Worst Case Copula-CVaR Performance based on Distance selection criterion

Fernando A. Boeira Sabino da Silva, Flavio A. Ziegelmann, Cristina Tessari

Abstract

Using data from the S&P 500 shares 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 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 Worst Optimal Mean-Variance model and a constant mix portfolio. Our empirical analysis shows that the Copula-CVaR approach is the best performing portfolio with respect to CVaR and yields higher returns than the benchmarks. Particularly, the Copula and benchmark methods show a mean annualized excess return as high as 24.42%, 15.06% and 15.61%. To cope with the dimensionality problem we employ the distance method of 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. To test the statistical significance of the excess returns and Sharpe Ratio we use the stationary bootstrap of Politis and Romano (1994) adopting the automatic block-length selection of Politis and White (2004).

Keywords: Asset Allocation; Copula; Distance; Portfolio Selection; Risk Management; S&P 500; Stationary Bootstrap; WCVaR.

JEL Codes: C15; C52; C53; C61; C63; G11.

1 Introduction

The theory of portfolio selection started with the seminal paper of Markowitz (1952). Markowitz considered the problem of a risk-averse agent who is a wealth maximizer, i.e., 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, i.e., an optimized portfolio that dominate all other feasible portfolios in terms of their risk-return trade-off.

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

---

*Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: fsabino@ufrrgs.br; Corresponding author.
†Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: flavioz@ufrrgs.br
‡Finance and Economics Division, Columbia Business School, Columbia University, New York, NY 10027, USA, e-mail: ct2759@columbia.edu
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 handle a portfolio with a large number of instruments.

Estimation errors of the expected returns of the assets and the covariance matrix of these returns can significantly affect the outcome of the optimization problem, which makes the weight solution sensitive to the input parameters. 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. The performance of these portfolios is compared to a Worst Optimal Mean-Variance model, a naive 1/N (equally-weighted) and the S&P 500 index. 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 12 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.

We look at an optimization model that somehow involves hedging of decisions to protect the investors against any market conditions. To be effective in dealing with uncertainty, we select, among all listed assets in each formation period, a set of 10 stocks based on the ranked sum of squared deviations (the five largest and the five smallest) between the normalized daily closing prices deviations (known as distance) of the S&P 500 index and all shares, adjusting them by dividends, stock splits and other corporate actions. By selecting a diversified set 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.

Returns are calculated in excess of a risk-free asset. Our results indicate that the Worst Case Copula-CVaR approach outperforms consistently the benchmark strategies in the long term in terms of wealth accumulation and downside risk. However, the Worst Case Copula-CVaR allocation has a higher weight on riskier constituents as measured with regard to the portfolios’ return volatilities, maximum drawdown, turnover and breakeven costs over the time period. The Worst Optimal Mean-Variance model exhibit a slightly higher risk-adjusted performance than the other strategies measured by Sharpe and Sortino ratios.

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 $F_Y(u) = P(Y \leq u)$. Let also $F^{-1}_Y(v) = \inf \{u : F_Y(u) \geq v\}$ be its right continuous inverse and assume that it has a probability density function represented by $p(y)^1$. Define the $VaR$ as the $\beta$-quantile by

$$VaR_\beta(Y) = \arg \min \{\alpha \in \mathbb{R} : P(Y \leq \alpha) \geq \beta\}$$

$$= F^{-1}_Y(\beta),$$

and the $CVaR$ as the solution to the following optimization problem (Pflug, 2000):

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

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.,

$$CVaR_\beta(Y) = \mathbb{E}(Y \mid Y \geq VaR_\beta(Y)).$$

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 $(\Omega, \mathcal{F}, P_F)$ having a probability distribution on $(\mathbb{R}, \mathcal{B}, P_B)$ induced by that of $y$, where $(\mathbb{R}, \mathcal{B}, P_B)$ stands for a Borel probability space including, therefore, the open and closed intervals in $\mathbb{R}$. The underlying distribution of $y \in \mathbb{R}^m$ is assumed to have a density $p(y)$ and let the probability of $f(x, y)$ not exceeding some threshold $\alpha$ be denoted by

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

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

$^1$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 $R$. 

3
Using the previously defined notation (2) we can write the CVaR function, at confidence level $\beta$, by

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

(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

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

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

$$\min_{\alpha \in \mathbb{R}} F_\beta(x,\alpha) = F_\beta(x,\text{VaR}_\beta(x)) = CVaR_\beta(x)$$

(7)

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

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

(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 CVaR Minimization with Finite Number of Scenarios

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

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

$$= \alpha + \frac{1}{1 - \beta} \int_{y \in \mathbb{R}} (f(x,y) - \alpha)^+ p(y)dy$$

(9)

by its discretized version

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

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

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

(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:

\[
\min_{x \in \mathbb{R}^n; z \in \mathbb{R}^J; \alpha \in \mathbb{R}} \alpha + \frac{1}{(1 - \beta) J} \sum_{j=1}^{J} z_j \\
\text{s.t. } x \in X, \\
z_j \geq f(x, y_j) - \alpha, \quad z_j \geq 0, \quad j = 1, \ldots, J,
\]

where \( z_j \) are indicator variables. By solving the LP problem above we find the optimal decision vector, \( x^* \), the optimal VaR, \( \alpha^* \), and consequently the optimal CVaR, \( F_\beta(x = x^*, \alpha = \alpha^*) \).

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., \( r = (r_1, \ldots, r_n)^\top \), with expected returns \( \mu = (\mu_1, \ldots, \mu_n)^\top \) and covariance matrix \( \Sigma_{n \times n} \). Let also \( r_f \) represent the risk-free asset returns and the decision (portfolio’s weights) vector by \( w = (w_1, \ldots, w_n)^\top \). Also assume that \( w \in \mathcal{W} \), where \( \mathcal{W} \) is a feasible set and that the portfolio return loss function \( f_L(w, r) \) is a convex (linear) function given by

\[
f_L(w, r) = w^\top r.
\]

By definition, the portfolio return is the negative of the loss, i.e., \( -w^\top r \).

### 2.3 Worst Case CVaR

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

\[
\mathcal{P} = \{ r | E(r) = \mu, \text{Cov}(r) = \Sigma \},
\]

where \( \Sigma \) 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 \( \pi \) 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 \( \beta, \beta \in (0,1) \), the worst-case CVaR for fixed \( w \in \mathcal{W} \) with respect to the uncertainty set \( \mathcal{P} \) is defined as

\[
WCVaR_\beta(w) \equiv \sup_{\pi \in \mathcal{P}} CVaR_\beta(w) = \sup_{\pi \in \mathcal{P}} \min_{\alpha \in \mathbb{R}} F_\beta(w, \alpha)
\]

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.,

\[
p(\cdot) \in \mathcal{P} \equiv \left\{ \sum_{i=1}^{d} \pi_i p^i(\cdot) : \sum_{i=1}^{d} \pi_i = 1, \quad \pi_i \geq 0, \quad i = 1, ..., d \right\},
\]

(16)

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

\[
G_{\beta}(w, \alpha, \pi) = \alpha + \frac{1}{1 - \beta} \int_{\mathbb{R}^n} \left( f(w, r) - \alpha \right)^+ \sum_{i=1}^{d} \pi_i p^i(r) dr, \quad j = 1, ..., J
\]

where

\[
G^j_{\beta}(w, \alpha) = \alpha + \frac{1}{1 - \beta} \int_{\mathbb{R}^n} \left( f(w, r) - \alpha \right)^+ p^j(r) dr, \quad i = 1, ..., d.
\]

(17)

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

\[
WCVaR_{\beta}(w) = \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} G_{\beta}(w, \alpha, \pi)
\]

\[
= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} \sum_{i=1}^{d} \pi_i G^i_{\beta}(w, \alpha).
\]

(19)

Thus, minimizing the worst-case CVaR over \( w \in W \) is equivalent to the following min-min-sup optimization problem:

\[
\min_{w \in W} WCVaR_{\beta}(w) = \min_{w \in W} \min_{\alpha \in \mathbb{R}} \sup_{\pi \in \Pi} G_{\beta}(w, \alpha, \pi).
\]

(20)

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

\[
C(u_1, ..., u_n) = P(U_1 \leq u_1, ..., U_n \leq u_n),
\]

(21)

where \( U_i \sim U[0, 1] \) and \( u_i \) are realizations of \( U_i, \quad i = 1, ..., n \). The margins \( u_i \) can be replaced by \( F_i(x_i) \), where \( x_i, \quad i = 1, ..., n \) 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 \( P(F(x) \leq u) = P(x \leq F^{-1}(u)) = F(F^{-1}(u)) = u \)). 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 and $n$-increasing;
2. $C$ has marginal distributions $C_k$, $k = 1, \ldots, n$, where $C_k(u) = u$ for every $u = (u_1, \ldots, u_n)$ in $[0, 1]^n$.

Equivalently, an $n$-copula is a function

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

with the following properties:

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

(ii) ($n$-increasing) For all $a$ and $b$ in $[0, 1]^n$ such that $a_i \leq b_i$, for every $i$, $V_C([a, b]) \geq 0$, 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 $X_1, \ldots, X_n$ be random variables with distribution functions $F_1, \ldots, F_n$, respectively. Then, there exists an $n$-copula $C$ such that,

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)), \quad (22)$$

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

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

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)), \quad (23)$$

where $F_i^{-1}$, $i = 1, \ldots, n$ 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 $f_1, \ldots, f_n$ univariate density functions of the margins $F_1, \ldots, F_n$. Assuming that $F(\cdot)$ and $C(\cdot)$ are differentiable, by (22) and (23)

$$\frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \equiv f(x_1, \ldots, x_n) = \frac{\partial^n C(F_1(x_1), \ldots, F_n(x_n))}{\partial x_1 \cdots \partial x_n} \quad (24)$$

$$= c(u_1, \ldots, u_n) \prod_{i=1}^n f_i(x_i). \quad (25)$$

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

\[
c(u_1, \ldots, u_n) = \frac{f(x_1, \ldots, x_n)}{\prod_{i=1}^{n} f_i(x_i)} \tag{26}
\]

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).


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

\[
C(u_1, \ldots, u_n) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2) + \ldots + \psi^{-1}(u_n)), \tag{27}
\]

for an appropriate generator \( \psi(\cdot) \), where \( \psi: [0, \infty] \to [0, 1] \) and satisfies (i) \( \psi(0) = 1 \) and \( \psi(\infty) = 0 \); (ii) \( \psi \) is d-monotone, i.e., \((−1)^k d^k \psi(s) ds^k \geq 0 \) for \( k \in \{0, \ldots, d-2\} \) and \((-1)^{d-2} d^{d-2} \psi(s) ds^{d-2} \) is decreasing and convex. Most of the copulas in this class has 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 \( w = (w_1, \ldots, w_n)^\top \), \( u = (u_1, \ldots, u_n) \) in \([0, 1]^n \) be an stochastic vector, \( h(w, u) = h(w, F(x)) \) the loss function and \( F(x) = (F_1(x_1), \ldots, F_n(x_n))^\top \) a set of marginal distributions. Also, assume that \( u \) follows a continuous distribution with copula \( C(\cdot) \).
Given a fixed $w \in W$, a random vector $x \in \mathbb{R}^n$ and the equation (25), the probability that $h(w, x)$ does not exceed a threshold $\alpha$ is represented by

$$
P(h(w, x) \leq \alpha) = \int_{h(w, x) \leq \alpha} f(x) \, dx$$

$$= \int_{h(w, x) \leq \alpha} c(F(x)) \prod_{i=1}^n f_i(x_i) \, dx$$

$$= \int_{h(w, F^{-1}(u)) \leq \alpha} c(u) \, du$$

$$= C \left( u \mid h(w, u) \leq \alpha \right),$$

where $f_i(x_i) = \frac{\partial F_i(x_i)}{\partial x_i}$, $F^{-1}(u) = (F_{i_1}^{-1}(u_1), ..., F_{i_n}^{-1}(u_n))^+$ and $\tilde{h}(w, u) = h(w, F^{-1}(u))$. Thus, we can represent the VaR$_\beta$ by

$$VaR_\beta(w) = \arg\min\{\alpha \in \mathbb{R} : C \left( u \mid h(w, u) \leq \alpha \right) \geq \beta \} \quad (28)$$

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

$$CVaR_\beta(w) = \frac{1}{1 - \beta} \int_{h(w, u) \geq VaR_\beta(w)} \tilde{h}(w, u) c(u) \, du, \quad (29)$$

and by (9) we can write

$$H_{\beta}(w, \alpha) = \alpha + \frac{1}{1 - \beta} \int_{u \in [0,1]^n} \left( \tilde{h}(w, u) - \alpha \right)^+ c(u) \, du. \quad (30)$$

Zhu and Fukushima (2009) derived the WCVaR considering a mixture of distributions in a prescribed set as we have seen in equations (15) – (20). Kakouris and Rustem (2014) extended their framework considering a set of copulas $C(\cdot) \in C$.

Let

$$C(\cdot) \in C_M = \left\{ \sum_{i=1}^d \pi_i C_i(\cdot) : \sum_{i=1}^d \pi_i = 1, \pi_i \geq 0, i = 1, ..., d \right\}, \quad (31)$$

and similarly to (17)

$$H_{\beta}(w, \alpha, \pi) = \alpha + \frac{1}{1 - \beta} \int_{u \in [0,1]^n} \left( \tilde{h}(w, u) - \alpha \right)^+ \sum_{i=1}^d \pi_i c_i(u) \, du, \quad i = 1, ..., d \quad (32)$$

$$= \sum_{i=1}^d \pi_i H_{\beta}(w, \alpha), \quad i = 1, ..., d, \quad (33)$$

where

$$H_{\beta}(w, \alpha) = \alpha + \frac{1}{1 - \beta} \int_{u \in [0,1]^n} \left( \tilde{h}(w, u) - \alpha \right)^+ c_i(u) \, du, \quad i = 1, ..., d. \quad (34)$$

Invoking theorem 1 of Zhu and Fukushima (2009) again we can state that for each fixed $w \in W$ the
\( \text{WCVaR} \) with respect to the set \( C \) is represented by
\[
\text{WCVaR}_\beta (w) = \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} \sum_{i=1}^{d} \pi_i \text{H}_\beta (w, \alpha).
\]

Thus, the Worst Case Copula-CVaR with respect to \( C \) is the mixture copula that produces the greatest CVaR, i.e., the worst performing copula combination in the set \( C \). Moreover, minimizing the worst-case Copula-CVaR over \( w \in W \) can be defined as the following optimization problem:
\[
\min_{w \in W} \text{WCVaR}_\beta (w) = \min_{w \in W} \min_{\alpha \in \mathbb{R}} \sup_{\pi \in \Pi} \text{H}_\beta (w, \alpha),
\]

Using the approach of Rockafellar and Uryasev (2000) the integral in (34) is approximated by sampling realizations from the copulas \( C_i (\cdot) \in C \) using as inputs the filtered uniform margins. If the sampling generates a collection of values \( \{u_{ij}^1, u_{ij}^2, ..., u_{ij}^J\} \), where \( u_{ij}^j \) 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 \( \text{H}_\beta (w, \alpha) \) by
\[
\tilde{H}_\beta (w, \alpha) = \alpha + \frac{1}{(1 - \beta)} \sum_{j=1}^{S_i} \left( h(w, u_{ij}^j) - \alpha \right)^+, i = 1, ..., d
\]

Assuming that the allowable set \( W \) is convex and the loss function \( \tilde{h} (w, u) \) is linear with respect to \( w \) then optimization problem
\[
\min_{w \in W, \alpha \in \mathbb{R}} \tilde{H}_\beta (w, \alpha).
\]

reduces to the following LP problem:
\[
\min_{w \in \mathbb{R}^n, v \in \mathbb{R}^m, \alpha \in \mathbb{R}} \alpha + \frac{1}{(1 - \beta)} \sum_{i=1}^{S_i} v_i
\]

s.t. \( w \in W \),
\[
v_i^j \geq \tilde{h}(w, u_{ij}^j) - \alpha, \quad v_i^j \geq 0, \quad j = 1, ..., J; \quad i = 1, ..., d.
\]

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

2.6 Worst Optimal Mean-Variance

2.6.1 Estimators of the Covariance Matrix

Portfolio optimization strategies based on the trade-off risk-return depend on accurate estimates of a covariance matrix of asset returns. There are many approaches for constructing such a covariance matrix, some using the sample covariance matrix as a starting point.

However, in practice, the Markowitz approach is quite sensitive to potential estimation errors which cause an impact in the asset allocations weights, no longer leading to an efficient portfolio.
The unbiased estimator of the covariance matrix of returns, i.e., the sample covariance matrix (\( S \)), seems to be an appropriate candidate for replacing the population moments. However, it is known that it is very unstable when the number of data points is smaller or comparable to the number of assets. Another alternative is to consider a structured covariance estimator, such as the single-factor model of Sharpe (1963). Such estimators contain relatively little estimation error but if the structure is misspecified the estimators can be severely biased.

Ledoit and Wolf (2003, 2004a,b) propose to estimate \( \Sigma \) by a convex combination of an estimated structured matrix, denoted by \( \hat{F} \), and the sample covariance matrix estimator \( S \):

\[
\hat{\Sigma} = \hat{\delta} \hat{F} + \left( 1 - \hat{\delta} \right) S,
\]

where \( \hat{\delta} \) is an estimator of the optimal shrinkage constant \( \delta \). The idea of shrinkage estimators is to combine two "utmost" estimators in order to obtain an estimator that takes advantage of the good properties of both estimators. Ledoit and Wolf (2003) propose to shrink towards one single-factor model (to the cross-sectional average of all random variables), while Ledoit and Wolf (2004b) propose to use one-parameter matrix for \( F \) (all variances are the same, all covariances are zero) and Ledoit and Wolf (2004a) to use a constant correlation matrix (all pairwise correlations are equal). In this paper, we also consider other two shrinkage estimators: shrinkage toward two-parameter matrix (all variances are the same, all covariances are the same) and toward a diagonal matrix\(^3\). We call these shrinkage estimators LW1, LW2, LW3, LW4 and LW5.

Finally, we also consider the RiskMetrics EWMA filter (Morgan et al. (1996)), since it is popular in most industry application entailing large-scale covariance matrix measurements. The RiskMetrics procedure is based on the exponentially weighted moving average estimator, which is a special case of the GARCH(1,1) model (Bollerslev (1986)). Thus, the volatility of the next period can be computed as a weighted average of the current volatility and squared returns, i.e.,

\[
\Sigma_{t+1} = (1 - \lambda) r_t r_t^\top + \lambda \Sigma_t
\]

where \( \lambda \) is a decay factor, \( 0 < \lambda < 1 \). RiskMetrics found, on a widely diversified international portfolio, that the value \( \lambda = 0.94 \) produces the best backtesting results for a daily data set (Morgan et al. (1996), pp. 97-101).

### 2.6.2 Mean-Variance Portfolio

Assume that a portfolio consists of \( N \) financial instruments and also the previously defined notations about the distribution of a random vector \( r \). Define the portfolio risk by the portfolio variance \( \sigma_w^2 = w^\top \Sigma w \), where \( \Sigma \) denotes a positive definite variance-covariance matrix of the assets returns, and the expected portfolio return by \( w^\top \mu \).

The optimal mean-variance portfolio for an risk-averse agent who is a wealth maximizer is the solution of the following convex optimization problem subject to linear constraints:

\[
\arg \min_w w^\top \Sigma w - \frac{1}{\gamma} E (r_p,t+1)
\]  

\(^3\)For a more detailed exposition the reader is referred to Ledoit and Wolf (2003, 2004a,b).
where $w \in \mathbb{R}^N$, $w^T\Sigma w$ is the sample portfolio variance, $\gamma$ is a risk aversion coefficient, $E(r_{p,t+1})$ is the sample mean of the portfolio return. The constraint $w^T \mathbf{1} = 1$ states that the sum of the weights should be 1 and the last constraint ($w_i \geq 0, i = 1, \ldots, N$) ensures that there is no short-selling.

To assess what we call Worst Optimal Mean-Variance (WOMV), we compute the sample portfolio variance using seven alternative specifications to estimate the covariance matrix: the sample covariance matrix, RiskMetrics EWMA and the five shrinkage estimators already mentioned. The WOMV is the one that produces the greatest portfolio variance, i.e., the worst performing method.

3 Data and Methodology

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 and log-returns are calculated in excess of a risk-free asset\(^4\). 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 (12-month formation period) and out-of-sample (6 months) periods are included in the analysis, i.e., around 500 stocks in each trading period.

We want a diversified set of stocks that can be useful during crises and tranquil periods. To attain this goal we select, among all listed stocks in each formation period, a set of 10 stocks based on the ranked sum of squared spreads (the five largest and the five smallest) between the normalized daily closing prices deviations (known as distance, Gatev, Goetzmann, and Rouwenhorst (2006)) of the S&P 500 index and all shares. Distances are computed using data for 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

$$\text{Spread}_t = N P_{i,t} - N P_{SP500,t},$$  \hspace{1cm} (48)$$

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

Our optimization strategy adopt a sliding window of calibration of $T=252$ observations, differently from Kakouris and Rustem (2014), which corresponds to approximate one year of daily data. Therefore, we use day 1 to 252 to estimate the parameters of all models and determine portfolio weights for day 253 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 $L - 1$ daily log-returns.

\(^4\)We use 3-month Treasury Bill obtained at https://fred.stlouisfed.org/series/TB3MS as a proxy for the risk-free rate.

\(^5\)Missing values have been interpolated.
3.1 Strategies under Analysis

To apply Worst Case Copula-CVar Portfolio Optimization we go through the following steps: (1) First we fit a GARCH(1,1) model with t-distributed innovations to each univariate time series selected from distance method; (2) Using the estimated parametric model, we construct the standardized residuals vectors given, for each $i = 1, \ldots, 10$ and $t = 1, \ldots, L - T - 1$, by

$$\frac{\hat{\epsilon}_{i,t}}{\hat{\sigma}_{i,t}}.$$  

The standardized residuals vectors are then converted to pseudo-uniform observations $z_{i,t} = \frac{2}{n+1} F_i (\hat{\epsilon}_{i,t})$, 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

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

where $\Theta = (\alpha, \beta, \delta)^T$ are the Clayton, Frank and Gumbel copula parameters, respectively, and $\pi_1, \pi_2 \in [0,1]$. 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 $\hat{\sigma}_{i,t}$ using the estimated GARCH model; (7) determine the simulated daily asset log-returns, i.e., determine the simulated daily log-returns as $r_{i,t}^{sim} = \hat{\mu}_t + \hat{\sigma}_{i,t} z_{i,t}$; (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.

For each of the three copulas, we run 10000 return scenarios from the estimated multivariate CFG Mixture Copula model for each of the ten 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 $(46) - (47)$. The constraint $(45)$ is replaced by $\hat{\mu}(r_{t}^{sim}) \geq \bar{\tau}$, $t = 1, \ldots, L - T - 1$, in order to make sure it is attended for each calibration day, where $\hat{\mu}(r_{t}^{sim})$ is the sample mean of the simulated portfolio returns and $\bar{\tau}$ is the minimum daily expected return for the portfolio. We solve all problems for $\bar{\tau} = 0, 0.025\%$ and without this constraint. The confidence level $\beta$ is set at $\beta = 0.95$.

We compute the optimal mean-variance portfolio considering a risk aversion coefficient $\gamma = 1$, following DeMiguel and Nogales (2009) and Santos and Tessari (2012), and replace $\mathbb{E}(r_{p,t+1})$ by $\hat{\mu}(r_{p,t})$, where $\hat{\mu}(r_{p,t})$ is the sample mean of the assets returns computed using a moving average of 252 previous days, and $\Sigma$ by $\hat{\Sigma}$, where $\hat{\Sigma}$ represents the estimated covariance matrix by the seven proposed estimators.

We also consider two other benchmarks: the equally weighted portfolio (EWP), where $\mathbf{w} = \left( \frac{1}{n}, \ldots, \frac{1}{n} \right)^T$ for any period $t$ 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 consec-
utive days and between two days within a period of maximum six months, Sharpe ratio, Sortino ratio, turnover, breakeven costs and CVaR.

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

$$\hat{r}_{p,t} = w_{t-1}^\top r_{p,t} - r_{f,t},$$

and the portfolio mean excess return as

$$\hat{\mu}_p = \frac{1}{L - T} \sum_{t=T+1}^{L-1} \hat{r}_{p,t}$$

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

$$\hat{\sigma}_p = \sqrt{\frac{1}{L - T - 2} \sum_{t=T}^{L-1} (w_{t} r_{p,t+1} - \hat{\mu}_p)^2},$$

and

$$SR = \frac{\hat{\mu}_p}{\hat{\sigma}_p}.$$  

Denote now by $r_p(w,t)$ the cumulative portfolio return at time $t$, where $w$ 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 $r_p(w,t)$ over the history preceding time $t$ and the value of this function at time $t$, i.e.,

$$D(w,t) = \max_{0 \leq \tau \leq t} \{ r_p(w,\tau) \} - r_p(w,t).$$

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

$$MaxDD(w) = \max_{0 \leq \tau \leq T} \{ D(w,\tau) \}.$$

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.,

$$SoR = \frac{\hat{\mu}_p}{\hat{\sigma}_{p,n}},$$

where

$$\hat{\sigma}_{p,n} = \sqrt{\frac{1}{L - T - 1} \sum_{t=T}^{L-1} \left( \min(0, w_t^\top r_{t+1} - r_{MAR}) \right)^2},$$

where $r_{MAR}$ is the value of a minimal acceptable return (MAR), usually zero or the risk-free rate.\(^6\)

\(^6\)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.
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:

\[
\text{Turnover} = \frac{1}{L - T - 1} \sum_{t=T}^{L-1} \sum_{j=1}^{N} (|w_{j,t} - w_{j,t+1}|),
\]

where \( w_{j,t+1} \) is the actual weight in asset \( j \) before rebalancing at time \( t+1 \), and \( w_{j,t} \) is the optimal weight in asset \( j \) at time \( t \). 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, \( \hat{\mu}_{TC} \), given by

\[
\hat{\mu}_{TC} = \frac{1}{L - T} \sum_{t=T}^{L-1} \left[ \left( 1 + w_t^\top r_{t+1} \right) \left( 1 - c \sum_{j=1}^{N} (|w_{j,t+1} - w_{j,t+1}|) \right) - 1 \right],
\]

where \( c \) is called breakeven cost when we solve \( \hat{\mu}_{TC} = 0 \).

One of the possible criticisms that can be done is that the conclusions are based on only one realization of the stochastic process of asset returns based on the observed series of prices, since among thousands of different strategies is very likely we find some that show superior performance in terms of excess return or Sharpe Ratio. In order to mitigate data-snooping criticisms, we use the stationary bootstrap of Politis and Romano (1994) to compute bootstrap p-values using the methodology proposed by Ledoit and Wolf (2008) and test if the differences between the average excess return and Sharpe ratio of the Worst Case Copula-CVaR and the benchmark strategies are significant.

In order to construct the distributions we bootstrapped 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.

To test the hypotheses that the average excess returns and Sharpe Ratios of the Worst Case Copula-CVaR strategy is equal to that of benchmark methods, that is,

\[
H_0 : \mu_{wc} = \mu_b \quad \text{and} \quad H_0 : \frac{\mu_{wc}}{\sigma_{wc}} = \frac{\mu_b}{\sigma_b} = 0,
\]

we compute, following Davison and Hinkley (1997), a two-sided p-value using \( B = 10000 \) (stationary) bootstrap resamples as follows:

\[
p_{\text{psboot}} = \begin{cases} 
2 \sum_{b=1}^{B} I_{\{0 < t^{*}(b) + 1\}} \frac{1}{B+1}, & \text{if } \text{median}\left\{ t^{*}(1), \ldots, t^{*}(B) \right\} > 0, \\
2 \sum_{b=1}^{B} I_{\{0 > t^{*}(b) + 1\}} \frac{1}{B+1}, & \text{otherwise,}
\end{cases}
\]

where \( I \) is the indicator function, \( t^{*}(b) \) are the values in each block stationary bootstrap replication and \( B \) denotes the number of bootstrap replications.
Tables 1 to 3 report 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) and CVaR_{0.95} of the different portfolio strategies from 1991/2 to 2015/2 involving daily, weekly and monthly rebalancing frequencies. Returns, standard deviation, Sharpe ratio and Sortino ratio are summarized in an annualized basis.

By analyzing Tables 1 to 3, it is clear that the allocations according to the Worst Case Copula-CVaR (WCCVaR) strategy consistently outperforms the other strategies in terms of profitability for any rebalancing frequency and return constraints. The 1/N naive diversification comes second with respect to excess returns in most of the scenarios, although the difference to Worst Optimal Mean-Variance (WOMV) is not large and decreases when the portfolio holding period increases.

WCCVaR is also the best performing portfolio with respect to the conditional risk measure. Judged by CVaR, the WCCVaR approach yields the less riskier trajectory. However, WCCVaR can be identified as the riskiest portfolio with respect to volatility and maximum drawdown measures. Overall, the WCCVaR standard deviations are twice as large as those for WOMV, affecting its risk-adjusted ranking performance measured by Sharpe and Sortino ratios across-the-board.

The numerical experiments show that the pictures out-of-sample stay very similar for these levels of return constraints. However, it should be mentioned that the WOMV allocations do not satisfy the \( r \geq 0\% \) and \( r \geq 0.0025\% \) constraints in 2 and 15 days, respectively. Instead of discarding these observations, we consider the allocations without a constraint for those 2 days that \( r \geq 0\% \) is not satisfied and those for the constraint \( r \geq 0\% \) for the other 13 days. The WCCVaR allocations satisfies all constraints during the whole sample period.

Furthermore, the results indicate that risk-return performance of the portfolios is not very sensitive to the rebalancing frequencies. The main exception is the drawdown witnessed during the global financial crisis, specially for the WCCVaR allocations for monthly rebalancing. Given the market adjustment during 2008, a daily rebalancing would save as much as 27% if a daily rebalancing would have been done

\footnote{DeMiguel and Nogales (2009) found that gains from an ‘optimal’ diversification may not be large enough to offset the loss arising from estimation errors. In this paper, the covariance matrix associated to the worst performing mean-variance portfolio is estimated by RiskMetrics EWMA filter in all scenarios.}

\footnote{For simplicity, we set \( r_{MAR} = 0\).}

\footnote{All of them between October and December, 2008. We have also considered the constraint \( r \geq 0.005\% \). Given the results are still very similar they are not presented here and are available under request. In this case, the WOMV portfolio does not satisfy the constraint in 29 days.}
compared to the monthly rebalancing.

**Table 1:** Excess returns of Worst Case Copula-CVaR (WCCVaR), Worst Optimal Mean Variance (WOMV) and Equal Weights portfolios without daily minimum expected return constraint

<table>
<thead>
<tr>
<th></th>
<th>WCCVaR</th>
<th>WOMV</th>
<th>1/N</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Daily Rebalancing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Return (%)</td>
<td>24.28***</td>
<td>13.85</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>41.18</td>
<td>19.75</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.59</td>
<td>0.70</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.99</td>
<td>1.16</td>
<td>0.96</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.349</td>
<td>0.0668</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>0.25</td>
<td>0.77</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-16.40</td>
<td>-9.97</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-59.51</td>
<td>-40.44</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR0.95</td>
<td>-0.0017</td>
<td>-0.0266</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
<tr>
<td><strong>Weekly Rebalancing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Return (%)</td>
<td>22.41**</td>
<td>14.68</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>40.88</td>
<td>20.02</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.55</td>
<td>0.73</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.90</td>
<td>1.21</td>
<td>0.96</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.097</td>
<td>0.035</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>0.83</td>
<td>1.56</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-16.14</td>
<td>-9.87</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-68.89</td>
<td>-43.25</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR0.95</td>
<td>-0.0017</td>
<td>-0.0269</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
<tr>
<td><strong>Monthly Rebalancing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Return (%)</td>
<td>24.41**</td>
<td>15.74</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>42.85</td>
<td>20.82</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.57</td>
<td>0.76</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.94</td>
<td>1.25</td>
<td>0.96</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.0365</td>
<td>0.02</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>2.38</td>
<td>2.91</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-24.39</td>
<td>-13.52</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-86.55</td>
<td>-38.80</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR0.95</td>
<td>-0.0017</td>
<td>-0.0277</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
</tbody>
</table>

*Note:* Out-of-sample performance statistics between July 1991 and December 2015 (6173 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.

***, **, * significant at 1%, 5% and 10% levels, respectively.
Table 2: Excess returns of Worst Case Copula-CVaR (WCCVaR), Worst Optimal Mean Variance (WOMV) and Equal Weights portfolios for a daily minimum expected return constraint $r \geq 0\%$

Daily Rebalancing

<table>
<thead>
<tr>
<th></th>
<th>WCCVaR</th>
<th>WOMV</th>
<th>1/N</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Return (%)</td>
<td>24.33***</td>
<td>13.65</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>41.18</td>
<td>20.03</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.59</td>
<td>0.68</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.99</td>
<td>1.12</td>
<td>0.97</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.349</td>
<td>0.0839</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>0.25</td>
<td>0.61</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-16.38</td>
<td>-9.97</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-59.51</td>
<td>-40.59</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR$_{0.95}$</td>
<td>-0.0017</td>
<td>-0.0272</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
</tbody>
</table>

Weekly Rebalancing

<table>
<thead>
<tr>
<th></th>
<th>WCCVaR</th>
<th>WOMV</th>
<th>1/N</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Return (%)</td>
<td>22.41**</td>
<td>14.51</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>40.88</td>
<td>20.24</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.55</td>
<td>0.72</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.90</td>
<td>1.18</td>
<td>0.97</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.0969</td>
<td>0.0392</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>0.83</td>
<td>1.37</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-16.15</td>
<td>-9.87</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-67.85</td>
<td>-45.26</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR$_{0.95}$</td>
<td>-0.0017</td>
<td>-0.0272</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
</tbody>
</table>

Monthly Rebalancing

<table>
<thead>
<tr>
<th></th>
<th>WCCVaR</th>
<th>WOMV</th>
<th>1/N</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Return (%)</td>
<td>24.42**</td>
<td>14.80</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>42.85</td>
<td>20.91</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.57</td>
<td>0.71</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.94</td>
<td>1.16</td>
<td>0.97</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.0364</td>
<td>0.0209</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>2.38</td>
<td>2.62</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-24.39</td>
<td>-13.52</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-86.56</td>
<td>-42.73</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR$_{0.95}$</td>
<td>-0.0017</td>
<td>-0.0282</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
</tbody>
</table>

Note: Out-of-sample performance statistics between July 1991 and December 2015 (6173 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. *** , **, * significant at 1%, 5% and 10% levels, respectively.
Table 3: Excess returns of Worst Case Copula-CVaR (WCCVaR), Worst Optimal Mean Variance (WOMV) and Equal Weights portfolios for a daily minimum expected return constraint $r \geq 0.0025\%$

<table>
<thead>
<tr>
<th></th>
<th>WCCVaR</th>
<th>WOMV</th>
<th>1/N</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Daily Rebalancing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Return (%)</td>
<td>24.11**</td>
<td>13.54</td>
<td>15.61</td>
<td>6.10</td>
</tr>
<tr>
<td>Standard Deviation (%)</td>
<td>41.15</td>
<td>19.99</td>
<td>26.86</td>
<td>18.70</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.59</td>
<td>0.68</td>
<td>0.58</td>
<td>0.33</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>0.98</td>
<td>1.12</td>
<td>0.97</td>
<td>0.52</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.349</td>
<td>0.0705</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Break-even (%)</td>
<td>0.25</td>
<td>0.71</td>
<td>429.45</td>
<td></td>
</tr>
<tr>
<td>MDD1 (%)</td>
<td>-16.38</td>
<td>-10.72</td>
<td>-14.24</td>
<td>-12.42</td>
</tr>
<tr>
<td>MDD2 (%)</td>
<td>-59.42</td>
<td>-44.79</td>
<td>-38.21</td>
<td>-47.69</td>
</tr>
<tr>
<td>CVaR$_{0.95}$</td>
<td>-0.0017</td>
<td>-0.0271</td>
<td>-0.0357</td>
<td>-0.0272</td>
</tr>
</tbody>
</table>

| **Weekly Rebalancing**|        |      |     |         |
| Mean Return (%)      | 22.29**| 14.37| 15.61| 6.10    |
| Standard Deviation (%)| 40.85  | 20.30| 26.86| 18.70   |
| Sharpe Ratio         | 0.55   | 0.71 | 0.58 | 0.33    |
| Sortino Ratio        | 0.90   | 1.15 | 0.97 | 0.52    |
| Turnover             | 0.0970 | 0.0360| 0.0001|         |
| Break-even (%)       | 0.82   | 1.48 | 429.45|         |
| MDD1 (%)             | -16.12 | -10.91| -14.24| -12.42 |
| MDD2 (%)             | -68.87 | -46.93| -38.21| -47.69 |
| CVaR$_{0.95}$        | -0.0017| -0.0275| -0.0357| -0.0272|

| **Monthly Rebalancing**|        |      |     |         |
| Mean Return (%)      | 23.79**| 15.06| 15.61| 6.10    |
| Standard Deviation (%)| 42.74  | 20.94| 26.86| 18.70   |
| Sharpe Ratio         | 0.56   | 0.72 | 0.58 | 0.33    |
| Sortino Ratio        | 0.92   | 1.18 | 0.97 | 0.52    |
| Turnover             | 0.0365 | 0.0204| 0.0001|         |
| Break-even (%)       | 2.32   | 2.73 | 429.45|         |
| MDD1 (%)             | -24.39 | -13.52| -14.24| -12.42 |
| MDD2 (%)             | -86.43 | -46.58| -38.21| -47.69 |
| CVaR$_{0.95}$        | -0.0017| -0.0282| -0.0357| -0.0272|

*Note:* Out-of-sample performance statistics between July 1991 and December 2015 (6173 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.

\*, \**, \*** significant at 1\%, 5\% and 10\% levels, respectively.

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.

The WCCVaR portfolio has consistently the highest turnover in all combinations analyzed, although the difference to the WOMV portfolio turnover decreases when the rebalancing frequency decreases. In
fact, the WCCVaR turnover is more than 5 times higher than WOMV turnover for daily rebalancing and less than 2 times higher when holding the portfolio for a month. As expected, the portfolio turnover decreases for both allocations when the holding portfolio period increases and the EWP (1/N) has the smallest turnover.

Finally, we report the break-even transaction cost in order to investigate if the profits are economically significant. The break-even values in Tables 1 to 3 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.

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 and WOMV allocations the break-even costs are between (0.25%-2.38%) and (0.61%-2.91%), respectively. Thus, for month rebalancing this means that if the trading costs are anything less than 230 basis points the excess profits for these allocations are still greater than zero. We can also note that the break-even costs for WCCVaR are higher than for WOMV, but the difference is relatively smaller when the holding period increases. Given the robustness of our results, it is clear that rebalancing the portfolio less frequently may be beneficial for investors who desire to maximize their wealth.

To test the statistical significance of the excess returns and Sharpe ratio performances among WCCVaR and benchmark portfolios we use a resampling procedure called stationary bootstrap (Politis and Romano (1994)) using the automatic block-length selection of Politis and White (2004) and 10000 bootstrap resamples. To compute the bootstrap p-values we use the methodology proposed by Ledoit and Wolf (2008). Since each portfolio has a different optimal block length we use the average block length for simplicity. The asterisks in Tables 1 to 3 indicate a statistically significant difference. We found that WCCVaR allocations are significantly more profitable than buy and hold the S&P 500 index.\(^\text{10}\)

Figure 1 depicts the wealth trajectories of the portfolios strategies for all rebalancing frequencies, without enforcing constraint (45) and assuming an initial wealth of $1 monetary unit. Panels (a) to (c) shows the excess returns for daily, weekly and monthly rebalancing, respectively.

\(^{10}\)We found bootstrap p-values between 0.008 and 0.0204.
Figure 1: Cumulative excess returns of the portfolio strategies without a return constraint

This figure shows how an investment of $1 evolves from July 1991 to December 2015 for each of the portfolios.

It is clear that the copula-based approach is the more profitable method, in particular during tranquil periods. However, the hump-shaped pattern during the dot-com bubble crisis (2000-2001) and the sub-prime mortgage crisis (2007-2008) are noticeable and much more pronounced for the WCCVaR than for the other portfolio strategies, reinforcing the maximum drawdown results. Therefore, our results indicate that the WCCVaR method may be the best option for an investor who tries to maximize his/her end-of-period wealth, specially during upward trends, but it is the riskiest among the approaches analyzed. The WOMV and the 1/N portfolio strategies have a similar wealth accumulation performance over time.

Figures 2 and 3 show similar patterns after enforcing the return constraints.
Figure 2: Cumulative excess returns of the portfolio strategies with daily target mean return $r \geq 0\%$

This figure shows how an investment of $1$ evolves from July 1991 to December 2015 for each of the portfolios.
Figure 3: Cumulative excess returns of the portfolio strategies with daily target mean return \( r \geq 0.0025\% \).

This figure shows how an investment of $1 evolves from July 1991 to December 2015 for each of the portfolios.

5 Concluding Remarks

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 the distance method of 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 shares from 1990 to 2015 we evaluate the performance of the WCCVaR (Worst Case Copula CVaR) portfolio, considering different rebalancing strategies and return constraints, and compare it against two competitive models: a proposed Worst Optimal Mean-Variance (WOMV) portfolio and a equally weighted portfolio (EWP).

Our empirical analysis shows that the Copula-CVaR approach is more profitable than the other portfolio solutions and yields the best long-term performance with respect to CVaR. However, it is a riskier choice for short-term investments, in particular during unstable periods. In addition, we compute the portfolio turnover and report the break-even transaction cost of each strategy in order to investigate if the profits are economically significant. Although WCCVaR solutions present the highest turnover and
thus the smallest break-even costs in all cases, we found that its out-of-sample performance is relatively better when rebalancing the portfolio less frequently.

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 non-linear association measurement 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. Furthermore, a regular or a D-vine copula could add flexibility when modeling the dependence structure of the portfolio assets (even in higher dimensions), since they can describe a wider array of dependence patterns.

Additional suggestions include relaxing the assumption of no short selling, incorporate transaction cost as an additional contraint in the optimization problem as in Krokhmal, Palmquist, and Uryasev (2002) and investigate higher return constraints.

References


