A Constrained Hierarchical Risk Parity Algorithm with Cluster-based Capital Allocation

Johann Pfitzinger[a], Nico Katzke[a,b]

[a] *Stellenbosch University, Stellenbosch, South Africa*

[b] *Prescient Securities, Cape Town, South Africa*

Abstract

Hierarchical Risk Parity (HRP) is a risk-based portfolio optimisation algorithm, which has been shown to generate diversified portfolios with robust out-of-sample properties without the need for a positive-definite return covariance matrix (Lopez de Prado 2016). The algorithm applies machine learning techniques to identify the underlying hierarchical correlation structure of the portfolio, allowing clusters of similar assets to compete for capital. The resulting allocation is both well-diversified over risk sources and intuitively appealing. This paper proposes a method of fully exploiting the information created by the clustering process, achieving enhanced out-of-sample risk and return characteristics. In addition, a practical approach to calculating HRP weights under box and group constraints is introduced. A comprehensive set of portfolio simulations over 6 equity universes demonstrates the appeal of the algorithm for portfolios consisting of $20 - 200$ assets. HRP delivers highly diversified allocations with low volatility, low portfolio turnover and competitive performance metrics.

Keywords: Risk Parity, Diversification, Portfolio Optimisation, Clustering

JEL classification G110

1. Introduction

Optimal capital allocation across assets is arguably one of the most widely discussed topics in quantitative finance. Since the advent of modern portfolio theory introduced by Markowitz (1952), portfolio optimisation has generally been framed as a quadratic programming problem considering the risk (variance) and return characteristics of the underlying assets. Mean-variance optimisation (MVO) forms the basis of most applied portfolio optimisation routines, due to its intuitive appeal and theoretical property as the pareto-optimal in-sample allocation (Kolm, Tütüncü, and Fabozzi 2014). Despite its popularity, MVO has been subjected to important critiques (see Steinbach (2001) and Rubinstein (2002) for more comprehensive reviews). Sensitivity to measurement error, leading to poor out-of-sample performance, as well as instability of the covariance matrix are important practical concerns. The sample dependence of MVO has precipitated the development of purely risk-based allo-

Email addresses: jpfitzinger@xmlh.de (Johann Pfitzinger), nfkatzke@gmail.com (Nico Katzke)

cation methods, which ignore error-prone return forecasts (Roncalli 2014). Risk parity and minimum-variance methods exemplify this strategy, but often remain vulnerable to the instability-critique in large N portfolios.

Hierarchical risk parity (HRP) is a risk parity allocation algorithm introduced by Lopez de Prado (2016), which aims to address the short-comings of MVO portfolios. At its core HRP proposes a hierarchical implementation of an inverse-variance allocation with weights calculated between clusters of correlated asset returns. The method benefits from the out-of-sample robustness of risk parity portfolios, and can be applied to singular covariance matrices, hence solving the stability problem. Lopez de Prado (2016) shows that HRP achieves lower out-of-sample volatility and higher risk-adjusted return than inverse-variance allocations. Apart from desirable quantitative properties, the method has the intuitive appeal of allowing groups of related assets (as opposed to individual assets) to compete for capital in the portfolio, improving diversification over risk sources. Apart from Lopez de Prado (2016), variants of HRP have been presented by Alipour et al. (2016) and Raffinot (2016).

This paper introduces three modifications to the HRP algorithm of Lopez de Prado (2016): (i) A method is proposed to impose weight constraints to individual assets or groups of assets, while maintaining cluster-optimal allocations; (ii) Hierarchical clustering algorithms are compared to an efficient genetic optimisation based alternative, demonstrating significantly improved performance under top-down and optimisation-based approaches to clustering; (iii) Finally, the naive allocation hierarchy proposed in Lopez de Prado (2016) is replaced by an algorithm capable of considering asset similarity at the cluster edges, permitting a more complete exploitation of the information generated by the cluster algorithm.

The HRP portfolio is benchmarked against conventional risk-based allocation techniques in a systematic manner using historical daily returns data from 6 equity universes. The portfolio backtest explores the performance of HRP in portfolios of varying degree of N-diversification, with particular emphasis placed on South African equity markets. HRP is found to achieve lower out-of-sample volatility and higher risk-adjusted return than inverse-variance as well as naive portfolios, and is a close competitor to the minimum-variance portfolio, but without the dependence on a positive-definite covariance matrix. Furthermore, the HRP portfolio achieves superior diversification and low turnover. Proposed modifications to the clustering and allocation approaches make better use of the hierarchical structure of the covariance matrix and consequently improve out-of-sample volatility and risk-adjusted return of the HRP allocation.

2. Hierarchical Risk Parity Algorithm

Conceptually, the HRP algorithm proposed by Lopez de Prado (2016) calculates inverse-variance weights for groups of similar assets, iteratively applied to ever smaller sub-groups until each asset forms its own cluster. The approach is operationalised in three steps: (i) clustering, (ii) quasi-diagonalisation and (iii) recursive bisection. In the following sub-sections, each step is briefly outlined alongside a few proposed modifications, with a more detailed mathematical treatment provided in Appendix A.

2.1. Clustering

Clustering is a statistical partitioning technique which groups a set of N data series based on their characteristics (Maimon and Rokach 2010). In the case of correlation clustering, the correlation coefficient is deemed to be the relevant measure of similarity between the data series. The intuitive appeal of a portfolio optimisation based on the clustered covariance matrix is the ability to allocate capital to groups of similar assets. Thus rather than allowing individual assets to act as direct substitutes, asset groups jointly compete for capital leading to a more compelling and robust allocation result. Hierarchical clustering in particular has been widely applied in the financial literature (see Tola et al. (2008) for an overview). Examples of applications include Tola et al. (2008) and Tumminello, Lillo, and Mantegna (2010) who show that the stability characteristics of the covariance matrix can be improved using linkage-based clustering filters, boosting the feasibility of MVO allocations for large N portfolios. Mantegna (1999), Di Matteo, Aste, and Mantegna (2004) use clustering techniques to infer the hierarchical structure of financial data series. Other applications include León et al. (2017), who present a cluster-based MVO allocation.

In general, hierarchical clustering approaches can be subdivided into two broad parent categories: agglomerative nesting (AGNES) and divisive analysis (DIANA) (Maimon and Rokach 2010). Agglomerative nesting is a bottom-up procedure where objects initially represent individual clusters, and are successively merged into larger clusters until the full hierarchical structure is obtained. Single-linkage clustering is a particular type of agglomerative nesting which calculates the distance between two clusters as the shortest distance between any member of one cluster to any member of another cluster. By contrast, divisive analysis is a top-down approach which begins with a single cluster containing all objects and successively sub-divides assets into smaller clusters until each cluster contains only a single object.

Lopez de Prado (2016) performs a single-linkage (AGNES) hierarchical clustering of the correlation matrix, which groups similar assets based on their correlation. Appendix A.1 outlines the necessary mathematical transformations. Figure 2.1 (left panel) presents the cluster dendrogram associated with the single-linkage clustering of a randomly drawn sample of JSE shares. Section 5.1 compares HRP performance employing divisive analysis, as well as two bottom-up ANGES clustering techniques (single-linkage and Ward). The top-down nature of divisive analysis is consistent with the allocation strategy, and achieves lower out-of-sample variance under certain conditions described in Section 5.1. The right panel of Figure 2.1 displays the cluster dendrogram for the same sample of stocks obtained using divisive analysis.

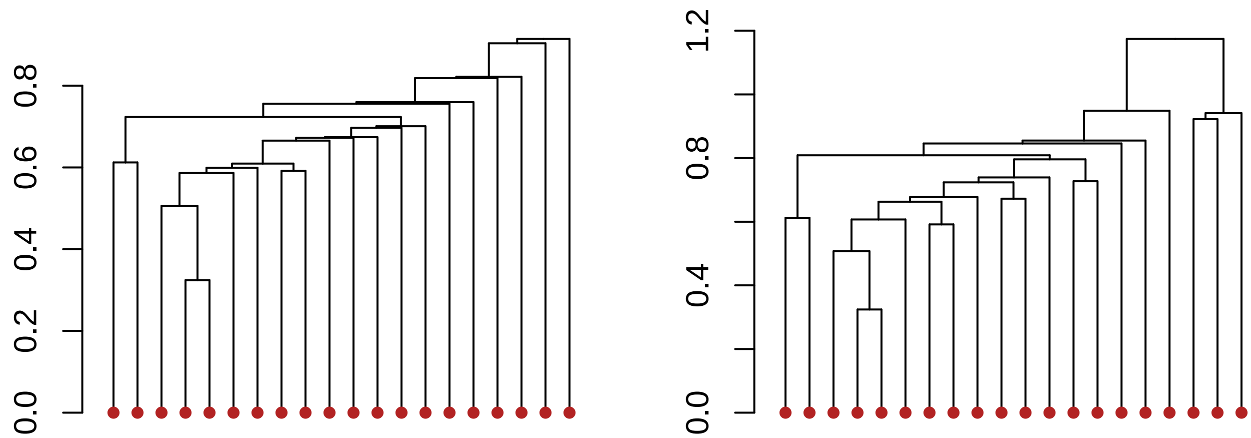

Figure 2.1: Example of Cluster Dendrograms, AGNES (left) and DIANA (right)

2.2. Quasi-diagonalisation

The most efficient (i.e. variance-minimising) combination of uncorrelated time series, is an inverse-variance weighting (Lopez de Prado 2016). Thus, an inverse-variance asset allocation is most appropriate for assets with an approximately diagonal correlation matrix. The second step of the HRP algorithm uses the information obtained from the clustering algorithm to transform (rearrange) the covariance matrix into a more diagonal representation. High correlations are placed adjacently and close to the matrix diagonal, achieving the desired quasi-diagonal structure. Figure 2.2 demonstrates the effect by displaying the naive correlation matrix alongside the reordered matrix using single-linkage clustering (middle) and divisive analysis (right) for the same random sample of assets used before:

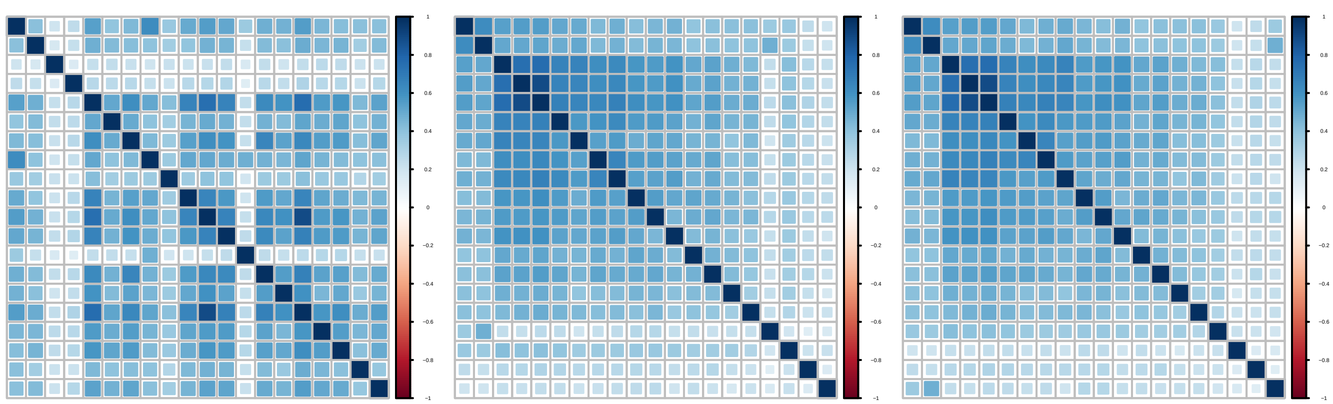

Figure 2.2: Diagonalised Correlation Matrix, untransformed (left), AGNES (middle) and DIANA (right)

Alipour et al. (2016) introduce an innovative modification to the HRP algorithm, which recasts the clustering and quasi-diagonalisation steps as an optimisation problem. The extent of diagonalisation can be quantified using a simple loss function $\mathcal{L}$ assigning larger weights to correlation coefficients further away from the matrix diagonal (see Appendix A.2). The authors solve the diagonalisation problem by minimising a penalised version of $\mathcal{L}$ with a quantum annealing optimiser, replacing the clustering and quasi-diagonalisation steps with a computationally costly N^2-parameter binary optimisation task, which uncovers the most diagonal structure attainable by 'brute force' (Alipour et al. 2016). An alternative proposed in this paper minimizes $\mathcal{L}$ using the genetic permutation algorithm outlined in Even (1973)[1]. Appendix A.2 describes the objective function in more detail, with the resulting HRP allocation denoted 'genetic HRP' (GHRP) in this context. While the genetic optimiser increases the computational cost vis-à-vis clustering techniques, it is possible to achieve convergence in as little as 150-200 iterations.

2.3. Recursive Bisection

Finally, the HRP weights are calculated as the inverse-variance allocation between two clusters applied recursively by bisecting the covariance matrix until only one asset is contained in each cluster. The full algorithm is provided in Appendix A.3. Figure 2.3 (top-left) visualises the bisection for the first iteration using the cluster dendrogram of the previously discussed stock sample. The assets are split into equally sized sub-groups with the weights calculated based on the aggregate cluster variance. An alternative risk-budgetting allocation approach is developed in Raffinot (2016), but not discussed further here.

As shown in Figure 2.3, the naive bisection rule can violate the intuitive character of the result, by placing similar assets into separate clusters for allocation purposes. While centered bisection yields a symmetric allocation tree, which results in well-diversified portfolio weights, the method does not respect empirical cluster boundaries and discards information about the hierarchical structure inferred in the cluster algorithm. Figure 2.3 (top-left) demonstrates the concern, with the bisection separating two closely related assets.

A modification is proposed in this paper, which permits the allocation to take the similarity of assets into account. The allocation trades off the symmetry objective against a minimum dissimilarity between border objects of the sub-clusters. The trade-off is characterised by a threshold τ, which measures the relative location on the y-axis of the dendrogram above which splits may occur. In the extreme case ($\tau = 1$), the dissimilarity is maximised and allocation is performed exactly along the cluster dendrogram. If $\tau = 0$ the symmetry objective dominates resulting in a naive bisection as in Lopez de Prado (2016). Figure 2.3 visualises different values of τ, with the extreme cases in the top-left and bottom-right. If $0 < \tau < 1$, the constraint is weakly enforced and the algorithm selects the most symmetrical allocation tree possible under the condition that assets with a certain degree of relative similarity are not split into different clusters.

[1] The implementation makes use of the `R`-package `gaoptim`.

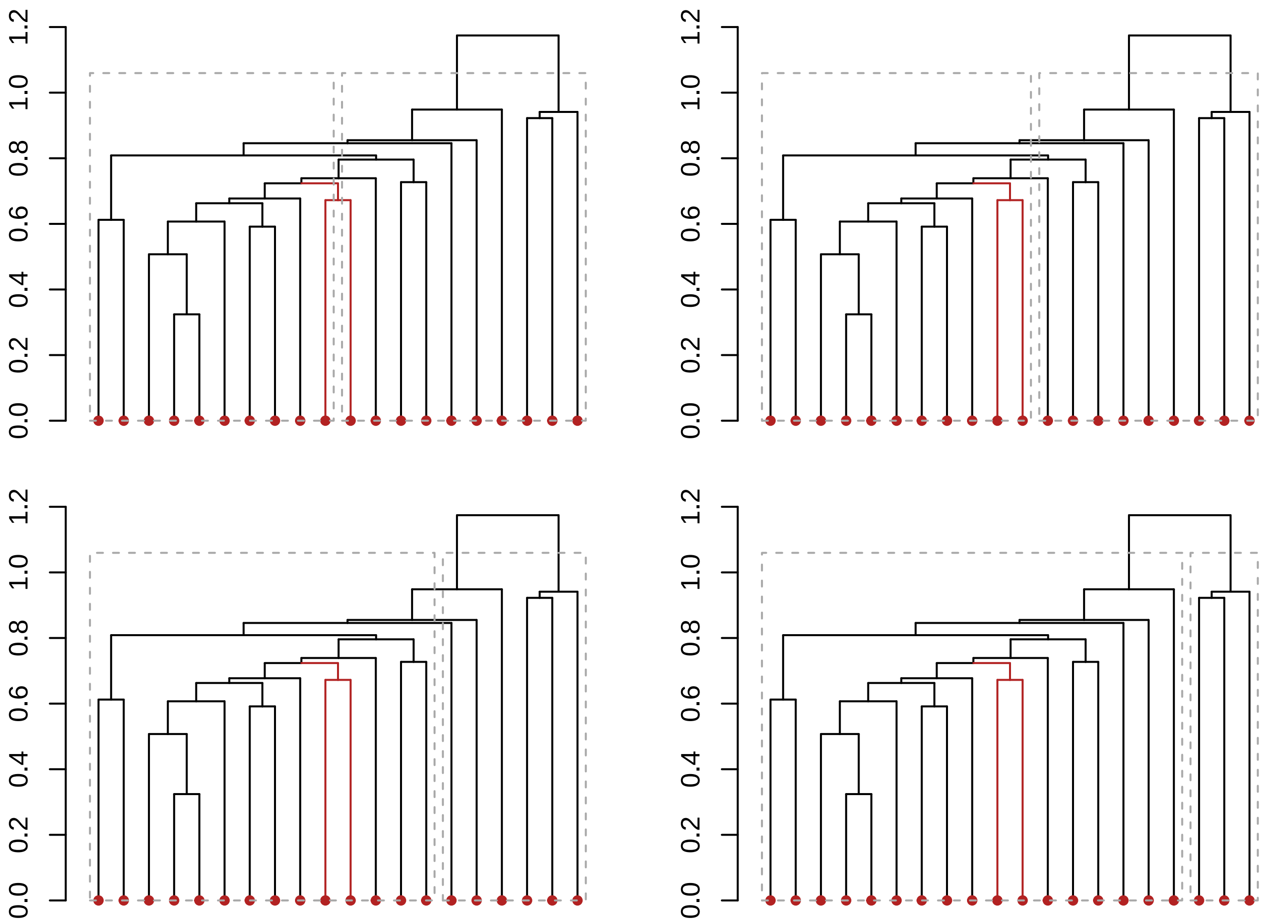

Figure 2.3: Hierarchical Tree Splits for $\tau = 0 - 1$

As a practical matter, the recursive bisection step permits portfolio weight constraints to be introduced, as recommended by Lopez de Prado (2016). The modification to the HRP algorithm proposed in this paper enforces box and group constraints in a bottom-up manner. Violations of the constraints are resolved at the lowest possible location on the hierarchical allocation tree, leaving preceding cluster allocations unaffected and retaining relative within-cluster proportions wherever possible. The approach is computationally highly efficient, adding negligible calculation time, and effectively identifying the weights-vector most resembling the unconstrained solution. The mathematical derivation is provided in Appendix A.3. Figure 2.4 illustrates the constrained allocation based on the random sample of JSE shares (weights are calculated using the portfolio backtest described in Section 4). Groups are defined as BICS[2] level 1 industries.

[2]Bloomberg Industry Classification System.

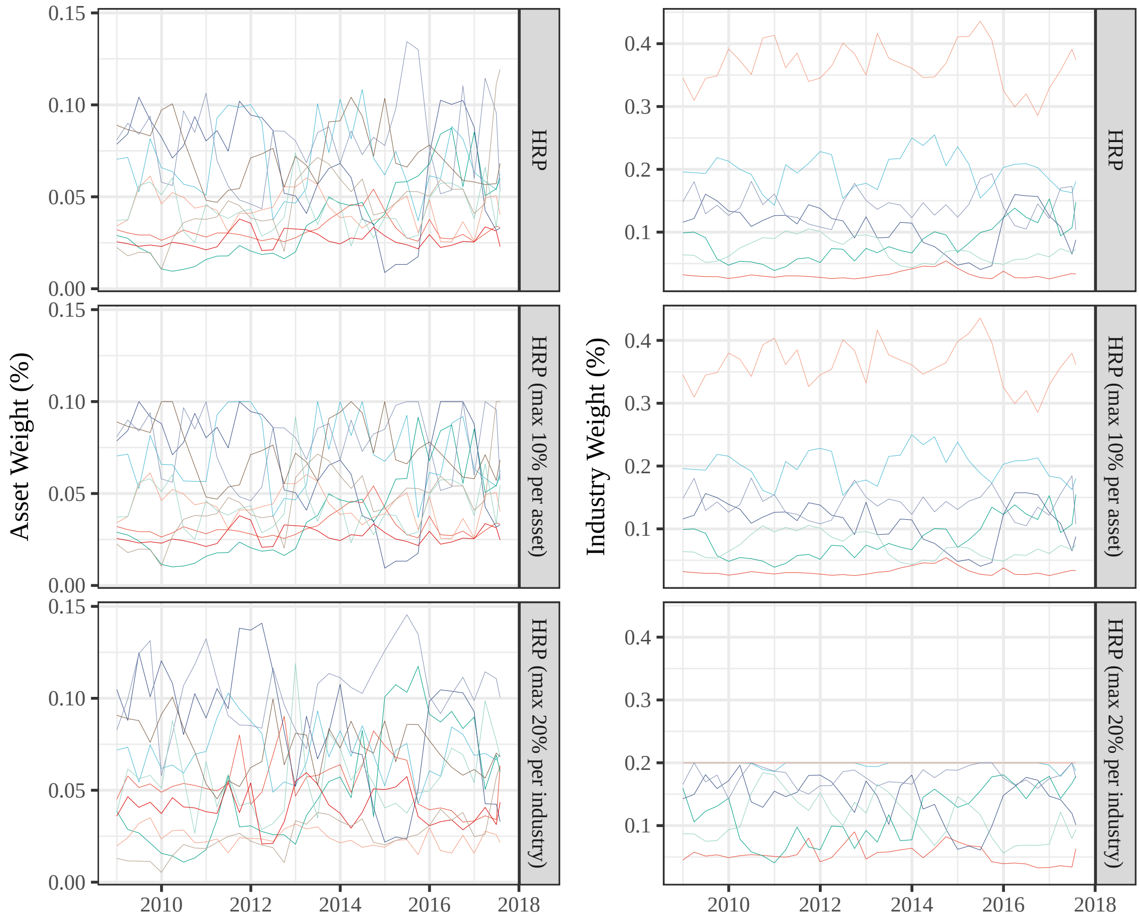

Figure 2.4: Box and Group Constraint Illustration

While the emphasis of the portfolio simulations in Section 5 is placed on the performance of unconstrained HRP, adjunct simulations reported in Appendix B demonstrate that the findings are not materially affected by the imposition of constraints, with the same relative performance outcomes as in Section 5.

3. Asset Universes[3]

Portfolio simulations are performed using daily total return index data from 6 equity universes. Data is obtained from Bloomberg (2017), for index samples with changing compostion based on index membership. Table 3.1 provides an overview of the data universes:

[3]Section to be provided by supervisor.

Universe	Country	Code	No. of stocks	Period
JSE All Share Index	South Africa	SA	255	2008-2017
S&P 500	United States	US	707	2008-2017
S&P Euro	Euro Area	EU	225	2008-2017
FTSE 100[4]	United Kingdom	UK	167	2008-2017
Canada	Canada	CA	408	2008-2017
Australia	Australia	AU	453	2008-2017

Table 3.1: Equity Return Universes

4. Portfolio Backtest

The out-of-sample performance of three HRP variants is benchmarked against conventional naive and risk-based portfolio optimisation techniques. The global minimum-variance portfolio (MV) is calculated using quadratic programming and is defined as the allocation w which satisfies:

$$\min_{w} w'\Sigma w \text{ s.t. } \sum_{i} w_i = 1 \text{ and } w \geq 0 \tag{4.1}$$

where Σ is the assets' covariance matrix and the constraints ensure full investment and no short sales. The inverse-variance portfolio (IV) is the closest comparable risk parity portfolio, and is computed as:

$$w_i = \frac{\sigma_i^{-2}}{\sum_{i=1}^{N} \sigma_i^{-2}} \tag{4.2}$$

where σ_i is the i-th asset return standard deviation.

In addition to the risk-based methods, two naive portfolios are benchmarked: The equal weighted portfolio is calculated with $w = 1/N$ and the random portfolio selects weights in a completely non-deterministic manner[5]. The HRP variants include single-linkage based HRP with $\tau = 0$ (identical to Lopez de Prado (2016)), DIANA clustering with $\tau = 1$ (DHRP), and genetic optimisation based clustering with $\tau = 1$ (GHRP).

The out-of-sample performance is compared using a portfolio backtest over 10 years of daily historical

[4] Note: The assumption is made that this is the underlying index. Can be updated otherwise.
[5] The random portfolio is constructed with the `R`-package `rportfolios`.

returns data. Weights are rebalanced at the beginning of each quarter, with covariances calculated over a trailing 12 month period. The backtest is repeated in a Monte Carlo experiment where 1000 N-sized portfolios are randomly sampled from each equity universe. The approach provides a robust method of evaluating the performance of the allocation algorithms in concentrated portfolios with $N = 20 - 50$, making the study applicable to the South African context where relatively small portfolio sizes are not uncommon. A global perspective is added by performing large N performance comparisons for stocks randomly sampled from the aggregate equity universe across all regions ($N = 50 - 200$).

Since the Monte Carlo experiment is invariably subject to survivorship bias with the sample composition remaining unadjusted over time[6], the results are additionally compared in full-index samples for 4 universes (S&P 500, JSE All Share, S&P Euro & FTSE 100) with changing composition. Bootstrap confidence intervals are calculated using iterated backtests over resampled index assets[7].

The study compares algorithms using various out-of-sample performance and concentration metrics. Risk measures include the volatility (standard deviation) of returns as well as maximum drawdown, with additional performance perspectives obtained from the portfolios' risk-adjusted return[8] and turnover (trading volume). Turnover is calculated as the absolute difference between the value of end-of-day positions prior to rebalancing and positions immediately after rebalancing, as a percentage of the total position. Portfolio concentration is measured using the Herfindahl index, and an industry concentration measure (Herfindahl index of industry weights). The effect of sample-based variation in industry concentration is removed by utilizing a stratified sampling technique in the Monte Carlo experiment, which draws assets from a minimum of 5 different randomly selected BICS level 1 industries, and by normalising concentration measures by the naive (equally-weighted) portfolio.

5. Discussion of results

5.1. Cluster Techniques

The objective of hierarchical clustering in the HRP algorithm, is to obtain a quasi-diagonal correlation matrix which is theoretically amenable to inverse-variance based allocation techniques. The merits of the clustering algorithm should thus be evaluated based on its success in generating diagonal correlation matrices. Recognising this fundamental purpose, Section 2.3 presented a version of the HRP algorithm, which converts the clustering and diagonalisation steps into an explicit optimisation

[6] The explicit objective is to test a large number of equally sized portfolios, requiring that a constant asset composition is chosen at each draw.

[7] Resampling with replacement is performed at each iteration. See Alipour et al. (2016) for a more detailed description of the bootstrap design.

[8] Calculated as the Sharpe Ratio, without risk-free benchmark (i.e. annualized return divided by annualized standard deviation).

problem, where the extent of diagonalisation (measured by $\mathcal{L}$) is maximised by permutating the order of the rows and columns in the correlation matrix. Since the objective function $\mathcal{L}$ is a numerical measure of the extent of diagonalisation of the correlation matrix, it can be used to benchmark different clustering algorithms against the optimally diagonalised correlation matrix. Figure 5.1 plots the value of $\mathcal{L}$ achieved by 3 different clustering methods in addition to the GHRP genetic permutation outcome[9]. The results are normalised by the mean of GHRP $\mathcal{L}$-scores. The GHRP optimisation achieves the highest extent of diagonalisation, while single-linkage clustering performs comparatively poorly. Ward clustering achieves promising diagonalisation results.

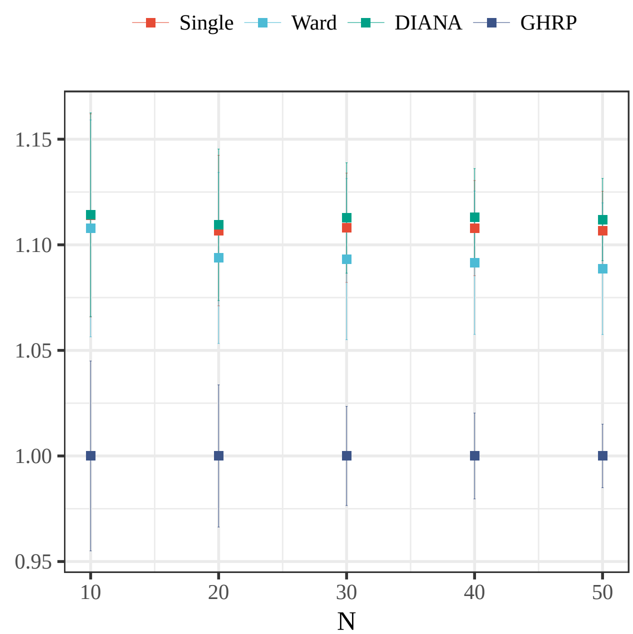

Figure 5.1: Normalised $\mathcal{L}$-scores of Cluster Algorithms

The out-of-sample volatility under different clustering techniques using JSE equity samples with $N = 20-50$ is compared in Figure 5.2 for two allocation scenarios: (i) naive allocation with $\tau = 0$ (left) and (ii) dissimilarity-based allocation with $\tau = 1$ (right). Under naive allocation, the choice of clustering technique is inconsequential. This result is not surprising since the information produced by the clustering algorithm is discarded in the allocation step. Conversely, when the cluster tree is utilised for weight allocation, GHRP, DIANA and Ward clustering all significantly outperform single-linkage clustering. DIANA clustering achieves almost identical results to GHRP for samples of $N \geq 40$. The top-down nature of the method makes it a good candidate for top-down dissimilarity-based allocation. For the remainder of the paper, DIANA clustering with $\tau = 1$ is denoted DHRP and benchmarked against conventional HRP and GHRP.

[9]Results from complete- and average-linkage clustering were too similar to single-linkage clustering to warrant their inclusion. Instead the focus is on the 3 approaches presented, which capture the most important aspects.

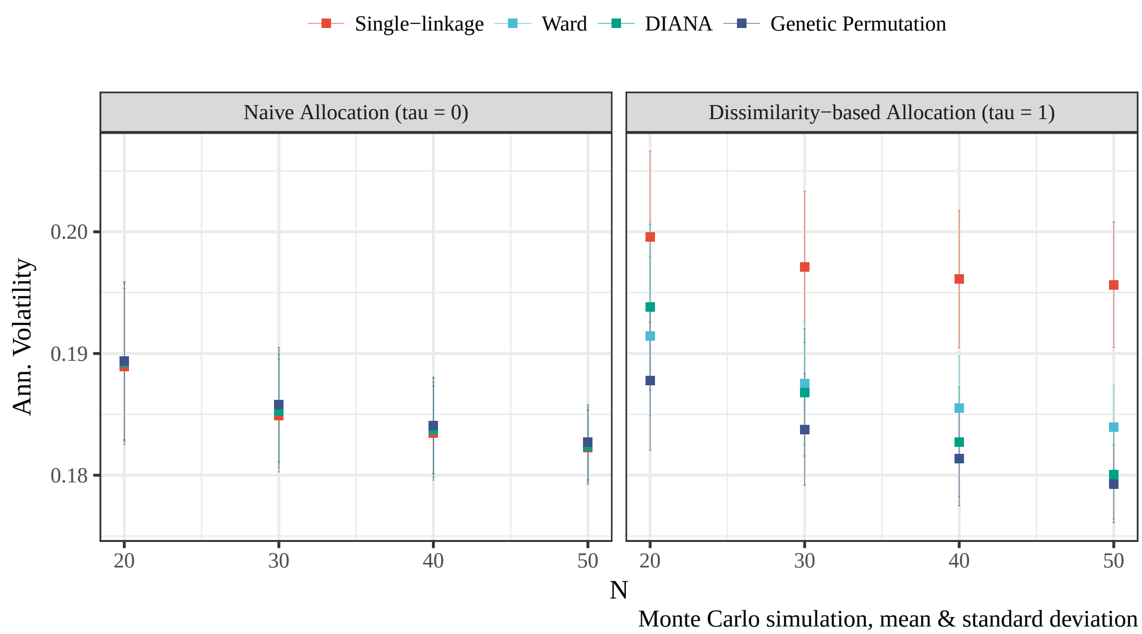

Figure 5.2: Out-of-sample Volatility of Cluster Algorithms

DIANA clustering furthermore achieves a fairly high risk-adjusted return (Figure 5.3) and low turnover (Figure 5.4). Despite its favourable $\mathcal{L}$-score, the Ward algorithm does not add significant improvement over DIANA on any metric, while single-linkage clustering performs well in terms of risk-adjusted return. GHRP achieves superior results on both volatility and return metrics, indicating that the hierarchical clusters extracted by genetic optimisation most completely exploit the information in the covariance matrix. The figures illustrate the advantage of dissimilarity-based allocation, with $\tau = 1$ conclusively achieving better results for portfolios with $N \geq 40$ (albeit with higher turnover). Ward clustering is omitted from further performance benchmarks, as it does not significantly improve outcomes.

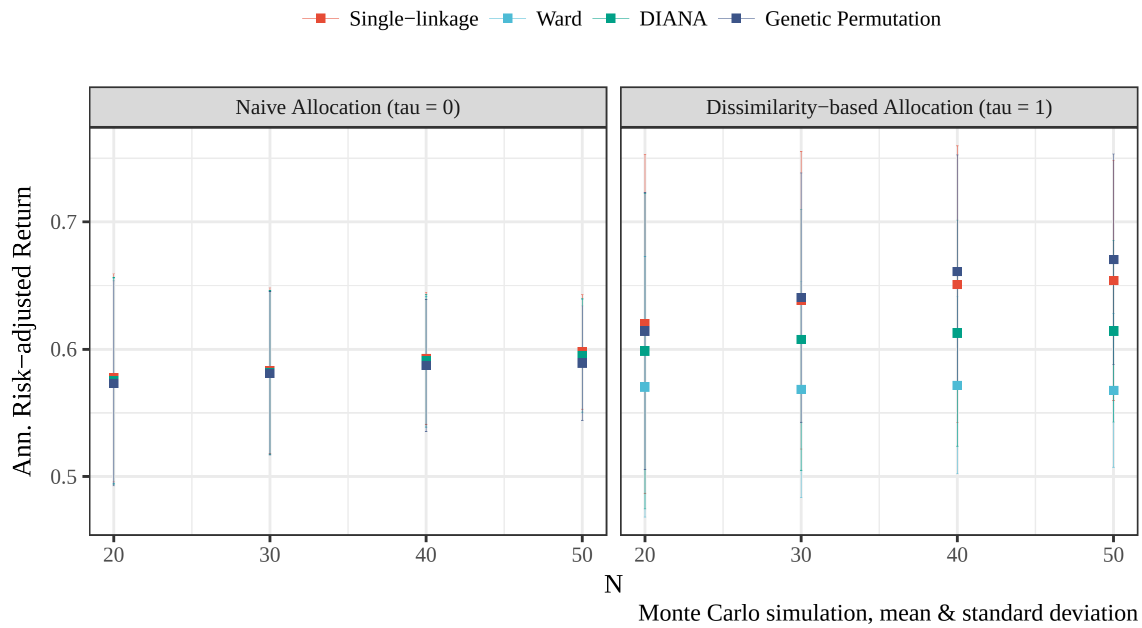

Figure 5.3: Out-of-sample Risk-adjusted Return of Cluster Algorithms

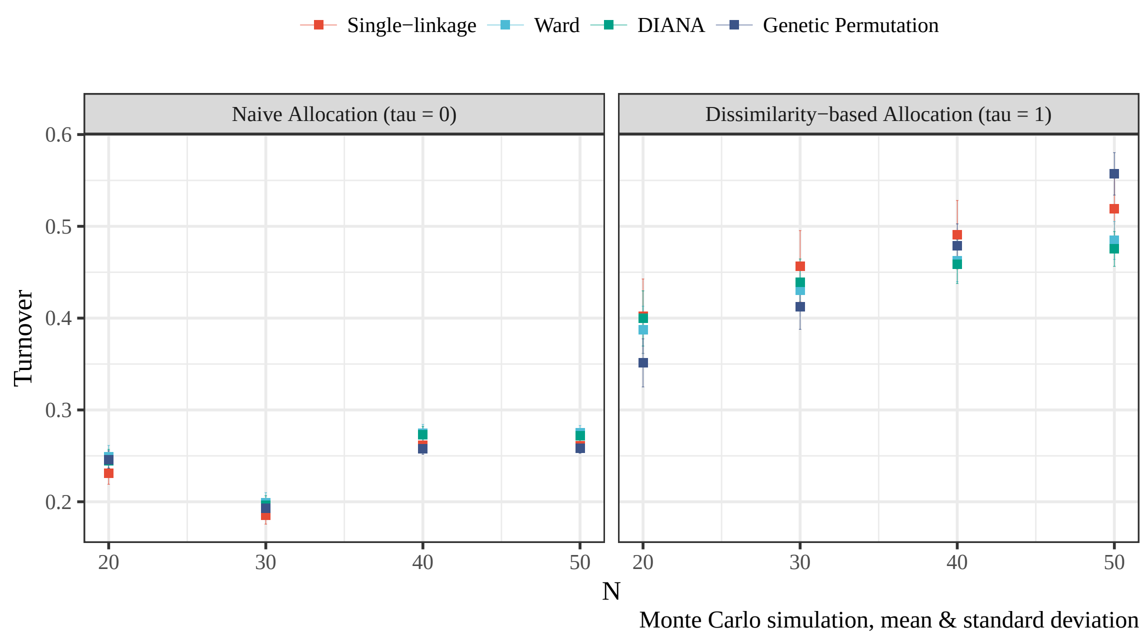

Figure 5.4: Out-of-sample Turnover of Cluster Algorithms

5.2. Portfolio Performance: Individual Universes

Figure 5.5 plots the annualised portfolio volatility and maximum drawdown for JSE stock portfolios with $N = 50$. The Monte Carlo backtest includes three variants of the HRP algorithm: (i) HRP as introduced in Lopez de Prado (2016), (ii) DHRP with top-down DIANA clustering and $\tau = 1$, (iii) GHRP with diagonalisation solved using genetic optimisation and $\tau = 1$. All HRP variants outperform the IV portfolio as well as equal and random weighted portfolios. DHRP and GHRP achieve superior risk outcomes relative to HRP, demonstrating the benefit of the modified algorithm over standard HRP. The close similarity of DHRP and GHRP suggests that DIANA clustering may be considered an adequate computationally efficient substitute for genetic optimisation. The MV portfolio achieves consistently low out-of-sample volatility — a finding that stands in contrast to Lopez de Prado (2016), who observe HRP to outperform MV using simulated returns. The sample dependence and concomittant poor out-of-sample results of MV appear to be limited when applied to real historical JSE data. Additional simulations (omitted here for brevity) demonstrate that the mean-variance portfolio does indeed perform poorly out-of-sample, indicating that (as argued by proponents of risk parity techniques) measurement error in returns as opposed to volatility is the primary source of underperformance.

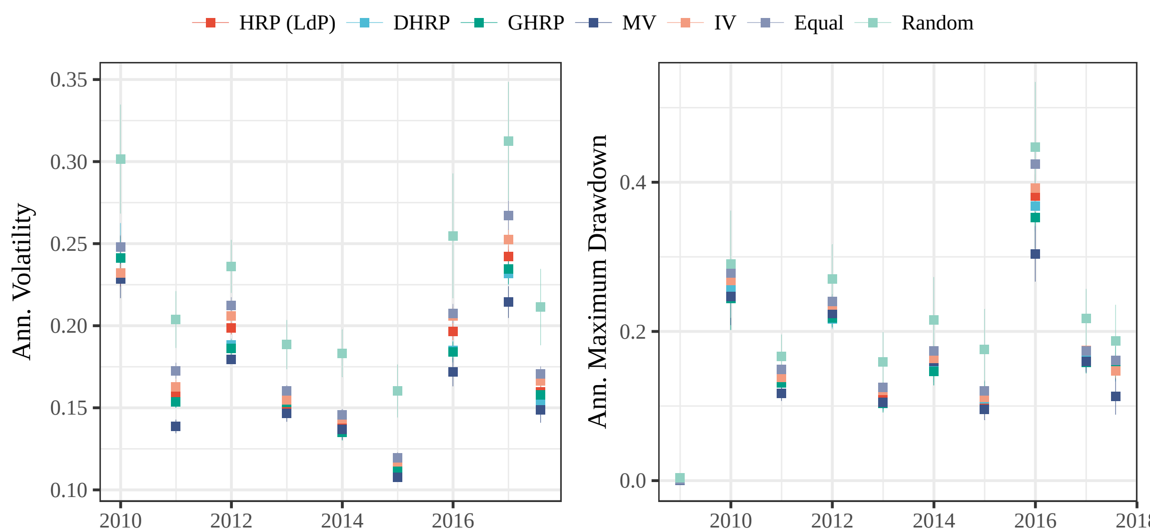

Figure 5.5: JSE Annualised Volatility & Maximum Drawdown ($N = 50$)

Figure 5.6 compares the risk-adjusted return and portfolio turnover to annualised volatility for the same portfolio simulation as above. In addition to reduced volatility, the DHRP and GHRP portfolios achieve higher risk-adjusted return than standard HRP, with GHRP performing roughly on par with MV. The additional return comes at the expense of portfolio turnover, where HRP significantly outperforms the MV portfolio. The naive portfolio achieves the lowest turnover (as expected), and the IV portfolio outperforms HRP in terms of turnover (but not risk-adjusted return and volatility).

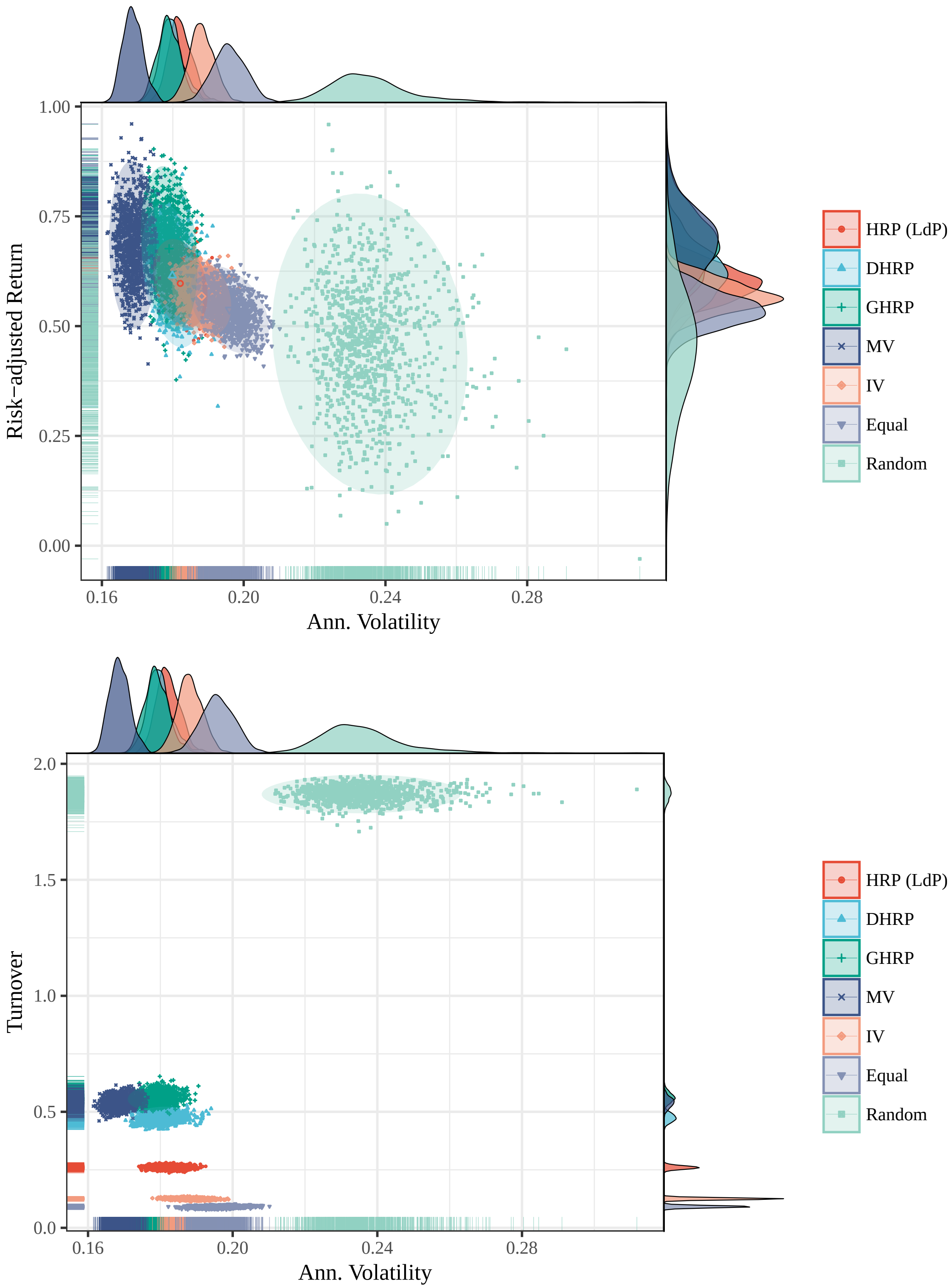

Figure 5.6: Portfolio Simulation Results ($N = 50$, JSE Equity Universe)

The generality of the findings is evaluated in Figure 5.7, plotting the overall portfolio volatility for

$N = 20 - 50$ for all 6 equity universes. The trend uncovered above is mostly confirmed: MV achieves the lowest volatility, while HRP and particularly GHRP and DHRP outperform the remaining benchmarked methods in most cases. The relationship is robust to different portfolio sizes, and DHRP and GHRP are close substitutes, once again confirming the usefulness of DIANA clustering as an alternative to direct optimisation.

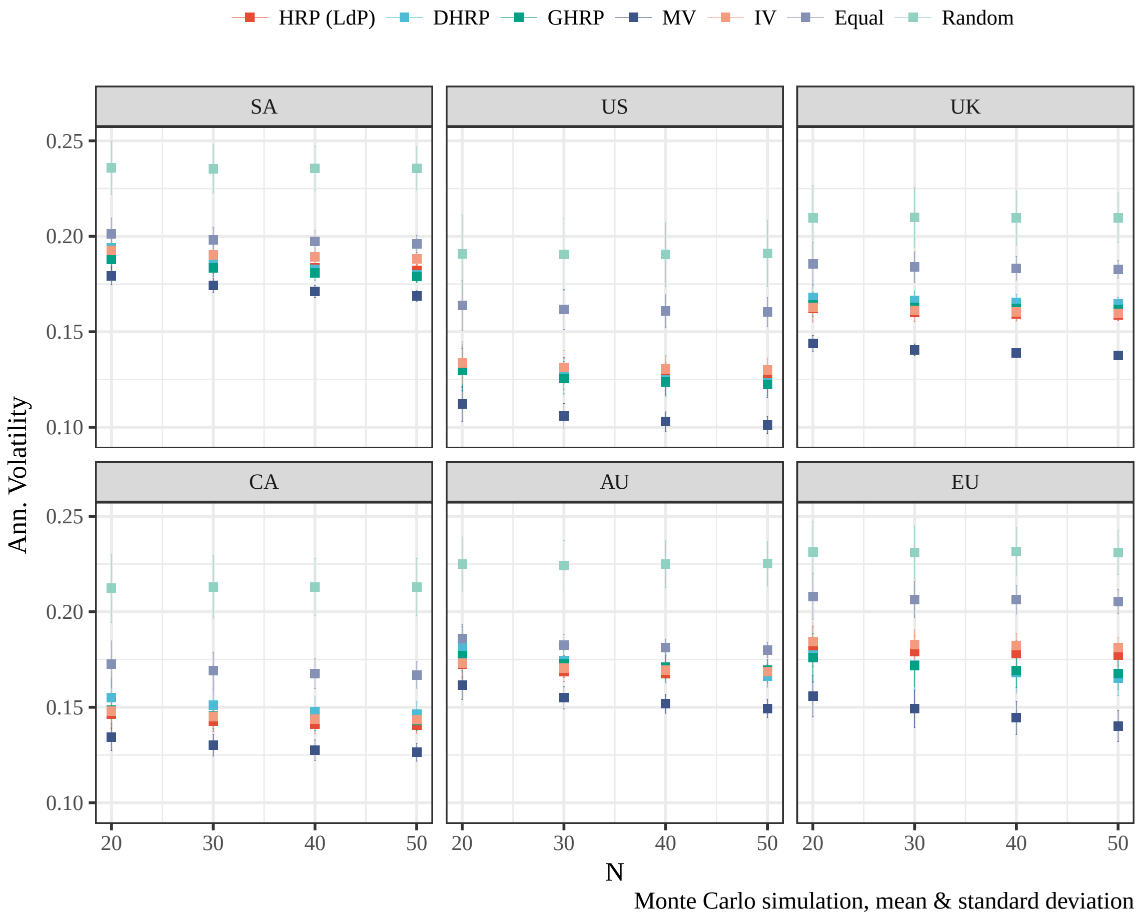

Figure 5.7: All Universes Annualised Volatility

The results for the out-of-sample risk-adjusted return in Figure 5.8 are somewhat more variable. MV achieves high returns on the JSE, with GHRP approximately equivalent (as discussed above). In the remaining universes, however, DHRP and GHRP outperform MV. The CA, AU and EU universes demonstrate the value added by the modified HRP algorithms, with standard HRP underperforming MV. IV and naive portfolios present inferior results in all cases apart from the UK, where the equally-weighted portfolio achieves robustly high results.

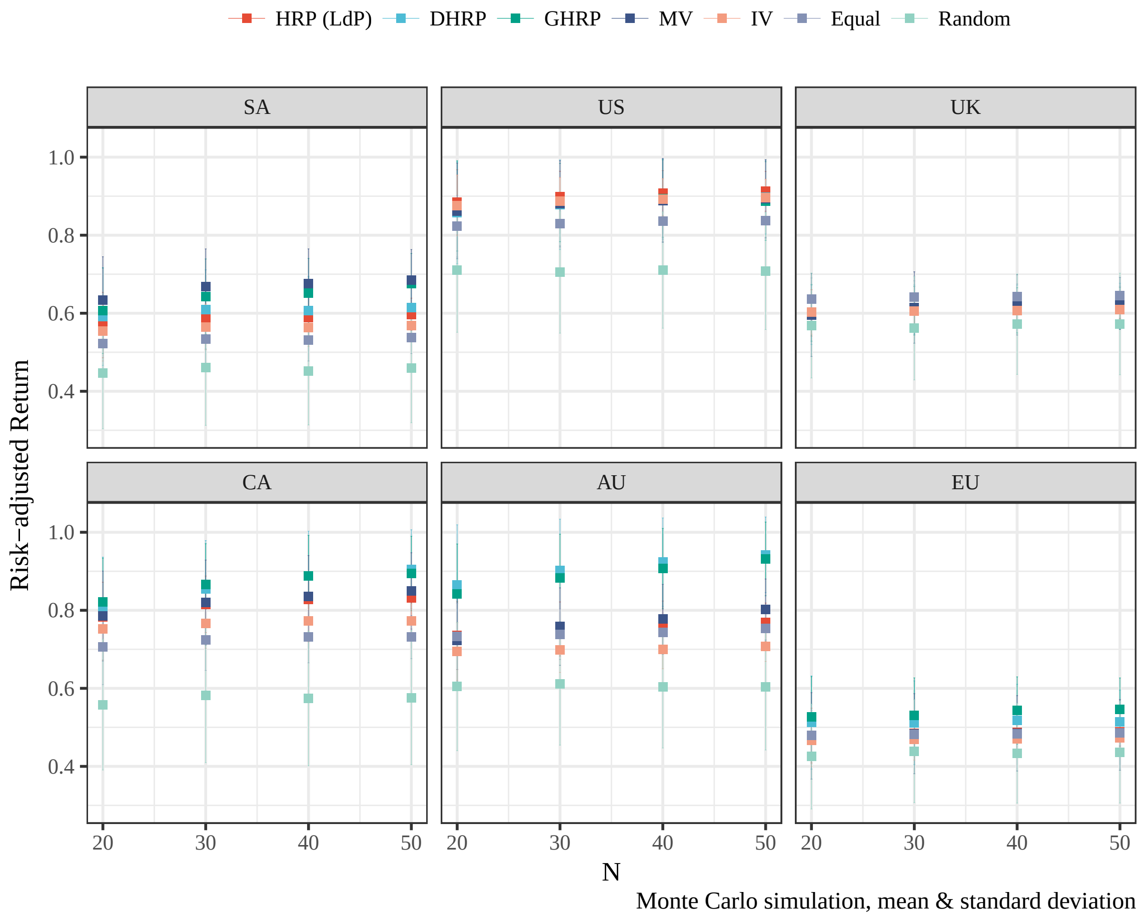

Figure 5.8: All Universes Annualised Risk-adjusted Return

Figure 5.9 plots portfolio turnover for the equity universes. The results confirm the finding that HRP significantly reduces turnover compared to MV. GHRP is outperformed by DHRP in all cases, which coupled with the approximately similar results on other performance metrics, confirms the usefulness of the DHRP algorithm. The significantly lower turnover of HRP in relation to DHRP and GHRP suggests that the symmetry-dissimilarity trade-off characterised by τ manifests as a trade-off between out-of-sample performance (risk, return) and asset cluster stability.

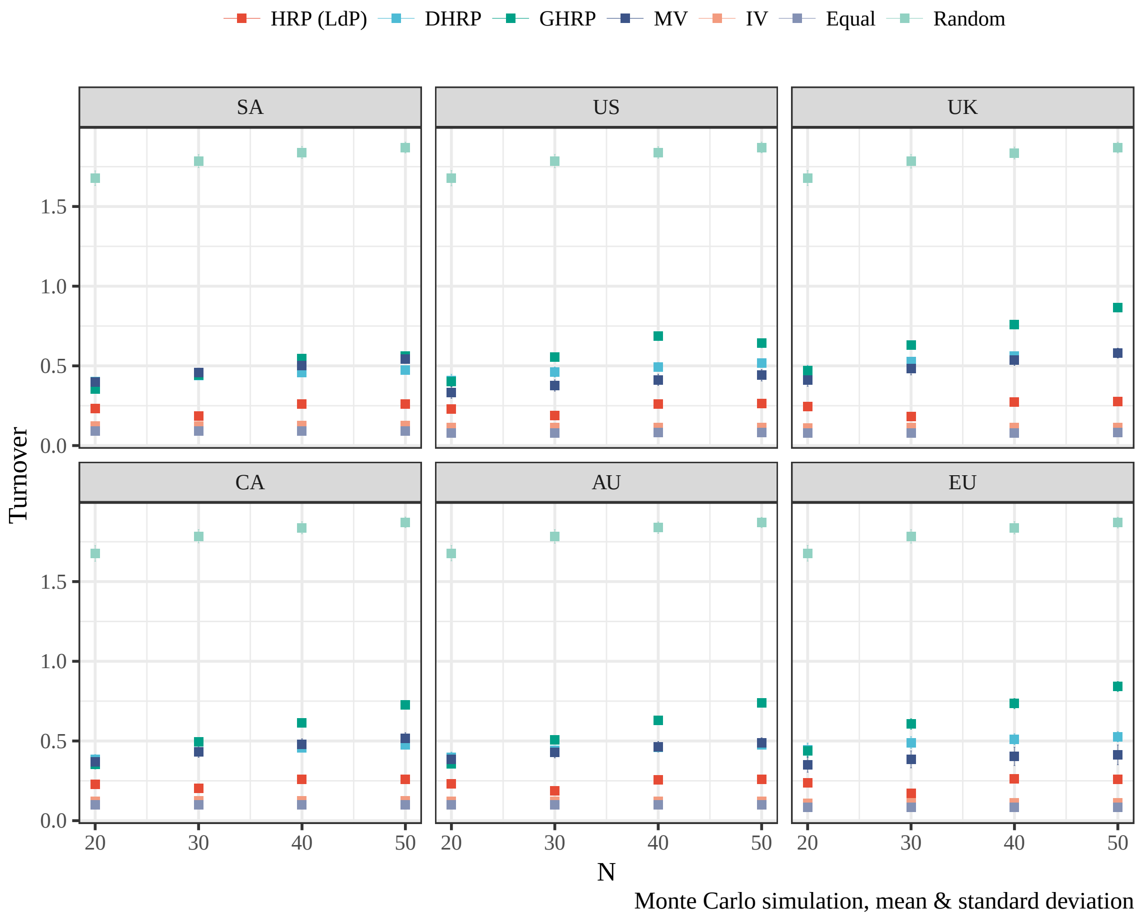

Figure 5.9: All Universes Portfolio Turnover

5.3. Portfolio Performance: Aggregate Universes

Figure 5.10 plots the volatility, risk-adjusted return and portfolio turnover for portfolios of $N = 50 - 200$ drawn in a Monte Carlo simulation from the aggregated 6 equity universes. GHRP is omitted due to its relative computational cost and similarity to DHRP. The results reaffirm the findings discussed above, indicating the superior performance of DHRP over HRP, while the MV portfolio continues to perform well in terms of volatility. The MV porftolio furthermore achieves high risk-adjusted returns for the large N portfolios, while turnover is once again signicantly lower for HRP than MV and DHRP. The feasibility of the MV portfolio is ensured by approximating the nearest invertible covariance matrix according to the algorithm provided by Higham (2002).

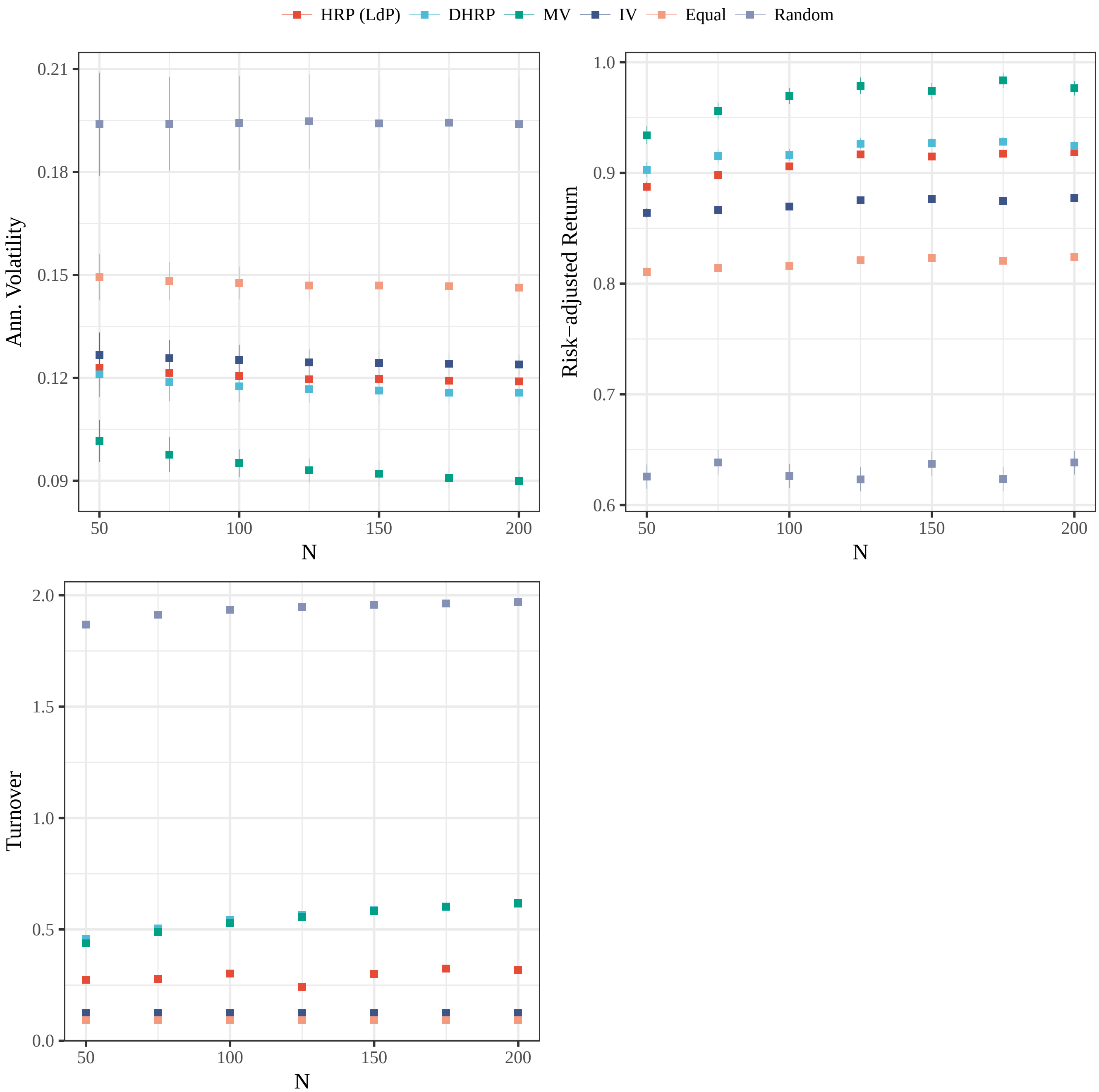

Figure 5.10: Global Annualised Volatility, Risk-adjusted Return and Turnover

5.4. Bootstrap results

A further robustness test is performed, which takes into accout the survivorship bias inherent in the Monte Carlo simulation. The full-index samples are utilized to calculate portfolio backtests for each method with sample adjustment over time to reflect index membership. A bootstrap confidence interval is computed for the results, and the market-capitalisation weighted index volatility is overlayed.

The outcomes are approximately identical to those in previous sections: DHRP outperforms HRP, both of which achieve lower volatility than the IV portfolio as well as naive allocations. Apart from the random portfolio, all allocations achieve lower volatility than the market-capitalisation weighted index and the MV portfolio achieves the lowest overall volatility.

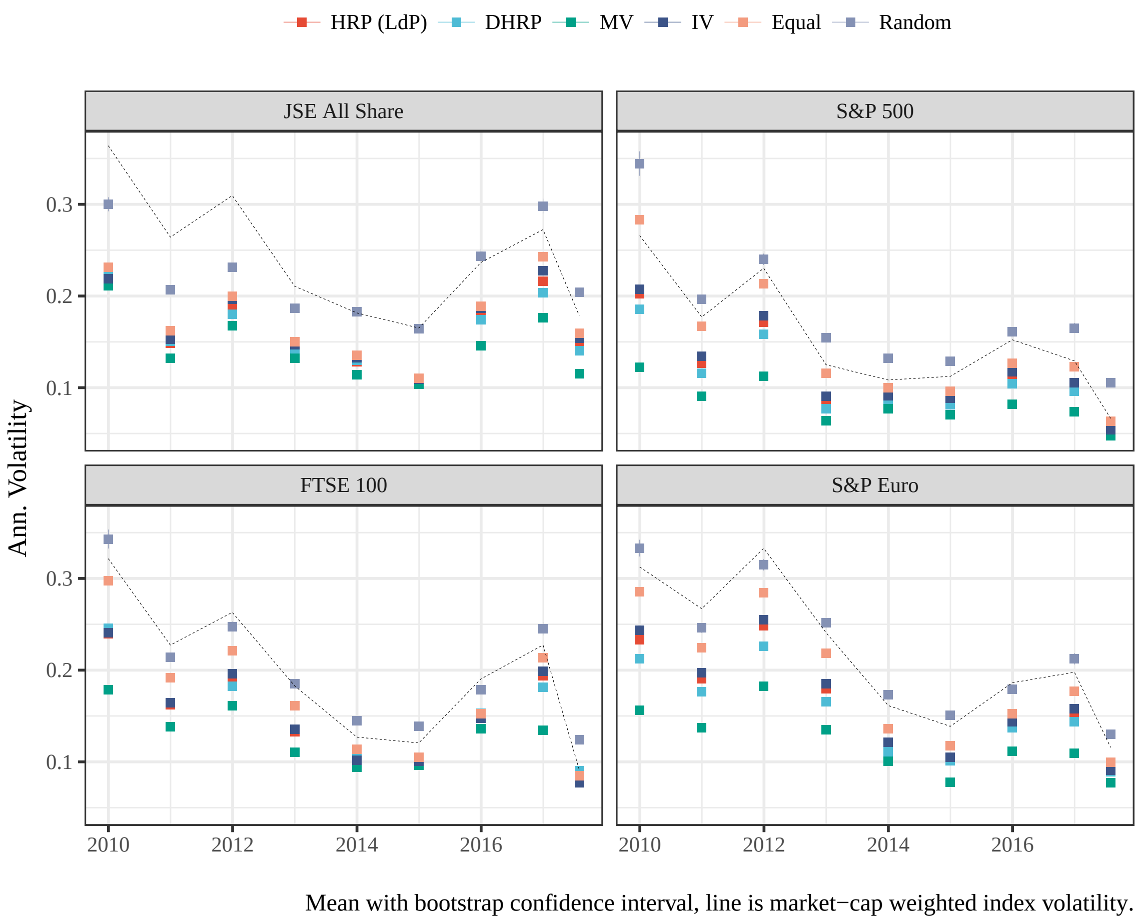

Figure 5.11: Annualised Index Volatility

5.5. Portfolio Concentration

The extent of portfolio concentration has been a concern for quadratic programming optimisers like the MV portfolio (DeMiguel, Garlappi, and Uppal 2009). To compare the performance of HRP against MV, the Herfindahl indices for the portfolio and industry weights are computed for each portfolio and plotted in Figure 5.12 (JSE universe, Monte Carlo simulation). As discussed in Section 4, the index is normalised by the equal weighted naive portfolio, to remove sample-specific concentration effects, highlighting instead the allocation-specific concentration. As expected, the MV portfolio performs poorly, becoming more concentrated at higher N. This demonstrates that adding more assets does not necessarily improve diversification. The HRP portfolio performs well on both measures, with the DHRP portfolio somewhat more concentrated. Industry concentration is significantly reduced at higher N for the DHRP portfolio, with both HRP variants again significantly outperforming the MV portfolio.

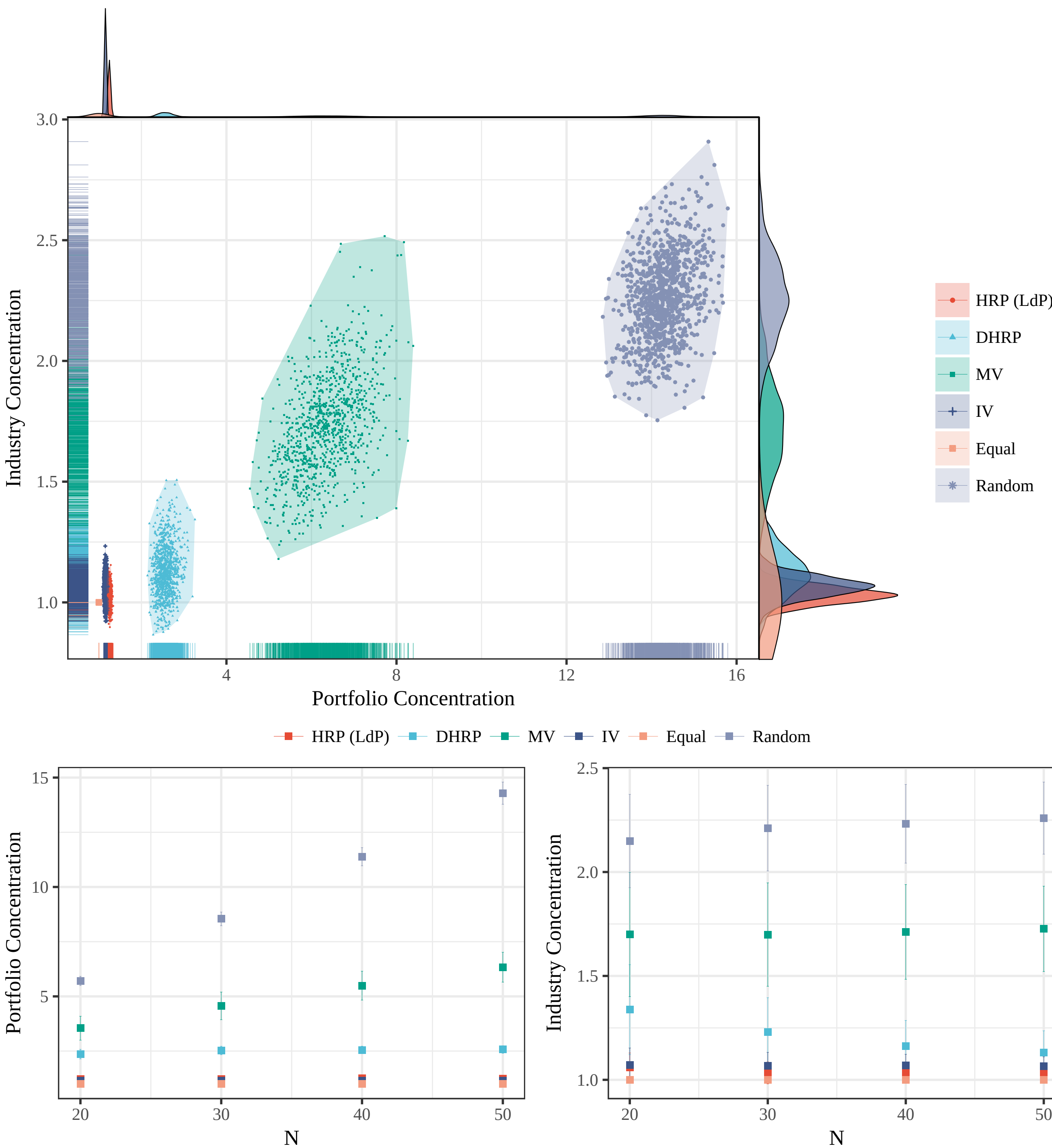

Figure 5.12: Portfolio Concentration Metrics (top: $N = 50$, bottom: $N = 20 - 50$)

The results underscore the effectiveness of the HRP algorithm as a diversification tool. The use of correlation clusters as the foundation for the weight-calculation ensures that concentration is stable and almost as low as the naive equally-weighted portfolio. Viewed in conjunction with the highly competitive out-of-sample performance of the algorithm, HRP offers an appealing alternative to quadratic programming based solutions.

6. Concluding Remarks

As with other risk-based optimisation techniques, the HRP algorithm computes weights exclusively based on the variance characteristics captured by the covariance matrix. Further enhancements to the quality and robustness of the information contained in the covariance matrix can improve the HRP allocation. While a detailed demonstrations exceeds the scope of this study, simulation results suggest that significant volatility reductions can be achieved by combining DHRP with an exponentially weighted multivariate GARCH estimator of the covariance matrix[10]. In addition, the temporal stability of clusters can be increased (lowering portfolio turnover) by applying a shrinkage estimate of the covariance matrix (e.g. Oracle-adjusted shrinkage). This is suggested in Raffinot (2016) and found to yield promising results. Allocation trees with $0 < \tau < 1$ and multiple splits per recursive iteration further enhance cluster stability over time compared to conventional HRP due to flatter hierarchies.

While the scope for additional enhancements naturally is vast, this paper suggests that the HRP algorithm delivers competitive out-of-sample results over a spectrum of relevant performance metrics. The method is useful for small and large N portfolios, as shown by the extensive simulations conducted here. The use of dissimilarity-based recursive allocation trees permits full utilisation of the information obtained from the clustering algorithm and unlocks additional performance improvements, positioning DHRP and GHRP as close competitors to the MV portfolio. While DHRP and GHRP do not attain the low levels of volatility of the MV portfolio, they generally exhibit higher return, lower turnover and better diversification, coupled with the attractive feature of being capable of calculating weights for singular covariance matrices. HRP with naive allocation is mostly invariant to the class of cluster algorithm applied, and attains superior results in terms of turnover and concentration metrics. The finding suggests that the allocation trade-off between symmetry and dissimilarity ($\tau = 0$ vs $\tau = 1$) manifests in the balance between higher diversification and better risk-return characteristics of the resulting portfolio.

[10]The RiskMetrics VECH model is used in simulations.

Appendix A: HRP Algorithm

The HRP algorithm is outlined mathematically below, including various modifications to the original method by Lopez de Prado (2016). The notation is borrowed from Lopez de Prado (2016).

A.1 Clustering

Let ρ_{ij} be the correlation coefficient between assets i and j, and $\rho = \{\rho_{ij}\}_{i,j=1,...N}$ represent the correlation matrix. The hierarchical clustering is performed on a distance transformation of ρ, such that $d_{ij} = \sqrt{\frac{1}{2}(1-\rho_{ij})}$. The Euclidean distance between columns in $\{d_{ij}\}_{i,j=1,...N}$ is utilized in the clustering algorithm. Lopez de Prado (2016) apply single-linkage agglomerative nesting to $\{d_{ij}\}_{i,j=1,...N}$, a review of which can be found in Maimon and Rokach (2010) and is omitted here. Alternative clustering techniques explored in this paper are the Ward method, and top-down divisive analysis (see Maimon and Rokach (2010)). The `R`-package `cluster` is used to calculate covariance clusters.

A.2 Quasi-Diagonalisation

The quasi-diagonalisation step reorders the rows and columns in ρ such that the largest values lie close to the diagonal. This is achieved by rearranging the matrix based on the ordering generated by the cluster algorithm. A detailed discussion is provided in Lopez de Prado (2016).

Following Alipour et al. (2016), the step can be recast as an optimisation problem. The loss function $\mathcal{L}$ is selected to measure the extent of diagonalisation of the covariance matrix, with correlation coefficients far away from the diagonal given higher weights:

$$\mathcal{L} = \sum_{ij} d_{ij}(i-j)^2 \quad \forall \quad i,j \in [1,...,N] \tag{6.1}$$

$\mathcal{L}$ takes on the value of zero, if ρ is a diagonal matrix. The use of $\{d_{ij}\}$ in $\mathcal{L}$, sets the approach apart from Alipour et al. (2016), who propose a version of $\mathcal{L}$ which uses $\{|\rho_{ij}|\}_{i,j=1,...,N}$ instead of $\{d_{ij}\}$. The specification presented here ensures that the information used for the clustering and quasi-diagonalisation steps is identical for GHRP and HRP, facilitating direct comparisons.

An allocation tree can be derived for the GHRP approach, by recursively searching over the range of N for the split that maximises the mean absolute correlation of off-diagonal cluster blocks in the matrix $\{\rho_{ij}\}_{i,j=1,...,N}$ reflecting the optimised order, as suggested in Alipour et al. (2016). The allocation tree is necessary when $\tau = 1$.

A.3 Recursive Bisection

The recursive algorithm consists of the following steps (see Lopez de Prado (2016)):

1. Initialise a list of assets in the portfolio, denoted L_0
2. Initialise a vector of unit weights, $w_i = 1 \quad \forall \quad i \in [1, ..., N]$
3. Stop if $|L_i| = 1 \quad \forall \quad L_i \in L$
4. For each $L_i \in L$ such that $|L_i| > 1$:
    a. Split L_i into two or more subsets $L_i^{(j)}$ preserving the order with $L_i^{(1)} \cup \cdots \cup L_i^{(J)} = L_i$
    b. Calculate the variance of $L_i^{(j)}$ such that $\tilde{V}_i^j = \tilde{w}_i^{j\prime} V_i^j \tilde{w}_i^j$, where V_i^j is the covariance matrix of elements within cluster j, and $\tilde{w}_i^j = \frac{\mathrm{tr}[V_i^j]^{-1}}{\sum_i \mathrm{tr}[V_i^j]^{-1}}$
    c. The unconstrained weight allocation is computed using the split factor $\alpha_i = \frac{[\tilde{V}_i^j]^{-1}}{\sum_j [\tilde{V}_i^j]^{-1}}$. The calculation of α_i is more general than Lopez de Prado (2016), in order to permit more than one split for each L_i in the hierarchy ($J \geq 2$).
5. Enforce weight constraints:
    a. Let $\tilde{\alpha}_i = \max[\min[\alpha_i, \sum_i \bar{\alpha}_i^{max}], \sum_i \bar{\alpha}_i^{min}]$, where $\bar{\alpha}_i^{max}$ and $\bar{\alpha}_i^{min}$ are the respective $1 \times N$ constraint vectors
    b. Now let $\tilde{\alpha}_k = \tilde{\alpha}_k + |1 - \sum_i \tilde{\alpha}_i| \frac{\alpha_k}{\sum_k \alpha_k}$, where k denotes elements in the subset of L_i with $\tilde{\alpha}_i = \alpha_i$
    c. Repeat step b until $\sum_i \tilde{\alpha}_i = 1$
6. Rescale w_i by $\tilde{\alpha}_i$
7. Loop to step 3

The resulting $\tilde{\alpha}_i$ satisfies two objectives: (i) it does not violate any constraints, and (ii) the weights within the bounds set by the constraints retain their relative proportions. As such, $\tilde{\alpha}_i$ is the constrained allocation that is closest in proportion to the unconstrained weights α_i. Furthermore, convergence is typically achieved instantaneously or after very few iterations (depending on the complexity of the weight constraints).

A useful feature of the method proposed above is that constraint violations are resolved as low in the hierarchy as possible. The allocation between higher level clusters is retained in unconstrained proportions as long as it is possible to resolve violations within sub-clusters.

The above method can easily be extended to facilitate group constraints by inserting the following steps before step 5:

1. Define $\bar{\alpha}_i^{max} = \alpha_i \frac{\sum_j \bar{g}_j^{max}}{\sum_j \alpha_j}$ and $\bar{\alpha}_i^{min} = \alpha_i \frac{\sum_j \alpha_j}{\sum_j \bar{g}_j^{min}}$, where $\bar{g}_i^{max}$ and $\bar{g}_i^{min}$ are the group constraints, and j denotes that subset of elements belonging to the same group as i.
2. Continue with 5. to calculate $\tilde{\alpha}_i$

Appendix B: Portfolio Backtest with Box Constraint

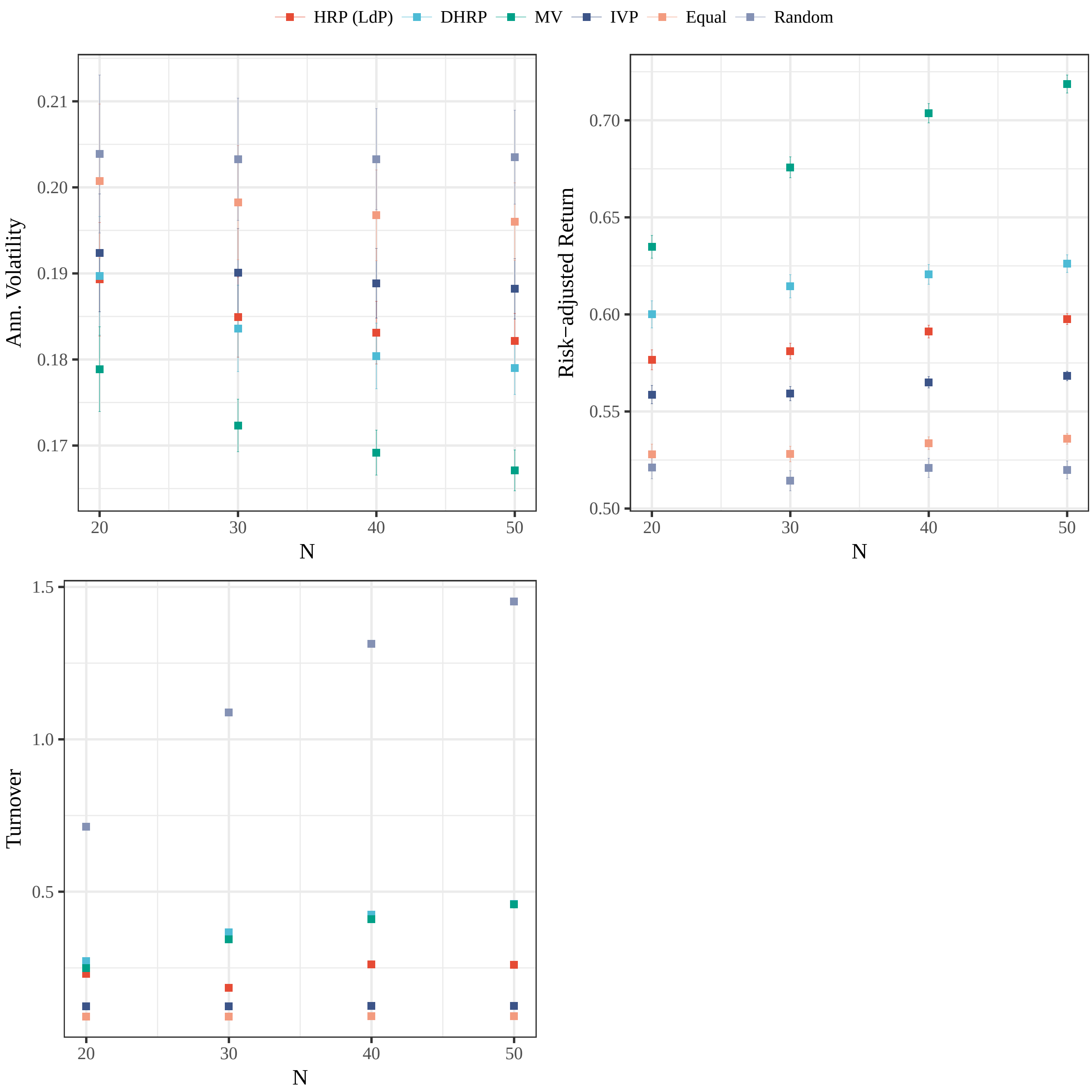

Figure 6.1: SA Constrained Annualised Volatility, Risk-adjusted Return and Turnover

References

Alipour, Elham, Clemens Adolphs, Arman Zaribafiyan, and Maxwell Rounds. 2016. "Quantum-Inspired Hierarchical Risk Parity." 1QBit.

Bloomberg. 2017. "Bloomberg Professional Subscription Service."

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. "Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?" *Review of Financial Studies* 22 (5): 1915–53.

Di Matteo, T., T. Aste, and R.N. Mantegna. 2004. "An Interest Rates Cluster Analysis." *Physica A: Statistical Mechanics and Its Applications* 339 (1-2): 181–88.

Even, S. 1973. *Algorithmic Combinatorics.* New York: The Macmillan Company.

Higham, Nicholas J. 2002. "Computing the Nearest Correlation Matrix—a Problem from Finance." *IMA Journal of Numerical Analysis* 22 (3): 329–43.

Kolm, Petter N., Reha Tütüncü, and Frank J. Fabozzi. 2014. "60 Years of Portfolio Optimization: Practical Challenges and Current Trends." *European Journal of Operational Research* 234 (2): 356–71.

León, Diego, Arbey Aragón, Javier Sandoval, Germán Hernández, Andrés Arévalo, and Jaime Niño. 2017. "Clustering Algorithms for Risk-Adjusted Portfolio Construction." *Procedia Computer Science* 108: 1334–43.

Lopez de Prado, Marcos. 2016. "Building Diversified Portfolios That Outperform Out-of-Sample." *SSRN Electronic Journal.*

Maimon, Oded, and Lior Rokach, eds. 2010. *Data Mining and Knowledge Discovery Handbook.* Boston, MA: Springer US.

Mantegna, Rosario N. 1999. "Hierarchical Structure in Financial Markets." *The European Physical Journal B-Condensed Matter and Complex Systems* 11 (1): 193–97.

Markowitz, Harry. 1952. "Portfolio Selection." *The Journal of Finance* 7 (1): 77–91.

Raffinot, Thomas. 2016. "Hierarchical Clustering Based Asset Allocation." *SSRN Electronic Journal.*

Roncalli, Thierry. 2014. *Introduction to Risk Parity and Budgeting.* Chapman & Hall/CRC Financial Mathematics Series. Boca Raton: Taylor & Francis.

Rubinstein, Mark. 2002. "Markowitz's 'Portfolio Selection': A Fifty-Year Retrospective." *The Journal* *of Finance* 57 (3).

Steinbach, Marc C. 2001. "Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis." *SIAM Review* 43 (1): 31–85.

Tola, Vincenzo, Fabrizio Lillo, Mauro Gallegati, and Rosario N. Mantegna. 2008. "Cluster Analysis for Portfolio Optimization." *Journal of Economic Dynamics and Control* 32 (1): 235–58.

Tumminello, Michele, Fabrizio Lillo, and Rosario N. Mantegna. 2010. "Correlation, Hierarchies, and Networks in Financial Markets." *Journal of Economic Behavior & Organization* 75 (1): 40–58.