Equally Diversified or Equally Weighted?

Gianluca Fusai[1]

Domenico Mignacca[2]

Andrea Nardon[3]

Ben Human[4]

June 16, 2020

Abstract

The aim of this paper is to shed new light on the concept of diversification showing that it is not necessarily related to the reduction of the volatility of a portfolio, as it is commonly perceived. We introduce a diversification index that exploits the decomposition of portfolio volatility into undiversified volatility and a diversification component. The diversification component offsets the undiversified part leaving as a final result the portfolio volatility itself. Our decomposition has a clear statistical interpretation because it relates the diversification component to the so-called partial covariances, i.e. the covariances between the residuals of the regressions of the weighted asset returns with respect to the portfolio return. On this basis, we advocate the construction of an equally diversified portfolio versus an equally weighted portfolio. An empirical analysis illustrates the superior performance of the equally diversified portfolios with respect to the equally weighted portfolio.

KeyWords: Diversification Measure, Portfolio Allocation, Risk Contribution, Euler decomposition **JEL code:** G10, G11, D81

[1]**Corresponding Author**. Professor of Financial Mathematics, Università del Piemonte Orientale, Novara, Italy, and Faculty of Finance, Cass Business School, City, University of London. Email: gianluca.fusai@uniupo.it and gianluca.fusai.1@city.ac.uk

[2]Director Risk Management Department Qatar Investment Authority

[3]Partner, Head of Quants, Sarasin & Partners LLP

[4]Quant Analyst, Sarasin & Partners LLP

Electronic copy available at: https://ssrn.com/abstract=3628585

1 Introduction

In this paper, we aim to shed new light on the concept of diversification showing that it is not necessarily related to the reduction of the volatility of a portfolio, as it is commonly perceived. In order to build a portfolio that "is not heavily exposed to individual shocks", one popular strategy is to select mean-variance efficient portfolios. Another popular suggestion is to build an equally weighted portfolio. This simple strategy has become popular also from an academic perspective due to DeMiguel et al. [3]. They show that the out-of-sample Sharpe ratio of the equally weighted portfolio is larger than the one achieved by mean-variance portfolios. The main reason of this being the estimation error affecting the construction of the efficient frontier: mean-variance portfolios are characterized by very extreme weights and disappointing out-of-sample performances, e.g. Michaud [13]. However, to hold an equally weighted portfolio is not a guarantee of diversification, see Pola [15]. For an extreme example, suppose an investor is optimizing a portfolio with only two assets: a broad market index, such as the SP500, and one of its constituents (e.g. the IBM stock). The (true) optimal portfolio would likely be very different from the equally weighted one. In a more realistic example, a portfolio equally weighted in 10 high-tech firms would be less diversified than one invested in ten stocks from very different industries. Whatever the case, there is the need to have a measure able to quantify how well diversified a portfolio is. Mean-variance portfolios appear to be too concentrated in few assets to be considered well diversified. However, no consensus on a unique diversification measure has been achieved so far. Among the different proposals, we recall the weighted average variance of assets, see Goetzmann and Kumar [4], the principal component analysis of the correlation matrix, Rudin and Morgan [18], the Diversification Ratio (DR), introduced by Choueifaty and Coignard [1] and the Effective Number of Bets (ENB), see Meucci [12].

The main contribution of the present paper is to introduce a new diversification index that can also be used in building an optimally diversified portfolio, and whose calculation only requires the knowledge of the portfolio weights and of the covariance matrix. This new index is obtained using a new decomposition of the portfolio volatility into undiversified volatility and a diversification component. The diversification component offsets the undiversified part leaving as a final result the portfolio volatility itself. Our decomposition has a clear statis-

Electronic copy available at: https://ssrn.com/abstract=3628585

tical interpretation because it relates the diversification component to the so called partial covariances, i.e. the covariances between the residuals of the regressions of the weighted asset returns with respect to the portfolio return. The undiversified component is related to the sum between the individual asset variance and the (partial) variance of the residual of the just mentioned regressions. Moreover, the risk that can be diversified is measured with reference to the partial variances, i.e. the risk orthogonal to the portfolio return, and not with respect to the variance of the assets. The idiosyncratic risk is completely offset by the component related to the partial covariances. This is discussed in detail in Section 2. In Section 3 we present an illustrative example showing that an equally weighted portfolio is not necessarily the most diversified portfolio and we introduce the idea of the equally diversified portfolio. Finally, in Section 4 we discuss an empirical application where we compare the out-of-sample performance of the equally weighted and of the equally diversified portfolios.

2 The Diversification Index

In this Section we introduce a new decomposition of the risk contribution associated with a specific asset into how much of it derives directly from that asset and how much of it is eliminated trough the interaction with other assets in the portfolio.

Let us consider a portfolio composed of n risky securities. We take the volatility σ of the portfolio return as risk measure

$$\sigma = \sqrt{\mathbf{w}'\mathbf{\Sigma}\mathbf{w}} \tag{1}$$

where $\mathbf{\Sigma}$ is the $n \times n$ covariance matrix with elements $\sigma_{ij}, i, j = 1, \cdots, n$ and $\mathbf{w}$ is the vector of portfolio weights with elements w_i.

Our diversification measure arises from considering the n linear regressions of each weighted asset return with respect to the portfolio return

$$w_j r_j = \alpha_j + \beta_j r_p + \varepsilon_j, j = 1, \cdots, n. \tag{2}$$

where α_j is the average return not related to the portfolio, β_j is the beta of the weighted

Electronic copy available at: https://ssrn.com/abstract=3628585

return with respect to the portfolio and ε_j is the diversifiable risk. At portfolio level, we have that $\sum_j \alpha_j = \sum_j \varepsilon_j = 0, \sum_j \beta_j = 1$. The coefficients of the least-square fit are

$$\hat{\beta}_j = \frac{w_j \sigma_j \rho_{jp}}{\sigma} \tag{3}$$

where ρ_{jp} is the correlation between the returns of asset j and the portfolio. It is convenient to introduce the quantity $\gamma_j = w_j \sigma_j \rho_{jp}$ and, given that $\sum_j \gamma_j = \sigma$, we can interpret it as measure of the risk contribution of asset j to the overall portfolio risk (and therefore $\hat{\beta}_j$ as percentage contribution). This decomposition of the portfolio volatility is very popular in the asset management industry. Indeed the risk contribution is also the basis for the construction of risk parity portfolios for which weights satisfy $\gamma_j = \sigma/n, \quad \forall j$, see for example Roncalli [17], i.e. portfolios are constructed so that the ex-ante risk contributions are equal to some given risk budget. The decomposition of the volatility as sum of the individual risk contributions holds because the adopted risk measure is homogeneous of degree one and we can apply the Euler's theorem[5]. Therefore, the risk contribution of asset j has the equivalent mathematical representation as weighted partial derivative of the portfolio volatility, i.e. $\gamma_j = w_j \frac{\partial \sigma}{\partial w_j}$, and, according to the regression, it also measures the exposure of each weighted return to the portfolio return.

Whilst the coefficients γ_j are a measure of the risk contribution of each asset, our idea is to consider the variance of the residuals of the above regressions as a measure of diversification. In other words, whilst γ_j is measuring the risk contribution of each asset to the overall portfolio risk, the variance of the residuals of the regressions is a measure of the risk unrelated to the portfolio risk and then can be taken as a measure of how well the portfolio is diversified. The reason for this choice is now detailed.

In particular, let us consider the covariance $\sigma_{i,j.r_p}$ between residuals of projections of $w_i r_i$ and $w_j r_j$ on the linear space spanned by r_p

$$\sigma_{i,j.r_p} = cov(w_i r_i - \hat{\beta}_i r_p, w_j r_j - \hat{\beta}_j r_p) \tag{4}$$

[5] A function $f : R^n \to R$ is said to be homogeneous of degree k if $f(\alpha \mathbf{x}) = \alpha^k f(\mathbf{x}) \forall$ non-zero $\alpha \in R$ and $\mathbf{x} \in R^n$. Positively homogeneous functions are characterized by Euler's homogeneous function theorem. If the function f is also continuously differentiable, then f is positively homogeneous of degree k if and only if $kf(x) = \mathbf{x} \nabla f(\mathbf{x}) = \sum_{i=1}^n x_i \frac{\partial f(x)}{\partial x_i}$.

Electronic copy available at: https://ssrn.com/abstract=3628585

This quantity is called partial covariance and, with some algebra, it turns out to be equal to

$$\sigma_{i,j.r_p} = w_i w_j \sigma_{i,j} - \gamma_i \gamma_j \tag{5}$$

Therefore, we can decompose the covariance between two weighted stock returns as follows

$$cov(w_i r_i, w_j r_j) = \gamma_i \gamma_j + \sigma_{i,j.r_p} \tag{6}$$

and, setting $i = j$, we also have the variance decomposition

$$var(w_i r_i) = \gamma_i^2 + \sigma_{i.r_p}^2 \tag{7}$$

where $\sigma_{i.r_p}^2$ is called partial variance.

Formulas (6) and (7) provide a decomposition of the covariances (and variances) of the weighted returns: the first component is the product $\gamma_i \gamma_j$, that, once summed across all i and j, builds to up to the portfolio variance; the second term is the idiosyncratic covariance, $\sigma_{i,j.r_p}$, that allows to diversify away, at stock and portfolio level, the partial variances. Indeed, summing over all indexes i and j, we have

$$\sigma^2 = \sum_{i,j} cov(w_i r_i, w_j r_j) = \sum_{i,j} \left(\gamma_i \gamma_j + \sigma_{i,j.r_p}\right) = \sum_i \gamma_i \sum_j \gamma_j + \sum_{i,j} \sigma_{i,j.r_p} = \sigma^2 + \sum_{i,j} \sigma_{i,j.r_p} \tag{8}$$

and it is clear that the sum of partial covariances must be equal to zero. In particular, we can split it in two components, and, after normalizing with respect to the portfolio volatility[6], we have

$$\sigma = \sigma + \frac{1}{\sigma}\left(\sum_i \sigma_{i.r_p}^2 + \sum_{i,j,i \neq j} \sigma_{i,j.r_p}\right) \tag{9}$$

that confirms that, at portfolio level, the partial variances are offset by the partial covariances. Due to the homogeneity of the risk contribution, a similar decomposition holds also

[6] The normalization allows us to have a diversification measure having the same unit of measure as the risk measure.

Electronic copy available at: https://ssrn.com/abstract=3628585

at asset level, indeed we have

$$\gamma_j = \gamma_j + \frac{1}{\sigma}\left(\sigma^2_{j.r_p} + \sum_{k,j\neq k} \sigma_{j,k.r_p}\right) \tag{10}$$

i.e. the partial variance of asset j is cancelled by the partial covariances with all the remaining assets. Therefore, according to (10) we can call $\gamma_j + \frac{\sigma^2_{j.r_p}}{\sigma}$ the undiversified risk contribution of asset i and γ_i simply the (diversified) risk contribution. According to (9), $\sigma + \frac{\sum_i \sigma^2_{i.r_p}}{\sigma}$ is the undiversified portfolio risk. Combining an asset with the remaining ones, the portfolio risk becomes just σ.

On the basis of this discussion, we introduce as the diversification contribution of asset j, the following quantity[7]

$$QDX_j = \frac{\sigma^2_{j.r_p}}{\sigma} \tag{11}$$

It is also convenient to introduce $DIV_j = \frac{1}{\sigma}\sum_{i,i\neq j}\sigma_{i,j.r_p}$. Therefore, we can rewrite (10) as

$$\gamma_j = \gamma_j + QDX_j + DIV_j \tag{12}$$

and (9) as

$$\sigma = \sigma + QDX + DIV \tag{13}$$

where we have defined $QDX = \sum_{i=1} QDX_i$ and $DIV = \sum_{i=1} DIV_i$. These decompositions are summarized in Table 1.

A few comments are useful.

First. The QDX_i measure has an easy to understand interpretation if we rewrite it as

$$QDX_i = \frac{w_i^2\sigma_i^2 - \gamma_i^2}{\sigma} \tag{14}$$

In particular, an asset having a low diversification contribution to portfolio is characterized by having a squared risk contribution γ_i^2 similar to the weighted variance $w_i^2\sigma_i^2$, and then a low value of QDX_i. Viceversa, in a well diversified portfolio the risk contributions of the different assets tend to be different from the weighted variance and the QDX_i take large values.

[7] QDX is itself an acronym for Quantitative Diversification Index.

Electronic copy available at: https://ssrn.com/abstract=3628585

Table 1: Decomposition of the risk contribution and portfolio volatility in terms of partial variances and covariances.

Asset	Risk Contribution	=	Undiversified	+ Diversified	=	Sum
1	γ_1	=	$\gamma_1 + QDX_1$	$+DIV_1$	=	γ_1
2	γ_2	=	$\gamma_2 + QDX_2$	$+DIV_2$	=	γ_2
$\cdots$	$\cdots$	=	$\cdots$	$+\cdots$	=	$\cdots$
n	γ_n	=	$\gamma_n + QDX_n$	$+DIV_n$	=	γ_n
Sum	σ	=	$\sigma + QDX$	$+ DIV$	=	σ

Therefore, on the basis of (12) we can define $\gamma_i + QDX_i$ as the undiversified risk contribution of asset i, whilst DIV_i is the diversification contribution of the individual holdings to the overall portfolio amount of diversification as measured by the $QDX = \sum_{i=1}^{n} QDX_i$. The novel idea here is that in measuring the diversification effect of an asset we need to control for the portfolio return, so that the diversification contribution of an asset is related to the partial covariance of that asset with the remaining assets rather than being measured by the covariances among weighted returns. This is also confirmed when we notice that in our decomposition, the covariances among residuals counter-balance the ~~idiosyncratic~~ specific variances: indeed we have at asset level that $QDX_i + DIV_i = 0$ and at portfolio level $QDX + DIV = 0$, so that $\sigma + QDX$ can be interpreted as the un-diversified portfolio risk.

Second. We can also rewrite our diversification measure in terms of the correlation of the asset with the portfolio

$$QDX_i = \frac{w_i^2 \sigma_i^2 \left(1 - \rho_{ip}^2\right)}{\sigma}, i = 1, \cdots, n \tag{15}$$

According to intuition, diversification should be related to the elimination of the risk unrelated to the portfolio. This is confirmed by the above expression: an asset having a low correlation with the portfolio provides diversification benefits. Viceversa, if the asset has a large correlation with the portfolio, it cannot help in diversifying away the risk of the remaining assets.

As third remark, it is also important to discuss the relationship between risk reduction and diversification. Indeed, according to formula (13) we can build portfolios with the same level

Electronic copy available at: https://ssrn.com/abstract=3628585

of risk σ but very different values of QDX, or viceversa portfolios having different levels of risk but similar diversification levels. In order to reduce the portfolio risk we exploit the covariances between asset returns. In order to control the level of diversification of the portfolio we need to control the partial covariances. This is discussed empirically also in Pola [15]. Moreover, our decomposition justifies the remark in Roll [16] that diversification is not related to the low correlation between assets, but to the variance of the residuals in a factor model. Here we use as the factor the portfolio return. In addition, working with uncorrelated residuals, we can express our diversification measure at portfolio level as sum of idiosyncratic variances.

As fourth remark, we stress that the decompositions in equations (9) and (10) hold if the regression is made with respect to the given portfolio. If, instead, we run the regressions with respect to some wide market portfolio, they do not hold. This should not be a surprise. For example, Sharpe [19] runs similar regressions, but with respect to a wide market portfolio and not, like we do here, with respect to the invested portfolio. Therefore (9) and (10) do not hold in the Sharpe approach. In addition, in the Sharpe approach, the asset manager must identify some "typical non market risk". So, our diversification measure is less subjective because it is independent on the choice of the benchmark and on the measurement of the "typical non market risk" as required in the Sharpe [19] formulation. This quantity indeed, without further assumptions, in general is not readily available. If it is identified with the portfolio volatility, and the market portfolio is exactly the portfolio at hand, the two approaches provide the same diversification measure.

As fifth remark, the use of the just introduced QDX can be motivated also from a mathematical point of view, using a second order Euler decomposition formula, whose importance has never been realized previously in the literature. This derivation is presented and discussed in detail in Mignacca and Fusai [14]. They show that if the risk measure $\mathcal{R}(\mathbf{w}) : R^n \to R$ is homogeneous of degree one, then we can obtain a generalized QDX measure via

$$QDX_j(\mathbf{w}) = w_j^2 \frac{\partial^2 \mathcal{R}(\mathbf{w})}{\partial w_j^2} \tag{16}$$

Electronic copy available at: https://ssrn.com/abstract=3628585

and similarly

$$DIV_j(\mathbf{w}) = \sum_{k=1,k\neq j}^{n} w_k \frac{\partial^2 \mathcal{R}(\mathbf{w})}{\partial w_k \partial w_j} \tag{17}$$

In the particular case where the risk measure is the portfolio volatility, i.e. $\mathcal{R}(\mathbf{w}) = \sqrt{\mathbf{w}'\Sigma\mathbf{w}}$, we obtain (14)[8]. In particular, the diversification measure is strictly related to the convexity of the risk function. Trivially, a risk function that is linear in the weights does not allow for diversification. In this case, the QDX takes the value of 0. More convex the risk function is, larger the second derivatives and therefore the QDX measure. The diversification component is related to the mixed second order partial derivatives of the risk function. In conclusion, our decomposition is not arbitrary and it is valid whenever the risk-measure is homogeneous of order one. In addition, this implies that our measure can promptly be extended to other risk measures, such as Value at Risk and Expected Shortfall.

Finally, we observe that our analysis complements the one in Hallerbach [5]. In particular, Hallerbach observes that the orthogonal projection of the weighted asset returns into the space spanned by the portfolio return is linear whenever the returns belong to the class of multivariate elliptical distributions, for which the Gaussian distribution is a member of. This is a rich family of distributions that includes, among the others, non-normal variance mixtures that exhibit tails heavier than that of a normal distribution and can be very useful for capturing extreme events. To be in an elliptical market is of paramount importance. Indeed, if we adopt an homogeneous risk-measure such as VaR and Expected Shortfall, the partial derivatives in (16) and (17) are related to conditional covariances and then can be computed analytically, see Jaworski and Pitera [6]. On the other side, the decomposition of the portfolio volatility in (12) and (13) holds whatever the joint distribution of asset returns, because it only requires the homogeneity of the risk measure, and its calculation is based on the partial covariances, i.e. the covariances between residuals of the linear regressions we have discussed above, rather than on conditional covariances[9].

[8] The calculation of the Hessian matrix returns $(\Sigma\sigma - \Sigma\mathbf{w}\mathbf{w}'\Sigma/\sigma)/\sigma^2$

[9] Except for the multivariate normal distribution, in the elliptical class conditional covariances are random, depending on the conditioning variable, whilst partial covariances are deterministic.

Electronic copy available at: https://ssrn.com/abstract=3628585

3 Diversification Parity

Given that the contribution of each asset to the overall portfolio diversification is measured by QDX_j, in the vein of Maillard et al. [10] and Meucci [12], we build a portfolio where the diversification contributions are similar across assets. Hence, the idea of minimizing the cross-section dispersion of the variances of the idiosyncratic risks. A low diversified portfolio is one with idiosyncratic variances concentrated in few assets. This is a different approach from risk parity. In risk parity, we equalize the risk contributions across assets. Here, we equalize the diversification contribution across assets. If we define R_j to be the ratio between the quantity of diversification relative to the single asset and their sum, i.e.

$$R_j = \frac{QDX_j}{QDX},$$

we look for a portfolio allocation such that $R_j = \frac{1}{n}, \forall j$. In order to build the QDX parity portfolio, we solve the following minimization problem

$$\hat{\mathbf{w}}_{QDX} = argmin_{\mathbf{w}} \quad \sum_{j=1}^{n} \left(R_j - \frac{1}{n} \right)^2 \tag{18}$$

subject to the balance constraint $\mathbf{1}'\mathbf{w} = 1$ and the no short sell constraint $\mathbf{w} \geq \mathbf{0}$. Additional constraints on the expected return as well as on the portfolio risk can be added. It is also useful to exploit the concept of effective number of bets introduced in Meucci [12]. For this, we define the (entropy) quantity

$$\mathcal{N}^{(\mathbf{w},\Sigma)} = \exp\left(-\sum_{i=1}^{n} R_j \ln(R_j) \right). \tag{19}$$

$\mathcal{N}^{(\mathbf{w},\Sigma)}$ achieves the maximum value of n when the diversification parity condition holds, and takes the minimum value of 1 for a portfolio fully invested in one asset, i.e. there exists j such that $R_j = 1$ (or $w_j = 1$) and therefore $R_i = 0$ (or $w_i = 0$) for $i \neq j$[10]. Notice also that if the weight of a subset made of m assets is zero, the entropy measure has a maximum value of $n-m$. Therefore, this entropy measure takes values in $[1, n]$ and its value can be interpreted

[10] If $R_j = 0$, we set $R_j ln(R_j) = 0$ in 19.

Electronic copy available at: https://ssrn.com/abstract=3628585

as the effective number of uncorrelated (specific) risks we are investing respect to the nominal number of constituents in a portfolio. We also remark that if we have homogeneous assets, i.e. they have the same variances and the same cross-correlations, the equally diversified portfolio and the equally weighted one coincide. In general, this is not the case as we can illustrate in the following example.

Let us consider the 3x3 covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 7.40 & 9.60 & 10.00 \\ 9.60 & 18.20 & 16.40 \\ 10.00 & 16.40 & 27.00 \end{bmatrix}$$

If we consider the equally weighted portfolio, the corresponding partial covariance matrix is given by

$$\mathbf{\Sigma_{.r_p}} = \begin{bmatrix} 0.17 & 0.00 & -0.17 \\ 0.00 & 0.28 & -0.28 \\ -0.17 & -0.28 & 0.46 \end{bmatrix}$$

where the elements on the main diagonal are the partial variances and the off-diagonal elements are the partial covariances. We can verify that the decomposition (13) is verified: the sum across rows is equal to zero. However, according to QDX, the portfolio is not well diversified, because the three assets are characterized by very different partial variances, i.e. 0.17, 0.28 and 0.46. Moreover, the idiosyncratic variance of asset 1 is near completely cancelled by the interaction with asset 3. Similarly, the idiosyncratic risk of the second asset is mainly balanced by the covariance with the third asset. The number of effective bets is 2.78 (92.6% in percentage of the total number of assets), an apparently large value not far from 3 but still there seems to be some room to try to improve with respect to the equally weighted portfolio. Indeed, the most diversified portfolio turns out to have weights equal to

Electronic copy available at: https://ssrn.com/abstract=3628585

46.86%, 35.41% and 17.73% and the corresponding partial covariance matrix is

$$\boldsymbol{\Sigma}_{.\mathbf{r_p}} = \begin{bmatrix} 0.22 & -0.11 & -0.11 \\ -0.11 & 0.22 & -0.11 \\ -0.11 & -0.11 & 0.22 \end{bmatrix}$$

Equal diversification is achieved because now the partial variances are equal across assets. Even more, the partial covariances are also equal: this means that each asset is contributing the same amount of diversification as the others.

The idea of building a diversification parity portfolio is not new in literature and has been introduced for example in Meucci [12]. Meucci considers the so called *principal portfolios*, i.e. portfolios built performing a principal component analysis of the covariance matrix and imposing that the variances of the principal portfolios are *approximately equal*. Therefore, his entropy measure is computed with respect to the risk contributions of the principal portfolios. The main difference respect to our approach is that our diversification measure arises quite naturally from the Euler's formula, and it can be easily extended to other risk measures not based on the covariance matrix. Indeed, the only assumption needed to justify the use of our diversification parity approach is the homogeneity of order one of the risk measure.

4 An Empirical Analysis

Does a portfolio satisfying the diversification parity constraint performs better respect to the equally weighted strategy? This is an empirical issue that is addressed in this section.

The main objective of the analysis presented in this section, is to understand from an empirical point of view if a diversified portfolio according to the QDX approach can provide some advantage with respect to the equally weigthed portfolio, advocated by De Miguel et al. [3] as the winner in an out-of-sample comparison. From a theoretical point of view, diversification should work by moderating some of the negative returns during adverse markets,

Electronic copy available at: https://ssrn.com/abstract=3628585

though not necessarily eliminating them and this can be relevant to long-term investment success. To better understand why this is true, we need to assess the relationship between diversification and market performance. For this reason, in the following, aside traditional performance measures such as annualized return and volatility, Sharpe ratio, Information Ratio we consider other metrics such as drawdown, empirical probability of beating the benchmark and the EW portfolio at different horizons. Finally, we also consider conditional performance measures: i.e. we measure the average return and volatility *conditional* to the benchmark experimenting adverse movements, such as going below the empirical percentiles (1%, 5% and 10%).

For sake of completeness, we also consider other popular risk-based portfolio strategies, such as b) Risk parity (RP), where each position has the same contribution to portfolio risk; c) Inverse volatility (IV) portfolio where assets are weighted in inverse proportion to their risk; d) Most Diversified (DR) Portfolio which consists in maximising the ratio between the weighted average volatilities and the portfolio volatility; e) market-wide index.

4.1 Data

The backtesting period goes from February 1, 1995 through December 31, 2019 and considers the S&P 500 index as the investable universe as this index offers a number of advantages[11]: our dataset spans a long history covering 25 years; its constituents are remarkably liquid; it is an index which has been used both by academics and investment practitioners; and all constituents are listed on a single exchange and trade in one single currency, the USD. The S&P 500 while having only 500 active securities at every single point in time, has held 1446 securities. We have only considered active securities, i.e. securities that at rebalancing times were part of the benchmark and have further removed new issues and corporate actions which do not have prices. Hence, the results presented below do not contain off-benchmark positions. All corporate actions, including stock splits, dividends, mergers and acquisitions, rights issues and spin-offs are reflected in the dataset and have been maintained by Bloomberg. To validate the QDX methodology we have also considered other benchmarks

[11] The datasets used to test and validate the QDX approach are sourced and maintained through Bloomberg by the Sarasin Quant Solutions team of Sarasin & Partners in London.

Electronic copy available at: https://ssrn.com/abstract=3628585

such as MSCI Europe, MSCI Japan and MSCI US[12].

4.2 Estimation of the covariance matrix

The implementation of the different strategies requires the estimation of the covariance matrix. The most common approach is to use a sample covariance matrix estimated on a rolling basis. The challenges occur when the number of variables (securities) is large or the number of observations is small, which results in an unstable forecast[13]. There are several techniques that considerably improve on this approach, the most popular being the shrinkage approach proposed by Ledoit and Wolf, [7] and [8]. This technique shrinks the extreme values of the sample covariance matrix closer to a reference well-structured matrix by taking a weighted average of the sample covariance matrix with Sharpe's single index model estimator, see [8], or with a constant correlation and constant variance matrix, see [7]. The results reported here are based on [8] using weekly returns over a rollinw sample of five years. We also performed a test using the sample covariance matrix and the Ledoit and Wolf [8] method and by considering a daily sampling frequency. We found that these choices had little impact on the results[14]. T

The portfolio is constructed and rebalanced with a semi-annual frequency, at the end of May and at the end of November. The portfolio is always fully invested and cash allocations as well as short positions are not permitted at any time and only securities with a positive weight at the rebalancing date can be included in the portfolio.

[12]Source: MSCI. The MSCI data contained herein is the property of MSCI Inc. and/or its affiliates (collectively, "MSCI"). MSCI and its information providers make no warranties with respect to any such data. The MSCI data contained herein is used under license and may not be further used, distributed or disseminated without the express written consent of MSCI.

[13]As a preliminary control, the following steps have been used to compute the sample covariance matrix: 1) Remove stocks which are not part of the benchmark from the investable universe; 2) Remove stocks which have not been in the benchmark for the full time window; 3) For the remaining n stocks, using a 5-year rolling window of weekly total returns, compute the $n \times n$ sample covariance matrix; 4) If the correlation matrix is not positive semidefinite, we re-scale the correlations, until there are no negative eigenvalues and then recompute the covariance matrix using the original standard deviations and the rescaled correlations; 5) Compute the covariance matrix using the shrinkage constant and constant correlation matrix as defined in Ledoit and Wolf

[14]The results are available upon request.

Electronic copy available at: https://ssrn.com/abstract=3628585

4.3 Results

Results are reported in Table 2 where different out-of-sample performance metrics are reported[15]. We can clearly identify a cluster, which contains the QDX, the EW, IV and the RP strategies: the correlation between their performances turns out to be above 0.99. We can explain this on the basis of formula (7). According to it, the QDX parity portfolio assigns the weights so that $\sigma^2_{i.r_p}$ is the same across assets; the IV portfolio requires $w_i^2\sigma_i^2$ to be equal across assets and the RP portfolio sets the portfolio weights so that assets have the same γ_i. The EW portfolio simply requires $w_i = 1/N$. In practice, these strategies are focusing on different components of the same formula and it is not a surprise that they have very similar metrics. In addition, they considerably improve with respect to the wide market SP500 index. However, the most important insight is that all considered portfolios beat the equally weighted portfolio in terms of larger realized return, lower volatility and larger Sharpe ratios, so confirming that we can achieve some benefit moving away from this basic strategy. It is also interesting to notice that the diversification parity strategy appears to generate portfolio returns that are less negatively skewed and less leptokurtic (i.e. lower skewness and kurtosis). The lowest turnover is achieved by the inverse volatility and the QDX strategies. In addition, among these strategies, the EW strategy has the largest maximum drawdown (in absolute value and relative to the realized volatility). We also notice that the ex-ante (maximized) level of defined diversification carries over to ex-post realized diversification: the QDX strategy turns out to be the most diversified according to different criteria, such as the effective number of bets, computed either according to Meucci and to our formula 19, and the Herfindalh Index, both in-sample and out-of-sample. Few additional comments about the DR strategy are also important. Indeed, this strategy appears to have a different investment style respect to the other strategies. It returns a larger Sharpe ratio, a larger tracking error volatility, a lower Information ratio, a very large turnover. The DR portfolio suffers from high levels of concentration, having the largest Herfindahl Index and the lowest number of effective bets (in and out-of sample). The probability of beating the benchmark over mid-long horizons is significantly lower respect to the other strategies. Indeed, there is

[15] In addition to average return, standard deviation and Sharpe ratio, we have: tracking error volatility (TEV) and the information ratio (excess return with respect to the benchmark divided by the TEV), the maximum drawdown scaled by the volatility, the beta of the strategy, the probability of beating the benchmark at different horizons, the average turnover, and different out-of-sample diversification measures.

Electronic copy available at: https://ssrn.com/abstract=3628585

some open question in the literature about the real features of this strategy, see for example Taliaferro [21].

The benefits of diversification need to be assessed in the long run, after experiencing bull and bear trends, and for this reason the empirical probability of beating the benchmark and the EW portfolio at different horizons can be an additional meaningful metric. In our backtest, the empirical probability of beating the benchmark over 1, 3 and 5 years looks very similar across the strategies and larger with respect to the equally weighted portfolio. In order to examine how a diversification strategy performs in adverse market conditions, Table 2 also reports performance measures conditional to the benchmark realizing extreme negative returns, i.e. below the 5% empirical percentile. In such a case, the three strategies, QDX, IV and RP, have very similar (conditional) average return, volatility, Sharpe ratio and probability of beating the benchmark and in general better with respect to the EW and DR portfolios.

Table 3 reports performance measures conditional to the benchmarks (MSCI US, MSCI JAPAN and MSCI Europe) realizing very extreme negative return, i.e. below the 1%, 5% and 10% empirical percentiles. Both at unconditional and conditional level, in all considered markets, the QDX portfolio has a better average return, a lower volatility and a higher probability of beating the benchmark respect to the EW. These results confirm that equally diversifying can give marginal benefits with respect to equally weighting.

A final note is on the fact that diversification measures are not invariant to the risk dimension. Suppose we have two sub-portfolios which have exactly the same composition and are invested in a large number of asset classes, so that each each sub-portfolio should be well diversified. Let us build a new portfolio investing 50% in each sub-portfolio. If we consider the diversification measure with respect to the asset classes, we still have a large number of effective bets. But if we compute the diversification measure with respect to the sub-portfolio dimension, given that they are perfectly correlated, we obtain a portfolio not diversified, i.e. the number of effective bets turns out to be equal to 1. So, we can expect that, depending on the perspective we take, a strategy can be interpreted as diversification based or as an active one. This has been remarked also from an empirical point of view by Scherer [20] and Leote de Carvalho et al. [9]. They show that different diversification strategies can be

Electronic copy available at: https://ssrn.com/abstract=3628585

Table 2: Performance measures of different diversification strategies considering as investable universe stocks belonging to the SP500. EW: equally weighted portfolio, MV: global minimum variance portfolio; QDX: QDX parity portfolio; RP: risk parity portfolio; IV: inverse volatility portfolio; DR: maximum diversification portfolio. The backtesting period goes from February 1, 1995 through December 31, 2019.

	EW	QDX	RP	IV	DR	S&P 500
Ann.Ret %	10.70%	10.96%	11.02%	10.92%	11.48%	9.62%
Vol %	15.12%	14.36%	13.53%	13.96%	13.68%	14.68%
Sharpe Ratio	0.7078	0.7628	0.8145	0.7820	0.8395	0.6553
Skewness	-0.8882	-0.8994	-0.9742	-0.9505	-0.9255	-0.8933
Kurtosis	5.9254	5.8490	6.2121	6.0139	6.1573	4.7607
TEV %	5.51%	5.58%	6.34%	5.75%	12.29%	0.00%
Inf. Ratio	0.1969	0.2400	0.2216	0.2256	0.1518	
Max. DD%	31.12%	30.28%	29.06%	29.61%	28.38%	27.72%
Max. DD/Vol. %	2.0579	2.1078	2.1469	2.1212	2.0750	1.8886
Beta	0.9603	0.9067	0.8319	0.8755	0.5839	1.0000
Prob 1y over S&P	51.90%	56.06%	58.48%	59.52%	57.79%	0.00%
Prob 3y over S&P	67.17%	72.08%	75.47%	73.96%	73.21%	0.00%
Prob 5y over S&P	83.40%	83.82%	83.82%	87.97%	85.06%	0.00%
Prob 1y over EW	0.00%	56.40%	51.90%	53.63%	50.17%	48.10%
Prob 3y over EW	0.00%	70.57%	66.79%	61.89%	61.51%	32.83%
Prob 5y over EW	0.00%	80.50%	77.59%	73.86%	80.50%	16.60%
Turnover	7.74	7.32	7.91	7.19	14.21	5.88
Herfindalh	0.2908	0.2766	0.2997	0.3127	4.9107	
Effective Bets (Meucci)	394	395	403	396	234	
Effective Bets (formula 19)	282	387	270	338	15	
SP500 has performance below the 5% percentile						
Conditional Ann. Ret. %	-117,28%	-109,78%	-100,24%	-106,14%	-76,20%	-120,51%
Conditional Vol. %	16,51%	16,07%	17,22%	16,40%	23,54%	10,75%
Prob. Over SP500	46,67%	53,33%	73,33%	60,00%	73,33%	0,00%

Electronic copy available at: https://ssrn.com/abstract=3628585

Table 3: The Table returns the unconditional and conditional monthly average return, monthly volatility and probability of beating the benchmark for the EW and QDX strategies and the benchmark itself. The conditioning is with respect to the benchmark having a return below its empirical percentiles (at level 1%, 3%, 5% and 10%). The same sample period as in Table has been considered

Percentile		EW	QDX	S&P 500	EW	QDX	MSUS
Unconditional	Mon. Ret. %	0.89%	0.91%	0.80%	0.75%	0.77%	0.66%
	Mon. Vol. %	4.37%	4.15%	4.24%	4.47%	4.18%	4.11%
	Prob. over bncmk	49.33%	52.33%	0.00%	54.17%	55.09%	0.00%
10%	Mon. Ret. %	-7.35%	-6.77%	-8.09%	-8.33%	-7.68%	-8.12%
	Mon. Vol. %	4.39%	4.27%	2.96%	3.78%	3.74%	2.93%
	Prob. over bncmk	53.33%	70.00%	0.00%	45.46%	68.18%	0.00%
5%	Mon. Ret. %	-9.77%	-9.15%	-10.04%	-10.58%	-9.93%	-9.91%
	Mon. Vol. %	4.77%	4.64%	3.10%	3.92%	3.86%	3.24%
	Prob. over bncmk	46.67%	53.33%	0.00%	36.36%	54.55%	0.00%
3%	Mon. Ret. %	-11.33%	-10.67%	-11.42%	-12.58%	-11.96%	-11.52%
	Mon. Vol. %	5.17%	5.00%	3.38%	4.44%	4.29%	3.74%
	Prob. over bncmk	44.44%	55.56%	0.00%	33.33%	33.33%	0.00%
1%	Mon. Ret. %	-15.72%	-14.69%	-15.17%	-16.38%	-15.52%	-15.36%
	Mon. Vol. %	5.30%	5.13%	3.46%	6.98%	6.45%	4.79%
	Prob. over bncmk	66.67%	66.67%	0.00%	50.00%	50.00%	0.00%
		EW	QDX	MSJP	EW	QDX	MSEU
Unconditional	Mon. Ret. %	0.58%	0.59%	0.39%	0.58%	0.60%	0.41%
	Mon. Vol. %	5.01%	4.82%	5.12%	4.63%	4.40%	4.25%
	Prob. over bncmk	57.41%	56.48%	0.00%	60.65%	63.89%	0.00%
10%	Mon. Ret. %	-9.04%	-8.71%	-9.69%	-8.71%	-8.26%	-8.36%
	Mon. Vol. %	3.70%	3.46%	3.58%	3.50%	3.30%	3.10%
	Prob. over bncmk	77.27%	86.36%	0.00%	45.46%	63.64%	0.00%
5%	Mon. Ret. %	-11.13%	-10.72%	-11.66%	-11.40%	-10.83%	-11.02%
	Mon. Vol. %	4.26%	3.91%	4.23%	2.97%	2.81%	1.96%
	Prob. over bncmk	72.73%	90.91%	0.00%	54.55%	72.73%	0.00%
3%	Mon. Ret. %	-13.22%	-12.68%	-13.52%	-13.06%	-12.27%	-12.37%
	Mon. Vol. %	4.91%	4.41%	5.15%	2.41%	2.28%	1.60%
	Prob. over bncmk	50.00%	83.33%	0.00%	33.33%	66.67%	0.00%
1%	Mon. Ret. %	-17.74%	-16.71%	-18.85%	-16.02%	-15.03%	-14.26%
	Mon. Vol. %	7.54%	6.80%	6.79%	1.64%	1.63%	0.86%
	Prob. over bncmk	100.00%	100.00%	0.00%	50.00%	50.00%	0.00%

Electronic copy available at: https://ssrn.com/abstract=3628585

viewed as simple active strategies with respect to a few factors. Following the above Authors, we have run a multi-factor regression of the excess returns of each strategy against different Fama-French (FF) factor portfolios (Excess Market Return, HML, SMB, High Beta-Low Beta Porfolios, Momentum, Variance of the residuals from the FF three-factor model). The regression coefficients can be interpreted as weights of a multi-factor portfolio that closely tracks excess returns. The results are reported in Table 4 and confirm the remarks in Leote de Carvalho et al. [9]: the DR is a defensive strategies with a negative exposure to the market; the remaining strategies, i.e. QDX, EW, RP and IV, look very similar in terms of factor exposures and this confirms that they represent a cluster. These results also confirm, as said earlier, that diversification is a relative concept depending on the dimension we are considering. For example, Mignacca and Fusai [14] show that in the single factor model the equally weighted portfolio has an effective number of bets, in terms of principal components extracted via a principal component analysis of the covariance matrix, equal to 1, whatever the portfolio size. Viceversa, a portfolio, that is equally weighted across principal components, turns out to be fully invested in one of the available stocks. In conclusion, the user of a diversification measure has to clarify if his/her aim is to build a balanced portfolio at asset or factor level. Forgetting one dimension, diversification benefits can be overstated.

Table 4: Factor Regression Coefficients for Risk-Based Strategies. Factors series have been downloaded from the French data library: MKT-RF is the market-cap index minus the U.S. one-month T-Bill rate, HML and SMB are the Value and Size factors, LBMHB and LRVMHRV are the Beta and residual Volatility Factors, MOM is the momentum factor. Significance levels at 0.1%, 1% and 5% are marked by a, b and c.

	EW	QDX	RP	IV	DR	S&P 500
Intercept	0.0007	0.0009	0.0010	0.0009	0.0027	-0.0001
MKT-RF	0.0391	0.0288	0.0092	0.0181	-0.0777	0.0052
HML	0.2831[a]	0.2681[a]	0.2508[a]	0.2511[a]	0.1515[b]	0.0339[b]
SMB	0.1273[a]	0.1091[a]	0.1395[a]	0.1202[a]	0.1718[b]	-0.1116[a]
LBMHB	0.1152[a]	0.1419[a]	0.2196[a]	0.1640[a]	0.5396[a]	0.0301[b]
MOMENTUM	-0.1379[a]	-0.1265[a]	-0.1073[a]	-0.1262[a]	-0.0826[c]	-0.0388[a]
LRVMHRV	0.0499[c]	0.0649[b]	0.0420[c]	0.0677[b]	-0.1078[c]	0.0691[a]
R^2	68%	73%	73%	74%	57%	81%

Electronic copy available at: https://ssrn.com/abstract=3628585

We also checked the impact of changing the sampling frequency, from daily to monthly as well as the procedure for estimating the covariance matrix moving from Ledoit and Wolf [7] to Ledoit and Wolf [8] i.e. shrinking the sample covariance matrix towards the identity matrix, instead of the single-factor Sharpe model, with the weights estimated from historical returns. We found that these choices had little impact on the results.

Acknowledgements: The Authors would like to thanks for the referee's valuable comments and their time spent in reviewing our paper. They also acknowledge for precious comments participants to the AFMathConf 2019 conference in Brussels, and to the Finance Research Seminars at Cass Business School, City, University of London, University of Roma Tor Vergata, University of Bergamo, Lazaridis School of Business and Economics, Wilfrid Laurier University, 43rd Annual Meeting of Amases in Perugia. All errors remain ours.

References

[1] **Choueifaty, Y., Y. Coignard, (2008)**, Toward Maximum Diversification, The Journal of Portfolio Management, pp. 40-51.

[2] **Clarke, R., H. De Silva, S. Thorley, (2013)**, Risk Parity, Maximum Diversification, and Minimum Variance: An Analytic Perspective, The Journal of Portfolio Management, pp. 39-53.

[3] **DeMiguel, V., Garlappi, L., R. Uppal (2009)**, Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?, The Review of Financial Studies, Vol. 22, No. 5, May 2009, pp. 1915-1953.

[4] **Goetzmann, W. N., A. Kumar (2008)**, Equity Portfolio Diversification, Review of Finance, Volume 12, Issue 3, pp. 433–463.

[5] **Hallerbach W. G, (2017)**, If you have said A, you must also say B: Calculating diversified asset returns, The Journal of Wealth Management Fall 2017, 20 (2) pp. 76-81.

Electronic copy available at: https://ssrn.com/abstract=3628585

[6] **Jaworski, P. and M. Pitera, (2017)** A note on conditional covariance matrices for elliptical distributions, w.p. arXiv:1703.00918.

[7] **Ledoit, O. and M. Wolf (2003)**, Honey, I Shrunk the Sample Covariance Matrix. The Journal of Portfolio Management Summer 2004, 30 (4) 110-119.

[8] **Ledoit, O. and M. Wolf (2004)**, A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis, 88(2), pp. 365–411.

[9] **R. Leote de Carvalho, X. Lu, P. Moulin (2012)**, Demystifying equity risk–based strategies: a simple alpha plus beta description, J. Portfolio Management, 38 (3), pp. 56-70

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

[11] **Meucci, A. (2007)**, Risk Contributions from Generic User-Defined Factors, Risk, June, pp. 84-88.

[12] **Meucci, A. (2009)**, Managing Diversification, Risk, May, pp. 74-79.

[13] **Michaud, R. (1989)**, The Markowitz optimization enigma: Is 'optimized' optimal?, Financial Analysts Journal, volume 45, issue 1, p. 31 - 42.

[14] **Mignacca, D. and G. Fusai (2019)**, Diversify Diversification, working paper.

[15] **Pola, G. (2016)**, On entropy and portfolio diversification, Journal of Asset Management, vol. 17(4), pp. 218-228.

[16] **Roll, R. (2013)**, Volatility, Correlation, and Diversification in a Multi-Factor World, J. of Portfolio Management, pp. 11-18.

[17] **Roncalli T. (2013)**, Introduction to Risk Parity and Budgeting, Chapman & Hall/CRC Financial Mathematics Series.

[18] **Rudin. A. M., J. S. Morgan (2006)**, A Portfolio Diversification Index, The Journal of Portfolio Management Winter, 32 (2) 81-89.

Electronic copy available at: https://ssrn.com/abstract=3628585

[19] **Sharpe, W. F. (1972)**, Risk, Market Sensitivity and Diversification, Financial Analysts Journal, Vol. 28, No. 1, pp. 74-79. Reprinted in Financial Analysts Journal, Vol. 51, No. 1, pp. 84-89 (1995).

[20] **Scherer, B., (2011)**, A note on the returns from minimum variance investing, J. Empirical Finance 18, pp. 652–660.

[21] **Taliaferro, R. (2012)**: Understanding risk-based portfolios, The Journal of Investment Strategies, 1, pp. 119–131.

Electronic copy available at: https://ssrn.com/abstract=3628585