Journal of Mathematical Finance, 2022, 12, 179-195
https://www.scirp.org/journal/jmf
ISSN Online: 2162-2442
ISSN Print: 2162-2434

Evaluating Hierarchical Equal Risk Contribution Portfolios in the Chinese Stock Market

Weige Huang[1], **Xiang Gao**[2*]

[1]Wenlan School of Business, Zhongnan University of Economics and Law, Wuhan, China
[2]Research Center of Finance, Shanghai Business School, Shanghai, China
Email: *gaoxiang@sbs.edu.cn

How to cite this paper: Huang, W.G. and Gao, X. (2022) Evaluating Hierarchical Equal Risk Contribution Portfolios in the Chinese Stock Market. *Journal of Mathematical Finance*, **12**, 179-195.
https://doi.org/10.4236/jmf.2022.121011

Received: January 20, 2022
Accepted: February 18, 2022
Published: February 21, 2022

Abstract

This paper investigates the usefulness of the Hierarchical Equal Risk Contribution algorithm to exploit correlation structure in China's equity market over 2001-2020. By running a horse race of different combinations of metrics and linkages, we demonstrate that the winner strategy always beats traditional portfolio construction techniques. Better-performing risk-based hierarchy strategies vary with stock-sorting methods by size, mean return, volatility, and Sharpe ratio. However, our treatment results in extremely imbalanced asset allocation, implying that we capture information other than the standard Chinese industrial classification.

Keywords

Hierarchical Equal Risk Contribution, Machine Learning, Hierarchical Risk Parity, Asset Allocation, Critical Line Algorithm, Inverse-Variance Portfolio

1. Introduction

Equity investment exhibits many attributes, such as size, liquidity, region, industry, etc. A key aspect of equity-market complexity is the hierarchical way via which stocks correlate with each other. While the correlation matrix lacks the notion of hierarchy [1], there exists an urgent need for characterizing such structure long recognized in other areas ([2] [3]) for the financial markets, especially for investment diversification in less efficient ones. At any level of the hierarchy in a portfolio, stocks within a given attribute group compete for allocation [4]. For example, when adjusting for capital allocated to a large-cap financial stock, managers will buy or sell another big bank, rather than some small-cap shares, foreign assets, or real estate holdings.

Given the above said, one strand of recent literature seeks to quantitatively exploit the complementary role of hierarchical correlations in conventional asset allocation techniques like the Critical Line Algorithm (CLA) [5]. However, these techniques still rely on classical portfolio optimization ([5] [6] [7]). Introducing hierarchy will completely change the classical setting. The ultimate goal is to develop more efficiently diversified portfolios that can address CLA caveats and thus outperform out-of-sample. The seminal work by López de Prado [1] introduces the Hierarchical Risk Parity (HRP) algorithm consisting of three implementation steps: hierarchical tree clustering, matrix seriation, and recursive bisection. Theoretically speaking, HRP could overcome the drawbacks of quadratic optimizers such as the instability, concentration, and under-performance problem ([4] [8]). Up-to-date HRP application research also suggests that the use of hierarchy identified by clustering is indeed helpful in achieving an optimal weight allocation ([9] [10] [11]). Particularly, Raffinot [11] proposes the Hierarchical Clustering based Asset Allocation (HCAA), which agrees with the waterfall idea of HRP and, is inspired by DeMiguel *et al.* [12], features in dividing capital equally among hierarchical clusters and computing an equal-weighted allocation within each stock cluster. Yet, without incorporating sophisticated risk measures, HCAA's naive treatment of equal distribution suffers from over simplicity and subjectivity.

Via integrating the HRP and HCAA method, the Hierarchical Equal Risk Contribution (HERC) algorithm adopts machine learning to allocate weights across and within asset clusters ([13] [14]). HERC resembles HRP since both of them start by reorganizing the covariance matrix to place similar investments together. But HERC differs from HRP in that HRP makes no further use of clustering after an inverse-variance weighting allocation based on the number of assets. HERC has the advantage of HCAA's double-layer weighting scheme and alternative risk metrics allowed for investors. Not limited to standard deviation and expected shortfall, one can extend HERC to include downside risk measures such as Conditional Value at Risk (CVaR) and Conditional Drawdown at Risk (CDaR). According to Raffinot [14], although the "Hierarchical 1/N" performance of HCAA is hard to beat on a return-only viewpoint, HERC-CDaR portfolios can attain better risk-adjusted returns statistically.

Extending from the general ideas of HERC, we attempt to digest in detail how HERC performs in China's equity market using variant implementations of HERC strategies built on portfolio quantiles of stocks sorted according to size, return, variance, and Sharpe ratio. We argue that the motivation of these endeavors is attributable to the fact that stocks are not only correlated hierarchically but also affected by investors' reliance on sorting and rotation to attain higher risk-adjusted returns. Following the line of research on hierarchical correlation with a focus on evaluating its usefulness for sorted Chinese equities, this paper first shows that the best HERC strategy is sensitive to different ways of constructing portfolios. Second, we discover that the clusters suggested by HERC are imbalanced due to many portfolios included in one stock cluster and the abnormal high weight assigned to a single portfolio. Thus, we conclude that the HERC-based portfolios are unlikely to capture the industry structures in China. This implies the clustering trees of the Chinese stock market depart considerably from the common industry classification, and data-driven approaches are needed for the China case. Another contribution of our research is that investors should expect distinct clusters compared to regional or sectoral categories when applying HERC approaches at the individual-stock level. While whether HERC is superior to simple diversification across sectors, regions, or style factors under all scenarios is still an open question, our finding adds to the literature by pointing out the limitations of frequently-used classifications employed over a top-down portfolio construction process.

The remainder is organized as follows. Section 2 introduces the methods and data used. Section 3 presents and discusses the corresponding empirical results. Section 4 concludes by summarizing the contribution and pointing out possible future extensions.

2. Methodology and Data

Aiming at diversifying risk allocation, HERC is an algorithmic approach of using tree clustering to explicitly consider hierarchical structure in the investment universe by grouping assets and determining weights. The implementation of HERC consists of four steps. First, we calculate hierarchical clusters that can graphically represent real-life interactions between assets in a portfolio [1]. Second, we identify the optimal number of stock clusters by trimming the tree formed previously according to the Gap Index ([15]) and Optimal Number of Clusters (ONC) algorithm ([8] [16]).[1] Third, we determine the weights of each cluster using Top-Down Recursive Bisection. At each level of recursion in a tree, the following equal risk contribution weights are adopted:

$$\alpha_1 = 1 - \frac{RC_1}{RC_1 + RC_2};\ \alpha_2 = 1 - \alpha_1,$$

where α_1 and α_2 are the weights of the left and right clusters, respectively, and RC_1 and RC_2 represent the corresponding risk contributions of these two clusters.[2]

Fourth, final weights are assigned to assets residing in every cluster under

[1]ONC detects the optimal number of K-Means clusters using a correlation matrix as input. See López de Prado [17] for a brief description of the logic behind the ONC algorithm and López de Prado [18] for the reason why the angular distance metric can be used to obtain distances between elements.

[2]Asset weights in the cluster are computed as $w = diag(\Sigma)^{-1} / tr\left(diag(\Sigma)^{-1}\right)$, where Σ denotes the covariance matrix for assets in the same cluster, *diag* the matrix diagonal, and *tr* the trace. Then, we can write the risk contribution of asset *i* as $RC_{w_i} = w_i(\Sigma w)_i / \sqrt{w'\Sigma w}$. Lastly, we have $RC_j = N \cdot RC_{w_i}$, where *N* means the total number of assets in cluster *j*. It merits a note that we adopt variance to be the risk measure for illustration. However, when implementing the HERC, the way to compute *w* can be different if alternative risk metrics other than the variance are selected.

naive risk parity allocation. Particularly, an initial set of asset weights w_{NRP} within the same cluster is adjusted in proportion to the inverse of assets' respective risk—higher risk assets will receive lower portfolio weights and lower risk assets will receive higher weights. The risk here can be quantified by a range of proxies such as variance, CVaR, CDaR, maximum daily loss, etc. These final asset weights are:

$$w_{\text{final}}^{i} = w_{\text{NRP}}^{i} C^{i},\ i \in \text{Clusters},$$

where w_{NRP}^{i} refers to native parity weights of assets in the ith clusters and C^{i} is the weight of the ith cluster.

Given the above-mentioned four steps, HERC allocates capital within and across the right number of clusters of assets at multiple hierarchical levels. What matters most in such a portfolio construction process are the linkage criteria employed and risk metrics chosen. Recall in the first step, two similar (or least dissimilar) clusters are merged into one to produce a single cluster at the next higher level. Therefore, hierarchical clustering requires a suitable distance measure of dissimilarity between two clusters. There are four common linkage criteria including single, complete, average, and ward's linkage ([11] [13]), this paper begins analysis with the simplest distance shown below [19]:

$$D_{i,j} = \sqrt{2\left(1-\rho_{i,j}\right)},$$

where $D_{i,j}$ is the correlation-distance index between the ith and jth asset, and $\rho_{i,j}$ is the respective Pearson's correlation coefficient.[3]

At last, there remain open questions of choosing what measures for equity risks and which performance evaluation criteria. On the one hand, concerning risk metrics needed for computing risk contribution and final asset weights in the above third and fourth steps, this paper considers variance, standard deviation, expected shortfall, and conditional drawdown risk to be risk metrics, and we adopt market value instead of equal asset allocation as the weighting scheme. On the other hand, to compare the out-of-sample performance of investment strategies established with variants of HERC algorithms, we resort to the adjusted Sharpe ratio (ASR), certainty-equivalent return (CER), and maximum drawdown (MDD) as comparison indicators. In specific, ASR incorporates a penalty factor for negative skewness and excess kurtosis [20]:

$$\text{ASR} = \text{SR}\left[1+\frac{\mu_3}{6}\text{SR}-\frac{\mu_4-3}{24}\text{SR}^2\right],$$

where μ_3 and μ_4 are, respectively, the skewness and kurtosis of return distribution and SR denotes the traditional Sharpe ratio with a risk-free rate set at 3.

Besides, CER is defined as [12]:

$$\text{CER} = \left(\mu - r_f\right) - \frac{\gamma}{2}\sigma^2,$$

where $\gamma = 1$ means risk aversion. The above quadratic utility represents the

[3]Other distance measures based on non-linear codependency can also be used, e.g., see López de Prado [8].

level of expected utility of a mean-variance investor, hence playing a crucial role in building profitable portfolios [21]. MDD indicates a permanent loss of capital, which purportedly measures the largest single drop from the peak to the bottom of portfolio value, *i.e.*, the worst-case scenario.

The historical daily prices of Chinese stocks used in this study are sourced from China Stock Market & Accounting Research Database. We collect the daily closing prices for all A-share Chinese stocks listed in the Shanghai and Shenzhen stock markets during the sample period from January 1st, 2000 to January 1st, 2020. Then, we create a "rebalancing window" which starts one year before the rebalancing day and ends one month after that day. The portfolio is rebalanced monthly on the first trading day of each month. On the rebalancing day, the last year's daily trading prices are used to calculate portfolio weights and those prices in the following month are used to compute the future portfolio return. Thus, we keep in our sample stocks that have at least one year's trading record before and a month's trading record after the portfolio rebalancing day.[4] The minimum and the maximum number of stocks in all rebalancing days during the twenty years are 132 and 2515, respectively.[5]

Turning to the sort procedure, we divide sample stocks into twenty quantiles by either their sizes (market values) evaluated on the trading day right before rebalancing or their mean returns, volatilities and Sharpe ratios during the last year prior to the rebalancing day. The value-weighted returns (equally-weighted returns are adopted as a robustness check; the corresponding main conclusions are similar thus not reported) are then computed at the quantile level. Finally, we apply variants of the HERC algorithm one by one to these twenty quantile portfolios. The reason HERC is not implemented directly at the individual stock level is that rotating among single stocks is unrealistic due to high transaction costs. Nevertheless, we run the exercise without sorting. And the results stay almost the same except for two differences. On the one hand, the optimal number of clusters becomes twice larger. On the other hand, the number of single stocks included in each cluster turns out to be extremely imbalanced. Specifically, in many cases, we observe one cluster including almost all stocks but there are only a few stocks and even one stock included in each of the other clusters. For these reasons, we study HERC performance using portfolios instead of individual stocks hereinafter. The present paper takes advantage of the HERC python package "portfoliolab" provided by Hudson & Thames. The "HierarchicalEqualRiskContribution" and "CriticalLineAlgorithm" class in this package are employed to compute the portfolio weights based on different combinations of risk measures and linkage criteria.[6]

[4]Note that before constructing the rebalancing window we fill forward the dataset for missing values. This could potentially lead to columns of many duplicates. We hence drop stocks with more than five duplicated prices in each of their rebalancing windows. This treatment is equivalent to dropping stocks with prices unchanged for a week.

[5]The minimum is very small because during 2006-2007 many stocks are not traded continuously. And given our selection criteria, these stocks are excluded.

[6]This paper adopts the 5% confidence level for calculating the expected shortfall and conditional draw-down at risk. For simplicity, the risk-free rate is set to be zero without the loss generality.

3. Empirical Results

In this section, we present and interpret the performance of HERC investment strategies relative to benchmark returns in Table 1, with HERC portfolio characteristics reported in Table 2. Notice that these strategies are built on twenty portfolios of stocks in Table 3 sorted by size, mean return, volatility, and Sharpe ratio. In these tables, different HERC strategies are denoted by letting the first several letters in their name represent the risk metric used (e.g., V, SD, ES, and CDR stand for variance, standard deviation, expected shortfall, and conditional drawdown risks, respectively) and letting the last letter represent the type of linkage used (e.g., S, C, A, and W stand for single, complete, average, and ward's linkage, respectively). Benchmark returns are calculated using CLA portfolios which are established based on weights from maximum Sharpe ratio, inverse-variance portfolios (IVP), equal-weighted (EW) portfolios, and Shanghai Stock Exchange (SSE) Composite Index.

Table 1. Comparing the performance of different investment strategies.

	Size			Mean Return			Volatility			Sharpe Ratio		
	ASR	CER	MDD	ASR	CER	MDD	ASR	CER	MDD	ASR	CER	MDD
CLA	**0.28**	0.04	0.90	−0.38	−0.18	0.97	0.14	−0.00	0.90	−0.30	−0.14	0.95
ESS	**0.24**	0.03	0.90	0.18	0.01	0.91	0.21	0.02	0.92	**0.25**	0.03	0.93
SDS	**0.23**	0.03	0.89	0.16	0.00	0.90	0.20	0.02	0.92	**0.23**	0.02	0.93
IVP	**0.23**	0.02	0.93	**0.23**	0.02	0.93	0.23	0.02	0.93	0.23	0.02	0.93
EW	**0.21**	0.02	0.93	0.21	0.02	0.93	0.21	0.02	0.93	0.21	0.02	0.93
VS	0.21	0.02	0.89	0.17	0.00	0.90	0.21	0.02	0.92	0.23	0.02	0.93
EWS	0.21	0.02	0.89	0.18	0.01	0.91	0.22	0.02	0.92	0.22	0.02	0.93
SSE	0.21	0.02	0.83	**0.21**	0.02	0.83	0.21	0.02	0.83	0.21	0.02	0.83
ESW	0.19	0.01	0.86	0.16	−0.00	0.91	0.18	0.01	0.91	0.19	0.01	0.94
CDRS	0.19	0.01	0.88	0.18	0.01	0.91	0.20	0.01	0.92	0.22	0.02	0.93
SDW	0.18	0.01	0.85	0.15	−0.00	0.91	0.16	0.00	0.90	0.18	0.01	0.93
EWW	0.17	0.01	0.85	0.15	−0.00	0.91	0.15	0.00	0.90	0.16	0.00	0.93
ESC	0.17	0.01	0.88	**0.21**	0.02	0.91	0.23	0.02	0.91	**0.24**	0.03	0.93
VW	0.16	0.01	0.84	0.16	0.00	0.91	0.17	0.01	0.90	0.17	0.01	0.93
SDC	0.16	0.01	0.87	**0.19**	0.01	0.91	0.22	0.02	0.91	0.23	0.02	0.93
SDRC	0.14	0.00	0.87	0.15	−0.00	0.90	0.20	0.02	0.91	0.23	0.02	0.93
CDRW	0.14	0.00	0.85	0.15	−0.00	0.91	0.17	0.01	0.91	0.20	0.02	0.93
EWC	0.14	0.00	0.87	**0.20**	0.01	0.91	**0.26**	0.03	0.92	0.22	0.02	0.92
VC	0.14	0.00	0.86	0.18	0.01	0.91	0.23	0.02	0.91	0.22	0.02	0.93
CDRA	0.10	−0.01	0.85	0.15	−0.00	0.89	0.22	0.02	0.92	0.23	0.02	0.93
SDA	0.09	−0.01	0.84	0.12	−0.01	0.88	**0.24**	0.03	0.92	**0.23**	0.03	0.93
EWA	0.09	−0.01	0.85	0.14	−0.00	0.88	**0.26**	0.03	0.92	0.23	0.02	0.93
ESA	0.06	−0.02	0.85	0.13	−0.01	0.88	**0.24**	0.03	0.92	**0.24**	0.03	0.93
VA	0.05	−0.02	0.86	0.12	−0.01	0.87	**0.24**	0.03	0.92	0.23	0.02	0.93

Note: The rows are sorted by values reported in the first column in descending order. Non-HERC benchmark strategies are in bold. For each performance indicator, the top five best-performing strategies according to the ASR criteria are underlined and in bold.

Table 2. Optimal clustering.

	Size				Mean Return				Volatility				Sharpe Ratio			
	ONC	N	W	N^{th}	ONC	N	W	N^{th}	ONC	N	W	N^{th}	ONC	N	W	N^{th}
CLA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ESS	3	16	0.4	16	5	11	0.2	10	5	12	0.3	7	5	15	0.5	9
SDS	3	16	0.5	17	5	11	0.2	10	4	13	0.3	8	5	15	0.5	9
IVP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VS	3	16	0.4	17	5	12	0.2	10	4	12	0.3	7	5	15	0.5	10
EWS	3	16	0.4	17	5	11	0.2	10	4	13	0.3	6	5	15	0.5	10
SSE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ESW	3	16	0.5	16	6	15	0.8	8	6	15	0.8	11	6	14	0.8	8
CDRS	3	16	0.4	16	5	11	0.2	8	4	13	0.3	9	4	15	0.4	9
SDW	3	16	0.5	16	6	15	0.8	8	6	15	0.8	11	6	15	0.8	8
EWW	3	16	0.5	16	6	15	0.8	8	6	15	0.8	11	6	15	0.7	8
ESC	3	17	0.6	17	4	15	0.5	8	4	16	0.6	8	4	12	0.3	10
VW	3	16	0.5	16	6	15	0.8	8	6	15	0.8	11	6	15	0.8	8
SDC	3	17	0.6	17	4	16	0.5	8	4	16	0.6	8	4	12	0.2	10
SDRC	3	17	0.5	17	4	15	0.4	8	4	16	0.4	8	4	12	0.2	8
CDRW	3	16	0.4	16	6	15	0.5	8	6	15	0.5	11	6	15	0.5	8
EWC	3	17	0.6	17	4	15	0.5	8	4	16	0.6	8	4	12	0.2	10
VC	3	17	0.6	17	4	15	0.5	8	4	16	0.6	8	4	12	0.2	10
CDRA	5	16	0.5	16	4	14	0.3	9	4	15	0.3	8	4	14	0.3	8
SDA	5	16	0.8	16	4	13	0.3	10	4	15	0.4	7	4	14	0.3	9
EWA	5	16	0.8	16	4	13	0.3	10	4	15	0.4	6	4	14	0.3	10
ESA	5	16	0.7	16	4	13	0.3	9	4	15	0.4	7	4	13	0.3	9
VA	5	16	0.8	16	4	13	0.3	10	4	15	0.4	7	4	14	0.3	10

Note: For each investment strategy in our sample period, ONC denotes the average optimal number of clusters; N is the total number of sorted portfolios included in the cluster with the maximal number of portfolios; N^{th} is the rank of the sorted portfolio with the maximal weight, and W represents the maximal weight associated to the N^{th} ranked portfolio.

Table 3. Performance of portfolios sorted by size, return, volatility, and Sharpe ratio.

	Size			Mean Return			Volatility			Sharpe Ratio		
	ASR	CER	MDD	ASR	CER	MDD	ASR	CER	MDD	ASR	CER	MDD
0	0.48	0.11	0.98	0.28	0.04	0.94	0.33	0.05	0.93	0.27	0.04	0.93
1	0.47	0.10	0.98	0.37	0.07	0.96	0.34	0.06	0.94	0.37	0.07	0.95
2	0.40	0.08	0.97	0.37	0.07	0.96	0.25	0.03	0.91	0.34	0.06	0.95
3	0.37	0.07	0.96	0.30	0.05	0.95	0.33	0.05	0.95	0.31	0.05	0.95
4	0.32	0.05	0.97	0.35	0.06	0.95	0.28	0.04	0.95	0.39	0.07	0.96
6	0.28	0.04	0.96	0.34	0.06	0.96	0.36	0.06	0.96	0.40	0.08	0.97
5	0.28	0.04	0.95	0.40	0.08	0.97	0.31	0.05	0.95	0.38	0.07	0.96
SSE	0.21	0.02	0.83	0.21	0.02	0.83	0.21	0.02	0.83	0.21	0.02	0.83

Continued

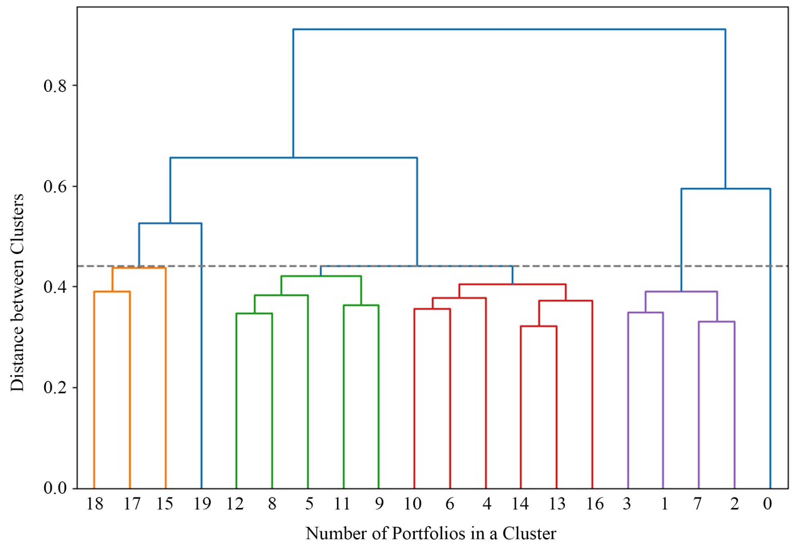
9	0.20	0.01	0.94	0.32	0.05	0.96	0.25	0.03	0.94	0.28	0.04	0.95
10	0.19	0.01	0.93	0.28	0.04	0.95	0.26	0.03	0.95	0.27	0.03	0.95
7	0.18	0.00	0.94	0.39	0.07	0.96	0.27	0.03	0.94	0.37	0.07	0.96
8	0.17	0.00	0.94	0.38	0.07	0.97	0.19	0.01	0.93	0.30	0.05	0.95
13	0.14	−0.01	0.89	0.17	0.00	0.92	0.19	0.01	0.93	0.14	−0.01	0.92
17	0.12	−0.01	0.97	0.02	−0.04	0.87	0.06	−0.04	0.85	−0.03	−0.06	0.87
12	0.10	−0.02	0.91	0.17	0.00	0.92	0.21	0.02	0.93	0.19	0.01	0.91
11	0.08	−0.03	0.86	0.23	0.02	0.93	0.18	0.01	0.92	0.28	0.04	0.95
14	0.07	−0.03	0.86	0.16	0.00	0.91	0.20	0.01	0.93	0.18	0.01	0.92
15	0.05	−0.03	0.86	0.07	−0.03	0.86	0.14	−0.01	0.90	0.05	−0.03	0.85
19	0.04	−0.02	0.85	−0.40	−0.18	0.97	0.03	−0.05	0.85	−0.30	−0.14	0.96
18	0.04	−0.03	0.85	−0.13	−0.09	0.87	0.01	−0.05	0.87	−0.13	−0.09	0.86
16	−0.00	−0.05	0.87	0.05	−0.03	0.87	0.04	−0.05	0.85	0.04	−0.04	0.87

Note: Recall that SSE denotes the Shanghai Stock Exchange Composite Index for China's stock market. "0" and "19" indicate the lowest- and highest-ranked portfolio quantile, respectively. "1" - "18" are portfolio quantiles lying between these two extremes.

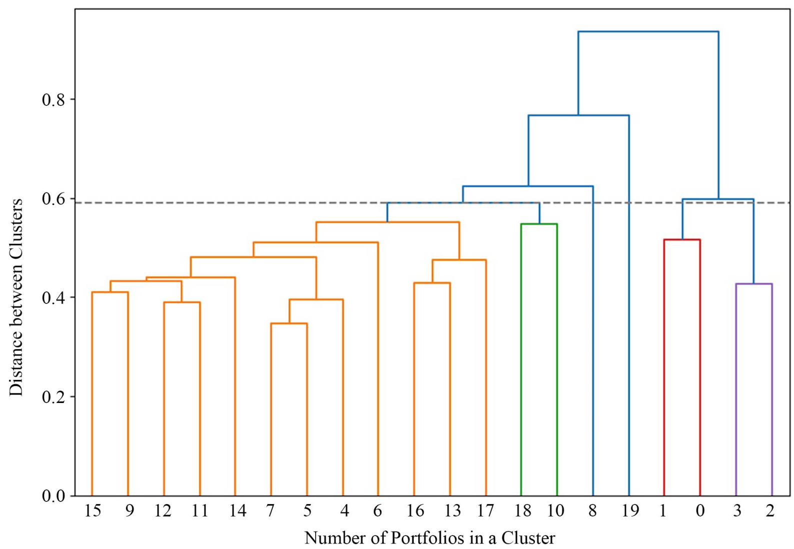
Figure 1 gives an illustration of the dendrograms in different sorting setups. In specific, the upper-left sub-graph shows the best way to allocate Chinese stocks to hierarchical clusters according to stock size. The key to interpreting the dendrogram is to focus on the height at which any two clusters are joined together. Take the up-per-right sub-graph for example. Stock clusters 5 and 7 are most similar in terms of return averaged across time, as the height of the link that joins them together is the smallest. Heights in an informative dendrogram can reflect the distance between the clusters. In the lower-left sub-graph of Figure 1, it shows us that the big volatility difference between clusters is between the stock cluster of 0 and 1 versus that of 2, 3, 4, and 5. Finally, it is important to appreciate that the dendrogram is a summary of the distance matrix, and, as occurs with most summaries, information is lost. For example, the dendrogram of clustering stocks according to Sharpe ratio suggests that stock clusters 17 and 11 are much closer to each other than is 17 to 18, but the original data may tell a different story. Hence, we should be careful when using dendrogram. It is accurate given that data satisfies the ultrametric tree inequality, and this is unlikely for any real-world data.

Focusing on the Size panel of all exhibits, Table 1 indicates that CLA performs the best and several HERC portfolios outperform the IVP, EW, and SSE benchmark. Table 2 tells us that better-performing HERC portfolios usually have a smaller Optimal Number of Clusters (ONC). To be more specific, The ONCs for the top four HERC portfolios, *i.e.*, ESS, SDS, VS, and EWS, (refer to the Size panel of Table 1) all equal to 3 (see the Size panel of Table 2), which are significantly smaller than the peer portfolios such as VA, ESA, EWA, and SDA which have an ONC of 5. Similarly, as for the Mean Return panel, the three best-performing HERC portfolios of ESC, EWC, and SDC (indicated by the

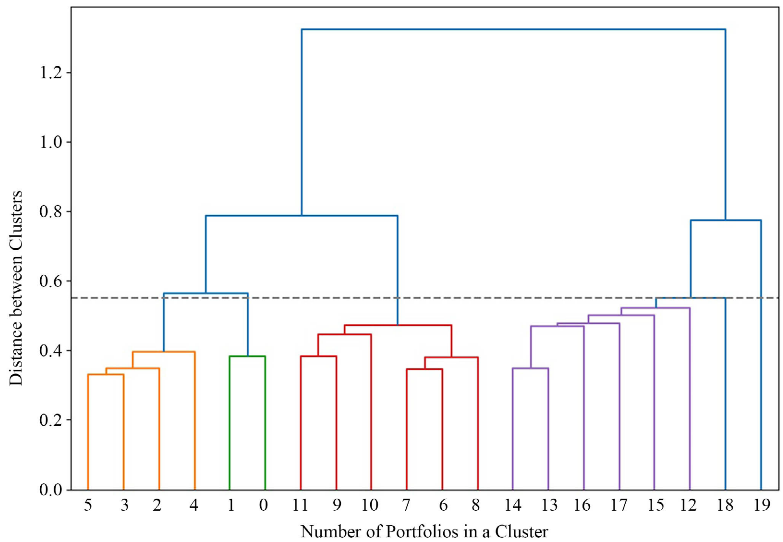
Mean Return panel of Table 1) only have an ONC as large as 4 (derived from the Mean Return panel of Table 2), but SEW and EWW which perform worse have a higher ONC of 6. The mean of N equates to roughly sixteen, meaning that about sixteen out of the twenty size-sorted portfolios are included in one cluster. On average, the maximum weight allocated to size portfolios is close to 60%. Larger maximum weights are assigned to portfolios of higher market value. For each HERC strategy, we use N^{th} to denote which one in the sequence of twenty size-sorted portfolios admits the highest weight W. As can be also seen in Table 2, worse-performing HERC strategies (e.g., VA, ESA, EWA, SDA, and CDRA) tend to allocate very large weights (up to 80%) to larger-size portfolios. Average linkage is used to calculate the present results. After repeating all the above exercises using alternative types of linkage, the single linkage turns out to deliver the best performance.



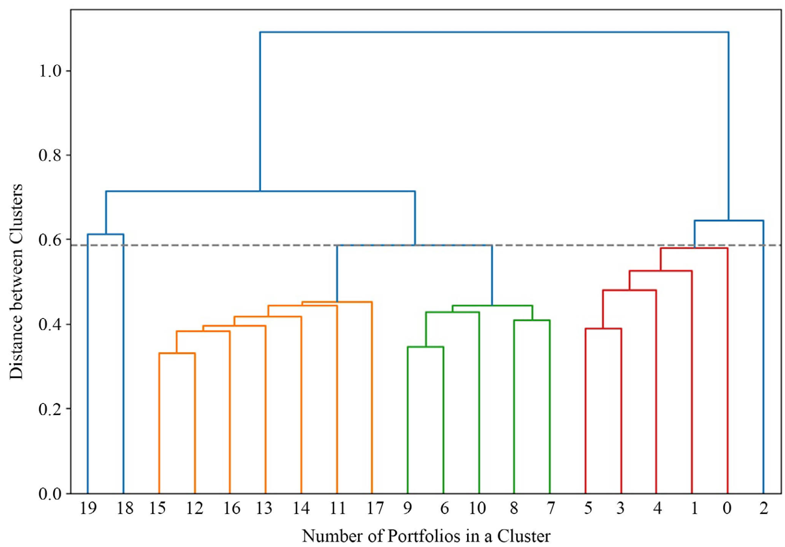
Size



Mean Return



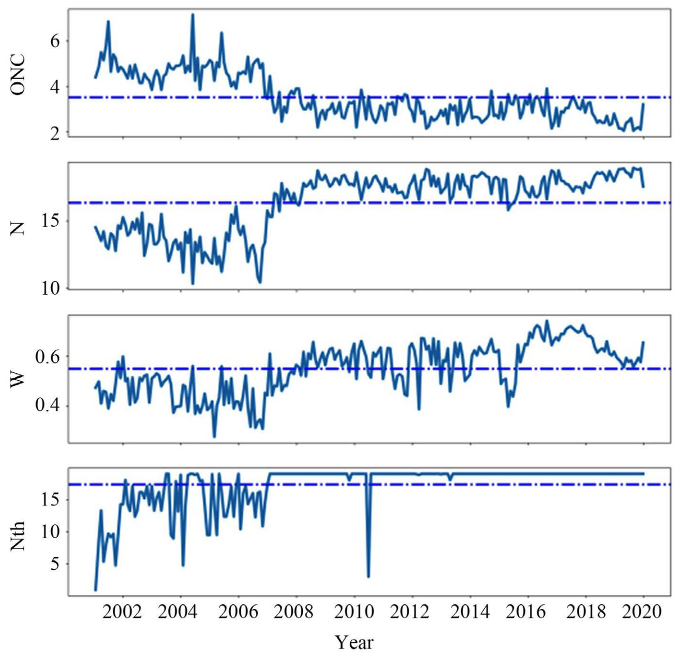
Volatility



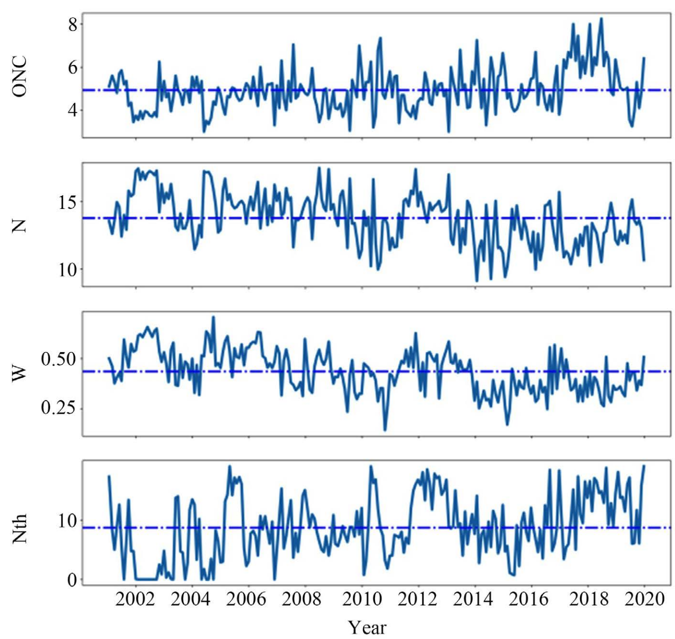
Sharpe Ratio

Figure 1. An illustration of dendrograms based on HERC algorithm. Note: This figure gives an example of HERC-based dendrograms using expected shortfall and ward's linkage calculated with data ranging from 2000-01-05 to 2000-12-29 (*i.e.*, the first year of our sample). The structure of dendrograms changes over time. The optimal number of clusters is set at 5 for different sorting methods for illustration. Graphically speaking, the number of intersections between the horizontal dashed line and all vertical lines should equal the ONC.

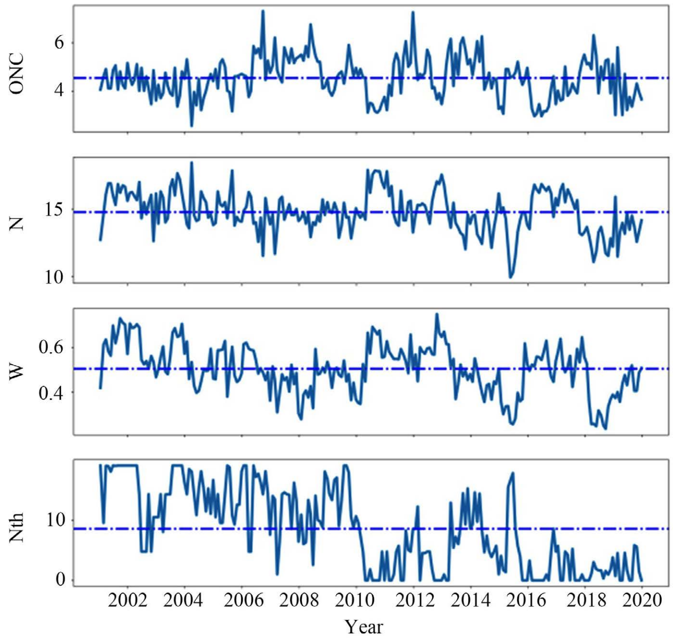
Note that the average ONC for all HERC portfolios across time is slightly below four as shown in the top-left of **Figure 2**. That is, the twenty size-sorted portfolios are grouped into four clusters on average. However, in these clusters there exists a large cluster that generally contains sixteen portfolios. This leads to unbalanced numbers of portfolios across clusters. Further analysis implies that maximum weights (more than 50% and close to 60%) are on average allocated to portfolios with a large size, making weights on size-sorted portfolios even more imbalanced. In conclusion, portfolios with smaller ONC, W, and N^{th} seem to generate higher returns than others with larger ONC, W, and N^{th}. Consider the smallest-size portfolio labeled as 0, it displays an ASR of 0.48, which is in comparison to the largest-size portfolio labeled as 19 which has an ASR of 0.04. For size-sorted portfolios, the ONC decreases over time as shown in Figure 2, but all N, W, and N^{th} indicators increase over time. These trends disappear when we sort portfolios by the mean return, volatility, or Sharpe ratio as we have done below.



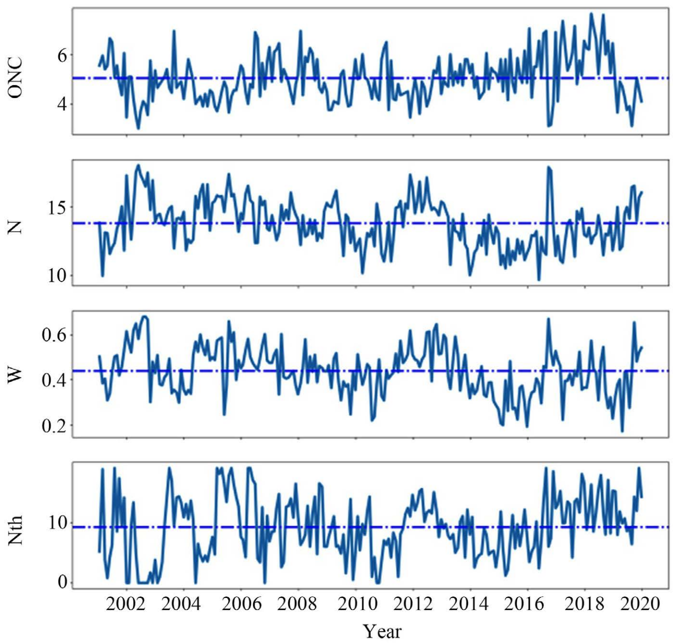
Size



Mean Return



Volatility



Sharpe Ratio

Figure 2. The optimal number of clusters and maximal weights of portfolios. Note: The horizontal dash-and-dot line denotes the mean.

Next, we continue to analyze the Mean Return panel of all exhibits. Results in **Table 1** show that all HERC portfolios could not beat IVP, EW, and SSE. We have to point out that CLA performs worst this time. In contrast to size-sorted portfolios, the average ONC increases from three to five, N decrease from sixteen to fourteen, and W also decreases as shown in **Figure 2**. Besides, the return-sorted portfolio with the maximum weight is not the portfolios with extreme values but not the ones with medium returns of about 10. The top three HERC portfolios allocated about 50% weights to the eighth-high-return portfolio. Such good performance is consistent with the attractive characteristics of the 8^{th} portfolio illustrated by **Table 3**.

In general, HERC portfolios with lower historical returns would outperform in the coming month, which follows the prediction of the mean-reverting theory. Like before, strategies using the average linkage perform not so well and those using either single or complete linkage tend to perform better. Then we look at the Volatility panel of all exhibits. In **Table 1**, there exist HERC portfolios that can outperform the IVP, EW, SSE, and CLA benchmarks. Strategies at the bottom have higher ONC and allocate very high maximum weights (80%) to portfolios of high volatility.

In **Table 3**, the eighteenth-high-volatility portfolio produces the worst performance by ASR, and portfolios with lower volatility, in general, have performed much better. **Table 2** says that HERC clusters are again imbalanced, especially so for top-performing strategies. This is evident that fifteen to sixteen volatility portfolios are included in one cluster. All in all, the ONC, N, W, and N^{th} of portfolios sorted by volatility are similar to those of portfolios sorted by mean returns. The only exception is that the N^{th} seems to decrease over time as can be seen from the Volatility panel of **Figure 2**. The maximum weights are allocated to portfolios with low volatility in the first ten years. Afterward, most weights are shifted to portfolios with high volatility in the second ten years. It merits a note that strategies with average linkage such as EWA, SDA, ESA, and VA stand out with excellent performance in this case. This is opposite to the case of return-sorted portfolios where strategies using the average linkage lie at the bottom of the ranking. Strategies with ward's linkage perform worst here.

Finally, results in the Sharpe Ratio panel of all exhibits are explained as follows. We notice in **Table 1** that many HERC portfolios, especially those established based on the expected shortfall and standard deviation risk metric, outperform the IVP, EW, SSE, and CLA. Strategies employing ward's linkage perform worst in the Sharpe ratio setup of sorting, just like the case for volatility sort. We can tell that N is high, implying imbalanced clusters (see **Table 2**). The bottom-right of **Figure 2** shows no trends for the time-series of the ONC, N, W, and N^{th}. **Table 3** shows that portfolios with a lower past Sharpe ratio tend to perform better in the future. Although volatility-sorted HERC strategies outperform their benchmarks, the best combination of risk metric and linkage varies a lot with portfolio sorting. The clusters resulting from the HERC algorithm turn out to be imbalanced once again. In particular, many portfolios are incorporated into one cluster and very high weights are assigned on one portfolio.

By integrating the results from all four sorting setups, we conclude that the portfolios sorted by different variables constitute a collection of "anomalies". The most noticeable anomaly is the imbalanced structure that emerged from portfolios sorted by size. These "anomalies" are meant for future exploration in the non-traditional classification of Chinese stocks rather than concurrent returns. Based on the above conclusions, we hence add to the current literature in the following three aspects. First, this paper tests the HERC allocation technique in an extensive horse race based on Chinese stock market data, ultimately providing empirical evidence on the superiority of 1/N investing or Ward-linkages over other linkage criteria in HRP-style allocations. Besides digesting a multitude of HERC variants, our results enlighten how such techniques have appeal in the context of the Chinese market. It is learned that there exist new forms in the Chinese market's hierarchical structure that we would otherwise miss, and, more importantly, we can exploit such knowledge in constructing more diversified portfolios by way of running corresponding HERC strategies.

Second, we enrich the benchmarks used in this research field. Whilst 1/N or inverse volatility acts as common choices for benchmarking, the most important market portfolio benchmark is also reported in the paper. Obviously, a market-cap weighted investment is a natural choice to invest in Chinese equities and it would be important to learn about the benefits of departing from such market cap weights in terms of a HERC allocation. In addition to reporting basic strategies performance figures such as annualized return or volatility statistics, this paper also juxtaposes the reported maximum drawdown figures to those of the markets (as measured by the SSE composite index). This comparison suggests the market has a considerably lower maximum drawdown than most of the presented risk-based strategies. In turn, this evidence is suggestive that a simple market portfolio might be a strong contender in market downtrends, warranting further comparisons under extreme scenarios. In a similar vein, we add the two classic strategies of HRP and HCAA to the list of benchmarks as the HERC strategies ultimately are derivatives thereof.

Third, this paper contributes by describing the nature and characteristics of various HERC strategies. As we all know, a common way to rationalize a given equity investment strategy is to look for its salient style factor exposures. For instance, a 1/N equity strategy typically shows a positive size bias, a negative momentum exposure, and usually, some value tilts as well. Conversely, minimum-variance strategies have been documented to mostly exploit the low-volatility effect [22]. Our findings document similar notions for the HERC strategies in Chinese markets. In other words, we show that HERC investing is not just a complex way of running existing equity factor strategies. It reveals a non-industry classification structure.

4. Conclusions

This article evaluates the effectiveness of HERC in the Chinese stock market. By applying this clustering and weighting algorithm with different combinations of risk metrics and linkage criteria to portfolios sorted by market value, return, volatility, and Sharpe ratio and compared to CLA, IVP, EW, and SSE benchmarks, we find the following regularities. First, HERC performance is sensitive to different portfolio sorting. That is, the best HERC strategy is portfolio-dependent and is not robust to alternative sorting methods. Second, HERC clusters are unbalanced regarding the fact that one cluster includes almost all sorted portfolios while others include just a few portfolios and the fact that a single portfolio may receive very high weight. Third, consider sorting portfolios by size, the return of HERC-based portfolios seems to be negatively associated with the optimal number of clusters and the maximal weight of portfolios. Lastly, the size- and volatility-sorted portfolios which have the maximal weight allocation vary over time. All four findings imply that the imbalanced HERC strategies have captured information different from the common Chinese industry structure where every sector contains many stocks.

As a result, our paper can be further improved in the following two directions. On the one hand, we suggest investigating alternative clustering trees of the Chinese stock market. If these new classifications considerably depart from the common sector or industry clusters, one could readily make the case for turning to such data-driven approaches. In addition, it is important to figure out the added benefit of using these techniques at the single-stock level. In other words, are HERC methods superior to simply diversifying across sectors, industries, or style factors? On the other hand, to make the topic of this paper relevant to practitioners, future studies need to report turnover numbers and strategy performance net of transaction costs associated with the turnover. This is particularly relevant given that the allocation step uses naïve risk parity allocation based on different risk measures including standard deviation, variance, expected shortfall, conditional drawdown risk. One would expect these measures to have different stability which will, in turn, translate into portfolio turnover.

Acknowledgment

This work is supported by the General Program of National Natural Science Foundation of China (No. 72071211) and "the Fundamental Research Funds for the Central Universities", Zhongnan University of Economics and Law (2722021BZ042).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] López de Prado, M. (2016) Building Diversified Portfolios that Outperform Out of Sample. *Journal of Portfolio Management*, **42**, 59-69. https://doi.org/10.3905/jpm.2016.42.4.059

[2] Anderson, P.W. (1972) More Is Different. *Science*, **177**, 393-396. https://doi.org/10.1126/science.177.4047.393

[3] Simon, H.A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society*, **106**, 467-482.

[4] López de Prado, M. (2018) Advances in Financial Machine Learning. John Wiley & Sons, New Jersey.

[5] Bailey, D.H. and López de Prado, M. (2013) An Open-source Implementation of the Critical-line Algorithm for Portfolio Optimization. *Algorithms*, **6**, 169-196. https://doi.org/10.3390/a6010169

[6] Black, F., and Litterman, R. (1992) Global Portfolio Optimization. *Financial Analysts Journal*, **48**, 28-43. https://doi.org/10.2469/faj.v48.n5.28

[7] Markowitz, H. and Todd, G.P. (2000) Mean-Variance Analysis in Portfolio Choice and Capital Markets. John Wiley & Sons, New Jersey.

[8] López de Prado, M. (2020) Machine Learning for Asset Managers. Cambridge University Press, London. https://doi.org/10.1017/9781108883658

[9] Jain, P. and Jain, S. (2019) Can Machine Learning-based Portfolios Outperform Traditional Risk-based Portfolios? The Need to Account for Covariance Misspecification. *Risks*, **7**, 74. https://doi.org/10.3390/risks7030074

[10] Lohre, H., Rother, C. and Schäfer, K.A. (2020) Hierarchical Risk Parity: Accounting for Tail Dependencies in Multi-asset Multi-factor Allocations. In: Jurczenko, E., Ed., *Machine Learning and Asset Management: New Developments and Financial Applications*, John Wiley & Sons, New Jersey, 329-368. https://doi.org/10.1002/9781119751182.ch9

[11] Raffinot, T. (2017) Hierarchical Clustering-based Asset Allocation. *The Journal of Portfolio Management*, **44**, 89-99. https://doi.org/10.3905/jpm.2018.44.2.089

[12] DeMiguel, V., Garlappi, L. and Uppal, R. (2009) Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, **22**, 1915-1953. https://doi.org/10.1093/rfs/hhm075

[13] Papenbrock, J. (2011) Asset Clusters and Asset Networks in Financial Risk Management and Portfolio Optimization. Unpublished Doctoral Dissertation, Karlsruher Institut für Technologie (KIT).

[14] Raffinot, T. (2018) The Hierarchical Equal Risk Contribution Portfolio. SSRN Working Paper. https://doi.org/10.2139/ssrn.3237540

[15] Tibshirani, R., Walther, G., & Hastie, T. (2001) Estimating the Number of Clusters in a Data Set via the Gap Statistic. *Journal of the Royal Statistical Society: Series B* (*Statistical Methodology*), **63**, 411-423. https://doi.org/10.1111/1467-9868.00293

[16] López de Prado, M. and Lewis, M.J. (2019) Detection of False Investment Strategies Using Unsupervised Learning Methods. *Quantitative Finance*, **19**, 1555-1565. https://doi.org/10.1080/14697688.2019.1622311

[17] López de Prado, M. (2020) Clustering (Presentation Slides). https://doi.org/10.2139/ssrn.3512998

[18] López de Prado, M. (2020) Codependence (Presentation Slides). https://doi.org/10.2139/ssrn.3512994

[19] Mantegna, R.N. (1999) Hierarchical Structure in Financial Markets. *The European Physical Journal B*, **11**, 193-197. https://doi.org/10.1007/s100510050929

[20] Pézier, J. and White, A. (2008) The Relative Merits of Alternative Investments in Passive Portfolios. *Journal of Alternative Investments*, **10**, 37-39. https://doi.org/10.3905/jai.2008.705531

[21] Levy, M. (2017) Measuring Portfolio Performance: Sharpe, Alpha, or the Geometric Mean? *Journal of Investment Management*, **15**, 1-17. https://joim.com/measuring-portfolio-performance-sharpe-alpha-or-the-geometric-mean/

https://doi.org/10.2139/ssrn.2837484

[22] Leote De Carvalho, R., Lu, X. and Moulin, P. (2012) Demystifying Equity Risk-Based Strategies: A Simple Alpha plus Beta Description. *Journal of Portfolio Management*, **38**, 56-70. https://doi.org/10.3905/jpm.2012.38.3.056