HỘI SINH VIÊN VIỆT NAM

HỌC VIỆN TÀI CHÍNH

ĐỀ TÀI NGHIÊN CỨU KHOA HỌC
THAM GIA HỘI THI KHOA HỌC SINH VIÊN TOÀN QUỐC
"OLYMPIC KINH TẾ LƯỢNG VÀ ỨNG DỤNG" LẦN THỨ VI, 2021

ĐỀ TÀI:

MAXIMAL PREDICTABILITY PORTFOLIO OPTIMIZATION IN US AND VIETNAM STOCK MARKET

Người hướng dẫn: TS. TẠ QUỐC BẢO

Tập thể sinh viên thực hiện:

1. Huỳnh Tấn Vũ (Trưởng nhóm): Khoa MA K18, ĐH Quốc Tế - ĐHQG TPHCM
2. Hồ Hữu Bình: Khoa MA K18, ĐH Quốc Tế - ĐHQG TPHCM
3. Nguyễn Ngọc Phụng: Khoa MA K18, ĐH Quốc Tế - ĐHQG TPHCM
4. Nguyễn Hưng Quang Khải: Khoa MA K18, ĐH Quốc Tế - ĐHQG TPHCM
5. Trần Hoàng Phi: Khoa MA K17, ĐH Quốc Tế - ĐHQG TPHCM

Hà Nội - tháng 05 năm 2021

ABSTRACT

The theory of portfolio allocation has tremendously experienced a breakthrough stage since Harry Markowitz introduced his paper named "Portfolio Selection" with one of the major parts is the Mean-Variance analysis in 1952. The Mean-Variance approach has been among criticism by its overstated portfolio's weight and the ability to capture the estimation errors. This research is an attempt to apply the Maximal Predictability Portfolio, which is proposed by Lo and MacKinlay, to create an investment portfolio that can avoid the forecasting errors which is a serious impediment to the Mean-Variance framework. After going under a close scrutiny, the result has been yielded and thus proving Maximal Predictability Portfolio Optimization's superiority to its counterpart, Mean-Variance Portfolio Optimization. It shows that the new method performed satisfactorily in both US and Vietnam market, thereby sufficiently showing its potential for further development in portfolio optimization progress.

ACKNOWLEDGEMENTS

This research was conducted at math department from October 2020 to March 2021 under the supervision of Dr.Ta Quoc Bao. First and foremost, we would like to deeply appreciate for the wholehearted and unstinting support from him. Your guidance and advice, together with the dynamism, motivation, enthusiasm, assistance at every stage of the research project have tremendously inspired us to a new horizon of mathematics and to successfully accomplish the research. Moreover, we are highly indebted to department of mathematics of International University for delivering many useful advices and imparting their knowledge in this research. We also give a sense of gratitude to our faculty for their facilities and sense of grateful for all the lecturers' from mathematics department of International University who have offered many insightful comments and suggestions for this study. The generosity and expertise of one and all have improved this research field in innumerable ways and deter us from many avoidable errors.

Contents

I. INTRODUCTION

Portfolio management presents the trade-off between return and risk of the investors, which is usually the case that we need to construct a portfolio that maximizes expected returns at a specific level of risk [1]. However, the assumption of Gaussian distribution for the assets' return is not fully sufficient in Mean-Variance framework. Moreover, the major reason for the reluctance of academics and practitioners in using the framework of the Markowitz model is that the freamework is extremely sensitive to the change of the input (the expected return of each asset, the variance of each asset and the covariance between two assets), which leads to the result of implausible weights in the portfolio. The absence of transaction costs in accompany with the market impact cost in the standard quadratic programming (the following optimization problem in Mean-Variance analysis) would not be suitable in practice.

Approximate two decades before, Lo and MacKinlay have devised a vigorous method in analyzing portfolio (1997) [2]. Maximal Predictability Portfolio (MPP, for short) seems to be a propitious point in giving insight into the risk-return trade-off among investors issues.

To obtain such a portfolio, we have to solve a nonconvex fractional programming problem in which the objective function is a ratio of two convex quadratic functions.

Despite its attractiveness to be applied in financial industry, MPP at that time failed to overcome two primary hurdles. The first thing to consider is that constructing an MPP needs asset returns by using factor models such as Frama-French's or Cahart's model. Secondly, the problem in accordance with MPP optimization problem is non-convex fractional quadratic programming problem, for which there exists no efficient method to solve such thing. Fortunately, advancement in global optimization and 0-1 integer programming approach proposed by Konno [9].

The scope of this research is limited to showing MPP outperforming the conventional MV approach in Viet Nam and US stock market. In the next section, we will briefly explain the required structure to create MPP portfolio and review the solution to the non-convex fractional problem. Section 3 and 4 will be devoted to showing the computational results, using Viet Nam and US market data from 2015 to 2019, and giving further suggestions to the problem. In section 5, final conclusion would be draw from the above results.

II. METHODOLOGY

2.1 Return of asset and portfolio

Firstly, the time horizon of later calculation will be introduced. Supposed that the interval $[0, T]$ is our historical time series and N is the number of equal sub-intervals between 0 and T. Denote P_{i,t_k}, where t_k is often measured in day (or any frequency like weekly, monthly, yearly,...) and index i represents the i^{th} asset in the portfolio, then the return of asset is calculated as below:

$$R_i(t_k) = \frac{P_{i,t_k} - P_{i,t_{k-1}}}{P_{i,t_k}}, \tag{2.1}$$

If there is a cash flow in the partition k of time t then return of asset will be adjusted below:

$$R_i(t_k) = \frac{P_{i,t_k} - P_{i,t_{k-1}} + C_{i,t_k}}{P_{i,t_k}},$$

where C_{i,t_k} is the cash flow at time t_k such as dividend, ...

In this report, the cumulative return of the asset i^{th} would be used to be the performance metrics. The cumulative return of asset i^{th} in N-periods between $[0, T]$ will be:

$$R_i = \prod_{k=1}^{N}(1 + R_{i,t_k}) - 1, \tag{2.2}$$

Supposed n weights of portfolio be denoted by the set $W = \{w_i \mid i = 1, 2, ..., n\}$ where n is the number of assets in that portfolio. Then, the portfolio return at period k, denoted by $R_{p,k}$:

$$R_{p,k} = \sum_{i=1}^{n} w_i R_{i,t_k}, \tag{2.3}$$

And the cumulative return of portfolio over $[0, T]$ will be:

$$R_p = \sum_{i=1}^{n} w_i R_i, \tag{2.4}$$

2.2 Beta and Sharpe ratio

2.2.1 Sharpe ratio

Modern Portfolio Theory postulates that rational investors can maximize their expected returns using diversified portfolios. A pioneerer in MPT theory is Harry Markowitz, who used the framework of choosing pairs of assets that raise the optimum trade-off between variance and expected return (for the cases of more than two assets, it will be the set of assets). We will discuss the classical technique of Harry Markowitz in section 2.3.

William F.Sharpe developed a model called *Capital Asset Pricing Model* or CAPM in short, which is followed by 7 assumptions [3]. Here, the first 5 assumptions deals with the decision making of investors. The other two assumptions about the properties of the market will be introduced below for the sake of constituting the model:

- There exists risk-free asset and there is no limitation in borrowing or lending at a same risk-free rate.
- The market is frictionless, meaning there appears no transaction cost and the market itself has to be competitive.

Recalling the Capital Market Line (CML), the model measures the linear relationship of expected return of portfolio and the combination of market portfolio (represent the choice of risky assets) and risk-free asset:

$$\mathbb{E}[\,R_p] \;= R_f + \sigma_p \frac{\mathbb{E}[\,R_M\,] - R_f}{\sigma_M},$$

Sharpe ratio:

$$\frac{\mathbb{E}[\,R_p\,] - R_f}{\sigma_p} \tag{2.5}$$

The ratio in the above is *Efficiency of investment* that measures the risk-adjusted return (the return an investor may receive for taking an amount of risk). R_f is the risk-free rate and R_p is the portfolio return and σ_p is the given risk of portfolio.

2.2.2 Beta

Total risk may be interpreted as a mixture of systematic risk and non-systematic risk. While non-systematic risk could be reduced by diversification, systematic risk are unavoidable. Hence, Sharpe defines the risk that investors accept to commit after choosing randomly a well-diversified portfolio. This risk metric is *Beta*. Beta indicates the strength that one asset covaries with the movement of market as a whole. The value is ranged from 0 to 1, and we could address it as a correlation.

Thus, if we denote the market benchmark returns in Viet Nam, VN index (VNI) and in the US (S&P500) R_{bm}, then:

$$\beta_p = \frac{Cov(R_p, R_{bm})}{Var(R_{bm})},$$

2.3 The Fama-French three-factor model

CAPM proposed by Sharpe, Lintner, Mossin and Treynor is tremendously used by managers due to its simplicity. Albeit, many assumptions that are unfavourable in the real world, especially the market portfolio requirement, which contains all assets in the world, are impractical.

In 1992, Kenneth French and Eugene Fama introduced a paper where they proposed a two new fundamental factors, combined with the first factor in the CAPM model (beta value) after investigating cross-sectional expected return in the market [4]. The equation of Fama-French 3 factors model (FF3M in short) in the case of market portfolio is given by:

$$E[R_p] = \alpha + R_f + \beta_1(E[R_M - R_f]) + \beta_2 E[R_{SMB}] + \beta_3 E[R_{HML}], \tag{2.6}$$

SMB factor illustrates the observation that concerns with size of firm, which has a negative affect on excess return. It is found that firms with small market capitalization outperform those with big cap. Computing this factor will require the difference between the former and the latter.
HML factor describes that high book-to-market ratio firm (usually it is referred to Value stock) and Growth stock is low book-to-market ratio. It is motivated that Value stock, i.e. a moderate stock price connected with the book value will make the earnings on assets low and conversely, Growth stock such as a high stock price compared to its book value will bring about high profitability.

2.4 The Markowitz Mean-Variance model

The framework of Mean-Variance has become the most influential and vital backbone for the quantitative management of portfolio. In this section, the scheme of MV model will shortly be provided, which is remarkably adopted by both practitioners and academics over the time and introduce the formula in an technical and straightforward way.
It is well-argued that the ingredients for Mean-Variance optimization are the estimate of expected return (showing the expected performance of the investment) and the portfolio variance (served as a measurement risk) together with many constraints. Markowitz stated that an averse investor would choose a portfolio that raise the highest

return at a favorable risk. Hence, solving the optimization problem will determine the efficient frontier which is the set of all possible portfolios and subsequently the optimal portfolio weights.

Suppose we have a set of n risky assets, a n-vector $\omega = (\omega_1, \omega_2, \omega_3, ..., \omega_n)$ where each w_i represents the percentage of capital invested in i^{th} asset in the portfolio. For simplicity and practical issue, no short selling is introduced in this model because some kinds of assets are difficult to sell short. Additionally, in each allocation, we restrict the amount by imposing a number α_i.

Assume that the risky assets' returns $\mathbf{R} = (R_1, R_2, R_3, ..., R_n)$ and its following expected returns $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_n)$ and let $\sigma_{ij} = Cov(R_i, R_j)$ be the covariance between the i^{th} and j^{th} return. Moreover, $\sigma_i^2 = Var(R_i)$ and $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, where ρ_{ij} is the correlation between asset $i^{th}$and asset j^{th} and $\sigma_{ij} = \sigma_{ji}$. Two more assumptions are returns of all the assets in portfolio follow the normal distribution (we will review this assumption in the MAD model section later), which express as $R_i \sim N(\mu_i, \sigma_i)$. The other one is based on the behavior of risk-averse investors, hence, it is necessary to add in a requirement for the minimal return they will receive and denoted by υ. Thus, the formulation of minimizing portfolio's risk subject to a targeted expected portfolio's return is shown as following:

$$
\begin{aligned}
\min_{\omega} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij}\omega_i\omega_j \\
\text{s.t.} \quad & \sum_{i=1}^{n} \omega_i = 1 \\
& \sum_{i=1}^{n} \omega_i\mu_i \geq \upsilon \\
& 0 \leq \omega_i \leq \alpha, \qquad i = 1, 2, ..., n
\end{aligned}
\tag{2.7}
$$

Writing in algebraic in the first hand could be simple and easier to treat. However, if the number of assets are above 4,5, presenting the problem using matrix notation will be more handy and conspicuous to calculate. The compact and elegant way to formulate the constrained optimization problem is illustrated [5]:

$$\begin{aligned} \min_{\omega} \quad & \omega' \mathbf{\Sigma} \omega \\ \text{s.t.} \quad & \omega' \mu \geq \upsilon \\ & \omega' \overrightarrow{1}, \qquad \overrightarrow{1} = [1, 1, ..., 1] \end{aligned} \tag{2.8}$$

Σ is the covariance matrix, in which σ_{ij} is the ij^{th} element and the main diagonal is $\sigma_i^2 = Var(R_i)$. Since the problem generally does not have an analytical (closed-form) solution, especially when our constraints are all inequality. The solution to Markowitz's optimization problem nowadays is addressed by numerical method.

2.5 Mean absolute deviation model

The quadratic programming problem corresponding to Mean-Variance analysis becomes such a huge obstacle when it comes to large scale of assets. Hence, many researchers have attempted to devise a linearized version of Markowitz's model (Sharpe, 1971 and Konno & Yamazaki, 1991), thereby reducing significantly the computational time of searching the optimal portfolio by the mean of proposing an alternative risk measure to the conventional one (the standard deviation) that is absolute deviation and subsequently develop a model with this risk metric-MAD model [6]. We will start this section by introducing the definition of the MAD model, presenting one properties that connects the result of optimal solution of this model to Markowitz's one and ,last but not least, the formulating process of portfolio's weight allocation problems using the so-called absolute deviation risk measure. Supposed that there will be n assets $S_i, i = 1, 2, ..., n$ and R_i be the corresponding returns of asset S_i. Then, recalling that the return of the portfolio with a set of n weights $\mathbf{x} = (x_1, x_2, x_3, ..., x_n)$ is computed as followed:

$$R_p(x) = \sum_{i=1}^{n} R_i x_i, \tag{2.9}$$

In conventional analysis of portfolio management, practitioners often use standard deviation as a measure of risk:

$$V(x) = E[(R_p(x) - E[R_p(x)])^2], \tag{2.10}$$

$$\Leftrightarrow \sigma(x) = \sqrt{V(x)}, \tag{2.11}$$

Konno and Yamazaki replaced the L_2 risk by L_1 risk by reasoning that these two risk metrics both measure the dispersion of sample around their mean and ,hence are not significantly different:

$$W(x) = E[|R_p(x) - E[R_p(x)]|], \tag{2.12}$$

If the random variable R is normally distributed, then MAD is proportional to standard deviation, as the below theorem presents.

Theorem 2.5.1 *If $R_p(x)$ follows a normal distribution with mean μ and standard deviation σ, then*

$$W(x) = \sqrt{\tfrac{2}{\pi}}\sigma(x),$$

Proof.

Using the definition of expectation of a random variable and let $R_p(x) - E[R_p(x)]$ be $r - \mu$. Since $R_p(x) \sim N(\mu, \sigma)$,

$$\begin{aligned} W(x) = E[|R_p(x) - E[R_p(x)]|] &= \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{\infty} |r-\mu| exp\{-\frac{(r-\mu)^2}{2\sigma(x)^2}\} dr \\ &= \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{\infty} (r-\mu) exp\{-\frac{(r-\mu)^2}{2\sigma(x)^2}\} dr \\ &= \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{\infty} texp\{-\frac{t^2}{2\sigma(x)^2}\} dt \\ &= \sqrt{\frac{2}{\pi}}\sigma(x) \end{aligned} \tag{2.13}$$

The above theorem also holds when R is multivariate normal random variable [7]. Moreover, this results in a crucial theorem which links MAD model with Mean-Variance model:

Theorem 2.5.2.

The MAD model generates the alike optimal portfolio to the corresponding MV model when $R = (R_1, R_2, ..., R_n)$ *follows a multivariate normal distribution.*

From now on, the establishment of the problem in terms of linear programming will be shown. The formulation plays a major role in constituting the problem in MPP model that will be demonstrated in the next section).

Assume that the vector of n returns from n assets $\mathbf{R} = (R_1, R_2, ..., R_n)$ and is distributed over a finite set of points $(r_{1t}, r_{2t}, ..., r_{nt}), t \in 1, 2, ..., T$. Additionally, $p_t = P[(R_1, R_2, ..., R_n) = (r_{1t}, r_{2t}, ..., r_{nt})]$ is assumed to be available through historical data. As denoted above, x_i is the percentage that asset i^{th} contributes to portfolio. The absolute deviation *W(x)* is defined as follows:

$$\begin{aligned} W(x) =& E[|R_p(x) - E[R_p(x)]|] \\ &= E[|\sum_{i=1}^{n} R_i x_i - E[\sum_{i=1}^{n} R_i x_i]|] \\ &= \sum_{t=1}^{T} p_t |\sum_{i=1}^{n} (r_{it} - r_i) x_i|, \end{aligned} \tag{2.14}$$

where expected value of R_i is $r_i = \sum_{t=1}^{T} p_t r_{it}$.

The Mean-Absolute-Deviation model is defined as follows:

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} r_i x_i \\ \text{s.t.} \quad & \mathbf{W(x)} \leq w, \\ & \mathbf{x} \in X, \end{aligned} \tag{2.15}$$

where w is an acceptable level of risk and $X \subset R_n$ is an investable set defined by:

$$X = \{\mathbf{x} \in R^n \mid \sum_{i=1}^{n} x_i = 1, 0 \leq x_i \leq \alpha, i = 1, 2, ..., n\}$$

where α is the upperbound for weight of asset j^{th}. Solving the above problem is absolutely hardship since the constraint is complicated to treat. However, thanks to the result of standard linear programming, the following conditions to decipher the issue with absolute value are presented:

Let introduce a set of nonnegative variables ϕ_t, ψ_t satisfying the following conditions

$$\begin{aligned} &\phi_t - \psi_t = p_t \sum_{i=1}^{n} (r_{it} - r_i)x_i, \;\; t = 1, 2, ..., T \\ &\phi_t \psi_t = 0, \quad t = 1, 2, ..., T \\ &\phi_t, \psi_t \geq 0, \quad t = 1, 2, ..., T \end{aligned} \tag{2.16}$$

It is neccessary to replace all occurrence of the term $p_t|\sum_{i=1}^{n}(r_{it} - r_i)x_i|$ to with ϕ_t, ψ_t in such a manner that for all $p_t \geq 0$ at any t, at an optimal solution to the reformulated problem and for each t, we must have either $\phi_t = 0$ or $\psi_t = 0$, because otherwise both ϕ_t and ψ_t would be reduced by the same amount and preserve feasibility, while maximizing the benefit(or reducing the cost in dual problem), which is in contradiction of optimality [8]. Having guaranteed that either $\phi_t = 0$ or $\psi_t = 0$, the desired relation $p_t|\sum_{i=1}^{n}(r_{it} - r_i)x_i| = \phi_t + \psi_t$ now follows. Then, the MAD model equation (2.14) turns into:

$$\begin{aligned} maximize \quad & \sum_{i=1}^{n} r_i x_i \\ \text{s.t.} \quad & \sum_{t=1}^{T} (\phi_t + \psi_t) \leq w \\ & \phi_t - \psi_t = p_t \sum_{i=1}^{n} (r_{it} - r_i)x_i, \;\; t = 1, 2, ..., T \\ & \phi_t \psi_t = 0, \quad t = 1, 2, ..., T \\ & \phi_t, \psi_t \geq 0, \quad t = 1, 2, ..., T \\ & \mathbf{x} \in X. \end{aligned} \tag{2.17}$$

The above problem could be reformulated by applying the result in linear programming from Chvatal (1983) [8], complementary conditions $\phi_t \psi_t = 0, t = 1, 2, ..., T$ can be removed so that the problem becomes a linear programming problem[8]:

$$
\begin{aligned}
maximize \quad & \sum_{i=1}^{n} r_i x_i \\
\text{s.t.} \quad & \sum_{t=1}^{T} (\phi_t + \psi_t) \leq w \\
& \phi_t - \psi_t = p_t \sum_{i=1}^{n} (r_{it} - r_i) x_i, \;\; t = 1, 2, ..., T \\
& \phi_t, \psi_t \geq 0, \quad t = 1, 2, ..., T \\
& \sum_{i=1}^{n} x_i = 1, \\
& 0 \leq x_i \leq \alpha, \quad i = 1, 2, ..., n.
\end{aligned}
\tag{2.18}
$$

In the last section of methodology, the Mean-Absolute-Deviation model will be applied to a so-called Maximal Predictability Portfolio optimization problems, reducing it into a linear complementary problem and solving it using 0-1 integer programming.

2.6 Maximal Predictability Portfolio

Maximal Predictability Portfolio is amongst one of the most potential methods to be exploited in portfolio analysis. The year of 2005 experienced a breakthrough stage in solving the resulting non-convex fractional quadratic programming from the problem by algorithm proposed by Yamamoto and Konno [6], which can resolve the MPP optimization problem up to hundred of assets and a maximum of 15 factors [9]. This section is devoted to what is MPP optimization problem and the way to formulate it from a non-convex quadratic programming into 0-1 linear programming. We will limit our discussion to the definition and formulation and not present the algorithm to solve it since it is out of scope of this report.

2.6.1 Definition

Consider the set of n risky assets with the following rate of returns during time period t; $r_t = (r_{1t}, r_{2t}, r_{3t}, ..., r_{nt}), t = 1, 2, ..., T$. Assume that r_t is jointly stationary and follows ergodic process with finite expectation $\mu = (\mu_1, \mu_2, ..., \mu_n)^\intercal$ and finite autocovariance matrix.

Now suppose that a factor model that represent the rate of return of r_i of i^{th} is presented:

$$r_i = \beta_{i0} + \beta_{i1}F_1 + \beta_{i2}F_2 + ... + \beta_{iK}F_K + \epsilon_i, \quad i = 1, 2, 3, ..., n, \tag{2.19}$$

where $F_1, F_2, ..., F_K$ are factors and ϵ_i's are independent from each other with $E[\epsilon_i] = 0$, $Cov(F_k, \epsilon_i) = 0, k = 1, 2, ..., K$ and we obtain the coefficient β_{ik}'s, $k = 1, 2, ..., K$ from the method of least squares.

Denoting realization of r_i and F_k during period t as below:

$$r_t = (r_{1t}, r_{2t}, ..., r_{nt})^\intercal \in R^n,$$
$$F_t = (F_{1t}, F_{2t}, ..., F_{Kt})^\intercal \in R^K,$$
$$\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, ..., \epsilon_{nt})^\intercal \in R^n.$$

Suppose that ϵ_t is a conditionally homoskedastic process with $E[\epsilon_t] = 0$ and the information set $\mathcal{F}_t$ is well behaved at least to make r_t is a stationary and ergodic process. Moreover, $\mathcal{F}_t$ is earlier observable economic variables ,such as dividend yield, interest rate spreads, or other leading economic indicators. Recalling that an investor is risk-averse and wishes to diversify portfolio allocation in order to achieve some minimal return at a given risk level. Then, the set of optimal portfolio weights, as denoted by **x**, will boost the coefficient of determination of an portfolio investment is the optimal (or suboptimal) of the following optimization problem:

$$\begin{aligned}
\max_{x} \quad & \frac{Var(\tilde{r_t}^{\intercal} x)}{Var(r_t^{\intercal} x)} \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = 1 \\
& \sum_{i=1}^{n} x_i \tilde{r}_i \geq \upsilon \\
& 0 \leq x_i \leq \alpha, \quad i = 1, 2, ..., n
\end{aligned} \tag{2.20}$$

where $\tilde{r_t} = E[r_t \mid \mathcal{F}_{t-1}]$ be the forecast return based on information that available at time $t-1$ and we easily deduce that $r_t = \tilde{r_t} + \epsilon_t$, υ is a constant represent the minimal return that an investor desires and α is the restriction of asset i^{th} in allocation among portfolio.
The interpretation of the model is fairly comprehensive: *a maximal predictability portfolio (MPP) is seen as a portfolio that achieve the best predictive power (measured by* R^2*) in terms of rate of return.*

2.6.2 Portfolio Optimization problems

By the definition of standard deviation of a random variable

$$\sigma(r) = \sqrt{Var(r)},$$

hence, when the quotient between variances of two random variables maximizes, so does the standard deviation. Applying theorem 2.5.1, if r_t and $\tilde{r_t}$ are normally distributed,

$$\text{maximize } \tfrac{Var(\tilde{r_t})}{Var(r_t)} \text{ is equivalent to maximize } \tfrac{W(\tilde{r_t})}{W(r_t)}$$

Doing a inverse transformation turns the problem into a minimization one,

$$\text{maximize } \tfrac{W(\tilde{r_t})}{W(r_t)} \text{ is equivalent to minimize } \tfrac{W(r_t)}{W(\tilde{r_t})}$$

Reformulating equation (2.20) by replace the variance by absolute deviation and arrive as follows:

$$\begin{aligned}
\min_{x} \quad & \frac{E[|r_t^\intercal x - E[r_t^\intercal x]|]}{E[|\tilde{r_t}^\intercal x - E[\tilde{r_t}^\intercal x]|]} \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = 1 \\
& \sum_{i=1}^{n} x_i \tilde{r}_i \geq \upsilon \\
& 0 \leq x_i \leq \alpha, \quad i = 1, 2, ..., n
\end{aligned} \tag{2.21}$$

Further simplifying is under requirement. Note that $r_t^\intercal x = \sum_{i=1}^{n} x_i r_{i,t}$, this relationship will be worked on sooner. Recall the Fama-French three factors model and let F_1, F_2 and F_3 be three factors in the model . Then, the weighted rate of return at period t could be represented with explanatory variables F_1, F_2, F_3 in FF3 model:

$$\begin{aligned}
x_i r_{i,t} &= x_i\beta_0 + x_i\beta_{i,1}F_{1,t-1} + x_i\beta_{i,2}F_{2,t-1} + x_i\beta_{i,3}F_{3,t-1} + x_i\epsilon_{i,t} \\
E[x_i r_{i,t}] &= x_i\beta_0 + x_i\beta_{i,1}E[F_{1,t-1}] + x_i\beta_{i,2}E[F_{2,t-1}] + x_i\beta_{i,3}E[F_{3,t-1}],
\end{aligned} \tag{2.22}$$

Taking the difference between the former and the latter:

$$\begin{aligned}
x_i r_{i,t}^\intercal - E[x_i r_{i,t}^\intercal] &= x_i\beta_{i,1}(F_{1,t-1} - \hat{F}_1) + x_i\beta_{i,2}(F_{2,t-1} - \hat{F}_2) + x_i\beta_{i,3}(F_{3,t-1} - \hat{F}_3) \\
&= \sum_{i=1}^{n} x_i \sum_{k=1}^{3} \beta_{i,k}(F_{k,t-1} - \hat{F}_k) + \sum_{i=1}^{n} x_i\epsilon_{i,t}
\end{aligned} \tag{2.23}$$

Hence, with the relation $r_t = \tilde{r}_t + \epsilon_t$, the denominator becomes:

$$\begin{aligned}
x_i \tilde{r_{i,t}}^\intercal - E[x_i \tilde{r_{i,t}}^\intercal] &= x_i\beta_{i,1}(F_{1,t-1} - \hat{F}_1) + x_i\beta_{i,2}(F_{2,t-1} - \hat{F}_2) + x_i\beta_{i,3}(F_{3,t-1} - \hat{F}_3) \\
&= \sum_{i=1}^{n} x_i \sum_{k=1}^{3} \beta_{i,k}(F_{k,t-1} - \hat{F}_k)
\end{aligned} \tag{2.24}$$

Assume that (F_1, F_2, F_3) be distributed over a finite set $(f_{1t}.f_{2t}, ..., f_{3t})$ with equal probability $p_t = \frac{1}{T}$.Then

$$E[|r_t^\intercal x - E[r_t^\intercal x]|] = \frac{1}{T}|\sum_{t=1}^{T}\sum_{k=1}^{3}\beta_k(x)(f_{k,t-1} - \hat{f}_k) + \epsilon_t(x)|, \tag{2.25}$$

$$E[|\tilde{r_t}^\intercal x - E[\tilde{r_t}^\intercal x]|] = \frac{1}{T}|\sum_{t=1}^{T}\sum_{k=1}^{3}\beta_k(x)(f_{k,t-1} - \hat{f}_k)|, \tag{2.26}$$

where

$$\beta_k(x) = \sum_{i=1}^{n}\beta_{i,k}x_i, \quad k = 1, 2, 3 \tag{2.27}$$

$$\epsilon_t(x) = \sum_{i=1}^{n}\epsilon_{i,t}x_i, \quad t = 1, 2, ..., T \tag{2.28}$$

$$\hat{f}_k = E[f_{k,t-1}] = \frac{1}{T}\sum_{t=1}^{T}f_{k,t-1}, \ \ k = 1, 2, 3 \tag{2.29}$$

$$\tag{2.30}$$

Then the maximal predictability portfolio construction problem becomes as follows:

$$\begin{aligned}
\underset{x}{minimize} \quad & \frac{\sum_{t=1}^{T}p_t|\sum_{k=1}^{3}\beta_k(x)(f_{k,t-1} - \hat{f}_k) + \epsilon_t(x)|}{\sum_{t=1}^{T}p_t|\sum_{k=1}^{3}\beta_k(x)(f_{k,t-1} - \hat{f}_k)|} \\
\text{subject to} \quad & \sum_{i=1}^{n}x_i = 1, \\
& \beta_k(x) = \sum_{i=1}^{n}\beta_{i,k}x_i, \quad k = 1, 2, 3 \\
& \epsilon_t(x) = \sum_{i=1}^{n}\epsilon_{i,t}x_i, \quad t = 1, 2, ..., T \\
& 0 \le x_i \le \alpha, \quad i = 1, 2, ..., n, \\
& \sum_{i=1}^{n}\tilde{r}_i x_i \ge \upsilon,
\end{aligned} \tag{2.31}$$

It is vital to convert the problem into 0-1 mixed integer programming, which can be solved faster by computational approach. First of all, applying the Charnes-Cooper linear fractional functionals transformation (1962) [10]. Denoted

$$y_0 = \frac{1}{\sum_{t=1}^{T} p_t |\sum_{k=1}^{3} \beta_k(x)(f_{k,t-1} - \hat{f}_k)|} \tag{2.32}$$

The purpose of introducing this new variable results in escaping from the quotient form and thus modifying the problem into linear programming.

equation (2.31) could be altered:

$$\begin{aligned} \underset{x}{minimize} \quad & \sum_{t=1}^{T} p_t |\sum_{k=1}^{3} \beta_k(x)(f_{k,t-1} - \hat{f}_k) + \epsilon_t(x)| \cdot y_0 \\ \text{s.t.} \quad & \sum_{t=1}^{T} p_t |\sum_{k=1}^{3} \beta_k(x)(f_{k,t-1} - \hat{f}_k)| \cdot y_0 = 1, \\ & \sum_{i=1}^{n} x_i = 1, \\ & \beta_k(x) = \sum_{i=1}^{n} \beta_{i,k} x_i, \quad k = 1, 2, 3 \\ & \epsilon_t(x) = \sum_{i=1}^{n} \epsilon_{i,t} x_i, \quad t = 1, 2, ..., T \\ & 0 \leq x_i \leq \alpha, \quad i = 1, 2, ..., n \\ & \sum_{i=1}^{n} \tilde{r}_i x_i \geq \upsilon, \end{aligned} \tag{2.33}$$

Let $y = y_0 \cdot x$. Then the equivalent problem stems from equation (2.33) will be

$$\begin{aligned}
\underset{x}{minimize} \quad & \sum_{t=1}^{T} p_t |\sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) + \epsilon_t(y)| \\
\text{s.t.} \quad & \sum_{t=1}^{T} p_t |\sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k)| = 1, \\
& \sum_{i=1}^{n} y_i = y_0, \\
& \beta_k(y) = \sum_{i=1}^{n} \beta_{i,k} y_i, \quad k = 1,2,3 \\
& \epsilon_t(y) = \sum_{i=1}^{n} \epsilon_{i,t} y_i, \quad t = 1,2,...T \\
& 0 \leq y_i \leq \alpha \cdot y_0, \quad i = 1,2,...,n, \\
& \sum_{i=1}^{n} \tilde{r}_i y_i \geq \upsilon \cdot y_0,
\end{aligned} \tag{2.34}$$

The optimal solution to equation (2.34) is a pair (y^*, y_0^*), thereby deriving the optimal weights is $x^* = \frac{y^*}{y_0^*}$ for the equation (2.31). Since there exists absolute function in the programming, repeating the same transformation as being introduced earlier in section 2.5 is inquisite. Let present 4 new variables

$$\begin{aligned}
\nu_t &= (\nu_1, \nu_2, ..., \nu_T), & \psi_t &= (\psi_1, \psi_2, ..., \psi_T), \\
\xi_t &= (\xi_1, \xi_2, ..., \xi_T), & \eta_t &= (\eta_1, \eta_2, ..., \eta_T),
\end{aligned}$$

The equation (2.34) could be reconstructed as follows:

$$\begin{aligned}
\min \quad & \sum_{t=1}^{T} p_t(\nu_t + \psi_t) \\
\text{s.t.} \quad & \sum_{t=1}^{T} p_t(\xi_t + \eta_t) = 1, \\
& \nu_t - \psi_t = \sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) + \epsilon_t(y), \quad t = 1, 2, ..., T, \\
& \xi_t - \eta_t = \sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) \quad , t = 1, 2, ..., T, \\
& \nu_t\psi_t = 0, \xi_t\eta_t = 0, \quad t = 1, 2, ..., T, \\
& \nu_t, \psi_t, \xi_t, \eta_t \geq 0, \quad t = 1, 2, ..., T, \\
& (y, y_0) \in Y.
\end{aligned} \tag{2.35}$$

where set Y contains all the following relations:

$$Y = \{(y, y_0) \mid \sum_{i=1}^{n} y_i = y_0, \quad \beta_k(y) = \sum_{i=1}^{n} \beta_{i,k} y_i,$$

$$\epsilon_t(y) = \sum_{i=1}^{n} \epsilon_{i,t} y_i, \quad 0 \leq y_i \leq \alpha \cdot y_0, \quad i = 1, 2, ..., n,$$

$$\sum_{i=1}^{n} \tilde{r}_i y_i \geq \upsilon \cdot y_0, \quad y_0 \geq 0\}$$

with the result of Chvatal (1983) and Konno et al. (2005), discarding the constraint $\nu_t\psi_t \geq 0$ is permitted. However, there is one set of complementary conditions which need manipulating. By introducing an integer 0-1 variable $z_t = \{0, 1\}$and considering these linear inequalities:

$$\xi_t \leq a_t z_t, \quad t = 1, 2, .., T,$$

$$\eta_t \leq b_t(1 - z_t) \quad t = 1, 2, .., T,$$

where

$$a_t = \max\{\max\{\sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) \mid (y, y_0) \in Y\}, 0\},$$
$$b_t = -\min\{\min\{\sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) \mid (y, y_0) \in Y\}, 0\},$$

The above system of inequalities forces ξ_t and η_t to be either 0 and the other one will be unconstrained, which satisfies the complementarity. Finally, the 0-1 mixed integer linear programming will be in the form of the equation (2.34) [11]:

$$\begin{aligned}
\min \quad & \sum_{t=1}^{T} p_t(\nu_t + \psi_t) \\
\text{s.t.} \quad & \sum_{t=1}^{T} p_t(\xi_t + \eta_t) = 1, \\
& \nu_t - \psi_t = \sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) + \epsilon_t(y), \quad t = 1, 2, ..., T, \\
& \xi_t - \eta_t = \sum_{k=1}^{3} \beta_k(y)(f_{k,t-1} - \hat{f}_k) \quad , t = 1, 2, ..., T, \\
& 0 \leq \xi_t \leq a_t z_t, \;\; 0 \leq \eta_t \leq b_t(1 - z_t), \quad t = 1, 2, ..., T, \\
& z_t \in \{0, 1\}, \quad , t = 1, 2, ..., T, \\
& \nu_t, \psi_t, \eta_t, \xi_t \geq 0, \quad t = 1, 2, ..., T, \\
& (y, y_0) \in Y.
\end{aligned} \tag{2.36}$$

III. DATA AND EMPIRICAL RESULTS

In this chapter, there will be 3 main parts. First, Data description part will provide important features (explanatory variables) of data and the steps of transforming raw data into suitable one for optimization problem. Second, optimization programmings in which methods, packages and programing applications will be proposed. After following all the above instruction, the empirical results (back-testing) for US market and Vietnam market will be provided and come through a scutiny before reaching a conclusion. The strategy is that, the portfolio will be purchased proportionally according the optimal weights of **Mean-Variance Portfolio** (MVP) and **Maximal Predictability Portfolio** (MPP) in Sep/2019 and assessed their performance in Dec/2019, a total of 12 weeks

3.1 Data description

The following words will be used in abbreviation for the purpose of convenience:

- MPP: Maximal Predictability Portfolio
- MVP: Mean-Variance Portfolio
- Index: the index market return of S&P500 (US market), VNIDEX (VN market)
- FF3: Fama–French three-factor model
- HOSE: HO CHI MINH stock exchanges

Data acquisition and time frame

In order to compare the performance between and MVP and MPP, the frame of references would include 2 markets: Vietnam and US market. In US market, as the year of 2020 has suffered with numerous uncertained variations, the weekly data would be collected, in *yahoo.com* [12], only for 6 years, from Jan/2014 to Dec/2019. On the other hand, the weekly data in Vietnam market [13] was only collected from Jan/2015 to Dec/2019 since there are some stocks that have not publicly offered on HOSE in 2014, which further became Bluechips, had enormous capitalization and maintained high stability in HOSE, thus being important for calculating SMB and HML factors in the framework of FF3

The data collected contained 2 main kinds: data of stocks, data for Fama-French 3 factor model (FF3) [14]. Moreover, such data was divided into 3 main parts: The first part, from Jan/2014 to April/2019 in US weekly data and from Jan/2015 to April/2019 in Vietnam weekly data, would be the training set. The training set was fitted on the FF3, thereby finding expected return and residuals, and beta coeffiecients. The second part, from May/2019 to August/2019, those parameters, expected return, residuals and beta coeffiencients, would be used as inputs for optimization problem (MPP), with T=12 and n=12, and then yielding optimal weights of each stock. The last part involved performance testing basing on given optimial weights for both MVP and MPP methods for the last 12-week-period between Sep/2019 and Dec/2019

3.1.1 US portfolio

The constructed portfolio in the US market consists of 12 stocks, weekly prices of which was also collected to further calculate weekly returns:

- APPLE: APP
- GOOGLE: GG
- FACEBOOK: FB
- AMAZON: AMZ

- TESLA: TSLA
- NETFLIX: NTFX
- DISNEY: DIS
- NIKE: NIKE
- MICROSOFT: MCFT
- STARBUCKS: SBUX
- ROYAL DUTCH SHEEL: RDS-B
- UNILEVER: UL

The FF3 weekly factors in the US market was obtained through the Kenneth French's website [14]. As some infographical pictures in terms of returns and co-variance between those stocks are provided below in order to help readers have more information about each stocks:

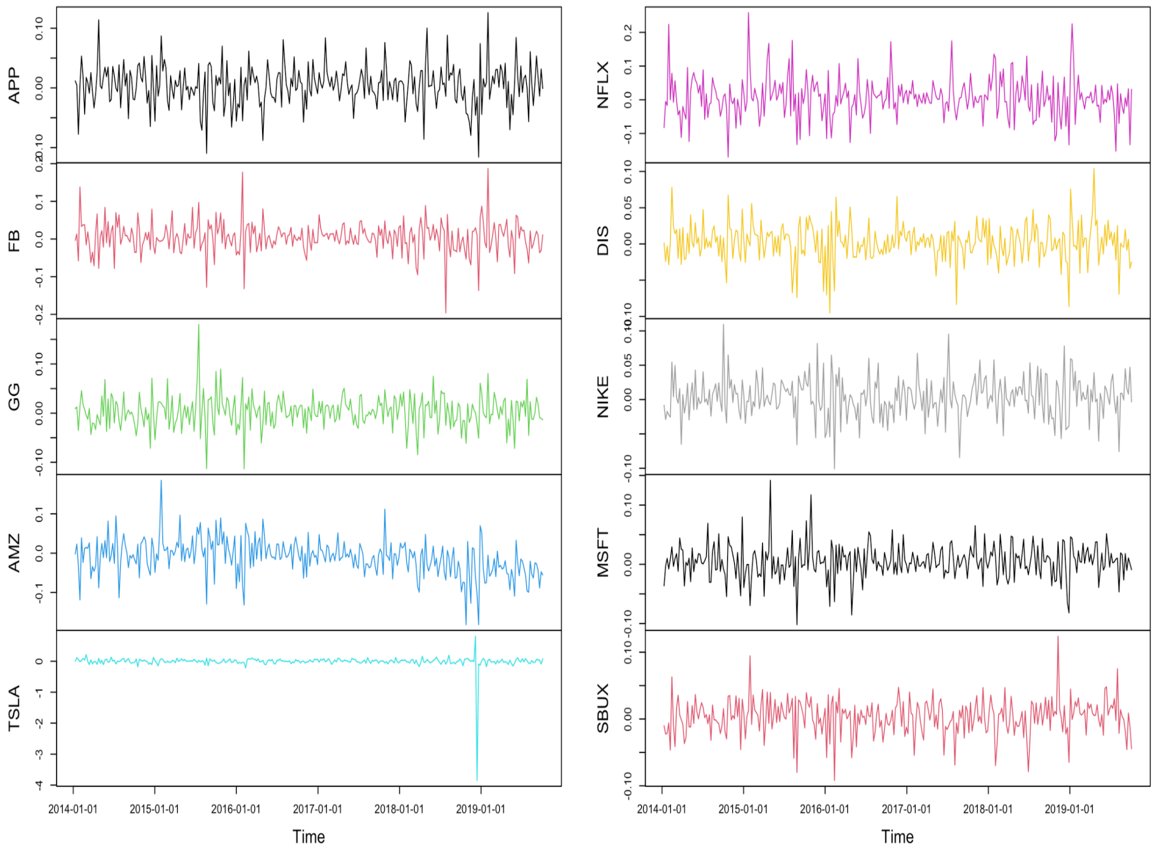

Figure 3.1: US stock returns

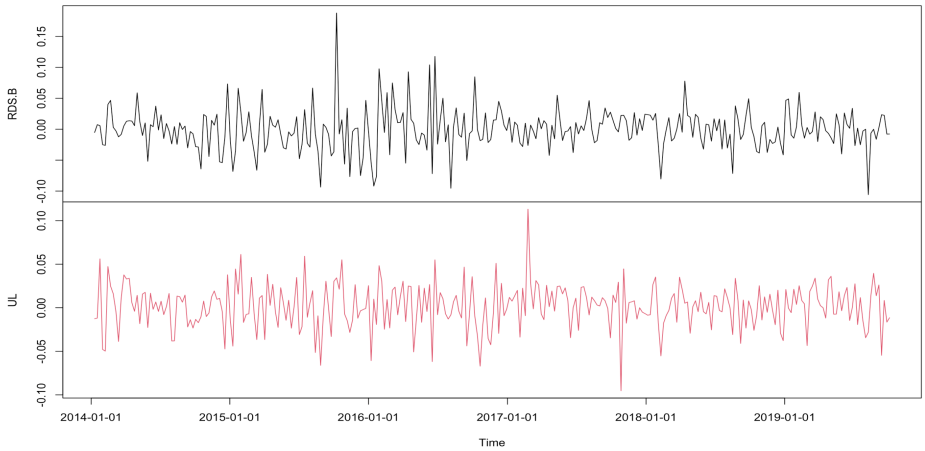

Figure 3.2: US stock returns

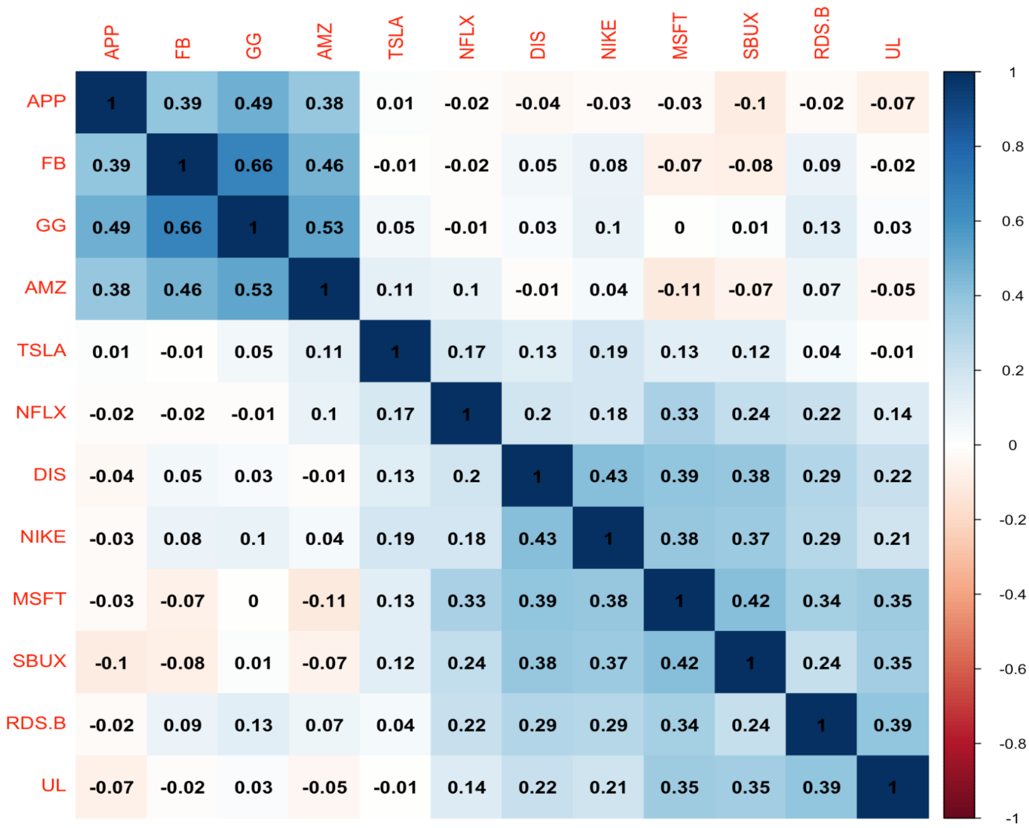

Figure 3.3: Covariance matrix

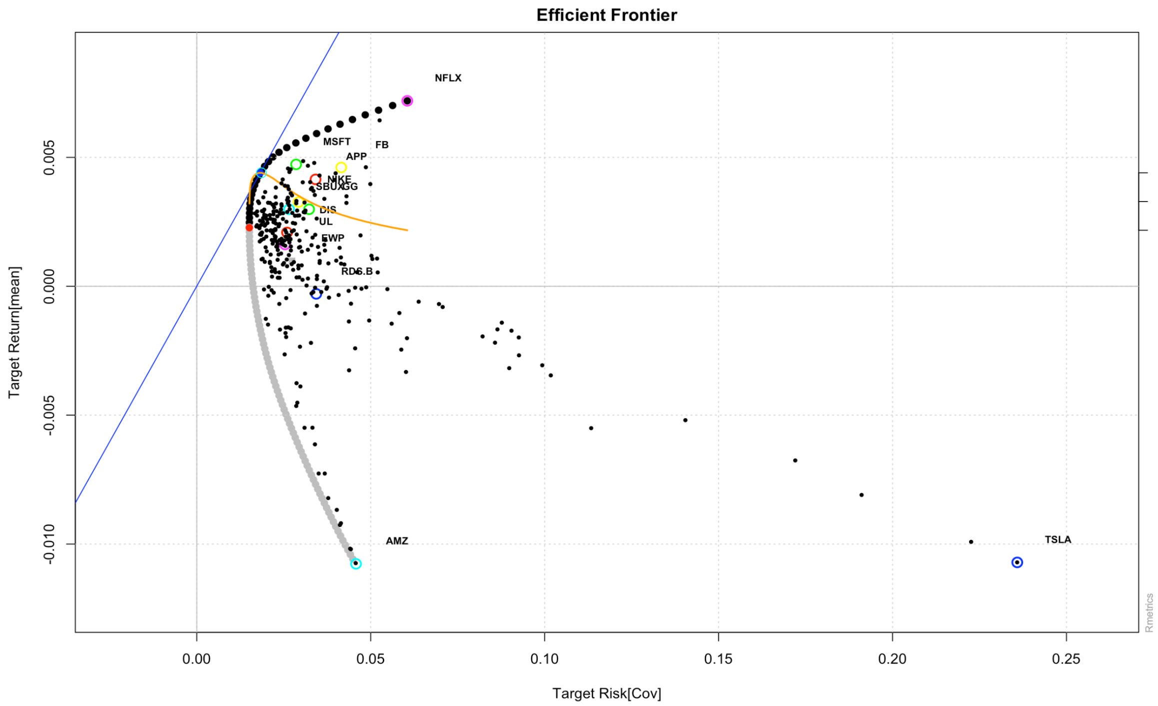

Figure 3.4: Efficient Frontier in US market

Surprisingly, for Tesla, as one of the largest company in the world at present, however, its return remained relatively unchanged prior 2020. with an exception of one week. Thus, from those data, the covariance matrix has been deduced, and shows that there are some potentials to use the diversification of stocks to minimize the risk of a portfolio. Therefore, MVP method would be expected to perform relatively well over the last 12 periods.

3.1.2 Vietnam portfolio

The constructed portfolio in Vietnam market also consists of 12 stocks that are the most stable and trustworthy companies in HOSE, weekly prices of those stocks were collected to calculate returns. Clearly, The following stock data was collected:

- Phu Nhuan Jewelry Joint Stock Company: PNJ
- Refrigeration Electrical Engineering Corporation: REE
- SSI Securities Corporation: SSI

- Sai Gon Thuong Tin Commercial Joint Stock Bank: STB
- Bank for Foreign Trade of Vietnam: VCB
- Vingroup Joint Stock Company: VIC
- Viet Nam Dairy Products Joint Stock Company: VNM
- DHG Pharmaceutical Joint Stock Company: DHG
- Hoa Sen Group: HSG
- Phuoc Hoa Rubber Joint Stock Company: PHR
- FPT Corporation: FPT
- PetroVietnam Gas Joint Stock Corporation: GAS

Unfortunately, the data for factors in FF3 is not available in Vietnam market, thus the data of 30 Bluechip companies that would be on behalf of the performance of Vietnam market was collected, processed and finally resulting with 3 neccessary factors for FF3 framework. Some data insights are also provided below:

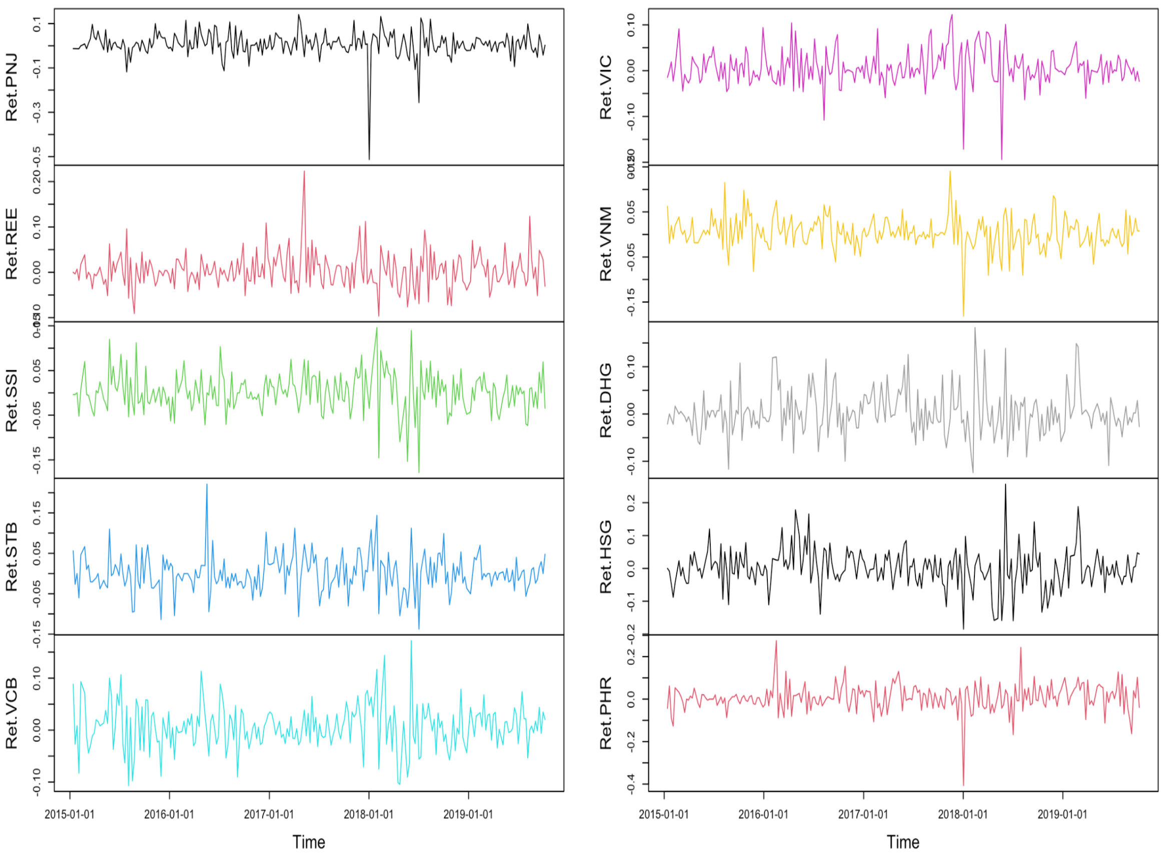

Figure 3.5: Vietnam stock returns

Figure 3.6: Vietnam stock returns

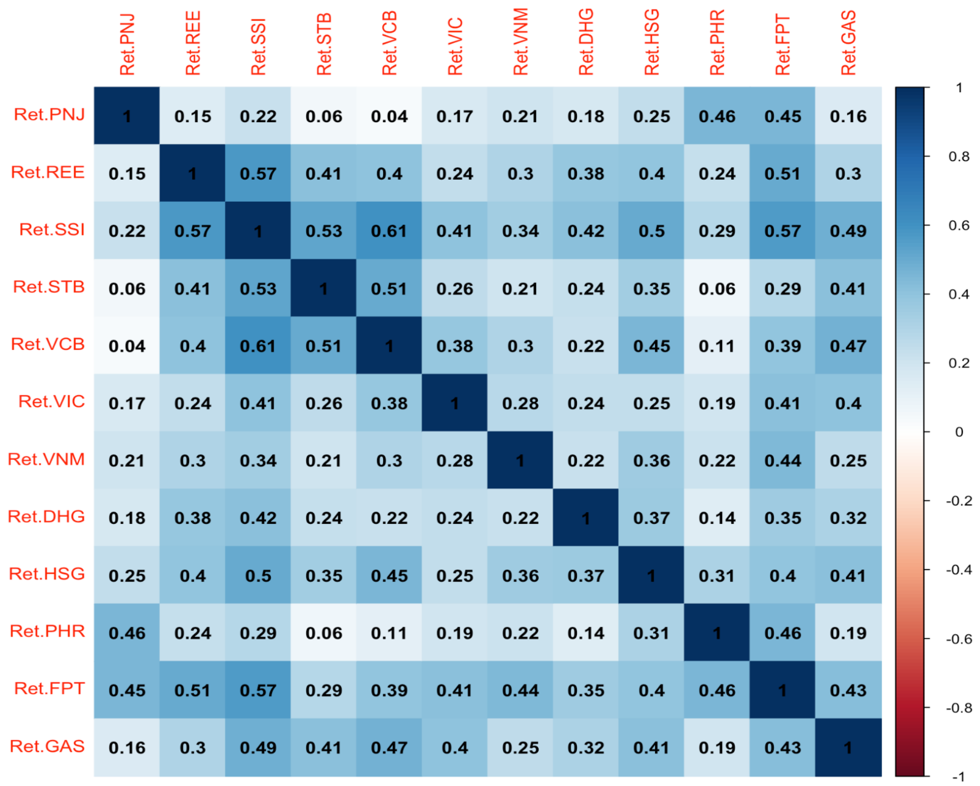

Figure 3.7: Covariance matrix

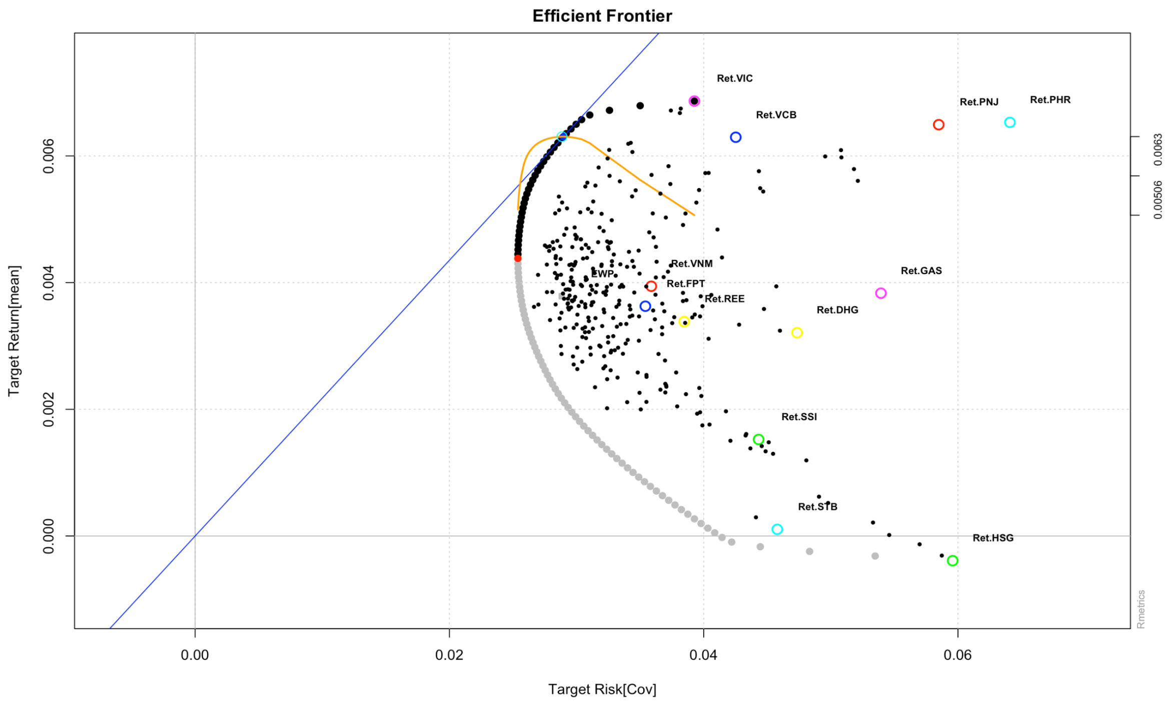

Figure 3.8: Efficient Frontier in Vietnam market

According to the covariance matrix, all the stocks are somehow positively correlated with each other. Therefore, It is quite hard to ultilize the potential of diversification effect, however, there are some stocks that are nearly non-correlated with others, MVP method can use this to build such portfolio with minimal risks.

3.2 Optimization programming

For the goal of seeking optimal weight for stocks, 3 main programming applications that have been used for data processing, optimization problem and back-testing analysis are Python, R programming and Excel (in lastest version)
As mentioned above, first, the raw data of each stock collected from online website just existed in the form of raw weekly price, and then Excel application has been used to transform those into weekly return and used for FF3 to find beta coeffiecients and residual errors that would be later used for MPP optimization problem. The second programming is used for maximal predictability portfolio optimization is Python. In specific, Python has offered a usesul and innovative package, namely *GUROBI*, that is continuously updated and facilitates solving linear and non-linear constraints simultaneously. Indeed, such package has proven their efficiency in solving this optimization problem by giving great solutions under 2 seconds. Lastly, in the scope of testing the performance of each portfolio after finding their optimal weights, R programming has provided a user-friendly environment for plotting and matrix calculating, and maximizing the visualization capacity for further investigations.

3.2.1 Portfolio assessment

In optimization problem, particularly MPP, there are 2 main parameter constraints, υ and α. In specific, υ is the minimum desired weekly rate of return, while α is the maximum weight allowed for each stock in portfolio. In this optimization problem, υ was set to be 0.15% and α was set to be 15%. However, there was no such parameter constraints for traditional optimization problem. Moreover, the following metric would be calculated for assessment.

- Single weekly return equation (2.1) and equation (2.3)
- Performance (Cumulative weekly return) equation (2.2) and equation (2.4)
- Sharpe ratio equation (2.5)

3.3 Empirical results

3.3.1 US market

As proposed above, the constructed portfolio is a combination of 12 stocks. the period between May/2019 and August.2019, 12 weeks, was the standard time frame for optimization problem, both **Mean-Variance Portfolio** (MVP) and **Maximal Predictability Portfolio** (MPP). The first step was to find input for MPP, particularly beta coefficients that are given below:

Stock	APP	FB	GG	AMZ	TSLA	NFLX	DIS	NIKE	MSFT	SBUX	RDS.B	UL
Beta 1	0.741	0.868	0.871	0.861	0.618	0.38	0.3468	0.3436	0.403	0.126	0.508	0.196
Beta 2	-0.215	-0.103	-0.332	-0.0059	1.741	0.354	0.029	0.1448	-0.314	0.065	0.029	-0.177
Beta 3	-0.545	-0.529	-0.444	-0.761	1.792	-0.078	0.295	0.221	0.0465	0.008	0.602	-0.126

After using the above input to solve and then the weight of each stock of each method has been yielded and shown in the following table:

Stock	APP	FB	GG	AMZ	TSLA	NFLX	DIS	NIKE	MSFT	SBUX	RDS.B	UL
MVP weight(%)	16.6	3.16	5.5	4.35	0.00	0.00	16.14	7.43	7.23	15.22	1.51	22.82
MPP weight(%)	15	3.77	15	0.00	14.339	0.00	0.00	15	15	15	0.00	6.89

As we can see from the table, Tesla was weighted 0% in the MVP, but it accounted for nearly 15% in MPP. Conversely, in MPP, there were several stocks, namely AMZ, NFLX, DIS, and RDS.B, with their weight being 0%. When it came to this result, their was a huge surprise since MPP "seemed" a little "biased" in terms of some stocks, and thus calling in question the efficiency of this novel method. Nonetheless, the performance of MPP was far expected, as it completely outperformed traditional method as well as index benchmark (S&P500). The following table will give some insights into single weekly return of each portfolio and cumulative weekly return of a 12-week-period as a whole:

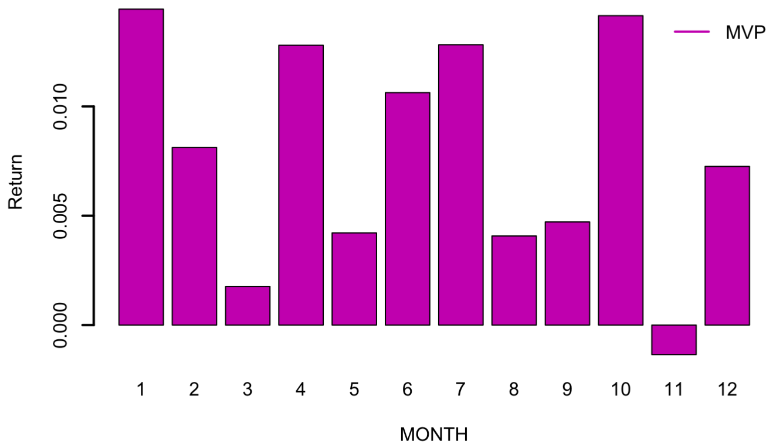

Figure 3.9: MVP weekly return

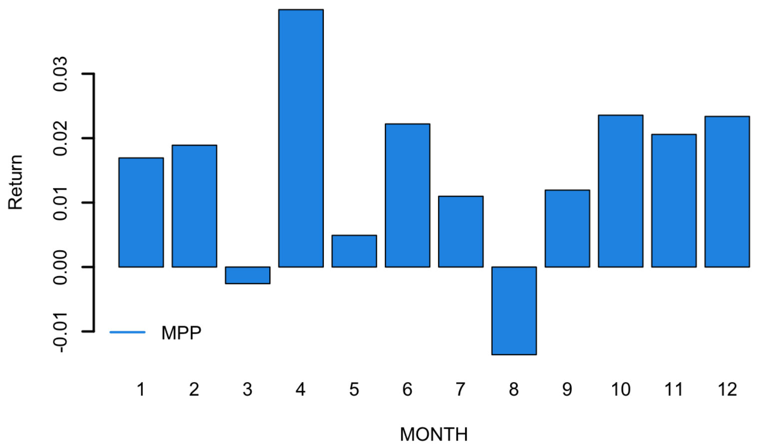

Figure 3.10: MPP weekly return

As it can be clearly observed that the return of both portfolios experienced an overall positive values in weekly return, and both had a tendency to remain positive throughout a period. However, at a closer look, the overall returns for MPP was higher compared to that of MVP. In specific, while the weekly return of MVP fluctuated around under 1%, that of MPP was far from 1%, particularly hovering around 1.5%, nearly doubled its counterpart portfolio. Such returns have to be compared in the same picture frame to have an overall view of each portfolio, as the below picture shows:

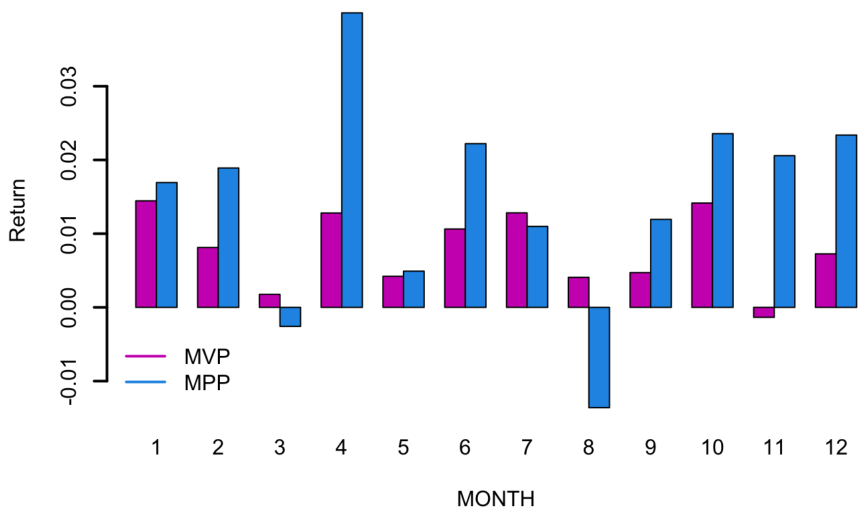

Figure 3.11: MVP and MPP weekly return

The weekly performance of MPP was outstanding compared to the that of MVP. In particular, throughout a period of 12 weeks, the MPP returns on a weekly basis were mostly higher than that of MVP. This, however, only demonstrated a one-side look into weekly returns of each porfolio. With purpose of having objective perspectives of both portfolio, the cumulative performance of a total of 12 weeks will be shown below:

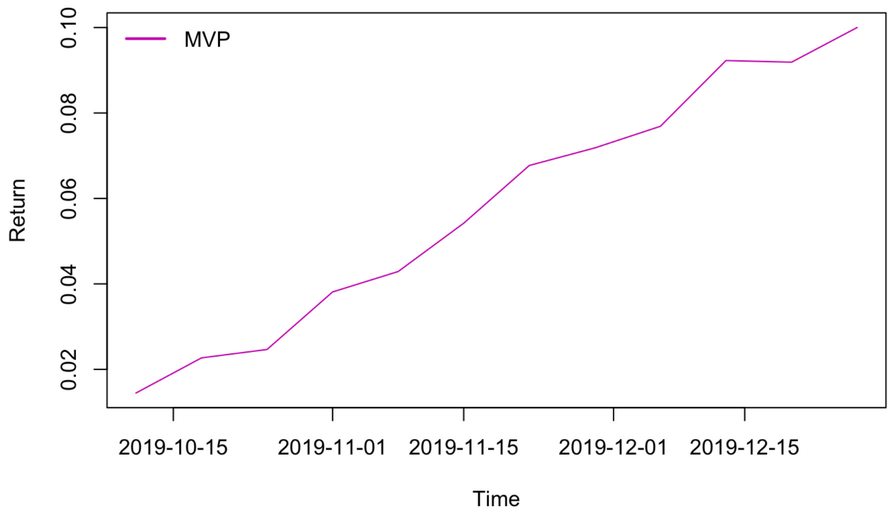

Figure 3.12: MVP cumulative performance

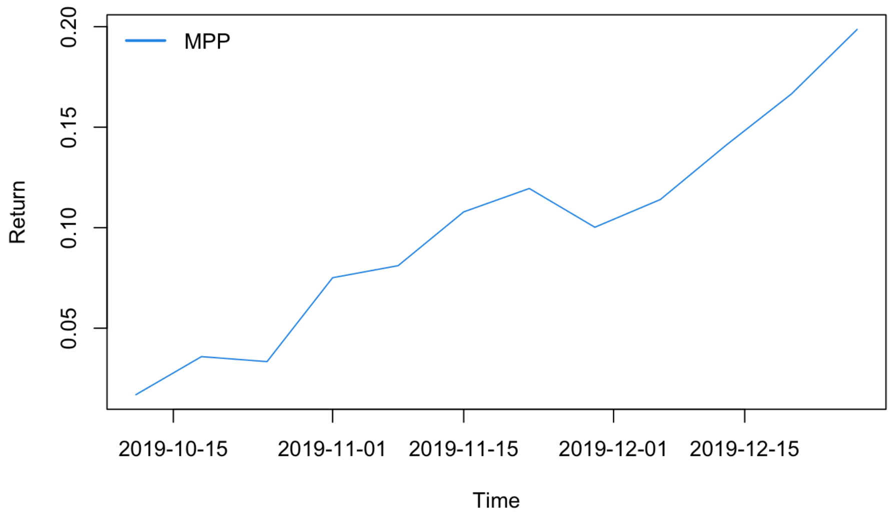

Figure 3.13: MPP cumulative performance

Having a glance at the two pictures, the cumulative returns of both portfolio followed an upward trend throughout a period of 12 weeks. Thus, a proposal of *INDEX* returns, on behalf of S&P500, would be used as a framework of reference.

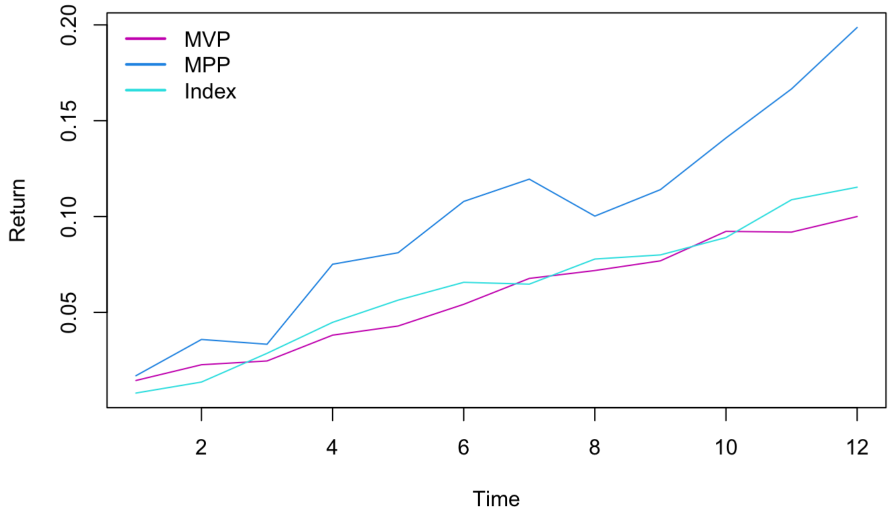

Figure 3.14: MVP, MPP and INDEX per

In this cumulative metric, it is more reasonable to evaluate the performance of both portfolios. This framework has significant provided more evidence for the outstanding of MPP portfolio in compare to its counterpart, namely MVP. Moreover, a subtle detail that has to be noted is that the MVP performed nearly the same with INDEX returns, thus calling into question the efficiency of this traditional method. The MPP method was indeed superior to all the reference frameworks, certainly outperforming and having higher returns over a 12-week-period. Notably, in the last week, the returns of MPP was relatively 2 times higher than MPP's and INDEX's returns, 20% and 10% respectively.

In addition, the **Sharpe ratio (investment efficiency)** will be shown to evaluate the investment effiency of each portfolio:

Week	1	2	3	4	5	6	7	8	9	10	11	12
MVP	0.8075	1.7416	1.6673	1.6072	1.5644	1.5490	1.5639	1.60564	1.6941	1.8440	2.0104	2.6473
MPP	0.8120	1.7786	1.6913	1.6617	1.6164	1.6163	1.6358	1.6894	1.8034	1.9824	2.2956	2.8742

Table 3.1: Investment Effieciency

The weekly-updated invesment effiency of both portfolio was postive over the period. Notably, however, the sharpe ratio of MPP was higher than that of MVP for all weeks.

In conclusion, in the US market, as the trends had been predicted to increase over the next 12 weeks, MPP has successfully shown its outstanding performance in portfolio optimization framework by maximizing the capacity of predictability of this portfolio that strongly fitted on the upward markets trend (shown by market index). This, consequently, would improve investors' sense of decision making on whether keeping or selling stocks and proposed a novel concept for future portfolio optimization.

3.3.2 Vietnam market

To begin, after rigorously finding the 3 main factors, namely excess market return, SMB and HML, of FF3 model. Weekly data return of each stock was fitted on this model and the yielded beta and residuals were stored and later used for MPP optimization, all of which are shown as follows:

Stock	PNJ	REE	SSI	STB	VCB	VIC	VNM	DHG	HSG	PHR	FPT	GAS
Beta 1	0.405	0.787	1.069	0.849	0.896	0.568	0.482	0.82	1.219	0.559	0.675	1.233
Beta 2	0.628	0.643	0.212	-0.019	-0.496	-0.270	-0.238	0.562	0.944	1.177	0.387	-0.160
Beta 3	-0.366	0.476	0.170	0.487	-0.119	-0.416	-0.422	-0.211	0.176	0.211	-0.017	-0.263

After using the above inputs, a result of each method in terms of stock weights was yielded and shown below:

Stock	PNJ	REE	SSI	STB	VCB	VIC	VNM	DHG	HSG	PHR	FPT	GAS
MVP weight(%)	7.96	14.70	0.00	9.88	7.17	19.08	26.50	9.07	0.00	1.93	3.70	0.00
MPP weight(%)	15	14.02	0.00	6.631	15	15	8.353	15	15	0.00	7.818	0.793

The amount of weight of each stock varied significantly between each portfolio, thus evoking a typical question of the performance of each portfolio. The strategy is that after finding optimal solutions, the corresponding weight of stocks will be purchased and then deducing their performance. Therefore, a backtesting period including 12 weeks was applied to check whether did the portfolio perform well or not. As the following bar chart will show the single weekly return of MVP and MPP respectively:

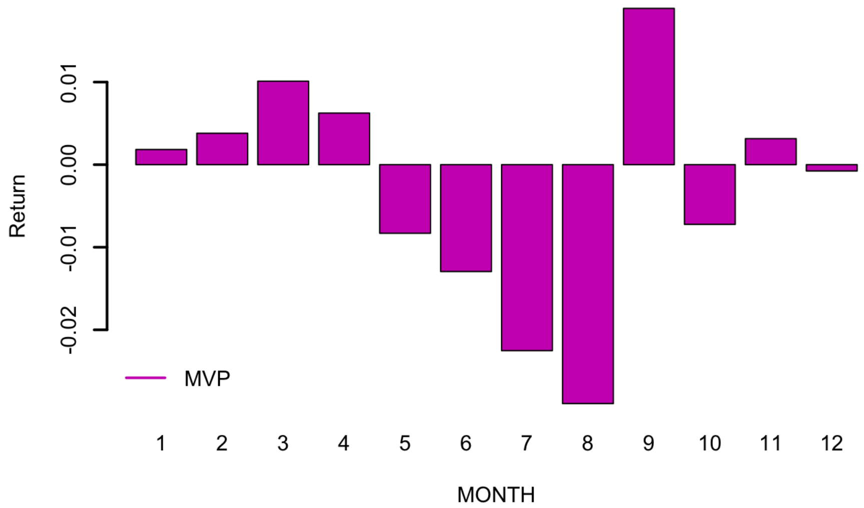

Figure 3.15: MVP single weekly return in Vietnam market

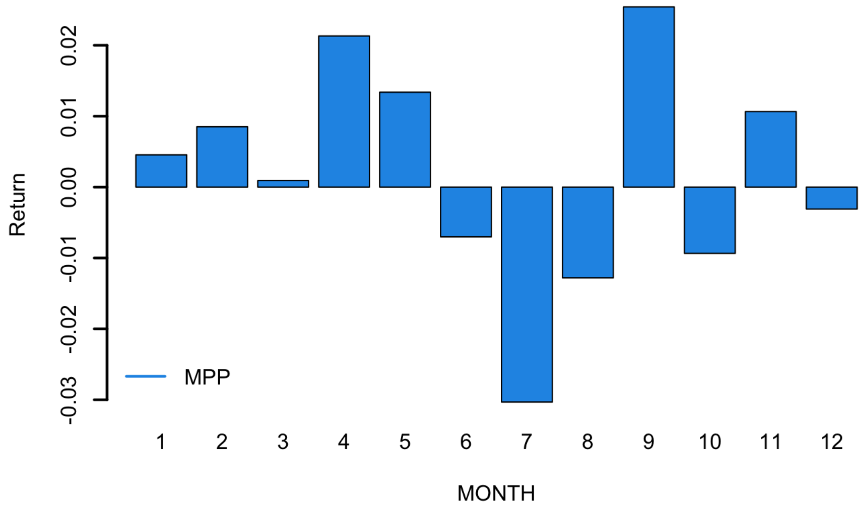

Figure 3.16: MPP single weekly return in Vietnam market

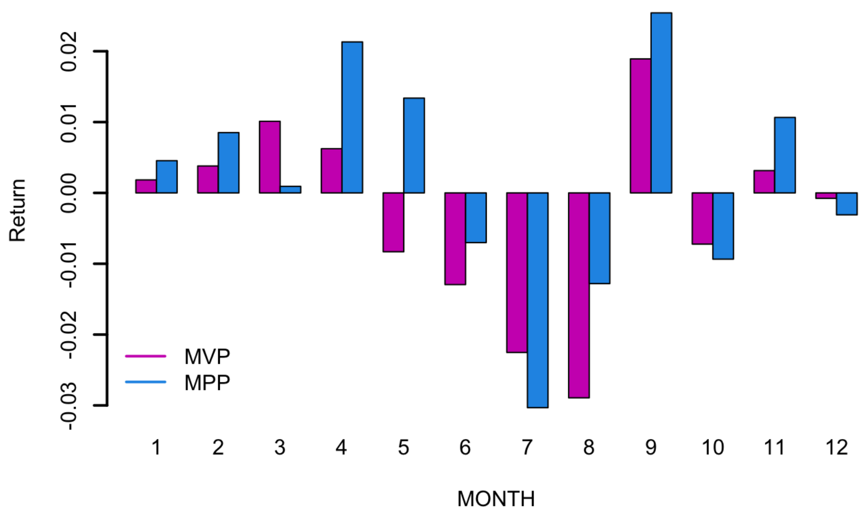

Figure 3.17: MVP and MPP single weekly return in Vietnam market

Looking at closer details of the single weekly return, over the first 4 weeks, Oct/2019, MPP showed good positive return compared to traditional method. However, there were drops in Nov/2019 of both portfolio, thus indicating that the market at the mean time was experiencing a downward trend. Fortunately, MPP still demonstrated its effieciency by having a lower magnitute of loss in comparison with MVP portfolio. The overall single weekly performance was thus assessed with a better operation.
The cumulative performance of each in conjuction with INDEX performance, *VNIDECT*, would be used to achieve more divergent perspectives:

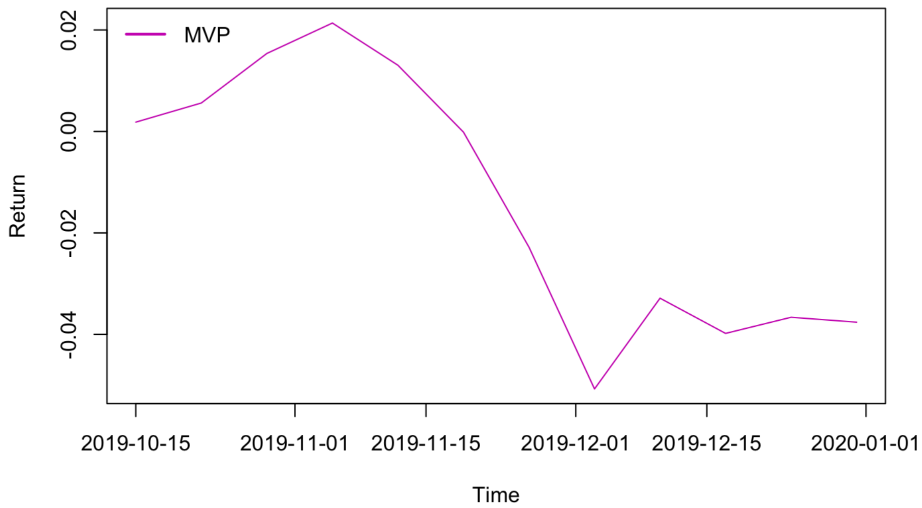

Figure 3.18: MVP cumulative performance in Vietnam market

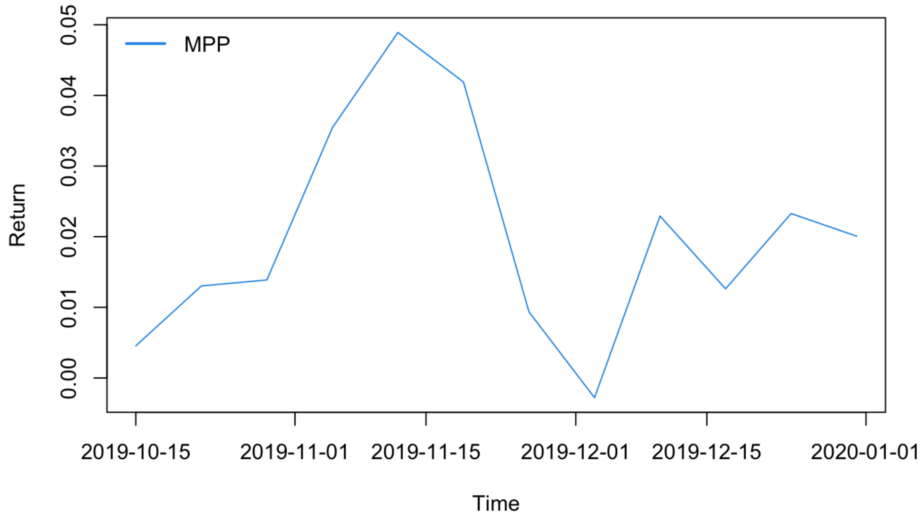

Figure 3.19: MPP cumulative performance in Vietnam market

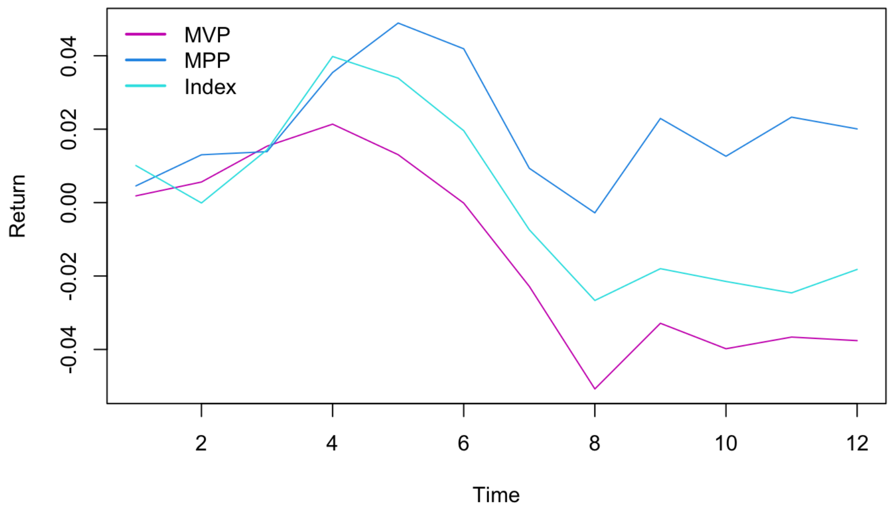

Figure 3.20: MVP and MPP cumulative performance in Vietnam market

Apparently, there was some variations in Nov/2019, which will be further discussed at the next chapter. Primarily, the cumulative performance of MPP was indeed outstanding compared to tradition method as well as the index cumulative returns. In particular, over the first 3 weeks, the predictability of MPP was strongly fit with an upward trend of the market. Significantly, in the 5th week of the period, while the index and MVP return started to experience a long decline, MPP still observed a rise in its return. The rest of period remained stagnant, but cumulative return of MPP still maintained at positive sign, thus showing the effectiveness of MPP method. The following **Sharpe ratio (invesment efficiency)** was yielded in order to make further analysis:

Week	1	2	3	4	5	6	7	8	9	10	11	12
MVP	0.0632	0.2592	0.7864	1.1052	0.6769	-0.0326	-1.1069	-2.494	-1.6611	-2.054	-1.8516	-1.9078
MPP	0.2008	0.6428	0.6931	1.8413	2.6413	2.2771	0.4237	-0.1603	1.1329	0.6083	1.1456	0.996

Table 3.2: Investment Effieciency

The sharpe ratio has indicated that the performance of MPP was much better than that of MVP. In specific, while investment efficiency of MVP remained negative in the second half of the period, that of MPP always hovered positively with only one exception.

In conclusion, in Vietnam market, although the effective of the new method was not as clear as the US market, MPP still performed greatly, therefore being superior to tradional MVP and INDEX at the mean time.

IV. DISCUSSION

4.1 Maximal predictability portfolio optimization

To begin, as the back-testing section has proved the performance of the two methods, MPP and MVP, and index market return. Some market information has to be discussed for the purpose of elucidating the effienciency of MPP method.
First, in the US market, as it has been clearly reflected through the cumulative return index, in the last 3 months of 2019, the market somehow experienced a significant rise of over 10%. This phenomenon has been partly explained by the control of FED, by decreasing the interest rates, thus stimulating the economy to thrive, forcing more money into stocks and significantly pulling up the market. The trade war, however, between US and China escalated but it only had transient effects on the stock market. Moreover, the year of 2019 also marked some important milestones of the two tech giants, namely Apple and Microsoft, that have raised their maket capitalization to more than $1 trillion dollars. Year of 2019 enjoyed a compounded effect of booming technology and supplemented money from the FED. Those contribution has somehow explained an upward trend of the market in late 2019. Regarding MPP, as well as index market return, this method has successfully captured increased trends of market and maximized its capacity of prediction. In specific, MPP's cumulative return was far higher than traditional method and market index's, with nearly double, approximately 20% for the former and only around 10% for the latter. Therefore, this method can be scintillatingly potential for a maximizing-profit strategy in the US market
Second, in Vietnam market, however, as can be observed from the chart, maximizing-

profit strategy was transient, MPP only significantly performed well over the first 5 weeks. At the mean time, the index of VNIndex just had reached the mile stone of 1000 points, and thus decreasing after which since investors' tended to predict the market will fall after continuously increases. Notably, MPP performance has a similar movement to that of market index. This, consequently, shows some clues. In particular, as the market was following the downward trends, MPP could possibly captured the declining overall trend. Most importantly, one significant detail is that even though overall performance of the market and traditional method observed negative results, MPP still maintained at a positive rate of returns over the rest of the period.

To conclude, MPP method has successfully maximized its predictability in both market, thus showing its superiority to MVP method. Moreover, back-testing section also hints that investors/users should stop investing when there are some decreasing signals of the market, and only invest when there is a postive market in order to capture the predictability and thereby maximing returns. Lastly, this section suggests that if there was some unexpected uncertainty from the market, MPP still guarantees a positive returns for investors.

4.2 Fama-French 3 factor model and suggestions for futher research

On the one hand, throughout the process of optimization, the proposed method was in line with FF3 regression. The FF3 method, with is an extention of traditional one, namely CAPM, has sufficiently suited the empirical data for the US market and Vietnam market. On the other hand, in the process of developing MPP method, Lo and MacKinley proposed 7-fator model. Perhaps, the more factors, the more predictive power. Therefore, 7-factor model will be suggested for further research, with the other 4 factors, the yielded expected return will be more correct. Moreover, as the market is becoming more uncertained, which might violate some assumptions of traditional multi-regression model (stationary), a more suitable alternative is *Autoregressive intergrated moving average* (ARIMA) model will be suggested. It can be used to achieve stationary time series through an amount of differences. The liminted time frame of this project also demands further analysis before reaching final conclusion on its effectiveness. A suggestion is thus to develop and test for several time frame including different period duration on daily data, or even more effective one is rolling window.

V. CONCLUSION

The goal of this project is to develop such a portfolio that can maximazing predictability, and thus being suitable for ultimate maximizing-profit purpose. The result has been tested in US and Vietnam market, and therefore proving it effectiveness. Moreover, these results demonstrate the robustness of maximal predictability portfolio optimization compared to Mean-Variance portfolio optimization. Most of the promising features of MPP has been fairly exploited through FF3 model, but further reasearch should conduct on other models beside FF3 to ultilize its predictive power. To conclude, basing on all of results that have been yielded on the basis of MPP and MVP, the novel concept of optimization performed satisfactory.

Bibliography

[1] Petter N. Kolm, Reha Tütüncü, and Frank J. Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. European Journal of Operational Research, 234(2):356 – 371, 2014. ISSN 0377-2217. http://www.sciencedirect.com/science/article/pii/S0377221713008898. 60 years following Harry Markowitz's contribution to portfolio theory and operations research.

[2] Andrew W Lo and A. Craig MacKinlay. Maximizing predictability in the stock and bond markets. Working Paper 5027, National Bureau of Economic Research, February 1995. http://www.nber.org/papers/w5027.

[3] William F. Sharpe. "Capital Asset Prices: A Theory Of Market Equilibrium Under Conditions Of Risk," Journal of Finance, American Finance Association, vol. 19(3), pages 425-442, September, 1964.

[4] Fama, E. F.; French, K. R. (1992). "The Cross-Section of Expected Stock Returns". The Journal of Finance. 47 (2): 427.

[5] Frank J. Fabozzi, Petter N. Kolm, Dessislava A. Pachamanova, Sergio M. Focardi. Robust Portfolio Optimization and Management. John Wiley & Sons, Inc. 2 January 2012

[6] Konno, H., Morita, Y. & Yamamoto, R. A maximal predictability portfolio using absolute deviation reformulation. Comput Manag Sci 7, 47 (2010).

[7] Hiroshi Konno & Tomoyuki Koshizuka (2005) Mean-absolute deviation model, IIE Transactions, 37:10, 893-900

[8] Chvatal, V. (1983) ´ Linear Programming. San Francisco: Freeman

[9] Rei Yamamoto, Daisuke Ishii, Hiroshi Konno (2007). A maximal predictability portfolio model: Algorithm and performance evaluation. International Journal of Theoretical and Applied Finance 10(06):1095-1109. September 2007

[10] A. Charnes W. W. Cooper (1962) Programming with Linear Fractional Functionals Naval Research Logistics Quarterly 9

[11] Constantin Zopounidis, Michael Doumpos, and Panos M. Pardalos. Handbook of Financial Engineering. Springer Publishing Company, Incorporated, 2011. ISBN 1441945733, 9781441945730

[12] Yahoo stock data (US data). https://finance.yahoo.com/lookup?s=DATA.

[13] Vietnam stock data https://www.cophieu68.vn/export.php

[14] Kenneth French. Data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.