Does the Hierachical Risk Parity Method Really Outperform Out-of-Sample?

Joo Hyung Kim *

August 15, 2020

Contents

*Department of Computational Mathematics, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1;

List of Figures

List of Tables

Abstract

Prado [5] recently proposed a machine learning-based portfolio strategy that does not involve matrix inversions. This new scheme can be applied to problems with singular covariance matrices and terminates in finite number of iterations, as opposed to many traditional risk-based portfolio strategies. This project is inspired by the work of Prado [5] and Jain at el., [4] to a great extent. In this project, we re-examine the strength of the HRP algorithm in two ways. First, we artificially generate the data and verify that the HRP algorithm is more robust against the out-of-sample performance. Second, we extend our experiment using real world data using the framework proposed by Jain at el., [4].

1 Introduction

Portfolio optimization is a task of selecting the best portfolio provided the objective function. Out of all the possible choices for portfolios, it strives to find the best portfolio so that no other portfolio gives a better objective value. Markowitz [2], back in 1952, is one of the first to bring this idea to the community. Since Markowitz's development on the minimum variance portfolio [2], many portfolio strategies have been proposed. Many of these strategies involve arithmetic computations involving matrix inversions and problems are often solved using iterative methods numerically. In case of having a badly conditioned covariance matrix, iterative schemes may cause numerical instability. Prado [5] provides a recipe for avoiding these difficulties. Prado [5] proposes a machine learning-based portfolio scheme that does not involve matrix inversions.

This new scheme can be applied to problems with singular covariance matrices and terminates in a finite number of iterations.

This project is inspired by the work of Prado [5] and Jain at el., [4] to a great extent. Both of the work concern the performance of the machine learning-based portfolio strategy, called the Hierachical Risk Parity (HRP) algorithm. In this project, we re-examine the strength of the HRP algorithm in two ways. First, we perform the Monte-Carlo experiment presented in Prado [5]. We artificially generate the data and verify that the HRP algorithm is more robust against the out-of-sample performance. Second, we extend our experiment using real world data. Jain et.al., [4] present a framework to examine the performance of the HRP algorithm using real world data. We use the framework suggested by Jain et.al., [4] to examine the performance of the HRP algorithm applied the US stock market.

1.1 Overview

This report is organized as follows. Section 2 introduces some preliminary results used in the project. We present some well-known traditional risk-based portfolio methods in Section 2.1. In Section 2.2, we introduce the machine learning based method, so called the HRP algorithm. In Section 2.3, we present three methods for forecasting covariance matrices. In the rest of this report, we present our numerical experiments using methods introduced in Section 2. In Section 3.1, we present the numerical experiments using artificial data. In Section 4.1, we present the numerical experiments using real world data. Finally, we present our conclusions in Section 5.

1.2 Notations

We use $\mathbb{R}^n$ to denote the usual Euclidean space of dimension n. We use $\mathbb{S}^n$ to denote the space of real symmetric matrices of order n. We use $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$, resp.) to denote the code of n-by-n positive semidefinite (definite, resp.) matrices. We let ℓ denote the vector of all 1's with its dimension trivially specified by the context. We use $\circ$ to denote the Hadamard (or element-wise) product of two matrices.

2 Preliminaries

In this section we introduce some preliminary results used in the analysis of this project. Section 2 introduces some preliminary results used in the project. We present some well-known traditional risk-based portfolio methods in Section 2.1 and introduce the machine learning based method, so called HRP algorithm in Section 2.2. In Section 2.3, we present three methods for forecasting covariance matrices.

2.1 Traditional Risk-Based Portfolios

In this section we introduce some well-known methods for traditional risk-based portfolios of N risky assets. In what follows, we use Σ to denote the covariance matrix of the return of the N risky assets. We assume that Σ is positive definite in this section. We impose no short-selling restrictions on our analysis. In other words, we consider the feasible region to be the intersection of the budget constraint and the no short-selling:

$$\mathcal{C} := \{x \in \mathbb{R}^N \ : \ \ell^T x = 1, \ x \geq 0\}. \tag{2.1}$$

Minimum Variance Portfolio (MVP) [2] The minimum variance portfolio (MVP, also known as Markowitz's portfolio) is one of the most well-known methods for traditional risk-based portfolio. The MVP is obtained by solving

$$x_{\mathrm{MVP}} = \operatorname{argmin}\{x^T \Sigma x \ : \ x \in \mathcal{C}\}.$$

Inverse Variance Portfolio (IVP) Let $\hat{\sigma} = \operatorname{diag}(\Sigma)$, i.e., σ holds the variance of each of the return. The inverse variance portfolio x_{IVP} assigns the weights in proportion to the inverse of the assets' volatilities:

$$x_{\mathrm{IVP}} = \begin{bmatrix} \frac{1}{\hat{\sigma}_1} \\ \vdots \\ \frac{1}{\hat{\sigma}_N} \end{bmatrix} \Bigg/ \sum_{i=1}^{N} \frac{1}{\hat{\sigma}_i}.$$

The normalization of the weights is performed to satisfy the budget constraint in (2.1).

Inverse Volatility Weighted Portfolio (IVWP) Let $\sigma = \sqrt{\operatorname{diag}(\Sigma)}$, i.e., σ holds the standard deviation of each of the return. The inverse volatility weighted portfolio (IVWP) and IVP are similar in nature. Instead of using the variance as the measure of risk, IVWP use standard deviation as the measure of risk:

$$x_{\mathrm{IVWP}} = \begin{bmatrix} \frac{1}{\sigma_1} \\ \vdots \\ \frac{1}{\sigma_N} \end{bmatrix} \Bigg/ \sum_{i=1}^{N} \frac{1}{\sigma_i}.$$

Equal Risk Contribution Portfolio (ERC) [1] The equal risk contribution (ERC) portfolio assigns weights such that each asset contributes equally to the overall portfolio volatility. Define $\%\mathrm{RC}_i := \frac{x_i(\Sigma x)_i}{x^T \Sigma x}$, the percentage risk contribution of the i-th asset: Then the ERC portfolio is obtained by solving the following loss-minimization:

$$x_{\mathrm{ERC}} = \operatorname*{argmin}_{x \in \mathcal{C}} \left\{ \sum_{i=1}^{N} \left(\%\mathrm{RC}_i - \frac{1}{N} \right)^2 \right\} = \operatorname*{argmin}_{x \in \mathcal{C}} \left\| \%\mathrm{RC} - \frac{1}{N}\ell \right\|^2.$$

Maximum Diversification Portfolio (MDP) Let $\sigma = \sqrt{\operatorname{diag}(\Sigma)}$, i.e., σ holds the standard deviation of each of the return. The diversification ratio is defined by the ratio of the weighted average of stock volatility and the portfolio volatility:

$$\mathrm{DR}(x) = \frac{\sigma^T x}{\sqrt{x^T \Sigma x}}. \tag{2.2}$$

The maximum diversification portfolio is then obtained by maximizing the diversification ratio and the following equivalent convex minimization problem can be used for its computation:

$$x_{\mathrm{MDP}} = \operatorname*{argmin}_{x \in \mathcal{C}} \{-\mathrm{DR}(x)\}.$$

Market-Capitalization-Weighted Portfolio (MCWP) Define

$$M_i := \text{ market capitalization of the } i\text{-th asset at the time of asset allocation.}$$

Then the market-capitalization-weighted portfolio (MCWP) assigns weights to be the asset market capitalization divided by the sum of all asset market capitalizations:

$$x_{\mathrm{MCWP}} = \frac{1}{\ell^T M} \begin{bmatrix} M_1 \\ \vdots \\ M_n \end{bmatrix}.$$

2.2 Machine Learning-Based Portfolio: Hierachical Risk Parity

Most of the traditional risk-based portfolios introduced in Section 2.1 rely on the covariance matrix. Namely, the covariance matrix plays an essential role in obtaining the optimal portfolio weights for various strategies. Some methods involve a course of solving a series of linear systems involving covariance matrices. In a numerical point of view, problems with ill-conditioned covariance matrices prone to have numerical errors (see [3, section 2.7].). The minimum variance portfolio strategy gets the criticism known as the Markowitz's curse in this regard; the more correlated the investments, the greater the need for diversification, and yet the more likely we will receive unstable solutions.

In order to overcome the numerical difficulties arising in the traditional risk-based portfolio computations, Prado [5] presents the *hierachical risk parity* (HRP) algorithm. The most differentiated feature of HRP algorithm from the risk-based traditional methods is that the HRP algorithm can be used on ill-conditioned matrices or even on singular covariance matrices. The HRP algorithm does not require covariance matrix inversions for its portfolio construction.

In the rest of this section, we study the hierachical risk parity (HRP) algorithm. The HRP algorithm can be summarized in three stages; tree clustering, quasi-diagonalization and recursive bisection.

Stage I: Tree Clustering In this stage[1], we determine the hierarchy among the assets using a recursive cluster formation scheme. Let each asset be a cluster containing only one element. Then, using the correlations between assets, we recursively group two closest clusters into a big cluster until only one cluster is remaining. Suppose that we are given a matrix $X \in \mathbb{R}^{T \times N}$, holding observations with N variables over T periods. Below, we illustrate the detailed implementation for the tree clustering.

1. [step 1]

   We use correlation coefficients to measure the similarity between time series. We are given an N-by-N Pearson's correlation matrix ρ. We construct the distance matrix $D \in \mathbb{R}^{n \times n}$ such that

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

   Here, $D_{i,j} = d(X_i, X_j)$, a function of X_i and X_j.

2. [step 2]

   We compute the Euclidean distance between a pair of columns of D, and construct the matrix $\tilde{D}$:

   $$\tilde{D}_{i,j} = \|D_i - D_j\|_2,$$

   where D_i is the i-th column of D. There are many choices for the distance metric.

[1] For a numerical example, a reader may refer to `https://en.wikipedia.org/wiki/Single-linkage_clustering` for a detail example.

3. [step 3]

   Recursively apply step 3 $(N-1)$ times; We stop when $\tilde{D} \in \mathbb{S}^2$. (Single Linkage - The other paper uses two more; Average Linkage, Ward's Method)

   (a) We form a cluster as follows.

   $$u[1] = (i^*, j^*) = \underset{(i,j), i \neq j}{\operatorname{argmin}} \{\tilde{D}_{i,j}\}.$$

   In other words, we are looking for an matrix index that gives us the smallest value among the off-diagonal elements in $\tilde{D}$.

   (b) We now define the distance vector $\dot{d}_{:,u[1]} \in \mathbb{R}^N$ by comparing the elements between a newly formed cluster $u[1]$ and each of the unclustered elements.

   $$\dot{d}_{i,u[1]} := \min\left\{ \{\tilde{d}_{i,j}\}_{j \in u[1]} \right\}, \quad \forall i \in \mathcal{I}_{\text{row}}.$$

   $\mathcal{I}_{\text{row}}$ denotes the current row indices of $\tilde{D}$.

   (c) Update $\tilde{D}$ by appending $\dot{d}_{:,u[1]}$ to $\tilde{D}$

   $$\tilde{D} \leftarrow \begin{bmatrix} \tilde{D} & \dot{d}_{:,u[1]} \\ \left(\dot{d}_{:,u[1]}\right)^T & 0 \end{bmatrix} \in \mathbb{S}^{N+1}$$

   and dropping the clustered columns and rows $j \in u[1]$.

   $$\tilde{D} \leftarrow \tilde{D}_{([N] \setminus u[1],\ [N] \setminus u[1])} \in \mathbb{S}^{N-1},$$

   where $[N] \setminus u[1]$ denotes the set $\{1, \ldots, N\} \setminus u[1]$. When we append row and columns, we use a new index. Similarly, when we drop rows and columns, we do not re-label the original indices.

Stage II: Quasi-Diagonalization This stage rearranges the row and columns of the covariance matrix based on the first stage so that the largest value lie along the diagonal and we call this procedure quasi-diagonalization. It bring a useful property: similar investments are place together and dissimilar investments are placed far apart.

Stage III: Recursive Bisection At the end of tree-clustering step, we have one cluster containing all assets with subsequent clusters nested with each one. Now, we use a top down approach to break each cluster into two sub-clusters.

First, we initialize weights for all assets as 1. Then, the weights in sub-clusters are distributed using inverse variance allocation. A weighting factor is calculated for each sub-cluster to re-scale the weighs. Each cluster is recursively bisected into two sub-clusters until a single asset is left in each sub-cluster. Algorithm 2.1 illustrates the implementation of this stage.

2.3 Covariance Forecasting

Given returns $r_0, \ldots, r_{T-1}$ of T periods, we wish to forecast the covariance matrix $\hat{\Sigma}_T$ of return r_T. We use the predicted covariance matrices in order to examine the out-of-sample performance of the HRP algorithm in our numerical experiments. In this section we present three methods for forecasting covariance matrices.

Algorithm 2.1 Recursive Bisection

1: **Initialize:** a list of assets in the portfolio with $L = \{L_0\}$ with $L_0 = \{1, \ldots, N\}$, a vector of weights $w_i = 1,\ \forall i \in [N]$
2: **if** $|L_i| = 1, \forall L_i \in L$ **then**
3: stop
4: **end if**
5: **for** each $L_i \in L$ such that $|L_i| > 1$ **do**
6: Bisect L_i into two subsets, $L_i^1 \dot{\cup} L_i^2 = L_i$, where $|L_i^1| = \lfloor \frac{1}{2}|L_i| \rfloor$
7: Calculate the variance of $L_i^j,\ j = 1, 2$, as $\tilde{V}_i^j = \langle \tilde{w}_i^j, V_i^j \tilde{w}_i^j \rangle$, where V_i^j is the covariance matrix of the elements within cluster j, and

$$\tilde{w}_i^j = \mathrm{diag}(V_i^j)^{-1} / \operatorname{tr}\Big(\mathrm{Diag}(V_i^j)^{-1}\Big),$$

which is the inverse volatility weight for the elements of the cluster.
8: Compute the split factor $\alpha_i = 1 - \frac{\tilde{V}_i^1}{\tilde{V}_i^1 + \tilde{V}_i^2}$
9: Rescale allocations w_n by a factor of $\alpha_i,\ \forall n \in L_i^1$
10: Rescale allocations w_n by a factor of $(1 - \alpha_i),\ \forall n \in L_i^2$
11: **end for**
12: loop to line 2

Sample-Based Covariance (SMPL) Assume that the vector of daily asset *log* return is given by $r_t = \mu_t + \epsilon_t$, where ϵ_t is the vector of white noise on the day t such that $\mathbb{E}[\epsilon_t | \text{prior}] = 0$. Given the historical returns $r_0, \ldots, r_{T-1}$ for the past T periods, we forecast the covariance matrix on the T-th period as follows:

$$\hat{\Sigma}_T = \mathbb{E}[\epsilon_T \epsilon_T^T] \approx \frac{1}{T} \sum_{i=0}^{T} \epsilon_i \epsilon_i^T, \tag{2.3}$$

i.e., it is the average of the hostorical white noise.

Exponentially Weighted Moving Average (EWMA) Covariance We note that SMPL generates the covariance matrix by simply averaging the historical covariance matrices. The exponentially weighted moving average (EWMA) is designed to focus more on the recent returns. The EWMA is estimated recursively as follows:

$$\hat{\Sigma}_T = (1 - \lambda)\epsilon_{T-1}\epsilon_{T-1}^T + \lambda \hat{\Sigma}_{T-1}, \tag{2.4}$$

with the multiplier $\lambda \in (0, 1)$. In our test, we use $\lambda = 0.94$[2].

Dynamic Conditional Correlation GARCH (DCC-GARCH) Covariance [13] The Autoregressive Conditional Heteroskedasticity (ARCH) method is developed by Engle [12] in 1982. The model is developed to capture the volatility clustering, namely, the volatile periods tend to be clustered. The model captures the difference between the unconditional and conditional variance allowing the latter to change over time as a function of past errors. The generalized ARCH (GARCH) is proposed by Bollerslev [14], as an extension work of Engle [12]. The GARCH model involves more parameter estimations than ARCH model; in addition to

[2]The decaying rate $\lambda = 0.94$ is recommended by Riskmetrics group.

the parameter estimation associated with the hostorical return in the ARCH model, it requires parameter estimation involving the historical covaraince matrices.

The dynamic conditional correlation GARCH (DCC-GARCH) model [13] engages more parameters than the GARCH model. Not only it requires the parameter estimations associated with the historical return and covariance matrices, but it also requires parameter estimation related to the correlation matrices. We give an algebraic description of this method as follows. Below, the parameters can be estimated via regression or log-likelihood estimations. First, with the estimated parameters

$$w = \mathrm{Diag}(w_1, \ldots, w_N),\ \kappa = \mathrm{Diag}(\kappa_1, \ldots, \kappa_N),\ \lambda = [\lambda_1, \ldots, \lambda_N],$$

we define

$$D_t^2 = w + \kappa \odot \epsilon_{t-1}\epsilon_{t-1}^T + \lambda \odot D_{t-1}^2.$$

We use this to accommodate the discrepancy between the variability of a variable and the range of values of other variables that predict it (conditional heteroskedasticity). Secondly, we model the conditional correlation. With the estimated parameters α, β, we define

$$\begin{aligned} Q_t &= (1-\alpha-\beta)\bar{Q} + \alpha\epsilon_{t-1}\epsilon_{t-1}^T + \beta Q_{t-1}, \\ R_t &= \mathrm{diag}(Q_t)^{-\frac{1}{2}} Q_t \,\mathrm{diag}(Q_t)^{-\frac{1}{2}}, \end{aligned}$$

where $\bar{Q}$ is the unconditional correlation matrix of ϵ. With D_t and R_t computed above, the DCC-GARCH model yields the forecasted covariance matrix as follows:

$$\hat{\Sigma}_t = D_t R_t D_t, \quad \text{for } t = 1, \ldots, T.$$

3 Numerical Experiment: Artificial Data

In this section we illustrate the strength of the HRP algorithm using artificial data. In Section 3.1, we examine how the HRP method assigns the weights compared to the MVP and IVP. In Section 3.2, we examine the strength of the HRP algorithm compared to the MVP, IVP and IVWP via Monte-Carlo simultation.

3.1 A Numerical Example

In this section we verify the characteristic of HRP algorithm via artificial data. Prado [5] proposed that the HRP algorithm output compromises between diversifying across all investments and diversifying across cluster.

Data and Methodology First, we generate two instances of matrix of size 5×100 using the Gaussian distribution with 0 mean, 10% and 5% standard deviation name X and Y respectively. We then choose arbitrary columns of X and add to columns in Y. Then, we form the dataset by appending X and Y horizontally.

Results and Discussions We compute the Weights of HRP, IVP and MVP using the dataset described above and the result is presented in Figure 3.1. The horizontal axis indicates the 10 assets we generated, whereas the vertical axis represents the weights each portfolio allocated. The MVP concentrates investments in few assets which is exposed to idiosyncratic shocks. In other words, negative impacts on a specific group of assets will be critical to this investment allocation. Conversely, the IVP distributes the investments equally to 10 assets, therefore unprotected towards systemic shocks. As we see in Figure 3.1, the HRP method finds a balance between the two methods.

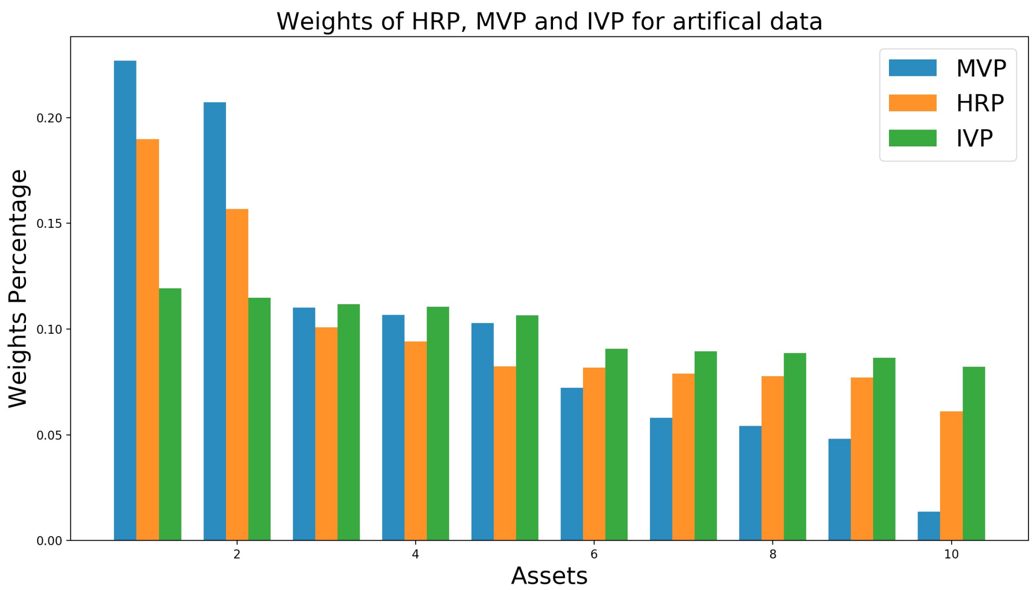

Figure 3.1.1: Weights for MVP, HRP and IVP on artificial data

3.2 Numerical Experiment: Monte-Carlo simulation

Prado [5] proposes that the HRP algorithm outputs portfolios that outperform out-of-sample. In this section, we illustrate the performance of the HRP algorithm via the Monte-Carlo simulation.

3.2.1 Data and Methodology

First, we generate 10 series of random standard Gaussian data, each with 520 observations (equivalent to 2 years of stock data). Second, we add random shocks and correlations to simulate real world stock market. Then, we use first 260 observations (equivalent to 1 year of stock data) to obtain portfolios weights and these portfolios weights are re-computed and re-balanced every 22 observations (monthly). Next, we compute out-of-sample returns associated with the HRP, MVP, IVP and IVWP. This procedure is repeated 10,000 times.

Prado [5] compares the HRP out-of-sample variance with MVP and IVP whereas Jain [4] uses IVWP rather than IVP. Therefore, all 4 methods are selected in the Monte-Carlo simulation in order to accommodate the discussions made in both [5] and [4].

3.2.2 Results and Discussions

Prado [5] concludes that the HRP has the lowest out-of-sample variances compared to the MVP and IVWP. Furthermore, Jain [4] claims that the IVWP has lower out-of-sample variances than the HRP. Using the methodology presented in Section 3.2.1, we obtain Table 3.2.1.

	Standard Deviation	Variance	% of Standard Deviation in terms of HRP
HRP	0.047883	0.002293	100%
MVP	0.116911	0.013668	596%
IVP	0.048526	0.002355	103%
IVWP	0.047648	0.002270	99%

Table 3.2.1: Results from Monte-Carlo Simulation

Although the objective of the MVP is minimizing the portfolio variance, we obtain $\sigma^2_{\text{MVP}} = 0.013668$ which is the highest among all the four methods. The remaining methods, HRP, IVP and IVWP, have variances close to 0.0023. In particular, We obtain

$$\sigma^2_{\text{IVWP}} < \sigma^2_{\text{HRP}} < \sigma^2_{\text{IVP}} < \sigma^2_{\text{MVP}}.$$

Therefore, we reproduced results for Monte-Carlo simulation for out-of-sample successfully.

4 Numerical Experiment: Real World Data

In this section we examine the strength of the HRP algorithm using the real world data by determining (1) which covariance forecasting method is superior and (2) which portfolio technique performs better than others in terms of risk and return.

4.1 Data and Methodology

Overview The methodology used in the experiment can be summarized as follows. As suggested in [4], we consider two data sets; estimation period and evaluation period. We obtain the estimated covariance matrix $\hat{\Sigma}_t$ using T observations made in the estimation period presented in Section 2.3. We then use $\hat{\Sigma}_t$ to generate various portfolios $\hat{w}_t$ presented in Section 2.1 and 2.2. We finally use the realized covariance matrix Σ_t using the data obtained in evaluation period and use appropriate the risk measures with the input $\hat{w}_t$ to get the loss function values.

Dataset There are differences in the data between our approach and Jain's [4]. For data used in Superior Predictive Testing Ability, which is discussed in later part of the report, Jain [4] takes Jan 2017-Dec 2017 intra-day prices. Per day, [4] take 400 points including the night time and then truncate the data from 9AM-3:30PM. They then use interpolation so that we have 200 observations a day.

For our experiment, two set of data is collected for different purposes,

(1) For Superior Predictive Ability (SPA) Test, we've used 390 minute by minute intraday US stock price from 9:30AM-4:30PM EST from 2016 to 2017. The information is organized in Table 4.1.1

(2) To determine which portfolio techniques are better in terms of risk and return, historical prices from 2010 to 2016 are collected to calculate in-sample weights and another set of prices from 2016 to 2017 are used to calculate the out-of-sample performance. The detailed information is provided in Table 4.1.2

Table 4.1.1 list the details of the data used for our experiment.

Purpose	Superior Predictive Ability Test
time period	$1, \ldots, n$
real time	January 1, 2016 - January 1, 2017
number of obs	390 per day, total of 97740
observed values	intraday minute by minute closing prices
usage	$\hat{\Sigma}_t$: covariance forecasting

Table 4.1.1: Data set used in Superior Predictive Ability Test

In our study we use the stock prices from the US stock market. The stocks used for our study is listed in Table 4.1.3.

	Estimation Period	Evaluation Period
time period t	$-T+1,\ldots,0$	$1,\ldots,n$
real time	November 1, 2010 - December 31,2015	January 1, 2016 - January 1, 2017
number of obs	1525 ($T=1525$)	252, 52, 12
observed values	daily adjusted closing prices	daily, weekly, monthly
usage	$\hat{w}_t$ from portfolio strategies	Rebalancing

Table 4.1.2: Information Used to Compute the Weights and Rebalancing

Universe	Ticker Name									
1	AAPL	MSFT	BRK-B	AMZN	GOOGL	WMT	UNH	INTC	PG	ADBE
2	WFC	GS	JPM	MFC	SPGI	AMT	V	MA	TD	BAC
3	XOM	CVX	RDS-A	NEE	PTR	TOT	BP	E	EPD	D
4	NTES	JNJ	TSM	RY	BUD	OESX	PEP	ODFL	CSCO	INO
5	HSBC	NVDA	MMM	MRK	CL	DIS	FISV	RIO	CRM	PFE

Table 4.1.3: Stocks Used in Our Study (in ticker name)

4.2 Superior Predictive Ability (SPA) Test

The Superior Predictive Ability (SPA) test was introduced by White in early 2000, where this test compares the performances of two or more forecasting models. Each forecast model is benchmark against the rest of the models based on the prespecified loss function, in order determine benchmarked model is inferior to the rest of the models. The test quantifies the inferiority check by computing p-value for the null hypothesis, which states the benchmark is not inferior to other models. For instance, if we get a low p-value then we reject the null hypothesis and that model cannot be the benchmark and may be inferior to others.

In our test, SPA test were performed on three covariance forecasting models under different portfolio strategies to identify which covariance forecasting model is superior for each strategy. We give a brief overview on how this is achieved,

Let $L_{0,t}$ be the loss of the benchmark strategy and $L_{k,t}$, for $k=1,\ldots,j$, be the loss of the remaining strategies. Then we define

$$X_t = [X_{1,t},\ldots,X_{j,t}], \text{ where } X_{k,t} := L_{0,t} - L_{k,t}, \text{ for } k=1,\ldots,j.$$

Then the SPA test concerns the null hypothesis $H_0 : \mathbb{E}[X_t] \leq 0$. In other words, this gives us a way to check that the benchmark strategy is *not inferior* to any of the remaining strategies in terms of minimizing the loss value given in Table 4.3.1. SPA test is based on a particular test statistics and additional re-sampling strategy (called stationary bootstrap) is applied. We leave its full discussion to [7, 10, 11].

4.3 Portfolio Risk Measures

To evaluate the out-of-sample performance of different portfolio techniques under different covariance forecasting method, following risk measures were used. Furthermore, these risk measures are associated with particular portfolio strategy due to the construction of the strategy and they are used as the loss function for the superior predictive ability.

We use the same risk measure suggested in [4] and it is presented in Table 4.3.1 below. In the rest of this section, we explain the meaning of the risk measures listed in Table 4.3.1.

Category	Portfolio Strategy	Risk Measure used for SPA
Traditional Risk-Based	MVP	$\sigma_t(\hat{w}_t)$
	IVWP	$H^*(\%\mathrm{RC}(\hat{w}_t))$
	ERC	$H^*(\%\mathrm{RC}(\hat{w}_t))$
	MDP	$-\mathrm{DR}(\hat{w}_t)$
Machine Learning	HRP (SL)	$\sigma_t(\hat{w}_t)$

Table 4.3.1: The portfolio allocation method and corresponding loss function to determine benchmark covariance forecast models

4.3.1 Portfolio Variance

The variance of portfolio is the most common metric to compute the risk of the portfolio, where it implies the scale and frequency of the movement of portfolio value. Daily variance of the portfolio which is computed by:

$$\sigma_t(\hat{w}_t) = \hat{w}_t^T \Sigma_t \hat{w}_t,$$

where $\hat{w}_t$ is the vector of weights obtained for a selected portfolio strategy using the forecasted covariance matrix $\hat{\Sigma}_t$, for $t = 1, ..., n$. The matrix Σ_t is the realized daily covariance matrix computed from intraday minute by minute stock prices. In general, higher volatility indicates there is higher likelihood of the value of the portfolio moving at a larger scale and often.

4.3.2 Herfindahl Index (H^*) of Percentage Risk Contribution

The Herfindahl index is a measure of concentration risk. Concentration risk is the possibility of decreasing in portfolio value when a group of underlying stocks moves in the unfavorable direction. The index ranges from 0 to 1, where 0 implies perfectly diversified portfolio. It is calculated using the following formula:

$$H^*(\%\mathrm{RC}(\hat{w}_t)) = \frac{\sum_{i=1}^{N}(\%\mathrm{RC}_i^2) - \frac{1}{N}}{1 - \frac{1}{N}}.$$

4.3.3 Diversification Ratio

The diversification ratio (DR) is introduced in Section 2.1 (see equation (2.2)). This ratio measures how well the portfolio is diversified, thus higher the ratio, it is better performance measure. For the purpose of superior predictive ability test, $-$DR is used for the loss function.

4.3.4 Superior Predictive Ability Test for Covariance Forecasting Methods

If we refer to Appendix A.1, the machine learning based HRP, DCC-GARCH and EWMA covariance forecasting models can be considered as the benchmark model as in all universes the p-values for DCC-GARCH and EWMA as benchmark are above the significance level of 1%. Amongst the traditional risk based allocation methods, EWMA is clearly considered as the benchmark model for MDP, because when EWMA is chosen as the benchmark model, this is only forecasting model with the null hypothesis cannot be rejected. In case of ERC, IVWP and MVP, all the covariance forecasting models have p-values above the significance level of 1% in all the universes. Overall, EWMA is the superior method for all portfolio strategies as it is only one that did not get rejected under five universes.

4.4 Performance Attribution Analysis

In this section, we explore which portfolio strategy is the best in terms of risk and return with respect to different covariance forescasting methods. For simplicity, we have combined the five universes presented in Table 4.1.3 into one universe and create one portfolio of 50 stocks. The out-of-sample performances include the daily, weekly and monthly rebalancing in order to confirm that the performance of individual portfolio allocation technique is consistent.

4.4.1 Portfolio Return

In this section, we make discussions on the portfolio return with Table A.2.1 presented in Appendix A.2. Table A.2.1 shows that, for the daily rebalancing with SMPL, the MVP has the highest annual return among other portfolio techniques. When the EWMA method is used for the covaraine forecasting method, the MDP has the highest return. When the DCC-GARCH is used for the covaraine forecasting method, the HRP has the highest return. We observe the same pattern in the weekly rebalancing and monthly rebalancing.

In addition, all the portfolio strategies have outperformed S&P 500 index. Due to survivorship bias, it is no surprise that all the portfolio strategies perform better than the market. However, it is very difficult to be included as one of the 500 underlying companies in S&P 500 index, as the company must have market capitalization of at least USD 8.2 billion and its most recent quarter's earnings and the sum of its trailing four consecutive quarters' earnings must be positive. Which means survivorship bias is also embedded in S&P 500 index. Given this fact, we are comparing apple to apple and so we cannot discredit the performance of portfolio strategies.

4.4.2 Portfolio Variance

In this section, we make discussions on the portfolio variance with Table A.3.1 in presented in Appendix A.3. Table A.3.1 shows that, for the daily and weekly rebalancing, the variance of HRP is lower than the variance of both MVP and IVP. This confirms the claim from Prado [5] that the HRP has the lowest out-of-sample variances when comparing with MVP and IVWP. However, in monthly rebalancing, the variance of HRP is higher than the variance of IVWP, thus it solidifies Jain [4]'s discovery that the IVWP has a lower out-of-sample variance compared to the HRP.

4.4.3 Portfolio Sharpe Ratio

In this section, Sharpe ratio of the portfolio strategies is used to identify which technique has the best risk-adjusted return with respect to three covariance forecasting method. The Sharpe ratio describes how much excess return is earned when holding riskier asset or portfolio in this case. In other word, we look for a portfolio with high return and low risk. For this experiment, we choose T-bill of 0.61% (as of 2016) to be the risk-free rate. We now refer to Table A.4.1 presented in Appendix A.4. In case of the daily, weekly and monthly rebalancing, the HRP has the highest Sharpe ratio under DCC-GARCH.

Conventionally, DCC-GARCH is considered to be one of the most powerful methods for forecasting covariance matrices of portfolios as it can recognize infrequent, extreme volatile movements in the stock prices. Hence, when DCC-GARCH is used, the volatile movements do not influence the forecasting of the covariance and true characteristic of the stocks is captured in the forecast. With this in mind, any result coming from DCC-GARCH is most accurate forecast compared to other covariance forecasting methods discussed in Section 2.3. It gives us more confidence that the HRP is a superior portfolio strategy among aforementioned strategies.

5 Conclusions

In this project, we re-examined the strength of the machine learning-based algorithm, HRP algorithm presented in [5], by engaging the comparative study with the traditional risk-based portfolio strategies. In the course of our experiment, we extensively used the methodologies presented in Jain [4].

There are questions that remain in regards to the reliability of forecasting the covariance matrices in the financial industry. However, forecasting covariance or correlation matrices is well-recognized and often used in quantitative finance. For instance, when pricing the derivative such as equity swap and basket option, covariance forecasting methods are alternative methods for computing the implied volatility instead of the Black-Scholes equation. In addition, there are no documents of investment managers using covariance forecasting to construct portfolios. However, it is likely that companies, such as Two Sigma and Renaissance Technology, have delved into covariance forecasting for more efficient trading methods or portfolio constructions. The demand from the financial industry for the covariance forecasting method is increasing, hence, additional research is required.

It is shown, in Section 4.3.4, that the EWMA is superior to other covariance forecasting methods, as it is the only one that did not get rejected under different portfolio strategies. We point out that this result may differ when we choose different significance level and/or with different data sets. Additionally, we have shown similar results presented in Jain [4]. The DCC-GARCH and EWMA can be considered as the benchmark models for the covariance forecasting methods, when the machine learning-based method (HRP) is used.

From Table A.3.1, we verified the results from Prado [5]; It is shown that the HRP has lower out-of-sample variances than the MVP when both artificial and real-world data are considered. However, there is a caveat to our experiments. There are assumptions that do not reflect real world setting. For instance, we ignored transaction costs and currency rates as these factors can sway our results. In particular, the daily rebalancing would have been an irrational endeavour to make if transaction costs were included in our experiments.

References

[1] Y. Choueifaty, Y. Coignard, Toward Maximum Diversification, Journal of Portfolio Management, 2008, vol. 35, no.1, 40-51

[2] H. Markowitz, Portfolio selection, The Journal of Finance, 1952, vol. 7, no. 1, 77-91

[3] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, 1996, 3E

[4] P. Jain, S.Jain, Can Machine Learning-Based Portfolios Outperform Traditional Risk-Based Portfolios? The Need to Account for Covariance Misspecification journal = Risks (Basel), 2019, MDPI AG, vol. 7, no. 3, 74-

[5] M.L. Prado, Building diversified portfolios that outperform out-of-sample, Journal of portfolio management, Pageant Media US, 2016, vol. 42, no. 4, 59-69

[6] R. Engle, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, 2002, Journal of Business & Economic Statistics, vol 20, no.3 ,339-350

[7] P.R. Hansen, A Test for Superior Predictive Ability, Journal of Business & Economic Statistics, 2005, vol. 23, no. 4, 365-380

[8] M. Martens, ,Measuring and forecasting s&p 500 index-futures volatility using high frequency data, Journal of Futures Markets: Futures, Options, and Other Derivative Products, vol. 22, no. 6, 497-518

[9] S.J. Koopman, E. Hol Uspensky, The stochastic volatility in mean model: empirical evidence from international stock markets, Journal of applied Econometrics, vol. 17, no. 6, 667-689

[10] S. Gonçalves, R. de Jong, Consistency of the stationary bootstrap under weak moment conditions, Economics Letters, 2003, vol. 81, no. 2, 273-278

[11] P.R. Hansen, A. Lunde, A forecast comparison of volatility models: does anything beat a GARCH(1,1)?, Journal of applied econometrics (Chichester, England), 2005, vol. 20, no. 7, 873-889

[12] R.F. Engle, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica, 1982, vol. 50, no. 4, 987–1007

[13] R.F. Engle, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, Journal of Business & Economic Statistics, , 2002, vol. 20, no. 3, 339-350

[14] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of econometrics, 1986, Vol. 31, 307–327

A Numerical Results using Real World Data

A.1 SPA test for Covariance Forecasting

	MVP				
	Universe 1	Universe 2	Universe 3	Universe 4	Universe 5
SMPL	0.352	0.352	0.176	0.26	0.176
EWMA	0.3	0.176	0.176	0.176	0.176
DCC-GARCH	0.032	0.172	0.26	0.26	0.26
	IVWP				
SMPL	0.052	0.176	0.052	0.052	0.136
EWMA	0.3	0.176	0.176	0.176	0.176
DCC-GARCH	0.032	0.172	0.26	0.26	0.26
	ERC				
SMPL	0.288	0.468	0.26	0.292	0.136
EWMA	0.052	0.272	0.26	0.824	0.136
DCC-GARCH	0.18	0.176	0.468	0.564	0.052
	MDP				
SMPL	0	0.044	0.076	0.044	0.032
EWMA	0.032	0.228	0.272	0.076	0.076
DCC-GARCH	0.044	0.336	0	0	0.176
	HRP				
SMPL	0	0.176	0.176	0.044	0.308
EWMA	0.044	0.176	0.176	0.352	0.42
DCC-GARCH	0.076	0.052	0.388	0.176	0.208

Table A.1.1: p-values from the SPA test when DCC-GARCH, SMPL and EWMA are the benchmarked model for the different portfolio strategies

A.2 Portfolio Return

	Daily Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	20.79%	20.96%	24.60%
MCWP	20.55%	20.55%	20.55%
MDP	22.39%	30.82%	17.62%
ERC	21.46%	22.25%	23.50%
IVWP	21.54%	21.48%	23.03%
MVP	23.50%	23.50%	23.50%
S&P 500	13.59%	13.59%	13.59%
	Weekly Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	25.01%	25.88%	30.64%
MCWP	27.12%	27.12%	27.12%
DR	27.30%	35.42%	23.71%
ERC	26.71%	27.68%	29.68%
IVWP	26.79%	26.86%	28.97%
MVP	29.68%	29.68%	29.68%
S&P 500	18.10%	18.10%	18.10%
	Monthly Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	21.75%	23.11%	28.21%
MCWP	25.05%	25.05%	25.05%
MDP	23.94%	31.32%	22.77%
ERC	23.95%	24.94%	27.53%
IVWP	24.03%	24.51%	26.66%
MVP	27.53%	27.53%	27.53%
S&P 500	17.18%	17.18%	17.18%

Table A.2.1: Annual Return of different portfolio strategies with different covariance forecasting method for different rebalancing horizons

A.3 Portfolio Variance

	Daily Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	12.70%	13.57%	14.56%
MCWP	14.97%	14.97%	14.97%
MDP	15.61%	15.27%	16.31%
ERC	14.03%	14.26%	15.36%
IVWP	14.07%	14.30%	14.94%
MVP	15.36%	15.36%	15.36%
	Weekly Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	25.09%	27.09%	31.46%
MCWP	28.73%	28.73%	28.73%
MDP	32.93%	33.62%	31.97%
ERC	29.17%	29.58%	32.93%
IVWP	29.11%	29.51%	31.84%
MVP	32.93%	32.93%	32.93%
	Monthly Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	32.34%	35.29%	35.55%
MCWP	45.81%	45.81%	45.81%
MDP	52.18%	59.46%	34.62%
ERC	33.39%	34.22%	35.79%
IVWP	31.53%	32.75%	34.71%
MVP	35.79%	35.79%	35.79%

Table A.3.1: Annualized variance of different portfolio strategies with different covariance forecasting method for different rebalancing horizons

A.4 Portfolio Sharpe Ratio

	Daily Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	1.58827879	1.49879225	1.64824651
MCWP	1.3312939	1.3312939	1.3312939
MDP	1.39562158	1.97881471	1.04272721
ERC	1.48527874	1.51772791	1.48995345
IVWP	1.48782369	1.4594841	1.50108799
MVP	1.48995345	1.48995345	1.48995345
	Weekly Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	0.97249739	0.93258469	0.95449452
MCWP	0.92280636	0.92280636	0.92280636
MDP	0.81065989	1.0352081	0.72257574
ERC	0.89491864	0.915173	0.88282863
IVWP	0.89941251	0.88968865	0.89088829
MVP	0.88282863	0.88282863	0.88282863
	Monthly Rebalancing		
	SMPL	EWMA	DCC-GARCH
HRP	0.65355847	0.63737919	0.77647095
MCWP	0.53349852	0.53349852	0.53349852
MDP	0.44703398	0.51639158	0.64014217
ERC	0.69905903	0.71105321	0.75233077
IVWP	0.74298315	0.72995643	0.75044818
MVP	0.75233077	0.75233077	0.75233077

Table A.4.1: Sharpe ratio of different portfolio strategies with different covariance forecasting method for different rebalancing horizons