A Modified Hierarchical Risk Parity Framework for Portfolio Management

Marat Molyboga

Marat Molyboga
is the director of research at Efficient Capital Management in Warrenville, IL.
molyboga@efficient.com

KEY FINDINGS

- The author introduces a modified hierarchical risk parity (MHRP) approach, which incorporates three popular portfolio management techniques into the hierarchical risk parity framework of López de Prado.
- He reports a striking improvement in the out-of-sample Sharpe ratio of 50%, on average, for portfolios of commodity trading advisors.
- The author argues that the MHRP framework has broad applications for portfolios of traditional and alternative investments.

ABSTRACT: *This article introduces a modified hierarchical risk parity (MHRP) approach that extends the HRP approach by incorporating three intuitive elements commonly used by practitioners. The new approach (1) replaces the sample covariance matrix with an exponentially weighted covariance matrix with Ledoit–Wolf shrinkage; (2) improves diversification across portfolio constituents both within and across clusters by relying on an equal volatility, rather than an inverse variance, allocation approach; and (3) improves diversification across time by applying volatility targeting to portfolios. The author examines the impact of the enhancements on portfolios of commodity trading advisors within a large-scale Monte Carlo simulation framework that accounts for the realistic constraints of institutional investors. The author finds a striking improvement in the out-of-sample Sharpe ratio of 50%, on average, along with a reduction in downside risk.*

TOPICS: ***Statistical methods, simulations, big data/machine learning****

*All articles are now categorized by topics and subtopics. **View at PM-Research.com.**

Portfolio selection has been a fruitful area of research since the introduction of the parsimonious theory of Markowitz (1952), which reduces a complex asset allocation problem to a simple calculation that relies solely on the vector of expected returns and the covariance matrix. Unfortunately, as discussed by Michaud (1989) and Best and Grauer (1991), this beautiful theory produces ugly results, as evidenced by poor out-of-sample performance and extreme, unstable portfolio weights. The issue of instability is driven by the high sensitivity of the Markowitz portfolio weights to estimation error.[1] This difficulty is further

[1] Markowitz portfolios are obtained by solving a quadratic problem, which requires the inversion of a covariance matrix. Bailey and López de Prado (2012) and Yang et al. (2016) showed that the magnitude of the sensitivity issue can be assessed using the condition number of the covariance matrix. If the covariance matrix is near singular, the condition number, defined as the ratio of the largest to the smallest eigenvalues of

compounded by the impossible problem of estimating the vector of expected returns with a high degree of precision, as noted by Merton (1980).

Although many articles have attempted to reduce the impact of estimation error, DeMiquel, Garlappi, and Uppal (2009) showed that even very sophisticated approaches[2] fail to outperform a naive 1/N portfolio and argued that the estimation window required to capture the potential gains of optimal portfolios is too long. The failure of mean–variance optimization has led to the emergence of the popular risk-parity approach discussed in detail by Maillard, Roncalli, and Teiletche (2010). The risk-parity approach ignores return forecasts and instead attempts to maximize diversification by allocating risk equally across portfolio constituents. Despite its popularity, the risk parity approach is subject to the same weaknesses as mean–variance optimization, namely unstable weights caused by an inversion of a near singular covariance matrix and a lack of alignment with the top-down perspective in portfolio management typically employed by sophisticated practitioners. The hierarchical risk parity (HRP) approach introduced by López de Prado (2016) elegantly addresses those two weaknesses, establishing it as the cutting-edge application of machine learning to portfolio management.

The purpose of this article is to examine whether the HRP approach can be further improved by adding three intuitive elements commonly used by practitioners. The modified approach, MHRP, (1) replaces the sample covariance matrix with an exponentially weighted covariance matrix with Ledoit–Wolf shrinkage; (2) improves diversification across portfolio constituents both within and across clusters by relying on an equal volatility approach rather than an inverse variance approach; and (3) improves diversification across the dimension of time by applying volatility targeting to portfolios.

Although the proposed improvements represent relatively minor changes to HRP from a theoretical perspective, the practical implications for investors can be meaningful because the individual components have been tested across many asset classes and market cycles and subsequently embraced by practitioners for their robustness and solid theoretical underpinnings. For example, exponential weighting takes into account the presence of heteroskedasticity in financial returns, which is at the heart of RiskMetrics' volatility estimation, as discussed by JPMorgan and Reuters (1996), while also being closely related to the illustrious generalized autoregressive conditional heteroskedasticity (GARCH) model of Bollerslev (1986). Similarly, diversification has been proclaimed by the Nobel Prize winner Harry Markowitz to be "the only free lunch in finance." Improvements in performance as a result of greater diversification across portfolio constituents with equal risk allocation and across time with portfolio volatility targeting have been reported in many previous studies, such as those by Hallerbach (2012), Kim, Tse, and Wald (2016), Daniel and Moskowitz (2016), and Moreira and Muir (2017).

I examine the marginal impact of each component and their combined impact on the performance of portfolios of commodity trading advisors (CTAs) within the large-scale simulation framework of Molyboga and L'Ahelec (2016) that accounts for the realistic constraints faced by institutional investors. The framework is imposed on a BarclayHedge CTA sample of 528 live and 1,113 defunct funds over the 2002–2016 period. I find that each enhancement improves out-of-sample Sharpe ratios of multi-CTA portfolios by 13% to 19%, on average. Moreover, when all three enhancements are combined into a unified MHRP approach, I observe a striking improvement in the out-of-sample Sharpe ratio of 50%, on average, with a meaningful reduction in downside risk.

Finally, I discuss whether CTA investors can capture the benefit of MHRP after accounting for real-life constraints such as trading frictions caused by monthly rebalancing and portfolio leverage changes. I show that although the frictions are potentially high for fund investors because of a time gap between a redemption from one fund and a subscription into another fund, the frictions can be mitigated if the investor uses managed accounts that employ notional funding and benefit from more favorable liquidity terms. Although the performance benefits of the MHRP approach should be carefully evaluated within other types of portfolios that have their own sets of investments and are subject to their own unique real-life constraints, the three

the covariance matrix, is large and the portfolio weights are highly sensitive to estimation error.

[2]DeMiquel, Garlappi, and Uppal (2009) considered 14 models that include the sample-based mean–variance model and its extensions designed to reduce estimation error. Two of the models they considered are the Pastor (2000) model, which produces Bayesian portfolios based on belief in an asset-pricing model, and the MacKinlay and Pastor (2000) approach, which produces portfolios implied by asset-pricing models with unobservable factors.

proposed enhancements are likely to have a positive impact on the portfolio performance of both traditional and alternative investments because they improve the quality of the covariance matrix estimates and the degree of diversification across portfolio constituents and time. The rest of the article is organized as follows. The first section describes the HRP approach and the three enhancements that comprise the MHRP approach. The second section describes the data and the empirical framework of Molyboga and L'Ahelec (2016). The third section presents empirical results, and the fourth section concludes.

FROM HRP TO MHRP

As discussed by López de Prado (2016), most bottom-up approaches to portfolio construction, such as mean–variance optimization or risk parity, tend to rely heavily on each value of a correlation matrix, implying that any two securities are potential substitutes for one another. This assumption is inconsistent with the investment practices of investment professionals, particularly in light of the work of Brinson, Hood, and Beebower (1995), who highlighted the importance of asset allocation decisions. If an investor attempts to build a 60–40 portfolio of stocks and bonds, which is a typical starting point for a US institutional investor according to Anson (2011), the investor would group stocks and bonds separately rather than consider individual stocks and bonds as substitutes for each other. López de Prado (2016) showed that considering a hierarchical tree structure allows for focusing on a small number of relationships that are consistent with the top-down perspective of investors who build their portfolios by starting at the asset-class level and then going down to the level of the individual securities.

As shown by López de Prado (2016), the HRP approach follows three steps[3]:

1. **Tree clustering.** Hierarchical clustering is performed using a sample correlation matrix. This step highlights important top-down relationships among individual portfolio constituents and their clusters.
2. **Quasi-diagonalization.** Similar investments are placed together and dissimilar investments are placed far apart. This procedure reshuffles rows and columns of the original correlation matrix so that the largest values are close to the diagonal and the smallest values are away from the diagonal.
3. **Recursive bisection.** This step allocates across and within clusters using an inverse variance allocation (IVA) approach.

In this article, I enhance the HRP approach by adding three intuitive elements commonly used by practitioners. This modified approach (1) replaces the sample covariance matrix with an exponentially weighted covariance matrix with Ledoit–Wolf shrinkage, (2) improves diversification across portfolio constituents both within and across clusters by relying on an equal volatility allocation (EVA) approach rather than an IVA approach, and (3) improves diversification across the dimension of time by applying volatility targeting to portfolios.

It is important to distinguish the volatility weighting of portfolio constituents (which closely relates to the concept of diversification across portfolio constituents) from targeting portfolio volatility (which is linked to diversification across time). Volatility weighting is a constant leverage approach highlighted by Maillard, Roncalli, and Teiletche (2010) and Hallerbach (2012) that allocates risk equally across portfolio constituents. By contrast, volatility targeting dynamically scales aggregate portfolio leverage to achieve constant expected portfolio volatility or equally allocates risk across time rather than among portfolio constituents.

Covariance Matrix Estimation

Exponential weighting is used pervasively in academia and industry to account for the presence of heteroskedasticity in financial returns. The exponentially weighted moving-average approach is at the heart of the volatility estimation employed by RiskMetrics, as discussed by JPMorgan and Reuters (1996), and is closely related to the GARCH model of Bollerslev (1986). Likewise, Pafka, Potters, and Kondor (2004) extensively examined applications of exponentially weighted covariance matrices to portfolio optimization problems.

One of the most common approaches used to improve the estimation of covariance matrixes is based on the shrinkage approach of Ledoit and Wolf (2004).

[3] See López de Prado (2016) for a detailed description of the HRP approach and the Python implementation.

I follow Molyboga (2019) and combine shrinkage estimation with exponential weighting of the covariance matrix.

Better Diversification across Portfolio Constituents

This study considers an IVA approach and an EVA approach. The two approaches can be contrasted as follows:

- IVA is a methodology in which an allocation weight to portfolio constituent i is inversely related to the estimated variance of the constituent, $\hat{\sigma}_i^2$, as follows:

$$w_i^{IVA} = \frac{1/\hat{\sigma}_i^2}{\sum_{j=1}^{N} 1/\hat{\sigma}_j^2}.$$

- EVA is an equal volatility (or inverse volatility) allocation wherein

$$w_i^{EVA} = \frac{1/\hat{\sigma}_i}{\sum_{j=1}^{N} 1/\hat{\sigma}_j}.$$

López de Prado (2016) suggested that IVA is optimal when the covariance matrix is diagonal because it produces the minimum variance portfolio. Because investors care about risk-adjusted performance measures such as the Sharpe ratio, IVA is likely suboptimal because it only focuses on minimizing risk rather than on optimizing the trade-off between return and risk. Section A in the online appendix considers two potential assumptions regarding portfolio constituents. It shows that if all portfolio constituents have the same expected return, minimization of variance is equivalent to Sharpe maximization and IVA maximizes the Sharpe ratio. However, if the portfolio constituents have similar Sharpe ratios, which is more likely than similar expected returns, EVA is optimal.

Practitioners often prefer EVA because it attempts to maximize diversification by equally allocating risk, measured in terms of volatility, across portfolio constituents, and it intuitively responds to changes in the volatility levels of portfolio constituents.[4] To demonstrate the utility of this approach, let us consider a simple hypothetical example in which fund A exhibits annualized volatility of 20% and fund B realizes annualized volatility of 10%. Because fund A is twice as risky as fund B, an investor who attempts to diversify risk equally across the two funds would allocate one-third of the investment amount to fund A and two-thirds to fund B ($1/3 \times 20\% = 2/3 \times 10\%$). By contrast, the IVA approach would give a one-fifth allocation to fund A and a four-fifths allocation to fund B—an allocation of twice as much risk to fund B as to fund A ($2 \times [1/5 \times 20\%] = 4/5 \times 10\%$).

CTAs commonly offer their trading programs at multiple volatility levels, capitalizing on the free leverage embedded in futures and forward contracts that are traded on margin. If manager A decides to offer a 10% version of the program, both funds run at the same volatility and both the EVA and IVA approaches will allocate risk equally between the funds. Thus, EVA is more attractive to practitioners because it preserves risk allocations across funds regardless of the choice of volatility levels while attempting to maximize diversification by allocating risk equally across portfolio constituents.

Better Diversification across the Dimension of Time

Previous research has suggested that managing portfolio volatility improves the performance of risk premiums and active strategies. Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016) showed that volatility targeting, or adjusting exposure to target a constant ex ante volatility, nearly doubles the Sharpe ratio of cross-sectional momentum. Kim, Tse, and Wald (2016) reported that the abnormal returns of time-series momentum are largely driven by scaling returns to volatility. Moreira and Muir (2017) found evidence for the benefit of managing the volatility of portfolios across a broad array of risk premiums, such as market, value, currency carry, and betting against beta.

Molyboga (2019) provided theoretical underpinnings for volatility targeting and investigated empirically whether, and under what conditions, it improves the performance of multi-CTA portfolios. Molyboga (2019) reported that volatility targeting improves the out-of-sample, net-of-fee Sharpe ratio of portfolios with 10 CTAs by approximately 13%–14%, on average. This improvement in the Sharpe ratio represents approximately 63–68

[4] As shown in Section A of the online appendix, this approach is optimal when portfolio constituents have equal Sharpe ratios.

bps of return per annum for a 15% volatility portfolio. As the number of managers in the portfolio varies between 5 and 20, volatility targeting consistently improves performance by 0.53% to 0.80% per annum, with larger portfolios achieving greater performance enhancements. Section B in the online appendix describes how volatility targeting is accomplished through dynamically adjusting portfolio leverage.

Portfolio Construction Approaches

To examine the marginal impact of each enhancement and their combined impact on performance, I consider the following four portfolio construction approaches:

- *HRP* is the original hierarchical risk parity approach of López de Prado (2016).
- HRP_C is an HRP approach that relies on the exponentially weighted covariance matrix with Ledoit–Wolf shrinkage.
- HRP_{CE} is an HRP approach that relies on the exponentially weighted covariance matrix with Ledoit–Wolf shrinkage and uses the EVA instead of the IVA to diversify within and across clusters.
- *MHRP* is a modified HRP approach that relies on the exponentially weighted covariance matrix with Ledoit–Wolf shrinkage, uses the EVA instead of the IVA to diversify within and across clusters, and targets a constant portfolio volatility of 12%.

The marginal impact of the covariance matrix estimation is evaluated by comparing the performance of $\mathrm{HRP_C}$ against HRP. The marginal improvement due to superior diversification across and within clusters is measured by comparing the performance of $\mathrm{HRP_{CE}}$ against that of $\mathrm{HRP_C}$. The marginal contribution of volatility targeting is measured by comparing the performance of MHRP against $\mathrm{HRP_{CE}}$. Finally, the combined impact of the three enhancements is captured by comparing the performance of MHRP against HRP.

DATA AND EMPIRICAL FRAMEWORK

In this section, I describe the data used in the study and the empirical framework of Molyboga and L'Ahelec (2016) used to evaluate the portfolio management approaches.

Data

I rely on the BarclayHedge database, the largest and highest quality publicly available database of CTAs, according to Joenvaara, Kosowski, and Tolonen (2012). I include the graveyard database of defunct funds to account for survivorship bias and adjust for backfill/incubation bias by using the first reported date, as recommended by Bhardwaj, Gorton, and Rowenhorst (2014), Foran et al. (2017), and Jorion and Schwarz (2017).[5] Bhardwaj, Gorton, and Rowenhorst (2014) showed that the common adjustment of removing 24 months of returns is inadequate for CTAs because the remaining bias is still more than 1% per annum in a value-weighted index and almost 3% in an equally weighted index. Because the first reported date in the BarclayHedge database is December 2002, our dataset is limited to December 2002 through December 2017. Although the selection bias that results from funds choosing not to report to public databases is the most difficult to avoid, Edelman, Fung, and Hsieh (2013) demonstrated that it is likely to be insignificant.[6] I use the liquidation estimate of 1% as suggested by Ackermann, McEnally, and Ravenscraft (1999). Because CTA investments are often chosen for their diversification benefit during crisis periods, as summarized by Molyboga and L'Ahelec (2016), I exclude fund managers that are most susceptible to market crisis, such as those that predominantly rely on options or arbitrage.[7]

The final bias-free dataset includes 576 active and 1,187 defunct funds with net returns between December

[5] The voluntary nature of self-reporting to hedge fund databases results in the backfill/inclubation bias. Before reporting to public databases, funds often go through an incubation period during which they build a track record using proprietary capital. If the track record is attractive, fund managers report their performance to a database and often backfill the returns generated before inclusion in the database. This practice results in an upward bias because funds with poor performance are unlikely to report their performance to the database.

[6] There are two types of funds that choose not to report in public databases: poorly performing funds that cannot attract new investors based on their track record and successful funds that do not rely on the commercial databases to raise assets. Edelman, Fung, and Hsieh (2013) showed that the two sources of nonreporting bias cancel each other out.

[7] I use the BarclayHedge classification field to remove funds with three classifications: option strategies, stock index/option strategies, and arbitrage. I repeat analysis after including the three classifications and find qualitatively similar results.

Exhibit 1
Summary Statistics of CTAs for 2002–2017

	Defunct	Live	All
Mean	0.05%	0.32%	0.14%
Median	–0.02%	0.27%	0.08%
St Dev	3.69%	3.85%	3.74%
Skewness	0.12	0.17	0.14
Excess kurtosis	1.09	1.10	1.09
Autocorrelation	–0.01	–0.07	–0.03
Average no. of managers	325	226	551
Average track record length	50	71	57
Average AUM	164	819	379
Total no. of managers	1,187	576	1,763
No. of observations	58,829	40,978	99,807

2002 and December 2017. I use monthly values of the three-month T-bill secondary market rate (ticker: TB3MS) from FRED as a proxy for the risk-free rate.

Exhibit 1 reports the average summary statistics for the defunct managers, live managers, and all managers in the CTA sample. The statistics include mean monthly excess return, median monthly excess return, monthly standard deviation, skewness, excess kurtosis, autocorrelation of order one, average number of managers, average track record length, average assets under management (AUM) in millions of US dollars, total number of managers, and number of observations.

Consistent with the findings of previous studies, live managers tend to perform better, stay in business longer, and attract more assets than defunct managers. The mean monthly excess return of live managers is 0.32%, which at the 99% significance level exceeds the 0.05% mean monthly excess return of the defunct managers. The average track record length of live and defunct managers is 71 and 57 months, respectively. The average AUM of live and defunct managers is US$819 million and US$164 million, respectively. The high mortality rate among CTA funds is indicated by the large share of defunct funds in the database. Out of 1,763 funds in the sample, 1,187 funds are defunct and only 576 are live.

Empirical Analysis

In this section, I describe the large-scale simulation framework of Molyboga and L'Ahelec (2016). As discussed in the previous section, this article attempts to improve on the original HRP approach of López de Prado (2016) by adding three intuitive enhancements commonly used by practitioners. The MHRP approach (1) replaces a sample covariance matrix with an exponentially weighted covariance matrix with Ledoit–Wolf shrinkage, (2) improves diversification across portfolio constituents within and across clusters by relying on an EVA rather than an IVA approach, and (3) improves diversification across the dimension of time by volatility targeting portfolios. To examine the marginal impact of each enhancement and their combined impact on performance, I consider the four portfolio construction approaches: HRP, HRP_C, HRP_{CE}, and MHRP.

In this article, I apply the large-scale framework of Molyboga and L'Ahelec (2016) with 2,000 simulations for the out-of-sample period between January 2006 and December 2017. The simulation framework is designed to evaluate portfolio construction approaches subject to the realistic constraints of institutional investors. The first allocation decision is made at the end of December 2005. Because of the delay in CTA reporting documented by Molyboga, Baek, and Bilson (2017), an institutional investor can rely on information only through November 2005 for the January 2006 allocation. The simulation accounts for two standard constraints that institutional investors face by considering only funds with at least 36 months of returns and excluding all funds in the bottom quintile of AUM among the funds considered. The simulation then randomly selects 15 funds from the available pool of funds and applies the four portfolio construction approaches. For each portfolio approach, monthly returns are recorded for January 2006 using the liquidation bias adjustment, if needed. At the end of January 2006, the available fund pool is updated to comply with the previously noted requirements of institutional investors. All funds from the original portfolio that no longer meet the requirements are randomly replaced with funds from the new pool. The portfolio turnover is relatively low because the funds that continue meeting the selection criteria remain in the portfolio. Each portfolio is rebalanced using one of the four methodologies, and the process is repeated for the entire out-of-sample period covering January 2006 through December 2017. A single simulation run produces four time series of out-of-sample performance, one for each portfolio construction approach. Distributions of out-of-sample returns are constructed based on the results of 2,000 simulations.

EXHIBIT 2
Annual Statistics of CTAs

Date	AUM	CTAs
2006	11.88	148
2007	10.54	156
2008	12.21	171
2009	12.78	200
2010	14.82	222
2011	14.02	244
2012	12.20	249
2013	10.87	242
2014	10.45	223
2015	10.01	213
2016	9.67	215
2017	10.21	223

Exhibit 2 reports the average AUM threshold level for each year (in millions of US dollars) and the average number of active funds that have at least 36 months of returns and meet the AUM threshold. Their time-series variation reflects the dynamic nature of the CTA universe.

EMPIRICAL RESULTS

In this section, I present empirical out-of-sample results obtained within the large-scale simulation framework. I start by measuring the performance impact of improved covariance matrix estimation, diversification across portfolio constituents, and volatility targeting. At this stage, I largely focus on the average impact on out-of-sample performance, but I also consider the quartiles of the distributions of Sharpe ratios. I then evaluate the combined improvement of the enhancements by comparing the MHRP and the HRP approaches. Finally, I discuss implications for CTA investors as well as broad applications to other types of traditional and alternative investments.

Marginal Impact of the New Covariance Matrix Estimation Approach

The large-scale simulation framework produces distributions of out-of-sample performance. Molyboga, Baek, and Bilson (2017) introduced a methodology for comparing distributions that relies on utility functions and stochastic dominance. In this article, I primarily focus on means and medians but also consider distribution quantiles based on this methodology. I start the analysis by comparing the performance of the original HRP approach that uses the sample correlation matrix for clustering and calculating volatilities with the HRP_C approach that relies on the correlation matrix estimated with exponential weighting and Ledoit and Wolf (2004) shrinkage.

EXHIBIT 3
Statistics of the Distributions of Sharpe Ratios for HRP and HRP_C

	HRP	HRP_C
Mean	0.217	0.248*
St Dev	0.255	0.227
Max	1.155	1.168
Third Quartile	0.386	0.403
Median	0.216	0.247
First Quartile	0.045	0.094
Min	–0.641	–0.474

*Notes: *Significant at 1% level.*

Exhibit 3 reports the mean, standard deviation, and quantiles of the distributions of Sharpe ratios, including the minimum, first quartile, median, third quartile, and maximum values. The distributions of Sharpe ratios are generated for the two portfolio construction approaches of HRP and HRP_C within the large-scale simulation framework. The out-of-sample period covers January 2006 to December 2017. The P-values of the differences in the means of the Sharpe ratios are calculated using the bootstrap approach of Molyboga and L'Ahelec (2016).

The improvement due to the superior covariance matrix estimation is meaningful in both economic and statistical terms. The four quartiles of the HRP_C method are consistently higher than those of the original HRP allocation of López de Prado (2016), and the standard deviation is lower. The mean Sharpe ratio of 0.248 for the HRP_C approach is approximately 14.2% higher than the mean Sharpe ratio of 0.217 for the standard HRP approach. This difference is significant at the 99% confidence level[8] and highlights the importance of de-noising correlation matrixes. Although the suggested covariance

[8] Statistical inference uses the bootstrap methodology of Molyboga and L'Ahelec (2016). Because the simulations are not independent, a bootstrap approach is required to eliminate the bias inherent in standard tests, as discussed in detail by Molyboga, Baek, and Bilson (2017).

EXHIBIT 4

Statistics of the Distributions of Sharpe Ratios for HRP_C and HRP_{CE}

	HRP_C	HRP_{CE}
Mean	0.248	0.294*
St Dev	0.227	0.146
Max	1.168	0.853
Third Quartile	0.403	0.390
Median	0.247	0.291
First Quartile	0.094	0.199
Min	–0.474	–0.203

*Notes: *Significance at 1% level.*

matrix estimation technique that relies on exponential weighting and Ledoit and Wolf (2004) shrinkage improves the out-of-sample performance of the original HRP approach, future research can extend these findings by investigating whether alternative approaches to de-noising the correlation matrixes, such as the random matrix theory approach of Pafka, Potters, and Kondor (2004), can further improve performance.

Marginal Impact of Improved Diversification across Portfolio Constituents

Having established the benefit of employing an exponentially weighted covariance matrix with shrinkage estimation, I turn to the performance impact of portfolio diversification across portfolio constituents. As discussed earlier, EVA is more attractive to practitioners than IVA because it preserves risk allocations across funds regardless of their choice of volatility levels and attempts to maximize diversification by allocating risk equally across portfolio constituents.

Exhibit 4 reports the mean, standard deviation, and quantiles of the distribution of Sharpe ratios, including the minimum, first quartile, median, third quartile, and maximum values. The distributions of Sharpe ratios are generated for the two portfolio construction approaches of HRP_C and HRP_{CE} within the large-scale simulation framework.

The improvement due to the superior diversification across portfolio constituents is also strong, and the standard deviation of the distribution decreases by about one-third; however, the improvement is less consistent across the distribution of Sharpe ratios. The performance improves for the bottom three quartiles, with a slight degradation for the top quartile. The mean Sharpe ratio of 0.294 for the HRP_{CE} approach that relies on EVA across and within clusters is approximately 18.6% higher than the mean Sharpe ratio of 0.248 for the HRP_C approach that uses IVA. This difference is significant at the 99% confidence level and suggests that an increase in portfolio diversification within the HRP approach leads to superior out-of-sample performance.

Marginal Impact of Improved Diversification across Time

Finally, I examine the impact of improving diversification across time with volatility targeting. As a reminder, it is important to distinguish volatility weighting of portfolio constituents, which closely relates to the concept of diversification across portfolio constituents, from targeting portfolio volatility, which is linked to diversification across time. Although EVA distributes risk equally across portfolio constituents while maintaining constant leverage at the portfolio level, volatility targeting dynamically scales aggregate portfolio leverage to achieve constant expected portfolio volatility by allocating risk equally across time rather than portfolio constituents (see section B in the online appendix for details).

Exhibit 5 reports the mean, standard deviation, and quantiles of the distribution of Sharpe ratios, including the minimum, first quartile, median, third quartile, and maximum values. The distributions of Sharpe ratios are generated for the two portfolio construction approaches of HRP_{CE} and MHRP within the large-scale simulation framework.

The improvement due to the superior diversification across time is consistent across the distribution of Sharpe ratios, except for a few outliers in the left tail. The mean Sharpe ratio of 0.333 for the MHRP approach that relies on volatility targeting to diversify across time is approximately 13.1% higher than the mean Sharpe ratio of 0.294 for the HRP_{CE} approach. This difference is significant at the 99% confidence level and suggests that volatility targeting improves performance. This finding is consistent with those of Kim, Tse, and Wald (2016) and Moreira and Muir (2017), who reported a positive impact from volatility targeting.

EXHIBIT 5

Statistics of the Distributions of Sharpe Ratios for HRP_{CE} and MHRP

	HRP_{CE}	MHRP
Mean	0.294	0.333
St Dev	0.146	0.149
Max	0.853	0.916
Third Quartile	0.390	0.431
Median	0.291	0.336
First Quartile	0.199	0.234
Min	–0.203	–0.211

Combined Impact of the Three Enhancements: MHRP versus HRP

Having established that each of the three enhancements improves out-of-sample performance by 13% to 19% on average, I examine their combined impact on performance by comparing the MHRP approach to the original HRP portfolio of López de Prado (2016).

Exhibit 6 reports the mean, standard deviation, and quantiles of the distribution of Sharpe ratios, including the minimum, first quartile, median, third quartile, and maximum values. The distributions of Sharpe ratios are generated for the two portfolio construction approaches of HRP and MHRP within the large-scale simulation framework. The improvement is striking, with the mean Sharpe ratio increasing by more than 50% and the standard deviation decreasing by approximately 40%.

Distributions of Sharpe ratios can be visualized using box and whiskers plots. The breadth of each distribution represents the role of chance in manager selection, which has been highlighted by Molyboga and L'Ahelec (2016).

Exhibit 7 shows the distributions of the Sharpe ratios of the hypothetical portfolios, generated within the large-scale simulation framework for the out-of-sample period of January 2006–December 2017. The plot shows the same pattern: The MHRP approach improves out-of-sample performance, on average, while reducing downside risk.

This finding is striking, yet intuitive, because the MHRP approach combines the structural benefits of the HRP approach of López de Prado (2016) and the practical ideas of improved covariance matrix estimation and diversification across time and portfolio constituents.

EXHIBIT 6

Statistics of the Distributions of Sharpe Ratios for HRP and MHRP

	HRP	MHRP
Mean	0.217	0.333*
St Dev	0.255	0.149
Max	1.155	0.916
Third Quartile	0.386	0.431
Median	0.216	0.336
First Quartile	0.045	0.234
Min	–0.641	–0.211

*Notes: *Significance at 1% level.*

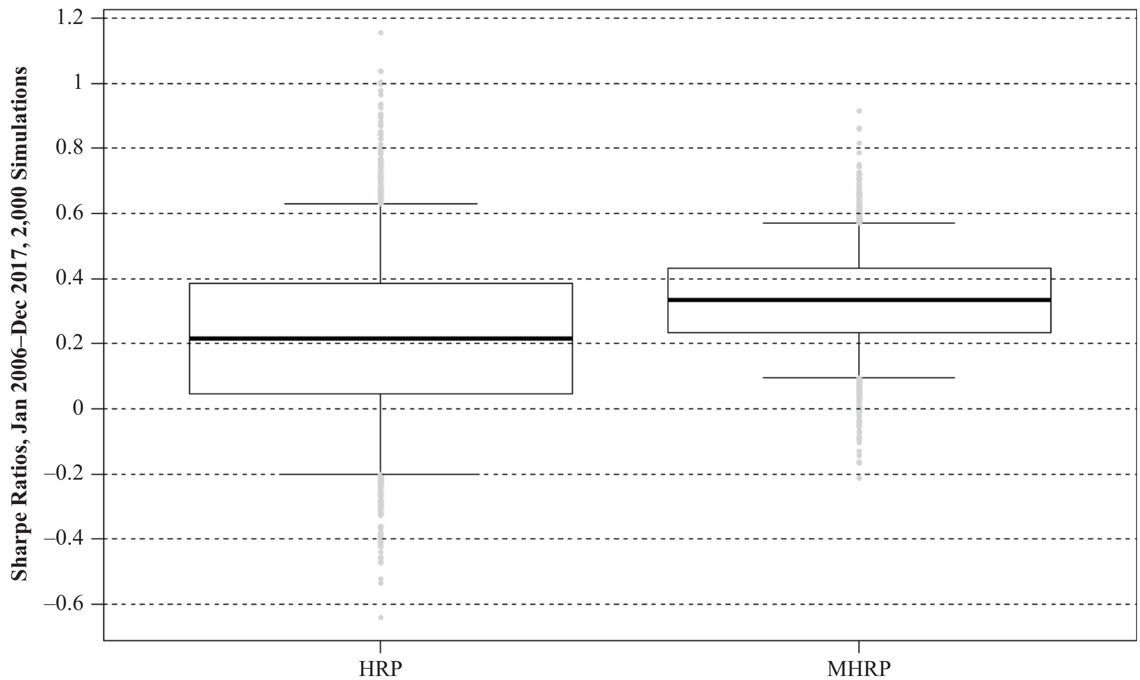
The HRP method is a cutting-edge application of machine learning to portfolio management that overcomes the issues of unstable weights caused by an inversion of a near singular covariance matrix and a lack of alignment with the top-down portfolio management typically employed by sophisticated practitioners. The practical ideas tested here are simple yet commonplace among practitioners because of their robust performance benefits. As a result, the MHRP approach can be a potentially attractive portfolio management technique for any investor, provided its benefits are robust to real-life constraints.

Implications for Investors

In this section, I discuss the implications of using MHRP for investors in CTAs and other types of investments. The empirical results are relevant for CTA investors because the methodology of Molyboga and L'Ahelec (2016) accounts for several important constraints that institutional investors experience, such as size and length of track record, while seeking to produce implementable results by adjusting for data biases and eliminating the look-ahead bias that is present in most hedge fund studies. However, as discussed in detail by Molyboga (2019), the monthly rebalancing and monthly leverage changes that are required for implementing the MHRP approach are available for managed account investors but not for fund investors.[9]

[9] Fund investors cannot easily adjust portfolio leverage and face time gaps between a redemption from one fund and a subscription into another fund.

EXHIBIT 7
Out-of-Sample Sharpe Ratios of HRP and MHRP



Although the performance benefits of the MHRP approach should be carefully evaluated for other types of portfolios with their own sets of investments and real-life constraints, the three proposed enhancements are likely to have a positive impact on the portfolio performance of traditional and alternative investments because they improve the quality of the covariance matrix estimation and the diversification across time and portfolio constituents.

CONCLUDING REMARKS

In this article, I extend the HRP approach of López de Prado (2016) by adding three intuitive elements commonly used by practitioners. The proposed MHRP approach (1) replaces a sample covariance matrix with an exponentially weighted covariance matrix with Ledoit–Wolf shrinkage, (2) improves diversification across portfolio constituents within and across clusters by relying on an equal volatility allocation approach rather than an IVA approach, and (3) improves diversification across time by volatility targeting portfolios.

I find that MHRP improves the out-of-sample Sharpe ratios of multi-CTA portfolios by more than 50%, on average, while reducing downside risk. This performance improvement is relevant for institutional investors because it is implementable within a managed account investment and the empirical methodology thoroughly accounts for real-life constraints.

Future research can be dedicated to investigating the performance benefits of MHRP for portfolios of traditional and alternative investments and whether it can be further improved by employing alternative approaches to de-noising the correlation matrixes, such as the random matrix theory approach of Pafka, Potters, and Kondor (2004).

ACKNOWLEDGMENTS

I am grateful to Andrew Berkin, Adam Duncan, Frank Fabozzi, Christophe L'Ahelec, Marcos López de Prado, David Mascio, Sanida Mustafovic, Kevin Townsend, and Stephan Zeisberger for helpful comments.

REFERENCES

Ackermann, C., R. McEnally, and D. Ravenscraft. 1999. "The Performance of Hedge Funds: Risk, Return, and Incentives." *The Journal of Finance* 54 (3): 833–874.

Anson, M. 2011. "The Evolution of Equity Mandates in Institutional Portfolios." *The Journal of Portfolio Management* 37 (4): 127–137.

Bailey, D., and M. López de Prado. 2012. "Balanced Baskets: A New Approach to Trading and Hedging Risks." *Journal of Investment Strategies* 1 (4): 21–62.

Barroso, P., and P. Santa-Clara. 2015. "Momentum has its Moments." *Journal of Financial Economics* 116: 111–120.

Best, M. J., and R. R. Grauer. 1991. "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results." *The Review of Financial Studies* 4: 315–342.

Bhardwaj, G., G. B. Gorton, and K. G. Rowenhorst. 2014. "Fooling Some of the People All of the Time: The Inefficient Performance and Persistence of Commodity Trading Advisors." *The Review of Financial Studies* 27: 3099–3132.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31 (3): 307–327.

Brinson, G. P., L. R. Hood, and G. L. Beebower. 1995. "Determinants of Portfolio Performance." *Financial Analysts Journal* 51 (1): 133–138.

Daniel, T., and T. J. Moskowitz. 2016. "Momentum Crashes." *Journal of Financial Economics* 122: 221–247.

DeMiquel, V., L. Garlappi, and R. Uppal. 2009. "Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?" *The Review of Financial Studies* 22: 1915–1953.

Edelman, D., W. Fung, and D. A. Hsieh. 2013. "Exploring Unchartered Territories of the Hedge Fund Industry: Empirical Characteristics of Mega Hedge Fund Returns." *Journal of Financial Economics* 109: 734–758.

Foran, J., M. C. Hutchinson, D. F. McCarthy, and J. O'Brien. 2017. "Just a One Trick Pony? An Analysis of CTA Risk and Return." *The Journal of Alternative Investments* 20 (2): 8–26.

Hallerbach, W. 2012. "A Proof of Optimality of Volatility Weighting over Time." *Journal of Investment Strategies* 1: 87–99.

Joenvaara, J., R. Kosowski, and P. Tolonen. 2012. "New 'Stylized Facts' About Hedge Funds and Database Selection Bias." Working paper, University of Oulu, 2012.

Jorion, P., and C. Schwarz. 2017. "The Fix Is In: Properly Backing Out Backfill Bias." Working paper.

JPMorgan and Reuters. *RiskMetrics—Technical Document*, 4th ed. New York: JPMorgan, 1996.

Kim, A. Y., Y. Tse, and J. K. Wald. 2016. "Time Series Momentum and Volatility Scaling." *Journal of Financial Markets* 30: 103–124.

Ledoit, O., and L. Wolf. 2004. "Honey, I Shrunk the Sample Covariance Matrix." *The Journal of Portfolio Management* 30: 110–119.

López de Prado, M. 2016. "Building Diversified Portfolios That Outperform Out of Sample." *The Journal of Portfolio Management* 42 (4): 56–69.

MacKinlay, A. C., and L. Pastor. 2000. "Asset Pricing Models: Implications for Expected Returns and Portfolio Selection." *The Review of Financial Studies* 13: 883–916.

Maillard, S., T. Roncalli, and J. Teiletche 2010. "The Properties of Equally Weighted Risk Contribution Portfolios." *The Journal of Portfolio Management* 36: 60–70.

Markowitz, H. M. 1952. "Portfolio Selection." *The Journal of Finance* 7: 77–91.

Merton, R. C. 1980. "On Estimating the Expected Return on the Markets: An Exploratory Investigation." *Journal of Financial Economics* 8: 323–361.

Michaud, R. O. 1989. "The Markowitz Optimization Enigma: Is Optimized Optimal? *Financial Analysts Journal* 45: 31–42.

Molyboga, M. 2019. "Portfolio Management of Commodity Trading Advisors with Volatility Targeting." *Journal of Investment Strategies*, forthcoming.

Molyboga, M., S. Baek, and J. Bilson. 2017. "Assessing Hedge Fund Performance with Institutional Constraints: Evidence from CTA Funds." *Journal of Asset Management* 18: 547–565.

Molyboga, M., and C. L'Ahelec. 2016. "A Simulation-Based Methodology for Evaluating Hedge Fund Investments." *Journal of Asset Management* 17: 434–452.

Moreira, A., and T. Muir. 2017. "Volatility-Managed Portfolios." *The Journal of Finance* 72: 1611–1643.

Pafka, S., M. Potters, and I. Kondor. "Exponential Weighting and Random-Matrix-Theory-Based Filtering of Financial Covariance Matrices for Portfolio Optimization." 2004, http://arxiv.org/abs/cond-mat/0402573.

Pastor, L. 2000. "Portfolio Selection and Asset Pricing Models." *The Journal of Finance* 55: 179–223.

Yang, Z., K. Han, M. Molyboga, and G. Molyboga. 2016. "A New Diagnostic Approach to Evaluating the Stability of Optimal Portfolios." *The Journal of Investing* 25 (1): 37–45.

To order reprints of this article, please contact David Rowe at d.rowe@pageantmedia.com or 646-891-2157.

ADDITIONAL READING

The Evolution of Equity Mandates in Institutional Portfolios

Mark Anson

The Journal of Portfolio Management

https://jpm.pm-research.com/content/37/4/127

ABSTRACT: *The asset allocation decision is one of the most important determinants of overall fund performance for institutions around the world. Pension funds, endowments, and foundations spend considerable time deciding how to split up the trillions of dollars they collectively control. Inevitably, the largest allocation is to the equity slice of the pie. The reasons are several: higher returns, largest asset class in which to invest, good liquidity, easiest to understand, and the cheapest to transact in. Anson traces the evolution of equity mandates as institutional investors have become more sophisticated in their approach to the equity markets including international investing, global mandates, Fama–French risk factors, and risk parity.*

Just a One-Trick Pony? *An Analysis of CTA Risk and Return*

Jason Foran, Mark C. Hutchinson, David F. McCarthy, and John O'Brien

The Journal of Alternative Investments

https://jai.pm-research.com/content/20/2/8

ABSTRACT: *Recently, a range of alternative risk premium products has been developed, promising investors hedge fund/Commodity Trading Advisor (CTA)-like returns with higher liquidity and transparency and relatively low fees. The attractiveness of these products rests on the assumption that they can deliver similar returns. Using a novel reporting bias–free sample of 3,419 CTA funds as a testing ground, the authors' results suggest that this assumption is questionable. They find that CTAs are not a homogenous group. They identify eight different CTA substrategies, each with very different sources of return and low correlation between substrategies. To illustrate the difficulty of modelling the strategies, they specify recently identified alternative risk premiums from the academic literature as factors to examine the sources of return of CTAs. They find that these premiums fail to explain between 56% and 86% of returns. Their results suggest that given the heterogeneity of CTAs, although these new products may deliver on liquidity, transparency, and fees, investors expecting hedge fund/CTA-like returns may be disappointed.*

Honey, I Shrunk the Sample Covariance Matrix

Olivier Ledoit and Michael Wolf

The Journal of Portfolio Management

https://jpm.pm-research.com/content/30/4/110

ABSTRACT: *The central message of this article is that no one should use the sample covariance matrix for portfolio optimization. It is subject to estimation error of the kind most likely to perturb a mean-variance optimizer. Instead, a matrix can be obtained from the sample covariance matrix through a transformation called shrinkage. This tends to pull the most extreme coefficients toward more central values, systematically reducing estimation error when it matters most. Statistically, the challenge is to know the optimal shrinkage intensity. Shrinkage reduces portfolio tracking error relative to a benchmark index, and substantially raises the manager's realized information ratio.*