Robust Portfolio Optimization with Multivariate Copulas: a Worst-Case CVaR approach

Fernando A. B. Sabino da Silva[*], Flavio A. Ziegelmann[†] and Cristina Tessari [‡]

Abstract

Using data from the S&P 500 stocks from 1990 to 2015, we measure the downside market risk by Conditional Value-at-Risk (CVaR) subject to return constraints following the approach of Rockafellar and Uryasev (2000, 2002) and the extended framework of Kakouris and Rustem (2014) through the use of multidimensional mixed archimedean copulas. We implement a dynamic investing viable strategy where the portfolios are optimized using three different length of rolling calibration windows. The out-of-sample performance is evaluated and compared against two benchmarks; a multidimensional gaussian copula model and a constant mix portfolio. Our empirical analysis shows that the Mixed Copula-CVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR and 1/N portfolios for daily and weekly rebalancing. To cope with the dimensionality problem we employ a similar approach to that used by Gatev, Goetzmann, and Rouwenhorst (2006) to select a set of assets that are the most diversified, in some sense, to the S&P 500 index in the constituent set. We find that copula-based approaches offer better hedges against losses than the 1/N portfolio. The accuracy of VaR forecasts is determined by how well they minimize a capital requirement loss function (CR). In order to mitigate data-snooping problems we apply a test for superior predictive ability (SPA) proposed by Hansen (2005) to determine which model significantly minimizes this expected loss function. The test is implemented via stationary bootstrap of Politis and Romano (1994) using the automatic block-length selection of Politis and White (2004) and 10,000 bootstrap resamples. We find that the minimum average loss of the mixed Copula-CVaR approach is smaller than the average performance of the Gaussian Copula-CVaR.

Keywords: Asset Allocation; Convex Programming; Copula; Portfolio Selection; Risk Management; S&P 500; WCCVaR.
JEL Codes: G11; G12; G17.

1 Introduction

The seminal article Portfolio Selection published by Markowitz (1952) introduces the foundation for modern portfolio theory (MPT) or mean-variance analysis. Markowitz considered the problem of an agent who wishes to find the maximum (expected) return for a given level of risk or minimize risk for a given level of return. He identified that, by diversifying a portfolio among investments that have different return patterns, investors can build such an efficient portfolio.

[*]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: fsabino@ufrgs.br; Corresponding author.
[†]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: flavioz@ufrgs.br
[‡]Finance Division, Columbia Business School, Columbia University, New York, NY 10027, USA, e-mail: ct2759@columbia.edu

Quantile functions are commonly used for measuring the market risk of models with parameter uncertainty and variability. Portfolio optimization involving a mean-value-at-risk (mean-VaR) portfolio and the CVaR have been analyzed by Alexander and Baptista (2002) and Rockafellar and Uryasev (2000), respectively. Akin to the classical Markowitz portfolio, in these approaches we want to determine the weights that maximize the portfolio return for a specified VaR or CVaR at a given confidence level or minimize these quantiles for a given confidence level subject to a fixed portfolio return.

Artzner, Delbaen, Eber, and Heath (1999) show that VaR has undesirable properties such as lack of sub-additivity and thus it is not a coherent measure. Furthermore, Uryasev and Rockafellar (2001) show that VaR may be a non-convex function with respect to portfolio weights, which can yield to multiple portfolio local solutions, but CVaR is coherent both for continuous and discrete distributions and it is a convex function. In addition, they show that an outright optimization with respect to CVaR is numerically difficult due to the dependence of the CVaR on VaR. However, Rockafellar and Uryasev (2000) show that VaR and CVaR can be computed simultaneously, introducing auxiliary risk measures, and it can be used in conjunction with scenario based optimization algorithms reducing the problem to a linear programming problem which allows us to optimize a portfolio with very large dimensions and stable numerical implementations.

A well-known problem of the Markowitz model is its sensitivity to the input parameters. In practice, the implementation of strategies based on the risk-return trade-off remains a fundamental challenge in many areas of financial management, since estimation errors of the expected returns of the assets and the covariance matrix of these returns can significantly affect the asset allocations weights, no longer leading to an efficient portfolio. This issue can be overcome by employing robust optimization and worst case techniques Zhu and Fukushima (2009) in which assumptions about the distribution of the random variable are relaxed, and thus, we obtain the optimal portfolio solution by optimizing over a prescribed feasible set and possible densities. Kakouris and Rustem (2014) show how copula-based models can be introduced in the Worst Case CVaR (WCVaR) framework. This approach is motivated by an investor's desire to be protected against the worst possible scenario.

In this paper, we employ a similar methodology to that of Kakouris and Rustem (2014) and investigate the advantage of such dependence structure through an empirical study. We evaluate the out-of-sample performance of the Worst Case Copula-CVaR and compare the relative performance of the strategy to the Gaussian Copula-CVaR, a naive 1/N (equally-weighted) and the S&P 500 index in the long term in terms of wealth accumulation and downside risk.

The main novel contribution of this paper is to select a diversified set of assets that can be useful during crises and tranquil periods, i.e., that somehow involves hedging of decisions to protect the investors against any market conditions and evaluate the approaches using a 50-dimensional archimedean and gaussian copula models without constructing hierarchical copulas.

Our data set consists of daily data of adjusted closing prices of all shares that belong to S&P 500 market index from July 2st, 1990 to December 31st, 2015. We obtain the adjusted closing prices from Bloomberg. The data set sample encompasses 6426 days and includes a total of 1100 stocks over the whole period. We consider only stocks that are listed throughout a 48 plus 6-month period in the analysis, i.e., close to 500 stocks in each trading period. Moreover, we consider alternative portfolio rebalancing frequencies and return constraints. Returns are calculated in excess of a risk-free asset.

To be effective in dealing with uncertainty, we select, among all listed assets in each formation period, a set of 50 stocks based on the ranked sum of absolute deviations (the twenty-five largest and smallest) between the normalized daily closing prices deviations of the S&P 500 index and all shares, adjusting them by dividends, stock splits and other corporate actions. By selecting a diversified cluster of assets that can be useful during crises and tranquil periods we address the issue of asset allocation taking into account the purpose of risk diversification. These stocks are then evaluated over the next six months.

Our backtest results indicate that the Worst Case Copula-CVaR approach outperforms consistently the benchmark strategies in the long term in terms of downside risk measured by VaR, CVaR and capital requirement loss.

The remainder of the paper is structured as follows. In Section 2 we present a general review, notations and definitions of the CVaR and mean-variance optimization methodologies and extends them to our Worst Case framework through the use of appropriate allowable and uncertainty sets. The data we use is briefly discussed in Section 3. Section 4 summarizes the empirical results of the analysis, while Section 5 concludes and provide further ideas for research.

2 Specifications of the Models under Analysis

2.1 Conditional Value-at-Risk

Let Y be a stochastic vector standing for market uncertainties and F_Y be its distribution function, i.e., $F_Y(u) = P(Y \leq u)$. Let also $F_Y^{-1}(v) = \inf\{u : F_Y(u) \geq v\}$ be its right continuous inverse and assume that it has a probability density function represented by $p(y)$[1]. Define the VaR_β as the β-quantile by

$$\begin{aligned} VaR_\beta(Y) &= \arg\min\{\alpha \in \mathbb{R} : P(Y \leq \alpha) \geq \beta\} \\ &= F_Y^{-1}(\beta), \end{aligned} \tag{1}$$

and the $CVaR_\beta$ as the solution to the following optimization problem (Pflug, 2000):

$$CVaR_\beta(Y) = \inf\left\{\alpha \in \mathbb{R} : \alpha + \frac{1}{1-\beta}E[Y-\alpha]^+\right\}, \tag{2}$$

where $[t]^+ = \max(t, 0)$.

Uryasev and Rockafellar (1999) have shown that the $CVaR$ is the conditional expectation of Y, given that $Y \geq VaR_\beta$, i.e.,

$$CVaR_\beta(Y) = \mathbb{E}(Y\,|Y \geq VaR_\beta(Y)). \tag{3}$$

Let $f(x, y)$ be a loss function depending upon a decision vector x that belongs to any arbitrarily chosen subset $X \in \mathbb{R}^n$ and a random vector $y \in \mathbb{R}^m$. In a portfolio optimization problem, the decision vector x can be a vector of portfolios' weights, X be a set of feasible portfolios, subject to linear constraints[2] and y a vector that stands for market variables that can affect the loss.

For each x, the loss function $f(x, y)$ is a random variable defined on a probability space $(\Omega, \mathcal{F}, P_{\mathcal{F}})$ having a probability distribution on $(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})$ induced by that of y, where $(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})$ stands for a Borel probability space including, therefore, the open and closed intervals in $\mathbb{R}$. The underlying distribution of

[1] This assumption can be relaxed (Uryasev, 2013).

[2] For example, we can assume a portfolio X that does not allow short-selling positions (all $x_i \geq 0$, for $i = 1, ..., n$), that be fully invested, i.e., the total portfolio weights sum up to unity and that the expected return be greater than an arbitrary value R.

$y \in \mathbb{R}^m$ is assumed to have a density $p(y)$ and let the probability of $f(x,y)$ not exceeding some threshold α be denoted by

$$F(x,\alpha) = \int_{f(x,y)\le\alpha} p(y)\, dy, \tag{4}$$

where $F(x,\alpha)$ is the cumulative distribution function for the loss function $f(x,y)$, non-decreasing and right-continuous with respect to α.

Using the previously defined notation (2) we can write the $CVaR$ function, at confidence level β, by

$$CVaR_\beta(x) = \frac{1}{1-\beta}\int_{f(x,y)\ge VaR_\beta(x)} f(x,y)p(y)dy, \tag{5}$$

The optimization of $CVaR$ is difficult because of the presence of the VaR in its definition (it requires the use of the nonlinear function max) in this infinite dimensional problem. The main contribution of Rockafellar and Uryasev (2000) was to define a simpler auxiliary function

$$F_\beta(x,\alpha) = \alpha + \frac{1}{1-\beta}\int_{f(x,y)\ge\alpha} (f(x,y)-\alpha)\, p(y)dy, \tag{6}$$

which can be used instead of $CVaR$ directly, without need to compute VaR first due to the following proposition (Pflug (2000)):

Proposition 1. *The function $F_\beta(x,\alpha)$ is convex with respect to α. In addition, minimizing $F_\beta(x,\alpha)$ with respect to α gives $CVaR$ and VaR is a minimum point of this function with respect to α, i.e.*

$$\min_{\alpha\in\mathbb{R}} F_\beta(x,\alpha) = F_\beta(x, VaR_\beta(x)) = CVaR_\beta(x) \tag{7}$$

Moreover, we can use $F_\beta(x,\alpha)$ to find the optimal weights, $CVaR$ and VaR simultaneously over an allowable feasible set, i.e.,

$$\min_{x\in X} CVaR_\beta(x) = \min_{(x,\alpha)\in X\times\mathbb{R}} F_\beta(x,\alpha). \tag{8}$$

Pflug (2000) show that under quite general conditions $F_\beta(x,\alpha)$ is a smooth function. In addition, if $f(x,y)$ is convex with respect to x, then $F_\beta(x,\alpha)$ is also convex with respect to x. Hence, if the allowable set X is also convex, we then have to solve a smooth convex optimization problem.

2.2 Optimization Problem

Assume that the analytical representation for the density $p(y)$ is not available, but we can approximate $F_\beta(x,\alpha)$ using J scenarios, y_j, $j = 1, ..., J$ which are sampled from the density function $p(y)$. Then, we approximate

$$\begin{aligned} F_\beta(x,\alpha) &= \alpha + \frac{1}{1-\beta}\int_{f(x,y)\ge\alpha} (f(x,y)-\alpha)\, p(y)dy \\ &= \alpha + \frac{1}{1-\beta}\int_{y\in\mathbb{R}^m} (f(x,y)-\alpha)^+ p(y)dy \end{aligned} \tag{9}$$

by its discretized version

$$\widetilde{F}_\beta(x,\alpha) = \alpha + \frac{1}{(1-\beta)J}\sum_{j=1}^{J} (f(x,y_j)-\alpha)^+ .$$

Assuming that the feasible set X and the loss function $f(x, y_j)$ are convex and the same for the loss function $f(x, y_j)$, we solve the following convex optimization problem:

$$\min_{x \in X, \alpha \in \mathbb{R}} \widetilde{F}_\beta (x, \alpha). \tag{10}$$

In addition, if the loss function $f(x, y)$ is linear with respect to x, then the optimization problem (10) reduces to the following linear programming (LP) problem:

$$\begin{aligned} \underset{x \in \mathbb{R}^n, z \in \mathbb{R}^J, \alpha \in \mathbb{R}}{\text{minimize}} \quad & \alpha + \frac{1}{(1-\beta)J} \sum_{j=1}^{J} z_j && (11) \\ \text{subject to} \quad & x \in X, && (12) \\ & z_j \geq f(x, y_j) - \alpha, && (13) \\ & z_j \geq 0, \ j = 1, \ldots, J. && (14) \end{aligned}$$

where z_j are indicator variables. By solving the LP problem above we find the optimal decision vector, x^*, the optimal VaR, α^*, and consequently the optimal $CVaR$, $\widetilde{F}_\beta (x = x^*, \alpha = \alpha^*)$.

Therefore, the optimization problem can be solved using algorithms that are capable of solving efficiently very large-scale problems with any distribution within reasonable time and high reliability as, for example, simplex or interior point methods.

In the next subsections we assume that there are n risky assets and denote by r their random (log)returns vector, i.e., $\mathbf{r}= (r_1, ..., r_n)^\top$, with expected returns $\mu= (\mu_1, ..., \mu_n)^\top$ and covariance matrix Σ_{nxn}. Let also r_f represent the risk-free asset returns and the decision (portfolio's weights) vector by $\mathbf{w}= (w_1, ..., w_n)^\top$. Also assume that $\mathbf{w} \in \mathcal{W}$, where $\mathcal{W}$ is a feasible set and that the portfolio return loss function $f_L (\mathbf{w}, \mathbf{r})$ is a convex (linear) function given by

$$f_L(\mathbf{w}, \mathbf{r}) = \mathbf{w}^\top \mathbf{r}.$$

By definition, the portfolio return is the negative of the loss, i.e., $-\mathbf{w}^\top \mathbf{r}$.

2.3 Worst Case CVaR

Assume now that we do not have precise information about the distribution of the random vector $\mathbf{r}$, but that its density belongs to a family of distributions $\mathcal{P}$ defined by

$$\mathcal{P} = \{\mathbf{r} \,|\mathbb{E}(\mathbf{r}) = \mu, Cov(\mathbf{r}) = \Sigma\}, \tag{15}$$

where Σ is a positive definite matrix.

Instead of assuming the precise distribution of the random vector r, Zhu and Fukushima (2009) consider the case where the probability distribution π is only known to belong to a certain set, say $\mathcal{P}$, defined the worst-case CVaR (WCVaR) as the CVaR when the worst-case probability distribution in the set $\mathcal{P}$ occurs.

Definition 1. *Given a confidence level β, $\beta \in (0,1)$, the worst-case CVaR for fixed $w \in \mathcal{W}$ with respect to the uncertainty set $\mathcal{P}$ is defined as*

$$\begin{aligned} WCVaR_{\beta}(\mathbf{w}) &\equiv \sup_{\pi \in \mathcal{P}} CVaR_{\beta}(\mathbf{w}) \\ &= \sup_{\pi \in \mathcal{P}} \min_{\alpha \in \mathbb{R}} F_{\beta}(\mathbf{w}, \alpha). \end{aligned} \tag{16}$$

Zhu and Fukushima (2009) further investigated the WCVaR risk measure with several structures of uncertainty in the underlying distribution. In particular, Zhu and Fukushima (2009) consider the case where the distribution of r belong to a set of distributions consisting of all mixtures of some possible distribution scenarios, i.e.,

$$p(\cdot) \in \mathcal{P}_M \equiv \left\{ \sum_{i=1}^{d} \pi_i p^i(\cdot) : \sum_{i=1}^{d} \pi_i = 1,\ \pi_i \geq 0,\ i = 1, ..., d \right\}, \tag{17}$$

where $p^i(\cdot)$ denotes the j-th likelihood distribution and define

$$\begin{aligned} G_{\beta}(\mathbf{w}, \alpha, \pi) &= \alpha + \frac{1}{1-\beta} \int_{\mathbf{r} \in \mathbb{R}^n} (f(\mathbf{w}, \mathbf{r}) - \alpha)^{+} \sum_{i=1}^{d} \pi_i p^i(\mathbf{r}) d\mathbf{r},\ j = 1, ..., J \\ &= \sum_{i=1}^{d} \pi_i G_{\beta}^{i}(\mathbf{w}, \alpha),\ i = 1, ..., d, \end{aligned} \tag{18}$$

where

$$G_{\beta}^{i}(\mathbf{w}, \alpha) = \alpha + \frac{1}{1-\beta} \int_{\mathbf{r} \in \mathbb{R}^n} (f(\mathbf{w}, \mathbf{r}) - \alpha)^{+} p^i(\mathbf{r}) d\mathbf{r},\ i = 1, ..., d. \tag{19}$$

By theorem 1 of Zhu and Fukushima (2009) we can state that for each fixed $\mathbf{w} \in \mathcal{W}$,

$$\begin{aligned} WCVaR_{\beta}(\mathbf{w}) &= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} G_{\beta}(\mathbf{w}, \alpha, \pi) \\ &= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} \sum_{i=1}^{d} \pi_j G_{\beta}^{i}(\mathbf{w}, \alpha). \end{aligned} \tag{20}$$

Thus, minimizing the worst-case CVaR over $w \in \mathcal{W}$ is equivalent to the following min-min-sup optimization problem:

$$\min_{\mathbf{w} \in \mathcal{W}} WCVaR_{\beta}(\mathbf{w}) = \min_{\mathbf{w} \in \mathcal{W}} \min_{\alpha \in \mathbb{R}} \sup_{\pi \in \Pi} G_{\beta}(\mathbf{w}, \alpha, \pi). \tag{21}$$

Zhu and Fukushima (2009) also proved that WCVaR is a coherent risk measure and it clearly satisfies $WCVaR_{\beta}(\mathbf{w}) \geq CVaR_{\beta}(\mathbf{w}) \geq VaR_{\beta}(\mathbf{w})$. Thus, $WCVaR_{\beta}(\mathbf{w})$ can be effectively used as a risk measure.

In the next subsections we provide a brief review of copulas and consider the augmented framework of Kakouris and Rustem (2014) to compute WCVaR through the use of multidimensional mixture archimedean copulas.

2.4 Copulas

Copulas are often defined as multivariate distribution functions whose marginals are uniformly distributed on $[0,1]$. In other words, a copula C is a function such that

$$C\left(u_1,...,u_n\right) = P\left(U_1 \leq u_1,...,U_n \leq u_n\right), \tag{22}$$

where $U_i \sim U[0,1]$ and u_i are realizations of U_i, $i = 1,...,n$. The margins u_i can be replaced by $F_i\left(x_i\right)$, where x_i, $i = 1,...,n$ is a realization of a (continuous) random variable, since they both belong to the domain $[0,1]$ and are uniformly distributed by its probability integral transform (note that $P\left(F\left(x\right) \leq u\right) = P\left(x \leq F^{-1}\left(u\right)\right) = F\left(F^{-1}\left(u\right)\right) = u$). Therefore, copulas can be used to model the dependence structure and margins separately, and therefore provide more flexibility.

Formally, we can define a copula function C as follows.

Definition 2. *An n-dimensional copula (or simply n-copula) is a function C with domain $[0,1]^n$, such that:*

1. C is grounded an n-increasing;
2. C has marginal distributions C_k, $k = 1,...,n$, where $C_k(\mathbf{u}) = u$ for every $\mathbf{u} = (u_1,...,u_n)$ in $[0,1]^n$.

Equivalently, an n-copula is a function

$$C : [0,1]^n \rightarrow [0,1]$$

with the following properties:

(i) (grounded) For all $\mathbf{u}$ in $[0,1]^n$, $C(\mathbf{u}) = 0$, if at least one coordinate of $\mathbf{u}$ is 0 and $C(\mathbf{u}) = u_k$, if all the coordinates of $\mathbf{u}$ are 1 except u_k;

(ii) (n-increasing) For all $\mathbf{a}$ and $\mathbf{b}$ in $[0,1]^n$ such that $a_i \leq b_i$, for every i, $V_C\left([a,b]\right) \geq 0$, where V_c is called $C-$volume.

One of the main results of the theory of copulas is Sklar's Theorem Sklar (1959).

Theorem 1. *(Sklar's Theorem) Let $X_1,...,X_n$ be random variables with distribution functions $F_1,...,F_n$, respectively. Then, there exists an n-copula C such that,*

$$F\left(x_1,...,x_n\right) = C\left(F_1\left(x_1\right),...,F_n\left(x_n\right)\right), \tag{23}$$

for all $\mathbf{x} = (x_1,...,x_n) \in \mathbb{R}^n$. If $F_1,...,F_n$ are all continuous, then the function C is unique; otherwise C is determined only on $\operatorname{Im} F_1 \times ... \times \operatorname{Im} F_n$. Conversely, if C is an n-copula and $F_1,...,F_n$ are distribution functions, then the function F defined above is an $n-$dimensional distribution function with marginals $F_1,...,F_n$.

Corollary 1.1. *Let F be an n-dimensional distribution function with continuous marginals $F_1,...,F_n$, and copula C. Therefore, for any $\mathbf{u} = (u_1,...,u_n)$ in $[0,1]^n$,*

$$C\left(u_1,...,u_n\right) = F\left(F_1^{-1}\left(u_1\right),...,F_n^{-1}\left(u_n\right)\right), \tag{24}$$

where F_i^{-1}, $i = 1,...,n$ are the quasi-inverses of the marginals.

Using the Sklar's theorem and its corollary we can derive an important relation between the probability density functions and copulas. Let f be a joint density function (of the $n-$dimensional distribution function F) and $f_1, ..., f_n$ univariate density functions of the margins $F_1, ..., F_n$. Assuming that $F(\cdot)$ and $C(\cdot)$ are differentiable, by (23) and (24)

$$\frac{\partial^n F(x_1, ..., x_n)}{\partial x_1 ... \partial x_n} \equiv f(x_1, ..., x_n) = \frac{\partial^n C(F_1(x_1), ..., F_n(x_n))}{\partial x_1 ... \partial x_n} \tag{25}$$

$$= c(u_1, ..., u_n) \prod_{i=1}^{n} f_i(x_i). \tag{26}$$

From a modelling perspective, Sklar's Theorem allow us to separate the modeling of the marginals $F_i(x_i)$ from the dependence structure, represented in C. The copula probability density function

$$c(u_1, ..., u_n) = \frac{f(x_1, ..., x_n)}{\prod_{i=1}^{n} f_i(x_i)} \tag{27}$$

is the ratio of the joint probability function to what it would have been under independence. Thus, we can interpret the copula as the adjustment that we need to make to convert the independence probability density function into the multivariate density function. In other words, copulas decompose the joint probability density function from its margins. Now we can estimate the multivariate distribution in two parts: (i) finding the marginal distributions; (ii) finding the dependency between the filtered data from (i).

Thereafter, copulas accommodate various forms of dependence through suitable choice of the copula "correlation matrix" since they conveniently separate marginals from dependence component. The methodology allows one to derive joint distributions from marginals, even when these are not normally distributed. In fact, copulas allow the marginal distributions to be modeled independently from each other, and no assumption on the joint behavior of the marginals is required, which provides a great deal of flexibility in modeling joint distributions.

The reader interested in the history of copulas is referred to Schweizer and Sklar (2011) and Nelsen (2006). For an extensive list of parametric copula families see Joe (1997) and Nelsen (2006) and references therein. After late 1990's the theory and application of copulas grew enormously. Cherubini, Luciano, and Vecchiato (2004) accounts for developments in finance and stochastic processes.

Here we focus in a special class of copulas called Archimedean. An Archimedean copula has the form

$$C(u_1, ..., u_n) = \psi\left(\psi^{-1}(u_1) + \psi^{-1}(u_2) + ... + \psi^{-1}(u_n)\right), \tag{28}$$

for an appropriate generator $\psi(\cdot)$, where $\psi : [0, \infty] \to [0, 1]$ and satisfies (i) $\psi(0) = 1$ and $\psi(\infty) = 0$; (ii) ψ is d-monotone, i.e., $(-1)^k \mathrm{d}^k \psi(s) \mathrm{d}s^k \geq 0$ for $k \in \{0, ..., d-2\}$ and $(-1)^{d-2} \mathrm{d}^{d-2} \psi(s) \mathrm{d}s^{d-2}$ is decreasing and convex. Most of the copulas in this class have a closed form. Moreover, each member has a single parameter that controls the degree of dependence, which allow modeling dependence in arbitrarily high dimensions with only one parameter.

The concept of tail dependence can be especially useful when dealing with co-movements of assets. Studies on financial markets show that financial assets have asymmetric distributions and heavy tails. Often elliptical copulas, like gaussian and t-student copulas are used in Finance. However, these copulas suffer from an absent or symmetric lower and upper tail dependence, respectively. Therefore, we need copulas that capture better the stylized facts and thus give a more complete description of the joint distribution.

In this paper, the worst case Copula-CVaR is achieved through the use of a convex linear combination of archimedean copulas consisting of the best mixture of Clayton, Frank and Gumbel copulas. This mixture is chosen because these archimedean copulas contain different tail dependence characteristics. It combines a copula with lower tail dependence, a copula with positive or negative dependence and a copula with upper tail dependence to produce a more flexible copula capable of modelling the multivariate log returns. Hence, by using a mixture copula we cover a wider range of possible dependencies structures within a single model capturing better the dependence between the individual assets which strongly influences the risk measures.

2.5 Worst Case Copula-CVaR

Similarly to Kakouris and Rustem (2014) and using the previously defined notations, let the decision vector be $\mathbf{w} = (w_1, ..., w_n)^\top$, $\mathbf{u} = (u_1, ..., u_n)$ in $[0,1]^n$ be an stochastic vector, $h(\mathbf{w}, \mathbf{u}) = h(\mathbf{w}, F(\mathbf{x}))$ the loss function and $F(\mathbf{x}) = (F_1(x_1), ..., F_n(x_n))^\top$ a set of marginal distributions. Also, assume that $\mathbf{u}$ follows a continuous distribution with copula $C(\cdot)$.

Given a fixed $\mathbf{w} \in \mathcal{W}$, a random vector $\mathbf{x} \in \mathbb{R}^n$ and the equation (26), the probability that $h(\mathbf{w}, \mathbf{x})$ does not exceed a threshold α is represented by

$$\begin{aligned} P(h(\mathbf{w}, \mathbf{x}) \leq \alpha) &= \int_{h(\mathbf{w}, \mathbf{x}) \leq \alpha} f(\mathbf{x})\, d\mathbf{x} \\ &= \int_{h(\mathbf{w}, \mathbf{x}) \leq \alpha} c(F(\mathbf{x})) \prod_{i=1}^{n} f_i(x_i)\, d\mathbf{x} \\ &= \int_{h(\mathbf{w}, \mathbf{F}^{-1}(\mathbf{u})) \leq \alpha} c(\mathbf{u})\, d\mathbf{u} \\ &= C\left(\mathbf{u} \middle| \widetilde{h}(\mathbf{w}, \mathbf{u}) \leq \alpha\right), \end{aligned}$$

where $f_i(x_i) = \frac{\partial F_i(x_i)}{\partial x_i}$, $\mathbf{F}^{-1}(\mathbf{u}) = \left(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)\right)^+$ and $\widetilde{h}(\mathbf{w}, \mathbf{u}) = h\left(\mathbf{w}, \mathbf{F}^{-1}(\mathbf{u})\right)$. Thus, we can represent the VaR_β by

$$VaR_\beta(\mathbf{w}) = \arg\min \left\{\alpha \in \mathbb{R} : C\left(\mathbf{u} \middle| \widetilde{h}(w, \mathbf{u}) \leq \alpha\right) \geq \beta\right\} \tag{29}$$

and following (5) we can define $CVaR_\beta$ by

$$CVaR_\beta(\mathbf{w}) = \frac{1}{1-\beta} \int_{\widetilde{h}(\mathbf{w}, \mathbf{u}) \geq VaR_\beta(\mathbf{w})} \widetilde{h}(\mathbf{w}, \mathbf{u})\, c(\mathbf{u})\, d\mathbf{u}, \tag{30}$$

and by (9) we can write

$$H_\beta(\mathbf{w}, \alpha) = \alpha + \frac{1}{1-\beta} \int_{\mathbf{u} \in [0,1]^n} \left(\widetilde{h}(\mathbf{w}, \mathbf{u}) - \alpha\right)^+ c(\mathbf{u})\, d\mathbf{u}. \tag{31}$$

Zhu and Fukushima (2009) derived the WCVaR considering a mixture of distributions in a prescribed set as we have seen in equations (16) – (21). Kakouris and Rustem (2014) extended their framework considering a set of copulas $C(\cdot) \in \mathcal{C}$.

Let

$$C(\cdot) \in \mathcal{C}_M \equiv \left\{ \sum_{i=1}^{d} \pi_i C_i(\cdot) : \sum_{i=1}^{d} \pi_i = 1,\ \pi_i \geq 0,\ i = 1, ..., d \right\}, \tag{32}$$

and similarly to (18)

$$\begin{aligned} H_\beta(\mathbf{w}, \alpha, \pi) &= \alpha + \frac{1}{1-\beta} \int_{\mathbf{u} \in [0,1]^n} \left(\widetilde{h}(\mathbf{w}, \mathbf{u}) - \alpha \right)^+ \sum_{i=1}^{d} \pi_i c_i(\mathbf{u})\, d\mathbf{u},\ i = 1, ..., d & (33) \\ &= \sum_{i=1}^{d} \pi_i H_\beta^i(\mathbf{w}, \alpha),\ i = 1, ..., d, & (34) \end{aligned}$$

where

$$H_\beta^i(\mathbf{w}, \alpha) = \alpha + \frac{1}{1-\beta} \int_{\mathbf{u} \in [0,1]^n} \left(\widetilde{h}(\mathbf{w}, \mathbf{u}) - \alpha \right)^+ c_i(\mathbf{u})\, d\mathbf{u},\ i = 1, ..., d. \tag{35}$$

Invoking theorem 1 of Zhu and Fukushima (2009), again we can state that for each fixed $\mathbf{w} \in \mathcal{W}$ the $WCVaR_\beta(\mathbf{w})$ with respect to the set $\mathcal{C}$ is represented by

$$\begin{aligned} WCVaR_\beta(\mathbf{w}) &= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} H_\beta(\mathbf{w}, \alpha, \pi) & \\ &= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} \sum_{i=1}^{d} \pi_i H_\beta^i(\mathbf{w}, \alpha). & (36) \end{aligned}$$

Thus, the Worst Case Copula-CVaR with respect to $\mathcal{C}$ is the mixture copula that produces the greatest CVaR, i.e., the worst performing copula combination in the set $\mathcal{C}$. Moreover, minimizing the worst-case Copula-CVaR over $\mathbf{w} \in \mathcal{W}$ can be defined as the following optimization problem:

$$\min_{\mathbf{w} \in \mathcal{W}} WCVaR_\beta(\mathbf{w}) = \min_{\mathbf{w} \in \mathcal{W}} \min_{\alpha \in \mathbb{R}} \sup_{\pi \in \Pi} H_\beta(\mathbf{w}, \alpha, \pi). \tag{37}$$

Using the approach of Rockafellar and Uryasev (2000) the integral in (35) is approximated by sampling realizations from the copulas $C_i(\cdot) \in \mathcal{C}$ using as inputs the filtered uniform margins. If the sampling generates a collection of values $\left(\mathbf{u}_i^{[1]}, \mathbf{u}_i^{[2]}, ..., \mathbf{u}_i^{[J]}\right)$, where $\mathbf{u}_i^{[j]}$ and S^i are the j-th sample drawn from copula $C_i(\cdot)$ of the mixture copula using as inputs the filtered uniform margins and its corresponding size, respectively, $i = 1, ..., d$, we can approximate $H_\beta^i(\mathbf{w}, \alpha)$ by

$$\widetilde{H}_\beta^i(\mathbf{w}, \alpha) = \alpha + \frac{1}{(1-\beta) S^i} \sum_{j=1}^{S^i} \left(\widetilde{h}(\mathbf{w}, \mathbf{u}_i^{[j]}) - \alpha \right)^+, i = 1, .., d \tag{38}$$

Assuming that the allowable set $\mathcal{W}$ is convex and the loss function $\widetilde{h}(\mathbf{w}, \mathbf{u})$ is linear with respect to $\mathbf{w}$ then optimization problem

$$\min_{\mathbf{w} \in \mathcal{W}, \alpha \in \mathbb{R}} \widetilde{H}_\beta(\mathbf{w}, \alpha). \tag{39}$$

reduces to the following LP problem:

$$\underset{\mathbf{w}\in\mathbb{R}^n,v\in\mathbb{R}^m,\alpha\in\mathbb{R}}{\text{minimize}}\ \alpha + \frac{1}{(1-\beta)\,S^i}\sum_{i=1}^{S^i} v_i \tag{40}$$

$$\text{subject to}\quad \mathbf{w} \in \mathcal{W}, \tag{41}$$

$$v_i^{[j]} \geq \widetilde{h}(\mathbf{w}, \mathbf{u}_i^{[j]}) - \alpha, \tag{42}$$

$$v_i^{[j]} \geq 0,\ j = 1, \ldots, J;\ i = 1, \ldots, d, \tag{43}$$

where v_i, $i = 1, \ldots, d$, are auxiliary indicator (dummy) variables and $m = \sum_{i=1}^{d} S^i$. By solving the LP problem we find the optimal decision vector, $\mathbf{w}^*$, and at "one shot" the optimal VaR, α^*, and the optimal $CVaR$, $\widetilde{H}_\beta\,(\mathbf{w} = \mathbf{w}^*, \alpha = \alpha^*)$.

3 Methodology

Our data set consists of daily data of adjusted closing prices of all stocks that belong to S&P 500 market index from July 2st, 1990 to December 31st, 2015. We obtain the adjusted closing prices from Bloomberg and log-returns are calculated in excess of a risk-free asset[3]. The data set sample period is made up of 6426 days and includes a total of 1100 stocks over all periods. Only stocks that are listed throughout in-sample (48-month formation period) and out-of-sample (6 months) periods are included in the analysis.

We want a diversified set of stocks that can be useful during crises and tranquil periods. To attain a robust portfolio construction, we select a set of 50 stocks (only S&P 500 constituents) in each formation period, based on the ranked sum of absolute spreads[4] (the twenty-five largest and smallest) between the normalized daily closing prices deviations of the S&P 500 index and all shares. The distances are computed using data from January to December or from July to June. Prices are scaled to 1 at the beginning of each formation period and then evolve using the return series [5]. Specifically, the spread between the normalized closing prices at time t is computed as

$$Spread_t = NP_{i,t} - NP_{SP500,t}, \tag{44}$$

where $NP_{i,t} = NP_{i,t-1}\,(1 + r_{i,t})\,,\ i = 1, ..., N,\ t = 2, ..., T$. We rebalance our portfolio every six months.

Our optimization strategy adopts a sliding window of calibration of T=4 years of daily data (1008 observations in average). Therefore, for example, we use day 1 to 1008 to estimate the parameters of all models and determine portfolio weights for day 1009 and then repeat the process including the latest observation and removing the oldest until reaching the end of the time series. We define L as the number of days in the data set and thus, we compute $L - T - 1$ daily log-returns.

3.1 Strategies under Analysis

To apply Worst Case Copula-CVar Portfolio Optimization we go through the following steps : (1) First we fit an AR(1)-GARCH(1,1) model with skew t-distributed innovations to each univariate time series

[3]We use 3-month Treasury Bill obtained at https://fred.stlouisfed.org/series/TB3MS as a proxy for the risk-free rate.

[4]A L1-norm minimization is used for increase robustness.

[5]Missing values have been interpolated.

selected from distance method; (2) Using the estimated parametric model, we construct the standardized residuals vectors given, for each $i = 1, ..., 50$ and $t = 1, ..., L - T - 1$, by

$$\frac{\widehat{\varepsilon}_{i,t}}{\widehat{\sigma}_{i,t}}.$$

The standardized residuals vectors are then converted to pseudo-uniform observations $z_{i,t} = \frac{n}{n+1} F_i\left(\widehat{\varepsilon}_{i,t}\right)$, where F_i is their empirical distribution function; (3) Estimate the copula model, i.e., fits the multivariate Clayton-Frank-Gumbel (CFG) Mixture Copula to data that has been transformed to [0,1] margins by

$$C^{CFG}(\Theta, \mathbf{u}) = \pi_1 C^C(\theta_1, \mathbf{u}) + \pi_2 C^F(\theta_2, \mathbf{u}) + (1 - \pi_1 - \pi_2) C^G(\theta_3, \mathbf{u}) \tag{45}$$

where $\Theta = (\alpha, \beta, \delta)^\top$ are the Clayton, Frank and Gumbel copula parameters, respectively, and π_1, $\pi_2 \in [0, 1]$. The estimates are obtained by the minimization of the negative log-likelihood consisting of the weighted densities of the Clayton, Frank and Gumbel copulas. Probability density function for multivariate Archimedean copula is computed as described in Mcneil and Neshelova (2009); (4) Use the dependence structure determined by the estimated copula for generating J scenarios. To simulate data from the three Archimedean copulas we use the sampling algorithms provided in Melchiori (2006); (5) compute t-quantiles for these Monte Carlo draws; (6) Compute the standard deviation $\widehat{\sigma}_{i,t}$ using the estimated GARCH model; (7) determine the simulated daily asset log-returns, i.e., determine the simulated daily log-returns as $r_{i,t}^{sim} = \widehat{\mu}_t + \widehat{\sigma}_{i,t} z_{i,t}$; (8) Finally, use the simulated data as inputs when optimizing portfolio weights by minimizing CVaR for a given confidence level and a given minimum expected return.

From Sklar's theorem the Gauss copula is given by

$$C_p\left(u_1, ..., u_d\right) = \mathbf{\Phi}_p\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right), \tag{46}$$

where Φ denotes the standard normal distribution function, and $\mathbf{\Phi}_p$ denotes the multivariate standard normal distribution function with correlation matrix P. We simulate realizations from a Gaussian Copula repeating the following steps J times. We first perform a Cholesky decomposition of P, and set A as the resulting lower triangular matrix. Next we generate a vector $Z = (Z_1, ..., Z_d)'$ of independent standard normal variables and set $X = AZ$. Finally, we compute $U = \left(\Phi(X_1), ..., \Phi(X_d)\right)'$.

For each of the four copulas, we run 1000 return scenarios from the estimated multivariate CFG Mixture Copula and multivariate Gaussian Copula models for each of the fifty assets. The weights are recalibrated at a daily, weekly and monthly basis. We assume that the feasible set $\mathcal{W}$ attends the budget and non-negativity constraints, i.e., the sum of the weights should be 1 ($\mathbf{w}^\top \mathbf{1} = 1$) and no short-selling is allowed ($\mathbf{w} \geq 0$).

The expected return should be bound below by an arbitrary value $\bar{r}$, i.e., $\mathbf{w}^\top \mathbb{E}\left(\mathbf{r}_{p,t+1}\right) \geq \bar{r}$, where $\bar{r}$ represents the target mean return.. The confidence level β is set at $\beta = 0.95$.

We also consider three benchmarks: the Gaussian Copula portfolio, the equally weighted naive or 1/N portfolio, where $\mathbf{w} = \left(\frac{1}{N}, ..., \frac{1}{N}\right)^\top$ for any rebalancing date t and the S&P 500 index as a proxy for market return.

3.2 Performance Measures

We assess out-of-sample portfolio allocation performance and its associated risks by means of the following statistics: mean excess returns, standard deviation, maximum drawdown between two consecutive days and between two days within a period of maximum six months, Sharpe ratio, Sortino ratio, turnover, breakeven costs, $\mathrm{VaR}_{0.95}$,$\mathrm{CVaR}_{0.95}$ and $\mathrm{CR}_{0.95}$.

For each strategy, we compute the optimal weights $\mathbf{w}_t$ for each t using the moving calibration window. We compute the portfolio excess return in time t, $t = T+1, ..., L-1$ by

$$\widehat{r}_{p,t} = \mathbf{w}_{t-1}^{\top} \mathbf{r}_{p,t} - r_{f,t}, \tag{47}$$

and the portfolio mean excess return as

$$\widehat{\mu}_p = \frac{1}{L-T} \sum\nolimits_{t=T+1}^{L-1} \widehat{r}_{p,t} \tag{48}$$

The portfolio standard deviation and Sharpe ratio are given, respectively, by

$$\widehat{\sigma}_p = \sqrt{\frac{1}{L-T-2} \sum\nolimits_{t=T}^{L-1} \left(w_t r_{p,t+1} - \widehat{\mu}_p\right)^2}, \tag{49}$$

and

$$SR = \frac{\widehat{\mu}_p}{\widehat{\sigma}_p}. \tag{50}$$

Denote now by $r_p(\mathbf{w},t)$ the cumulative portfolio return at time t, where $\mathbf{w}$ are asset weights in the portfolio. The drawdown function at time t (Unger (2014)) is the defined as the difference between the maximum of the function $r_p(\mathbf{w},t)$ over the history preceding time t and the value of this function at time t, i.e.,

$$D(\mathbf{w},t) = \max_{0 \le \tau \le t} \{r_p(\mathbf{w},t)\} - r_p(\mathbf{w},t). \tag{51}$$

The maximum drawdown on the interval $[0, T]$ is defined as

$$MaxDD(\mathbf{w}) = \max_{0 \le \tau \le t} \{D(\mathbf{w},t)\}. \tag{52}$$

In other words, the maximum drawdown over a period is the maximum loss from worst peak to a trough of a portfolio drop from the start to the end of the period.

The Sortino's ratio (Sortino and Price (1994)) is the ratio of the mean excess return to the standard deviation of negative asset returns, i.e.,

$$SoR = \frac{\widehat{\mu}_p}{\widehat{\sigma}_{p,n}}, \tag{53}$$

where

$$\widehat{\sigma}_{p,n} = \sqrt{\frac{1}{L-T-1} \sum\nolimits_{t=T}^{L-1} \left(\min\left(0, \mathbf{w}_t^{\top} \mathbf{r}_{t+1} - \mathbf{r}_{MAR}\right)\right)^2}, \tag{54}$$

where $\mathbf{r}_{MAR}$ is the value of a minimal acceptable return (MAR), usually zero or the risk-free rate[6].

We define the portfolio turnover from time t to time $t+1$ as the sum of the absolute changes in the N risky asset weights, i.e., in the optimal values of the investment fractions:

$$Turnover = \frac{1}{L-T-1} \sum\nolimits_{t=T}^{L-1} \sum\nolimits_{j=1}^{N} \left(|w_{j,t+} - w_{j,t+1}| \right), \tag{55}$$

where $w_{j,t+}$ is the actual weight in asset j before rebalancing at time $t+1$, and $w_{j,t+1}$ is the optimal weight in asset j at time $t+1$. Turnover measures the amount of trading required to implement a particular portfolio strategy and can be interpreted as the average fraction of wealth traded in each period.

We also report the break-even transaction cost proposed by Bessembinder and Chan (1995), which is the level of transaction costs leading to zero net profits, i.e., the maximum transaction cost that can be imposed before making the strategies less desirable than the buy-and-hold strategy (see Han (2006)). Following Santos and Tessari (2012) we consider the average net returns on transaction costs, $\widehat{\mu}_{TC}$, given by

$$\widehat{\mu}_{TC} = \frac{1}{L-T} \sum\nolimits_{t=T}^{L-1} \left[\left(1 + \mathbf{w}_t^\top \mathbf{r}_{t+1}\right) \left(1 - c \sum\nolimits_{j=1}^{N} \left(|w_{j,t+1} - w_{j,t+}| \right) \right) - 1 \right], \tag{56}$$

where c is called breakeven cost when we solve $\widehat{\mu}_{TC} = 0$. Those portfolios that achieve a higher break-even cost are preferable, since the level required to make these portfolios non-profitable are higher.

The Capital requirement loss function (CR) is a regulatory loss function to evaluate VaR forecasts, given by:

$$CR_t = \max \left[\left(\left(\frac{3+\delta}{60} \right) \sum_{i=0}^{59} VaR_{\beta,t-i} \right), VaR_{\beta,t} \right], \tag{57}$$

where $\delta = \begin{cases} 0, \text{if } \zeta \leq 4 \\ 0.3 + 0.1\,(\zeta - 4)\,, \text{if } 5 \leq \zeta \leq 6 \\ 0.65, \text{if } \zeta = 7 \\ 0.65 + 0.1\,(\zeta - 7)\,, \text{if } 8 \leq \zeta \leq 9 \\ 1, \text{if } \zeta \geq 10 \end{cases}$

is a multiplicative factor that depends on the number of violations of the VaR in the previous 250 trading days (ζ). Note that the CR is based on the larger out of the current VaR estimate and a multiple of the average estimate over the past 60 days.

We compute the capital requirement losses based on the 95% daily VaR forecasts, and in order to mitigate data-snooping problems we apply a test for superior predictive ability (SPA) proposed by Hansen (2005) to test which model significantly minimizes the expected loss function (or has a superior predictive ability, in other words). The relative performance of the Mixed Copula-CVaR portfolio to the Gaussian Copula-CVaR portfolio may be defined as

$$d_{m,t} = CR_{g,t} - CR_{m,t},$$

[6]Sortino Ratio is an improvement on the Sharpe Ratio since it is more sensitive to extreme risks or downside than measures Sharpe Ratio. Sortino contends that risk should be measured in terms of not meeting the investment goal. The Sharpe ratio penalizes financial instruments that have a lot of upward jumps, which investors usually view as a good thing.

where m and g stands for Mixed and Gaussian Copula CVaR portfolios, respectively. The null hypothesis is that the average performance of the Gaussian Copula CVaR portfolio is as small as the minimum average performance of the Worst Case Copula CVaR portfolio. Provided that $E\left(d_{m,t}\right) = \mu_{m,t}$ is well defined, we can formulate the null hypothesis as

$$\begin{cases} H_0 : \mu_{m,t} \leq 0 \\ H_1 : \mu_{m,t} > 0. \end{cases}$$

The test statistic is given by

$$T^{SPA} = \max\left[\frac{\sqrt{T}\overline{d}_{k,t}}{\widehat{w}_k}, 0\right],$$

where $\overline{d}_{k,t} = T^{-1}\sum_{t=1}^{T} d_{k,t}$ e $\widehat{w}_k^2$ is some consistent estimator of w_k^2, the variance of $\sqrt{T}\overline{d}_{k,t}$. To construct the distributions we bootstraped the original time series $B = 10000$ times and select the optimal block length for the stationary bootstrap following Politis and White (2004). Our bootstrapped null distributions result from Theorem 2 of Politis and Romano (1994). Since the optimal bootstrap block length is different for each strategy we average the block lengths found to proceed the comparisons between the strategies.

4 Empirical Results

Table 1 reports out-of-sample mean excess return, standard deviation, Sharpe ratio, Sortino ratio, turnover, break-even transaction costs, maximum drawdown between two consecutive days (MDD1) and between two days within a period of maximum six months (MDD2), $\mathrm{VaR}_{0.95}$, $\mathrm{CVaR}_{0.95}$ and $\mathrm{CR}_{0.95}$ of the different portfolio strategies from 1991/2 to 2015/2 involving daily, weekly and monthly rebalacing frequencies. Returns, standard deviation, Sharpe ratio and Sortino ratio are summarized in an annualized basis.

By selecting a diversified set of assets over a long-term period we found that copula-based approaches offer better hedges against losses than the 1/N portfolio. The WCCVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR for daily and weekly rebalancing.

By analyzing Table 1, it is clear that the copula-based allocations consistently outperform the other benchmarks in terms of risk-return adjusted metrics, volatility and maximum drawdown, although the results are somewhat sensitive to the rebalancing window. We can also note that the Worst Case Copula-CVaR has a superior performance than the Gaussian Copula-CVaR in terms of the downside (tail) risk metrics as VaR, CVaR and CR.

Although our main objective is to protect investors from the worst possible scenarios it is worthwhile noticing that the WCCVaR strategy consistently outperforms the other strategies since July 1994 in terms of cumulative excess returns for any daily and weekly rebalancing frequencies without a minimum expected return constraint [7]. However, the breakeven costs show that the profits of the copula strategies

[7] The numerical experiments show that the performances out-of-sample stay very similar when we consider a daily constraint $\overline{r} \geq VaR_{10\%}$. Since the resuls are very much alike they are not presented here and are available under request.

do not survive for transactions costs greater than 6.2 basis points.

Table 1: Excess returns of Worst Case Copula-CVaR (WCCVaR), Gaussian Copula-CVaR (GCCVaR), Equal Weights (1/N) portfolios and S&P 500 market index without a minimum expected return constraint

	WCCVaR	GCCVaR	1/N	S&P 500
Daily Rebalancing				
Mean Return (%)	11.29	10.59	10.52	6.61
Standard Deviation (%)	15.35	15.05	23.86	19.01
Sharpe Ratio	0.70	0.67	0.42	0.34
Sortino Ratio	1.14	1.09	0.68	0.54
Turnover	0.6868	0.6425	0.0001	
Break-even (%)	0.0618	0.0622	275.09	
MDD1 (%)	-13.20	-12.89	-18.00	-13.20
MDD2 (%)	-38.94	-41.66	-71.84	-52.61
$VaR_{0.95}$	-0.0091	-0.0109		
$CVaR_{0.95}$	-0.0109	-0.0135		
$CR_{0.05}$(%)	13.00***	15.63		
Weekly Rebalancing				
Mean Return (%)	10.57	10.04	10.52	6.61
Standard Deviation (%)	15.48	15.21	23.86	19.01
Sharpe Ratio	0.65	0.63	0.42	0.34
Sortino Ratio	1.06	1.02	0.68	0.54
Turnover	0.1535	0.1564	0.0001	
Break-even (%)	0.2596	0.2427	275.09	
MDD1 (%)	-12.59	-10.39	-18.00	-13.20
MDD2 (%)	-39.78	-41.32	-71.84	-52.61
$VaR_{0.95}$	-0.0091	-0.0109		
$CVaR_{0.95}$	-0.0109	-0.0135		
$CR_{0.05}$ (%)	12.86***	15.57		
Monthly Rebalancing				
Mean Return (%)	9.22	9.42	10.52	6.61
Standard Deviation (%)	15.69	15.55	23.86	19.01
Sharpe Ratio	0.56	0.58	0.42	0.34
Sortino Ratio	0.91	0.94	0.68	0.54
Turnover	0.0447	0.0499	0.0001	
Break-even (%)	0.7829	0.7155	275.09	
MDD1 (%)	-13.21	-13.22	-18.00	-13.20
MDD2 (%)	-38.58	-46.13	-71.84	-52.61
$VaR_{0.95}$	-0.0091	-0.0109		
$CVaR_{0.95}$	-0.0109	-0.0134		
$CR_{0.05}$ (%)	12.22***	14.67		

Note: Out-of-sample performance statistics between July 1994 and December 2015 (5414 observations). The rows labeled MDD1 and MDD2 compute the largest drawdown in terms of maximum percentage drop between two consecutive days and between two days within a period of maximum six months, respectively. Returns, standard deviation, Sharpe ratio and Sortino ratio are annualized.
*** , **, * indicate a rejection of the null hypothesis of equal predictive ability at 1%, 5% and 10% levels, respectively, according to the Hansen (2005)'s SPA test.

Up to now we did not consider transaction costs when we purchase and sell the assets (or "turns over" our portfolio). But, if we want to use our strategies for tactical asset allocation, transaction costs play a non-trivial role and must not be neglected. With this in mind, we compute the portfolio turnover of each strategy. The higher the turnover, the higher the transaction cost that the portfolio incurs at each rebalancing day. We can note that the copula-based approaches present a similar turnover in the rebalancing frequencies analyzed. However, the WCCVaR has a relatively better performance when the rebalancing frequency decreases. As expected, the portfolio turnover decreases for both allocations when the holding portfolio period increases and the EWP (1/N) has the smallest turnover.

In addition, we report the break-even transaction cost in order to investigate if the profits are economically significant. The break-even values in Table 1 represent the level of transaction costs leading to zero excess return. Thus, those portfolios that achieve a higher break-even cost are preferable, since the level required to make these portfolios non-profitable are higher[8]. For the WCCVaR and GCCVaR allocations the break-even costs are between [0.062%-0.783%] and [0.062%-0.716%], respectively. Thus, for month rebalancing this means that if the trading costs are anything less than 78 basis points the excess profits

[8] Jegadeesh and Titman (1993) consider a conservative transaction cost of 0.5% per trade, while Allen and Karjalainen (1999) consider 0.1%, 0.25% and 0.5%.

for the WCCVaR allocations are still greater than zero. We can also note that the break-even costs for WCCVaR and GCCVaR are similar, but that the difference increases when the holding period lengthens, as expected after analyzing the portfolio turnover.

Finally, we reject for all rebalancing frequencies the null hypothesis of the Hansen (2005)'s SPA test (p-values=0.0000) that the average performance of the Gaussian Copula CVaR portfolio is as small as the minimum average performance of the Worst Case Copula CVaR portfolio. In other words, we find that WCCVaR allocations have a superior predictive ability when compared to the GCCVaR allocations.

Figure 1 depicts the wealth trajectories of the portfolios strategies for all rebalancing frequencies assuming an initial wealth of \$1 monetary unit. Panels (a) to (c) show the excess returns for daily, weekly and monthly rebalancing, respectively.

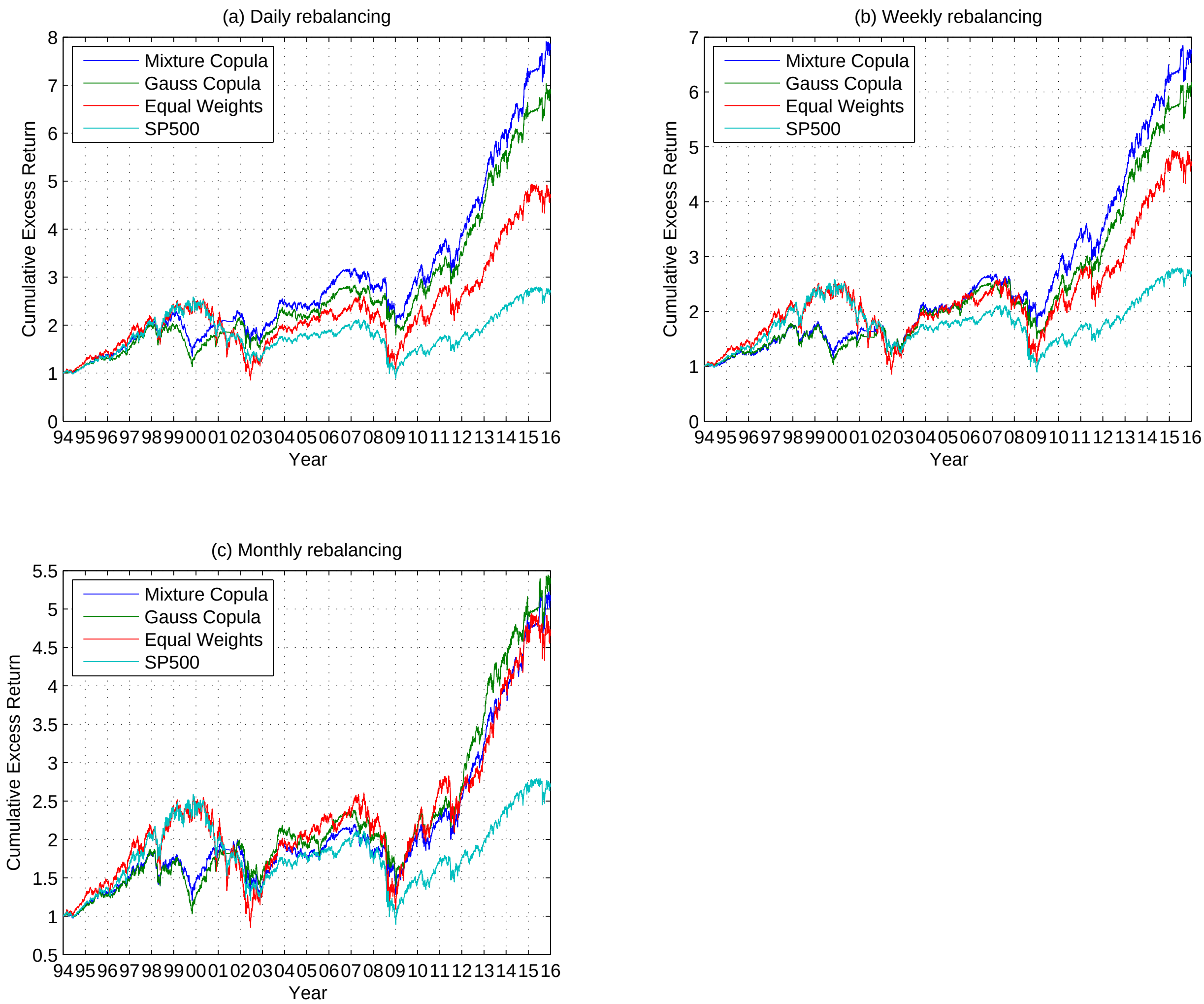

Figure 1: Cumulative excess returns of the portfolio strategies without daily target mean return

This figure shows how an investment of \$1 evolves from July 1994 to December 2015 for each of the portfolios.

The copula-based approaches are more profitable for daily and weekly rebalancing over time, especially after 2009. We can note the copula methods present a hump-shaped pattern in 1999, while the other benchmarks show a sharp decline in the subperiod that corresponds to the bear market that comprises the dotcom crisis and the September 11th terrorist attack (2000-2002). All portfolios show a hump-shaped pattern during the subprime mortgage financial crisis in 2007-2008. Overall, we can observe that after 2002 the patterns are similar, but the figure indicates that the copula methods, even though the objective function is the minimization of CVaR under a constraint on expected return, preserve more wealth in the long-term period, particularly the WCCVaR portfolio, for daily and weekly rebalancing [9].

5 Conclusions

In this paper we combine robust portfolio optimization and copula-based models in a Worst Case CVaR framework. To cope with the large number of financial instruments we employ a procedure similar to that used by Gatev, Goetzmann, and Rouwenhorst (2006) to select a set of diversified assets that can be useful during crises and tranquil periods. Using data from the S&P 500 stocks from 1990 to 2015 we evaluate the performance of the WCCVaR (Worst Case Copula-CVaR) portfolio, considering different rebalancing strategies, and compare it against three benchmarks: a Gaussian Copula-CVaR (GCCVaR) portfolio, an equally weighted portfolio (1/N) and the S&P 500 index.

By selecting a diversified set of assets over a long-term period we found that copula-based approaches offer better hedges against losses than the 1/N portfolio. Moreover, the WCCVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR for daily and weekly rebalancing.

Finally, we offer some suggestions for future research for improving our discoveries. First, we could improve the method of asset selection. We suspect that if we use a measure of nonlinear dependence between random variables of arbitrary dimension such as the randomized dependency coefficient (Lopez-Paz, Hennig, and Schölkopf (2013)) or a procedure based on data mining tools as random forest Giovanni De Luca and Zuccolotto (2010) to select the stocks the portfolio performances would be even better. Additional suggestions include relaxing the assumption of no short selling and incorporate transaction cost as an additional constraint in the optimization problem as in Krokhmal, Palmquist, and Uryasev (2002).

References

Alexander, G. J., and A. M. Baptista. 2002. Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis. *Journal of Economic Dynamics and Control* 26:1159–1193.

Allen, F., and R. Karjalainen. 1999. Using genetic algorithms to find technical trading rules. *Journal of financial Economics* 51:245–271.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. 1999. Coherent measures of risk. *Mathematical finance* 9:203–228.

Bessembinder, H., and K. Chan. 1995. The Profitability of Technical Trading Rules in the Asian Stock Markets. *Pacific-Basin Finance Journal* 3:257–284.

Cherubini, U., E. Luciano, and W. Vecchiato. 2004. *Copula methods in finance.* John Wiley & Sons.

[9] Transaction costs should not be neglected as we could notice after the breakeven analysis

Gatev, E., W. N. Goetzmann, and K. G. Rouwenhorst. 2006. Pairs Trading: Performance of a Relative-Value Arbitrage Rule. *Review of Financial Studies* 19:797–827.

Giovanni De Luca, G. R., and P. Zuccolotto. 2010. Combining random forest and copula functions: a heuristic approach for selecting assets from a financial crisis perspective. *Intelligent Systems in Accounting, Finance and Management* 17.2:91–109.

Han, Y. 2006. Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model. *Review of Financial Studies* 19:237–271.

Hansen, P. R. 2005. A test for superior predictive ability. *Journal of Business & Economic Statistics* 23:365–380.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance* 48:65–91.

Joe, H. 1997. *Multivariate models and multivariate dependence concepts*. CRC Press.

Kakouris, I., and B. Rustem. 2014. Robust portfolio optimization with copulas. *European Journal of Operational Research* 235:28–37.

Krokhmal, P., J. Palmquist, and S. Uryasev. 2002. Portfolio Optimization with Conditional Value-At-Risk Objective and Constraints. *Journal of Risk* 4:43–68.

Lopez-Paz, D., P. Hennig, and B. Schölkopf. 2013. The Randomized Dependence Coefficient. In *Advances in Neural Information Processing Systems*, pp. 1–9.

Markowitz, H. 1952. Portfolio selection. *The journal of finance* 7:77–91.

Mcneil, A. J., and J. Neshelova. 2009. Multivariate Archimedean Copulas, d-Monotone Functions and ℓ1-norm Symmetric Distributions. *Annals of statistics* 37:3059–3097.

Melchiori, M. R. 2006. Tools for sampling multivariate archimedean copulas .

Nelsen, R. B. 2006. *An introduction to copulas, 2nd*. New York: Springer Science Business Media.

Pflug, G. C. 2000. Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization*, pp. 272–281. Springer.

Politis, D. N., and J. P. Romano. 1994. The stationary bootstrap. *Journal of the American Statistical association* 89.428:1303–1313.

Politis, D. N., and H. White. 2004. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23.1:53–70.

Rockafellar, R. T., and S. Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of risk* 2:21–42.

Rockafellar, R. T., and S. Uryasev. 2002. Conditional value-at-risk for general loss distributions. *Journal of banking & finance* 26:1443–1471.

Santos, A. A. P., and C. Tessari. 2012. Técnicas Quantitativas de Otimização de Carteiras Aplicadas ao Mercado de Ações Brasileiro (Quantitative Portfolio Optimization Techniques Applied to the Brazilian Stock Market). *Revista Brasileira de Finanças* 10:369.

Schweizer, B., and A. Sklar. 2011. *Probabilistic Metric Spaces*. Courier Corporation.

Sklar, M. 1959. *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8.

Sortino, F. A., and L. N. Price. 1994. Performance Measurement in a Downside Risk Framework. *The Journal of Investing* 3:59–64.

Unger, A. 2014. *The Use of Risk Budgets in Portfolio Optimization*. Springer.

Uryasev, S. 2013. *Probabilistic Constrained Optimization: Methodology and Applications*, vol. 49. Springer Science & Business Media.

Uryasev, S., and R. T. Rockafellar. 1999. *Optimization of Conditional Value-At-Risk*. Department of Industrial & Systems Engineering, University of Florida.

Uryasev, S., and R. T. Rockafellar. 2001. Conditional value-at-risk: optimization approach. In *Stochastic optimization: algorithms and applications*, pp. 411–435. Springer.

Zhu, S., and M. Fukushima. 2009. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations research* 57:1155–1168.