Constructing Optimal Portfolios under Risk Budgeting

Vali Asimit
Bayes Business School, City, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK, asimit@city.ac.uk

Liang Peng
Georgia State University, 33 Gilmer Street SE Atlanta, GA 30303, USA, lpeng@gsu.edu

Radu Tunaru
University of Sussex Business School, Brighton, BN1 9SN, UK, r.tunaru@sussex.ac.uk

Feng Zhou
University of Cambridge, Cambridge, CB2 0SR, UK, feng.zhou@mrc-bsu.cam.ac.uk

We provide a mathematical characterization for risk parity/budgeting portfolio construction problems for general risk preferences. For the general problem when distribution of returns is not known, we demonstrate the existence of a solution to the risk budgeting problem for any convex and homogeneous risk preferences. Statistical inferences are determined for those portfolios when risk preferences are ordered by variance or Conditional Value-at-Risk. A novel Conditional Value-at-Risk estimator is proposed, which is shown to perform very well on non i.i.d observations, based on simulated and real-life data, especially during periods of bull market and irrational exuberance. Our numerical results show superior performance of risk parity portfolios in terms of various measure of performance such as Sharpe ratio and diversification when comparing with other benchmark portfolios including the equally weighted portfolio. We also find that the risk parity portfolios with an opportunity set selected via socially responsible investment attributes have good performance.

Key words: Risk budgeting/parity; Portfolio selection; Non-parametric estimation; Socially responsible investment.

December 5, 2023

1. Introduction

One of the most important activities in financial markets is robust portfolio construction. Mean-variance portfolio optimization and its numerous variants have dominated the finance and operations research literature over the last seven decades; e.g., see Cornuéjols et al. (2018) and references therein. The rich operations research literature has been proposing a wide variety of theoretically and computationally sound methods over time. One strand of research is to create robust decisions by including distributional ambiguity; the underlying distribution is assumed to depart from a reference probability measure through a probability metric constraint (Pflug and Wozabal 2007, Blanchet et al. 2022), though other distributional ambiguity settings exist in the literature (e.g., see Goh and Sim (2010), Wiesemann et al. (2014), Bertsimas et al. (2018)). A second strand of research is to consider various uncertainty sets for asset returns and/or their dependence structure; e.g., see (Goldfarb and Iyengar 2003, Tütüncü and Koenig 2004, Popescu 2007, Bertsimas et al. 2011). Other approaches have been considered, e.g., (Glasserman and Xu 2014) where the effect of model error is accounted for and (DeMiguel and Nogales 2009) where robust estimators are employed, but all models discussed so far have one common goal, namely to generate innovative solutions to robust investment decisions. Often, this is achieved through robust optimisation (Ben-Tal et al. 2009) and distributionally robust optimisation (Delage and Ye 2010) formulations.

Parsimonious investment strategies that are more agnostic to how uncertainty is modelled and measured have been considered, and one example is the *equally weighted (EW)* portfolio that shows very good out-of-sample performance; for details, see the seminal paper of DeMiguel et al. (2009b).

Shrinking estimators are non-standard efficient estimators, and the econometrics field measures their efficiency by their out-of-sample performance. This is often achieved in portfolio construction by using a linear or non-linear combination of some estimators, or by imposing weights constraints. In other words, shrinking methods reconcile the trade-off between robust and parsimonious traits of an investment strategy, and these methods have been very popular in the portfolio construction literature; e.g., covariance matrix (Jagannathan and Ma 2003, Ledoit and Wolf 2017) and weights (DeMiguel et al. 2009b, Tu and Zhou 2011, Lassance et al. 2022) shrinking methods have been proposed. Such shrinking methods showed very good performance when compared to various benchmark portfolios.

An alternative robust investment strategy is the so-called *Risk Parity (RP)*, which is also known as *Equal Risk Contribution (ERC)* portfolio (Roncalli 2013, Ang 2014). Such portfolios achieve diversification through imposing equal individual risk contributions. The first RP formulation can be traced back to Qian (2005), and the main idea has evolved from a paradigm applied by a well-known hedge fund (Bridgewater) in the 1990s. The initial RP implementation makes simplified assumptions for which the weights are inversely proportional to the asset-class risk position, and risk preferences are ordered by the standard deviation. This portfolio is not the standard RP portfolio with equal risk contributions, and in fact only approximates RP portfolios (Qian 2005), which was the practical way to perform RP-like evaluations before bespoke RP algorithms became available. The first contribution in that respect appears in Maillard et al. (2010), which was followed by many other important contributions (Roncalli 2013, Spinu 2013, Bai et al. 2016, Mausser and Romanko 2018, Bellini et al. 2021).

RB/RP portfolios reflect the preference of asset managers to diversify the portfolios with respect to the individual risk contributions and not with respect to the asset's weights. Such portfolios have shown good performance when risk preferences are ordered by various risk measures (Maillard et al. 2010, Bai et al. 2016, Mausser and Romanko 2018, Bellini et al. 2021), though the variance-based RB/RP portfolios have been the most prominent choice. The standard RB/RP formulation disregard the mean asset returns, which has been argued to be a good choice in the wider context beyond the RB/RP literature; e.g., see (Jagannathan and Ma 2003, DeMiguel et al. 2009a). One could add the mean asset returns in RB/RP portfolio construction though a different numerical implementation is required (Haugh et al. 2017).

The RP literature has been widely focusing on equal risk contributions across the assets' returns, but the same methodology could be adapted so that equal risk contributions across uncorrelated factors is imposed (Ang 2014). The variance-based RP with respect to the underlying factors extracted via *Principal Component Analysis (PCA)* is provided in Meucci (2009); Roncalli and Weisang (2016) investigate the same problem by being agnostic on how the factors are obtained. PCA is a widely used data reduction technique, but the PCA decomposition is designed to extract uncorrelated underlying factors, which is unsatisfactory as independence and lack of correlation are two different concepts. A remedy of such drawback is proposed in Lassance et al. (2022) where parity is achieved over a set of risk factors obtained via Independent Component Analysis; while the previous papers focus on variance factor RP, Lassance et al. (2022) consider other risk preferences based on higher moments.

This paper makes three contributions. Our first contribution is a theoretical framework of RB/RP portfolios for which a mathematical characterization of such portfolios is provided under general risk preferences. We also show that all RP portfolios are less risky than EW portfolios, but many other interesting properties of RB portfolios are shown in Section 3.1; e.g., we found that elliptically distributed asset returns (for which multivariate Gaussian and t-distribution are special cases) lead to RP and RB portfolios that are invariant with respect to a large class of risk measures. Our theoretical results confirm and generalize previous results that have been found in the literature for specific risk preferences; e.g., *Conditional Value-at-Risk (*CVaR*)*(Mausser and Romanko 2018), expectiles (Bellini et al. 2021), *standard deviation (SD)* or *variance (var)* (Maillard et al. 2010, Roncalli 2013). Our findings are quite general and apply to many risk measures, but we do not discuss the risk measure choice, which is a multifaceted problem investigated by He et al. (2022).

Secondly, statistical inferences are emphasized as vital to the portfolio construction process, which according to our knowledge, have not been previously attempted in the RB/RP literature. We discuss the asymptotic distribution of RB/RP estimators corresponding to two popular risk preferences used in practice, variance and CVaR. Asymptotic results are obtained for general dependent data assumptions. In addition, a novel CVaR estimator is proposed, which is shown to perform very well on simulated and real-life data, especially during periods of irrational exuberance.

Thirdly, we provide further empirical evidence in support of RP strategies, which complements the evidence in the literature. Our research shows how to build portfolios with RP targets and *Socially Responsible Investment (SRI)* preferences, and we provide empirical evidence that low and high ranked SRI portfolios outperform the EW portfolio. Besides the rich portfolio selection literature driven by risk-based arguments, there is an increasing demand to integrate SRI factors into investment decisions due to the international societal demand that is pertinent to the *Environmental, Social and Governance (ESG)* agenda (Hallerbach et al. 2004, Ballestero et al. 2015). Our findings are in line with previous studies that ESG driven portfolios may produce higher returns and enhanced risk positions than portfolios constructed without considering ESG attributes (Verheyden et al. 2016). Note that the ESG agenda is a component of the wider SRI agenda that is now a top priority for policymakers. We believe that portfolio construction under SRI driven constraints will open up a new strand of literature that will interest regulators and government sponsored sovereign funds, financial institutions and also the academic world.

The paper is organized as follows. Section 2 introduces all definitions and notations, and essential background. Section 3 contains the main theoretical results of the RB/RP portfolios and their statistical inferences. Section 4 contains extensive empirical evidence of why a fund manager should consider investing in RP portfolios. The main conclusions and recommendations are provided in Section 5. All proofs and ancillary results are included in the electronic companion.

2. Problem Formulation

We first introduce some generic notations used throughout this paper. Let Δ_d be the unit d-simplex $\Delta_d := \{\mathbf{x} \in \Re^d_+ : \mathbf{1}^T\mathbf{x} = 1\}$ for any positive integer d, where $\Re^d_+ := \{\mathbf{x} \in \Re^d : \mathbf{x} \geq 0\}$ is the standard polyhedral cone of the positive quadrant of $\Re^d$. We also use the notation $\Re^d_{++} := \{\mathbf{x} \in \Re^d : \mathbf{x} > 0\}$.

The financial field is represented by $(\Omega, \mathcal{F}, \mathbb{P})$, an atomless probability space, endowed with $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$, the set of all real-valued random variables on this probability space. Let L^q, $q \in [0, \infty)$, be the set of random variables with finite q^{th} moment, and L^∞ be the set of bounded random variables. A risk measure φ is a function that maps an element of L^0 to the real set, i.e. $\varphi : L^0 \to \Re \cup \{\pm\infty\}$. We recall below some properties for a generic risk measure and generic random variable Y that represents the future loss of a financial asset. These properties are well-known in the literature, see Föllmer and Schied (2011).

Convexity: $\varphi(aY_1 + (1-a)Y_2) \leq a\varphi(Y_1) + (1-a)\varphi(Y_2)$ for any $Y_1, Y_2 \in L^0$ and $a \in [0, 1]$;

Homogeneous of order $\tau > 0$: $\varphi(cY) = c^\tau \varphi(Y)$ for any $Y \in L^0$ and $c \geq 0$;

Shift invariance: $\varphi(Y+c)=\varphi(Y)$ for any $Y\in L^0$ and $c\in\Re$;

Translation invariance: $\varphi(Y+c)=\varphi(Y)+c$ for any $Y\in L^0$ and $c\in\Re$.

Four risk measures are often recalled in this paper: SD, var, *Value-at-Risk (*VaR*)* and CVaR. For any $p\in(0,1)$, VaR at probability level p is $\mathrm{VaR}_p(Y):=\inf_x\left\{\mathbb{P}(Y\le x)\ge p\right\}$, while CVaR at probability level p is $\mathrm{CVaR}_p(Y):=\min_\theta\left\{\theta+\frac{1}{1-p}\mathbb{E}(Y-\theta)_+\right\}$ with $(\cdot)_+:=\max(\cdot,0)$ on $\Re$. There are other risk measures interrelated to those four choices; e.g., *Median Shortfall (MS)* (median of the tail distribution, i.e., a VaR-type risk measure) which is shown in He et al. (2022) to have superior robustness properties to CVaR, but we are agnostic of the risk measure choice, since this is a discussion that goes beyond the scope of this paper.

The investor is assumed to invest in a given opportunity portfolio set, $\mathbf{X}=(X_1,\ldots,X_d)$ containing d assets, and the investment strategy is uniquely determined by a vector of proportions $\boldsymbol{\alpha}\in\Delta_d$; that is, the portfolio loss is $\boldsymbol{\alpha}^T\mathbf{X}$. Note that short sales are not allowed in our portfolio selection models, which is recommended in the financial literature (Jagannathan and Ma 2003, DeMiguel et al. 2009b). Furthermore, we assume that the risk preferences of an investor are represented by the risk measure φ and therefore, the investor's perception of risk is given by $\mathcal{R}(\boldsymbol{\alpha}):=\varphi\big(\boldsymbol{\alpha}^T\mathbf{X}\big)$. In our paper, we rely on the mathematical properties of various risk measures assumed to be homogeneous of order $\tau\in\mathbb{R}$. Hence, by Euler's Homogeneous Function Theorem, we have that

$$\mathcal{R}(\boldsymbol{\alpha})=\frac{1}{\tau}\sum_{k=1}^{d}\alpha_k\frac{\partial\mathcal{R}(\boldsymbol{\alpha})}{\partial\alpha_k}=\sum_{k=1}^{d}\mathcal{RC}_k(\boldsymbol{\alpha}),\quad\text{where}\quad\mathcal{RC}_k(\boldsymbol{\alpha}):=\frac{\alpha_k}{\tau}\frac{\partial\varphi(\boldsymbol{\alpha}^T\mathbf{X})}{\partial\alpha_k}.\tag{1}$$

By definition, $\mathcal{RC}_k(\boldsymbol{\alpha})$ is the *risk contribution* of the k^{th} individual risk.

Definition 1. Let $\mathbf{b}:=(b_1,\ldots,b_d)^T$ be a given constant vector such that $\mathbf{b}\in\Delta_d\cap\Re^d_{++}$. An investing strategy $\boldsymbol{\alpha}\in\Delta_d\cap\Re^d_{++}$ is a solution to the RB problem based on the risk measure φ if

$$\mathcal{RC}_k(\boldsymbol{\alpha})=b_k\varphi\left(\boldsymbol{\alpha}^T\mathbf{X}\right),\quad\text{for all}\quad k\in\{1,2,\ldots,d\}\text{, where }\mathcal{RC}_k(\boldsymbol{\alpha})\text{ is given in (1).}\tag{2}$$

For any $\mathbf{b}\in\Delta_d$, define $\mathcal{RB}(\mathbf{b}):=\{\boldsymbol{\alpha}\in\Delta_d:\boldsymbol{\alpha}\text{ is RB}\}$ as the set of RB portfolios for a given budgeting allocation vector $\mathbf{b}$ and a general risk measure φ. The *risk contribution target* b_k in

(2) represents the pre-specified risk contribution proportion of the k^{th} risk to the overall portfolio risk. In particular, if $b_k = \frac{1}{d}$, for all $k \in \{1, 2, \ldots, d\}$, the allocation strategy is called RP. To specify the specific risk preference, e.g., $\varphi = SD$, to an RB (or RP) portfolio, we say that the portfolio is $RB - SD$ (or $RP - SD$).

Table 1 summarizes the closed-form risk contributions for the four previously-mentioned risk measures. Note that RB-SD and RB-var strategies are always equivalent. Further, we implicitly assume that the VaR risk allocations are well-defined, and a sufficient condition is for $\mathbf{X}$ to admit a joint probability density function.

φ	$\mathcal{RC}_k$
Standard deviation	$\frac{\mathrm{Cov}\left(\alpha_k X_k, \boldsymbol{\alpha}^T \mathbf{X}\right)}{\sqrt{\mathrm{Var}\left(\boldsymbol{\alpha}^T \mathbf{X}\right)}}$
Variance	$\mathrm{Cov}\left(\alpha_k X_k, \boldsymbol{\alpha}^T \mathbf{X}\right)$
Value-at-Risk at level $p \in (0,1)$	$\mathbb{E}\left[\alpha_k X_k \middle| \boldsymbol{\alpha}^T \mathbf{X} = \mathrm{VaR}_p\left(\boldsymbol{\alpha}^T \mathbf{X}\right)\right]$
Conditional Value-at-Risk at level $p \in (0,1)$	$\mathbb{E}\left[\alpha_k X_k \middle| \boldsymbol{\alpha}^T \mathbf{X} \geq \mathrm{VaR}_p\left(\boldsymbol{\alpha}^T \mathbf{X}\right)\right]$

Table 1 Individual risk contributions for some well-known risk measures.

Early versions of RB portfolios were reduced to approximations of RP-SD portfolios known in the literature as the inverse volatility weighted portfolio; for further details, see Qian (2005) and our Section 4.1. Spinu (2013) showed that (2) could be written as an efficient convex optimization problem, which is a much simpler numerical problem than solving the system of non-linear equations in (2), whenever the aggregate risk position is measured by SD. Similar formulations are provided by Roncalli (2013) for general homogeneous risk functionals. Finding RP portfolios under CVaR risk preferences is discussed in Mausser and Romanko (2018), while Bellini et al. (2021) investigate the RP portfolios for expectiles. Both articles provide excellent computationally efficient algorithms that make their proposed investment strategies implementable even for a large number of assets.

3. Main Results

The main theoretical results are included in this section and all proofs are relegated to Appendix EC.1. The mathematical characterization of the long-only RB portfolio solutions (defined in (2)) is provided in Section 3.1. This section also explains how to compute RB portfolios and describes the ancillary validatory statistical inferences. This is achieved in Section 3.2 for two risk measures, SD and CVaR, where time dependent data are allowed in our model.

3.1. RB/RP Characterization

We assume a specific parametric distribution of portfolio loss (not returns) $\mathbf{X}$, namely the elliptical family due to its tractability of aggregating the risks (McNeil et al. 2015). The elliptical class includes multivariate Gaussian and multivariate t- families of distributions.

We work with a multivariate random vector $\mathbf{X}$ that is elliptically distributed. This is signified by $\mathbf{X} \sim E_d\big(\boldsymbol{\mu}, \Sigma, \psi\big)$, where $\boldsymbol{\mu}$ is the location vector, Σ is the covariance matrix, and ψ is its generator. This means that $\mathbf{X}$ and $\boldsymbol{\mu} + A\mathbf{Z}$ have the same joint distribution, where $A \in \Re^{d\times k}$ such that $\Sigma = AA^T$, and $\mathbf{Z}$ is an k-dimensional spherical random vector with generator ψ, i.e. $\mathrm{E}\big(\exp\{i\mathbf{t}^T\mathbf{Z}\}\big) = \psi(\mathbf{t}^T\mathbf{t})$ for all $\mathbf{t} \in \Re^k$ (McNeil et al. 2015). Without loss of generality, we assume that all variances are finite, and in turn, the elliptical distribution is precisely determined by the triplet $\big(\boldsymbol{\mu}, \Sigma, \psi\big)$.

Proposition 1 provides an extension of Theorem 8.28 of McNeil et al. (2015) that determines closed-form risk measurements for elliptically distributed risks.

Proposition 1. *Let $\boldsymbol{X} \sim E_d\big(\boldsymbol{\mu}, \Sigma, \psi\big)$. If φ is a homogeneous risk measure of order $\tau > 0$ that is shift invariant, then*

$$\varphi(\boldsymbol{\alpha}^T\boldsymbol{X} + c) = \big(\boldsymbol{\alpha}^T\Sigma\boldsymbol{\alpha}\big)^{\tau/2}\varphi(Z_1) \quad \text{for any} \quad c \in \Re, \tag{3}$$

and if φ is a homogeneous risk measure of order $\tau > 0$ that is translation invariant, then $\tau = 1$ and

$$\varphi(\boldsymbol{\alpha}^T\boldsymbol{X} + c) = c + \boldsymbol{\alpha}^T\boldsymbol{\mu} + \big(\boldsymbol{\alpha}^T\Sigma\boldsymbol{\alpha}\big)^{1/2}\varphi(Z_1) \quad \text{for any} \quad c \in \Re, \tag{4}$$

where Z_1 is a spherical random variable with generator ψ.

If $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ with φ being a shift invariant and homogeneous risk measure of order $\tau > 0$ such that $\varphi(Z_1) \neq 0$, then (3) implies that finding RB portfolio strategies relative to φ is equivalent to finding $\boldsymbol{\alpha} \in \Delta_d \cap \Re^d_{++}$ such that

$$\alpha_k \sum_{i=1}^{d} \alpha_i \Sigma_{ik} = b_k \boldsymbol{\alpha}^T \Sigma \boldsymbol{\alpha} \quad \text{for all} \quad k \in \{1, 2, \ldots, d\}, \tag{5}$$

for any given $\mathbf{b} \in \Delta_d \cap \Re^d_{++}$, where Σ_{ik} represents the $(i,k)^{th}$ entry of Σ. Equation (5) tells us that all shift invariant and homogeneous risk measures of order $\tau > 0$ lead to the same set of RB portfolio strategies for a fixed $\mathbf{b} \in \Delta_d \cap \Re^d_{++}$. If φ is a translation invariant and homogeneous risk measure, then the latter conclusion holds under the condition that the aggregated risk position does not change. These are summarized in Corollary 1 below.

Corollary 1. *Let $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ and a given $\boldsymbol{b} \in \Delta_d \cap \Re^d_{++}$. Further, let φ and $\tilde{\varphi}$ be two homogeneous risk measures of order $\tau > 0$ and $\tilde{\tau} > 0$, respectively such that $\varphi(Z_1) \neq 0$ and $\tilde{\varphi}(Z_1) \neq 0$.*

a) Assume that φ and $\tilde{\varphi}$ are shift invariant risk measures. If $\boldsymbol{\alpha}^* \in \mathcal{RB}(\boldsymbol{b})$ based on φ, then $\boldsymbol{\alpha}^* \in \mathcal{RB}(\boldsymbol{b})$ based on $\tilde{\varphi}$.

b) Assume that φ and $\tilde{\varphi}$ are translation invariant risk measures such that $\tau = \tilde{\tau} = 1$ and $\varphi(Z_1) = \tilde{\varphi}(Z_1)$. If $\boldsymbol{\alpha}^* \in \mathcal{RB}(\boldsymbol{b})$ based on φ, then $\boldsymbol{\alpha}^* \in \mathcal{RB}(\boldsymbol{b})$ based on $\tilde{\varphi}$.

Corollary 1 can be helpful to provide a computational shortcut and it suggests focusing on one risk measure of choice whenever the returns are jointly elliptically distributed. For example, RB portfolios based on either variance, standard deviation, *skewness (skew)*, *kurtosis (kurt)*, excess risk (measured by either MS/VaR or CVaR) over the expected return, or any other risk measure that is a function of centred moments would lead to the same RB portfolios, i.e., if

$$\varphi \in \left\{var, SD, skew, kurt, \mathrm{VaR}_p - \mathbb{E}, \mathrm{CVaR}_q - \mathbb{E}\right\} \quad \text{for any } 0 < p, q < 1. \tag{6}$$

This invariance result shows that RB/RP portfolios balance the risk across individual (or group of) assets in the same way across all shift invariant and homogeneous risk preferences. The same result follows when the risk preferences are modelled by translation invariant and homogeneous risk measures of order $\tau = 1$ provided that the aggregate risk for these RB portfolios are equal. A similar result was shown in Asimit et al. (2019) in the context of capital allocation, where VaR and CVaR based capital allocations are found to be equivalent if the same total amount of capital ought to be allocated. Further discussions about Corollary 1 are included in Appendix EC.2.

We are now ready to provide two methods of finding and characterizing RB portfolios for a large class of risk measures without making any assumption on the underlying asset returns distribution. Two methods are investigated, which are known in the literature (e.g., see Roncalli (2013) and Bellini et al. (2021)) as the *logarithmic barrier* formulation (see (8) below) and *logarithmic constraint* RB formulation (see (9) below).

THEOREM 1. *Let* $\boldsymbol{b} \in \Delta_d \cap \Re^d_{++}$*, and* φ *be a convex, homogeneous risk measure of order* $\tau \geq 1$ *. Further, assume that*

$$\min_{\mathbf{0} < \boldsymbol{x} \leq \frac{1}{d}\mathbf{1}} \mathcal{R}(\boldsymbol{x}) > 0. \tag{7}$$

a) For any given $\lambda > 0$*, the following instance*

$$\min_{\boldsymbol{x} \in \Re^d_{++}} \frac{1}{\tau}\mathcal{R}(\boldsymbol{x}) - \lambda \sum_{k=1}^{d} b_k \log x_k \tag{8}$$

admits a unique solution, denoted as $\boldsymbol{x}^*(\lambda, \boldsymbol{b})$*, that is an interior point of* $\Re^d_{++}$*. If* $\mathcal{R}(\boldsymbol{\alpha})$ *is differentiable at* $\boldsymbol{x}^*(1, \boldsymbol{b})$*, then* $\boldsymbol{\alpha}^*(\boldsymbol{b}) \in \mathcal{RB}(\boldsymbol{b})$*, where* $\boldsymbol{\alpha}^*(\boldsymbol{b}) = \boldsymbol{x}^*(\lambda, \boldsymbol{b}) / \mathbf{1}^T \boldsymbol{x}^*(\lambda, \boldsymbol{b})$*. Moreover,*

$$\boldsymbol{\alpha}^*(\boldsymbol{b}) = \boldsymbol{x}^*(\lambda^*, \boldsymbol{b}) = (\lambda^*)^{1/\tau} \boldsymbol{x}^*(1, \boldsymbol{b}), \text{ where } \lambda^* = \left(\mathbf{1}^T \boldsymbol{x}^*(1, \boldsymbol{b})\right)^{-\tau}.$$

b) For any given $c \in \Re$*, the following instance*

$$\min_{\boldsymbol{x} \in \Re^d_{++}} \mathcal{R}(\boldsymbol{x}) \quad \text{such that} \sum_{k=1}^{d} b_k \log x_k \geq c \quad \text{with} \quad c \in \Re \tag{9}$$

admits a unique solution, denoted as $\boldsymbol{x}^{**}(c, \boldsymbol{b})$*, that is an interior point of the feasibility set. If* $\mathcal{R}(\boldsymbol{\alpha})$ *is differentiable at* $\boldsymbol{x}^{**}(1, \boldsymbol{b})$*, then* $\boldsymbol{\alpha}^{**}(\boldsymbol{b}) \in \mathcal{RB}(\boldsymbol{b})$*, where* $\boldsymbol{\alpha}^{**}(\boldsymbol{b}) = \boldsymbol{x}^{**}(c, \boldsymbol{b}) / \mathbf{1}^T \boldsymbol{x}^{**}(c, \boldsymbol{b})$*. Moreover,*

$$\boldsymbol{\alpha}^{**}(\boldsymbol{b}) = \boldsymbol{x}^{**}(c^*, \boldsymbol{b}) = e^{c^*-1} \boldsymbol{x}^{**}(1, \boldsymbol{b}), \text{ where } c^* = 1 - \log\left(\mathbf{1}^T \boldsymbol{x}^{**}(1, \boldsymbol{b})\right).$$

Furthermore, strong duality holds in (9).

c) For any $\boldsymbol{b}$, we have that $\boldsymbol{\alpha}^*(\boldsymbol{b}) = \boldsymbol{\alpha}^{**}(\boldsymbol{b})$,

$$\min_{\boldsymbol{x}\in\Delta_d\cap\Re^d_{++}} \mathcal{R}(\boldsymbol{x}) \leq \mathcal{R}\big(\boldsymbol{\alpha}^*(\boldsymbol{b})\big) \leq \mathcal{R}(\boldsymbol{b}) \tag{10}$$

and $\mathcal{R}\big(\boldsymbol{\alpha}^*(\boldsymbol{b})\big) \leq \mathcal{R}$.

d) For any $\boldsymbol{b}$, we have that

$$\prod_{k=1}^{d}\left(\frac{\alpha^*_k(\boldsymbol{b})}{\tilde{\alpha}_k}\right)^{b_k} \geq 1, \ \text{where } \tilde{\boldsymbol{\alpha}} = \underset{\boldsymbol{x}\in\Delta_d\cap\Re^d_{++}}{\arg\min}\ \mathcal{R}(\boldsymbol{x}). \tag{11}$$

Further, for any $\boldsymbol{b}$, there exists $\epsilon > 0$ such that

$$\mathcal{R}\big(\boldsymbol{\alpha}^*(\boldsymbol{b})\big) < \mathcal{R}(\boldsymbol{\alpha}), \ \text{for any } \boldsymbol{\alpha}\in\Delta_d\cap\Re^d_{++} \text{ satisfying } \min_{1\leq k\leq d}\alpha_k \leq \epsilon. \tag{12}$$

Note 2. Theorem 1 summarizes a series of very interesting results and we outline them in a non-technical language below:

i) Theorem 1 tells us through (8) and (9) that RB portfolios could be found under some mild conditions (mainly if (7) holds). This means that we show the existence of RB portfolios for any homogeneous risk preferences, but the uniqueness remains an open problem.

ii) The technical condition in (7) is sufficient (but not necessary) to ensure that our RB portfolios are found without major computational issues, since (7) guarantees finite optimal solutions in the surrogate instances (8) and (9). If (7) does not hold, one could bluntly apply the results of Theorem 1, if there is sufficient empirical evidence that (8) and (9) have finite optimal solutions. If the latter is not evident, then RB portfolios could be still found, and a method of approximating RB portfolios in this extreme setting is possible. This requires a longer discussion, and thus, the entire discussion is put in Appendix EC.3.

iii) Condition (7) is always satisfied for *var* and *SD* risk preferences if the covariance matrix is positive definite. Note that the covariance matrix empirical estimator is guaranteed to be positive semidefinite if there are no missing values and the assets are observed over the same horizon.

iv) The *logarithmic barrier* and *logarithmic constraint* RB formulations in (8) and (9), respectively, lead to the same RB portfolio that does not depend upon the normalizing parameters λ and c. This means that $\lambda = 1$ and $c = 0$ are recommended for numerical implementations.

v) We found in (10) that $\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) \leq \mathcal{R}(\mathbf{b})$ for any risk contribution target vector $\mathbf{b}$. This means that our RB portfolios with target vector $\mathbf{b}$ are less risky (with a lower aggregate risk position measured through φ) than the portfolio with weights given by $\mathbf{b}$. Particularly, the RP portfolio is always less risky than the EW portfolio.

vi) We also found that $\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) \leq \mathcal{R}\left(\frac{1}{d}\mathbf{1}\right)$ for any risk contribution target vector $\mathbf{b}$. This means that our RB portfolio (obtained by either (8) or (9)) is always less risky than the EW portfolio, irrespective of the risk preferences. This confirms a similar property found by Roncalli (2013) and Bellini et al. (2021) for some particular risk measures choices.

vii) We found in (12) that a substantial reduction in at least one asset of the portfolio would increase the overall portfolio's risk position.

Theorem 1 extends previous results for var/SD portfolios (Theorem 1.1 of Spinu (2013) and Lemma 2.2 of Bai et al. (2016)) and expectiles-based portfolios (Theorem 4 of Bellini et al. (2021)).

3.2. Statistical Inferences

The previous section explains how to find RB portfolios with the help of Theorem 1. Even though variants of Theorem 1 exist in the literature, there are no statistical inferences for RB portfolios according to our knowledge, which is the main aim of this section. In practice, the sample covariance matrix is often employed, but for high-dimensional data, it leads to inference problems. Over time, more advanced methods to estimate the covariance matrix have been introduced (Fan et al. 2008, Ledoit and Wolf 2008, Bailey et al. 2019). Our statistical inferences are focused on two risk preferences, CVaR and var, which are popular choices in practice.

In this section, we observe $\big\{\boldsymbol{X}_t = (X_{t,1},\dots,X_{t,d})^T\big\}_{t=1}^{n}$ from the strictly stationary α-mixing sequence of $\big\{\boldsymbol{X}_t = (X_{t,1},\dots,X_{t,d})^T\big\}_{t=-\infty}^{\infty}$ satisfying

$$\alpha_{\boldsymbol{X}}(k) = \sup\big\{|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{-\infty}^{i}, B \in \mathcal{F}_{i+k}^{\infty}, -\infty < i < \infty\big\} \to 0 \quad \text{as } k \to \infty,$$

where $\mathcal{F}_a^b$ denotes the σ-field generated by $\{\boldsymbol{X}_t : a \leq t \leq b\}$. For statistical inferences, Theorem 1 suggests searching for a non-parametric estimator for $\mathcal{R}(\boldsymbol{x})$, which is convex, homogeneous, and differentiable.

First, we consider CVaR_p risk preferences with $0 < p < 1$, for which the portfolio risk position is measured as follows:

$$\inf_{\theta} \left\{ \theta + \frac{1}{1-p} \mathbb{E}\big((\boldsymbol{x}^T \boldsymbol{X}_t - \theta)_+ \big) \right\};$$

see Rockafellar and Uryasev (2002). Hence, the simple non-parametric estimator is

$$\widehat{\mathcal{R}}_{cvar}^{emp}(\boldsymbol{x}) := \inf_{\theta} \left\{ \theta + \frac{1}{n(1-p)} \sum_{t=1}^{n} \left(\boldsymbol{x}^T \boldsymbol{X}_t - \theta \right)_+ \right\},$$

which is convex, homogeneous, but not differentiable, though differentiable almost everywhere implied by the convexity. To have a differentiable estimator, one can use the smooth non-parametric estimation in Scaillet (2004) and Chen (2008), defined as

$$\widehat{\mathcal{R}}_{cvar}^{KD}(\boldsymbol{x}) := \frac{1}{n(1-p)} \sum_{t=1}^{n} \boldsymbol{x}^T \boldsymbol{X}_t \left\{ 1 - K\left(\frac{\theta - \boldsymbol{x}^T \boldsymbol{X}_t}{h} \right) \right\},$$

where $\theta = \theta(\boldsymbol{x})$ solves

$$\frac{1}{n} \sum_{t=1}^{n} K\left(\frac{\theta - \boldsymbol{x}^T \boldsymbol{X}_t}{h} \right) = p,$$

$K(\cdot)$ is a smooth distribution function on $\Re$, and $h = h(n) > 0$ is the kernel bandwidth. Unfortunately, we cannot ensure $\widehat{\mathcal{R}}_{cvar}^{KD}(\boldsymbol{x})$ to be convex and homogeneous. By writing that

$$\mathbb{E}\big((\boldsymbol{x}^T \boldsymbol{X}_t - \theta)_+ \big) = \int (\boldsymbol{x}^T \boldsymbol{s} - \theta)_+ f_{\boldsymbol{X}}(s_1, \cdots, s_d)\, d\mathbf{s},$$

where $\boldsymbol{s} = (s_1, \cdots, s_d)^T$ and $f_{\boldsymbol{X}}(\boldsymbol{s})$ is the density function of $\boldsymbol{X}_t$, we propose the following smooth non-parametric estimator

$$\widehat{\mathcal{R}}_{cvar}(\boldsymbol{x}) := \inf_{\theta} \left\{ \theta + \frac{1}{n(1-p)} \sum_{t=1}^{n} \int (\boldsymbol{x}^T \boldsymbol{s} - \theta)_+ \prod_{i=1}^{d} h_i^{-1} k\left(\frac{s_i - X_{t,i}}{h_i} \right) d\mathbf{s} \right\},$$

where $k(\cdot) = K'(\cdot)$ on $\Re$, and $h_i = h_i(n) > 0$ is a bandwidth for all $i \in \{1, 2, \ldots, d\}$. It is straightforward to verify that $\widehat{\mathcal{R}}_{cvar}(\boldsymbol{x})$ is convex, homogeneous with order one, and differentiable everywhere. Also,

$$\widehat{\mathcal{R}}_{cvar}(\boldsymbol{x}) = \frac{1}{n(1-p)} \sum_{t=1}^{n} \int I(\boldsymbol{x}^T \boldsymbol{s} > \theta) \boldsymbol{x}^T \boldsymbol{s} \prod_{i=1}^{d} h_i^{-1} k\left(\frac{s_i - X_{t,i}}{h_i} \right) d\mathbf{s}, \tag{13}$$

with $\theta = \theta(\boldsymbol{x})$ satisfying

$$1 - \frac{1}{n(1-p)} \sum_{t=1}^{n} \int I(\boldsymbol{x}^T \boldsymbol{s} > \theta) \prod_{i=1}^{d} h_i^{-1} k\left(\frac{s_i - X_{t,i}}{h_i}\right) d\mathbf{s} = 0 \tag{14}$$

and I denoting the indicator function with $I(A) = 1$ if A is true, and $I(A) = 0$ otherwise. Hence, using $\tau = 1$ for CVaR risk measure and taking $\lambda = 1$ in Theorem 1, we estimate $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ by

$$\widehat{\boldsymbol{x}}_{cvar} = \arg\min_{\boldsymbol{x} \in \Re^d_{++}} \widehat{\mathcal{R}}_{cvar}(\boldsymbol{x}) - \sum_{i=1}^{d} b_i \log x_i \quad \text{and} \quad \widehat{\boldsymbol{\alpha}}_{cvar} = \widehat{\boldsymbol{x}}_{cvar} / \mathbf{1}^T \widehat{\boldsymbol{x}}_{cvar}.$$

That is, $\hat{\boldsymbol{x}}_{cvar}$ and $\hat{\theta}_{cvar} = \theta(\hat{\boldsymbol{x}}_{cvar})$ solve the system of equations for $\boldsymbol{x} > \mathbf{0}$:

$$\begin{cases} \dfrac{1}{n(1-p)} \displaystyle\sum_{t=1}^{n} \int I(\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})) s_i \prod_{j=1}^{d} h_j^{-1} k\left(\frac{s_j - X_{t,j}}{h_j}\right) d\mathbf{s} - \dfrac{b_i}{x_i} = 0 \text{ for } i \in \{1, 2, \ldots, d\}, \\ 1 - \dfrac{1}{n(1-p)} \displaystyle\sum_{t=1}^{n} \int I(\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})) \prod_{j=1}^{d} h_j^{-1} k\left(\frac{s_j - X_{t,j}}{h_j}\right) d\mathbf{s} = 0. \end{cases} \tag{15}$$

On the other hand, the true values $\boldsymbol{x}_0$ and $\theta_0 = \theta(\boldsymbol{x}_0)$ solve

$$\mathbb{E}[\bar{\boldsymbol{Z}}_t(\boldsymbol{x}, \theta)] = \mathbf{0} \quad \text{for } \boldsymbol{x} > \mathbf{0}, \tag{16}$$

where $\bar{\boldsymbol{Z}}_t(\boldsymbol{x}, \theta) = \big(\bar{Z}_{t,1}(\boldsymbol{x}, \theta), \ldots, \bar{Z}_{t,d+1}(\boldsymbol{x}, \theta)\big)^T$ is given by

$$\begin{cases} \bar{Z}_{t,i}(\boldsymbol{x}, \theta) = \frac{1}{1-p} X_{t,i} I\big(\boldsymbol{x}^T \boldsymbol{X}_t > \theta(\boldsymbol{x})\big) - \frac{b_i}{x_i} \text{ for all } i \in \{1, 2, \ldots, d\}, \\ \bar{Z}_{t,d+1}(\boldsymbol{x}, \theta) = 1 - \frac{1}{1-p} I\big(\boldsymbol{x}^T \boldsymbol{X}_t > \theta(\boldsymbol{x})\big). \end{cases}$$

Define $\bar{\Gamma}(\boldsymbol{x}, \theta) = \mathbb{E}\bar{\boldsymbol{Z}}_1(\boldsymbol{x}, \theta)$ and denote the partial derivatives of $\bar{\Gamma}(\cdot, \cdot)$ by $\dot{\bar{\Gamma}}(\cdot, \cdot)$ on $\Re_d \times \Re$. We assume the following regularity conditions to derive the asymptotic limits of $\hat{\boldsymbol{x}}_{cvar}$, $\hat{\theta}_{cvar}$, and $\hat{\boldsymbol{\alpha}}_{cvar}$:

C1) $\{\boldsymbol{X}_t\}_{t=-\infty}^{\infty}$ is a strictly stationary α-mixing sequence with $\alpha_{\boldsymbol{X}}(m) = O\big(a^m\big)$ for some $a \in (0, 1)$ as $m \to \infty$. Furthermore, assume $\mathbb{E}||\boldsymbol{X}_t||^{2+\delta} < \infty$ for some $\delta > 0$, where $||\cdot||$ is the l_2 norm.

C2) $(\boldsymbol{x}_0^T, \theta_0)^T$ is the unique solution to (16).

C3) The probability density function of $\boldsymbol{X}_t$ has bounded second partial derivatives on the closure of $\Omega = \cup_{(\boldsymbol{x}^T, \theta)^T \in \Omega_0} \big\{\boldsymbol{s} \in \Re_r : \boldsymbol{x}^T \boldsymbol{s} \geq \theta\big\}$, where Ω_0 is an open set covering $(\boldsymbol{x}_0^T, \theta_0)^T$. For any $s \geq 1$, the joint density of $\boldsymbol{X}_t$ and $\boldsymbol{X}_{t+s}$ has bounded second partial derivatives on the closure of $\Omega \times \Omega$.

C4) $k(\cdot)$ is a symmetric density function on $[-1, 1]$. For each $i \in \{1, 2, \ldots, d\}$, $h_i = c_i n^{-1/3}$ for some positive constant c_i.

The next theorem provides the result for deriving inference when risk preferences are ordered by the CVaR_p risk measure.

THEOREM 3. *Assume conditions C1)–C4) hold and consider the case in which* $\varphi = \mathrm{CVaR}_p$ *with* $0 < p < 1$. *Then, there is a positive definite matrix* $\bar{\Sigma}$ *such that*

$$\mathbb{E}\left\{\bar{\boldsymbol{Z}}_1(\boldsymbol{x}_0,\theta_0)\bar{\boldsymbol{Z}}_1^T(\boldsymbol{x}_0,\theta_0)\right\} + 2\lim_{n\to\infty}\sum_{m=1}^{n-1}\mathbb{E}\left\{\bar{\boldsymbol{Z}}_1(\boldsymbol{x}_0,\theta_0)\bar{\boldsymbol{Z}}_{1+m}^T(\boldsymbol{x}_0,\theta_0)\right\} = \bar{\Sigma}. \tag{17}$$

Furthermore, as $n\to\infty$,

$$\sqrt{n}\left(\hat{\boldsymbol{x}}_{cvar}^T - \boldsymbol{x}_0^T, \hat{\theta}_{cvar} - \theta_0\right)^T \xrightarrow{w} N\left(\mathbf{0}, \dot{\bar{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\bar{\Sigma}\left(\dot{\bar{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\right)^T\right), \tag{18}$$

$$\sqrt{n}(\hat{\boldsymbol{\alpha}}_{cvar} - \boldsymbol{\alpha}_0) \xrightarrow{w} N\left(\mathbf{0}, \frac{\bar{\Sigma}_0}{(\mathbf{1}^T\boldsymbol{x}_0)^2} - \frac{2\boldsymbol{x}_0\mathbf{1}^T\bar{\Sigma}_0}{(\mathbf{1}^T\boldsymbol{x}_0)^3} + \frac{\boldsymbol{x}_0\mathbf{1}^T\bar{\Sigma}_0\mathbf{1}\boldsymbol{x}_0^T}{(\mathbf{1}^T\boldsymbol{x}_0)^4}\right), \tag{19}$$

where $\bar{\Sigma}_0$ *is the first* $d\times d$ *matrix of* $\dot{\bar{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\bar{\Sigma}\left(\dot{\bar{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\right)^T$.

Next, we study the variance risk preferences. Clearly, the portfolio risk position is measured by the following non-parametric estimator:

$$\widehat{\mathcal{R}}_v(\boldsymbol{x}) = \frac{1}{n}\sum_{t=1}^n \boldsymbol{x}^T\boldsymbol{X}_t\boldsymbol{X}_t^T\boldsymbol{x} - \left(\frac{1}{n}\sum_{t=1}^n \boldsymbol{x}^T\boldsymbol{X}_t\right)^2,$$

which is convex, homogeneous, and differentiable. Using $\tau = 2$ for the variance risk measure and taking $\lambda = 1$ in Theorem 1, we estimate $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ by

$$\widehat{\boldsymbol{x}}_v = \arg\min_{\boldsymbol{x}\in\Re_{++}^d} \frac{1}{2}\widehat{\mathcal{R}}_v(\boldsymbol{x}) - \sum_{i=1}^d b_i\log x_i \quad \text{and} \quad \widehat{\boldsymbol{\alpha}}_v = \widehat{\boldsymbol{x}}_v/\mathbf{1}^T\widehat{\boldsymbol{x}}_v.$$

That is, $\hat{\boldsymbol{x}}_v$ and $\hat{\theta}_v = \theta(\hat{\boldsymbol{x}}_v)$ solve the system of equations for $\boldsymbol{x} > \mathbf{0}$:

$$\begin{cases} \frac{1}{n}\sum_{t=1}^n X_{t,i}\{\boldsymbol{x}^T\boldsymbol{X}_t - \theta\} - \frac{b_i}{x_i} = 0 \text{ for } i\in\{1,2,\ldots,d\}, \\ \frac{1}{n}\sum_{t=1}^n \boldsymbol{x}^T\boldsymbol{X}_t = \theta. \end{cases} \tag{20}$$

On the other hand, the true values $\boldsymbol{x}_0$ and $\theta_0 = \theta(\boldsymbol{x}_0)$ solve

$$\mathbb{E}[\tilde{\boldsymbol{Z}}_t(\boldsymbol{x},\theta)] = \mathbf{0} \quad \text{for } \boldsymbol{x} > \mathbf{0}, \tag{21}$$

where $\tilde{\boldsymbol{Z}}_t(\boldsymbol{x},\theta)=\left(\tilde{Z}_{t,1}(\boldsymbol{x},\theta),\ldots,\tilde{Z}_{t,d+1}(\boldsymbol{x},\theta)\right)^T$ is given by

$$\begin{cases} \tilde{Z}_{t,i}(\boldsymbol{x},\theta)=X_{t,i}\{\boldsymbol{x}^T\boldsymbol{X}_t-\theta\}-\frac{b_i}{x_i} \text{ for all } i=k\in\{1,2,\ldots,d\}, \\ \tilde{Z}_{t,d+1}(\boldsymbol{x},\theta)=\boldsymbol{x}^T\boldsymbol{X}_t. \end{cases}$$

Define $\tilde{\Gamma}(\boldsymbol{x},\theta)=E\tilde{\boldsymbol{Z}}_1(\boldsymbol{x},\theta)$ and denote the partial derivatives of $\tilde{\Gamma}(\cdot,\cdot)$ by $\dot{\tilde{\Gamma}}(\cdot,\cdot)$ on $\Re_d\times\Re$.

The following regularity conditions are required for deriving the asymptotic behavior of our estimators, namely $\hat{\boldsymbol{x}}_v$, $\hat{\theta}_v$, and $\hat{\boldsymbol{\alpha}}_v$. These conditions are formalized below:

C5) $\{\boldsymbol{X}_t\}_{t=-\infty}^{\infty}$ is a strictly stationary α-mixing sequence with $\alpha_{\boldsymbol{X}}(m)=O\big(a^m\big)$ for some $a\in(0,1)$ as $m\to\infty$. Furthermore, assume $E||\boldsymbol{X}_t||^{4+\delta}<\infty$ for some $\delta>0$.

C6) $(\boldsymbol{x}_0^T,\theta_0)^T$ is the unique solution to (21).

The next theorem provides the result for deriving inference when risk preferences are ordered by the variance as a risk measure.

Theorem 4. *Assume conditions C5) and C6) hold and consider the variance risk measure, i.e., $\varphi=var$. Then, there is a positive definite matrix $\tilde{\Sigma}$ such that*

$$\mathbb{E}\left\{\tilde{\boldsymbol{Z}}_1(\boldsymbol{x}_0,\theta_0)\tilde{\boldsymbol{Z}}_1^T(\boldsymbol{x}_0,\theta_0)\right\}+2\lim_{n\to\infty}\sum_{m=1}^{n-1}\mathbb{E}\left\{\tilde{\boldsymbol{Z}}_1(\boldsymbol{x}_0,\theta_0)\tilde{\boldsymbol{Z}}_{1+m}^T(\boldsymbol{x}_0,\theta_0)\right\}=\tilde{\Sigma}. \tag{22}$$

Furthermore, as $n\to\infty$,

$$\sqrt{n}\big(\hat{\boldsymbol{x}}_v^T-\boldsymbol{x}_0^T,\hat{\theta}_v-\theta_0\big)^T\xrightarrow{w}N\Big(\boldsymbol{0},\dot{\tilde{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\tilde{\Sigma}\big(\dot{\tilde{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\big)^T\Big), \tag{23}$$

and

$$\sqrt{n}(\hat{\boldsymbol{\alpha}}_v-\boldsymbol{\alpha}_0)\xrightarrow{w}N\left(\boldsymbol{0},\frac{\tilde{\Sigma}_0}{(\boldsymbol{1}^T\boldsymbol{x}_0)^2}-\frac{2\boldsymbol{x}_0\boldsymbol{1}^T\tilde{\Sigma}_0}{(\boldsymbol{1}^T\boldsymbol{x}_0)^3}+\frac{\boldsymbol{x}_0\boldsymbol{1}^T\tilde{\Sigma}_0\boldsymbol{1}\boldsymbol{x}_0^T}{(\boldsymbol{1}^T\boldsymbol{x}_0)^4}\right), \tag{24}$$

where $\tilde{\Sigma}_0$ is the first $d\times d$ matrix of $\dot{\tilde{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\tilde{\Sigma}\big(\dot{\tilde{\Gamma}}^{-1}(\boldsymbol{x}_0,\theta_0)\big)^T$.

A small simulation study for our main results in Theorem 3 is provided in Appendix EC.4, where we show that our CVaR estimator performs very well.

4. RP Portfolio Performance

We provide a comprehensive empirical analysis based on our previous methodologies. Section 4.1 includes some background necessary for the next two sections. Section 4.2 presents extensive comparative portfolio performance of RP portfolios and other benchmark portfolios. Section 4.3 investigates the performance of RP portfolios when the opportunity set is defined by SRI meanings.

The financial data and SRI attributes used in this section are described in Appendix EC.5.1, while the structural breaks are detailed in Appendix EC.5.2. To this end, we follow 408 firms from year 2001 to 2020. The analyses in Sections 4.1 and 4.2 are using the top and bottom US firms, totaling 100; the analysis in Section 4.3 is based on the entire sample, where additional non-monetary (SRI) preferences are applied for selecting the opportunity set. All details regarding the RB/RP implementations are provided in Appendix EC.5.3.

4.1. Background and Preliminary Results

This section first describes the six portfolios based on RP principles and also a benchmark portfolio, they are further examined in Sections 4.2 and 4.3. We then provide a real data analysis that illustrates the differences between the $RP-\text{CVaR}$ portfolios by comparing the *independent and identically distributed (i.i.d.)* and non-i.i.d. CVaR non-parametric estimators.

The following portfolios are compared in our numerical analyses:

1. *Equal weight (EW)* portfolio, denoted as $\boldsymbol{\alpha}^{(EW)} = \frac{1}{d}\mathbf{1}$;

2. *RP-SD*, where the variance is estimated via the $\widehat{\mathcal{R}}_v$ estimator; recall that RP portfolios based on SD and var are the same;

3. RP-$\text{CVaR}^{iid}_{95\%}$, where $\text{CVaR}_{95\%}$ is estimated via the i.i.d. $\widehat{\mathcal{R}}^{emp}_{cvar}$ estimator, which is further recalled as $\boldsymbol{\alpha}^{(RP-\text{CVaR}^{iid}_{95\%})}$;

4. RP-$\text{CVaR}^{niid}_{95\%}$, where $\text{CVaR}_{95\%}$ is estimated via the non-i.i.d $\widehat{\mathcal{R}}_{cvar}$ estimator, which is further recalled as $\boldsymbol{\alpha}^{(RP-\text{CVaR}^{niid}_{95\%})}$;

5. IWP-SD is the portfolio with the following weights:

$$\alpha_k^{(IWP-SD)} = \frac{1/\widehat{\mathcal{R}}_v(X_k)}{\sum_{k=1}^{d} 1/\widehat{\mathcal{R}}_v(X_k)} \quad \text{for all } k \in \{1,2,\ldots,d\};$$

6. IWP-CVaR$_{95\%}^{iid}$ is the portfolio with the following weights:

$$\alpha_k^{(IWP-\text{CVaR}_{95\%}^{iid})} = \frac{1/\widehat{\mathcal{R}}_{cvar}^{emp}(X_k)}{\sum_{k=1}^{d} 1/\widehat{\mathcal{R}}_{cvar}^{emp}(X_k)} \quad \text{for all } k \in \{1, 2, \ldots, d\};$$

7. *S&P 500*, the benchmark S&P 500 portfolio.

EW is a standard benchmark portfolio in the portfolio construction literature, see DeMiguel et al. (2009b); the other benchmark portfolio in our analysis is S&P 500. The remaining five portfolios are selected by searching for RP under various risk preferences. That is, Portfolios 2, 3, and 4 are standard RP portfolios for SD and CVaR risk preferences. Portfolios 5 and 6 are inverse volatility weighted and inverse CVaR weighted, which are early implementations of RP portfolios; essentially, these are heuristic methods with assets being weighted in inverse proportion to their risk (SD and CVaR, respectively). The inverse volatility weighted portfolio is a very popular portfolio due to its simple implementation, but these two portfolios are not the same; it is true that RP-SD and IWP-SD coincide if the asset correlations are equal (Roncalli 2013), but equivalent results are unknown for other risk preferences. For these reasons, the inverse weighted portfolios are (computationally) simplified RP portfolios (Qian 2005, Chaves et al. 2011, Ang 2014). In summary, only the (standard) RP portfolios require bespoke algorithms, which are explained in Appendix EC.5.3.

The second part of this section is to understand the effect of using our proposed non-i.i.d. CVaR estimator instead of the classical i.i.d. CVaR estimator over the RP-CVaR portfolio performance. That is, we use financial data to evaluate and compare RP-CVaR$_{95\%}^{iid}$ and RP-CVaR$_{95\%}^{niid}$ that are SOCP implementations as in Mausser and Romanko (2018); these are showcased in Figure 1. The opportunity set consists of the top 50 firms (ranked on *All Factors Score (AFS)*) from the US-subsample consisiting of 100 firms.

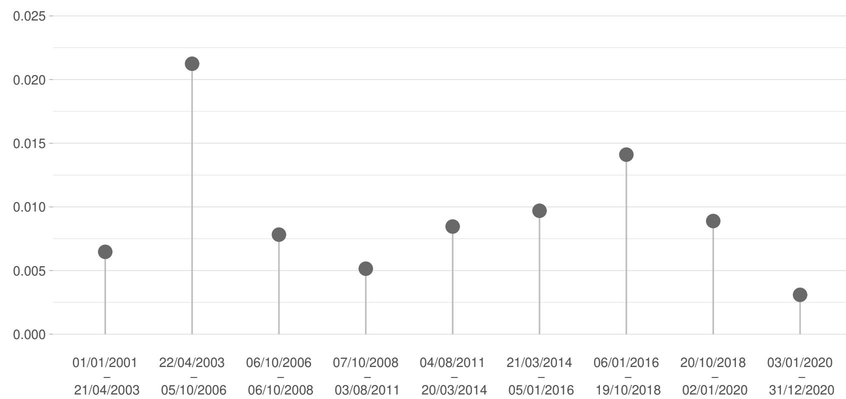

Figure 1 The l_1 distances between $\boldsymbol{\alpha}^{(RP-\mathrm{CVaR}^{iid}_{95\%})}$ and $\boldsymbol{\alpha}^{(RP-\mathrm{CVaR}^{niid}_{95\%})}$ estimates based on the values of top 50 (AFS ranked) US firms during the respective periods.

Figure 1 shows the difference between the RP-CVaR$^{iid}_{95\%}$ and RP-CVaR$^{niid}_{95\%}$ portfolios by plotting the l_1 distances between $\boldsymbol{\alpha}^{(RP-\mathrm{CVaR}^{iid}_{95\%})}$ and $\boldsymbol{\alpha}^{(RP-\mathrm{CVaR}^{niid}_{95\%})}$, which are computed so that we understand how the market conditions may affect the behavior of the two estimators. The estimators are applied only to the data in that particular period. This plot shows that the two portfolios in Periods 2 and 7 are more different than other periods; these periods are characterized by irrational exuberance on bull market periods when the views of the majority of investors are highly correlated driven by the self-fulfilling prophecy that stocks have an ascending trend. Quite small differences are observed in Periods 1, 4 and 9 that are associated with market disruptions and crises. Further analyses for various values of the serial dependence parameter ρ are discussed in the Appendix in Figure EC.4.

4.2. Risk Parity Portfolio Performance

The comparative portfolios performance is analyzed from 2001 to 2020. We characterize the financial stock market by the S&P 500 index and we identify several structural break points using the Bai-Perron test (Bai and Perron 1998, 2003). The structural breaks are illustrated in Figure EC.5 in Appendix EC.5.2. For the first structural break (denoted as ‘Period 1’, which is between 01/01/2002 and 21/04/2003), we used the historical prices between 01/01/2001 and 30/12/2001 that is prior to ‘Period 1’ to calculate the asset weights for each portfolio. After that, we used all historical prices before each structural break period to get the portfolio weights to be used in an out-of-sample fashion for the next period; e.g., for the 'Period 2', the historical data between 01/01/2001 and 21/04/2003 are used.

Our empirical application employs a recursive estimation framework up to the last structural break for the calculation of portfolio weights. The portfolios constructed in this manner are then compared for their financial performance over the next period up to the next structural break. Comparing different portfolios strategies based on the out-of-sample results reflects the important questions of "what was believed and projected, at what time, using information available at the time", as advocated in Diebold (2015). Our portfolios are compared over the nine periods, with main results summarized in Tables 2 and 3.

Several performance measures are reported in Table 2, including *Sharpe ratio (SR)*, *skew-Adjusted Sharpe ratio (skew-Adj SR)* and *Calmar ratio* in addition to the standard expected return and SD. It can be noted that skew-Adj SR incorporates a penalty factor for negative skewness, while the Calmar ratio is defined as the ratio of annualized return over the absolute value of the maximum drawdown of an investment computed over the last 36 months.

Portfolio performance has been directly linked with the idea of diversification for a long time in the literature. A recent review on diversification measures appears in Koumou (2020), but two choices are more popular. The first one is given by the p-norms for which EW is the most diversified portfolio. A second approach (that we employ in this paper) is the *Diversification Index (DI)*; by definition, the DI for a linear portfolio $\boldsymbol{\alpha}^T\mathbf{X}$ with risk preferences order by the risk measure φ is given by $DI(\boldsymbol{\alpha}) := \varphi(\boldsymbol{\alpha}^T\mathbf{X})/\sum_{k=1}^{d}\alpha_k\varphi(X_k)$.

Four performance measures are provided in Table 3. The first two are SR-like measurements, where the CVaR-based and VaR-based SRs replace the SD measure of market risk with CVaR and VaR, respectively; the two DIs performance measures are computed with $\varphi \in \{SD, \text{CVaR}\}$ risk preferences. Note that all CVaR estimates in Table 3 rely on the standard non-parametric estimator, $\text{CVaR}_{95\%}^{iid}$; further, all computations in Table 3 are not annualized due to the *PerformanceAnalytics*

Table 2 Various Performance Measures for Our Portfolios and All Periods

Structural Break	Portfolio	Mean	SD	SD-SR	skew-Adj SR	Calmar Ratio
Period 1	EW	-0.0661	0.2425	-0.2727	-0.3626	-0.3098
01/01/2002	RP-SD	-0.0477	0.1845	-0.2585	-0.3303	-0.2835
-	RP-CVaR$^{iid}_{95\%}$	**-0.0145**	**0.1660**	**-0.0871**	**-0.1653**	**-0.1374**
21/04/2003	RP-CVaR$^{niid}_{95\%}$	**-0.0142**	**0.1662**	**-0.0852**	**-0.1636**	**-0.1360**
	IWP-SD	-0.0568	0.2206	-0.2576	-0.3427	-0.3055
	IWP-CVaR$^{iid}_{95\%}$	-0.0559	0.2220	-0.2517	-0.3378	-0.3015
	S&P500	-0.1455	0.2539	-0.5733	-0.6042	-0.5090
Period 2	EW	0.1780	0.1125	1.5830	1.6541	2.5022
22/04/2003	RP-SD	0.1783	**0.0922**	**1.9344**	**1.9696**	**3.8350**
-	RP-CVaR$^{iid}_{95\%}$	**0.1787**	0.0936	1.9098	1.9365	3.3021
05/10/2006	RP-CVaR$^{niid}_{95\%}$	**0.1788**	**0.0935**	**1.9103**	**1.9370**	**3.3037**
	IWP-SD	0.1698	0.1053	1.6125	1.6820	2.6904
	IWP-CVaR$^{iid}_{95\%}$	0.1702	0.1063	1.6004	1.6699	2.6320
	S&P500	0.1458	0.1139	1.2797	1.2998	2.0814
Period 3	EW	-0.0737	0.1980	-0.3721	-0.4458	-0.3168
06/10/2006	RP-SD	**-0.0473**	0.1610	**-0.2936**	**-0.3649**	**-0.2402**
-	RP-CVaR$^{iid}_{95\%}$	-0.0489	**0.1572**	-0.3113	-0.3812	-0.2458
06/10/2008	RP-CVaR$^{niid}_{95\%}$	**-0.0488**	**0.1574**	**-0.3107**	**-0.3807**	**-0.2457**
	IWP-SD	-0.0668	0.1955	-0.3418	-0.4163	-0.3036
	IWP-CVaR$^{iid}_{95\%}$	-0.0682	0.1960	-0.3480	-0.4221	-0.3064
	S&P500	-0.0814	0.2051	-0.3969	-0.4769	-0.3246
Period 4	EW	**0.1684**	0.2886	0.5836	0.4488	0.3896
07/10/2008	RP-SD	0.1461	0.2307	0.6332	0.5164	0.4128
-	RP-CVaR$^{iid}_{95\%}$	0.1467	**0.2281**	**0.6430**	**0.5282**	**0.4245**
03/08/2011	RP-CVaR$^{niid}_{95\%}$	0.1467	**0.2280**	**0.6431**	**0.5283**	**0.4246**
	IWP-SD	0.1550	0.2685	0.5775	0.4496	0.3801
	IWP-CVaR$^{iid}_{95\%}$	**0.1555**	0.2692	0.5777	0.4496	0.3798
	S&P500	0.1112	0.2854	0.3896	0.2528	0.2016
Period 5	EW	**0.2400**	0.1694	1.4173	0.6401	1.9833
04/08/2011	RP-SD	0.2030	0.1351	1.5025	0.6512	2.4106
-	RP-CVaR$^{iid}_{95\%}$	0.2040	**0.1313**	**1.5529**	0.6464	**2.5333**
20/03/2014	RP-CVaR$^{niid}_{95\%}$	0.2042	**0.1314**	**1.5533**	0.6452	**2.5343**
	IWP-SD	0.2137	0.1563	1.3668	0.6810	2.1216
	IWP-CVaR$^{iid}_{95\%}$	**0.2146**	0.1569	1.3680	**0.6817**	2.1120
	S&P500	0.1895	0.1692	1.1203	**0.7079**	1.5981
Period 6	EW	0.0566	0.1306	0.4331	0.3649	0.3837
21/03/2014	RP-SD	0.0729	0.1093	0.6667	0.5875	0.6214
-	RP-CVaR$^{iid}_{95\%}$	**0.0751**	**0.1067**	**0.7044**	**0.6237**	**0.6535**
05/01/2016	RP-CVaR$^{niid}_{95\%}$	**0.0753**	**0.1065**	**0.7072**	**0.6261**	**0.6558**
	IWP-SD	0.0704	0.1259	0.5591	0.4873	0.5414
	IWP-CVaR$^{iid}_{95\%}$	0.0696	0.1263	0.5509	0.4793	0.5302
	S&P500	0.0748	0.1374	0.5446	0.4714	0.5822
Period 7	EW	**0.1720**	0.1168	1.4728	**0.6437**	1.9575
06/01/2016	RP-SD	0.1655	0.0947	1.7474	0.6273	2.1749
-	RP-CVaR$^{iid}_{95\%}$	**0.1716**	**0.0924**	**1.8558**	0.6387	**2.4257**
19/10/2018	RP-CVaR$^{niid}_{95\%}$	0.1714	**0.0925**	**1.8535**	**0.6415**	**2.4231**
	IWP-SD	0.1562	0.1090	1.4326	0.6076	1.6828
	IWP-CVaR$^{iid}_{95\%}$	0.1566	0.1093	1.4329	0.6080	1.6895
	S&P500	0.1625	0.1167	1.3923	0.5717	1.7320
Period 8	EW	0.1337	0.1440	0.9289	0.7841	0.6949
20/10/2018	RP-SD	0.1394	0.1168	1.1939	0.9430	0.9244
-	RP-CVaR$^{iid}_{95\%}$	**0.1456**	**0.1133**	**1.2847**	**0.9986**	**1.0016**
02/01/2020	RP-CVaR$^{niid}_{95\%}$	**0.1453**	**0.1134**	**1.2807**	**0.9958**	**0.9982**
	IWP-SD	0.1299	0.1351	0.9616	0.7864	0.7347
	IWP-CVaR$^{iid}_{95\%}$	0.1308	0.1354	0.9655	0.7897	0.7374
	S&P500	0.1253	0.1482	0.8456	0.6972	0.6493
Period 9	EW	0.2095	0.3664	0.5718	0.3863	0.4165
03/01/2020	RP-SD	0.2153	0.2939	0.7324	0.5298	0.6022
-	RP-CVaR$^{iid}_{95\%}$	**0.2222**	**0.2775**	**0.8008**	**0.5672**	**0.6693**
31/12/2020	RP-CVaR$^{niid}_{95\%}$	0.2213	**0.2780**	**0.7960**	**0.5651**	**0.6640**
	IWP-SD	0.1931	0.3474	0.5559	0.3814	0.4106
	IWP-CVaR$^{iid}_{95\%}$	0.1943	0.3486	0.5574	0.3822	0.4117
	S&P500	**0.2279**	0.3443	0.6618	0.4570	0.5633

Notes: Various annualized portfolio performance measurements (within each period and performance criterion) – the **"best"** portfolio is in bold and underlined, the **"second best"** portfolio is only in bold. (Column 1) Nine structural break periods are identified with a Bai-Perron test (Bai and Perron 1998, 2003). (Column 2) Six portfolios are compared with the S&P500 benchmark (i.e., EW: Equal Weighted portfolios, RP-SD: Risk Parity portfolios based on Standard Deviation (SD), RP-CVaR$^{iid}_{95\%}$: Risk Parity portfolios based on the i.i.d. Conditional Value-at-Risk (CVaR) estimator at 95%, RP-CVaR$^{niid}_{95\%}$: Risk-Parity portfolios based on the non-i.i.d. CVaR estimator at 95%, IWP-SD: Inverse Weighted Portfolios (IWP) based on SD, IWP-CVaR$^{iid}_{95\%}$: IWP based on the i.i.d. CVaR estimator at 95%). (Column 3) Annualized Mean of each portfolio. (Column 4) Annualized SD of each portfolio. (Column 5) Annualized Sharpe Ratio (SR) based on SD. (Column 6) Annualized skew-adjusted SR, which adjusts for negative skewness. (Column 7) Calmar Ratio is the ratio of annualized return over the absolute value of the maximum drawdown of each portfolio.

R package limitations that is used for Modified $\text{VaR}_{95\%}$ computations, but this does not affect the trend captured by our analysis.

Overall, the results in Table 2 show that all RP portfolios (RP-SD, RP-$\text{CVaR}_{95\%}^{iid}$ and RP-$\text{CVaR}_{95\%}^{niid}$) are superior to the other four portfolios over a wide set of market scenarios as spanned by the structural breakpoints; as expected, these numerical experiments confirm our Note 2 vi). The annualized SD of RP portfolios is smaller than the other four portfolios, and RP portfolios perform better than all other portfolios in terms of SR. The RP-$\text{CVaR}_{95\%}^{iid}$ and RP-$\text{CVaR}_{95\%}^{niid}$ portfolios show similar performance, though there is a slight advantage for using our proposed (non-i.i.d.) estimator that does not ignore the serial correlation in the data. This confirms the results of our simulated data analysis depicted in Figure EC.4. All six portfolios outperform the S&P500 benchmark.

Table 3 indicates that RP portfolios perform really well compared to EW and IWP portfolios. This is reassuring in the sense that the RP portfolios are not loss making vis-a-vis other strategies with respect to SR or DI performance; note that the larger the DI value is, the least diversified the portfolio is. By comparing the DI measurements across all nine periods, we observe that there is a substantial increase in DI levels during Periods 4 and 9. Thus, our empirical results imply that the degree of diversification decreases during severe crises (which is equivalent to increased DI levels). Our results are qualitatively similar to those reported in Capponi and Rubtsov (2022), where it is inferred that investors tend to prefer less diversified portfolios during periods of market distress so that systemic risk is reduced. Our portfolio construction method though is different than theirs which aims at building portfolios focusing on systemic risk. All other results in Table 3 confirm the same conclusions drawn from Table 2.

4.3. Portfolio Selection and Social Responsibility Investment

An emerging investment demand is from investors who have a strong set of non-monetary preferences that are expected to be reflected in their investments. For example, pension and insurance funds in France and most of the Scandinavian countries must consider SRI portfolios; this is not only driven by compliance meanings, but also by internal policies of such funds due to an increased

Table 3 Modified Sharpe Ratios and Diversification Index Values for Our Portfolios and All Periods

Period	Portfolio	Modified SR ($CVaR_{95\%}$)	Modified SR ($VaR_{95\%}$)	DI (SD)	DI ($CVaR_{95\%}$)
Period 1	EW	-0.0091	-0.0114	0.5796	0.5218
01/01/2002	RP-SD	-0.0079	-0.0106	0.4876	0.4741
-	RP-$CVaR^{iid}_{95\%}$	**-0.0026**	**-0.0036**	**<u>0.4416</u>**	**<u>0.4391</u>**
21/04/2003	RP-$CVaR^{niid}_{95\%}$	**<u>-0.0025</u>**	**<u>-0.0036</u>**	**0.4418**	**0.4394**
	IWP-SD	-0.0082	-0.0107	0.5756	0.5357
	IWP-$CVaR^{iid}_{95\%}$	-0.0081	-0.0105	0.5756	0.5331
Period 2	EW	0.0500	0.0655	0.4787	0.4611
22/04/2003	RP-SD	**<u>0.0609</u>**	**<u>0.0811</u>**	0.3959	0.3792
-	RP-$CVaR^{iid}_{95\%}$	0.0600	0.0799	**<u>0.3929</u>**	**<u>0.3756</u>**
05/10/2006	RP-$CVaR^{niid}_{95\%}$	**0.0601**	**0.0799**	**0.3930**	**0.3757**
	IWP-SD	0.0509	0.0668	0.4769	0.4575
	IWP-$CVaR^{iid}_{95\%}$	0.0504	0.0662	0.4786	0.4595
Period 3	EW	-0.0097	-0.0140	0.5803	0.6186
06/10/2006	RP-SD	**<u>-0.0076</u>**	**<u>-0.0109</u>**	0.4931	0.5274
-	RP-$CVaR^{iid}_{95\%}$	-0.0081	-0.0114	**<u>0.4868</u>**	**<u>0.5214</u>**
06/10/2008	RP-$CVaR^{niid}_{95\%}$	**-0.0080**	**-0.0114**	**0.4873**	**0.5220**
	IWP-SD	-0.0090	-0.0130	0.5837	0.6167
	IWP-$CVaR^{iid}_{95\%}$	-0.0091	-0.0132	0.5831	0.6164
Period 4	EW	0.0152	0.0265	0.6476	0.6935
07/10/2008	RP-SD	0.0170	0.0296	0.5517	0.5660
-	RP-$CVaR^{iid}_{95\%}$	**0.0174**	**<u>0.0300</u>**	**<u>0.5410</u>**	**0.5514**
03/08/2011	RP-$CVaR^{niid}_{95\%}$	**<u>0.0174</u>**	**0.0300**	**0.5410**	**<u>0.5513</u>**
	IWP-SD	0.0151	0.0265	0.6634	0.7019
	IWP-$CVaR^{iid}_{95\%}$	0.0151	0.0264	0.6620	0.7009
Period 5	EW	0.0377	0.0560	0.6074	0.6455
04/08/2011	RP-SD	0.0400	0.0596	0.5199	0.5463
-	RP-$CVaR^{iid}_{95\%}$	**0.0410**	**0.0617**	**<u>0.5016</u>**	**<u>0.5314</u>**
20/03/2014	RP-$CVaR^{niid}_{95\%}$	**<u>0.0410</u>**	**<u>0.0617</u>**	**0.5020**	**0.5318**
	IWP-SD	0.0362	0.0540	0.6327	0.6717
	IWP-$CVaR^{iid}_{95\%}$	0.0362	0.0541	0.6311	0.6701
Period 6	EW	0.0116	0.0164	0.5565	0.5904
21/03/2014	RP-SD	0.0182	0.0254	0.4803	0.5003
-	RP-$CVaR^{iid}_{95\%}$	**0.0193**	**0.0269**	**0.4664**	**0.4835**
05/01/2016	RP-$CVaR^{niid}_{95\%}$	**<u>0.0194</u>**	**<u>0.0270</u>**	**<u>0.4656</u>**	**<u>0.4825</u>**
	IWP-SD	0.0152	0.0214	0.5805	0.6034
	IWP-$CVaR^{iid}_{95\%}$	0.0150	0.0211	0.5794	0.6025
Period 7	EW	0.0375	0.0555	0.4766	0.5135
06/01/2016	RP-SD	0.0458	0.0674	0.4035	0.4160
-	RP-$CVaR^{iid}_{95\%}$	**<u>0.0494</u>**	**<u>0.0724</u>**	**<u>0.3925</u>**	**<u>0.3999</u>**
19/10/2018	RP-$CVaR^{niid}_{95\%}$	**0.0493**	**0.0723**	**0.3925**	**0.4000**
	IWP-SD	0.0366	0.0538	0.4836	0.5095
	IWP-$CVaR^{iid}_{95\%}$	0.0367	0.0538	0.4830	0.5095
Period 8	EW	0.0235	0.0359	0.5173	0.5648
20/10/2018	RP-SD	0.0305	0.0472	0.4359	0.4640
-	RP-$CVaR^{iid}_{95\%}$	**<u>0.0335</u>**	**<u>0.0516</u>**	**<u>0.4224</u>**	**<u>0.4409</u>**
02/01/2020	RP-$CVaR^{niid}_{95\%}$	**0.0334**	**0.0514**	**0.4232**	**0.4423**
	IWP-SD	0.0238	0.0371	0.5209	0.5754
	IWP-$CVaR^{iid}_{95\%}$	0.0239	0.0372	0.5206	0.5750
Period 9	EW	0.0144	0.0230	0.7124	0.7717
03/01/2020	RP-SD	0.0183	0.0286	0.6094	0.6589
-	RP-$CVaR^{iid}_{95\%}$	**<u>0.0199</u>**	**<u>0.0307</u>**	**<u>0.5741</u>**	**<u>0.6228</u>**
31/12/2020	RP-$CVaR^{niid}_{95\%}$	**0.0197**	**0.0305**	**0.5753**	**0.6238**
	IWP-SD	0.0140	0.0225	0.7167	0.7721
	IWP-$CVaR^{iid}_{95\%}$	0.0140	0.0226	0.7171	0.7728

Notes: Various modified Sharpe Ratio (SR) that are based on the Conditional Value-at-risk (CVaR) and Value-at-risk (VaR) at level 95% risk measurements, as well as Diversification Index (DI) based on Standard Deviation (SD) and $CVaR_{95\%}$ portfolio performance measurements (within each period and performance criterion) – the **<u>"best"</u>** portfolio is in bold and underlined, the **"second best"** portfolio is only in bold. (Column 1) Nine structural break periods are identified with a Bai-Perron test (Bai and Perron 1998, 2003). (Column 2) Six portfolios are presented and compared (i.e. EW: Equal Weighted portfolios, RP-SD: Risk Parity portfolios based on SD, RP-$CVaR^{iid}_{95\%}$: Risk Parity portfolios based on the i.i.d. CVaR estimator at 95%, RP-$CVaR^{niid}_{95\%}$: Risk-Parity portfolios based on the non-i.i.d. CVaR estimator at 95%, IWP-SD: Inverse Weighted Portfolios (IWP) based on SD, IWP-$CVaR^{iid}_{95\%}$: IWP based on the i.i.d. CVaR estimator at 95%). (Column 3) Modified un-annualized SR based on the i.i.d. CVaR estimator at 95%. (Column 4) Modified un-annualized SR based on $VaR_{95\%}$. (Column 5) DI based on SD. (Column 6) DI based on the i.i.d. CVaR estimator at 95%.

demand for green assets and assets with high business ethics standards. Hence, the question is how to design such strategies that provide portfolios with certain exposures to SRI factors, but remain viable investment opportunities. A possible methodology and subsequent empirical results are described in the current section.

The SRI data analysis relies once again on the data described in Appendix EC.5.1. We only used the US subsample (that have been AFS ordered) in the previous empirical analyses. Here, the SRI analysis includes all (US and non US) 408 firms, which are ranked again based on AFS. Specifically, we create two sets of portfolios in this section: i) *High-SRI* that consists of the top 100 ranked firms, and ii) *Low-SRI* that consists of the bottom 100 ranked firms. For each of those two sets of portfolios, six portfolios are constructed as in Section 4.1, but since the opportunity sets may include non-US companies, S&P 500 is not included in the SRI analysis summarized in Table 4. As before, the performance is monitored for a buy and hold portfolio for Period k, with the portfolio constructed based on the information from Period 0 to the end of Period $k-1$.

We now compare in Table 4 the performance of six High-SRI and Low-SRI portfolios when the opportunity set is chosen through SRI meanings. In general, the standard RP portfolios, either based on SD or CVaR preferences, have a better SR than the benchmark EW portfolio and simplified RP portfolios (IWPs). In addition, the Low-SRI portfolios show superior SR performance in all periods following some type of crisis, such as Period 2 following the dot com crisis, Period 5 following the global financial crisis, and Period 9 which covers the COVID-19 period. This conjecture also holds for Period 6 following events associated to the European sovereign debt crisis (second bailout in March 2012 for Greece and rescue packages for Spain and Cyprus in June 2012).

The evolution of portfolios with SRI preferences suggests that an RP-SD type portfolio would perform "best" between 2002 and 2008, followed by an $RP-CVaR^{iid}$ for the periods 2008-2016, then followed again by an RP-SD portfolio for 2016-2018, and finally, by an $RP-CVaR^{niid}$ for the last two periods between 2018-2020. Our analysis is *ex post* but it does highlight that portfolio construction with different risk-measures may be superior in different periods. More sophisticated

Table 4 Annualized Sharpe Ratios of High-SRI versus Low-SRI Portfolios Over Time

Period	Portfolio	High-SRI	Low-SRI	Period	Portfolio	High-SRI	Low-SRI
Period 1				**Period 6**			
01/01/2002	EW	**-0.3987**	-0.8550	21/03/2014	EW	1.0067	**1.1529**
-	RP-SD	**-0.2825**	-0.7951	-	RP-SD	1.2934	**1.4828**
21/04/2003	RP-CVaR$_{95\%}^{iid}$	**-0.3156**	-0.8819	05/01/2016	RP-CVaR$_{95\%}^{iid}$	1.2870	**1.4890**
	RP-CVaR$_{95\%}^{niid}$	**-0.3168**	-0.8825		RP-CVaR$_{95\%}^{niid}$	1.2829	**1.4867**
	IWP-SD	**-0.3463**	-0.8067		IWP-SD	0.8769	**1.0963**
	IWP-CVaR$_{95\%}^{iid}$	**-0.3377**	-0.8000		IWP-CVaR$_{95\%}^{iid}$	0.8850	**1.1035**
Period 2				**Period 7**			
22/04/2003	EW	2.2504	**2.5359**	06/01/2016	EW	**0.7130**	0.7027
-	RP-SD	2.6018	**2.8093**	-	RP-SD	**0.9106**	0.8228
05/10/2006	RP-CVaR$_{95\%}^{iid}$	2.5095	**2.6900**	19/10/2018	RP-CVaR$_{95\%}^{iid}$	**0.9032**	0.7988
	RP-CVaR$_{95\%}^{niid}$	2.5069	**2.6902**		RP-CVaR$_{95\%}^{niid}$	**0.9003**	0.7991
	IWP-SD	2.3251	**2.6106**		IWP-SD	0.6654	**0.7869**
	IWP-CVaR$_{95\%}^{iid}$	2.3206	**2.5970**		IWP-CVaR$_{95\%}^{iid}$	0.6752	**0.7969**
Period 3				**Period 8**			
06/10/2006	EW	**0.8152**	0.7364	20/10/2018	EW	**0.3312**	0.0275
-	RP-SD	**1.0088**	0.6446	-	RP-SD	**0.5242**	0.0942
06/10/2008	RP-CVaR$_{95\%}^{iid}$	**0.9684**	0.7299	02/01/2020	RP-CVaR$_{95\%}^{iid}$	**0.5388**	0.0984
	RP-CVaR$_{95\%}^{niid}$	**0.9651**	0.7285		RP-CVaR$_{95\%}^{niid}$	**0.5392**	0.0968
	IWP-SD	**0.8148**	0.6776		IWP-SD	**0.3836**	0.1973
	IWP-CVaR$_{95\%}^{iid}$	**0.8209**	0.6808		IWP-CVaR$_{95\%}^{iid}$	**0.3845**	0.1924
Period 4				**Period 9**			
07/10/2008	EW	**0.2427**	0.2302	03/01/2020	EW	-0.5199	**-0.4854**
-	RP-SD	**0.2355**	0.2302	-	RP-SD	-0.3142	**-0.2961**
03/08/2011	RP-CVaR$_{95\%}^{iid}$	0.2245	**0.2435**	31/12/2020	RP-CVaR$_{95\%}^{iid}$	-0.3073	**-0.2960**
	RP-CVaR$_{95\%}^{niid}$	0.2247	**0.2424**		RP-CVaR$_{95\%}^{niid}$	-0.3063	**-0.2955**
	IWP-SD	**0.2292**	0.2010		IWP-SD	-0.4913	**-0.3799**
	IWP-CVaR$_{95\%}^{iid}$	**0.2294**	0.1988		IWP-CVaR$_{95\%}^{iid}$	-0.4780	**-0.3815**
Period 5							
04/08/2011	EW	0.6673	**0.7357**				
-	RP-SD	0.7707	**0.9316**				
20/03/2014	RP-CVaR$_{95\%}^{iid}$	0.7735	**0.9486**				
	RP-CVaR$_{95\%}^{niid}$	0.7740	**0.9486**				
	IWP-SD	**0.7572**	0.6682				
	IWP-CVaR$_{95\%}^{iid}$	**0.7581**	0.6759				

Notes: Annualized Sharpe Ratios (SR) of High-SRI (top 100 ranked firms based on All Factors Score) vs Low-SRI (bottom 100 All Factors Score ranked firms) portfolios (within each period) – the **"better"** SR (High-SRI vs Low-SRI) is in bold and the **"best"** portfolio (among the 12 portfolios in each period) is in bold and underlined. (Columns 1 and 5) Nine structural break periods are identified with a Bai-Perron test (Bai and Perron 1998, 2003). (Columns 2 and 6) Six portfolios are presented and compared (i.e. EW: Equal Weighted portfolios, RP-SD: Risk Parity (RP) portfolios based on Standard Deviation (SD), RP-CVaR$_{95\%}^{iid}$: RP portfolios based on the i.i.d. Conditional Value-at-Risk (CVaR) estimator at 95%, RP-CVaR$_{95\%}^{niid}$: RP portfolios based on the non-i.i.d. CVaR estimator at 95%, IWP-SD: Inverse Weighted Portfolios based on SD, IWP-CVaR$_{95\%}^{iid}$: Inverse Weighted Portfolios (IWP) based on the i.i.d. CVaR estimator at 95%). (Columns 3 and 7) Annualized SR based on SD of each portfolio with a High SRI score. (Columns 4 and 8) Annualized SR based on SD of each portfolio with a Low SRI score.

risk measures and more robust estimators would benefit investors during difficult market conditions while in normal market conditions standard deviation is preferable since by design it captures also the positive tail of returns. It is also evident that in bull market periods such as Periods 2, 5, and 6, the low-SRI portfolio performs better, consistent with the literature presenting evidence that in those times investors "forget" about SRI preferences. Hong and Kacperczyk (2009) demonstrated that low-SRI stocks are akin to counter-cyclical instruments, producing relatively better investment performance than comparable high-SRI stocks during low states of the economy. Period 9 is more of an anomaly period associated with the Covid-19 pandemic, where the low SRI associated preferences would produce the portfolios with smaller negative returns than high SRI portfolios.

5. Conclusions

The portfolio construction literature has considered various RB/RP formulations over the last two decades. We provide a mathematical characterization of such long-only portfolios for general risk preferences and we derive important multiple theoretical properties of RB/RP portfolios. We showed that RP portfolios are less risky than the EW benchmark portfolio for a general risk measure. In addition, we found that elliptically distributed asset returns (that include the multivariate Gaussian and t-distributed cases) lead to RP and RB investment strategies that are invariant with respect to a large class of risk measures.

Asymptotic normality for two popular RB portfolios (with variance and CVaR risk preferences) are provided, which according to our knowledge, is the very first attempt in the literature. Further, our statistical inferences introduce a novel CVaR estimator that is shown to perform very well on (serial) dependent data, which is not surprising since the asymptotic distribution of our novel estimator is found under some fairly general data dependent assumptions.

References

Ang, A. (2014). *Asset management: A systematic approach to factor investing.* Oxford University Press.

Asimit, A. V., Furman, E., Tang, Q., and Vernic, R. (2011). Asymptotics for risk capital allocations based on conditional tail expectation. *Insurance: Mathematics and Economics*, 49(3):310–324.

Asimit, V., Peng, L., Wang, R., and Yu, A. (2019). An efficient approach to quantile capital allocation and sensitivity analysis. *Mathematical Finance*, 29(4):1131–1156.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18:1–22.

Bai, X., Scheinberg, K., and Tütüncü, R. (2016). Least-square approach to risk parity in portfolio selection. *Quantitative Finance*, 16(3):357–376.

Bailey, N., Pesaran, M., and Smith, L. (2019). A multiple testing approach to the regularisation of large sample correlation matrices. *Journal of Econometrics*, 208:507–534.

Ballestero, E., Perez-Gladish, B., and Garcia-Bernabeu, A. (2015). *Socially Responsible Investment. A Multicriteria Decision Making Approach.* International Series in Operations Research & Management Science. Springer.

Bellini, F., Cesarone, F., Colombo, C., and Tardella, F. (2021). Risk parity with expectiles. *European Journal of Operational Research*, 291(3):1149–1163.

Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust Optimization*, volume 28. Princeton University Press.

Bertsimas, D., Brown, D. B., and Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501.

Bertsimas, D., Gupta, V., and Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming*, 167:235–292.

Blanchet, J., Chen, L., and Zhou, X. Y. (2022). Distributionally robust mean-variance portfolio selection with Wasserstein distances. *Management Science*, 68(9):6382–6410.

Capponi, A. and Rubtsov, A. (2022). Systemic risk-driven portfolio selection. *Operations Research*, 70(3):1598–1612.

Chaves, D., Hsu, J., Li, F., and Shakernia, O. (2011). Risk parity portfolio vs. other asset allocation heuristic portfolios. *Journal of Investing*, 20(1):108.

Chen, S. (2008). Nonparametric estimation of expected shortfall. *Journal of Financial Econometrics*, 6:67–107.

Cornuéjols, G., Peña, J., and Tütüncü, R. (2018). *Optimization Methods in Finance (2nd ed.).* Cambridge University Press.

Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612.

DeMiguel, V., Garlappi, L., Nogales, F., and Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norm. *Management Science*, 55:798–812.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22:1915–1953.

DeMiguel, V. and Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research*, 57(3):560–577.

Dhaene, J., Vanduffel, S., Goovaerts, M. J., Kaas, R., Tang, Q., and Vyncke, D. (2006). Risk measures and comonotonicity: A review. *Stochastic Models*, 22(4):573–606.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano tests. *Journal of Business & Economic Statistics*, 33(1):1–9.

Fan, J., Fan, Y., and Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147:186–197.

Föllmer, H. and Schied, A. (2011). *Stochastic Finance: An Introduction in Discrete Time*. Third ed., Walter de Gruyter.

Glasserman, P. and Xu, X. (2014). Robust risk measurement and model risk. *Quantitative Finance*, 14(1):29–58.

Goh, J. and Sim, M. (2010). Distributionally robust optimization and its tractable approximations. *Operations research*, 58(4):902–917.

Goldfarb, D. and Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38.

Hallerbach, W., Spronk, J., and Soppe, A. (2004). A framework for managing a portfolio of socially responsible investments. *European Journal of Operational Research*, 153(2):517–529.

Haugh, M., Iyengar, G., and Song, I. (2017). A generalized risk budgeting approach to portfolio construction. *Journal of Computational Finance*, 21(2):29–60.

He, X. D., Kou, S., and Peng, X. (2022). Risk measures: robustness, elicitability, and backtesting. *Annual Review of Statistics and Its Application*, 9:141–166.

Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1):15–36.

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1683.

Koumou, G. B. (2020). Diversification and portfolio theory: a review. *Financial Markets and Portfolio Management*, 34(3):267–312.

Lassance, N., DeMiguel, V., and Vrins, F. (2022). Optimal portfolio diversification via independent component analysis. *Operations Research*, 70:55–72.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Ledoit, O. and Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12):4349–4388.

Maillard, S., Roncalli, T., and Teïletche, J. (2010). The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, 36(4):60–70.

Mausser, H. and Romanko, O. (2018). Long-only equal risk contribution portfolios for CVaR under discrete distributions. *Quantitative Finance*, 18(11):1927–1945.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools-revised edition*. Princeton University Press.

Meucci, A. (2009). Managing diversification. *Risk*, pages 74–79.

Pflug, G. and Wozabal, D. (2007). Ambiguity in portfolio selection. *Quantitative Finance*, 7(4):435–442.

Popescu, I. (2007). Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112.

Qian, E. E. (2005). Risk parity portfolios: Efficient portfolios through true diversification. Technical report, Panagora.

Rockafellar, R. and Uryasev, S. (2002). Conditional Value-at-Risk for general loss distributions. *Journal of Banking & Finance*, 26:1443–1471.

Roncalli, T. (2013). *Introduction to risk parity and budgeting*. CRC Press.

Roncalli, T. and Weisang, G. (2016). Risk parity portfolio with risk factors? *Quantitative Finance*, 16:377–388.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the Natural Academy of Sciences*, 42:43–47.

Scaillet, O. (2004). Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance*, 14:115–129.

Spinu, F. (2013). An algorithm for computing risk parity weights. SSRN.

Tu, J. and Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, 99:204–215.

Tütüncü, R. H. and Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132:157–187.

Verheyden, T., Eccles, R., and Feiner, A. (2016). ESG for all? The impact of ESG screening on return, risk, and diversification. *Journal of Applied Corporate Finance*, 28:47–55.

Wiesemann, W., Kuhn, D., and Sim, M. (2014). Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376.

Electronic companion to "Constructing Optimal Portfolios under Risk Budgeting" authored by Vali Asimit, Liang Peng, Radu Tunaru and Feng Zhou

EC.1. Proofs

EC.1.1. Proof of Proposition 1

Theorem 8.28 (1) of McNeil et al. (2015) gives that $\boldsymbol{\alpha}^T\mathbf{X}$ and $||\boldsymbol{\alpha}^T\Sigma\boldsymbol{\alpha}||Z_1+\boldsymbol{\alpha}^T\boldsymbol{\mu}$ have the same distribution, which immediately justifies (3) if φ is shift invariant.

The other case, when φ is translation invariant, is also true as long as $\tau=1$. The latter follows from the fact that

$$\varphi(tY+tc)=t^\tau\varphi(Y+c)=t^\tau(\varphi(Y)+c) \quad \text{and} \quad \varphi(tY+tc)=\varphi(tY)+tc=t^\tau\varphi(Y)+tc$$

hold for any $t>0$ and $c\in\Re$, which in turn implies that $\tau=1$.

EC.1.2. Proof of Theorem 1

We first prove part a). Let $F(\mathbf{x};\lambda)$ be the objective function in (8). The first step is to show that the optimal solution in (8) exists and is an interior point of the feasible set. Now, for any $\mathbf{x}\in\Re^d_{++}$

$$\begin{aligned} F(\mathbf{x};\lambda) &= \frac{1}{\tau}\mathcal{R}\left(\frac{1}{d\max\limits_{1\le k\le d}x_k}\mathbf{x}\right)d^\tau\Big(\max_{1\le k\le d}x_k\Big)^\tau-\lambda\sum_{k=1}^d b_k\log x_k \\ &\geq \frac{\delta^* d^\tau}{\tau}\Big(\max_{1\le k\le d}x_k\Big)^\tau-\lambda\log\Big(\max_{1\le k\le d}x_k\Big), \end{aligned} \tag{EC.1}$$

since φ is homogeneous of order τ, where $\delta^*>0$ that does not depend upon $\mathbf{x}$ and its existence is guaranteed by (7). Since $\lim_{t\to\infty}\delta t^\tau-\lambda\log t=\infty$ for any $\delta,\lambda,\tau>0$, then (EC.1) implies that

$$F(\mathbf{x};\lambda)\to\infty,\ \text{whenever}\ \max_{1\le k\le d}x_k\to\infty. \tag{EC.2}$$

We now prove that (EC.2) holds on the boundary (of the feasibility set) regions away from infinity. That is, let $M>\epsilon>0$; note that for any $\mathbf{x}\in\Re^d_{++}$ such that $\min_{1\le k\le d}x_k\le\epsilon$ and $\max_{1\le k\le d}x_k\le$

M, there exists $0 < b^* \leq 1$ such that $F(\mathbf{x};\lambda) \geq -\lambda b^* \log \epsilon$ by using similar arguments as in (EC.1). Thus,

$$F(\mathbf{x};\lambda) \to \infty, \text{ whenever } \min_{1\leq k\leq d} x_k \downarrow 0 \text{ and } \max_{1\leq k\leq d} x_k \leq M \text{ for any finite } M > 0. \tag{EC.3}$$

Equations (EC.2) and (EC.3) imply that there exist an $a > 0$ and an $\epsilon \in (0, a]$ such that

$$\inf_{\mathbf{x}\in\Re^d_{++}} F(\mathbf{x};\lambda) = \inf_{\mathbf{x}\in B_{a,\epsilon}} F(\mathbf{x};\lambda), \quad \text{where } B_{a,\epsilon} := \{\mathbf{x}\in B_a : \min_{1\leq k\leq d} x_k \geq \epsilon\}$$

with $B_a := \{\mathbf{x}\in\Re^d_{++} : ||\mathbf{x}|| \leq a\}$ and $||\cdot||$ being the Euclidean distance. Since $B_{a,\epsilon}$ is a compact set, the global minimum of $F(\cdot;\lambda)$ on $\Re^d_{++}$ (denoted as $\mathbf{x}^*(\lambda,\mathbf{b})$) is an interior point of the feasibility set for any given $\lambda > 0$. Thus, (8) must have an optimal solution that is an interior point of the feasible set.

The objective function in (8) is strictly convex in $\mathbf{x}$ over the convex cone $\Re^d_{++}$ for any given $\lambda > 0$, since the logarithmic barrier term $(-\lambda\sum_{k=1}^d b_k \log x_k)$ has the same property and the fact that φ is a convex risk measure. Thus, (8) admits a unique solution.

It only remains to prove for part a) the relationship between the unique solution in (8) for various penalty parameters λ. Note that

$$F(\mathbf{x};\lambda) = \lambda F\left(\lambda^{-1/\tau}\mathbf{x};1\right) - \frac{\lambda}{\tau}\log\lambda, \text{ for any } \mathbf{x}\in\Re^d_{++} \text{ and } \lambda > 0,$$

and any given $\mathbf{b}\in\Delta_d\cap\Re^d_{++}$, since φ is a homogeneous risk measure of order τ, and in turn, $\mathbf{x}^*(\lambda,\mathbf{b}) = \lambda^{1/\tau}\mathbf{x}^*(1,\mathbf{b})$ for any $\lambda > 0$ and $\mathbf{b}\in\Delta_d\cap\Re^d_{++}$. Further,

$$F(t\mathbf{x};\lambda) = F(\mathbf{x};\lambda)t^\tau + \lambda\big(t^\tau - 1\big)\sum_{k=1}^d b_k \log x_k - \lambda\log t, \text{ for any } \mathbf{x}\in\Re^d_{++} \text{ and } \lambda, t > 0,$$

and any given $\mathbf{b}\in\Delta_d\cap\Re^d_{++}$, since φ is a homogeneous risk measure of order τ, and in turn, $F(;\lambda)$ is differentiable at $\mathbf{x}^*(\lambda,\mathbf{b})$ as $\mathcal{R}$ is differentiable at $\mathbf{x}^*(1,\mathbf{b})$ and $\mathbf{x}^*(\lambda,\mathbf{b}) = \lambda^{1/\tau}\mathbf{x}^*(1,\mathbf{b})$.

Now, the first-order conditions in (8) and the homogeneity of φ imply that $\mathcal{RC}_k$ is also homogeneous of order τ, and thus,

$$\mathcal{RC}_k\big(t\mathbf{x}^*(\lambda,\mathbf{b})\big) = b_k\mathcal{R}\big(t\mathbf{x}^*(\lambda,\mathbf{b})\big), \quad \text{for all } k\in\{1,2,\ldots,d\} \quad \text{and any } \lambda, t > 0, \tag{EC.4}$$

and any $\mathbf{b}\in\Delta_d\cap\Re^d_{++}$. Therefore, $\boldsymbol{\alpha}^*(\mathbf{b})=\mathbf{x}^*(\lambda,\mathbf{b})/\mathbf{1}^T\mathbf{x}^*(\lambda,\mathbf{b})\in\mathcal{RB}(\mathbf{b})$ since $\boldsymbol{\alpha}^*(\mathbf{b})\in\Delta_d$ by construction.

We finish the proof of part a) by noting that

$$\boldsymbol{\alpha}^*(\mathbf{b})=\frac{\mathbf{x}^*(\lambda,\mathbf{b})}{\mathbf{1}^T\mathbf{x}^*(\lambda,\mathbf{b})}=\frac{\mathbf{x}^*(1,\mathbf{b})}{\mathbf{1}^T\mathbf{x}^*(1,\mathbf{b})}\quad\text{and}\quad\mathbf{x}^*(\lambda^*,\mathbf{b})=\left(\lambda^*\right)^{1/\tau}\mathbf{x}^*(1,\mathbf{b})=\frac{\mathbf{x}^*(1,\mathbf{b})}{\mathbf{1}^T\mathbf{x}^*(1,\mathbf{b})},$$

since $\mathbf{x}^*(\lambda,\mathbf{b})=\lambda^{1/\tau}\mathbf{x}^*(1,\mathbf{b})$.

We now prove part b). As before, we initially show that (9) admits a unique solution that is an interior point.

Due to the homogeneity of $\mathcal{R}$, the objective function in (9) is unbounded in the neighbourhood of infinity. Further, for any $M>\epsilon>0$ and ϵ sufficiently small, any $\mathbf{x}\in\Re^d_{++}$ such that $\min_{1\le k\le d}x_k\le\epsilon$ and $\max_{1\le k\le d}x_k\le M$ is not feasible in (9), and in turn, the optimal solutions of (9) are interior points of the feasibility set.

We now show the uniqueness in (9). The homogeneity property of the objective function in (9) implies that any optimal solution of (9) satisfies

$$\sum_{k=1}^{d}b_k\log x_k^{**}(c,\mathbf{b})=c. \tag{EC.5}$$

One could show (EC.5) by assuming that (EC.5) does not hold, which implies that $(1-\epsilon)\mathbf{x}^{**}(c,\mathbf{b})$ is feasible for any $\epsilon>0$ sufficiently small; further,

$$\mathcal{R}\big((1-\epsilon)\mathbf{x}^{**}(c,\mathbf{b})\big)=(1-\epsilon)^\tau\mathcal{R}\big(\mathbf{x}^{**}(c,\mathbf{b})\big)<\mathcal{R}\big(\mathbf{x}^{**}(c,\mathbf{b})\big)$$

due to the homogeneity of φ and the fact $\mathcal{R}\big(x_k^{**}(c,\mathbf{b})\big)>0$ (due to (7)), which contradicts our assumption and concludes (EC.5). The optimal solution in (9) is unique, since the inequality constraint in (9) is strictly concave due to (EC.5). One could show that by assuming a case in which there are two optimal solutions, $\mathbf{x}^{**}(c,\mathbf{b})$ and $\mathbf{y}^{**}(c,\mathbf{b})$. The latter implies that

$$\mathbf{z}^{**}(c,\mathbf{b}):=\gamma\mathbf{x}^{**}(c,\mathbf{b})+(1-\gamma)\mathbf{y}^{**}(c,\mathbf{b})$$

is another optimal solution of (9) for any $0<\gamma<1$, since φ is a convex risk measure. Moreover,

$$\sum_{k=1}^{d}b_k\log z_k^{**}<\gamma\sum_{k=1}^{d}b_k\log x_k^{**}+(1-\gamma)\sum_{k=1}^{d}b_k\log y_k^{**}=c,$$

since the log function is strictly concave, which in turn contradicts that $\mathbf{z}^{**}(c,\mathbf{b})$ must satisfy (EC.5). Therefore, (9) admits a unique optimal solution that is an interior point of the feasibility set.

It only remains to prove for part b) the relationship between the unique solution in (9) for various penalty parameters c. We first show that

$$\mathbf{x}^{**}(c,\mathbf{b}) = e^{c-1}\mathbf{x}^{**}(1,\mathbf{b}) \quad \text{for any given } \mathbf{b}\in\Delta_d. \tag{EC.6}$$

Again, we show this claim by contradiction and assume that $\mathbf{x}^{**}(1,\mathbf{b})$ solves (9) when $c=1$, but there exists $c_0\neq 1$ such that $e^{c-1}\mathbf{x}^{**}(1,\mathbf{b})$ does not solve (9) whenever $c=c_0$. Therefore, there exists $\mathbf{y}\in\Re^d_{++}$ such that

$$\mathcal{R}(\mathbf{y}) < \mathcal{R}\big(e^{c_0-1}\mathbf{x}^{**}(1,\mathbf{b})\big) \quad \text{and} \quad \sum_{k=1}^{d} b_k \log y_k = c_0.$$

Clearly, the above imply that $e^{1-c_0}\mathbf{y}$ is feasible in (9) when $c=1$, and

$$\mathcal{R}\big(e^{1-c_0}\mathbf{y}\big) = e^{(1-c_0)\tau}\mathcal{R}\big(\mathbf{y}\big) < e^{(1-c_0)\tau}\mathcal{R}\big(e^{c_0-1}\mathbf{x}^{**}(1,\mathbf{b})\big) = \mathcal{R}\big(\mathbf{x}^{**}(1,\mathbf{b})\big)$$

by keeping in mind that φ is a homogeneous risk measure of order τ, which in turn contradicts our assumption and concludes (EC.6). Now, (EC.6), the homogeneity of φ and the fact that $\mathcal{R}$ is differentiable at $\mathbf{x}^{**}(1,\mathbf{b})$ imply that $\mathcal{R}(\boldsymbol{\alpha})$ is differentiable at $\mathbf{x}^{**}(c;\mathbf{b})$. All other relationships among various optimal solutions stated in part b) could be easily shown as in part a). Finally, the Slater's condition is clearly satisfied in (9), and therefore, the strong duality holds in (9). The proof of part b) is fully argued.

We show the claims from part c). Note that $\boldsymbol{\alpha}^{*}(\mathbf{b}) = \boldsymbol{\alpha}^{**}(\mathbf{b})$, which is true since

$$\boldsymbol{\alpha}^{*}(\mathbf{b}) = \mathbf{x}^{*}(\lambda,\mathbf{b})/\mathbf{1}^T\mathbf{x}^{*}(\lambda,\mathbf{b}), \quad \boldsymbol{\alpha}^{**}(\mathbf{b}) = \mathbf{x}^{**}(c,\mathbf{b})/\mathbf{1}^T\mathbf{x}^{**}(c,\mathbf{b}),$$

and the fact that there exists $\gamma^*>0$ such that $\mathbf{x}^{**}(c,\mathbf{b}) = \gamma^*\mathbf{x}^{*}(\lambda,\mathbf{b})$ for all $\lambda>0$ and any $c\in\Re$. The latter is a direct consequence of the fact that solving the primal optimal in (9) is the same as solving (8) with $\lambda=\gamma^*/\tau$, where γ^* is the dual optimal in (9) corresponding to the logarithmic constraint $\sum_{k=1}^{d} b_k\log x_k \geq c$.

The left-hand side inequality in (10) is trivial, and thus, we show now the right-hand side inequality in (10). The proof of part a) allows us to say that $\boldsymbol{\alpha}^*(\mathbf{b})$ solves

$$\min_{\mathbf{x}\in\Delta_d\cap\Re^d_{++}} \frac{1}{\tau}\mathcal{R}(\mathbf{x}) - \lambda^* \sum_{k=1}^d b_k \log x_k, \tag{EC.7}$$

which implies that

$$\frac{1}{\tau}\Big(\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) - \mathcal{R}(\mathbf{b})\Big) \le \lambda^* \sum_{k=1}^d b_k \log\left(\frac{\boldsymbol{\alpha}_k^*(\mathbf{b})}{b_k}\right) = -\lambda^* \times D_{KL}\big(\mathbf{b}||\boldsymbol{\alpha}^*(\mathbf{b})\big) \le 0$$

where $D_{KL}\big(\mathbf{b}||\boldsymbol{\alpha}^*(\mathbf{b})\big)$ is the Kullback-Leibler divergence between the probability distributions induced by (the probability vectors) $\mathbf{b}$ and $\boldsymbol{\alpha}^*(\mathbf{b})$. Therefore, $\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) \le \mathcal{R}(\mathbf{b})$ for any $\mathbf{b}$.

The very last step is to show $\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) \le \mathcal{R}\left(\frac{1}{d}\mathbf{1}\right)$. From (EC.7) we get that

$$\begin{aligned}
\frac{1}{\tau}\left(\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) - \mathcal{R}\left(\frac{1}{d}\mathbf{1}\right)\right) &\le \lambda^*\left(\sum_{k=1}^d b_k \log \boldsymbol{\alpha}_k^*(\mathbf{b}) - \sum_{k=1}^d b_k \log\left(\frac{1}{d}\right)\right) \\
&\le \lambda^*\left(\max_{\mathbf{1}^T\mathbf{x}=1} \sum_{k=1}^d b_k \log x_k + \log d\right) \\
&= \lambda^*\left(\sum_{k=1}^d b_k \log b_k + \log d\right) \\
&\le \lambda^*\left(\max_{\mathbf{1}^T\mathbf{x}=1} \sum_{k=1}^d x_k \log x_k + \log d\right) \\
&= \lambda^*\left(\sum_{k=1}^d \frac{1}{d}\log\left(\frac{1}{d}\right) + \log d\right) \\
&= 0,
\end{aligned}$$

which implies that $\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) \le \mathcal{R}\left(\frac{1}{d}\mathbf{1}\right)$, and in turn, it concludes part c).

Finally, we show the claims from part d). Relation (11) follows from (EC.7), where as before,

$$\frac{1}{\tau}\Big(\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) - \mathcal{R}\big(\tilde{\boldsymbol{\alpha}}\big)\Big) \le \lambda^* \sum_{k=1}^d b_k \log\left(\frac{\boldsymbol{\alpha}_k^*(\mathbf{b})}{\tilde{\alpha}_k}\right).$$

The left hand side is non-negative since $\tilde{\boldsymbol{\alpha}}$ is portfolio minimizer under φ risk preferences, which implies (11). Similarly, one may show (12) by noting

$$\frac{1}{\tau}\Big(\mathcal{R}\big(\boldsymbol{\alpha}^*(\mathbf{b})\big) - \mathcal{R}\big(\boldsymbol{\alpha}\big)\Big) \le \lambda^* \sum_{k=1}^d b_k \log\left(\frac{\boldsymbol{\alpha}_k^*(\mathbf{b})}{\alpha_k}\right) \text{ for any } \boldsymbol{\alpha}\in\Delta_d\cap\Re^d_{++},$$

which is negative if $\min_{1\le k\le d}\alpha_k \le \epsilon$ for a sufficiently small ϵ. This completes the proof of Theorem 1.

EC.1.3. Proof of Theorem 3

For simplicity, assume $h_i = h$ for all $i \in \{1, \dots, d\}$. Let $f_{1,1+r}(\boldsymbol{x}, \bar{\boldsymbol{x}})$ denote the joint density function of $\left(\boldsymbol{X}_t, \boldsymbol{X}_{t+r}\right)$. Put $\boldsymbol{s}, \bar{\boldsymbol{s}}, \boldsymbol{y}, \bar{\boldsymbol{y}}, \boldsymbol{x} \in \Re_d$, $\boldsymbol{Z}_t(\boldsymbol{x}, \theta) = \left(Z_{t,1}(\boldsymbol{x}, \theta), \dots, Z_{t,d+1}(\boldsymbol{x}, \theta)\right)^T$,

$$\begin{cases} Z_{t,i}(\boldsymbol{x}, \theta) = \dfrac{1}{1-p} \displaystyle\int I\left(\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})\right) s_i \prod_{j=1}^{d} h_j^{-1} k\left(\frac{X_{t,j} - s_j}{h_j}\right) d\mathbf{s} - \dfrac{b_i}{x_i} \quad \text{for all } i = 1, \dots, d, \\ Z_{t,d+1}(\boldsymbol{x}, \theta) = 1 - \dfrac{1}{1-p} \displaystyle\int I(\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})) \prod_{j=1}^{d} h_j^{-1} k\left(\frac{X_{t,j} - s_j}{h_j}\right) d\mathbf{s}. \end{cases}$$

Then, (15) becomes

$$\frac{1}{n} \sum_{t=1}^{n} \boldsymbol{Z}_t(\hat{\boldsymbol{x}}_{cvar}, \hat{\theta}_{cvar}) = \mathbf{0}. \tag{EC.8}$$

Define

$$\gamma_i(s; \boldsymbol{x}, \theta) = E\left\{\left(Z_{t,i}(\boldsymbol{x}, \theta) - \bar{Z}_{t,i}(\boldsymbol{x}, \theta)\right)\left(Z_{t+s,i}(\boldsymbol{x}, \theta) - \bar{Z}_{t+s,i}(\boldsymbol{x}, \theta)\right)\right\} - \left\{E\left(Z_{t,i}(\boldsymbol{x}, \theta) - \bar{Z}_{t,i}(\boldsymbol{x}, \theta)\right)\right\}^2$$

for $i = 1, \cdots, d+1$ and nonnegative integer s. Write

$$Z_{t,d+1}(\boldsymbol{x}, \theta) - \bar{Z}_{t,d+1}(\boldsymbol{x}, \theta) = \frac{1}{1-p} \int \left\{\prod_{j=1}^{d} k(s_j)\right\} \left\{I\left(\boldsymbol{x}^T \boldsymbol{X}_t + h\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})\right) - I\left(\boldsymbol{x}^T \boldsymbol{X}_t > \theta(\boldsymbol{x})\right)\right\} d\mathbf{s},$$

$$\begin{aligned} Z_{t,i}(\boldsymbol{x}, \theta) - \bar{Z}_{t,i}(\boldsymbol{x}, \theta) = {} & \frac{1}{1-p} \int \left\{\prod_{j=1}^{d} k(s_j)\right\} \left\{I\left(\boldsymbol{x}^T \boldsymbol{X}_t + h\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})\right) s_i h \right. \\ & \left. + X_{t,i} I\left(\boldsymbol{x}^T \boldsymbol{X}_t + h\boldsymbol{x}^T \boldsymbol{s} > \theta(\boldsymbol{x})\right) - X_{t,i} I\left(\boldsymbol{x}^T \boldsymbol{X}_t > \theta(\boldsymbol{x})\right)\right\} d\mathbf{s} \end{aligned}$$

for $i = 1, \cdots, d$. Then,

$$E\{Z_{t,i}(\boldsymbol{x}, \theta) - \bar{Z}_{t,i}(\boldsymbol{x}, \theta)\} = \mathcal{O}(h^2), \ E\{Z_{t,i}(\boldsymbol{x}, \theta) - \bar{Z}_{t,i}(\boldsymbol{x}, \theta)\}^2 = \mathcal{O}(h^2) \tag{EC.9}$$

hold uniformly in $\left\{\left(\boldsymbol{x}^T, \theta\right)^T : ||\boldsymbol{x} - \boldsymbol{x}_0|| + |\theta - \theta_0| \le n^{-1/3}\right\}$ for all $i \in \{1, \dots, d+1\}$, implying that

$$|\gamma_i(0; \boldsymbol{x}, \theta)| = \mathcal{O}(h^2) \tag{EC.10}$$

uniformly in $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$ for all $i \in \{1,\ldots,d+1\}$. Here, $\mathcal{O}(h^2)$ means less than a constant times h^2. Using C2), we have that for any $r \ge 1$,

$$E\left\{\int I(\boldsymbol{x}^T\boldsymbol{s} > \theta(\boldsymbol{x}))\prod_{j=1}^d h^{-1}k\left(\frac{X_{t,j}-s_j}{h}\right)d\mathbf{s} \times \int I(\boldsymbol{x}^T\boldsymbol{s} > \theta(\boldsymbol{x}))\prod_{j=1}^d h^{-1}k\left(\frac{X_{t+r,j}-s_j}{h}\right)d\mathbf{s}\right\}$$
$$= \int I(\boldsymbol{x}^T\boldsymbol{s} > \theta(\boldsymbol{x}))I(\boldsymbol{x}^T\bar{\boldsymbol{s}} > \theta(\boldsymbol{x}))\left\{\prod_{j=1}^d h^{-1}k\left(\frac{y_j-s_j}{h}\right)\right\}\left\{\prod_{j=1}^d h^{-1}k(\frac{\bar{y}_j-\bar{s}_j}{h})\right\}$$
$$\times f_{1,1+r}(\boldsymbol{y},\bar{\boldsymbol{y}})\,d\mathbf{s}\,d\bar{\mathbf{s}}\,d\mathbf{y}\,d\bar{\mathbf{y}}$$
$$= \int I(\boldsymbol{x}^T\boldsymbol{s} > \theta(\boldsymbol{x}))I(\boldsymbol{x}^T\bar{\boldsymbol{s}} > \theta(\boldsymbol{x}))\left\{\prod_{j=1}^d k(y_j)\right\}\left\{\prod_{j=1}^d k(\bar{y}_j)\right\}f_{1,1+r}(\boldsymbol{s}+h\boldsymbol{y},\bar{\boldsymbol{s}}+h\bar{\boldsymbol{y}})\,d\mathbf{y}\,d\bar{\mathbf{y}}\,d\mathbf{s}\,d\bar{\mathbf{s}}$$
$$= \int I(\boldsymbol{x}^T\boldsymbol{s} > \theta(\boldsymbol{x}))I(\boldsymbol{x}^T\bar{\boldsymbol{s}} > \theta(\boldsymbol{x}))\left\{\prod_{j=1}^d k(y_j)\right\}\left\{\prod_{j=1}^d k(\bar{y}_j)\right\}$$
$$\times\left\{f_{1,1+r}(\boldsymbol{s},\bar{\boldsymbol{s}}) + h\sum_{j=1}^d \frac{\partial}{\partial s_j}f_{1,1+r}(\boldsymbol{s},\bar{\boldsymbol{s}})y_j + h\sum_{j=1}^d \frac{\partial}{\partial \bar{s}_j}f_{1,1+r}(\boldsymbol{s},\bar{\boldsymbol{s}})\bar{y}_j + \mathcal{O}(h^2)\right\}d\mathbf{y}\,d\bar{\mathbf{y}}\,d\mathbf{s}\,d\bar{\mathbf{s}}$$
$$= \int I(\boldsymbol{x}^T\boldsymbol{s} > \theta(\boldsymbol{x}))I(\boldsymbol{x}^T\bar{\boldsymbol{s}} > \theta(\boldsymbol{x}))\left\{\prod_{j=1}^d k(y_j)\right\}\left\{\prod_{j=1}^d k(\bar{y}_j)\right\}f_{1,1+r}(\boldsymbol{s},\bar{\boldsymbol{s}})\,d\mathbf{s}\,d\bar{\mathbf{s}} + \mathcal{O}(h^2)$$

holds uniformly in $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$. Similarly, we can show that

$$|\gamma_i(s;\boldsymbol{x},\theta)| = \mathcal{O}(h^2) \tag{EC.11}$$

holds uniformly in positive integer s and $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$ for all $i \in \{1,\cdots,d+1\}$. Using C1) and the Davydov inequality, we have

$$|\gamma_i(s;\boldsymbol{x},\theta)| = \mathcal{O}(\{\alpha(s)\}^{1-2/(2+\delta)}) \tag{EC.12}$$

uniformly in nonnegative integer s and $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$ for all $i \in \{1,\cdots,d+1\}$. Hence, it follows from (EC.10), (EC.11), and (EC.12) that for any given $\xi \in (1/2,1)$,

$$|\gamma_i(s;\boldsymbol{x},\theta)| = \mathcal{O}(h^{2\xi}\{\alpha(s)\}^{1-\xi-2(1-\xi)/(2+\delta)}) \tag{EC.13}$$

uniformly in nonnegative integer s and $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$ for all $i \in \{1,\cdots,d+1\}$. It follows from (EC.9), (EC.10), (EC.13), and C1) that

$$E\left\{\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^n\left(Z_{t,i}(\boldsymbol{x},\theta) - \bar{Z}_{t,i}(\boldsymbol{x},\theta)\right)\right\}^2$$
$$= \gamma_i(0;\boldsymbol{x},\theta) + 2\textstyle\sum_{m=1}^{n-1}(1-m/n)\gamma_i(m;\boldsymbol{x},\theta) + n\left\{E\left(Z_{1,i}(\boldsymbol{x},\theta) - \bar{Z}_{1,i}(\boldsymbol{x},\theta)\right)\right\}^2$$
$$= \mathcal{O}(h^2) + h^{2\xi}\mathcal{O}\left(\textstyle\sum_{m=1}^{n-1}\{\alpha(m)\}^{1-\xi-2(1-\xi)/(2+\delta)}\right) + \mathcal{O}(nh^4)$$
$$= o(1)$$

uniformly in $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$ for all $i=1,\cdots,d+1$, implying that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\{\boldsymbol{Z}_t(\boldsymbol{x},\theta)-\bar{\boldsymbol{Z}}_t(\boldsymbol{x},\theta)\right\}=o_p(1) \quad \text{as } n\to\infty \tag{EC.14}$$

uniformly in $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$.

For any constant $\boldsymbol{\lambda}\in\Re_{d+1}\backslash\{\mathbf{0}\}$, it follows from C1) that $\{\boldsymbol{\lambda}^T\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)\}$ is a strictly stationary α-mixing sequence with $\alpha_{\boldsymbol{\lambda}^T\bar{\boldsymbol{Z}}}(m)=\mathcal{O}(a^m)$ as $m\to\infty$. Hence, using the Central Limit Theorem for α-mixing sequence (e.g., see Rosenblatt (1956)), (17) in Theorem 3 holds and

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\boldsymbol{\lambda}^T\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)\xrightarrow{w}N\left(0,\boldsymbol{\lambda}^T\bar{\Sigma}\boldsymbol{\lambda}\right) \quad \text{as } n\to\infty$$

Using the Cramér-Wold device, we have that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)\xrightarrow{w}N(\mathbf{0},\bar{\Sigma}) \quad \text{as } n\to\infty. \tag{EC.15}$$

Decomposing

$$\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n}\left\{\bar{\boldsymbol{Z}}_t(\boldsymbol{x},\theta)-\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)-\Gamma(\boldsymbol{x},\theta)+\Gamma(\boldsymbol{x}_0,\theta_0)\right\}\\
&\quad=\frac{1}{n}\sum_{t=1}^{n}\left\{\bar{\boldsymbol{Z}}_t(\boldsymbol{x},\theta)-\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta)-\Gamma(\boldsymbol{x},\theta)+\Gamma(\boldsymbol{x}_0,\theta)\right\}\\
&\qquad+\frac{1}{n}\sum_{t=1}^{n}\left\{\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta)-\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)-\Gamma(\boldsymbol{x}_0,\theta)+\Gamma(\boldsymbol{x}_0,\theta_0)\right\}\\
&\quad:=I_1+I_2,
\end{aligned}$$

similar to the proofs of Lemmas 1 and 2 in Chen (2008), one can show that

$$I_1=o_p\left(||\boldsymbol{x}-\boldsymbol{x}_0||+|\theta-\theta_0|\right) \quad \text{and} \quad I_2=o_p\left(||\boldsymbol{x}-\boldsymbol{x}_0||+|\theta-\theta_0|\right) \quad \text{as } n\to\infty$$

uniformly in $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$. That is,

$$\frac{1}{n}\sum_{t=1}^{n}\left\{\bar{\boldsymbol{Z}}_t(\boldsymbol{x},\theta)-\bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)-\Gamma(\boldsymbol{x},\theta)+\Gamma(\boldsymbol{x}_0,\theta_0)\right\}=o_p(||\boldsymbol{x}-\boldsymbol{x}_0||+|\theta-\theta_0|) \quad \text{as } n\to\infty \tag{EC.16}$$

uniformly in $\{(\boldsymbol{x}^T,\theta)^T : ||\boldsymbol{x}-\boldsymbol{x}_0|| + |\theta-\theta_0| \le n^{-1/3}\}$. Therefore, it follows from (EC.8)–(EC.16) that

$$\begin{aligned}
\mathbf{0} &= \tfrac{1}{\sqrt{n}}\sum_{t=1}^{n} \boldsymbol{Z}_t(\hat{\boldsymbol{x}}_{cvar},\hat{\theta}_{cvar}) \\
&= \tfrac{1}{\sqrt{n}}\sum_{t=1}^{n} \bar{\boldsymbol{Z}}_t(\hat{\boldsymbol{x}}_{cvar},\hat{\theta}_{cvar}) + o_p(1) \\
&= \tfrac{1}{\sqrt{n}}\sum_{t=1}^{n} \bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0) + \sqrt{n}\frac{1}{n}\sum_{t=1}^{n}\big\{\bar{\boldsymbol{Z}}_t(\hat{\boldsymbol{x}}_{cvar},\hat{\theta}_{cvar}) - \bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0)\big\} + o_p(1) \\
&= \tfrac{1}{\sqrt{n}}\sum_{t=1}^{n} \bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0) + \sqrt{n}\big\{\bar{\Gamma}(\hat{\boldsymbol{x}}_{cvar},\hat{\theta}_{cvar}) - \bar{\Gamma}(\boldsymbol{x}_0,\theta_0)\big\} + o_p(1) \\
&= \tfrac{1}{\sqrt{n}}\sum_{t=1}^{n} \bar{\boldsymbol{Z}}_t(\boldsymbol{x}_0,\theta_0) + \sqrt{n}\dot{\bar{\Gamma}}(\boldsymbol{x}_0,\theta_0)\big(\hat{\boldsymbol{x}}_{cvar}^T - \boldsymbol{x}_0^T, \hat{\theta}_{cvar}-\theta_0\big)^T + o_p(1),
\end{aligned}$$

which implies (18). Equation (19) follows from (18) and the fact that

$$\sqrt{n}(\hat{\boldsymbol{\alpha}}_{cvar}-\boldsymbol{\alpha}_0) = \sqrt{n}\frac{\hat{\boldsymbol{x}}_{cvar}-\boldsymbol{x}_0}{\mathbf{1}^T\boldsymbol{x}_0} - \frac{\boldsymbol{x}_0}{(\mathbf{1}^T\boldsymbol{x}_\mathbf{0})^2}\mathbf{1}^T\sqrt{n}(\hat{\boldsymbol{x}}_{cvar}-\boldsymbol{x}_0) + o_p(1).$$

EC.1.4. Proof of Theorem 4

It follows from the same arguments after (EC.14) in the proof of Theorem 3, and thus, no specific derivations are further required.

EC.2. Further Discussions about Corollary 1

The idealized elliptical assumption in Corollary 1 is a good illustration of the inherit properties of RP portfolios. Two examples are provided in the next section and their implementations are outlined in Appendix EC.2.2.

EC.2.1. Examples for Corollary 1

Example EC.1. Let $\mathbf{X} = (X_1, X_2, X_3)$ be a trivariate normally distributed random vector with correlation coefficients $corr(X_1, X_2) = 0.75$, $corr(X_1, X_3) = 0.5$ and $corr(X_2, X_3) = \phi$. We consider $\phi \in [-0.1978, 0.9478]$ so that the correlation matrix is positive definite. The standard deviations and expected returns are assumed to be as follows:

i) $SD(X_1) = SD(X_2) = SD(X_3) = 1$ and $\mathbb{E}(X_1) = \mathbb{E}(X_2) = \mathbb{E}(X_3) = 1$;

ii) $SD(X_1) = SD(X_2) = 1$, $\mathbb{E}(X_1) = \mathbb{E}(X_2) = 1$, $SD(X_3) = 2$ and $\mathbb{E}(X_3) = 2$.

We then compute (for details, see Appendix EC.2.2) the RP-SD and RP-CVaR$_{95\%}$ portfolios in Figure EC.1 for various values of ϕ.

First, we observe that RP portfolios are not too sensitive to the correlation matrix even if RP targets are based on SD and CVaR$_{95\%}$, which are non-tail and tail risk measures, respectively. Thus, RP investment strategies might be more robust to out-of-sample portfolio performance.

Second, we note in the left plots that equal variances does not imply RP portfolios to be equally weighted, even though the individual risk positions are identical. This occurs in spite of the correlation matrix having a wide range of values.

Third, the left plots show that the correlation matrix has a slightly more impact on RP-SD portfolios than to RP-CVaR$_{95\%}$ portfolios, which is a consequence of how risk preferences are ordered for non-tail risk measures (SD) vs tail risk measures (CVaR$_{95\%}$); this is also explained by a weak strength of dependence of the Gaussian dependence that implies less variability for tail risk measures. Further, changing the individual risk position has a similar effect on RP-SD and RP-CVaR$_{95\%}$ portfolios; e.g., top plots in Figure EC.1 show that the riskiest asset (Asset 3) tends to have a lower proportion in RP-SD in the right-hand-side, and in turn, more is invested in Asset 2 than in Asset 1, which shows that RP portfolios have also a diversification effect besides its equal risk contribution mechanism.

Note that RP-SD and RP-CVaR$_{95\%}$ are not identical in Example EC.2, which does not contradict Corollary 1 a) since SD/var are shift invariant risk measures while CVaR$_{95\%}$ does not have such a property. If we would have replaced CVaR$_{95\%}$ by CVaR$_{95\%} - \mathbb{E}$, then the two RP portfolios would have been identical since Corollary 1 a) holds in this case.

Recall that Corollary 1 does not hold if $\mathbf{X}$ is not elliptically distributed. This could be seen in Example EC.2 below, which shows that one would need to be cautious if parametric assumptions are imposed in RB/RP modeling.

Example EC.2. Assume that $d = 2$ and $X_2 = X_1^3$ almost surely and $X_1 \sim N(0,1)$, i.e. X_1 and X_2 are *comonotonic*; by definition, X_1 and X_2 are *comonotonic* if there exists a non-decreasing function f such that $\Pr(X_2 = f(X_1)) = 1$. By construction, (X_1, X_2) is not elliptically distributed.

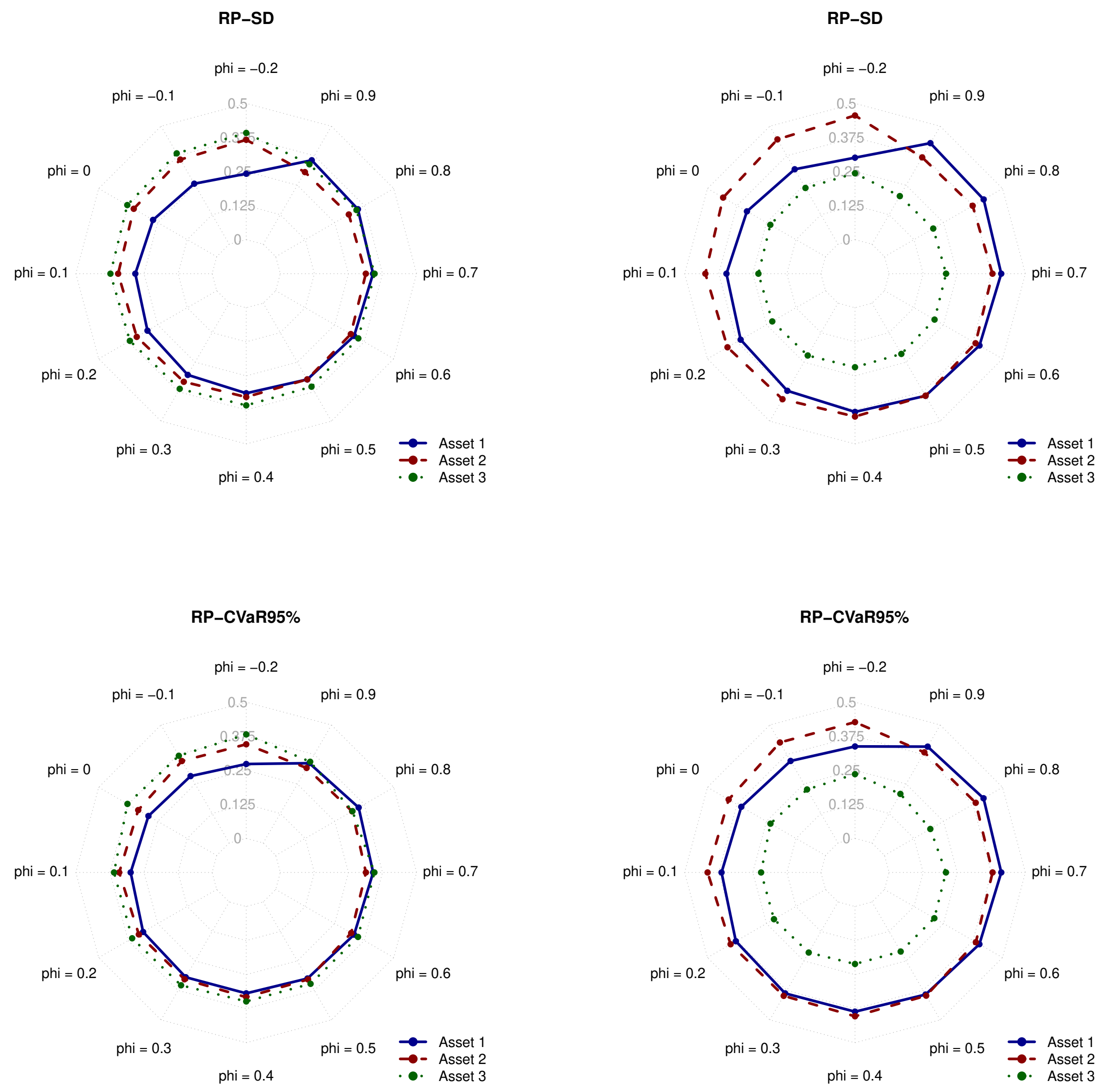

Figure EC.1 Radar plots for RP-SD (top) and RP-CVaR$_{95\%}$ (bottom) for settings i) (left) and setting ii) (right) as defined in Example EC.1 for various values of ϕ.

If the budgeting target vector is $\mathbf{b} = (b, 1-b)$ with $0 < b < 1$, then the var/SD RB portfolio, denoted as $\boldsymbol{\alpha}^{*SD}(\mathbf{b})$, is the solution of $x_1^2 + 3x_1x_2 = b\big(x_1^2 + 6x_1x_2 + 15x_2^2\big)$ such that $\mathbf{x} \in \Delta_2 \in \Re^2_{++}$, since $\mathrm{Cov}(X_1, X_2) = 3$ and $var(X_2) = 15$. Thus,

$$\alpha_1^{*SD}(\mathbf{b}) = \frac{24b + 3 - \sqrt{-24b^2 + 24b + 9}}{20b + 4} \quad \text{and} \quad \alpha_2^{*SD}(\mathbf{b}) = 1 - \alpha_1^{*SD}(\mathbf{b}), \quad \text{for any } 0 < b < 1.$$

Particularly, the RP-SD portfolio is achieved with $\alpha_1^{*SD}(1/2, 1/2) = 0.7948$.

Recall that $\varphi \in \{\text{VaR}, \text{CVaR}\}$ are comonotonic additive risk measures, i.e. $\varphi(X_1+X_2)=\varphi(X_1)+\varphi(X_2)$ for any comonotonic risks X_1 and X_2; for details, see Dhaene et al. (2006). Thus, the RB-CVaR$_{95\%}$ portfolio, denoted as $\boldsymbol{\alpha}^{*\text{CVaR}_{95\%}}(\mathbf{b})$, is the solution of

$$x_1\text{CVaR}_{95\%}(X_1)=b\Big(x_1\text{CVaR}_{95\%}(X_1)+x_2\text{CVaR}_{95\%}(X_2)\Big) \quad \text{such that} \quad \mathbf{x}\in\Delta_2\in\Re^2_{++},$$

since CVaR is a comonotonic additive risk measure. Particularly, the RP-CVaR$_{95\%}$ portfolio is achieved with $\alpha_1^{*\text{CVaR}_{95\%}}(1/2,1/2)=0.8247$, since $\text{CVaR}_{95\%}(X_1)=\psi\big(\text{VaR}_{95\%}(X_1)\big)$ and

$$\text{CVaR}_{95\%}(X_2)=\frac{1}{0.05}\int_{0.95}^{1}\big(\text{VaR}_s(X_1)\big)^3\,ds=\psi\big(\text{VaR}_{95\%}(X_1)\big)\Big(2+\big(\text{VaR}_{95\%}(X_1)\big)^2\Big),$$

where $\psi(\cdot):=\frac{1}{0.05\sqrt{2\pi}}e^{-\frac{\cdot^2}{2}}$ on $\Re$. Similarly, one may find that the RP-VaR$_{95\%}$ portfolio, denoted as $\boldsymbol{\alpha}^{*\text{VaR}_{95\%}}(1/2,1/2)$, is achieved with $\alpha_1^{*\text{VaR}_{95\%}}(1/2,1/2)=0.7301$.

Since $\mathbb{E}[X_1]=\mathbb{E}[X_2]=0$, then $\boldsymbol{\alpha}^{*\text{CVaR}_{95\%}}(1/2,1/2)$ and $\boldsymbol{\alpha}^{*\text{VaR}_{95\%}}(1/2,1/2)$ are RP portfolios if the risk preferences are also ordered by $\text{CVaR}_{95\%}-\mathbb{E}$ and $\text{VaR}_{95\%}-\mathbb{E}$, respectively. Note that X_2 is a non-linear function of X_1, and for this reason, SD does not capture the perfect dependence between the two assets ($corr(X_1,X_2)=0.7746$) though the RP-SD weights are closer to those of RP-CVaR$_{95\%}$ than RP-VaR$_{95\%}$. Further, CVaR$_{95\%}$ is tail sensitive, and the strong associations between the assets (due to comonotonic (X_1,X_2)) is better captured by CVaR$_{95\%}$ (than VaR$_{95\%}$) that invests 82.47% (instead of 73.01%) in the less risky X_1; similarly, RP-SD balances the risk better than RP-VaR$_{95\%}$ by partially capturing the strong associations as $corr(X_1,X_2)=0.7746$.

These show how different RP-SD is from the RP portfolios based on $\text{CVaR}_{95\%}-\mathbb{E}$ and $\text{VaR}_{95\%}-\mathbb{E}$ (which are two shift invariant risk measures) and thus, Corollary 1 a) does not hold since the elliptical assumption is not valid.

EC.2.2. Implementations for Example EC.1

The implementations for RP-SD and RP-CVaR$_{95\%}$ computations in Example EC.1 are now provided. These implementations rely on the formulation in (8).

Solving for RP-SD portfolios requires first finding

$$\mathbf{y}^* = \underset{\mathbf{y}\in\Re^d_{++}}{\arg\min} \quad \mathbf{y}^T\Sigma\mathbf{y} - \sum_{k=1}^{d}\log y_k, \tag{EC.17}$$

and in turn, the RP-SD portfolio is $\mathbf{x}^* = \mathbf{y}^*/\mathbf{1}^T\mathbf{y}^*$.

By keeping (4) in mind, solving for RP-CVaR$_{95\%}$ portfolios requires first finding

$$\mathbf{y}^{**} = \underset{\mathbf{y}\in\Re^d_{++}}{\arg\min} \quad \delta\sqrt{\mathbf{y}^T\Sigma\mathbf{y}} + \sum_{k=1}^{d} y_k E[X_k] - \sum_{k=1}^{d}\log y_k, \tag{EC.18}$$

where $\delta = \text{CVaR}_{95\%}(Z_1) = 2.06271$ and Z_1 being a standard normal distribution. Then, the RP-CVaR$_{95\%}$ portfolio is $\mathbf{x}^{**} = \mathbf{y}^{**}/\mathbf{1}^T\mathbf{y}^{**}$.

Note that (EC.17) and (EC.18) are convex instances that could be solved by any general purpose convex solver, though both instances are SOCP representable by recasting $-\sum_{k=1}^{d}\log y_k$ as a set of hyperbolic constraints; e.g., see Mausser and Romanko (2018). Note that we employ MATLAB (Optimization Toolbox) Quadratic programming ‘quadprog’ for solving (EC.17) and SOCP solver ‘coneprog’ for solving (EC.18).

EC.3. Technical Details regarding Note 2 ii)

An extended discussion to Note 2 ii) is now provided. Recall that Theorem 1 is based on condition (7), and its rationale is explained in Note 2 ii), where it is highlighted that the optimal solutions may be on the boundary of the feasibility set (that is, $(0,\infty)^{d-d_0} \times \infty^{d_0}$ with $0 < d_0 \le d$) with a $-\infty$ optimal objective value. Clearly, (7) ensures that the latter does not happen, which is pointed out in Bellini et al. (2021) as well, though a slightly different equivalent condition to (7) is considered. Now, if (7) is not satisfied then one may approximate RB solutions, which is the main aim of this discussion. The technical details are provided in Theorem EC.1, which is an extension of Theorem 1, and its proof is skipped because the proofs are quite similar.

THEOREM EC.1. *Let $\boldsymbol{b} \in \Delta_d \cap \Re^d_{++}$, and φ be a convex, homogeneous of order $\tau \ge 1$ risk measure. Further, assume that there exists $M < 0$ such that*

$$\min_{\boldsymbol{0}<\boldsymbol{x}\le\frac{1}{d}\boldsymbol{1}} \mathcal{R}(\boldsymbol{x}) - M\sum_{k=1}^{d} x_k^\tau > 0. \tag{EC.19}$$

a) For any given $\lambda > 0$*, the following problem*

$$\min_{\boldsymbol{x}\in\Re^d_{++}} \frac{1}{\tau}\left(\mathcal{R}(\boldsymbol{x}) - M\sum_{k=1}^{d} x_k^{\tau}\right) - \lambda\sum_{k=1}^{d} b_k \log x_k \tag{EC.20}$$

admits a unique solution, denoted as $\boldsymbol{x}^*(\lambda, \boldsymbol{b}, M)$*, that is an interior point of* $\Re^d_{++}$*. If* $\mathcal{R}(\boldsymbol{\alpha})$ *is differentiable at* $\boldsymbol{x}^*(1, M)$*, then* $\boldsymbol{\alpha}^*(\boldsymbol{b}, M) = \boldsymbol{x}^*(\lambda, \boldsymbol{b}, M)/\boldsymbol{1}^T\boldsymbol{x}^*(\lambda, \boldsymbol{b}, M)$ *satisfies*

$$\mathcal{RC}_k(\boldsymbol{\alpha}) = b_k\varphi\left(\boldsymbol{\alpha}^T\boldsymbol{X}\right) + \epsilon_k(M), \quad \epsilon_k(M) := M\left(\left(\alpha_k^*(M)\right)^{\tau} - b_k\sum_{l=1}^{d}\left(\alpha_l^*(M)\right)^{\tau}\right) \tag{EC.21}$$

for all $k \in \{1, 2, \dots, d\}$*, where* $\mathcal{RC}_k(\boldsymbol{\alpha})$ *is given in* (1)*. Moreover,*

$$\boldsymbol{\alpha}^*(\boldsymbol{b}, M) = \boldsymbol{x}^*\left(\lambda^*(M), \boldsymbol{b}, M\right) = \left(\lambda^*(M)\right)^{1/\tau}\boldsymbol{x}^*(1, \boldsymbol{b}, M), \quad \text{where } \lambda^*(M) = \left(\boldsymbol{1}^T\boldsymbol{x}^*(1, \boldsymbol{b}, M)\right)^{-\tau}.$$

b) For any given $c \in \Re$*, the following problem*

$$\min_{\boldsymbol{x}\in\Re^d_{++}} \mathcal{R}(\boldsymbol{x}) - M\sum_{k=1}^{d} x_k^{\tau} \quad \text{such that} \sum_{k=1}^{d} b_k \log x_k \geq c \quad \text{with} \quad c \in \Re \tag{EC.22}$$

admits a unique solution, denoted as $\boldsymbol{x}^{**}(c, \boldsymbol{b}, M)$*, that is an interior point of the feasibility set. If* $\mathcal{R}(\boldsymbol{\alpha})$ *is differentiable at* $\boldsymbol{x}^{**}(1, \boldsymbol{b}, M)$*, then* $\boldsymbol{\alpha}^{**}(\boldsymbol{b}, M)$ *satisfies* (EC.21)*, where* $\boldsymbol{\alpha}^{**}(\boldsymbol{b}, M) = \boldsymbol{x}^{**}(c, \boldsymbol{b}, M)/\boldsymbol{1}^T\boldsymbol{x}^{**}(c, \boldsymbol{b}, M)$*. Moreover,*

$$\boldsymbol{\alpha}^{**}(\boldsymbol{b}, M) = \boldsymbol{x}^{**}\left(c^*(M), \boldsymbol{b}, M\right) = e^{c^*(M)-1}\boldsymbol{x}^{**}(1, \boldsymbol{b}, M), \text{ where } c^*(M) = 1 - \log\left(\boldsymbol{1}^T\boldsymbol{x}^{**}(1, \boldsymbol{b}, M)\right).$$

Furthermore, strong duality holds in (EC.22)*.*

c) For any $\boldsymbol{b}$*, we have that* $\boldsymbol{\alpha}^*(\boldsymbol{b}, M) = \boldsymbol{\alpha}^{**}(\boldsymbol{b}, M)$ *and*

$$\min_{\boldsymbol{x}\in\Delta_d\cap\Re^d_{++}} \mathcal{R}(\boldsymbol{x}) \leq \mathcal{R}\big(\boldsymbol{\alpha}^*(\boldsymbol{b}, M)\big) \leq \mathcal{R}(\boldsymbol{b}) + M\left(\sum_{k=1}^{d}\left(\alpha_k^*(M)\right)^{\tau} - \sum_{k=1}^{d} b_k^{\tau}\right). \tag{EC.23}$$

If $\tau = 1$*, then* $\mathcal{R}\big(\boldsymbol{\alpha}^*(\boldsymbol{b}, M)\big) \leq \mathcal{R}\left(\frac{1}{d}\boldsymbol{1}\right)$ *for any* $\boldsymbol{b}$*.*

It should be noted that $\sum_{k=1}^{d}\epsilon_k(M) = 0$ for any M and τ, which justifies why $\boldsymbol{\alpha}^*(\mathbf{b}, M)$ and $\boldsymbol{\alpha}^{**}(\mathbf{b}, M)$ are RB approximations. As expected, Theorem EC.1 recovers the results in Theorem 1 if $M = 0$. In fact, most of the RB properties shown in Theorem 1 are shared by the RB approximations in Theorem EC.1. Note that the right-hand side of (EC.23) is the same as its counterpart in (10) for any M when $\tau = 1$, since $\sum_{k=1}^{d}\alpha_k^*(M) - \sum_{k=1}^{d} b_k = 0$.

One may wonder how to find the "'best" possible M, for which a practical solution is available. Clearly, if there exists $M_0 < 0$ such that (EC.19) is satisfied, then (EC.19) holds for any $M < M_0$. At the same time, there is no theoretical justification to conclude that $\big(\epsilon_1(M), \dots, \epsilon_d(M)\big)$ becomes closer (in any distance choice) to $\mathbf{0}$ if M takes smaller values, though this is what would be expected. Numerical solutions would be needed to find the "'best" possible penalty parameter M, and a small numerical example is provided to illustrate our point.

Example EC.3. Let (X_1, X_2) be a synthetic two-asset portfolio such that each asset may take one of the $m = 23$ values from $\mathbf{s} = \{-20, -19, \dots, 0, 1, 2\}$ with probabilities

$$\Pr\big(X_1 = s_{i_1}, X_2 = s_{i_1}\big) = m^{-2}, \quad \text{for all } 1 \le i_1 \le m$$

$$\Pr\big(X_1 = s_{i_1}, X_2 = s_{i_2}\big) = m^{-2} - m^{-2.2}, \quad \text{for all } 1 \le i_1 < i_2 \le m$$

$$\Pr\big(X_1 = s_{i_1}, X_2 = s_{i_2}\big) = m^{-2} + m^{-2.2}, \quad \text{for all } 1 \le i_2 < i_1 \le m.$$

Thus, there are $n = 23^2 = 529$ states of the world. We apply a SOCP formulation similar to the one in Mausser and Romanko (2018) to evaluate RP-CVaR$^{iid}_{0.9}$ portfolios based on Theorem EC.1, i.e., (EC.22) with $c = 0$. Our results are showcased in Figure EC.2.

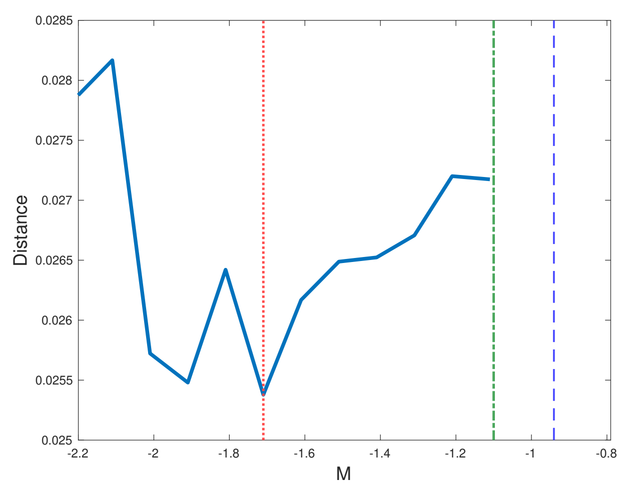

Figure EC.2 The l_1 distance $|\epsilon_1(M)| + |\epsilon_2(M)|$ of SOCP solutions of (EC.22) (as explained in Example EC.3) for various values of M.

Note that condition (7) is not satisfied, since the minimal objective value of that *Linear Programming (LP)* instance is -0.94, which is signified by the blue dashed vertical line in Figure EC.2. Some simple LP implementations show that the largest negative M such that condition (EC.19) holds is achieved when $M_0^* = -1.10$, which is signified by the green dashed vertical line in Figure EC.2. We then evaluate the SOCP solutions of (EC.22) for various values of $M \leq M_0^*$, and display the l_1 distance $|\epsilon_1(M)| + |\epsilon_2(M)|$ in Figure EC.2. We empirically find that $M^* = -1.71$ is the "best" possible choice for M, which is signified by the red dotted vertical line in Figure EC.2. This is only a simple example for the rule of thumb we recommend using in the extreme case where (7) is not satisfied, and solving of either (8) or (9) may lead to optimal solutions in the neighbourhood of infinity. The latter is not the case in this simple example, and our example could be viewed as a proof of concept for Theorem EC.1.

EC.4. Simulation Study

A simulation study for Theorem 3 is provided in this section by assuming first the case of i.i.d. data (in Figure EC.3) and then considering strictly stationary α-mixing data (in Figure EC.4). The main aim is to compare the performance of the two $\text{CVaR}_{95\%}$ estimators, $\widehat{\mathcal{R}}_{cvar}$ and $\widehat{\mathcal{R}}_{cvar}^{emp}$. Figure EC.3 compares the weights of assets between the two portfolios and EC.4 focuses on the portfolio risks.

We first generate $m = 100$ i.i.d. samples of stock returns of size $n = 500$ from a d-dimensional Gaussian distribution, and also from a multivariate t-distribution with 3 degrees of freedom, with zero mean vectors and a randomly generated covariance matrix. The first set of estimates relies on our novel CVaR estimator ($\widehat{\mathcal{R}}_{cvar}$) to compute $RP - \text{CVaR}_{95\%}^{niid}$ portfolios. Specifically, we compute $RP - \text{CVaR}_{95\%}^{niid}$ portfolio weights, $\hat{\boldsymbol{\alpha}}^{*(s)}$ with $1 \leq s \leq m$, for each sample of stock returns of size n; the second set of estimates relies on the standard CVaR estimator ($\widehat{\mathcal{R}}_{cvar}^{emp}$) to compute $RP - \text{CVaR}_{95\%}^{iid}$ weights, $\hat{\hat{\boldsymbol{\alpha}}}^{*(s)}$ with $1 \leq s \leq m$. We then compute $l_1^{(s)} = \sum_{k=1}^{d} |\hat{\alpha}_k^{*(s)} - \hat{\hat{\alpha}}_k^{*(s)}|$, which is the l_1 distance between the non-i.i.d. and i.i.d. vector estimates of the two sets of weights.

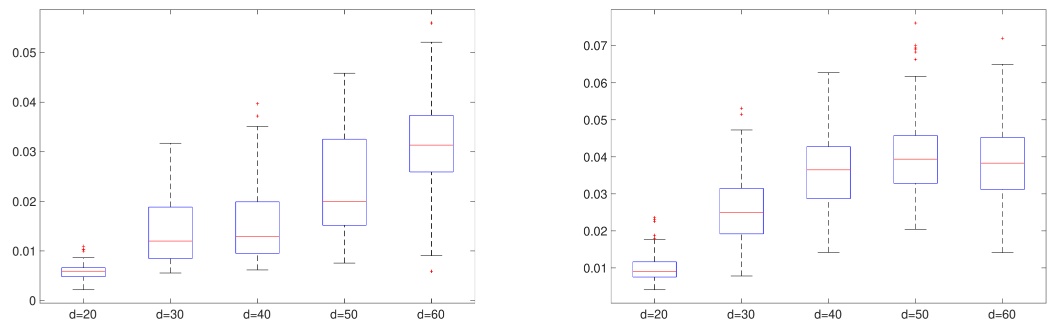

Figure EC.3 The boxplots for various number of assets d of the l_1 distances between the RP-CVaR$^{iid}_{95\%}$ and RP-CVaR$^{niid}_{95\%}$ estimates (i.e. the weights of assets in the portfolio) based on $m = 100$ samples of size $n = 500$ from the d-dimensional multivariate Gaussian distribution (left) and multivariate t-distribution with 3 degrees of freedom (right). Asset returns have zero mean vectors and randomly generated covariance matrices.

As expected, Figure EC.3 shows very few outliers for lighter tailed distribution (Gaussian returns). The l_1 distance increases as the number of assets increases, but the errors are very small. Note that the increase in the l_1 distances is reaching a plateau with the increase in the number of assets for the distribution with a heavier tail (see the right-hand side plot). This is explained by the fact that CVaR is a tail risk measure that is more sensitive in risk aggregation when the likelihood of concomitant extreme events is reduced (Asimit et al. 2011); recall that multivariate Gaussian distributions exhibit tail independence, while multivariate t-distributions exhibit tail dependence for which there is a significant likelihood of observing concomitant extreme events. This explains why the median l_1 distance increases by almost 6 times for Gaussian returns and 4 times for the t-distributed case when the number of assets increases by 3 times (from 20 to 60 assets).

In a nutshell, Figure EC.3 shows that the two estimators ($\widehat{\mathcal{R}}_{cvar}$ and $\widehat{\mathcal{R}}^{emp}_{cvar}$) are not significantly different for i.i.d. data, which could be seen by comparing the RP-CVaR$_{95\%}$ weights. We next show that the two estimators lead to similar CVaR$_{95\%}$ estimates for strictly stationary α-mixing data, though $\widehat{\mathcal{R}}_{cvar}$ is shown to be slightly more advantageous than the standard estimator ($\widehat{\mathcal{R}}^{emp}_{cvar}$). Note that we do not construct portfolios in the second part, and Figure EC.4 is aimed to better understand the difference between the two CVaR estimators by comparing them to the "true" CVaR value.

The second set of simulations is for strictly stationary α-mixing data that are non-i.i.d. data. Specifically, we generate $m = 100$ samples of size $n = 500$ from a multivariate normal distribution with zero mean vectors and certain covariance matrix so that the resulting data are strictly stationary α-mixing. That is, we assume $d = 1$ and $\mathbf{X} = (X_1, \ldots, X_n) \sim MVN(\mathbf{0}, \Sigma(\rho))$ with the $(i,j)^{th}$ entry of Σ given by $\Sigma_{ij} = \rho^{|i-j|}$ for all $1 \le i, j \le n$ and $-1 < \rho < 1$. Therefore, for each sample of size n, we estimate $\mathrm{CVaR}_{95\%}$ via $\widehat{\mathcal{R}}_{cvar}$ and $\widehat{\mathcal{R}}_{cvar}^{emp}$, denoted as $\hat{\beta}^{(s)}$ and $\hat{\hat{\beta}}^{(s)}$, respectively for all $1 \le s \le m$. We then compute the MSE ratio and the variance ratio between the two estimators as follows:

$$MSE(\hat{\hat{\beta}}||\hat{\beta}) = \frac{\sum_{s=1}^{m} \left(\hat{\hat{\beta}}^{(s)} - \beta\right)^2}{\sum_{s=1}^{m} \left(\hat{\beta}^{(s)} - \beta\right)^2} \quad \text{and} \quad var(\hat{\hat{\beta}}||\hat{\beta}) = \frac{var\left(\hat{\hat{\beta}}^{*}\right)}{var\left(\hat{\beta}^{*}\right)}, \tag{EC.24}$$

where β is the true value of $\mathrm{CVaR}_{95\%}(X) = 2.06271$ that is computed as in Example EC.1.

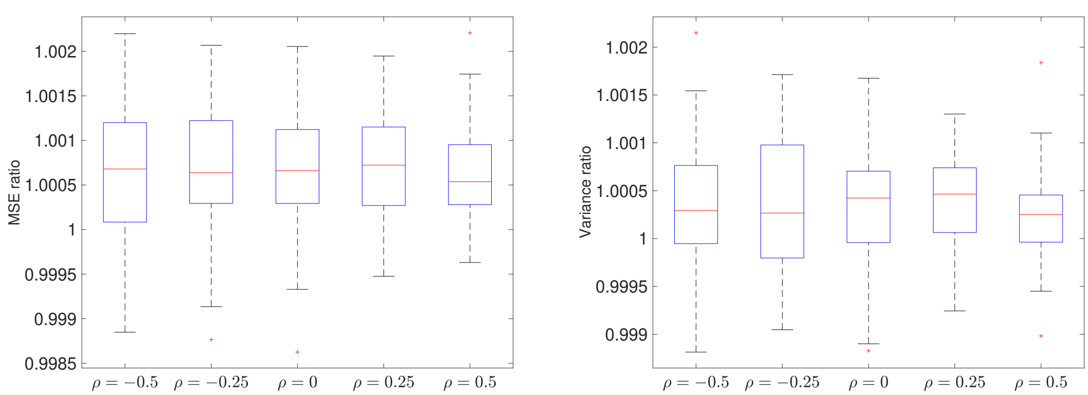

Figure EC.4 The boxplots for the MSE ratios (left) and Variance ratios (right) of portfolio risks $\mathrm{CVaR}_{95\%}$ via $\widehat{\mathcal{R}}_{cvar}$ and $\widehat{\mathcal{R}}_{cvar}^{emp}$ with the true value of $\mathrm{CVaR}_{95\%}(X) = 2.06271$ (based on 50 replications of $m = 100$ samples of size $n = 500$ that are strictly stationary α-mixing). Boxplots are shown for various serial dependence parameter values ρ.

Recall that MSE/variance ratios greater than 1 in Figure EC.4 mean that our novel estimator $\widehat{\mathcal{R}}_{cvar}$ shows better performance than the standard estimator $\widehat{\mathcal{R}}_{cvar}^{emp}$. Figure EC.4 suggests that $\widehat{\mathcal{R}}_{cvar}$ is consistently superior to $\widehat{\mathcal{R}}_{cvar}^{emp}$.

EC.5. Data and Computational Details

Here we describe the data behind the SRI factors analyzed in this paper and how we identify the timestamps of the structural breaks for the time series of stock returns over our study period. In addition, we provide some clarifications on the computational side of portfolio construction.

EC.5.1. Data Description of SRI Factors

We describe the bespoke dataset that is used across our data analysis, which is the same as the dataset in Hallerbach et al. (2004). According to our knowledge, Hallerbach et al. (2004) is the earliest reference showing how to build an investment portfolio with investment preferences ordered by a *set of attributes* that characterize the societal effects of that business.

The raw dataset came from Triodos Bank, one of the first green banks, and it contains SRI scores for a set of companies constructed from the questionnaires gathered by the SiRi research group in year 2000. In our paper, we focus on 11 SRI factor variables: Business Ethics, Community, Corporate Governance, Customer, Contract Relations, Labor Rights, Labor Care, Environment Policies and Programmes, Environment Activities, Contractors, and Sin Activities. Each SRI factor variable is constructed as an averaging index, while the aggregate score (denoted as AFS) combines all SRI factors variables based on equal weighting. The dataset consists of 590 firms that are SRI ordered by AFS, but some companies were delisted during the 20-year period. Specifically, we have an opportunity set consisting of 408 firms from 2000 to 2020 and 100 firms are based in the US. In addition to these SRI factors and rating scores, we collected historical daily stock prices for all of these companies between year 2001 and 2020, from various sources: Datastream, WRDS-CRSP, Compustat, IBES and Yahoo!Finance.

EC.5.2. Determine the Structural Breaks

We would like to determine the periods in the financial stock market represented by S&P 500 that were delineated by various important events. The structural break points are identified with a Bai-Perron test (Bai and Perron 1998, 2003). The structural breaks are time stamped in Figure EC.5, which considers the S&P 500 daily returns from 01/01/2001 until 31/12/2020.

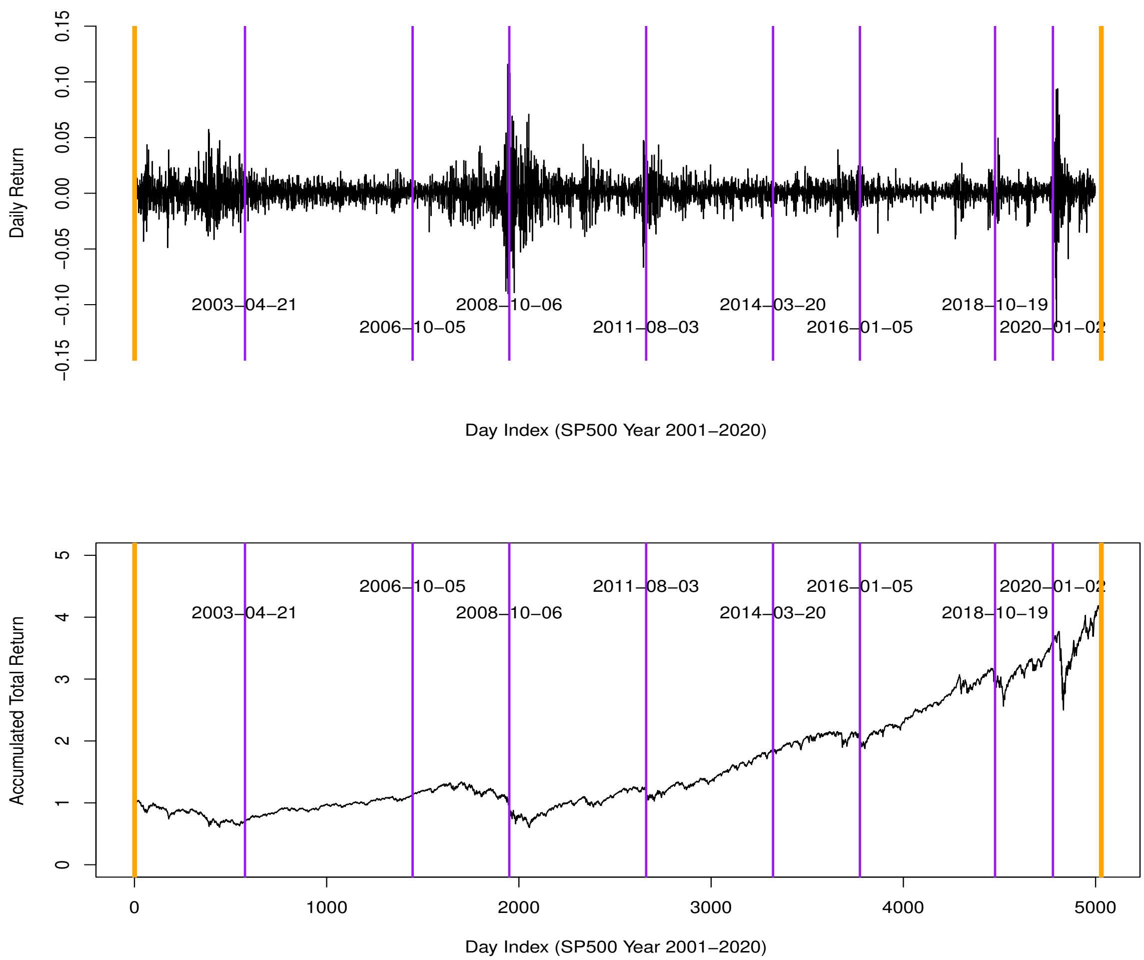

Figure EC.5 Structural break points for the S&P 500 time series based on a Bai-Perron test.

Based on the Bai-Perron test, we identify the following structural breaks in our data as illustrated in Figure EC.5: April 21, 2003, that is the end of the dot-com aftermath period; October 5, 2006 shortly after the housing prices fell by more than 6% in 20 large metropolitan areas, according to Standard & Poor's/Case-Shiller indexes; October 6, 2008, about two weeks after the Lehman Brothers collapse; August 3, 2011 when US and global stock markets crashed upon Standard and Poor's credit rating downgrade of the US sovereign debt from AAA to AA+, the first time in history the United States was downgraded; March 20, 2014 marking the Maidan revolution in Ukraine; January 5, 2016 marking the period August 2015-2016 stock market sell off when S&P 500 and DJIA dropped more than 10% twice; October 19, 2018 associated with the loss of nearly 2 trillion dollars in the U.S. stock markets leading to S&P 500 losing about 20% by the end of that year; and finally January 2, 2020 marking the beginning of the COVID-19 period.

EC.5.3. Portfolio Computations

Specific numerical methods are required for finding the RP-SD and RP-CVaR$_{95\%}$ portfolios. Recall that Theorem 1 provides two general methods called– *logarithmic barrier* and *logarithmic constraint* – that could be applied to finding RB/RP portfolios based on general risk preferences. Solving the instances from Theorem 1 with high efficiency would require using the specific properties of the optimization problems from either (8) or (9). The logarithmic barrier and logarithmic constraint formulations in (8) and (9), respectively are described in the literature (Roncalli 2013, Bellini et al. 2021). Spinu (2013) shows that RB-SD portfolios could be found via an efficient convex algorithm; alternatively, an efficient algorithm for RP-SD portfolios is suggested in Bai et al. (2016), which could be efficiently solved via the Alternating Linearisation Method with backtracking (ALM-bktr) given as Algorithm 3 in Bai et al. (2016). The specific RP-CVaR portfolios (not the general RB-CVaR portfolios) involve a hyperbolic constraint such as $c - \frac{1}{d}\sum_{k=1}^{d}\log(\alpha_k) \leq 0$, which is Second-order cone (SOC) representable; thus, the logarithmic barrier and logarithmic constraint formulations for CVaR-based portfolios can be reduced to solving an SOCP instance, that can be solved efficiently (Mausser and Romanko 2018).

Note that the multiple integrals in the $RP - \text{CVaR}_{95\%}^{niid}$ estimator are approximated via the Monte-Carlo method by generating $N = 50$ samples; further, the Epanechnikov kernel function and a bandwidth choice of $h_k = 0.2n^{-1/3}$ for all $k \in \{1, 2, \ldots, d\}$ is used for $\widehat{\mathcal{R}}_{cvar}$ estimations. We compute $RP - \mathrm{CVaR}_{95\%}^{iid}$ and $RP - \mathrm{CVaR}_{95\%}^{niid}$ portfolios by solving SOCP instances as in (Mausser and Romanko 2018), since the non-i.i.d. $\mathrm{CVaR}_{95\%}$ estimator is replaced by a linear sum via the Monte-Carlo method so that the multiple integrals are approximated.