). It was not until the collapse of Long Term Capital Management (LTCM) in 1998 and the Nasdaq crash (2000), however, that hedge fund investors were aware of the need for implementing risk management techniques. Jorion [2] was the first to apply the value-at-risk (VaR) methodology to hedge fund returns for the special case of LTCM and, more recently, Gupta and Liang [3] analyzed capital requirements of nearly 1500 hedge funds.

* Correspondence to: Universidad de Salamanca. Campus Miguel de Unamuno (Edif. FES). 37007 Salamanca, Spain. Tel.: +(34) 923294400 Ext. 3512.
E-mail address: perote@usal.es (J. Perote).

0378-4371/$ – see front matter
http://dx.doi.org/10.1016/j.physa.2014.01.037

VaR techniques are based on quantile methods [4], but among them there is not a consensus on the more appropriate methodology, despite the fact that a wrong choice of VaR model (or an incorrect parameter estimation of the model) might yield a hedge fund failure, especially in a crisis scenario. Alexander and Sarabia [5] proposed an interesting solution in which quantile estimates are adjusted for model risk, relative to an institutional benchmark, but still traditional VaR techniques are employed by hedge fund managers. Specifically, the former VaR models, the so-called parametric VaR, have been widely criticized since they assume that returns follow a normal distribution and this hypothesis involves risk underestimation in highly volatile scenarios. For this reason, alternative skewed and heavy-tailed distributions [6–8], or volatility models [9–12] have been proposed.

Hedge fund returns, as well as most asset returns, also exhibit time-varying variances, heavy-tails and skewness, and thus different authors have investigated into these issues. For example, Refs. [13,14] found that a large proportion of the hedge fund return volatility can be explained by market-related factors, such as equity market indices, book-to-market, 'momentum' or commodity indices. Other authors focused on capturing tail risk by implementing the extreme value theory (EVT) approach [15] or the semi-nonparametric (SNP) expansions based on the Cornish–Fisher approximation [16]. We compare the VaR performance of both methodologies, as well as other parametric distributions, in a VaR backtesting framework [17]. As far as we know, such a comparison has not previously been studied for hedge fund returns. A related contribution was conducted by Güner et al. [18], which also implemented a two-step backtesting procedure: In the first step they estimated an ARMA(1, 1)–GARCH(1, 1) model with Student-t distributed innovations and in the second step fitted an α-stable distribution to the standardized residuals. They concluded that the conditional stable model performs well in predicting VaR.

The current paper implements a similar backtesting procedure for two different hedge funds (Equity Hedge and Event Driven strategies) and compares the VaR forecasting performance obtained by the normal distribution (benchmark) to four natural candidates that account for stock returns heavy tails: The Student's t, a skewed variant of the Student's t distribution [19] (skewed-t hereafter), the EVT approach [20] and the SNP approach based on the Gram–Charlier (GC) density, which is an expansion around the normal density allowing for skewness and excess kurtosis [21–24]. The theoretical advantages of the SNP methodology rely on the asymptotic properties of the GC (Type A) series to approximate any frequency function under certain regularity conditions. For this reason this approach has been applied in different sciences, e.g. physics [25] or astronomy [26], but it has scarcely been used for financial risk forecasting purposes. Furthermore, we analyze the non-linear dependence between the two hedge fund indices using the concept of copulas [27]. For this purpose we derive an explicit formula for GC copula that encompasses the Gaussian copula and incorporates non-linear dependencies through the Hermite polynomial structure. We compare the VaR performance of this meta GC to alternative meta distributions, including the meta Gaussian and meta t, for an equally weighted portfolio. This is the main contribution of this paper, as well as the application to forecast hedge funds during the recent subprime and sovereign debt crises.

The univariate results show that the both normal- and Student's t-based VaR forecasts are inadequate for high confidence levels and/or high volatility periods, although the skewed-t performs better than the Student's t. On the other hand, GC and EVT produce accurate VaR forecasts in these contexts. Therefore the risk management and hedging strategies should implement these methods; particularly the SNP approach that is very flexible to improve VaR measures and is not sensitive to arbitrary decisions as the threshold choice in the EVT methodology. These results may be extended to the multivariate framework, where the GC copula and the corresponding meta GC seem to adequately assess portfolio risk for the traditionally employed confidence levels.

The rest of the paper is organized as follows: Section 2 reviews the VaR methodology and the theoretical models. Section 3 shows the backtesting technique and the VaR forecasting performance of the models for the univariate case. Section 4 introduces the GC copula and Section 5 compares portfolio VaR quantified by such a copula to alternative meta distributions. Section 6 summarizes the main results of the article.

2. Univariate case

2.1. VaR methodology

Stock returns, and consequently hedge fund returns, usually present a small predictable component in the conditional mean that has been traditionally modeled according to simple ARMA structures. On the other hand, squared returns exhibit particular dynamics (conditional heteroskedasticity, volatility clustering and long memory) that have extensively been studied since Engle [28] and Bollerslev [29] introduced ARCH and GARCH models. In this article, we focus on VaR performance due to the distributional hypotheses described in next section, and thus our model incorporates a traditional ARMA(1, 1)–GARCH(1, 1) for conditional mean and variance, following the common use in most risk management applications.

We present the basic model for log returns, r_t, in Eqs. (1)–(4) below.

$$r_t = \mu_t + \sigma_t z_t, \tag{1}$$

$$\mu_t = \varphi + \phi \mu_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t, \tag{2}$$

$$z_t = \varepsilon_t / \sigma_t, \qquad z_t \sim G(z_t | \Omega_{t-1}), \tag{3}$$

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{4}$$

where μ_t and σ_t^2 denote the conditional mean and conditional variance, respectively, of r_t taken on the set of information available up to time $t-1$, denoted as Ω_{t-1}, and z_t is a martingale difference conditionally distributed according to certain density function $G(z_t|\Omega_{t-1})$ and satisfying $E(z_t|\Omega_{t-1}) = 0$ and $E(z_t^2|\Omega_{t-1}) = 1$.

2.1.1. Parametric VaR

In the former model, the estimated VaR with a confidence level α is computed as the estimated α-quantile, $\hat{q}_\alpha(z)$, of the assumed $G(z_t|\Omega_{t-1})$ distribution. Therefore, the predicted VaR for the variable r at the time horizon $t+1$ and with confidence level α is given in Eq. (5).

$$\mathrm{VaR}_{t+1}^{\alpha} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\hat{q}_\alpha(z), \tag{5}$$

where $\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$ are the predictions for the mean and standard deviation conditioned on the available information at time t, Ω_t, based on the ARMA–GARCH model described in Eqs. (2) and (4).

Note that under the Gaussian assumption $\hat{q}_\alpha(z)$ is fixed for a particular α (given the estimates for $\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$) but this quantile depends on other density parameters if non-Gaussian distributions are considered.

2.1.2. Extreme value theory

VaR can also be computed by the EVT methodology through two different approximations: block maxima and POT. We implement the latter method following the two-step procedure proposed in Ref. [20].

In the first step, the ARMA(1, 1)–GARCH(1, 1) model is fitted by quasi maximum likelihood (QML) and in the second step the so-obtained standardized residuals are used to implement the peaks over threshold (POT) methodology using the 10% of the tail of the distribution as threshold. Thus, $\hat{q}_\alpha(Z)$ is given by

$$\hat{q}_\alpha(Z) = u + \frac{\beta}{\xi}\left(\left(\frac{1-\alpha}{N_u/n}\right)^{-\xi} - 1\right), \tag{6}$$

where u is the estimated threshold, N_u is the number of exceedances over the threshold, n is the sample size (thus N_u/n is a non-parametric estimator of the empirical distribution tail) and β and ξ are the scale and shape parameters of the generalized Pareto distribution (GPD). The cumulative distribution function (cdf) of the GPD distribution [30–32] is given by

$$F(x) = \begin{cases} 1-(1+\xi x/\beta)^{-1/\xi}, & \xi \neq 0, \\ 1-\mathrm{e}^{-x/\beta}, & \xi = 0. \end{cases} \tag{7}$$

The weakness of the EVT lies on the threshold selection, which involves a tradeoff between bias and variance in the estimation of the parameters, especially the shape parameter ξ. This parameter may be estimated by bootstrapping or graphical techniques, but there is not an optimal method to choose the appropriate threshold. According to Ref. [20], we choose 10% of the data in the tail of the distribution as an accurate threshold.

2.2. Alternative distributions

2.2.1. Parametric distributions

Since the early work of Mandelbrot [33] the normality assumption of stock returns is deemed inappropriate (see Refs. [34,35] for a revision of the asset returns empirical regularities). However, we implement the standard normal distribution as a benchmark, whose well-known probability density function (pdf) in given below.

$$\phi(z) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{z^2}{2}}. \tag{8}$$

To account for the leptokurtosis implied by the heavy-tails, the Student's t pdf in Eq. (9) seems to be the more natural candidate to use.

$$t(z) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{z^2}{n-2}\right)^{-\frac{\nu+1}{2}}, \tag{9}$$

Γ being the gamma function and ν the degrees of freedom parameter.

For the sake of capturing asymmetries different skewed versions of the Student's t have been employed [36–41]. Particularly, we implement the skewed-t pdf in Ref. [36]:

$$g(z) = \begin{cases} -\dfrac{2}{\gamma+\frac{1}{\gamma}}t(\gamma z) & \text{for } x < 0, \\ \dfrac{2}{\gamma+\frac{1}{\gamma}}t\left(\dfrac{z}{\gamma}\right) & \text{for } x \geq 0, \end{cases} \tag{10}$$

where γ is the shape parameter, which incorporates the skewness, and $t(z)$ is the Student's t pdf in Eq. (9).

2.2.2. Semi-nonparametric approach

We compare the VaR forecasting performance of the parametric distributions (Section 2.2.1), the EVT (Section 2.1.2) and the SNP approach based on the GC approximation. The validity of GC expansion is due to Charlier [23] and Edgeworth [24] who proved that, under weak regularity conditions [42], any frequency function, $p(z)$, could be expanded on an (infinite) series of derivatives of the standard normal density, $\phi(z)$, (see Ref. [21, pp. 168–172]) as follows:

$$p(z) = \left(1 + \sum_{s=1}^{\infty} k_s H_s(z)\right)\phi(z), \tag{11}$$

where $k_s = \frac{1}{s!}\int_{-\infty}^{\infty} H_s(z)\phi(z)\mathrm{d}z$ and $H_s(z)$ is the Hermite polynomial (HP) of order s, which can be defined in terms of the derivatives of $\phi(z)$ as,

$$\frac{\mathrm{d}^s\phi(z)}{\mathrm{d}z^s} = (-1)^s H_s(z)\phi(z). \tag{12}$$

In particular, the first eight HP are:

$$H_1(z) = z, \tag{13}$$

$$H_2(z) = z^2 - 1, \tag{14}$$

$$H_3(z) = z^3 - 3z, \tag{15}$$

$$H_4(z) = z^4 - 6z^2 + 3, \tag{16}$$

$$H_5(z) = z^5 - 10z^3 + 15z, \tag{17}$$

$$H_6(z) = z^6 - 15z^4 + 45z^2 - 15, \tag{18}$$

$$H_7(z) = z^7 - 21z^5 + 105z^3 - 105z, \tag{19}$$

$$H_8(z) = z^8 - 28z^6 + 210z^4 - 420z^2 + 105. \tag{20}$$

For empirical purposes the asymptotic expansion in Eq. (11) has to be truncated at the n-th term and then the GC density is defined in Eq. (21),

$$f(z, \boldsymbol{d}) = \left(1 + \sum_{s=1}^{n} d_s H_s(z)\right)\phi(z), \tag{21}$$

where $\boldsymbol{d}' = (d_1, \ldots, d_n) \in \mathbb{R}^n$ denotes the vector of parameters. It can be proved that d_1 and d_2 capture mean and variance, respectively, the odd parameters incorporate skewness and the even parameters the tail behavior. In particular, if $d_1 = d_2 = 0$ then d_4 represents the excess kurtosis.

Despite the fact that the main advantage of this distribution lies in its flexibility to incorporate different shapes (including jumps in the probabilistic mass) with a variable number of parameters, most of the empirical studies truncate the expansion in $n = 4$ and employ only d_3 and d_4 to account for skewness and excess kurtosis, respectively. For comparison purposes, we initially estimate VaR taking into account the GC expansion truncated at the fourth order estimated via method of moments (MM) and maximum likelihood (ML). We follow the procedure proposed in Ref. [43] and estimate the density parameters in two steps. In the first step, we estimate the conditional mean and variance by QML and obtain the standardized residuals and, in the second step, we estimate the d_s parameters for the standardized residuals by either MM or ML. The former method is based on the following relations among density parameters and (estimated) moments:

$$\hat{d}_1 = \hat{\mu}_1, \tag{22}$$

$$\hat{d}_2 = \frac{1}{2}(\hat{\mu}_2 - 1), \tag{23}$$

$$\hat{d}_3 = \frac{1}{6}(\hat{\mu}_3 - 3\hat{\mu}_1), \tag{24}$$

$$\hat{d}_4 = \frac{1}{24}(\hat{\mu}_4 - 6\hat{\mu}_2 + 3), \tag{25}$$

$$\hat{d}_5 = \frac{1}{120}(\hat{\mu}_5 - 10\hat{\mu}_3 + 15\hat{\mu}_1), \tag{26}$$

$$\hat{d}_6 = \frac{1}{720}(\hat{\mu}_6 - 15\hat{\mu}_4 + 45\hat{\mu}_2 - 15), \tag{27}$$

$$\hat{d}_7 = \frac{1}{5040}(\hat{\mu}_7 - 21\hat{\mu}_5 + 105\hat{\mu}_3 - 105), \tag{28}$$

$$\hat{d}_8 = \frac{1}{40\,320}(\hat{\mu}_8 - 28\hat{\mu}_6 + 210\hat{\mu}_4 - 420\hat{\mu}_2 + 105), \tag{29}$$

$\hat{\mu}_s = \frac{1}{n}\sum_{t=1}^{T} z_t^s$, being a consistent estimate of the non-central moment of order s.

Table 1
Descriptive statistics for hedge fund returns.

	Equity Hedge	Event Driven
Mean	0.0004	0.0108
Median	0.0351	0.0341
Standard deviation	0.4317	0.3113
Variance	0.1864	0.0969
Excess kurtosis	5.3728	12.1327
Skewness	−0.7835	−1.1468
Range	5.5797	5.6929
Minimum	−3.0204	−3.1182
Maximum	2.5593	2.5747

Furthermore, we also consider a second GC model expanded up to the eighth term that is computed by ML (second step). For this model, we identify the "optimal" truncation order using the Akaike Information Criterion (AIC). Therefore, within the GC family we consider three cases: GC4-MM (GC model expanded to the 4th term estimated by MM), GC4-ML (GC model expanded to the 4th term estimated by ML) and GC8-ML (GC model expanded to the 8th term estimated by ML).

Most of the financial literature about SNP methodologies is devoted to pricing derivatives [44,45] and only few articles use the GC expansion for VaR forecasting purposes [46,47]. Furthermore, most authors [48–50] implement positive transformations of the GC distribution by squaring the weighted sum of HP in Eq. (21), since the truncated GC series does not necessarily ensure a positive domain for the whole parametric space. Nevertheless, we decided not to implement these positive transformations since the resulting densities are too complex for backtesting applications and the consistent estimation (ML or MM) are sufficient conditions to achieve positivity.

3. Empirical application: VaR performance of univariate hedge fund series

3.1. Data

We use daily data from two main hedge funds based on different strategies: Equity Hedge and Event Driven. The strategy of the former mainly consists of large and short positions in equity. Event driven strategy involves capitalization on anomalies related to corporate transactions. Data ranged from July 2006 (one year before the date when Bear Stearns hedge funds reported massive losses triggering the subprime crisis) to first quarter of 2013. Table 1 presents the descriptive statistics for continuously compounded returns of these series, defined as $r_t = 100 \log(P_t/P_{t-1})$ where P_t represents the corresponding hedge fund index. The statistics show that the selected hedge fund returns present negative skewness and high excess kurtosis as is usually found in the related literature.

3.2. In-sample results

Fig. 1 displays the autocorrelation function (ACF) of the return series (upper graphs) and the absolute return series (lower graphs) using the total sample. The ACF of the return series show that there is a slightly autoregressive structure in the data and thus either an AR(1) or ARMA(1, 1) structure might be identified. The ACF of the absolute return series reveal a strong presence of conditional heteroskedasticity in the data that can be adequately captured by a GARCH(1, 1) process.

Next we proceed to select among the three plausible models for conditional mean – white noise, AR(1) and ARMA(1, 1) – according to accuracy criteria. Table 2 gathers the log-likelihood values of these three models combined with a GARCH(1, 1) for modeling conditional variance and under different distributional hypotheses, either normal, Student's t or skewed-t. The results show clear evidence in favor of the autoregressive models but it does not strongly support the ARMA(1, 1) versus the AR(1) model. We choose the ARMA(1, 1) since it has slightly higher log-likelihood values and nests the AR(1).

Table 3 presents the Maximum Likelihood (ML) estimates of the parameters of the ARMA(1, 1)–GARCH(1, 1) model under the five distributional hypotheses. p-values for testing the significance of every parameter are given in parentheses. These values show that the GARCH(1, 1) parameters are statistically significant but not all the parameters of the ARMA(1, 1) are statistically different from zero, especially for Equity Hedge Fund returns. This fact is in line with the 'small predictable component of the conditional mean' stylized fact featured by stock returns. Moreover GARCH(1, 1) processes exhibit persistent behavior since they are close to the non-stationarity (i.e., $\alpha + \beta$ is close to one). This fact captures the 'long memory' or the 'persistence of conditional variance' usually found in this type of data. Consequently, we decided to use the ARMA(1, 1)–GARCH(1, 1) model when implementing the backtesting technique to investigate the performance of the different distributional hypothesis on VaR computation.

Table 3 (Panels B and C) also includes the estimates for the shape parameter (degrees of freedom) and the skew parameter of the Student's t distributions. Both parameters are significant, which reflects that the distribution is leptokurtic and asymmetric. Furthermore, the parameters of the EVT and GC distributions are displayed in Panels D and E, F and G, respectively. In these cases two-step estimation was implemented (i.e. returns were filtered according to the ARMA(1, 1)–GARCH(1, 1) estimated in the first step by QML). The shape parameter for EVT, ξ, is not significantly different from

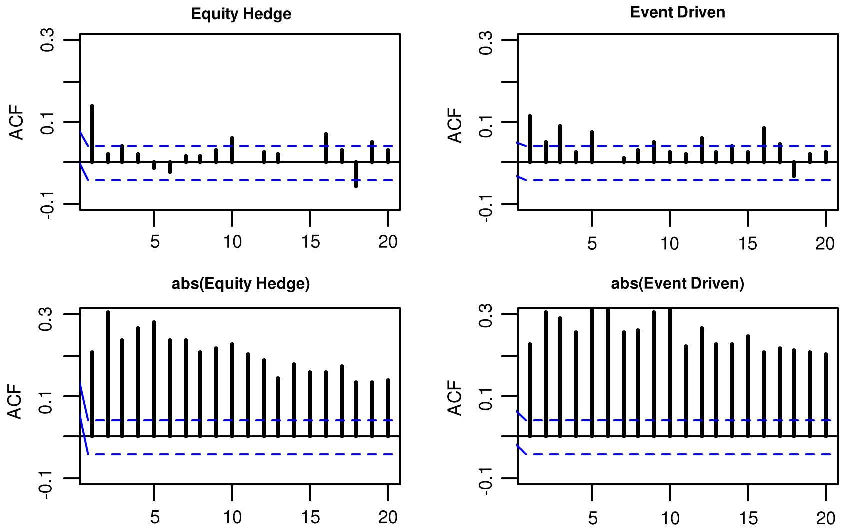


Fig. 1. Autocorrelation functions of hedge fund returns levels and absolute values.

Table 2
Log-likelihood for different conditional mean–variance models and under different distributional assumptions.

	Equity Hedge	Event Driven
	Panel A: Normal	
GARCH(1, 1)	−946.807	−80.404
AR(1)-GARCH(1, 1)	−921.565	−65.201
ARMA(1, 1)–GARCH(1, 1)	−921.447	−62.336
	Panel B: Student's t	
GARCH(1, 1)	−909.326	−31.577
AR(1)-GARCH(1, 1)	−881.268	−12.582
ARMA(1, 1)–GARCH(1, 1)	−880.946	−9.727
	Panel C: Skewed Student's t	
GARCH(1, 1)	−891.788	−22.084
AR(1)-GARCH(1, 1)	−868.763	−5.132
ARMA(1, 1)–GARCH(1, 1)	−868.747	−3.283

zero, which means that, after filtering the returns by an ARMA(1, 1)–GARCH(1, 1) model, the standardized residuals exhibit medium-tailed distributions (when $\xi = 0$, the GPD takes the shape of an exponential distribution). Regarding the GC density parameters, d_3 and d_4 confirm the presence of (negative) skewness and leptokurtosis. Nevertheless, not all the parameters of the larger expansion (Panel G) are significantly different from zero. In particular, d_3 seems to capture the whole skewness pattern of the density (i.e. further odd parameters seem to be unnecessary) but the inclusion of other even parameters seem to be useful to account for extreme values. However, the flexibility of GC densities represents a clear advantage in order to incorporate different features along the backtesting procedure implemented in Section 3.3.

3.3. Backtesting

To test the validity of the distributional assumptions of stock returns, (normal, Student's t, skewed-t, GC and EVT) the historical series, $r_1, \ldots, r_m$, are compared with the VaR_t^{α} predicted for the day $t = \{n+1, \ldots, m\}$ by using a time window of the n previous days. We consider a time window of 500 days for computing every one-step-ahead VaR prediction (the first in-sample rolling window being July 2004–July 2006) and a total period of 1750 days as the backtesting or out-of-sample period (July 2006–1st quarter of 2013).

The predicted VaR is compared to the observed return at 99% confidence level. Therefore, when calculating VaR at 99% we expect that in 1% of the backtesting days, the negative returns exceed VaR predictions. These values are referred to as 'violations' or 'exceptions'. More specifically, if I_t is the indicator defined in Eq. (30),

$$I_t := 1_{\{r_{t+1} > \mathrm{VaR}_t^{\alpha}\}}, \tag{30}$$

a 'violation' occurs when $r_{t+1} > \mathrm{VaR}_t^{\alpha}$, and then the indicator function takes value 1. Otherwise, I_t takes value 0 (whenever $r_{t+1} \leq \mathrm{VaR}_t^{\alpha}$). Therefore, if the VaR methodology is adequate, it is expected that the violation indicator function values

Table 3
Estimation of the parameters of an ARMA(1, 1)–GARCH(1, 1).

	Equity Hedge	Event Driven
	Panel A: Normal	
φ	0.0036 (0.5962)	0.0131 (0.0179)
ϕ	0.2296 (0.1193)	0.5626 (0.0006)
θ	−0.0724 (0.6384)	−0.4510 (0.0114)
ω	0.0039 (0.0001)	0.0014 (0.0001)
α	0.1059 (0.0000)	0.0984 (0.0000)
β	0.8722 (0.0000)	0.8861 (0.0000)
	Panel B: Student's t	
φ	0.0036 (0.5837)	0.0168 (0.0068)
ϕ	0.2699 (0.0582)	0.5433 (0.0005)
θ	−0.1140 (0.4507)	−0.4251 (0.0118)
ω	0.0036 (0.0007)	0.0012 (0.0007)
α	0.1115 (0.0000)	0.0939 (0.0000)
β	0.8703 (0.0000)	0.8918 (0.0000)
ν	7.5050 (0.0000)	7.0846 (0.0000)
	Panel C: Skewed Student's t	
φ	0.0036 (0.6462)	0.0161 (0.0057)
ϕ	0.1130 (0.5451)	0.4763 (0.0039)
θ	0.0316 (0.8685)	−0.3628 (0.0401)
ω	0.0039 (0.0004)	0.0012 (0.0005)
α	0.1093 (0.0000)	0.0898 (0.0000)
β	0.8705 (0.0000)	0.8939 (0.0000)
ν	7.4784 (0.0000)	7.4184 (0.0000)
γ	0.8677 (0.0000)	0.8937 (0.0000)
	Panel D: EVT	
ξ	−0.0268 (0.3260)	0.0503 (0.2388)
β	0.4818 (0.0000)	0.4616 (0.0000)
	Panel E: GC4MM	
d_3	−0.0762 (0.0000)	−0.0744 (0.0000)
d_4	0.0836 (0.0000)	0.1061 (0.0000)
	Panel F: GC4ML	
d_3	−0.0511 (0.0000)	−0.0352 (0.0006)
d_4	0.0338 (0.0000)	0.0362 (0.0000)
	Panel G: GC8ML	
d_3	−0.0585 (0.0000)	−0.0468 (0.0012)
d_4	0.0392 (0.0000)	0.0543 (0.0000)
d_5	−0.0059 (0.1295)	−0.0064 (0.1281)
d_6	0.0036 (0.0596)	0.0076 (0.0027)
d_7	−0.0013 (0.0288)	−0.0011 (0.0822)
d_8	0.0007 (0.0009)	0.0006 (0.0174)

p-values for the t-test in parentheses.

behave as realizations of independent and identically distributed (i.i.d.) Bernoulli experiments with success probability equal to $1-\alpha$, i.e. $\sum_{t=1}^{m} I_t \sim \text{Bin}(m, 1-\alpha)$. Thus, the null hypothesis that 'the model adequately estimates VaR' can be tested by a straightforward one-sided binomial hypothesis test (see Ref. [51] for different VaR forecasting tests). The alternative hypothesis suggests that the method underestimates or overestimates the VaR calculation depending on the number of expected violations.

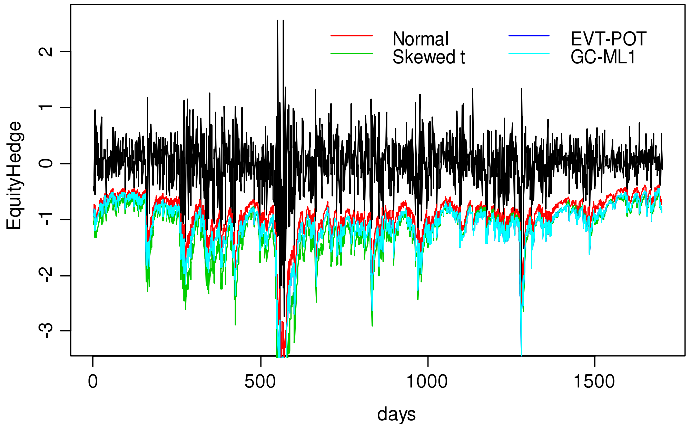
Table 4 displays the number of exceptions and the p-value for the binomial test (in parentheses) for the seven models (ARMA–GARCH-normal, ARMA–GARCH-t, ARMA–GARCH-skewed-t, ARMA–GARCH-EVT, ARMA–GARCH-GC4-MM, ARMA–GARCH-GC4-ML, ARMA–GARCH-GC8-ML) at 99% confidence level and for the two hedge fund returns. These results show that normal and Student's t significantly under predict the VaR for both series, whilst the method based on skewed-t is also rejected for the Equity Hedge strategy. Nonetheless, the EVT approach, and GC models perform well for capturing risk in both hedge funds since these methods focus on modeling the extreme values. These results are consistent with the evidence usually found in stock returns VaR performance and extend it to hedge fund returns calculation [52]. Furthermore, we also find that the larger SNP expansions do not seem to improve the forecasting performance of GC densities and that the MM estimation procedures do not provide more accurate VaR forecasts compared to ML methods.

Finally, the hedge fund returns and their corresponding forecasted VaR at 99% confidence is plotted in Fig. 2for both strategies. It is clear that the normal (red line) is the distribution that produces the less conservative VaR measures (lower values) and the EVT (dark blue line) and GC (light blue line) the ones that involve higher VaR predictions.

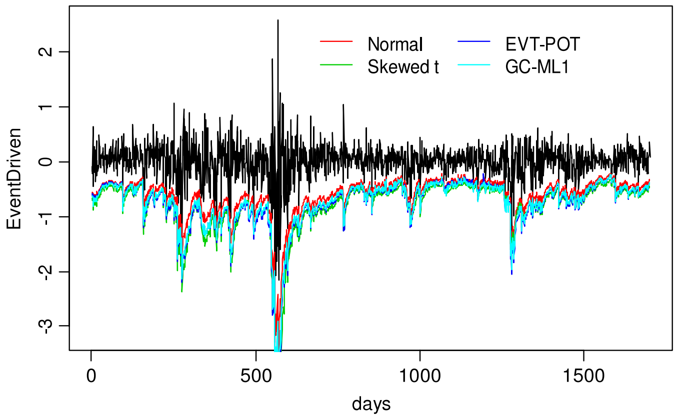
Table 4
Backtesting for hedge fund returns.

VaR 99% Expected number of exceptions = 17 Data: July 2006–first quarter of 2013		
1750 days	Equity Hedge	Event Driven
ARMA–GARCH-normal	38 (0.0000)	38 (0.0000)
ARMA–GARCH-*t*	25 (0.0398)	29 (0.0047)
ARMA–GARCH-skewed-*t*	9 (0.0256)	17 (0.5640)
ARMA–GARCH-EVT	17 (0.5640)	20 (0.2628)
ARMA–GARCH-GC4-MM	11 (0.0836)	14 (0.2796)
ARMA–GARCH-GC4-ML	17 (0.5640)	21 (0.1935)
ARMA–GARCH-GC8-ML	16 (0.4672)	21 (0.1935)

p-values for the binomial test are in parentheses. EVT considers a 10% threshold.



A. Equity Hedge.



B. Event Driven.

Fig. 2. VaR at 99% in the backtesting period under different specifications. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In the next two sections we model the possible non-linear dependences between the two hedge funds by using the concept of copulas. We extend the SNP approach to the multivariate case by introducing a novel expression for the GC copula and comparing its performance on capturing VaR to other meta distributions.

4. Multivariate case

Measuring the risk of hedge fund or portfolio positions requires the consideration of non-linear dependence among asset returns. For this purpose the use of copulas represents a useful tool since it allows defining an enormous variety of multivariate distributions by linking variables with different marginal distributions. Consistently with the SNP approach, we derive an analytical expression for the GC copula and its density based on the multivariate GC density provided by

Perote [53]. The density of this GC copula can be straightforwardly obtained from the density of the Gaussian copula but incorporates non-linear dependence depending on the Hermite polynomial parameter structure. In what follows we revise the basic concepts of copulas and apply them to obtain the GC copula and other related meta distributions. For the sake of simplicity, but without loss of generality, we focus on the bivariate case and only the third and fourth Hermite polynomials are considered.

4.1. Copula basics

A bivariate copula is a function $C : [0, 1][0, 1] \to [0, 1]$ such that (i) $C(u_1, u_2)$ is increasing in both u_1 and u_2, (ii) $C(u_1, 0) = C(0, u_2) = 0$, $C(u_1, 1) = u_1$ and $C(1, u_2) = u_2$, (iii) $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0\, \forall u_1, u_2, v_1, v_2 \in [0, 1]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$.

Sklar's theorem [54] states that if $F(x_1, x_2)$ is a multivariate distribution function with marginals $F_1(x_1)$ and $F_2(x_2)$ then there exists a copula C such that

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{31}$$

If margins are continuous, then the copula is unique. The Sklar's Theorem can be rewritten as

$$C(u_1, u_2) = F\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right), \tag{32}$$

where $F_i^{-1}(u_i)$ with $i \in \{1, 2\}$ is the inverse of F_i, and u_i are standard uniform variates. From (32) a copula can be extracted from a joint distribution function F and continuous margins F_i.

4.2. GC copula

Let $\boldsymbol{X} = (x_1 x_2)$ be a random vector with GC joint distribution function, F^{GC}, and GC margins $F_1(x_1)$ and $F_2(x_2)$, where $x_i = F_i^{-1}(u_i)$ with $i = 1, 2$, then the GC copula can be obtained as

$$C^{GC}(u_1, u_2) = F^{GC}\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right). \tag{33}$$

Thus, the GC copula is given by (see the proof in Appendix A):

$$\begin{aligned} C^{GC}(u_1, u_2) = {} & \boldsymbol{\Phi}\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right) - \Phi\left(F_2^{-1}(u_2)\right)\phi\left(F_1^{-1}(u_1)\right)\tilde{q}_1\left(F_1^{-1}(u_1)\right) \\ & - \Phi\left(F_1^{-1}(u_1)\right)\phi\left(F_2^{-1}(u_2)\right)\tilde{q}_2\left(F_2^{-1}(u_2)\right), \end{aligned} \tag{34}$$

where $\tilde{q}_i(x_i) = \frac{d_{3i}}{\sqrt{3}} H_2(x_i) + \frac{d_{4i}}{\sqrt{4}} H_3(x_i)$, $\forall i = 1, 2$, $\boldsymbol{\Phi}$ is the multivariate normal cdf, Φ is the univariate normal cdf and ϕ is univariate normal pdf.

4.3. The density of the GC copula density and the ML estimation

The density of a copula $C(u_1, u_2)$ can be obtained as

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}. \tag{35}$$

Similarly, if $f(x_1, x_2)$ is the pdf of the random vector $\boldsymbol{X} = (x_1 x_2)'$, with marginal pdfs $f_1(x_1)$ and $f_2(x_2)$, then the density of the copula is

$$c(u_1, u_2) = \frac{f\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right)}{f_1\left(F_1^{-1}(u_1)\right) f_2\left(F_2^{-1}(u_2)\right)}, \tag{36}$$

since $f\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right) = c(u_1, u_2) f_1\left(F_1^{-1}(u_1)\right) f_2\left(F_2^{-1}(u_2)\right)$, known as the canonical representation. Thus, the expression for the density of the GC copula is (see the proof in Appendix B):

$$c^{GC}(u_1, u_2) = \frac{c^{Ga}(u_1, u_2) + q_1\left(F^{-1}(u_1)\right) + q_2\left(F^{-1}(u_2)\right)}{\left[1 + q_1\left(F^{-1}(u_1)\right)\right]\left[1 + q_2\left(F^{-1}(u_2)\right)\right]}, \tag{37}$$

where c^{Ga} is the Gaussian copula density, F^{-1} is the inverse of the standard GC distribution function, and $q_i(x_i) = d_{3i} H_3(x_i) + d_{4i} H_4(x_i)$, $\forall i = 1, 2$.

The estimation of the vector of parameters of the copula, $\boldsymbol{\theta}$, has traditionally been estimated by ML techniques. For such purpose the log-likelihood function may be defined as

$$\ln L\left(\boldsymbol{\theta}, \hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_n\right) = \sum_{t=1}^{n} \ln c_{\boldsymbol{\theta}}\left(\hat{\boldsymbol{U}}_t\right), \tag{38}$$

where c_θ denotes the copula density and $\hat{\boldsymbol{U}}_t$ a pseudo-observation from the copula. Thus, the log-likelihood function for the density of the GC copula is given by (see Appendix B):

$$\ln L\left(\boldsymbol{\theta}; \hat{\boldsymbol{Y}}_{\mathbf{1}}, \hat{\boldsymbol{Y}}_{\mathbf{2}}\right) = \sum_{t=1}^{n} \ln\{\boldsymbol{\phi}(Y_{1t}, Y_{2t}) + \phi(Y_{1t})\,\phi(Y_{2t})\,[q_1(Y_{1t}) + q_2(Y_{2t})]\}$$

$$- \sum_{t=1}^{n} \ln(1 + q_1(Y_{1t})) - \sum_{t=1}^{n} \ln(1 + q_2(Y_{2t})) - \sum_{t=1}^{n} \ln(\phi(Y_{1t})) - \sum_{t=1}^{n} \ln(\phi(Y_{2t}))\,, \tag{39}$$

where $Y_1 = F^{-1}(u_1)$ and $Y_2 = F^{-1}(u_2)$, and $\boldsymbol{\phi}$ denotes the multivariate density of a normal distribution.

4.4. Fitting the GC copula

The ML estimation of the copula parameters are based on pseudo-sample observations, since copula data are not directly observed. This procedure [55] is divided in two steps:

First, margins are estimated from the sample $\{x_t\}_{t=1}^{T}$ via the empirical distribution function

$$\hat{F}_T\left(x_j\right) = \frac{1}{T+1} \sum_{t=1}^{T} I_{\{x_t \le x_{j,T}\}}, \tag{40}$$

where $I_{\{.\}}$ denotes the indicator function and thus $\hat{F}_T\left(x_j\right)$ is the frequency of observations below or equal to x_j. The term $T+1$ ensures that data from the pseudo-sample lies in the interior of the unit square.

Second, the pseudo-sample of observations from the copula is generated

$$\hat{\boldsymbol{U}}_t = (U_{1t}\,, U_{2t})' = \left(\hat{F}_1(x_{1t})\,, \hat{F}_2(x_{2t})\right)'. \tag{41}$$

Once the pseudo-sample is obtained, expression (38) is maximized (numerically) in order to estimate the copula parameters.

4.5. Simulation of the GC meta distribution

The Sklar's Theorem is a useful tool to define multivariate distributions with arbitrary copulas and margins. The well-known Li model [56] is a particular case of the Gaussian copula with exponential margins traditionally used to model default times in companies. This model implements the so-called Gaussian meta distribution. Similarly, we introduce the GC meta distribution whose simulation algorithm is described below:

(1) Generate $\boldsymbol{X} \sim F_{d=2}^{\mathrm{GC}}(\mathbf{0}, P)$.
  (i) Perform a Cholesky decomposition to obtain $P^{1/2}$.
  (ii) Generate $\boldsymbol{Z} = (Z_1, Z_2)'$ of GC variates, where $z_i \sim \mathrm{GC}(0, 1, d_{3i}, d_{4i})\,,\ i = 1, 2$.
  (iii) Set $\boldsymbol{X} = \mu + P^{1/2}\boldsymbol{Z}$.
(2) Return $\boldsymbol{U} = (F_1(X_1)\,, F_2(X_2))'$, where F_1 and F_2 are GC margins. Then the random vector $\boldsymbol{U}$ has C_P^{GC} distribution.
(3) Return $\boldsymbol{R} = G^{-1}(\boldsymbol{U})$. Then $\boldsymbol{R}$ is a GC meta distribution, in other words, a random vector with F_1 and F_2 margins and multivariate distribution $C_P^{\mathrm{GC}}(G_1(X_1)\,, G_2(X_2))$.

5. Empirical application: VaR quantification using the GC meta distribution

This section develops an application of the GC copula to compute the VaR of a portfolio composed of the two hedge fund returns previously analyzed in Section 3. We compare the VaR performance of the meta distribution based on this copula to other meta distributions. For this purpose we proceed as follows:

First, the daily hedge fund index returns described in Table 1 are standardized and the parameters of the GC, Gaussian and Student t copulas are estimated, as well as the marginal distribution which fits the data the best. Table 5 displays the estimated parameters for the copulas and their log-likelihood values. The correlation coefficient (ρ) is very similar for the three specifications but the parameters of the GC (d_{3i} and d_{4i}, $\forall i = 1, 2$) induce non-linear dependence. For this reason these parameters (d_{4i}) are negative and the Student's t seems to be the best fitted copula according to the log-likelihood value. Table 6 shows the parameters estimated for the three marginal distributions (normal, Cauchy and Student's t) and their loglikelihood values. Note that both the Cauchy and the normal are limited cases of the Student's t as the degrees of freedom (ν) tend to one and infinity, respectively. Thus the Student's t marginal distribution is chosen in terms of accuracy.

Second, bivariate series, $\boldsymbol{R}_t = (R_{1t}R_{2t})'$, of size T are simulated from the algorithm detailed in Section 4.5, where T is the length of the hedge fund index returns, then we construct different meta distributions and compare their VaR performance. Particularly, we combine the three copulas in Table 5 (Gaussian, GC and Student t) with the best marginal in Table 6 (Student's t). We denote these meta distributions as M1, M2 and M3, respectively. Furthermore, we include the multivariate normal (meta Gaussian with normal margins or M4) and the multivariate GC (meta GC with GC margins or M5). The information about models M1–M5 is summarized in Table 7.

Table 5
Fitted copulas.

Copula	Loglikelihood value	Parameters estimated
Gaussian	865.3603	$\rho = 0.7393$
Gram–Charlier	916.3513	$\rho = 0.7209$, $d_{31} = 0.0196$, $d_{41} = -0.0108$, $d_{32} = 0.0132$, $d_{42} = -0.0104$
Student *t*	954.2817	$\rho = 0.7490$, $\nu = 3.7443$

Table 6
Fitted marginal densities.

Distribution	Equity Hedge		Event Driven	
	Loglikelihood	Parameters	Loglikelihood	Parameters
Normal	−3121.165	$\mu = 0.0000$, $\sigma = 0.9998$	−3121.165	$\mu = 0.0000$, $\sigma = 0.9998$
Cauchy	−3114.102	Location = 0.0985, scale = 0.4606	−2992.852	Location = 0.0970, scale = 0.4399
Student's *t*	−2930.764	$\nu = 3.2971$	−2808.314	$\nu = 3.0928$

Table 7
Meta distributions analyzed.

Meta distribution	Copula	Marginal distribution
M1-Meta Gauss (marginals *t*)	Gaussian, $\rho = 0.7393$	Student's *t* ($\nu_1 = 3.2971$, $\nu_2 = 3.0928$)
M2-Meta *t* (marginals *t*)	Gram–Charlier, $\rho = 0.7209$, $d_{31} = 0.0196$, $d_{41} = -0.0108$, $d_{32} = 0.0132$, $d_{42} = -0.0104$	Student's *t* ($\nu_1 = 3.2971$, $\nu_2 = 3.0928$)
M3-Meta GC (marginals *t*)	*t*, $\rho = 0.7490$, $\nu = 3.7443$	Student's *t* ($\nu_1 = 3.2993$, $\nu_2 = 3.0915$)
M4-Meta Gauss (marginals normal)	Gaussian, $\rho = 0.7393$	Standard normal
M5-Meta GC (marginals GC)	Gram–Charlier, $\rho = 0.7209$, $d_{31} = 0.0196$, $d_{41} = -0.0108$, $d_{32} = 0.0132$, $d_{42} = -0.0104$	Standard GC

Table 8
Descriptive statistics for hedge fund returns.

	Portfolio	Standardized portfolio
Mean	0.0056	0.0000
Median	0.0374	0.0813
Standard deviation	0.3504	0.9416
Variance	0.1228	0.8865
Excess kurtosis	8.8094	9.5089
Skewness	−1.0459	−1.0829
Range	5.6337	15.6006
Minimum	−3.0693	−8.5246
Maximum	2.5644	7.0760

Third, we form a portfolio $R_{pt} = w_1 R_{1t} + w_2 R_{2t}$, where $0 \leq w_i \leq 1\ \forall i = 1, 2$ and $w_1 + w_2 = 1$. Particularly, we consider an equally weighted portfolio, i.e. $w_1 = w_2 = 0.5$. Table 8 gathers the main descriptive statistics for the formed portfolio.

The portfolio VaR for the confidence level α, VaR_α, is calculated as the α-quantile of R_p. This procedure is repeated M times (M large) and the mean and standard error are computed according to the Eqs. (42) and (43), respectively.

$$\overline{\mathrm{VaR}_\alpha} = \frac{1}{M} \sum_{i=1}^{M} \mathrm{VaR}_{\alpha,i}, \tag{42}$$

$$\sigma(\mathrm{VaR}_\alpha) = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} \left(\mathrm{VaR}_{\alpha,i} - \overline{\mathrm{VaR}_\alpha}\right)^2}. \tag{43}$$

The VaR of each meta distribution is compared with the empirical VaR, calculated as the quantile from the original (hedge fund) standardized returns. The results are shown in Table 9.

The average VaR values estimated by M5 are relatively close to the empirical VaR when α is equal to 0.05 and 0.01; however, the result is far from the empirical VaR when $\alpha = 0.001$. In the latter case, the VaR estimated by M2 is closer to the empirical VaR. M4 performs poorly, which is a Meta Gauss copula with normal margin distributions. Therefore the GC copula seems to provide accurate VaR measures for the traditionally employed confidence levels.

Table 9
VaR performance of meta distributions.

Model	A		
	0.05	0.01	0.001
Empirical (EW)	−1.5274	−2.8924	−5.7409
M1-Meta Gauss (marginals *t*)	−1.3193 (0.0659)	−2.4610 (0.2051)	−4.9737 (1.0749)
M2-Meta *t* (marginals *t*)	−1.3057 (0.0675)	−2.5042 (0.2241)	−5.2359 (1.1460)
M3-Meta GC (marginals *t*)	−1.3059 (0.0638)	−2.4504 (0.2019)	−4.8635 (0.9913)
M4-Meta Gauss (marginals normal)	−1.5312 (0.0471)	−2.1592 (0.0833)	−2.8076 (0.1800)
M5-Meta GC (marginals GC)	−1.6533 (0.0762)	−2.7067 (0.1259)	−4.1470 (1.3305)

Standard deviation in parenthesis.

6. Conclusions

Traditional VaR measures, based on the normal distribution, have been criticized because of their inability to adequately capture market risk, particularly for high confidence levels. For this purpose, the use of alternative thick-tailed and skewed distributions has been proposed. However, still there is not a consensus about the more appropriate methodology for VaR forecasting. On the other hand, hedging strategies require accurate VaR measures to quantify potential losses and avoid disasters as the LTCM bankruptcy. In this article, we investigate on these issues by comparing the relative performance of three parametric models (normal, Student's *t* and skewed-*t*), the EVT approach and a SNP model (GC).

Our results confirm the strongly rejection of both the normal- and the Student's-based VaR measures for high confidence levels and volatile scenarios [24]. The skewed-*t* is an alternative to capture skewness although might not be the best option when leptokurtosis is severe due to the high occurrence of outliers. The EVT and the GC densities involve accurate market risk measures since the former focus on extreme values and the latter is very flexible to adapt different scenarios with a variable number of parameters. These results are consistent with other articles that have implemented SNP [47,57] and EVT [58,59] methodologies for other asset returns, but this is the first evidence for hedge fund returns.

The EVT, however, is very sensitive to the threshold selection [59] and thus our study particularly focuses on and recommends the SNP approach. Within the SNP framework we analyze the performance of distributions based on GC series, unlike other related contributions that implement the Cornish–Fisher [15], Laguerre [60] or Positive Edgeworth–Sargan expansions [47]. Furthermore, the article compares the VaR forecasting performance of the GC series traditionally expanded to the fourth term for option prices purposes [44,45] to larger expansions proposed to quantify portfolio risk [34]. Our results support the use of shorter expansions for risk forecasting since simpler models seem to provide better (or at least as good as) outcomes. This result reinforces the fact that the best good in-sample fit does not guarantee the best out-of-sample performance [61], we also explore the advantages of using different estimation techniques for estimating GC densities: MM versus ML. Unlike [43], we find that efficient ML techniques seem to outperform MM when forecasting risk of hedge fund strategies.

The article also goes one step further into the multivariate case by providing a direct expression for the GC copula and its density based on the multivariate GC distribution in Ref. [53]. This expression allows us to directly obtaining the GC copula from the Gaussian copula, which is a particular case, and introduces non-linear dependence by means of the Hermite polynomial structures. We compare the VaR performance of different meta distributions based on GC, Gaussian and Student's *t* copulas by computing portfolio VaR at different confidence levels (95%, 99% and 99.9%) for an equally weighted portfolio formed by the two hedge fund returns. Our results show that the portfolio VaR estimated by the multivariate GC (meta GC) is close to the empirical VaR and seems to outperform other multivariate specifications when $\alpha = 0.01$, the regulatory VaR for market risk. However, for a VaR beyond the 99% confidence level, the result obtained by the meta *t* works better. For all these reasons we recommend the use of SNP distributions for risk managing purposes.

Acknowledgment

Financial support from the Junta de Castilla y León, through the project SA218A11-1 is gratefully acknowledged.

Appendix A. GC copula

Let Eq. (A.1) be the standardized multivariate GC density introduced in Ref. [53]. Without loss of generality we consider the bivariate case and expansions including only the third and fourth Hermite polynomials for both dimensions.

$$\boldsymbol{f}(x_1, x_2) = \boldsymbol{\phi}(x_1, x_2) + [q_1(x_1) + q_2(x_2)]\,\phi(x_1)\,\phi(x_2), \tag{A.1}$$

where $\boldsymbol{\phi}$ is the standard multivariate normal density with correlation coefficient ρ, ϕ denotes the univariate normal density in Eq. (8) and $q_i(x_i) = d_{3i}H_3(x_i) + d_{4i}H_4(x_i)$, $\forall i = 1, 2$. The marginal distributions of this density are univariate GC distributed and thus such distribution inherits the good performance of the SNP approach in Section 2.2.2. and allows for non-linear between x_1 and x_2.

The corresponding multivariate GC cdf can be easily obtained by integration:

$$\begin{aligned}
\boldsymbol{F}(x_1,x_2) &= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\{\boldsymbol{\phi}(s,t)\,\mathrm{d}t\mathrm{d}s + (q_1(s)+q_2(t))\,\phi(s)\phi(t)\}\,\mathrm{d}t\mathrm{d}s\\
&= \boldsymbol{\Phi}(x_1,x_2) + \int_{-\infty}^{x_1} q_1(s)\,\phi(s)\mathrm{d}s\int_{-\infty}^{x_2}\phi(t)\,\mathrm{d}t + \int_{-\infty}^{x_1}\phi(s)\mathrm{d}s\int_{-\infty}^{x_2} q_2(t)\phi(t)\mathrm{d}t\\
&= \boldsymbol{\Phi}(x_1,x_2) + \Phi(x_2)\int_{-\infty}^{x_1}[d_{31}H_3(s)+d_{41}H_4(s)]\,\phi(s) + \Phi(x_1)\int_{-\infty}^{x_2}[d_{32}H_3(t)+d_{42}H_4(t)]\,\phi(t)\\
&= \boldsymbol{\Phi}(x_1,x_2) + \Phi(x_2)\left[-\frac{d_{31}}{\sqrt{3}}H_2(x_1)\,\phi(x_1) - \frac{d_{41}}{\sqrt{4}}H_3(x_1)\,\phi(x_1)\right]\\
&\quad + \Phi(x_1)\left[-\frac{d_{32}}{\sqrt{3}}H_2(x_2)\,\phi(x_2) - \frac{d_{42}}{\sqrt{4}}H_3(x_2)\,\phi(x_2)\right]\\
&= \boldsymbol{\Phi}(x_1,x_2) - \Phi(x_2)\,\phi(x_1)\,\tilde{q}_1(x_1) - \Phi(x_1)\,\phi(x_2)\,\tilde{q}_2(x_2)
\end{aligned}\tag{A.2}$$

where $\tilde{q}_i(x_i) = \frac{d_{3i}}{\sqrt{3}}H_2(x_i) + \frac{d_{4i}}{\sqrt{4}}H_3(x_i)\ \forall i = 1, 2$ and Φ stands for the univariate normal cdf. Note that to solve the integral we applied a well-known Hermite polynomial property:

$$\int_{-\infty}^{x} H_j(s)\phi(s)\mathrm{d}s = -\frac{1}{\sqrt{j}}H_{j-1}(x)\phi(x). \tag{A.3}$$

Thus, the GC copula with GC margins $F_1(x_1)$ and $F_2(x_2)$ can be straightforwardly extracted by the Sklar's theorem—see Eq. (32):

$$\begin{aligned}
C^{\mathrm{GC}}(u_1,u_2) &= \boldsymbol{\Phi}\left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right) - \Phi\left(F_2^{-1}(u_2)\right)\phi\left(F_1^{-1}(u_1)\right)\tilde{q}\left(F_1^{-1}(u_1)\right)\\
&\quad - \Phi\left(F_1^{-1}(u_1)\right)\phi\left(F_2^{-1}(u_2)\right)\tilde{q}\left(F_2^{-1}(u_2)\right).
\end{aligned}\tag{A.4}$$

Appendix B. GC copula density and its log-likelihood function

The density of the GC copula provided in Appendix A, c^{GC}, can be obtained in terms of the density of the Gaussian copula, c^{Ga}, as a direct application of the canonical representation theorem. Particularly, the multivariate normal pdf and the multivariate GC pdf can be expressed as in Eqs. (B.1) and (B.2), respectively:

$$\boldsymbol{\phi}\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right) = c^{\mathrm{Ga}}(u_1,u_2)\,\phi\left(\Phi^{-1}(u_1)\right)\phi\left(\Phi^{-1}(u_2)\right) \tag{B.1}$$

$$\begin{aligned}
\boldsymbol{f}\left(F^{-1}(u_1), F^{-1}(u_2)\right) &= \boldsymbol{\phi}\left(F^{-1}(u_1)\right)\phi\left(F^{-1}(u_2)\right) + \left[q\left(F^{-1}(u_1)\right) + q\left(F^{-1}(u_2)\right)\right]\\
&\quad\times\phi\left(F^{-1}(u_1)\right)\phi\left(F^{-1}(u_2)\right).
\end{aligned}\tag{B.2}$$

By substituting (B.1) in (B.2) and dividing by $f_1\left(F^{-1}(u_1)\right)f_2\left(F^{-1}(u_2)\right)$ we directly obtain the density of the GC copula:

$$\begin{aligned}
c^{\mathrm{GC}}(u_1,u_2) &= \frac{\left[c^{\mathrm{Ga}}(u_1,u_2) + q_1\left(F^{-1}(u_1)\right) + q_2\left(F^{-1}(u_2)\right)\right]\phi\left(F^{-1}(u_1)\right)\phi\left(F^{-1}(u_2)\right)}{f_1\left(F^{-1}(u_1)\right)f_2\left(F^{-1}(u_2)\right)}\\
&= \frac{c^{\mathrm{Ga}}(u_1,u_2) + q_1\left(F^{-1}(u_1)\right) + q_2\left(F^{-1}(u_2)\right)}{\left[1+q_1\left(F^{-1}(u_1)\right)\right]\left[1+q_2\left(F^{-1}(u_2)\right)\right]},
\end{aligned}\tag{B.3}$$

since the terms of the univariate Gaussian pdfs cancel out. Therefore the log-likelihood function for a couple of observations is

$$\ln c^{\mathrm{GC}}(u_1,u_2) = \ln c^{\mathrm{Ga}}(u_1,u_2) + q_1(Y_1) + q_2(Y_2) - \ln[1+q_1(Y_1)] - \ln[1+q_2(Y_2)],$$

provided that $Y_1 = F^{-1}(u_1)$ and $Y_2 = F^{-1}(u_2)$.

For a whole sample of n observations the log-likelihood becomes:

$$\ln L\left(\boldsymbol{\theta}; \hat{Y}_1, \hat{Y}_2\right) = \sum_{i=1}^{n}\left\{\ln c^{\mathrm{Ga}}_{(u_1,u_2)} + q_1(Y_{1i}) + q_2(Y_{2i})\right\} - \sum_{i=1}^{n}\ln[1+q_1(Y_{1i})] - \sum_{i=1}^{n}\ln[1+q_2(Y_{2i})].$$

By replacing Gaussian copula, $c^{\mathrm{Ga}}_{(u_1,u_2)} = \frac{\boldsymbol{\phi}(Y_{1i},Y_{2i})}{\phi(Y_{1i})\phi(Y_{2i})}$ and after straightforward manipulation:

$$\begin{aligned}
\ln L\left(\boldsymbol{\theta}; \hat{\boldsymbol{Y}}_1, \hat{\boldsymbol{Y}}_2\right) &= \sum_{i=1}^{n}\{\ln\boldsymbol{\phi}(Y_{1i},Y_{2i}) + \phi(Y_{1i})\,\phi(Y_{2i})\,[q_1(Y_{1i}) + q_2(Y_{2i})]\}\\
&\quad - \sum_{i=1}^{n}\ln\phi(Y_{1i}) - \sum_{i=1}^{n}\phi(Y_{2i}) - \sum_{i=1}^{n}\ln[1+q_1(Y_{1i})] - \sum_{i=1}^{n}\ln[1+q_2(Y_{2i})].
\end{aligned}\tag{B.4}$$

References

[1] V. Agarwal, N.Y. Naik, Hedge Funds, in: Found. Trends Financ., vol. 1, 2005, pp. 103–170.
[2] P. Jorion, Risk management lessons from long term capital management, Eur. Financ. Manage. 6 (2000) 277–300.
[3] A. Gupta, B. Liang, Do hedge funds have enough capital? A value-at-risk approach, J. Financ. Econ. 77 (2005) 219–253.
[4] K. Dowd, D. Blake, After VaR: the theory, estimation, and insurance applications of quantile-based risk measures, J. Risk Insurance 73 (2006) 193–229.
[5] C. Alexander, J.M. Sarabia, Quantile uncertainty and value-at-risk model risk, Risk Anal. 32 (2012) 1293–1308.
[6] G. Bormetti, E. Cisana, G. Montagna, O. Nicrosini, A non-Gaussian approach to risk measures, Physica A 376 (2007) 532–542.
[7] W.C. Chin, Heavy-tailed value-at-risk analysis for Malaysian stock exchange, Physica A 387 (2008) 4285–4298.
[8] T.B. Bali, P. Theodossiou, Risk measurement performance of alternative distribution functions, J. Risk Insurance 75 (2008) 411–437.
[9] S. Pafka, I. Kondor, Evaluating the RiskMetrics methodology in measuring volatility and value-at-risk in financial markets, Physica A 299 (2001) 305–310.
[10] M. Kozaki, A.-H. Sato, Application of the Beck model to stock markets: value-at-risk and portfolio risk assessment, Physica A 387 (2007) 1225–1246.
[11] Y. Wei, W. Chen, Y. Lin, Measuring daily value-at-risk of SSEC index: a new approach based on multifractal analyisis and extreme value theory, Physica A 388 (2013) 3536–3542.
[12] T.-L. Tang, S.-J. Shieh, Long memory in stock index futures markets: a value-at-risk approach, Physica A 366 (2006) 437–448.
[13] W. Fung, D.A. Hsieh, N.Y. Naik, T. Ramadorai, Hedge funds: performance, risk and capital formation, J. Finance 63 (2008) 1777–1803.
[14] D. Giannikis, I.D. Vrontos, A Bayesian approach to detecting nonlinear risk exposures in hedge fund strategies, J. Bank. Finance 35 (2011) 1399–1414.
[15] R.D.F. Harris, M. Mazibas, Dynamic hedge fund portfolio construction: a semi-parametric approach, J. Bank. Finance 37 (2013) 139–149.
[16] L. Favre, J.A. Galeano, Mean-modified value-at-risk optimization with hedge funds, J. Altern. Invest. 6 (2002) 21–25.
[17] G. Zumbach, Backtesting risk methodologies from one day to one year, J. Risk 9 (2006) 55–91.
[18] B. Güner, S.T. Rachev, B. Edelman, F.J. Fabozzi, Bayesian inference for hedge funds with stable distribution of returns, in: Klaus Böcker (Ed.), Rethinking Risk Measurement and Reporting: Volume II, 2010, pp. 95–136.
[19] B.E. Hansen, Autoregressive conditional density estimation, Internat. Econom. Rev. 35 (1994) 705–730.
[20] A.J. McNeil, R. Frey, Estimation of tail-related risk measures for heteroskedastic financial time series: an extreme value approach, J. Empir. Finance 7 (2000) 271–300.
[21] M. Kendall, A. Stuart, The Advanced Theory of Statistics, Griffin & Co., London, 1977.
[22] J.P. Gram, Om Raekkeudviklinger Bestempte Ved Hjaelp av de Mindste Kvadraters Methode (Kobenhavn), 1879.
[23] C.V. Charlier, Uber die darstellung willkurlicher funktionen, Ark. Mat. Astron. Fys. 9 (1905) 1–13.
[24] F.Y. Edgeworth, On the representation of statistical frequency by series, J. Roy. Statist. Soc. Ser. A 70 (1907) 102–106.
[25] M.B. Kock, H.C. Eggers, J. Schimiegel, Edgeworth versus Gram–Charlier series: x-Cumulant and probability density tests, Phys. Particles Nuclei 8 (2011) 1023–1027.
[26] S. Blinnikov, R. Moessner, Expansions for nearly Gaussian distributions, Astron. Astrophys. Suppl. Ser. 130 (1998) 193–205.
[27] V. Fernandez, Copula-based measures of dependence structure in assets returns, Physica A 387 (2008) 3615–3628.
[28] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50 (1982) 987–1007.
[29] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, J. Econometrics 31 (1986) 307–327.
[30] R.L. Smith, Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone, Statist. Sci. 4 (1989) 367–393.
[31] A.C. Davison, R.L. Smith, Models for exceedances over high thresholds, J. R. Stat. Soc. Ser. B Stat. Methodol. 52 (1990) 393–442.
[32] M.R. Leadbetter, On a basis for 'peaks over thresholds' modeling, Statist. Probab. Lett. 4 (1991) 357–362.
[33] B. Mandelbrot, The variation of certain speculative prices, J. Bus. 36 (1963) 394–419.
[34] I. Mauleon, J. Perote, Testing densities with financial data: an empirical comparison of the Edgeworth–Sargan density to the Student's t, Eur. J. Finance 6 (2000) 225–239.
[35] R. Cont, Empirical properties of asset returns: stylized facts and statistical issues, Quant. Finance 1 (2001) 223–236.
[36] C. Fernández, M.F.J. Steel, On Bayesian modeling of fat tails and skewness, J. Amer. Statist. Assoc. 93 (1998) 359–371.
[37] P. Theodossiou, Financial data and the skewed generalized t distribution, Manage. Sci. 44 (1998) 1650–1661.
[38] M.D. Branco, D.K. Dey, A general class of multivariate skew-elliptical distributions, J. Multivariate Anal. 79 (2001) 99–113.
[39] M.C. Jones, M.J. Faddy, A skew extension of the T-distribution, with applications, J. R. Stat. Soc. Ser. B 65 (2003) 159–174.
[40] P. Giot, S. Laurent, Value-at-risk for long and short trading positions, J. Appl. Ecol. 18 (2003) 641–663.
[41] A. Azzalini, A. Capitanio, Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution, J. R. Stat. Soc. Ser. B 65 (2003) 367–389.
[42] H. Cramér, On some classes of series used in mathematical statistics, in: Sixth Scandinavian Congr. of Hath., Kobenhaven, 1925, p. 399.
[43] E.B. Del Brio, J. Perote, Gram–Charlier densities: maximum likelihood versus the method of moments, Insurance Math. Econom. 51 (2012) 531–537.
[44] R. Jarrow, A. Rudd, Approximate option valuation for arbitrary stochastic processes, J. Financ. Econ. 10 (1982) 347–369.
[45] B. Baliero, R. Rosenfeld, Testing option pricing with Edgeworth expansion, Physica A 34 (2004) 484–490.
[46] J. Perote, Value at risk of non-normal portfolios, Span. Rev. Financ. Account. 115 (2003) 290–310.
[47] T.M. Ñíguez, J. Perote, Forecasting heavy-tailed densities with positive Edgeworth and Gram–Charlier expansions, Oxford Bull. Econ. Stat. 74 (2012) 600–627.
[48] A.R. Gallant, D.W. Nychka, Seminonparametric maximum likelihood estimation, Econometrica 55 (1987) 363–390.
[49] E. Jondeau, M. Rockinger, Gram–Charlier densities, J. Econom. Dynam. Control 25 (2001) 1457–1483.
[50] A. León, J. Mencía, E. Sentana, Parametric properties of semi-nonparametric distributions, J. Bus. Econom. Statist. 27 (2009) 176–192.
[51] P. Christoffersen, Evaluating interval forecasts, Internat. Econom. Rev. 39 (1998) 841–862.
[52] T.G. Bali, S. Gokcan, Alternative approaches to estimating VaR for hedge fund portfolios, in: B. Schachter (Ed.), Intelligent Hedge Fund Investing, Risk Books, Incisive Media PLC, 2004, pp. 253–277.
[53] J. Perote, The multivariate Edgeworth–Sargan density, Span. Econ. Rev. 6 (2004) 77–96.
[54] B. Schweizer, A. Sklar, Operations on distribution functions not derivable from operations on random variables, Studia Math. 52 (1974) 43–52.
[55] C. Genest, K. Ghoudi, L.-P. Rivest, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, Biometrika 82 (1995) 543–552.
[56] D. Li, On default correlation: a copula function approach, in: Working Paper, RiskMetrics Group, New York, 1999.
[57] E.B. Del Brio, T.M. Ñíguez, J. Perote, Multivariate semi-nonparametric distributions with dynamic conditional correlations, Int. J. Forecast. 27 (2011) 347–364.
[58] A. Cifter, Value-at-risk estimation with wavelet-based extreme value theory: evidence from emerging markets, Physica A 390 (2011) 2356–2367.
[59] S.T. Rachev, B. Racheva-Iotova, S. Stoyanov, Capturing fat tails, Risk Mag. (2010) 72–77.
[60] K. Marumo, R. Wolff, Expansion methods applied to asset return distributions, J. Risk 10 (2007) 3–24.
[61] P.R. Hansen, In-sample fit and out-of-sample fit: their joint distribution and its implications for model selection, in: Working Paper, Standford University, 2009.