MAD risk parity portfolios

Çağın Ararat[1] · **Francesco Cesarone[2]** · **Mustafa Çelebi Pınar[1]** ·
Jacopo Maria Ricci[2,3]

Abstract
In this paper, we investigate the features and the performance of the risk parity (RP) portfolios using the mean absolute deviation (MAD) as a risk measure. The RP model is a recent strategy for asset allocation that aims at equally sharing the global portfolio risk among all the assets of an investment universe. We discuss here some existing and new results about the properties of MAD that are useful for the RP approach. We propose several formulations for finding MAD-RP portfolios computationally, and compare them in terms of accuracy and efficiency. Furthermore, we provide extensive empirical analysis based on three real-world datasets, showing that the performances of the RP approaches generally tend to place both in terms of risk and profitability between those obtained from the minimum risk and the Equally Weighted strategies.

Keywords Mean absolute deviation · Risk parity · Portfolio optimization · Risk diversification

1 Introduction

The 2008 subprime financial crisis led many scholars and practitioners to strong criticism of classical risk-gain models and to the consequent development of new portfolio selection strategies based on the concept of risk allocation. Indeed, even though risk-gain models often have nice features in terms of formulation and of tractability, they often show several draw-

✉ Çağın Ararat
  cararat@bilkent.edu.tr

  Francesco Cesarone
  francesco.cesarone@uniroma3.it

  Mustafa Çelebi Pınar
  mustafap@bilkent.edu.tr

  Jacopo Maria Ricci
  jacopomaria.ricci@uniroma3.it

[1] Bilkent University - Department of Industrial Engineering, Ankara, Turkey

[2] Roma Tre University - Department of Business Studies, Rome, Italy

[3] Bergamo University - Department of Economics, Bergamo, Italy

backs, such as their high sensitivity to estimation errors of the input parameters (in particular, of expected returns, see, e.g., Best and Grauer 1991a, Best and Grauer 1991b, Chopra and Ziemba 1993, DeMiguel et al. 2009, and Michaud and Michaud 1998), or their lack of risk diversification. A straightforward method to deal with this issue could be the choice of the Equally Weighted (EW) portfolio, where the invested capital is equally distributed among the assets that belong to the investment universe. However, if the investment universe consists of assets with very different intrinsic risks, then the resulting portfolio has limited total risk diversification.

A more refined risk-focused method is the risk parity (RP), also called Equal Risk Contribution. It consists of selecting a portfolio where each asset equally contributes to the total portfolio risk, regardless of their estimated expected returns (Maillard & Teiletche, 2010). The RP strategy has its roots in the practitioner world (see, e.g., Fabozzi et al. 2021 and Liu et al. 2020), indeed it is often considered as a heuristic method. However, several recent theoretical and empirical findings justify the growing popularity of the RP strategy in the practice of asset allocation.

From an empirical viewpoint, risk parity portfolios show smaller sensitivity to estimation errors of the input parameters compared to portfolios based on risk minimization and on other optimization strategies (Cesarone et al., 2020a). Furthermore, risk parity portfolios seem to show promising out-of-sample performance (Fisher et al., 2015; Jacobsen & Lee, 2020).

From a theoretical viewpoint, since an RP portfolio can be found by solving a minimization problem with logarithmic constraints on the portfolio weights, the RP strategy can be interpreted as a minimum risk approach with a constraint on the minimum level of diversification, which can be seen as a sort of regularization. Furthermore, an example of the theoretical justification of the Risk Parity approach can be found in Oderda (2015), where the author shows that the analytic form of the optimal portfolio solution obtained by maximizing portfolio relative logarithmic wealth at a fixed tracking risk level with respect to the market-capitalization-weighted index, consists of a linear combination of this index, the global minimum variance portfolio, the EW portfolio, the RP portfolio, and the high cash flow rate of return portfolio.

The risk measure commonly used in the RP approach is volatility (Maillard & Teiletche, 2010; Roncalli, 2013). Other risk measures have been considered in the literature such as the Conditional Value-at-Risk (Boudt et al., 2013; Cesarone & Colucci, 2018; Mausser & Romanko, 2018) and the Expectiles (Bellini et al., 2021), both belonging to the class of coherent risk measures. However, the long-only risk parity model has a unique solution when the risk is positive, convex and positively homogeneous (Cesarone et al., 2020b), and for Conditional Value-at-Risk and Expectiles, positivity is not always guaranteed (Bellini et al., 2021; Cesarone & Colucci, 2018; Cesarone & Tardella, 2017). An alternative risk measure to volatility, which is, by definition, positive for nonconstant market returns, is the Mean Absolute Deviation (MAD) belonging to the class of deviation risk measures (Rockafellar et al., 2006). In risk-gain portfolio optimization analysis, it was introduced by Konno and Yamazaki (1991) as an alternative to the Markowitz model.

In this paper, the main goal is to investigate the theoretical properties and the performance of the MAD risk parity portfolios, providing, to the best of our knowledge, multiple contributions to the literature. First, we revisit existing theoretical results on MAD, discuss its differentiability, and provide a characterization of multivariate random market returns for which MAD is additive. We show the conditions that determine this characterization, which can be seen as strong positive dependence among the asset returns. Furthermore, under these conditions, we provide a closed-form solution for the long-only MAD-RP portfolio. Second, we establish the existence and uniqueness of the MAD-RP portfolio, thus extending the the-

oretical results of Cesarone et al. (2020b). Third, we propose several formulations to find the MAD-RP portfolios practically, thus comparing their performances both in terms of accuracy and efficiency. Finally, we provide an extensive empirical analysis on three real-world datasets by comparing the out-of-sample performance obtained from the global minimum volatility and MAD portfolios, the volatility and MAD risk parity portfolios, and the Equally Weighted portfolio.

The rest of the paper is organized as follows. Section 2 introduces the Mean Absolute Deviation risk measure, providing a discussion of MAD properties relevant to our context. In Sect. 3, we show how to formulate the RP approach with MAD mathematically and how to find the MAD-RP portfolios in practice. Section 4 provides an extensive empirical analysis based on three real-world datasets. More precisely, in Sect. 4.2.1, we test and compare all the MAD-RP formulations in terms of accuracy and efficiency, while in Sects. 4.2.2 and 4.2.3, we report a graphical comparison of some portfolio selection approaches and out-of-sample results, respectively. Finally, in Sect. 5, we draw some overall conclusions.

2 Measuring the portfolio risk by MAD

Let $n \in \mathbb{N}$. We denote by $\mathbb{R}^n$ the n-dimensional Euclidean space and define the cones $\mathbb{R}^n_+ := \{\boldsymbol{x} \in \mathbb{R}^n \mid x_i \geq 0 \text{ for each } i \in \{1, \ldots, n\}\}$, $\mathbb{R}^n_{++} := \{\boldsymbol{x} \in \mathbb{R}^n \mid x_i > 0 \text{ for each } i \in \{1, \ldots, n\}\}$. We also denote by Δ^{n-1} the set of all $\boldsymbol{x} \in \mathbb{R}^n_+$ with $\sum_{i=1}^{n} x_i = 1$.

To introduce the probabilistic setup, let us fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For an event $A \in \mathcal{F}$, we define its indicator function by $\mathbb{1}_A(\omega) = 1$ for $\omega \in A$, and by $\mathbb{1}_A(\omega) = 0$ for $\omega \in \Omega \backslash A$. We denote by $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$ the space of all $\mathcal{F}$-measurable and real-valued random variables, where two elements are distinguished up to $\mathbb{P}$-almost sure (a.s.) equality. We denote by $L^1 := L^1(\Omega, \mathcal{F}, \mathbb{P})$ the set of all $X \in L^0$ with $\mathbb{E}[|X|] < +\infty$.

The *mean absolute deviation* (MAD) of a random variable is a measure of its dispersion and is defined as the expected value of the absolute deviations from a reference value that is generally represented by the mean of the random variable:

$$\text{MAD}(X) := \mathbb{E}[|X - \mathbb{E}[X]|], \quad X \in L^1. \tag{1}$$

MAD was first used in the field of portfolio optimization by Konno and Yamazaki (1991) as a symmetric risk measure alternative to variance. An upside version of MAD is the *mean semi-absolute deviation* defined by

$$\text{MSAD}(X) := \mathbb{E}[(X - \mathbb{E}[X])^+], \quad X \in L^1, \tag{2}$$

where $a^+ := \max\{a, 0\}$ denotes the positive part of $a \in \mathbb{R}$. As it is well-known, we have $\text{MSAD}(X) = \frac{1}{2}\text{MAD}(X)$.

2.1 MAD as a deviation measure

The next proposition summarizes the properties of MAD as a *deviation measure*. We omit its straightforward proof for brevity and refer the reader to Rockafellar et al. (2006) for a general treatment of deviation measures.

Proposition 1 *The functional* $\text{MAD} : L^1 \to [0, \infty)$ *is a deviation measure, that is, it satisfies the following properties for every* $X, Y \in L^1$:

(i) *Translation invariance:* $\text{MAD}(X + \alpha) = \text{MAD}(X)$ *for every* $\alpha \in \mathbb{R}$.

(ii) Positive homogeneity: $\mathrm{MAD}(\lambda X) = \lambda \, \mathrm{MAD}(X)$ *for every* $\lambda \geq 0$.

(iii) Subadditivity: $\mathrm{MAD}(X + Y) \leq \mathrm{MAD}(X) + \mathrm{MAD}(Y)$.

(iv) Normalization: $\mathrm{MAD}(0) = 0$.

(v) Strict positivity: It holds $\mathrm{MAD}(X) > 0$ *if and only if* X *is not equal to a constant* $\mathbb{P}$-*a.s.*

2.2 Additivity of MAD

Next, we provide a property that allows for finding a closed form solution to MAD-RP portfolios, namely, a characterization of random variables for which MAD is additive. To the best of our knowledge, this basic observation is new to the current paper.

Lemma 2 *(Additivity of MAD) Let* $X, Y \in L^1$. *Then,*

$$\mathrm{MAD}(X + Y) = \mathrm{MAD}(X) + \mathrm{MAD}(Y) \tag{3}$$

if and only if

$$(X - \mathbb{E}[X])(Y - \mathbb{E}[Y]) \geq 0 \quad \mathbb{P}\text{-a.s.} \tag{4}$$

Proof See Appendix A. ◻

Just as the portfolio volatility is additive for asset returns that are perfectly correlated and the portfolio Conditional Value-at-Risk (CVaR) is additive for asset returns that are comonotone (Tasche, 2002), in view of Lemma 2, MAD is additive for random returns X, Y if and only if the deviations of X and Y from their respective means have the same sign almost surely. Hence, Condition (4) can be interpreted as a kind of positive dependence. Furthermore, we point out that in the case of additivity of homogeneous and subadditive risk measures, the sum of the risks of X and Y is an upper bound for the risk of $X + Y$. Such an upper bound represents the worst portfolio risk, and it occurs for volatility when the linear correlation between X and Y is equal to 1, for CVaR when X and Y are comonotone, and for MAD when X and Y satisfy Condition (4).

Remark 3 Clearly, if X, Y are square-integrable random variables that satisfy (4), then they are positively correlated (i.e., nonnegative covariance). Moreover, if X, Y are square-integrable random variables with perfect positive correlation, then (4) holds. Indeed, having perfect positive correlation implies that $Y = aX + b$ for some $a > 0$ and $b \in \mathbb{R}$ so that $(X - \mathbb{E}[X])(Y - \mathbb{E}[Y]) = a(X - \mathbb{E}[X])^2 \geq 0$ $\mathbb{P}$-a.s. In the next remark and example, we prove that the converse is not true in general.

Remark 4 Let $\alpha \in (0, 1)$. For a square-integrable random variable X, its α-expectile is defined as

$$e_\alpha(X) := \underset{r \in \mathbb{R}}{\mathrm{argmin}} \, \mathbb{E}[\alpha((X - r)^+)^2 + (1 - \alpha)((X - r)^-)^2].$$

In Bellini et al. (2021, Theorem 3), it is shown that the additivity of the α-expectiles at two square-integrable random variables X, Y, i.e., $e_\alpha(X + Y) = e_\alpha(X) + e_\alpha(Y)$, is equivalent to the condition

$$(X - e_\alpha(X))(Y - e_\alpha(Y)) \geq 0 \quad \mathbb{P}\text{-a.s.} \tag{5}$$

whenever $\alpha \in (\frac{1}{2}, 1)$. In Bellini et al. (2021 Example 2), it is assumed that $\mathbb{P}\{X = 0, Y = 0\} = \mathbb{P}\{X = 0, Y = 1\} = \mathbb{P}\{X = 1, Y = 0\} = \mathbb{P}\{X = 1, Y = 1\} = \frac{1}{6}$ and $\mathbb{P}\{X = 2, Y = 2\} = \frac{1}{3}$. In this case, they show that (5) holds for $\alpha = \frac{3}{4}$ and X, Y are not comonotonic. In the

same example, note that we have $\mathbb{E}[X] = \mathbb{E}[Y] = 1$, $\mathbb{E}[XY] = \frac{3}{2}$, $\mathrm{MAD}(X) = \mathrm{MAD}(Y) = \frac{2}{3}$, and $\mathrm{MAD}(X + Y) = \frac{4}{3}$ so that (4) also holds. In particular, this example shows that (4) does not imply comonotonicity. Since $\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{2}{3}$ and $\mathrm{Cov}(X, Y) = \frac{1}{2}$, X and Y are not perfectly positively correlated. Hence, (4) does not imply perfect positive correlation either. As another example, assuming that $\mathbb{P}\{X = 0, Y = 0\} = \mathbb{P}\{X = 1, Y = 1\} = \mathbb{P}\{X = 2, Y = 1\} = \frac{1}{3}$, they show that (5) does not hold for $\alpha = \frac{3}{4}$ but X, Y are comonotonic. Note that $\mathbb{E}[X] = 1$, $\mathbb{E}[Y] = \frac{2}{3}$, $\mathrm{MAD}(X) = \frac{2}{3}$, $\mathrm{MAD}(Y) = \frac{4}{9}$, and $\mathrm{MAD}(X + Y) = \frac{11}{9}$ so that (4) does not hold either. In particular, comonotonicity does not imply (4).

As shown in Remarks 3, 4, two square-integrable random variables that satisfy (4) are positively correlated but not necessarily perfectly correlated. Indeed, the correlation coefficient of such random variables can be any value in [0, 1] as the next example shows.

Example 5 Let $\mu_X, \mu_Y \in \mathbb{R}$ and $\rho \in [0, 1]$. Suppose that there exist two square-integrable positive random variables R_X, R_Y such that

$$\frac{\mathbb{E}[R_X R_Y]}{\sqrt{\mathbb{E}[R_X^2]\mathbb{E}[R_Y^2]}} = \rho.$$

Suppose also that there exists an event $A \in \mathcal{F}$ such that

$$\mathbb{E}[R_X \mathbb{1}_A] = \frac{\mathbb{E}[R_X]}{2}, \quad \mathbb{E}[R_Y \mathbb{1}_A] = \frac{\mathbb{E}[R_Y]}{2}. \tag{6}$$

(For instance, any event $A \in \mathcal{F}$ that is independent of R_X, R_Y and has $\mathbb{P}(A) = \frac{1}{2}$ satisfies this condition.) Let $S := \mathbb{1}_A - \mathbb{1}_{A^c}$ and define

$$X := \mu_X + S R_X, \quad Y := \mu_Y + S R_Y.$$

Thanks to (6), we have $\mathbb{E}[S R_X] = \mathbb{E}[S R_Y] = 0$ so that $\mathbb{E}[X] = \mu_X$, $\mathbb{E}[Y] = \mu_Y$. Hence,

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[S^2 R_X R_Y] = \mathbb{E}[R_X R_Y] = \rho\sqrt{\mathbb{E}[R_X^2]\mathbb{E}[R_Y^2]},$$

which implies that the correlation coefficient of X, Y is ρ.

The next theorem is a generalization of Lemma 2 for the multivariate setting. It provides a list of equivalent statements that can be seen as the definition of additivity for MAD when multiple random variables are considered.

Theorem 6 Let $Y_1, \ldots, Y_n \in L^1$. Then, the following are equivalent.

(i) $\mathrm{MAD}(\sum_{i=1}^n x_i Y_i) = \sum_{i=1}^n x_i \mathrm{MAD}(Y_i)$ for every $\boldsymbol{x} \in \mathbb{R}_+^n$.
(ii) $\mathrm{MAD}(\sum_{i=1}^n x_i Y_i) = \sum_{i=1}^n x_i \mathrm{MAD}(Y_i)$ for every $\boldsymbol{x} \in \Delta^{n-1}$.
(iii) $\mathrm{MAD}(\sum_{i \in I} Y_i) = \sum_{i \in I} \mathrm{MAD}(Y_i)$ for every nonempty subset $I \subseteq \{1, \ldots, n\}$.
(iv) $(Y_i - \mathbb{E}[Y_i])(Y_j - \mathbb{E}[Y_j]) \geq 0$ $\mathbb{P}$-a.s. for every $i, j \in \{1, \ldots, n\}$ with $i \neq j$.

Proof See Appendix A. □

2.3 Subdifferential of MAD

In order to introduce the marginal risk contributions of portfolios, we provide a theoretical treatment of the subdifferential of MAD as a function of the portfolio weights in this section. Suppose that there are n assets in the market. We write $\boldsymbol{r} = (r_1, \ldots, r_n)^\mathsf{T}$ for the random

vector of asset returns where we assume $r_i \in L^1$ for each $i \in \{1, \ldots, n\}$. We also write $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\mathsf{T}$ for the corresponding mean vector of $\boldsymbol{r}$, i.e., $\mathbb{E}[r_i] = \mu_i$ for each $i \in \{1, \ldots, n\}$. Furthermore, we indicate by $\boldsymbol{x} = (x_1, \ldots, x_n)^\mathsf{T} \in \mathbb{R}^n$ a vector of portfolio weights, which are the decision variables of the problems tackled in this work. For such $\boldsymbol{x}$, we have $\sum_{i=1}^n x_i = 1$ and we denote by $R(\boldsymbol{x})$ the random portfolio return. Note that we have

$$R(\boldsymbol{x}) = \boldsymbol{r}^\mathsf{T} \boldsymbol{x} = \sum_{i=1}^n r_i x_i \,.$$

Then, with some abuse of notation, the risk of the portfolio measured by MAD is given by

$$\mathrm{MAD}(\boldsymbol{x}) := \mathrm{MAD}(R(\boldsymbol{x})) = \mathbb{E}\left[\left|(\boldsymbol{r} - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{x}\right|\right] = \mathbb{E}\left[\left|\sum_{i=1}^n (r_i - \mu_i) x_i\right|\right]. \qquad (7)$$

In (7), we observe that the function $\boldsymbol{x} \mapsto \mathrm{MAD}(\boldsymbol{x})$ is generally not differentiable at some points in its domain due to the absolute value function. However, it is a finite, convex, hence also a continuous function on $\mathbb{R}^n$; thus, it has a nonempty subdifferential $\partial \mathrm{MAD}(\boldsymbol{x})$ at every given $\boldsymbol{x} \in \mathbb{R}^n$ as calculated by the next proposition.

Proposition 7 *Let $\boldsymbol{x} \in \mathbb{R}^n$. Then, the subdifferential of* MAD *at $\boldsymbol{x}$ is given by*

$$\partial \mathrm{MAD}(\boldsymbol{x}) = \left\{ \mathbb{E}\left[\left(\mathrm{sgn}\left((\boldsymbol{r} - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{x}\right) + \eta \mathbb{1}_{\{(\boldsymbol{r}-\boldsymbol{\mu})^\mathsf{T}\boldsymbol{x}=0\}}\right)(\boldsymbol{r} - \boldsymbol{\mu})\right] \mid \eta \in L^0, \ \eta \in [-1, 1] \ \mathbb{P}\text{-a.s.} \right\}.$$

Proof By viewing MAD as an integral functional with respect to the probability measure $\mathbb{P}$, we may calculate its subdifferential using Rockafellar (1973, Theorem 23), which works under the assumption that the probability space is complete. (See also Rockafellar and Wets 1982, Corollary, p. 179 for a similar result with slightly different assumptions.) Note that we may write

$$\mathrm{MAD}(\boldsymbol{x}) = \int_\Omega h(\omega, \boldsymbol{x}) \mathbb{P}(d\omega), \quad \boldsymbol{x} \in \mathbb{R}^n,$$

where $h \colon \Omega \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$h(\omega, \boldsymbol{x}) := \left|(\boldsymbol{r}(\omega) - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{x}\right|, \quad \omega \in \Omega, \boldsymbol{x} \in \mathbb{R}^n.$$

For every $\boldsymbol{x} \in \mathbb{R}^n$, the function $\omega \mapsto h(\omega, \boldsymbol{x})$ on Ω is measurable since $\boldsymbol{r}$ is a random vector; for every $\omega \in \Omega$, the function $\boldsymbol{x} \mapsto h(\omega, \boldsymbol{x})$ is convex and continuous on $\mathbb{R}^n$. In particular, h is measurable on $\Omega \times \mathbb{R}^n$. Let $\boldsymbol{\xi} \colon \Omega \to \mathbb{R}^n$ be a bounded random variable, we can find $M > 0$ such that $|\xi_i| \le M$ for every $i \in \{1, \ldots, n\}$ $\mathbb{P}$-a.s. Then,

$$\int_\Omega h(\omega, \boldsymbol{\xi}(\omega)) \mathbb{P}(d\omega) = \int_\Omega \left|(\boldsymbol{r}(\omega) - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{\xi}(\omega)\right| \mathbb{P}(d\omega) \le M \sum_{i=1}^n \int_\Omega |r_i(\omega) - \mu_i| \mathbb{P}(d\omega) < +\infty$$

since $r_1, \ldots, r_n \in L^1$. Hence, all assumptions in Rockafellar (1973, Theorem 23) are satisfied and we obtain that

$$\partial \mathrm{MAD}(\boldsymbol{x}) = \{\mathbb{E}[\boldsymbol{\xi}] \mid \boldsymbol{\xi} \in (L^1)^n, \ \mathbb{P}\{\omega \in \Omega \mid \boldsymbol{\xi}(\omega) \in \partial_{\boldsymbol{x}} h(\omega, \boldsymbol{x})\} = 1\}, \qquad (8)$$

where, for each $\omega \in \Omega$, $\partial_{\boldsymbol{x}} h(\omega, \boldsymbol{x})$ denotes the subdifferential of the function $\boldsymbol{x} \mapsto h(\omega, \boldsymbol{x})$ at $\boldsymbol{x}$. By standard rules of finite-dimensional subdifferential calculus (see, e.g., Rockafellar 1970, Theorem 23.9), we can calculate this set as

$$\partial_{\boldsymbol{x}} h(\omega, \boldsymbol{x}) = (\boldsymbol{r}(\omega) - \boldsymbol{\mu}) \partial |\cdot| ((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{x}), \qquad (9)$$

where $\partial |\cdot|$ denotes the (set-valued) subdifferential mapping of the absolute value function and is given by $\partial |\cdot| (s) = \{\mathrm{sgn}(s)\}$ for $s \neq 0$, and by $\partial |\cdot| (s) = [-1, +1]$ for $s = 0$, which can be written compactly as

$$\partial |\cdot| (s) = \{\mathrm{sgn}(s) + t \mathbb{1}_{\{0\}}(s) \mid t \in [-1, +1]\}, \quad s \in \mathbb{R}. \tag{10}$$

Combining (9), (10), we get

$$\partial_{\boldsymbol{x}} h(\omega, \boldsymbol{x}) = \{(\mathrm{sgn}((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x}) + t \mathbb{1}_{\{0\}}((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x}))(\boldsymbol{r}(\omega) - \boldsymbol{\mu}) \mid t \in [-1, +1]\}.$$

Combining this with (8) yields that $\partial \, \mathrm{MAD}(\boldsymbol{x})$ is the set of all expectations of the form $\mathbb{E}[\boldsymbol{\xi}]$, where $\boldsymbol{\xi} \in (L^1)^n$ is such that

$$\mathbb{P}\{\omega \in \Omega \mid \exists t \in [-1, 1] \colon \boldsymbol{\xi}(\omega) = (\mathrm{sgn}((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x}) + t \mathbb{1}_{\{(\boldsymbol{r} - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x} = 0\}}(\omega))(\boldsymbol{r}(\omega) - \boldsymbol{\mu})\} = 1. \tag{11}$$

Here, t depends on ω but its measurability is not guaranteed. To resolve this measurability issue, we show that condition (11) holds if and only if there exists $\eta \in L^0$ such that $-1 \leq \eta \leq 1$ $\mathbb{P}$-a.s. and

$$\mathbb{P}\{\omega \in \Omega \mid \boldsymbol{\xi}(\omega) = (\mathrm{sgn}((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x}) + \eta(\omega) \mathbb{1}_{\{(\boldsymbol{r} - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x} = 0\}}(\omega))(\boldsymbol{r}(\omega) - \boldsymbol{\mu})\} = 1. \tag{12}$$

Indeed, the "if" part of the claim is trivial. For the "only if" part, suppose that (11) holds. Let $A \in \mathcal{F}$ be the event in the condition. Hence, the set

$$T(\omega) := \{t \in [-1, 1] \mid \boldsymbol{\xi}(\omega) = (\mathrm{sgn}((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x}) + t \mathbb{1}_{\{(\boldsymbol{r} - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x} = 0\}}(\omega))(\boldsymbol{r}(\omega) - \boldsymbol{\mu})\}$$

is nonempty for every $\omega \in A$. Moreover,

$$(t, \omega) \mapsto (\mathrm{sgn}((\boldsymbol{r}(\omega) - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x}) + t \mathbb{1}_{\{(\boldsymbol{r} - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x} = 0\}}(\omega))(\boldsymbol{r}(\omega) - \boldsymbol{\mu})$$

is a Carathéodory function, i.e., it is continuous in $t \in [-1, 1]$ and measurable in $\omega \in \Omega$. For definiteness, let us set $t(\omega) = 0$ for $\omega \in \Omega \backslash A$. The measurability of $\boldsymbol{\xi}$ together with Rockafellar and Wets (1997, Corollary 14.6, Example 14.15) imply that there exists a random variable η such that $\eta(\omega) \in T(\omega)$ for almost every $\omega \in \Omega$. Therefore, (12) holds for η, which yields the desired subdifferential formula for MAD. $\qquad \square$

3 Risk parity approach with MAD

The risk parity (RP) approach aims to choose the portfolio weights such that the risk contributions of all assets are equal. In this section, we discuss the existence, uniqueness, and calculation of such portfolio weights.

3.1 MAD-RP portfolios: existence and uniqueness

The standard method for decomposing a risk measure (in the broad sense) into additive components is based on *Euler's homogenous function theorem*. Since $\boldsymbol{x} \mapsto \mathrm{MAD}(\boldsymbol{x})$ is a homogeneous function of degree 1, if it were continuously differentiable, then we would have

$$\mathrm{MAD}(\boldsymbol{x}) = \nabla \, \mathrm{MAD}(\boldsymbol{x})^{\mathsf{T}} \boldsymbol{x} = \sum_{i=1}^{n} x_i \frac{\partial \, \mathrm{MAD}(\boldsymbol{x})}{\partial x_i}. \tag{13}$$

Similar to the case of volatility treated as a risk measure (Maillard & Teiletche, 2010), we could interpret the quantity $x_i \frac{\partial \text{MAD}(x)}{\partial x_i}$ as the risk contribution of the asset i to the global MAD of the portfolio.

We need a suitable replacement of (13) since MAD is not differentiable at some points of its domain in general. Nevertheless, by the generalized version of Euler's homogeneous function theorem for the subdifferentials of convex functions Yang and Wei (2008, Theorem 3.1), we still have

$$\text{MAD}(x) = s^\mathsf{T} x, \quad s \in \partial \text{MAD}(x). \tag{14}$$

Given two vectors $a, b \in \mathbb{R}^n$, we write $a \cdot b = (a_1 b_1, \dots, a_n b_n)^\mathsf{T}$ for their Hadamard product. Let us introduce the set

$$\mathcal{RC}(x) := x \cdot \partial \text{MAD}(x) := \{x \cdot s(x) \mid s(x) \in \partial \text{MAD}(x)\}. \tag{15}$$

Suppose that $x \in \Delta^{n-1}$ is a portfolio vector. For each $\text{RC}(x) = (\text{RC}_i(x), \dots, \text{RC}_n(x))^\mathsf{T} \in \mathcal{RC}(x)$, (14) yields that

$$\text{MAD}(x) = \sum_{i=1}^{n} \text{RC}_i(x).$$

For this reason, we call $\text{RC}(x)$ a *feasible risk contribution vector* for portfolio x. For instance, it is easy to verify that $x \cdot \mathbb{E}[\text{sgn}((r - \mu)^\mathsf{T} x)(r - \mu)]$ is a feasible risk contribution vector for x; see Proposition 7.

A MAD-RP portfolio is characterized by the requirement of having the same risk contribution for each asset. More precisely, a long-only portfolio $x \in \Delta^{n-1}$ is called a *MAD-RP portfolio* if there exists $\text{RC}(x) \in \mathcal{RC}(x)$ such that

$$\text{RC}_i(x) = \text{RC}_j(x) \tag{16}$$

for each $i, j \in \{1, \dots, n\}$ with $i \neq j$. Then, a long-only MAD-RP portfolio can be found by solving the system

$$\begin{cases} \text{RC}_i(x) = \lambda, & i \in \{1, \dots, n\}, \\ x \in \Delta^{n-1}, \\ \text{RC}(x) \in \mathcal{RC}(x), \\ \lambda \in \mathbb{R}. \end{cases} \tag{17}$$

For the results of this section, we work under the following nondegeneracy assumption.

Assumption 8 The only vector $x \in \mathbb{R}^n_+$ such that $(r - \mu)^\mathsf{T} x = 0$ $\mathbb{P}$-a.s. is $x = 0$.

We discuss the mildness of Assumption 8 in the next two examples.

Example 9 Suppose that $r - \mu$ has distinct values $v_1, \dots, v_m \in \mathbb{R}^n$ with respective probabilities $p_1, \dots, p_m > 0$, where $m \geq n$ and $\sum_{j=1}^{m} p_j = 1$. In this case, the condition $(r - \mu)^\mathsf{T} x = 0$ $\mathbb{P}$-a.s. is equivalent to the linear system $Ax = 0_m$, where $A \in \mathbb{R}^{m \times n}$ is the matrix whose respective rows are $v_1^\mathsf{T}, \dots, v_m^\mathsf{T}$ and 0_m is the m-dimensional zero vector. Then, the following are equivalent:

(i) Assumption 8 holds.
(ii) The only solution of the system $Ax = 0_m$, $x \in \mathbb{R}^n_+$ is $x = 0$.
(iii) There exists $\lambda \in \mathbb{R}^m$ such that $A^\mathsf{T} \lambda \in \mathbb{R}^n_{++}$.
(iv) There exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that $\sum_{j=1}^{m} \lambda_j v_j \in \mathbb{R}^n_{++}$.

Here, the equivalence between (ii) and (iii) is by Gordan's Alternative Theorem; see Gordan (1873) and Güler (2010, Theorem 3.14). In particular, Assumption 8 holds if the rank of A is n, i.e., there are n linearly independent vectors among $v_1, \ldots, v_m$.

Example 10 Suppose that $r - \mu$ has the multivariate centered Gaussian distribution with covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. Then, for every $x \in \mathbb{R}^n$, $(r - \mu)^\mathsf{T} x$ has the univariate centered Gaussian distribution with variance $x^\mathsf{T} \Sigma x$. In this case, the condition $(r - \mu)^\mathsf{T} x = 0$ $\mathbb{P}$-a.s. is equivalent to $x^\mathsf{T} \Sigma x = 0$. Hence, Assumption 8 holds if and only if Σ is strictly copositive. In particular, Assumption 8 holds if Σ is positive definite, i.e., Σ has full rank or, equivalently, Σ is nonsingular.

The next proposition establishes the existence and uniqueness of a MAD-RP portfolio.

Proposition 11 *Suppose that Assumption 8 holds.*

(i) Let $\lambda > 0$ and consider the problem

$$\min_{x \in \mathbb{R}^n_{++}} \text{MAD}(x) - \lambda \sum_{i=1}^n \ln x_i. \tag{18}$$

Then, there exists a unique optimal solution of this problem.
(ii) There exists a unique MAD-RP portfolio, that is, the system (17) has a unique solution.

Proof (i) The supposition ensures that $\text{MAD}(x) = \mathbb{E}[|(r - \mu)^\mathsf{T} x|] > 0$ for every $x \in \mathbb{R}^n_+ \setminus \{0\}$, that is, MAD is a positive function as defined in Cesarone et al. (2020b, Section 2). Moreover, MAD is also convex and positively homogeneous. Although MAD is not continuously differentiable at some points of its domain, the same arguments as in the proof of Cesarone et al. (2020b, Section 2) are applicable, and it follows that the given problem has a unique optimal solution.

(ii) Let us fix some $\lambda > 0$. By (i), there exists a unique optimal solution $\bar{x}$ of the problem in (18). By Cesarone et al. (2020b, Proposition 2), $x^* := \frac{\bar{x}}{\sum_{i=1}^n \bar{x}_i}$ is the unique optimal solution of the problem in (18) but with λ replaced with $\lambda^* := \frac{\lambda}{\sum_{i=1}^n \bar{x}_i}$. By the first order condition for optimality, we have

$$0 \in \partial \text{MAD}(x^*) - \lambda^* \left[\frac{1}{x_i^*} \right]_{i=1}^n, \tag{19}$$

that is,

$$(\lambda^*, \ldots, \lambda^*)^\mathsf{T} \in x^* \cdot \partial \text{MAD}(x^*).$$

Hence, there exists $s(x^*) \in \partial \text{MAD}(x^*)$ such that $x_i^* s_i(x^*) = \lambda^*$ for every $i \in \{1, \ldots, n\}$. In particular, $\text{RC}(x) := x^* \cdot s(x^*) \in \mathcal{RC}(x^*)$ and $\text{RC}_i(x) = \lambda^*$ for every $i \in \{1, \ldots, n\}$. Therefore, x^* is a MAD-RP portfolio. Its uniqueness follows from the strict convexity of the logarithmic objective function in (18). $\qquad \square$

Remark 12 In Cesarone et al. (2020b, Theorem 2), the existence and uniqueness of an RP portfolio are shown for a continuously differentiable risk measure. Their proof has two main components: (1) arguments based on coercivity and Weierstrass theorem to prove that the associated logarithmic problem (similar to (18)) has a unique optimal solution, where the positivity, convexity, and positive homogeneity of the risk measure are used but its differentiability is not used, (2) their Proposition 1, where it is shown that an optimal solution of the logarithmic problem gives rise to an RP portfolio and the proof uses first order conditions for

the minimization of differentiable functions. In the present paper, (1) still works when the risk measure is replaced with MAD, as stated in the proof of Proposition 11(i). However, in the proof of Proposition 11(ii), we extend (2) by using the more general first order condition based on subdifferentials.

Assuming that the portfolio MAD is additive (see Theorem 6(iv)), i.e., the asset returns satisfy the condition

$$(r_i - \mu_i)(r_j - \mu_j) \geq 0 \quad \mathbb{P}\text{-a.s.} \tag{20}$$

for every $i, j \in \{1, \ldots, n\}$ with $i \neq j$, we show in the next proposition that the weights of the assets in the MAD-RP portfolio are proportional to the reciprocals of the MADs of the individual asset returns.

We point out that taking a portfolio with weights proportional to the inverse of the risks of the individual asset returns is a naïve approach frequently used in practice. This approach is often called *naïve risk parity* due to the strong implicit assumption on the dependence of the asset returns (see, e.g., Qian 2011, Qian 2017, Clarke et al. 2013, and Haesen et al. 2017).

Proposition 13 (*Closed-form solution for the long-only MAD-RP portfolio*) *Suppose that Assumption 8 holds. Furthermore, suppose that the additivity condition* (20) *holds. Then, the unique MAD-RP portfolio is given by*

$$x_i = \frac{\mathbb{E}[|r_i - \mu_i|]^{-1}}{\sum_{j=1}^{n} \mathbb{E}[|r_j - \mu_j|]^{-1}}, \quad i \in \{1, \ldots, n\}. \tag{21}$$

Proof The first supposition ensures that there is a unique MAD-RP portfolio by Proposition 11. Moreover, as checked in the proof of Proposition 11, we have $\text{MAD}(r_i) = \mathbb{E}[|r_i - \mu_i|] > 0$ for each $i \in \{1, \ldots, n\}$. Hence, the expression in (21) is well-defined.

Under (20), MAD is additive for long-only portfolios by Theorem 6. Hence, we have

$$\text{MAD}(\boldsymbol{x}) = \sum_{i=1}^{n} x_i \, \text{MAD}(r_i) \tag{22}$$

for every $\boldsymbol{x} \in \mathbb{R}_+^n$, i.e., MAD is linear on $\mathbb{R}_+^n$. It follows that MAD is differentiable at every $\boldsymbol{x} \in \mathbb{R}_{++}^n$ and the gradient is given by $\nabla \text{MAD}(\boldsymbol{x}) = (\text{MAD}(r_1), \ldots, \text{MAD}(r_n))^{\mathsf{T}}$. Hence, $\mathcal{RC}(\boldsymbol{x})$ is a singleton and the only feasible risk contribution vector is given by

$$\text{RC}(\boldsymbol{x}) = (x_1 \, \text{MAD}(r_1), \ldots, x_n \, \text{MAD}(r_n))^{\mathsf{T}}$$

for each $\boldsymbol{x} \in \mathbb{R}_{++}^n$. Imposing condition (16), we can write

$$x_i \mathbb{E}[|r_i - \mu_i|] = x_j \mathbb{E}[|r_j - \mu_j|], \quad i, j \in \{1, \ldots, n\}. \tag{23}$$

Thus, for a fixed $i \in \{1, \ldots, n\}$, considering the full investment constraint $\sum_{j=1}^{n} x_j = 1$, we obtain

$$\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} \frac{\mathbb{E}[|r_i - \mu_i|]}{\mathbb{E}[|r_j - \mu_j|]} x_i = \mathbb{E}[|r_i - \mu_i|] x_i \sum_{j=1}^{n} \mathbb{E}[|r_j - \mu_j|]^{-1} = 1.$$

From this, (21) follows immediately. Hence, the portfolio given in (21) is a MAD-RP portfolio, and this is the only such portfolio as established above. $\square$

3.2 MAD-RP portfolios: calculations

In the following sections, we provide three methods for finding MAD-RP portfolios in practice. For these methods, as usual in portfolio optimization, we use a lookback approach where the possible realizations of the discrete random returns arise from historical data. This will ensure that our optimization problems are finite-dimensional. Let $T \in \mathbb{N}$ be the length of the time series of the prices of the assets. For each $i \in \{1, \ldots, n\}$ and $t \in \{0, \ldots, T\}$, we denote by p_{it} the realized price of asset i at time t, and for $t \geq 1$, we denote by $r_{it} = \frac{p_{it} - p_{i(t-1)}}{p_{i(t-1)}}$ the realized (linear) return of asset i for the period ending at t. We assume that there are no ties of the outcomes and, therefore, that each historical scenario is equally likely with probability $\frac{1}{T}$ (see, e.g., Carleo et al. 2017, Cesarone 2020, and references therein). In particular, for each $i \in \{1, \ldots, n\}$, the random return r_i of asset i is a discrete random variable with the uniform distribution over the set $\{r_{i1}, \ldots, r_{iT}\}$ and we have $\mu_i = \frac{1}{T} \sum_{t=1}^{T} r_{it}$. Furthermore, we assume that the only vector $\boldsymbol{x} \in \mathbb{R}_+^n$ such that $(\boldsymbol{r} - \boldsymbol{\mu})^{\mathsf{T}} \boldsymbol{x} = 0$ $\mathbb{P}$-a.s. is $\boldsymbol{x} = 0$, which guarantees the existence and uniqueness of an MAD-RP portfolio by Proposition 11.

3.2.1 Logarithmic formulations

The first formulation that we present consists of minimizing MAD with a logarithmic barrier term in the objective function (Bai et al., 2016) as already given in (18):

$$\begin{cases} \min_{\boldsymbol{x} \in \mathbb{R}_{++}^n} \ \mathrm{MAD}(\boldsymbol{x}) - \lambda \sum_{i=1}^{n} \ln x_i, \end{cases} \tag{24}$$

where $\lambda > 0$ is a constant. We name this strictly convex optimization problem **log_obj**. Suppose that **log_obj** has an optimal solution $\bar{\boldsymbol{x}}$ (due to the strict convexity of the objective function, it must be the unique optimal solution). Then, following the arguments in the proof of Proposition 11, the *normalized* portfolio $\boldsymbol{x}^* = \frac{\bar{\boldsymbol{x}}}{\sum_{i=1}^{n} \bar{x}_i}$ is the MAD-RP portfolio and it is the unique optimal solution of **log_obj** with λ replaced with $\lambda^* := \frac{\lambda}{\sum_{i=1}^{n} \bar{x}_i}$.

The second formulation has a logarithmic constraint and we call it **log_constr** (see Spinu 2013, Cesarone and Colucci 2018, Bellini et al. 2021). It consists of solving the following convex problem:

$$\begin{cases} \min_{\boldsymbol{x} \in \mathbb{R}_{++}^n} \ \mathrm{MAD}(\boldsymbol{x}) \\ \sum_{i=1}^{n} \ln x_i \geq c, \end{cases} \tag{25}$$

where $c \in \mathbb{R}$ is a constant. Moreover, if $\bar{\boldsymbol{x}}$ is an optimal solution of **log_constr**, then $\frac{\bar{\boldsymbol{x}}}{\sum_{i=1}^{n} \bar{x}_i}$ is the MAD-RP as checked in the next remark.

Remark 14 Since the function $\boldsymbol{x} \mapsto \mathrm{MAD}(\boldsymbol{x})$ is not differentiable everywhere, we can apply the general Karush-Kuhn-Tucker conditions with subdifferentials as formulated in Rockafellar (1970, Theorem 28.3). Note that the Lagrangian of **log_constr** is given by

$$L(\boldsymbol{x}, \lambda) = \mathrm{MAD}(\boldsymbol{x}) + \lambda \left(c - \sum_{i=1}^{n} \ln(x_i) \right) \tag{26}$$

for each $\boldsymbol{x} \in \mathbb{R}^n_{++}$, $\lambda \geq 0$. Then, for a fixed $\lambda \geq 0$, the subdifferential of $\boldsymbol{x} \mapsto L(\boldsymbol{x}, \lambda)$ at $\boldsymbol{x} \in \mathbb{R}^n_{++}$ is given by

$$\partial_{\boldsymbol{x}} L(\boldsymbol{x}, \lambda) = \partial \operatorname{MAD}(\boldsymbol{x}) - \lambda \left[\frac{1}{x_i} \right]_{i=1}^n .$$

In particular, the problem $\inf_{\boldsymbol{x} \in \mathbb{R}^n_{++}} L(\boldsymbol{x}, \lambda)$ of calculating the dual objective function at a given $\lambda > 0$ is equivalent to **log_obj**. Let $\bar{\boldsymbol{x}}$ be an optimal solution of **log_constr** with $\bar{\lambda} \geq 0$ denoting the Lagrange multiplier of the logarithmic constraint at optimality. Similar to (19), we have $0 \in \partial_{\boldsymbol{x}} L(\bar{\boldsymbol{x}}, \bar{\lambda})$. Arguing as in the proof of Proposition 11, we obtain the existence of $s(\bar{\boldsymbol{x}}) \in \partial \operatorname{MAD}(\bar{\boldsymbol{x}})$ such that $\bar{x}_i s_i(\bar{\boldsymbol{x}}) = \bar{\lambda}$ for every $i \in \{1, \ldots, n\}$. In particular, $\bar{\boldsymbol{x}} \cdot s(\bar{\boldsymbol{x}}) \in \operatorname{RC}(\bar{\boldsymbol{x}})$ and $\operatorname{MAD}(\bar{\boldsymbol{x}}) = n\bar{\lambda}$. Since $\operatorname{MAD}(\bar{\boldsymbol{x}}) > 0$ by the assumption stated at the beginning of this section, we have $\bar{\lambda} > 0$. On the other hand, by complementary slackness, we have $\bar{\lambda}(c - \sum_{i=1}^n \ln(\bar{x}_i)) = 0$. Hence, $c = \sum_{i=1}^n \ln(\bar{x}_i)$. It follows that the normalized version $\boldsymbol{x}^* := \frac{\bar{\boldsymbol{x}}}{\sum_{i=1}^n \bar{x}_i}$ is the MAD-RP portfolio and it is an optimal solution of **log_constr** with c replaced with $c^* := \sum_{i=1}^n \ln(x_i^*) = \sum_{i=1}^n \ln(\bar{x}_i) - n \ln(\sum_{i=1}^n \bar{x}_i) = c - n \ln(\sum_{i=1}^n \bar{x}_i)$.

Formulation (25) allows us to obtain some features of the risk parity portfolio $\boldsymbol{x}^{MAD-RP}$ related to a global minimizer $\boldsymbol{x}^{MinMAD}$ of MAD (i.e., the MinMAD portfolio) and the EW portfolio $\boldsymbol{x}^{EW}$, as shown in Remarks 15 and 16, respectively.

Remark 15 (MinMAD vs. MAD-RP) From the formulation **log_constr** in (25), it is clear that

$$\operatorname{MAD}(\boldsymbol{x}^{MinMAD}) \leq \operatorname{MAD}(\boldsymbol{x}^{MAD-RP}).$$

Remark 16 (EW vs. MAD-RP) Let us consider a slightly modified version of Problem (25), where we add the budget constraint $\sum_{i=1}^n x_i = 1$. Due to Jensen inequality, we have

$$\frac{1}{n} \sum_{i=1}^n \ln(x_i) \leq \ln\left(\frac{1}{n} \sum_{i=1}^n x_i \right) = -\ln(n) \tag{27}$$

for every feasible solution $\boldsymbol{x}$ of the new problem. Now, fixing $c = -n \ln(n)$ in Problem (25), the unique feasible solution is given by $x_i = \frac{1}{n}$, i.e., $\boldsymbol{x} = \boldsymbol{x}^{EW}$. As a consequence, we have

$$\operatorname{MAD}(\boldsymbol{x}^{MAD-RP}) \leq \operatorname{MAD}(\boldsymbol{x}^{EW}).$$

3.2.2 System-of-equation formulations

An alternative method, named **soe_1**, consists of solving the system (17) directly. Following the expression of the set $\mathcal{RC}(\boldsymbol{x})$ in (15) and that of the subdifferential $\partial \operatorname{MAD}(\boldsymbol{x})$ in Proposition 7, we can rewrite this system as

$$\begin{cases} \frac{x_i}{T} \sum_{t=1}^T s_t (r_{it} - \mu_i) = \lambda, & i \in \{1, \ldots, n\}, \\ \sum_{i=1}^n x_i = 1, & \\ x_i \geq 0, & i \in \{1, \ldots, n\}, \\ \sum_{i=1}^n (r_{it} - \mu_i) x_i \neq 0 \Rightarrow s_t = \operatorname{sgn}(\sum_{i=1}^n (r_{it} - \mu_i) x_i), & t \in \{1, \ldots, T\}, \\ -1 \leq s_t \leq +1, & t \in \{1, \ldots, T\}, \\ \lambda \in \mathbb{R}, \ x_i \in \mathbb{R}, \ s_t \in \mathbb{R}, & i \in \{1, \ldots, n\}, t \in \{1, \ldots, T\}. \end{cases} \tag{28}$$

Note that s_t is an auxiliary decision variable in this formulation and it stands for a subgradient of the absolute value function at the point $\sum_{i=1}^{n}(r_{it}-\mu_i)x_i$. The logical implication constraint in the above formulation can be converted into linear inequalities by introducing additional variables. We rewrite s_t as a convex combination of $+1$ and -1, i.e., $s_t = \alpha_t - (1-\alpha_t) = 2\alpha_t - 1$ for some variable α_t with values in $[0,1]$, where α_t is enforced to be 1 when $\sum_{i=1}^{n}(r_{it}-\mu_i)x_i > 0$, and it is enforced to be 0 when $\sum_{i=1}^{n}(r_{it}-\mu_i)x_i < 0$. This is achieved by the following formulation with the help of two additional binary decision variables:

$$
\begin{cases}
\frac{x_i}{T}\sum_{t=1}^{T}(2\alpha_t - 1)(r_{it}-\mu_i) = \lambda, & i \in \{1,\ldots,n\}, \\
\sum_{i=1}^{n} x_i = 1, & \\
x_i \geq 0, & i \in \{1,\ldots,n\}, \\
-Mu_t \leq \sum_{i=1}^{n}(r_{it}-\mu_i)x_i \leq Mv_t, & t \in \{1,\ldots,T\}, \\
1 - M(1-v_t) \leq \alpha_t \leq M(1-u_t), & t \in \{1,\ldots,T\}, \\
0 \leq \alpha_t \leq 1, & t \in \{1,\ldots,T\}, \\
u_t, v_t \in \{0,1\}, & t \in \{1,\ldots,T\}, \\
\lambda \in \mathbb{R};\ x_i \in \mathbb{R};\ \alpha_t, u_t, v_t \in \mathbb{R}, & i \in \{1,\ldots,n\},\ t \in \{1,\ldots,T\},
\end{cases}
\tag{29}
$$

where $M > 0$ is a suitably large constant. This reformulation is processed numerically using the global optimization software BARON (Tawarmalani and Sahinidis 2005).

Remark 17 As a simplification of (29), one may consider the following system:

$$
\begin{cases}
\frac{x_i}{T}\sum_{t=1}^{T}(v_t - u_t)(r_{it}-\mu_i) = \lambda, & i \in \{1,\ldots,n\}, \\
\sum_{i=1}^{n} x_i = 1, & \\
x_i \geq 0, & i \in \{1,\ldots,n\}, \\
-Mu_t \leq \sum_{i=1}^{n}(r_{it}-\mu_i)x_i \leq Mv_t, & t \in \{1,\ldots,T\}, \\
v_t + u_t \leq 1, & t \in \{1,\ldots,T\}, \\
u_t, v_t \in \{0,1\}, & t \in \{1,\ldots,T\}, \\
\lambda \in \mathbb{R};\ x_i \in \mathbb{R};\ u_t, v_t \in \mathbb{R}, & i \in \{1,\ldots,n\},\ t \in \{1,\ldots,T\}.
\end{cases}
\tag{30}
$$

In this system, $v_t - u_t$ stands for a subgradient of the absolute value function at a point where we have $\sum_{i=1}^{n}(r_{it}-\mu_i)x_i$; however, it is restricted to the set $\{-1,0,+1\}$ when $\sum_{i=1}^{n}(r_{it}-\mu_i)x_i = 0$. Hence, for some corner cases, the system (30) may be infeasible even though the MAD-RP portfolio exists. On the other hand, if x^* is part of a solution of (30), then it is the MAD-RP portfolio. For this reason, we will use (30) in the empirical analysis of Sect. 4 when implementing **soe_1**.

Finally, we also consider the following formulation, which we call **soe_2** and consists of a further simplified version of the system (17)

$$
\begin{cases}
q_i^2 \mathbb{E}\left[\text{sgn}\left[\sum_{j=1}^{n}(r_j - \mu_j)q_j^2 \right](r_i - \mu_i) \right] = \frac{z}{n}, & i \in \{1,\ldots,n\}, \\
\sum_{i=1}^{n} q_i^2 = 1, &
\end{cases}
\tag{31}
$$

where we remove the sign restrictions for x.

3.2.3 Least-squares formulations

The last method, called **ls_rel**, exploits the fact that we have $\frac{RC_i(x)}{MAD(x)} = \frac{1}{n}$ for each $i \in \{1,\ldots,n\}$ whenever x is the MAD-RP portfolio (see Cesarone and Colucci 2018). More

precisely, it consists of the following least-squares formulation:

$$\begin{cases} \min \sum_{i=1}^{n} \left(\dfrac{RC_i}{MAD(\boldsymbol{x})} - \dfrac{1}{n} \right)^2 \\ \mathbf{RC} \in \mathcal{RC}(\boldsymbol{x}), \\ \sum_{i=1}^{n} x_i = 1, \\ x_i \geq 0, \qquad\qquad i \in \{1, \ldots, n\}, \\ \boldsymbol{x} \in \mathbb{R}^n, \ \mathbf{RC} \in \mathbb{R}^n. \end{cases} \qquad (32)$$

A variant of this least-squares method is given by

$$\begin{cases} \min \sum_{i=1}^{n} \sum_{j=1}^{n} (RC_i - RC_j)^2 \\ \mathbf{RC} \in \mathcal{RC}(\boldsymbol{x}), \\ \sum_{i=1}^{n} x_i = 1, \\ x_i \geq 0, \qquad\qquad i \in \{1, \ldots, n\}, \\ \boldsymbol{x} \in \mathbb{R}^n, \ \mathbf{RC} \in \mathbb{R}^n, \end{cases} \qquad (33)$$

named **ls_abs** (see, e.g., Maillard and Teiletche 2010). The functional constraint $\mathbf{RC} \in \mathcal{RC}(\boldsymbol{x})$ can be reformulated in terms of linear inequalities with additional binary variables as in Sect. 3.2.2. If an optimal solution for **ls_rel** or **ls_abs** is found with zero optimal value, then this solution is the MAD-RP portfolio. Nevertheless, numerical methods can only guarantee local optimality due to the nonconvex nature of these problems. Hence, using the least-squares formulations should be considered as a heuristic approach.

4 Empirical analysis

In this section, we provide an extensive empirical analysis based on three investment universes which are described in Sect. 4.1. More precisely, in Sect. 4.2.1, we test and compare all the MAD-RP formulations in terms of accuracy and efficiency. Section 4.2.2 shows some graphical examples of the portfolio selection strategies analyzed in terms of portfolio weights and relative risk distribution among assets, using both volatility and MAD as risk measures. In Sect. 4.2.3, we report and discuss their out-of-sample performance results based on a Rolling Time Window scheme of evaluation.

4.1 Datasets

We give here a brief description of the three real-world datasets used in the empirical analysis. These datasets were obtained from Refinitiv and are publicly available on https://www. francescocesarone.com/.

- *ETF Emerging countries* (ETF-EC): it consists of the total return index,[1] expressed in US dollars, of the ETFs of 24 emerging countries.

---

[1] It consists in the price of the asset including dividends, assuming these dividends are reinvested in the company.
See, e.g., https://www.msci.com/eqb/methodology/meth_docs/MSCI_May12_IndexCalcMethodology.pdf.

Table 1 List of the datasets analyzed

Dataset	# Assets	Days	From-To
ETF-EC	24	1042	1/2015-12/2018
EuroBonds	62	1564	1/2013-12/2018
CIB-mix	11	1564	1/2013-12/2018

- *Euro bonds* (EuroBonds): it consists of the total return index in euros of the government bonds of 11 countries belonging to the Eurozone, with maturities ranging from 1 year to 30 years.
- *Commodities and Italian bonds mix* (CIB-mix): it consists of the total return index, expressed in euros, of a mixture of Italian government bonds, whose maturity ranges from 2 to 30 years, and commodities (agriculture, gold, energy and industrial metals).

In Table 1, we provide additional information about these datasets.

4.2 Computational results

4.2.1 Accuracy and efficiency of the MAD-RP formulations

In this section, we test and compare all the MAD-RP formulations, described in Sect. 3.2, in terms of accuracy and efficiency using the following quantities:

- the square root of the value of the objective function of Problem (32), named $\sqrt{F(\boldsymbol{x})}$;
- the portfolio MAD obtained from each method, named $\text{MAD}(\boldsymbol{x})$;
- the empirical mean of the absolute deviations of relative risk contributions from $\frac{1}{n}$,

$$\text{MeanAbsDev} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\text{RC}_i(\boldsymbol{x})}{\text{MAD}(\boldsymbol{x})} - \frac{1}{n} \right| ;$$

- the maximum of the absolute deviations of relative risk contributions from $\frac{1}{n}$,

$$\text{MaxAbsDev} = \max_{1 \le i \le n} \left| \frac{\text{RC}_i(\boldsymbol{x})}{\text{MAD}(\boldsymbol{x})} - \frac{1}{n} \right| ;$$

- the inverse of the number of assets selected, namely $\frac{1}{n}$;
- the computational times (in seconds) required to solve each MAD-RP formulation considered.

All the experiments have been implemented on a workstation with Intel Core CPU (i7-6700, 3.4 GHz, 16 Gb RAM) under MS Windows 10, using MATLAB 9.1. More precisely, we solve models **log_obj** (24), **log_constr** (25), **ls_rel** (32), **ls_abs** (33) by means of the built-in function fmincon; model **soe_1** (30) by using the global optimization software BARON (version 15.9.22), which is also called from MATLAB by means of the MATLAB/BARON toolbox (Tawarmalani & Sahinidis, 2005); finally, we solve **soe_2** (31) by means of the built-in function fsolve.

Since in Sect. 4.2.3 we perform an extensive out-of-sample analysis using in-sample windows of 2 years, for the following accuracy and efficiency analysis of the MAD-RP formulations we set the number of observations $T = 500$ days (i.e., the first two years of each dataset). For Problems (24) and (25), we also perform a sensitivity analysis by varying

Table 2 Experimental results for the MAD-RP methods using ETF-EC, where $1/n = 0.0417$

Formulation	$\sqrt{F(x)}$	MAD(x)	MeanAbsDev	MaxAbsDev	Time (secs.)
ls_rel	3.33e−03	6.25e−03	5.19e−04	1.74e−03	4.76
ls_abs	2.52e−03	6.24e−03	4.42e−04	9.37e−04	4.74
log_constr	1.49e−03	6.23e−03	1.96e−04	8.36e−04	5.92
log_obj	1.49e−03	6.23e−03	1.95e−04	8.36e−04	2.51
soe_1	-	-	-	-	> 1 day
soe_2	7.87e−03	6.29e−03	8.16e−04	7.28e−03	37.08

Table 3 Experimental results for the MAD-RP methods using EuroBonds, where $1/n = 0.0161$

Formulation	$\sqrt{F(x)}$	MAD(x)	MeanAbsDev	MaxAbsDev	Time (secs.)
ls_rel	2.54e−04	6.48e−04	2.36e−05	1.11e−04	62.18
ls_abs	5.19e−04	6.50e−04	4.87e−05	2.36e−04	65.22
log_constr	8.94e−05	6.47e−04	7.43e−06	6.15e−05	28.32
log_obj	9.02e−05	6.47e−04	7.75e−06	6.01e−05	5.43
soe_1	-	-	-	-	> 1 day
soe_2	5.88e−02	1.19e−03	6.57e−03	1.45e−02	131.86

**Table 4** Experimental results for the MAD-RP methods using CIB-mix, where $1/n = 0.0909$

Formulation	$\sqrt{F(x)}$	MAD(x)	MeanAbsDev	MaxAbsDev	Time (secs.)
ls_rel	2.48e−03	1.40e−03	6.16e−04	1.54e−03	2.66
ls_abs	2.32e−03	1.40e−03	5.63e−04	1.54e−03	2.09
log_constr	7.75e−04	1.40e−03	2.11e−04	3.48e−04	2.52
log_obj	7.74e−04	1.40e−03	2.13e−04	3.51e−04	1.29
soe_1	-	-	-	-	> 1 day
soe_2	7.98e−02	1.79e−03	1.83e−02	5.44e−02	18.65

the constants $\lambda > 0$ and $c \in \mathbb{R}$, respectively. Such a sensitivity analysis leads us to set $\lambda = 0.001$ and $c = -40$, which guarantee a reasonable level of accuracy for all experiments. Tables 2, 3, and 4 report the computational results obtained for the three datasets analyzed. Both **soe_1** and **soe_2** seem to be the least accurate and efficient formulations. Note that, with this experimental setup, the most computationally demanding formulation **soe_1** fails to find a solution in a reasonable time (1 day).

Methods based on least-squares formulations, i.e., **ls_rel** and **ls_abs**, achieve intermediate performances for all the datasets.

Finally, on one hand, the two logarithmic formulations essentially lead to the same level of accuracy. On the other hand, **log_obj** seems to be the most efficient method for the three datasets considered.
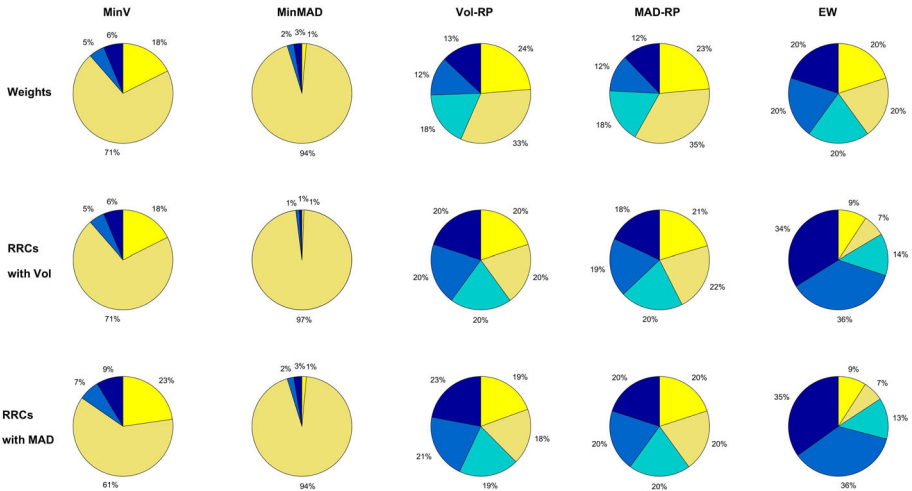
Fig. 1 Weights (top), RRCs with volatility (middle), and RRCs with volatility (bottom) pie charts for 5 assets belonging to CIB-mix

4.2.2 Comparison of some portfolio selection approaches

To better illustrate the concept of risk allocation, we report here a graphical comparison of the five portfolio selection strategies listed below.

1. MAD-RP: the portfolio obtained by solving the **log_obj** formulation.
2. MinMAD: the global minimum MAD portfolio, obtained by implementing problem (3.6) of Konno and Yamazaki (1991) without constraining the portfolio expected return.
3. MinV: the global minimum variance portfolio, as in Markowitz (1952).
4. Vol-RP: the portfolio obtained by means of the risk parity approach using volatility as a risk measure (Problem (7) of Maillard and Teiletche 2010).
5. EW: the Equally Weighted portfolio, where the capital is uniformly distributed among all the assets in the investment universe, i.e., $x_i = \frac{1}{n}$ for each $i \in \{1, \ldots, n\}$.

More precisely, we consider an investment universe of 5 assets and we evaluate weights and Relative Risk Contributions (RRCs), obtained from all the portfolio strategies analyzed. Note that for each portfolio approach, RRCs are computed by considering both volatility and MAD as risk measures. In Fig. 1, we report the results obtained for 5 assets belonging to CIB-mix.[2]

We can observe that the Vol-RP and MAD-RP portfolios generally show a slightly different portfolio composition but both appear to be well-diversified similar to the EW portfolio. A converse situation can be pointed out for the minimum risk portfolios. Indeed, as shown on the top of Fig. 1, both the MinV and MinMAD approaches concentrate 71% and 94% of the capital in only one asset, respectively. In the middle (bottom) of Fig. 1, for each portfolio strategy, we can see how much each asset contributes (in percentage) to the whole portfolio volatility (MAD). Clearly, for Vol-RP (MAD-RP), each asset equally contributes to the portfolio volatility (MAD). Furthermore, by construction, for a fixed risk measure, the

---

[2] IT Benchmark 10 years DS Govt. Index, IT Benchmark 5 years DS Govt. Index, IT Benchmark 15 years DS Govt. Index, S&P GSCI Agriculture Total Return, and S&P GSCI Industrial Metals Total Return for the period 1/2013-12/2018.

asset weights of the minimum risk portfolio coincide with their RRCs. It is also interesting to observe that, for the EW portfolio, RRCs are uneven regardless of the risk measure used.

4.2.3 Out-of-sample analysis

In this section, we discuss the main results of the out-of-sample performance analysis for all the portfolio selection strategies listed in Sect. 4.2.2. For this analysis, we adopt a Rolling Time Window scheme of evaluation: we allow for the possibility of rebalancing the portfolio composition during the holding period, at fixed intervals. To calibrate the portfolio selection models, we use 2 years (i.e., 500 days) for the in-sample window, while we choose 20 days for the out-of-sample window with rebalancing allowed every 20 days. To evaluate the out-of-sample performances of the five portfolios analyzed, we use the following quantities typically adopted in the literature (see, e.g., Bacon 2008; Rachev et al. 2008; Bruni et al. 2017; Cesarone et al. 2015, 2016, 2019, 2022). Let R^{out} denote the out-of-sample portfolio return, and $W_t = W_{t-1}(1 + R^{out})$ the portfolio wealth at time t.

- **Mean** is the annualized average portfolio return, i.e., $\mu^{out} = \mathbb{E}[R^{out}]$. Obviously, higher values are associated to higher performances.
- **Volatility (Vol)** is the annualized standard deviation, computed as $\sigma^{out} = \sqrt{\mathbb{E}[(R^{out} - \mu^{out})^2]}$. Since it measures the portfolio risk, lower values are preferable.
- **Mean Absolute Deviation (MAD)** of R^{out}, namely $\text{MAD}(R^{out}) := \mathbb{E}[|R^{out} - \mu^{out}|]$.
- **Maximum Drawdown (MaxDD)**: denoting the *drawdowns* by

$$dd_t = \frac{W_t - \max_{1 \leq \tau \leq t} W_\tau}{\max_{1 \leq \tau \leq t} W_\tau}, \quad t \in \{1, \ldots, T\},$$

MaxDD is defined as $MaxDD = \min_{1 \leq t \leq T} dd_t$.

It measures the loss from the observed peak in the returns; hence, it will always have negative sign (or, in the best case scenario, it will be equal to 0), meaning that the higher the better.

- **Ulcer** index, i.e.,

$$UI = \sqrt{\frac{\sum_{t=1}^{T} dd_t^2}{T}}.$$

It evaluates the depth and the duration of drawdowns in prices over the out-of-sample period. Lower **Ulcer** values are associated to better portfolio performances.

- **Sharpe** ratio measures the gain per unit risk and it is defined as

$$SR^{out} = \frac{\mu^{out} - r_f}{\sigma^{out}},$$

where we choose $r_f = 0$. The larger this value, the better the performance.

- **Sortino** ratio is similar to the Sharpe ratio but uses another risk measure, i.e., the Target Downside Deviation, $TDD = \sqrt{\mathbb{E}[((R^{out} - r_f)^-)^2]}$. Therefore, the Sortino Ratio is defined as

$$SoR = \frac{\mu^{out} - r_f}{TDD},$$

where $r_f = 0$. Similar to the Sharpe ratio, the higher it is, the better the portfolio performance.

Table 5 Out-of-sample results for ETF-EC. (Color figure online)

	MinV	MinMAD	Vol-RP	MAD-RP	EW
Mean	0.0485	0.0605	0.0666	0.0666	0.0703
Vol	0.0816	0.0801	0.0907	0.0904	0.1027
MAD	1.0087	0.9777	1.0990	1.0941	1.2485
MaxDD	-0.1532	-0.1514	-0.1959	-0.1989	-0.2218
Ulcer	0.0627	0.0601	0.0819	0.0838	0.0960
Sharpe	0.0376	0.0477	0.0465	0.0466	0.0433
Sortino	0.0527	0.0671	0.0641	0.0643	0.0597
Turnover	0.1184	0.1477	0.0242	0.0329	0
Rachev-5%	0.8473	0.8493	0.8381	0.8360	0.8270

- **Turnover**, i.e.,

$$Turn = \frac{1}{Q} \sum_{q=1}^{Q} \sum_{k=1}^{n} |x_{q,k} - x_{q-1,k}|,$$

  where Q represents the number of rebalances, $x_{q,k}$ is the portfolio weight of asset k after rebalancing, and $x_{q-1,k}$ is the portfolio weight before rebalancing at time q. Lower turnover values imply better portfolio performance.
  We point out that this definition of portfolio turnover is a proxy of the effective one, since it evaluates only the amount of trading generated by the models at each rebalance, without considering the trades due to changes in asset prices between one rebalance and the next. Thus, by definition, the turnover of the EW portfolio is zero.
- **Rachev-5%** ratio measures the upside potential, comparing right and left tail. Mathematically, it is computed as

$$\frac{CVaR_\alpha(r_f - R^{out})}{CVaR_\beta(R^{out} - r_f)},$$

  where we choose $\alpha = \beta = 5\%$ and $r_f = 0$.

In Tables 5, 6 and 7, we report the experimental results obtained for the three dataset, ETF-EC, EuroBonds and CIB-mix, respectively. For each performance measure, we show with different colors the rank of the results of the five portfolio selection strategies analyzed. More precisely, the colors range from deep-green to deep-red, where deep-green represents the best performance while deep-red the worst one.

As shown in Table 5, the MinMAD portfolio seems to have better performance with respect to the other portfolios with two exceptions: **Turnover** and **Mean**. Indeed, **Turnover** of the Vol-RP and MAD-RP portfolios are much smaller than that of MinMAD. The EW portfolio guarantees the highest **Mean**, followed by the Vol-RP and MAD-RP portfolios. We can observe that these two RP strategies provide very similar values for all the performance measures considered, and their out-of-sample performances, both in terms of risk and of gain, are located in the middle between those of the minimum risk portfolios and those of the EW portfolio, as theoretically expected in the in-sample case (see Remarks 15 and 16). The MinV and MinMAD portfolios appear to be the least risky, both in terms of **Vol**, **MAD**, **MaxDD**, and **Ulcer**. Furthermore, even though MAD-RP tends to be riskier than the

Table 6 Out-of-sample results for EuroBonds. (Color figure online)

	MinV	MinMAD	Vol-RP	MAD-RP	EW
Mean	-0.0004	-0.0006	0.0082	0.0085	0.0272
Vol	0.0046	0.0043	0.0117	0.0120	0.0379
MAD	0.0442	0.0425	0.1323	0.1346	0.4334
MaxDD	-0.0122	-0.0119	-0.0204	-0.0206	-0.0716
Ulcer	0.0047	0.0049	0.0064	0.0062	0.0223
Sharpe	-	-	0.0443	0.0447	0.0454
Sortino	-	-	0.0604	0.0611	0.0631
Turnover	0.1119	0.1153	0.0199	0.0401	0
Rachev-5%	0.8212	0.8297	0.8387	0.8465	0.9054

Table 7 Out-of-sample results for CIB-mix. (Color figure online)

	MinV	MinMAD	Vol-RP	MAD-RP	EW
Mean	0.0025	0.0023	0.0042	0.0073	0.0028
Vol	0.0201	0.0208	0.0290	0.0298	0.0605
MAD	0.1110	0.1198	0.3125	0.3296	0.7194
MaxDD	-0.0488	-0.0498	-0.0541	-0.0559	-0.1335
Ulcer	0.0080	0.0086	0.0261	0.0244	0.0657
Sharpe	0.0079	0.0069	0.0092	0.0155	0.0030
Sortino	0.0097	0.0084	0.0125	0.0214	0.0042
Turnover	0.0471	0.0382	0.0288	0.0886	0
Rachev-5%	0.8123	0.8326	0.9074	0.9217	0.9924

minimum risk strategies, it achieves the second-largest values of **Sharpe** and **Sortino**. Note that we do purposely leave the EW portfolio out of the comparison concerning **Turnover**, since its value is 0 by construction.

In Table 6, we report the computational results for EuroBonds. Again as expected, the minimum risk portfolios show the best risk performances but with the worst values of out-of-sample expected return. Conversely, the EW portfolio has the best values in terms of Gain-Risk ratios and the worst in terms of risk. Also in this case, the performances of the RP strategies generally lie between those of EW and those of the minimum risk portfolios. Furthermore, the Vol-RP portfolio shows the lowest value of **Turnover**, followed by MAD-RP.

In Table 7, we report the computational results for CIB-mix. We observe here that the MAD-RP portfolio presents very good performances: it has the highest **Mean**, **Sharpe** and **Sortino** values (followed by the Vol-RP portfolio), and the second-highest **Rachev-5%**, below that of EW. Again, the minimum risk portfolios, particularly MinV, have the best **Vol**, **MAD**, **MaxDD**, and **Ulcer**. Furthermore, the Vol-RP portfolio provides the lowest **Turnover**, followed by MinMAD.
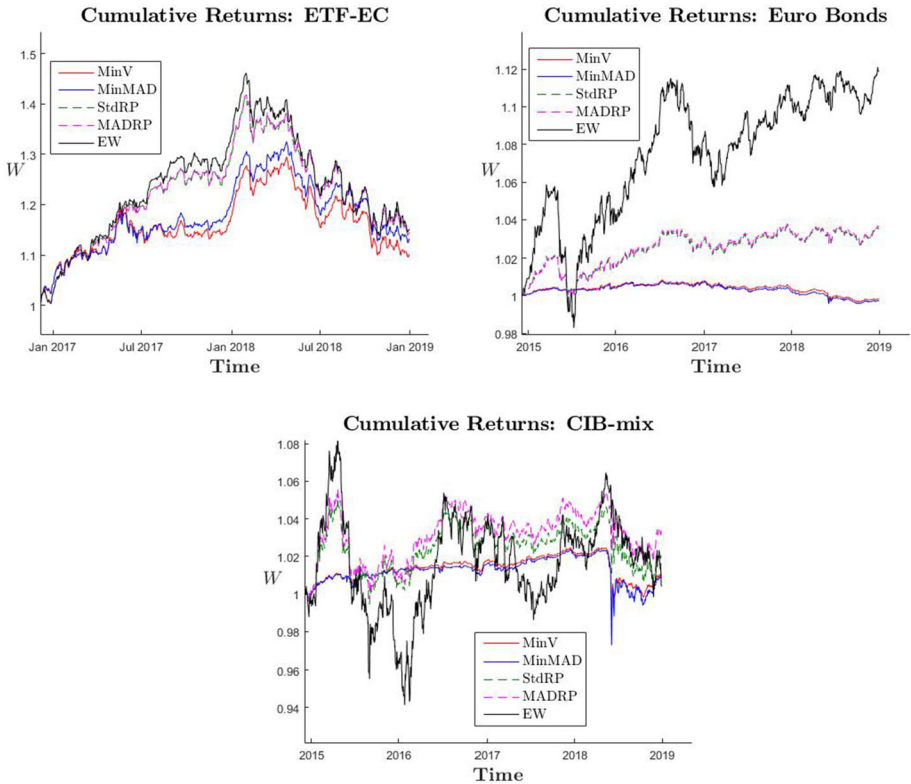
Fig. 2 Time evolution of the portfolio wealth for ETF-EC, EuroBonds and CIB-mix

Table 8 Analysis of the differences in the portfolio weights of the Risk Parity and minimum risk strategies

	ETF-EC	EuroBonds	CIB-mix
Vol-RP vs. MAD-RP	0.0023	0.0020	0.0197
MinV vs. MinMAD	0.0097	0.0038	0.0154

In Fig. 2, for ETF-EC, EuroBonds and CIB-mix, we report the time evolution of the portfolio wealth, investing one unit of the currency at the beginning of the investment horizon. We can note that the EW portfolio often provides the highest values of wealth and that the risk parity portfolios are systematically better than the minimum risk ones. The Vol-RP and MAD-RP portfolios tend to perform similarly, with the only exception of the CIB-mix dataset, where the MAD-RP portfolio performs significantly better and is often the best among all.

Since the out-of-sample performance measures for the risk parity portfolios using volatility and MAD generally tend to appear very close, we compute a simple metric to better highlight the differences between the portfolio weights selected by these two approaches. More precisely, for each dataset listed in Table 1, we use the following metric named Overall Weight Differences (OWD):

$$OWD = \frac{1}{Q} \sum_{q=1}^{Q} \left(\frac{1}{n} \sum_{k=1}^{n} |x_{q,k}^{Vol-RP} - x_{q,k}^{MAD-RP}| \right), \tag{34}$$

where, again, Q represents the number of rebalances, $x_{q,k}^{Vol-RP}$ and $x_{q,k}^{MAD-RP}$ are the weights of asset k obtained, at time q, by the RP strategy using volatility and MAD, respectively. Table 8 shows the results obtained for all datasets considered, in which the analysis of the differences in the portfolio weights of the MinV and MinMAD approaches is also included. For each dataset, we can observe that the OWD values computed for the risk parity and minimum risk strategies appear to be of the same order of magnitude. Furthermore, as highlighted for the out-of-sample performance results, it seems that the greatest portfolio allocation differences can be observed for the CIB-mix dataset.

5 Conclusions

In this paper, we have proposed an extension of the now popular Risk Parity approach for volatility to the Mean Absolute Deviation (MAD) as a portfolio risk measure.

From a theoretical viewpoint, we have discussed the subdifferentiability and additivity features of MAD, highlighting their usefulness in the risk parity strategy, and we have established the existence and uniqueness results for the MAD-RP portfolio. Furthermore, we have presented several mathematical formulations for finding the MAD-RP portfolios practically and we have tested them on three real-world datasets both in terms of efficiency and accuracy. We have observed that the system-of-equation formulations seem to be the least accurate and efficient. Methods based on least-squares formulations show intermediate performances for all the datasets, while those based on logarithmic formulations appear to be the best.

From an experimental viewpoint, we have presented an extensive empirical analysis comparing the out-of-sample performance obtained from the global minimum volatility and MAD portfolios, the volatility and MAD risk parity portfolios, and the Equally Weighted portfolio. In terms of out-of-sample risk, these results confirm the theoretical (in-sample) properties of the RP portfolio, namely its risk lies between that of the minimum risk portfolio and the risk of the Equally Weighted portfolio. The RP portfolios always show positive annual expected returns for all datasets, even though they are smaller than those of the Equally Weighted portfolio, which is however much riskier. Furthermore, in terms of Gain-Risk ratios, the risk parity portfolios tend to provide medium-high performances, often between those of EW and those of minimum risk portfolios; however, in few cases, they are better. We point out that the two risk parity portfolios show similar values for all performance measures. However, when directly comparing Vol-RP with MAD-RP, the latter usually provides slightly higher profitability at the expense of slightly higher risk and turnover. Furthermore, for each dataset, we have observed that the portfolio allocation differences computed through the risk parity and minimum risk strategies using volatility and MAD seem to be of the same order of magnitude.

Each model tested tends to respond to different requirements related to different risk attitudes of the investors. On one hand, as expected, the minimum risk models are advisable for risk-averse investors, avoiding as much as possible any shock represented by deep drawdowns. On the other hand, the RP strategies seem to be appropriate for investors mildly adverse to the total portfolio risk. These investors could be willing to waive a bit of safety (w.r.t. that of the minimum risk portfolios) and a bit of gain (w.r.t. that of the EW portfolio) to achieve a more balanced risk allocation and a more diversified portfolio. Furthermore, although the Equally Weighted approach embodies the concept of high diversification and achieves good out-of-sample expected returns, it generates portfolios with very high out-of-sample risk both in terms of volatility, MAD, Maximum Drawdown, and Ulcer index.

Therefore, according to our findings, the EW portfolio seems to be advisable for sufficiently risk-seeking investors who try to maximize gain without worrying about periods of deep drawdowns.

Declarations

Conflict of interest We declare that we do not have any conflicting interests related to this article.

Appendix A: Proofs of the results in Sect. 2.2

Proof of Lemma 2 By the triangle inequality on $\mathbb{R}$, we have that

$$|X - \mathbb{E}[X]| + |Y - \mathbb{E}[Y]| - |(X - \mathbb{E}[X]) + (Y - \mathbb{E}[Y])| \geq 0 \quad \mathbb{P}\text{-a.s.} \tag{35}$$

Note that the additivity condition (3) is equivalent to

$$\mathbb{E}\left[|(X - \mathbb{E}[X]) + (Y - \mathbb{E}[Y])| - |X - \mathbb{E}[X]| - |Y - \mathbb{E}[Y]|\right] = 0 \tag{36}$$

By (35), the random variable $|X - \mathbb{E}[X]| + |Y - \mathbb{E}[Y]| - |(X - \mathbb{E}[X]) + (Y - \mathbb{E}[Y])|$ is $\mathbb{P}$-a.s. nonnegative. Hence, (36) is equivalent to

$$|(X - \mathbb{E}[X]) + (Y - \mathbb{E}[Y])| - |X - \mathbb{E}[X]| - |Y - \mathbb{E}[Y]| = 0 \quad \mathbb{P}\text{-a.s.} \tag{37}$$

Now, for each $a, b \in \mathbb{R}$, we have $|a + b| = |a| + |b|$ if and only if $ab \geq 0$. It follows that (37) is equivalent to (4), which completes the proof.

Proof of Theorem 6 It is clear that (i) implies (ii) and (iii). Next, we show that (ii) implies (i). Let $x \in \mathbb{R}_+^n$. If $x = 0$, then the equality in (i) becomes trivial. Suppose that $x \neq 0$. In particular, $a := \sum_{i=1}^n x_i > 0$. Then, by (ii) and the positive homogeneity of MAD, we have

$$\text{MAD}\left(\sum_{i=1}^n x_i Y_i\right) = a\,\text{MAD}\left(\sum_{i=1}^n \frac{x_i}{a} Y_i\right) = a\sum_{i=1}^n \frac{x_i}{a}\,\text{MAD}(Y_i) = \sum_{i=1}^n x_i\,\text{MAD}(Y_i).$$

Hence, (i) follows.

To prove that (iii) implies (iv), let us fix $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and set $I := \{i, j\}$. Then, (iii) yields $\text{MAD}(Y_i + Y_j) = \text{MAD}(Y_i) + \text{MAD}(Y_j)$. Hence, by Theorem 2, we have $(Y_i - \mathbb{E}[Y_i])(Y_j - \mathbb{E}[Y_j]) \geq 0$ $\mathbb{P}$-a.s., that is, (iv) holds.

To complete the proof, we show that (iv) implies (i). We prove the equality in (i) by induction on $k(x) := \min\{i \geq 0 \mid x_{i+1} = \ldots = x_n = 0\}$. (We assume that $k(x) = n$ if

$x_n \neq 0$.) Note that $0 \leq k(\boldsymbol{x}) \leq n$. The base case $k(\boldsymbol{x}) = 0$ is trivial since $x_1 = \ldots = x_n = 0$ in this case. Let $k \leq n - 1$ and suppose that the equality in (i) holds for every $\bar{\boldsymbol{x}} \in \mathbb{R}^n_+$ with $k(\bar{\boldsymbol{x}}) = k$. Let $\boldsymbol{x} \in \mathbb{R}^n_+$ with $k(\boldsymbol{x}) = k + 1$. Then,

$$\text{MAD}\left(\sum_{i=1}^n x_i Y_i\right) = \text{MAD}\left(\sum_{i=1}^{k+1} x_i Y_i\right) = \text{MAD}\left(\sum_{i=1}^k x_i Y_i + x_{k+1} Y_{k+1}\right). \quad (38)$$

Note that we have

$$\left(\sum_{i=1}^k x_i Y_i - \mathbb{E}\left[\sum_{i=1}^k x_i Y_i\right]\right)(x_{k+1} Y_{k+1} - x_{k+1} \mathbb{E}[Y_{k+1}])$$

$$= x_{k+1} \sum_{i=1}^k x_i (Y_i - \mathbb{E}[Y_i])(Y_{k+1} - \mathbb{E}[Y_{k+1}]) \geq 0$$

thanks to (iv) and the fact that $\boldsymbol{x} \in \mathbb{R}^n_+$. Hence, by Theorem 2,

$$\text{MAD}\left(\sum_{i=1}^k x_i Y_i + x_{k+1} Y_{k+1}\right) = \text{MAD}\left(\sum_{i=1}^k x_i Y_i\right) + \text{MAD}(x_{k+1} Y_{k+1}). \quad (39)$$

On the other hand, we may write $\sum_{i=1}^k x_i Y_i = \sum_{i=1}^n \bar{x}_i Y_i$, where $\bar{\boldsymbol{x}} := (x_1, \ldots, x_k, 0, \ldots, 0)$. Since $k(\bar{\boldsymbol{x}}) = k$, we may use the induction hypothesis and obtain

$$\text{MAD}\left(\sum_{i=1}^k x_i Y_i\right) = \text{MAD}\left(\sum_{i=1}^n \bar{x}_i Y_i\right) = \sum_{i=1}^n \bar{x}_i \, \text{MAD}(Y_i) = \sum_{i=1}^k x_i \, \text{MAD}(Y_i). \quad (40)$$

Combining (38), (39), (40) and using the positive homogeneity of MAD give that

$$\text{MAD}\left(\sum_{i=1}^n x_i Y_i\right) = \sum_{i=1}^k x_i \, \text{MAD}(Y_i) + x_{k+1} \, \text{MAD}(Y_{k+1}) = \sum_{i=1}^n x_i \, \text{MAD}(Y_i),$$

which concludes the induction argument.

References

Bacon, C. A. (2008). *Practical portfolio performance measurement and attribution* (2nd ed.). Wiley.

Bai, X., Scheinberg, K., & Tutuncu, R. (2016). Least-squares approach to risk parity in portfolio selection. *Quantitative Finance, 16*(3), 357–376.

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

Best, M., & Grauer, R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies, 4*(2), 315–342.

Best, M., & Grauer, R. (1991). Sensitivity analysis for mean-variance portfolio problems. *Management Science, 37*(8), 980–989.

Boudt, K., Carl, P., & Peterson, B. (2013). Asset allocation with conditional value-at-risk budgets. *Journal of Risk, 15*, 39–68.

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2017). On exact and approximate stochastic dominance strategies for portfolio selection. *European Journal of Operational Research, 259*(1), 322–329.

Carleo, A., Cesarone, F., Gheno, A., & Ricci, J. M. (2017). Approximating exact expected utility via portfolio efficient frontiers. *Decisions in Economics and Finance, 40*(1–2), 115–143.

Cesarone, F. (2020). *Computational finance: MATLAB® oriented modeling*. Routledge.

Cesarone, F., & Colucci, S. (2018). Minimum risk versus capital and risk diversification strategies for portfolio construction. *Journal of the Operational Research Society, 69*(2), 183–200.

Cesarone, F., & Tardella, F. (2017). Equal risk bounding is better than risk parity for portfolio selection. *Journal of Global Optimization, 68*(2), 439–461.

Cesarone, F., Scozzari, A., & Tardella, F. (2015). Linear vs. quadratic portfolio selection models with hard real-world constraints. *Computational Management Science, 12*(3), 345–370.

Cesarone, F., Moretti, J., Tardella, F., et al. (2016). Optimally chosen small portfolios are better than large ones. *Economics Bulletin, 36*(4), 1876–1891.

Cesarone, F., Lampariello, L., & Sagratella, S. (2019). A risk-gain dominance maximization approach to enhanced index tracking. *Finance Research Letters, 29*, 231–238.

Cesarone, F., Mango, F., Mottura, C. D., Ricci, J. M., & Tardella, F. (2020). On the stability of portfolio selection models. *Journal of Empirical Finance, 59*, 210–234.

Cesarone, F., Scozzari, A., & Tardella, F. (2020). An optimization-diversification approach to portfolio selection. *Journal of Global Optimization, 76*(2), 245–265.

Cesarone, F., Martino, M. L., & Carleo, A. (2022). Does ESG impact really enhance portfolio profitability? *Sustainability, 14*(4), 2050.

Chopra, V., & Ziemba, W. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management, 19*(2), 6–11.

Clarke, R., De Silva, H., & Thorley, S. (2013). Risk parity, maximum diversification, and minimum variance: An analytic perspective. *The Journal of Portfolio Management, 39*(3), 39–53.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies, 22*(5), 1915–1953.

Fabozzi, F. A., Simonian, J., & Fabozzi, F. J. (2021). Risk parity: The democratization of risk in asset allocation. *The Journal of Portfolio Management, 47*, 41–50.

Fisher, G. S., Maymin, P. Z., & Maymin, Z. G. (2015). Risk parity optimality. *The Journal of Portfolio Management, 41*(2), 42–56.

Gordan, P. (1873). Über die auflösung linearer gleichungen mit reelen coefficienten (on the solution of linear inequalities with real coefficients). *Mathematische Annalen, 6*(1), 23–28.

Güler, O. (2010). *Fundamentals of optimization*. Springer.

Haesen, D., Hallerbach, W. G., Markwat, T., & Molenaar, R. (2017). Enhancing risk parity by including views. *The Journal of Investing, 26*(4), 53–68.

Jacobsen, B., & Lee, W. (2020). Risk-parity optimality even with negative Sharpe ratio assets. *The Journal of Portfolio Management, 46*(6), 110–119.

Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo Stock Market. *Management Science, 37*(5), 519–531.

Liu, B., Brzenk, P,, Cheng, T. (2020). Indexing risk parity strategies. S&P Global, S&P Dow Jones Indices, October 2020 Available at https://www.spglobal.com/spdji/en/documents/research/research-indexing-risk-parity-strategies.pdf?force_download=true (October 2020).

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

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.

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

Michaud, R., & Michaud, R. (1998). *Efficient asset management*. Harvard Business School Press.

Oderda, G. (2015). Stochastic portfolio theory optimization and the origin of rule-based investing. *Quantitative Finance, 15*(8), 1259–1266.

Qian, E. (2011). Risk parity and diversification. *The Journal of Investing, 20*(1), 119–127.

Qian, E. (2017). How naïve is naïve risk parity? Panagora Asset Management.

Rachev, S., Stoyanov, S., & Fabozzi, F. (2008). *Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures*. Wiley.

Rockafellar, R., & Wets, R. J. B. (1982). On the interchange of subdifferentiation and conditional expectation for convex functionals. *Stochastics, 7*, 173–182.

Rockafellar, R., & Wets, R. J. B. (1997). *Variational analysis*. Springer.

Rockafellar, R., Uryasev, S., & Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics, 10*(1), 51–74.

Rockafellar, R. T. (1970). *Convex analysis*. Princeton University Press.

Rockafellar, R. T. (1973). *Conjugate duality and optimization*. SIAM.

Roncalli, T. (2013). Introducing expected returns into risk parity portfolios: A new framework for tactical and strategic asset allocation. Available at SSRN: http://ssrncom/abstract=2321309

Spinu, F. (2013). An algorithm for computing risk parity weights. Available at SSRN: http://ssrncom/abstract=2297383

Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking & Finance, 26*(7), 1519–1533.

Tawarmalani, M., & Sahinidis, N. V. (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming, 103*(2), 225–249.

Yang, F., & Wei, Z. (2008). Generalized Euler identity for subdifferentials of homogeneous functions and applications. *Journal of Mathematical Analysis and Applications, 337*, 516–523.