# Maximizing Positive Portfolio Diversification

Phil Maguire,
Philippe Moser,
Kieran O' Reilly,
Conor McMenamin,
Robert Kelly
Department of Computer Science
NUI Maynooth, Ireland
Email: pmaguire@cs.nuim.ie

Rebecca Maguire
School of Business
National College of Ireland
IFSC, Dublin 1, Ireland

***Abstract*—In this article we introduce a new strategy for optimal diversification which combines elements of Diversified Risk Parity [1], [2] and Diversification Ratio [3], with emphasis on positive risk premiums. The Uncorrelated Positive Bets strategy involves the identification of reliable, independent sources of randomness and the quantification of their positive risk premium. We use principal component analysis to identify the most significant sources of randomness contributing to the market and then apply the Randomness Deficiency Coefficient metric [4] and principal portfolio positivity to identify a set of reliable uncorrelated positive bets. Portfolios are then optimized by maximizing their diversified positive risk premium. We contrast the performance of a range of diversification strategies for a portfolio held for a two-year out-of-sample period with a 30 stock constraint. In particular, we introduce the notion of diversification inefficiency to explain why diversification strategies might outperform the market.**

*"But divide your investments among many places, for you do not know what risks might lie ahead"* – Ecclesiastes 11:2

## I. Introduction

Perhaps the only predictable thing about financial markets is that, over the long term, we expect things to get better. The population of the earth will steadily increase beyond 2050, developing nations will become industrialised and new technologies will unlock previously inaccessible resources and enhance efficiency. These future scenarios will provide new opportunities for the creation of wealth, leading to the expectation that, over the long term, the value of global stock markets will rise. Investors who hold stocks for long enough can expect to be rewarded. A downside of investing in the stock market is that unpredictable events will cause the value of the investment to fluctuate over the short and medium term. Investors' aversion to risk explains the difference between the risk-free returns that can be earned by holding safe assets such as government treasury bills and the larger gains anticipated for investors holding equities.

According to the Capital Asset Pricing Model (CAPM; see [5]), a competitive investment environment only awards a risk premium for the holding of risk that cannot be diluted in any way by holding other securities. This concept of 'undiversifiable risk' is commonly associated with the market index, the assumption being that it is not possible to create a portfolio which is more diversified than one which includes every company.

However, it may be the case that the market index does not represent the most diversified portfolio. If it were possible to create a more diversified portfolio, of which relatively few investors were aware, then this would set up the opportunity of holding an investment with a higher risk-to-reward ratio than that of the market. This scenario would present a form of market inefficiency, with risk premiums transferred from naive diversifiers to sophisticated diversifiers. In this article we investigate the efficacy of various strategies for constructing portfolios which are more diversified than the market.

## II. Diversification Strategies

Imagine throwing a single dice. In this case you are equally likely to get any number between 1 and 6. However, if you roll 100 dice and take the average, this is likely to be close to 3.5, with an extreme value being vanishingly improbable. The more dice you roll, the more likely it is that the average will be right in the middle.

Now imagine you have a very slightly biased dice, which favours the higher numbers. What are your odds of beating the average expected roll of 3.5? If you roll only one dice, the odds are not far from 50/50. However, if you roll 100 biased dice you can be nearly certain of beating the average each time.

What diversification strategies seek to do is identify multiple uncorrelated sources of risk with a positive risk premium. The more of these that can be combined simultaneously, the more reliably the portfolio will outperform the risk-free rate. If the market is indeed diversification-efficient, then there should be only a single source of undiversifiable risk, with no investor being able to diversify better than any other.

Ever since its introduction in the 1960s the CAPM has come under scrutiny, with numerous diversified alternatives being offered to the market capitalization weighted index [6]. According to the theory, the tangency portfolio is the only efficient one and should produce the greatest returns relative to risk [7]. However, many empirical studies have shown that investing in a minimum variance portfolio yields better out-of-sample results than does an investment in the tangency portfolio, thus supporting the idea that the market is diversification inefficient (see [8], [9]). For example, over the 1972 to 1989 period, the minimum variance portfolio delivered equal or greater returns than a broad market cap-weighted index of U.S. stocks, while achieving lower volatility [6]. Lohre, Neugebauer and Zimmer [1] identify the existence of a large literature demonstrating minimum-variance strategies to be far more efficient than capitalization-weighted benchmarks. These surprising findings have been attributed to the high estimation risk associated with expected returns [7].

Given perfect foresight, Markowitz's mean-variance optimization approach provides the rationale of choice for generating efficient portfolios with an optimal risk to return trade-off. However, because of the estimation risk that confounds expected returns, these supposedly optimal weights are rarely put into practice [10]. DeMiguel, Garlappi and Uppal [11] investigated the out-of-sample performance of the mean-variance model relative to the naive 1/N portfolio. Of the 14 models they evaluated across seven empirical datasets, none was consistently better than the 1/N rule in terms of Sharpe ratio, certainty-equivalent return, or turnover. These results indicate that, out-of-sample, the gain from mean-variance optimization is more than offset by estimation error. In fact, DeMiguel et al. estimate that, in order for the modern portfolio theory approach and its extensions to outperform the 1/N benchmark, 500 years of data would be required to optimize a portfolio containing 50 assets [11].

One way to circumvent the problem is simply to avoid estimating returns and instead use risk-based allocation techniques. In this article we explore and contrast a range of such techniques for evaluating the diversification of a portfolio, including Diversification Ratio (DR), Diversified Risk Parity (DRP), Uncorrelated Positive Bets (UPB) and Minimum Variance (MV).

## III. Diversification Ratio

Choueifaty and Coignard [3] originally proposed a measure of portfolio diversification called Diversification Ratio (DR), defined as the ratio of a portfolio's weighted average volatility to its overall volatility. The DR of a portfolio $w$ is given by

$$DR(w) = \frac{\langle w|\sigma \rangle}{\sigma(w)}$$

where $w = (w_i)$ is the weights of a long-only portfolio, $\sigma(w)$ its volatility and $\sum_i w_i \sigma_i$ its average volatility.

This concept encapsulates the core feature of diversification, which is that the volatility of a diversified portfolio should be less than that of the individual assets' volatilities. The DR of a single asset portfolio is 1, while that of a portfolio containing $N$ independent assets is $\sqrt{N}$. According to Choueifaty et al. [6], the DR$^2$ of a portfolio is equal to the number of independent risk factors, or degrees of freedom, represented in the portfolio. It can be interpreted as the number of independent risk factors necessary for a portfolio that allocates equal risk to independent risk factors to achieve the same DR.

Choueifaty et al. [6] define the Most Diversified Portfolio (MDP) as the long-only portfolio that maximizes the Diversification Ratio. The MDP supports several elegant properties. First, any stock not held by the MDP is more correlated with the portfolio than any of the stocks that do belong to it. In addition, all stocks belonging to the MDP share the same correlation to the portfolio. These properties illustrates that all assets in the universe are effectively represented in the MDP, even if the portfolio does not physically hold them [6]. For example, even though an MDP portfolio might hold only 50 stocks from the S&P 500, the 450 stocks not included are more correlated to it than the ones it actually holds, supporting the view that it represents the most undiversifiable portfolio. A further core property of the MDP is that the more diversified a given long-only portfolio, the greater its correlation with the MDP [6].

### A. Limitations of the Diversification Ratio

The DR approach aims to deliver a portfolio with an overall volatility which is much lower than that of its individual constituents. One problem with this concept is that it fails to distinguish between different forms of volatility. Not all diversification is good diversification. For example, if you go to the roulette wheel in a casino you can bet on red and black at the same time. Betting on red or black in isolation is a risky investment, but betting on both at the same time eliminates this volatility. Despite boasting a great DR, one cannot expect to profit from this strategy because the reduction in volatility is due to oppositional bets which sum to zero, as opposed to uncorrelated bets where the reduction in volatility does not affect the expected sum.

Not only do we need uncorrelated sources of randomness to profit from diversification, we also need to ensure that those risk sources are associated with a long-term positive risk premium. The type of independent risk sources investors want to hold are those that are expected, over the long term, to increase in value (like the biased dice). Even though it might reduce volatility, there is no point holding a risk source without an associated positive risk premium, or worse, a negative risk premium. The DR measure fails to recognize whether risk sources are associated with positive or negative risk premiums. For example, going long and going short on the same portfolio both have the same DR.

Choueifaty et al. [6] get around these two problems by defining the Most Diversified Portfolio (MDP) in terms of a long-only portfolio, which minimizes the potential for oppositional relationships and negative risk premiums. However, this ad hoc constraint suggests that the MDP is unlikely to represent the final word in diversification. Without the ability to short securities it may be impossible to unlock the full range of uncorrelated risk sources present in the market.

Consider, for instance, a pharmaceutical company involved in the development of untested new drugs. Whether the research succeeds or fails is completely independent of any other event in the financial universe. It is also associated with a positive risk premium (the drug company would not invest in the research otherwise). However, the risk of drug development might only account for a fraction of the overall volatility of the company's stock. In order to separate out the independent risk source, we must short the market to cancel out its influence. MDP's imposition of long-only constraints reduces the variety of independent risk sources that can be exploited, suggesting that the diversification the strategy provides can be further enhanced.

A further issue of MDP is its susceptibility to over-fitting. When a portfolio is optimised to maximize DR for a large pool of securities with limited historical data, it is likely that spurious data will rise to the surface. There will often be groups of stocks which, by chance, happen to appear uncorrelated for the training period, hence dominating the optimization process. The hidden risk is that, out-of-sample, these supposedly independent time series will immediately return to being correlated (see [12] for a demonstration of such overfitting).

In the following section we review an alternative strategy for diversification which identifies the largest sources of independent risk in the market, and allows them to be isolated by combining long and short positions.

## IV. Diversified Risk Parity

Diversified Risk Parity, which is based on the work of Meucci [2] and Lohre et al. [1], involves using principal components analysis (PCA) to identify the largest uncorrelated risk sources in the investment universe. The most diversified portfolio is the one which follows the risk parity approach and spreads risk evenly across these independent sources [10]. Uncorrelated components can be constructed by applying a PCA to the variance-covariance matrix of the portfolio assets. Following the spectral decomposition theorem the covariance matrix can be expressed as the product

$$\Sigma = E\Lambda E'$$

where $\Lambda = \text{diag}(\lambda 1, \ldots, \lambda N)$ is a diagonal matrix consisting of $\Sigma$'s eigenvalues that are assembled in descending order, $\lambda 1 \ldots$, and the columns of matrix $E$ represent the eigenvectors of $\Sigma$ [1]. These eigenvectors define a set of $N$ uncorrelated principal portfolios with variance $(\lambda_1, \ldots, \lambda_N)$. A given portfolio can be either expressed in terms of its original weights $w$ or in terms of weights in the principal portfolios, $\tilde{w} = E'w$. Given that the principal portfolios are uncorrelated, the total portfolio variance in the returns $Rw$ can be expressed as the weighted average over the principal portfolios' variance $\lambda_i$ using weights $\tilde{w}_i^2$

$$\text{var}(R_w) = \sum_{i=1}^{n} \tilde{w}_i^2 \lambda_i$$

The principal portfolio's contribution can then be normalized in terms of the overall portfolio variance to yield a diversification distribution which sums to one.

$$p_i = \frac{\tilde{w}_i^2 \lambda_i}{\text{var}(R_w)}, i = 1, \ldots, N$$

Meucci considers a portfolio to be well-diversified when the $p_i$ terms are approximately equal and the diversification distribution is close to uniform, in other words, when the portfolio variance is equally distributed across each of the uncorrelated principal portfolios. To quantify the overall diversification, Meucci applies a dispersion metric to the diversification distribution, namely the exponential of its entropy [2].

$$N_{Ent} = \exp\left(-\sum_{i=1}^{n} p_i \ln p_i\right)$$

$N_{Ent}$ delivers values between 1, for a portfolio whose variance is concentrated on a single principal portfolio, and $N$, for a portfolio whose variance is equally distributed across all principal portfolios. Lohre et al. [1] define the Diversified Risk Parity (DRP) portfolio as the one whose risk parity weights maximize $N_{Ent}$ under a set of investment constraints (e.g. long-only investment, limited size portfolio, weight granularity).

One issue with the use of PCA is that only the first few components are reliable. In theory, it is possible to identify as many principal portfolios as assets that enter the PCA decomposition. For example, a PCA analysis of the S&P 500 yields 500 supposedly uncorrelated risk sources. Table 1 shows that the first principal portfolio accounts for an overwhelming majority of the overall variance in the market, with subsequent portfolios accounting for diminishing residues. This makes it necessary to determine a cut-off point which separates genuine components from statistical artefacts. Lohre et al. [10] use the $PC_{p1}$ and $PC_{p2}$ criteria of Bai and Ng [13] for determining the number of principal portfolios to use. For the data analysed in Table 1, these criteria identify a cut-off after the sixth principal portfolio. The values for what we henceforth refer to as 'positivity' reflect the balance between long and short positions. For example, the first principal portofolio consists of 100% long positions, while the variance-normalized long positions of the second principal portfolio exceed the short positions by only 1.4%.

TABLE I. Principal Portfolio contributions to market variance for S&P 500 July 2009 to July 2011

| PP | Variance % | Positivity % |
|---|---|---|
| 1 | 99.59 | 100 |
| 2 | 0.0006 | 1.4 |
| 3 | 0.12 | 18.1 |
| 4 | 0.19 | 28.4 |
| 5 | 0.03 | 13.1 |
| 6 | 0.003 | 6.6 |
| 7 | 0.0002 | 1.3 |
| 8 | 0.00005 | 0.7 |
| 9 | 0.01 | 12.1 |
| 10 | 0.00003 | 0.6 |

### A. Limitations of $N_{Ent}$

As previously stated, the goal of diversification is not only to identify uncorrelated risk sources but to identify uncorrelated risk sources with a positive risk premium. Each principal portfolio has two degrees of freedom, in that investors can go short or go long on it. Intuitively, one would rather go with the direction which has an excess of long market positions. Lohre et al. [1] recommend flipping the principal portfolio weights such that they align with the direction of the historical risk premium. However, as can be seen in the case of the second principal portfolio in Table 1, long and short positions can be closely matched, making it impossible to decide the direction in which to invest. Decisions which are effectively taken on the basis of a coin toss cannot be expected to yield a long-term return.

This raises the question of whether it might be best to avoid investing in principal portfolios which have no clear directionality. Not all sources of risk are associated with long-term growth. For example, although the weather may be an independent source of risk in the market, good weather can affect stocks positively (ice-cream sales) as well as negatively (umbrella sales). Because they are easily diversified, such directionless sources of risk do not attract a risk premium. A limitation associated with using the Bai and Ng cut-off [13] is that it fails to take into account the positivity of the principal portfolios.

Another issue that emerges with DRP is that, when investment constraints are imposed, portfolios retain a significant amount of idiosyncratic risk that is not accounted for by the principal portfolios being used. The effect of this idiosyncratic risk on the level of diversification is not taken into account by the $N_{Ent}$ value. For example, an extremely independent stock might have 1% variance across the first six principal portfolios and 94% idiosyncratic variance, yielding an $N_{Ent}$ of 6. The problem here is that the massive exposure to idiosyncratic risk renders the stock almost totally undiversified.

In a personal communication Lohre states "I have seen optimizations returning a 100% allocation to some special stock... Any long-only combination of assets will quickly pick up Principal Portfolio 1 risk, while single stock portfolios represent a lot of idiosyncratic risk that can translate into seemingly more bets, at least statistically. As a result, the optimizer leans towards stocks that exhibit the weirdest return history relative to the remaining universe. Of course, these portfolios should not be put into practice. For taming the optimizer, we introduce upper stock bounds of 2%." The fact that stock bounds must be deployed to tame the optimizer suggests that the $N_{Ent}$ formula does not provide a comprehensive reflection of diversification where investment constraints are involved. As will subsequently be discussed, this problem can be resolved by recognizing the impact that idiosyncratic risk has on the overall level of diversification.

### *B. Comparison of DR and DRP*

Comparing Choueifaty and Coignard's DR [3] with Lohre et al.'s DRP [10], we can see that both generate a maximum value when risk is spread evenly across a set of uncorrelated risk sources. Both measures have a minimum value of 1 when risk is concentrated on a single source. Because the DR value is wholly determined by the stocks in the portfolio, investment in a single stock always yields a DR of 1. However, a single stock can have a DRP well above 1 if the variance of that stock is well spread across the principal portfolios. Unlike DR, which considers the number of effective uncorrelated risk sources present in a given portfolio, DRP identifies its uncorrelated risk sources by decomposing covariances in the universe of investments.

In sum, the weaknesses of DR as a measure of diversification are that it does not discriminate between good and bad diversification, cannot handle short investments, and bases its value on relationships within rather than outside the portfolio, leaving it vulnerable to overfitting (see [12]). The weaknesses of DRP as a measure of diversification are that it does not account for idiosyncratic risk where investment constraints are applied, and imposes an arbitrary cut-off for identifying investable principal portfolios, without regard to their positivity.

In the following section we develop a strategy for diversification we henceforth refer to as Uncorrelated Positive Bets, which combines the desirable features of DR with DRP, while counteracting the identified weaknesses.

## V. Uncorrelated Positive Bets

The principle idea of Uncorrelated Positive Bets (UPB) is that we want to identify uncorrelated risk sources that carry a positive risk premium and then to construct a portfolio that spreads investment as beneficially as possible over these risk sources, while minimizing exposure to idiosyncratic risk. Following the approach of Meucci [2] and Lohre et al. [10] we perform PCA on the universe of stocks. However, rather than applying the Bai and Ng cut-off [13] we consider the positivity of the principal components. The UPB approach is founded on the assumption that the positivity of a principal portfolio quantifies its risk premium. Figure 1 shows the absolute positivity of the 100 first principal portfolios for the period July 2009 to July 2011.

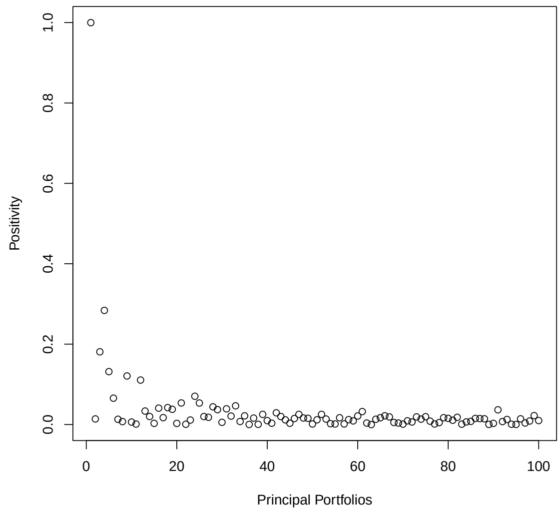


Fig. 1. Positivity by principal portfolio for the S&P 500 during the period July 2009 to July 2011

As can be seen, the first few principal portfolios are the most positive. This is not a coincidence. The first principal portfolio, which aims to maximize the variance accounted for, is expected to contain weights in a single direction (i.e. all companies move with the market, none move against the market). For the later principal portfolios, a greater proportion of the investment is eaten up by the need to isolate the influences of the earlier principal portfolios.

The positivity value reflects the proportion of the investment that is actually being invested in the random risk source.

For example, if we want to isolate the part of a pharmaceutical stock's volatility that reflects the development of new drugs, we must short other stocks to counteract its correlation with the market. A substantial amount of the investment is tied up by the process of isolating the source, as opposed to being invested in the source itself. The lower its positivity, the less likely it is that a given principal portfolio will match the risk-free rate. As can be seen in Table 1, higher principal portfolios require nearly 100% of the investment to go into isolating the risk source, completely drowning out any risk premium. Because they are market neutral and exhibit little volatility, these later principal portfolios are certain to fail to match the risk-free rate (like betting on red and black at the same time).

We want as little of our equity as possible to be tied up in isolating risk sources. Accordingly, we should choose the principal portfolios with the highest positivity. As shown in Figure 1, the first few principal portfolios tend to be the most positive, raising the question of where the cut-off should be applied. As more principal portfolios are included, the level of risk falls but so do returns (see Table 2).

UPB orders principal portfolios by positivity and applies a cut-off such that the Randomness Deficiency Coefficient (RDC; see [4]) of the resulting portfolio is maximized. RDC is a risk to reward measure which, in general, is closely related to the Sharpe Ratio [14], though expressed from a sceptical perspective. While Sharpe's ratio seeks to quantify the performance of an investment over the risk-free rate, RDC instead assumes that there is no performance over the risk-free rate. Instead, its value reflects the number of random walks with the same volatility profile that would have to be considered before finding one as good as the one being evaluated. For example, if a time series has an RDC of 10, this means that we would expect to have to consider 10 random time series like this one before finding one with the same level of growth [4].

Positive RDC values fall in the range of 2 to infinity. A negative sign is used for investments that can be shorted to generate profits above the risk-free rate, while an RDC value of 0 reflects an investment that does not exceed the risk-free rate whether longed or shorted. Maguire et al. [4] show that RDC holds several advantages over the Sharpe Ratio, including the fact that it can be used with small datasets, is time-frame independent and can be easily adjusted to take into account the familywise error rate which results from selection bias.

Table 2 shows the RDC of the diversified portfolios created by applying different cut-offs to the positivity-ordered principal portfolios for the period July 2009 to July 2011. For the purpose of identifying a cut-off, it is assumed that the risk-free rate is zero, with no cost involved in isolating risk sources. Initially, the RDC rises as the inclusion of more risk sources with positive risk premiums enhances diversification and improves the risk to reward profile. However, as more principal portfolios are added, the lower positivity leads to lower returns and more noise, with the result that RDC drops. For the example shown, UPB selects the top 3 positivity-ordered principal portfolios (1, 4 and 3) as these yield the historical portfolio with the highest RDC (i.e. 92.9).

TABLE II. PRINCIPAL PORTFOLIO CONTRIBUTIONS TO MARKET VARIANCE FOR S&P 500 JULY 2009 TO JULY 2011

| Principal Portfolios used | RDC(0) | Volatility % | Return % |
|---|---|---|---|
| 1 | 43.3 | 1.96 | 82.4 |
| 1,4 | 31.6 | 0.45 | 19.2 |
| 1,4,3 | 92.9 | 0.41 | 21.2 |
| 1,4,3,5 | 20.3 | 0.36 | 13.7 |
| 1,4,3,5,9 | 8.0 | 0.32 | 8.3 |
| 1,4,3,5,9,12 | 12.6 | 0.30 | 9.5 |
| 1,4,3,5,9,12,24 | 9.5 | 0.26 | 7.6 |
| 1,4,3,5,9,12,24,6 | 20.1 | 0.27 | 10.3 |
| 1,4,3,5,9,12,24,6,21 | 5.5 | 0.25 | 5.2 |
| 1,4,3,5,9,12,24,6,21,25 | 4.2 | 0.24 | 3.9 |

When constraints are imposed on a portfolio (e.g. long-only investment, limited size portfolio, weight granularity) then it is no longer possible to balance investments equally across the selected principal portfolios. As a result, the portfolio will have exposure to idiosyncratic risk sources, reducing the overall level of diversification. If these are not taken into account then portfolios which have a small yet evenly spread risk across the principal portfolios, and a large remaining undiversified idiosyncratic risk, will be incorrectly identified as being well diversified. Consequently, the spread of the idiosyncratic risk must also be taken into account.

The UPB strategy acknowledges both principal portfolio dispersion and the remaining idiosyncratic risk dispersion in quantifying the overall level of diversification. While the principal portfolio risk is treated as having a positive risk premium, the idiosyncratic risk is expected to match the risk-free rate, in line with the CAPM (e.g. [5]). In essence, the principal portfolios are considered as separate sources of undiversifiable risk while the remaining idiosyncratic risk sources are treated as diversifiable and thus without associated risk premium.

UPB draws on Choueifaty and Coignard's DR formula [3] to quantify the number of equivalent risk factors and the expected impact on returns. Choueifaty et al. [6] decompose the DR into a volatility-weighted average correlation $\rho(w)$ and a volatility-weighted Concentration Ratio $CR(w)$:

$$DR(w) = [\rho(w)\,(1 - CR(w)) + CR(w)]^{-\frac{1}{2}}$$

Because the principal portfolios and idiosyncratic risk sources are all uncorrelated, we can set $\rho(w)$ to 0. In such a case the DR is given by the inverse square root of Choueifaty et al.'s Concentration Ration $CR(w)$ [6], where $w = (w_i)$ is the weights of a long-only portfolio, $\sigma(w)$ its volatility and $\sum_i w_i \sigma_i$ its average volatility.

$$DR(w) = CR(w)^{-\frac{1}{2}} = \frac{\Sigma_i(w_i\sigma_i)}{\sqrt{\Sigma_i(w_i\sigma_i)^2}}$$

By slightly adapting this formula, we can define a Positive Diversification Ratio (PDR) which takes into account the three flavours of diversification, namely valenced diversification (going long or short on the principal portfolios), neutral diversification (idiosyncratic risk which is expected to match the risk-free rate) and oppositional diversification (investing in opposing positions with an expected return of zero). The original DR formula represents the expected enhancement in the risk to return profile of a portfolio relative to the market benchmark. The numerator quantifies the expected returns, while the denominator quantifies the reduction in risk.

Given that the principal portfolios are associated with both a positive risk premium (negative if shorted) and a lowering of volatility, we include these components in both the numerator and denominator. While separate sources of idiosyncratic risk serve to lower volatility, they are not expected to deliver returns above the risk-free rate. Hence, we include exposure to sources of idiosyncratic risk in the denominator only.

The final ingredient in the PDR formula is to adjust for the reduction in overall equity premium caused by shorting stocks. Because the market is expected to increase in value over the long term, any short positions in the portfolio will eat into long-term profits. If a portfolio consists of 60% long positions and 40% short positions, then the proportion attracting an equity premium is only 20%. The other 80% of the investment is effectively neutralised through oppositional bets. In the case that the risk-free rate is zero, then this is not a problem, as the investment can be leveraged for free to deliver the same level of returns as a long-only portfolio. However, if the risk-free rate is higher than zero, this leverage has an associated cost which consumes profits and must be taken into account. The cost is based on the difference between the risk-free rate $r_f$ and the expected market return $r_m$. The narrower the gap between the market return and the risk-free rate, the greater the burden of financing the leverage, and the greater the level of diversification needed to justify it.

Accordingly, we define the Positive Diversification Ratio (PDR) as follows:

$$PDR(w) = \frac{\Sigma_i(p_i)}{\sqrt{\Sigma_i(p_i)^2 + \Sigma_i(s_i)^2}} \left[\frac{r_m - r_f P^{-1}}{r_m - r_f}\right]$$

where $p_i$ is the set of principal portfolio volatilities and $s_i$ is the set of residual idiosyncratic equity volatilities that together make up the volatility of portfolio $w$, $P$ is the volatility-normalized positivity of the portfolio $w$, $r_m$ is the market return and $r_f$ is the risk-free rate.

The PDR value reflects the factor by which the portfolio is expected to outperform the risk to reward profile of the market. For example, a PDR of 2 implies that the portfolio will deliver twice as much return per unit of risk as will the market. Similar to Choueifaty and Coignard's DR [3], the PDR$^2$ is the number of degrees of freedom represented in the portfolio, that is, the effective number of positive independent risk factors necessary for a portfolio that allocates equal risk to independent risk factors to achieve the same PDR.

The principal difference between DR [3] and PDR is that PDR takes into account the impact of different types of diversification, namely valenced, neutral and oppositional. As these distinctions are achieved via a principal components analysis of the universe of stocks, our Uncorrelated Positive Bets (UPB) approach can be viewed as combining Choueifaty and Coignard's DR [3] with Lohre et al.'s DRP [1], while simultaneously quantifying the associated positive risk premium.

## VI. Results

We investigated the performance of maximized Diversification Ratio (DR), maximized Diversified Risk Parity (DRP), maximized Uncorrelated Positive Bets (UPB) and minimized Variance (MV) against the market. A 30 stock constraint was imposed, with full investment. The four strategies were optimised based on historical S&P 500 data for the training period July 2009 to July 2011 and tested on the period July 2011 to July 2013. The constraints and test periods were not cherry-picked to suit any particular strategy, with the earlier sections of this paper having been written before running the experiment.

As the $N_{Ent}$ formula does not recognize any difference between going short or long on the principal portfolios or the market, we applied the optimization constraints that all portfolio investment weights must be positive and all exposures to the principal portfolios must be in the same direction as the in-sample historical risk premium. Similarly, we applied the positive weight constraint for optimising MV and DR. The PDR formula was the only strategy which was allowed the choice of holding short equity positions. The fact that no constraints needed to be applied reinforces our claim that the PDR value provides a comprehensive measure of diversification. For the variables $r_m$ and $r_f$ we applied Damadoran's estimations of 6.95 and 1.23 respectively [15]. All formulas were optimized through numerical computation.

Table 3 shows how the in-sample historical portfolios selected by each strategy were evaluated by the various diversification measures associated with those strategies, including Diversification Ratio, Diversified Risk Parity (i.e. $N_{Ent}$), Uncorrelated Positive Bets, variance, and market beta (with the market defined as 1/N). It also provides the percentage of portfolio variance explained by the first principal portfolio (i.e. the market), the drawdown, overall returns and RDC. Figure 2 shows how the in-sample performance of the top-performing $N_{Ent}$ strategy compares to the market. Because the S&P 500 index is sampled in July 2011, it includes stocks which qualified for the index because of a rise over the training period and excludes those which dropped in value and were removed from the index. This future bias explains why the market benchmark achieves an unrealistically high RDC.

TABLE III. Cross-comparison of in-sample optimized portfolio performance for the diversification strategies

| | Max DR | Max $N_{Ent}$ | Max PDR | Min Variance | 1/N |
|---|---|---|---|---|---|
| DR | 3.07 | 1.52 | 1.63 | 2.26 | 1.50 |
| $N_{Ent}$ | 2.95 | 3.63 | 2.42 | 3.45 | 1.55 |
| PDR | 0.55 | 0.82 | 1.13 | 0.76 | 1.08 |
| Volatility | 1.41% | 1.84% | 1.65% | 0.97% | 1.64% |
| $\beta$ | 0.33 | 0.46 | 0.44 | 0.23 | 1.00 |
| 1st PP % | 59.9 | 82.9 | 88.5 | 68.1 | 99.6 |
| Drawdown | 15.0% | 20.1% | 20.0% | 22.0% | 17.0% |
| Return | 83.6% | 107% | 70.1% | 18.4% | 73.7% |
| RDC | 96.9 | 168 | 28.7 | 3.59 | 38.9 |

Table 4 shows how the portfolios selected by each strategy performed over the test period July 2011 to July 2013. Figure 3 compares the out-of-sample portfolio performance of the top performing $N_{Ent}$ strategy with the market benchmark.

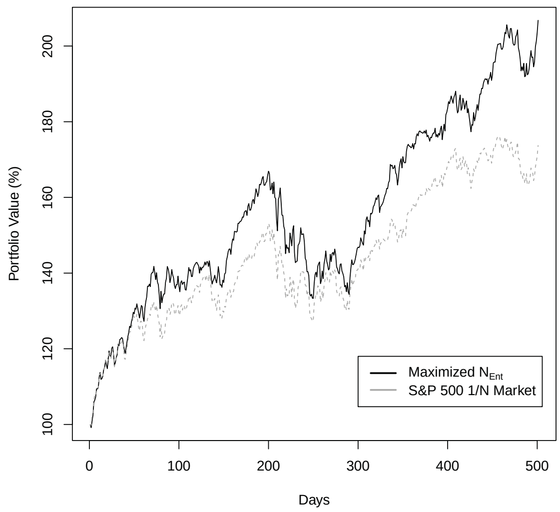


Fig. 2. In-sample performance of maximized $N_{Ent}$ against the market benchmark

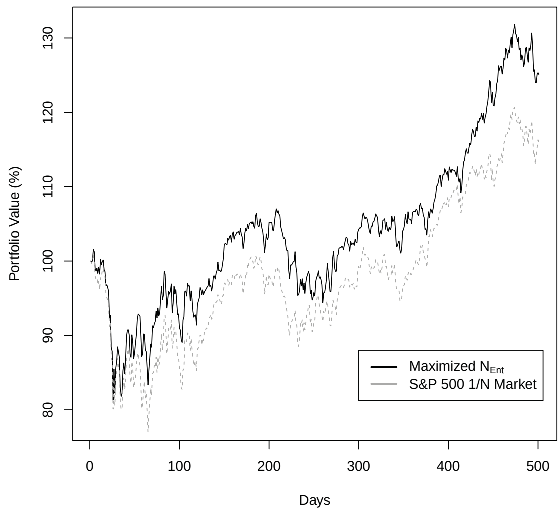


Fig. 3. Out-of-sample performance of maximized $N_{Ent}$ against the market benchmark

All of the diversified strategies have higher absolute returns than the market benchmark. In addition, they all have a higher RDC, supporting the idea that the market features diversification inefficiency. The maximized DR portfolio retains the highest DR in the test period. The maximized $N_{Ent}$ strategy also retains a higher $N_{Ent}$ than the market. Though the market beta has risen towards 1 for all strategies, the exposure to the first principal portfolio remains significantly below 100%, suggesting that some element of diversification has been sustained into the test period. However, the PDR values, which we have argued provide the most accurate quantification of positive diversification, are all below that of the market. These results indicate that the type of diversification on display involves idiosyncratic risk, which should not outperform the risk-free rate over the longer-term. The fact that the diversified strategies have outperformed the market may either be down to luck, or it may be the case that PDR is too conservative in its distinction between undiversifiable and idiosynractic risk. Further testing is required to differentiate between these possibilities. Either way, it should be noted that, as the PDR formula is the only one that can be optimized without constraints, it is the one that most closely matches the intuitive concept of diversification.

TABLE IV. CROSS-COMPARISON OF OUT-OF-SAMPLE OPTIMIZED PORTFOLIO PERFORMANCE FOR THE DIVERSIFICATION STRATEGIES

| | Max DR | Max $N_{Ent}$ | Max PDR | Min Variance | 1/N |
|---|---|---|---|---|---|
| DR | 3.04 | 2.16 | 2.05 | 2.64 | 2.11 |
| $N_{Ent}$ | 4.64 | 2.98 | 2.41 | 2.73 | 1.62 |
| PDR | 0.56 | 0.85 | 0.95 | 0.81 | 1.06 |
| Volatility | 1.21% | 1.21% | 1.26% | 1.17% | 1.21% |
| $\beta$ | 0.86 | 0.93 | 1.00 | 0.89 | 1.00 |
| 1st PP % | 75.6 | 86.5 | 92.9 | 85.8 | 99.7 |
| Drawdown | 28.1% | 20.0% | 22.4% | 21.5% | 23.8% |
| Return | 17.6% | 25.1% | 17.6% | 20.9% | 16.0% |
| RDC | 2.86 | 3.84 | 2.85 | 3.30 | 2.74 |

## VII. CONCLUSION

If the stock market was truly efficient, with people investing their money as cleverly as possible, then there would be no possibility of gaining an edge in diversification. With everybody competing to diversify as much as possible, no constructible portfolio of stocks would be more diversified than any other. Successful diversification allows investors to enhance their risk to reward ratio at the expense of others. In an efficient market, this should not occur.

The results from our experiment provide some evidence that the stock market contains diversification inefficiency. All of the strategies DR, DRP, UPB and MV outperformed the market, which supposedly represents the limit of undiversifiable risk. How can this be explained?

Much of the focus on earning profit from the stock market is based on attempts to predict its direction. The efficient market hypothesis (EMH) asserts that financial markets are informationally efficient (e.g. [16]). Informational efficiency entails that current prices reflect all publicly available information, and that prices instantly change to reflect new public information. Accordingly, all stocks, and indeed all portfolios of stocks, should follow a random walk, where knowledge of past events has no value for predicting future prices changes.

However, the notion of prediction efficiency advanced by the EMH is separate to that of diversification efficiency. Even if a market is prediction-efficient it may not be diversification-efficient: Sophisticated diversifiers may be drawing down a larger proportion of the long-term risk premium the market provides, leaving other naive diversifiers to shoulder risk without the expected rewards. Even if the market is totally unpredictable from day to day, the results from our experiment hint that investors who simply buy the market index may be transferring their risk premiums to more sophisticated diversifiers.

It may be the case that diversification inefficiency is harder to spot than prediction inefficiency. While predictable patterns in individual stocks can be recognized by individual traders without complex analysis, the diversification strategies discussed in this article have only recently come to light. Another possible explanation for the observation of diversification inefficiency in the stock market is that, while prediction inefficiency can be exploited to generate immediate profits, diversification inefficiency only delivers over the longer term, and hence may be overlooked by active traders.

If diversification inefficiency is real, then, by definition, sophisticated diversifiers are gaining more reward per unit of risk than naive diversifiers, who are taking on risk but failing to achieve the expected rewards. An interesting hypothesis is that the growth of diversification inefficiency may explain why stock market returns have fallen short of expectations in the 21st century. Maguire et al. [4] found that, when adjusted for selection bias, none of the stocks in the S&P 500 provided evidence of outperforming the risk-free rate over the last 10 years.

While U.S. equities delivered an average of 4.3% inflation-adjusted annual return during the 20th century, the S&P 500 has since failed to match the risk-free rate. If diversification inefficiency is being exploited by a subset of investors, this would explain the reduction in risk premium for investors who are not diversifying optimally (e.g. those who are simply buying the market index). It seems conceivable that increasing sophistication of diversification practiced by a subset of investors may be drawing down the bulk of the expected risk premium from the market, reducing the market index to a form of idiosyncratic risk.

## REFERENCES

[1] H. Lohre, H. Opfer, and G. Ország, "Diversifying risk parity," *Available at SSRN 1974446*, 2012.

[2] A. Meucci, "Managing diversification," *Risk*, vol. 2, pp. 74–79.

[3] Y. Choueifaty and Y. Coignard, "Toward maximum diversification," *The Journal of Portfolio Management*, vol. 35, no. 1, pp. 40–51, 2008.

[4] P. Maguire, P. Moser, J. McDonnell, R. Kelly, S. Fuller, and R. Maguire, "A probabilistic risk-to-reward measure for evaluating the performance of financial securities," in *Computational Intelligence for Financial Engineering & Economics (CIFEr)*. IEEE, 2013, pp. 1–8.

[5] W. F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The Journal of Finance*, vol. 19, no. 3, pp. 425–442, 1964.

[6] Y. Choueifaty, T. Froidure, and J. Reynier, "Properties of the most diversified portfolio," *Journal of Investment Strategies*, vol. 2, no. 2, pp. 49–70, 2011.

[7] C. Memmel and A. Kempf, "Estimating the global minimum variance portfolio," *Schmalenbach Business Review*, vol. 58, 2006.

[8] P. Jorion, "The pricing of exchange rate risk in the stock market," *Journal of Financial and Quantitative Analysis*, vol. 26, no. 3, pp. 363–376, 1991.

[9] V. K. Chopra and W. T. Ziemba, "The effect of errors in means, variances, and covariances on optimal portfolio choice," *The Journal of Portfolio Management*, vol. 19, no. 2, pp. 6–11, 1993.

[10] H. Lohre, U. Neugebauer, and C. Zimmer, "Diversified risk parity strategies for equity portfolio selection," *Journal of Investing*, vol. 21, no. 3, p. 111, 2012.

[11] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?" *Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 2009.

[12] P. Maguire, D. O'Sullivan, P. Moser, and G. Dunne, "Risk-adjusted portfolio optimisation using a parallel multi-objective evolutionary algorithm," in *Computational Intelligence for Financial Engineering & Economics (CIFEr)*. IEEE, 2012, pp. 1–8.

[13] J. Bai and S. Ng, "Determining the number of factors in approximate factor models," *Econometrica*, vol. 70, no. 1, pp. 191–221, 2002.

[14] W. F. Sharpe, "Mutual fund performance," *The Journal of Business*, vol. 39, no. 1, pp. 119–138, 1966.

[15] A. Damodaran, "Equity risk premiums (erp): Determinants, estimation and implications–the 2012 edition," *Stern School of Business, New York University*, 2012.

[16] B. G. Malkiel and E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.