Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento Economia e Statistica Cognetti de Martiis

Working Paper Series

08/22

KURTOSIS-BASED RISK PARITY: METHODOLOGY AND PORTFOLIO EFFECTS

MARIA DEBORA BRAGA, CONSUELO RUBINA NAVA and MARIA GRAZIA ZOIA

UNIVERSITÀ DEGLI STUDI DI TORINO

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Kurtosis-based Risk Parity: Methodology and Portfolio Effects

Maria Debora Braga[a,b], Consuelo Rubina Nava[c], Maria Grazia Zoia[d,*]

[a] *SDA Bocconi School of Management, Bocconi University, Milan, Italy*
[b] *Department of Economics and Political Sciences, University of Aosta Valley, Italy*
[c] *Department of Economics and Statistics "Cognetti de Martiis", University of Turin, Italy*
[d] *Department of Economic Policy, Università Cattolica del Sacro Cuore, Milan, Italy*

Abstract

In this paper, a risk parity strategy based on portfolio kurtosis as reference measure is introduced. This strategy allocates the asset weights in a portfolio in a manner that allows an homogeneous distribution of responsibility for portfolio returns' huge dispersion, since portfolio kurtosis puts more weight on extreme outcomes than standard deviation does. Therefore, the goal of the strategy is not the minimization of kurtosis, but rather its "fair diversification" among assets. An original closed-form expression for portfolio kurtosis is devised to set up the optimization problem for this type of risk parity strategy. The latter is then compared with the one based on standard deviation by using data from a global equity investment universe and implementing an out-of-sample analysis. The kurtosis-based risk parity strategy has interesting portfolio effects, with lights and shadows. It outperforms the traditional risk parity according to main risk-adjusted performance measures. In terms of asset allocation solutions, it provides extremely unbalanced and more erratic portfolio weights (albeit without excluding any component) in comparison to those pertaining the traditional risk parity strategy.

Keywords: Kurtosis, Risk parity, Risk diversification, Asset allocation
JEL: G110, C10, C610

*Corresponding author
Email address: maria.zoia@unicat.it (Maria Grazia Zoia)

2010 MSC: 62P20, 91-08

1. Introduction

Portfolio construction, the process of allocating wealth among different asset classes, has a long tradition in the academic literature since the seminal works by [1, 2]. For over 60 years, the approach based on the Mean-Variance Optimisation (MVO) has provided the solution to investors' asset allocation problem suggesting the portfolio that maximizes the expected return for a given volatility level or the portfolio that minimizes volatility for a given level of expected return. Despite its elegance and rationality, it is widely acknowledged that MVO suffers from several drawbacks when put into practice [3, 4, 5, 6]. The prominent problems are its excessive concentration in a limited number of the asset classes of the investment universe and the high sensitivity of the suggested asset class exposures to small changes in the estimated input parameters, most notably in expected returns. Alternative portfolio construction methodologies have been proposed in order not to run into these pitfalls. A notable contribution in this sense comes from a class of approaches, called risk-based strategies or, alternatively, μ-free strategies [7, 8], which do not require expected returns estimates as inputs. Since the global financial crisis of 2008, a version of these strategies, known as risk parity, has gained popularity [9]. The distinguishing feature of portfolios based on risk-parity strategy is that of allocating wealth among asset classes in such a way that each of them contributes to the portfolio volatility to the same extent, so they are also called "equally weighted risk contribution portfolios" (or ERC portfolios). The risk parity strategy is driven by the powerful idea of risk diversification (instead of dollars /capital diversification) emphasized by Qian first [10, 11] and reported as "true diversification". The notion of risk contribution defined by [12] forms the basis for the formal development of the risk parity strategy.

This paper aims to expand the research on risk parity with a new version of the strategy where volatility of the portfolio's return is replaced by the portfolio kurtosis as reference measure. The existing literature has already dealt with other risk measures as appropriate for risk parity, provided they are homogeneous of degree one. Precisely, interesting variations use measures of downside risk such as semi-volatility, value at risk [13] and conditional value at risk [14, 15]. Unlike these studies, the central and exclusive place given to kurtosis implies that this study is not looking for a "new metric" but instead it wants to focus on a specific feature regarding the shape of the distribution of portfolio returns. It is well known that kurtosis, depending on whether it is higher or lower than 3, indicates fatter (or thinner) tails and stronger (or weaker) "peakedness" of the distribution when compared to the normal distribution. Therefore, as indicator of distributional characteristic, kurtosis concentrates on the movement of probability mass from the shoulders of the distribution towards its tails and center. A kurtosis-based risk parity strategy, relying on fourth portfolio moment, puts more weight on extreme values/outcomes (either positive or negative) than standard deviation does. Therefore, when investors set up a portfolio by disseminating the responsibility for portfolio kurtosis equally among asset classes, they still focus on dispersion, but the huge dispersion. So, this work brings a novelty while remaining adherent to a symmetric framework and considering the entire returns distribution. In a certain sense, a psychological explanation can be suggested for choosing kurtosis as reference risk measure: simplifying, since kurtosis pays more attention to high uncertainty and captures higher probability of extremely significant changes in returns, by setting up a risk parity strategy based on portfolio kurtosis, the investor can avoid a sense of regret or escape the responsibility that could arise from taking net "bets" on kurtosis (instead of splitting it homogeneously), that would appear more daring than "bets" regarding volatility. This also becomes an economic rationale behind the design of the new version for the risk parity strategy. Investors would dislike the fat tail on the left, in contrast the fat tail on the right would be preferred. In presence of possible contradictory effects, investor's sentiment for uncertainty can lead to rejecting a speculative stance in favour, instead, of a democratic distribution of responsibility for the portfolio kurtosis.

The importance recognized to kurtosis in investment decision is well motivated although it is frequently omitted. Numerous empirical studies reveal that most financial asset returns are not normally distributed. They exhibit fatter tails than a gaussian distribution (they are leptokurtic) and asymmetry to the left, more occasionally to the right [16, 17, 18, 19, 20].

It is also worth considering the findings, based on a large set of combinations of portfolios of several sizes considered over a set of possible holding periods, that diversification does not reduce kurtosis which becomes a persistent presence [21]. Other authors [22] raised doubts about the benefit of international portfolio diversification when higher order moments are considered in determining the benefit itself.

Kurtosis is something which investors care about once distribution of returns deviates from normality. [23], while studying the investor's expected utility, identified the direction of his preference for moments including kurtosis. They concluded that usual investors dislike even central moments and, therefore, kurtosis as well, while they prefer odd central moments.

In the field of risk management, these results have motivated the use of Modified VaR by [24] to estimate Value at Risk in a way that corrects Gaussian VaR to consider skewness and kurtosis to allow a better risk monitoring and alignment with investor preferences. Furthermore, the specific direction of preference for kurtosis has encouraged several attempts to model density functions to account with flexibility of kurtosis detected in the empirical distributions [25, 26].

In previous studies on portfolio construction, the acknoweledged relevance of kurtosis was used in different ways. [27] developed a portfolio frontier resulting from the goal of minimizing kurtosis for a given expected return and skewness. Other authors [28, 29, 19, 30, 31] applied the polynomial goal programming (PGP) approach. Starting from the stated preferences for common investors, the method requires, at a first stage, to find a solution for each legitimate individual goal separately: maximize expected return, maximize skewness, minimize variance and minimize kurtosis subject to classical long-only and budget constraints. At the second stage, the PGP method needs to find a solution for an optimization problem consisting in minimizing the sum of the deviations of each individual goal from its optimal value, with each moment deviation weighted by a parameter expressing the investor's importance for that moment. An alternative approach to deal with higher moments in portfolio construction is that of maximizing expected utility when it is approximated by Taylor series expansion up to a given order, precisely the fourth order [20].

This paper aims at exploring the use of kurtosis in the portfolio construction framework represented by the risk parity. This differentiates its contribution from previous studies: the goal becomes the homogeneous distribution of portfolio kurtosis, rather than its minimization. As a related goal, the characteristics of the "Equally Weighted Kurtosis Contribution Portfolio" are investigated and compared to those of the traditional risk parity portfolio denominated "Equally Weighted Risk Contribution Portfolio" [9]. An exam of the similarity or diversity between the two strategies proves useful to gain awareness of the ability or not of the risk parity based on kurtosis to avoid the most important criticisms of the mean-variance efficient portfolios.

In developing the new risk parity strategy, the paper contributes to the existing literature with two methodological advances. First, a more tractable calculation formula for portfolio kurtosis is provided. Unlike [20, 32, 33], who used tensor matrices to compute portfolio kurtosis, a novel formula for the fourth-order moment of the portfolio returns is established, The latter, besides being easier to interpret statistically, allows to work out convenient closed-form expressions for the portfolio kurtosis. These formulas, highlighting the role played by asset kurtosis and co-kurtosis, are employed to derive analytical expressions for the marginal risk contributions to the portfolio kurtosis.

The paper implements both the traditional risk parity portfolio and the new one based on kurtosis relying on sample estimates for input parameters. It is well known the latter are sensitive to estimation errors, however it is not a task of this paper to deal with the issue of sampling errors as it does not belong the line of studies dedicated to improved and/or parsimonious estimates for financial assets and portfolio parameters [34, 35].

The new framework of the risk parity strategy and the traditional one are then applied to a global equity investment universe of 7 MSCI indices using real market data from January 2001 to December 2020. For this time interval, a first dataset that contains monthly returns is considered, then combined with a second dataset that contains weekly returns to perform a robustness check. Using these datasets, a study to investigate the portfolio effects of a strategy that aims at the "democratization" of kurtosis in asset allocation and no longer to the "democratization" of volatility [36] is carried out. To this end, a rolling sample approach like that employed by [37] is implemented. This allows the computation of the portfolio weights for the Equally Weighted Kurtosis Contribution Portfolios with different estimation window lengths and applying alternative portfolio rebalancing frequencies.

As just stated, the investingation of the behavior of the new risk parity strategy is carried out considering the asset classes as the primary "building blocks" of a portfolio as it happens traditionally and practically for many asset managers. There is awareness that, in the last decade or so, the interest for a new type of "building blocks" has emerged. For this reason, Kurtosis-based Risk Parity in a risk-factor-based asset allocation framework is a topic which surely deserves to be also investigated. Dealing with assets, and then with risk factors, corresponds to the sequential steps adopted for the traditional risk parity by [9, 38].

The empirical results for the current work reveal that a kurtosis-based risk parity strategy, compared to the classic risk parity, produces asset allocation solutions characterized by extremely unbalanced portfolio weights. The non-exclusion of any asset class does not prevent phenomena of excessive concentration. Alongside this weakness, however the novel strategy proves effective in terms of financial performance. The results suggest that the adoption of the kurtosis-based risk parity, rather than the traditional risk parity, enhances the risk-return profile of the portfolio providing higher Sharpe ratio, Sortino ratio and Omega ratio.

The remainder of this paper is organized as follows. In Section 2 the new risk parity strategy based on kurtosis is developed. To this end, original formulas for portfolio kurtosis and its gradient, which prove useful for the determination of the marginal risk contributions of the asset classes to portfolio kurtosis, are established. Section 3 describes the dataset used for the empirical application (subsection 3.1), provides an out-of-sample comparative analysis between the new and the traditional risk parity strategy (subsections 3.2 and 3.3), and performs a robustness test (subsection 3.4). Section 4 concludes the paper. Appendix A and Appendix B provides the proofs of methodological results carried out in Section 2, while Appendix C includes some graphs of the series employed in the empirical application.

2. Setting up Risk Parity based on portfolio kurtosis

As well known, the distinguishing feature of risk parity is that of recommending portfolio weights such that portfolio risk is equally distributed among asset classes. In the traditional risk parity approach, portfolio volatility serves as reference measure for risk. Consequently, the portfolio weights are identified in such a way that each portfolio component contributes equally to portfolio standard deviation. In this study, portfolio volatility is replaced with portfolio kurtosis as reference measure. Both these metrics capture dispersion. However, the fourth moment puts much more weight on extreme movements and less weight on small movements than standard deviation does. Reference to kurtosis, therefore, reflects the idea of spreading the responsibility for huge dispersion around the mean homogeneously among the asset classes in the investment universe. The aim underlying this replacement is not to assert or demonstrate the superiority of those that can be called "Equally Weighted Kurtosis Contribution Portfolios" over the traditional "Equally Weighted Risk Contribution Portfolios" [9], but investigate what they can mean in terms of different portfolio effects.

The new risk parity set-up hinges on the portfolio kurtosis, for which it is here provided a novel closed-formula which proves useful for the computation of the the marginal kurtosis contribution of each asset class exposure. The latter, which defines the sensitivity of the portfolio kurtosis to a small change on an asset class weight, must be determined together with the total kurtosis contribution. After that, portfolio kurtosis can be viewed as the sum of all total kurtosis contributions. Starting from their definition, the optimization problem for the Equally Weighted Risk Contribution Portfolio Kurtosis under the traditional budget constraint and long-only constraint can be formulated.

2.1. An original representation of portfolio kurtosis - Rethinking portfolio kurtosis and its representation

Let $\boldsymbol{R}$ be the matrix of N returns observed over T periods

$$\underset{(N,T)}{\boldsymbol{R}} = \begin{bmatrix} \underset{(1,T)}{\boldsymbol{r}_1^{'}} \\ \vdots \\ \underset{(1,T)}{\boldsymbol{r}_N^{'}} \end{bmatrix} = [\, \underset{(N,1)}{\boldsymbol{x}_1}, \ldots, \underset{(N,1)}{\boldsymbol{x}_T} \,] \tag{1}$$

where $\boldsymbol{r}_i^{'}$ is the set of T observations on the i-th return.

Dropping the pedix t, let $\boldsymbol{x}$ be the t-th column of the matrix $\boldsymbol{R}$. The portfolio p, at time t, can be written as

$$p = \boldsymbol{w}^{'}\boldsymbol{x} \tag{2}$$

where $\boldsymbol{w}$ is a N dimensional vector of non-negative weights such that $\boldsymbol{w}'\boldsymbol{u} = 1$ and $\boldsymbol{u}$ denotes a vector of $1's$. The portfolio mean, μ_p, and variance, σ_p^2, are given by

$$\mu_p = \boldsymbol{w}'\mathbb{E}[\boldsymbol{x}] \tag{3}$$

$$\sigma_p^2 = \mathbb{E}[\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}] \tag{4}$$

where $\mathbb{E}[\cdot]$ stands for expectation, and $\boldsymbol{\Sigma} = \mathbb{E}[(\boldsymbol{x} - \mathbb{E}(\boldsymbol{x}))(\boldsymbol{x} - \mathbb{E}(\boldsymbol{x}))'] = \mathbb{E}[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}']$ is the portfolio dispersion matrix. With this premise, the portfolio fourth-order moment, $\mu_{4,p}$, and kurtosis K_p, can be expressed as follows

$$\mu_{4,p} = \mathbb{E}[\boldsymbol{w}'\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\boldsymbol{w}]^2 \tag{5}$$

$$K_p = \frac{\mathbb{E}[\boldsymbol{w}'\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\boldsymbol{w}]^2}{[\mathbb{E}[\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}]]^2} = \frac{\mu_{4,p}}{(\sigma_p^2)^2} \tag{6}$$

The portfolio kurtosis can also be expressed in terms of the kurtosis (K) and co-kurtosis (CK) of single assets $x_s, s = 1, \ldots, N$

$$K_s = \frac{\mathbb{E}(\tilde{x}_s^4)}{\sigma_{x_s}^4} = \frac{\mathbb{E}[x_s - \mathbb{E}(x_s)]}{\sigma_{x_s}^4} \tag{7}$$

$$CK_{i,j,\gamma,r}(x_i^k, x_j^l, x_\gamma^g, x_r^f) = \frac{\mathbb{E}(\tilde{x}_i^k \tilde{x}_j^l \tilde{x}_\gamma^g \tilde{x}_r^f)}{\sigma_{x_i}^k \sigma_{x_j}^l \sigma_{x_\gamma}^g \sigma_{x_r}^f}, \;\; k+l+g+f = 4, \;\; k,l,g,f \in N \tag{8}$$

as follows

$$\begin{aligned} K_p = &\frac{1}{\sigma_p^4} \sum_{j=1}^{N} \Bigg(w_j^4 \sigma_{x_j}^4 K(x_j) + 4 \sum_{i=1, i\neq j}^{N} w_j^3 \sigma_{x_j}^3 w_i \sigma_{x_i} CK(x_j^3, x_i) \\ &+ 6 \sum_{i=j+1}^{N} w_j^2 \sigma_{x_j}^2 w_i^2 \sigma_{x_i}^2 CK(x_j^2, x_i^2) \\ &+ 12 \sum_{i=1, i\neq j}^{N} \sum_{r=i+1, r\neq j}^{N} w_j^2 \sigma_{x_j}^2 w_i \sigma_{x_i} w_r \sigma_{x_r} CK(x_j^2, x_i, x_r) \\ &+24 \sum_{i=j+1}^{N} \sum_{r=i+1}^{N} \sum_{\gamma=r+1}^{N} w_j \sigma_{x_j} w_i \sigma_{x_i} w_r \sigma_{x_r} w_\gamma \sigma_{x_\gamma} CK(x_j, x_i, x_r, x_\gamma) \Bigg). \end{aligned} \tag{9}$$

Eq. (9) follows from the multinomial theorem [see e.g., 39]. Note that the terms involved in the equation are composed of N, $4N(N-1)$, $6\binom{N}{2}$, $12N\binom{N-1}{2}$ and $24\binom{N}{4}$ addends respectively, for a total of N^4 elements which are the entries of the fourth order moment matrix of the returns $\boldsymbol{\Psi} = \mathbb{E}[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}']$, where $\otimes$ denotes the Kronecker product. In particular, the last two terms in Eq. (9), involving the co-kurtosis of triplets and quadruplets of the portfolio assets, $CK(x_j^2, x_i, x_r)$ and $CK(x_j, x_i, x_r, x_\gamma)$, are missing when $N \leq 2$, while the last one, $CK(x_j, x_i, x_r, x_\gamma)$, is missing when $N \leq 3$.

The following lemma provides an expression of the portfolio kurtosis which proves useful in the following.

Lemma 1. *The portfolio kurtosis has the following closed-form expression*

$$K_p = \frac{(\boldsymbol{w}' \otimes \boldsymbol{w}')\boldsymbol{\Psi}(\boldsymbol{w} \otimes \boldsymbol{w})}{(\boldsymbol{w}' \otimes \boldsymbol{w}')(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})(\boldsymbol{w} \otimes \boldsymbol{w})} \tag{10}$$

Proof. See Appendix A. □

Eq. (10), which expresses the portfolio kurtosis in terms of the $N^2 \times N^2$ fourth order moment matrix of the returns $\boldsymbol{\Psi}$, is of easier interpretability from a statistical point of view than the one which hinges on the $N \times N^3$ tensor matrix $\mathbb{E}[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}']$ as in [20, 32]. In Appendix A the equivalence of the two type of formulations is proved and it is explained how the entries of $\boldsymbol{\Psi}$ are specified in terms of the fourth-order moments of the returns.

2.2. The specification of the Equally Weighted Kurtosis Contribution Portfolios

As well known, the RP approach hinges on the idea that portfolio risk must be equally distributed among asset classes. In other words, risk parity prevents a single component or a few portfolio components from assuming a prominent role in guiding portfolio risk. This goal can be pursued for any risk measure RM, provided it is homogeneous of degree one in the weights. In such a case, the risk measure RM admits the Euler decomposition and it can be expressed as sum of contributions, say $c_i(\boldsymbol{w})$, of its constituents

$$RM = \sum_{i=1}^{N} c_i(\boldsymbol{w}) = \sum_{i=1}^{N} w_i \frac{\partial RM}{\partial w_i} = \boldsymbol{c}(\boldsymbol{w})'\boldsymbol{u}_N \tag{11}$$

where $\boldsymbol{c}(\boldsymbol{w}) = [c_i(\boldsymbol{w})] = \boldsymbol{w} * \frac{\partial RM}{\partial \boldsymbol{w}}$, with $*$ denoting the Hadamard product, and where $\boldsymbol{u}_N$ is a N-dimensional vector of 1's.

In light of Eq. (11), the risk parity portfolio can be identified by the condition

$$c(\boldsymbol{w}) = \boldsymbol{u}_N c; \quad \text{s.t.} \quad \boldsymbol{w}'\boldsymbol{u}_N = 1 \text{ and } w_i > 0\, \forall i = 1, \dots, n. \tag{12}$$

where c denotes a constant.

The portfolio standard deviation is a first order homogeneous function of the weights with marginal risks given by

$$\boldsymbol{\rho}_{\sigma_p} = \frac{\partial \sigma_p}{\partial \boldsymbol{w}} = \frac{\boldsymbol{\Sigma w}}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma w}}}. \tag{13}$$

In order to determine the marginal risk contributions to the portfolio kurtosis, we establish the following preliminary result.

Theorem 1. *The gradient of the portfolio fourth moment with respect to the weight vector* $\boldsymbol{w}$ *is given by*

$$\boldsymbol{\rho}_{\mu_{4,p}} = \frac{\partial \mu_{4,p}}{\partial \boldsymbol{w}} = 2(\boldsymbol{I}_N \otimes \boldsymbol{w}' + \boldsymbol{w}' \otimes \boldsymbol{I}_N)\boldsymbol{\Psi}(\boldsymbol{w} \otimes \boldsymbol{w}). \tag{14}$$

Proof. By the chain rule [see, e.g. 40, Sec. 10.7 p. 203] and taking into account (5), we have

$$\frac{\partial \mu_{4,p}}{\partial \boldsymbol{w}'} = \frac{\partial \mu_{4,p}}{\partial (\boldsymbol{w} \otimes \boldsymbol{w})'} \cdot \frac{\partial (\boldsymbol{w} \otimes \boldsymbol{w})}{\partial \boldsymbol{w}'} \tag{15}$$

where

$$\frac{\partial \mu_{4,p}}{\partial (\boldsymbol{w} \otimes \boldsymbol{w})'} = 2(\boldsymbol{w} \otimes \boldsymbol{w})'\boldsymbol{\Psi} \tag{16}$$

by a well-known formula of matrix differential calculus, and

$$\frac{\partial (\boldsymbol{w} \otimes \boldsymbol{w})}{\partial \boldsymbol{w}'} = \boldsymbol{K}_{1N} \otimes \boldsymbol{w} + (\boldsymbol{K}_{1N}\boldsymbol{w} \otimes \boldsymbol{I}_N) = \boldsymbol{I}_N \otimes \boldsymbol{w} + \boldsymbol{w} \otimes \boldsymbol{I}_N \tag{17}$$

by Theorem 11 in [41] with the commutation matrix, $\boldsymbol{K}_{1N}$ tallying, in this case, with the identity matrix $\boldsymbol{I}_N$. Eventually, Eq. (14) follows from Eq. (16) and (17). □

Now, we can state the following result:

Theorem 2. *The marginal risk contributions to the portfolio kurtosis with respect to the weight vector* $\boldsymbol{w}$ *are the entries of the gradient vector*

$$\boldsymbol{\rho}_{K_p} = \frac{\partial K_p}{\partial \boldsymbol{w}} = \frac{\boldsymbol{\rho}_{\mu_{4,p}}\sigma_p - 4\mu_{4,p}\boldsymbol{\rho}_{\sigma_p}}{\sigma_p^5} \tag{18}$$

with $\boldsymbol{\rho}_{\mu_{4,p}}$ *and* $\boldsymbol{\rho}_{\sigma_p}$ *given by Eq.* (14) *and* (13)*, respectively.*

Proof. Eq. (18) follows from Eq. (6) by simple differentiation rules. □

Now, let $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\Psi}}$ be sample estimates of the variance-covariance and fourth-order moment matrices of the N assets and let $\widehat{\sigma}_p^2 = \boldsymbol{w}'\hat{\boldsymbol{\Sigma}}\boldsymbol{w}$, and $\widehat{\mu}_{4,p} = (\boldsymbol{w}' \otimes \boldsymbol{w}')\widehat{\boldsymbol{\Psi}}(\boldsymbol{w} \otimes \boldsymbol{w})$ be estimates of the portfolio variance and fourth-order moment. Furthermore, let $\widehat{\boldsymbol{\rho}}_{\mu_{4,p}} = 2(\boldsymbol{I}_N \otimes \boldsymbol{w}' + \boldsymbol{w}' \otimes \boldsymbol{I}_N)\widehat{\boldsymbol{\Psi}}(\boldsymbol{w} \otimes \boldsymbol{w})$ and $\widehat{\boldsymbol{\rho}}_{\sigma_p} = \frac{\widehat{\boldsymbol{\Sigma}}\boldsymbol{w}}{\sqrt{\boldsymbol{w}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{w}}}$ be estimates of $\boldsymbol{\rho}_{\mu_{4,p}}$ and $\boldsymbol{\rho}_{\sigma_p}$, respectively.

Corollary 1. *A RP portfolio based on kurtosis satisfies the condition in Eq.* (12) *with* $\widehat{\boldsymbol{c}}(\boldsymbol{w})$ *given by*

$$\widehat{\boldsymbol{c}}(\boldsymbol{w}) = \boldsymbol{w} * \widehat{\boldsymbol{\rho}}_{K_p} \tag{19}$$

where $*$ *is the Hadamard product and*

$$\widehat{\boldsymbol{\rho}}_{K_p} = \frac{\partial \widehat{K_p}}{\partial \boldsymbol{w}} = \frac{\widehat{\boldsymbol{\rho}}_{\mu_{4,p}} \widehat{\sigma}_p - 4\widehat{\mu}_{4,p} \widehat{\boldsymbol{\rho}}_{\widehat{\sigma}_p}}{\widehat{\sigma}_p^5} \tag{20}$$

is a vector including estimates of the portfolio kurtosis marginal risks.

Proof. The proof is simple and it is omitted. □

An optimal RP strategy, based on portfolio kurtosis as risk measure, consists in finding a set of weights $\boldsymbol{w}_i$ such that the contribution of each asset class to portfolio kurtosis is, at least approximately, the same. According to this argument, the optimization problem for the new risk parity strategy can be formulated as follows:

$$\begin{aligned} \boldsymbol{w}^* &= \underset{\boldsymbol{w}}{argmin} \sum_{i=1}^{N} \sum_{j=1}^{N} (\widehat{c}_i(\boldsymbol{w}) - \widehat{c}_j(\boldsymbol{w}))^2 \\ &= \underset{\boldsymbol{w}}{argmin}\, tr[(\widehat{\boldsymbol{c}}(\boldsymbol{w})\boldsymbol{u}'_N - \boldsymbol{u}_N \widehat{\boldsymbol{c}}(\boldsymbol{w})')(\boldsymbol{u}_N \widehat{\boldsymbol{c}}(\boldsymbol{w})' - \widehat{\boldsymbol{c}}(\boldsymbol{w})\boldsymbol{u}'_N)] \\ S.T. &\quad \boldsymbol{w}'\boldsymbol{u} = 1 \\ &\quad 0 \leq w_i \leq 1. \end{aligned} \tag{21}$$

where $\widehat{c_i}(\boldsymbol{w})$ is the i-th element of $\widehat{\boldsymbol{c}}(\boldsymbol{w})$.

The optimization problem (21) solves the risk-parity problem by using the least-squares approach, as proposed by [42]. The authors also suggests the following alternative formulation

$$\begin{aligned} \boldsymbol{w}^* &= \underset{\boldsymbol{w}}{argmin} \sum_{i=1}^{N} (\widehat{c}_i(\boldsymbol{w}) - \theta)^2 \\ S.T. &\quad \boldsymbol{w}'\boldsymbol{u} = 1 \\ &\quad 0 \leq w_i \leq 1. \end{aligned} \tag{22}$$

where $\theta = \frac{\sum_i^N w_i \widehat{c}_i(\boldsymbol{w})}{N}$. They observed that the least-squares formulations (21) and (22) engender the same solutions, but the latter is less demanding computationally than the long-only risk parity via convex optimization given by

$$\begin{aligned}
\boldsymbol{w}^* = & \underset{\boldsymbol{w}}{argmin} \ \widehat{K}_p(\boldsymbol{w}) - c\sum_{i=1}^{N} ln(w_i) \\
& S.T. \boldsymbol{w}'\boldsymbol{u} = 1 \\
& w_i > 0
\end{aligned} \tag{23}$$

where $\widehat{K}_p(\boldsymbol{w}) = \frac{(\boldsymbol{w}'\otimes\boldsymbol{w}')\widehat{\boldsymbol{\Psi}}(\boldsymbol{w}\otimes\boldsymbol{w})}{(\boldsymbol{w}'\otimes\boldsymbol{w}')(\widehat{\boldsymbol{\Sigma}}\otimes\widehat{\boldsymbol{\Sigma}})(\boldsymbol{w}\otimes\boldsymbol{w})}$ with $\widehat{\boldsymbol{\Psi}}$ and $\widehat{\boldsymbol{\Sigma}}$ denoting sample estimates of $\boldsymbol{\Psi}$ and where $\boldsymbol{\Sigma}$ respectively, and c is an arbitrary constant. The use of the logarithmic barrier term in (23) was also proposed by [9], [43] and [44]. Since $\boldsymbol{\Psi}$ is positive definite and the logarithm function is strictly concave, the objective function in (23) turns out to be strictly convex. This assures the uniqueness of the solution. When the objective function of the optimization problem is non-smooth and/or with potentially many local minima, as it occurs in the problem under study, random initialization and multiple restarts proves to be suitable [45]. Therefore, we have also taken advantage of the GOSOLNP function in the R package RSOLNP.

3. Empirical applications

3.1. Data and methodology

In this section, the kurtosis-based risk parity strategy, KRP hereafter, presented in Section 2 is applied to real data in order to understand and compare its empirical implications with results obtained from the traditional standard deviation-based risk parity, SRP hereafter. The dataset considered for the scope includes seven equity indices for the period from January 2001 to December 2020, i.e. a complete 20-year time interval. All the selected indices are from Morgan Stanley Capital International and can appropriately represent the investment universe for a global equity portfolio manager. The indexes are: MSCI EMU, MSCI UK, MSCI USA, MSCI CANADA, MSCI JAPAN, MSCI PACIFIC EX-JAPAN, MSCI EMERGING MARKETS. These indices are considered in the total return version and in the euro denomination.

In the empirical analysis, monthly returns first and then weekly returns are used. Daily returns are not considered, although they can exhibit stronger kurtosis than lower-frequency ones, because they are less suitable and rarely used in practice for building medium-term or long-term strategic asset allocation solutions. The main summary statistics for the investment universe are displayed for both monthly and weekly dataset in Table 1. It provides evidence of the well-known stylized facts on negative skewness and positive excess kurtosis that are typically documented for financial assets. Not surprisingly, these characteristics are more pronounced for the weekly dataset. For all series, with the sole exception of the MSCI Japan, the Jarque-Bera test significantly rejects the normality assumption. Table 1 also shows that the returns of the dataset selected for the empirical analysis, being entirely focused on equity investments, do not exhibit particularly varied or disparate levels of volatility whatever the frequency of returns considered is. Conversely, they show a much more pronounced degree of differentiation in the levels of kurtosis, especially the weekly returns, rather than the monthly ones. Accordingly, it can be argued that an asset management company would act, in the case of adopting the traditional SRP strategy, within a market context that is indifferent to the frequency of data used to represent it. On the contrary, if the same asset management company adopted the new KRP strategy, the reference market context would change from the monthly to the weekly dataset, despite the presence of the same components.

To implement the out-of-sample study comparing KRP and SRP strategies, a rolling window estimation procedure as in [37] is employed. Specifically, given a T-period-long dataset of indexes returns, an estimation window of M-period length is selected to estimate the input parameters needed to compute the optimal portfolio weights according to the two strategies. From the above premises, in the following periods may be expressed either in months or in weeks. Then, the estimated weights are used to calculate the out-of-sample portfolio performance during the following L periods, where L depends on the rebalancing frequency applied. The process is repeated by moving the estimation window L periods forward and computing again the optimal portfolio weights. In this

Table 1: Summary statistics for the investment universe for monthly and weekly returns, respectively.

	Monthly data				
Asset	Mean	Volatitily	Skewness	Kurtosis	JB test (p-value)
MSCI EMU	0,00390	0,05208	-0,51241	4,47812	0,00000
MSCI UK	0,00249	0,04215	-0,51972	4,32250	0,00000
MSCI USA	0,00596	0,04309	-0,54427	3,72277	0,00020
MSCI CANADA	0,00574	0,05210	-0,63800	4,77753	0,00000
MSCI JAPAN	0,00312	0,04501	0,04854	3,47027	0,31572
MSCI PACIFIC EX JP F	0,00719	0,04995	-0,70870	4,71309	0,00000
MSCI EM	0,00836	0,05520	-0,49942	3,95109	0,00007
	Weekly data				
Asset	Mean	Volatitily	Skewness	Kurtosis	JB test (p-value)
MSCI EMU	0,00101	0,02905	-0,88397	9,47673	0,00000
MSCI UK	0,00075	0,02725	-0,73907	12,75220	0,00000
MSCI USA	0,00149	0,02608	-0,45126	6,71352	0,00000
MSCI CANADA	0,00144	0,02911	-0,77275	10,14544	0,00000
MSCI JAPAN	0,00082	0,02640	-0,15708	5,45686	0,00000
MSCI PACIFIC EX JP F	0,00173	0,02705	-0,80730	10,49454	0,00000
MSCI EM	0,00194	0,02836	-0,19367	8,46072	0,00000

way, each estimation window adds the returns for the next L periods and drops the equivalent number of earliest returns. The process is continued repeatedly until the end of the dataset is reached. The valuable outcome of this approach is twofold: the computation of a series of $T - M$ out-of-sample returns for each of the risk parity strategies considered and the implementation of $((T - M)/L) + 1$ asset allocation experiments under realistic conditions, i.e. without relying on forward-looking information.

The two strategies, KRP and SRP, are assessed along two directions. First, the asset allocations obtained from the two strategies are investigated by examining their levels of concentration/diversification through the Shannon Entropy measure. Second, the out-of-sample financial efficiency of the two strategies is assessed by using some risk-adjusted performance measures, including metrics that are not conditioned by the assumption of normality of the distribution of returns.

From the computational point of view, the proposed empirical applications have been conducted using an own written code in R.

3.2. Kurtosis-based risk parity at work

The empirical analysis first focuses on the monthly dataset. The KRP, introduced in Section 2, has been implemented by employing alternative estimation windows (60 and 120 months) and alternative rebalancing frequencies (6 and 3 months). This allows to carry out a variable number of asset allocation experiments, precisely equal to 31 (with an estimation window of 60 months and semi-annual rebalancing frequency), 21 (with and estimation window of 120 months and semi-annual rebalancing frequency), 61 (with an estimation window of 60 months and a quarterly rebalancing frequency) and 41 (with and estimation window of 120 months and a quarterly rebalancing frequency). Portfolio weights obtained following the KRP strategy and the SRP strategy for the above mentioned combinations of estimation window and rebalancing frequency are represented, respectively, in: Figure 1-left panel, Figure 1-right panel, Figure 2-left panel and Figure 2-right panel. It is worth noting that the weights, especially in the KRP case, even if small and difficult to grasp from the graphs, are never equal to zero.

The results clearly show an important difference between KRP and SRP. Although both risk parity strategies do not exclude any asset class from the portfolio composition, they appear significantly different in the way portfolio weights are distributed within the investable universe. In the case of the KRP strategy, portfolio weights are extremely unbalanced. In particular, only one or two asset classes play a visible dominant role in each optimized portfolio, while the remaining asset classes are rarely given a weight above one per cent. This findings suggest that implementing KRP in practice can be more difficult than implementing the traditional SRP, although the investment universe is made up of asset classes of the same nature, particularly for discretionary mandates rather than investment funds. In addition, the upper section of Figures 1 and 2 shows moments of violent substitution of the dominant asset classes with others. In such circumstances, an increase in transaction costs can obviously arise, at least as regards proportional transaction costs. The different behavior of the two risk parity strategies in terms of weight allocation to the asset classes also emerges from Figure 3, which provides the dynamics of the Shannon Entropy measure across the asset allocation experiments. It moves very close to the maximum admissible value in the case of the SRP strategy, while for the new KRP strategy it is close to zero for more than 50% of the asset allocation experiments that use a 60-month estimation window and for almost 50 per cent of the asset allocation experiments which make use of a 120-month estimation window.

An explanation of the remarkable heterogeneity in the distribution of portfolio weights characterizing the KRP strategy can be offered by Eq. (9). According to this formula, within each estimation window, the contribution of the $j-th$ asset class's returns to portfolio kurtosis can be expressed as

$$K_{p,x_j} = \frac{1}{\sigma_p^4} \sum_{i=0}^{3} C_{K_{p,x_j},i} \, ; \qquad C_{K_{p,x_j},i} = \delta_i w_j^{4-i} \sigma_{x_j}^{4-i} f_j \tag{24}$$

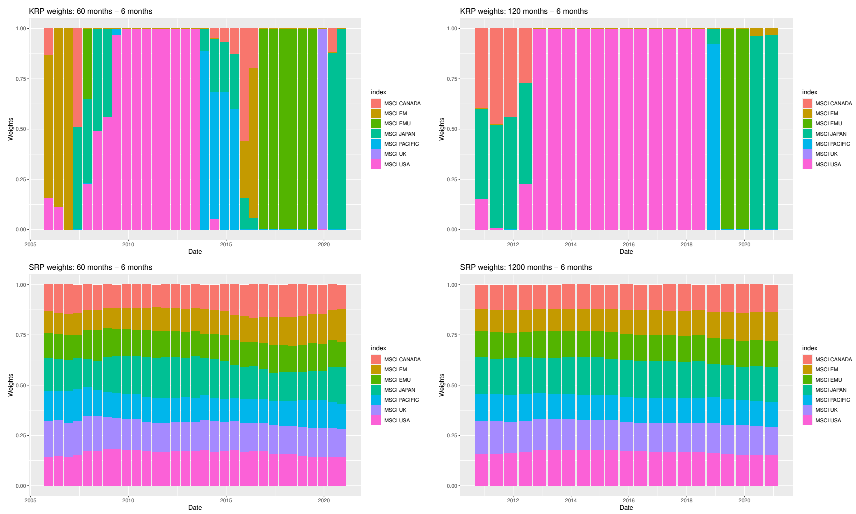

Figure 1: Bar-charts representing the weights worked out using the KRP (top panels) and the SRP strategy (bottom panels). The left panels refer to an estimation windows of 60 months while the right ones to 120 months. The rebalancing frequency is of 6 months.

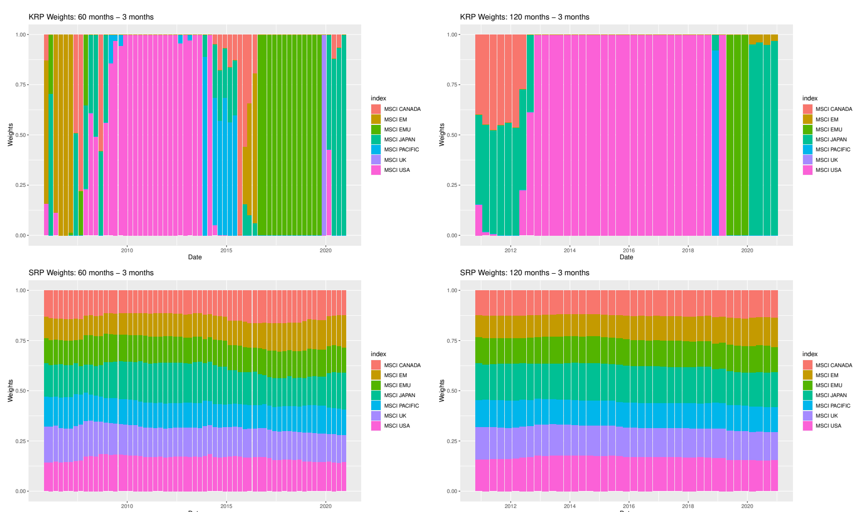

Figure 2: Bar-charts representing the weights worked out using the KRP (top panels) and the SRP strategy (bottom panels). The left panels refer to an estimation windows of 60 months while the right ones to 120 months. The rebalancing frequency is of 3 months.

Figure 3: Entropy of the portfolio weights when the SRP and KRP strategies are applied to monthly data, using different estimation windows and rebalancing frequencies

where δ_i are constant and the terms f_j depend on the kurtosis, and co-kurtoses of the $j-th$ asset class's returns (properly weighted with (powers) of the standard deviations and weights of the other asset classes). As highlighted in Eq. (24), the greater the term K_{p,x_j}, the higher the contribution of the asset class j to portfolio kurtosis and the lower its associated weight determined in accordance to a kurtosis-based risk parity. To provide evidences of this argument, the contributions of the different asset classes $K_{p,x_j}, j = 1, \ldots, 7$, have been computed, for each pair of estimation window length and rebalancing frequency They are reported in Table 2 together with the corresponding weight in the portfolio allocation. Looking at Table 2 it is worth noting that the lower the contribution K_{p,x_j} of an asset class to portfolio kurtosis, the higher its associated portfolio weight. Thus, Eq. (9) and Eq. (24), examined together, have the merit of highlighting how optimization within the KRP strategy is much harder than optimization within the traditional SRP. The number of relevant interactions among asset classes is so high that a decision making process consistent with the new KRP strategy turns out to be computationally intensive.

3.3. Comparative analysis of out-of sample results

This section analyzes comparatively the behavior of the risk parity strategies using time series of monthly out-of-sample returns generated by each of them as starting point. This choice reflects reality, as it avoids considering an optimization framework where the portfolio manager has perfect forecasts for the expected returns of all asset classes in the investment universe.

First, the traditional and new risk parity strategy are compared in terms of second and fourth statistical moments of the returns. Table 3 reports the portfolio volatility and the portfolio kurtosis for each combination of estimation window length and rebalancing frequency considered. The main and unequivocal observation is that systematically the KRP strategy exhibits lower kurtosis of its out-of-sample returns while it is the SRP strategy that systematically shows a lower standard deviation of out-of-sample returns. Interestingly, both strategies, despite the effect of unmanaged estimation errors as out-of-sample

Table 2: Asset contributions $K_{p,x}$ to portfolio kurtosis and asset class weights. In bold higher weights corresponding to assets with lower C_{K_p}

	60 months - 6 months		**120 months - 6 months**	
Asset	$K_{p,x}$	Weights	$K_{p,x}$	Weights
	EW - 4		EW - 3	
MSCI EMU	0,00770	1,62E-09	0,00858	3,3698E-09
MSCI UK	0,00770	1,23E-09	0,00858	1,1312E-08
MSCI USA	0,00770	7,28E-09	0,00858	8,1281E-08
MSCI CANADA	**0,00214**	**0,49081**	**0,00321**	**0,43920658**
MSCI JAPAN	**0,00192**	**0,50919**	**0,00202**	**0,56079332**
MSCI PACIFIC EX JP F	0,00770	1,4E-10	0,00858	3,5941E-10
MSCI EM	0,00770	7,41E-09	0,00858	2,307E-09
Asset	$K_{p,x}$	Weights	$K_{p,x}$	Weights
	EW - 5		EW - 4	
MSCI EMU	**0,00152**	**0,35361**	0,00511	1,5273E-10
MSCI UK	0,00341	1,1E-10	0,00511	1,061E-08
MSCI USA	**0,00190**	**0,23023**	**0,00348**	**0,22566061**
MSCI CANADA	0,00331	6,65E-08	**0,00301**	**0,27154602**
MSCI JAPAN	**0,00172**	**0,41617**	**0,00122**	**0,50279335**
MSCI PACIFIC EX JP F	0,00307	1,61E-09	0,00511	5,1079E-09
MSCI EM	0,00307	3,2E-09	0,00511	1,1591E-09
	60 months - 3 months		**120 months - 3 months**	
Asset	$K_{p,x}$	Weights	$K_{p,x}$	Weights
	EW - 2		EW - 6	
MSCI EMU	**0,00622**	**0,2957**	0,00854	4,9423E-10
MSCI UK	0,00672	8,35E-08	0,00854	4,7294E-09
MSCI USA	0,00672	1,44E-07	0,00854	1,2544E-07
MSCI CANADA	0,00672	2,44E-08	**0,00289**	**0,46318896**
MSCI JAPAN	**0,00327**	**0,7043**	**0,00217**	**0,5368109**
MSCI PACIFIC EX JP F	0,00672	1,38E-10	0,00854	1,0173E-10
MSCI EM	0,00672	6,13E-09	0,00854	6,4262E-09
Asset	$K_{p,x}$	Weights	$K_{p,x}$	Weights
	EW - 3		EW - 7	
MSCI EMU	0,01664	6,05E-08	0,00526	1,0166E-09
MSCI UK	0,01664	6,22E-08	0,00526	1,2113E-08
MSCI USA	**0,01374**	**0,113221**	**0,00359**	**0,22566081**
MSCI CANADA	0,01664	2,41E-08	**0,00311**	**0,27154585**
MSCI JAPAN	0,01664	5,41E-09	**0,00125**	**0,50279332**
MSCI PACIFIC EX JP F	0,01664	3,75E-10	0,00526	6,2523E-09
MSCI EM	**0,00458**	**0,886779**	0,00526	2,4183E-09

returns are considered, confirm they perform better where one would expect by construction.

Second, the financial efficiency of the SRP and the KRP strategy are evaluated by directly comparing different risk-adjusted performance measures. Initially, the Sharpe ratio is computed for each pair of estimation window and rebalancing frequency. The Barclays 3 month Euribor Cash Index is used as as risk-free rate. Results, presented in Table 4, reveal that the new KRP strategy persistently achieves better return above the risk-free returns for unit of risk. This is mainly attributable to its visibly higher mean return, since the volatility is less favourable for this strategy. In this regard, it is worth noting that the difference of the Sharpe Ratios of the two strategies is statistically significant whent the 120-month estimation window is used. The test is based on a studentized bootstrap approach with [46]'s block bootstrap [47]. In addition to the Sharpe ratio, alternative risk-adjusted performance measures, such as the Sortino Ratio and the Omega measure are computed, always in an annualized version. They are appropriate because they do not require the normal distribution assumption of out-sample returns which is rejected, for both strategies, by the Jarque Bera test (Table 5). As well known, Sortino ratio modifies the Sharpe ratio using the downside risk as denominator and considering the mean return over a chosen minimum acceptable return, MAR, as excess return. With the Omega measure, the out-of-sample returns are partitioned into losses and gains compared to a minimum acceptable return and then the probability weighted ratio of returns above and below this threshold is considered.The minimum acceptable return is identified with the risk-free rate .

The results for the Sortino ratio and the Omega measure, also represented in Table 4, are largely equivalent to those based on the Sharpe ratio. This implies that the KRP strategy exhibits systematically an improvement in the financial efficiency relative to a SRP strategy. Accordingly, it results that a portfolio strategy based on the equal distribution of responsibility for portfolio kurtosis does not lead to a penalty on the financial side compared to a more traditional risk parity approach.

Table 3: Out of sample portfolio kurtosis and volatility for the KRP and SRP strategies applied to monthly data, using different estimation windows and rebalancing frequencies

Estimation window	Rebalancing frequency	Risk Parity Method	Out of sample Portfolio Kurtosis	Out of sample Portfolio Volatility
60	6	KRP	4,506	0,041
		SRP	5,185	0,040
120	6	KRP	5,902	0,040
		SRP	6,366	0,036
60	3	KRP	4,494	0,043
		SRP	5,186	0,040
120	3	KRP	6,338	0,038
		SRP	6,344	0,036

Table 4: Sharpe, Omega and Sortino ratios for monthly data obtained with the RKP and SRP strategies by using different estimation windows and rebalancing frequencies (* indicate a statistically significant difference between KRP and SRP Sharpe ratio with $\alpha = 0.1$)

Estimation window	Rebalancing frequency	Risk Parity strategy	Sharpe ratio	Omega ratio	Sortino ratio
60	6	KRP	0,544	1,506	0,765
		SRP	0,412	1,384	0,569
60	3	KRP	0,505	1,465	0,699
		SRP	0,409	1,381	0,565
120	6	KRP	0,722*	1,767	1,095
		SRP	0,652*	1,679	0,953
120	3	KRP	0,672*	1,718	1,011
		SRP	0,651*	1,677	0,951

Table 5: Juarque-Bera test of monthly data obtained with the KRP and SRP strategies by using different estimation windows and rebalancing frequencies

Estimation window	Rebalancing frequency	Risk Parity method	JB test Statistic	JB test p-value
60	6	KRP	38,75	$<$0.0001
		SRP	49,82	$<$0.0001
120	6	KRP	51,37	$<$0.0001
		SRP	69,58	$<$0.0001
60	3	KRP	41,55	$<$0.0001
		SRP	49,98	$<$0.0001
120	3	KRP	48,20	$<$0.0001
		SRP	68,48	$<$0.0001

3.4. Robustness check with weekly data

In this section, the results of additional analyses aimed at testing the robustness of the previous findings are discussed. The portfolio effects and the performance of KRP and SRP strategies using the weekly frequency for the returns of the same investment universe are examined.

The shorter holding period implied by these returns makes it reasonable to identify portfolio solutions for investors with a shorter investment horizon and willing to make more frequent rebalancing. Therefore, estimation windows of different length are considered: 5 years (60 months), 10 years (120 months), 1 year (52 weeks) and 2 years (104 weeks). Then, portfolios with monthly (i.e. every 4 weeks) and bi-monthly (i.e. every 8 weeks) rebalancing, and no longer with quarterly or semi annual rebalancing, are computed. Figures 4 and 5 show the results for the alternative estimation windows and alternative rebalancing frequencies considered. They confirm the base case results obtained with the monthly dataset. Once again, a much more unbalanced and erratic portfolio structure emerges from the application of the KRP strategy. The Shannon Entropy measure again shows a different behavior depending on the risk parity strategy adopted. Figure 6 shows that the index keeps close to the maximum value (close to 2) across the asset allocation experiments when the SRP strategy is adopted, while in case of use of the KRP strategy it shows a very unstable and oscillating behavior, also including numerous cases in which it is very close to zero. Table 6, through the consideration of some asset allocation experiments, provides evidence that suggests exposures to asset classes within the KRP-based portfolios are inversely linked to their contribution to the portfolio kurtosis, just as was the case for the monthly dataset.

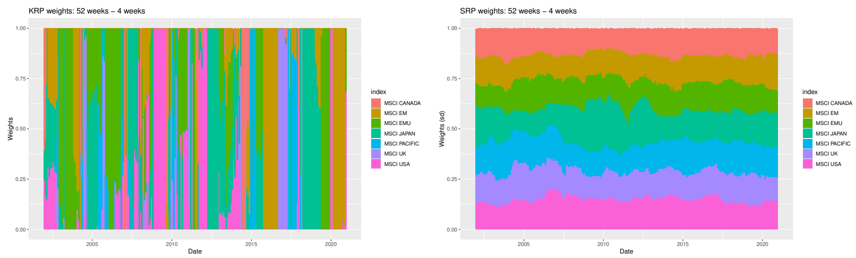

Figure 4: Bar-charts representing the weights worked out using the KRP (left panels) and the SRP strategy (right panels). The estimation windows is 52 weeks long and the rebalancing frequency is of 4 weeks.

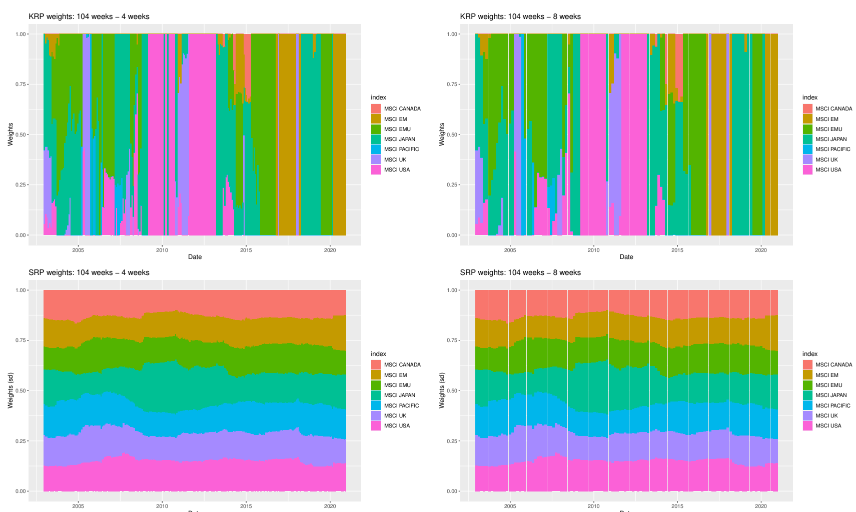

Figure 5: Bar-charts representing the weights worked out using the KRP (top panels) and the SRP strategy (bottom panels). The estimation windows is 104 weeks long. The left panels refer to a rebalancing frequency of 4 week, while the right ones to 8 weeks.

Table 6: Asset contributions C_{K_p} to portfolio kurtosis and asset class weights. In bold higher weights corresponding to assets with lower $C_{K_p,j}$

	52 weeks - 4 weeks		**104 weeks - 4 weeks**		**104 weeks - 8 weeks**	
Asset	C_{K_p}	Weights	C_{K_p}	Weights	C_{K_p}	Weights
	EW - 2		EW - 30		EW - 3	
MSCI EMU	0,00527	1,3487E-08	**0,00247**	**0,56205**	0,00436	1,9602E-08
MSCI UK	0,00527	1,04539E-10	0,00845	1,186E-08	**0,00274**	**0,32691**
MSCI USA	**0,00408**	**0,16978**	0,00845	7,8023E-07	**0,00415**	**0,05855**
MSCI CANADA	**0,00106**	**0,54854**	0,00845	1,6348E-09	0,00461	4,4547E-09
MSCI JAPAN	**0,00317**	**0,28168**	0,00845	8,2236E-08	**0,00156**	**0,50781**
MSCI PACIFIC EX JP F	0,00527	1,28801E-10	**0,00533**	**0,43795**	0,00485	3,9509E-09
MSCI EM	0,00527	2,49953E-08	0,00845	6,4183E-10	**0,00395**	**0,10673**
Asset	C_{K_p}	Weights	C_{K_p}	Weights	C_{K_p}	Weights
	EW - 3		EW - 31		EW - 4	
MSCI EMU	**0,00392**	**0,27927**	0,01377	2,1895E-08	**0,00173**	**0,36005**
MSCI UK	0,00472	8,49275E-10	**0,00246**	**0,58624**	0,00307	3,2395E-09
MSCI USA	**0,003717**	**0,15058**	**0,00489**	**0,41376**	**0,00209**	**0,16164**
MSCI CANADA	0,00469	4,29311E-09	0,01598	7,6962E-10	0,00292	1,5635E-08
MSCI JAPAN	**0,00168**	**0,57015**	0,01598	3,4095E-08	**0,00170**	**0,38974**
MSCI PACIFIC EX JP F	0,00461	7,83091E-10	0,01598	1,7021E-07	0,00300	1,149E-08
MSCI EM	0,00461	4,39269E-08	0,01598	2,1924E-08	**0,00241**	**0,08857**

As done before in Section 3.3, the relevant characteristics of the out-of-sample returns provided by the two risk parity strategies are assessed. As expected, they don't follow a normal distribution: the null hypothesis of the Jarque-Bera test is always rejected (Table 9.). The results shown in Table 7 for the riskiness of the two strategies are clear: regardless of the estimation window and the rebalancing frequency considered, the KRP strategy systematically implies a lower kurtosis of out-of-sample returns. We can afford to say that the discrepancy between the out of sample kurtosis relating to the KRP strategy and that relating to the SRP strategy becomes more visible than it was when the monthly dataset was used. This motivates the impression that the "democratization" of kurtosis also helps its mitigation at least in comparison with a strategy that neglects attention for this statistical moment when it is more pronounced.

From a financial point of view, the risk-adjusted performance of the competing risk parity strategies can be measured again by the Sharpe ratio, the Sortino ratio and the Omega ratio. Table 8 reports the results obtained using the weekly out-of-sample returns. Looking at it, the conclusion is that the KRP

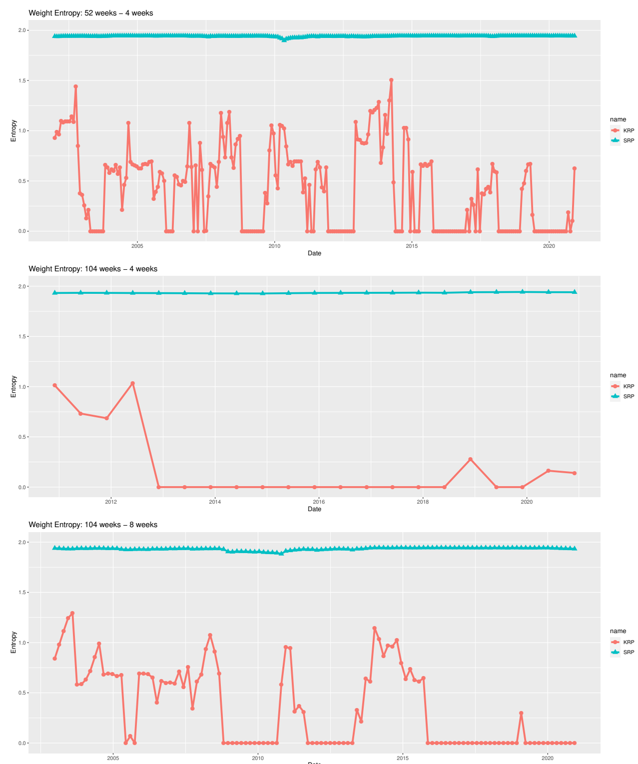

Figure 6: Entropy of the portfolio weights when the KRP and SRP strategy are applied to weekly data, using different estimation windows and rebalancing frequencies.

strategy provides persistently better reward per unit of risk taken, whatever the way the risk is interpreted. Indeed, KRP strategy produces higher Sharpe ratios, Sortino ratios and Omega ratios. In sum, the results previously obtained for the KRP strategy also hold for alternative data frequencies, alternative estimation windows and alternative rebalancing frequencies. Furthermore, with the weekly dataset, the test for differences in the Sharpe Ratio of the two strategies is statistically significant for any combination of estimation window and rebalancing frequency considered.

Table 7: Out of sample portfolio kurtosis and volatility for the KRP and SRP strategies applied to weekly data, using different estimation windows and rebalancing frequencies

Estimation window	Rebalancing frequency	Risk Parity Method	Out of sample Portfolio Kurtosis	Volatility
52	4	KRP	8,209	0,025
		SRP	10,850	0,023
104	4	KRP	9,971	0,023
		SRP	11,552	0,023
104	8	KRP	10,215	0,024
		SRP	11,594	0,023

Table 8: Sharpe, Omega and Sortino ratios for weekly data obtained with the KRP and SRP strategies by using different estimation windows and rebalancing frequencies (* indicate a statistically significant difference between KRP and SRP Sharpe ratio with $\alpha = 0.1$)

Estimation window	Rebalancing frequency	Risk Parity strategy	Sharpe ratio	Omega ratio	Sortino ratio
52	4	KRP	0,349*	1,145	0,476
		SRP	0,220*	1,087	0.301
104	4	KRP	0,490*	1,211	0,670
		SRP	0,411*	1,168	0,566
104	8	KRP	0,490*	1,211	0,670
		SRP	0,449*	1,185	0,621

4. Conclusion

There exists a wide literature illustrating the drawbacks of the well-known Mean-Variance Optimization when put into practice. Consequently, considerable effort has been devoted both to improve the Markowitz model and to seek

Table 9: Juarque-Bera test of weekly out-of-sample returns obtained with the KRP and SRP strategies by using different estimation windows and rebalancing frequencies

Estimation window	Rebalancing frequency	Risk Parity method	Out of sample JB statistic	p-value
52	4	KRP	1.204,76	<0.0001
		SRP	2.580,42	<0.0001
104	4	KRP	2.012,39	<0.0001
		SRP	3.022,33	<0.0001
104	8	KRP	2.151,84	<0.0001
		SRP	3.050,55	<0.0001

alternative portfolio construction methodologies. The latter include the risk-based strategies and, therefore, also the risk parity strategy on which this contribution is focused. Precisely, this paper provides a new version of the strategy where portfolio volatility is replaced by portfolio kurtosis as a reference measure. The goal of the strategy, called KRP (Kurtosis-based Risk Parity), is to identify portfolio weights in such a way that its constituents contribute homogeneously to portfolio kurtosis, which is different from an objective of kurtosis minimization found in already existing works. In the effort to set up the new KRP, an original closed-form expression for portfolio kurtosis and an equally original decomposition of the same in its marginal and total kurtosis contributions have been established. The new strategy is implemented, and compared with the traditional standard deviation-based risk parity, using real data with both monthly and weekly frequency, from January 2001 to December 2020, in relation to a global equity investment universe. In order to appreciate potential similarities or differentiations of the two strategies under realistic conditions in terms of portfolio implications, out-of-sample analyses exploiting the rolling window estimation procedure described by [37] have been implemented.

The empirical results reveal lights and shadows of the KRP strategy.The portfolio allocations driven by an idea of "democratization" of kurtosis are much more unbalanced and erratic than those originated by an idea of "democratization" of volatility, although in both cases no asset class is excluded from the portfolio weights vector. This finding suggests that implementing KRP strategy in practice can be more difficult than implementing the traditional risk parity. Nor does it allow us to consider ourselves completely protected from a criticism of portfolio structures commonly attributed to Markowitz's Mean-Variance Optimization, i.e. their unstable behavior. However, the empirical results are very appealing from a financial point of view. Using different risk adjusted performance measures, we show that "Equally Weighted Kurtosis Contribution Portfolios" resulting from the KRP strategy have a better risk-return profile than portfolios deriving from the classic risk parity.

This work allows for future extensions. An interesting development could be to test the KRP strategy with an effort for the issue of handling estimation errors of higher-order co-moments. However, the most obvious extension for this work consists in performing the proposed new risk parity strategy using risk factors, instead of asset classes, as building blocks of an asset allocation solution, as proposed by [48] and [49].

Data Availability

The data that support the findings of this study are available from the corresponding author, Prof. Maria Grazia Zoia, upon reasonable request. Restrictions apply to the availability of these data, which were used under license for this study.

Appendix A. Proof of Lemma 1

Eq. (10) can be worked out from Eq. (6) with simple algebra as follows

$$\mu_{4,p} = E(\boldsymbol{w}'\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\boldsymbol{w})^2 = E[vec(\boldsymbol{w}'\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\boldsymbol{w}\boldsymbol{w}'\boldsymbol{x}\boldsymbol{x}'\boldsymbol{w})] = E[(\boldsymbol{w}'\otimes\boldsymbol{w}')vec(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\boldsymbol{w}\boldsymbol{w}'\boldsymbol{x}\boldsymbol{x}')]$$

$$= E[(\boldsymbol{w}'\otimes\boldsymbol{w}')(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\otimes(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')vec(\boldsymbol{w}\boldsymbol{w}')] = E[(\boldsymbol{w}'\otimes\boldsymbol{w}')(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\otimes(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')(\boldsymbol{w}\otimes\boldsymbol{w})] = \tag{A.1}$$

where $\boldsymbol{\Psi} = (\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}'\otimes(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')$ and use has been done of the formula linking the vec and the Kronecker operator

$$vec\boldsymbol{A}\boldsymbol{B}\boldsymbol{C} = (\boldsymbol{C}'\otimes\boldsymbol{A})vec(\boldsymbol{B}) \tag{A.2}$$

which holds for matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ which are conformable for the product.
In the same manner, it can be worked out the expression $(\boldsymbol{w}'\otimes\boldsymbol{w}')(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})(\boldsymbol{w}\otimes\boldsymbol{w})$ for the square of the portfolio variance.

$$\sigma_p^2 = E[(\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w})^2] = E[vec(\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w})] = E[(\boldsymbol{w}'\otimes\boldsymbol{w}')vec(\boldsymbol{\Sigma}\boldsymbol{w}\boldsymbol{w}'\boldsymbol{\Sigma})]$$

$$= E[(\boldsymbol{w}'\otimes\boldsymbol{w}')(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})vec(\boldsymbol{w}\boldsymbol{w}')] = E[(\boldsymbol{w}'\otimes\boldsymbol{w}')(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})(\boldsymbol{w}\otimes\boldsymbol{w}')] \tag{A.3}$$

Specification of $\boldsymbol{\Psi}$

The entries, $\psi_{i,j}$, of the matrix $\boldsymbol{\Psi}$, can be expressed in terms of the fourth-order moments $\mathbb{E}(\tilde{x}_\gamma\tilde{x}_p x_k\tilde{x}_l)$ of the N assets $\tilde{x}_f = \mathbb{E}(x_f - \mathbb{E}(x_f))$ with $f = 1, 2, \ldots, N$, as follows

$$\psi_{i,j} = \mathbb{E}(\tilde{x}_\gamma\tilde{x}_p\tilde{x}_k\tilde{x}_l) \tag{A.4}$$

where

$$\gamma = \begin{cases} 1 & \text{if } 1 \le i \le N \\ 2 & \text{if } N+1 \le i \le 2N \\ \cdots & \\ N & \text{if } (N-1)N+1 \le i \le N^2 \end{cases} \qquad k = \begin{cases} 1 & \text{if } i = 1+g \\ 2 & \text{if } i = 2+g \\ \cdots & \\ N & \text{if } i = N+g \end{cases}, \tag{A.5}$$

where $g = 0, N, \ldots, (N-1)N$. Similarly, the indexes p and l are defined as γ and k, with the row-index i replaced by the row-column j

$$p = \begin{cases} 1 & \text{if } 1 \leq j \leq N \\ 2 & \text{if } N+1 \leq j \leq 2N \\ \cdots & \\ N & \text{if } (N-1)N+1 \leq N^2 \end{cases} \qquad l = \begin{cases} 1 & \text{if } j = 1+g \\ 2 & \text{if } j = 2+g \\ \cdots & \\ N & \text{if } j = N+g \end{cases}. \tag{A.6}$$

For instance, if $N = 3$, the entry $\psi_{5,7} = \mathbb{E}(\tilde{x}_2\tilde{x}_3\tilde{x}_2\tilde{x}_1) = \mathbb{E}(\tilde{x}_1\tilde{x}_2^2\tilde{x}_3)$ and can be estimated by $\frac{1}{T-1}\sum_{i=1}^{T}(\tilde{x}_1\tilde{x}_2^2\tilde{x}_3)$ where T denotes the sample size.

Appendix B. Equivalence of the fourth order moment $\boldsymbol{\mu_p}$ with the formula based on tensors

The representation of portfolio kurtosis based on tensor rests on the following portfolio fourth moment representation

$$\mu_{p,tensor} = \boldsymbol{w}'\boldsymbol{\Theta}(\boldsymbol{w} \otimes \boldsymbol{w} \otimes \boldsymbol{w}) \tag{B.1}$$

with $\boldsymbol{\Theta} = E(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}')$.

The equivalence of $\mu_{p,tensor}$ with

$$\mu_p = (\boldsymbol{w}' \otimes \boldsymbol{w}')\boldsymbol{\Psi}(\boldsymbol{w} \otimes \boldsymbol{w}) \tag{B.2}$$

with $\boldsymbol{\Psi} = E(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')$ can be proved as follows. Taking the vec of $\mu_{p,tensor}$ and μ_p yields

$$vec(\boldsymbol{w}'\boldsymbol{\Theta}(\boldsymbol{w} \otimes \boldsymbol{w} \otimes \boldsymbol{w})) = (\boldsymbol{w}' \otimes \boldsymbol{w}' \otimes \boldsymbol{w}' \otimes \boldsymbol{w}')vec\boldsymbol{\Theta}, \tag{B.3}$$

$$vec(\boldsymbol{w}'\boldsymbol{\Psi}(\boldsymbol{w} \otimes \boldsymbol{w} \otimes \boldsymbol{w})) = (\boldsymbol{w}' \otimes \boldsymbol{w}' \otimes \boldsymbol{w}' \otimes \boldsymbol{w}')vec\boldsymbol{\Psi} \tag{B.4}$$

As for $vec\boldsymbol{\Theta}$ and $vec\boldsymbol{\Psi}$, by using result by [50], some computations prove that

$$\begin{aligned} vec\boldsymbol{\Theta} &= (\boldsymbol{I}_N \otimes \boldsymbol{K}_{N^2,N})[vec(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}') \otimes vec(\tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}')] \\ vec\boldsymbol{\Psi} &= (\boldsymbol{I}_N \otimes \boldsymbol{K}_{N,N} \otimes \boldsymbol{I}_N)[vec(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}') \otimes vec(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')] \end{aligned} \tag{B.5}$$

and the right-hand sides of the above formulas are the same upon noting that $(\boldsymbol{I}_N \otimes \boldsymbol{K}_{N^2,N}) = (\boldsymbol{I}_N \otimes \boldsymbol{K}_{N,N} \otimes \boldsymbol{I}_N)$ and that $vec(\tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}') = \tilde{\boldsymbol{x}} \otimes \tilde{\boldsymbol{x}} = vec(\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}')$, as some computations prove.

Appendix C. Plots

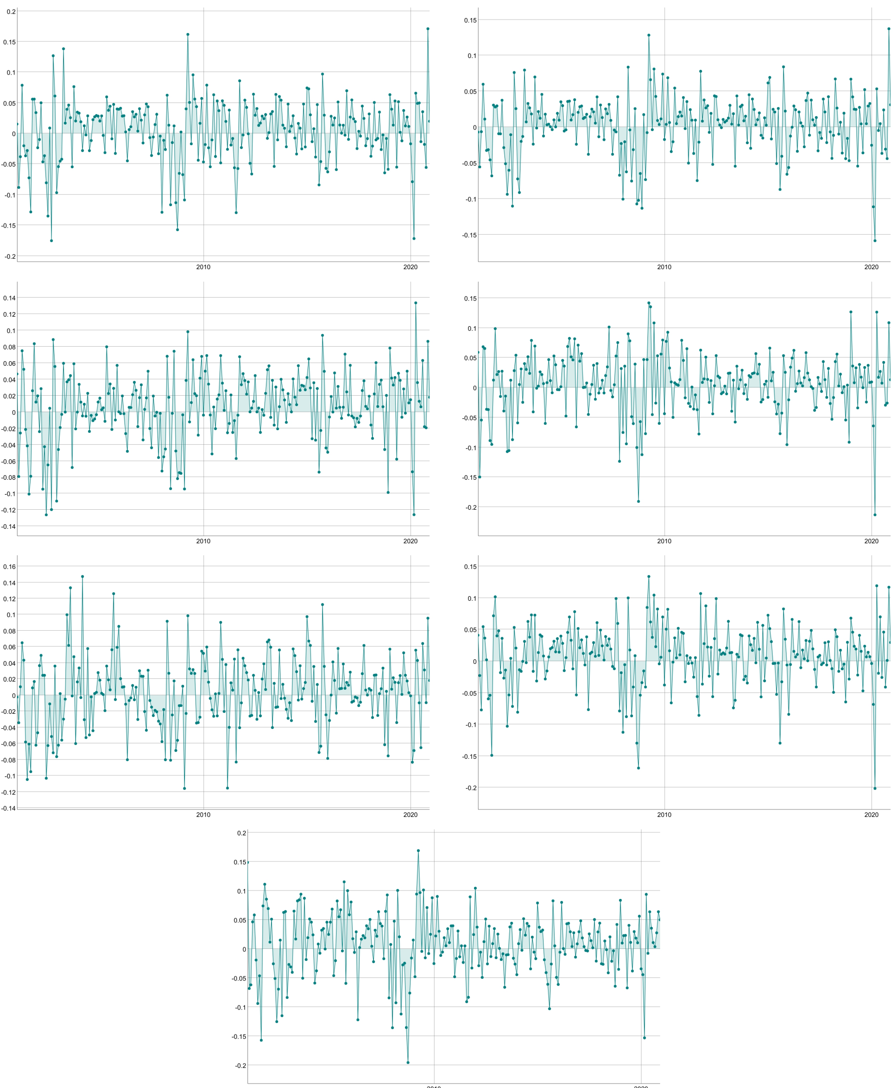

Figure C.1: Monthly total returns of MSCI EMU, MSCI UK, MSCI USA, MSCI CANADA, MSCI JAPAN, MSCI PACIFIC EX JAPAN, MSCI EMERGING MARKETS, respectively from left to right.

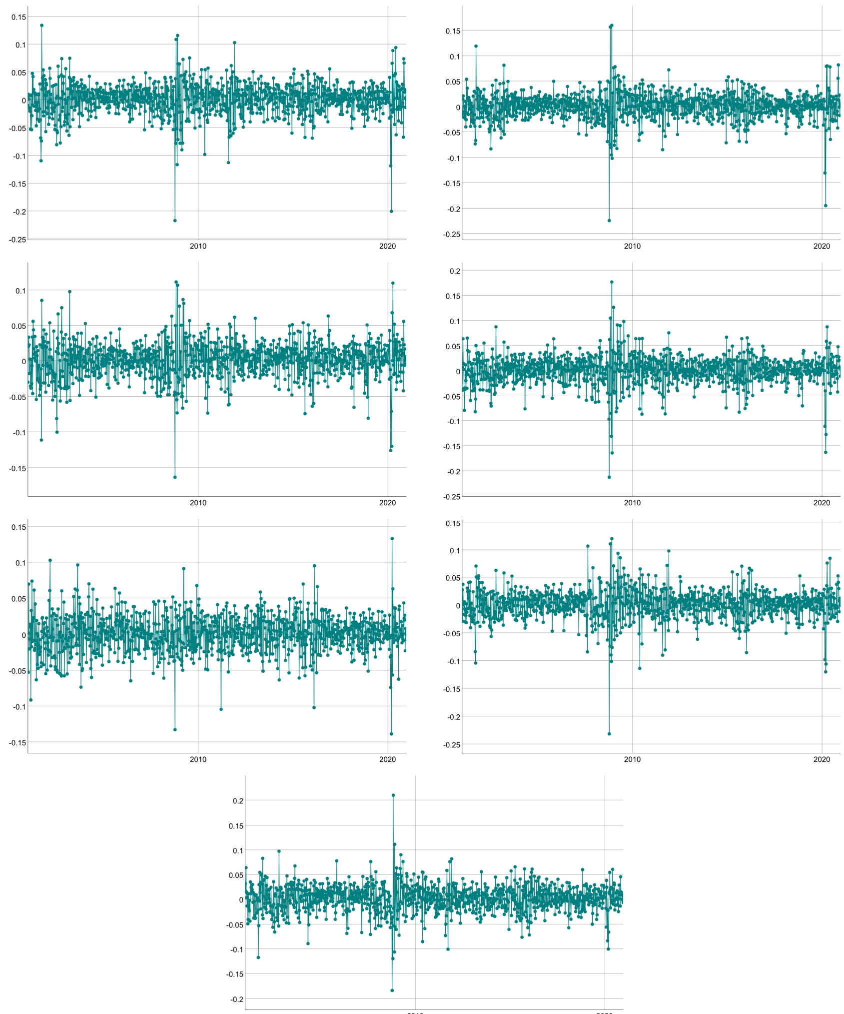

Figure C.2: Weekly total returns of MSCI EMU, MSCI UK, MSCI USA, MSCI CANADA, MSCI JAPAN, MSCI PACIFIC EX JAPAN, MSCI EMERGING MARKETS, respectively from left to right.

References

[1] H. Markowitz, Portfolio selection, journal of finance, Markowitz HM—1952.—№ (1952) 77–91.

[2] H. Markowitz, Portfolio selection (1959).

[3] M. J. Best, R. R. Grauer, On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results, The review of financial studies 4 (2) (1991) 315–342.

[4] V. K. Chopra, W. T. Ziemba, The effect of errors in means, variances, and covariances on optimal portfolio choice, in: Handbook of the fundamentals of financial decision making: Part I, World Scientific, 2013, pp. 365–373.

[5] R. O. Michaud, The markowitz optimization enigma: Is 'optimized'optimal?, Financial analysts journal 45 (1) (1989) 31–42.

[6] R. O. Michaud, R. O. Michaud, Efficient asset management: a practical guide to stock portfolio optimization and asset allocation, Oxford University Press, 2008.

[7] M. D. Braga, Risk-based approaches to asset allocation: Concepts and practical applications, Springer, 2015.

[8] E. Jurczenko, M. Thierry, J. Teiletche, A unified framework for risk-based investing, Tech. rep., EHL (2015).

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

[10] E. Qian, Risk parity portfolios: Efficient portfolios through true diversification, Panagora Asset Management (2005).

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

[12] R. Litterman, Hot spots and hedges, Journal of Portfolio Management (1996) 52.

[13] L. Martellini, V. Milhau, A. Tarelli, Toward conditional risk parity: Improving risk budgeting techniques in changing economic environments, The Journal of Alternative Investments 18 (1) (2015) 48–64.

[14] K. Boudt, P. Carl, B. G. Peterson, Asset allocation with conditional value-at-risk budgets, Journal of Risk 15 (3) (2012) 39–68.

[15] H. Mausser, O. Romanko, Long-only equal risk contribution portfolios for cvar under discrete distributions, Quantitative Finance 18 (11) (2018) 1927–1945.

[16] E. F. Fama, Mandelbrot and the stable paretian hypothesis, The journal of business 36 (4) (1963) 420–429.

[17] F. M. Longin, The asymptotic distribution of extreme stock market returns, Journal of business (1996) 383–408.

[18] A. Peiro, Skewness in financial returns, Journal of Banking & Finance 23 (6) (1999) 847–862.

[19] A. J. Prakash, C.-H. Chang, T. E. Pactwa, Selecting a portfolio with skewness: Recent evidence from us, european, and latin american equity markets, Journal of Banking & Finance 27 (7) (2003) 1375–1390.

[20] E. Jondeau, M. Rockinger, Optimal portfolio allocation under higher moments, European Financial Management 12 (1) (2006) 29–55.

[21] G. Y. Tang, D. F. Choi, Impact of diversification on the distribution of stock returns: international evidence, Journal of Economics and Finance 22 (2) (1998) 119–127.

[22] L. You, R. T. Daigler, Is international diversification really beneficial?, Journal of Banking & Finance 34 (1) (2010) 163–173.

[23] R. C. Scott, P. A. Horvath, On the direction of preference for moments of higher order than the variance, The Journal of finance 35 (4) (1980) 915–919.

[24] P. Zangari, A var methodology for portfolios that include options, Risk-Metrics Monitor 1 (1996) 4–12.

[25] M. G. Zoia, Tailoring the gaussian law for excess kurtosis and skewness by hermite polynomials, Communications in Statistics-Theory and Methods 39 (1) (2009) 52–64.

[26] M. G. Zoia, P. Biffi, F. Nicolussi, Value at risk and expected shortfall based on gram-charlier-like expansions, Journal of Banking & Finance 93 (2018) 92–104.

[27] G. M. De Athayde, R. G. Flôres, Incorporating skewness and kurtosis in portfolio optimization: A multidimensional efficient set, in: Advances in portfolio construction and implementation, Elsevier, 2003, pp. 243–257.

[28] T.-Y. Lai, Portfolio selection with skewness: a multiple-objective approach, Review of Quantitative Finance and Accounting 1 (3) (1991) 293–305.

[29] P. Chunhachinda, K. Dandapani, S. Hamid, A. J. Prakash, Portfolio selection and skewness: Evidence from international stock markets, Journal of Banking & Finance 21 (2) (1997) 143–167.

[30] M. Mhiri, J.-L. Prigent, International portfolio optimization with higher moments, International Journal of Economics and Finance 2 (5) (2010) 157–169.

[31] M. Aksaraylı, O. Pala, A polynomial goal programming model for portfolio optimization based on entropy and higher moments, Expert Systems with Applications 94 (2018) 185–192.

[32] L. Martellini, V. Ziemann, Improved estimates of higher-order comoments and implications for portfolio selection, The Review of Financial Studies 23 (4) (2010) 1467–1502.

[33] K. Boudt, W. Lu, B. Peeters, Higher order comoments of multifactor models and asset allocation, Finance Research Letters 13 (2015) 225–233.

[34] N. Lassance, F. D. Vrins, Portfolio selection with parsimonious higher comoments estimation, Journal of Banking and Finance 126 (9) (2021) 106–115.

[35] K. Boudt, D. Cornilly, T. Verdonck, A coskewness shrinkage approach for estimating the skewness of linear combinations of random variables, Journal of Financial Econometrics 18 (1) (2020) 1–23.

[36] F. A. Fabozzi, J. Simonian, F. J. Fabozzi, Risk parity: The democratization of risk in asset allocation, The Journal of Portfolio Management 47 (5) (2021) 41–50.

[37] V. DeMiguel, L. Garlappi, R. Uppal, Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?, The review of Financial studies 22 (5) (2009) 1915–1953.

[38] T. Roncalli, G. Weisang, Risk parity portfolios with risk factors, Quantitative Finance 16 (3) (2016) 377–388.

[39] L. Rade, B. Westergren, Mathematics handbook for science and engineering, Vol. 4, Springer, 1995.

[40] H. Lutkepohl, Handbook of matrices., Computational statistics and Data analysis 2 (25) (1997) 243.

[41] J. R. Magnus, H. Neudecker, Matrix differential calculus with applications to simple, hadamard, and kronecker products, Journal of Mathematical Psychology 29 (4) (1985) 474–492.

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

[43] H. Kaya, W. Lee, Demystifying risk parity, Available at SSRN 1987770 (2012).

[44] F. Spinu, An algorithm for computing risk parity weights, Available at SSRN 2297383 (2013).

[45] X. Hu, R. Shonkwiler, M. C. Spruill, Random restarts in global optimization (2009).

[46] D. N. Politis, J. P. Romano, A circular block-resampling procedure for stationary data, Purdue University. Department of Statistics, 1991.

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

[48] V. Bhansali, Beyond risk parity, The Journal of Investing 20 (1) (2011) 137–147.

[49] R. Deguest, L. Martellini, A. Meucci, Risk parity and beyond-from asset allocation to risk allocation decisions, Available at SSRN 2355778 (2013).

[50] H. Neudecker, T. Wansbeek, Some results on commutation matrices, with statistical applications, Canadian Journal of Statistics 11 (3) (1983) 221–231.