UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

PAULO FERREIRA NAIBERT

ESSAYS IN PORTFOLIO OPTIMIZATION

Porto Alegre

2019

PAULO FERREIRA NAIBERT

ESSAYS IN PORTFOLIO OPTIMIZATION

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Doutor em Economia.

Orientador: Prof. Dr. João Frois Caldeira

Porto Alegre

2019

CIP - Catalogação na Publicação

PAULO FERREIRA NAIBERT

ESSAYS IN PORTFOLIO OPTIMIZATION

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Doutor em Economia.

Aprovada em: Porto Alegre, 17 de setembro de 2019.

BANCA EXAMINADORA:

Prof. Dr. João Fróis Caldeira – Orientador
UFRGS

Prof. Dr. Márcio Poletti Laurini
USP–RP

Prof. Dr. Tiago Pascoal Filomena
UFRGS

Dr. Leonardo Riegel Sant'anna
Sicredi

ACKNOWLEDGEMENTS

ABSTRACT

This thesis presents three essays in the topic of portfolio optimization and index tracking. The first essay is a critique of the Tangency Portfolio (TP). The TP has paramount theoretical importance in the Modern Portfolio Theory, however, its performance is far from satisfactory. The TP presents unstable weights, what increases the portfolio turnover and, consequently, its transaction costs. Furthermore, the denominator of the TP is frequently close to zero, what can result in extreme weights, precluding the portfolio from being well diversified. There is also the possibility that the TP's denominator is negative, causing an inversion of the allocation vector's signal and the delivery of a poor portfolio. This work compares the TP with other portfolios and finds that the TP always comes at the bottom. The work also offers propositions that show that the TP can be outperformed (in terms of utility) by other portfolios. The second essay provides an explicit derivation to the equivalence between the minimum variance portfolio of excess returns on a benchmark and the minimum Tracking Error volatility portfolio. This result relies on the Sherman-Morrison formula. The essay also presents an equivalence of those results to an OLS regression with constrained beta. Further, the essay uses the first equivalence result to find a tracking portfolio using the approach of Kempf and Memmel (2006). The third essay compares the performance of passive investment portfolio with a small number of assets (sparse index tracking portfolios) using different feature selection algorithms. To that end we provide an empirical examination with two datasets: one American and one Brazilian. To isolate the effect of the selection methods, we separate the asset selection and the asset allocation phase. In the asset allocation phase, we follow Liu (2009), and select minimum Tracking Error volatility portfolios. The selection methods used are the backward stepwise selection, forward stepwise selection, and the lasso. Our results show that, in the Brazilian case, the lasso selection method is the best tracker. In the American case, the lasso presents better risk-adjusted performance, but this is due to higher mean returns, not lower volatility. This is undesirable in our case. One highlight of this essay is that the best tracker for the American Dataset uses the backward iteration algorithm (a simple method that receive little attention in the literature).

Key-words: Mean-Variance Analysis. Portfolio Optimization. Index Tracking. Performance Evaluation.

RESUMO

A presente tese apresenta três ensaios com o tema de otimização de carteiras e *index tracking*. O primeiro ensaio é uma critica ao Portfolio de Tangência (TP). O TP tem grande importância teórica na Moderna Teoria do Portfólio; porém, o seu desempenho está longe de ser satisfatório. O TP apresenta pesos instáveis, o que aumenta o *turnover* da carteira e, consequentemente, os seus custos de transação. Além disso, frequentemente, o denominador do TP é perto de zero, o que pode resultar em pesos extremos, o que impede a boa diversificação dos pesos da carteira. Também há a possibilidade de o denominador do TP ser negativo, o que causa a inversão do sinal do vetor de alocação, resultando em um portfólio ruim. Este trabalho compara o TP com outras carteiras e chega a conclusão que o TP tem um desempenho pífio. O trabalho também oferece proposições que mostram que o TP pode ser superado (em termos de utilidade) por outras carteiras. O segundo ensaio oferece uma derivação explícita da equivalência entre a carteira de variância mínima e a carteira de mínima volatilidade do *Tracking Error*. Esse resultado depende da fórmula de Sherman-Morrison . O ensaio também apresenta uma equivalência desses resultados com uma regressão restringida de Mínimos Quadrados. Além disso, o ensaio usa a primeira equivalência para achar uma carteira de *tracking* usando a abordagem de Kempf and Memmel (2006). O terceiro ensaio compara a performance de carteiras de investimento passivo com um número pequeno de ativos (carteiras esparsas de *index tracking* usando diferentes algoritmos de *feature selection*. Para isso, oferecemos um estudo empírico com duas bases de dados, uma americana e outra brasileira. Para isolarmos o efeito dos métodos de seleção, nós separamos a fase de seleção de ativos da fase de alocação de ativos. Na fase de alocação de ativos, seguimos Liu (2009) e selecionamos a carteira com a menor volatilidade de *Tracking Error*. Os métodos de seleção utilizados são o *backward stepwise selection*, o *forward stepwise selection* e o lasso. Nossos resultados mostram que, no caso brasileiro, a seleção pelo método lasso é o melhor *tracker*. No caso americano, o lasso apresenta melhor desempenho ajustado pelo risco, porém isto é devido a maiores retornos, não a menor volatilidade, o que não é desejável no nosso caso. Um destaque deste ensaio é que o melhor *tracker* para o caso americano é uma carteira que utiliza o algoritmo de iteração *backward* (métodos simples que recebem pouca atenção na literatura).

Palavras-chaves: Média-Variância. Otimização de Carteiras. *Index Tracking.* Avaliação de Desempenho.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

It is reasonable to imagine that investors prefer to elevate their wealth while minimizing the risks associated to any potential gain. To that end, professional wealth managers use portfolio optimization on investment decisions. This optimization is an important process of diversification in which the proportion of each financial asset kept on a portfolio is determined to maximize the return contingent to a desired risk level. The classic approach was introduced by Markowitz (1952, 1959), and perfected by Merton (1972). It is known as the Modern Portfolio Theory (MPT) and it aims to optimize the trade-off between risk and return of a portfolio. This mean-variance framework is the basis to the subsequent development of portfolio optimization.

Since the seminal contributions of Markowitz (1952, 1959), mean-variance portfolio optimization has become the cornerstone to several areas in quantitative finance, including asset allocation, portfolio management of stocks, bonds, and other assets. This framework defines the appropriate portfolio weights by solving an optimization problem. This problem presents itself in three main forms: the first is to maximize return given a level of risk; the second is to minimize risk given a level of return; and the third is to optimize an objective (or utility) function. Even though this approach presents some deficiencies, one of them being the sensibility of the optimal weights to the inputs (estimated means and variances), the mean-variance optimization still is the mainstay of the Modern Portfolio Theory. Hence, a solid understanding of this method allied with some intuition is essential to the comprehension and to the successful implementation of any quantitative investment strategies.

Before Markowitz's work, the risk return interaction was treated in a rather *ad hoc* manner, as Kolm, Tütüncü and Fabozzi (2014) point out. Since then, portfolio selection and optimization is a fundamental problem of finance, both from the academic point of view and from the market practitioner point of view. The MPT offers an answer to one of finance's most fundamental question: "How should an investor allocate his funds amongst all the possible assets"?

Underlying Markowitz's optimization problem there is a known economic problem: the utility maximization, which arises from the definition of Economics. Markowitz's portfolio theory presupposes that the investor has a concave utility function that expresses a risk-averse behavior. So, Markowitz formulates the portfolio selection with two criteria at its base: mean and variance. One of the basic assumptions in economics is that the economic agents want to allocate its resources in the most efficient way possible. Hence, it is also legitimate that finance supposes that the investors want to allocate wealth in the most efficient manner. They do this by trying to maximize return and minimize risk. Ergo, the portfolio selection problem is a genuine economic problem that deserves the researcher's attention.

However, it is prudent to highlight that the process of portfolio selection proposed by Markowitz has its pitfalls. For example, it presents a high sensibility to the uncertainty of its inputs. This is observed by Best and Grauer (1991), who show how changes in the mean asset returns affect the return, risk, and composition of a given portfolio. Chopra and Ziemba (1993) do something similar, they examine the relative impact on the portfolio of estimation errors in the mean, variance, and covariance of assets.

This thesis is inserted in the context of portfolio selection and optimization, it offers three different essays with this common theme. However, the essays differ in their approaches and subthemes. Because this thesis is comprised of three different essays with a common theme, it is inevitable to have some content overlap. Also, because we wanted to make the essays the most self contained as possible, we had to allow those redundancies in content throughout the text.

Finally, this thesis is comprised of five chapters. Beyond this introduction, there are three essays, one for each chapter, and the thesis finishes in Chapter 5, with some final considerations. The first essay offers a critique of the Tangency Portfolio (TP). [1] This critique is due to the widespread misconception that the TP is the most efficient mean-variance portfolio, as can be seen in DeMiguel, Garlappi and Uppal (2009) and Kempf and Memmel (2006) among others. The second essay provides an explicit derivation of the equivalence between the minimum variance portfolio and the minimum Tracking Error volatility portfolio. This equivalence is much talked about in intuitive terms but seldom demonstrated analytically, see Liu (2009) and DeMol (2016). The main contribution of this essay, besides the direct analytical derivation, is to present the connection among general portfolio optimization, index tracking portfolio optimization and OLS regressions. The third essay is about feature selection in the context of index tracking. In it, we make an empirical comparison of feature selection methods in the context of portfolio optimization. The study of such methods is important, in view of the observation made by Liu (2009) and Tu and Zhou (2011) that the majority of institutional managers are benchmarked by an external index like the SP500.

---

[1] A preliminary version of this essay was accepted to the XIX Econtro Brasileiro de Finanças with the title "Tangency Portfolio: A Critique"

2 TANGENCY PORTFOLIO: A CRITIQUE

In the present essay we will offer evidence as to why the Tangency Portfolio (TP) should not be used for empirical evaluations. We present an analytical study of the TP, where we offer some propositions about the ex-ante performance (in terms of utility) of the TP. Even though the propositions do not take uncertainty into account they reveal an important fact: the TP can be outperformed (in terms of utility) in sample by other portfolios. We also perform an empirical evaluation using the Kenneth French Data Library. This data was also used in DeMiguel, Garlappi and Uppal (2009), Kirby and Ostdiek (2012), but with a different time frame. Further, we use a Brazilian dataset of asset prices with daily frequency from January 1999 to December 2010 (2970 days).

Our results show that the TP rarely outperforms other portfolios on in and out of sample evaluations. This corroborates the work in Kirby and Ostdiek (2012). Further, our propositions show that the TP can be outperformed in sample. This underperformance is exacerbated in out of sample evaluations of the TP, as shown here and in other papers. Hence, it is not that the TP is optimal with a bad out of sample performance. The TP is not optimal in sample as well. To a lesser extent, our results also corroborates the work of Clarke, Silva and Thorley (2006) about the performance of the Global Minimum Variance Portfolio.

In Portfolio Theory, the Tangency Portfolio (TP) is of paramount importance and has several meaningful properties, as stated by Roll (1977), Brandt (2010), Constantinides and Malliaris (1995), and Best (2010). However, it presents a poor out of sample performance, as shown in Kirby and Ostdiek (2012), Kritzman, Page and Turkington (2010), and DeMiguel, Garlappi and Uppal (2009). Sometimes the TP is presented as the best mean-variance efficient portfolio as in Kempf and Memmel (2006), Britten-Jones (1999), and DeMiguel, Garlappi and Uppal (2009). This is inaccurate. The TP is optimal under the criterion of Sharpe Ratio optimality, not mean-variance utility optimality. For a portfolio to be optimal under a mean-variance utility, the optimization problem has to take into account the investor's risk-aversion, which the TP does not do. Furthermore, there are other alternatives for mean-variance efficient portfolios that present less uncertainty with more stable and less extreme weights, as Kirby and Ostdiek (2012) show.

Markowitz (1952) pioneered the research of the Modern Portfolio Theory by establishing rules to select portfolios. One of its greatest innovations was the ability to provide a quantitative method through which the diversification of investments could be performed. With this method, the investor was able to directly solve the risk-return trade off of a portfolio by maximizing the return given an acceptable level of risk, or minimizing the risk given a wanted return. Merton (1972) derived the efficient frontier of portfolio returns from Markowitz's work. He did it by creating a function that finds the optimal variance given an expected return of the combination of assets that form the optimal

portfolio.

Further problems with the Tangeny Portfolio are aired by Kirby and Ostdiek (2012) when they point out that: (a) if its denominator ($A = \mathbf{e}'\Sigma^{-1}\mu$) is small, then the TP can display extreme weights and hence extreme returns; (b) if the denominator is negative ($A < 0$), it reverses the direction of the allocation vector, which can bring disastrous consequences to the investor using such portfolio; (c) those undesirable values of A happen more often than we would like. So, we can see that the values of A play a large role in what the value of the TP will be.

With this in mind, Jobson and Korkie (1980) analyze the distribution of $\mathbf{e}'\Sigma^{-1}\mu$ and show that it is very volatile. This volatility directly affects the TP's weights, making them unstable and extreme, which raises the TP's turnover and, with it, its transaction costs. This high transaction costs worsen the TP's performance.

One of the components of A is μ. Several papers on parameter uncertainty like Merton (1980) and Chopra and Ziemba (1993) indict μ as the main culprit of the volatility of portfolio weights. They also point out that μ is one of the most volatile quantities to be estimated in Portfolio Theory.

Dealing with parameter and estimation uncertainty, Kan and Zhou (2007) derive rules for portfolio allocation that minimize estimation risk. In the beginning of their paper, Kan and Zhou (2007) derive the optimal portfolio allocation when the investor knows, with certainty, the vector of expected returns (μ) and the covariance matrix of returns (Σ), and find that the maximum utility is given when $\mathbf{w} = \dfrac{\Sigma^{-1}\mu}{\gamma}$. Later, they relax the assumption of knowing μ and Σ with certainty and find rules for optimal allocation when those quantities are not known with certainty. The derivation of those rules are based on loss and risk functions, which are widely used in decision analysis and Bayesian choice theory.

Kan and Zhou (2007) never mention the TP per se, because the it is not the solution of their maximization problem. But as Okhrin and Schmid (2006) point out, the TP does not have moments higher or equal to one. This means that even if we wanted, we could not compute the ex-post expected utility of the TP.

The rest of the essay is organized as follows. Section 2.1 presents the basics of portfolio selection and the portfolios that we will use throughout this essay. Section 2.2 offers an analytical study with some Propositions that show in what situations the TP is overperformed by other portfolios ex-ante. Section 2.3 describes our empirical evaluation of the TP against other portfolios. Section 2.4 shows the results of our empirical evaluation. Finally, Section 2.5 concludes.

2.1 PORTFOLIO SELECTION

In this section, we present the basics of portfolio selection and the portfolios that we will use throughout this essay. We observe T N-tuples of returns throughout time. We denote the vector of returns of those assets in time $t = 1, \ldots, T$ by:

$$R_t = [R_{t1}, R_{t2}, \ldots, R_{tN}]'.$$

This results in the following $T \times N$ matrix:

$$\mathbf{R} = (R_{ti})_{t=1,\ldots,T, i=1,\ldots,N} = \begin{bmatrix} R_{11} & R_{12} & \ldots & R_{1N} \\ R_{21} & R_{22} & \ldots & R_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ R_{T1} & R_{T2} & \ldots & R_{TN} \end{bmatrix}.$$

We assume that the returns follow a N-variate distribution with mean vector $\mu = E[R_t]$ and covariance matrix $\Sigma = \mathrm{Var}[R_t]$, denoted as follows:

$$\mathrm{E}(R_t) = \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix}, \qquad \mathrm{Var}(R_t) = \Sigma = (\sigma_{ij})_{i,j=1,\ldots,N} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \ldots & \sigma_{NN} \end{bmatrix}.$$

The covariance matrix of returns $\Sigma = \mathrm{Var}[R_t]$ has elements $\sigma_{ii} = \sigma_i^2$ and $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ (for $i \neq j$), where σ_i is the standard deviation of R_{it} and ρ_{ij} is the correlation coefficient of assets S_i and S_j (for $i \neq j$). Throughout this essay, we will assume that Σ is positive definite so it can be invertible.

Portfolios

A portfolio is represented by a N-dimensional vector $\mathbf{w}$,

$$\mathbf{w} = [w_1, w_2, \ldots, w_N]',$$

where w_i is the fraction of total wealth invested in asset i. The vector $\mathbf{w}$ can also be called the allocation vector. We use the following portfolios: Equally Weighted (EW), Optimal Unconstrained (OU), Tangency Portfolio (TP), Minimum Variance (MV), and Optimal Constrained (OC), with formulas:

$$\mathbf{w}_{\text{ew}} = \mathbf{e}/N, \tag{2.1}$$

$$\mathbf{w}_{\text{ou}} = \frac{\Sigma^{-1}\mu}{\gamma}, \tag{2.2}$$

$$\mathbf{w}_{\text{tp}} = \frac{\Sigma^{-1}\mu}{\mathbf{e}'\Sigma^{-1}\mu}, \tag{2.3}$$

$$\mathbf{w}_{\text{mv}} = \frac{\Sigma^{-1}\mathbf{e}}{\mathbf{e}'\Sigma^{-1}\mathbf{e}}, \tag{2.4}$$

$$\mathbf{w}_{\text{oc}} = \frac{C\mu_0 - A}{D}\Sigma^{-1}\mu + \frac{B - A\mu_0}{D}\Sigma^{-1}\mathbf{e}, \tag{2.5}$$

where A, B, C, and D are Merton's constants, which are equal to:

$$A = \mathbf{e}'\Sigma^{-1}\mu, \quad B = \mu'\Sigma^{-1}\mu, \quad C = \mathbf{e}'\Sigma^{-1}\mathbf{e}, \quad D = BC - A^2. \tag{2.6}$$

(MERTON, 1972, p. 1853) points out the fact that $B > 0$, $C > 0$ and $D > 0$, if Σ is Positive Definite. For more information about the above portfolios, the reader should consult Kirby and Ostdiek (2012), Brandt (2010), Constantinides and Malliaris (1995), Merton (1972).

Portfolio Returns

Holding the portfolio $\mathbf{w}_{\text{p,t}}$ for one period gives the out of sample return in $t+1$,

$$R_{p,t+1} = \mathbf{w}'_{\text{p,t}} R_{t+1}. \tag{2.7}$$

With this formula, we find portfolio's expected return μ_p, portfolio variance σ_p^2, and standard deviation σ_p:

$$\mu_p = \text{E}[R_p] = \text{E}[\mathbf{w}'R_t] = \mathbf{w}'\mu$$
$$\sigma_p^2 = \text{Var}[R_p] = \text{Var}[\mathbf{w}'R_t] = \mathbf{w}'\Sigma\mathbf{w}$$
$$\sigma_p = \sqrt{\mathbf{w}'\Sigma\mathbf{w}}.$$

Also, with those definitions, we can define the mean-variance utility function of portfolio $\mathbf{w}$, as:

$$U(\mathbf{w}) = \mu_p - \frac{\gamma}{2}\sigma_p^2 = \mathbf{w}'\mu - \frac{\gamma}{2}\mathbf{w}'\Sigma\mathbf{w}. \tag{2.8}$$

2.2 ANALYTICAL STUDY

In this section, we will present the ex-ante utilities of the following portfolios: Equally-Weighted Portfolio (EW), Optimal Unconstrained Portfolio (OU), Tangency

Portfolio (TP), Minimum Variance Portfolio (MV), and the Optimal Constrained Portfolio (OC). Next, we will compare the Utility of the TP with other portfolios to find under what conditions the TP has higher utility.

Even though the ideal comparison would be to proceed as Kan and Zhou (2007) and compute ex-post expected utilities for the TP, Okhrin and Schmid (2006) point out that the TP does not have moments higher or equal to one. Hence, finding the expected utility of the TP is an impossible task. Still, our propositions present an interesting result: In sample, the TP can be outperformed by other portfolios in terms of utility. Consequently, the TP might not be the best portfolio choice.

Using equation (2.8), we may find the utilities of the portfolios listed above. This results in:

$$U_{\text{ew}} = \frac{\mathbf{e}'\mu}{N} - \frac{\gamma}{2N^2}\mathbf{e}'\Sigma\mathbf{e}, \tag{2.9}$$

$$U_{\text{ou}} = \frac{B}{2\gamma}, \tag{2.10}$$

$$U_{\text{tp}} = \frac{B}{2A^2}(2A - \gamma), \tag{2.11}$$

$$U_{\text{mv}} = \frac{1}{2C}(2A - \gamma), \tag{2.12}$$

$$U_{\text{oc}} = \frac{-\gamma^2 + 2A\gamma + D}{2\gamma C}, \tag{2.13}$$

where $\mathbf{e}$ is a conformable vector of ones, and A, B, C, and D are Merton's constants, as shown in (2.6).

TP and OU

Comparing (2.11) with (2.10), we find:

$$U_{\text{tp}} \geq U_{\text{ou}}$$

$$\frac{B}{2A^2}(2A - \gamma) \geq \frac{B}{2\gamma}$$

$$\gamma(2A - \gamma) \geq A^2$$

$$-\gamma^2 + 2A\gamma - A^2 \geq 0$$

$$-(\gamma - A)^2 \geq 0$$

Unless $\gamma = A$, $U_{\text{tp}} \leq U_{\text{ou}}$. This leads to our first proposition:

Proposition 1: *Ex-ante, the utility of the OU is higher than the TP, except when $\gamma = A$, in which case $U_{ou} = U_{tp}$.*

TP and OC

Comparing (2.11) with (2.13), we find:

$$U_{\text{tp}} \geq U_{\text{oc}}$$

$$\frac{B}{2A^2}(2A - \gamma) \geq \frac{-\gamma^2 + 2A\gamma + D}{2\gamma C}$$

$$BC\gamma(2A - \gamma) \geq -\gamma^2 A^2 + 2A^3\gamma + A^2 D$$

$$-\gamma^2(BC - A^2) + \gamma A(2BC - 2A^2) - A^2 D \geq 0$$

$$-D\gamma^2 + 2AD\gamma - A^2 D \geq 0$$

D is always positive, so we can divide both sides by it.

$$-(\gamma^2 - 2A\gamma + A^2) \geq 0$$

$$-(\gamma - A)^2 \geq 0$$

Again, unless $\gamma = A$, $U_{tp} \leq U_{oc}$. This leads to our second proposition:

Proposition 2: *Ex-ante, the utility of the OC is higher than the TP, except when $\gamma = A$, in which case $U_{oc} = U_{tp}$.*

TP and MV

Comparing (2.11) with (2.12), we find:

$$U_{\text{mv}} \geq U_{\text{tp}}$$

$$\frac{1}{2C}(2A - \gamma) \geq \frac{B}{2A^2}(2A - \gamma)$$

$$A^2(2A - \gamma) \geq BC(2A - \gamma)$$

$$(A^2 - BC)(2A - \gamma) \geq 0$$

$$-D(2A - \gamma) \geq 0$$

D is always positive, so we can divide both sides by it.

$$-2A + \gamma \geq 0$$

$$\gamma \geq 2A$$

So for $\gamma \geq 2A$, $U_{\text{mv}} \geq U_{\text{tp}}$. Otherwise ($\gamma < 2A$), $U_{\text{tp}} > U_{\text{mv}}$. This leads to our third proposition:

Proposition 3: *The ex-ante utilities U_{mv} and U_{tp} depend on the values of γ with $U_{\text{mv}} > U_{\text{tp}}$ for $\gamma > 2A$ and $U_{\text{mv}} < U_{\text{tp}}$ for $\gamma < 2A$.*

The propositions above demonstrate that, even without uncertainty, the utility of the TP can be lower than the other portfolios. They also demonstrate that the utilities depend on what values we use for γ in relation to A. This is significant because the portfolio literature treats the TP as being the best mean-variance portfolio (KEMPF; MEMMEL, 2006; BRITTEN-JONES, 1999; DEMIGUEL; GARLAPPI; UPPAL, 2009). This is falsified by our propositions above.

Further comparisons of utility values for different values of γ can be seen in the Figure 2.1 below. In the figure, we can see that when $\gamma = A$, the OC, the OU, and the TP have the same utilities. For values of γ different from A, the TP has lower utility than the OU and the OC. The figures also show that the MV has lower utility than the TP when γ is lesser than $2A$, but the MV has higher utility when γ is greater than $2A$.

2.3   EMPIRICAL EVALUATION

In this section, we will do an empirical evaluation to test the performance of the Tangency Portfolio against alternative portfolios. This empirical evaluation consists in forming the optimal portfolios from Section 2.1 in pseudo real time and evaluating the out of sample performance of those portfolios against each other.

2.3.1   Data

The data for this empirical analysis consists of a total of three datasets. The first two are of monthly raw returns (not in excess of the risk-free asset) drawn from the Kenneth French library [1]. They are the 10 Industry portfolios (FF10) and the 25 Size and Book to Market portfolios (FF25). Versions of those datasets were used in DeMiguel, Garlappi and Uppal (2009) and Kirby and Ostdiek (2012). The sample period used here is from July 1963 to June 2018 ($T = 660$ months). The third dataset is a Brazilian one and it consists of daily closing prices for 61 stocks traded in the Bovespa. The sample period used here is from January 1999 to December 2010 (2970 days).

Rolling Windows Estimators

To estimate the parameters inputs, $\hat{\mu}$ and $\hat{\Sigma}$ used in the portfolio policies, we adopt a recursive estimation approach based on a *rolling window* time series of returns. With this approach, at each point in time $t = J, \ldots, T$, we use use the last J points of data and estimate:

---

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

$$\hat{\mu}_t = J^{-1} \sum_{j=0}^{J-1} R_{t-j}$$

$$\hat{\Sigma}_t = (J-1)^{-1} \sum_{j=0}^{J-1} (R_{t-j} - \hat{\mu}_t)(R_{t-j} - \hat{\mu}_t)'$$

This results in $T - J - 1$ estimates of expected return vectors and covariance matrices. For the FF10 and FF25 datasets, we use window length equal to $J = 120$ (10 years of monthly data). So, we have 479 estimates of covariance matrices and expected returns. For the Brazilian dataset, we use window length equal to $J = 250$ (approximately one year of daily data). So, we have 2719 estimates.

Allocation

Using the estimates above, we find the allocation vector using the policies of Section 2.1. They are the EW (equation (2.1)), TP (equation (2.3)), MV (equation (2.4)), and the OC (equation (2.5)). For this last portfolio, Kirby and Ostdiek (2012) set the target return, μ_0, equal to the return of the EW to let the aggressiveness of the mean-variance portfolio comparable to it. In this essay however, we use 1.1 times the mean return of the MV as our target return. This happens because, in our data, the mean return of the EW is smaller than the mean return of the MV , consequently, it is not efficient to use μ_{ew} as a target return. Also, in this way, we always have $\mu_0 > A/C$. By the end of this process, we generate $T - J - 1$ allocation vectors for each policy, in each dataset.

2.3.2   Evaluation Measures

What we will evaluate are the out of sample returns of the portfolios. Using the formula (2.7), we generate the time series $\{R_{p,t}\}_{t=J+1}^{T}$ of out of sample returns for each policy in each dataset. This time series is the subject of our analysis. Since the risk-return tradeoff is at the core of the Portfolio Theory, any evaluation of portfolio performance has to take this tradeoff into account. We do this by computing two measures: the Sharpe Ratio and the certainty equivalent (CEQ), as defined below.

Sharpe Ratio (SR)

The first measure is the *Sharpe Ratio* (SR), that measures the amount of additional return for each unit of risk incurred by the portfolio. The SR has the following formula:

$$\hat{SR}_p = \frac{\hat{\mu}_p}{\hat{\sigma}_p}.$$

Where $\hat{\mu}_p$ is the average of portfolio returns, and $\hat{\sigma}_p$ is the standard deviation (volatility) of portfolio returns, calculated as follows:

$$\hat{\mu}_p = (T - J - 1)^{-1} \sum_{t=J+1}^{T} R_{pt},$$

$$\hat{\sigma}_p = \sqrt{(T - J - 2)^{-1} \sum_{t=J+1}^{T} (R_{pt} - \hat{\mu}_p)^2},$$

where R_{pt} is the portfolio return calculated with formula (2.7).

However, the SR does not take into account the investor's specific risk aversion, nor the amount of leverage the portfolio has. Hence, we use alternative measures to mitigate this problem.

Certainty Equivalent (CEQ)

The second measure that we use is the *certainty equivalent* (CEQ) *return.* This measure expresses the risk-free rate (return without risk) that the investor is willing to accept in lieu of a particular risky portfolio strategy. It is also a measure of the investor's quadratic utility.

$$\hat{CEQ}_p = \hat{U}_p = \hat{\mu}_p - \frac{\gamma}{2} \hat{\sigma}_p^2.$$

Here, we used $\gamma = 1$ and $\gamma = 5$ to evaluate the portfolios' performance.

Transaction Costs

To get a sense of the amount of trading required to implement each portfolio strategy, we compute the portfolio turnover. As Liu (2009) points out, the turnover can be understood in the following way: in the moment prior to rebalancing, each dollar invested in asset S_i in the portfolio has changed its value from $w_{i,t}$ to $w_{i,t}(1 + R_{i,t+1})$. So, in general, each dollar invested in the whole portfolio has changed from w_t to $\mathbf{w}_t(1 + R_{t+1})$, where $w_{i,t}$ is the i-th element of $\hat{w}_t$ and the i-th element of R_t is denoted as $R_{i,t}$. Therefore, prior to rebalancing, the weight on asset S_i has changed from $w_{i,t}$ to $w_{i,t}\dfrac{1 + R_{i,t+1}}{1 + \mathbf{w}_t' R_{t+1}}$, meaning the turnover rate in time t after rebalancing the portfolio can be defined as:

$$\text{TO}_t = \sum_{i=1}^{N} \left| w_{i,t+1} - w_{i,t}\frac{1 + R_{i,t+1}}{1 + R_{p,t}} \right|.$$

We report the average of this rate throughout time.

Associated with the turnover there is the associated transaction costs of rebalancing a portfolio. If the rebalancing is too frequent or too extreme, it can lead to overtrading. According to Barber and Odean (2000), overtrading has negative effects on portfolio

performance due to its associated transaction costs. To take into account the impact of such costs we also evaluate the returns net of transaction costs, which have formula:

$$R_{p,t}^{net} = (1 + R_{p,t})(1 - c \times \text{TO}),$$

where c is the fee that must be paid for each transaction. We use $c = 50$bps like in DeMiguel, Garlappi and Uppal (2009) and Kirby and Ostdiek (2012).

Implied Risk Aversion

Bodnar, Parolya and Schmid (2013) explore the relation between μ_0 (the target return) and γ (the investor's risk aversion). They show for what values of γ those values are the same, and the answer is:

$$\gamma = \begin{cases} \dfrac{s}{\mu_0 - R_{\text{gmv}}}, & \text{for } \mu_0 > R_{\text{gmv}} \\ \infty, & \text{for } \mu_0 = R_{\text{gmv}}, \end{cases}$$

where,

$$R_{\text{gmv}} = \frac{A}{C}, \quad s = \mu' Q \mu, \quad Q = \Sigma^{-1} - \frac{\Sigma^{-1} \mathbf{e} \mathbf{e}' \Sigma^{-1}}{C}.$$

If $\mu_0 < R_{\text{gmv}} = A/C$, the solution does not lie on the efficient frontier. Representing γ in terms of Merton's constants:

$$\gamma = \frac{D}{C\mu_0 - A}, \text{ for } \mu_0 > \frac{A}{C}, \tag{2.14}$$

$$\mu_0 = \frac{A}{C} + \frac{D}{\gamma C}. \tag{2.15}$$

With equations (2.14) and (2.15), we can find an implied risk aversion of the investor who seeks target return μ_0.

Weight Statistics

To gauge the amount of leverage in each portfolio, we report their *Gross Exposure* and their *Short Interest*. The Gross Exposure is expressed in proportion of the notional value of the portfolio (e.g. when the Gross Exposure is equal to one, it means that there is no leverage; but a Gross Exposure of 1.1 means that the portfolio is leveraged with 10% of its notional value). The Short Interest expresses the size of the short positions (weights smaller than zero) in the portfolio, it is expressed in percentage (e.g. when the Short Interest is equal to zero, it means that there is no short-selling happening; but when the Short Interest is equal to 10%, it means that the sum of of short positions are 10%). These measures have the following formulas:

$$\text{Gross}_t = \sum_{i=1}^{N} |w_{i,t}|$$

$$\text{Short}_t = \frac{\sum_{i=1}^{N} |w_{i,t}| - 1}{2}.$$

Hypothesis Testing

To verify if the Sharpe Ratio and the CEQ of each policy are statiscally different from the benchmark (the TP), we use the framework laid out in Ledoit and Wolf (2008, 2011, 2018). Their framework relies on the fact that those performance measures are smooth functions of population moments. We use the circular block bootstrap of Politis and Romano (1992) with block length $2, 4, 6, 8$, and 10 to find p-values of the differences in performance measures.

2.4   RESULTS

In this section we present the results of our empirical analysis. We do not annualize any performance measure due to the problem of time aggregation pointed out in Lo (2002). So the performance measures are in monthly terms for the FF10 and FF25 datasets, and in daily terms for the Brazilian dataset.

2.4.1   In Sample Results

Table 2.1 presents the in sample results for the Tangency Portfolio (TP), Minimum Variance Portfolio (MV), Optimal Constrained Portfolio (OC), and the Equally Weighted Portfolio (EW) for the FF10 (Panel A), the FF25 (Panel B), and for the Brazilian (Panel C) datasets. The statistics shown are: the mean of the portfolio return (μ), the standard deviation of the portfolio return (σ), the Sharpe Ratio of the portfolio returns (SR), the quadratic utility of the portfolio returns for $\gamma = 1$ and $\gamma = 10$ (CEQ 1 and CEQ 10), and the in sample implied risk-aversion of the portfolios (γ). The performance measures are in monthly terms for the FF10 and FF25 datasets, and in daily terms for the Brazilian dataset. The sample size is 479 for the FF10 and FF25 dataset and 2719 for the Brazilian dataset.

The in sample results can be summarized as follows: the TP strategy has the biggest SR in all datasets, but it does not have bigger CEQs, as in the case of the FF10 dataset. This fact can be explained by our propositions. The results also show the high values of gross exposure and short interest of the TP strategy.

2.4.2 Out of Sample Results

Table 2.2 presents out of sample results for the Tangency Portfolio (TP), Minimum Variance Portfolio (MV), Optimal Constrained Portfolio (OC), and the Equally Weighted Portfolio (EW) for the FF10 (Panel A), FF25 (Panel B), and Brazilian (Panel C) datasets. The statistics shown are the same as in Table 2.1 with the exclusion of γ, and the inclusion of the average turnover of the portfolio (TO). Also, the statistics to the right of the TO column are for the returns net of transaction costs with 50bp. The performance measures are in monthly terms for the FF10 and FF25 datasets, and in daily terms for the Brazilian dataset. The sample size is 479 for the FF10 and FF25 dataset and 2719 for the Brazilian dataset.

Tables 2.3, 2.4, and 2.5 show the p-values of the aforementioned hypothesis tests of the SR, CEQ 1, and CEQ 10 statistics of the portfolios returns of each strategy. Tables 2.6, 2.7, and 2.8 show the p-values of the same statistics but it uses the net portfolio returns. The benchmark of all tables is the TP strategy.

2.4.2.1 Without Transaction Costs

Without transaction costs the TP's SR is the lowest of all datasets (0.10, 0.18, -0.02 for the FF10, FF25, and Brazilian datasets), meanwhile the MV (0.29, 0.37, 0.10) and OC (0.28, 0.38, 0.10) have the highest SRs. To verify the difference in the SRs, we consult Table 2.3, in it we can see that the p-values hover around 5% for the MV and the OC for the FF10 and FF25 datasets and 2% for the Brazilian dataset. Such a feature does not happen to the EW strategy.

In terms of the CEQs, the results are not so clear. Without the transactions costs, in the FF25 dataset, the TP is tied with the OC in terms of the CEQ 1 statistic with a value of 1.39. This kind of result does not happen in the FF10 and the Brazilian dataset where the TP has the lowest CEQ 1. In terms of CEQ 10, maybe because of the higher risk aversion, the TP has the lowest values. However, it is important to point out that the results are not statistically significant for any dataset or block length that we might choose (see Tables 2.4 and 2.5).

2.4.2.2 With Transaction Costs

With transaction costs, the TPs SR and CEQ has the worst statistics with negative values for all of them. Negative value for a statistic only happens again for the specific case of the EW strategy in the Brazilian data for the CEQ 10 statistic. Also, the difference of the statistics with the other strategies increase drastically.

Table 2.6 reports the statistical significance for the SR, with transaction costs. In this table we can see that for the Brazilian data, the difference has significance below the

level of 1% for the MV and OC strategies for all block lenghts. The EW strategy has significance below the 5% level for all block lenghts. The FF10 and FF25 datasets do not have such clear results. The FF10 dataset has significance below 10% for all block lengths for the MV strategy, and below 15% for the OC strategy with some block lenghts registering levels below 10%. The FF25 has the OC strategy with statistical levels below 5% for all block lengths and the MV strategy below 10% with some block lenghts below 5%.

The statistical significance for the CEQs with transaction costs are reported in Tables 2.7 and 2.8. None of the results is below 30%. It is a noteworthy fact that such differences are not statistically significant, this has to be due to the variance of the TP's CEQs.

2.5 CONCLUDING REMARKS

In our empirical evaluation, we showed that in the FF10 dataset, the MV portfolio wins in the out of sample evaluation and the TP almost always come at the bottom. In the FF25 dataset, the OC portfolio wins in the out of sample evaluation, and the TP almost always come at the bottom. We also observe that, although the TP has the higher SR in sample, this might not be true for the CEQ. We also observe that almost certainly this is not true out of sample with and without transaction costs. The transaction costs greatly degrade the TP's performance, because it has the highest turnover of the portfolios.

The poor out of sample performance of the TP is already common sense in the portfolio selection literature, what does not seem to be common sense is the results shown in the Section 2.2. Those results show that the TP might be overperformed in sample by other portfolios, they also show in what occasions they overperform the TP. Through those propositions and the empirical illustration, we wish to show that it might be undesirable to use the TP to practical ends. The TP's weights not only present instability and extreme weights, but they also might not be optimal in sample. In its stead we should use alternative portfolios like the MV and the OC, depending on the investor's risk aversion.

Table 2.1 – In Sample Results

	Mean	SD	SR	CEQ 1	CEQ 10	Gamma	Gross	Short
Panel A: FF10 Dataset								
TP	1.09	3.63	0.30	1.03	0.44	8.31	1.98	49.23
MV	1.02	3.50	0.29	0.96	0.41	*Inf*	1.67	33.32
OC	1.12	3.74	0.30	1.05	0.42	6.01	2.15	57.27
EW	1.02	4.23	0.24	0.93	0.13	147.48	1.00	0.00
Panel B: FF25 Dataset								
TP	2.63	4.84	0.54	2.51	1.46	11.24	10.80	489.77
MV	1.31	3.41	0.38	1.25	0.73	*Inf*	6.15	257.63
OC	1.44	3.43	0.42	1.38	0.85	113.39	6.28	263.88
EW	1.20	5.05	0.24	1.08	−0.07	−141.46	1.00	0.00
Panel C: IBOV Dataset								
TP	0.21	1.25	0.17	0.20	0.13	13.44	3.38	119.11
MV	0.11	0.90	0.12	0.10	0.07	*Inf*	1.96	47.89
OC	0.12	0.90	0.13	0.12	0.08	125.34	1.98	49.15
EW	0.10	1.51	0.07	0.09	−0.01	−229.44	1.00	0.00

Source: The author.

Table 2.2 – Out of Sample Results

	Mean	SD	SR	CEQ 1	CEQ 10	TO	Gross	Short	SR	CEQ 1	CEQ 10
Panel A: FF10 Dataset											
TP	0.72	7.48	0.10	0.44	−2.08	146.56	3.92	145.86	−0.01	−0.98	−8.83
MV	1.04	3.57	0.29	0.98	0.41	16.16	2.34	67.00	0.27	0.90	0.32
OC	1.03	3.62	0.28	0.96	0.37	18.63	2.36	67.81	0.26	0.87	0.28
EW	1.02	4.23	0.24	0.93	0.13	2.46	1.00	0.00	0.24	0.92	0.12
Panel B: FF25 Dataset											
TP	2.04	11.44	0.18	1.39	−4.50	1276.13	20.05	952.41	−0.05	−163.02	−1555.59
MV	1.40	3.82	0.37	1.33	0.67	79.32	7.48	323.83	0.26	0.92	0.27
OC	1.46	3.83	0.38	1.39	0.73	80.77	7.65	332.50	0.27	0.98	0.32
EW	1.20	5.05	0.24	1.08	−0.07	1.89	1.00	0.00	0.24	1.07	−0.08
Panel C: IBOV Dataset											
TP	−3.16	187.97	−0.02	−179.83	−1769.79	52.45	24.32	1165.84	−0.02	−180.12	−1770.33
MV	0.10	1.00	0.10	0.10	0.05	4.34	2.32	66.20	0.08	0.08	0.03
OC	0.10	1.00	0.10	0.10	0.05	4.35	2.32	66.18	0.08	0.08	0.03
EW	0.10	1.50	0.06	0.08	−0.02	0.41	1.00	0.00	0.06	0.08	−0.02

Source: The author.

Table 2.3 – p-values for Difference of SR of Portfolio Returns

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: FF10 Dataset						
MV	19.58	0.82	1.30	2.26	3.02	3.70
OC	18.82	1.58	2.76	3.62	5.20	6.84
EW	14.60	4.88	6.18	6.72	8.10	8.74
Panel B: FF25 Dataset						
MV	18.77	2.82	3.98	4.12	4.64	5.76
OC	20.34	2.30	2.72	3.00	3.90	3.54
EW	5.96	45.72	46.94	46.22	50.20	49.30
Panel C: IBOV Dataset						
MV	11.96	1.18	1.56	1.54	1.76	1.52
OC	12.02	1.56	1.82	1.56	1.48	1.62
EW	8.07	7.44	8.68	8.10	8.04	8.74

Source: The author.

Table 2.4 – p-values for Difference of CEQ 1 of Portfolio Returns

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: FF10 Dataset						
MV	0.54	17.58	20.36	20.88	22.24	22.08
OC	0.52	21.52	22.20	23.04	24.56	24.76
EW	0.49	18.92	20.12	21.16	22.56	22.00
Panel B: FF25 Dataset						
MV	-0.06	91.34	91.52	91.48	90.32	91.64
OC	0.00	99.84	99.54	99.66	99.68	99.90
EW	-0.31	60.14	60.84	60.80	63.18	61.42
Panel C: IBOV Dataset						
MV	179.92	44.54	44.88	44.48	44.68	44.96
OC	179.92	44.78	45.20	44.24	45.92	45.32
EW	179.91	45.42	44.50	45.28	44.42	44.66

Source: The author.

Table 2.5 – p-values for Difference of CEQ 10 of Portfolio Returns

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: FF10 Dataset						
MV	2.48	18.06	19.00	21.12	24.98	26.66
OC	2.45	18.72	18.98	23.14	25.24	26.82
EW	2.21	20.18	21.82	23.90	28.42	29.32
Panel B: FF25 Dataset						
MV	5.17	14.00	14.74	16.86	21.58	23.28
OC	5.23	14.54	14.12	17.44	20.56	23.08
EW	4.43	16.84	16.38	19.00	21.86	23.42
Panel C: IBOV Dataset						
MV	1769.85	44.50	44.96	43.88	45.32	45.46
OC	1769.85	43.90	45.78	44.86	46.30	44.38
EW	1769.78	43.78	44.82	44.14	43.76	45.32

Source: The author.

Table 2.6 – *p*-values for Difference of SR of Net Portfolio Returns

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: FF10 Dataset						
MV	27.76	3.28	4.40	7.18	7.38	8.58
OC	26.68	6.70	9.16	11.42	13.22	14.56
EW	24.77	2.70	3.80	5.12	5.96	5.82
Panel B: FF25 Dataset						
MV	30.80	2.12	3.98	5.30	5.94	7.28
OC	32.21	1.40	2.78	3.96	5.24	4.38
EW	28.35	13.76	14.14	17.54	18.46	19.84
Panel C: IBOV Dataset						
MV	9.90	0.48	0.64	0.68	0.70	0.84
OC	9.95	0.76	0.84	0.74	0.74	0.78
EW	8.07	2.88	3.46	3.24	2.86	3.14

Source: The author.

Table 2.7 – *p*-values for Difference of CEQ 1 of Net Portfolio Returns

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: FF10 Dataset						
MV	1.88	37.14	37.92	39.04	38.94	39.80
OC	1.85	37.86	38.54	38.02	37.60	39.04
EW	1.91	38.38	38.30	38.26	38.80	38.16
Panel B: FF25 Dataset						
MV	163.95	44.78	44.36	44.42	43.76	45.52
OC	164.00	44.62	43.72	45.48	43.72	44.46
EW	164.09	44.40	44.84	44.04	44.68	44.50
Panel C: IBOV Dataset						
MV	180.19	44.54	44.88	44.48	44.68	44.96
OC	180.19	44.78	45.20	44.24	45.92	45.32
EW	180.20	45.42	44.50	45.28	44.42	44.66

Source: The author.

Table 2.8 – *p*-values for Difference of CEQ 10 of Net Portfolio Returns

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: FF10 Dataset						
MV	9.15	39.52	39.34	39.18	39.72	38.04
OC	9.11	40.32	39.20	41.08	39.94	38.80
EW	8.95	40.36	41.02	40.20	41.20	40.16
Panel B: FF25 Dataset						
MV	1555.85	44.22	44.42	44.78	45.48	44.70
OC	1555.91	45.32	45.38	44.96	44.14	43.70
EW	1555.50	45.06	44.70	45.38	44.30	43.74
Panel C: IBOV Dataset						
MV	1770.36	44.50	44.96	43.88	45.32	45.46
OC	1770.36	43.90	45.78	44.86	46.30	44.38
EW	1770.31	43.78	44.82	44.14	43.76	45.32

Source: The author.

Figure 2.1 – Portfolio Utility for Different Risk-Aversions of the FF10 Data

(a) FF10 Data
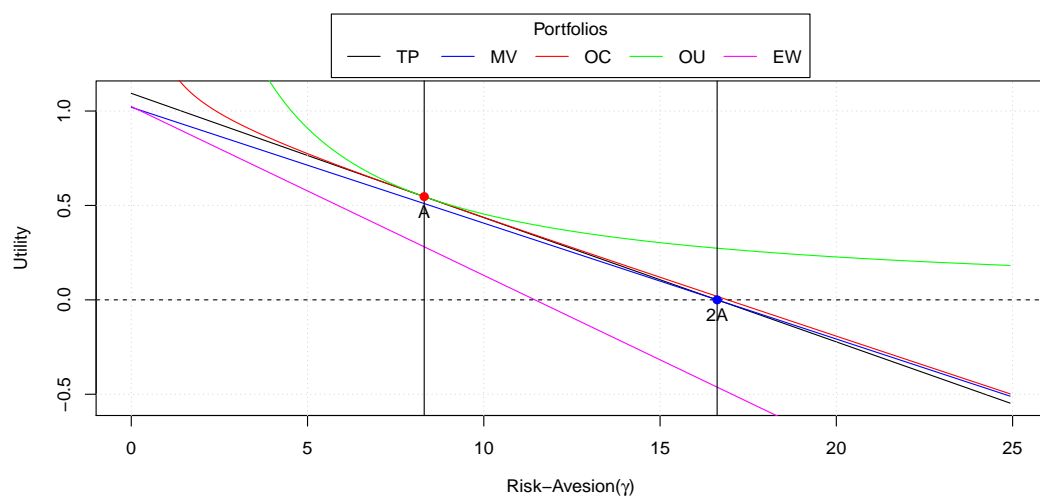
(b) FF25 Data
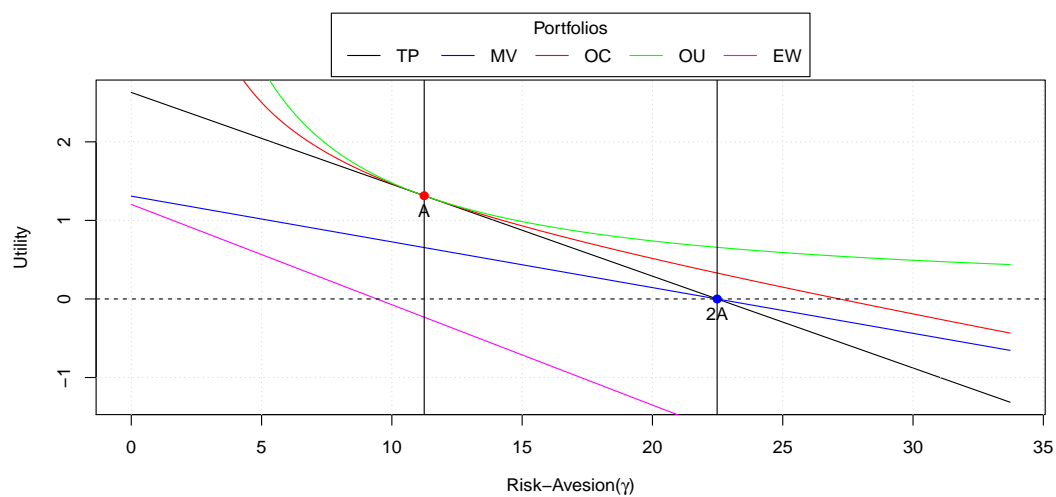
(c) Brazilian Data
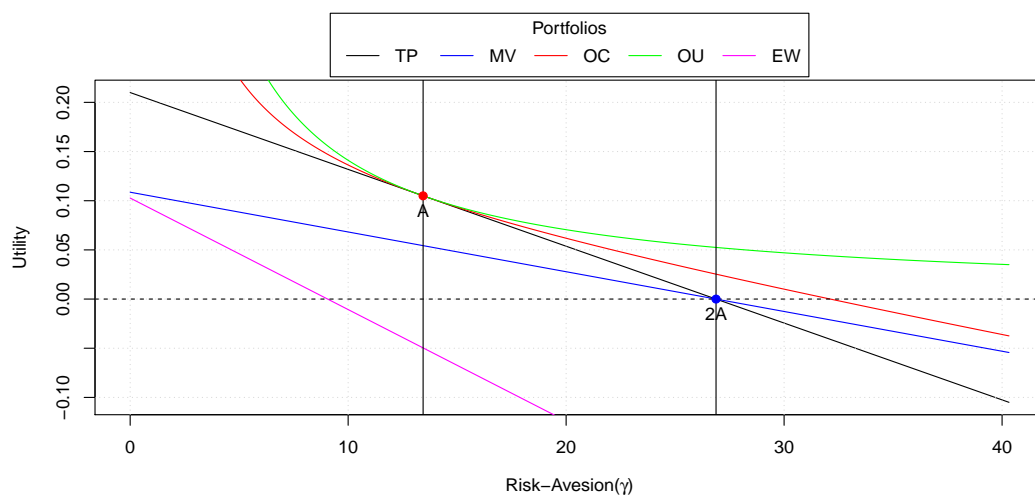
Source: The author.

3 ON THE EQUIVALENCE OF MINIMUM TRACKING ERROR AND MINIMUM VARIANCE PORTFOLIOS

Minimum variance (MV) and minimum Tracking Error (MT) optimal portfolios of excess returns on a benchmark are among the most used portfolio policies considered by market practitioners and academics; see, for instance, Chan, Karceski and Lakonishok (1999), Liu (2009), Brodie et al. (2009) and DeMol (2016). In short, the MV policy consists in finding optimal portfolios that minimize the portfolio variance, whereas the MT policy finds portfolios that minimize the variance of the difference with respect to a given benchmark.

Existing literature often refers to these two portfolio policies as interchangeable, that is, they are treated as being equivalent. However, to the best of our knowledge there is no single reference that establishes an analytical derivation to find this equivalence between the two portfolio problems. We, on the other hand, provide this derivation by using the **Sherman-Morrison** formula (PETERSEN; PEDERSEN, 2012).

Furthermore, we provide a formal connection of our analytical result with the OLS regression of the returns on a benchmark against the returns of risky assets. Also, we take advantage of the equivalence between MV and MT portfolios to use the approach used in Kempf and Memmel (2006) to find a Global Minimum Variance Portfolio (GMVP) portfolio by means of a regression.

The essay is organized as follows. Section 3.1 describes the MV and MT portfolio policies and provides an analytical expression for the equivalence between the two. Section 3.2 provides an equivalence result to OLS regression and an application to the result of Kempf and Memmel (2006). Section 3.3 concludes.

3.1 TWO OF A KIND: MINIMMUM TRACKING ERROR AND MINIMUM VARIANCE PORTFOLIOS

In this section, we establish the equivalence between two commonly used problems in the portfolio selection literature: the minimum Tracking Error and minimum variance portfolios of excess returns on a benchmark index. For that purpose, let us define the excess returns of the risky assets on the benchmark as:

$$R_t^e = R_t - ey_t,$$

where R is the $T \times K$ matrix of asset returns, y is the T-dimensional vector of returns of the benchmark index, and e is a conformable vector of ones. The unconditional variance of the excess returns is:

$$V(r) = \Omega = V(X_t) + V(ey_t) - 2cov(X_t, ey_t)$$
$$V(r) = \Sigma_{XX} + \sigma_{yy}ee' - 2\sigma_{Xy}e',$$

whereas the portfolio variance is:

$$V(r_p) = w'V(r)w = w'\Omega w$$
$$= w'(\Sigma_{XX} + \sigma_y^2 ee' - 2\sigma_{Xy}e')w$$
$$= w'\Sigma_{XX}w + \sigma_y^2 w'ee'w - 2w'\sigma_{Xy}e'w,$$

where w is the vector of portfolio weights.

Remark: Even though the classic optimization problem considers the risk-free asset as a benchmark, Tu and Zhou (2011) argue that portfolio managers are benchmarked by a market index, like the S&P500. In that case, the return on the market index can be viewed as the riskless asset to apply the same framework. It is also important to note that in the case where the investor uses the GMVP with returns in excess of the risk-free asset, he is tracking the risk-free asset.

We now turn our attention to formulation of the minimum variance (MV) portfolio problem. The MV problem is defined as:

$$\operatorname*{Min}_{w} \quad w'\Sigma_{XX}w + \sigma_y^2 w'ee'w - 2w'\sigma_{Xy}e'w \tag{3.1}$$
$$s.t. \quad e'w = 1.$$

And has FOC:

$$2\Sigma_{XX}w - 2(\sigma_{Xy}e' + e\sigma'_{Xy})w + 2\sigma_y ee'w = \lambda e$$
$$e'w = 1.$$

On the other hand, the minimum Tracking Error (MT) problem consists in minimizing the variance of the difference between the portfolio $R_{p,t} = w'X_t$ and the benchmark, i.e.:

$$Var(R_{p,t} - y_t) = Var(w'X_t) + Var(y) - 2Cov(w'X_t, y_t)$$
$$= w'\Sigma_{XX}w + \sigma_y^2 - 2w'\sigma_{Xy}.$$

Therefore, the MT problem can be stated as:

$$\operatorname*{Min}_{w} \quad w'\Sigma_{XX}w + \sigma_y^2 - 2w'\sigma_{Xy} \tag{3.2}$$
$$s.t. \quad e'w = 1.$$

And has FOC:

$$2\Sigma_{XX}w - 2\sigma_{Xy} = \lambda e$$
$$e'w = 1.$$

It is widely known in the literature (CHAN; KARCESKI; LAKONISHOK, 1999; LIU, 2009; BRODIE et al., 2009; DEMOL, 2016) that the problems (3.1) and (3.2) are equivalent once they are both minimization of a squared sum. However, they do not have the same objective function, neither the same FOC. Here, we will show that it is only because of the **Sherman-Morrison** formula (PETERSEN; PEDERSEN, 2012) that they will have the same solution. Proposition 1 states the equivalence between the results of the optimization problems in (3.1) and (3.2).

Proposition 1: *The minimum variance portfolio and minimum Tracking Error portfolios obtained upon solving optimization problems in* (3.1) *and* (3.2), *respectively, are equivalent.*

Proof:

The **Sherman-Morrison** formula (PETERSEN; PEDERSEN, 2012) states that:

$$(A + uv')^{-1} = A^{-1} - \frac{A^{-1}uv'A^{-1}}{1 + v'A^{-1}u}$$

where A is a $N \times N$ matrix and u, v are column vectors of dimension N.

The solution to the MV problem in (3.1) is:

$$w^*_{MV} = \frac{A^{-1}e}{e'A^{-1}e}[1 - (e'A^{-1}\sigma_{Xy})] + A^{-1}\sigma_{Xy},$$

where $A = \Sigma_{XX} - e\sigma'_{Xy}$, and we find $A^{-1} = \Sigma_{XX}^{-1} + \frac{\Sigma_{XX}^{-1}e\sigma'_{Xy}\Sigma_{XX}^{-1}}{1 - \sigma'_{Xy}\Sigma_{XX}^{-1}e}$ using the **Sherman-Morrison** formula (PETERSEN; PEDERSEN, 2012). After rearranging terms, we arrive at:

$$w^*_{MV} = w_{ou} + w_{mv}(1 - e'w_{ou}), \tag{3.3}$$

where $w_{ou} = \Sigma_{XX}^{-1}\sigma_{Xy}$ and $w_{mv} = \frac{\Sigma_{XX}^{-1}e}{e'\Sigma_{XX}^{-1}e}$. w_{ou} is the vector of weights of the optimal mean-variance unconstrained portfolio whereas w_{mv} is the vector weights of the minimum variance portfolio of the K assets.

The solution to the MT problem in (3.2) is:

$$w^*_{MT} = \Sigma_{XX}^{-1}\sigma_{Xy} + (1 - e'\Sigma_{XX}^{-1}\sigma_{Xy})\frac{\Sigma_{XX}^{-1}e}{e'\Sigma_{XX}^{-1}e} \tag{3.4}$$

$$= w_{ou} + (1 - e'w_{ou})w_{mv}, \tag{3.5}$$

which is equivalent to the solution of the MV portfolios in (3.3). ∎

3.2 REGRESSION BASED PORTFOLIOS

Another way to find an index-tracking portfolio is to regress the benchmark returns against the risky assets returns. This problem is a regression with stochastic regressors,

more details about this problem can be found in Rao et al. (2008) and Rencher and Schaalje (2007). With this in mind, we now consider a regression-based approach to the portfolio selection problem. To do this we define the following regression with intercept α, slope β and constrain $e'\beta = 1$:

$$\underset{\beta}{Min} \quad (y_t - \alpha - \beta X_t)'(y_t - \alpha - \beta X_t) \quad \text{s.t.} \quad e'\beta = 1. \tag{3.6}$$

The solution to (3.6) is:

$$\beta^* = \beta_{ols} + (1 - \beta_{ols})\frac{(X'X)^1 e}{e'(X'X)^{-1}e},$$

where $\beta_{ols} = (X'X)^{-1}X'y$. If we use centered quantities of X and y, we arrive at:

$$\beta^* = \Sigma_{XX}^{-1}\sigma_{Xy} + (1 - e'\Sigma_{XX}^{-1}\sigma_{Xy})\frac{\Sigma_{XX}^{-1}e}{e'\Sigma_{XX}^{-1}e}$$

$$= w_{ou} + (1 - e'w_{ou})w_{mv},$$

which is equivalent to the result obtained in Proposition 1. Therefore, the equivalence of the minimum variance and minimum Tracking Error portfolios can also be established by means of a regression-based approach.

3.2.1 Kempf and Memmel (2006)

Because the index-tracking portfolio is equivalent to the GMVP, we can use the approach presented in Kempf and Memmel (2006) to determine the weights w_{MV}. In their original approach, they rewrite the weights of the GMVP as regression coefficients.

First, they set up the problem by choosing the return of stock N to be the dependent variable:

$$r_{tN} = \alpha + \beta_1(r_{tN} - r_{t1}) + \cdots + \beta_{N-1}(r_{tN} - r_{t,N-1}) + \varepsilon_t, \quad t = 1, \ldots, T. \tag{3.7}$$

Then, to solve this problem, they define $\beta^{ex}, w_{MW}^{ex}, r_t^{ex}$ as $(N - 1) \times 1$ vectors. The superscript ex indicates that the vector has no entry for asset N. Further, they define Ω as the covariance matrix of regressors of equation (3.7):

$$\Omega = var(er_{t,N} - r_t^{ex})$$

The regression coefficients β^{ex} are the standardized covariances of the regressors and the dependent variable:

$$\beta^{ex} = \Omega^{-1}cov(er_{tN} - r_t^{ex}, r_{tN}) \tag{3.8}$$

Now, we have to show that the weights w_{MV}^{ex} of the GMVP correspond to the regression coefficient β^{ex}. We can compute the weight $w_{MV,N}$ as $1 - e'w_{MV}^{ex}$.

To prove $\beta^{ex} = w_{MV}^{ex}$, they consider an arbitrary portfolio P. Its return is determined by the weght vector $w_{MV}^{ex} = (w_{P,1}, \ldots, w_{P,K})'$ and the stock returns r_t^{ex} and $r_{t,N}$:

$$r_{tP} = w_P^{ex\prime} r_t^{ex} + (1 - w_P^{ex\prime} e) r_{tN}$$
$$= r_{tN} + w_P^{ex\prime} (r_t^{ex} - e r_{tN})$$

The return variance of this arbitrary portfolio P

$$\sigma_P^2 = \sigma_N^2 + w_P^{ex\prime} \Omega w_P^{ex} - 2 w_p^{ex\prime} cov(r_{tN} e - r^{ex}, r_{tN}) \tag{3.9}$$

is a function of the weights w_P^{ex}. To find the weights of the GMVP, we minimize equation (3.9) with relation to portolio weights w_P^{ex}. This miminization leads to:

$$2 \Omega w_P^{ex} - 2 cov(r_{tN} e - r_t^{ex}, r_{tN}) = 0$$
$$\Omega w_P^{ex} = cov(r_{tN} e - r_t^{ex}, r_{tN})$$

and they find:

$$w_P^{ex} = \Omega^{-1} cov(r_{tN} e - r_t^{ex}, r_{tN}) \tag{3.10}$$

The weights (3.10) correspond to the regression coefficients (3.8).

If we substitute the returns of asset N at time t by the excess returns of asset N on the benchmark y at time t, the solution will be a tracking portfolio and the results of Kempf and Memmel (2006) will hold for our portfolio.

3.3 FINAL REMARKS

In this essay, we have shown the equivalence of a Tracking Error portfolio and a minimum variance portfolio of excess returns on a benchmark. Here, we provided an analytical derivation of the solutions to the problems above. The solution is a 2-fund portfolio, very common in portfolio theory. The solutions are related both to OLS regression with stochastic regressors and the Quadratic Utility formulation of portfolio optimization.

4 INDEX TRACKING WITH FEATURE SELECTION

In this essay, we compare the performance of different feature selection algorithms applied to index tracking and portfolio optimization. More specifically, we study the case where an investor has to track an Index and he rebalances his portfolio monthly, bimonthly and quarterly (20, 40, and 60 days rebalancing). First, we select the assets with the following strategies: backward iteration, forward iteration, lasso regression. Here, we highlight the use of the backward and forward iteration algorithms that receive very little attention in the literature. With the selected assets, we form portfolios by minimizing the Tracking Error volatility. Then, we take the out of sample portfolio returns and analyze those returns in comparison to the selected Index.

We use two different datasets. The first is a Brazilian dataset of asset prices with daily frequency from January 1999 to December 2010 (2970 days). The second is an American dataset of asset prices with daily frequency from January 2010 to October 2017 (1950 days).

Our results show that, in general more active positions in the portfolio helps index tracking performance. That was already expected. Other general conclusion is that the quarterly rebalancing delivers better performance. Now for the more specific results, in the Brazilian case, the lasso selection method, specially the portfolio with at most 20 active positions, better tracks the Bovespa index. Note that this is due only to asset selection, not to the shrinkage of the portfolio weights that results from the constraints of the problem. This happens because the asset allocation method is the same for all portfolios. More on the shrinkage of portfolio weights can be found in Fan, Zhang and Yu (2012), Jagannathan and Ma (2003), Brodie et al. (2009). In the American case, the backward iteration portfolio with at most 40 assets and quarterly rebalancing is the portfolio that better tracks the benchmark. And even though the lasso method tracks the benchmark well in terms of mean returns, it presents more volatility. But, because our primary goal is to minimize Tracking Error volatility, this kind of risk-return tradeoff is not desirable in our case. This has to be highlighted, because those simple methods of feature selection receive too little attention in portfolio literature, even though they might bring good results.

Tu and Zhou (2011) and Liu (2009) point out that the majority of institutional investors are benchmarked by an index, e.g. the SP500. So, their performance has to be evaluated in relation to that index. Consequently, having to choose index tracking portfolio is closer to the practical life of an institutional investor. Index tracking can be defined as a quantitative method of trading and of passive portfolio management (MCWILLIAMS; MONTANA, 2010; WU; YANG; LIU, 2014). The most popular way to use this strategy is to form a portfolio of assets to mimic the performance of an externally specified index that, usually, is a market index such as the SP500. This performance takes

into account the risk-return profile of the specified benchmark.

There are two main ways of maintaining a fund that tracks an index; one is the full replication and the other is the partial replication. Full replication is the simplest one; it consists on maintaining all the assets with the same relative weights of the index being tracked. However, despite being possible to achieve a perfect replication with it, the full replication presents some disadvantages: (i) it results in small weights, which increase transaction costs; (ii) it presents instability of the weights, which means frequent revisions of them; (iii) its portfolio rebalancing can become complicated; and (iv) price fluctuation bring too many additions and subtractions of assets to the index.

It is also possible to try to replicate an index with a portfolio that contains only a small proportion of all the available assets, this is the partial replication. But this method also presents some problems: (i) the initial selection of the small number of assets to be included in the portfolio can be hard, and (ii) the need to estimate optimal weights (allocation) complicates the problem.

To implement an index tracking strategy, the investor has to solve two interconnected problems associated with it. The first is the asset selection problem, where the investor selects p assets from a universe of N assets. The second is the asset allocation problem, where the investor decides the relative weights of each selected asset.

The benefits of this two phase method are twofold. First, it isolates the effect of the selection methods, so the differences in performance may be attributed to differences in asset selection methods. Second, we make use of the Global Minimum Variance Portfolio (GMVP) in the second phase. The GMVP has several desirable properties discussed in Jagannathan and Ma (2003) and Clarke, Silva and Thorley (2006). As the goal of this essay is to compare the performance of the selection algorithms applied to index tracking and portfolio selection, this separation of phases is very suitable to our needs.

Next, we list papers that evaluated different methods of asset selection and portfolio optimization to implement index tracking and we compare their works to ours. Liu (2009) selects portfolios that track the SP500 by minimizing Tracking Error volatility without regard for the size of mean of the Tracking Error. In that paper, the asset selection phase is solved by only using the 30 assets in the Dow Jones Industrial Average (DJIA). Jansen and Dijk (2002) use index tracking with small portfolios by asset selection methods. McWilliams and Montana (2010) mention the possibility that the problems of asset selection and portfolio optimization can be put in the form of variable selection. They implement the asset selection with methods of penalized regression. Thus, we can use other methods of variable selection to choose the assets that will enter our portfolio. Here, we will use some methods listed in James et al. (2014).

Santos (2015) selects minimum variance portfolios using cardinality constraints to form portfolios with few assets and assesses their performance. Sant'Anna, Filomena and Borenstein (2014) select Index Tracking portfolios with cardinality constraints to

form portfolios with few assets and track the Bovespa Index. This paper resembles those, because we will form portfolios with few assets; however, we will use a different method. We will not use cardinality constraints, but feature selection algorithms. This can be justified by the lower computational cost of those algorithms. We will use only regressions to select the portfolio assets, the optimization of portfolios weights will occur in a different phase.

Other related works are Wu, Yang and Liu (2014) e Wu and Yang (2014), which use non-negative lasso and elastic net regressions to select assets. Those works use non-negative least-squares for portfolio optimization (weight selection). Yet another related work is Medeiros, Vasconcelos and Freitas (2015), which uses lasso regression to select variables to forecast inflation. This work differs from that in terms of the goal for which we use the lasso regression. Here we use it to form portfolios to track a financial Index.

Beyond this introduction, this essay is organized as follows. Section 4.1 presents the basics of portfolio selection that will be used throughout the essay. Section 4.2 presents the algorithms used to apply feature selection in the dataset. Section 4.3 presents the methodology of the empirical study. Section 4.4 shows the results obtained in the empirical study. Finally, Section 4.5 concludes.

4.1   BASIC DEFINITIONS

In this section, we present the basics of portfolio selection that we will use throughout this essay. We start by considering an investment universe with N risky assets. Each asset has a price in time t denoted by $P_{t,i}$. By holding asset i from $t-1$ to t, the investor earns the return

$$X_{t,i} = \frac{P_{t,i}}{P_{t-1,i}} - 1. \tag{4.1}$$

Let's denote the $N \times 1$ vector of future and uncertain returns of those assets in time t by X_t, where

$$X_t = [X_{t1}, X_{t2}, \ldots, X_{tN}].$$

We will assume that the returns have mean vector $E[X_t] = \mu_X$ and covariance matrix $V[X_t] = \Sigma$. We also denote the future and uncertain return of the index by y_t. We assume that the index return have mean $E[y_t] = \mu_y$ and variance $V[y_t] = \sigma_y^2$.

With those definitions, we can define the *excess returns* of the risky assets on the index as:

$$X_t^e = X_t - \mathbf{1}y_t.$$

where $\mathbf{1}$ is a conformable vector of ones. We assume that X_t^e has mean vector and covariance matrix

$$E[X_t^e] = \mu_X - \mathbf{1}\mu_y = \mu$$
$$V[X_t^e] = \Sigma + \sigma_y^2 \mathbf{1}\mathbf{1}' - 2Cov(X_t, y_t)\mathbf{1}' = \Omega.$$

Porfolio Returns

Next, we define portfolio and portfolio returns. A portfolio of the N risky assets is represented by a $N \times 1$ vector w, where

$$w = [w_1, w_2, \ldots, w_N]',$$

and w_i is the fraction of total wealth invested in asset i. The vector w can also be called the allocation vector.

Holding the portfolio w_t from t to $t+1$ yields the out of sample return in $t+1$,

$$X_{p,t+1} = w_t' X_{t+1}. \tag{4.2}$$

$X_{p,t+1}$ is a weighted average of the asset returns selected to the portfolio with weights w_i for $i = 1, \ldots, N$.

$$X_{p,t+1} = w_{t,1} X_{t+1,1} + \cdots + w_{t,N} X_{t+1,N} = \sum_{i=1}^{N} w_{t,i} X_{t+1,i} = w_t' X_{t+1}.$$

Evolution of weights

Liu (2009) points out that in the moment prior to rebalancing, each dollar invested in asset i in the portfolio has changed its value from $w_{t,i}$ to $w_{t,i}(1 + X_{t+1,i})$, where $w_{t,i}$ is the i-th element of w_t and the i-th element of X_t is denoted as $X_{t,i}$. Also, in general, each dollar invested in the whole portfolio has changed from w_t to $w_t(1 + X_{p,t+1})$. Therefore, prior to rebalancing, the weight on asset i has changed from $w_{i,t}$ to

$$w_{t,i}^{+1} = w_{t,i} \frac{1 + X_{t+1,i}}{1 + w_t' X_{p,t+1}}.$$

More generally, we can express equation the changed vector of weights as

$$w_t^{+1} = w_t \odot \frac{\mathbf{1} + X_{t+1}}{1 + X_{p,t+1}}, \tag{4.3}$$

where $\odot$ denotes the Hadamard (direct) product. Compounding another period, the portfolio w_t in $t+2$ will have changed to

$$w_t^{+2} = w_t \odot \frac{(\mathbf{1} + X_{t+1}) \odot (\mathbf{1} + X_{t+2})}{(1 + X_{p,t+1}) \times (1 + X_{p,t+2})}.$$

Following this logic, we reach a equation for h periods:

$$w_t^{+h} = w_t \odot \frac{(\mathbf{1} + X_{t+1}) \odot \cdots \odot (\mathbf{1} + X_{t+h})}{(1 + X_{p,t+1}) \times \cdots \times (1 + X_{p,t+h})}. \tag{4.4}$$

Tracking Error

The Tracking Error may be defined as a measure of the difference between the index tracking portfolio return, $X_{p,t}$ and the index return, y_t. So, we have the following expression for the Tracking Error:

$$TE_{t+1} = w_t' X_{t+1} - y_{t+1} = X_{p,t+1} - y_{t+1}, \tag{4.5}$$

where w_t is the allocation vector of our portfolio.

With the expression of the Tracking Error above, we can see that, if the weights are restricted to sum up to one ($w'\mathbf{1} = 1$), then the excess return of a portfolio on a benchmark is the same as the Tracking Error:

$$\begin{aligned} X_{p,t+1}^e = w_t' X_{t+1}^e &= w'(X_{t+1} - \mathbf{1}y_{t+1}) \\ &= w_t' X_{t+1} - w'\mathbf{1}y_{t+1} \\ X_{p,t+1}^e = X_{p,t+1} - y_{t+1} &= TE_{t+1}. \end{aligned} \tag{4.6}$$

Hence, if the weights sum up to one ($w'\mathbf{1} = 1$), minimizing the variance of the Tracking Error is equivalent to minimizing the variance of the excess returns of a portfolio on a benchmark. Note that the restriction is important.

4.2   FEATURE SELECTION ALGORITHMS

Here, we present the methods through which we will select the assets, they are: forward stepwise selection, backward stepwise selection, and lasso regression. More information about these methods can be found in James et al. (2014). The methods of asset selection used here depend heavily on the Ordinary Least Squares (OLS) Regression. Below, we offer more details about these methods, beginning with some notes on the OLS Regression.

4.2.1   Ordinary Least Squares (OLS) Regression

By regressing the index returns against the asset returns with the portfolio weights being parameters to estimate, we have a regression problem with stochastic regressors, more details about this problem can be found in Rao et al. (2008) and Rencher and Schaalje (2007). We may represent this problem as:

$$y_t = \alpha + X_t\beta + \varepsilon_t.$$

Note that the intercept α is necessary to let $E[\varepsilon_t] = 0$, so we can have $V[\varepsilon_t] = E[\varepsilon_t^2]$.

The OLS regression has the goal to minimize the sum of squared residuals, which is expressed as:

$$RSS(\beta) = \sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} (y_t - \alpha - X_t \beta)^2. \tag{4.7}$$

where ε_t is the residual of the regression.

By minimizing $RSS(\beta)$ in equation (4.7), we find β_{ols}:

$$\beta_{ols} = (X'X)^{-1}(X'y) \tag{4.8}$$

where X is the $T \times N$ matrix of asset returns, and y is the $T \times 1$ vector of index returns.

4.2.2 Forward Stepwise Selection

The forward stepwise selection uses a series of OLS regressions to select the assets that will enter the investor's portfolio, i.e. the columns that will enter the matrix X. It starts with a null model with no predictor, then we add one predictor at a time to the model until all predictors are added. More specifically, at each step the variable that provides the best fit is added to the model.

4.2.3 Backward Stepwise Selection

In opposition to the forward stepwise selection, the backward stepwise selection starts with the full OLS model, which contains all N predictors available. Then the least useful predictors are removed, one at a time, until we reach a model with the desired p predictors.

The forward and backward stepwise selections result in a set of models, each of which contains a subset of p predictors. Usually, the researcher determines which of the models in the set of models is the best. Here, however, we want a predefined number of assets in our portfolios. So we just choose the model with the predefined number p of assets with the smaller RSS and bigger R^2. To implement both of those algorithms, we use the R software (R Core Team, 2019) with the leaps package of Lumley (2017).

4.2.4 Lasso

To fit the least squares regression, we estimate the values of $\alpha, \beta_1, \cdots, \beta_p$ that minimize the RSS, as in equation (4.7). The lasso regression has a similar form, with its coefficients minimizing the RSS plus a penalty:

$$\sum_{t=1}^{T} \left(y_t - \alpha - \sum_{j=1}^{p} \beta_j x_{tj} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| = RSS + \lambda \sum_{j=1}^{p} |\beta_j|. \tag{4.9}$$

That is, the lasso is a penalized regression. In the equation above, $\lambda \geq 0$ controls the amount of regularization in the regression. When $\lambda = 0$, we have the least squares regression. The bigger the value of λ, the bigger the amount of regularization and tighter the variable selection will be. Thus, we have to fit λ in a convenient way to suit our objectives. In this essay we use the `glmnet` package of Friedman, Hastie and Tibshirani (2010) of the R Core Team (2019) software. Because the betas in the algorithms above might not sum to one, we only use the algorithms to select the assets. Wu, Yang and Liu (2014) and Wu and Yang (2014) also segregate the asset selection phase and the asset allocation phase by doing different regressions for each phase. Another reason to do this is that by using the same equation for all the portfolios' allocation, we can isolate the effect of the asset selection strategies.

4.3   EMPIRICAL EVALUATION

In this section, we present the methodology of our empirical evaluation. The goal of such evaluation is to evaluate the performance of the alternative asset selection techniques against each other and the benchmark index. It is done in three steps. The first step is the asset selection phase, which is done by using the algorithms from Section 4.2. The second step is asset allocation phase, where we form optimal portfolios in pseudo real time. The third step is the out of sample performance evaluation of those portfolios against each other and the benchmark.

4.3.1   Data

In this empirical analysis we have two different datasets. The first is the Brazilian dataset; it consists of daily closing prices for 61 stocks traded in the Bovespa and a series of daily closing prices for the Bovespa Index (IBovespa). The sample period used here is from January 1999 to December 2010 (2970 days). The second dataset is the American dataset; it consists of daily closing prices for 462 stocks and a series of daily closing prices for the SP500 index. The sample period used here is from January 2010 to October 2017 (1950 days).

From the price and index series, we take the simple daily raw return of each asset, according to equation (4.1). That will leave us with a total of $N + 1$ (N stocks plus one index) series of raw returns, with length of T, each. With the raw return series (X_t), we use the algorithms from Section 4.2 to select which assets will enter in our portfolios (asset selection phase).

Rolling Windows Estimators

Before we enter the asset allocation phase, we need to estimate the parameter input, $\hat{\Omega}$ used in the portfolio policies, we adopt a recursive estimation approach based on a *rolling window* time series of returns. With this approach, at each point in time $t = J, \ldots, T$, we use use the last J points of data and estimate:

$$\hat{\mu}_t = J^{-1} \sum_{j=0}^{J-1} X_{t-j}$$

$$\hat{\Omega}_t = (J-1)^{-1} \sum_{j=0}^{J-1} (X_{t-j} - \hat{\mu}_t)(X_{t-j} - \hat{\mu}_t)'$$

This results in $T - J - 1$ estimates of covariance matrices. For the Brazilian dataset, we use window length equal to $J = 250$ (approximately one year of daily data). So, we have 2719 estimates of covariance matrices. For the American dataset, we use window length equal to $J = 500$ (approximately 2 years of daily data), because the American dataset has more assets, and we need $J > N$ for the covariance matrix to be invertible. So, we have 1449 estimates of covariance matrices.

4.3.2 Allocation

After the covariance matrix estimation above, we enter the asset allocation phase. To that end, we follow Liu (2009) and select portfolios that minimize the volatility of the Tracking Error, without any regard for the size of the mean of the Tracking Error. To achieve this, we find the portfolio that minimizes the variance of the excess returns on the benchmark with a fully bought contraint ($e'w = 1$). The answer to the problem is:

$$w_t = \frac{\hat{\Omega}_t^{-1} e}{e' \hat{\Omega}_t^{-1} e} \tag{4.10}$$

where e is a $N \times 1$ vector of ones and $\hat{\Omega}_t$ is the covariance matrix estimator of Ω using information untill time t.

One clarification has to be made here, what the formula (4.10) finds is the portfolio with the least variance of excess returns, without any regard for the mean. So, following the index exactly and zeroing the Tracking Error is not a concern, only minimizing the volatility of the Tracking Error. Using the equation (4.10) can be defended on the grounds that it is the Global Minimum Variance Portfolio (GMVP). This portfolio achieves just what we want and it offers several desirable properties as presented by Jagannathan and Ma (2003) and Clarke, Silva and Thorley (2006). If we wanted to zero the Tracking Error, we would have to optimize a different objective function and use different restrictions in our portfolio. However, this portfolio would not have the properties of the GMVP highlighted above.

By the end of this process, we generate $T - J - 1$ allocation vectors for each asset selection algorithm, in each dataset. With the allocation vectors, we compute the *out of sample portfolio returns* with equation (4.2), which generates the time series $\{X_{p,t}\}_{t=J+1}^{T}$. This time series is the subject of our analysis.

4.3.3 Evaluation Measures

What we will evaluate are the out of sample portfolio returns in relation to the benchmark index. Since the goal of index tracking portfolio is to closely follow the series of interest, our evaluation has to consider how our portfolio performs in relation to the benchmark. Note that, as Tu and Zhou (2011) and Liu (2009) point out, the majority of institutional investors are benchmarked by an index, e.g. the SP500. So, their performance has to be evaluated in relation to that index. Consequently, having to choose index tracking portfolio is closer to the practical life of an institutional investor.

Transaction Costs

To get a sense of the amount of trading required to implement each portfolio strategy, we compute the portfolio turnover and report its average through time. Turnover in time t has the following formula

$$TO_t = |w_{t+1} - w_t^{+1}|, \tag{4.11}$$

where w_t^{+1} is the portfolio prior to rebalancing as defined in equation (4.3).

Associated with the turnover there is the associated transaction costs of rebalancing a portfolio. If the rebalancing is too frequent or too extreme, it can lead to overtrading. According to Barber and Odean (2000), overtrading has negative effects on portfolio performance due to its associated transaction costs. To take into account the impact of such costs we also evaluate the returns net of transaction costs, which have formula:

$$X_{p,t}^{net} = (1 + X_{p,t})(1 - cTO) - 1, \tag{4.12}$$

where c is the fee that must be paid for each transaction. We use $c = 50bp$ like in DeMiguel, Garlappi and Uppal (2009) and Kirby and Ostdiek (2012).

Weight Statistics

To gauge the amount of leverage in each portfolio, we report their *Gross Exposure* and their *Short Interest.* The Gross Exposure is expressed in proportion of the notional value of the portfolio (e.g. when the Gross Exposure is equal to one, it means that there is no leverage; but a Gross Exposure of 1.1 means that the portfolio is leveraged with

10% of its notional value). The Short Interest expresses the size of the short positions (weights smaller than zero) in the portfolio, it is expressed in percentage (e.g. when the Short Interest is equal to zero, it means that there is no short-selling happening; but when the Short Interest is equal to 10%, it means that the sum of of short positions are 10%). These measures have the following formulas:

$$\text{Gross}_t = \sum_{i=1}^{N} |w_{i,t}|$$

$$\text{Short}_t = \frac{\sum_{i=1}^{N} |w_{i,t}| - 1}{2}.$$

To assess how many assets our strategies select, we also report the *average number of assets in the portfolio*.

Hypothesis Testing

To verify if the Mean, Variance, and Sharpe Ratio of each policy are statiscally different from the benchmark index, we use the framework laid out in Ledoit and Wolf (2008, 2011, 2018). Their framework relies on the fact that those performance measures are smooth functions of population moments. We use the circular block bootstrap of Politis and Romano (1992) with block length $2, 4, 6, 8$, and 10 to find p-values of the differences in performance measures.

4.4 RESULTS

In this section we present the results of our empirical analysis. We do not annualize any performance measure due to the problem of time aggregation pointed out in Lo (2002). So the performance measures are all in daily terms. Each row of the tables below presents the statistics of a portfolio, and the first column of the rows presents the name of the portfolios. In the tables, IBOV is the Bovespa Index and SP500 is the SP500 Index. Also, in each table, `bwd` stands for the portfolio formed using the Backward Selection Algorithm, `fwd` stands for the portfolio formed using the Forward Selection Algorithm, and `las` stands for the portfolio formed using the Lasso Algorithm. The numbers after the letters denote the maximum number of active positions allowed in each portfolio.

Tables 4.1 and 4.5 show statistics of the Brazilian and American datasets, respectively. The presented statistics are: Mean, Standard Deviation (SD), Sharpe Ratio (SR), and correlation with the benchmark index (CORR) for the raw and net portfolio returns of each policy and each rebalancing frequency: monthly (Panel A), bimonthly (Panel B), and quarterly (Panel C). For the net portfolio returns, we used transaction costs equal to $50bp$ as in DeMiguel, Garlappi and Uppal (2009). The Tables also show the Turnover (TO), Gross Exposure (Gross), Short Interest (Short), and Active Positions (Active).

The ideal result of an index tracking portfolio is to track the benchmark index perfectly with zero difference in the mean and in the volatility. That would mean that our SR would also match the index's SR and we would have mimicked the index perfectly. Our optimization problem, however, does not take the mean returns into account. It only minimizes the mean squared differences of returns. So our results might have more differences in the mean returns, but not in the volatility.

Apropos of those mean results, we have to remember that trading and rebalancing the stocks in our portfolios generate transaction costs and those transaction costs will diminish our mean returns (BARBER; ODEAN, 2000). In this way, we have to decide what we really want to evaluate, raw portfolio returns, or net portfolio returns. On the one hand, if we are evaluating raw portfolio returns and we have a difference of zero with the benchmark, we must remember that the transaction costs will change that difference to a negative number, when evaluating net portfolio returns. But if we have a positive difference with raw returns, the transaction costs might be enough to bring that difference to zero with net returns. Or, in a unlucky case, bring that difference to a negative number. On the other hand, if we are evaluating net returns, we have more of a "what you see is what you get" type o scenario, because we already accounted for the transaction costs. With that said we only evaluate the p-value tables of the net return portfolios.

Tables 4.2, 4.3, and 4.4 show the p-values for the differences in the Sharpe Ratio, Log-Variance, and Mean of the net portfolio returns of the Brazilian dataset using the methodology developed in Ledoit and Wolf (2008, 2011, 2018). Tables 4.6, 4.7, and 4.8 show the p-values of those statistics for the American dataset.

Now, for the results that we are seeking to avoid. For the mean of returns, we do not want a significant negative difference. Actually, we do not want any negative difference at all, because it would mean that our portfolio has lower mean returns than the benchmark's. Acceptable results are nonnegative and nonsignificant differences. For the volatility of returns, we do not want a significant positive difference, because that would mean a higher volatitily than the benchmark's. Acceptable results are nonpositive and nonsignificant differences. For the Sharpe Ratio of returns, we do not want a significant negative difference, because that would mean a lower Sharpe Ratio than the benchmark's. Acceptable results are nonnegative nonsignificant differences.

4.4.1 Brazilian Dataset

The test of differences in Sharpe Ratios are presented in Table 4.2. Regarding this differences, there is no evidence of statistical significance for any portfolio. But we can observe that the only portfolios to never present a negative value of the difference are the `las.15` and `las.20` portfolios. Also, note that the `bwd.10` and `fwd.10` are never positive.

The test of difference with log variances are presented in Table 4.3. In it we can

observe that the `las.20` portfolio is the only portfolio that does not present evidence of statistical significance. The `las.20` portfolio present negative (lower variance than the benchmark's) nonsignificant differences. All the other portfolios present p-values lower than 10%. Noteworthy portfolios are the `bwd.15`, `fwd.15`, `bwd.20`, and `fwd.20` portfolios. These portfolios present negative statistically significant differences, which means that they have statistically significant lower variance than the benchmark.

The test of differences in mean portfolio net returns are presented in Table 4.4. In this table no portfolio has p-value under 15%, which means that ther is no evidence of statistical significance in the difference of mean net returns. Notwithstanding, it might be worth to mention that the `bwd.10` and `fwd.10` portfolios do present negative (lower mean than the benchmark's) differences; whereas all other differences are nonnegative.

It would be interesting to see figures of the returns and volatility of all the portfolios. However, that would take an inordinate amount of space. So, for the Brazilian dataset, we only show a few selected figures of portfolios. The `las.20` portfolio (Figure 4.1), because this is the only portfolio to not have any evidence of statistical significance in any difference. The `bwd.10` portfolio (Figure 4.2), because it has the highest negative difference in Sharpe Ratio. Finally, the `bwd.20` portfolio (Figure 4.3), because it has the highest negative difference in variance.

In summary, for the Brazilian dataset, the index tracking with asset selection did not find portfolios with significant statistical difference in the means, but did find portfolios with significant statistical differences in the log-variances. This, however, did not translate into significant statistical differences in the Sharpe Ratio. The only portfolio that did not present any statistically significant differences is the `las.20` portfolio. Also, we should point out that the quarterly rebalancing, for all portfolios, generally has a higher mean (due to the diminished transaction costs) and lower variance.

4.4.2 American Dataset

The test of differences in Sharpe Ratios are presented in Table 4.6. In it, the only portfolios that are consistently nonsignificant are the lasso portfolios, with the `las.20` portfolio also being the only portfolio to have a consistently positive difference. In the quarterly rebalancing frequency, all the differences are nonsignificant.

The test of difference with log variances are presented in Table 4.7. In it, we can observer that all differences are positive (bigger variance than the benchmark's). Also, the only portfolios to have a p-value above 10% are the `bwd.40` in the bimonthly and quarterly rebalancing frequency and the `bwd.30` in the quarterly rebalancing frequency.

The test of differences in mean portfolio net returns are presented in Table 4.8. The lasso portfolios are consistently nonsignificant and nonnegative (mean portfolio net return at least as good as the benchmark). In the quarterly rebalancing frequency, all

differences are nonsignificant.

Once more, we only show the figures of a few selected portfolios to save space. The `las.20` portfolio (Figure 4.4), because it has the highest positive difference in the variance and the Sharpe Ratio. The `fwd.20` portfolio (Figure 4.5), because of its poor performance. The `bwd.40` portfolio (Figure 4.6), because in its quarterly rebalancing form it is the one of the two portfolios to not show any evidence of statistical significance (in this rebalancing frequency).

For the American dataset, we can sum up the results as follows. Again, the quarterly rebalancing, generally has a higher mean (due to the diminished transaction costs) and lower variance and, in general, more assets in the portfolio enhance performance. The `bwd.40` portfolio with quarterly rebalancing tracks the SP500 index well, which is not repeated by the other rebalancing frequencies. The lasso portfolios generally surpass the performance of the index, but costing a higher volatility and tracking error.

4.5   CONCLUDING REMARKS

In this essay, we presented portfolios formed by methods of variable selection to track a benchmark index. We used two different datasets for two different countries. One is an American dataset to track the SP500 Index; the other is a Brazilian dataset to track the Bovespa Index. In the empirical exercise, we segregated the asset selection and the asset allocation phases to isolate the effect of asset selection for each method. In the asset allocation phase we used the Global Minimum Variance Portfolio with returns in excess of the benchmark to minimize the Tracking Error volatility of the portfolio, as in Liu (2009).

It should be pointed out that this optimization does not take the mean returns into account. It only minimizes the mean squared differences of returns. So our results might have more differences in the mean returns, but less differences in the volatility. Further research should study the case where we predict the actual transaction costs and stablish a margin of error just enough to zero the return difference.

In the results section we can observe that, for the Brazilian dataset, the `las.20` portfolio is the portfolio that best tracks the Bovespa index, while for the American dataset, the `bwd.40` with quarterly rebalancing is the best tracker, but with an honorable mention to the lasso portfolios. The positive performance of the portfolios leads us to believe that simple variable selection methods can add value to index tracking portfolios. By simple methods we mean methods that are not computationally costly as Integer Quadratic Programming like in Santos (2015), Sant'Anna, Filomena and Borenstein (2014).

Table 4.1 – Out of Sample Rebalancing for the Index Tracking Portfolios (IBOV)

	Mean	SD	SR	CORR	TO	Gross	Short	Active	Mean	SD	SR	CORR
index	0.07	1.99	0.04	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Panel A: Monthly Rebalancing												
bwd.10	0.07	2.02	0.04	0.96	2.13	1.00	0.04	10.00	0.06	2.02	0.03	0.96
fwd.10	0.08	2.02	0.04	0.96	2.18	1.00	0.12	10.00	0.07	2.02	0.03	0.96
las.10	0.08	2.09	0.04	0.96	1.40	1.01	0.37	9.99	0.08	2.09	0.04	0.96
bwd.15	0.08	1.96	0.04	0.96	1.99	1.02	1.25	15.00	0.07	1.96	0.04	0.96
fwd.15	0.08	1.96	0.04	0.96	2.08	1.02	0.83	15.00	0.07	1.96	0.03	0.96
las.15	0.08	2.02	0.04	0.96	1.41	1.01	0.27	14.93	0.08	2.02	0.04	0.96
bwd.20	0.08	1.93	0.04	0.97	1.96	1.08	4.02	20.00	0.07	1.93	0.04	0.97
fwd.20	0.08	1.94	0.04	0.97	2.00	1.07	3.30	20.00	0.07	1.94	0.04	0.97
las.20	0.08	1.99	0.04	0.97	1.44	1.01	0.57	19.95	0.07	1.99	0.04	0.97
Panel B: Bimonthly Rebalancing												
bwd.10	0.08	2.02	0.04	0.96	1.46	1.00	0.00	10.00	0.07	2.02	0.03	0.96
fwd.10	0.08	2.03	0.04	0.96	1.41	1.00	0.24	10.00	0.07	2.03	0.03	0.96
las.10	0.09	2.08	0.04	0.96	1.01	1.01	0.48	9.99	0.08	2.08	0.04	0.96
bwd.15	0.08	1.95	0.04	0.96	1.36	1.03	1.36	15.00	0.07	1.95	0.04	0.96
fwd.15	0.08	1.96	0.04	0.96	1.41	1.02	1.04	15.00	0.07	1.96	0.04	0.96
las.15	0.09	2.02	0.04	0.96	0.98	1.01	0.30	14.97	0.08	2.02	0.04	0.96
bwd.20	0.08	1.93	0.04	0.97	1.35	1.09	4.29	20.00	0.07	1.93	0.04	0.97
fwd.20	0.08	1.93	0.04	0.97	1.34	1.07	3.56	20.00	0.07	1.93	0.04	0.97
las.20	0.08	1.98	0.04	0.97	1.01	1.01	0.49	19.93	0.08	1.98	0.04	0.97
Panel C: Quarterly Rebalancing												
bwd.10	0.08	2.02	0.04	0.96	1.17	1.00	0.00	10.00	0.07	2.02	0.03	0.96
fwd.10	0.07	2.02	0.04	0.96	1.11	1.00	0.00	10.00	0.07	2.01	0.03	0.96
las.10	0.08	2.08	0.04	0.96	0.81	1.01	0.40	10.00	0.08	2.07	0.04	0.96
bwd.15	0.08	1.95	0.04	0.96	1.11	1.02	1.10	15.00	0.07	1.94	0.04	0.96
fwd.15	0.08	1.95	0.04	0.96	1.06	1.02	0.84	15.00	0.07	1.95	0.04	0.96
las.15	0.08	2.01	0.04	0.96	0.79	1.00	0.22	15.00	0.08	2.01	0.04	0.96
bwd.20	0.08	1.92	0.04	0.97	1.11	1.08	4.00	20.00	0.07	1.92	0.04	0.97
fwd.20	0.08	1.92	0.04	0.97	1.05	1.07	3.29	20.00	0.07	1.92	0.04	0.97
las.20	0.08	1.98	0.04	0.97	0.77	1.01	0.54	19.96	0.08	1.98	0.04	0.97

Source: The author.

Table 4.2 – p-values for Difference of SR of Portfolio Net Returns for the Brazilian Dataset

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: Monthly Rebalancing						
bwd.10	-0.63	16.34	16.38	15.56	17.02	16.28
fwd.10	-0.46	31.38	31.06	30.08	30.94	30.60
las.10	-0.03	95.30	95.06	94.66	95.10	95.00
bwd.15	-0.13	74.64	74.50	74.32	74.96	74.72
fwd.15	-0.21	60.28	62.36	62.76	61.94	60.94
las.15	0.05	91.22	91.30	91.10	91.14	89.92
bwd.20	-0.12	76.92	75.58	75.32	75.42	75.88
fwd.20	-0.09	81.42	82.04	81.48	81.06	81.62
las.20	0.01	96.80	96.94	96.78	97.20	97.10
Panel B: Bimonthly Rebalancing						
bwd.10	-0.31	47.72	48.48	47.14	48.02	48.90
fwd.10	-0.20	66.22	67.04	66.74	66.04	66.86
las.10	0.16	71.76	73.24	72.46	73.12	72.34
bwd.15	0.16	69.14	68.00	69.52	67.76	68.84
fwd.15	0.04	91.18	91.38	91.58	91.58	90.92
las.15	0.41	33.40	32.92	31.88	32.14	33.76
bwd.20	0.09	82.84	83.52	82.90	83.16	83.42
fwd.20	0.19	61.44	62.28	63.12	61.44	60.96
las.20	0.27	49.30	50.50	51.74	50.12	51.08
Panel C: Quarterly Rebalancing						
bwd.10	-0.24	57.60	58.20	58.28	59.34	58.48
fwd.10	-0.40	35.48	35.98	36.82	36.08	35.94
las.10	0.17	72.14	71.72	71.64	71.84	72.20
bwd.15	0.09	82.70	82.10	82.50	82.28	82.50
fwd.15	0.07	86.82	86.26	86.76	86.72	86.58
las.15	0.23	60.42	60.06	61.46	58.70	59.64
bwd.20	0.10	82.36	80.68	81.48	80.62	81.14
fwd.20	0.16	68.12	69.46	70.00	71.32	70.04
las.20	0.17	67.24	66.68	66.36	67.28	67.76

Source: The author.

Table 4.3 – p-values for Difference of Log Variance of Portfolio Net Returns for the Brazilian Dataset

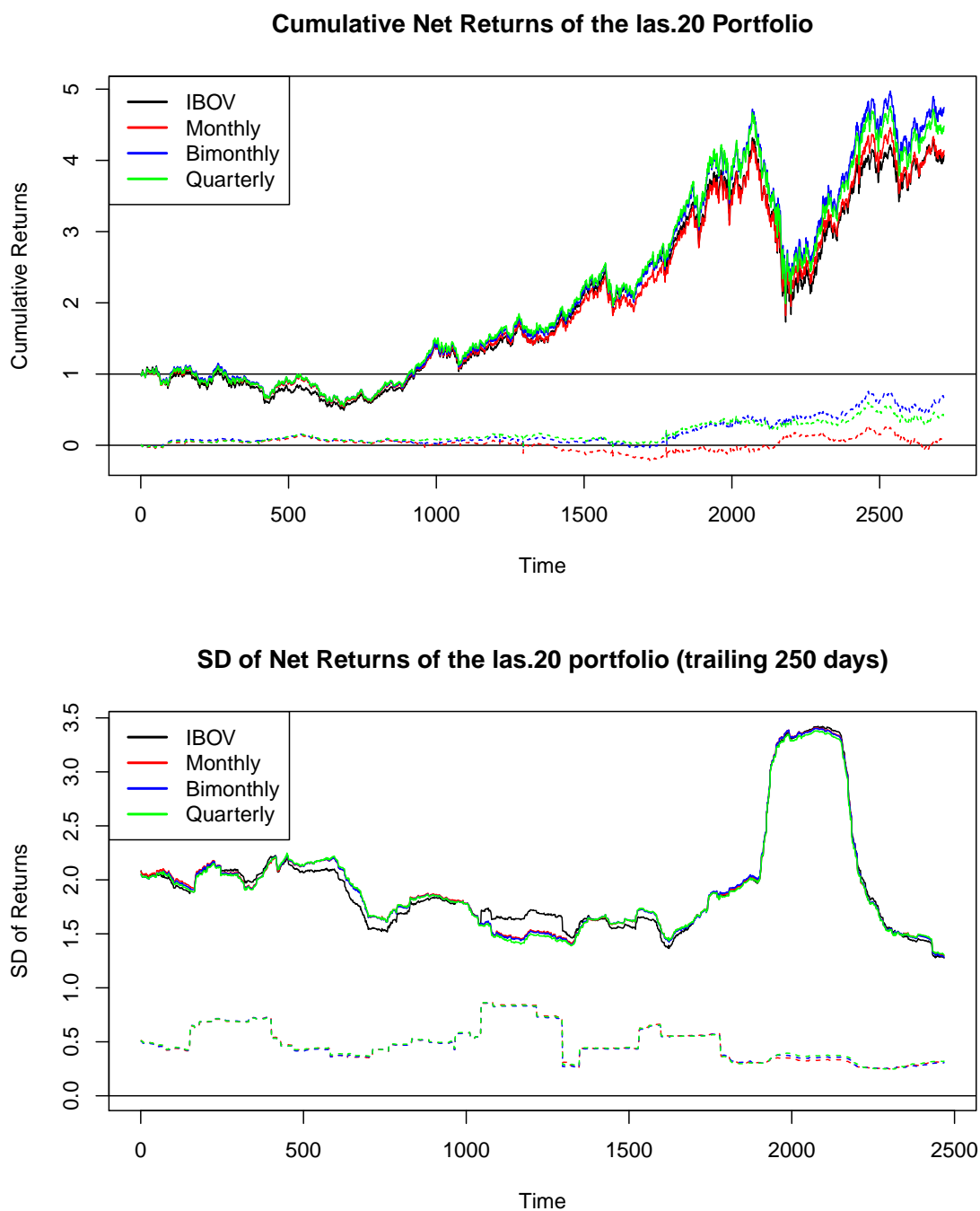|	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: Monthly Rebalancing						
bwd.10	3.20	3.74	3.90	3.44	3.50	3.24
fwd.10	3.89	0.92	1.10	1.18	1.14	1.44
las.10	9.92	0.02	0.02	0.02	0.02	0.02
bwd.15	−2.91	4.26	4.62	5.62	4.48	5.02
fwd.15	−2.40	9.02	8.84	8.68	8.68	8.62
las.15	3.56	1.34	1.84	1.68	1.68	1.60
bwd.20	−5.59	0.02	0.02	0.04	0.02	0.02
fwd.20	−5.05	0.10	0.16	0.10	0.12	0.12
las.20	0.18	89.26	88.36	88.88	88.96	89.26
Panel B: Bimonthly Rebalancing						
bwd.10	3.09	5.22	5.18	5.46	5.18	5.36
fwd.10	4.64	0.28	0.18	0.50	0.30	0.36
las.10	9.68	0.02	0.02	0.02	0.02	0.02
bwd.15	−3.74	1.20	1.18	1.34	1.22	1.52
fwd.15	−2.36	8.26	8.88	8.72	8.14	8.26
las.15	3.23	2.20	2.24	2.30	2.46	2.24
bwd.20	−5.94	0.04	0.06	0.02	0.02	0.02
fwd.20	−5.60	0.02	0.02	0.04	0.02	0.02
las.20	−0.29	82.74	82.64	82.00	83.52	82.16
Panel C: Quarterly Rebalancing						
bwd.10	3.03	5.20	5.28	5.50	6.22	5.16
fwd.10	2.82	5.00	5.48	5.68	5.46	5.78
las.10	8.72	0.02	0.02	0.02	0.02	0.02
bwd.15	−4.21	0.70	0.84	0.78	1.00	0.92
fwd.15	−3.77	1.30	1.24	1.16	1.50	1.34
las.15	2.60	6.16	6.32	6.24	6.50	6.42
bwd.20	−6.77	0.02	0.02	0.02	0.04	0.02
fwd.20	−6.74	0.02	0.02	0.02	0.02	0.02
las.20	−0.98	46.32	46.44	46.96	47.16	46.40

Source: The author.

Table 4.4 – p-values for Difference of Mean of Portfolio Net Returns for the Brazilian Dataset

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: Monthly Rebalancing						
bwd.10	−0.01	19.96	20.96	19.34	18.94	19.86
fwd.10	−0.01	37.82	38.34	38.52	39.12	38.14
las.10	0.00	75.66	74.94	75.28	75.38	74.86
bwd.15	0.00	66.42	64.64	64.70	64.78	65.52
fwd.15	0.00	53.42	52.96	54.16	53.84	53.46
las.15	0.00	77.28	78.12	76.68	77.66	76.98
bwd.20	0.00	56.54	57.50	56.86	56.42	56.74
fwd.20	0.00	63.78	63.48	63.98	63.04	64.16
las.20	0.00	95.96	96.30	96.10	95.96	96.32
Panel B: Bimonthly Rebalancing						
bwd.10	−0.01	55.88	53.20	56.84	56.60	57.24
fwd.10	0.00	81.60	80.02	80.50	81.00	80.98
las.10	0.01	46.96	47.22	47.22	48.42	47.68
bwd.15	0.00	81.20	82.38	80.90	81.84	82.12
fwd.15	0.00	99.76	99.82	99.94	99.94	99.86
las.15	0.01	23.80	24.78	24.94	25.42	25.72
bwd.20	0.00	95.44	94.94	95.26	95.44	95.54
fwd.20	0.00	81.10	81.66	81.52	82.88	82.54
las.20	0.01	48.18	46.68	48.64	49.08	48.90
Panel C: Quarterly Rebalancing						
bwd.10	0.00	66.88	67.38	66.12	67.68	67.34
fwd.10	−0.01	43.44	44.24	43.98	43.12	43.38
las.10	0.01	48.92	47.16	46.92	48.02	47.98
bwd.15	0.00	96.80	97.22	97.18	97.06	96.76
fwd.15	0.00	99.66	99.62	99.56	99.66	99.66
las.15	0.01	52.30	51.26	51.24	50.74	51.48
bwd.20	0.00	95.12	95.38	95.52	95.50	96.10
fwd.20	0.00	93.72	93.84	93.96	94.60	94.02
las.20	0.00	68.74	69.74	68.62	68.40	69.50

Source: The author.

Figure 4.1 – `las.20` Portfolio (Brazilian Data)

Cumulative Net Returns of the las.20 Portfolio

SD of Net Returns of the las.20 portfolio (trailing 250 days)

Source: The author.

Figure 4.2 – `bwd.10` Portfolio (Brazilian Data)

Cumulative Net Returns of the bwd.10 Portfolio

SD of Net Returns of the bwd.10 portfolio (trailing 250 days)
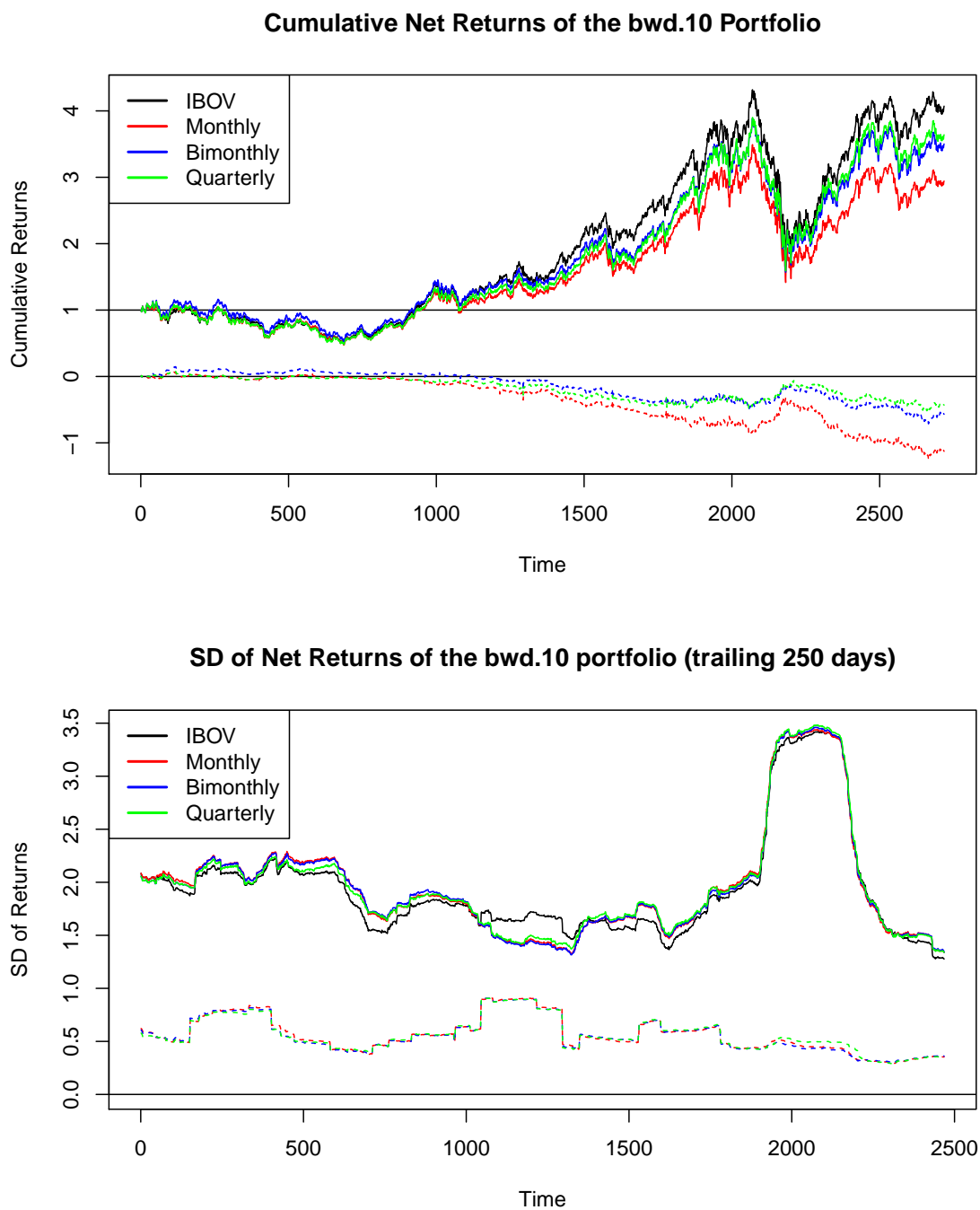
Source: The author.

Figure 4.3 – `bwd.20` Portfolio (Brazilian Data)

Cumulative Net Returns of the bwd.20 Portfolio

SD of Net Returns of the bwd.20 portfolio (trailing 250 days)
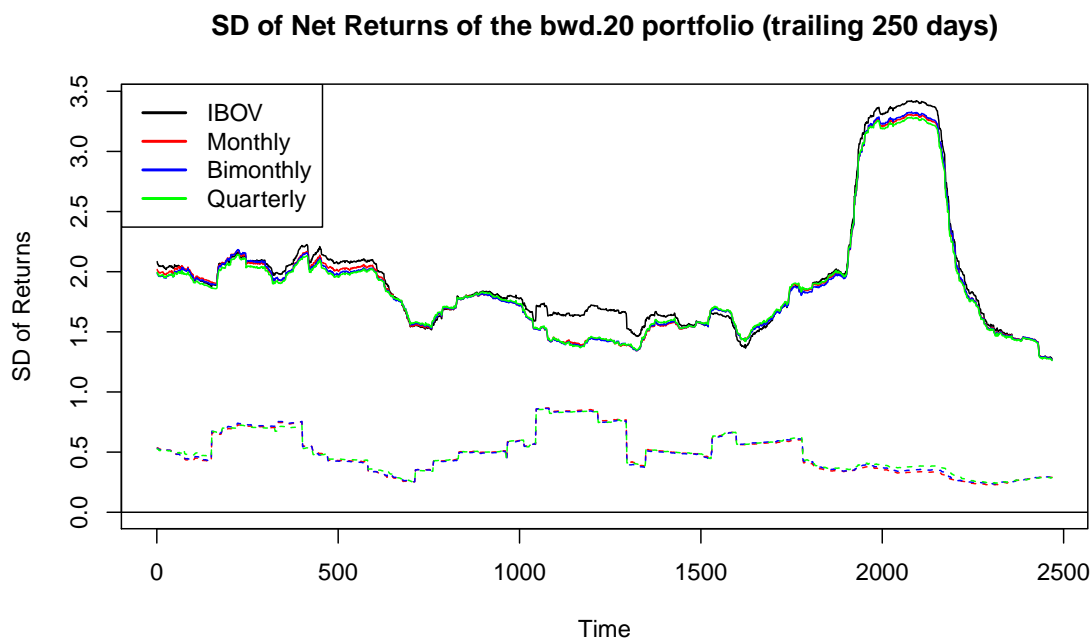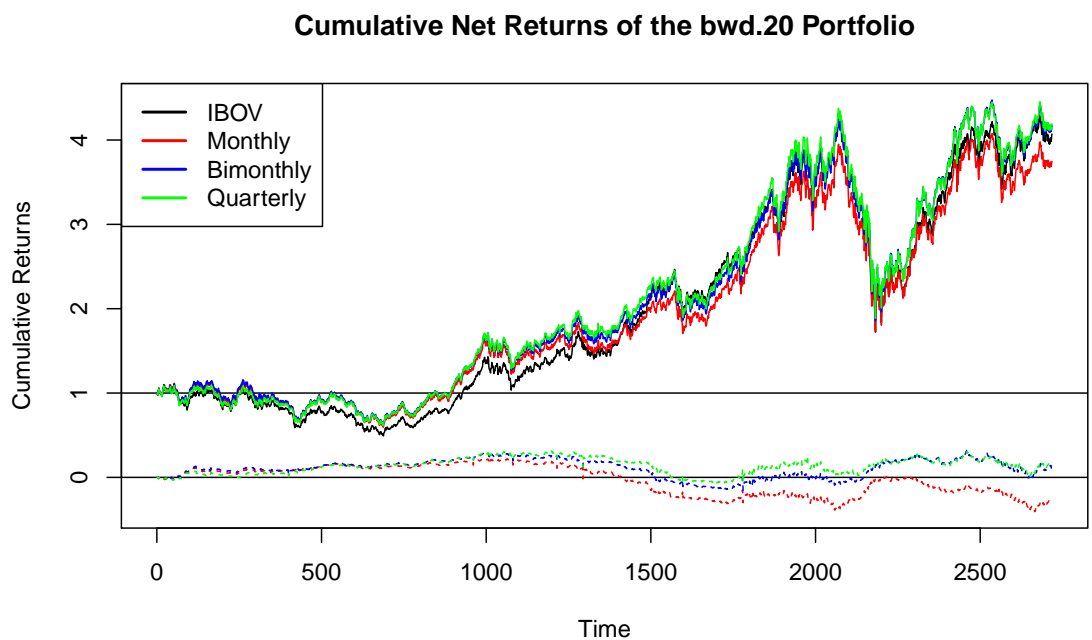
Source: The author.

Table 4.5 – Out of Sample Rebalancing for the Index Tracking Portfolios (SP500)

	Mean	SD	SR	CORR	TO	Gross	Short	Active	Mean	SD	SR	CORR
index	0.05	0.77	0.07	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Panel A: Monthly Rebalancing												
bwd.20	0.06	0.79	0.08	0.98	6.42	1.00	0.00	20.00	0.03	0.81	0.03	0.96
fwd.20	0.05	0.80	0.07	0.97	5.39	1.00	0.01	20.00	0.03	0.81	0.03	0.96
las.20	0.06	0.81	0.08	0.95	1.82	1.06	2.93	19.93	0.06	0.81	0.07	0.95
bwd.30	0.05	0.78	0.07	0.98	6.07	1.00	0.00	30.00	0.02	0.80	0.03	0.97
fwd.30	0.06	0.79	0.07	0.98	5.43	1.00	0.04	30.00	0.03	0.80	0.04	0.97
las.30	0.06	0.80	0.08	0.96	1.89	1.09	4.51	29.88	0.05	0.80	0.07	0.96
bwd.40	0.06	0.78	0.07	0.99	5.66	1.00	0.00	40.00	0.03	0.79	0.04	0.98
fwd.40	0.06	0.79	0.07	0.99	5.48	1.00	0.06	40.00	0.03	0.80	0.04	0.97
las.40	0.06	0.79	0.07	0.97	2.07	1.09	4.46	39.88	0.05	0.79	0.06	0.97
Panel B: Bimonthly Rebalancing												
bwd.20	0.05	0.78	0.07	0.97	3.43	1.00	0.00	20.00	0.04	0.79	0.05	0.96
fwd.20	0.06	0.79	0.07	0.98	2.82	1.00	0.02	20.00	0.04	0.80	0.05	0.97
las.20	0.06	0.82	0.08	0.95	1.18	1.06	2.95	19.92	0.06	0.82	0.07	0.95
bwd.30	0.05	0.78	0.07	0.98	3.33	1.00	0.00	30.00	0.03	0.78	0.04	0.97
fwd.30	0.06	0.79	0.08	0.98	2.84	1.00	0.03	30.00	0.05	0.79	0.06	0.98
las.30	0.06	0.79	0.08	0.96	1.32	1.09	4.34	29.92	0.06	0.80	0.07	0.96
bwd.40	0.05	0.77	0.07	0.99	3.09	1.00	0.00	40.00	0.04	0.78	0.05	0.98
fwd.40	0.06	0.78	0.08	0.99	2.81	1.00	0.04	40.00	0.05	0.79	0.06	0.98
las.40	0.06	0.78	0.08	0.97	1.45	1.09	4.67	39.92	0.05	0.79	0.07	0.97
Panel C: Quarterly Rebalancing												
bwd.20	0.06	0.78	0.08	0.98	2.49	1.00	0.00	20.00	0.05	0.79	0.07	0.97
fwd.20	0.06	0.80	0.07	0.97	2.06	1.00	0.00	20.00	0.05	0.80	0.06	0.97
las.20	0.06	0.81	0.08	0.95	1.05	1.06	2.87	19.96	0.06	0.81	0.07	0.95
bwd.30	0.06	0.77	0.08	0.98	2.30	1.00	0.00	30.00	0.05	0.77	0.07	0.98
fwd.30	0.06	0.79	0.08	0.98	2.08	1.00	0.01	30.00	0.05	0.79	0.06	0.98
las.30	0.06	0.80	0.08	0.96	1.10	1.09	4.52	29.92	0.06	0.80	0.07	0.96
bwd.40	0.06	0.77	0.08	0.99	2.11	1.00	0.00	40.00	0.05	0.77	0.07	0.98
fwd.40	0.06	0.78	0.08	0.99	2.04	1.00	0.02	40.00	0.05	0.78	0.06	0.98
las.40	0.06	0.79	0.08	0.97	1.17	1.08	4.22	39.88	0.05	0.79	0.07	0.97

Source: The author.

Table 4.6 – p-values for Difference of SR of Portfolio Net Returns for the American Dataset

	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: Monthly Rebalancing						
bwd.20	−3.29	0.02	0.02	0.02	0.02	0.02
fwd.20	−3.34	0.02	0.02	0.02	0.02	0.02
las.20	0.17	84.14	84.60	85.12	85.14	85.32
bwd.30	−3.63	0.02	0.02	0.02	0.02	0.02
fwd.30	−3.08	0.02	0.02	0.02	0.02	0.02
las.30	−0.11	87.54	88.58	88.54	86.60	88.04
bwd.40	−3.07	0.02	0.02	0.02	0.02	0.02
fwd.40	−2.90	0.02	0.02	0.02	0.02	0.02
las.40	−0.73	27.12	27.68	28.80	28.12	27.62
Panel B: Bimonthly Rebalancing						
bwd.20	−1.88	0.94	0.84	0.94	1.12	0.86
fwd.20	−1.21	7.80	7.36	7.66	7.20	8.28
las.20	0.46	61.96	60.94	61.38	61.34	61.56
bwd.30	−2.32	0.04	0.04	0.10	0.02	0.10
fwd.30	−0.74	21.58	21.76	21.22	22.14	21.48
las.30	0.26	74.52	74.86	73.88	73.40	74.54
bwd.40	−1.82	0.08	0.16	0.18	0.20	0.20
fwd.40	−0.74	15.68	16.22	15.06	16.42	16.34
las.40	−0.04	95.40	95.74	96.00	95.42	95.46
Panel C: Quarterly Rebalancing						
bwd.20	−0.03	96.22	96.10	96.16	95.82	96.26
fwd.20	−0.55	43.12	41.20	41.42	41.68	41.60
las.20	0.38	67.72	66.94	68.04	68.26	69.38
bwd.30	0.01	98.82	98.64	98.60	98.50	98.56
fwd.30	−0.18	76.02	76.48	74.96	75.80	76.54
las.30	0.29	69.66	70.06	70.18	70.52	69.64
bwd.40	0.20	68.48	69.76	68.68	69.24	68.00
fwd.40	−0.24	62.20	64.40	63.10	64.02	63.88
las.40	0.12	84.96	85.00	85.82	85.12	84.38

Source: The author.

Table 4.7 – *p*-values for Difference of Log Variance of Portfolio Net Returns for the American Dataset

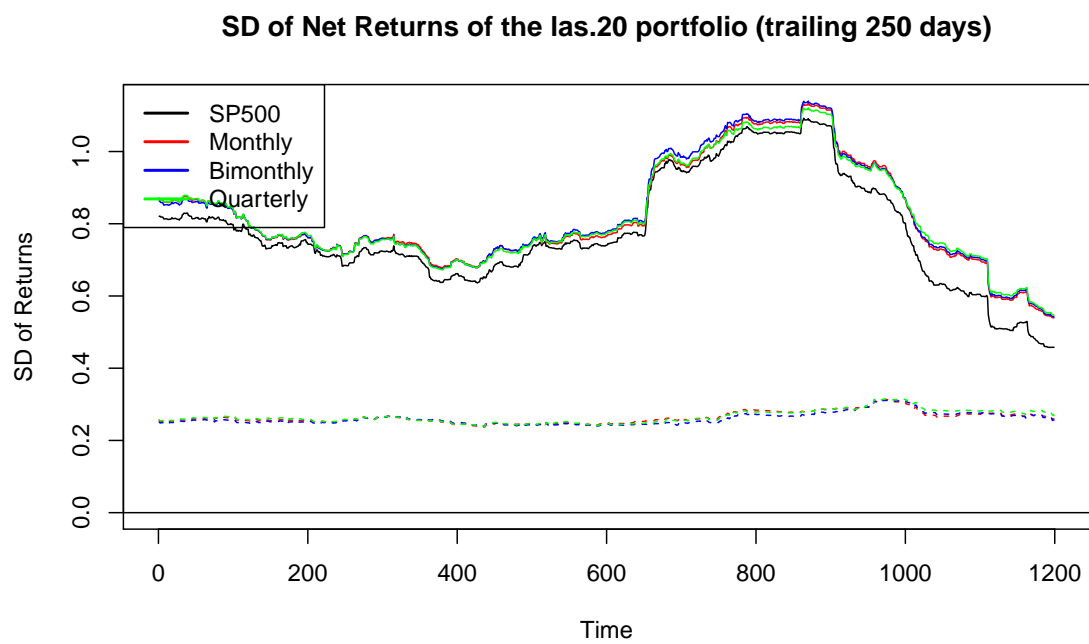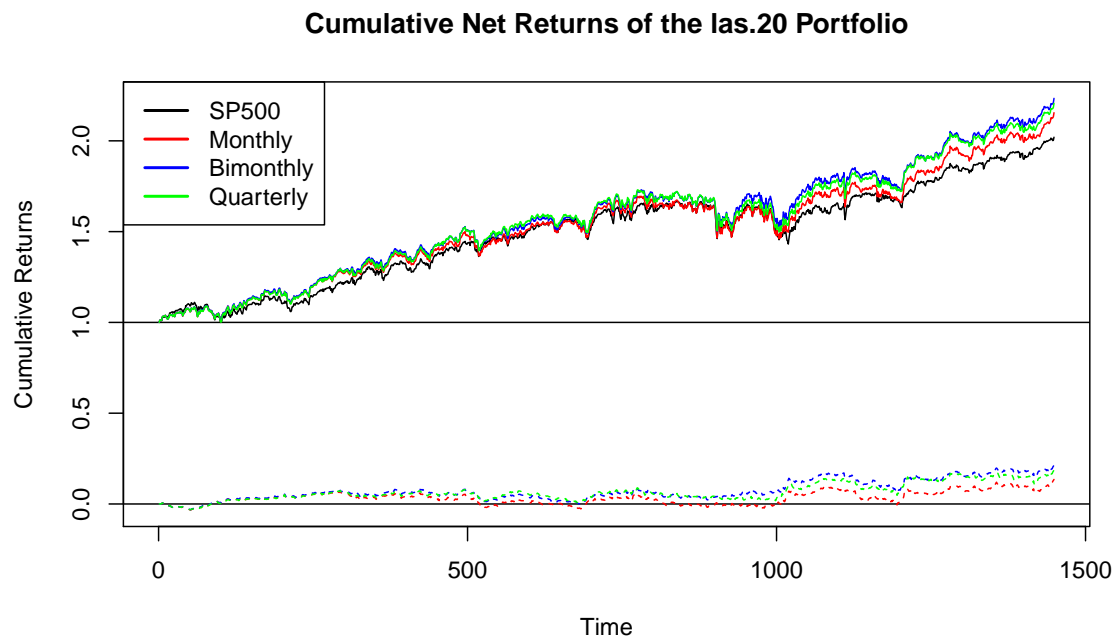|	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: Monthly Rebalancing						
bwd.20	9.29	0.02	0.02	0.02	0.02	0.02
fwd.20	11.09	0.02	0.02	0.02	0.02	0.02
las.20	11.28	0.02	0.02	0.02	0.02	0.02
bwd.30	6.84	0.02	0.02	0.02	0.04	0.04
fwd.30	8.47	0.02	0.02	0.02	0.02	0.02
las.30	6.90	0.02	0.02	0.02	0.04	0.02
bwd.40	5.04	0.14	0.16	0.08	0.16	0.12
fwd.40	7.17	0.02	0.02	0.02	0.02	0.02
las.40	5.56	0.04	0.04	0.08	0.04	0.08
Panel B: Bimonthly Rebalancing						
bwd.20	5.08	0.30	0.22	0.26	0.36	0.42
fwd.20	7.58	0.02	0.02	0.02	0.02	0.02
las.20	11.89	0.02	0.02	0.02	0.02	0.02
bwd.30	3.50	2.14	1.98	1.70	1.46	2.48
fwd.30	5.64	0.02	0.04	0.02	0.04	0.02
las.30	6.66	0.02	0.08	0.06	0.12	0.02
bwd.40	2.00	14.32	14.10	15.72	15.08	15.02
fwd.40	4.28	0.04	0.06	0.06	0.14	0.14
las.40	3.98	0.84	0.82	1.30	1.02	1.36
Panel C: Quarterly Rebalancing						
bwd.20	4.30	0.52	0.58	0.56	0.48	0.60
fwd.20	8.68	0.02	0.02	0.02	0.02	0.02
las.20	11.13	0.02	0.02	0.02	0.02	0.02
bwd.30	1.07	41.98	42.88	43.08	42.66	42.48
fwd.30	5.73	0.02	0.02	0.02	0.02	0.02
las.30	7.30	0.02	0.02	0.02	0.02	0.02
bwd.40	0.67	56.16	57.38	56.04	55.98	56.94
fwd.40	3.55	0.10	0.08	0.06	0.16	0.10
las.40	5.29	0.04	0.06	0.08	0.08	0.02

Source: The author.

Table 4.8 – p-values for Difference of Mean of Portfolio Net Returns for the American Dataset

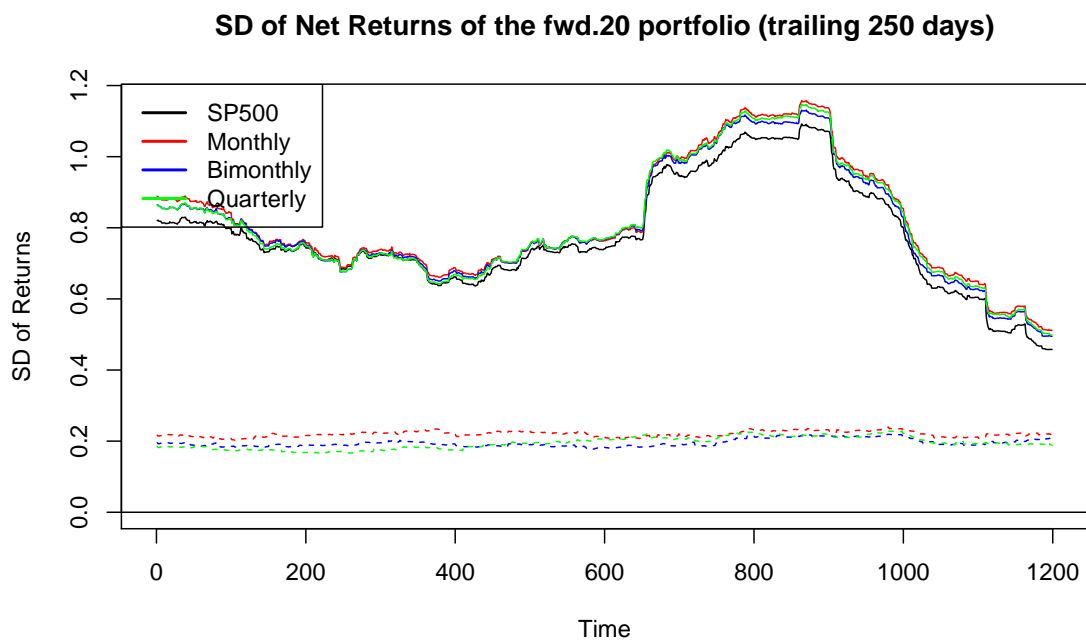|	Difference	$b = 02$	$b = 04$	$b = 06$	$b = 08$	$b = 10$
Panel A: Monthly Rebalancing						
bwd.20	−0.02	0.02	0.02	0.02	0.02	0.02
fwd.20	−0.02	0.02	0.02	0.02	0.02	0.02
las.20	0.00	53.82	55.42	54.70	55.34	55.46
bwd.30	−0.03	0.02	0.02	0.02	0.02	0.02
fwd.30	−0.02	0.02	0.02	0.02	0.02	0.02
las.30	0.00	87.22	87.20	87.86	86.38	87.54
bwd.40	−0.02	0.02	0.02	0.02	0.02	0.02
fwd.40	−0.02	0.02	0.02	0.02	0.02	0.02
las.40	0.00	40.02	39.84	41.18	40.56	40.00
Panel B: Bimonthly Rebalancing						
bwd.20	−0.01	1.82	1.56	1.40	1.22	1.68
fwd.20	−0.01	13.62	13.80	13.30	14.84	13.92
las.20	0.01	32.84	33.04	33.06	33.12	34.38
bwd.30	−0.02	0.06	0.04	0.02	0.06	0.12
fwd.30	0.00	34.08	34.30	34.40	32.98	33.02
las.30	0.00	52.00	51.08	51.04	52.24	53.70
bwd.40	−0.01	0.06	0.08	0.16	0.10	0.14
fwd.40	0.00	25.62	24.66	25.30	26.04	27.30
las.40	0.00	88.58	89.02	88.68	88.84	88.96
Panel C: Quarterly Rebalancing						
bwd.20	0.00	86.54	87.54	87.02	87.04	86.86
fwd.20	0.00	68.94	69.34	68.92	68.00	68.72
las.20	0.01	40.46	41.84	40.26	41.96	42.48
bwd.30	0.00	93.24	93.64	93.50	93.58	94.02
fwd.30	0.00	98.40	98.32	98.58	98.26	98.32
las.30	0.00	46.04	46.76	45.66	47.14	47.20
bwd.40	0.00	65.12	64.92	65.78	64.08	65.34
fwd.40	0.00	79.76	80.98	81.14	80.96	81.14
las.40	0.00	63.78	63.36	64.44	64.40	63.82

Source: The author.
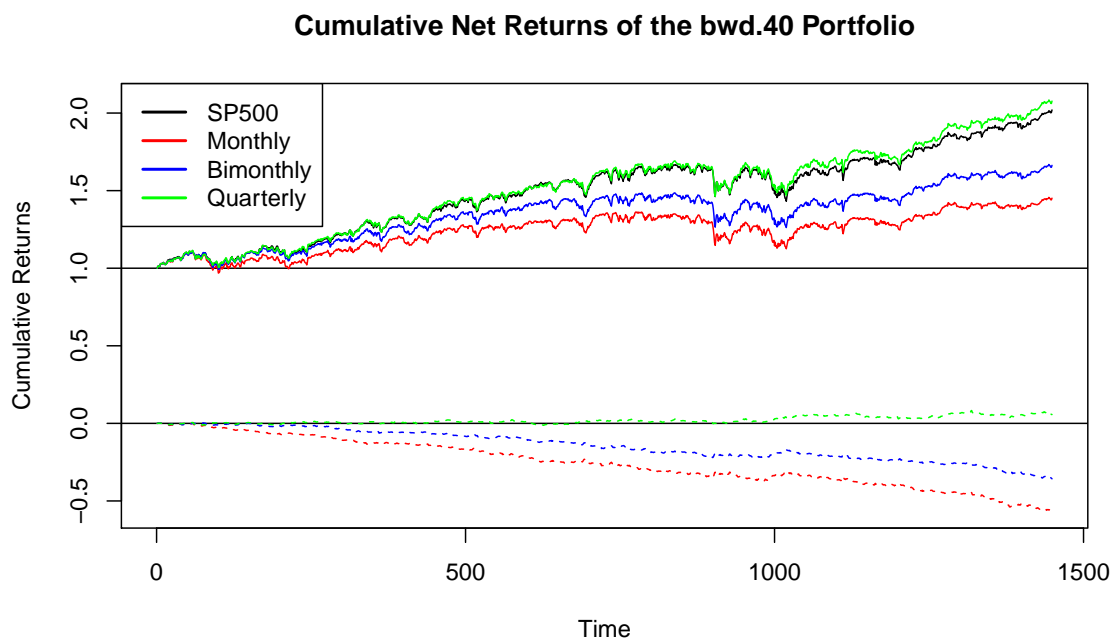
Figure 4.4 – `las.20` Portfolio (American Data)

Cumulative Net Returns of the las.20 Portfolio

SD of Net Returns of the las.20 portfolio (trailing 250 days)

Source: The author.

Figure 4.5 – `fwd.20` Portfolio (American Data)

Cumulative Net Returns of the fwd.20 Portfolio

SD of Net Returns of the fwd.20 portfolio (trailing 250 days)
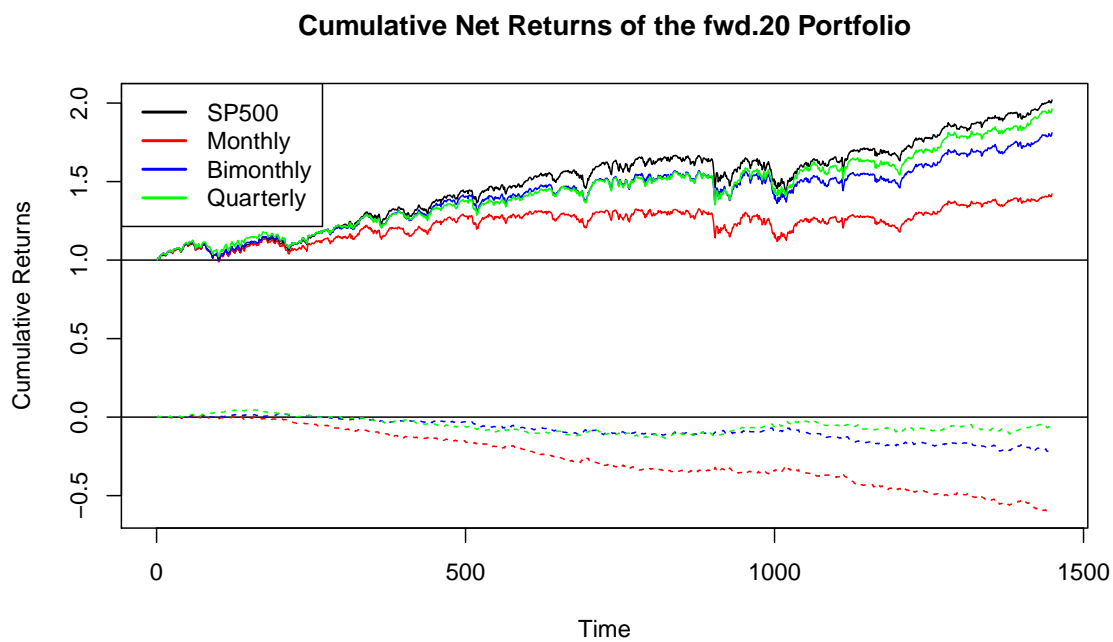
Source: The author.

Figure 4.6 – `bwd.40` Portfolio (American Data)

Cumulative Net Returns of the bwd.40 Portfolio

SD of Net Returns of the bwd.40 portfolio (trailing 250 days)
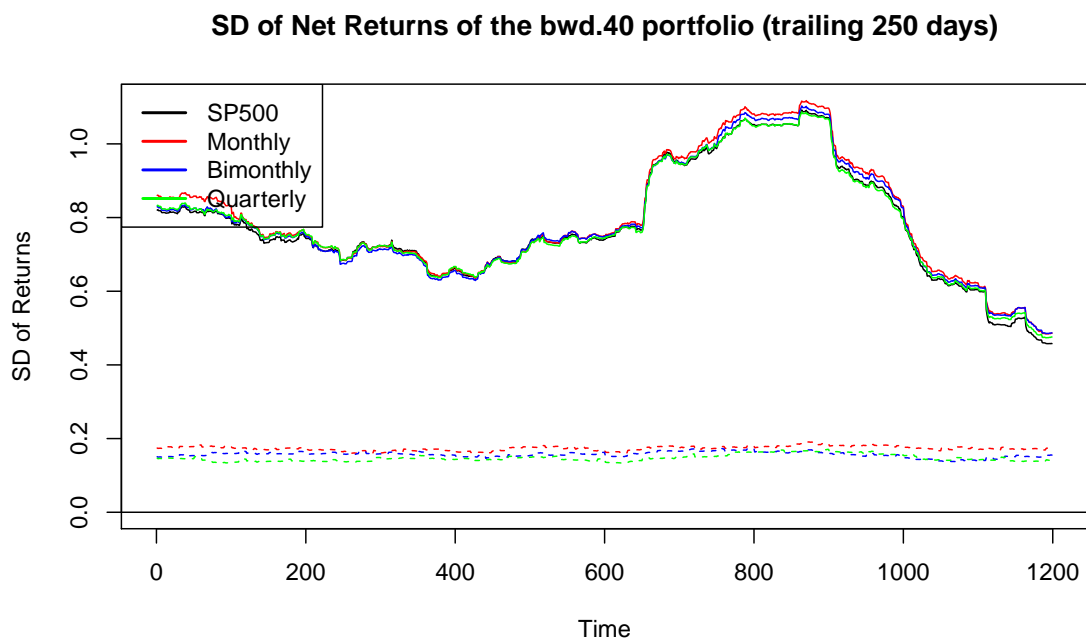
Source: The author.

5 FINAL CONSIDERATIONS

In this thesis, we presented three different essays in the context of portfolio optimization and index tracking. The first essay offers a critique of the Tangency Portfolio (TP).[1] This critique is due to the widespread misconception that the TP is the most efficient mean-variance portfolio as can be seen in DeMiguel, Garlappi and Uppal (2009) and Kempf and Memmel (2006) among others. In that essay, we present three propositions that show how the TP can be outperformed by other optimal portfolios. We also develop an empirical evaluation by testing the performance (in and out of sample) of the TP against other portfolios. The data used is from the Kenneth French database in the period from July 1963 to June 2018. The result is that the TP delivers poor performance and is outperformed by the other portfolios. This corroborates the work in Kirby and Ostdiek (2012) and, to a lesser extent, the work of Clarke, Silva and Thorley (2006) about the Global Minimum Variance Portfolio. The ultimate takeaway of that essay is that the TP should not be used in empirical comparisons, seen as it will always come at the bottom of the reap.

The second essay provides an explicit derivation of the equivalence between the minimum variance portfolio and the minimum Tracking Error volatility portfolio. This equivalence is much talked about in intuitive terms but seldom demonstrated analytically (LIU, 2009; DEMOL, 2016). This derivation relies on the Sherman-Morison formula, that shows how to invert the sum of two matrices with a certain structure. A further equivalence of those portfolios to a restricted OLS regression is shown. Finally, a 2-fund portfolio result of those portfolios is presented. The main contribution of this essay, besides the direct analytical derivation, is to present the connection among general portfolio optimization, index tracking portfolio optimization and OLS regressions.

The third essay is about feature selection in the context of index tracking. Index tracking is an important subject in portfolio optimization because, as Liu (2009) and Tu and Zhou (2011) point out, the majority of managers are benchmarked by broad market indices like the SP500. In that essay, we provide an empirical study comparing different feature selection algorithms in the context of portfolio optimization. The algorithms we use are the backward iteration, forward iteration, and lasso regression. To isolate the effect of the selection algorithms, we separate the asset selection phase from the asset allocation phase. In the asset allocation phase we follow Liu (2009) and find portfolios that minimize the Tracking Error volatility from a chosen benchmark index, without regard for the size of the Tracking Error mean. We use two different datasets for the empirical study; one is an American dataset from January 2010 to October 2017 where we try to follow the SP500 Index; the other is a Brazilian dataset from January 1999 to

---

[1] A preliminary version of this essay was accepted to the XIX Econtro Brasileiro de Finanças with the title "Tangency Portfolio: A Critique"

December 2010 where we try to follow the Bovespa Index. The results we find is that for the Brazilian case the lasso is the better tracker of the Bovespa Index. For the American case, the lasso also delievers better risk-adjusted performance, but with higher volatility than the other methods. So the better tracker is the backward iteration method with quarterly rebalancing. It is important to highlight that the other methods do not receive much attention in the literature, even though they deliver good results and have rather simple ideas in their basis.

REFERENCES

BARBER, B. M.; ODEAN, T. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, Malden, v. 55, n. 2, p. 773–806, 2000.

BEST, M. *Portfolio Optimization*. Boca Raton: Taylor & Francis, 2010. (Chapman & Hall/CRC finance series).

BEST, M. J.; GRAUER, R. R. Sensitivity Analysis for Mean-Variance Portfolio Problems. *Management Science*, Hanover, v. 37, n. 8, p. 980–989, Aug 1991.

BODNAR, T.; PAROLYA, N.; SCHMID, W. On the equivalence of quadratic optimization problems commonly used in portfolio theory. *European Journal of Operational Research*, Amsterdam, v. 229, n. 3, p. 637 – 644, 2013.

BRANDT, M. W. Portfolio choice problems. In: AÏT-SAHALIA, Y.; HANSEN, L. P. (ed.). *Handbook of Financial Econometrics*. San Diego: North-Holland, 2010, (Handbooks in Finance, v. 1). cap. 5, p. 269 – 336.

BRITTEN-JONES, M. The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance*, Malden, v. 54, n. 2, p. 655–671, 1999.

BRODIE, J. et al. Sparse and stable markowitz portfolios. *Proceedings of the National Academy of Sciences*, Allahabad, v. 106, n. 30, p. 12267–12272, 2009.

CHAN, L.; KARCESKI, J.; LAKONISHOK, J. On portfolio optimization: forecasting covariances and choosing the risk model. *Review of Financial Studies*, Oxford, v. 12, n. 5, p. 937–974, 1999.

CHOPRA, V. K.; ZIEMBA, W. T. The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, New York, v. 19, n. 2, p. 6–11, 1993.

CLARKE, R. G.; SILVA, H. de; THORLEY, S. Minimum-variance portfolios in the u.s. equity market. *Journal of Portfolio Management*, New York, v. 33, n. 1, p. 10–24, 2006.

CONSTANTINIDES, G.; MALLIARIS, A. Portfolio theory. In: JARROW, R.; MAKSIMOVIC, V.; ZIEMBA, W. (ed.). *Finance*. Amsterdam: Elsevier, 1995, (Handbooks in Operations Research and Management Science, v. 9). cap. 1, p. 1 – 30.

DEMIGUEL, V.; GARLAPPI, L.; UPPAL, R. Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies*, New York, v. 22, n. 5, p. 1915–1953, May 2009.

DEMOL, C. Sparse markowitz portfolios. In: AKANSU, A. N.; KULKARNI, S. R.; MALIOUTOV, D. (ed.). *Financial Signal Processing and Machine Learning*. Hoboken: Wiley-IEEE Press, 2016. cap. 2, p. 11–22.

FAN, J.; ZHANG, J.; YU, K. Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association*, New York, v. 107, n. 498, p. 592–606, 2012.

FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, California, v. 33, n. 1, p. 1–22, 2010.

JAGANNATHAN, R.; MA, T. Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance*, Malden, v. 58, n. 4, p. 1651–1684, 08 2003.

JAMES, G. et al. *An Introduction to Statistical Learning: With Applications in R*. New York: Springer, 2014.

JANSEN, R.; DIJK, R. van. Optimal benchmark tracking with small portfolios. *The Journal of Portfolio Management*, New York, v. 28, n. 2, p. 33–39, 2002.

JOBSON, J. D.; KORKIE, B. Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association*, v. 75, n. 371, p. 544–554, 1980.

KAN, R.; ZHOU, G. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, Seattle, v. 42, n. 3, p. 621–656, 2007.

KEMPF, A.; MEMMEL, C. Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, Düsseldorf, v. 58, n. 4, p. 332–348, Oct 2006.

KIRBY, C.; OSTDIEK, B. It's all in the timing: Simple active portfolio strategies that outperform naïve diversification. *Journal of Financial and Quantitative Analysis*, Seattle, v. 47, n. 2, p. 437–467, 2012.

KOLM, P. N.; TüTüNCü, R.; FABOZZI, F. J. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, Amsterdam, v. 234, n. 2, p. 356 – 371, 2014.

KRITZMAN, M.; PAGE, S.; TURKINGTON, D. In defense of optimization: The fallacy of 1/n. *Financial Analysts Journal*, New York, v. 66, n. 2, p. 31–39, 2010.

LEDOIT, O.; WOLF, M. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, Amsterdam, v. 15, n. 5, p. 850–859, Dec 2008.

LEDOIT, O.; WOLF, M. Robust performance hypothesis testing with the Variance. *Wilmott Magazine*, Chichester, p. 86–89, Sept 2011.

LEDOIT, O.; WOLF, M. *Robust Performance Hypothesis Testing with Smooth Functions of Population Moments*. Zürich, 2018.

LIU, Q. On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics*, Chichester, v. 24, n. 4, p. 560–582, 2009.

LO, A. W. The statistics of sharpe ratios. *Financial Analysts Journal*, Abingdon, v. 58, n. 4, p. 36–52, 2002.

LUMLEY, T. *leaps: Regression Subset Selection*. [S.l.], 2017. R package version 3.0.

MARKOWITZ, H. Portfolio Selection. *Journal of Finance*, New York, v. 7, n. 1, p. 77–91, 03 1952.

MARKOWITZ, H. *Portfolio Selection: Efficient Diversification of Investments*. New Haven: Yale University Press, 1959. (Cowles Commission for Research in Economics: Monographs).

MCWILLIAMS, B.; MONTANA, G. Sparse partial least squares regression for on-line variable selection with multivariate data streams. *Statistical Analysis and Data Mining*, Hoboken, v. 3, n. 3, p. 170–193, 2010.

MEDEIROS, M. C.; VASCONCELOS, G.; FREITAS, E. Forecasting brazilian inflation with high-dimensional models. *Brazilian Review of Econometrics*, Rio de Janeiro, v. 99, n. 2, p. nil, 2015.

MERTON, R. C. An Analytic Derivation of the Efficient Portfolio Frontier. *Journal of Financial and Quantitative Analysis*, Seattle, v. 7, n. 04, p. 1851–1872, Sept 1972.

MERTON, R. C. On estimating the expected return on the market : An exploratory investigation. *Journal of Financial Economics*, v. 8, n. 4, p. 323–361, December 1980.

OKHRIN, Y.; SCHMID, W. Distributional properties of portfolio weights. *Journal of Econometrics*, Amsterdam, v. 134, n. 1, p. 235 – 256, 2006.

PETERSEN, K. B.; PEDERSEN, M. S. *The Matrix Cookbook*. Lyngby: Technical University of Denmark, 2012. Version 20121115.

POLITIS, D. N.; ROMANO, J. P. A circular block-resampling procedure for stationary data. In: LEPAGE, R.; BILLARD, L. (ed.). *Exploring the limits of bootstrap*. New York: John Wiley, 1992.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, 2019.

RAO, C. R. et al. *Linear Models and Generalizations*. 3rd. ed. Berlin; New York: Springer, 2008. (Springer Series in Statistics).

RENCHER, A. C.; SCHAALJE, G. B. *Linear Models in Statistics*. 2nd. ed. New York; Chichester: Wiley, Interscience, 2007.

ROLL, R. A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of Financial Economics*, v. 4, n. 2, p. 129 – 176, 1977.

SANT'ANNA, L.; FILOMENA, T.; BORENSTEIN, D. Index tracking with control on the number of assets. *Brazilian Review of Finance*, v. 12, n. 1, p. 89–119, 2014.

SANTOS, A. A. Beating the market with small portfolios: Evidence from brazil. *EconomiA*, Brasília, v. 16, n. 1, p. 22 – 31, 2015.

TU, J.; ZHOU, G. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, v. 99, n. 1, p. 204–215, January 2011.

WU, L.; YANG, Y. Nonnegative elastic net and application in index tracking. *Applied Mathematics and Computation*, New York, v. 227, p. 541 – 552, 2014.

WU, L.; YANG, Y.; LIU, H. Nonnegative-lasso and application in index tracking. *Computational Statistics & Data Analysis*, Amsterdam, v. 70, p. 116 – 126, 2014.